

(11) 공개번호 10-2020-0095534
(43) 공개일자 2020년08월10일

(71) 출원인
지머젠 인코포레이티드
미국 94608 캘리포니아 에머리빌 스위트 105 호턴
스트리트 5980

(72) 발명자
트레이스웰 캐라
미국 94608 캘리포니아주 에머리빌 호턴 스트리트
5980 넘버105

에드가 스티븐
미국 94608 캘리포니아주 에머리빌 호턴 스트리트
5980 넘버105
(뒷면에 계속)

(74) 대리인
특허법인코리아나

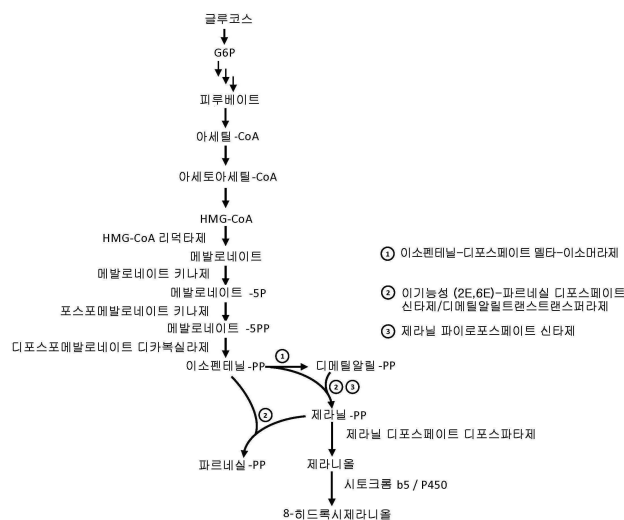
전체 청구항 수 : 총 44 항

(54) 발명의 명칭 발효에 의한 (6E)-8-히드록시제라니올의 제조를 위해 조작된 생합성 경로

(57) 요약

본 개시는 (6E)-8-히드록시제라니올의 발효 생산을 위한 미생물 세포의 조작을 기술하고, 신규한 조작된 미생물 세포 및 배양물을 비롯하여 관련 (6E)-8-히드록시제라니올 제조 방법을 제공한다.

대표도



(52) CPC특허분류

C12N 9/16 (2013.01)

C12N 9/90 (2013.01)

C12P 7/18 (2013.01)

C12Y 503/03002 (2013.01)

C12Y 114/13152 (2015.07)

C12Y 301/07011 (2015.07)

(72) 발명자

티모센코 스테판

미국 94608 캘리포니아주 에머리빌 호턴 스트리트
5980 넘버105

시어러 알렉산더

미국 94608 캘리포니아주 에머리빌 호턴 스트리트
5980 넘버105

명세서

청구범위

청구항 1

조작된 미생물 세포로서, 조작된 미생물 세포는

- (a) 비천연 제라닐 디포스페이트 디포스파타제 (제라니올 신타제); 및
- (b) 비천연 제라니올-8-히드록실라제

를 발현하고; 조작된 미생물 세포는 (6E)-8-히드록시제라니올을 생산하는 것인 조작된 미생물 세포.

청구항 2

제1항에 있어서, 조작된 미생물 세포는 하나 이상의 상류 (6E)-8-히드록시제라니올 경로 효소(들) 또는 상류 경로 활성의 조절인자의 증가된 활성을 포함하고, 상기 증가된 활성은 대조군 세포에 비해서 증가되는 것인 조작된 미생물 세포.

청구항 3

제2항에 있어서, 하나 이상의 상류 (6E)-8-히드록시제라니올 경로 효소(들)는 ATP-시트레이트 신타제, 아세틸-CoA 신싸타제, 티올라제, 히드록시메틸글루타릴 조효소 A 신타제 (HMG-CoA 신타제), 히드록시메틸글루타릴 조효소 A 리덕타제 (HMG-CoA 리덕타제), 메발로네이트 키나제, 포스포메발로네이트 키나제, 디포스포메발로네이트 디카복실라제, 이소펜테닐-디포스페이트 델타-이소머라제, 및 제라닐 디포스페이트 신타제로 이루어진 군으로부터 선택되는 것인 조작된 미생물 세포.

청구항 4

제3항에 있어서, 하나 이상의 상류 (6E)-8-히드록시제라니올 경로 효소(들)는 이소펜테닐-디포스페이트 델타-이소머라제를 포함하는 것인 조작된 미생물 세포.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 조작된 미생물 세포는 하나 이상의 (6E)-8-히드록시제라니올 경로 전구체를 소모하는 하나 이상의 효소(들)의 감소된 활성을 포함하고, 상기 감소된 활성은 대조군 세포에 비해서 감소되는 것인 조작된 미생물 세포.

청구항 6

제5항에 있어서, 하나 이상의 (6E)-8-히드록시제라니올 경로 전구체를 소모하는 하나 이상의 효소(들)는 이기능성 (2E,6E)-파르네실 디포스페이트 신타제/디메틸알릴트랜스트랜스퍼라제 및/또는 제라닐 파이로포스페이트 신타제를 포함하는 것인 조작된 미생물 세포.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서, 조작된 미생물 세포는 피드백-탈조절 HMG-CoA 리덕타제를 추가로 발현하는 것인 조작된 미생물 세포.

청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서, 조작된 미생물 세포는 대조군 세포에 비해서, 더 높은 속도의 아세틸-CoA 합성 및/또는 더 낮은 속도의 아세틸-CoA 분해에 기인하여 아세틸-CoA의 증가된 이용률을 포함하는 것인 조작된 미생물 세포.

청구항 9

조작된 미생물 세포로서, 조작된 미생물 세포는

- (a) 비천연 천연 제라닐 디포스페이트 디포스파타제 (제라니올 신타제)를 발현시키기 위한 수단; 및
 - (b) 비천연 제라니올-8-히드록실라제를 발현시키기 위한 수단
- 을 포함하고; 조작된 미생물 세포는 (6E)-8-히드록시제라니올을 생산하는 것인 조작된 미생물 세포.

청구항 10

제9항에 있어서, 조작된 미생물 세포는 하나 이상의 상류 (6E)-8-히드록시제라니올 경로 효소(들) 또는 상류 경로 활성의 조절인자의 활성을 증가시키기 위한 수단을 포함하는 것인 조작된 미생물 세포.

청구항 11

제10항에 있어서, 하나 이상의 상류 (6E)-8-히드록시제라니올 경로 효소(들)는 ATP-시트레이트 신타제, 아세틸-CoA 신싸타제, 티올라제, 히드록시메틸글루타릴 조효소 A 신타제 (HMG-CoA 신타제), 히드록시메틸글루타릴 조효소 A 리덕타제 (HMG-CoA 리덕타제), 메발로네이트 키나제, 포스포메발로네이트 키나제, 디포스포메발로네이트 디카복실라제, 이소펜테닐-디포스페이트 델타-이소머라제, 및 제라닐 디포스페이트 신타제로 이루어진 군으로부터 선택되는 것인 조작된 미생물 세포.

청구항 12

제11항에 있어서, 하나 이상의 상류 (6E)-8-히드록시제라니올 경로 효소(들)는 이소펜테닐-디포스페이트 델타-이소머라제를 포함하는 것인 조작된 미생물 세포.

청구항 13

제9항 내지 제12항 중 어느 한 항에 있어서, 조작된 미생물 세포는 하나 이상의 (6E)-8-히드록시제라니올 경로 전구체를 소모하는 하나 이상의 효소(들)의 활성을 감소시키기 위한 수단을 포함하고, 상기 감소된 활성은 대조군 세포에 비해서 감소되는 것인 조작된 미생물 세포.

청구항 14

제13항에 있어서, 하나 이상의 (6E)-8-히드록시제라니올 경로 전구체를 소모하는 하나 이상의 효소(들)는 이기능성 (2E,6E)-파르네실 디포스페이트 신타제/디메틸알릴트랜스트랜스퍼라제 및/또는 제라닐 파이로포스페이트 신타제를 포함하는 것인 조작된 미생물 세포.

청구항 15

제9항 내지 제14항 중 어느 한 항에 있어서, 조작된 미생물 세포는 피드백-탈조절 HMG-CoA 리덕타제를 발현시키기 위한 수단을 추가로 포함하는 것인 조작된 미생물 세포.

청구항 16

제9항 내지 제15항 중 어느 한 항에 있어서, 조작된 미생물 세포는 대조군 세포에 비해서, 더 높은 속도의 아세틸-CoA 합성 및/또는 더 낮은 속도의 아세틸-CoA 분해에 기인하여 아세틸-CoA의 이용률을 증가시키기 위한 수단을 포함하는 것인 조작된 미생물 세포.

청구항 17

제1항 내지 제16항 중 어느 한 항에 있어서, 조작된 미생물 세포는 진균 세포를 포함하는 것인 조작된 미생물 세포.

청구항 18

제17항에 있어서, 조작된 미생물 세포는 효모 세포를 포함하는 것인 조작된 미생물 세포.

청구항 19

제18항에 있어서, 효모 세포는 사카로마이세스 (*Saccharomyces*) 속의 세포인 조작된 미생물 세포.

청구항 20

제19항에 있어서, 효모 세포는 세레비지아 (*cerevisiae*) 종의 세포인 조작된 미생물 세포.

청구항 21

제18항에 있어서, 효모 세포는 야로위아 (*Yarrowia*) 속의 세포인 조작된 미생물 세포.

청구항 22

제21항에 있어서, 효모 세포는 리폴리티카 (*lipolytica*) 종의 세포인 조작된 미생물 세포.

청구항 23

제1항 내지 제22항 중 어느 한 항에 있어서, 비천연 제라니올 신타제는 페릴라 세토이엔시스 (*Perilla setoyensis*) 유래 제라니올 신타제와 적어도 70% 아미노산 서열 동일성을 갖는 제라니올 신타제를 포함하는 것인 조작된 미생물 세포.

청구항 24

제1항 내지 제22항 중 어느 한 항에 있어서, 비천연 제라니올 신타제는 비티스 비니페라 (*Vitis vinifera*) 유래 제라니올 신타제와 적어도 70% 아미노산 서열 동일성을 갖는 제라니올 신타제를 포함하는 것인 조작된 미생물 세포.

청구항 25

제1항 내지 제23항 중 어느 한 항에 있어서, 비천연 제라니올-8-히드록실라제는 파세올러스 안굴라리스 (*Phaseolus angularis*) 유래 제라니올-8-히드록실라제와 적어도 70% 아미노산 서열 동일성을 갖는 제라니올-8-히드록실라제를 포함하는 것인 조작된 미생물 세포.

청구항 26

제4항 또는 제12항 내지 제25항 중 어느 한 항에 있어서, 이소펜테닐-디포스페이트 델타-이소머라제의 증가된 활성은 이소펜테닐-디포스페이트 델타-이소머라제를 이중으로 발현시켜서 획득되는 것인 조작된 미생물 세포.

청구항 27

제26항에 있어서, 이중성 이소펜테닐-디포스페이트 델타-이소머라제는 사카로마이세스 세레비지아 (*Saccharomyces cerevisiae*) 유래 이소펜테닐-디포스페이트 델타-이소머라제와 적어도 70% 아미노산 서열 동일성을 갖는 이소펜테닐-디포스페이트 델타-이소머라제를 포함하는 것인 조작된 미생물 세포.

청구항 28

제6항, 또는 제14항 내지 제27항 중 어느 한 항에 있어서, 하나 이상의 (6E)-8-히드록시제라니올 경로 전구체를 소모하는 하나 이상의 효소(들)는 이기능성 (2E,6E)-파르네실 디포스페이트 신타제/디메틸알릴트랜스트랜스퍼라제를 포함하는 것인 조작된 미생물 세포.

청구항 29

제28항에 있어서, 이기능성 (2E,6E)-파르네실 디포스페이트 신타제/디메틸알릴트랜스트랜스퍼라제는 에스케리치아 콜라이 (*Escherichia coli*) 유래 이기능성 (2E,6E)-파르네실 디포스페이트 신타제/디메틸알릴트랜스트랜스퍼라제와 적어도 70% 아미노산 동일성을 가지고, 아미노산 치환 S80F를 포함하는 것인 조작된 미생물 세포.

청구항 30

제7항, 또는 제15항 내지 제27항 중 어느 한 항에 있어서, HMG-CoA 리덕타제는 에스. 세레비지아 HMG-CoA 리덕타제의 변이체인 조작된 미생물 세포.

청구항 31

제1항 내지 제30항 중 어느 한 항에 있어서, 배양했을 때, 조작된 미생물 세포는 (6E)-8-히드록시제라니올을 100 μ g/배양 배지 L 초과로 생산하는 것인 조작된 미생물 세포.

청구항 32

제1항 내지 제31항 중 어느 한 항에 따른 조작된 미생물 세포의 배양물.

청구항 33

제32항에 있어서, 기질은 우레아, 암모늄 염, 암모니아 및 이의 임의 조합으로 이루어진 군으로부터 선택되는 질소원, 및 탄소원을 포함하는 것인 배양물.

청구항 34

제32항 또는 제33항에 있어서, 조작된 미생물 세포는 배양물이 10-500의 600 nm에서의 광학 밀도를 갖게 하는 농도로 존재하는 것인 배양물.

청구항 35

제32항 내지 제34항 중 어느 한 항에 있어서, 배양물은 (6E)-8-히드록시제라니올을 포함하는 것인 배양물.

청구항 36

제32항 내지 제35항 중 어느 한 항에 있어서, 배양물은 (6E)-8-히드록시제라니올을 100 μ g/배양 배지 L 초과 수준으로 포함하는 것인 배양물.

청구항 37

제1항 내지 제31항 중 어느 한 항에 따른 조작된 미생물 세포를 배양하는 방법으로서, 방법은 (6E)-8-히드록시제라니올 생산에 적합한 조건 하에서 세포를 배양하는 단계를 포함하는 것인 배양 방법.

청구항 38

제37항에 있어서, 방법은 1-100 g/L 범위의 초기 글루코스 수준에 이어서, 제어된 당 공급이 후속되는 유가 배양을 포함하는 것인 배양 방법.

청구항 39

제37항 또는 제38항에 있어서, 발효 기질은 우레아, 암모늄 염, 암모니아, 및 이의 임의 조합으로 이루어진 군으로부터 선택되는 질소원, 및 글루코스를 포함하는 것인 배양 방법.

청구항 40

제37항 내지 제39항 중 어느 한 항에 있어서, 배양물은 배양 단계 동안 pH-제어되는 것인 배양 방법.

청구항 41

제37항 내지 제40항 중 어느 한 항에 있어서, 배양물은 배양 단계 동안 통기되는 것인 배양 방법.

청구항 42

제37항 내지 제41항 중 어느 한 항에 있어서, 조작된 미생물 세포는 (6E)-8-히드록시제라니올을 100 μ g/배양 배지 L 초과 수준으로 생산하는 것인 배양 방법.

청구항 43

제37항 내지 제42항 중 어느 한 항에 있어서, 방법은 배양물로부터 (6E)-8-히드록시제라니올을 회수하는 단계를 추가로 포함하는 것인 배양 방법.

청구항 44

(6E)-8-히드록시제라니올을 생산하도록 조작된 미생물 세포를 사용하여 (6E)-8-히드록시제라니올을 제조하기 위한 방법으로서, 방법은

(a) 미생물 세포에서 비천연 제라닐 디포스페이트 디포스파타제 (제라니올 신타제)를 발현시키는 단계;

(b) 미생물 세포에서 비천연 제라니올-8-히드록실라제를 발현시키는 단계;

(c) 미생물 세포가 (6E)-8-히드록시제라니올을 생산할 수 있게 하는 조건 하의 적합한 배양 배지에서 미생물 세포를 배양시키는 단계로서, (6E)-8-히드록시제라니올이 배양 배지로 방출되는 것인 단계; 및

(d) (6E)-8-히드록시제라니올을 배양 배지로부터 분리시키는 단계

를 포함하는 것인 제조 방법.

발명의 설명

기술 분야

[0001] 관련 출원에 관한 교차 참조

[0002] 본 출원은 2017년 12월 7일 출원된 미국 가출원 제62/596,013호의 우선권을 청구하고, 이의 전문을 참조로 본 명세서에 편입시킨다.

[0003] 정부 지원 연구 및 개발 하에 수행된 본 발명의 권리에 관한 진술

[0004] 본 발명은 DARPA가 수여하는 협약 번호 HR0011-15-9-0014 하의 정부 지원으로 수행되었다. 정부는 본 발명의 일정 권리를 갖는다.

[0005] 서열 목록의 참조에 의한 편입

[0006] 본 출원은 ASCII 형식으로 전자 제출된 서열 목록을 포함하고 그 전문을 참조로 본 명세서에 편입시킨다. 2018년 12월 5일 생성된 이 ASCII 사본은 명칭이 2018-12-05_ZMGNP007WO_SeqList_ST25.txt 이고 크기는 327,680 바이트이다.

[0007] 기술분야

[0008] 본 개시는 일반적으로 발효에 의한 (6E)-8-히드록시제라니올의 제조를 위한 조작된 미생물 분야에 관한 것이다.

배경 기술

[0009] (6E)-8-히드록시제라니올 (8-히드록시제라니올)은 자연계에 존재하는 것으로 알려진 비환식 모노테르펜이다. 화학 합성에 의한 테르펜의 제조 방법은 공지되어 있다 (미국 특허 제4,107,219호). (6E)-8-히드록시제라니올은 핵심 대사산물 전구체 아세틸-CoA를 기반으로, 메발로네이트 생합성 경로로부터 유래된다 (도 1). 메발로네이트 생합성 경로에서 HMG-CoA 리덕타제는 피드백 억제를 겪는다 [1, 2]. 8-히드록시제라니올은 모노테르펜 인돌 알칼로이드 [3] 및 모노테르펜 글리코시드에 대한 전구체이다 (JP2013158298A [4]; 미국 특허 제9,518,282호). 테르펜은 신규한 폴리에스테르 물질을 제조하는데 사용되어 왔다 [5]. 열-불합성 테르펜 중합체 필름이 제조되었다 (미국 특허 제3,278,646호). 수첨 테르펜은 중합체 블렌드에서 사용되었다 (미국 특허 제3,361,849호). 테르펜 수지 및 테르펜-페놀 수지가 제조되었고 자동차 산업용 코팅 보호 필름으로서 사용되었다 (미국 특허 제5,643,676호). 테르펜은 포장 필름 재료로서 사용을 위해 고수분 장벽 성질을 갖는 연신 폴리프로필렌 필름에 혼입되었다 (미국 특허 제5,500,282호).

발명의 내용

[0010] 본 개시는 하기를 포함하여, 조작된 미생물 세포, 미생물 세포의 배양물, 및 (6E)-8-히드록시제라니올의 제조 방법을 제공한다:

[0011] 실시형태 1: 조작된 미생물 세포로서, 조작된 미생물 세포는 (a) 비천연 제라닐 디포스페이트 디포스파타제 (제라니올 신타제); 및 (b) 비천연 제라니올-8-히드록실라제를 발현하고; 조작된 미생물 세포는 (6E)-8-히드록시제라니올을 생산한다.

[0012] 실시형태 2: 실시형태 1의 조작된 미생물 세포로서, 조작된 미생물 세포는 하나 이상의 상류 (6E)-8-히드록시제라니올 경로 효소(들) 또는 상류 경로 활성의 조절인자의 증가된 활성을 포함하고, 상기 증가된 활성은 대조군 세포에 비해서 증가되는 것이다.

[0013] 실시형태 3: 실시형태 2의 조작된 미생물 세포로서, 하나 이상의 상류 (6E)-8-히드록시제라니올 경로 효소(들)는 ATP-시트레이트 신타제, 아세틸-CoA 신씨타제, 티올라제, 히드록시메틸글루타릴 조효소 A 신타제 (HMG-CoA

신타제), 히드록시메틸글루타릴 조효소 A 리덕타제 (HMG-CoA 리덕타제), 메발로네이트 키나제, 포스포메발로네이트 키나제, 디포스포메발로네이트 디카복실라제, 이소펜테닐-디포스페이트 델타-이소머라제, 및 제라닐 디포스페이트 신타제로 이루어진 군으로부터 선택된다.

- [0014] 실시형태 4: 실시형태 3의 조작된 미생물 세포로서, 하나 이상의 상류 (6E)-8-히드록시제라니올 경로 효소(들)는 이소펜테닐-디포스페이트 델타-이소머라제를 포함한다.
- [0015] 실시형태 5: 실시형태 1-4 중 어느 하나의 조작된 미생물 세포로서, 조작된 미생물 세포는 하나 이상의 (6E)-8-히드록시제라니올 경로 전구체를 소모하는 하나 이상의 효소(들)의 감소된 활성을 포함하고, 상기 감소된 활성은 대조군 세포에 비해서 감소되는 것이다.
- [0016] 실시형태 6: 실시형태 5의 조작된 미생물 세포로서, 하나 이상의 (6E)-8-히드록시제라니올 경로 전구체를 소모하는 하나 이상의 효소(들)는 이기능성 (2E,6E)-파르네실 디포스페이트 신타제/디메틸알릴트랜스트랜스퍼라제 및/또는 제라닐 파이로포스페이트 신타제를 포함한다.
- [0017] 실시형태 7: 실시형태 1-6 중 어느 하나의 조작된 미생물 세포로서, 조작된 미생물 세포는 피드백-탈조절 HMG-CoA 리덕타제를 추가로 발현한다.
- [0018] 실시형태 8: 실시형태 1-7 중 어느 하나의 조작된 미생물 세포로서, 조작된 미생물 세포는 대조군 세포에 비해서, 더 높은 속도의 아세틸-CoA 합성 및/또는 더 낮은 속도의 아세틸-CoA 분해에 기인하여, 아세틸-CoA의 증가된 이용률을 포함한다.
- [0019] 실시형태 9: 조작된 미생물 세포로서, 조작된 미생물 세포는 (a) 비천연 천연 제라닐 디포스페이트 디포스파타제 (제라니올 신타제)를 발현시키기 위한 수단; 및 (b) 비천연 제라니올-8-히드록실라제를 발현시키기 위한 수단을 포함하고, 조작된 미생물 세포는 (6E)-8-히드록시제라니올을 생산한다.
- [0020] 실시형태 10: 실시예 9의 조작된 미생물 세포로서, 조작된 미생물 세포는 하나 이상의 상류 (6E)-8-히드록시제라니올 경로 효소(들) 또는 상류 경로 활성의 조절인자의 활성을 증가시키기 위한 수단을 포함한다.
- [0021] 실시형태 11: 실시형태 10의 조작된 미생물 세포로서, 하나 이상의 상류 (6E)-8-히드록시제라니올 경로 효소(들)는 ATP-시트레이트 신타제, 아세틸-CoA 신췌타제, 티올라제, 히드록시메틸글루타릴 조효소 A 신타제 (HMG-CoA 신타제), 히드록시메틸글루타릴 조효소 A 리덕타제 (HMG-CoA 리덕타제), 메발로네이트 키나제, 포스포메발로네이트 키나제, 디포스포메발로네이트 디카복실라제, 이소펜테닐-디포스페이트 델타-이소머라제, 및 제라닐 디포스페이트 신타제로 이루어진 군으로부터 선택된다.
- [0022] 실시형태 12: 실시형태 11의 조작된 미생물 세포로서, 하나 이상의 상류 (6E)-8-히드록시제라니올 경로 효소(들)는 이소펜테닐-디포스페이트 델타-이소머라제를 포함한다.
- [0023] 실시형태 13: 실시형태 9-12 중 어느 하나의 조작된 미생물 세포로서, 조작된 미생물 세포는 하나 이상의 (6E)-8-히드록시제라니올 경로 전구체를 소모하는 하나 이상의 효소(들)의 활성을 감소시키기 위한 수단을 포함하고, 상기 감소된 활성은 대조군 세포에 비해 감소되는 것이다.
- [0024] 실시형태 14: 실시형태 13의 조작된 미생물 세포로서, 하나 이상의 (6E)-8-히드록시제라니올 경로 전구체를 소모하는 하나 이상의 효소(들)는 이기능성 (2E,6E)-파르네실 디포스페이트 신타제/디메틸알릴트랜스트랜스퍼라제 및/또는 제라닐 파이로포스페이트 신타제를 포함한다.
- [0025] 실시형태 15: 실시형태 9-14 중 어느 하나의 조작된 미생물 세포로서, 조작된 미생물 세포는 피드백-탈조절 HMG-CoA 리덕타제를 발현시키기 위한 수단을 추가로 포함한다.
- [0026] 실시형태 16: 실시형태 9-15 중 어느 하나의 조작된 미생물 세포로서, 조작된 미생물 세포는 대조군 세포에 비해서, 더 높은 속도의 아세틸-CoA 합성 및/또는 더 낮은 속도의 아세틸-CoA 분해에 기인하여 아세틸-CoA의 이용률을 증가시키기 위한 수단을 포함한다.
- [0027] 실시형태 17: 실시형태 1-16 중 어느 하나의 조작된 미생물 세포로서, 조작된 미생물 세포는 진균 세포를 포함한다.
- [0028] 실시형태 18: 실시형태 17의 조작된 미생물 세포로서, 조작된 미생물 세포는 효모 세포를 포함한다.
- [0029] 실시형태 19: 실시형태 18의 조작된 미생물 세포로서, 효모 세포는 사카로마이세스 (*Saccharomyces*) 속의 세포이다.

- [0030] 실시형태 20: 실시형태 19의 조작된 미생물 세포로서, 효모 세포는 세레비지아 (*cerevisiae*) 종의 세포이다.
- [0031] 실시형태 21: 실시형태 18의 조작된 미생물 세포로서, 효모 세포는 야로위아 (*Yarrowia*) 속의 세포이다.
- [0032] 실시형태 22: 실시형태 21의 조작된 미생물 세포로서, 효모 세포는 리폴리티카 (*lipolytica*) 종의 세포이다.
- [0033] 실시형태 23: 실시형태 1-22 중 어느 하나의 조작된 미생물 세포로서, 비천연 제라니올 신타제는 페릴라 세토이엔시스 (*Perilla setoyensis*) 유래 제라니올 신타제와 적어도 70% 아미노산 서열 동일성을 갖는 제라니올 신타제를 포함한다.
- [0034] 실시형태 24: 실시형태 1-22 중 어느 하나의 조작된 미생물 세포로서, 비천연 제라니올 신타제는 비티스 비니페라 (*Vitis vinifera*) 유래 제라니올 신타제와 적어도 70% 아미노산 서열 동일성을 갖는 제라니올 신타제를 포함한다.
- [0035] 실시형태 25: 실시형태 1-23 중 어느 하나의 조작된 미생물 세포로서, 비천연 제라니올-8-히드록실라제는 파세올러스 안굴라리스 (*Phaseolus angularis*) 유래 제라니올-8-히드록실라제와 적어도 70% 아미노산 서열 동일성을 갖는 제라니올-8-히드록실라제를 포함한다.
- [0036] 실시형태 26: 실시형태 4 또는 12-25 중 어느 하나의 조작된 미생물 세포로서, 이소펜테닐-디포스페이트 델타-이소머라제의 증가된 활성은 이소펜테닐-디포스페이트 델타-이소머라제를 이중으로 발현시켜 획득된다.
- [0037] 실시형태 27: 실시형태 26의 조작된 미생물 세포로서, 이중성 이소펜테닐-디포스페이트 델타-이소머라제는 사카로마이세스 세레비지아 (*Saccharomyces cerevisiae*) 유래 이소펜테닐-디포스페이트 델타-이소머라제와 적어도 70% 아미노산 서열 동일성을 갖는 이소펜테닐-디포스페이트 델타-이소머라제를 포함한다.
- [0038] 실시형태 28: 실시형태 6, 및 14-27 중 어느 하나의 조작된 미생물 세포로서, 하나 이상의 (6E)-8-히드록시제라니올 경로 전구체를 소모하는 하나 이상의 효소(들)는 이기능성 (2E,6E)-파르네실 디포스페이트 신타제/디메틸알릴트랜스트랜스퍼라제를 포함한다.
- [0039] 실시형태 29: 실시형태 28의 조작된 미생물 세포로서, 이기능성 (2E,6E)-파르네실 디포스페이트 신타제/디메틸알릴트랜스트랜스퍼라제는 에스케리치아 콜라이 (*Escherichia coli*) 유래 이기능성 (2E,6E)-파르네실 디포스페이트 신타제/디메틸알릴트랜스트랜스퍼라제와 적어도 70% 아미노산 동일성을 가지고, 아미노산 치환 S80F를 포함한다.
- [0040] 실시형태 30: 실시형태 7, 또는 15-27 중 어느 하나의 조작된 미생물 세포로서, HMG-CoA 리덕타제는 에스. 세레비지아 (*S. cerevisiae*) HMG-CoA 리덕타제의 변이체이다.
- [0041] 실시형태 31: 실시형태 1-30 중 어느 하나의 조작된 미생물 세포로서, 배양했을 때, 조작된 미생물 세포는 (6E)-8-히드록시제라니올을 100 μ g/배양 배지 L 를 초과하는 수준으로 생산한다.
- [0042] 실시형태 32: 실시형태 1-31 중 어느 하나에 따른 조작된 미생물 세포의 배양물.
- [0043] 실시형태 33: 실시형태 32의 배양물로서, 기질은 우레아, 암모늄 염, 암모니아, 및 이의 임의 조합으로 이루어진 군으로부터 선택되는 질소원, 및 탄소원을 포함한다.
- [0044] 실시형태 34: 실시형태 32-33 중 어느 하나의 배양물로서, 조작된 미생물 세포는 배양물이 10-500의 600 nm에서의 광학 밀도를 갖는 농도로 존재한다.
- [0045] 실시형태 35: 실시형태 32-34 중 어느 하나의 배양물로서, 배양물은 (6E)-8-히드록시제라니올을 포함한다.
- [0046] 실시형태 36: 실시형태 32-35 중 어느 하나의 배양물로서, 배양물은 (6E)-8-히드록시제라니올을 100 μ g/배양 배지 L를 초과하는 수준으로 포함한다.
- [0047] 실시형태 37: 실시형태 1-31 중 어느 하나에 따른 조작된 미생물 세포를 배양하는 방법으로서, 방법은 (6E)-8-히드록시제라니올을 생산하는데 적합한 조건 하에서 세포를 배양하는 단계를 포함한다.
- [0048] 실시형태 38: 실시형태 37의 방법으로서, 방법은 1-100 g/L 범위의 초기 글루코스 수준에 이어서, 제어된 당 공급이 후속되는, 유가 배양을 포함한다.
- [0049] 실시형태 39: 실시형태 37 또는 실시형태 38의 방법에서, 발효 기질은 우레아, 암모늄 염, 암모니아, 및 이의 임의 조합으로 이루어진 군으로부터 선택되는 질소원, 및 글루코스를 포함한다.

- [0050] 실시형태 40: 실시형태 37-39 중 어느 하나의 방법에서, 배양물은 배양 단계 동안 pH-제어된다.
- [0051] 실시형태 41: 실시형태 37-40 중 어느 하나의 방법에서, 배양물은 배양 단계 동안 통기된다.
- [0052] 실시형태 42: 실시형태 37-41 중 어느 하나의 방법에서, 조작된 미생물 세포는 (6E)-8-히드록시제라니올을 100 μg /배양 배지 L를 초과하는 수준으로 생산한다.
- [0053] 실시형태 43: 실시형태 37-42 중 어느 하나의 방법에서, 방법은 (6E)-8-히드록시제라니올을 배양물로부터 회수하는 단계를 추가로 포함한다.
- [0054] 실시형태 44: (6E)-8-히드록시제라니올을 생산하도록 조작된 미생물 세포를 사용하여 (6E)-8-히드록시제라니올을 제조하는 방법으로서, 방법은 (a) 비천연 제라닐 디포스페이트 디포스파타제 (제라니올 신타제)를 미생물 세포를 발현시키는 단계; (b) 비천연 제라니올-8-히드록실라제를 미생물 세포에서 발현시키는 단계; (c) 미생물 세포를 미생물 세포가 (6E)-8-히드록시제라니올을 생산할 수 있게 하는 조건 하의 적합한 배양 배지에서 배양시키는 단계로서, (6E)-8-히드록시제라니올이 배양 배지로 방출되는 것인 단계; 및 (d) (6E)-8-히드록시제라니올을 배양 배지로부터 분리하는 단계를 포함한다.

도면의 간단한 설명

- [0055] 도 1: 발효에 의해 (6E)-8-히드록시제라니올을 생산하기 위한 경로.
- 도 2: 1차 라운드의 조작된 숙주 사카로마이세스 세레비지아에 의한 발효 후 세포의 액체배지에서 측정된 (6E)-8-히드록시제라니올 역가 (또한 실시예 1, 표 1 참조).
- 도 3: 2차 라운드의 조작된 숙주 에스. 세레비지아에 의한 발효 후 세포의 액체배지에서 측정된 (6E)-8-히드록시제라니올 역가 (또한 실시예 1, 표 2 참조).
- 도 4: 1차 라운드의 조작된 숙주 야로위아 리폴리티카에 의한 발효 후 세포의 액체배지에서 측정된 (6E)-8-히드록시제라니올 역가 (또한 실시예 2, 표 4 참조).
- 도 5: 숙주 평가를 위해 에스. 세레비지아에서 발효 후 세포의 액체배지에서 측정된 (6E)-8-히드록시제라니올 역가 (또한 실시예 3, 표 5 참조).
- 도 6: 4차 (개선) 라운드 균주에 대한 에스. 세레비지아에서 발효 후 세포의 액체배지에서 측정된 (6E)-8-히드록시제라니올 역가 (또한 실시예 1, 표 3 참조).
- 도 7: 에스. 세레비지아 및 야로위아 리폴리티카로 프로모터-유전자-종결인자의 통합.
- 도 8: 에스. 세레비지아 및 와이. 리폴리티카에서 프로모터 치환.
- 도 9: 에스. 세레비지아 및 와이. 리폴리티카에서 표적화 유전자 결실.

발명을 실시하기 위한 구체적인 내용

- [0056] 생물학적 발효에 의한 (6E)-8-히드록시제라니올과 같은 디-알콜의 생산은 단량체를 오래된 것뿐만 아니라, 새롭게 확인된 재료 응용분야에 경제적으로 접근가능하게 만들 수 있다. 디-알콜 함유 중합체는 신규 재료 응용분야에 매력적인 성질을 갖는다.
- [0057] 우리는 산업적 숙주 유기체에서 물질대사 경로가 (6E)-8-히드록시제라니올을 생산할 수 있게 하는 효소를 동정하기 위해 물질대사 검색 [1]을 수행하였다 (표 1 참조). (6E)-8-히드록시제라니올의 생산을 조작하기 위해서 산업 미생물은 DNA 서열을 조작하기 위한 방법 및 유전자 조작 도구를 요구하였다 (도 7-9 참조). 다음으로, 미생물 물질대사는 단일-유전자 및 단일-유전자 변형을 사용한 반복적인 고속 대량 (HTP) 균주 조작을 통해서, 모든 생화학적 반응 또는 물질대사 조절 방식이 특징규명되지 않은 산업 숙주를 포함하여, (6E)-8-히드록시제라니올을 생산하도록 조직적으로 재조작되었다 (예를 들어, HTP 균주 조작 방법의 경우, 공동 소유의 공개류 중인 미국 공개 특허 출원 제0170159045호를 참조하고, 또한, 티라민 생산을 위한 미생물의 조작을 예시하는 공동 소유의 공개류 중인 미국 특허 출원 제62/455,428호 참조).
- [0058] 상기 언급된 바와 같이, (6E)-8-히드록시제라니올은 테르페노이드 경로로부터 물질대사적으로 생산되는 모노테르펜이다. 미생물에는 2종의 테르페노이드 생합성 경로: 메발로네이트 경로 및 비-메발로네이트 경로가 존재한다. 사카로마이세스 세레비지아 및 야로위아 리폴리티카 둘 모두는 테르펜의 생산을 위해 메발로네이트

경로를 사용한다 [2].

- [0059] 본 개시는 (6E)-8-히드록시제라니올의 발효 생산을 위한 미생물 세포의 조작을 기술하고 신규의 조작된 미생물 세포 및 배양물을 비롯하여, 관련된 (6E)-8-히드록시제라니올 제조 방법을 제공한다.
- [0060] **정의**
- [0061] 청구항 및 명세서에서 사용되는 용어는 달리 명시하지 않으면 하기 기재된 대로 정의된다.
- [0062] 용어 "발효"는 임의의 화학적 전환 단계를 필요로 하지 않고, 하나 이상의 생물학적 전환 단계를 통해서, 미생물 세포가 하나 이상의 기질(들)을 바람직한 생성물 (예컨대 (6E)-8-히드록시제라니올)로 전환시키는 과정을 의미하고자 본 명세서에서 사용된다.
- [0063] 용어 "조작된"은 천연 발생 세포와 조작된 세포를 구별하는 사람에 의해 도입된 적어도 하나의 표적화된 유전자 변형을 세포가 함유한다는 것을 의미하고자, 세포를 참조하여, 본 명세서에서 사용된다.
- [0064] 용어 "천연"은 특정한 세포에서 천연적으로 존재하는, 폴리뉴클레오티드 또는 폴리펩티드와 같은 세포 성분을 의미하고자 본 명세서에서 사용된다. 천연 폴리뉴클레오티드 또는 폴리펩티드는 세포에 내생적이다.
- [0065] 폴리뉴클레오티드 또는 폴리펩티드를 참조하여 사용될 때, 용어 "비천연"은 특정 세포에서 천연적으로 존재하지 않은 폴리뉴클레오티드 또는 폴리펩티드를 의미한다.
- [0066] 유전자가 발현되는 상황을 참조하여 사용될 때, 용어 "비천연"은 천연적으로 발현되는 게놈 및 세포 상황 이외의 임의 상황에서 발현되는 유전자를 의미한다. 비천연 방식으로 발현되는 유전자는 숙주 세포에서 상응하는 유전자와 동일한 뉴클레오티드 서열을 가질 수 있지만, 천연 유전자의 유전자좌와 상이한 게놈 내 통합점 또는 벡터로부터 발현될 수 있다.
- [0067] 용어 "이종성"은 숙주 세포로 도입되는 폴리뉴클레오티드 또는 폴리펩티드를 설명하고자 본 명세서에서 사용된다. 이 용어는 숙주 세포와 상이한 유기체, 종 또는 군주로부터 유래되는, 각각 폴리뉴클레오티드 또는 폴리펩티드를 포괄한다. 이러한 경우에, 이종성 폴리뉴클레오티드 또는 폴리펩티드는 동일한 숙주 세포에 존재하는 임의 서열(들)과 상이한 서열을 갖는다. 그러나, 이 용어는 또한 숙주 세포에 존재하는 서열과 동일한 서열을 갖는 폴리뉴클레오티드 또는 폴리펩티드를 포괄하고, 여기서 폴리뉴클레오티드 또는 폴리펩티드는 천연 서열과 상이한 상황에서 존재한다 (예를 들어, 이종성 폴리뉴클레오티드는 상이한 프로모터에 연결될 수 있고 천연 서열과 상이한 게놈 위치에 삽입될 수 있음). 따라서 "이종성 발현"은 숙주 세포에 비천연인 서열의 발현을 비롯하여, 비천연 상황에서 숙주 세포에 천연인 서열의 발현을 포괄한다.
- [0068] 폴리뉴클레오티드 또는 폴리펩티드를 참조하여 사용시, 용어 "야생형"은 분자의 공급원과 무관하게, 뉴클레오티드 서열을 갖는 임의의 폴리뉴클레오티드, 또는 아미노산을 갖는 폴리펩티드, 천연 발생 유기체 유래 폴리뉴클레오티드 또는 폴리펩티드에 존재하는 서열을 의미하고; 다시 말해서, 용어 "야생형"은 분자가 천연 공급원으로 정제되거나; 재조합적으로 발현 후 정제되거나; 또는 합성되거나 여부와 무관한, 서열 특징을 의미한다. 용어 "야생형"은 또한 천연 발생 세포를 의미하고자 사용된다.
- [0069] "대조군 세포"는 조작된 세포와 동일한 속 및 종인 것을 포함하여, 시험되는 조작된 세포와 동일하지만, 조작된 세포에서 시험되는 특이적 유전자 변이(들)는 결여된 세포이다.
- [0070] 효소는 본 명세서에서 그들이 촉매하는 반응으로 식별되고, 달리 표시하지 않으면, 식별되는 반응을 촉매할 수 있는 임의의 폴리펩티드를 의미한다. 달리 표시하지 않으면, 효소는 임의의 유기체로부터 유래될 수 있고 천연 또는 돌연변이 아미노산 서열을 가질 수 있다. 충분히 공지된 바와 같이, 효소는 때때로 그들이 유래되는 유기체 공급원에 따라서, 복수의 기능 및/또는 복수의 명칭을 가질 수 있다. 본 명세서에서 사용되는 효소 명칭은 하나 이상의 추가적인 기능 또는 상이한 명칭을 가질 수 있는 효소를 포함하여, 오솔로그를 포괄한다.
- [0071] 용어 "피드백-탈조절"은 특정 세포에서 정상적으로 효소 경로의 하류 생성물에 의해 음성적으로 조절 (즉, 피드백-억제)되는 효소를 참조하여 본 명세서에서 사용된다. 이러한 문맥에서, "피드백-탈조절" 효소는 세포에 천연인 천연 효소에 비해 피드백-억제에 덜 민감한 효소의 형태이다. 피드백-탈조절 효소는 천연 효소에 하나 이상의 돌연변이를 도입시켜 생성시킬 수 있다. 대안적으로, 피드백-탈조절 효소는 단순히, 특정한 미생물 세포에 도입되었을 때, 천연의, 천연 효소만큼 피드백-억제에 민감하지 않은 이종성, 천연 효소일 수 있다. 일부 실시형태에서, 피드백-탈조절 효소는 미생물 세포에서 피드백-억제를 보이지 않는다.

- [0072] 용어 "(6E)-8-히드록시제라니올"은 (2E,6E)-2,6-디메틸-2,6-옥타디엔-1,8-디올 (CAS#26488-97-1)을 의미한다.
- [0073] 둘 이상의 아미노산 또는 뉴클레오타이드 서열의 문맥에서 용어 "서열 동일성"은 서열 비교 알고리즘 또는 육안 검사를 사용하여 측정시, 최대 상응도로 비교 및 정렬했을 때, 동일한 아미노산 잔기 또는 뉴클레오타이드의 명시된 백분율을 갖거나 또는 동일한 둘 이상의 서열을 의미한다.
- [0074] 뉴클레오타이드 또는 아미노산 서열 동일성의 백분율을 결정하기 위한 서열 비교의 경우에, 전형적으로 한 서열은 "시험" 서열이 비교되는 "기준 서열"로서 작용한다. 서열 비교 알고리즘을 사용하는 경우에, 시험 및 기준 서열을 컴퓨터에 입력하고, 하위서열 좌표를 필요하다면 지정하며, 서열 알고리즘 프로그램 매개변수를 지정한다. 그 다음으로 서열 비교 알고리즘이 지정된 프로그램 매개변수를 기반으로, 기준 서열에 대한 시험 서열의 서열 동일성 백분율을 계산한다. 비교를 위한 서열의 알고리즘은 디폴트 매개변수로 설정된 BLAST를 사용하여 수행될 수 있다.
- [0075] 본 명세서에서 사용되는 용어 "역가"는 미생물 세포의 배양물을 배양 부피로 나눈 생성된 생성물 (예를 들어, (6E)-8-히드록시제라니올)의 질량을 의미한다.
- [0076] 세포 배양물로부터 (6E)-8-히드록시제라니올의 회수를 참조하여 본 명세서에서 사용시, "회수하는"은 세포 배양 배지의 적어도 하나의 다른 성분으로부터 (6E)-8-히드록시제라니올의 분리를 의미한다.
- [0077] **(6E)-8-히드록시제라니올 생산을 위해 조작된 미생물**
- [0078] **(6E)-8-히드록시제라니올 생합성 경로**
- [0079] (6E)-8-히드록시제라니올은 핵심 대사산물 전구체 아세틸-CoA를 기반으로, 메발로네이트 생합성 경로로부터 유래된다. 이 경로는 도 1에 예시되어 있다. 메발로네이트 생합성 경로에서 HMG-CoA 리덕타제는 피드백 억제를 겪는다. 많은 미생물, 예컨대 사카로마이세스 세레비지아는 이 경로의 최종 2개 단계를 촉매하는 효소, 즉 제라닐 디포스페이트 디포스파타제 (제라니올 신타제) 및 제라니올-8-히드록실라제를 촉매하는 효소가 결여되어 있다. 이러한 미생물 숙주에서 (6E)-8-히드록시제라니올의 생산은 적어도 하나의 이중성 제라니올 신타제 효소 및 적어도 하나의 이중성 제라니올-8-히드록실라제의 첨가를 요구한다.
- [0080] **미생물 (6E)-8-히드록시제라니올 생산을 위한 조작**
- [0081] 조작되는 미생물 세포에서 활성인 임의의 제라니올 신타제 및 제라니올-8-히드록실라제는 전형적으로 표준 유전자 조작 기술을 사용하여 효소를 코딩하는 유전자를 도입시키고 발현시켜서 세포로 도입될 수 있다. 적합한 제라니올 신타제 및 제라니올-8-히드록실라제는 식물, 고세균, 진균, 그람 양성 박테리아, 및 그람 음성 박테리아 공급원을 포함한, 임의의 공급원으로부터 유래될 수 있다. 예시적인 공급원은 제한없이, 페릴라 세토이엔시스 (*Perilla setoyensis*), 파세올러스 안굴라리스 (*Phaseolus angularis*), 비티스 비니페라 (*Vitis vinifera*) (포도), 스위티아 무소티 (*Swertia mussotii*) (용담), 포폴러스 트리코카르파 (*Populus trichocarpa*) (서부 발삼 포플러) (포폴러스 발사미페라 (*Populus balsamifera*) subsp. 트리코카르파 (*Trichocarpa*)), 파파베르 쉼니페럼 (*Papaver somniferum*), 페트로셀리눔 크리스툼 (*Petroselinum crispum*), 오리자 사티바 (*Oryza sativa*), 메타노스파에룰라 팔루스트리스 (*Methanosphaerula palustris*), 메타노칼도코커스 잔나스키 (*Methanocaldococcus jannaschii*), 자이코사카로마이세스 바일리 (*Zygosaccharomyces bailii*), 페니실리움 마르네페이 (*Penicillium marneffeii*), 탈라로마이세스 스티피타터스 (*Talaromyces stipitatus*), 트리코피톤 에퀴눔 (*Trichophyton equinum*), 프로피오니박테리움 (*Propionibacterium*) sp. 경구균, 엔테로코커스 패시움 (*Enterococcus faecium*), 스트렙토마이세스 히그로스코피커스 (*Streptomyces hygroscopicus*), 스트렙토마이세스 스비세우스 (*Streptomyces sviveus*), 모데스토박터 마리너스 (*Modestobacter marinus*), 슈도모나스 푸티다 (*Pseudomonas putida*), 시노리조비움 프레디 (*Sinorhizobium fredii*), 카타탄터스 로세아우스 (*Cathanthus roseaus*), 제아 마이스 (*Zea mays*), 카타란터스 로세우스 (*Catharanthus roseus*) (마다가스카르 페리윙클 (*Madagascar periwinkle*)) (일일초 (*Vinca rosea*)), 페릴라 프루테센스 var. 크리스파 (*Perilla frutescens* var. *crispa*), 페릴라 프루테센스 var. 힐텔라 (*Perilla frutescens* var. *hirtella*), 신나모뎀 테누이필 (*Cinnamomum tenuipile*) (알세오다프테 몰리스 (*Alseodaphne mollis*)), 오시엄 바실리쿰 (*Ocimum basilicum*) (스위트 바질), 페릴라 시트리오도라 (*Perilla citriodora*), 올레아 유로파에아 (*Olea europaea*) (일반 올리브), 필라 둘시스 (*Phyla dulcis*) (아즈텍 스위트 허브) (리피아 둘시스 (*Lippia dulcis*)), 로사 루고사 (*Rosa rugosa*) (해당화), 캄프토테카 아쿠미니타 (*Camptotheca acuminata*) (행운목), 시트러스 잠비리 (*Citrus jambhiri*) (러프 레몬), 피크로리자 쿠루아 (*Picrorhiza kurrooa*), 아라비도시스 탈리아나 (*Arabidopsis thaliana*) (애기장대), 글리신 막스 (*Glycine max*) (대두) (글리신 히스피다 (*Glycine*

hispida)), 베타 불가리스 (*Beta vulgaris*) (사탕무), 몰루고 베르티실라타 (*Mollugo verticillata*) (녹색 큰 식류풀), 암보렐라 트리코포다 (*Amborella trichopoda*), 솔라넘 튜버로섬 (*Solanum tuberosum*) (감자), 글리신 소자 (*Glycine soja*) (돌콩), 반다 코에룰레아 (*Vanda coerulea*), 오리자 바르티 (*Oryza barthii*), 히페리컴 안드로사에뎀 (*Hypericum androsaemum*) (투스산 (*Tutsan*)), 솔라넘 리코퍼시컴 (*Solanum lycopersicum*) (토마토) (리코퍼시콘 에스쿨렌툼 (*Lycopersicon esculentum*)), 및 코페아 카네포라 (*Coffea canephora*) (로부스타 커피)를 포함한다.

[0082] 제라니올 신타제 및 제라니올-8-히드록실라제 유전자 각각의 하나 이상의 카피가 선택된 미생물 숙주 세포에 도입될 수 있다. 하나 초과 유전자 카피가 도입되면, 그 카피들은 동일하거나 또는 상이한 뉴클레오티드 서열을 가질 수 있다. 일부 실시형태에서, 이중성 유전자(들) 중 하나 또는 둘 모두는 강력한, 항상성 프로모터로부터 발현된다. 일부 실시형태에서, 이중성 제라니올 신타제 및/또는 이중성 제라니올-8-히드록실라제 유전자는 유도성 프로모터로부터 발현된다. 이중성 유전자는 선택된 미생물 숙주 세포에서의 발현을 증강시키기 위해 임의로는 코돈 최적화될 수 있다.

[0083] 실시예 1에서, 에스. 세레비지아는 페틸라 세토이엔시스 유래 제라니올 신타제 (UniProt ID COKWV4) (SEQ ID NO:5) 및 파세올러스 안굴라리스 유래 제라니올-8-히드록실라제 (UniProt ID C6J436) (SEQ ID NO:11)를 발현시키도록 조작되어, 유전자 조작의 1차 라운드에서 37.5 $\mu\text{g/L}$ 의 (6E)-8-히드록시제라니올 역가를 산출하였다 (하기 표 1). 이러한 역가는 이소펜테닐-디포스페이트 델타3-델타2-이소머라제 (UniProt ID P15496) (SEQ ID NO:25)의 3개 카피를 추가로 발현시킨 균주에서 2차 라운드에 122.9 $\mu\text{g/L}$ 로 증가되었다.

[0084] 실시예 2에서, 와이. 리폴리티카는 페틸라 세토이엔시스 유래 제라니올 신타제 (UniProt ID COKWV4) (SEQ ID NO:99), 파세올러스 안굴라리스 유래 제라니올-8-히드록실라제 (UniProt ID AOAOL9UT99) (SEQ ID NO:115), 및 에스 세레비지아 유래 이소펜테닐-디포스페이트 델타3-델타2-이소머라제 (UniProt ID P15496) (SEQ ID NO:126)를 발현시키도록 조작되어, 310 마이크로그램/L의 (6E)-8-히드록시제라니올 역가를 산출하였다.

[0085] 실시예 3에서, 에스. 세레비지아는 페틸라 세토이엔시스 유래 제라니올 신타제 (UniProt ID COKWV4) (SEQ ID NO:99), 파세올러스 안굴라리스 유래 제라니올-8-히드록실라제 (UniProt ID C6J436), 및 에스. 세레비지아 유래 이소펜테닐-디포스페이트 델타3-델타2-이소머라제 (UniProt ID P15496) (SEQ ID NO:126)를 발현하도록 조작되어, 217 마이크로그램/L의 역가를 산출하였다.

[0086] 상류 효소의 활성 증가

[0087] 이러한 생산을 할 수 있는 미생물 세포에서 (6E)-8-히드록시제라니올 생산을 증가시키기 위한 한가지 접근법은 (6E)-8-히드록시제라니올 생합성 경로에서 하나 이상의 상류 효소의 활성을 증가시키는 것이다. 상류 경로 효소는 공급원료로부터 모든 마지막 천연 대사산물 (즉, 에스. 세레비지아의 경우 제라닐 디포스페이트)로의 전환에 관여되는 모든 효소를 포함한다. 일정 실시형태에서, 상류 경로 효소는 특별히 (6E)-8-히드록시제라니올로 이어지는 경로에서 핵심 전구체의 제라닐 디포스페이트로의 전환에 관여하는 효소를 의미한다. 이러한 유전자는 ATP-시트레이트 신타제, 아세틸-CoA 신썬타제, 티올라제, 히드록시메틸글루타릴 조효소 A 신타제 (HMG-CoA 신타제), 히드록시메틸글루타릴 조효소 A 리덕타제 (HMG-CoA 리덕타제), 메발로네이트 키나제, 포스포메발로네이트 키나제, 디포스포메발로네이트 디카복실라제, 이소펜테닐-디포스페이트 델타-이소머라제 (체계명: 이소펜테닐-디포스페이트 델타2-델타3-이소머라제), 및 제라닐 디포스페이트 신타제를 코딩하는 것들을 포함한다. 적합한 상류 경로 유전자는 예를 들어 이중성 제라니올 신타제 또는 제라니올-8-히드록실라제 유전자에 대한 공급원으로서 상기 기술된 것들을 포함한, 임의 공급원으로부터 유래될 수 있다.

[0088] 일부 실시형태에서, 하나 이상의 상류 경로 효소의 활성은 천연 효소(들)의 발현 또는 활성을 조절하여 증가된다. 이러한 접근법의 예는 (1) (6E)-8-히드록시제라니올 경로 전구체 제라닐 디포스페이트의 수준을 증가시키기 위한 HMG-CoA 리덕타제의 과발현 및/또는 제라닐 디포스페이트 신타제의 항상적 과발현, 및/또는 (2) 아세틸-CoA의 이용률을 개선시키기 위한 ATP-시트레이트 신타제 (P53396) 및/또는 아세틸-CoA 신썬타제 (ACS, Q8ZKF6)의 과발현을 포함한다.

[0089] 천연 상류 경로 효소의 발현은 상류 경로 활성의 하나 이상의 천연 조절인자를 통해서 증가될 수 있다. 예를 들어, 이소프레노이드 경로 효소의 발현을 개선시키기 위해서, 스테롤 흡수 제어 단백질, UPC2 (UniProt ID Q12151, 사카로마이세스 세레비지아 S288c 유래, G888D 또는 G888R 함유 [이들 명칭은 아미노산에 대한 표준 1-글자 코드를 사용한 아미노산 치환을 의미하는 것으로서, 첫번째 글자는 야생형 잔기를 표시하고 마지막 글자를 치환 잔기를 표시하고, 번호는 번역된 단백질에서 아미노산 치환의 위치를 표시함]) [6, 7]의 변이체를 도입시

킬 수 있고, 임의로는 과발현시킬 수 있다. 스테롤 흡수 제어 단백질 UPC2는 스테롤 합성을 조절하고 C-말단 아미노산 치환은 이러한 전사 인자의 활성을 증가시킨다. 예를 들어, ERG8의 상류 UPC2 결합 구성요소는 다른 스테롤 생합성 경로 유전자 이외에도 ERG8의 UPC2 전사를 활성화시킬 수 있다 [7].

[0090] 대안적으로, 또는 추가로, 하나 이상의 프로모터는 예를 들어 도 5에 예시된 바와 같은 기술을 사용하여 천연 프로모터를 대체할 수 있다. 일정 실시형태에서, 대체 프로모터는 천연 프로모터보다 강력하고/하거나 항상성 프로모터이다.

[0091] 일부 실시형태에서, 하나 이상의 상류 경로 효소의 활성은 제라니올 신타제- 및 제라니올-8-히드록실라제-발현 미생물 숙주 세포로 하나 이상의 상응하는 유전자를 도입시켜 보충된다. 실시예 1은 이소펜테닐-디포스페이트 델타-이소머라제를 코딩하는 도입된 유전자와 함께, 이중성 제라니올 신타제 및 이중성 제라니올-8-히드록실라제를 발현하기 위한 미생물 숙주 세포의 성공적인 조작을 설명한다.

[0092] 도입된 상류 경로 유전자는 숙주 세포 이외의 유기체로부터 유래될 수 있거나 또는 단순히 천연 유전자의 추가 카피일 수 있다. 일부 실시형태에서, 하나 이상의 이러한 유전자는 (6E)-8-히드록시제라니올을 생산할 수 있는 미생물 숙주 세포로 도입되고 강력한 항상성 프로모터로부터 발현되고/되거나 선택된 미생물 숙주 세포에서 발현을 증강시키기 위해 임의로는 코돈 최적화될 수 있다.

[0093] 다양한 실시형태에서, 하나 이상의 상류 경로 효소의 활성을 증가시키기 위한 (6E)-8-히드록시제라니올-생산 미생물 세포의 조작은 (6E)-8-히드록시제라니올 역가를 적어도 10, 20, 30, 40, 50, 60, 70, 80, 또는 90% 또는 적어도 2배, 2.5배, 3배, 3.5배, 4배, 4.5배, 5배, 5.5배, 6배, 6.5배, 7배, 7.5배, 8배, 8.5배, 9배, 9.5배, 10배, 11배, 12배, 13배, 14배, 15배, 16배, 17배, 18배, 19배, 20배, 21배, 22배, 23배, 24배, 25배, 30배, 35배, 40배, 45배, 50배, 55배, 60배, 65배, 70배, 75배, 80배, 85배, 90배, 95배, 또는 100배까지 증가시킨다. 다양한 실시형태에서, (6E)-8-히드록시제라니올 역가의 증가는 10% 내지 100배, 2배 내지 50배, 5배 내지 40배, 10배 내지 30배의 범위이다 (본 명세서에서 범위는 그들 종결점을 포함함). 이들 증가는 상류 경로 효소 활성의 임의의 증가가 결여된 (6E)-8-히드록시제라니올-생산 미생물 세포에서 관찰되는 (6E)-8-히드록시제라니올 역가에 대해서 결정된다. 이러한 기준 세포는 (6E)-8-히드록시제라니올 생산 증가를 목적으로 하는 하나 이상의 다른 유전자 변경을 가질 수 있으며, 예를 들어 세포는 피드백-탈조절 효소를 발현할 수 있다.

[0094] 다양한 실시형태에서, 하나 이상의 상류 경로 유전자의 활성을 증가시켜 획득된 (6E)-8-히드록시제라니올 역가는 적어도 100, 200, 300, 400, 500, 600, 700, 800, 또는 900 $\mu\text{g/L}$, 또는 적어도 1, 10, 50, 75, 100, 200, 300, 400, 500, 600, 700, 800, 또는 900 mg/L 또는 적어도 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 또는 10 gm/L이다. 다양한 실시형태에서, 역가는 100 $\mu\text{g/L}$ 내지 10 gm/L, 200 $\mu\text{g/L}$ 내지 5 gm/L, 500 $\mu\text{g/L}$ 내지 4 gm/L, 1 mg/L 내지 3 gm/L, 500 mg/L 내지 2 gm/L의 범위 또는 상기 열거된 임의의 값으로 경계된 임의의 범위이다.

[0095] 피드백-탈조절 효소의 도입

[0096] (6E)-8-히드록시제라니올 생합성이 피드백 억제를 겪으므로, (6E)-8-히드록시제라니올을 생산하도록 조작된 미생물 세포에서 (6E)-8-히드록시제라니올을 증가시키기 위한 다른 접근법은 정상적으로 피드백 조절을 겪는 하나 이상의 효소의 피드백-탈조절 형태를 도입시키는 것이다. HMG-CoA 리덕타제가 그러한 효소이다. 피드백-탈조절 형태는 특정 미생물 숙주 세포에서 천연 효소에 비해서 피드백 억제에 덜 민감한 이중성, 천연 효소일 수 있다. 대안적으로, 피드백-탈조절 형태는 상응하는 천연 효소에 비해서 피드백 억제에 덜 민감하게 만드는 하나 이상의 돌연변이 또는 절단을 갖는 천연 또는 이중성 효소의 변이체일 수 있다. 후자의 예는 N-말단 절단을 갖는 변이체 HMG-CoA 리덕타제 (에스. 세레비지아 유래) (SEQ ID NO:27)를 포함한다. 숙주 세포에서 이러한 피드백-탈조절 HMG-CoA 리덕타제의 발현은 에스. 세레비지아 [3] 및 다른 유기체에서 메발로네이트 경로 유동성을 개선시키는 것으로 확인되었다 (예를 들어, 식물의 유전자 조작을 기술하는 PCT 공개 특허 출원 번호 W02001031027A1을 참조함).

[0097] 다양한 실시형태에서, 피드백-탈조절 효소를 발현시키기 위한 (6E)-8-히드록시제라니올-생산 미생물 세포의 조작은 (6E)-8-히드록시제라니올 역가를 적어도 10, 20, 30, 40, 50, 60, 70, 80, 또는 90% 또는 적어도 2배, 2.5배, 3배, 3.5배, 4배, 4.5배, 5배, 5.5배, 6배, 6.5배, 7배, 7.5배, 8배, 8.5배, 9배, 9.5배, 10배, 11배, 12배, 13배, 14배, 15배, 16배, 17배, 18배, 19배, 20배, 21배, 22배, 23배, 24배, 25배, 30배, 35배, 40배, 45배, 50배, 55배, 60배, 65배, 70배, 75배, 80배, 85배, 90배, 95배, 또는 100배까지 증가시킨다. 다양한

실시형태에서, (6E)-8-히드록시제라니올 역가의 증가는 10% 내지 100배, 2배 내지 50배, 5배 내지 40배, 10배 내지 30배 범위, 또는 상기 열거된 임의 값을 경계로하는 임의 범위이다. 이들 증가는 피드백-탈조절 효소를 발현하지 않는 (6E)-8-히드록시제라니올-생산 미생물 세포에서 관찰된 (6E)-8-히드록시제라니올 역가에 대해서 결정된다. 이러한 기준 세포는 (6E)-8-히드록시제라니올 생산을 증가시키는 목적으로 다른 유전자 변경을 (그럴 필요는 없지만) 가질 수 있고, 다시 말해서 세포는 피드백-불감성 이외의 일부 수단에 의한 상류 경로 효소의 증가된 활성을 가질 수 있다.

[0098] 다양한 실시형태에서, (6E)-8-히드록시제라니올 생합성 경로를 통해 유동성을 증가시키기 위해 피드백-탈조절 효소를 사용해 획득된 (6E)-8-히드록시제라니올 역가는 적어도 100, 200, 300, 400, 500, 600, 700, 800, 또는 900 $\mu\text{g/L}$, 또는 적어도 1, 10, 50, 75, 100, 200, 300, 400, 500, 600, 700, 800, 또는 900 mg/L 또는 적어도 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 또는 10 g/L이다. 다양한 실시형태에서, 역가는 100 $\mu\text{g/L}$ 내지 10 g/L, 200 $\mu\text{g/L}$ 내지 5 g/L, 500 $\mu\text{g/L}$ 내지 4 g/L, 1 mg/L 내지 3 g/L, 500 mg/L 내지 2 g/L의 범위 또는 상기 열거된 임의 값으로 경계되는 임의 범위이다.

[0099] 하나 이상의 천연 효소의 활성을 보충하고/하거나 하나 이상의 피드백-탈조절 효소를 도입하는 접근법을 제라니올 신타제-발현 미생물 세포에서 조합하여 보다 더 높은 (6E)-8-히드록시제라니올 생산 수준을 달성할 수 있다.

[0100] 전구체 소모 감소

[0101] 이러한 생산을 할 수 있는 미생물 세포에서 (6E)-8-히드록시제라니올 생산을 증가시키기 위한 다른 접근법은 (6E)-8-히드록시제라니올 생합성의 하나 이상의 전구체를 하나 이상의 부경로 (즉, (6E)-8-히드록시제라니올 이외의 다른 생성물로 이르게 하는 경로)로 이동하게 하는 하나 이상의 효소의 활성을 감소시키는 것이다. 일부 실시형태에서, 하나 이상의 부-경로 효소의 활성은 천연 효소(들)의 발현 또는 활성을 조절하여 감소된다.

예시적인 부-경로 효소는 이기능성 (2E,6E)-파르네실 디포스페이트 신타제/ 디메틸알릴트랜스퍼라제, 제라닐제라닐 파이로포스페이트 신타제, 및 아세틸 Co-A를 소모하는 임의의 부-경로 효소를 포함한다. 이러한 효소의 활성은 예를 들어 상응하는 유전자(들)의 천연 프로모터를 덜 활성화되거나 또는 불활성인 프로모터로 치환시키거나 또는 상응하는 유전자(들)를 결실시켜 감소시킬 수 있다. 예를 들어 에스. 세레비지아에서 각각 프로모터 치환 및 표적화된 유전자 결실의 계획에 대해 도 5 및 6을 참조한다.

[0102] 천연 이기능성 (2E,6E)-파르네실 디포스페이트 신타제/ 디메틸알릴트랜스퍼라제는 이기능성이고 제라닐 디포스페이트를 형성할 수 있고 이후에 제2 반응은 제라닐 디포스페이트를 (2E,6E)-파르네실 디포스페이트로 전환시킬 수 있다. 천연 효소는 이들 2종의 활성을 보유하고 중간체가 (6E)-8-히드록시제라니올 경로 대사산물이기 때문에, 천연 이기능성 (2E,6E)-파르네실 디포스페이트 신타제/ 디메틸알릴트랜스퍼라제의 발현을 낮추는 것이 유익하다. 그러나, S80F 아미노산 치환을 보유하는 효소의 발현은 측정가능한 분량의 모노테르페노이드를 생산한다 [4] (또한 미국 특허 제8,715,962호 참조). 예를 들어, 스테롤로의 추가 유동을 방지하기 위해서, 에스. 세레비지아에서 ERG9에 의해 코딩되는 이기능성 파르네실-디포스페이트 파르네실트랜스퍼라제 (EC 2.5.1.21) 및/또는 에스. 세레비지아에서 ERG20에 의해 코딩되는 (2E,6E)-파르네실 디포스페이트 신타제 (EC 2.5.1.10)의 발현 또는 활성은 (6E)-8-히드록시제라니올 생합성을 위한 제라닐 디포스페이트 풀을 최대화하도록 저해된다.

[0103] 에스. 세레비지아의 예시적인 실시형태에서, (1) 이기능성 (2E,6E)-파르네실 디포스페이트 신타제/디메틸알릴트랜스퍼라제 (ERG20, YJL167W)를 위한 프로모터는 (6E)-8-히드록시제라니올 경로 대사산물 제라닐 디포스페이트를 소모하는 이러한 천연 효소의 발현을 저하시키도록 에스. 세레비지아 pRnr1 프로모터로 치환될 수 있다; (2) 제라닐제라닐 파이로포스페이트 신타제 (Bts1, YPL069C)를 위한 프로모터는 역시 제라닐 디포스페이트를 소모하는 이러한 천연 효소의 발현을 저하시키도록 에스. 세레비지아 pPsp2로 치환될 수 있다; 그리고/또는 Pdc5 (YLR134W), Pdc6 (YGR087C) 및 Pdc1 (YLR044C) 중 하나 이상은 아세틸-CoA 소모를 감소시키도록 결실될 수 있다.

[0104] 다양한 실시형태에서, 하나 이상의 부경로에 의해 전구체 소모를 감소시키기 위한 (6E)-8-히드록시제라니올-생산 미생물 세포의 조작은 (6E)-8-히드록시제라니올 역가를 적어도 10, 20, 30, 40, 50, 60, 70, 80, 또는 90% 또는 적어도 2배, 2.5배, 3배, 3.5배, 4배, 4.5배, 5배, 5.5배, 6배, 6.5배, 7배, 7.5배, 8배, 8.5배, 9배, 9.5배, 10배, 11배, 12배, 13배, 14배, 15배, 16배, 17배, 18배, 19배, 20배, 21배, 22배, 23배, 24배, 25배, 30배, 35배, 40배, 45배, 50배, 55배, 60배, 65배, 70배, 75배, 80배, 85배, 90배, 95배, 또는 100배까지 증가시킨다. 다양한 실시형태에서, (6E)-8-히드록시제라니올 역가의 증가는 10% 내지 100배, 2배 내지 50배, 5배 내지 40배, 10배 내지 30배 범위, 또는 상기 열거된 임의 값이 경계가 되는 임의 범위이다. 이들 증가는

전구체 소모를 감소시키기 위한 유전자 변형을 포함하지 않는 (6E)-8-히드록시제라니올-생산 미생물 세포에서 관찰된 (6E)-8-히드록시제라니올 역가에 대해 결정된다. 이러한 기준 세포는 (그럴 필요는 없지만) (6E)-8-히드록시제라니올 생산을 증가시키기 위한 목적으로 다른 유전자 변형을 가질 수 있으며, 다시 말해 세포는 상류 경로 효소의 증가된 활성을 가질 수 있다.

[0105] 다양한 실시형태에서, 하나 이상의 부경로에 의한 전구체 소모의 감소를 통해 획득된 (6E)-8-히드록시제라니올 역가는 적어도 100, 200, 300, 400, 500, 600, 700, 800, 또는 900 $\mu\text{g/L}$, 또는 적어도 1, 10, 50, 75, 100, 200, 300, 400, 500, 600, 700, 800, 또는 900 mg/L 또는 적어도 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 또는 10 g/L 이다. 다양한 실시형태에서, 역가는 100 $\mu\text{g/L}$ 내지 10 g/L , 200 $\mu\text{g/L}$ 내지 5 gm/L , 500 $\mu\text{g/L}$ 내지 4 g/L , 1 mg/L 내지 3 g/L , 500 mg/L 내지 2 g/L 범위 또는 상기 열거된 임의 값이 경계되는 임의 범위이다.

[0106] 하나 이상의 천연 효소의 활성을 증가시키고/시키거나 하나 이상의 피드백-탈조절 효소를 도입시키고/시키거나 하나 이상의 부경로에 의해 전구체 소모를 감소시키는 접근법을 조합하여 더 높은 (6E)-8-히드록시제라니올 생산 수준을 획득할 수 있다.

[0107] 미생물 숙주 세포

[0108] 도입된 유전자를 발현시키는데 사용할 수 있는 임의의 미생물은 상기 기술된 바와 같이 (6E)-8-히드록시제라니올의 발효 생산을 위해 조작될 수 있다. 일정 실시형태에서, 미생물은 (6E)-8-히드록시제라니올의 발효 생산을 천연적으로 할 수 없는 것이다. 일부 실시형태에서, 미생물은 쉽게 배양되는 것으로서, 예컨대, 예를 들어 관심 화합물의 발효 생산에서 숙주 세포로서 유용하다고 알려진 미생물이다. 그람 양성 또는 그람 음성 박테리아를 포함한, 박테리아 세포가 상기 기술된 대로 조작될 수 있다. 예에는 *씨. 글루타미쿰* (*C. glutamicum*) 세포 이외에도, 바실러스 서브틸리스 (*Bacillus subtilis*), 비. 리체니포르미스 (*B. licheniformis*), 비. 렌터스 (*B. lentus*), 비. 브레비스 (*B. brevis*), 비. 스테아로써모필러스 (*B. stearothermophilus*), 비. 알칼로필러스 (*B. alkalophilus*), 비. 아밀로리켄파시엔스 (*B. amyloliquefaciens*), 비. 클라우시 (*B. clausii*), 비. 할로두란스 (*B. halodurans*), 비. 메가테리움 (*B. megaterium*), 비. 코아굴란스 (*B. coagulans*), 비. 실쿨란스 (*B. circulans*), 비. 라우터스 (*B. lautus*), 비. 투링지엔시스 (*B. thuringiensis*), 에스. 알버스 (*S. albus*), 에스. 리비단스 (*S. lividans*), 에스. 코엘리콜로르 (*S. coelicolor*), 에스. 그리세우스 (*S. griseus*), 슈도모나스 (*Pseudomonas*) *sp.*, 피. 알칼리체네스 (*P. alcaligenes*), 피. 시트레아 (*P. citrea*), 락토바실리스 (*Lactobacillus*) *spp.* (예컨대 엘. 락티스 (*L. lactis*), 엘. 플란타럼 (*L. plantarum*)), 엘. 그라이 (*L. grayi*), 이. 콜라이 (*E. coli*), 이. 패시움 (*E. faecium*), 이. 갈리나룸 (*E. gallinarum*), 이. 카셀리플라버스 (*E. casseliflavus*), 및/또는 이. 패칼리스 (*E. faecalis*) 세포가 포함된다.

[0109] 본 명세서에 기술된 방법에서 미생물 숙주 세포로서 사용할 수 있는 수많은 유형의 혐기성 세포가 존재한다. 일부 실시형태에서, 미생물 세포는 절대 혐기성 세포이다. 절대 혐기성균은 전형적으로 산소가 존재하는 조건이라면 전혀 잘 성장하지 못한다. 소량의 산소가 존재할 수도 있고, 다시 말해 절대 혐기성균이 낮은 수준의 산소에 대해 가지는 내성 수준의 어느 정도 수준이 존재한다는 것을 이해해야 한다. 상기 기술된 바와 같이 조작된 절대 혐기성균은 실질적으로 산소-무함유 조건 하에서 성장될 수 있으며, 존재하는 산소의 양은 혐기성균의 성장, 유지 및/또는 발효에 유해하지 않다.

[0110] 대안적으로, 본 명세서에 기술된 방법에서 사용되는 미생물 숙주 세포는 통성 혐기성 세포일 수 있다. 통성 혐기성균은 산소가 존재한다면 호기성 호흡 (예를 들어, TCA 사이클의 활용)에 의해 세포의 ATP를 발생시킬 수 있다. 그러나, 통성 혐기성균은 또한 산소 부재 하에서도 성장할 수 있다. 상기 기술된 바와 같이 조작된 통성 혐기성균은 실질적으로 산소-무함유 조건 하에서 성장할 수 있으며, 존재하는 산소의 양이 혐기성균의 성장, 유지 및/또는 발효에 유해하지 않거나, 또는 더 많은 양의 산소 존재 하에서 대안적으로 성장할 수 있다.

[0111] 일부 실시형태에서, 본 명세서에 기술된 방법에서 사용되는 미생물 숙주 세포는 사상 진균 세포이다 (예를 들어, [Berka & Barnett, Biotechnology Advances, (1989), 7(2):127-154] 참조). 예에는 트리코더마 론지프라키아툼 (*Trichoderma longibrachiatum*), 티. 비리데 (*T. viride*), 티. 코닝지 (*T. koningii*), 티. 하르지아눔 (*T. harzianum*), 페니실리움 (*Penicillium*) *sp.*, 휴미콜라 인솔레스 (*Humicola insolens*), 에이치. 라누기노스 (*H. lanuginose*), 에이치. 그리세아 (*H. grisea*), 크리소스포리움 (*Chrysosporium*) *sp.*, 씨. 루크누웬스 (*C. lucknowense*), 글리오클라디움 (*Gliocladium*) *sp.*, 아스퍼질러스 (*Aspergillus*) *sp.* (예컨대 에이. 오리자 (*A. oryzae*), 에이. 니거 (*A. niger*), 에이 소자 (*A. sojae*), 에이. 자포니쿠스 (*A. japonicus*), 에이. 니둘란스 (*A. nidulans*), 또는 에이. 아와모리 (*A. awamori*)), 푸사리움 (*Fusarium*) *sp.* (예컨대 에프. 로세움

(*F. roseum*), 예프. 그라미넘 (*F. graminum*), 예프. 세레알리스 (*F. cerealis*), 예프, 옥시스포루임 (*F. oxysporum*), 또는 예프. 베네나툼 (*F. venenatum*)), 뉴로스포라 (*Neurospora*) *sp.* (예컨대 엔. 크라싸 (*N. crassa*) 또는 히포크레아 (*Hypocrea*) *sp.*), 무코르 (*Mucor*) *sp.* (예컨대 엠. 미에헤이 (*M. miehei*)), 리조퍼스 (*Rhizopus*) *sp.*, 및 에메리셀라 (*Emericella*) *sp.* 세포를 포함한다. 특정 실시형태에서, 상기 기술된 바와 같이 조작되는 진균 세포는 에이. 니둘란스, 에이. 아와모리, 에이. 오리자, 에이. 아쿨레아터스, 에이. 니거, 에이. 자포티커스, 티. 레에세이 (*T. reesei*), 티. 비리데, 예프. 옥시스포럼 (*F. oxysporum*), 또는 예프. 솔라니 (*F. solani*)이다. 이러한 숙주에서 사용을 위한 예시적인 플라스미드 또는 플라스미드 성분은 미국 공개 특허 출원 제2011/0045563호에 기술된 것을 포함한다.

[0112] 효모가 또한 본 명세서에 기술된 방법에서 미생물 숙주 세포로서 사용될 수 있다. 예로는 사카로마이세스 (*Saccharomyces*) *sp.*, 스킴조사카로마이세스 (*Schizosaccharomyces*) *sp.*, 피키아 (*Pichia*) *sp.*, 한세놀라 폴리모르파 (*Hansenula polymorpha*), 피키아 스티피테스 (*Pichia stipites*), 클루이베로마이세스 마르시아너스 (*Kluyveromyces marxianus*), 클루이베로마이세스 (*Kluyveromyces*) *spp.*, 야로위아 리폴리티카 (*Yarrowia lipolytica*) 및 칸디다 (*Candida*) *sp.*가 포함된다. 일부 실시형태에서, 사카로마이세스 *sp.* 는 에스. 세레비지아이다 ([Romanos et al., Yeast, (1992), 8(6):423-488] 참조). 이러한 숙주와 사용을 위한 예시적인 플라스미드 또는 플라스미드 성분은 미국 특허 제7,659,097호 및 미국 공개 특허 출원 제2011/0045563호에 기술된 것을 포함한다.

[0113] 일부 실시형태에서, 숙주 세포는 예를 들어, 녹조류, 홍조류, 회조류, 클로라라크니오피트 (*chlorarachniophyte*), 유글레니드 (*euglenid*), 크로미스타 (*chromista*), 또는 와편모충류 (*dinoflagellate*)로부터 유래된 조류 세포일 수 있다 (예를 들어, [Saunders & Warmbrodt, "Gene Expression in Algae and Fungi, Including Yeast," (1993), National Agricultural Library, Beltsville, Md.] 참조). 조류 세포에서 사용을 위한 예시적인 플라스미드 또는 플라스미드 성분은 미국 공개 특허 출원 제2011/0045563호에 기술된 것을 포함한다.

[0114] 다른 실시형태에서, 숙주 세포는 시아노박테리움 (*cyanobacterium*), 예컨대 형태학을 기반으로 다음의 그룹 중 어느 하나로 분류되는 시아노박테리움이다: 클로로코칼레스 (*Chlorococcales*), 플레우로카프살레스 (*Pleurocapsales*), 오실라토리아레스 (*Oscillatoriales*), 노스토칼레스 (*Nostocales*), 시네코시ست릭 (*Synechosystic*) 또는 스티고네마탈레스 (*Stigonematales*) (예를 들어, [Lindberg et al., Metab. Eng., (2010) 12(1):70-79] 참조). 시아노박테리아 세포에서 사용을 위한 예시적인 플라스미드 또는 플라스미드 성분은 미국 공개 특허 출원 제2010/0297749호 및 제2009/0282545호, 및 국제 공개 특허 출원 번호 WO 2011/034863에 기술된 것을 포함한다.

[0115] 유전자 조작 방법

[0116] 미생물 세포는 당업자에게 속하는 분자 생물학 (재조합 기술 포함), 미생물학, 세포 생물학 및 생화학의 통상적인 기술을 사용하여 발효적 (6E)-8-히드록시제라니올 생산을 위해 조작될 수 있다. 이러한 기술은 문헌에 완전히 설명되어 있고, 예를 들어, 다음의 문헌들을 참조한다: "Molecular Cloning: A Laboratory Manual," fourth edition (Sambrook et al., 2012); "Oligonucleotide Synthesis" (M. J. Gait, ed., 1984); "Culture of Animal Cells: A Manual of Basic Technique and Specialized Applications" (R. I. Freshney, ed., 6th Edition, 2010); "Methods in Enzymology" (Academic Press, Inc.); "Current Protocols in Molecular Biology" (F. M. Ausubel et al., eds., 1987, and periodic updates); "PCR: The Polymerase Chain Reaction," (Mullis et al., eds., 1994); Singleton et al., Dictionary of Microbiology and Molecular Biology 2nd ed., J. Wiley & Sons (New York, N.Y. 1994).

[0117] 벡터는 유전 물질을 세포에 도입시키는데 사용되는 폴리뉴클레오티드 비히클이다. 본 명세서에 기술된 방법에서 유용한 벡터는 선형일 수 있거나 또는 원형일 수 있다. 벡터는 숙주 세포의 표적 게놈으로 통합될 수 있거나 또는 숙주 세포에서 독립적으로 복제될 수 있다. 많은 적용분야의 경우에, 안정한 형질전환체를 생산하는 통합 벡터가 바람직하다. 벡터는 예를 들어 복제 기원, 다중 클로닝 부위 (MCS), 및/또는 선별 마커를 포함할 수 있다. 발현 벡터는 전형적으로 특정 숙주 세포에서 폴리뉴클레오티드 서열 (중중 코딩 서열)의 발현을 촉진하는 조절 구성요소를 함유하는 발현 카세트를 포함한다. 벡터는 제한없이 통합 벡터, 원형 생물 플라스미드, 에피솜, 바이러스 벡터, 코스미드, 및 인공 염색체를 포함한다.

[0118] 발현 카세트에서 사용될 수 있는 예시적인 조절 구성요소는 프로모터, 인핸서, 내부 리보솜 진입 부위 (IRES), 및 다른 발현 제어 구성요소 (예를 들어, 전사 종결 신호, 예컨대 폴리아데닐화 신호 및 폴리-U 서열)를 포함한다.

다. 이러한 조절 구성요소는 예를 들어 [Goeddel, Gene Expression Technology: Methods In Enzymology 185, Academic Press, San Diego, Calif. (1990)]에 기술되어 있다.

[0119] 일부 실시형태에서, 박터는 게놈 편집, 예컨대 CRISPR 시스템을 수행할 수 있는 시스템을 도입시키는데 사용될 수 있다. 예를 들어, 2014년 3월 6일 공개된 미국 공개 특허 출원 제2014/0068797호를 참조하고, 또한 [Jinek M., *et al.*, "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity," Science 337:816-21, 2012]를 참조한다. II형 CRISPR-Cas9 시스템에서, Cas9는 부위 지정 엔도뉴클레아제, 즉 2개의 별개 엔도뉴클레아제 도메인 (HNH 및 RuvC/RNase H-유사 도메인)을 사용하여 특정한 표적 서열에서 폴리뉴클레오티드를 절단하도록 유도하거나, 또는 유도할 수 있는 효소이다. Cas9는 Cas9가 RNA에 의한 이의 절단 부위로 유도되기 때문에 임의의 바람직한 부위에서 DNA를 절단하도록 조작될 수 있다. 그러므로 Cas9는 또한 "RNA-가이드된 뉴클레아제"로서 기술된다. 보다 특히, Cas9는 표적 폴리뉴클레오티드 내 특이적 서열과 RNA 분자(들)의 적어도 일부분의 하이브리드화를 기반으로 특이적 폴리뉴클레오티드 표적으로 Cas9를 가이드하는, 하나 이상의 RNA 분자와 회합된다. Ran, F.A. 등 ("In vivo genome editing using *Staphylococcus aureus* Cas9," Nature 520(7546):186-91, 2015, Apr 9], 모든 확장 데이터 포함)은 8종의 II형 CRISPR-Cas9 시스템의 crRNA/tracrRNA 서열 및 2차 구조를 제시한다. Cas9-유사 합성 단백질이 또한 당분야에 공지되어 있다 (2014년 10월 23일 공개된, 미국 공개 특허 출원 제2014-0315985호 참조).

[0120] 실시예 1은 폴리뉴클레오티드 및 다른 유전적 변형을 에스. 세레비지아 세포의 게놈으로 도입시키기 위한 예시적인 통합 접근법을 기술한다.

[0121] 박터 또는 다른 폴리뉴클레오티드는 임의의 다양한 표준 방법, 예컨대 형질전환, 접합, 전기천공, 핵 미세주입, 형질도입, 형질감염 (예를 들어, 리포펙션 매개 또는 DEAE-텍스트린 매개 형질감염 또는 재조합 파지 바이러스 사용 형질감염), 칼슘 포스페이트 DNA 침전물과의 인큐베이션, DNA-코팅 미세투사기를 사용한 고속 충돌법, 및 원형질 융합법을 통해 미생물 세포에 도입될 수 있다. 형질전환체는 당분야에 공지된 임의의 방법으로 선별될 수 있다. 형질전환체를 선별하기 위한 적합한 방법은 미국 공개 특허 출원 제2009/0203102호, 제 2010/0048964호, 및 제2010/0003716호, 및 국제 공개 특허 출원 번호 WO 2009/076676, WO 2010/003007, 및 WO 2009/132220에 기술되어 있다.

[0122] 조작된 미생물 세포

[0123] 상기 기술된 방법은 (6E)-8-히드록시제라니올을 생산하고 일정 실시형태에서, 과생산하는 조작된 미생물 세포를 생성시키는데 사용된다. 조작된 미생물 세포는 천연 미생물 세포, 예컨대 본 명세서에 기술된 임의의 미생물 숙주 세포와 비교하여, 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개, 또는 그 이상의 유전자 변경, 예컨대 30-40개의 변경을 가질 수 있다. 하기 실시예에 기술된 조작된 미생물 세포는 1개, 2개 또는 3개의 유전자 변경을 가지지만, 당업자는 본 명세서에 기재된 지침에 따라서, 추가의 변경을 갖는 미생물 세포를 디자인할 수 있다. 일부 실시형태에서, 조작된 미생물 세포는 천연 미생물 세포와 비교하여, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 또는 4개 이하의 유전자 변경을 갖는다. 다양한 실시형태에서, (6E)-8-히드록시제라니올 생산을 위해 조작된 미생물 세포는 임의의 하기 예시적인 범위 내에 속하는 많은 유전자 변경을 가질 수 있다: 1-10, 1-9, 1-8, 2-7, 2-6, 2-5, 2-4, 2-3, 3-7, 3-6, 3-5, 3-4 등.

[0124] 일부 실시형태에서, 예컨대 (6E)-8-히드록시제라니올을 천연적으로 생산하지 않는 미생물 숙주 세포의 경우에, 조작된 미생물 세포는 적어도 하나의 이중성 제라닐 디포스페이트 디포스파타제 (제라니올 신타제)를 발현한다. 다양한 실시형태에서, 미생물 세포는 예를 들어, 다음을 포함하고 발현시킬 수 있다: (1) 단일 이중성 제라니올 신타제 유전자, (2) 동일할 수 있거나 또는 상이할 수 있는 둘 이상의 이중성 제라니올 신타제 (달리 말해서, 동일 이중성 제라니올 신타제 유전자의 다수 카피가 도입될 수 있거나 또는 다수의 상이한 이중성 제라니올 신타제 유전자가 도입될 수 있음), (3) 세포에 천연적이지 않은 단일 이중성 제라니올 신타제 유전자, 및 천연 제라니올 신타제 유전자의 하나 이상의 추가 카피, 또는 (4) 동일할 수 있거나 또는 상이할 수 있는 둘 이상의 비천연 제라니올 신타제 유전자, 및 천연 제라니올 신타제 유전자의 하나 이상의 추가 카피.

[0125] 일부 실시형태에서, 예컨대 제라니올-8-히드록실라제 효소를 갖지 않는 미생물 숙주 세포의 경우에, 조작된 미생물 세포는 적어도 하나의 이중성 제라니올 신타제 이외에도, 적어도 하나의 이중성 제라니올-8-히드록실라제를 발현한다. 다양한 실시형태에서, 미생물 세포는 예를 들어, 다음을 포함하고 발현시킬 수 있다: (1) 단일 이중성 제라니올-8-히드록실라제 유전자, (2) 동일할 수 있거나 또는 상이할 수 있는 둘 이상의 이중성 제라니올-8-히드록실라제 유전자 (달리 말해서, 동일한 이중성 제라니올-8-히드록실라제 유전자의 다수 카피가 도입될 수 있거나, 또는 다수의 상이한 이중성 제라니올-8-히드록실라제 유전자가 도입될 수 있음), (3) 세포에 천

연적이지만 않은 단일 이중성 제라니올-8-히드록실라제 및 천연 제라니올-8-히드록실라제 유전자의 하나 이상의 추가 카피, 또는 (4) 동일할 수 있거나 또는 상이할 수 있는 둘 이상의 비천연 제라니올-8-히드록실라제, 및 천연 제라니올-8-히드록실라제의 하나 이상의 추가 카피.

[0126] 이러한 조작된 숙주 세포는 제라닐-PP ((6E)-8-히드록시제라니올의 중간 전구체)의 생산으로 이어지는 경로를 통해서 유동성을 증가시키는 적어도 하나의 추가적 유전자 변경을 포함할 수 있다. 경로에서 이들 "상류" 효소는 이들 효소 활성을 갖는 임의의 이소폼, 파라로그 또는 오솔로그 (당업자가 쉽게 이해하게 되는 바와 같이 상이한 명칭으로 공지될 수 있음)를 포함하여, ATP-시트레이트 신타제, 아세틸-CoA 신씨타제, 티올라제, 히드록시메틸글루타릴 조효소 A 신타제 (HMG-CoA 신타제), 히드록시메틸글루타릴 조효소 A 리덕타제 (HMG-CoA 리덕타제), 메발로네이트 키나제, 포스포메발로네이트 키나제, 디포스포메발로네이트 디카복실라제, 이소펜테닐-디포스페이트 델타-이소머라제, 및 제라닐 디포스페이트 신타제를 포함한다. 적어도 하나의 추가적인 변경은 임의의 이용가능한 수단, 예를 들어 (1) 천연 효소(들)의 발현 또는 활성의 조절, (2) 천연 효소에 대한 유전자의 하나 이상의 추가 카피의 발현, 또는 (3) 하나 이상의 비천연 효소에 대한 유전자의 하나 이상의 카피의 발현을 통해서, 상류 경로 효소(들)의 활성을 증가시킬 수 있다.

[0127] 일부 실시형태에서, 경로를 통해 증가된 유동성은 상기 기술된 바와 같이, 피드백-탈조절 효소를 코딩하는 하나 이상의 유전자를 발현시켜 획득될 수 있다. 예를 들어, 조작된 숙주 세포는 하나 이상의 피드백-탈조절 HMG-CoA 리덕타제 유전자를 포함하고 발현시킬 수 있다.

[0128] 조작된 미생물 세포는 천연 뉴클레오티드 서열을 갖거나 또는 천연의 것과 상이한 도입된 유전자를 함유할 수 있다. 예를 들어, 천연 뉴클레오티드 서열은 특정 숙주 세포에서 발현을 위해 코돈 최적화될 수 있다. 임의의 이들 도입된 유전자에 의해 코딩된 아미노산 서열은 천연일 수 있거나 또는 천연의 것과 상이할 수 있다. 다양한 실시형태에서, 아미노산 서열은 천연 아미노산 서열과 적어도 0%, 70%, 75%, 80%, 85%, 90%, 95% 또는 100% 아미노산 서열 동일성을 갖는다.

[0129] 일부 실시형태에서, (6E)-8-히드록시제라니올에 대한 전구체의 증가된 이용률은 하나 이상의 부-경로 효소, 예컨대 이기능성 (2E,6E)-파르네실 디포스페이트 신타제/디메틸알릴트랜스퍼라제, 제라닐제라닐 파이로포스페이트 신타제, 및 아세틸 Co-A를 소모하는 임의의 부-경로 효소의 발현 또는 활성을 감소시켜서 획득될 수 있다. 예를 들어, 조작된 숙주 세포는 임의의 이들 효소의 발현을 하향조절시키도록 하나 이상의 프로모터 교환을 포함할 수 있고/있거나 이들 발현을 전체로 제거하기 위해 이들 유전자가 결실될 수 있다.

[0130] 본 명세서에 기술된 접근법은 진균 세포, 즉 효모 예스. 세레비지아 (진핵생물), 및 박테리아 세포, 즉 씨. 글루타미컴 (원핵생물)에서 수행되었다 (실시예 1 참조).

[0131] 예시적인 조작된 진균 세포

[0132] 일정 실시형태에서, 조작된 효모 (예를 들어, 예스. 세레비지아) 세포는 페틸라 세토이엔시스의 제라니올 신타제 (예를 들어, SEQ ID NO:5)와 적어도 70%, 75%, 80%, 85%, 90%, 95% 또는 100% 아미노산 서열 동일성을 갖는 이중성 제라니올 신타제를 발현한다. 특정 실시형태에서, 페틸라 세토이엔시스 제라니올 신타제는 SEQ ID NO:5를 포함한다.

[0133] 조작된 효모 (예를 들어, 예스. 세레비지아) 세포는 또한 일정 실시형태에서, 파세올러스 안굴라리스 유래 제라니올-8-히드록실라제 (예를 들어, SEQ ID NO:11)와 적어도 70%, 75%, 80%, 85%, 90%, 95% 또는 100% 아미노산 서열 동일성을 갖는 이중성 제라니올-8-히드록실라제를 발현한다. 특정 실시형태, 파세올러스 안굴라리스 제라니올-8-히드록실라제는 SEQ ID NO:11을 포함할 수 있다.

[0134] 이들은 조작된 효모 세포의 유일한 유전자 변경일 수 있거나, 또는 효모 세포는 보다 일반적으로 상기 기술된 바와 같이 하나 이상의 추가적인 유전자 변경을 포함할 수 있다.

[0135] 하나 이상의 추가적인 유전자 변경을 갖는 예시적인 효모 (예를 들어, 예스. 세레비지아) 세포는 대조군 세포에 비해서, 예를 들어 천연 예스. 세레비지아 이소펜테닐-디포스페이트 델타-이소머라제 (SEQ ID NO:25) 유전자의 추가적인 카피를 세포에 도입시키거나 또는 천연 예스. 세레비지아 이소펜테닐-디포스페이트 델타-이소머라제와 적어도 70%, 75%, 80%, 85%, 90%, 또는 95% 아미노산 서열 동일성을 갖는 이소펜테닐-디포스페이트 델타-이소머라제를 코딩하는 유전자를 도입시켜 생산된, 상류 경로 효소, 예컨대 이소펜테닐-디포스페이트 델타-이소머라제의 증가된 활성을 가질 수 있다.

[0136] 특정 실시형태에서, 조작된 효모 (예를 들어, 예스. 세레비지아) 세포는 전형적으로 천연 예스. 세레비지아

HMG-CoA 리덕타제와 적어도 70%, 75%, 80%, 85%, 90%, 또는 95% 아미노산 서열 동일성을 갖는 에스. 세레비지아 HMG-CoA 리덕타제의 변이체, 또는 아미노산 잔기 1-529가 결실된 에스. 세레비지아 HMG-CoA 리덕타제의 절단형 변이체를 추가로 발현한다. 특정 실시형태에서, 에스. 세레비지아 HMG-CoA 리덕타제 변이체는 SEQ ID NO:27을 포함할 수 있다.

[0137] **조작된 미생물 세포의 배양**

[0138] 본 명세서에 기술된 임의의 미생물 세포는 예를 들어, 유지, 성장, 및/또는 (6E)-8-히드록시제라니올 생산을 위해 배양될 수 있다.

[0139] 일부 실시형태에서, 배양물은 10-500의 600 nm에서의 광학 밀도, 예컨대 50-150의 광학 밀도까지 성장된다.

[0140] 다양한 실시형태에서, 배양물은 생산된 (6E)-8-히드록시제라니올을 적어도 10, 25, 50, 75, 100, 200, 300, 400, 500, 600, 700, 800, 또는 900 $\mu\text{g/L}$, 또는 적어도 1, 10, 50, 75, 100, 200, 300, 400, 500, 600, 700, 800, 또는 900 mg/L 또는 적어도 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 또는 10 g/L의 역가로 포함한다. 다양한 실시형태에서, 역가는 10 $\mu\text{g/L}$ 내지 10 g/L, 25 $\mu\text{g/L}$ 내지 10 g/L, 100 $\mu\text{g/L}$ 내지 10 g/L, 200 $\mu\text{g/L}$ 내지 5 g/L, 500 $\mu\text{g/L}$ 내지 4 g/L, 1 mg/L 내지 3 g/L, 500 mg/L 내지 2 g/L 범위 또는 상기 열거된 임의 값으로 경계된 임의 범위이다.

[0141] **배양 배지**

[0142] 미생물 세포 제한없이 최소 배지, 즉 세포 성장을 위해 가능한 최소 영양소를 함유하는 것을 포함한, 임의의 적합한 배지에서 배양될 수 있다. 전형적으로 최소 배지는 (1) 미생물 성장을 위한 탄소원; (2) 특정 미생물 세포 및 성장 조건에 의존적일 수 있는 염; 및 (3) 물을 포함한다. 적합한 배지는 또는 하기의 임의 조합을 포함할 수 있다: 성장 및 생성물 형성을 위한 질소원, 성장을 위한 황원, 성장을 위한 포스페이트원, 성장을 위한 금속염, 성장을 위한 비타민, 및 성장을 위한 다른 보조인자.

[0143] 임의의 적합한 탄소원은 숙주 세포를 배양시키는데 사용될 수 있다. 용어 "탄소원"은 미생물 세포에 의해 물질대사될 수 있는 하나 이상의 탄소-함유 화합물을 의미한다. 다양한 실시형태에서, 탄소원은 탄수화물 (예컨대 단당류, 이당류, 올리고당류, 또는 다당류), 또는 전화당 (예를 들어, 효소 처리 수크로스 시럽)이다. 예시적인 단당류는 글루코스 (덱스트로스), 프룩토스 (레볼로스), 및 갈락토스를 포함하고; 예시적인 올리고당은 텍스트란 또는 글루칸을 포함하며, 예시적인 다당류는 전분 및 셀룰로스를 포함한다. 적합한 당류는 C6 당류 (예를 들어, 프룩토스, 만노스, 갈락토스, 또는 글루코스) 및 C5 당류 (예를 들어, 자일로스 또는 아라비노스)를 포함한다. 다른 덜 비싼 탄소원에는 사탕수수 주스, 비트 주스, 수수 주스 등이 포함되며, 임의의 이들은 그럴 필요는 없지만, 완전히 또는 부분적으로 탈이온화될 수 있다.

[0144] 배양 배지 중 염은 세포가 단백질 및 핵산을 합성할 수 있도록, 일반적으로 필수 원소, 예컨대 마그네슘, 질소, 인, 및 황을 제공한다.

[0145] 최소 배지는 하나 이상의 선별제, 예컨대 항생제가 보충될 수 있다.

[0146] (6E)-8-히드록시제라니올을 생산하기 위해서, 배양 배지는 글루코스 및/또는 질소원 예컨대 우레아, 암모늄 염, 암모니아, 또는 이의 임의 조합을 포함할 수 있고/있거나, 배양 동안 보충된다.

[0147] **배양 조건**

[0148] 미생물 세포의 유지 및 성장을 위해 적합한 재료 및 방법은 당분야에서 충분히 공지되어 있다. 예를 들어, 미국 공개 특허 출원 제2009/0203102호, 제2010/0003716호, 및 제2010/0048964호, 및 국제 공개 특허 출원 번호 WO 2004/033646, WO 2009/076676, WO 2009/132220, 및 WO 2010/003007, [Manual of Methods for General Bacteriology Gerhardt et al., eds), American Society for Microbiology, Washington, D.C. (1994)] 또는 [Brock in Biotechnology: A Textbook of Industrial Microbiology, Second Edition (1989) Sinauer Associates, Inc., Sunderland, Mass]을 참조한다.

[0149] 일반적으로, 세포는 적절한 온도, 가스 혼합물, 및 pH (예컨대 약 20°C 내지 약 37°C, 약 6% 내지 약 84% CO₂, 및 약 5 내지 약 9의 pH)에서 성장 및 유지된다. 일부 양태에서, 세포는 35°C에서 성장된다. 일정 실시형태에서, 예컨대 호열성 박테리아가 숙주 세포로서 사용되는 경우에, 더 높은 온도 (예를 들어, 50°C-75°C)가 사용될 수 있다. 일부 양태에서, 발효를 위한 pH 범위는 약 pH 5.0 내지 약 pH 9.0 (예컨대 약 pH 6.0 내지 약 pH 8.0 또는 약 6.5 내지 약 7.0)이다. 세포는 특정 세포의 요건을 기반으로 호기성, 무산소, 또는 혐기

성 조건 하에서 성장될 수 있다.

[0150] 사용될 수 있는 표준 배양 조건 및 발효 방식, 예컨대 회분식, 유가식, 또는 연속 발효는 미국 공개 특허 출원 제2009/0203102호, 제2010/0003716호, 및 제2010/0048964호, 및 국제 공개 특허 출원 번호 WO 2009/076676, WO 2009/132220, 및 WO 2010/003007에 기술되어 있다. 회분식 및 유가식 발효는 일반적으로 충분히 당분야에 공지되어 있고, 예는 [Brock, Biotechnology: A Textbook of Industrial Microbiology, Second Edition (1989) Sinauer Associates, Inc]에서 확인할 수 있다.

[0151] 일부 실시형태에서, 세포는 제한된 당류 (예를 들어, 글루코스) 조건 하에서 배양된다. 다양한 실시형태에서, 부가되는 당류의 양은 세포가 소모할 수 있는 당류 양의 약 105% 또는 그 미만 (예컨대 약 100%, 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 또는 10%)이다. 특정 실시형태에서, 배양 배지에 부가되는 당류의 양은 특정 시기 동안 세포에 의해 소모되는 당류의 양과 대략 동일하다. 일부 실시형태에서, 세포 성장 속도는 세포가 세포 배지 중 당류의 양에 의해 지원될 수 있는 속도로 성장되도록 부가되는 당류의 양을 제한하여 제어된다. 일부 실시형태에서, 당류는 세포가 배양되는 시간 동안 축적되지 않는다. 다양한 실시형태에서, 세포는 약 1, 2, 3, 5, 10, 15, 20, 25, 30, 35, 40, 50, 60 또는 70시간 이상의 시간 동안 또는 심지어 최대 약 5-10일 동안 제한된 당 조건 하에서 배양된다. 다양한 실시형태에서, 세포는 세포가 배양되는 총 시간 길이의 약 5, 10, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 95, 또는 100% 이상 동안 제한된 당 조건 하에서 배양된다. 임의의 특정 이론에 국한하려는 의도는 아니지만, 제한된 당 조건은 세포의 보다 호의적인 조절을 허용할 수 있다고 여겨진다.

[0152] 일부 양태에서, 세포는 회분식 배양으로 성장된다. 세포는 또한 유가 배양 또는 연속 배양으로 성장될 수 있다. 추가로, 세포는 제한없이 상기 기술된 임의의 최소 배지를 포함한, 최소 배지에서 배양될 수 있다. 최소 배지는 1.0% (w/v) 이하의 글루코스 (또는 임의의 다른 6탄당)가 더 보충될 수 있다. 특히, 최소 배지는 1% (w/v), 0.9% (w/v), 0.8% (w/v), 0.7% (w/v), 0.6% (w/v), 0.5% (w/v), 0.4% (w/v), 0.3% (w/v), 0.2% (w/v), 또는 0.1% (w/v)의 글루코스가 보충될 수 있다. 일부 배양에서, 당 (예를 들어, 글루코스)의 유의하게 더 높은 수준, 예를 들어 적어도 10% (w/v), 20% (w/v), 30% (w/v), 40% (w/v), 50% (w/v), 60% (w/v), 70% (w/v), 또는 배지 중 당의 용해 한도까지 사용된다. 일부 실시형태에서, 당 수준은 상기 값 중 어느 2개의 범위 내, 예를 들어 0.1-10% (w/v), 1.0-20% (w/v), 10-70% (w/v), 20-60 % (w/v), 또는 30-50% (w/v)이다. 뿐만 아니라, 상이한 당 수준이 상이한 배양 시기에 대해 사용될 수 있다. 유가 배양 (예를 들어, 에스. 세레비시아 또는 씨. 글루타미쿰 경우)에서, 당의 수준은 회분식 시기에는 약 100-200 g/L (10-20% (w/v))일 수 있고 그 다음으로 최대 약 500-700 g/L (공급 시 50-70%)일 수 있다.

[0153] 추가로, 최소 배지는 0.1% (w/v) 이하의 효모 추출물이 보충될 수 있다. 특히 최소 배지는 0.1% (w/v), 0.09% (w/v), 0.08% (w/v), 0.07% (w/v), 0.06% (w/v), 0.05% (w/v), 0.04% (w/v), 0.03% (w/v), 0.02% (w/v), 또는 0.01% (w/v)의 효모 추출물이 보충될 수 있다. 대안적으로, 최소 배지는 1% (w/v), 0.9% (w/v), 0.8% (w/v), 0.7% (w/v), 0.6% (w/v), 0.5% (w/v), 0.4% (w/v), 0.3% (w/v), 0.2% (w/v), 또는 0.1% (w/v)의 글루코스, 및 0.1% (w/v), 0.09% (w/v), 0.08% (w/v), 0.07% (w/v), 0.06% (w/v), 0.05% (w/v), 0.04% (w/v), 0.03% (w/v), 또는 0.02% (w/v)의 효모 추출물이 보충될 수 있다. 일부 배양에서, 효모 추출물의 유의하게 더 높은 수준, 예를 들어 적어도 1.5% (w/v), 2.0% (w/v), 2.5% (w/v), 또는 3% (w/v)가 사용될 수 있다. 일부 배양 (예를 들어, 에스. 세레비시아 또는 씨. 글루타미쿰의 경우)에서, 효모 추출물 수준은 상기 값 중 어느 2개의 범위 내, 예를 들어 0.5-3.0% (w/v), 1.0-2.5% (w/v), 또는 1.5-2.0% (w/v)이다.

[0154] 본 명세서에 기술된 조작된 미생물 세포의 유지 및 성장에 적합한 예시적인 재료 및 방법은 하기 실시예 1에서 확인할 수 있다.

[0155] (6E)-8-히드록시제라니올 생산 및 회수

[0156] 본 명세서에 기술된 임의의 방법은 (6E)-8-히드록시제라니올을 회수하는 단계를 더 포함할 수 있다. 일부 실시형태에서, 소위 수확 스트림에 함유되는 생산된 (6E)-8-히드록시제라니올은 생산 용기로부터 회수/수확된다. 수확 스트림은 생산 용기 내 휴지기 세포에 의한 생산 기질의 전환으로 인한 결과로서 (6E)-8-히드록시제라니올을 함유하는, 생산 용기로부터 비롯된, 예를 들어 세포-무함유 또는 세포-함유 수용액을 포함할 수 있다. 수확 스트림에 여전히 존재하는 세포는 당분야에 공지된 임의의 작업, 예컨대 예를 들어 여과, 원심분리, 경사법, 막 직교류 한외여과 또는 미세여과, 접선 유동 한외여과 또는 미세여과 또는 전량 여과를 통해 (6E)-8-히드록시제라니올로부터 분리될 수 있다. 이러한 세포 분리 작업 후에, 수확 스트림은 본질적으로 세포가 없다.

- [0157] 수확 스트림에 함유된 다른 성분으로부터 생산된 (6E)-8-히드록시제라니올의 분리 및/또는 정제의 추가 단계, 즉, 소위 하류 처리 단계가 임의로는 수행될 수 있다. 이들 단계는 당업자에게 공지된 임의 수단, 예컨대, 예를 들어 농축, 추출, 결정화, 침전, 흡착, 이온 교환 및/또는 크로마토그래피를 포함할 수 있다. 임의의 이들 과정은 (6E)-8-히드록시제라니올을 정제하기 위해서 조합하여 또는 단독으로 사용될 수 있다. 추가 정제 단계는 예를 들어, 농축, 결정화, 침전, 세척 및 건조, 활성탄 처리, 이온 교환, 나노여과 및/또는 재결정화 중 하나 이상을 포함할 수 있다. 적합한 정제 프로토콜의 디자인은 세포, 배양 배지, 배양의 크기, 생산 용기 등에 의존적일 수 있고, 당업자의 수준 내에 속한다.
- [0158] 하기 실시예는 본 개시의 다양한 실시형태를 예시하려는 목적으로 제공되며 임의 방식으로 본 개시를 제한하려는 것을 의미하지 않는다. 청구항의 범주로 정의된, 본 개시의 의도에 포괄되는 이의 변화 및 다른 용도는 당업자가 확인할 수 있을 것이다.
- [0159] **실시예 1 - (6E)-8-히드록시제라니올을 생산하도록 조작된 사카로마이세스 세레비지아 균주의 제작 및 선별**
- [0160] **플라스미드/DNA 디자인**
- [0161] 이러한 작업에서 시험된 모든 균주들은 사유 소프트웨어를 사용하여 디자인된 플라스미드 DNA로 형질전환되었다. 플라스미드 디자인은 이 작업에서 조작된 2종의 숙주 유기체 중 하나에 특이적이었다. 플라스미드 DNA는 표준 DNA 조립 방법에 의해 물리적으로 제작하였다. 이어서, 이 플라스미드 DNA를 사용하여 각각 하기에 기술된 2종의 숙주-특이적 방법 중 하나를 통해 물질대사 경로를 통합시켰다.
- [0162] **에스. 세레비지아 경로 통합**
- [0163] "분할-마커, 이중-교차" 게놈 통합 전략이 에스. 세레비지아 균주를 조작하기 위해 개발되었다. 도 2는 에스. 세레비지아에서 상보성, 분할-마커 플라스미드의 게놈 통합 및 콜로니 PCR을 통한 올바른 게놈 통합의 검증을 예시한다. 상보성 5' 및 3' 상동성 부분 및 URA3 선별 마커의 중복된 절반을 갖는 2종의 플라스미드 (직접 반복부는 해시된 막대로 표시됨)를 메가뉴클레아제로 분해하였고 선형 단편으로서 형질전환시켰다. 삼중-교차 사건은 바람직한 이중성 유전자를 표적화된 유전자좌에 통합시켰고 전체 URA3 유전자를 재구성시켰다. 이러한 통합 사건으로부터 유래된 콜로니는 5' 및 3' 접합부 둘 모두를 확인하기 위해 2개 3-프라이머 반응을 사용하여 어레이하였다 (UF/IF/wt-R 및 DR/IF/wt-F). 추가의 조작이 바람직한 균주의 경우에, 균주는 본래 직접 반복부의 소형 단일 카피가 남겨진, URA3의 제거에 대해 선별하기 위한 5-FOA 평판 상에 플레이팅하였다. 이러한 게놈 통합 전략은 동일한 작업 흐름으로 유전자 녹-아웃, 유전자 녹-인, 및 프로모터 적정에 사용될 수 있다.
- [0164] **세포 배양**
- [0165] 에스. 세레비지아에 대해 확립된 작업 흐름은 평판 전체에서 균주를 임의 추출하는 자동화 작업흐름을 사용해 성공적으로 구축된 균주를 강화시키는 적중-선발 (hit-picking) 단계를 포함한다. 성공적으로 구축된 각 균주의 경우에, 콜로니 대 콜로니 변이 및 다른 과정 변이를 시험하기 위해서 별개 콜로니로부터 최대 4개 복제물을 시험하였다. 4개보다 적게 콜로니가 수득되었다면, 존재하는 콜로니를 복제하여 각각의 바람직한 유전자형에 대해 적어도 4개 웰을 시험하였다.
- [0166] 콜로니는 선별 배지 (에스. 세레비지아의 경우 SD-ura)가 존재하는 96웰 평판에서 강화시켰고 포화까지 2일 동안 배양시킨 후에 16.6% 글리세롤과 함께 저장을 위해 -80℃에서 냉동시켰다. 냉동된 글리세롤 스톱은 이후에 냉동으로부터의 회복 및 성장을 돕기 위해 낮은 수준의 아미노산이 존재하는 최소 배지에서 씨드 단계를 접종하는데 사용되었다. 씨드 평판은 30℃에서 1-2일 동안 성장시켰다. 그 다음으로 씨드 평판은 최소 배지가 존재하는 주배양 평판을 접종시키는데 사용되었고 48-88시간 동안 성장시켰다. 평판은 바람직한 시점에 제거하였고 세포 밀도 (OD600), 생존능, 및 글루코스에 대해 시험하였고, 상청액 샘플은 관심 생성물에 대한 LC-MS 분석을 위해 저장하였다.
- [0167] **세포 밀도**
- [0168] 세포 밀도는 600 nm에서 각 웰의 흡광도를 검출하는 분광광도 어레이를 사용해 측정하였다. 로봇공학을 사용하여 각 배양 평판에서 어레이 평판으로 고정량의 배양물을 전달한 후에, 175 mM 소듐 포스페이트 (pH 7.0)를 혼합하여 10배 희석물을 생성시켰다. 어레이 평판은 Tecan M1000 분광광도계를 사용해 측정되었고 어레이 데이터는 LIMS 데이터베이스에 업로드하였다. 비접종 대조군을 사용하여 배경 흡광도를 차감하였다. 세포 성장은 각 단계에서 다수 평판을 접종하여 모니터링하였고, 그 이후 각 시점에 온전한 평판을 희생시켰다.

[0169] (측정 동안 비대표적 샘플을 야기시킬 수 있는) 대량의 평판을 취급하면서 세포 침강을 최소화시키기 위해서, 각 평판은 각각의 판독 전에 10-15초 동안 진탕시켰다. 평판 내 세포 밀도의 폭넓은 편차가 또한 선형 검출 범위 밖에서 흡광도 측정을 초래하여, 더 높은 OD 배양물의 과소추산을 야기시킬 수 있다. 일반적으로, 지금까지 시험된 균주들은 이것이 우려될 정도로 유의하게 충분히 가변적이지 않았다.

[0170] **액체-고체 분리**

[0171] LC-MS를 통한 분석을 위해 세포의 샘플을 수확하기 위해서, 액체상 및 고체상을 원심분리를 통해 분리하였다. 배양 평판을 2000 rpm으로 4분 동안 원심분리하였고, 상청액을 로봇공학을 사용해 목적 평판으로 전달하였다. 75 μ l의 상청액을 각 평판에 전달하였고, 하나는 4℃에 저장하였고, 두번째는 장기간 저장을 위해 80℃에 저장하였다.

[0172] **1차 라운드 유전자 조작 결과**

[0173] 유전자 조작 및 스크리닝의 1차 라운드는 숙주 세포로서 에스. 세레비지아를 사용하여 수행하였다. 라이브러리 접근법이 숙주 유기체에서 기능성 효소를 동정하는데 선택되었다. 제라니올 신타제 서열의 광범위 검색으로 총 13종의 오솔로그 서열이 확인되었다. 이중성 제라니올 신타제는 일부 경우에 이중성 제라니올-8-히드록실라제와 함께 숙주 세포에서 발현시켰다. 일부 경우에서, 제라니올 신타제 및/또는 제라니올-8-히드록실라제 뉴클레오티드 서열은 에스. 세레비지아에 대해 코돈-최적화되었다. 균주는 상기 기술된 바와 같이 생산하였고 배양하였으며, 배양 배지에서 (6E)-8-히드록시제라니올 역가는 LC-MS로 측정하였다. 균주 및 결과는 표 1 및 도 2에 표시되어 있다. 최고 성능 1차 라운드 균주는 파세올러스 안굴라리스 제라니올-8-히드록실라제와 함께, 페틸라 세토이엔시스 제라니올 신타제를 발현하는 에스. 세레비지아 균주로서, 37.5 μ g/배양 배지 L의 (6E)-8-히드록시제라니올 역가를 제공하였다. 이러한 균주는 유전자 조작 및 스크리닝의 2차 라운드에 선택되었다.

표 1

표 1. 사카로마이세스 세레비지아에서 1차 라운드 결과

군주 명칭	역가 (mg/L)	E1 Uniprot ID	호소 1- 활성 명칭	호소 1- 근원 유기체	E1 코돈 최적화 축약어	E2 Uniprot ID	호소 2- 활성 명칭	호소 2-근원 유기체	E2 코돈 최적화 축약어
Sc80 HGER _02	9.4	E5GAH8	제라니 올 신타제 (SEQ ID NO:1) (SEQ ID NO:2)	비티스 비니페라 (포도)	변형 Cg 코돈 용법	D1MI46	제라니올 8- 히드록실 라제 (SEQ ID NO:7) (SEQ ID NO:8)	스위티아 무스타(웅담)	변형 Cg 코돈 용법
Sc80 HGER _08	6.5	D7SQP4	제라니 올 신타제 (SEQ ID NO:3) (SEQ ID NO:4)	비티스 비니페라 (포도)	변형 Cg 코돈 용법	B9GW31	제라니올 8- 히드록실 라제 (SEQ ID NO:9) (SEQ ID NO:10)	포폴리스 트리코카르파(서부 발삼 포폴라) (포폴리스 발사미페라 subsp. 트리코카르파)	변형 Cg 코돈 용법
Sc80 HGER _11	37.5	C0KWW4	제라니 올 신타제	페릴라 세토이엔 시스	변형 Cg 코돈 용법	A0A0L9 UT99	제라니올 8- 히드록실	파세올리스 안콜라리스(팔) (비그나 안콜라리스)	변형 Cg 코돈 용법

[0174]

주 균 명	역가 (mg/L)	E1 Uniprot ID	호소 1- 활성 명칭	호소 1- 근원 유기체	E1 코돈 최적화 축약어	E2 Uniprot ID	호소 2- 활성 명칭	호소 2- 근원 유기체	E2 코돈 최적화 축약어
			(SEQ ID NO:5) (SEQ ID NO:6)				라제 (SEQ ID NO:11) (SEQ ID NO:12)		

2차 라운드 유전자 조작 결과

최고 성능의 1차 라운드 균주가 조합 라이브러리 접근법을 사용한 유전자 조작의 2차 라운드를 위한 출발 숙주로 사용되었다. 이 라운드에서, 1-3개 상류 경로 유전자의 추가 카피가 강력한, 항상성 프로모터의 제어 하에, 별개 "딸" 균주에 도입되었다 (표 2). 상류 경로 유전자는 (6E)-8-히드록시제라니올로 이어지는 경로의 마지막 천연 대사산물로 핵심 전구체 (예를 들어, 아세틸-CoA)의 전환에 관여되는 모든 유전자를 의미한다. 조합 라이브러리 접근법으로 균주에서 시험하려고 선택된 효소는 메발로네이트 경로 다이어그램에 도시되어 있다 (도 1). 이 라운드의 최고 성능 균주는 에스. 세레비지아로부터 이소펜테닐-디포스페이트 텔타-이소머라제를 과발현시켜서, 123 μg /배양 배지 L의 (6E)-8-히드록시제라니올 역가를 제공하였다.

표 2

표 2. 사카로마이세스 세레비지아의 2차 라운드 결과

관주 명칭	역가 (mg/L)	E1 Uniprot ID	호소 1-활성명칭	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성명칭	호소 3-근원 유기체
사카로마이세스 세레비지아										
Ss80 HGER _19	39.5	P07277	ATP-(R)- 메발로네이트 5- 포스포트랜스퍼라 제 (SEQ ID NO:13) (SEQ ID NO:14)	사카로마이세스 세레비지아						
Ss80 HGER _20	10.9	P08524	제라닐제라닐 파이로포스페이트 신타제 (SEQ ID NO:15) (SEQ ID NO:16)	사카로마이세스 세레비지아						
Ss80 HGER _21	42.8	P54839	아세틸- CoA:아세토아세틸- CoA C- 아세틸트랜스퍼라 제	사카로마이세스 세레비지아						

[0178]

균주 명칭	역가 (mg/L)	E1 Uniprot ID	호소 1-활성 명칭	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	호소 3-근원 유기체
			(SEQ ID NO:17) (SEQ ID NO:18)							
Sc80 HGER _23	40.5	P41338	아세틸-CoA:아세틸-CoA C-아세틸트랜스퍼라제 (SEQ ID NO:19) (SEQ ID NO:20)	사카로마이세스 세레비지아						
Sc80 HGER _24	36.6	P12684	(R)-메발로네이트:NADP + 옥시도리덕타제 (CoA 아실화) (SEQ ID NO:21) (SEQ ID NO:22)	사카로마이세스 세레비지아	P41338	아세틸-CoA:아세틸-CoA C-아세틸트랜스퍼라제 (SEQ ID NO:19) (SEQ ID NO:20)	사카로마이세스 세레비지아	P24521	ATP:(R)-5-포스포메발로네이트 포스포트랜스퍼라제 (SEQ ID NO:23) (SEQ ID NO:24)	사카로마이세스 세레비지아
Sc80 HGER _29	37.8	P12684	(R)-메발로네이트:NADP + 옥시도리덕타제 (CoA 아실화) (SEQ ID NO:21) (SEQ ID NO:22)	사카로마이세스 세레비지아	P07277	ATP:(R)-메발로네이트 5-포스포트랜스퍼라제 (SEQ ID NO:13) (SEQ ID NO:14)	사카로마이세스 세레비지아	P54839	아세틸-CoA:아세토아세틸-CoA C-아세틸트랜스퍼라제 (SEQ ID NO:17)	사카로마이세스 세레비지아

[0179]

근주 명칭	역가 (mg/L)	E1 Uniprot ID	호소 1·활성 명칭	호소 1·근원 유기체	E2 Uniprot ID	호소 2·활성 명칭	호소 2·근원 유기체	E3 Uniprot ID	호소 3·활성 명칭	호소 3·근원 유기체
Sc80 HGER _30	34.9	P24521	ATP:(R)-5- 포스포메탈로네이 트 포스포트랜스퍼라 제 (SEQ ID NO:23) (SEQ ID NO:24)	사카로마이세 스 세레비지아					(SEQ ID NO:18)	
Sc80 HGER _32	36.2	P12684	(R)- 메탈로네이트:NADP + 옥시도리덕타제 (CoA 아실화) (SEQ ID NO:21) (SEQ ID NO:22)	사카로마이세 스 세레비지아	P08524	제라닐제라닐 파일로포스페이 트 신타제 (SEQ ID NO:15) (SEQ ID NO:16)	사카로마이세 스 세레비지아	P15496	이소펜테닐- 디포스페이트 델타 3-델타 2- 이소머라제 (SEQ ID NO:25) (SEQ ID NO:26)	사카로마이세 스 세레비지아
Sc80 HGER _33	13.9	P12684	(R)- 메탈로네이트:NADP + 옥시도리덕타제 (CoA 아실화) (SEQ ID NO:21) (SEQ ID NO:22)	사카로마이세 스 세레비지아	P08524	제라닐제라닐 파일로포스페이 트 신타제 (SEQ ID NO:15) (SEQ ID NO:16)	사카로마이세 스 세레비지아	P24521	ATP:(R)-5- 포스포메탈로네이 트 포스포트랜스퍼라 제 (SEQ ID NO:23) (SEQ ID NO:24)	사카로마이세 스 세레비지아

[0180]

균주 명칭	역가 (mg/L)	E1 Uniprot ID	호소 1-활성 명칭	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	호소 3-근원 유기체
Sc80 HGER _36	11.7	P12684	(R)- 메발로네이트:NADP + 옥시도리덕타제 (CoA 아실화) (SEQ ID NO:21) (SEQ ID NO:22)	사카로마이세스 세레비지아	P08524	제라닐제라닐 파일로포스페이트 신타제 (SEQ ID NO:15) (SEQ ID NO:16)	사카로마이세스 세레비지아	P07277	ATP-(R)- 메발로네이트 5- 포스포트랜스퍼라 제 (SEQ ID NO:13) (SEQ ID NO:14)	사카로마이세스 세레비지아
Sc80 HGER _38	122.9	P15496	이소펜테닐- 디포스페이트 델타 3-델타 2- 이소머라제 (SEQ ID NO:25) (SEQ ID NO:26)	사카로마이세스 세레비지아	P15496	이소펜테닐- 디포스페이트 델타 3-델타 2- 이소머라제 (SEQ ID NO:25) (SEQ ID NO:26)	사카로마이세스 세레비지아	P15496	이소펜테닐- 디포스페이트 델타 3-델타 2- 이소머라제 (SEQ ID NO:25) (SEQ ID NO:26)	사카로마이세스 세레비지아
Sc80 HGER _39	35.6	P12684	(R)- 메발로네이트:NADP + 옥시도리덕타제 (CoA 아실화) (SEQ ID NO:21) (SEQ ID NO:22)	사카로마이세스 세레비지아						
Sc80 HGER _40	38.7	P15496	ATP-(R)-5- 디포스포메발로네 이트 카르복시-	사카로마이세스 세레비지아						

[0181]

균주 명칭	여가 (mg/L)	E1 Uniprot ID	호소 1·활성 명칭	호소 1·근원 유기체	E2 Uniprot ID	호소 2·활성 명칭	호소 2·근원 유기체	E3 Uniprot ID	호소 3·활성 명칭	호소 3·근원 유기체
			리아제 (SEQ ID NO:25) (SEQ ID NO:26)							

상기 호소 이외에도, 표 2 균주는 또한 1차 DBTAL 리온드의 최고 호소를 함유하였다. 사카로마이세스 세레비지아 속주는 제라니올 신타제 (UniProt ID COKWV4) 및 제라니올 8-하이드록실라제 (UniProt ID A0A0L9UT99)를 함유한다. 2차 라운드에서 시험된 모든 호소는 사카로마이세스 세레비지아 및 코리네박테리아 클루터미쿰에 대한 변형된 코돈 용법을 갖는다.

3차 및 4차 라운드의 유전자 조작 결과

유전자 조작의 3차 라운드는 개선된 균주가 생산되지 않았는데, 아마도 균주 제작의 오류 때문인 듯하다. 4차 (개선) 라운드의 균주 디자인 및 결과는 표 3에 표시되어 있다.

표 3

표 3. 사카로마이세스 세레비지아에서 (6P)-8-히드록시제라니올 생산을 개선시키기 위해 1-3 이종성 효소를 시험하기 위한 4 차 (개선) 라운드 균주 디자인 (개선 라운드)

균주 명칭	역가 (마이 크로그 램/L)	E1 Uniprot ID	효소 1- 활성 명칭	E1 변형	효소 1- 근원 유기체	E2 Uniprot ID	효소 2- 활성 명칭	효소 2- 근원 유기체	E3 Uniprot ID	효소 3- 활성 명칭	E3 변형	효소 3- 근원 유기체
에스. 세레비지아												
Ss80 HGER _38	113.597	P15496	이소펜테닐- 디포스페이트 델타 3-델타 2- 이소머라제 (SEQ ID NO25) (SEQ ID NO26)		에스. 세레비지아 (균주 ATCC 204508 / S288c) (발효도)							
Ss80 HGER	6.043	P12683	3-히드록시-3- 메틸글루타릴-	절단형	에스. 세레비지아							

군주 명칭	역가 (마이 크로그 램/L)	E1 Uniprot ID	호소 1-활성 명칭	E1 변형	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	E3 변형	호소 3-근원 유기체
_73			조호소 A 리덕타제 1 (HMG-CoA 리덕타제 1) (EC 1.1.1.34) (SEQ ID NO:27) (SEQ ID NO:28)		S288c							
Sc8O HGER _74	0	P12684	(R)-메발로네이트:NADP+ 옥시도리덕타제 (CoA 아실화) (SEQ ID NO:21) (SEQ ID NO:22)		에스. 세레비지아 S288c	P12684	(R)- 메발로네이트:NA DP+ 옥시도리덕타제 (CoA 아실화) (SEQ ID NO:21) (SEQ ID NO:22)	에스. 세레비지아 S288c				
Sc8O HGER _75	0	Q9FD86	3-히드록시-3-메틸글루타릴 조호소 A 리덕타제 (HMG- CoA 리덕타제) (EC 1.1.1.88) (SEQ ID NO:29) (SEQ ID NO:30)		스타필로코 커스 아우레우스							

[0186]

군주 명칭	역가 (마이크로그램/L)	E1 Uniprot ID	호소 1-활성 명칭	E1 변형	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	E3 변형	호소 3-근원 유기체
Sc8O HGER _76	0	V4HIU8	3-히드록시-3- 메틸글루타릴- CoA 리덕타제 (SEQ ID NO:31) (SEQ ID NO:32)		칸디다타리스 할로모넨 티렐렌시스 G22							
Sc8O HGER _77	0	A0A0A0 KA96	3-히드록시-3-메틸글루타릴 조효소 A 리덕타제 (HMG- CoA 리덕타제) (EC 1.1.34) (SEQ ID NO:33) (SEQ ID NO:34)		쿠쿠미스 사티버스 (오이)							
Sc8O HGER _78	0	A0A0F8 XPA3	3-히드록시-3-메틸글루타릴 조효소 A 리덕타제 (HMG- CoA 리덕타제) (SEQ ID NO:35) (SEQ ID NO:36)		로키알카에 움 sp. GC14_75							
Sc8O HGER _79	0	G3HXP6	3-히드록시-3- 메틸글루타릴- 조효소 A 리덕타제		크리세롤러 스 그리세우스							

[0187]

군주 명칭	역가 (마이 크로그 램/L)	E1 Uniprot ID	호소 1-활성 명칭	E1 변형	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	E3 변형	호소 3-근원 유기체
			(SEQ ID NO:37) (SEQ ID NO:38)									
Ss80 HGER _80	0	A0A0V8 RU49	3-히드록시-3-메틸글루타릴 조효소 A 리덕타제 (HMG- CoA 리덕타제) (EC 1.1.1.34) A 리덕타제 (SEQ ID NO:39) (SEQ ID NO:40)		피로딕티움 옥쿨럼							
Ss80 HGER _82	0	P08524	파르네실 파일로포스페 이트 신타제 (FPP 신타제) (FPS) (EC 2.5.1.10) ((2E,6E)- 파르네실 디포스페이트 신타제) (디메틸알릴트 랜스트랜스퍼	K197G	에스. 세레비지아 S288c							

[0188]

관주 명칭	역가 (마이 크로그 램/L)	티 Uniprot ID	호소 1-활성 명칭	티변형	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	E3변형	호소 3-근원 유기체
			라제) (EC 2.5.1.1) (파르네실 디포스페이트 신타제) (제라닐트랜스 트랜스퍼라제) (SEQ ID NO:41) (SEQ ID NO:42)									
Sc80 HGER _83	0	Q8LKJ2	제라닐디포스 페이트 신타제 (SEQ ID NO:43) (SEQ ID NO:44)		아비에스 그란디스							
Sc80 HGER _84	0	B1A9K6	제라닐디포스 페이트 신타제 2 (SEQ ID NO:45) (SEQ ID NO:46)		피세아 아비에스							
Sc80 HGER _85	0	BAP822 33.1	활성 명칭 불명 (SEQ ID NO:47) (SEQ ID NO:48)		0							

[0189]

군주 명칭	역가 (마이 크로그 램/L)	E1 Uniprot ID	호소 1-활성 명칭	E1 변형	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	E3 변형	호소 3-근원 유기체
Sc80 HGER _86	0	Q1A746	제라닐파일로 포스페이트 신타제 (SEQ ID NO:49) (SEQ ID NO:50)		솔라넘 리코페르시 컴							
Sc80 HGER _87	0	B2MV87	제라닐파일로 포스페이트 신타제 (EC 2.5.1.1) (SEQ ID NO:51) (SEQ ID NO:52)		카타린터스 로세우스							
Sc80 HGER _88	4,109	A0A160 PQU3	폴리프레닐 신씨타제 (SEQ ID NO:53) (SEQ ID NO:54)		씨. 글루타미컴							
Sc80 HGER _89	0	P54383	파르네실 디포스페이트 신타제 (FPP 신타제) (EC	S81F	바실러스 서브틸리스 (균주 168)							

[0190]

군주 명칭	역가 (마이크로그 램/L)	E1 Uniprot ID	호소 1-활성 명칭	E1 변형	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	E3 변형	호소 3-근원 유기체
			2.5.1.10) (2E.8E)- 파르네실 디포스페이트 신타제) (제라닐트랜스 트랜스퍼라제) (SEQ ID NO:55) (SEQ ID NO:56)									
S80 HGER _90	0	P07277	ATP:(R)- 메발로네이트 5- 포스포- 트랜스퍼라제 (SEQ ID NO:13) (SEQ ID NO:14)		에스. 세레비지아 S288c							
S80 HGER _91	0	P24521	ATP:(R)-5- 포스포- 메발로네이트 포스포- 트랜스퍼라제		에스. 세레비지아 S288c							

[0191]

균주 명칭	역가 (마이크로 그램/L)	E1 Uniprot ID	호소 1-활성 명칭	E1 변형	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	E3 변형	호소 3-근원 유기체
			(SEQ ID NO:23) (SEQ ID NO:24)									
S80 HGER _92	0	P41338	아세틸- CoA:아세틸-CoA C- 아세틸트랜스 퍼라제 (SEQ ID NO:29) (SEQ ID NO:20)		에스. 세레비지아 S288c							
S80 HGER _93	0	P32377	ATP(R)-5- 디포스포메발로네이트 카르복시-리아제 (ATP 첨가; 이소펜테닐-디포스페이트- 형성); (SEQ ID NO:57) (SEQ ID NO:58)		에스. 세레비지아 S288c							
S80 HGER _94		P54839	아세틸-CoA:아세토아세틸- CoA C-아세틸트랜스퍼라제 (티오에스테르-기수분해,		에스. 세레비지아 S288c							

[0192]

균주 명칭	역가 (마이 크로그 램/L)	E1 Uniprot ID	호소 1-활성 명칭	E1 변형	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	E3 변형	호소 3-근원 유기체
			카복시메틸-활성) (SEQ ID NO:17) (SEQ ID NO:18)									
Sc80 HGER _95		P07277	ATP-(R)- 메발로네이트 5- 포스포트랜스 퍼라제 (SEQ ID NO:13) (SEQ ID NO:14)		에스. 세레비지아 S288c	P12684	(R)- 메발로네이트:NA DP+ 옥시도리덕타제 (CoA 아실화) (SEQ ID NO:21) (SEQ ID NO:22)	에스. 세레비지아 S288c	P08524	파르네실 파이로포스페 이트 신타제 (EC 2.5.1.10), 디메틸알릴- 트랜스트랜스 퍼라제 (EC 2.5.1.1) (SEQ ID NO:77) (SEQ ID NO:78)	F96W, N127W	에스. 세레비지아 S288c
Sc80 HGER _96		P24521	ATP-(R)-5- 포스포- 메발로네이트 포스포- 트랜스퍼라제 (SEQ ID NO:23)		에스. 세레비지아 S288c	P12684	(R)- 메발로네이트:NA DP+ 옥시도리덕타제 (CoA 아실화) (SEQ ID NO:21)	에스. 세레비지아 S288c	P08524	파르네실 파이로포스페 이트 신타제 (EC 2.5.1.10), 디메틸알릴- 트랜스트랜스	F96W, N127W	에스. 세레비지아 S288c

[0193]

근주 명칭	역가 (마이크로그 램/L)	E1 Uniprot ID	호소 1-활성 명칭	E1 변형	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	E3 변형	호소 3-근원 유기체
			(SEQ ID NO:24)				(SEQ ID NO:22)			퍼라제(EC 2.5.1.1) (SEQ ID NO:77) (SEQ ID NO:78)		
Sc80 HGER _97		P41338	아세틸- CoA:아세틸-CoA C- 아세틸트랜스 퍼라제 (SEQ ID NO:19) (SEQ ID NO:20)		에스. 세레비지아 S288c	P12684	(R)- 메 발로네이트:NA DP+ 옥시도리덕타제 (CoA 아실화) (SEQ ID NO:21) (SEQ ID NO:22)	에스. 세레비지아 S288c	P08524	파르네실 파일로포스페 이트 신타제 (EC 2.5.1.10), 디메틸알릴- 트랜스트랜스 퍼라제(EC 2.5.1.1) (SEQ ID NO:77) (SEQ ID NO:78)	F96W, N127W	에스. 세레비지아 S288c
Sc80 HGER _98		P32377	ATP:(R)-5- 디포스포메 발로네이트 카르복시-리아제 (ATP 첨가; 이소펜테닐-디포스페이트- 형성);		에스. 세레비지아 S288c	P12684	(R)- 메 발로네이트:NA DP+ 옥시도리덕타제 (CoA 아실화)	에스. 세레비지아 S288c	P08524	파르네실 파일로포스페 이트 신타제 (EC 2.5.1.10), 디메틸알릴-	F96W, N127W	에스. 세레비지아 S288c

[0194]

균주 명칭	역가 (마이크로 그램/L)	E1 Uniprot ID	호소 1-활성 명칭	E1 변형	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	E3 변형	호소 3-근원 유기체
			(SEQ ID NO:57) (SEQ ID NO:58)				(SEQ ID NO:21) (SEQ ID NO:22)			트랜스트랜스 퍼라제(EC 2.5.1.1) (SEQ ID NO:77) (SEQ ID NO:78)		
Sc8O HGER _99		P54839	아세틸-CoA:아세토타세틸- CoA C-아세틸트랜스퍼라제 (티오에스테르-가수분해, 카복시메틸-형성) (SEQ ID NO:17) (SEQ ID NO:18)		에스. 세레비지아 S288c	P12884	(R)- 메탈로네이트.NA DP+ 옥시도리덕타제 (CoA 아실화) (SEQ ID NO:21) (SEQ ID NO:22)	에스. 세레비지아 S288c	P08524	파르네실 파이로포스페 이트신타제 (EC 2.5.1.10), 디메틸알릴- 트랜스트랜스 퍼라제(EC 2.5.1.1) (SEQ ID NO:77) (SEQ ID NO:78)	F96W, N127W	에스. 세레비지아 S288c
Sc8O HGER _100		Q8C3H5	YAL10E34793p (SEQ ID NO:59) (SEQ ID NO:60)		야로위아 리플리티카 CLIB 122 / E 150							

[0195]

군주 명칭	역가 (마이 크로그 램/L)	E1 Uniprot ID	호소 1-활성 명칭	E1 변형	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	E3 변형	호소 3-근원 유기체
Sc80 HGER _101		Q8ZKF6	아세틸-조호소 A 신생 타제 (AcCoA 신생 타제) (Acs) (EC 6.2.1.1) (아세틸 이트-- CoA 리가제) (아실-활성화 효소) (SEQ ID NO 61) (SEQ ID NO 62)	L641P	실모넬라 티피뮤리움 LT2							
Sc80 HGER _102		P54115	마그네슘-활성화 알데히드 디히드로게나제, 시토크롬 (EC 1.2.1.4) (Mg(2+)-활성화 아세틸알데히드 디히드로게나제) (Mg(2+)- ACDH) (SEQ ID NO 63) (SEQ ID NO 64)	에스. 세레비지아 S288c								

[0196]

관주 명칭	역가 (마이 크로그 램/L)	E1 Uniprot ID	호소 1- 활성 명칭	E1 변형	호소 1- 근원 유기체	E2 Uniprot ID	호소 2- 활성 명칭	호소 2- 근원 유기체	E3 Uniprot ID	호소 3- 활성 명칭	E3 변형	호소 3- 근원 유기체
Sc80 HGER _103		Q8ZKF6	아세틸-조효소 A 신생 타제 (AcCoA 신생 타제) (Acs) (EC 6.2.1.1) (아세틸이트-- CoA 리가제) (아실-활성화 효소) (SEQ ID NO 61) (SEQ ID NO 62)	L641P	살모넬라 티피뮤리움 LT2	P00331	알콜 디히드로게나제 2 (EC 1.1.1.1) (알콜 디히드로게나제 II) (YADH-2) (SEQ ID NO 65) (SEQ ID NO 66)	에스. 세레비지아 S288c				
Sc80 HGER _104		Q8ZKF6	아세틸-조효소 A 신생 타제 (AcCoA 신생 타제) (Acs) (EC 6.2.1.1) (아세틸이트-- CoA 리가제) (아실-활성화 효소)	L641P	살모넬라 티피뮤리움 LT2	P54115	마그네슘-활성화 알데히드 디히드로게나제, 시토솔 (EC 1.2.1.4) (Mg(2+)- 활성화 아세트알데히드 디히드로게나제)	에스. 세레비지아 S288c				

[0197]

관주 명칭	역가 (마이크로 그램/L)	E1 Uniprot ID	호소 1-활성 명칭	E1 변형	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	E3 변형	호소 3-근원 유기체
			호소 (SEQ ID NO:61) (SEQ ID NO:62)				(Mg(2+)-ACDH) (SEQ ID NO:63) (SEQ ID NO:64)					
Sc8O HGER _105		P00331	알콜 디히드로게나제 II (EC 1.1.1.1) (알콜 디히드로게나제 II) (YADH-2) (SEQ ID NO:65) (SEQ ID NO:66)		에스. 세레비지아 S288c	P54115	마그네슘-활성화 알데히드 디히드로게나제, 시토크롬 (EC 1.2.1.4) (Mg(2+)- 활성화 아세트알데히드 디히드로게나제) (Mg(2+)-ACDH) (SEQ ID NO:63) (SEQ ID NO:64)	에스. 세레비지아 S288c				
Sc8O HGER _106		P00331	알콜 디히드로게나제 II (EC 1.1.1.1) (알콜 디히드로게나제 II) (YADH-2) (SEQ ID NO:65) (SEQ ID NO:66)		에스. 세레비지아 S288c	P54115	마그네슘-활성화 알데히드 디히드로게나제, 시토크롬 (EC 1.2.1.4) (Mg(2+)- 활성화 아세트알데히드 디히드로게나제) (Mg(2+)-ACDH) (SEQ ID NO:63) (SEQ ID NO:64)	에스. 세레비지아 S288c	P41338	아세트알- CoA:아세트알-CoA C- 아세트알트랜스 퍼라제		에스. 세레비지아 S288c

[0198]

균주 명칭	역가 (마이 크로그 램/L)	E1 Uniprot ID	호소 1-활성 명칭	E1 변형	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	E3 변형	호소 3-근원 유기체
							활성화 아세트알데히드 디히드로게나제) (Mg(2+)-ACDH) (SEQ ID NO:63) (SEQ ID NO:64)			(SEQ ID NO:19) (SEQ ID NO:20)		
Sc8O HGER _107		P12684	(R)-메발로네이트:NADP+ 옥시도리덕타제 (CoA 아실화) (SEQ ID NO:21) (SEQ ID NO:22)		에스. 세레비리아 S288c	P40312	시토크롬 b5 (SEQ ID NO:71) (SEQ ID NO:72)	에스. 세레비리아 S288c	P22939	파르네실 파이로포스페 이트 신타제 (EC 2.5.1.10), 디메틸알릴- 트랜스트랜스 퍼라제 (EC 2.5.1.1) (SEQ ID NO:79) (SEQ ID NO:80)	S80F	이. 콜라이 (균주 K12)
Sc8O HGER _108		Q6C7Y1	YAL10D24431p (SEQ ID NO:67) (SEQ ID NO:68)		야로쿼아 리 폴리티카 CLIB 122 / E	Q5BAJ 5	시트레이트 리아제 서브유닛 (Eurofung)	에메리셀라 니틀란스 ATCC 38163				

[0199]

균주 명칭	역가 (마이크로 그램/L)	E1 Uniprot ID	호소 1-활성 명칭	E1 변형	호소 1-근원 유기체	E2 Uniprot ID	호소 2-활성 명칭	호소 2-근원 유기체	E3 Uniprot ID	호소 3-활성 명칭	E3 변형	호소 3-근원 유기체
			150				(SEQ ID NO:73) (SEQ ID NO:74)					
S80 HGER -109		P0AFG8	피루베이트 디하이드로게나제 E1 성분 (PDH E1 성분) (EC 1.2.4.1) (SEQ ID NO:69) (SEQ ID NO:70)		이. 콜라이 (균주 K12)	P06959	피루베이트 디하이드로게나제 복합체의 디하이드로리포일 리신-잔기 아세틸트랜스퍼 라제 성분 (EC 2.3.1.12) (피루베이트 디하이드로게나제 복합체 의 디하이드로리포아 미드 아세틸트랜스퍼 라제 성분) (E2) (SEQ ID NO:75) (SEQ ID NO:76)	이. 콜라이 (균주 K12)	E01VR5	디하드로리포 일 디하드로게나 제 (EC 1.8.1.4) (SEQ ID NO:81) (SEQ ID NO:82)	G187A, G191A,E 205V/M2 06R,F20 7K,D208 H,P212R	이. 콜라이 ATCC 9637

[0200]

균주 명칭	역가 (마이 크로그 램/L)	E1 Uniprot ID	효소 1-활성 명칭	E1 변형	효소 1-근원 유기체	E2 Uniprot ID	효소 2-활성 명칭	효소 2-근원 유기체	E3 Uniprot ID	효소 3-활성 명칭	E3 변형	효소 3-근원 유기체
Ss80 HGER _110		P0AF68	피루베이트 디하이드로게나제 E1 성분 (PDH E1 성분) (EC 1.2.4.1) (SEQ ID NO:69) (SEQ ID NO:70)		이. 콜라이 (균주 K12)	P06859	피루베이트 디하이드로게나제 복합체의 디하이드로리 포일 라신-진기 아세틸트랜스퍼 라제 성분 (EC 2.3.1.12) (피루베이트 디하이드로게나제 복합체의 디하이드로리 포아 미드 아세틸트랜스퍼 라제 성분) (E2) (SEQ ID NO:76) (SEQ ID NO:76)	이. 콜라이 (균주 K12)	E01YR5	디하이드로리 포 일 디하이드로게나 제 (EC 1.8.1.4) (SEQ ID NO:83) (SEQ ID NO:84)		이. 콜라이 ATCC 9637

주: 이들 균주가 의도한 대로 처음 2개 라운드에서 유전자를 갖지 않는다는 증거가 존재한다. 또한, 모든 효소 유전자는 Cg 및 Sc에 대해 변형된 코돈 용법을 가졌다.

실시예 2 - (6E)-8-히드록시제라니올을 생산하도록 조작된 야로위아 리폴리티카 균주의 제작 및 선별

야로위아 리폴리티카는 에스. 세레비지아에 대해 상기 기술된 접근법을 사용해 조작하였다. 유전자 조작의 1차 라운드에서 제작된 균주 및 그들 (6E)-8-히드록시제라니올 역가는 표 4에 표시되어 있다.

표 4

표 4. 아로위아 리플리티카에서 (6E)-8-히드록시제라니올 생산을 위한 1 차 라운드 결과

역가 ($\mu\text{g/L}$)	E1 Uni- prot ID	호소 1- 활성 명칭	E1 변형	호소 1- 근원 유기체	E1 코돈 최적화 축약어	E2 Uni- prot ID	호소 2- 활성 명칭	호소 2- 근원 유기체	E2 코돈 최적화 축약어	E3 Uni- prot ID	호소 3-활성 명칭	호소 3-근원 유기체	E3 코돈 최적화 축약어
와이. 리플리 티카													
Y180 HGE R_01	48.4 P22 939	파르네실 디포스페 이트 신타제 (SEQ ID NO:85) (SEQ ID NO:86)	S80F	이. 올라이 K12	비. 서브틸 리스	C0K WV4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:94)	페릴라 세토이엔 시스	비. 서브틸 리스	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:104)	파세올리스 인콜라리스	비. 서브틸 리스
Y180 HGE R_02	53.8 P22 939	파르네실 디포스페 이트	S80F	이. 올라이 K12	Cg 윗 sc에 대해	C0K WV4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:94)	페릴라 세토이엔 시스	Cg 윗 sc에 대해	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:103)	파세올리스 인콜라리스	Cg 윗 sc에 대해

역 가 (μ g /L)	E1 Uni- prot ID	호소 1- 활성 명칭	E1 변형	호소 1- 근원 유기체	E1 코돈 최적화 축약어	E2 Uni- prot ID	호소 2- 활성 명칭	호소 2- 근원 유기체	E2 코돈 최적화 축약어	E3 Uni- prot ID	호소 3- 활성 명칭	호소 3- 근원 유기체	E3 코돈 최적화 축약어
		신타제 (SEQ ID NO:85) (SEQ ID NO:87)			변형된 코돈 용법		NO:92) (SEQ ID NO:93)		변형된 코돈 용법				변형된 코돈 용법
Y18O HGE R_03	P22 939	파르네실 디포스페 이트 신타제 (SEQ ID NO:85) (SEQ ID NO:88)	S80F	이. 콜라이 K12	에스. 세레비 지아	C0K WV4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:95)	페렐라 세토이엔 시스	에스. 세레비 지아	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:105)	파세올러스 안콜라리스	에스. 세레비 지아
Y18O HGE R_04	P22 939	파르네실 디포스페 이트 신타제 (SEQ ID NO:85) (SEQ ID NO:88)	S80F	이. 콜라이 K12	와이. 리폴리 티카	C0K WV4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:96)	페렐라 세토이엔 시스	와이. 리폴리 티카	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:106)	파세올러스 안콜라리스	와이. 리폴리 티카

[0205]

역가 (μ g/L)	E1 Uni- prot ID	호소 1- 활성 명칭	E1 변형	호소 1- 근원 유기체	E1 코돈 최적화 축약어	E2 Uni- prot ID	호소 2- 활성 명칭	호소 2- 근원 유기체	E2 코돈 최적화 축약어	E3 Uni- prot ID	호소 3-활성 명칭	호소 3-근원 유기체	E3 코돈 최적화 축약어
		NO:89)											
Y180 HGE R_05	P22 939	파르네실 디포스페 이트 신타제 (SEQ ID NO:85) (SEQ ID NO:86)	S80F	이. 클라이 K12	비. 서브릴 리스	E5G AH8	제라니올 신타제 (SEQ ID NO:97) (SEQ ID NO:98)	비티스 비니페라	비. 서브릴 리스	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:104)	파세올러스 안굴라리스	비. 서브릴 리스
Y180 HGE R_06	P22 939	파르네실 디포스페 이트 신타제 (SEQ ID NO:85) (SEQ ID NO:87)	S80F	이. 클라이 K12	Cg 및 sc에 대해 변형된 코돈 용법	E5G AH8	제라니올 신타제 (SEQ ID NO:97) (SEQ ID NO:99)	비티스 비니페라	Cg 및 sc에 대해 변형된 코돈 용법	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:103)	파세올러스 안굴라리스	Cg 및 sc에 대해 변형된 코돈 용법
Y180 HGE R_07	P22 939	파르네실 디포스페	S80F	이. 클라이 K12	에스. 세레비	E5G AH8	제라니올 신타제 (SEQ ID	비티스 비니페라	에스. 세레비	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:105)	파세올러스 안굴라리스	에스. 세레비

[0206]

역 가 (μ g /L)	E1 Uni- prot ID	호소 1- 활성 명칭	E1 변형	호소 1- 근원 유기체	E1 코돈 최적화 촉약어	E2 Uni- prot ID	호소 2- 활성 명칭	호소 2- 근원 유기체	E2 코돈 최적화 촉약어	E3 Uni- prot ID	호소 3-활성 명칭	호소 3-근원 유기체	E3 코돈 최적화 촉약어
		이트 신타제 (SEQ ID NO:85) (SEQ ID NO:88)			지아		NO:97) (SEQ ID NO:100)		지아				지아
Y180 HGE R_08	P22 939	파르네실 디포스페 이트 신타제 (SEQ ID NO:85) (SEQ ID NO:89)	S80F	이. 콜라이 K12	와이. 리폴리 티카	E5G AH8	제라니올 신타제 (SEQ ID NO:97) (SEQ ID NO:101)	비티스 비니페라	와이. 리폴리 티카	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:106)	파세올러스 안글라리스	와이. 리폴리 티카
Y180 HGE R_09	P22 939	파르네실 디포스페 이트 신타제 (SEQ ID NO:85)	S80F	이. 콜라이 K12	비. 서브틸 리스	C0K WV4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:94)	펠라 세토이엔 시스	비. 서브틸 리스	D1M46	제라니올 8- 히드록실라제 (SEQ ID NO:107) (SEQ ID NO:108)	스워티아 무스티	비. 서브틸 리스

[0207]

역가 (μg /L)	E1 Uni-prot ID	호스 1-활성 명칭	E1 변형	호스 1-근원 유기체	E1 코돈 최적화 축약어	E2 Uni-prot ID	호스 2-활성 명칭	호스 2-근원 유기체	E2 코돈 최적화 축약어	E3 Uni-prot ID	호스 3-활성 명칭	호스 3-근원 유기체	E3 코돈 최적화 축약어
		(SEQ ID NO:86)											
Y180 HGE R_10 45.2	P22 939	파르네실 디포스페이트 신타제 (SEQ ID NO:85) (SEQ ID NO:88)	S80F	이. 클라이 K12	에스. 세레비 지아	C0K W4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:95)	페릴라 세토이엔 시스	에스. 세레비 지아	D1M46	제라니올 8-히드록실라제 (SEQ ID NO:107) (SEQ ID NO:109)	스위티아 무스티	에스. 세레비 지아
Y180 HGE R_11 62.9	P22 939	파르네실 디포스페이트 신타제 (SEQ ID NO:85) (SEQ ID NO:89)	S80F	이. 클라이 K12	와이. 리폴리 티카	C0K W4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:96)	페릴라 세토이엔 시스	와이. 리폴리 티카	D1M46	제라니올 8-히드록실라제 (SEQ ID NO:107) (SEQ ID NO:110)	스위티아 무스티	와이. 리폴리 티카
Y180 HGE 61.7	P08 524	파르네실 파이로포	F96W, N127	에스. 세레비	비. 서브틸	C0K W4	제라니올 신타제	페릴라 세토이엔	비. 서브틸	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102)	파세올리스 안굴라리스	비. 서브틸

[0208]

역 가 (μ g /L)	E1 Uni- prot ID	호소 1- 활성 명칭	E1 변형	호소 1- 근원 유기체	E1 코돈 최적화 축약어	E2 Uni- prot ID	호소 2- 활성 명칭	호소 2- 근원 유기체	E2 코돈 최적화 축약어	E3 Uni- prot ID	호소 3-활성 명칭	호소 3-근원 유기체	E3 코돈 최적화 축약어
R_12		스펙이트 신타제 (SEQ ID NO:77) (SEQ ID NO:90)	W	지아 S288c	리스		(SEQ ID NO:92) (SEQ ID NO:94)	시스	리스		(SEQ ID NO:104)		리스
Y180 HGE R_13	P08 524	파르네실 파일로포 스펙이트 신타제 (SEQ ID NO:77) (SEQ ID NO:78)	F96W, N127 W	에스. 세레비 지아 S288c	에스. 세레비 지아	C0K WV4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:95)	페릴라 세토이엔 시스	에스. 세레비 지아	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:105)	파세올러스 안쿨라리스	에스. 세레비 지아
Y180 HGE R_14	P08 524	파르네실 파일로포 스펙이트 신타제 (SEQ ID NO:77)	F96W, N127 W	에스. 세레비 지아 S288c	와이. 리폴리 티카	C0K WV4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:96)	페릴라 세토이엔 시스	와이. 리폴리 티카	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:106)	파세올러스 안쿨라리스	와이. 리폴리 티카

[0209]

역가 (μ g/L)	E1 Uni- prot ID	호소 1- 활성 명칭	E1 변형	호소 1- 근원 유기체	E1 코돈 최적화 축약어	E2 Uni- prot ID	호소 2- 활성 명칭	호소 2- 근원 유기체	E2 코돈 최적화 축약어	E3 Uni- prot ID	호소 3- 활성 명칭	호소 3- 근원 유기체	E3 코돈 최적화 축약어
		(SEQ ID NO:91)											
Y18O HGE R_15	P08 524	파르네실 파이로포 스페이트 신타제 (SEQ ID NO:77) (SEQ ID NO:90)	F96W, N127 W	에스. 세레비 지아 S288c	비. 서브릴 리스	E5G AH8	제라니올 신타제 (SEQ ID NO:97) (SEQ ID NO:98)	비티스 비니페라	비. 서브릴 리스	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:104)	파세올러스 안쿨라리스	비. 서브릴 리스
Y18O HGE R_16	P08 524	파르네실 파이로포 스페이트 신타제 (SEQ ID NO:77) (SEQ ID NO:78)	F96W, N127 W	에스. 세레비 지아 S288c	에스. 세레비 지아	E5G AH8	제라니올 신타제 (SEQ ID NO:97) (SEQ ID NO:100)	비티스 비니페라	에스. 세레비 지아	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:105)	파세올러스 안쿨라리스	에스. 세레비 지아
Y18O HGE 8	P08 524	파르네실 파이로포	F96W, N127	에스. 세레비	와이. 리폴리	E5G AH8	제라니올 신타제	비티스 비니페라	와이. 리폴리	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102)	파세올러스 안쿨라리스	와이. 리폴리

[0210]

역가 (μg /L)	E1 Uni-prot ID	호소 1-활성 명칭	E1 변형	호소 1-근원 유기체	E1 코돈 최적화 축약어	E2 Uni-prot ID	호소 2-활성 명칭	호소 2-근원 유기체	E2 코돈 최적화 축약어	E3 Uni-prot ID	호소 3-활성 명칭	호소 3-근원 유기체	E3 코돈 최적화 축약어
R_17		스페이트 신타제 (SEQ ID NO:77) (SEQ ID NO:91)	W	지아 S288c	티카		(SEQ ID NO:97) (SEQ ID NO:101)		티카		(SEQ ID NO:106)		티카
Y18O HGE R_18	P08 524	파르네실 파이로포 스페이트 신타제 (SEQ ID NO:77) (SEQ ID NO:90)	F96W, N127 W	에스. 세레비 지아 S288c	비. 서브릴 리스	C0K W4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:94)	페릴라 세 토이엔 시스	비. 서브릴 리스	D1M46	제라니올 8-히드록실라제 (SEQ ID NO:107) (SEQ ID NO:108)	스워티아 무스티	비. 서브릴 리스
Y18O HGE R_19	P08 524	파르네실 파이로포 스페이트 신타제 (SEQ ID NO:77)	F96W, N127 W	에스. 세레비 지아 S288c	에스. 세레비 지아	C0K W4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:95)	페릴라 세 토이엔 시스	에스. 세레비 지아	D1M46	제라니올 8-히드록실라제 (SEQ ID NO:107) (SEQ ID NO:109)	스워티아 무스티	에스. 세레비 지아

[0211]

역가 (μg/L)	E1 Uni-prot ID	호소 1-활성 명칭	E1 변형	호소 1-근원 유기체	E1 코돈 최적화 축약어	E2 Uni-prot ID	호소 2-활성 명칭	호소 2-근원 유기체	E2 코돈 최적화 축약어	E3 Uni-prot ID	호소 3-활성 명칭	호소 3-근원 유기체	E3 코돈 최적화 축약어
Y180 HGE R_20	P08 524	(SEQ ID NO:78) 피르네실 파이로포 스페이트 신타제 (SEQ ID NO:77) (SEQ ID NO:91)	F66W, N127 W	에스. 세레비지아	와이. 리폴리티카	C0K WV 4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:102) (SEQ ID NO:103)	페릴라 세토이엔시스	와이. 리폴리티카	D1M146	제라니올 8-히드록실라제 (SEQ ID NO:107) (SEQ ID NO:110)	스웨티아 무스티	와이. 리폴리티카
Y180 HGE R_21	C0K WV 4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:93)		페릴라 세토이엔시스	Cg 및 sc에 대해 변형된 코돈 용법	A0A 0L9U T99	제라니올 8-히드록실라제 (SEQ ID NO:102) (SEQ ID NO:103)	파세올러스 안굴라리스	Cg 및 sc에 대해 변형된 코돈 용법	P15496	이소펜테닐-디포스페이트 델타 3-델타 2-이소머라제 (SEQ ID NO:25) (SEQ ID NO:111)	에스. 세레비지아 S288c	Cg 및 sc에 대해 변형된 코돈 용법

실시예 3 - (6E)-8-히드록시제라니올 생산을 위한 숙주 평가

최고 성능 효소는 다음 4종의 숙주에서 시험하였다: 야로위아 리폴리티카, 바실러스 서브틸리스, 코리네박테리아 글루타미컴, 사카로마이세스 세레비지아. 에스. 세레비지아에 대한 결과는 하기 표 5에 표시되어 있다.

와이. 리폴리티카의 최고 성능 균주는 310 마이크로그램/L의 역가를 생산하였다. 이 와이. 리폴리티카 균주는 페릴라 세토이엔시스 유래 제라니올 신타제 (UniProt ID C0KVV4), 파세올러스 안굴라리스 유래 제라니올 8-히드록실라제 (UniProt ID A0A0L9UT99) 및 에스. 세레비지아 유래 이소펜테닐-디포스페이트 델타3-델타2-이소머라제 (UniProt ID P15496)를 발현하였다. 차선 성능 와이. 리폴리티카 균주는 200 마이크로그램/L를 생산하였고, 이 균주는 아미노산 치환 S80F [4]를 보유하는 에스케리치아 콜라이 K12 유래 파르네실 디포스페이트 신타제 (UniProt ID P22939), 비티스 비니페라 유래 제라니올 신타제 (UniProt ID E5GAH8), 및 파세올러스 안굴라리스 유래 제라니올 8-히드록실라제 (UniProt ID A0A0L9UT99)를 발현하였다.

에스. 세레비지아에서 최고 성능 균주는 217 마이크로그램/L의 역가를 생산하였다. 이 에스. 세레비지아 균주는 페릴라 세토이엔시스 유래 제라니올 신타제 (UniProt ID C0KVV4), 파세올러스 안굴라리스 유래 제라니올

8-히드록실라제 (UniProt ID A0A0L9UT99) 및 에스. 세레비지아 유래 이소펜테닐-디포스페이트 델타3-델타2-이소머라제 (UniProt ID P15496)를 발현하였다.

비. 서브틸리스 또는 씨 글루타미컴 균주에 의해 생산된 역가는 없었다.

표 5

표 5. 에스. 세레비지아에서 시험된 (6E)-8-히드록시제라니올 생산을 위한 숙주 평가 디자인

역가 (μ g/L)	E1 Uni- prot ID	호소 1- 활성 명칭	E1 변형	호소 1- 근원 유기체	E1 코돈 최적화 축약어	E2 Uni- prot ID	호소 2- 활성 명칭	호소 2- 근원 유기체	E2 코돈 최적화 축약어	호소 2- 활성 명칭	호소 2- 근원 유기체	E2 코돈 최적화 축약어	E3 Uni- prot ID	호소 3-활성 명칭	호소 3-근원 유기체	E3 코돈 최적화 축약어
에스. 세레비지아																
Sc8 OHG ER ₁₁₂	P22 939	파르네실 디포스페 이트 신타제 (SEQ ID NO:85) (SEQ ID NO:86)	S80F	OI. 콜라이 K12	비. 서브틸 리스	COK WV4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:94)	페릴라 세토이엔 시스 SEQ ID NO:99SEQ ID NO:101	비. 서브틸 리스	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:104)	파세올리스 안콜라리스	비. 서브틸 리스	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:104)	파세올리스 안콜라리스	비. 서브틸 리스
Sc8 OHG ER ₁₁₃	P22 939	파르네실 디포스페 이트 신타제 (SEQ ID NO:85) (SEQ ID NO:86)	S80F	OI. 콜라이 K12	에스. 세레비 지아	COK WV4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:95)	페릴라 세토이엔 시스 SEQ ID NO:99SEQ ID NO:101	에스. 세레비 지아	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:105)	파세올리스 안콜라리스	에스. 세레비 지아	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:105)	파세올리스 안콜라리스	에스. 세레비 지아

Sc8	23.4	OHG	ER_114	NO:85) (SEQ ID NO:88)	S80F	이. 콜라이 K12	와이. 리폴리 티카	C0K WV4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:96)	페릴라 세토이엔 시스	ID NO:102	와이. 리폴리 티카	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:106)	파세올러스 안콜라리스	와이. 리폴리 티카
Sc8	14.7	OHG	ER_115	NO:85) (SEQ ID NO:86)	S80F	이. 콜라이 K12	비. 서브틸 리스	E5G AH8	제라니올 신타제 (SEQ ID NO:97) (SEQ ID NO:98)	비티스 비니페라	비. 서브틸 리스	비. 서브틸 리스	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:104)	파세올러스 안콜라리스	비. 서브틸 리스
Sc8	11.3	OHG	ER_116	NO:85) (SEQ ID NO:86)	S80F	이. 콜라이 K12	에스. 세레비 지아	E5G AH8	제라니올 신타제 (SEQ ID NO:97) (SEQ ID NO:98)	비티스 비니페라	에스. 세레비 지아	에스. 세레비 지아	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:105)	파세올러스 안콜라리스	에스. 세레비 지아

[0219]

Sc8	15.5	P22	939	NO:85) (SEQ ID NO:88)	S80F	오. 클라이 K12	와이. 리폴리 티카	E5G AH8	NO:100)	비티스 비니페라	와이. 리폴리 티카	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:106)	파세올러스 인쿨라리스		
Sc8	9.2	P22	939	파르네실 디포스페 이트 신타제 (SEQ ID NO:85) (SEQ ID NO:89)	S80F	오. 클라이 K12	비. 서브틸 리스	C0K WV4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:94)	페릴라 세토이엔 시스	비. 서브틸 리스	D1M46	제라니올 8- 히드록실라제 (SEQ ID NO:107) (SEQ ID NO:108)	스워티아 무스티	비. 서브틸 리스	
Sc8	15.3	P22	939	파르네실 디포스페 이트 신타제 (SEQ ID NO:86)	S80F	오. 클라이 K12	Cg 및 sc에 대해 변형된	C0K WV4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:94)	페릴라 세토이엔 시스	Cg 및 sc에 대해 변형된	D1M46	제라니올 8- 히드록실라제 (SEQ ID NO:107) (SEQ ID NO:112)	스워티아 무스티	Cg 및 sc에 대해 변형된	

[0220]

Sc8 OHG ER_120	20.5	P22 939	NO:85) (SEQ ID NO:87)	S80F	이. 콜리이 K12	에스. 세레비 지아	C0K WW4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:95)	페릴라 세토이엔 시스	에스. 세레비 지아	D1M146	제라니올 8- 히드록실라제	스위티아 무스티	에스. 세레비 지아	코론 용법
Sc8 OHG ER_121	10.6	P22 939	파르네실 디포스페 이트 신타제 (SEQ ID NO:85) (SEQ ID NO:88)	S80F	이. 콜리이 K12	와이. 리폴리 티카	C0K WW4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:96)	페릴라 세토이엔 시스	와이. 리폴리 티카	D1M146	제라니올 8- 히드록실라제 (SEQ ID NO:107) (SEQ ID NO:109)	스위티아 무스티	와이. 리폴리 티카	코론 용법
Sc8 OHG ER_122	85.4	P08 524	파르네실 파일로포 스페이트 신타제 (SEQ ID	F96W, N127W	에스. 세레비 지아 S288c	비. 서브틸 리스	C0K WW4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID	페릴라 세토이엔 시스	비. 서브틸 리스	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:104)	파세올러스 안쿨라리스	비. 서브틸 리스	코론 용법

[0221]

Sc8	78.6	P08	524	NO:77) (SEQ ID NO:90)	F96W, N127W	에스. 세레비 지아 S288c	와이. 리폴리 티카	COK WV4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:96)	페릴라 세토이엔 시스	와이. 리폴리 티카	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:106)	파세올러스 안굴라리스	와이. 리폴리 티카
Sc8	54.4	P08	524	파르네실 파이로포 스페이트 신타제 (SEQ ID NO:77) (SEQ ID NO:91)	F96W, N127W	에스. 세레비 지아 S288c	비. 서브틸 리스	E5G AH8	제라니올 신타제 (SEQ ID NO:97) (SEQ ID NO:98)	비티스 비니페라	비. 서브틸 리스	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:104)	파세올러스 안굴라리스	비. 서브틸 리스
Sc8	52.7	P08	524	파르네실 파이로포 스페이트 신타제 (SEQ ID	F96W, N127W	에스. 세레비 지아 S288c	에스. 세레비 지아	E5G AH8	제라니올 신타제 (SEQ ID NO:97) (SEQ ID	비티스 비니페라	에스. 세레비 지아	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:105)	파세올러스 안굴라리스	에스. 세레비 지아

[0222]

Sc8	94.4	P08	524	NO:77) (SEQ ID NO:78)	F96W, N127W	에스. 세레비 지아 S288c	와이. 리폴리 티카	E5G AH8	제라니올 신타제 (SEQ ID NO:97) (SEQ ID NO:101)	비티스 비니페라	와이. 리폴리 티카	A0A0L9U T99	불확정 단백질 (SEQ ID NO:102) (SEQ ID NO:106)	파세올러스 안굴라리스	와이. 리폴리 티카
Sc8	106.6	P08	524	파르네실 파일로포 스페이트 신타제 (SEQ ID NO:77) (SEQ ID NO:91)	F96W, N127W	에스. 세레비 지아 S288c	에스. 세레비 지아	C0K WV4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:95)	페릴라 세토이엔 시스	에스. 세레비 지아	D1M46	제라니올 8- 히드록실라제 (SEQ ID NO:107) (SEQ ID NO:109)	스워티아 무스티	에스. 세레비 지아
Sc8	49.4	P08	524	파르네실 파일로포 스페이트 신타제 (SEQ ID	F96W, N127W	에스. 세레비 지아 S288c	와이. 리폴리 티카	C0K WV4	제라니올 신타제 (SEQ ID NO:92) (SEQ ID	페릴라 세토이엔 시스	와이. 리폴리 티카	D1M46	제라니올 8- 히드록실라제 (SEQ ID NO:107) (SEQ ID NO:110)	스워티아 무스티	와이. 리폴리 티카

[0223]

Sc8	217.0	C0 KW V4	NO:77) (SEQ ID NO:91)	제라니올 신타제 (SEQ ID NO:92) (SEQ ID NO:93)	페클라 세프이 엔시스	Cg 및 sc에 대해 변경된 코드 용어	A0A0 L9UT 99	NO:96)	제라니올 8- 히드록시 라제 (SEQ ID NO:102) (SEQ ID NO:103)	파세올러 스 안클라리 스	Cg 및 sc에 대해 변경된 코드 용어	P15496	이소펜테닐- 디포스페이트 델타 3-델타 2- 이소머라제 (SEQ ID NO:25) (SEQ ID NO:111)	에스 세레비자아 S288c	Cg 및 sc에 대해 변경된 코드 용어
-----	-------	----------------	-----------------------------	---	-------------------	--------------------------------------	--------------------	--------	--	------------------------	--------------------------------------	--------	--	----------------------	--------------------------------------

[0224]

[0225]

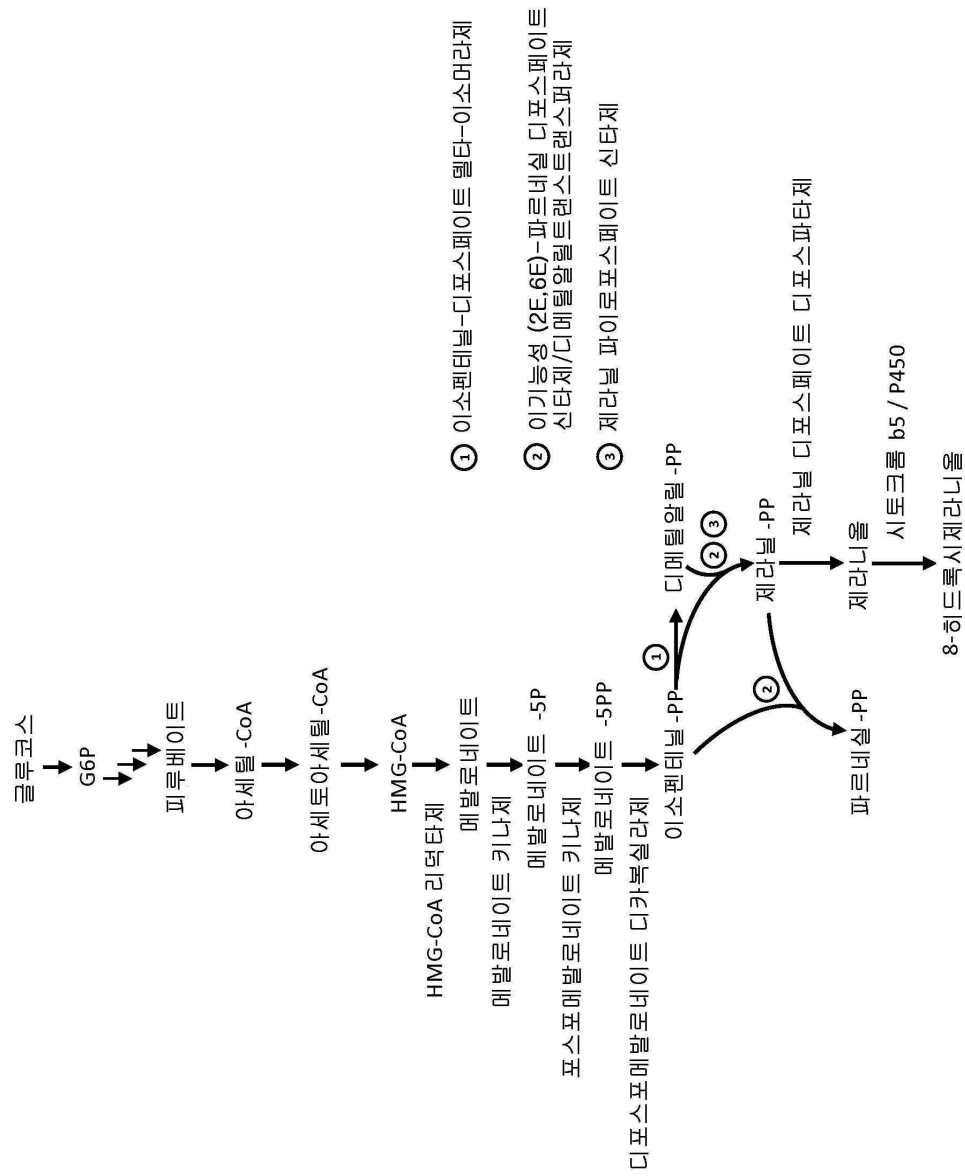
참조문헌

1. Polakowski, T., U. Stahl, and C. Lang, *Overexpression of a cytosolic hydroxymethylglutaryl-CoA reductase leads to squalene accumulation in yeast*. Appl Microbiol Biotechnol, 1998. **49**(1): p. 66-71.
2. Dimster-Denk, D., M.K. Thorsness, and J. Rine, Feedback regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase in *Saccharomyces cerevisiae*. Mol Biol Cell, 1994. **5**(6): p. 655-65.
3. Billingsley, J.M., et al., Engineering the biocatalytic selectivity of iridoid production in *Saccharomyces cerevisiae*. Metab Eng, 2017. **44**: p. 117-125.
4. Eiichiro Ono, N.T., Method for utilizing monoterpene glycosidation enzyme. 2015.
5. Roth, S., et al., *Chemoenzymatic Synthesis of a Novel Borneol-Based Polyester*. ChemSusChem, 2017. **10**(18): p. 3574-3580.
6. Crowley, J.H., et al., A mutation in a purported regulatory gene affects control of sterol uptake in *Saccharomyces cerevisiae*. J Bacteriol, 1998. **180**(16): p. 4177-83.
7. Vik, A. and J. Rine, Upc2p and Ecm22p, dual regulators of sterol biosynthesis in *Saccharomyces cerevisiae*. Mol Cell Biol, 2001. **21**(19): p. 6395-405.
1. Kanehisa, M. and S. Goto, *KEGG: kyoto encyclopedia of genes and genomes*. Nucleic Acids Res, 2000. **28**(1): p. 27-30.
2. Oswald, M., et al., *Monoterpenoid biosynthesis in Saccharomyces cerevisiae*. FEMS Yeast Res, 2007. **7**(3): p. 413-21.
3. Polakowski, T., U. Stahl, and C. Lang, Overexpression of a cytosolic hydroxymethylglutaryl-CoA reductase leads to squalene accumulation in yeast. Appl Microbiol Biotechnol, 1998. **49**(1): p. 66-71.
4. Reiling, K.K., et al., *Mono and diterpene production in Escherichia coli*. Biotechnol Bioeng, 2004. **87**(2): p. 200-12.

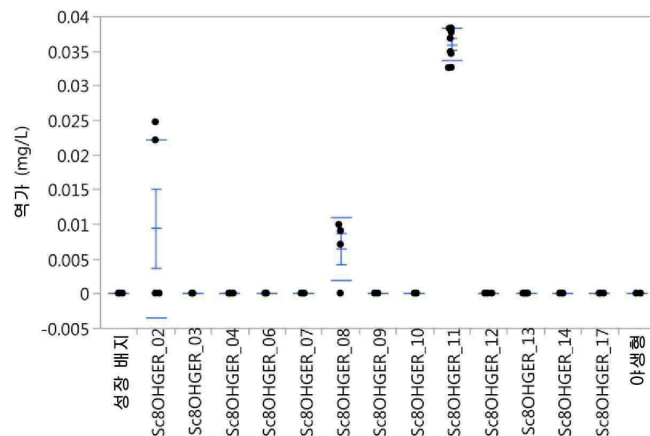
[0226]

도면

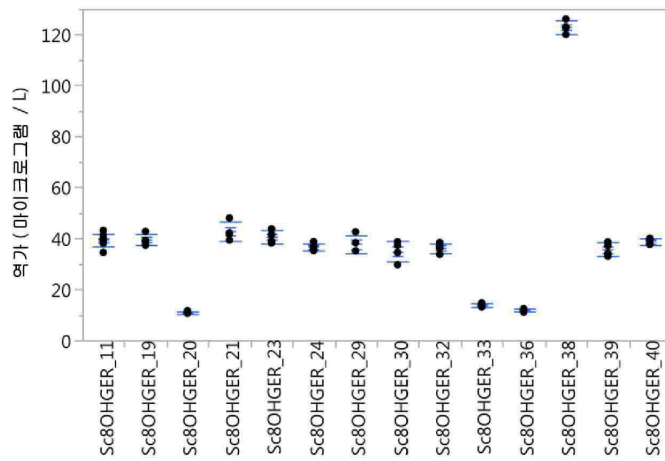
도면1



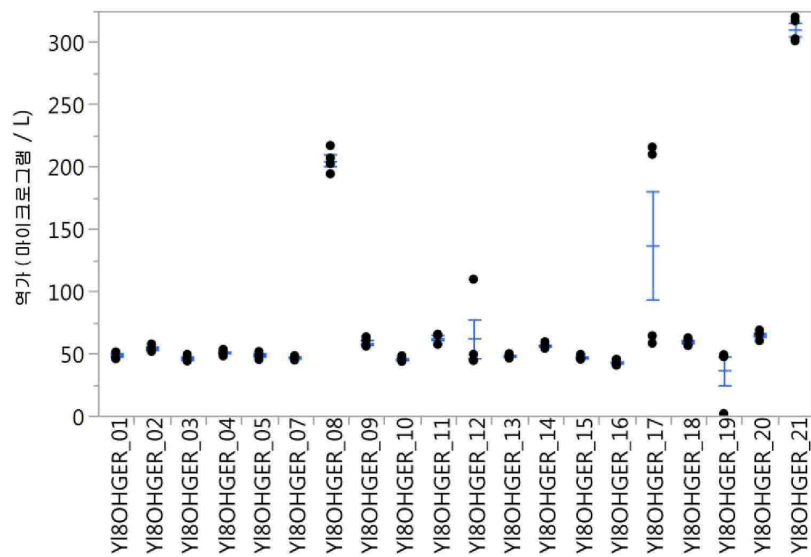
도면2



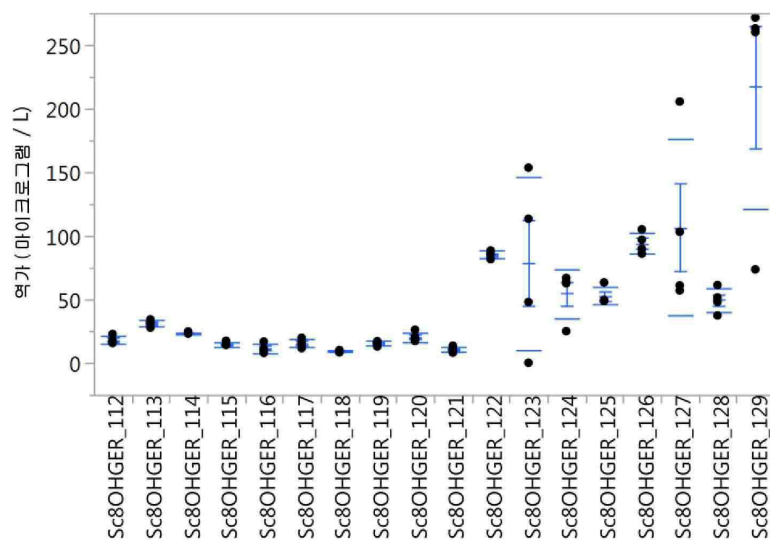
도면3



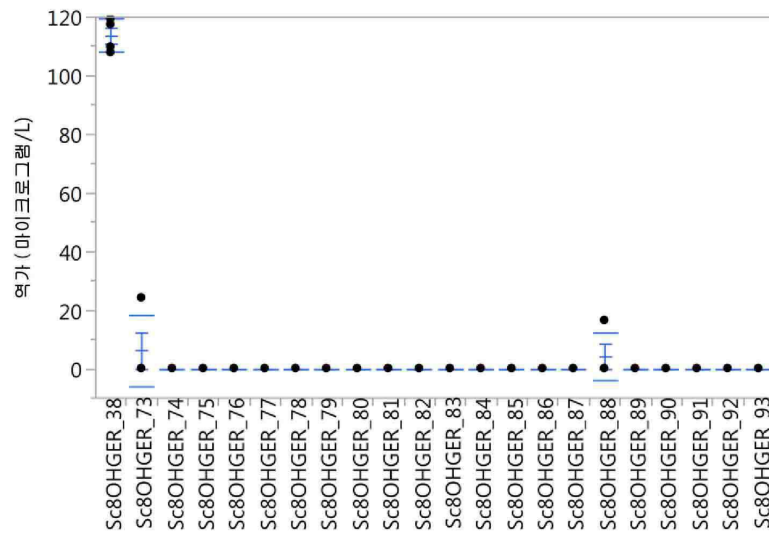
도면4



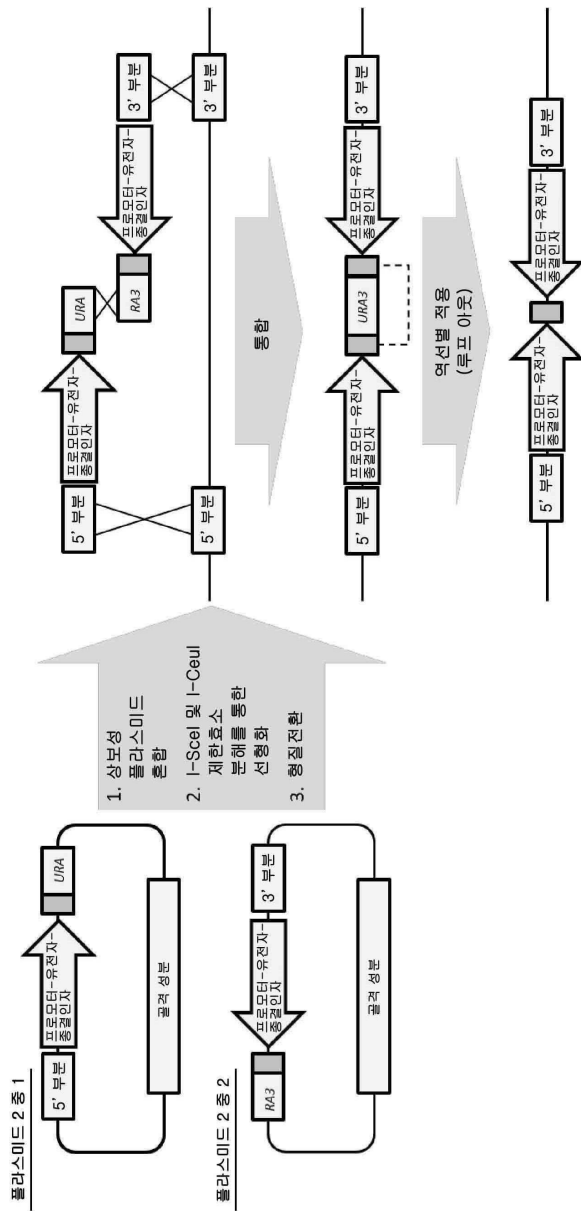
도면5



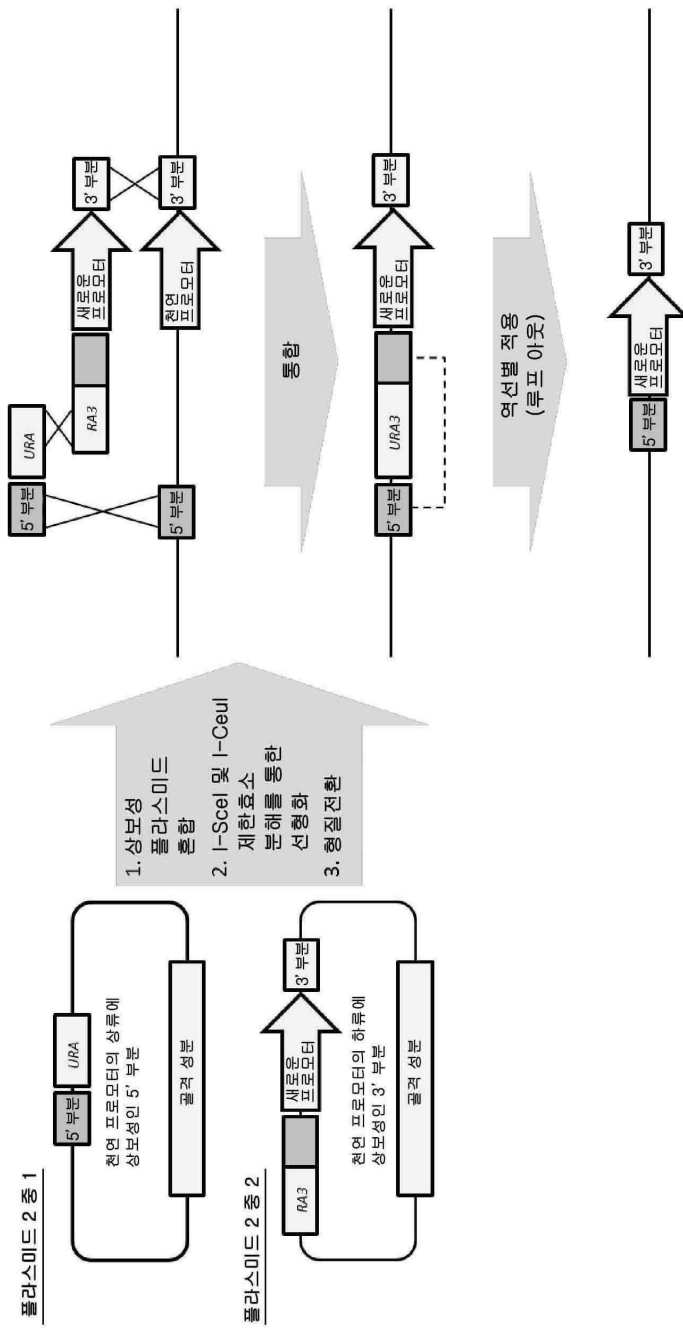
도면6



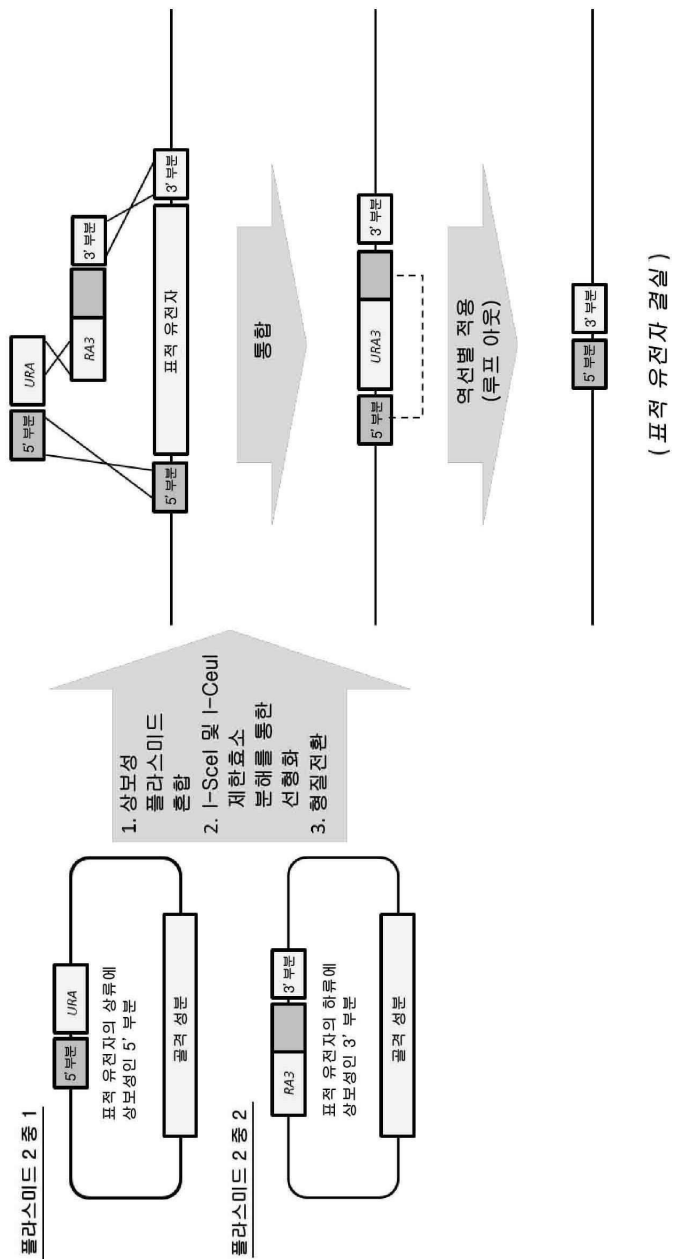
도면7



도면8



도면9



서열 목록

SEQUENCE LISTING

- <110> Zymergen Inc.
- <120> ENGINEERED BIOSYNTHETIC PATHWAYS FOR PRODUCTION OF
(6E)-8-HYDROXYGERANIOL BY FERMENTATION
- <130> ZMGNP007PUS/ZYM0027-WO-00
- <140> Not Assigned
- <141> 2018-12-06

<150> US 62/596,013

<151> 2017-12-07

<160> 112

<170> PatentIn version 3.5

<210> 1

<211> 591

<212> PRT

<213> Vitis vinifera (Grape)

<400> 1

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

1 5 10 15

Val Ala Pro Ser Leu Gln Val Ser Ser Ser Pro Cys Ser Trp Arg Thr

20 25 30

Arg Pro Ser Pro Cys Thr Ser Cys His Leu Ser Pro Ser Ser Ser Ser

35 40 45

Lys Pro Leu Leu Gly Ser His Asp Tyr Ser Leu Leu Lys Ser Leu Thr

50 55 60

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

65 70 75 80

Met Lys Glu Val Lys Glu Arg Thr Trp Glu Ala Phe Tyr Arg Ala Trp

85 90 95

Asp Ser Arg Ala Ala Met Glu Met Val Asp Thr Val Glu Arg Leu Gly

100 105 110

Leu Ser Tyr His Phe Glu Asp Glu Ile Asn Ala Leu Leu Gln Arg Phe

115 120 125

Cys Asp Trp Asn Ala Ser Glu Asp Leu Phe Thr Thr Ala Leu Arg Phe

130 135 140

Arg Leu Leu Arg Gln Asn Gly Phe Pro Thr His Ser Asp Val Phe Gly

145 150 155 160

Lys Phe Met Asp Lys Asn Gly Lys Phe Lys Glu Ser Leu Thr Glu Asp

165 170 175

Ile Trp Gly Met Leu Ser Leu His Glu Ala Ser His Leu Gly Ala Lys

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

Thr Gly Gly Thr Tyr Met Ala Leu Val His Ala Phe Phe Leu Met Gly

435 440 445

Gln Gly Val Thr Arg Glu Asn Met Ala Met Leu Lys Pro Tyr Pro Asn

450 455 460

Ile Phe Ser Cys Ser Gly Lys Ile Leu Arg Leu Trp Asp Asp Leu Gly

465 470 475 480

Thr Ala Arg Glu Glu Gln Glu Arg Gly Asp Asn Ala Ser Ser Ile Glu

485 490 495

Cys Tyr Lys Arg Glu Arg Glu Met Asp Thr Val Leu Glu Asp Glu Ala

500 505 510

Cys Arg Lys His Ile Arg Gln Met Ile Gln Ser Leu Trp Val Glu Leu

515 520 525

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

530 535 540

Ala Ala Phe Asn Leu Ser Arg Thr Ala Gln Val Ile Tyr Gln His Gly

545 550 555 560

Asp Asp Asn Lys Thr Ser Ser Val Glu Asp His Val Gln Ala Leu Leu

565 570 575

Phe Arg Pro Val Ser Ser Asn Gly His Ala Gln Ile Thr Met His

580 585 590

<210> 2

<211> 1773

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 2

atgtcgcgtt ttgtccat gccctctcac gtctgcctt catcattgt cgcccatcc	60
cttcaagttt ccagctgcc ctgcagctgg cgtaccgac cgtcgccctg caccagctgc	120
caccttcgc catcgtcgag ctctaaaccg ttgttgggta gccacgatta cagcctgttg	180
aagagcctga ccttttcgcc acacgcggtg aattccgaag cagactcttc gactcgccgt	240
atgaaagaag tgaagaacg aacctgggag gcattctacc gcgcttggga ttgcgcgcga	300

gcaatggaaa tggtagatac tgttgaacga ctgggacttt cttaccactt cgaagacgaa 360

atcaacgcac tcttccaacg cttttgtgac tggaacgcta gcgaggattt gttcacaact 420

gcgctgcgct ttcgcctgct tcgcaaaaac ggtttccga cccacagcga cgtcttcggc 480

aaatttatgg ataaaaatgg caaatcaaa gaatcactga ccgaagacat ttggggtatg 540

ttgtcccttc acgaggcgct tcaccttggg gctaaaaacg aagaagtcct tgcggaggct 600

aaagaattta ctgcaccca tctgattcag tcgatgccac acatggagcc tcatttttcg 660

tcgcacgtgg gccgagctct ggagctgccg cgccacctgc gtatggtcg tttggaagca 720

cgtaactata tcggcgagta ttgcgcgag tccaatccaa acctgcatt cctggagttg 780

gccaaattgg atttcgatat ggtgcagtct ttgcaccaga aagaattggc cgaaatcgtg 840

cgctggttga aacagctggg cctcgtggac aagttggatt tcgcacgtga ccgccaatg 900

gaatgtttcc tctggaccgt tggcatcttt ccggatcccc gacactctc gtgccgcatc 960

gaacttacca aagctattgc aattctgttg gtgatcgacg acatctatga ttcctacggc 1020

tcgctggatg aactcgtct ctttaccgat gccgtcaagc gctgggatct gggagccatg 1080

gatcaactcc cagaatacat gaagatttgt tacatggcgc tgtacaacac caccaatgat 1140

atcgcttacc gcaccccaa agagcatggt tggtcctca tcgaggactt gaaacgaacg 1200

tggatggata tctttggcgc attccttgct gaggcattat gcttcaaagg aggccatgtg 1260

ccttcgcttg aggaatatat gaccaacgca gttacgactg gcggtaccta tatggcactg 1320

gtgcacgct tctttcttat gggacagggc gtgaccgcg agaatatggc aatgctgaag 1380

ccctatccta atatcttttag ctgctcagga aagatcctcc gtctttggga tgatcttggc 1440

acagcgcgtg aggagcagga gcgtggcgac aacgccagct ccatcgaatg ctacaagcgt 1500

gaacgtgaga tggacaccgt cctcgaagat gaggcattgcc gcaaacacat ccgtcaaatg 1560

attcagtcct tgtgggttga attgaatggc gagcttgtgg cttcatctgc attgccgctg 1620

tccatcatca aggacgcgtt caatctctcc cgcaccgcac aagtatttta tcagcacggt 1680

gacgataaca aaacgtcctc tgttgaggat catgttcaag ccttgctttt ccgtcctgtt 1740

tctccaatg gtcacgcaca aattacatg cac 1773

<210> 3

<211> 292

<212> PRT

<213> Vitis vinifera (Grape)

<400> 3

Met Glu Cys Phe Leu Trp Thr Val Gly Ile Phe Pro Asp Pro Arg His

1	5	10	15
Ser Ser Cys Arg Ile Glu Leu Thr Lys Ala Ile Ala Ile Leu Leu Val			
	20	25	30
Ile Asp Asp Ile Tyr Asp Ser Tyr Gly Ser Leu Asp Glu Leu Ala Leu			
	35	40	45
Phe Thr Asp Ala Val Lys Arg Trp Asp Leu Gly Ala Met Asp Gln Leu			
	50	55	60
Pro Glu Tyr Met Lys Ile Cys Tyr Met Ala Leu Tyr Asn Thr Thr Asn			
65	70	75	80
Asp Ile Ala Tyr Arg Ile Leu Lys Glu His Gly Trp Ser Val Ile Glu			
	85	90	95
Asp Leu Lys Arg Thr Trp Met Asp Ile Phe Gly Ala Phe Leu Ala Glu			
	100	105	110
Ala Tyr Cys Phe Lys Gly Gly His Val Pro Ser Leu Glu Glu Tyr Leu			
	115	120	125
Thr Asn Ala Val Thr Thr Gly Gly Thr Tyr Met Ala Leu Val His Ala			
	130	135	140
Phe Phe Leu Met Gly Gln Gly Val Thr Arg Glu Asn Met Ala Met Leu			
145	150	155	160
Lys Pro Tyr Pro Asn Ile Phe Ser Cys Ser Gly Lys Ile Leu Arg Leu			
	165	170	175
Trp Asp Asp Leu Gly Thr Ala Arg Glu Glu Gln Glu Arg Gly Asp Asn			
	180	185	190
Ala Ser Ser Ile Glu Cys Tyr Lys Arg Glu Arg Glu Met Asp Thr Val			
	195	200	205
Leu Glu Asp Glu Ala Cys Arg Lys His Ile Arg Gln Met Ile Gln Ser			
	210	215	220
Leu Trp Val Glu Leu Asn Gly Glu Leu Val Ala Ser Ser Ala Leu Pro			
225	230	235	240
Leu Ser Ile Ile Lys Ala Ala Phe Asn Leu Ser Arg Thr Ala Gln Val			
	245	250	255

Ile Tyr Gln His Gly Asp Asp Asn Lys Thr Ser Ser Val Glu Asp His
 260 265 270
 Val Gln Ala Leu Leu Phe Arg Pro Val Ser Ser Asn Gly His Ala Gln
 275 280 285

Ile Thr Met His
 290

<210> 4

<211> 876

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 4

atggaatgtt ttctgtggac cggttggtatt ttcccagacc cgcgacactc ttctgtcgt 60

atcgaactta cgaaagcgat tgcaatcctt ttggtgattg atgatatcta cgactcgta 120

ggctcattgg acgaacttgc gctcttcacc gatgcagtaa aacgttgga ttgggcgcc 180

atggatcaat tgccggaata catgaagatc tgctatatgg cattgtacaa caccacgaac 240

gacatcgcat atcgatcct caaagagcac ggttggtctg taatcgaaga cctcaagcgc 300

acctggatgg atatcttcgg agcgttcctc gcagaagcgt attgcttcaa gggtggacac 360

gttccttccc tcgaagaata cctgacaaac gcggttacta ccggaggtag atacatggct 420

ctcgtgatg ctttctttct tatgggtcag ggtgttacc gcgaaaacat ggcatgctt 480

aaaccgatc caaacatctt ctctgttca ggcaagatcc tgcgcttggt ggatgatttg 540

ggaacagctc gcgaagagca ggagcgtggt gataacgcaa gctctatcga atgttacaaa 600

cgtgaacgcg aaatggacac agtcttgag gatgaagcct gccgcaagca cattcgccag 660

atgatccaat ccctctgggt tgaactgaac ggcaactcg ttgcaagctc cgcgctgccg 720

ctttcgatca ttaaggcggc cttcaatctc tcgcgcactg cacaggttat ttaccagcat 780

ggcgtatgaca ataaaacgtc ctccgtagaa gatcatgtgc aggctctgtt gtttcgcca 840

gtgagcagca acggacacgc acagatcacg atgcac 876

<210> 5

<211> 603

<212> PRT

<213> Perilla setoyensis

<400> 5

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

245	250	255	
Arg Ile Gln Arg Leu Glu Ala Arg Trp Phe Leu Asp Ala Tyr Ala Thr			
260	265	270	
Arg His Asp Met Asn Pro Ile Ile Leu Glu Leu Ala Lys Leu Asp Phe			
275	280	285	
Asn Ile Ile Gln Ala Thr His Gln Glu Glu Leu Lys Asp Val Ser Arg			
290	295	300	
Trp Trp Gln Asn Thr Arg Leu Ala Glu Lys Leu Pro Phe Val Arg Asp			
305	310	315	320
Arg Leu Val Glu Ser Tyr Phe Trp Ala Ile Ala Leu Phe Glu Pro His			
325	330	335	
Gln Tyr Gly Tyr Gln Arg Arg Val Ala Ala Lys Ile Ile Thr Leu Ala			
340	345	350	
Thr Ser Ile Asp Asp Val Tyr Asp Ile Tyr Gly Thr Leu Asp Glu Leu			
355	360	365	
Gln Leu Phe Thr Asp Asn Phe Arg Arg Trp Asp Thr Glu Ser Leu Gly			
370	375	380	
Gly Leu Pro Tyr Ser Met Gln Leu Phe Tyr Met Val Ile His Asn Phe			
385	390	395	400
Val Ser Glu Leu Ala Tyr Glu Ile Leu Lys Glu Lys Gly Phe Ile Ala			
405	410	415	
Ile Pro Tyr Leu Gln Arg Ser Trp Val Asp Leu Ala Glu Ser Phe Leu			
420	425	430	
Lys Glu Ala Asn Trp Tyr Tyr Ser Gly Tyr Thr Pro Ser Leu Glu Glu			
435	440	445	
Tyr Ile Asp Asn Gly Ser Ile Ser Ile Gly Ala Val Ala Val Leu Ser			
450	455	460	
Gln Val Tyr Phe Thr Leu Ala Asn Ser Ile Glu Lys Pro Lys Ile Glu			
465	470	475	480
Ser Met Tyr Lys Tyr His His Ile Leu Arg Leu Ser Gly Leu Leu Val			
485	490	495	

Arg Leu His Asp Asp Leu Gly Thr Ser Leu Phe Glu Lys Lys Arg Gly
 500 505 510
 Asp Val Pro Lys Ala Val Glu Ile Cys Met Lys Glu Arg Asn Asp Thr
 515 520 525
 Glu Glu Glu Ala Glu Glu His Val Lys Tyr Leu Ile Arg Glu Ala Trp
 530 535 540

Lys Glu Met Asn Thr Ala Thr Ala Ala Ala Gly Cys Pro Phe Met Asp
 545 550 555 560
 Glu Leu Asn Val Ala Ala Ala Asn Leu Gly Arg Ala Ala Gln Phe Val
 565 570 575
 Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Lys Ile His Gln
 580 585 590
 Gln Met Gly Gly Leu Met Phe Lys Pro Tyr Val
 595 600

<210> 6

<211> 1809

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 6

atgtgcagca tctctcaaaa ggtcgttatt ggattgaata aggcagcggc caataactgc	60
ctccaaaacc ttgatcgccg aggtttttaa acacgccgtg tttcactctc cgaagcggct	120
tcttgccctc gagcctcaag ctgcgtgcag ctgatgtca agccgggtga ggaaggccga	180
cgttccggta actatcagcc atcaatctgg gattttaatt atgtccagtc cctcaacacc	240
ccctacaaag aggaacgtta ccttaccgga cacgcggaac tcatcggtca ggtcaagccg	300
cttctcgaaa aaaagatgga ggctactcag cagttggagc tcatcgatga ccttaacaac	360
ctcgggtcga gctacttttt ccaagaccgc atcaagcaga ttctctcatt catctatgat	420
gaaaatcagt gtttccactc caatatcaat gatcaagctg aaaagcgaga cctgtacttt	480
accgcccttg gatttcgtct ccttcgccag cacggcttta acgtgtctca ggagggtgtc	540
gattgcttca aaaacgataa aggtccgat tttaaagcgt ccctgtcttg aaacaccaag	600
ggcctcctcc agctctacga agcctccttc ttggtccttg aggggtgaaga caccctcgaa	660
ctggcccgctc agttcgcgac caagttcctg cgccgcaagc ttgacgaaat cgacgacaac	720

catctgctct cacgaattca ccactcactc gaaatcccgt tgcactggcg catccagcgt 780

ctggaggcgc gctggtttct ggacgcgtac gctacacgcc acgatatgaa tcccattatc 840

ttggagctcg caaaactgga ctttaatat attcaagcaa cccaccaaga agagctcaag 900

gacgtgagcc gtgggtggca gaatacgcgc ctggccgaaa agctcccgtt tgtgcgcgat 960

cgacttgtcg agtcatactt ttgggccatc gcaacttttg agccacatca atatggatac 1020

cagcgccgtg ttgctgcca gattatcacc ctgcgcgat ccatcgatga cgtgtacgac 1080

atctacggca ctctcgacga gctccaactt ttaccgata acttccgtcg ctgggacacg 1140

gaatccttgg gcggtcttcc ttactcgatg cagttgttct acatggtgat tcacaacttt 1200

gtctctgagt tggcctatga aatcctgaaa gaaaaagggt ttatcgccat ccctacctc 1260

caacgctcct gggttgattt ggcagaatcc tttttgaaag aggcaactg gtactactct 1320

ggttatacac cctccctcga agaatacatc gacaacggct ccatctcaat tggcgagta 1380

gctgtgttgt cccaagtta tttactttg gcaaattcta tcgaaaagcc aaagattgaa 1440

tctatgtaca agtaccacca tattttgcgt ctgtcgggtt tgttggtgcg cctgcacgat 1500

gacttgggaa cttccttggt cgagaagaag cgcggcgatg ttcttaaggc ggtcgaaatt 1560

tgcataaag aacgcaacga tacggaggag gaagcagaag aacacgttaa gtacttgatt 1620

cgcaagcct ggaaggagat gaacactgct acggcagctg caggatgtcc ttttatggat 1680

gagctcaacg tggctgctgc caatctgggt cgcgcagccc agttcgtgta cctcgacgga 1740

gacggtcatg gcgtgcagca ctccaaaatc catcaacaga tgggtggcct tatgtttaaa 1800

ccgtatgtg 1809

<210> 7

<211> 495

<212> PRT

<213> Swertia mussotii (Felwort)

<400> 7

Met Asp Phe Asp Phe Leu Thr Ile Ala Ile Gly Phe Leu Phe Thr Ile

1 5 10 15

Thr Leu Tyr Gln Ala Leu Asn Phe Phe Ser Arg Lys Ser Lys Asn Leu

20 25 30

Pro Pro Gly Pro Ser Pro Leu Pro Leu Ile Gly Asn Leu His Leu Leu

35 40 45

Gly Asp Gln Pro His Lys Ser Leu Ala Lys Leu Ala Lys Lys His Gly

50					55					60									
Pro	Ile	Met	Gly	Leu	Gln	Leu	Gly	Gln	Val	Thr	Thr	Ile	Val	Val	Thr				
65					70					75					80				
Ser	Ser	Gly	Met	Ala	Lys	Glu	Val	Leu	Gln	Lys	Gln	Asp	Leu	Ala	Phe				
85					90					95									
Ser	Ser	Arg	Ser	Ile	Pro	Asn	Ala	Ile	His	Ala	His	Asp	Gln	Tyr	Lys				
100					105					110									
Tyr	Ser	Val	Ile	Trp	Leu	Pro	Val	Ala	Ser	Arg	Trp	Arg	Gly	Leu	Arg				
115					120					125									
Lys	Ala	Leu	Asn	Ser	Asn	Met	Phe	Ser	Gly	Asn	Arg	Leu	Asp	Ala	Asn				
130					135					140									
Gln	His	Leu	Arg	Ser	Arg	Lys	Val	Gln	Glu	Leu	Ile	Ala	Tyr	Cys	Arg				
145					150					155					160				
Lys	Ser	Ser	Gln	Thr	Gly	Asp	Ala	Ile	Asp	Val	Gly	Arg	Ala	Ala	Phe				
165					170					175									
Arg	Thr	Ser	Leu	Asn	Leu	Leu	Ser	Asn	Thr	Met	Phe	Ser	Lys	Asp	Leu				
180					185					190									
Thr	Asp	Pro	Tyr	Ser	Asp	Ser	Ala	Lys	Glu	Phe	Lys	Asp	Leu	Val	Trp				
195					200					205									
Asn	Val	Met	Val	Glu	Ala	Gly	Lys	Pro	Asn	Leu	Val	Asp	Tyr	Phe	Pro				
210					215					220									
Leu	Leu	Asp	Lys	Val	Asp	Pro	Gln	Gly	Ile	Arg	Lys	Arg	Met	Thr	Ile				
225					230					235					240				
His	Phe	Gly	Lys	Ile	Leu	Glu	Leu	Phe	Gly	Gly	Leu	Ile	Asp	Glu	Arg				
245					250					255									
Leu	Gln	Gln	Lys	Lys	Ala	Lys	Gly	Val	Asn	Asp	Asp	Val	Leu	Asp	Val				
260					265					270									
Leu	Leu	Thr	Thr	Ser	Glu	Glu	Ser	Pro	Glu	Glu	Ile	Asp	Arg	Thr	His				
275					280					285									
Ile	Gln	Arg	Met	Cys	Leu	Asp	Leu	Phe	Val	Ala	Gly	Thr	Asp	Thr	Thr				
290					295					300									

Ser Ser Thr Leu Glu Trp Ala Met Ser Glu Met Leu Lys Asn Pro Glu
 305 310 315 320
 Lys Met Lys Ala Ala Gln Ala Glu Leu Ala Gln Val Ile Gly Lys Gly
 325 330 335

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

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

Gly Gly Ile Ala Pro Lys Asp Leu Asp Met Glu Glu Lys Phe Gly Ile
 465 470 475 480
 Thr Leu Gln Lys Ala His Pro Leu Arg Ala Val Ala Thr Pro Leu
 485 490 495

<210> 8

<211> 1486

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 8

atggatttcg acttccttac tatcgcgatt gggtttctgt ttactatcac ttgtaccaa

60

gcgcttaact tcttttcccg taaatcaaag aatctccac caggtcgctc cccattgcc 120

ctcattggta acttgcatct gctcggtagac caacccata agtcactggc caaactcgca 180

aagaaacatg gtcctatcat gggactgcaa ctgggtcaag tcaccacat cgttgtaacg 240

tcttctggta tggcgaagga ggttctccag aagcaggacc tggccttttc aagccgttcc 300

atcccaaagc caatccacgc ccatgaccaa tataagtata gcgtgatttg gctcccggtg 360

gcgtcccgat ggcgaggctc gcgtaaggcg ctcaactcca atatgttctc tggaaatcg 420

ctggatgcaa atcagcactt gcgttcacga aaggtccagg aactgatcgc ctattgccgc 480

aagtcatctc aaactggcga tgcattgat gttggtcgcg cagcattccg cacttcattg 540

aacctgctta gcaacacat gttcagcaag gaccttacag acccttattc tgactccgca 600

aaagaattta aggatctcgt ttggaacgtg atggtggagg caggcaaacc caatttggtt 660

gactacttcc ctctgttga taaagttgac cccaaggca ttcgaaagcg catgaccatt 720

catttcggta aaatcctcga attgtttggc ggacttatcg acgaacgatt gcagcaaaaa 780

aaagcaaagg gagtgaacga cgacgtgctg gacgtgctcc ttaccacttc agaggaaagc 840

ccggaggaaa tcgaccgtac ccatatccag cgcattgtcc tggacttggt cgtcgcgggc 900

acggacacta cgtcgtcgac gctcgaatgg gccatgtctg aaatgcttaa gaaccccgaa 960

aaaatgaagg cggcccaggc agaactggcc caggttattg gaaaaggcaa ggctgttgaa 1020

gaagcagact tggcgcgcct gccatacttg cgtcgcgcta ttaagagac ctttcgaatc 1080

catctccgg ttccactgct gattcctcga cgcacggaac aagaggtcga ggtgtgcggt 1140

tatacggttc cgaagaactc acaggttttg gtgaacgttt gggccatctc tcgagacgat 1200

gcaatttga aggatccact ctctttaag cctgaacgtt tcttgagtc tgaattggag 1260

atgcgttgga aggatttcga gcttatcca ttcggtgcgg gccgacgat ttgccgggc 1320

cttctctgg cggtagcgtat ggtccagtg atgcttggat ctcttttgaa ctcatcgat 1380

tggaagctgg agggcggcat cgcgctaag gacttgaca tggaagaaaa gttcggcatc 1440

acgtgcaga aagccaccc tctcgcgct gtcgcaactc cactgt 1486

<210> 9

<211> 505

<212> PRT

<213> Populus trichocarpa (Western balsam poplar) (Populus balsamifera subsp. trichocarpa)

<400> 9

Met Asn Phe Phe Ile Ser Val Leu Leu Tyr Phe Leu Leu Thr Phe Ala

1 5 10 15

Val Ile Gln Ser Leu Asp Tyr Ile Leu Arg Arg Ser Lys Arg Lys Ser
20 25 30

Gly Lys Leu Pro Pro Gly Pro Ser Arg Leu Pro Ile Val Gly Asn Leu
35 40 45

Leu Asp Leu Gly Asp Lys Pro His Lys Ser Leu Ala Lys Leu Ala Lys
50 55 60

Thr His Gly Gln Leu Met Ser Leu Lys Leu Gly Gln Val Thr Thr Ile
65 70 75 80

Val Val Ser Ser Ala Thr Met Ala Lys Glu Val Leu Gln Lys His Asp
85 90 95

Leu Thr Phe Cys Asn Arg Thr Val Val Asp Ala Val Arg Ala Leu Asp
100 105 110

His His Glu Ala Gly Ile Ala Trp Leu Pro Val Ala Thr Arg Trp Arg
115 120 125

Asn Leu Arg Lys Ile Cys Asn Ser His Ile Phe Thr Ser Gln Lys Leu
130 135 140

Asp Ala Asn Gln Asp Leu Arg Arg Lys Lys Val Gln Asp Leu Leu Ala
145 150 155 160

Glu Val Gln Glu Arg Cys Leu Val Gly Glu Ala Val Asp Leu Arg Gln
165 170 175

Ala Ala Phe Thr Ala Thr Leu Asn Ala Leu Ser Asn Thr Val Leu Ser
180 185 190

Leu Asp Leu Thr Asp Leu Ser Ser Asp Ile Ala Arg Glu Phe Lys Glu
195 200 205

His Ile Ser Cys Ile Met Asp Glu Ala Gly Lys Pro Asn Leu Val Asp
210 215 220

Tyr Phe Pro Leu Leu Arg Arg Ile Asp Pro Gln Gly Ile Arg Arg Arg
225 230 235 240

Thr Ala Ile His Phe Gly Lys Val Phe Asp Leu Phe Asp Arg Leu Ile
245 250 255

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

260	265	270
Asp Met Leu Asp Thr Leu Leu Thr Ile Ser Glu Val Asn Asn Glu Glu		
275	280	285
Met Asp Ala Thr Arg Ile Lys His Phe Phe Leu Asp Leu Phe Gly Ala		
290	295	300
Gly Thr Asp Thr Thr Ser Ser Thr Leu Glu Trp Ala Met Ala Glu Leu		
305	310	315
Leu His Ser Pro Lys Thr Leu Leu Lys Ala Arg Ala Glu Leu Glu Arg		
325	330	335
Thr Ile Gly Glu Gly Asn Leu Leu Glu Glu Ser Asp Ile Thr Arg Leu		
340	345	350
Pro Tyr Leu Gln Ala Val Ile Lys Glu Thr Leu Arg Leu His Pro Ala		
355	360	365
Val Pro Phe Leu Leu Pro His Lys Ala Gly Ala Asp Ala Glu Ile Gly		
370	375	380
Gly Phe Thr Val Pro Lys Asn Ala Gln Val Leu Val Asn Val Trp Ala		
385	390	395
Ile Gly Arg Asp Pro Ser Met Trp Glu Asp Pro Asn Ser Phe Val Pro		
405	410	415
Glu Arg Phe Leu Glu Ser Gly Ile Asp His Arg Gly Gln Asn Phe Glu		
420	425	430
Phe Ile Pro Phe Gly Ser Gly Arg Arg Ile Cys Pro Gly Leu Pro Leu		
435	440	445
Ala Met Arg Met Leu Pro Leu Met Leu Gly Ser Leu Ile Leu Ser Phe		
450	455	460
Asp Trp Lys Leu Ala Asp Gly Val Thr Pro Glu Asn Leu Asn Met Asp		
465	470	475
Asp Lys Phe Gly Leu Thr Leu Leu Lys Ala Gln Pro Leu Arg Ala Ile		
485	490	495
Pro Ile Thr Arg Glu Leu Lys His Gly		
500	505	

<210> 10

<211> 1515

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 10

```

atgaactttt tcactctgt gctgtgtat ttctttctga cattcgcggt gattcaatcc      60
ctggactaca tccttcgccc ctccaagcgt aagtcggca aactgccgcc aggccttca      120
cgacttccga ttgttgaaa cctgctggac ctgggagaca agccccaca atccctcgca      180

aagctggcga aaaccacgg ccaattgat tcaactaagc tcggacaagt gaccacgac      240
gtggtttcaa gcctaccat ggccaaggag gtgttcgaga agcatgatct gactttctgt      300
aaccgcactg ttgtgacgc tgttcgcgcg ctggaccacc atgaggccgg cattgcgtgg      360
ctgccagtgg ctaccggtg gcgcaacctg cgtaaaattt gcaatagcca tatctttaca      420
tccagaaaac tcgacgctaa tcaagacctg cgacgcaaaa aagtcagga tctgctggca      480
gaagtgaag agcgctgcct ggtaggtgag gctgttgact tgcgtcaagc agcatttacc      540
gcgactctta atgccctgag caacactgtt ttgtcgtcg acctaccga tctgtcatca      600

gatatcgctc gcgagtttaa agagcacatc tctgcatta tggatgaagc gggcaagcca      660
aacctggttg attactttcc ttgtcttcgc cgcacgacc cgcagggcat tcgtcgccgc      720
actgcaattc acttcggcaa ggtgttcgat ctcttcgac gcttgattat cgaacgtttg      780
cagctccgca aagtcaaggg ttacattccc ctggacgaca tgcctggacac ttgtctcacg      840
atctcagaag tcaacaacga agagatggac gctacacgta tcaagcatt ttttttgat      900
ttgttcggcg caggcaccga tacgacttca tcgactcttg agtgggcat ggcagagctc      960
ttgactcgc cgaaaacct cctcaaagcc cgcgctgaac tggaacgcac catcggcgag      1020

ggcaatcttt tggaggaatc cgacatcacc cgtctccct acctgcagc tgtcatcaag      1080
gagactctcc gactccatcc agccgtgcc ttctctctc cacacaaggc aggcgtgat      1140
gctgaaatcg gcggttttac tgtccaaaa aacgcgcagg tcttggtgaa tgtctgggcg      1200
attggccgtg atccctcaat gtgggaagac ccaattcct tcgtccctga acgttcttg      1260
gaatcaggta tcgatcaccg aggacaaaac ttgaattta tccctttcgg cagcggacgt      1320
cgaatttgcc caggctctcc gctggcgatg cgtatgctcc ctctcatgct cggttctctc      1380
atcttgctct tcgactggaa gttggcagat ggtgtgacgc cagaaaactt gaacatggac      1440

gataagttcg gtttgactct cctgaaagct cagccccttc gtgcgatccc catcacccga      1500

```

gaattgaaac acggt

1515

<210> 11

<211> 501

<212> PRT

<213> Phaseolus angularis (Azuki bean) (Vigna angularis)

<400> 11

Met Glu Thr Asn Ser Thr Leu Phe Phe Thr Phe Thr Val Leu Leu Ile

1 5 10 15

Thr Ile Ile Thr Phe Leu Lys Ala Leu Lys Ser Leu Phe Thr Pro Asn

20 25 30

Lys Ser Lys Ser Asn Leu Pro Pro Gly Pro Lys Gly Leu Pro Leu Val

35 40 45

Gly Asn Leu Leu Gln Leu Gly Ala Lys Pro His Gln Thr Leu Ala Thr

50 55 60

Leu Ala Asn Ile His Gly Ser Ile Met Ser Leu Lys Leu Gly Gln Glu

65 70 75 80

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

85 90 95

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

100 105 110

Gly Ser Ser His Asp His Phe Ser Leu Pro Phe Met Pro Val Ser Gln

115 120 125

Gln Trp Arg Glu Leu Arg Lys Leu Cys Asn Glu Leu Leu Phe Ser Asn

130 135 140

Lys Asn Leu Asp Ala Thr Gln Gly Leu Arg Ser Lys Lys Val Arg Glu

145 150 155 160

Leu Tyr Ser Asp Ile His Arg Ser Ser Leu Lys Gly Glu Pro Val Asn

165 170 175

Ile Gly Arg Leu Ala Phe Lys Thr Thr Ile Asn Gln Leu Ser Asn Thr

180 185 190

Ile Tyr Ser Glu Asp Phe Leu Gln Ser Ala Glu Lys Ala Gly Glu Met

195 200 205

Lys Glu Leu Val Thr Asn Ile Met Lys Glu Val Gly Arg Pro Asn Leu
 210 215 220

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

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

 Pro Tyr Leu Gln Ala Ile Val Lys Glu Thr Phe Arg Leu His Pro Ala
 355 360 365
 Val Pro Leu Leu Leu Pro Arg Lys Ala Glu Val Glu Phe Glu Met His
 370 375 380
 Gly Tyr Thr Ile Pro Lys Gly Ala Gln Val Leu Ile Asn Val Trp Ala
 385 390 395 400
 Ile Gly Arg Asp Pro Asn Leu Trp Glu Lys Pro Arg Leu Phe Trp Pro
 405 410 415

 Glu Arg Phe Leu Glu Ser Glu Ile Asp Phe Lys Gly Arg Ser Phe Glu
 420 425 430
 Leu Thr Pro Phe Gly Gly Gly Arg Arg Ile Cys Pro Gly Leu Pro Leu
 435 440 445
 Ala Ile Arg Leu Val Phe Leu Met Leu Gly Leu Phe Ile Asn Ser Phe

450	455	460	
Asp Trp Glu Leu Gln Asp Ile Gln Pro Glu Asp Met Asn Met Asp Glu			
465	470	475	480
Asn Phe Gly Leu Thr Leu Glu Lys Ala Gln Pro Val Leu Ala Ile Pro			
485	490	495	
Ile Ile Pro Lys His			
500			
<210>	12		
<211>	1503		
<212>	DNA		
<213>	Artificial		
<220><223>	mutated and/or codon-optimized sequence		
<400>	12		
atggagacta	actctacctt	tttctttaca	ttcacctgcc
tctctgatcac	aatcattacg		60
ttcctgaagg	cgttgaaatc	tttgtttacc	cctaacaagt
caaaatctaa	cctccccacc		120
ggtcccaagg	gccttccitt	ggttggcaac	ctgcttcagt
tgggcgcgaa	gccgcaccag		180
accctcgcaa	ccctcgcaaa	catccacgga	tccatcatgt
ccctcaaatt	gggtcaagag		240
acaaccatcg	tgatgtcttc	cgctgaggcc	gcaaaggag
ttctgcagat	tcatgatcat		300
ttctctcaa	accgcaaaat	ccccgacgca	atgcgcggct
cctcccacga	ccacttctcc		360
cttcttttca	tgctgttttc	ccaacagtgg	cgagaattgc
gaaagctttg	caacgaactc		420
ctgttctcca	acaagaacct	ggacgctaca	caaggcctgc
gctccaaaaa	ggttcgagaa		480
ctgtacagcg	atattcaccg	cagctcgctg	aagggcgagc
ccgtgaacat	cggtcgtctg		540
gcttttaaaa	cgaccatcaa	ccaactttct	aacaccatct
actcggaaga	cttcttgcaa		600
tccgccgaga	aagcgggcga	aatgaaggaa	ctcgttacca
atattatgaa	ggaggttggc		660
cgcccgaacc	ttgccgactg	cttccccgtt	ctgaaaatga
tcgaccccca	cggcatccgc		720
cgtcgaaccg	gctcttactt	cagcaagctg	cttaatatct
tcaaatccct	gatccacaag		780
cgtctggaac	ttcgtaaaga	tgctgcgggc	tactgtacta
aaaaggatat	gcttgacgcg		840
atgctgaacg	acgcccaca	caaaatggac	atcgттаага
tccagcgctc	ctctcttgat		900
ttgttcgttg	caggaacgga	taccgtcacc	tccactgtcg
agtgggctat	ggctgaactg		960
cttcacaacc	cccacgttat	gtcaaaggct	aaagaggaa
tcgaacgc	tatcggtaa		1020
gataatctgg	tggaggagtc	agacatcgcc	aagttgcctt
acttgcaggc	tatcggtgaag		1080

gagaccttc gtctccacc agcagtcctt cttttgtgc cagcaaagc cgaggttgaa 1140
 ttgaaatgc acggttacac aatccccaaa ggcgccagg tcctgatcaa tgtctgggcc 1200
 attggtcgcg acccgaactt gtgggaaaag ccacgcctgt tttggcccga gcgatttctt 1260
 gaatccgaga tcgacttcaa gggccgatcg ttcgagctca ccccgttcgg cggcggccgt 1320
 cgcactgcc ctggactccc attggcgatc cgacttggtt ttctgatgct cggcctgttt 1380
 attaactctt tcgattggga acttcaggac attcagccgg aagacatgaa tatggacgaa 1440

aacttcggtc tgactcttga aaaggcccag ccagttctcg caattccaat cattccaaaa 1500
 cac 1503

<210> 13

<211> 443

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 13

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

1 5 10 15

Gly Glu His Ser Ala Val Tyr Asn Lys Pro Ala Val Ala Ala Ser Val

20 25 30

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

35 40 45

Thr Ile Glu Leu Asp Phe Pro Asp Ile Ser Phe Asn His Lys Trp Ser

50 55 60

Ile Asn Asp Phe Asn Ala Ile Thr Glu Asp Gln Val Asn Ser Gln Lys

65 70 75 80

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

85 90 95

Ser Leu Leu Asp Pro Leu Leu Ala Gln Leu Ser Glu Ser Phe His Tyr

100 105 110

His Ala Ala Phe Cys Phe Leu Tyr Met Phe Val Cys Leu Cys Pro His

115 120 125

Ala Lys Asn Ile Lys Phe Ser Leu Lys Ser Thr Leu Pro Ile Gly Ala

130 135 140

Gly Leu Gly Ser Ser Ala Ser Ile Ser Val Ser Leu Ala Leu Ala Met

145 150 155 160
 Ala Tyr Leu Gly Gly Leu Ile Gly Ser Asn Asp Leu Glu Lys Leu Ser

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

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

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

 355 360 365
 Glu Gln Ile Asp Ser Phe Lys Lys Lys Leu Gln Asp Asp Phe Ser Tyr
 370 375 380
 Glu Thr Phe Glu Thr Asp Leu Gly Gly Thr Gly Cys Cys Leu Leu Ser
 385 390 395 400

Ala Lys Asn Leu Asn Lys Asp Leu Lys Ile Lys Ser Leu Val Phe Gln

405

410

415

Leu Phe Glu Asn Lys Thr Thr Thr Lys Gln Gln Ile Asp Asp Leu Leu

420

425

430

Leu Pro Gly Asn Thr Asn Leu Pro Trp Thr Ser

435

440

<210> 14

<211> 1329

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 14

atgtccctgc cattcctcac gtccgcacca ggaaaagtca ttatctttgg cgaacattcc 60

gcagtgtaca acaagcctgc cgtggctgcc tccgtgagcg cctccgcac ctacttgctg 120

attagcgaga gctctgcccc agataccatc gagctggatt ttccagacat ctctttaac 180

cataaatggt ccattaacga ttttaacgcg attaccgaag atcaagtcaa cagccagaag 240

ctcgcaaagg cacaacagcg aacagatggc ctctcccagg agcttgtttc cctcctggat 300

ccactgcttg cacagctctc agaatccttt cactaccatg ccgctttctg tttcctgtac 360

atgttcgtgt gtctctgccc ccacgcgaag aacatcaagt tcagcctgaa gtcgacactc 420

cctatcgggtg ctggtctcgg atcgtccgcg agcatctcgg tctcgttggc acttgcaatg 480

gcatacctcg gtggtcttat tgggtcaaac gatctggaaa agttgtcgga aaacgataaa 540

catattgtga accagtgggc attcatcggc gaaaaatgta tccacggcac cccagcgga 600

atcgataacg cggttgcaac ctacggaaat gctttgctgt tcgagaagga ttcccacaat 660

ggaacgatta acacgaataa ctttaagtgc ctgcacgatt tccccgaat tcccatgac 720

ctcacctaca cccgtatccc tcgttcgacc aaagacttgg tcgctcgcgt acgcgtgttg 780

gtcacggaaa agtttcaga ggtcatgaaa cctattcttg atgcatggg tgagtgcgcc 840

cttcagggcc tggagattat gactaagctg tccaagtga agggtagcga cgacgaggca 900

gtcgaaacta acaatgaact ctacgagcag cttctggaat tgatccgcat caaccacggt 960

ctgctcgtgt ccatcggcgt gtctcaccca ggactggaac tgattaaaaa cctctcggat 1020

gacctccgta tcggctctac caagctgacc ggccgccggag gcggtggatg ctctctcacc 1080

ttgctccgac gcgatatcac ccaggagcag atcgattctt tcaaaaagaa actccaagat 1140

gacttctctt acgagacttt tgagactgac ctcggtggca ccggtgctg cttgctttcc 1200
gcgaagaact tgaataaaga tctcaaaatc aaatctctcg tcttcagct cttcgaaaac 1260
aagaccacca ccaagcaaca gatcgatgac ttgctccttc caggcaacac taaccttctc 1320
tggaacaagc 1329

<210> 15

<211> 352

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 15

Met Ala Ser Glu Lys Glu Ile Arg Arg Glu Arg Phe Leu Asn Val Phe

1 5 10 15

Pro Lys Leu Val Glu Glu Leu Asn Ala Ser Leu Leu Ala Tyr Gly Met

20 25 30

Pro Lys Glu Ala Cys Asp Trp Tyr Ala His Ser Leu Asn Tyr Asn Thr

35 40 45

Pro Gly Gly Lys Leu Asn Arg Gly Leu Ser Val Val Asp Thr Tyr Ala

50 55 60

Ile Leu Ser Asn Lys Thr Val Glu Gln Leu Gly Gln Glu Glu Tyr Glu

65 70 75 80

Lys Val Ala Ile Leu Gly Trp Cys Ile Glu Leu Leu Gln Ala Tyr Trp

85 90 95

Leu Val Ala Asp Asp Met Met Asp Lys Ser Ile Thr Arg Arg Gly Gln

100 105 110

Pro Cys Trp Tyr Lys Val Pro Glu Val Gly Glu Ile Ala Ile Trp Asp

115 120 125

Ala Phe Met Leu Glu Ala Ala Ile Tyr Lys Leu Leu Lys Ser His Phe

130 135 140

Arg Asn Glu Lys Tyr Tyr Ile Asp Ile Thr Glu Leu Phe His Glu Val

145 150 155 160

Thr Phe Gln Thr Glu Leu Gly Gln Leu Met Asp Leu Ile Thr Ala Pro

165 170 175

Glu Asp Lys Val Asp Leu Ser Lys Phe Ser Leu Lys Lys His Ser Phe

180 185 190

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

195 200 205

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

210 215 220

Ala Arg Asp Val Leu Ile Pro Leu Gly Glu Tyr Phe Gln Ile Gln Asp

225 230 235 240

Asp Tyr Leu Asp Cys Phe Gly Thr Pro Glu Gln Ile Gly Lys Ile Gly

245 250 255

Thr Asp Ile Gln Asp Asn Lys Cys Ser Trp Val Ile Asn Lys Ala Leu

260 265 270

Glu Leu Ala Ser Ala Glu Gln Arg Lys Thr Leu Asp Glu Asn Tyr Gly

275 280 285

Lys Lys Asp Ser Val Ala Glu Ala Lys Cys Lys Lys Ile Phe Asn Asp

290 295 300

Leu Lys Ile Glu Gln Leu Tyr His Glu Tyr Glu Glu Ser Ile Ala Lys

305 310 315 320

Asp Leu Lys Ala Lys Ile Ser Gln Val Asp Glu Ser Arg Gly Phe Lys

325 330 335

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

340 345 350

<210> 16

<211> 1056

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 16

atggcatccg agaaggaaat ccgtcgtgaa cgattcctga acgtattccc aaagctcgtc 60

gaagagctta atgcttcctt gcttgcgtac ggaatgccga aggaagcgtg cgactggtat 120

gcgcattctc tcaactataa cacacctggc ggtaagctga atcgcggtct cagcgtggtg 180

gacacgtacg caatcctcag caataagacc tggaacagc tgggacagga ggagtacgag 240

aaagttgcta tccttggatg gtgtatcgaa ctgctgcaag cctattgggtt ggtagctgat 300
gacatgatgg ataagagcat cactcgacgt ggtcaacat gctgtataaa ggtgccggag 360
gttggcgaaa tcgcgatctg ggatgcgttc atgttggagg cgcgaattta taagcttctc 420
aagtccact ttcgtaacga aaagtactat attgacatca cagaactctt tcatgaagtg 480
accttccaaa ccgagctcgg ccaactcatg gatttgatca ccgccccga agataaggtg 540

gacctttcga agttctccct caagaagcac tcgttcattg ttaccttcaa gacagcatac 600
tactcathtt acctgccagt cgccctcgca atgtacgtgg cgggcatcac tgatgaaaag 660
gatctgaagc aggccccga cgtccttatt cctcttggag agtacttcca gatccaggac 720
gattacttgg attgttttgg caccacagag cagattggta aaatcggtac tgatatccaa 780
gacaacaaat gttcctgggt aattaacaag gctctcgagt tggcttccgc ggaacaacgc 840
aagactcttg acgaaaatta cggtaagaaa gattcagtcg ctgaagccaa gtgcaagaaa 900
atctttaacg atcttaaaat cgaacagttg taccacgagt atgaagagtc tatcgctaaa 960

gacctgaaag ctaaaatttc ccaggtcgat gagtcacgcg gcttcaaagc ggatgtgctc 1020
accgccttcc ttaacaaggt gtataagcgc tcaaaa 1056

<210> 17

<211> 491

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 17

Met Lys Leu Ser Thr Lys Leu Cys Trp Cys Gly Ile Lys Gly Arg Leu

1 5 10 15

Arg Pro Gln Lys Gln Gln Gln Leu His Asn Thr Asn Leu Gln Met Thr

20 25 30

Glu Leu Lys Lys Gln Lys Thr Ala Glu Gln Lys Thr Arg Pro Gln Asn

35 40 45

Val Gly Ile Lys Gly Ile Gln Ile Tyr Ile Pro Thr Gln Cys Val Asn

50 55 60

Gln Ser Glu Leu Glu Lys Phe Asp Gly Val Ser Gln Gly Lys Tyr Thr

65 70 75 80

Ile Gly Leu Gly Gln Thr Asn Met Ser Phe Val Asn Asp Arg Glu Asp

85 90 95

Ile Tyr Ser Met Ser Leu Thr Val Leu Ser Lys Leu Ile Lys Ser Tyr

100 105 110

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

115 120 125

Leu Ile Asp Lys Ser Lys Ser Val Lys Ser Val Leu Met Gln Leu Phe

130 135 140

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

145 150 155 160

Gly Gly Thr Asn Ala Leu Phe Asn Ser Leu Asn Trp Ile Glu Ser Asn

165 170 175

Ala Trp Asp Gly Arg Asp Ala Ile Val Val Cys Gly Asp Ile Ala Ile

180 185 190

Tyr Asp Lys Gly Ala Ala Arg Pro Thr Gly Gly Ala Gly Thr Val Ala

195 200 205

Met Trp Ile Gly Pro Asp Ala Pro Ile Val Phe Asp Ser Val Arg Ala

210 215 220

Ser Tyr Met Glu His Ala Tyr Asp Phe Tyr Lys Pro Asp Phe Thr Ser

225 230 235 240

Glu Tyr Pro Tyr Val Asp Gly His Phe Ser Leu Thr Cys Tyr Val Lys

245 250 255

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

260 265 270

Gly Leu Val Ser Asp Pro Ala Gly Ser Asp Ala Leu Asn Val Leu Lys

275 280 285

Tyr Phe Asp Tyr Asn Val Phe His Val Pro Thr Cys Lys Leu Val Thr

290 295 300

Lys Ser Tyr Gly Arg Leu Leu Tyr Asn Asp Phe Arg Ala Asn Pro Gln

305 310 315 320

Leu Phe Pro Glu Val Asp Ala Glu Leu Ala Thr Arg Asp Tyr Asp Glu

325 330 335

Ser Leu Thr Asp Lys Asn Ile Glu Lys Thr Phe Val Asn Val Ala Lys

340 345 350
Pro Phe His Lys Glu Arg Val Ala Gln Ser Leu Ile Val Pro Thr Asn

355 360 365
Thr Gly Asn Met Tyr Thr Ala Ser Val Tyr Ala Ala Phe Ala Ser Leu

370 375 380
Leu Asn Tyr Val Gly Ser Asp Asp Leu Gln Gly Lys Arg Val Gly Leu

385 390 395 400
Phe Ser Tyr Gly Ser Gly Leu Ala Ala Ser Leu Tyr Ser Cys Lys Ile

405 410 415
Val Gly Asp Val Gln His Ile Ile Lys Glu Leu Asp Ile Thr Asn Lys

420 425 430
Leu Ala Lys Arg Ile Thr Glu Thr Pro Lys Asp Tyr Glu Ala Ala Ile

435 440 445
Glu Leu Arg Glu Asn Ala His Leu Lys Lys Asn Phe Lys Pro Gln Gly

450 455 460
Ser Ile Glu His Leu Gln Ser Gly Val Tyr Tyr Leu Thr Asn Ile Asp

465 470 475 480
Asp Lys Phe Arg Arg Ser Tyr Asp Val Lys Lys

485 490

<210> 18

<211> 1473

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 18

atgaaattgt ctaccaagct ctgttggtgt ggcatcaaag gccgtcttcg cccacagaaa	60
cagcaacagc tgcacaatac caacctccag atgaccgagc ttaaaaagca aaagacggca	120
gagcagaaga ctcgccctca gaatgtcggc attaaggga ttcaaataca catccaacc	180
cagtgcgtaa atcagagcga gcttgaaaag ttcatggag tatcccagg aaagtataca	240
attggactgg gccagaccaa catgtcgttc gtgaacgacc gcgaagacat ttattcgatg	300
tcgcttactg tgttgagcaa attgatcaaa tctacaaca ttgatactaa taaaattggt	360

cgctcgagg tcggcaccga aacctgatac gataagtcga agagcgtgaa aagcgtgctg 420
atgcaacttt tcggtgagaa caccgatgtg gaaggaatcg atacccttaa cgcattgtac 480
ggtggaacca acgccttggt taattcgctt aactggatcg aatcgaacgc atgggatggc 540
cgcatgcaa tcgtgggttg cggagataac gccatttacg acaagggcgc cgcgcgcccc 600
acgggcggcg cgggcacagt ggcaatgtgg atcggctctg atgcaccaat tgtgttcgat 660
tcggtccgag caagctatat ggagcatgcc tatgatttct acaaaccaga ttttaccagc 720
gaataccctt acgtcgatgg tcacttctcc ttgacatgct acgtcaaggc gcttgaccag 780

gtgtacaaga gctactccaa aaaggcgatc tccaaaggcc tgggtgccga tcttcaggt 840
tcagatgctc ttaatgttct taagtatttc gattacaacg ttttcacagt gcccacgtgc 900
aaactcgtga cgaagtcata tggtcgcctc ttgtataacg attttcgagc taatccacag 960
ctttttccgg aagtggacgc agagctggcc acccgcgact acgacgagtc tctcacggat 1020
aagaatatcg aaaaaacttt cgtgaatgtg gctaaacat tccataagga acgcgtcgca 1080
caatctctca ttgttccac gaacaccgga aacatgtaca cggcctcagt gtatgtcgca 1140
ttcgctcac tgcttaacta tgttggatct gatgacctgc aaggtaaacg cgttggactc 1200

ttctcttacg gticgggttt ggcagcttcg ctgtactcct gcaagattgt cggagacgtc 1260
cagcatatca ttaaagaact ggacattaca aacaagctcg cgaaacgcat cactgaaacg 1320
cccaaggatt acgaagcggc cattgaattg cgcgagaacg cacatttgaa gaaaaacttt 1380
aaaccgcaag gatccatcga acacctgcaa tcgggtgtct attatctcac caacattgac 1440
gataagtttc gacgttccta tgatgttaaa aag 1473

<210> 19

<211> 398

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 19

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

1	5	10	15
Ser	Phe	Gln	Gly
Ser	Leu	Ser	Ser
Lys	Thr	Ala	Val
Glu	Leu	Gly	Ala
20	25	30	
Val	Ala	Leu	Lys
Gly	Ala	Leu	Ala
Lys	Val	Pro	Glu
Leu	Asp	Ala	Ser
35	40	45	
Lys	Asp	Phe	Asp
Glu	Ile	Ile	Phe
Gly	Asn	Val	Leu
Ser	Ala	Asn	Leu
50	55	60	

Gly Gln Ala Pro Ala Arg Gln Val Ala Leu Ala Ala Gly Leu Ser Asn

65 70 75 80

His Ile Val Ala Ser Thr Val Asn Lys Val Cys Ala Ser Ala Met Lys

85 90 95

Ala Ile Ile Leu Gly Ala Gln Ser Ile Lys Cys Gly Asn Ala Asp Val

100 105 110

Val Val Ala Gly Gly Cys Glu Ser Met Thr Asn Ala Pro Tyr Tyr Met

115 120 125

Pro Ala Ala Arg Ala Gly Ala Lys Phe Gly Gln Thr Val Leu Val Asp

130 135 140

Gly Val Glu Arg Asp Gly Leu Asn Asp Ala Tyr Asp Gly Leu Ala Met

145 150 155 160

Gly Val His Ala Glu Lys Cys Ala Arg Asp Trp Asp Ile Thr Arg Glu

165 170 175

Gln Gln Asp Asn Phe Ala Ile Glu Ser Tyr Gln Lys Ser Gln Lys Ser

180 185 190

Gln Lys Glu Gly Lys Phe Asp Asn Glu Ile Val Pro Val Thr Ile Lys

195 200 205

Gly Phe Arg Gly Lys Pro Asp Thr Gln Val Thr Lys Asp Glu Glu Pro

210 215 220

Ala Arg Leu His Val Glu Lys Leu Arg Ser Ala Arg Thr Val Phe Gln

225 230 235 240

Lys Glu Asn Gly Thr Val Thr Ala Ala Asn Ala Ser Pro Ile Asn Asp

245 250 255

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

260 265 270

Asn Leu Lys Pro Leu Ala Ile Ile Lys Gly Trp Gly Glu Ala Ala His

275 280 285

Gln Pro Ala Asp Phe Thr Trp Ala Pro Ser Leu Ala Val Pro Lys Ala

290 295 300

Leu Lys His Ala Gly Ile Glu Asp Ile Asn Ser Val Asp Tyr Phe Glu

305 310 315 320

Phe Asn Glu Ala Phe Ser Val Val Gly Leu Val Asn Thr Lys Ile Leu

325 330 335

Lys Leu Asp Pro Ser Lys Val Asn Val Tyr Gly Gly Ala Val Ala Leu

340 345 350

Gly His Pro Leu Gly Cys Ser Gly Ala Arg Val Val Val Thr Leu Leu

355 360 365

Ser Ile Leu Gln Gln Glu Gly Gly Lys Ile Gly Val Ala Ala Ile Cys

370 375 380

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

385 390 395

<210> 20

<211> 1194

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 20

atgtcacaga atgtctacat cgtctctacc gccgcaccc cgattgggtc cttccagggt	60
tctcttagct cgaaaacagc ggttgaactc ggcgcggtg cactgaaagg agccttggct	120
aaagtaccag aattggacgc gtccaaggac ttcgatgaaa ttattttcgg taacgtactg	180
tcagctaatic tcggccaagc tccagctcgc caggctcgcg tcgctcgggg cctctccaat	240
cacattgtgg ccagcactgt gaataaggtc tgcgcatcgg cgatgaaggc gattatcctg	300

ggcgcccaga gcatataaatg cggtaatgca gatgttgtgg ttgccggtgg ctgtgagtcg	360
atgaccaacg ctccgtacta catgcccgct gcacgcgctg gcgctaaatt cggtcaaacc	420
gtcctggtcg atggtgtcga gcgcgatggc ctcaacgatg cgtacgacgg cctcgcgatg	480
ggagtacacg ctgagaagtg tgcccgcgac tgggatatca cccgagagca gcaggacaat	540
tttgccatcg agagctatca gaagtcacaa aagtcacaga aggagggtaa attcgataac	600
gaaatcgttc cagtgcgatg caaaggcttc cgcggtaagc cagatactca agtaacccaa	660
gatgaggaac ctgcccgttt gcacgtagag aagctccgct cgcccccgcac tgtcttcag	720

aaagaaaatg gaaccgtgac ggcagctaac gcaagcccta ttaacgacgg tgcggcagcg 780
gtcattctcg tgtcagaaaa ggtgctgaaa gaaaagaacc tcaaacctct ggccatcatc 840

aagggtctggg gtgaggccgc tcaccaacct gctgacttca cctgggcacc atccctggct 900
 gtacaaaag cactcaagca cgctggtatt gaagatatca actccgtcga ctacttcgaa 960
 tttaatgagg cattctcggg tgtcggcctt gttaacacta agattcttaa gctcgatcca 1020
 tccaaggtca atgtatatgg cgggtcgggt gctttgggcc atcctctcgg ttgctcagga 1080
 gcccgctgg ttgttactct cctcagcatt cttcagcagg aggggtggaaa gatcggcgtc 1140

gctgctatatt gcaacggcgg tgggtggtgcc tctcaattg ttattgaaaa gatt 1194

<210> 21

<211> 1045

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 21

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

1 5 10 15

Cys Thr Ala Arg Phe Ser Ala Arg Tyr Pro Ile His Val Ile Val Val

20 25 30

Ala Val Leu Leu Ser Ala Ala Tyr Leu Ser Val Thr Gln Ser Tyr

35 40 45

Leu Asn Glu Trp Lys Leu Asp Ser Asn Gln Tyr Ser Thr Tyr Leu Ser

50 55 60

Ile Lys Pro Asp Glu Leu Phe Glu Lys Cys Thr His Tyr Tyr Arg Ser

65 70 75 80

Pro Val Ser Asp Thr Trp Lys Leu Leu Ser Ser Lys Glu Ala Ala Asp

85 90 95

Ile Tyr Thr Pro Phe His Tyr Tyr Leu Ser Thr Ile Ser Phe Gln Ser

100 105 110

Lys Asp Asn Ser Thr Thr Leu Pro Ser Leu Asp Asp Val Ile Tyr Ser

115 120 125

Val Asp His Thr Arg Tyr Leu Leu Ser Glu Glu Pro Lys Ile Pro Thr

130 135 140

Glu Leu Val Ser Glu Asn Gly Thr Lys Trp Arg Leu Arg Asn Asn Ser

145 150 155 160

Asn Phe Ile Leu Asp Leu His Asn Ile Tyr Arg Asn Met Val Lys Gln

	165		170		175
Phe Ser Asn Lys Thr Ser Glu Phe Asp Gln Phe Asp Leu Phe Ile Ile					
	180		185		190
Leu Ala Ala Tyr Leu Thr Leu Phe Tyr Thr Leu Cys Cys Leu Phe Asn					
	195		200		205
Asp Met Arg Lys Ile Gly Ser Lys Phe Trp Leu Ser Phe Ser Ala Leu					
	210		215		220
Ser Asn Ser Ala Cys Ala Leu Tyr Leu Ser Leu Tyr Thr Thr His Ser					
225		230		235	240
Leu Leu Lys Lys Pro Ala Ser Leu Leu Ser Leu Val Ile Gly Leu Pro					
	245		250		255
Phe Ile Val Val Ile Ile Gly Phe Lys His Lys Val Arg Leu Ala Ala					
	260		265		270
Phe Ser Leu Gln Lys Phe His Arg Ile Ser Ile Asp Lys Lys Ile Thr					
	275		280		285
Val Ser Asn Ile Ile Tyr Glu Ala Met Phe Gln Glu Gly Ala Tyr Leu					
	290		295		300
Ile Arg Asp Tyr Leu Phe Tyr Ile Ser Ser Phe Ile Gly Cys Ala Ile					
305		310		315	320
Tyr Ala Arg His Leu Pro Gly Leu Val Asn Phe Cys Ile Leu Ser Thr					
	325		330		335
Phe Met Leu Val Phe Asp Leu Leu Leu Ser Ala Thr Phe Tyr Ser Ala					
	340		345		350
Ile Leu Ser Met Lys Leu Glu Ile Asn Ile Ile His Arg Ser Thr Val					
	355		360		365
Ile Arg Gln Thr Leu Glu Glu Asp Gly Val Val Pro Thr Thr Ala Asp					
	370		375		380
Ile Ile Tyr Lys Asp Glu Thr Ala Ser Glu Pro His Phe Leu Arg Ser					
385		390		395	400
Asn Val Ala Ile Ile Leu Gly Lys Ala Ser Val Ile Gly Leu Leu Leu					
	405		410		415

Leu Ile Asn Leu Tyr Val Phe Thr Asp Lys Leu Asn Ala Thr Ile Leu
 420 425 430

 Asn Thr Val Tyr Phe Asp Ser Thr Ile Tyr Ser Leu Pro Asn Phe Ile
 435 440 445
 Asn Tyr Lys Asp Ile Gly Asn Leu Ser Asn Gln Val Ile Ile Ser Val
 450 455 460
 Leu Pro Lys Gln Tyr Tyr Thr Pro Leu Lys Lys Tyr His Gln Ile Glu
 465 470 475 480
 Asp Ser Val Leu Leu Ile Ile Asp Ser Val Ser Asn Ala Ile Arg Asp
 485 490 495

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

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

 Lys Leu Pro Leu Tyr Ser Leu Glu Lys Lys Leu Glu Asp Thr Thr Arg
 625 630 635 640
 Ala Val Leu Val Arg Arg Lys Ala Leu Ser Thr Leu Ala Glu Ser Pro
 645 650 655
 Ile Leu Val Ser Glu Lys Leu Pro Phe Arg Asn Tyr Asp Tyr Asp Arg

660	665	670
Val Phe Gly Ala Cys Cys Glu Asn Val Ile Gly Tyr Met Pro Ile Pro		
675	680	685
Val Gly Val Ile Gly Pro Leu Ile Ile Asp Gly Thr Ser Tyr His Ile		
690	695	700
Pro Met Ala Thr Thr Glu Gly Cys Leu Val Ala Ser Ala Met Arg Gly		
705	710	715
Cys Lys Ala Ile Asn Ala Gly Gly Gly Ala Thr Thr Val Leu Thr Lys		
725	730	735
Asp Gly Met Thr Arg Gly Pro Val Val Arg Phe Pro Thr Leu Ile Arg		
740	745	750
Ser Gly Ala Cys Lys Ile Trp Leu Asp Ser Glu Glu Gly Gln Asn Ser		
755	760	765
Ile Lys Lys Ala Phe Asn Ser Thr Ser Arg Phe Ala Arg Leu Gln His		
770	775	780
Ile Gln Thr Cys Leu Ala Gly Asp Leu Leu Phe Met Arg Phe Arg Thr		
785	790	795
Thr Thr Gly Asp Ala Met Gly Met Asn Met Ile Ser Lys Gly Val Glu		
805	810	815
Tyr Ser Leu Lys Gln Met Val Glu Glu Tyr Gly Trp Glu Asp Met Glu		
820	825	830
Val Val Ser Val Ser Gly Asn Tyr Cys Thr Asp Lys Lys Pro Ala Ala		
835	840	845
Ile Asn Trp Ile Glu Gly Arg Gly Lys Ser Val Val Ala Glu Ala Thr		
850	855	860
Ile Pro Gly Asp Val Val Lys Ser Val Leu Lys Ser Asp Val Ser Ala		
865	870	875
Leu Val Glu Leu Asn Ile Ser Lys Asn Leu Val Gly Ser Ala Met Ala		
885	890	895
Gly Ser Val Gly Gly Phe Asn Ala His Ala Ala Asn Leu Val Thr Ala		
900	905	910

Leu Phe Leu Ala Leu Gly Gln Asp Pro Ala Gln Asn Val Glu Ser Ser
 915 920 925
 Asn Cys Ile Thr Leu Met Lys Glu Val Asp Gly Asp Leu Arg Ile Ser
 930 935 940
 Val Ser Met Pro Ser Ile Glu Val Gly Thr Ile Gly Gly Gly Thr Val
 945 950 955 960
 Leu Glu Pro Gln Gly Ala Met Leu Asp Leu Leu Gly Val Arg Gly Pro
 965 970 975
 His Pro Thr Glu Pro Gly Ala Asn Ala Arg Gln Leu Ala Arg Ile Ile
 980 985 990
 Ala Cys Ala Val Leu Ala Gly Glu Leu Ser Leu Cys Ser Ala Leu Ala
 995 1000 1005

Ala Gly His Leu Val Gln Ser His Met Thr His Asn Arg Lys Thr
 1010 1015 1020
 Asn Lys Ala Asn Glu Leu Pro Gln Pro Ser Asn Lys Gly Pro Pro
 1025 1030 1035
 Cys Lys Thr Ser Ala Leu Leu
 1040 1045

<210> 22

<211> 3135

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 22

atgtccttgc ctcttaagac catcgtgcac ctggttaaagc ctttcgcctg cactgcccgc 60
 ttctccgcgc gctaccccat ccacgttata gtggtggcgg tgctgctctc ggcggcagct 120
 tacttgagcg taacccaate ctacttgaat gaatggaagc tgcactcgaa ccagtactcc 180
 acctacctta gcatcaagcc tgatgaattg ttcgaaaaat gtaccacta ctaccgtctt 240
 ccagtgtctg atacttggaa actccttagc tctaaggaag ctgcagatat ttacaccccc 300
 ttccattatt acctgtccac catttccttc cagtccaagg acaactcgac caccctgccc 360
 agcctggatg acgttatcta ctccgtagac catacccgct acttgctcag cgaggaaccg 420

aaaatcccga cggaactcgt atccgaaaac ggaaccaagt ggcgtctgcg taataactct	480
aactttatcc ttgatctcca caacatttat cgcaatatgg ttaagcagtt ctcaaacaaa	540
acctcggaat ttgaccaatt cgacttggtc atcattctcg cagcatattt gactttgttt	600
tacacctctt gcgtctcttt taatgacatg cgcaaaatcg gctccaagtt ctggctttcc	660
ttctcagctc ttccaacacg cgctgcgcga ctttatttga gcttgtaac tacccactcc	720
ctccttaaga aaccagcttc cctcctctcc ctgcgtcatg gcttgccatt catcgtggtt	780
atcattgggt tcaaacataa ggtgcgcctc gctgcgttct ccctgcagaa gtttcaccga	840
attagcatcg ataaaaagat cagggtctct aacattatct acgaagccat gttccaagaa	900
ggcgcgtacc ttatccgaga ttacctgttc tacatttcca gcttcatcgg ttgtgctatc	960
tacgcacgcc accttccagg tctggtgaat ttctgcattc tctccacctt tatgcttgtg	1020
tttgatctcc ttttgagcgc cactttctat agcgcgatcc tctcaatgaa actggaaatc	1080
aacatcattc accgcagcac cgtcattcgc cagacccttg aagaagatgg cgtgggtgccg	1140
accaccgccg atattattta taaagatgaa accgcacccg aaccacattt cctgcgttcc	1200
aacgtggcga ttatccttgg taaggcgtcg gtcattggac tgctgctgct gattaatttg	1260
tacgttttca cagataagct caacgcgacg attctgaata ctgtttactt cgattctacc	1320
atctattccc tcccgaactt catcaactat aaggatatcg gaaacctttc taatcaagtc	1380
attattagcg ttctgccaaa gcagtactac acccctttga aaaagtatca tcagattgag	1440
gactctgtgc tccttatcat cgactccgtg agcaatgcta ttcgcgatca atttatctcc	1500
aagctcttgt ttttcgcgtt cgctgtttca atttccatca acgtgtacct gttgaacgca	1560
gctaagattc acacgggcta catgaatttc cagccccaat ccaacaagat tgatgacctt	1620
gtcgtgcaac agaagctcgc aaccattgaa ttctccgaga ccgatccat gcctgcatcc	1680
tccggcctcg agaccccgat taccgcgaag gacatcatta tttcagaaga aatccagaat	1740
aatgaatgcg tatacgcgct gtcattctag gatgaacca ttcgcccgtt gtcgaaacctt	1800
gttgaactga tggagaaaaga gcaactgaaa aatatgaaca acacagaagt ttctaacttg	1860
gtggtgaacg gtaaattgcc gttgtactcc ctggaaaaga aactcgaaga taccacccga	1920
gccgttcttg ttcgacgcaa ggcgctttcc accctggccg aatctccaat ccttgitttc	1980
gaaaaactcc cttttcgcaa ctacgattac gatcgtgtgt tcggcgcgtg ttgcgagaac	2040
gtaattggct acatgcctat ccagtagga gtgatcggcc ctctcatcat cgatggcacc	2100
agctaccata ttccaatggc aaccacagaa ggttgcttgg tagcaagcgc gatgcgagga	2160

tgcaaagcaa tcaacgctgg aggcgggtgct accaccgttc tgaccaagga tggaatgact 2220
 cgcgcccccg tagtccgctt cctactctc atccgctctg gtgcgtgcaa gatttggtc 2280
 gattccgagg aaggccagaa ctctatcaaa aaagcgttca acagcacctc ccgcttcgcc 2340
 cgtttgcaac acattcagac ttgcctggca ggcgacctcc tcttcatgcg ttccgcacc 2400
 actactggcg acgcaatggg tatgaacatg atctccaaag gagttgagta ctgcgtcaaa 2460
 cagatgggtg aagagtatgg ctgggaagac atggaagtig tgtccgtttc cggttaactac 2520
 tgcacggaca agaaaccagc tgcgatcaat tggattgagg gacgcggcaa atctgtcgtc 2580

gcggaggcaa ccatcccagg tgacgtcgtc aaatccgttc tcaaactga cgtctccga 2640
 ctggttgaac tgaacatctc taagaacctt gtgggttctg ctatggctgg cagcgtcgg 2700
 ggctttaacg cgcatgcggc gaaccttggt accgcgtctt tcctggcgct cggacaggat 2760
 cccgcgcaga atgtggagtc atcaaatgac attacactga tgaaggaggt ggacggcgat 2820
 ctgcgtatct ctgtttccat gccgtccatt gaggtgggta ccattggtgg cggcacagtc 2880
 ttggagcctc agggtgccat gctggatctg ttgggcgtac gtggccccca tcctaccgaa 2940
 ccaggcgcaa acgcacgcca gctcgcgcgc attatcgcat gcgccgtgct ggcaggtgaa 3000

ttgtccctct gtcagcctt ggccgcgggt catttggtcc aaagccacat gaccataac 3060
 cgcaaaacga ataaggccaa cgaacttctt caaccttcaa acaagggtcc accttgcaaa 3120
 acctcgcgc tgttg 3135

<210> 23

<211> 451

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 23

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

1 5 10 15

Gly Gly Tyr Leu Val Leu Asp Thr Lys Tyr Glu Ala Phe Val Val Gly

20 25 30

Leu Ser Ala Arg Met His Ala Val Ala His Pro Tyr Gly Ser Leu Gln

35 40 45

Gly Ser Asp Lys Phe Glu Val Arg Val Lys Ser Lys Gln Phe Lys Asp

50 55 60

Gly Glu Trp Leu Tyr His Ile Ser Pro Lys Ser Gly Phe Ile Pro Val

65 70 75 80

Ser Ile Gly Gly Ser Lys Asn Pro Phe Ile Glu Lys Val Ile Ala Asn

85 90 95

Val Phe Ser Tyr Phe Lys Pro Asn Met Asp Asp Tyr Cys Asn Arg Asn

100 105 110

Leu Phe Val Ile Asp Ile Phe Ser Asp Asp Ala Tyr His Ser Gln Glu

115 120 125

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

130 135 140

His Arg Ile Glu Glu Val Pro Lys Thr Gly Leu Gly Ser Ser Ala Gly

145 150 155 160

Leu Val Thr Val Leu Thr Thr Ala Leu Ala Ser Phe Phe Val Ser Asp

165 170 175

Leu Glu Asn Asn Val Asp Lys Tyr Arg Glu Val Ile His Asn Leu Ala

180 185 190

Gln Val Ala His Cys Gln Ala Gln Gly Lys Ile Gly Ser Gly Phe Asp

195 200 205

Val Ala Ala Ala Ala Tyr Gly Ser Ile Arg Tyr Arg Arg Phe Pro Pro

210 215 220

Ala Leu Ile Ser Asn Leu Pro Asp Ile Gly Ser Ala Thr Tyr Gly Ser

225 230 235 240

Lys Leu Ala His Leu Val Asp Glu Glu Asp Trp Asn Ile Thr Ile Lys

245 250 255

Ser Asn His Leu Pro Ser Gly Leu Thr Leu Trp Met Gly Asp Ile Lys

260 265 270

Asn Gly Ser Glu Thr Val Lys Leu Val Gln Lys Val Lys Asn Trp Tyr

275 280 285

Asp Ser His Met Pro Glu Ser Leu Lys Ile Tyr Thr Glu Leu Asp His

290 295 300

Ala Asn Ser Arg Phe Met Asp Gly Leu Ser Lys Leu Asp Arg Leu His

305 310 315 320

Glu Thr His Asp Asp Tyr Ser Asp Gln Ile Phe Glu Ser Leu Glu Arg

325 330 335
 Asn Asp Cys Thr Cys Gln Lys Tyr Pro Glu Ile Thr Glu Val Arg Asp

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

 405 410 415
 Gln Thr Ala Asn Asp Lys Arg Phe Ser Lys Val Gln Trp Leu Asp Val
 420 425 430
 Thr Gln Ala Asp Trp Gly Val Arg Lys Glu Lys Asp Pro Glu Thr Tyr
 435 440 445
 Leu Asp Lys
 450
 <210> 24
 <211> 1353
 <212> DNA
 <213> Artificial
 <220><223> mutated and/or codon-optimized sequence
 <400> 24
 atgtccgaac tgcgcgcatt ttccgcgcct ggtaaggcgc tgctcgccgg cggtacatt 60

 gtgttggaca ctaaatacga ggccttcgtg gtgggccttt ccgcacgtat gcacgccgtc 120
 gctcatcctt atggttcgtt gcaaggttcc gataagtttg aggtgcgcgt gaaatccaaa 180
 cagtttaaag atggatgaat gctgtaccac atcagcccaa agtcaggctt cattccagta 240
 tccattgggtg gtccaagaa tccatttatc gagaaggcca ttgccaacgt cttttcatac 300
 ttcaaaccaa acatggatga ttattgtaat cgcaacttgt tcgtgatcga tatcttctct 360
 gacgacgctt atcactcgca ggaagattcc gttacggagc atcgtggaaa ccgacgtctg 420
 tcatttcatt cgcaccgcat cgaggaagtt cccaagaccg gcctgggctc gtctgccggc 480

cttgtcactg ttttgactac cgccctcgca tctttcttcg tgtctgacct tgaatacaac 540
gttgataagt accgcgaggt gattcacaac ctggcgagg tagcgactg ccaggcgag 600
ggaaaaattg gcagcggtt tgatgtggct gctgctgcat acggcagcat ccgttaccgc 660
cgcttccac cagcgcttat tagcaacctg ccggacatcg gctccgcgac ttacggttcc 720
aaattggccc acctgtcga cgaggaagac tggaacatca caatcaagtc taaccacctc 780
ccttcggcc tgacgctctg gatggcgat atcaagaatg gctctgaaac cgtgaagctc 840
gtgcagaagg ttaagaactg gtacgactca cacatgcctg aatcgttgaa aatctacacc 900

gagcttgatc acgcgaactc ccgtttcatg gacggattgt cgaagttgga ccgtctgcat 960
gaaacgcagc acgattattc cgaccaaacc tttgaatcgc tggaacgtaa cgattgtact 1020
tgtcagaagt atcccgaat caccgaggtc cgcgacgagg tcgcgacaat tcgccgtca 1080
ttccgcaaaa tctaagga atccggtgct gacattgagc ccccggttca gacctcctg 1140
ctcgatgatt gccagactct caagggtgtc ttgacgtgtt tgatccccgg agccggcggc 1200
tacgatgcca tcgcagttat caccaagcag gatgtggatc tccgcgcgca gacagcaaac 1260
gataagcgtt tttcaaaggt gcagtggctg gacgtgacct aagctgattg gggcggttcgt 1320

aaagaaaaag atccagagac atatctggat aag 1353

<210> 25

<211> 288

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 25

Met Thr Ala Asp Asn Asn Ser Met Pro His Gly Ala Val Ser Ser Tyr

1 5 10 15

Ala Lys Leu Val Gln Asn Gln Thr Pro Glu Asp Ile Leu Glu Glu Phe

20 25 30

Pro Glu Ile Ile Pro Leu Gln Gln Arg Pro Asn Thr Arg Ser Ser Glu

35 40 45

Thr Ser Asn Asp Glu Ser Gly Glu Thr Cys Phe Ser Gly His Asp Glu

50 55 60

Glu Gln Ile Lys Leu Met Asn Glu Asn Cys Ile Val Leu Asp Trp Asp

65 70 75 80

Asp Asn Ala Ile Gly Ala Gly Thr Lys Lys Val Cys His Leu Met Glu

85 90 95

Asn Ile Glu Lys Gly Leu Leu His Arg Ala Phe Ser Val Phe Ile Phe
100 105 110

Asn Glu Gln Gly Glu Leu Leu Leu Gln Gln Arg Ala Thr Glu Lys Ile
115 120 125

Thr Phe Pro Asp Leu Trp Thr Asn Thr Cys Cys Ser His Pro Leu Cys
130 135 140

Ile Asp Asp Glu Leu Gly Leu Lys Gly Lys Leu Asp Asp Lys Ile Lys
145 150 155 160

Gly Ala Ile Thr Ala Ala Val Arg Lys Leu Asp His Glu Leu Gly Ile
165 170 175

Pro Glu Asp Glu Thr Lys Thr Arg Gly Lys Phe His Phe Leu Asn Arg
180 185 190

Ile His Tyr Met Ala Pro Ser Asn Glu Pro Trp Gly Glu His Glu Ile
195 200 205

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

Asn Pro Asn Val Asn Glu Val Arg Asp Phe Lys Trp Val Ser Pro Asn
225 230 235 240

Asp Leu Lys Thr Met Phe Ala Asp Pro Ser Tyr Lys Phe Thr Pro Trp
245 250 255

Phe Lys Ile Ile Cys Glu Asn Tyr Leu Phe Asn Trp Trp Glu Gln Leu
260 265 270

Asp Asp Leu Ser Glu Val Glu Asn Asp Arg Gln Ile His Arg Met Leu
275 280 285

<210> 26

<211> 864

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400

> 26

atgacagcag ataacaattc catgcctcac ggtgctgtgt cttcatacgc gaaactggtg

60

cagaacaaaa cacctgagga tattctcgag gaattccccc aaatcatccc tctccagcag 120
 cggccgaaca cccgctcctc tgaaacctct aatgacgaat ccggcgagac atgtttctcc 180
 ggccacgacg aagagcaaat taagctgatg aacgaaaatt gcatcgtgtt ggactgggat 240
 gacaacgcca ttggagccgg taccaaaaag gtgtgccatc tgatggagaa cattgaaaag 300
 ggactgctcc atcgagcttt ctcagtgttc attttcaatg aacagggtga gctcctgctg 360
 caacagcgcg caacggaaaa gattaccttc cccgacctct ggactaacac ctgctgctca 420

cacccattgt gtatcgatga cgaacttggc ttgaaggga agcttgatga taagattaag 480
 ggcgcaatca cggccgctgt tcgtaaaact gatcacgaac ttggaattcc tgaagacgaa 540
 accaagactc gcggcaaat tcaattcttg aatcgatatcc actatatggc gccttccaac 600
 gagccttggg gcgagcacga gatcgactat attttgttct acaagatcaa tgcaaaggag 660
 aacctgactg tgaacccgaa tgtgaatgaa gtgcgtgact tcaagtgggt ctcaccgaac 720
 gatctgaaga ccatgtttgc cgatccatcc tacaaattca ccccggtggt caaaattatc 780
 tgtgaaaact atctgttcaa ttggtgggaa caactcgacg atttgtccga ggtcgagaac 840

gategtcaaa tcacacgtat gctt 864

<210> 27

<211> 527

<212> PRT

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 27

Met Ala Ala Asp Gln Leu Val Lys Thr Glu Val Thr Lys Lys Ser Phe

1 5 10 15

Thr Ala Pro Val Gln Lys Ala Ser Thr Pro Val Leu Thr Asn Lys Thr

20 25 30

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

35 40 45

Ser Ser Gly Pro Ser Ser Ser Ser Glu Glu Asp Asp Ser Arg Asp Ile

50 55 60

Glu Ser Leu Asp Lys Lys Ile Arg Pro Leu Glu Glu Leu Glu Ala Leu

65 70 75 80

Leu Ser Ser Gly Asn Thr Lys Gln Leu Lys Asn Lys Glu Val Ala Ala

85 90 95

Leu Val Ile His Gly Lys Leu Pro Leu Tyr Ala Leu Glu Lys Lys Leu

100 105 110
Gly Asp Thr Thr Arg Ala Val Ala Val Arg Arg Lys Ala Leu Ser Ile

115 120 125
Leu Ala Glu Ala Pro Val Leu Ala Ser Asp Arg Leu Pro Tyr Lys Asn

130 135 140
Tyr Asp Tyr Asp Arg Val Phe Gly Ala Cys Cys Glu Asn Val Ile Gly

145 150 155 160
Tyr Met Pro Leu Pro Val Gly Val Ile Gly Pro Leu Val Ile Asp Gly

165 170 175
Thr Ser Tyr His Ile Pro Met Ala Thr Thr Glu Gly Cys Leu Val Ala

180 185 190
Ser Ala Met Arg Gly Cys Lys Ala Ile Asn Ala Gly Gly Gly Ala Thr

195 200 205
Thr Val Leu Thr Lys Asp Gly Met Thr Arg Gly Pro Val Val Arg Phe

210 215 220
Pro Thr Leu Lys Arg Ser Gly Ala Cys Lys Ile Trp Leu Asp Ser Glu

225 230 235 240
Glu Gly Gln Asn Ala Ile Lys Lys Ala Phe Asn Ser Thr Ser Arg Phe

245 250 255
Ala Arg Leu Gln His Ile Gln Thr Cys Leu Ala Gly Asp Leu Leu Phe

260 265 270
Met Arg Phe Arg Thr Thr Thr Gly Asp Ala Met Gly Met Asn Met Ile

275 280 285
Ser Lys Gly Val Glu Tyr Ser Leu Lys Gln Met Val Glu Glu Tyr Gly

290 295 300
Trp Glu Asp Met Glu Val Val Ser Val Ser Gly Asn Tyr Cys Thr Asp

305 310 315 320
Lys Lys Pro Ala Ala Ile Asn Trp Ile Glu Gly Arg Gly Lys Ser Val

325 330 335
Val Ala Glu Ala Thr Ile Pro Gly Asp Val Val Arg Lys Val Leu Lys

340 345 350
Ser Asp Val Ser Ala Leu Val Glu Leu Asn Ile Ala Lys Asn Leu Val

355 360 365
Gly Ser Ala Met Ala Gly Ser Val Gly Gly Phe Asn Ala His Ala Ala

370 375 380
Asn Leu Val Thr Ala Val Phe Leu Ala Leu Gly Gln Asp Pro Ala Gln

385 390 395 400
Asn Val Glu Ser Ser Asn Cys Ile Thr Leu Met Lys Glu Val Asp Gly

405 410 415
Asp Leu Arg Ile Ser Val Ser Met Pro Ser Ile Glu Val Gly Thr Ile

420 425 430
Gly Gly Gly Thr Val Leu Glu Pro Gln Gly Ala Met Leu Asp Leu Leu

435 440 445
Gly Val Arg Gly Pro His Ala Thr Ala Pro Gly Thr Asn Ala Arg Gln

450 455 460
Leu Ala Arg Ile Val Ala Cys Ala Val Leu Ala Gly Glu Leu Ser Leu

465 470 475 480
Cys Ala Ala Leu Ala Ala Gly His Leu Val Gln Ser His Met Thr His

485 490 495
Asn Arg Lys Pro Ala Glu Pro Thr Lys Pro Asn Asn Leu Asp Ala Thr

500 505 510
Asp Ile Asn Arg Leu Lys Asp Gly Ser Val Thr Cys Ile Lys Ser

515 520 525
<210> 28

<211> 1581

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 28

atggcggctg atcagctcgt taaaacggaa gtaacaaaga aatccttcac cgctcccgta 60

cagaaagctt ctactcctgt cttgaccaac aagaccgtta tctccggctc taaagtgaag 120

tcgtctcttt ccgcccagag cagctcatcc ggtccgtcgt cttcctccga agaggacgac 180
 tcacgtgata ttgaatcact ggataaaaag attcgaccgc tcgaagagct tgaggcgctt 240
 ttgagctccg gcaataccaa acagttgaag aacaaggagg ttgccgcctt tgtgattcac 300
 ggcaaacttc cgctctatgc cttggagaag aagttgggag acaccacgcg cgcagtggcc 360
 gtgcgccgca aggcactgtc tattctcgcg gaggcacccg tcctggcttc cgaccgttg 420
 ccctacaaga actacgatta cgaccgtgtc ttcggcgcgt gttgcgagaa tgtaatcgga 480

tatatgcccc ttcccgttgg cgtgatcggc ccacttgtca tcgatggcac atcataccat 540
 attccaatgg cgaccaccga gggttgcctc gtggcgctcg ccatgcgcgg ttgtaaggcc 600
 atcaacgcgc gaggtggagc aactaccgta cttactaagg acggcatgac tcgtggcccc 660
 gtggtgcgtt tcccgaactc taaacgtcc ggtgcgtgca agatctggtt ggattccgag 720
 gagggacaga acgcaatcaa aaaagctttc aacagcactt ctgccttcgc tcgacttcag 780
 cacattcaga ccgtctcgc tgggtgacttg ttgttcatgc gttttcgac caccactgga 840
 gacgcgatgg gcatgaatat gattagcaaa ggagtggagt actccctcaa acaaatggtg 900

gaagaatatg gctgggaaga catggaagtg gtttccgtgt ctggttaatta ttgtaccgac 960
 aagaagccgg cagccatcaa ctggattgag ggccgtggca agtccgtggt gccggaggcg 1020
 acaattccgg gtgatgtcgt gcgcaaagt ctcaagtctg acgtgtcagc gttggtagaa 1080
 ctgaacatcg caaaaaacct tgtaggatcg gcaatggctg gttcagtggg cggttttaat 1140
 gcgcacgcag ccaacctcgt caccgtgtg ttcccttgccc tgggtcagga tccagcacag 1200
 aatgtggaaa gctcgaactg catcacctg atgaaagaag tggatggcga ctgtcgatc 1260
 tccgtctcga tgccatcaat cgaagtgggt accatcgggt gcggaaccgt gttggaacca 1320

caggagcaaa tgcttgatct cctcggcgtt cgaggccgc acgccacgc gcccgccacc 1380
 aacgcacgtc agctggcacg catcgtggtt tgcgcggttc tcgcgggtga attgagcctc 1440
 tgcgcagccc tggcagcggg tcacctggtg caatctcaca tgacacataa ccgtaaacca 1500
 gccgagccca ccaagccgaa taacctcgac gccactgata ttaaccgctt gaaggatggt 1560
 tccgtcacct gcatcaaatc t 1581

<210> 29

<211> 425

<212> PRT

<213> Staphylococcus aureus

<400> 29

Met Gln Ser Leu Asp Lys Asn Phe Arg His Leu Ser Arg Gln Gln Lys

1 5 10 15
 Leu Gln Gln Leu Val Asp Lys Gln Trp Leu Ser Glu Asp Gln Phe Asp
 20 25 30
 Ile Leu Leu Asn His Pro Leu Ile Asp Glu Glu Val Ala Asn Ser Leu
 35 40 45
 Ile Glu Asn Val Ile Ala Gln Gly Ala Leu Pro Val Gly Leu Leu Pro
 50 55 60
 Asn Ile Ile Val Asp Asp Lys Ala Tyr Val Val Pro Met Met Val Glu

 65 70 75 80
 Glu Pro Ser Val Val Ala Ala Ala Ser Tyr Gly Ala Lys Leu Val Asn
 85 90 95
 Gln Thr Gly Gly Phe Lys Thr Val Ser Ser Glu Arg Ile Met Ile Gly
 100 105 110
 Gln Ile Val Phe Asp Gly Val Asp Asp Thr Glu Lys Leu Ser Ala Asp
 115 120 125
 Ile Lys Ala Leu Glu Lys Gln Ile His Lys Ile Ala Asp Glu Ala Tyr

 130 135 140
 Pro Ser Ile Lys Ala Arg Gly Gly Gly Tyr Gln Arg Ile Ala Ile Asp
 145 150 155 160
 Thr Phe Pro Glu Gln Gln Leu Leu Ser Leu Lys Val Phe Val Asp Thr
 165 170 175
 Lys Asp Ala Met Gly Ala Asn Met Leu Asn Thr Ile Leu Glu Ala Ile
 180 185 190
 Thr Ala Phe Leu Lys Asn Glu Ser Pro Gln Ser Asp Ile Leu Met Ser

 195 200 205
 Ile Leu Ser Asn His Ala Thr Ala Ser Val Val Lys Val Gln Gly Glu
 210 215 220
 Ile Asp Val Lys Asp Leu Ala Arg Gly Glu Arg Thr Gly Glu Glu Val
 225 230 235 240
 Ala Lys Arg Met Glu Arg Ala Ser Val Leu Ala Gln Val Asp Ile His
 245 250 255

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

260 265 270
Val Leu Ala Thr Gly Asn Asp Thr Arg Gly Ala Glu Ala Ser Ala His
275 280 285
Ala Tyr Ala Ser Arg Asp Gly Gln Tyr Arg Gly Ile Ala Thr Trp Arg
290 295 300
Tyr Asp Gln Lys Arg Gln Arg Leu Ile Gly Thr Ile Glu Val Pro Met
305 310 315 320
Thr Leu Ala Ile Val Gly Gly Gly Thr Lys Val Leu Pro Ile Ala Lys

325 330 335
Ala Ser Leu Glu Leu Leu Asn Val Asp Ser Ala Gln Glu Leu Gly His
340 345 350
Val Val Ala Ala Val Gly Leu Ala Gln Asn Phe Ala Ala Cys Arg Ala
355 360 365
Leu Val Ser Glu Gly Ile Gln Gln Gly His Met Ser Leu Gln Tyr Lys
370 375 380
Ser Leu Ala Ile Val Val Gly Ala Lys Gly Asp Glu Ile Ala Gln Val

385 390 395 400
Ala Glu Ala Leu Lys Gln Glu Pro Arg Ala Asn Thr Gln Val Ala Glu
405 410 415
Arg Ile Leu Gln Glu Ile Arg Gln Gln
420 425

<210> 30

<211> 1275

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 30

atgcaatccc tcgataaaaa tttccgtcac ctttctcgac agcagaaatt gcagcaactt 60

gttgacaagc aatggctctc ggaggatcaa ttcgatatcc tgc tcaatca cccactgac 120

gatgaggagg tcgccaattc cctgattgag aacgtcatcg cgcagggtgc actgcccggtg 180

ggccttctgc ctaacatcat cgtcgtatgac aaggcctacg tagtacctat gatggtcgaa 240
gaaccgtccg tgggtggcagc tgcttcctat ggtgccaaagc tggtaacca gaccggcgga 300
ttcaagactg tgtcttctga acgcattatg atcgacaga ttgtcttcga tgggtgtgat 360
gatacggaaa agctctctgc agacatcaag gcattggaga aacagatcca taaaatcgcg 420
gatgaggcgt acccctccat caaagcgcg gcggcggt atcagcgaat cgctattgac 480
accttcctg aacagcagct gctctcctg aaagtcttcg tggatactaa ggacgcatg 540

ggtgcaaata tgttgaacac catcctggag gccatcacgc cgttcctgaa gaacgagagc 600
cctcagtcgg atattttgat gtcaatcctg tccaatcacg ccactgcac tgttgtaaaa 660
gtgcaaggtg aaatcgatgt caaagatctc gcccgcgag aacgcacggg cgaagaggtc 720
gcaaaacgaa tggagcgcg ctctgttctt gcgcaggtgg atatccaccg cgctgcgacc 780
cacaataaag gtgtgatgaa cggcatccat gcagtcgtac ttgccaccgg caacgatact 840
cgcggtgcag aagcgtctgc ccatgcatat gcgtcgcgag atggtcaata ccgtggcatt 900
gctacatggc gttacgatca aaaacgccag cgtctcattg gtacgatcga ggttcctatg 960

accttggcca tcgtcggcgg aggcacaaaa gtctcccg tgcgaaaggc ctgcgtcgaa 1020
cttttgaatg tcgatagcgc acaagaactc ggtcacgtcg tagcagccgt gggtctggcg 1080
cagaacttcg cagcgtgtcg agcacttgc tccgagggca ttcagcaggg tcacatgtcc 1140
cttcagtaca aaagcttggc aattgttgtg ggtgcaaagg gtgatgagat tgcccaggtc 1200
gcagaagcac tcaacaaga gccacgtgcg aacacgcagg ttgccgagcg tatccttcaa 1260
gaaattcgac aacaa 1275

<210> 31

<211> 446

<212> PRT

<213> Candidatus Halobonum tyrrellensis G22

<400> 31

Met Ser Asp Ala Asp Thr Glu Arg Leu Ala Ala Arg Val Arg Glu Gly

1 5 10 15

Asp Leu Arg Leu Tyr Glu Leu Glu Glu His Ala Asp Ala Asp Thr Ala

20 25 30

Val Ala Ala Arg Arg Leu Leu Val Glu Glu Ala Ser Gly Ala Glu Leu

35 40 45

Glu Thr Thr Gly Glu Tyr Ala Phe Pro Ala Glu Met Ala Thr Gly Thr

50	55	60
Asn Val Glu Asn Met Val Gly Ala Val Gln Val Pro Val Gly Val Val		
65	70	75
Gly Pro Val Thr Ile Asp Gly Gly Ala Leu Ser Gly Glu Arg His Val		
	85	90
Pro Met Ala Thr Thr Glu Gly Ala Leu Leu Ala Ser Val Asn Arg Gly		
	100	105
Cys Ser Ala Leu Asn Thr Ala Gly Gly Ala Ser Ala Arg Val Leu Lys		
	115	120
		125
Asn Arg Met Thr Arg Ala Pro Val Phe Arg Val Asp Gly Val Val Glu		
	130	135
		140
Ala Glu Ala Leu Ala Ser Trp Thr Arg Asp Asn Val Glu Ala Leu Arg		
145	150	155
		160
Ala Val Ala Glu Ala Thr Thr Asn His Gly Glu Leu Arg Glu Val Arg		
	165	170
		175
Pro Tyr Ala Val Gly Asn Asn Val Tyr Leu Arg Phe Ala Tyr Asp Thr		
	180	185
		190
Lys Asp Ala Met Gly Met Asn Met Ala Thr Ile Ala Thr Glu Ala Ala		
	195	200
		205
Cys Glu Val Val Glu Asp Glu Thr Gly Ala Glu Leu Val Ala Leu Ser		
	210	215
		220
Gly Asn Leu Cys Ser Asp Lys Lys Pro Ala Ala Val Asn Ala Val Glu		
225	230	235
		240
Gly Arg Gly Arg Thr Val Ala Ala Asp Val Thr Val Pro Arg Glu Thr		
	245	250
		255
Val Arg Glu Thr Leu Lys Thr Thr Pro Glu Ala Val Ala Glu Val Asn		
	260	265
		270
Thr Arg Lys Asn Leu Val Gly Ser Ala Lys Ala Gly Ser Leu Gly Phe		
	275	280
		285
Asn Ala His Val Ala Asn Ala Val Ala Ala Val Phe Leu Ala Thr Gly		
	290	295
		300

Gln Asp Ala Ala Gln Val Val Glu Gly Ala Asn Ala Ile Thr Thr Ala
305 310 315 320

Glu Val Val Ala Gly Asp Ala Pro Ser Arg Thr Val Asp Arg Asn Gly
325 330 335

Ala Ala Glu Ser Gly Val Gly Pro Asp Gly Asp Ala Pro Asp Gly Asp
340 345 350

Ala Pro Asp Gly Asp Gly Ala Ala Ala Gly Arg Gly Gly Asp Leu Tyr
355 360 365

Val Ser Val Thr Leu Ala Ser Leu Glu Val Gly Thr Val Gly Gly Gly
370 375 380

Thr Lys Leu Pro Thr Gln Ala Glu Ala Leu Asp Val Leu Gly Val Arg
385 390 395 400

Gly Gly Gly Asp Pro Ala Gly Ser Asn Ala Asp Ala Leu Ala Glu Ala
405 410 415

Val Ala Val Thr Ala Leu Ala Gly Glu Leu Ser Leu Leu Ser Ala Leu
420 425 430

Ala Ser Arg Asn Leu Ser Ser Ala His Ala Glu Leu Gly Arg
435 440 445

<210> 32

<211> 1338

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 32

atgtcagacg ctgacaccga gcgtctggct gcgcgcgttc gcgagggtga tctgcgcctg 60
tatgagctgg aggagcacgc ggatgcagat actgctgtcg ccgctcgtcg tctcttggtg 120
gaagaggcct ccggtgcaga acttgagact accggtgaat atgctttccc tgcagaaatg 180
gcaaccggca ccaatgtcga gaacatggtc ggtgctgttc aggttccagt gggcgtagtg 240
ggctctgtta cgattgatgg cggtagcgtg tctggcgagc gtcacgtccc aatggctaca 300
accgaaggtg cgctcctcgc gtccgtcaat cgcggctgct cggccctcaa cacagctggt 360

ggtgcatcag ctctgttctt gaagaatcgc atgactcgtg cccccgtctt tcgcgtcgat 420

ggcgtcgtcg aagccgaggc gcttgcttcg tggacccgcg acaacgtgga agcgtccgt 480
gcagtggctg aagctaccac caaccacgga gaactccgcg aagtgcgccc atacgtgtg 540
ggcaacaacg tctaccttcg attcgcttat gataccaagg atgcaatggg catgaatatg 600
gccactatcg ccaccgaggc ggcgtgcgag gtggtggaag atgaaaccgg cgctgagctg 660
gttgactga gcggtaacct ttgctctgat aagaagccag ccgctgtgaa cgcagtggaa 720
ggtcgcggtc gcacagttgc tgctgatgtg acggttcctc gcgagaccgt tcgagagacc 780

ctgaagacca cccctgaggc cgtggccgag gttaacacgc gcaagaactt ggtgggaagc 840
gcgaaggccg gtcattggg atttaatgct catgtggcaa acgccgtggc agcagctttt 900
cttgcaacag gtcaggatgc tgcgcaggtc gtggaagggtg ctaacgcat cactacggcg 960
gaagtcgtcg caggcgatgc tccatcccg accgtggatc gcaatggcg tgctgaatcc 1020
ggcgtaggtc cggatgtga tgcacctgac ggtgatgcac cggatggtga cggcgagca 1080
gcaggtcgtg gcggcgacct ctatgttagc gtgacctgg cctcactgga ggtcggtacc 1140
gtggcgggcg gtacgaaact gccacccaa gcggaagcat tggacgtgct cgggtgtacgt 1200

ggaggaggtg atcctgctgg tagcaacgca gatgctctcg ctgaagccgt tgctgtgacc 1260
gctttggctg gtgagttgtc tcttctttcc gctctcgcca gccgtaacct ctcgtcgga 1320
cacgtgaac tgggacgc 1338

<210> 33

<211> 665

<212> PRT

<213> Cucumis sativus (Cucumber)

<400> 33

Met Thr His Cys Ser Gln Leu Pro Thr His Lys Ala Thr Thr Thr Glu

1 5 10 15

Lys Asn Asp Glu Pro Ile Pro Leu Gly Pro Val Gln Leu Ser Trp Pro

20 25 30

Gly Ser Val Val Glu Lys Gly Glu Lys Ser Ser Phe His Tyr Lys Pro

35 40 45

Pro Phe Ala Ser Arg Ser Gly Ala Gly Val Phe His Phe Leu Phe Phe

50 55 60

Ser Cys Leu Thr Met Asp Ala Arg Arg Arg Arg Ser Ser Thr Leu Val

65 70 75 80

Lys Lys Leu Thr Ala Asp Glu Pro Pro Val Lys Ser Met Asp Lys Asn

85 90 95

Pro Leu Lys Val Lys Met Leu Glu Arg Gln His Val His Asp Asp Ala

100 105 110

Val Lys Ala Ser Asp Val Leu Pro Leu Pro Ile Tyr Leu Thr Asn Ala

115 120 125

Ala Phe Phe Thr Leu Phe Phe Ser Val Val Tyr Phe Leu Leu Thr Arg

130 135 140

Trp Arg Glu Lys Ile Arg Ser Ser Thr Pro Leu His Val Val Thr Leu

145 150 155 160

Ser Glu Met Val Ala Ile Ser Ala Phe Ile Ala Ser Phe Ile Tyr Leu

165 170 175

Leu Gly Phe Phe Gly Ile Asp Phe Val Gln Ser Ile Phe Arg Pro Ser

180 185 190

His Asp Val Trp Thr Ser Glu Asp Asp Glu Val Val Ile Ile Lys Glu

195 200 205

Asp Thr Arg Lys Val Pro Cys Gly Ala Gly Ile Asp Cys Ser Ile Pro

210 215 220

Ile Leu Ala Pro Pro Met Pro Ser Val Pro Lys Val Val Asp Pro Leu

225 230 235 240

Pro Val Ser Ile Asp Leu Thr Glu Glu Asp Glu Glu Ile Val Lys Ser

245 250 255

Val Val Asp Gly Ser Thr Pro Ser Tyr Ser Leu Glu Ser Lys Leu Gly

260 265 270

Asp Cys Gly Arg Ala Ala Ala Ile Arg Arg Val Ala Leu Gln Arg Val

275 280 285

Thr Gly Lys Ser Leu Ser Gly Leu Pro Leu Glu Gly Phe Asp Tyr Ala

290 295 300

Ser Ile Leu Gly Gln Cys Cys Glu Met Pro Ile Gly Tyr Val Gln Ile

305 310 315 320

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

325 330 335
 Val Pro Met Ala Thr Thr Glu Gly Cys Leu Val Ala Ser Thr Asn Arg

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

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

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

 530 535 540
 Tyr Ile Ala Thr Gly Gln Asp Pro Ala Gln Asn Val Glu Ser Ser His
 545 550 555 560
 Cys Ile Thr Met Met Glu Ala Val Asn Asp Gly Gln Asp Leu His Val
 565 570 575

Ser Val Thr Met Pro Ser Ile Glu Val Gly Thr Val Gly Gly Gly Thr

580

585

590

Gln Leu Ala Ser Gln Ser Ala Cys Leu Asn Leu Leu Gly Val Lys Gly

595

600

605

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

610

615

620

Val Ala Ala Ser Val Leu Ala Gly Glu Leu Ser Leu Met Ser Ala Ile

625

630

635

640

Ser Ala Gly Gln Leu Val Arg Ser His Met Lys Tyr Asn Arg Ser Ser

645

650

655

Arg Asp Ile Thr Lys Ala Ser Ser Ser

660

665

<210> 34

<211> 1995

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 34

atgacccatt gctcccagtt gccgaccac aaggcgacga cgactgaaaa aaacgatgag 60
cctattcccc ttggccccgt gcaactgagc tggccgggtt ccgttgttga aaagggagaa 120
aagtcacgt tccattataa gccccattc gcgtctcgca gcggcgagg cgtatttcat 180
tttctctttt tctcgtgctt gactatggac gcccgtcgcc gtcgctcttc gactctggtg 240
aagaaactta ctgcggatga gcctccagta aaatctatgg ataaaaatcc tcttaaagtc 300
aagatgcttg aacgacagca tgtccatgac gatgccgtaa aggcgtcgga tgtgtccca 360

ctgccaatct acctactaa cgctgcattc ttcactctgt tctttccgt ggtatatattc 420
cttcttacgc gctggcgtga gaagattcgt agctccacgc cctgcacgt agtaaccctc 480
tcagaaatgg tggtatctc tgccttcac gccctgttca ttacctgct tggtttttc 540
ggcattgatt tcgtgcagtc catcttcgc cctccacg atgtctggac ctcagaggac 600
gacgaagtgg tcatcatcaa agaagacact cgcaaggtcc catgtggcgc cggaattgac 660
tgttccattc caattctggc cctcctatg cctcgggtgc cgaaggtggt tgatccgctg 720
cctgtgtcca ttgatctcac cgaggaggat gaagagattg tgaagtcagt cgtcgacggc 780

tcgacacctt cttactcatt ggagtcctaag ctcgagatt gtggccgagc ggctgctatc 840
cgactgttg cacttcagcg cgtgactggc aagtcgctgt ctggcctccc gctggagggc 900
ttcgactatg cttccatcct cggacagtgt tgcgagatgc ccattggcta cgtgcaaatt 960
cccgtgggaa ttgccggccc attgcttttg gatgtaagg aatacagcgt ccccatggcg 1020
acgaccgaag gatgccttgt ggcctccacc aaccgaggct gcaaggcgat catgatctcc 1080
ggaggcgcta actccgtcct ctttcgtgat gccatgacac gtgcgccagt gatgcgtttt 1140
gcgactgcta agcgcgccgc ggaactcaaa ttctacgttg aagaccggc taacttcgat 1200

acccttgect ccgtatttaa caaatctagc cgctttggac gcctgcagag catcaaatgt 1260
gccatcgccg gcaaaaacct ctatatgca ttttcgtgct ccactggcga cgctatggga 1320
atgaatatgg ttccaaggg cgtacagaac gtcttgact ttctccagaa tgattttcca 1380
gacatggatg taatcggaat cagcggaaac tattgttcag ataagaagcc agccgccgtt 1440
aattggatcg aaggacgcgg caagtcctc gtgtgcgagg ttaccatcaa aggcgatgtt 1500
gtccgaagg tctgaagac agatgtccag gactgggtgg aactcaacat gttgaaaaat 1560
cttacgggtt ccgcaatggc tggcgactt ggtggcttta acgtcatgc atccaacatc 1620

gtttcagcca tctacatgc gacgggacag gatcccgctc aaaacgtgga aagctcccat 1680
tgcatcacta tgatggagcg cgtaaacgat ggtcaggatc ttacgtgag cgtgaccatg 1740
ccatccatcg aggtgggaac cgtaggcggt ggacgcaac tggcatcca gtccgcctgc 1800
ctgaacctct tgggtgttaa gggcgcaat cgcgaagctc caggttcaaa ttctcgcttg 1860
ctcgccacca ttgtggccgc atcagttctt gcaggcgaac ttagcctcat gtcagcgatc 1920
tccgccggcc aactcgccg cagcccatg aaatacaatc gttcatcag cgatatcacc 1980
aaggcttcat catcc 1995

<210> 35

<211> 428

<212> PRT

<213> Lokiarchaeum sp. GC14_75

<400> 35

Met Ile Met Lys Asp Phe His Ser Asp Ile Ser Gly Phe Tyr Lys Leu

1 5 10 15

Ser Ile Asp Glu Arg Gln Lys Leu Leu Ser Lys Leu Val Asn Leu Asn

20 25 30

Pro Glu Asp Leu Glu Ile Leu Lys Glu Leu Gly Tyr Phe Thr Pro Thr

35 40 45

Gln Ile Asp Thr Leu Ile Glu Asn Val Val Gly Ser Tyr Gln Leu Pro

50

55

60

Leu Gly Leu Ala Phe Asn Phe Arg Ile Asn Ser Lys Asp Tyr Ile Ile

65

70

75

80

Pro Met Val Ile Glu Glu Pro Ser Val Val Ala Ala Ala Ser Asn Ala

85

90

95

Ala Lys Met Ala Arg Lys His Gly Gly Phe His Ser Glu Asp Val Pro

100

105

110

Pro Ile Met Ile Ser Gln Ile Gln Ile Thr Lys Leu Lys Asp Ile Glu

115

120

125

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

130

135

140

Ala Asn Glu Gln Asp Pro Leu Leu Asn Lys Leu Gly Gly Gly Ala Gln

145

150

155

160

Asn Ile Glu Ile His Glu Ile Gln Thr Asn Lys Gly Lys Met Ile Ile

165

170

175

Leu His Leu Leu Val Asn Val Leu Asp Ala Met Gly Ala Asn Ile Val

180

185

190

Asn Thr Met Ala Glu Ala Ile Ser Pro Tyr Ile Glu Glu Ile Cys Gly

195

200

205

Gly Lys Ile Tyr Leu Arg Ile Val Ser Asn Leu Ala Thr His Arg Val

210

215

220

Ala Arg Ser Lys Ala Thr Phe Asp Lys Glu Met Leu Gly Gly Glu Glu

225

230

235

240

Val Val Glu Gly Ile Met Asn Ala Tyr Glu Phe Ala Leu Ala Asp Pro

245

250

255

Tyr Arg Ala Thr Thr His Asn Lys Gly Ile Met Asn Gly Val Val Ala

260

265

270

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

275

280

285

His Ser Tyr Ala Ser Leu Ser Gly Arg Tyr Leu Pro Leu Thr Lys Phe

290 295 300
Asp Val Asp Ser Gln Gly Asn Leu Val Gly Glu Ile Glu Ile Pro Leu

305 310 315 320
Ala Leu Gly Ile Ile Gly Gly Met Thr Lys Ile His Pro Met Ala Arg

325 330 335
Leu Ala Leu Lys Ile Leu Asn Leu Lys Ser Ser Ser Glu Leu Ala Gln

340 345 350
Val Gly Ala Ala Leu Gly Leu Ala Gln Asn Val Ala Ala Leu Arg Ala

355 360 365
Leu Ala Ala Glu Gly Ile Gln Lys Gly His Met Ala Leu His Ser Arg

370 375 380
Asn Ile Ala Lys Val Ala Gly Val Pro Asp Glu Leu Ile Glu Lys Val

385 390 395 400
Ser Lys Lys Met Val Ala Asp Lys Lys Ile Arg Val Asp Tyr Ala Lys

405 410 415
Glu Ile Phe Gln Lys Ile Lys Asp Gly Glu Asn Ile

420 425

<210> 36

<211> 1284

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 36

atgatcatga aagacttcca ctccgatatt tccggtttct acaagctgtc catcgatgag	60
cgccagaagc ttttgagcaa gcttgtaaac ctgaaccggg aagatctgga aattcttaag	120
gaactgggct atttcacccc gactcaaatt gacaccctga tcgagaacgt tgtcggctcc	180
tatcaacttc ccttgggatt ggccttcaac ttctgcatca attctaagga ttacatcatc	240
cctatggtga tcgaagagcc atccgtggta gctgcagcat ccaatgcggc taagatggct	300
cgcaaacacg gcggttttca ctctgaagat gttctccaa tcatgatctc acagattcaa	360
atcaccaaat tgaaagatat cgaggcagct aaagatgtcc tcgaggcgaa gaagcaggat	420
ttgctcaaga ttgccaacga acaagatccg ctgctgaaca aacttggcgg tggcgcccag	480

aacatcgaaa tccacgaaat ccaaactaac aagggaaga tgatcatcct ccattctgctc 540
 glaaacgttc tcatgcatc ggggtccaac atcgtcaata ccatggcgga ggctatctct 600
 ccatacatcg aagagatctg cgggtggcaag atctaccttc gcatgttttc caacctggca 660
 acccaccgtg tggcgcgttc caaggcaacc ttcgataagg aaatgctggg cggcgaagag 720
 gtggtggaag gcatcatgaa tgcataatgaa ttgcgcgtgg cggacccta ccgcgtacc 780
 acccataaca agggatatcat gaatgggtgtt gtggctttga ctctggcgac gggtaatgac 840

acccgcgcaa tcgagagcgg agcccactct tacgcatccc tttccggccg ttacctgccc 900
 ctgaccaagt ttgacgtaga ttctcaaggt aacctgtgtg gagaaatcga gattcctctt 960
 gccctgggca ttatcgggtg tatgaccaag atccatccaa tggcccgtct ggctttgaag 1020
 attcttaacc ttaaatcatc ttcagaactt gcacaggtag gtgcggcgct gggacttgcc 1080
 caaaacgttg cagctctcgc gcgcctcgcc gctgaaggta tccagaaagg tcacatggca 1140
 ctccactcac gtaatatgac aaaggctcga ggcgtccctg acgaactgat tgaataagt 1200
 tctaaaaaga tggctgccga taaaagatt cgtgtggatt atgcaaagga aatcttcag 1260

aagatcaagg acggcgaaaa catt 1284

<210> 37

<211> 931

<212> PRT

<213> Cricetulus griseus

<400> 37

Met Thr Pro Asp Met Arg Lys Thr Gly Arg Cys Glu Ala Phe Leu Glu

1 5 10 15

Gln Arg Ser Ile Leu Leu Pro Asn Arg Phe Leu Phe Ser Gly His Pro

20 25 30

His Ala Ala Thr Leu Thr Leu Pro Gln Glu Gly Pro Ser Gln Val Arg

35 40 45

Phe Arg Gly Arg Gly Thr Pro Pro Ala Gly Ser Pro Ile Ala Arg Glu

50 55 60

Gly Gly Val Leu Ala Gly Ala Pro His Gly Ser Gly Thr Asn Lys Lys

65 70 75 80

Val Val Met Leu Glu Leu Asp Gln Leu Leu Val Gly Ser Ala Val Val

85 90 95

Arg Asp Gly Ala Val Pro Val Leu Arg Pro Gly Ser Arg Asp Arg Val

100	105	110	
Ala Thr Met Leu Ser Arg Leu Phe Arg Met His Gly Leu Phe Val Ala			
115	120	125	
Ser His Pro Trp Glu Val Ile Val Gly Thr Val Thr Leu Thr Ile Cys			
130	135	140	
Met Met Ser Met Asn Met Phe Thr Gly Asn Asn Lys Ile Cys Gly Trp			
145	150	155	160
Asn Tyr Glu Cys Pro Lys Phe Glu Glu Asp Val Leu Ser Ser Asp Ile			
165	170	175	
Ile Ile Leu Thr Ile Thr Arg Cys Ile Ala Ile Leu Tyr Ile Tyr Phe			
180	185	190	
Gln Phe Gln Asn Leu Arg Gln Leu Gly Ser Lys Tyr Ile Leu Gly Ile			
195	200	205	
Ala Gly Leu Phe Thr Ile Phe Ser Ser Phe Val Phe Ser Thr Val Val			
210	215	220	
Ile His Phe Leu Asp Lys Glu Leu Thr Gly Leu Asn Glu Ala Leu Pro			
225	230	235	240
Phe Phe Leu Leu Leu Ile Asp Leu Ser Arg Ala Ser Ala Leu Ala Lys			
245	250	255	
Phe Ala Leu Ser Ser Asn Ser Gln Asp Glu Val Arg Glu Asn Ile Ala			
260	265	270	
Arg Gly Met Ala Ile Leu Gly Pro Thr Phe Thr Leu Asp Ala Leu Val			
275	280	285	
Glu Cys Leu Val Ile Gly Val Gly Thr Met Ser Gly Val Arg Gln Leu			
290	295	300	
Glu Ile Met Cys Cys Phe Gly Cys Met Ser Val Leu Ala Asn Tyr Phe			
305	310	315	320
Val Phe Met Thr Phe Phe Pro Ala Cys Val Ser Leu Val Leu Glu Leu			
325	330	335	
Ser Arg Glu Ser Arg Glu Gly Arg Pro Ile Trp Gln Leu Ser His Phe			
340	345	350	

Ala Arg Val Leu Glu Glu Glu Glu Asn Lys Pro Asn Pro Val Thr Gln
355 360 365

Arg Val Lys Met Ile Met Ser Leu Gly Leu Val Leu Val His Ala His
370 375 380

Ser Arg Trp Ile Ala Asp Pro Ser Pro Gln Asn Ser Thr Thr Glu His
385 390 395 400

Ser Lys Val Ser Leu Gly Leu Asp Glu Asp Val Ser Lys Arg Ile Glu
405 410 415

Pro Ser Val Ser Leu Trp Gln Phe Tyr Leu Ser Lys Met Ile Ser Met
420 425 430

Asp Ile Glu Gln Val Val Thr Leu Ser Leu Ala Phe Leu Leu Ala Val
435 440 445

Lys Tyr Ile Phe Phe Glu Gln Ala Glu Thr Glu Ser Thr Leu Ser Leu
450 455 460

Lys Asn Pro Ile Thr Ser Pro Val Val Thr Pro Lys Lys Ala Pro Asp
465 470 475 480

Asn Cys Cys Arg Arg Glu Pro Leu Leu Val Arg Arg Ser Glu Lys Leu
485 490 495

Ser Ser Val Glu Glu Glu Pro Gly Val Ser Gln Asp Arg Lys Val Glu
500 505 510

Val Ile Lys Pro Leu Val Val Glu Thr Glu Ser Ala Ser Arg Ala Thr
515 520 525

Phe Val Leu Gly Ala Ser Gly Thr Ser Pro Pro Val Ala Ala Arg Thr
530 535 540

Gln Glu Leu Glu Ile Glu Leu Pro Ser Glu Pro Arg Pro Asn Glu Glu
545 550 555 560

Cys Leu Gln Ile Leu Glu Ser Ala Glu Lys Gly Ala Lys Phe Leu Ser
565 570 575

Asp Ala Glu Ile Ile Gln Leu Val Asn Ala Lys His Ile Pro Ala Tyr
580 585 590

Lys Leu Glu Thr Leu Met Glu Thr His Glu Arg Gly Val Ser Ile Arg

595 600 605
 Arg Gln Leu Leu Ser Thr Lys Leu Pro Glu Pro Ser Ser Leu Gln Tyr
 610 615 620

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

 Gly Gly Gly Ala Ser Ser Arg Val Leu Ala Asp Gly Met Thr Arg Gly
 690 695 700
 Pro Val Val Arg Leu Pro Arg Ala Cys Asp Ser Ala Glu Val Lys Ala
 705 710 715 720
 Trp Leu Glu Thr Pro Glu Gly Phe Ala Val Ile Lys Asp Ala Phe Asp
 725 730 735
 Ser Thr Ser Arg Phe Ala Arg Leu Gln Lys Leu His Val Thr Met Ala
 740 745 750

 Gly Arg Asn Leu Tyr Ile Arg Phe Gln Ser Lys Thr Gly Asp Ala Met
 755 760 765
 Gly Met Asn Met Ile Ser Lys Gly Thr Glu Lys Ala Leu Leu Lys Leu
 770 775 780
 Gln Glu Phe Phe Pro Glu Met Gln Ile Leu Ala Val Ser Gly Asn Tyr
 785 790 795 800
 Cys Thr Asp Lys Lys Pro Ala Ala Ile Asn Trp Ile Glu Gly Arg Gly
 805 810 815

 Lys Thr Val Val Cys Glu Ala Val Ile Pro Ala Lys Val Val Arg Glu
 820 825 830
 Val Leu Lys Thr Thr Thr Glu Ala Met Ile Asp Val Asn Ile Asn Lys
 835 840 845

Asn Leu Val Gly Ser Ala Met Ala Gly Ser Ile Gly Gly Tyr Asn Ala
850 855 860

His Ala Ala Asn Ile Val Thr Ala Ile Tyr Ile Ala Cys Gly Gln Met
865 870 875 880

Leu Gly Val Gln Gly Ala Cys Lys Asp Asn Pro Gly Glu Asn Ala Arg
885 890 895

Gln Leu Ala Arg Ile Val Cys Gly Thr Val Met Ala Gly Glu Leu Ser
900 905 910

Leu Met Ala Ala Leu Ala Ala Gly His Leu Val Arg Ser His Met Val
915 920 925

His Asn Arg
930

<210> 38

<211> 2793

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 38

atgacccctg acatgcgcaa gactggccgc tgcgaggcat tcttgagca acgtccatc 60
ttgtcccta atcgcttctt cttttcagga caccacacg cagcaacctt gacacttctt 120
caggaaggcc caagccaggt ccgctttcgc ggacgaggca cgccgcccgc cggaagcccc 180
atcgcacgcg aaggcgggtgt cctcgaggt gctctcacg gctctggcac taataaaaaa 240
gtcgtgatgt tggaattgga ccagttgctg gtgggctctg ccgtcgtgcg tgacggtgct 300
gtacctgtgt tgcgtctgg ctcacgcgac cgcgtcgcca cgatgctgtc tcgcttggtc 360
cgcatgcatg gactttttgt ggctttctac ccgtgggaag tgatcgtggg tacggtaaca 420

ctcaccatct gcatgatgtc catgaatatg ttaccggta acaataaaat ttgtggctgg 480
aattatgagt gccc aaagt cgaagaggac gtcttgcct ctgatcatc cattttgacc 540
atcacccgtt gcatcgtat tctgtacatc tatttcaggt tccaaaacct gcgtcaactg 600
ggttcaaagt acatccttgg tatcgcggtt cttttcaca ttttttcgtc cttcgtgttc 660
agcactgtgg ttatccactt tttggacaag gaattgacag gacttaatga ggcgtccct 720
ttcttctcc tgcgtattga cttttcccga gcgtcagcat tggcgaaatt tgcattgtcg 780

tctaattctc aagatgaagt ccgagagaac atcgctcgcg gtatggcgat cctgggtccg	840
accttcacgc tcgatgctct ggtagagtgc ctgtcatcg gttaggcac catgagcggg	900
gtccgccagt tggagattat gtgttgcttt ggttgatgt ctgtgctcg caattacttc	960
gtattcatga ctttttccc tgcctgcgtg tctctcgcc tcgaactctc ccgagagtc	1020
cgcgagggcc gtccaatttg gcagctgtca cacttcgcac gagtcccgga ggaagaggaa	1080
aataagccaa acccagttac acagcgtgtt aagatgatca tgagcctggg cctggttctc	1140
gttcacgctc acagccgttg gatcgctgat ccgtctcccc aaaactccac gaccgaacac	1200
tccaaggttt ctcttggttt ggatgaagac gtttcaaagc gcatcgagcc ctccgtttcc	1260
ctctggcagt ttiacctgac caagatgatt tcgatggata tcgagcaggt cgtgaccctt	1320
agcttggcgt tccttttggc agttaagtac atcttcttcg agcaggcaga aacggagtct	1380
acattgtcac tcaagaacct tatcacgtca ccagtggatg ccccaaaaaa agcacctgat	1440
aattgttgcc gccgtgaacc gctgcttgtt cgccgttccg agaagttgag cagcgtggaa	1500
gaggaaccgg gcgtctccca agaccgcaaa gttgaagtga ttaagccgct cgtcgtcgag	1560
acagaatccg ctcccgccgc tacttttgtc ttggcgcat ctggaacatc cctcctgtg	1620
gcggcacgta cacaggagt ggagattgag ctgccatctg aaccacgtcc caacgaagag	1680
tgcttgacga tcttgagtc cgccgaaaag ggccgcaagt tctctctga cgccgaaatc	1740
atccagcttg tgaatgcaaa acacattcca gcatacaaat tggaaacgt gatggaaact	1800
catgaacgcg gcgtctctat tcgccgacag ctctcttagca ctaaaactccc ggagccatcg	1860
tcgttgacgt atctccata tcgagactac aactactcct tggatgatgg agcatgctgt	1920
gaaaacgtca tcggctacat gccgatccca gtaggcgtgg caggtccact ttgccttgat	1980
ggcaaggagt accaggtccc catgctacc acggagggt gccttggtgc atccacaaac	2040
cgtggctgtc gcgcgatcgg tcttggcgga ggagcatcat cgcgtgtact tgctgatggt	2100
atgacccgcg gccctgtagt tcgtctcccc cgcgctgctg atagcgtga ggttaaggcc	2160
tggttggaaa cacctgaggg atttgacgtg atcaaggacg cctttgactc tacctcacgc	2220
ttcgcgctg tgcagaagct ccacgttacc atggcgggtc gcaacttgta tatccgttc	2280
cagtcaaaaa ccggagacgc tatgggaatg aatatgatct ccaaaggtac cgaaaaagct	2340
ctctgaagc tccaggaatt ctccccagaa atgcaaattc tcgcggttag cggaactat	2400
tgcaccgaca agaagccagc agccatcaac tggatcgagg gtcgaggtaa gaccgtgtc	2460
tgcgaggccg ttattccggc taagtggtc cgtgaggtgc tgaagaccac gacggaagcg	2520

atgatcgatg tgaacatcaa caagaacctc gtcggctccg cgatggcggg ttccattggc 2580
 ggttacaatg ctcacgtgc taatatcgta actgctatgt atattgcttg cggtcagatg 2640
 ctggcggttc aaggtgcatg caaggataac ccaggtgaga acgcccgtca gttggcacgt 2700
 atcgtaatcg gcaccgtgat ggcaggcgaa ctttccctta tggccgcact ggcggctggt 2760
 cacctcgtcc gttcacacat ggttcataac cgc 2793

<210> 39

<211> 428

<212> PRT

<213> Pyrodictium occultum

<400> 39

Met Gly Glu Glu Arg Ser Leu Arg Asp Arg Leu Glu Glu Val Val Gln

1 5 10 15
 Gly Ile Val Glu Gly Arg Ile Arg Leu His Glu Ala Asp Arg Leu Leu
 20 25 30
 Gly Asn Ala Asn Ala Ala Ala Leu Ala Arg Arg Leu Ala Leu Glu Arg
 35 40 45
 Met Leu Gly Val Gly Leu Ser Ser Ile Gly Ser Thr Ile Leu Asp Phe
 50 55 60
 Glu Glu Leu Val Gly Arg Asn Ile Glu Asn Pro Ile Gly Ala Val Gln

65 70 75 80
 Ile Pro Leu Gly Val Ala Gly Pro Leu Arg Val His Gly Glu Tyr Ala
 85 90 95
 Arg Gly Asp Phe Tyr Ile Pro Leu Ala Thr Thr Glu Gly Ala Leu Val
 100 105 110
 Ala Ser Val Asn Arg Gly Ala Lys Ala Val Thr Leu Ser Gly Gly Ala
 115 120 125
 Arg Ala Arg Val Leu Arg Asp Gly Met Ala Arg Ala Pro Val Phe Trp

130 135 140
 Thr Pro Gly Val Glu Glu Ala Ala Arg Leu Ala Glu Trp Val Gln Glu
 145 150 155 160
 His Met Glu Glu Val Arg Arg Glu Ala Glu Ser Thr Thr Arg His Gly
 165 170 175

Arg Leu Leu Glu Ile Gln Pro Phe Ile Ala Gly Asn Ile Val Trp Leu
180 185 190

Arg Phe Val Tyr Ser Thr Gly Asp Ala Met Gly Met Asn Met Ala Thr
195 200 205

Ile Ala Thr Asp Lys Ala Ala Glu Trp Ile Leu Arg Asn Tyr Pro Gly
210 215 220

Glu Ala Lys Leu Ile Ala Ile Ser Gly Asn Leu Cys Thr Asp Lys Lys
225 230 235 240

Pro Ala Leu Leu Asn Ala Ile Leu Gly Arg Gly Lys Thr Val Val Ala
245 250 255

Glu Ala Thr Ile Lys Arg Asp Ile Ala Leu Arg Val Leu Lys Ala Ala
260 265 270

Pro Glu Asp Ile Asp Thr Val Asn Arg Val Lys Asn Leu Leu Gly Ser
275 280 285

Ala Arg Ala Gly Ser Pro Ser Phe Asn Ala His Tyr Ala Asn Ile Ile
290 295 300

Ala Ala Ile Phe Ile Ala Thr Gly Gln Asp Val Ala Gln Val Val Glu
305 310 315 320

Ser Ser Met Gly Tyr Thr Trp Thr Glu Val Arg Asn Gly Asp Leu Tyr
325 330 335

Ile Ser Val Thr Leu Pro Ser Leu Glu Leu Gly Thr Val Gly Gly Gly
340 345 350

Thr Arg Leu Pro Thr Gln Arg Glu Ala Leu Ala Leu Gly Ala Ala
355 360 365

Gly Gly Gly Asn Pro Pro Gly Ala Asn Ala Lys Lys Leu Ala Glu Val
370 375 380

Thr Ala Ala Ala Val Leu Ala Gly Glu Leu Asn Leu Leu Ala Ala Leu
385 390 395 400

Ala Ala Asn Glu Leu Ala Arg Ala His Arg Leu Leu Gly Arg Gly Glu
405 410 415

Ala Lys Thr Gly Asn Lys Asn Pro Thr Asn His Gly

420

425

<210> 40

<211> 1284

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 40

atgggagaag agcgctctct gcgcgatcgc ctggaagagg tagtccaagg catcgtagag 60

ggtcgcatcc gcctgcataga agcagaccgc ctgctgggaa atgcaaatgc agcagcgttg 120

gctcgccgcc tggcactgga acgcatgctt ggagttggct tgtccagcat tggatcgacg 180

atattggact tcgaagagct tgttggccgt aatattgaaa acccgatcgg tgccgtgcag 240

atccccctcg gtgtcgccgg cccgctgcgc gtccacggcg agtatgctcg cggtgacttc 300

tacattccac ttgctaccac tgaaggcgca cttgtggcat cagtgaatcg cggcgcaaag 360

gctgtgacgc tctctggcgg tgcacgcgcc cgcgtacttc gagatggcat ggcccagaca 420

ccggtgtttt ggactcctgg tgtggaagaa gcagcgcgac tcgcggaatg ggtgcaagag 480

catatggaag aagtcctgctg tgaagcagag tccacgacgc gtcacggacg tctcctcgaa 540

atccagccgt tcacgctgg caacatcggt tggcttcgct tcgtctattc taccggcgac 600

gccatgggca tgaacatggc aaccattgca accgataagg cggcagagtg gatcctgcgc 660

aattaccgga gagaagcgaa actcatcgca atctccggca acctctgcac agacaagaag 720

cctgctctct tgaacgctat tctgggccgc ggtaaaaccg tggtcgcgga ggcaactatt 780

aagcgcgaca tcgattcgcg cgtgtttaa gcagctcccg aggacatcga tactgtgaac 840

cgtgtgaaaa atcttttggg cagcgtcgc gccggttcac cttcgtttaa cgcacactat 900

gctaacaatc ttgcagcaat cttcatcgcg acaggacagg atgtggcgca ggtagtggag 960

tccagcatgg gatacacctg gaccgaagtt cgcaatggag acctttacat ttccgtgacg 1020

ctgccctcct tggaactggg caccgttggc ggtggtaccc gcctgcccac acagcgcgag 1080

gcactcgccc tcctgggagc agcaggcggc ggtaaccac cgggtgcaa tgcgaagaag 1140

ttggcagagg ttacagccgc agccgtgctg gccggcgaac tcaatttgtt ggccgctctt 1200

gcagcgaatg aactcgccc cgcgcaccgt ttgctgggtc gtggcgaagc gaagactggt 1260

aataaaaatc caaccaacca cggt 1284

<210> 41

<211> 352

<212> PRT

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 41

```

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

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

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

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

```

Ala Arg Asp Val Leu Ile Pro Leu Gly Glu Tyr Phe Gln Ile Gln Asp
 225 230 235 240
 Asp Tyr Leu Asp Cys Phe Gly Thr Pro Glu Gln Ile Gly Lys Ile Gly

245 250 255
 Thr Asp Ile Gln Asp Asn Lys Cys Ser Trp Val Ile Asn Lys Ala Leu
 260 265 270
 Glu Leu Ala Ser Ala Glu Gln Arg Lys Thr Leu Asp Glu Asn Tyr Gly
 275 280 285
 Lys Lys Asp Ser Val Ala Glu Ala Lys Cys Lys Lys Ile Phe Asn Asp
 290 295 300
 Leu Lys Ile Glu Gln Leu Tyr His Glu Tyr Glu Glu Ser Ile Ala Lys

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

<210> 42

<211> 1056

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 42

atggcgtcgg aaaaagaaat ccgccgcgag cgcttcctga atgtgtttcc aaagctggtc 60

gaggaattga acgcttcgct gcttgcgtat ggcatgccga aagaggcgtg cgattggtac 120

gcacattccc tgaactataa caccacaggc ggtaaactca accgcggtct gtccgttgta 180

gatacatatg ccattcttag caacaagacc gtagaacagt tgggacaaga ggaatatgaa 240

aaggtggcta tcttgggctg gtgtatcgaa ctgttgcaag cctacttcct ggtcgccgat 300

gatatgatgg ataagtccat cacacgacgt ggacagccat gttggtacaa ggtccctgaa 360

gtcggcgaaa tcgctatcaa cgacgcattc atgctcgagg cagcaattta caaacitttg 420

aaaagccatt tccgaaacga gaagtattac atcgacatta ccgagctctt ccatgaagtc 480

accttcaga ccgaactcgg ccaattgatg gatctgatta ccgccccgga ggacaaggtc 540

gatctttcca aattcagcct gaagaacat tccttcacg taacctttgg tacagcatac 600
tattcattct atcttctgt ggcaactcga atgtacgtag ccggtatcac ggatgaaaag 660
gatctgaaac aagcacgcga cgtactcatt cccctgggtg agtactttca gattcaagat 720
gactacctcg attgcttcgg cacgcctgaa caaatcggca aaatcggtag tgacatccag 780
gataataaat gctcctgggt gatcaataag gccctcgagt tggcgtcagc cgagcagcgc 840
aagactctgg acgaaaatta cggtaaaaaa gattcagttg ccgaggcaaa gtgcaaaaag 900

atctttaacg atcttaagat tgaacagctt taccacgaat acgaagaaag cattgctaag 960
gatctgaaag ctaagatttc tcaggtcgat gaatcacgcg gattcaaggc cgatgtgctc 1020
actgctttcc tgaacaaggt ctacaagcgc tcaaaa 1056

<210> 43

<211> 296

<212> PRT

<213> Abies grandis

<400> 43

Met Phe Asp Phe Asn Lys Tyr Met Asp Ser Lys Ala Met Thr Val Asn

1 5 10 15

Glu Ala Leu Asn Lys Ala Ile Pro Leu Arg Tyr Pro Gln Lys Ile Tyr

20 25 30

Glu Ser Met Arg Tyr Ser Leu Leu Ala Gly Gly Lys Arg Val Arg Pro

35 40 45

Val Leu Cys Ile Ala Ala Cys Glu Leu Val Gly Gly Thr Glu Glu Leu

50 55 60

Ala Ile Pro Thr Ala Cys Ala Ile Glu Met Ile His Thr Met Ser Leu

65 70 75 80

Met His Asp Asp Leu Pro Cys Ile Asp Asn Asp Asp Leu Arg Arg Gly

85 90 95

Lys Pro Thr Asn His Lys Ile Phe Gly Glu Asp Thr Ala Val Thr Ala

100 105 110

Gly Asn Ala Leu His Ser Tyr Ala Phe Glu His Ile Ala Val Ser Thr

115 120 125

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

130 135 140

Gly Arg Ala Thr Gly Ser Glu Gly Val Met Gly Gly Gln Met Val Asp
145 150 155 160

Ile Ala Ser Glu Gly Asp Pro Ser Ile Asp Leu Gln Thr Leu Glu Trp
165 170 175

Ile His Ile His Lys Thr Ala Met Leu Leu Glu Cys Ser Val Val Cys
180 185 190

Gly Ala Ile Ile Gly Gly Ala Ser Glu Ile Val Ile Glu Arg Ala Arg
195 200 205

Arg Tyr Ala Arg Cys Val Gly Leu Leu Phe Gln Val Val Asp Asp Ile
210 215 220

Leu Asp Val Thr Lys Ser Ser Asp Glu Leu Gly Lys Thr Ala Gly Lys
225 230 235 240

Asp Leu Ile Ser Asp Lys Ala Thr Tyr Pro Lys Leu Met Gly Leu Glu
245 250 255

Lys Ala Lys Glu Phe Ser Asp Glu Leu Leu Asn Arg Ala Lys Gly Glu
260 265 270

Leu Ser Cys Phe Asp Pro Val Lys Ala Ala Pro Leu Leu Gly Leu Ala
275 280 285

Asp Tyr Val Ala Phe Arg Gln Asn
290 295

<210> 44

<211> 888

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 44

atgttcgact tcaacaaata catggacagc aaggcaatga ccgtgaacga ggcgctgaat 60
aaagctattc ctttgcgtta tccgcagaag atctacgaat ccatgcgata ctctttgttg 120
gcaggtggta agcgagtgcg ccccgctgctt tgcatcgcg cttgagagct ggctcggtggt 180
actgaggaac tggcgatccc aaccgcttgt gctatcgaga tgattcacac gatgtccctt 240
atgcacgatg acctgccctg tatcgacaac gatgatttgc gccgaggcaa gccaacaaat 300

cataagatct tcggcgagga caccgccgtc actgccggca atgctctgca ctcgatatgcc 360
 ttcgagcaca ttgccgtgtc tacttccaaa acggtgggtg cagatcgcat tctgcgcatg 420
 gtatccgagt tgggtcgcg c aactggctcg gagggcgta tgggtggta aatggtcgac 480
 atcgcatctg aaggcgatcc atcgatcgac ctgcagacat tggaatggat ccacatccac 540
 aagactgcga tgccttttga atgttccgtg gtgtgcggag ctattattgg cggagcatct 600
 gagatcgtga tcgagcgcg tcgccgttat gctcgttgcg tgggtttgct gtttcaggtc 660
 gtggacgaca ttctcgatgt caccaaatcc tccgatgagt tgggtaaaac tgccggcaaa 720

gacttgattt ctgacaaggc gacctacccc aagcttatgg gtctcgaaaa agccaaggag 780
 ttctcagatg aactcctcaa ccgcgcaaag ggagaactct catgctttga tccagtgaag 840
 gcagcccat tgcttggttt ggctgattac gtggccttcc gccagaat 888

<210> 45

<211> 297

<212> PRT

<213> Picea abies

<400> 45

Met Glu Phe Asp Phe Asp Lys Tyr Met His Ser Lys Ala Ile Ala Val

1 5 10 15

Asn Glu Ala Leu Asp Lys Val Ile Pro Pro Arg Tyr Pro Gln Lys Ile

20 25 30

Tyr Glu Ser Met Arg Tyr Ser Leu Leu Ala Gly Gly Lys Arg Val Arg

35 40 45

Pro Ile Leu Cys Ile Ala Ala Cys Glu Leu Met Gly Gly Thr Glu Glu

50 55 60

Leu Ala Met Pro Thr Ala Cys Ala Ile Glu Met Ile His Thr Met Ser

65 70 75 80

Leu Ile His Asp Asp Leu Pro Tyr Ile Asp Asn Asp Asp Leu Arg Arg

85 90 95

Gly Lys Pro Thr Asn His Lys Val Phe Gly Glu Asp Thr Ala Ile Ile

100 105 110

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

115 120 125

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

130 135 140
Ile Gly Arg Ala Thr Gly Ser Glu Gly Val Met Gly Gly Gln Val Val
145 150 155 160

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

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

Ala Asp Tyr Ile Ala Ser Arg Gln Asn
290 295

<210> 46

<211> 891

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 46

atggaatttg atttcgataa gtacatgcac agcaaagcaa ttgccgtcaa cgaggetctc 60
gataaagtta tcccacctcg ttaccacacag aaaatctacg aaagcatgcg ctattccctg 120
ctggcaggcg gcaaacgcgt acgccaatc ctctgcatcg cggcctgcga acttatgggc 180
ggcaccgagg agctcgctat gccaaaccga tgcgcgattg agatgatcca caccatgagc 240

cttattcacg atgatctgcc atatatcgac aacgacgatt tgcgccgtgg aaagcctacc 300

aaccacaaag tcttcggaga ggacaccgcc atcattgctg gcgacgtctt gctgtccctg 360

gcttttgagc acgttgagct cagcacctcc cgcacgtctg gcacggatat taccctgcgc 420

ctcctcagcg aaatcggctg cgctaccggt tccgaagggt tgatgggcgg tcagggtggtg 480

gatatcgaat ctgaaggcga tcccagcatc gaccttgaaa cgctggaatg ggtccatatac 540

cataagaccg cggttctgct ggaatgctct gttgtttgcg gtgcaatcat gggcggcgca 600

tcagaggatg acatcgagcg cgcacgtcgc tacgctcggt gcgtaggctt gttgtttcaa 660

gtcgtggacg atatcctcga tgtatcccag tcttcgaag agctgggtaa gactgcgggc 720

aaggatctga tcagcgataa agccacctac ccaaaactta tgggtcttga gaaagcgaag 780

gaattcgctg acgaactgct gaaccgcggc aagcaggagc tgtcctgctt tgatccaacc 840

aaagccgcac cgtctcttcg tctggcagat tacattgcga gccgtcaaaa c 891

<210> 47

<211> 355

<212> PRT

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 47

Met Thr Pro Asp Val Thr Asn Asp Ala Ser Gln Glu Thr Val Asp Leu

1 5 10 15

Val Arg Arg Thr Ser Ala Asp Leu Leu Arg Arg Val Glu Ala Arg Ile

20 25 30

Gln Ser Leu Leu Ser Ala Gly His Asp Thr Trp Ala Ala Val Asp Glu

35 40 45

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

50 55 60

Gly Lys Arg Ile Arg Pro Ala Phe Cys Ile Ala Gly His Leu Ala Ala

65 70 75 80

Gly Gly Asp Pro Asp Asp Ser Gly Val Val Ala Ala Gly Ala Ala Leu

85 90 95

Glu Met Leu His Ala Ser Leu Leu Val His Asp Asp Ile Leu Asp Asp

100 105 110

Ser Ser Gln Arg Arg Gly Leu Pro Thr Val His Ala Lys His Thr Ala
 115 120 125
 Leu His Gln Asp Leu Gly Trp Gln Gly Glu Ser Arg Arg Phe Gly Glu
 130 135 140
 Gly Val Gly Leu Leu Val Gly Gly Leu Ala Leu Thr Tyr Ala Asp Glu

 145 150 155 160
 Leu Met Cys Glu Ala Pro Ala Thr Ala Ile Ala Glu Trp Asn Lys Leu
 165 170 175
 Arg Ser Glu Val Leu Ile Gly Gln Tyr Met Asp Val Val Ala Ala Ala
 180 185 190
 Glu Phe Ser Val Asp Pro Arg Leu Ser Arg Leu Ile Ala Val Ile Lys
 195 200 205
 Ser Gly Arg Tyr Thr Ile His Arg Pro Leu Val Ile Gly Ala Asn Val

 210 215 220
 Ala Gly Arg Thr Asp Leu Asp Ala Ala Phe Ala Glu Tyr Gly Glu Ala
 225 230 235 240
 Val Gly Glu Ala Phe Gln Leu Arg Asp Asp Leu Leu Asp Ala Phe Gly
 245 250 255
 Asp Ser Ala Ala Thr Gly Lys Pro Ala Gly Leu Asp Phe Thr Gln His
 260 265 270
 Lys Met Thr Leu Leu Leu Gly Trp Ala Met Gln Arg Asp Glu Ser Ile

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

 340 345 350
 Arg Asp Ala

355

<210> 48

<211> 1065

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 48

atgacccag acgtgaccaa cgatgcctca caagaaaccg tcgatctggt tcgccgaacc	60
tctgcggatc tctccgccg cggtgaagct cgcattcagt ccttctctc cgcgggccac	120
gatactggg ctgccgtcga tgagcgagcc gttgttccaa ttgatgcct cgcagagctc	180
atggcttcg gtggaaaacg catccgccg gcattctgca ttgccggcca ctggccgcc	240
ggtggagatc cagacgattc cggcgtctc gcggctggcg ccgcgctcga gatgcttcac	300
gcatcccttc tggcgacga cgatatactg gatgattcct ctacgcgcg cggcctccca	360
accgtccag caaagcacac ggccctgcat caggatttgg gatggcaagg agagtcacgt	420
cgctttggag aggggtgtgg actcttgggt ggtggccttg cacttaccta cgccgacgaa	480
ttgatgtgcg aagcgccagc aacagccatt gccgagtga acaagctgcg atccgaagtg	540
cttatcgcc agtacaatga tgtgtgtcga gctgcggagt ttagcgtgga cccacgcctt	600
tctcgcttga tcgtgtcat taaatcgggc cgctacacta ttcaccgtcc actggtcatt	660
ggtgctaacg tcgcgggccg caccgacttg gacgtgcgt tcgtgaata cggatgaagct	720
gtggcgagg ctttccaact gcgcgatgac ctctcgatg catttggtga tagcgccgca	780
accgtaagc cagcaggtct tgactttacc caacataaga tgaccctctt gttgggctgg	840
gcaatgcaac gcgacgagtc cattcgagcc ctattaccg agccggcca ctcggccgat	900
gacgttcgtc gccacctgct tgacaccaag gtccagacg atgttgagca gcacattgcc	960
ggtttggtgg agcgaggatg caaagctatc acagacgcac caattgcgc cgtgtggcgc	1020
gaagaactta tggaatggc gaatcgcgtt gcatatcgcg atgct	1065

<210> 49

<211> 415

<212> PRT

<213> Solanum lycopersicum

<400> 49

Met Ile Phe Ser Lys Gly Leu Ala Gln Ile Ser Arg Asn Arg Phe Ser

1	5	10	15
---	---	----	----

Arg Cys Arg Trp Leu Phe Ser Leu Arg Pro Ile Pro Gln Leu His Gln
 20 25 30
 Ser Asn His Ile His Asp Pro Pro Lys Val Leu Gly Cys Arg Val Ile
 35 40 45
 His Ser Trp Val Ser Asn Ala Leu Ser Gly Ile Gly Gln Gln Ile His
 50 55 60
 Gln Gln Ser Thr Ala Val Ala Glu Glu Gln Val Asp Pro Phe Ser Leu
 65 70 75 80
 Val Ala Asp Glu Leu Ser Leu Leu Thr Asn Arg Leu Arg Ser Met Val
 85 90 95
 Val Ala Glu Val Pro Lys Leu Ala Ser Ala Ala Glu Tyr Phe Phe Lys
 100 105 110
 Leu Gly Val Glu Gly Lys Arg Phe Arg Pro Thr Val Leu Leu Leu Met
 115 120 125
 Ala Thr Ala Leu Asn Val Gln Ile Pro Arg Ser Ala Pro Gln Val Asp
 130 135 140
 Val Asp Ser Phe Ser Gly Asp Leu Arg Thr Arg Gln Gln Cys Ile Ala
 145 150 155 160
 Glu Ile Thr Glu Met Ile His Val Ala Ser Leu Leu His Asp Asp Val
 165 170 175
 Leu Asp Asp Ala Asp Thr Arg Arg Gly Ile Gly Ser Leu Asn Phe Val
 180 185 190
 Met Gly Asn Lys Leu Ala Val Leu Ala Gly Asp Phe Leu Leu Ser Arg
 195 200 205
 Ala Cys Val Ala Leu Ala Ser Leu Lys Asn Thr Glu Val Val Cys Leu
 210 215 220
 Leu Ala Thr Val Val Glu His Leu Val Thr Gly Glu Thr Met Gln Met
 225 230 235 240
 Thr Thr Ser Ser Asp Glu Arg Cys Ser Met Glu Tyr Tyr Met Gln Lys
 245 250 255
 Thr Tyr Tyr Lys Thr Ala Ser Leu Ile Ser Asn Ser Cys Lys Ala Ile

260 265 270
Ala Leu Leu Ala Gly His Ser Ala Glu Val Ser Val Leu Ala Phe Asp
275 280 285
Tyr Gly Lys Asn Leu Gly Leu Ala Phe Gln Leu Ile Asp Asp Val Leu
290 295 300
Asp Phe Thr Gly Thr Ser Ala Thr Leu Gly Lys Gly Ser Leu Ser Asp

305 310 315 320
Ile Arg His Gly Ile Val Thr Ala Pro Ile Leu Tyr Ala Met Glu Glu
325 330 335
Phe Pro Gln Leu Arg Thr Leu Val Asp Arg Gly Phe Asp Asp Pro Val
340 345 350
Asn Val Glu Ile Ala Leu Asp Tyr Leu Gly Lys Ser Arg Gly Ile Gln
355 360 365
Arg Thr Arg Glu Leu Ala Arg Lys His Ala Ser Leu Ala Ser Ala Ala

370 375 380
Ile Asp Ser Leu Pro Glu Ser Asp Asp Glu Glu Val Gln Arg Ser Arg
385 390 395 400
Arg Ala Leu Val Glu Leu Thr His Arg Val Ile Thr Arg Thr Lys
405 410 415

<210> 50

<211> 1245

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 50

atgattttca gcaaaggatt ggcgcaaatt tctcgcaacc gtttttcacg ctgccgttgg 60

ctgttttcgc tgcgaccaat ccttcagctc caccagagca atcatattca tgatccaccc 120

aaggttcttg gttgccgagt aatccacagc tgggtttcca atgcgctgtc cggatatcggc 180

cagcagatcc accagcagtc taccgcagtg gcggaggaac aagtggatcc tttctccctc 240

gtggctgatg aactctccct cctgaccaat cgattgcgct ctatggctgt cgccgaagtc 300

cccaaattgg ctccgcggc ggagtacttt ttcaagttgg gcgtggaggg aaaacgcttc 360

cgaccaaccg tcttgttgct catggctact gcgctgaatg tccagatccc ccgtccgca 420

ccgcaagttg atgtggatag cttctcaggc gatcttcgta cccgacagca gtgtatcgcc 480

gagattaccg agatgatcca cgttgcttca ttgctccatg atgacgtgct tgatgatgct 540

gatacccgtc gaggcacgga ctactgaat ttcggttatgg gcaataagct tgctgttctc 600

gtgggtgatt ttttgctgtc ccgcgcgtgc gtggcattgg cctctctgaa aaacaccgag 660

gtgggttgcc ttcttgcaac agttgttgag catcttgtaa ccggtgagac catgcagatg 720

accacctcgt ccgatgaacg atgcagcatg gactactaca tgcaaaagac ctactataaa 780

acggcgctcc tgatctcgaa ctcttgtaaa gctattgcgc tcctggctgg ccactctgcg 840

gaagtttcgg tgttggcatt tgattacggt aaaaatctcg gcttggcttt ccaattgatc 900

gatgatgttc ttgatttcac tggaacttcc gcaaccttgg gcaaagggtc cctgtccgat 960

attcgccacg gcatcgtcac cgccccaatc ttgtatgcga tggaagagtt ccctcagctc 1020

cgtacctcgc ttgatcggtg cttcgatgat ccggtaaacg tggaaattgc cctcgactac 1080

ttgggtaagt ctgcggcat ccaacgcaca cgcgagcttg cacgtaaaca cgcgtctctc 1140

gcgagcgagg cgatcgactc cctgccggaa tccgatgatg aagaggtgca gcgcagccgc 1200

cgtgcgcttg ttgaattgac tcaccgcgtt atcacccgta ctaag 1245

<210> 51

<211> 420

<212> PRT

<213> Catharanthus roseus

<400> 51

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

1 5 10 15

Arg Ser Arg Leu Leu Tyr Pro Leu Gln Ser Gln Ser Pro Glu Leu Leu

20 25 30

Gln Ser Phe Gln Phe Arg Ser Pro Ile Gly Ser Ser Gln Lys Val Ser

35 40 45

Gly Phe Arg Val Ile Tyr Ser Trp Val Ser Ser Ala Leu Ala Asn Val

50 55 60

Gly Gln Gln Val Gln Arg Gln Ser Asn Ser Val Ala Glu Glu Pro Leu

65 70 75 80

Asp Pro Phe Ser Leu Val Ala Asp Glu Leu Ser Ile Leu Ala Asn Arg

85 90 95

Leu Arg Ser Met Val Val Ala Glu Val Pro Lys Leu Ala Ser Ala Ala
 100 105 110
 Glu Tyr Phe Phe Lys Leu Gly Val Glu Gly Lys Arg Phe Arg Pro Thr
 115 120 125

 Val Leu Leu Leu Met Ala Thr Ala Ile Asp Ala Pro Ile Ser Arg Thr
 130 135 140
 Pro Pro Asp Thr Ser Leu Asp Thr Leu Ser Thr Glu Leu Arg Leu Arg
 145 150 155 160
 Gln Gln Thr Ile Ala Glu Ile Thr Lys Met Ile His Val Ala Ser Leu
 165 170 175
 Leu His Asp Asp Val Leu Asp Asp Ala Glu Thr Arg Arg Gly Ile Gly
 180 185 190

 Ser Leu Asn Phe Val Met Gly Asn Lys Leu Ala Val Leu Ala Gly Asp
 195 200 205
 Phe Leu Leu Ser Arg Ala Cys Val Ala Leu Ala Ser Leu Lys Asn Thr
 210 215 220
 Glu Val Val Ser Leu Leu Ala Thr Val Val Glu His Leu Val Thr Gly
 225 230 235 240
 Glu Thr Met Gln Met Thr Thr Thr Ser Asp Gln Arg Cys Ser Met Glu
 245 250 255

 Tyr Tyr Met Gln Lys Thr Tyr Tyr Met Thr Ala Ser Leu Ile Ser Asn
 260 265 270
 Ser Cys Lys Ala Ile Ala Leu Leu Ala Gly Gln Thr Ser Glu Val Ala
 275 280 285
 Met Leu Ala Tyr Glu Tyr Gly Lys Asn Leu Gly Leu Ala Phe Gln Leu
 290 295 300
 Ile Asp Asp Val Leu Asp Phe Thr Gly Thr Ser Ala Ser Leu Gly Lys
 305 310 315 320

 Gly Ser Leu Ser Asp Ile Arg His Gly Ile Val Thr Ala Pro Ile Leu
 325 330 335
 Phe Ala Ile Glu Glu Phe Pro Glu Leu Arg Ala Val Val Asp Glu Gly

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

Leu Ala Ser Asp Ala Ile Asp Ser Leu Pro Val Thr Asp Asp Glu His
385 390 395 400
Val Leu Arg Ser Arg Arg Ala Leu Val Glu Leu Thr Gln Arg Val Ile
405 410 415
Thr Arg Arg Lys
420

<210> 52

<211> 1260

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 52

atgctgtttt cagcggcct ctatcgcatc gcgcgaacct cccigaaccg tagccgcctc 60
ttgtacccat tgcaatcaca aagccccgag ttgctccagt cgtttcagtt ccgctcgctt 120

attggatcat cccaaaaggc atccggattc cgtgttatct actcctgggt gtcctcggca 180
ttggccaacg tcggccagca ggtacaacga caaagcaact ccgtagcgga ggagcccctg 240
gatccatttt cactggttgc cgatgaactt tccattctcg ctaaccgtct ccgctcaatg 300
gtcgtcgctg aagtgcctaa gctcgcttcc gcagcagagt acttcttcaa gttgggtgtg 360
gagggcaagc gtttcgacc caccgtcctt ctgttgatgg cgaccgtat cgacgcaccc 420
atctctcgca caccacctga tacctcgttg gatacccttt ctactgagct tcgctgcgc 480
caacagacca ttgccgagat caccaaaatg attcacgttg cttcgctcct ccacgatgat 540

gtgctcgatg acgtgagac ccgccgagga attggttcac ttaattttgt aatgggtaac 600
aagcttgctg tgctcgagg tgatttcctg ctctcacgcg catgtgtggc cctggcttca 660
ctgaaaaata ccgaggtggt ttcgttgctg gccactgttg tagagcacct cgtcactggc 720
gagactatgc aaatgacaac cagtcggac cagcgtgca gcatggagta ttatatgcag 780
aaaacttatt acatgaccgc gtctctgatt tcgaactcct gcaaggccat tgcgttgctc 840
gccggacaaa cctcggaggc cgccatgctt gcatatgaat acggtaaaaa cctgggattg 900

gccttcacgc tcacgcacga tgtgtggat ttactggaa caagcgctag cctgggtaag 960

ggttccttga gcgatatccg ccatggcatc gtaacggccc caattctctt cgcgattgaa 1020

gagttccctg agttgcgtgc tgtagttgat gaaggctttg agaaccata taacgtagac 1080

cttgctttgc attacctcgg taaatctcgt ggtattcaac gtacccgcga actggcgatt 1140

aaacacgcca atcttgcac agatgcgac gattctttgc cggttaccga tgacgaacac 1200

gtgtgcgct ctcgtcgcgc acttgtggaa cttaccacgc gtgtgatcac gcgtcgcaaa 1260

<210> 53

<211> 367

<212> PRT

<213> C. glutamicum

<400> 53

Met Ser Thr Phe Asp Ala His Asp Leu Asp Leu Asp Thr Phe Pro Glu

1 5 10 15

Val Val Arg Asp Arg Leu Thr Gln Phe Leu Asp Ala Gln Ala Thr Thr

20 25 30

Ile Ala Gly Ile Gly Asp Pro Val Thr Glu Ala Val Ser His Leu Arg

35 40 45

Ser Phe Val Leu Asn Gly Gly Lys Arg Ile Arg Pro Leu Tyr Ala Trp

50 55 60

Ala Gly Phe Leu Ala Ala Gln Gly Leu Glu Asn Ser Gly Glu Lys Ile

65 70 75 80

Glu Ala Val Leu Asp Ala Ala Ser Ser Leu Glu Phe Ile Gln Ala Tyr

85 90 95

Phe Leu Ile His Asp Asp Pro Ala Ile Asp Ser Ser Asp Thr Arg Arg

100 105 110

Gly Ala Pro Thr Val His Arg Ala Val Glu Ala Asn His Arg Ala His

115 120 125

Asn Leu Glu Gly Asp Ser Glu His Phe Gly Glu Ser Val Ser Ile Leu

130 135 140

Ala Gly Asp Met Ala Leu Val Trp Ala Glu Asp Met Leu Gln Asp Ser

145 150 155 160

Gly Leu Ser Ala Glu Ala Leu Ala Arg Thr Arg Asp Ala Trp Arg Gly
165 170 175

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

Ala His Ala Asn Glu Ala Val Glu Leu Ala Asp Ser Val Asn Arg Phe
195 200 205

Lys Thr Ala Ala Tyr Thr Ile Ala Arg Pro Leu His Leu Gly Ala Ser
210 215 220

Ile Ala Gly Gly Ser Gln Glu Leu Ile Asp Ala Leu Leu His Tyr Gly
225 230 235 240

His Asp Ile Gly Ile Ala Phe Gln Leu Arg Asp Asp Leu Leu Gly Val
245 250 255

Phe Gly Asp Pro Lys Val Thr Gly Lys Pro Ala Gly Asp Asp Ile Arg
260 265 270

Glu Gly Lys Arg Thr Val Leu Leu Ala Leu Ala Leu Gln Arg Ala Asp
275 280 285

Lys Asn Ser Pro Glu Ala Ala Ala Ala Ile Arg Ala Gly Ile Gly Lys
290 295 300

Val Ser Thr Pro Glu Asp Ile Ala Val Ile Ala Lys His Ile Arg Ala
305 310 315 320

Thr Gly Ala Glu Glu Glu Val Glu Gln Arg Ile Thr Gln Leu Thr Glu
325 330 335

Ser Gly Leu Ala His Leu Asp Asp Val Glu Ile Pro Asp Glu Val Arg
340 345 350

Thr Gln Leu Arg Ala Leu Ala Ile Arg Ser Thr Glu Arg Arg Met
355 360 365

<210> 54
<211> 1101
<212> DNA
<213> Artificial
<220><223> mutated and/or codon-optimized sequence
<400> 54

atgtctacat tcgatgctca tgacttggac ctggatacct tccctgaagt tgtgcgcgac 60

cgctgacgc agtttctcga tgcacaggca acgacgattg cgggtatcgg agaccccgtt 120

accgaggcgc tatctcactt gcgttcgttt gtcctcaacg gaggtaaagc tatccgcccc 180

ctttatgcct gggccggctt tttggccgca caaggcctgg agaatagcgg cgaagagatt 240

gaggcggtgt tggatgcagc gaggcagctt gaattcatcc aagcttattt tcttattcac 300

gatgaccag caattgattc ctccgacacc cgtcgcggcg cacctaccgt ccaccgcgt 360

gtggaagcca accaccgcgc gcacaacttg gagggagaca gcgaacactt cggtgagtcg 420

gtttccatcc ttgcaggaga tatggccctg gtttgggcag aggacatgct tcaggattct 480

ggcctttccg cagaagcgtt ggctcgtacc cgtgacgcct ggcgcgccat gcgcaccgag 540

gtcatcggtg gccagctgct ggacatctat ctggaagctc atgctaata agccgtcgaa 600

ctggcagaca gcgtcaaccg ttttaagact gccgcatata ctattgccg ccctcttcat 660

cttgggtgct caatcgctgg cggctcccag gaattgatcg atgcactgct ccactacgga 720

cacgatatcg gtatgcatt ccagctccga gatgatctgc ttggtgtttt tggtgaccg 780

aaggtaacag gaaaaccagc cggcgatgac atccgtgaag gcaaacgcac agtgcttctc 840

gcattggcgt tgcagcgcgc agacaagaac agccctgagg cagctgccgc aatccgcgcg 900

ggcatcgcca aggtgtccac cccggaagat attgctgtca tcgctaaaca tattcgtgcg 960

acaggcgcgg aagaagaagt agagcaacgc atcactcaac tgactgaatc aggcttggcg 1020

catttggatg atgtggaaat tcttgacgag gtgcgtaccc aactgcgcgc gctcgcaatc 1080

cgctccacgg agcgccgaat g 1101

<210> 55

<211> 296

<212> PRT

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 55

Met Thr Asn Lys Leu Thr Ser Phe Leu Ala Asp Arg Lys Lys Thr Ile

1 5 10 15

Glu Asn Gln Leu Ser Val Tyr Thr Glu Lys Leu Asp Met Pro Asp Ser

20 25 30

Leu Lys Lys Ser Met Leu Tyr Ser Leu Gln Ala Gly Gly Lys Arg Leu

35 40 45

Arg Pro Leu Ile Val Leu Ala Val Leu Asn Ala Tyr Gly Lys Ser Glu
50 55 60
Lys Asp Gly Ile Pro Val Gly Cys Ala Val Glu Met Ile His Thr Tyr
65 70 75 80
Phe Leu Ile His Asp Asp Leu Pro Cys Met Asp Asp Asp Asp Leu Arg
85 90 95
Arg Gly Lys Pro Thr Asn His Lys Val Phe Gly Glu Ala Thr Ala Val
100 105 110
Leu Ala Gly Asp Gly Leu Leu Thr Glu Ser Phe Lys Leu Ile Thr Ser
115 120 125
His Val Ser Asp Glu Val Ser Ala Glu Lys Arg Leu Arg Leu Val Asn
130 135 140
Glu Leu Ile Ser Ala Ala Gly Thr Glu Gly Met Val Gly Gly Gln Val
145 150 155 160
Ala Asp Met Glu Ala Glu Asn Arg Gln Val Thr Leu Glu Glu Leu Glu
165 170 175
Ser Ile His Glu Arg Lys Thr Ala Lys Leu Leu Gly Phe Cys Val Ile
180 185 190
Ala Gly Ala Ile Leu Ala Asp Ala Pro Glu Glu Asp Ile Glu Thr Leu
195 200 205
Arg Thr Phe Ser Ser His Ile Gly Ile Gly Phe Gln Ile Arg Asp Asp
210 215 220
Ile Leu Asp Leu Glu Gly Ser Glu Glu Lys Ile Gly Lys Arg Val Gly
225 230 235 240
Ser Asp Thr Thr Asn Asp Lys Ser Thr Tyr Pro Ser Leu Leu Ser Leu
245 250 255
Glu Gly Ala Lys His Lys Leu Asp Val His Ile Lys Glu Ala Lys Arg
260 265 270
Leu Ile Gly Gly Leu Ser Leu Gln Lys Asp Leu Leu Tyr Glu Leu Cys
275 280 285
Asp Leu Ile Ala Ala Arg Asp His

290 295

<210> 56

<211> 888

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 56

atgactaata agctgacgtc cttcctcgca gaccgtaaaa agactatcga aaatcaattg 60

tctgtgtaca cagaaaagct cgatatgcct gattccctca agaagtctat gctttactcc 120

cttcaggcag gtggcaagcg tttgcgcccc ctcacgtgc ttgcggttct caacgcctac 180

ggtaagtcag agaaagatgg tatcccgtt ggatgcgcgc tggaaatgat ccacacctat 240

tttcttatcc acgacgatct tccgtgtatg gacgatgacg accttcgcgc tggtaagcca 300

actaaccaca aggttttttg tgaggcaact gcggttcttg caggatgatgg tctgcttacc 360

gaatccttta aacttatcac ttcctatgtt tctgatgaag tgtctgcaga aaaacgtctg 420

cgcttggtga acgaactcat cagcgccgcg ggcaccgaag gtatggtagg aggacaagtg 480

gctgatatgg aagccgagaa ccgtcagggt acgcttgaag aactcgagtc tattcacgaa 540

cgcaagaccg ccaagctctt gggcttctgt gtattgtctg gagccatcct tgcagatgca 600

cctgaagagg atattgaaac attgcgtacc ttttcatccc acatcggcac cggttttcag 660

atccgagacg acattcttga tctcgagggc tccgaagaaa agattggtaa acgcgtgggc 720

tccgacacca ccaatgacaa gtctacttac ccgtctcttc ttagccttga aggcgcaaag 780

cataagttgg acgttcatat taaggaggcg aagcgtttga ttggtggctt gaggcctccag 840

aaagacttgc tctacgagct gtgcgatctc atcgcagcac gcgatcac 888

<210> 57

<211> 396

<212> PRT

<213> S. cerevisiae S288c

<400> 57

Met Thr Val Tyr Thr Ala Ser Val Thr Ala Pro Val Asn Ile Ala Thr

1 5 10 15

Leu Lys Tyr Trp Gly Lys Arg Asp Thr Lys Leu Asn Leu Pro Thr Asn

20 25 30

Ser Ser Ile Ser Val Thr Leu Ser Gln Asp Asp Leu Arg Thr Leu Thr

35 40 45
 Ser Ala Ala Thr Ala Pro Glu Phe Glu Arg Asp Thr Leu Trp Leu Asn
 50 55 60
 Gly Glu Pro His Ser Ile Asp Asn Glu Arg Thr Gln Asn Cys Leu Arg

 65 70 75 80
 Asp Leu Arg Gln Leu Arg Lys Glu Met Glu Ser Lys Asp Ala Ser Leu
 85 90 95
 Pro Thr Leu Ser Gln Trp Lys Leu His Ile Val Ser Glu Asn Asn Phe
 100 105 110
 Pro Thr Ala Ala Gly Leu Ala Ser Ser Ala Ala Gly Phe Ala Ala Leu
 115 120 125
 Val Ser Ala Ile Ala Lys Leu Tyr Gln Leu Pro Gln Ser Thr Ser Glu

 130 135 140
 Ile Ser Arg Ile Ala Arg Lys Gly Ser Gly Ser Ala Cys Arg Ser Leu
 145 150 155 160
 Phe Gly Gly Tyr Val Ala Trp Glu Met Gly Lys Ala Glu Asp Gly His
 165 170 175
 Asp Ser Met Ala Val Gln Ile Ala Asp Ser Ser Asp Trp Pro Gln Met
 180 185 190
 Lys Ala Cys Val Leu Val Val Ser Asp Ile Lys Lys Asp Val Ser Ser

 195 200 205
 Thr Gln Gly Met Gln Leu Thr Val Ala Thr Ser Glu Leu Phe Lys Glu
 210 215 220
 Arg Ile Glu His Val Val Pro Lys Arg Phe Glu Val Met Arg Lys Ala
 225 230 235 240
 Ile Val Glu Lys Asp Phe Ala Thr Phe Ala Lys Glu Thr Met Met Asp
 245 250 255
 Ser Asn Ser Phe His Ala Thr Cys Leu Asp Ser Phe Pro Pro Ile Phe

 260 265 270
 Tyr Met Asn Asp Thr Ser Lys Arg Ile Ile Ser Trp Cys His Thr Ile
 275 280 285

Asn Gln Phe Tyr Gly Glu Thr Ile Val Ala Tyr Thr Phe Asp Ala Gly
290 295 300

Pro Asn Ala Val Leu Tyr Tyr Leu Ala Glu Asn Glu Ser Lys Leu Phe
305 310 315 320

Ala Phe Ile Tyr Lys Leu Phe Gly Ser Val Pro Gly Trp Asp Lys Lys

325 330 335
Phe Thr Thr Glu Gln Leu Glu Ala Phe Asn His Gln Phe Glu Ser Ser

340 345 350
Asn Phe Thr Ala Arg Glu Leu Asp Leu Glu Leu Gln Lys Asp Val Ala

355 360 365
Arg Val Ile Leu Thr Gln Val Gly Ser Gly Pro Gln Glu Thr Asn Glu

370 375 380
Ser Leu Ile Asp Ala Lys Thr Gly Leu Pro Lys Glu

385 390 395
<210> 58

<211> 1188

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 58

atgaccgttt acaccgcctc agtcaccgca ccagttaaca ttgcaaccct taaatactgg 60
ggcaagcgcg atactaagct taaccttctt accaactcat cgatctcggt gaccctgtct 120
caggatgata tgcgaacgct tacttctgcg gcaacagcac cagaattcga gcgagatacg 180
ctctggctga atggcggaacc gcattccatt gacaacgagc gcaccagaa ctgcttgctt 240
gatcttcgcc aactccgtaa ggagatggaa agcaaagatg caagcctccc aacctctctt 300

caatggaaac ttacatcgt ctctgaaaac aacttccta cggcggctgg cctggcgctc 360
tcggcagctg gcttcgctgc attggtttcc gccattgcaa agctttatca gtttcgcaa 420
tcgacctcgg agattagcgg catcgccgcg aagggttcag gttcggcatt tcgctctctc 480
tttggaggct acgtggcctg ggaaatggga aaggccgaag atggtcacga ttccatggcg 540
gtgcaaatcg ccgatatgctc agattggcct caaatgaagg cctgtgtgtt ggtggtatcc 600
gacatcaaaa aagacgtttc ctcgactcaa ggaatgcaac ttacggtagc tacgagcgaa 660

ctgtttaagg agcgattga gcatgtggcg cccaaacgct ttgaagtcac gcgtaaggca 720

atcgtggaga aggacttcgc gaccttcgcg aaggaaacta tgatggatag caacagcttc 780

cacgtacgt gtctggattc gttccctccg attttctata tgaatgacac gtctaaacgc 840

atcatttcct gggtcacac gatcaaccaa ttctacggcg aaactatcgt agcatacact 900

tttgacgcag gaccgaacgc cgttctctac tacctcgccg aaaacgagtc aaaattgttc 960

gccttcactc acaaactctt tggatccgtc cctggttggg acaagaagtt cacgactgaa 1020

cagttggaag cattcaacca tcagtttgaa agctctaact tcacagcccg tgagctggac 1080

ctcgaactgc agaaagatgt ggcacgcgtc attcttacc cagtggtc cggaccccaa 1140

gagacgaatg agtctctgat tgacgcaaaa acgggtctcc ccaaagaa 1188

<210> 59

<211> 650

<212> PRT

<213> Yarrowia lipolytica CLIB 122 / E 150

<400> 59

Met Ser Ala Asn Glu Asn Ile Ser Arg Phe Asp Ala Pro Val Gly Lys

1 5 10 15

Glu His Pro Ala Tyr Glu Leu Phe His Asn His Thr Arg Ser Phe Val

20 25 30

Tyr Gly Leu Gln Pro Arg Ala Cys Gln Gly Met Leu Asp Phe Asp Phe

35 40 45

Ile Cys Lys Arg Glu Asn Pro Ser Val Ala Gly Val Ile Tyr Pro Phe

50 55 60

Gly Gly Gln Phe Val Thr Lys Met Tyr Trp Gly Thr Lys Glu Thr Leu

65 70 75 80

Leu Pro Val Tyr Gln Gln Val Glu Lys Ala Ala Ala Lys His Pro Glu

85 90 95

Val Asp Val Val Val Asn Phe Ala Ser Ser Arg Ser Val Tyr Ser Ser

100 105 110

Thr Met Glu Leu Leu Glu Tyr Pro Gln Phe Arg Thr Ile Ala Ile Ile

115 120 125

Ala Glu Gly Val Pro Glu Arg Arg Ala Arg Glu Ile Leu His Lys Ala

130	135	140	
Gln Lys Lys Gly Val Thr Ile Ile Gly Pro Ala Thr Val Gly Gly Ile			
145	150	155	160
Lys Pro Gly Cys Phe Lys Val Gly Asn Thr Gly Gly Met Met Asp Asn			
	165	170	175
Ile Val Ala Ser Lys Leu Tyr Arg Pro Gly Ser Val Ala Tyr Val Ser			
	180	185	190
Lys Ser Gly Gly Met Ser Asn Glu Leu Asn Asn Ile Ile Ser His Thr			
	195	200	205
Thr Asp Gly Val Tyr Glu Gly Ile Ala Ile Gly Gly Asp Arg Tyr Pro			
	210	215	220
Gly Thr Thr Phe Ile Asp His Ile Leu Arg Tyr Glu Ala Asp Pro Lys			
225	230	235	240
Cys Lys Ile Ile Val Leu Leu Gly Glu Val Gly Gly Val Glu Glu Tyr			
	245	250	255
Arg Val Ile Glu Ala Val Lys Asn Gly Gln Ile Lys Lys Pro Ile Val			
	260	265	270
Ala Trp Ala Ile Gly Thr Cys Ala Ser Met Phe Lys Thr Glu Val Gln			
	275	280	285
Phe Gly His Ala Gly Ser Met Ala Asn Ser Asp Leu Glu Thr Ala Lys			
290	295	300	
Ala Lys Asn Ala Ala Met Lys Ser Ala Gly Phe Tyr Val Pro Asp Thr			
305	310	315	320
Phe Glu Asp Met Pro Glu Val Leu Ala Glu Leu Tyr Glu Lys Met Val			
	325	330	335
Ala Lys Gly Glu Leu Ser Arg Ile Ser Glu Pro Glu Val Pro Lys Ile			
	340	345	350
Pro Ile Asp Tyr Ser Trp Ala Gln Glu Leu Gly Leu Ile Arg Lys Pro			
	355	360	365
Ala Ala Phe Ile Ser Thr Ile Ser Asp Asp Arg Gly Gln Glu Leu Leu			
370	375	380	

Tyr Ala Gly Met Pro Ile Ser Glu Val Phe Lys Glu Asp Ile Gly Ile
 385 390 395 400
 Gly Gly Val Met Ser Leu Leu Trp Phe Arg Arg Arg Leu Pro Asp Tyr
 405 410 415
 Ala Ser Lys Phe Leu Glu Met Val Leu Met Leu Thr Ala Asp His Gly
 420 425 430

 Pro Ala Val Ser Gly Ala Met Asn Thr Ile Ile Thr Thr Arg Ala Gly
 435 440 445
 Lys Asp Leu Ile Ser Ser Leu Val Ala Gly Leu Leu Thr Ile Gly Thr
 450 455 460
 Arg Phe Gly Gly Ala Leu Asp Gly Ala Ala Thr Glu Phe Thr Thr Ala
 465 470 475 480
 Tyr Asp Lys Gly Leu Ser Pro Arg Gln Phe Val Asp Thr Met Arg Lys
 485 490 495

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

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

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

625	630	635	640
Gly Asp Val Ser Lys Ala Lys Thr Arg Ser			
	645	650	

<210> 60

<211> 1950

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 60

atgtctgcaa atgaaaacat ttcccgcttt gatgcacctg tgggtaagga gcatccagcc	60
tatgagctgt tccataacca caccgctcc ttctctacg gtctccagcc tcgcgcgtgc	120
cagggtatgc tcgactttga ttctatttgc aagcgcgaga acccctcggt tgccggtgtt	180
atctaccctt tcggcgggtca gtctgtgact aagatgtact ggggtaccaa ggagaccctg	240
ctcccggtgt accagcaagt tgaaaaagca gcagcgaaac atcccgaagt cgacgtcgtc	300
gtgaacttcg cgtcgtcccg cagcgtctac agctccacaa tggaactttt ggaataccca	360
caattccgca ccatcgctat catcgcagag ggcgtcccag aacgccgtgc gcgcgaaatc	420
ctgcacaagg ctcagaaaaa gggcgtgacg attatcggac cggcaaccgt cggcggcatt	480
aagcccgggt gctttaagt tggtataatc ggtggcatga tggacaacat cgtggcgtcc	540
aagctgtacc gccaggttc cgttgcttac gtctccaagt ccggtgggtat gtcaaatgag	600
cttaataaca tcatttccca caccaccgat ggcgtgtacg agggaaatcgc aattggtggc	660
gaccgctacc cgggcacaaac ttttattgat cacatcttgc gttacgaagc ggatccaaag	720
tgcaagatca ttgtgctctt gggcgaagta ggtggcgtgg aagagtaccg cgttatcgag	780
gcggttaaaa acggtcaaat taagaagcca atcgtagcgt gggcaatcgg cacctgcgcc	840
tccatgttta agacggaagt acagtttggc catgctggct caatggctaa ctcagacctc	900
gagacggcta aagcaaaaaa tgcggcaatg aaatcagcag gcttttacgt gccggatacc	960
ttcgaggata tgcccgaagt cttggccgag ctctacgaga agatggtggc gaaaggcgag	1020
ctcagccgca tctccgaacc ggaggttctt aagatcccga tcgactatag ctgggcgcag	1080
gaacttggtt tgatccgcaa gccagccgca ttcattctta ccatctcaga cgaccgcggc	1140
caagagctcc ttacgcggg catgcccatc tcagaggtct tcaaggagga catcggcac	1200
ggtggcggtta tgtccctgct gtggtttcgt cgcgctctgc ctgattatgc atcgaaatc	1260
cttgaatgg tgcttatgct gacggccgac caccgacctg ccgtttccgg tgcgatgaat	1320
acgatcatca cactcgagc gggtaaggac ctcatcttct ccttgttgc cggccttctt	1380

acaatttgta cccgcttcgg cggagcattg gacggcgcgg ccacagagtt tactaccgcc 1440

tacgataagg gcttgtcgcc ccgccaattt gtcgatacaa tgcgcaagca aaacaagctc 1500

atcccaggca tcggccaccg agtgaagtcc cgtataaacc ccgatttcgg tgtggaactg 1560

gttaaggact ttgtgaagaa gaatttcgcc tccaccaccg tgttggatta cgcgctcgcc 1620

gttgaggagg tgaccacatc taaaaaggat aacttgatcc tcaacgttga tgggtcgatt 1680

gcggtttcgt tcgtcgatct tatgcgctcg tgtggtgctt ttaccgttga agaaacagaa 1740

gattatctga aaaatggcgt ttgtaacggt ttgtttgtcc tcggtcgttc cattggcctg 1800

attgcgcacc acctggacca gaaacgtctt aaaaccggtc ttaccgccca tccttgggat 1860

gacattacgt acctcggttg ccaagaggca attcagaaga agcgtgtgga gatctctgca 1920

ggcgacgttt caaaggcgaa gactcgaagc 1950

<210> 61

<211> 652

<212> PRT

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 61

Met Ser Gln Thr His Lys His Ala Ile Pro Ala Asn Ile Ala Asp Arg

1 5 10 15

Cys Leu Ile Asn Pro Glu Gln Tyr Glu Thr Lys Tyr Lys Gln Ser Ile

20 25 30

Asn Asp Pro Asp Thr Phe Trp Gly Glu Gln Gly Lys Ile Leu Asp Trp

35 40 45

Ile Thr Pro Tyr Gln Lys Val Lys Asn Thr Ser Phe Ala Pro Gly Asn

50 55 60

Val Ser Ile Lys Trp Tyr Glu Asp Gly Thr Leu Asn Leu Ala Ala Asn

65 70 75 80

Cys Leu Asp Arg His Leu Gln Glu Asn Gly Asp Arg Thr Ala Ile Ile

85 90 95

Trp Glu Gly Asp Asp Thr Ser Gln Ser Lys His Ile Ser Tyr Arg Glu

100 105 110

Leu His Arg Asp Val Cys Arg Phe Ala Asn Thr Leu Leu Asp Leu Gly

115	120	125	
Ile Lys Lys Gly Asp Val Val Ala Ile Tyr Met Pro Met Val Pro Glu			
130	135	140	
Ala Ala Val Ala Met Leu Ala Cys Ala Arg Ile Gly Ala Val His Ser			
145	150	155	160
Val Ile Phe Gly Gly Phe Ser Pro Glu Ala Val Ala Gly Arg Ile Ile			
165	170	175	
Asp Ser Ser Ser Arg Leu Val Ile Thr Ala Asp Glu Gly Val Arg Ala			
180	185	190	
Gly Arg Ser Ile Pro Leu Lys Lys Asn Val Asp Asp Ala Leu Lys Asn			
195	200	205	
Pro Asn Val Thr Ser Val Glu His Val Ile Val Leu Lys Arg Thr Gly			
210	215	220	
Ser Asp Ile Asp Trp Gln Glu Gly Arg Asp Leu Trp Trp Arg Asp Leu			
225	230	235	240
Ile Glu Lys Ala Ser Pro Glu His Gln Pro Glu Ala Met Asn Ala Glu			
245	250	255	
Asp Pro Leu Phe Ile Leu Tyr Thr Ser Gly Ser Thr Gly Lys Pro Lys			
260	265	270	
Gly Val Leu His Thr Thr Gly Gly Tyr Leu Val Tyr Ala Ala Thr Thr			
275	280	285	
Phe Lys Tyr Val Phe Asp Tyr His Pro Gly Asp Ile Tyr Trp Cys Thr			
290	295	300	
Ala Asp Val Gly Trp Val Thr Gly His Ser Tyr Leu Leu Tyr Gly Pro			
305	310	315	320
Leu Ala Cys Gly Ala Thr Thr Leu Met Phe Glu Gly Val Pro Asn Trp			
325	330	335	
Pro Thr Pro Ala Arg Met Cys Gln Val Val Asp Lys His Gln Val Asn			
340	345	350	
Ile Leu Tyr Thr Ala Pro Thr Ala Ile Arg Ala Leu Met Ala Glu Gly			
355	360	365	

Asp Lys Ala Ile Glu Gly Thr Asp Arg Ser Ser Leu Arg Ile Leu Gly
 370 375 380
 Ser Val Gly Glu Pro Ile Asn Pro Glu Ala Trp Glu Trp Tyr Trp Lys
 385 390 395 400
 Lys Ile Gly Lys Glu Lys Cys Pro Val Val Asp Thr Trp Trp Gln Thr
 405 410 415

 Glu Thr Gly Gly Phe Met Ile Thr Pro Leu Pro Gly Ala Ile Glu Leu
 420 425 430
 Lys Ala Gly Ser Ala Thr Arg Pro Phe Phe Gly Val Gln Pro Ala Leu
 435 440 445
 Val Asp Asn Glu Gly His Pro Gln Glu Gly Ala Thr Glu Gly Asn Leu
 450 455 460
 Val Ile Thr Asp Ser Trp Pro Gly Gln Ala Arg Thr Leu Phe Gly Asp
 465 470 475 480

 His Glu Arg Phe Glu Gln Thr Tyr Phe Ser Thr Phe Lys Asn Met Tyr
 485 490 495
 Phe Ser Gly Asp Gly Ala Arg Arg Asp Glu Asp Gly Tyr Tyr Trp Ile
 500 505 510
 Thr Gly Arg Val Asp Asp Val Leu Asn Val Ser Gly His Arg Leu Gly
 515 520 525
 Thr Ala Glu Ile Glu Ser Ala Leu Val Ala His Pro Lys Ile Ala Glu
 530 535 540

 Ala Ala Val Val Gly Ile Pro His Ala Ile Lys Gly Gln Ala Ile Tyr
 545 550 555 560
 Ala Tyr Val Thr Leu Asn His Gly Glu Glu Pro Ser Pro Glu Leu Tyr
 565 570 575
 Ala Glu Val Arg Asn Trp Val Arg Lys Glu Ile Gly Pro Leu Ala Thr
 580 585 590
 Pro Asp Val Leu His Trp Thr Asp Ser Leu Pro Lys Thr Arg Ser Gly
 595 600 605

 Lys Ile Met Arg Arg Ile Leu Arg Lys Ile Ala Ala Gly Asp Thr Ser

610	615	620	
Asn Leu Gly Asp Thr Ser Thr Leu Ala Asp Pro Gly Val Val Glu Lys			
625	630	635	640
Pro Leu Glu Glu Lys Gln Ala Ile Ala Met Pro Ser			
	645	650	

<210> 62

<211> 1956

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 62

atgtctcaga ctcacaaaca cgctatccct gctaacattg ccgaccgctg tcttattaac	60
cctgagcaat atgaaactaa atacaagcaa tcgattaacg acccagatac cttttgggga	120
gagcagggca aaatcctcga ctggatcacc ccataccaaa aggtaaaaaa cacctcattt	180
gccccaggta acgtgtccat caagtggtag gaagatggta ctttgaacct ggccggccaac	240
tgccttgacc gacacctgca agagaacggt gaccgcaccg caatcatctg ggaggcgac	300
gacacctcac aatccaaaca catctcatat ctggaacttc atcgcgacgt ttgccgtttc	360
gcaaacacgt tgctcgatct gggcatcaag aagggtgatg ttgtggcaat ctacatgccg	420
atggttccag aagcagcagt cgcgatgctt gcatgcgcac gcatcggcgc ggttcactca	480
gtgatcttcg gcggattttc ccccgaagcc gttgctggcc gtattatcga tagctcttcc	540
cgcttggtga tctactgtga cgagggtgtg cgcgcaggtc gttccattcc gttgaagaag	600
aacgtcgatg atgccttgaa gaatcccaat gtgacctccg tggaacacgt tatcgtgttg	660
aaacgcactg gtcagacat cgactggcag gaaggccgag atttgtggtg gcgagacctc	720
atcgaaaagg cgtccccgga acatcaacct gaggcaatga acgcggaaga tccgtgtttt	780
attttgtaca cctcgggaag cacgggtaag ccaaaagggtg tgctccatac caccggcgggt	840
tatctcgtgt acgcagcgac tactttcaaa tatgttttcg actatcaccc cggagatatac	900
tactggtgca cagcagacgt tggttgggtt accggacatt cctacctgct gtacggtcct	960
ctggcatgcg gcgcaactac gttgatgttt gaagggtgtg ctaactggcc tactccagcg	1020
cgcatgtgtc aagtcgtgga caagcaccag gtaaacatcc tctatacagc acctaccgcc	1080
atccgagccc tcatggctga aggcgataag gccattgagg gtacagatcg ttcctccttg	1140
cgatatcctg gtcagtgagg cgaaccaatc aacctgaag catgggaatg gtactggaag	1200
aagattggaa aggagaagtg tccagttgtg gatacctggt ggagacaga gactggtggt	1260

ttcattgatta cccctctccc gggagcgatt gagttgaaag ccggttccgc gacccgcccc 1320

ttcttcggag tacagccagc gcttctggac aacgaaggac acccacagga aggcgccacc 1380

gaaggcaacc tggatcatcac ggacagctgg ccaggtcagg cacgtaccct gttcggcgac 1440

catgagcgct tcgagcaaac gtacttctct accttcaaga acatgtatit ctcgggtgac 1500

ggtgcccgtc gcgatgaaga tggctactac tggattacag gccgcgttga tgacgttctc 1560

aacgttagcg gacaccgttt gggtagcgca gaaatcgaat ccgctctggt agcacacccg 1620

aaaatcgctg aagcagccgt ggtgggtatc ccgcacgcga tcaagggtca agctatctac 1680

gcatacgtca cgcttaatca cggagaagaa ccacgccccg agttgtacgc cgaggtgcgc 1740

aattgggtcc gaaaagaaat tggccctacg gcaacccccg atgtgctcca ttggactgat 1800

agcctgcccc agactcgttc tggtaaaatt atgcgccgca tcctgcgcaa gattgcagct 1860

ggagacacct ccaatttggg cgatacctcc actttggccg atccaggagt ggtggagaaa 1920

ccgctcgaag aaaaacaggc aatcgctatg ccttcc 1956

<210> 63

<211> 500

<212> PRT

<213> S. cerevisiae S288c

<400> 63

Met Thr Lys Leu His Phe Asp Thr Ala Glu Pro Val Lys Ile Thr Leu

1 5 10 15

Pro Asn Gly Leu Thr Tyr Glu Gln Pro Thr Gly Leu Phe Ile Asn Asn

20 25 30

Lys Phe Met Lys Ala Gln Asp Gly Lys Thr Tyr Pro Val Glu Asp Pro

35 40 45

Ser Thr Glu Asn Thr Val Cys Glu Val Ser Ser Ala Thr Thr Glu Asp

50 55 60

Val Glu Tyr Ala Ile Glu Cys Ala Asp Arg Ala Phe His Asp Thr Glu

65 70 75 80

Trp Ala Thr Gln Asp Pro Arg Glu Arg Gly Arg Leu Leu Ser Lys Leu

85 90 95

Ala Asp Glu Leu Glu Ser Gln Ile Asp Leu Val Ser Ser Ile Glu Ala

100 105 110

Leu Asp Asn Gly Lys Thr Leu Ala Leu Ala Arg Gly Asp Val Thr Ile
 115 120 125
 Ala Ile Asn Cys Leu Arg Asp Ala Ala Ala Tyr Ala Asp Lys Val Asn
 130 135 140

 Gly Arg Thr Ile Asn Thr Gly Asp Gly Tyr Met Asn Phe Thr Thr Leu
 145 150 155 160
 Glu Pro Ile Gly Val Cys Gly Gln Ile Ile Pro Trp Asn Phe Pro Ile
 165 170 175
 Met Met Leu Ala Trp Lys Ile Ala Pro Ala Leu Ala Met Gly Asn Val
 180 185 190
 Cys Ile Leu Lys Pro Ala Ala Val Thr Pro Leu Asn Ala Leu Tyr Phe
 195 200 205

 Ala Ser Leu Cys Lys Lys Val Gly Ile Pro Ala Gly Val Val Asn Ile
 210 215 220
 Val Pro Gly Pro Gly Arg Thr Val Gly Ala Ala Leu Thr Asn Asp Pro
 225 230 235 240
 Arg Ile Arg Lys Leu Ala Phe Thr Gly Ser Thr Glu Val Gly Lys Ser
 245 250 255
 Val Ala Val Asp Ser Ser Glu Ser Asn Leu Lys Lys Ile Thr Leu Glu
 260 265 270

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

 Gly Asn Pro Phe Asp Lys Ala Asn Phe Gln Gly Ala Ile Thr Asn Arg
 340 345 350
 Gln Gln Phe Asp Thr Ile Met Asn Tyr Ile Asp Ile Gly Lys Lys Glu

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

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

Asp Ser Arg Val Pro Phe Gly Gly Val Lys Gln Ser Gly Tyr Gly Arg
 465 470 475 480
 Glu Met Gly Glu Glu Val Tyr His Ala Tyr Thr Glu Val Lys Ala Val
 485 490 495
 Arg Ile Lys Leu
 500

<210> 64

<211> 1500

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 64

atgaccaagc ttcacttcga caccgcagaa ccggttaaga tcaccctgcc aaacggcctc 60

acttatgaac agcccaccgg cctgttcac aacaacaaat tcatgaaggc gcaggatggc 120

aagacctacc cggctcgagga cccatcgacg gagaatactg tttgtgaagt gtcctcggcg 180

acgactgaag atgtcgagta tgcaatcgag tgcgctgac gagcattcca tgatactgaa 240

tgggcgacac aagatcctcg tgagcgtggt cgtctgctct ccaagcttgc ggacgagctt 300

gaatcccaaa ttgatctggt gtcttctatt gaagctctcg ataatggaaa aaccctcgct 360

ctggcccgcg gtgacgtgac catcgctatt aactgcctgc gcgacgcggc agcttacgca 420

gataagg tca atggccgcac cattaacaca ggcgatgggt atatgaattt tacaaccttg 480
gagcctattg gagtgtgcgg ccagattatc ccatggaact tccccatcat gatgctcgt 540

tggaagattg cccagccct tgcaatggga aacgtttgca ttctgaaacc ggccgcagtg 600
acgcccc tga atgctctgta ttctgcatct ctgtgtaaga aggtaggtat ccccgctggc 660
gtcgttaaca ttgttctggg tccaggtcgt actgtaggag cggcgctcac caacgatccg 720
cgcatccgta aactggcgtt caccggatcg acggaagtgg gtaagtccgt cgcagtcgac 780
tcttccgaat ctaatctcaa gaagattacg ctgcaactcg gcggcaaacc tgcacacctc 840
gtcttcgatg atgctaatat caagaaaacc ttgccgaacc tcgtcaacgg tattttcaag 900
aatgcggggc aaatctgctc ctccggcagc cgtatctacg ttcaggaagg tatttacgac 960

gagctgctcg cagccttcaa ggcgtatctg gagacggaaa ttaaggtggg caaccattt 1020
gataaagcca actttcaggg cgctattact aatgccagc agtttgacac cattatgaac 1080
tacattgata ttgtaagaa ggaaggcgtt aagattctca ctggaggaga aaaggttggt 1140
gacaaaggat actttatccg tccactgta ttctacgatg tgaacgaaga catgcgcatt 1200
gtgaaggaag aaattttcgg tccagtgtt acagtggcaa agttcaaac cctcgaggaa 1260
ggtgttgaga tggctaatag cagcgagttc ggctgggat caggcattga gacagaatcc 1320
ctttcgaccg gcctgaaggt ggccaagatg ttgaaggctg gcacagtttg gatcaatacc 1380

tataatgatt ttgactcccg tggtccgttt ggcggcgtga aacaatccgg ttacggccgc 1440
gaaatgggcg aggaggtgta ccacgcatac accgaagtga aggcagttcg aatcaaattg 1500

<210> 65
<211> 348
<212> PRT
<213> S. cerevisiae S288c
<400> 65

Met Ser Ile Pro Glu Thr Gln Lys Ala Ile Ile Phe Tyr Glu Ser Asn
1 5 10 15
Gly Lys Leu Glu His Lys Asp Ile Pro Val Pro Lys Pro Lys Pro Asn
20 25 30
Glu Leu Leu Ile Asn Val Lys Tyr Ser Gly Val Cys His Thr Asp Leu
35 40 45
His Ala Trp His Gly Asp Trp Pro Leu Pro Thr Lys Leu Pro Leu Val
50 55 60

Gly Gly His Glu Gly Ala Gly Val Val Val Gly Met Gly Glu Asn Val
 65 70 75 80
 Lys Gly Trp Lys Ile Gly Asp Tyr Ala Gly Ile Lys Trp Leu Asn Gly
 85 90 95
 Ser Cys Met Ala Cys Glu Tyr Cys Glu Leu Gly Asn Glu Ser Asn Cys
 100 105 110
 Pro His Ala Asp Leu Ser Gly Tyr Thr His Asp Gly Ser Phe Gln Glu
 115 120 125
 Tyr Ala Thr Ala Asp Ala Val Gln Ala Ala His Ile Pro Gln Gly Thr
 130 135 140
 Asp Leu Ala Glu Val Ala Pro Ile Leu Cys Ala Gly Ile Thr Val Tyr
 145 150 155 160
 Lys Ala Leu Lys Ser Ala Asn Leu Arg Ala Gly His Trp Ala Ala Ile
 165 170 175
 Ser Gly Ala Ala Gly Gly Leu Gly Ser Leu Ala Val Gln Tyr Ala Lys
 180 185 190
 Ala Met Gly Tyr Arg Val Leu Gly Ile Asp Gly Gly Pro Gly Lys Glu
 195 200 205
 Glu Leu Phe Thr Ser Leu Gly Gly Glu Val Phe Ile Asp Phe Thr Lys
 210 215 220
 Glu Lys Asp Ile Val Ser Ala Val Val Lys Ala Thr Asn Gly Gly Ala
 225 230 235 240
 His Gly Ile Ile Asn Val Ser Val Ser Glu Ala Ala Ile Glu Ala Ser
 245 250 255
 Thr Arg Tyr Cys Arg Ala Asn Gly Thr Val Val Leu Val Gly Leu Pro
 260 265 270
 Ala Gly Ala Lys Cys Ser Ser Asp Val Phe Asn His Val Val Lys Ser
 275 280 285
 Ile Ser Ile Val Gly Ser Tyr Val Gly Asn Arg Ala Asp Thr Arg Glu
 290 295 300
 Ala Leu Asp Phe Phe Ala Arg Gly Leu Val Lys Ser Pro Ile Lys Val

305 310 315 320
Val Gly Leu Ser Ser Leu Pro Glu Ile Tyr Glu Lys Met Glu Lys Gly

 325 330 335
Gln Ile Ala Gly Arg Tyr Val Val Asp Thr Ser Lys

 340 345

<210> 66

<211> 1044

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 66

atgtccatcc	cggaaaccca	gaaagccatt	atctttttag	agagcaacgg	caagctcgaa	60
cataaggata	ttccagtgcc	aaaaccgaag	ccaaacgagc	tcctgatcaa	cgttaagtac	120
tccggagtgt	gccacacgga	tctgcacgct	tggcatggcg	attggccctt	cctactaag	180
ctgccactgg	tcggcggaca	cgaaggcgct	ggtgtggtgg	tgggaatggg	tgagaacgtt	240
aaaggctgga	agattggcga	ctacgctggc	atcaaattggc	tgaatggttc	ctgcatggct	300
tgtgaatact	gtgagcttgg	taatgaaagc	aactgcccgc	acgccgattt	gagcggatac	360
acacatgatg	gtctcttcca	ggagtacgct	accgccgacg	ctgtgcaagc	agcgcacatt	420

ccacaaggca ccgacctcgc tgaagttgct ccaatcctgt gtgctggcat caccgtatac	480
aaggctctca aatccgcaaa cctgcgcgcg ggccactggg cagcgatttc aggcgctgcc	540
ggcggctctg gttccttggc tgtgcaatac gcaaaaggcta tgggataccg agtactggga	600
attgacggcg gaccaggaaa agaggaactt ttcacgtccc tcggcggcga ggctcttcac	660
gattttacca aagagaagga catcgtctcc gctgtcgtga aggcaactaa cggcggcgct	720
cacggtatca tcaacgtgtc tgtttcggag gcggctatcg aggctagcac ccgatactgc	780
cgtgcaaatt gcacggtcgt gtttggttgc ctgccagcgg gcgcaaagtg ctctttgat	840

gtcttcaacc atgttgtaa gtccatttcg attgtgggct catabgttgg caatcgcgct 900
gacactcgtg aggcgttgga cttcttcgcg cgcgccctcg taaagtcgc tatcaaggtt 960
gtcggtttga gcagcctgcc cgaaatctac gaaaaaatgg aaaagggtca aattgccggc 1020
cgatatgtcg tcgatacttc taaa 1044

<210> 67

<211> 497

<212> PRT

<213> Yarrowia lipolytica CLIB 122 / E 150

<400> 67

Met Ser Ala Lys Ser Ile His Glu Ala Asp Gly Lys Ala Leu Leu Ala
1 5 10 15

His Phe Leu Ser Lys Ala Pro Val Trp Ala Glu Gln Gln Pro Ile Asn
20 25 30

Thr Phe Glu Met Gly Thr Pro Lys Leu Ala Ser Leu Thr Phe Glu Asp
35 40 45

Gly Val Ala Pro Glu Gln Ile Phe Ala Ala Ala Glu Lys Thr Tyr Pro
50 55 60

Trp Leu Leu Glu Ser Gly Ala Lys Phe Val Ala Lys Pro Asp Gln Leu
65 70 75 80

Ile Lys Arg Arg Gly Lys Ala Gly Leu Leu Val Leu Asn Lys Ser Trp
85 90 95

Glu Glu Cys Lys Pro Trp Ile Ala Glu Arg Ala Ala Lys Pro Ile Asn
100 105 110

Val Glu Gly Ile Asp Gly Val Leu Arg Thr Phe Leu Val Glu Pro Phe
115 120 125

Val Pro His Asp Gln Lys His Glu Tyr Tyr Ile Asn Ile His Ser Val
130 135 140

Arg Glu Gly Asp Trp Ile Leu Phe Tyr His Glu Gly Gly Val Asp Val
145 150 155 160

Gly Asp Val Asp Ala Lys Ala Ala Lys Ile Leu Ile Pro Val Asp Ile
165 170 175

Glu Asn Glu Tyr Pro Ser Asn Ala Thr Leu Thr Lys Glu Leu Leu Ala
180 185 190

His Val Pro Glu Asp Gln His Gln Thr Leu Leu Asp Phe Ile Asn Arg
195 200 205

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

Pro Leu Val Val Ile Pro Thr Ala Gln Gly Val Glu Val His Tyr Leu

225 230 235 240
 Asp Leu Ala Gly Lys Leu Asp Gln Thr Ala Glu Phe Glu Cys Gly Pro
 245 250 255
 Lys Trp Ala Ala Ala Arg Ser Pro Ala Ala Leu Gly Gln Val Val Thr
 260 265 270

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

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

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

 Ser Gly Ile Val Pro Leu Ala Leu Leu Gly Lys Arg Pro Lys Asn Val
 465 470 475 480

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

485

490

495

Val

<210> 68

<211> 1491

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 68

atgagcgcca agagcatcca cgagctgat ggcaaagcac tccttgcca cttctgtcc 60
aaagcaccag tctgggcaga gcaacagccc atcaatacct ttgagatggg tacccccaaa 120

ctggcaagcc tgactttcga ggatgggtga ggcctgagc aaatcttcgc cgctgcagaa 180
aaaacctacc cgtggctgct tgaatcaggt gctaagtctg tcgccaagcc tgatcagctg 240
attaagcgac gtggttaaggc gggcttgctt gtgcttaaca aatcctggga agaataag 300
ccctggattg ccgagcgccg agcgaagcct attaatgtag agggaatcga cggcgtgttg 360
cgtacctttc tctgtgaacc atttgtcccc catgatcaga agcatgaata ttacatcaac 420
atccatagcg tccgcgaagg tgattggatc ctgttttacc acgaaggtgg tgcgacgtg 480
ggcgacgttg atgctaaggc agcaaaaatc ctgatcccg ttgatattga aaacgagtac 540

ccatcaaatg caacctcac caaggagcct ttggctcatg tgcctgaaga tcagcaccaa 600
accttgctcg atttcatcaa ccgctcttac gcggtatag tagatctga attcacctac 660
ttggaaatta acccgctcgt ggtgattcct acggctcagg gactcgaggt gcactatctc 720
gatttggcag gtaagttgga ccagaccgca gaattcgaat gtggacctaa atggcgggcc 780
gctcgttccc ccgcgccctt gggacaggtc gttacgatc acgcggttag caccaaggtc 840
tcgattgacg ctggtccagc catggtgttc ccagccccgt ttggccgcga actgagcaaa 900
gaagaagcat acatcgcgga actggattct aagacgggtg cctctctcaa gcttactgtc 960

ttgaacgcaa agggccgcat ctggaccctt gtcgcaggtg gaggagcaag cgtggtctac 1020
gcagacgca tcgcctccgc tggatttgc gatgaactgg ctaattacgg agagtactcg 1080
ggtgcaccga acgagacca gacttatgaa tatgccaaga ccgtcctcga tctcatgacc 1140
cgtggagacg cgacccaga gggaaaggtg ctgttcacg gcggaggcat cgccaacttc 1200
accaggtag gctccacttt caaaggtatt atccgcccct tccgtgacta ccaatcatcc 1260
cttcacaatc acaaggttaa gatctacgtc cgcccggtg gtccaaactg gcaggaaggt 1320

ctgcgactta tcaaattctgc aggcgcagcag ctgaatctcc caatggagat ttacggccca 1380

gatatgcacg tcagcggcat cgtgccctc gctctgttg gtaaagccc taagaatgtg 1440

aagccattcg gaaccggccc ttccactgag gcatcgactc cactgggtgt g 1491

<210> 69

<211> 887

<212> PRT

<213> E. coli (strain K12)

<400> 69

Met Ser Glu Arg Phe Pro Asn Asp Val Asp Pro Ile Glu Thr Arg Asp

1 5 10 15

Trp Leu Gln Ala Ile Glu Ser Val Ile Arg Glu Glu Gly Val Glu Arg

20 25 30

Ala Gln Tyr Leu Ile Asp Gln Leu Leu Ala Glu Ala Arg Lys Gly Gly

35 40 45

Val Asn Val Ala Ala Gly Thr Gly Ile Ser Asn Tyr Ile Asn Thr Ile

50 55 60

Pro Val Glu Glu Gln Pro Glu Tyr Pro Gly Asn Leu Glu Leu Glu Arg

65 70 75 80

Arg Ile Arg Ser Ala Ile Arg Trp Asn Ala Ile Met Thr Val Leu Arg

85 90 95

Ala Ser Lys Lys Asp Leu Glu Leu Gly Gly His Met Ala Ser Phe Gln

100 105 110

Ser Ser Ala Thr Ile Tyr Asp Val Cys Phe Asn His Phe Phe Arg Ala

115 120 125

Arg Asn Glu Gln Asp Gly Gly Asp Leu Val Tyr Phe Gln Gly His Ile

130 135 140

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

145 150 155 160

Glu Gln Leu Asp Asn Phe Arg Gln Glu Val His Gly Asn Gly Leu Ser

165 170 175

Ser Tyr Pro His Pro Lys Leu Met Pro Glu Phe Trp Gln Phe Pro Thr

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

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

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

 355 360 365
 Lys Ile Tyr Ala Ala Phe Lys Lys Ala Gln Glu Thr Lys Gly Lys Ala
 370 375 380
 Thr Val Ile Leu Ala His Thr Ile Lys Gly Tyr Gly Met Gly Asp Ala
 385 390 395 400
 Ala Glu Gly Lys Asn Ile Ala His Gln Val Lys Lys Met Asn Met Asp
 405 410 415
 Gly Val Arg His Ile Arg Asp Arg Phe Asn Val Pro Val Ser Asp Ala

 420 425 430

Asp Ile Glu Lys Leu Pro Tyr Ile Thr Phe Pro Glu Gly Ser Glu Glu
 435 440 445
 His Thr Tyr Leu His Ala Gln Arg Gln Lys Leu His Gly Tyr Leu Pro
 450 455 460
 Ser Arg Gln Pro Asn Phe Thr Glu Lys Leu Glu Leu Pro Ser Leu Gln
 465 470 475 480
 Asp Phe Gly Ala Leu Leu Glu Glu Gln Ser Lys Glu Ile Ser Thr Thr
 485 490 495
 Ile Ala Phe Val Arg Ala Leu Asn Val Met Leu Lys Asn Lys Ser Ile
 500 505 510
 Lys Asp Arg Leu Val Pro Ile Ile Ala Asp Glu Ala Arg Thr Phe Gly
 515 520 525
 Met Glu Gly Leu Phe Arg Gln Ile Gly Ile Tyr Ser Pro Asn Gly Gln
 530 535 540
 Gln Tyr Thr Pro Gln Asp Arg Glu Gln Val Ala Tyr Tyr Lys Glu Asp
 545 550 555 560
 Glu Lys Gly Gln Ile Leu Gln Glu Gly Ile Asn Glu Leu Gly Ala Gly
 565 570 575
 Cys Ser Trp Leu Ala Ala Ala Thr Ser Tyr Ser Thr Asn Asn Leu Pro
 580 585 590
 Met Ile Pro Phe Tyr Ile Tyr Tyr Ser Met Phe Gly Phe Gln Arg Ile
 595 600 605
 Gly Asp Leu Cys Trp Ala Ala Gly Asp Gln Gln Ala Arg Gly Phe Leu
 610 615 620
 Ile Gly Gly Thr Ser Gly Arg Thr Thr Leu Asn Gly Glu Gly Leu Gln
 625 630 635 640
 His Glu Asp Gly His Ser His Ile Gln Ser Leu Thr Ile Pro Asn Cys
 645 650 655
 Ile Ser Tyr Asp Pro Ala Tyr Ala Tyr Glu Val Ala Val Ile Met His
 660 665 670
 Asp Gly Leu Glu Arg Met Tyr Gly Glu Lys Gln Glu Asn Val Tyr Tyr

675 680 685
Tyr Ile Thr Thr Leu Asn Glu Asn Tyr His Met Pro Ala Met Pro Glu
690 695 700
Gly Ala Glu Glu Gly Ile Arg Lys Gly Ile Tyr Lys Leu Glu Thr Ile
705 710 715 720
Glu Gly Ser Lys Gly Lys Val Gln Leu Leu Gly Ser Gly Ser Ile Leu
725 730 735
Arg His Val Arg Glu Ala Ala Glu Ile Leu Ala Lys Asp Tyr Gly Val

740 745 750
Gly Ser Asp Val Tyr Ser Val Thr Ser Phe Thr Glu Leu Ala Arg Asp
755 760 765
Gly Gln Asp Cys Glu Arg Trp Asn Met Leu His Pro Leu Glu Thr Pro
770 775 780
Arg Val Pro Tyr Ile Ala Gln Val Met Asn Asp Ala Pro Ala Val Ala
785 790 795 800
Ser Thr Asp Tyr Met Lys Leu Phe Ala Glu Gln Val Arg Thr Tyr Val

805 810 815
Pro Ala Asp Asp Tyr Arg Val Leu Gly Thr Asp Gly Phe Gly Arg Ser
820 825 830
Asp Ser Arg Glu Asn Leu Arg His His Phe Glu Val Asp Ala Ser Tyr
835 840 845
Val Val Val Ala Ala Leu Gly Glu Leu Ala Lys Arg Gly Glu Ile Asp
850 855 860
Lys Lys Val Val Ala Asp Ala Ile Ala Lys Phe Asn Ile Asp Ala Asp

865 870 875 880
Lys Val Asn Pro Arg Leu Ala
885

<210> 70

<211> 2661

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 70

atgagcgagc gcttcccaaa tgacgttgat cccattgaga cccgcgattg gttgcaggcg	60
atcgagtcgg tcatccgtga ggagggagtc gaacgagccc agtatattgat cgatcaactg	120
cttgacagaag cccgcaaggg aggcgtgaat gtcgcggctg gtaccggcat tagcaattac	180
atcaacacca tcccagtcga ggagcaacca gactaccctg gtaatcttga actcgaacgc	240
cgtattcgca ggcatttcg ctggaacgca attatgaccg ttctgcgcgc gtcaaagaaa	300
gacctcgaac tgggaggcca tatggcatcc tttaactct cagccaccat ctacgacgtt	360
tgttttaacc acttcttcg cgcgcgtaat gaacaagatg gcggagacct ggtttatatt	420
cagggtcata tctgcgggg cgtctacgca cgcgccttc tggaaggacg cctgacgcaa	480
gaacagctcg acaatttcg acaagaagtg cacggcaacg gactgtcctc ttaccgcat	540
ccaaaactga tgcctgagtt ctggcaattt ccaacggtgt ccatgggtct cggcccaatc	600
ggtgcaatct accaggcaaa gtttctgaag tacctggagc accgcggcct caaagatacg	660
tccaagcaaa ccgtctacgc gtttctgggt gacggagaaa tggatgagcc tgaatccaag	720
ggcgctatta cgattgccac tcgcgaaaag ctgcacaacc tggctctcgt catcaattgc	780
aacctgcaac gattggacgg tccggttacg ggcaatggca agatcattaa tgagctcgaa	840
ggcatcttcg agggcgctgg ctggaacgtg attaaggtga tgtggggttc tcgctgggac	900
gaactgctgc gcaaagacac cagcggcaag ctgattcagc tgatgaacga gaccgtggat	960
ggtgattacc agaccttcaa gtccaaggat ggagcctacg tccgtgagca cttttttgga	1020
aaatatcctg aaacggcagc attggtggca gattggaccg acgagcagat ctgggcgctg	1080
aatcgcggtg gccatgacct taagaaaatc tacgcggcct tcaagaaggc acaagaaacc	1140
aagggcaaa cactgtcat cttggcgcat accatcaaag gctacggtat gggcgatgcc	1200
gctgagggca aaaacatcgc tcatcaagtt aagaagatga acatggatgg tgtccgccac	1260
atccgcgacg gctttaacgt cccgtctct gacgcagata ttgagaagct gccatacatc	1320
accttcccag aaggctccga ggaacatacc tatctccag cacagcgcca gaagctccat	1380
ggttaccttc cgtcgcgtca gccaaatttc accgaaaaat tggaattgcc ctcatgcaa	1440
gattttggcg cgctcttgga ggaacagtct aaagagatct ctacaacat cgcgtttgtg	1500
cgtgcccetca acgtcatgtt gaagaacaag tctatcaagg atcgattggt gccattatc	1560
gctgatgaag cccgcacctt cggtatggag ggtttgtttc gccagattgg catctactct	1620
cccaatggcc agcaatacac cccacaggat cgcgaacaag ttgcctacta taaggaagac	1680
gagaaaggcc agatccttca agaaggcatc aacgaacttg gcgcgggatg cagctggctc	1740
gccgctgcca caagctactc aactaacaac ctcccgatga tccctttcta tatctactat	1800

tcgatgttcg gcttccaacg cattggcgac ttgtgctggg cggcaggcga tcaacaagcc 1860
 cgcggctttc tgattggcgg tacttcgggt cgcactacgc tcaacggcga aggtctccag 1920

catgaggatg gtcattcgca catccaatca ctccaccattc caaactgtat ttcctatgac 1980
 cccgcatacg cttacgaggt tgcggtcatc atgcacgatg gtcctggaacg tatgtacggc 2040
 gagaagcagg aaaaatgtga ttactacatc accaccctga atgaaaatta ccatatgcct 2100
 gccatgcctg aggggtgcga agagggaatt cgtagggtta tctacaagct ggaaacaatt 2160
 gaaggatcca agggtaaggt acagcttctg ggatctggct ccattcttcg tcatgtgcgc 2220
 gaagcagcgg aaatcctcgc aaaagattac ggcgtaggca ggcacgttta ctccgttacg 2280
 agcttcactg aactggcccg cgacggccag gactgtgagc gttggaacat gctgcatccc 2340

cttgaacccc cagcgtgcc gtatatcgca caagtaatga atgacgcccc tgccgtcgcc 2400
 tcgactgact atatgaagct cttcgccgaa caggctcgca cgtacgttcc agcagatgat 2460
 taccggttc tgggcaccga cggctttggc cgctccgatt cagcgagaa cctgcgacac 2520
 cacttcgagg tcgacgttc ctacgtggta gtactgcgc tcggagaact ggcaaagcgt 2580
 ggcgagatcg acaaaaaggt agttgtgat gctatcgcta agttcaacat tgacgccgac 2640
 aaggtcaatc cccgcttggc a 2661

<210> 71
 <211> 120
 <212> PRT
 <213> S. cerevisiae S288c

<400> 71

Met Pro Lys Val Tyr Ser Tyr Gln Glu Val Ala Glu His Asn Gly Pro

1 5 10 15

Glu Asn Phe Trp Ile Ile Ile Asp Asp Lys Val Tyr Asp Val Ser Gln

20 25 30

Phe Lys Asp Glu His Pro Gly Gly Asp Glu Ile Ile Met Asp Leu Gly

35 40 45

Gly Gln Asp Ala Thr Glu Ser Phe Val Asp Ile Gly His Ser Asp Glu

50 55 60

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

65 70 75 80

Ser Glu Arg Val Ser Val Glu Lys Val Ser Thr Ser Glu Asn Gln Ser

85 90 95
Lys Gly Ser Gly Thr Leu Val Val Ile Leu Ala Ile Leu Met Leu Gly
100 105 110

Val Ala Tyr Tyr Leu Leu Asn Glu
115 120

<210> 72

<211> 360

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 72

atgccaaagg tctattctta tcaggaagtc gcagaacaca acggcccaga aaacttctgg 60
atcattatcg atgataaggt ttacgatgtg tcgcagtcca aagacgaaca cccaggtgga 120
gatgaaatta ttatggatct tggcggacaa gatgcgacgg aaagctttgt tgatatcggt 180
cactcagacg aagccttgcg cctcctcaag ggcctttaca tcggtgacgt agataaaacc 240
tcggagcgcg tttccgtcga gaaagtttcc acctcagaga accaatccaa aggtccggt 300
accctcgtgg tgatcctcgc tattctgatg ttgggtgtcg cctattatct tctgaatgag 360

<210> 73

<211> 485

<212> PRT

<213> Emericella nidulans ATCC 38163

<400> 73

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

50 55 60
Pro Trp Leu Leu Thr Pro Gly Ser Lys Phe Val Ala Lys Pro Asp Gln
65 70 75 80

Leu Ile Lys Arg Arg Gly Lys Ser Gly Leu Leu Ala Leu Asn Lys Thr
 85 90 95
 Trp Ala Glu Ala Arg Glu Trp Ile Glu Ala Arg Ala Thr Lys Glu Gln
 100 105 110
 Gln Val Glu Thr Val Val Gly Val Leu Arg His Phe Leu Val Glu Pro
 115 120 125
 Phe Val Pro His Pro Gln Glu Thr Glu Tyr Tyr Ile Asn Ile His Ser
 130 135 140
 Val Arg Glu Gly Asp Trp Ile Leu Phe Thr His Glu Gly Gly Val Asp
 145 150 155 160
 Val Gly Asp Val Asp Ala Lys Ala Glu Lys Leu Leu Ile Pro Val Asn
 165 170 175
 Leu Lys Asn Tyr Pro Ser Asn Glu Glu Ile Ala Ser Ala Leu Leu Ser
 180 185 190
 Lys Val Pro Lys Gly Ile His Asn Val Leu Val Asp Phe Ile Ser Arg
 195 200 205
 Leu Tyr Ala Val Tyr Val Asp Cys Gln Phe Thr Tyr Leu Glu Ile Asn
 210 215 220
 Pro Leu Val Val Ile Pro Asn Ala Asp Ala Thr Ser Ala Asp Val His
 225 230 235 240
 Phe Leu Asp Leu Ala Ala Lys Leu Asp Gln Thr Ala Glu Phe Glu Cys
 245 250 255
 Gly Thr Lys Trp Ala Val Ala Arg Ser Pro Ala Asn Leu Gly Leu Ala
 260 265 270
 Ala Leu Pro Thr Ser Asp Lys Val Asn Ile Asp Ala Gly Pro Pro Met
 275 280 285
 Glu Phe Pro Ala Pro Phe Gly Arg Glu Leu Ser Lys Glu Glu Lys Phe
 290 295 300
 Ile Ser Asp Met Asp Ala Lys Thr Gly Ala Ser Leu Lys Leu Thr Val
 305 310 315 320
 Leu Asn Pro Asn Gly Arg Val Trp Thr Leu Val Ala Gly Gly Gly Ala

325 330 335
 Ser Val Val Tyr Ala Asp Ala Ile Ala Ser Ala Gly Phe Val Ser Glu
 340 345 350
 Leu Ala Asn Tyr Gly Glu Tyr Ser Gly Ala Pro Thr Glu Thr Gln Thr
 355 360 365
 Phe Asn Tyr Ala Arg Thr Ile Leu Asp Leu Met Leu Arg Ser Pro Ile

370 375 380
 His Pro Asp Gly Lys Val Leu Phe Ile Gly Gly Gly Ile Ala Asn Phe
 385 390 395 400
 Thr Asn Val Ala Ser Thr Phe Lys Gly Val Ile Arg Ala Leu Arg Glu
 405 410 415
 Val Ala Pro Val Leu Asn Glu His Lys Val Gln Ile Trp Val Arg Arg
 420 425 430
 Ala Gly Pro Asn Tyr Gln Glu Gly Leu Lys Asn Ile Lys Ala Val Gly

435 440 445
 Glu Glu Leu Gly Leu Asn Met His Val Tyr Gly Pro Glu Met His Val
 450 455 460
 Ser Gly Ile Val Pro Leu Ala Leu Gln Gly Lys Gln Thr Asp Ile Lys
 465 470 475 480
 Glu Phe Gly Thr Ala
 485

<210> 74

<211> 1455

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 74

atgtcggcca agtcgatctt cgaagcagac ggaaaagcaa ttttgaatta ccacctgacc 60

cgcgaccag tgatcaaacc aaccctctg ccaccttcca acactcataa cccccctccg 120

aaattggctt ctctctactt cccggacgac ttgtccgtga aggacgttct cgatcaagca 180

gaagttacat atccgtggct tcttacccca ggtccaaat ttgtggccaa acctgatcag 240

ttgatcaaac gtcgcggaac atccggcctt ctggcgttga acaagacatg ggccgaagca 300

cgtgagtggg ttgaagctcg agcaactaag gaacaacagg tggaaactgt agtgggagtt 360
 ctccgtcact tcttggttga accttttgtg cctcaccctc aggagacgga gtactacatt 420
 aacatccact ccgtgcgcga ggggtgactgg atcctcttta cccacgaggg cgggtgtcgat 480

 gtcggcgatg tcgacgcaa agcagagaag ctgttgatcc cggatgaatct caagaactac 540
 ccctccaacg aggagatcgc atctgcattg ttgtccaagg tccctaaggg aatccataac 600
 gtccttgtgg actttatctc ccgtctgtac gcgggtctatg ttgattgtca atttacgtac 660
 ctcgaaatca acccattggt cgtgattccc aacgcggacg caacttccgc tgatgttcat 720
 tttcttgatc tcgtgcgaa gctggatcag acagccgaat tcgaatgcgg taccaagtgg 780
 gccgtcgctc gtcttccgc caacttgggc ctgcgacct tgcctacgag cgataaggtg 840
 aacattgacg cgggacctcc gatggaattt cctgccccct ttggtcgcga gctgtcaaaa 900

 gaagaaaaat ttatttccga catggatgct aagaccggag cgagccttaa gctcacgctc 960
 cttaatccga acggccgctg gtggaccctc gtcgccggcg gaggtgcctc tgtggtatac 1020
 gcagacgcga ttgcttctgc aggtctctgtg tccgaattgg ctaactacgg tgagtactcg 1080
 ggcgcacga ccgaaaccca aaccttcaac tacgcccgca ccatcttgga tctgatgctg 1140
 cgctcccaa ttcacccgga cggtaagggt cttttcattg gcggtggaat tgccaacttc 1200
 acaaacgttg cgagcacctt caaaggagtg atccgtgctc tgcgtgaagt ggcaccagtt 1260
 ttgaacgaac acaaagtga gatctgggtg cgccgcgcag gtccaaatta ccaagaaggt 1320

 ctgaaaaata tcaaagcggg ttggtgaggaa ctcgactta acatgcatgt ctacggccca 1380
 gaaatgcacg tcagcggcat tgtaccgtg gcactgcaag gcaagcaaac cgatattaaa 1440
 gaattcggca ccgcg 1455

<210> 75

<211> 630

<212> PRT

<213> E. coli (strain K12)

<400> 75

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

1 5 10 15

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

20 25 30

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

35 40 45

Pro Gln Ala Gly Ile Val Lys Glu Ile Lys Val Ser Val Gly Asp Lys
50 55 60

Thr Gln Thr Gly Ala Leu Ile Met Ile Phe Asp Ser Ala Asp Gly Ala
65 70 75 80

Ala Asp Ala Ala Pro Ala Gln Ala Glu Glu Lys Lys Glu Ala Ala Pro
85 90 95

Ala Ala Ala Pro Ala Ala Ala Ala Ala Lys Asp Val Asn Val Pro Asp
100 105 110

Ile Gly Ser Asp Glu Val Glu Val Thr Glu Ile Leu Val Lys Val Gly
115 120 125

Asp Lys Val Glu Ala Glu Gln Ser Leu Ile Thr Val Glu Gly Asp Lys
130 135 140

Ala Ser Met Glu Val Pro Ala Pro Phe Ala Gly Thr Val Lys Glu Ile
145 150 155 160

Lys Val Asn Val Gly Asp Lys Val Ser Thr Gly Ser Leu Ile Met Val
165 170 175

Phe Glu Val Ala Gly Glu Ala Gly Ala Ala Ala Pro Ala Ala Lys Gln
180 185 190

Glu Ala Ala Pro Ala Ala Ala Pro Ala Pro Ala Ala Gly Val Lys Glu
195 200 205

Val Asn Val Pro Asp Ile Gly Gly Asp Glu Val Glu Val Thr Glu Val
210 215 220

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

Val Glu Gly Asp Lys Ala Ser Met Glu Val Pro Ala Pro Phe Ala Gly
245 250 255

Val Val Lys Glu Leu Lys Val Asn Val Gly Asp Lys Val Lys Thr Gly
260 265 270

Ser Leu Ile Met Ile Phe Glu Val Glu Gly Ala Ala Pro Ala Ala Ala
275 280 285

Pro Ala Lys Gln Glu Ala Ala Ala Pro Ala Pro Ala Ala Lys Ala Glu

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

 Lys Gly Arg Ile Leu Arg Glu Asp Val Gln Ala Tyr Val Lys Glu Ala
 355 360 365
 Ile Lys Arg Ala Glu Ala Ala Pro Ala Ala Thr Gly Gly Gly Ile Pro
 370 375 380
 Gly Met Leu Pro Trp Pro Lys Val Asp Phe Ser Lys Phe Gly Glu Ile
 385 390 395 400
 Glu Glu Val Glu Leu Gly Arg Ile Gln Lys Ile Ser Gly Ala Asn Leu
 405 410 415

 Ser Arg Asn Trp Val Met Ile Pro His Val Thr His Phe Asp Lys Thr
 420 425 430
 Asp Ile Thr Glu Leu Glu Ala Phe Arg Lys Gln Gln Asn Glu Glu Ala
 435 440 445
 Ala Lys Arg Lys Leu Asp Val Lys Ile Thr Pro Val Val Phe Ile Met
 450 455 460
 Lys Ala Val Ala Ala Ala Leu Glu Gln Met Pro Arg Phe Asn Ser Ser
 465 470 475 480

 Leu Ser Glu Asp Gly Gln Arg Leu Thr Leu Lys Lys Tyr Ile Asn Ile
 485 490 495
 Gly Val Ala Val Asp Thr Pro Asn Gly Leu Val Val Pro Val Phe Lys
 500 505 510
 Asp Val Asn Lys Lys Gly Ile Ile Glu Leu Ser Arg Glu Leu Met Thr
 515 520 525
 Ile Ser Lys Lys Ala Arg Asp Gly Lys Leu Thr Ala Gly Glu Met Gln
 530 535 540

Gly Gly Cys Phe Thr Ile Ser Ser Ile Gly Gly Leu Gly Thr Thr His
 545 550 555 560
 Phe Ala Pro Ile Val Asn Ala Pro Glu Val Ala Ile Leu Gly Val Ser
 565 570 575
 Lys Ser Ala Met Glu Pro Val Trp Asn Gly Lys Glu Phe Val Pro Arg
 580 585 590
 Leu Met Leu Pro Ile Ser Leu Ser Phe Asp His Arg Val Ile Asp Gly
 595 600 605

Ala Asp Gly Ala Arg Phe Ile Thr Ile Ile Asn Asn Thr Leu Ser Asp
 610 615 620

Ile Arg Arg Leu Val Met

625 630

<210> 76

<211> 1890

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 76

atggccatcg aaatcaaggt gccagacatc ggtgccgatg aagtcgaaat caccgagatt 60
 ctgctcaaag tcggcgataa agtgaagca gaacagtcct tgatcaccgt agagggtgat 120
 aaggctagca tggaagttcc tagccgcag gctggcattg tgaaggaaat taaggtttca 180

gttgggtgaca aaactcaaac gggagccctt attatgatct tcgactcggc agacggagct 240
 gccgatgcgg cgccggccca agccgaagaa aagaaagaag cggctccagc agcagcaccc 300
 gccgctgccg cagctaagga tgtaacgtg cctgatatcg gctctgacga agtggaaagtc 360
 accgaaatct tggtaaggt ggggtgataag gtggaagcag aacagtcttt gattacggtt 420
 gaaggagata aggctagcat ggaagtgcc aacccttcg cgggcaccgt taaggaaatc 480
 aaagtgaatg tcggcgataa ggtatctacc ggctctctca ttatggtttt cgaagtcgcg 540
 ggtgaagcag gagccgccgc accagccgca aagcaagagg cagcaccagc ggcagcacct 600

gctccggctg caggcgtgaa ggaggtaaac gtccctgata ttggtggcga tgaagttgag 660
 gtgaccgagg tgatggtgaa agttggagat aaggttgag ctgaacaatc cttgattact 720
 gtggagggtg ataaggcgtc tatggaagtt cccgtccat ttgccggcgt cgtcaaggag 780
 ctttaaggtga acgttgaga caaggtgaag accggaagcc tcacatgat ttgcgaagtt 840

gaggggtgcag cccctgccgc agctccagcc aagcaagagg cagccgctcc cgcgccagcg 900
gcgaaagctg aggtctctgc agcggccccc gccgctaaag ccgaaggcaa atccgaattc 960
gctgagaatg acgtttatgt acacgaaca cctctgatcc gtcgcctggc acgcaattt 1020

ggagtcaacc tcgcaaaggt caagggcaca ggccgcaagg gtcgaatcct tcgcaagac 1080
gttcaagcat atgttaagga agccatcaag cgcgcggaag cggcaccggc agccaccggt 1140
ggtggaatcc cgggcatgtt gccctggccc aaagtggact tctcgaagtt tggcgagatt 1200
gaggaagtgg agctgggtcg catccagaaa atttccggcg caaacctgtc ccgcaactgg 1260
gtcatgattc ctcacgtgac ccatttcgat aaaaccgata ttaccgagtt ggaagcgttt 1320
cgtaaacac aaaacgaaga agcggctaag cgtgaagctg acgttaaaat taccctcggt 1380
gtctttatca tgaaggcggg agctgctgct ctcgaacaaa tgccgcgctt caactcttcc 1440

cttagcgagg acggacaacg ccttacattg aagaagtaca tcaatcgcg cgttgcgtgc 1500
gatactccca acggcctcgt ggtccagtg ttttaaggatg ttaacaagaa gggtatcatt 1560
gagctcagcc gtgagctgat gacgatcagc aagaaggcac gcgatggtaa gctgaccgcg 1620
ggagaaatgc aagggtgctg cttcactatc tcgtctatcg gaggccttgg aacaacgcac 1680
ttcgcgccga tcgtcaacgc accagagggtg gcgatcctgg gcgtttccaa atctgcgatg 1740
gagcctgtct ggaatggcaa agagtctgta cctcgactta tgcttccaat ttccctgagc 1800
ttcgaccatc gcgtgatcga tggagcagat ggtgcacgct tcattacat tatcaataac 1860

acgtgtcgg acatccgacg attggttatg 1890

<210> 77

<211> 352

<212> PRT

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 77

Met Ala Ser Glu Lys Glu Ile Arg Arg Glu Arg Phe Leu Asn Val Phe

1 5 10 15

Pro Lys Leu Val Glu Glu Leu Asn Ala Ser Leu Leu Ala Tyr Gly Met

20 25 30

Pro Lys Glu Ala Cys Asp Trp Tyr Ala His Ser Leu Asn Tyr Asn Thr

35 40 45

Pro Gly Gly Lys Leu Asn Arg Gly Leu Ser Val Val Asp Thr Tyr Ala

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

 100 105 110
 Pro Cys Trp Tyr Lys Val Pro Glu Val Gly Glu Ile Ala Ile Trp Asp
 115 120 125
 Ala Phe Met Leu Glu Ala Ala Ile Tyr Lys Leu Leu Lys Ser His Phe
 130 135 140
 Arg Asn Glu Lys Tyr Tyr Ile Asp Ile Thr Glu Leu Phe His Glu Val
 145 150 155 160
 Thr Phe Gln Thr Glu Leu Gly Gln Leu Met Asp Leu Ile Thr Ala Pro

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

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

 290 295 300

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

<210> 78

<211> 1056

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 78

atggcatcgg agaaggaaat ccgtcgtgaa cgattcctga acgtattccc aaagctcgtc	60
gaagagctta atgcttcctt gcttgcgtag ggaatgccga aggaagcgtg cgactggtat	120
gcgcattctc tcaactataa cacacctggc ggtaagctga atcgcggtct cagcgtgggtg	180
gacacgtacg caatcctcag caataagacc gtggaacagc tgggacagga ggagtacgag	240
aaagttgcta tccttggatg gtgtatcgaa ctgctgcaag cctattggtt ggtagctgat	300
gacatgatgg ataagagcat cactcgacgt ggtcaacat gctggtacaa ggtgccggag	360
gttggcgaaa tcgcatctg ggatgcgttc atgttgagg cgcgaattta taagcttctc	420
aagtccact ttcgtaacga aaagtactat attgacatca cagaactctt tcatgaagtg	480
accttccaaa ccgagctcgg ccaactcatg gatttgatca ccgccccga agataaggtg	540
gaccttctga agttctccct caagaagcac tcgttcattg ttaccttcaa gacagcatac	600
tactcathtt acctgccagt cgccctcgca atgtacgtgg cgggcatcac tgatgaaaag	660
gatctgaagc aggccccga cgtccttatc cctcttggag agtacttcca gatccaggac	720
gattacttgg attgttttgg caccacagag cagattggta aaatcggtac tgatatccaa	780
gacaacaaat gttcctgggt aattaacaag gctctcgagt tggcttccgc ggaacaacgc	840
aagactcttg acgaaaatta cggtgaagaa gattcagtcg ctgaagccaa gtgcaagaaa	900
atctttaacg atcttaaaat cgaacagttg taccacgagt atgaagagtc tatcgctaaa	960
gacctgaaag ctaaaatttc ccaggtcgat gagtcacgcg gcttcaaagc ggatgtgctc	1020
accgccttcc ttaacaaggt gtataagcgc tccaaa	1056

<210> 79

<211> 299

<212> PRT

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 79

Met Asp Phe Pro Gln Gln Leu Glu Ala Cys Val Lys Gln Ala Asn Gln

1 5 10 15

Ala Leu Ser Arg Phe Ile Ala Pro Leu Pro Phe Gln Asn Thr Pro Val

20 25 30

Val Glu Thr Met Gln Tyr Gly Ala Leu Leu Gly Gly Lys Arg Leu Arg

35 40 45

Pro Phe Leu Val Tyr Ala Thr Gly His Met Phe Gly Val Ser Thr Asn

50 55 60

Thr Leu Asp Ala Pro Ala Ala Ala Val Glu Cys Ile His Ala Tyr Phe

65 70 75 80

Leu Ile His Asp Asp Leu Pro Ala Met Asp Asp Asp Asp Leu Arg Arg

85 90 95

Gly Leu Pro Thr Cys His Val Lys Phe Gly Glu Ala Asn Ala Ile Leu

100 105 110

Ala Gly Asp Ala Leu Gln Thr Leu Ala Phe Ser Ile Leu Ser Asp Ala

115 120 125

Asp Met Pro Glu Val Ser Asp Arg Asp Arg Ile Ser Met Ile Ser Glu

130 135 140

Leu Ala Ser Ala Ser Gly Ile Ala Gly Met Cys Gly Gly Gln Ala Leu

145 150 155 160

Asp Leu Asp Ala Glu Gly Lys His Val Pro Leu Asp Ala Leu Glu Arg

165 170 175

Ile His Arg His Lys Thr Gly Ala Leu Ile Arg Ala Ala Val Arg Leu

180 185 190

Gly Ala Leu Ser Ala Gly Asp Lys Gly Arg Arg Ala Leu Pro Val Leu

195 200 205

Asp Lys Tyr Ala Glu Ser Ile Gly Leu Ala Phe Gln Val Gln Asp Asp

210 215 220
 Ile Leu Asp Val Val Gly Asp Thr Ala Thr Leu Gly Lys Arg Gln Gly
 225 230 235 240
 Ala Asp Gln Gln Leu Gly Lys Ser Thr Tyr Pro Ala Leu Leu Gly Leu
 245 250 255
 Glu Gln Ala Arg Lys Lys Ala Arg Asp Leu Ile Asp Asp Ala Arg Gln

260 265 270
 Ser Leu Lys Gln Leu Ala Glu Gln Ser Leu Asp Thr Ser Ala Leu Glu
 275 280 285
 Ala Leu Ala Asp Tyr Ile Ile Gln Arg Asn Lys
 290 295

<210> 80

<211> 897

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 80

atggacttcc cgcagcagct cgaggcctgc gttaagcagg ccaatcaggc gttgtcacgt 60
 ttcacgcgtc cctgccttt ccagaacaca ccagtagtgg aaaccatgca atacggcgcc 120

ctgctgggtg gaaaacgctt gcgcccattc ttggtgtacg caaccggcca catgttcggc 180
 gtcttacta acacccttga tgcacctgca gcggccgtgg agtgtatcca cgcctacttc 240
 ttgatccacg acgatttgcc ggctatggat gatgatgac tccgacgagg cctccctacc 300
 tgccatgtaa aattcggcga agcgaatgcg atcttggcag gtgatgcgt gcagacactt 360
 gccttttcga ttctgagcga cgcgatatg ccggaagtct ccgaccgaga ccgcattccc 420
 atgatcagcg agttggcttc tgctccggt atcgcgggca tgtgcggagg tcaggctctt 480
 gaccttgacg cagagggcaa gcacgttccc ttggacgctt tggagcgtat ccaccgcat 540

aagaccggcg ccttgatccg cgcccggtc cgcttggcg cgctgtctgc gggcgataag 600
 ggacgccg cactccagc gtcgataag tacgccgaat ccattggcct tgcttttcag 660
 gtgcaggatg acatccttga ttagtcggc gataccgcaa ctctcggcaa acgccaggcg 720
 gcggaccaac aattgggtaa gtcaacctac ccagctctgc tgggcttgga gcaagcacga 780
 aagaaagccc gcgacctgat cgacgatgcc cgccaatctc ttaacaact ggccgaacaa 840
 tctctcgata ccagcgcaact cgaggcactc gcagattaca tcatccaacg taataag 897

<210> 81

<211> 474

<212> PRT

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 81

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

1 5 10 15

Ala Gly Tyr Ser Ala Ala Phe Arg Cys Ala Asp Leu Gly Leu Glu Thr

20 25 30

Val Ile Val Glu Arg Tyr Asn Thr Leu Gly Gly Val Cys Leu Asn Val

35 40 45

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

50 55 60

Glu Ala Lys Ala Leu Ala Glu His Gly Ile Val Phe Gly Glu Pro Lys

65 70 75 80

Thr Asp Ile Asp Lys Ile Arg Thr Trp Lys Glu Lys Val Ile Asn Gln

85 90 95

Leu Thr Gly Gly Leu Ala Gly Met Ala Lys Gly Arg Lys Val Lys Val

100 105 110

Val Asn Gly Leu Gly Lys Phe Thr Gly Ala Asn Thr Leu Glu Val Glu

115 120 125

Gly Glu Asn Gly Lys Thr Val Ile Asn Phe Asp Asn Ala Ile Ile Ala

130 135 140

Ala Gly Ser Arg Pro Ile Gln Leu Pro Phe Ile Pro His Glu Asp Pro

145 150 155 160

Arg Ile Trp Asp Ser Thr Asp Ala Leu Glu Leu Lys Glu Val Pro Glu

165 170 175

Arg Leu Leu Val Met Gly Gly Gly Ile Ile Ala Leu Glu Met Ala Thr

180 185 190

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

195 200 205
 Gln Val Ile Arg Ala Ala Asp Lys Asp Ile Val Lys Val Phe Thr Lys
 210 215 220
 Arg Ile Ser Lys Lys Phe Asn Leu Met Leu Glu Thr Lys Val Thr Ala
 225 230 235 240
 Val Glu Ala Lys Glu Asp Gly Ile Tyr Val Thr Met Glu Gly Lys Lys

 245 250 255
 Ala Pro Ala Glu Pro Gln Arg Tyr Asp Ala Val Leu Val Ala Ile Gly
 260 265 270
 Arg Val Pro Asn Gly Lys Asn Leu Asp Ala Gly Lys Ala Gly Val Glu
 275 280 285
 Val Asp Asp Arg Gly Phe Ile Arg Val Asp Lys Gln Leu Arg Thr Asn
 290 295 300
 Val Pro His Ile Phe Ala Ile Gly Asp Ile Val Gly Gln Pro Met Leu

 305 310 315 320
 Ala His Lys Gly Val His Glu Gly His Val Ala Ala Glu Val Ile Ala
 325 330 335
 Gly Lys Lys His Tyr Phe Asp Pro Lys Val Ile Pro Ser Ile Ala Tyr
 340 345 350
 Thr Glu Pro Glu Val Ala Trp Val Gly Leu Thr Glu Lys Glu Ala Lys
 355 360 365
 Glu Lys Gly Ile Ser Tyr Glu Thr Ala Thr Phe Pro Trp Ala Ala Ser

 370 375 380
 Gly Arg Ala Ile Ala Ser Asp Cys Ala Asp Gly Met Thr Lys Leu Ile
 385 390 395 400
 Phe Asp Lys Glu Ser His Arg Val Ile Gly Gly Ala Ile Val Gly Thr
 405 410 415
 Asn Gly Gly Glu Leu Leu Gly Glu Ile Gly Leu Ala Ile Glu Met Gly
 420 425 430
 Cys Asp Ala Glu Asp Ile Ala Leu Thr Ile His Ala His Pro Thr Leu

 435 440 445

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

450

455

460

Asp Leu Pro Asn Pro Lys Ala Lys Lys Lys

465

470

<210> 82

<211> 1422

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 82

atgtcgactg agatcaagac ccaagtcgta gtcctgggcg caggccctgc cggatactcc 60

gccgcgttcc gctgcgcgga cctcggctctt gaaacctga ttgtcgaacg ttacaacaca 120

ctgggcggtg tatgccttaa cgtgggctgc atcccccca aagcactgtt gcacgttgcg 180

aaggtgattg aagaggcaaa ggcgctcgcg gaacacggca tcgtcttcgg cgaacctaa 240

actgacatcg ataaaattcg aacttggaag gaaaaagtga ttaaccagct taccggagga 300

ctggcaggta tggcaaaggg ccgaaaagtc aaagtgtga atggtttggg taagttcacc 360

ggcgccaaca ccttggaagt ggaggcgaa aacggcaaaa cggtcattaa tticgacaac 420

gcaattattg cagcaggttc gcgcctatc cagctgcctt tcattccaca tgaggacccc 480

cgtatctggg atagcacaga cgactggaa ctcaaggaag tcccagaacg tctgctggtg 540

atgggtggcg gtattatcg cactgaaatg gccacagtct accacgcgtt gggtttctcag 600

atcgatgtgg ttgtacgcaa gcaccaagtg atccgcgccg cggacaaaga tatcgtaag 660

gtatttacca agcgcatctc aaagaagttc aacttgatgc tcgaaaccaa ggttaccgct 720

gttgaagcta aggaagacgg aatctacgtg accatggaag gaaagaaggc tcccgcagag 780

ccacagcgct acgatgcagt gcttgttgca attggacgcg tgcctaacgg caaaaatctc 840

gatgccggtg aagctggcgt agaggtagat gaccgtgggt ttatccgcgt tgataagcaa 900

ttgcgaacga atgtgccaca ctttttgcgt atcggtgaca ttgtgggtca accgatgctt 960

gcccacaagg gcgtccacga aggccactg gccgccgagg ttatcgccgg taagaaacac 1020

tactttgatc caaaagtat cccgtccatt gcctacactg aaccggaagt ggcatgggtg 1080

ggcttgaccg aaaaagaagc aaaagaaaag ggcattagct atgaaacagc aacctttccg 1140

tgggccgcat ccggccgcgc cattgcgagc gactgcgcgg acggcatgac taaattgatt 1200

tttgacaagg aaagccatcg cgtgatcgga ggtgcaatcg tgggtaccaa cggtagcgag 1260

ttgttgggtg agatcggcct cgctatcgaa atgggctgcg acgccgaaga tatcgctctg 1320

accatttcacg cacacccaac cctgcatgaa tctgttggcc tggtgcgga ggtcttcgag 1380

ggctccatca ccgatctccc aaatccaaaa gcaaagaaaa aa 1422

<210> 83

<211> 474

<212> PRT

<213> E. coli ATCC 9637

<400> 83

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

1 5 10 15

Ala Gly Tyr Ser Ala Ala Phe Arg Cys Ala Asp Leu Gly Leu Glu Thr

20 25 30

Val Ile Val Glu Arg Tyr Asn Thr Leu Gly Gly Val Cys Leu Asn Val

35 40 45

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

50 55 60

Glu Ala Lys Ala Leu Ala Glu His Gly Ile Val Phe Gly Glu Pro Lys

65 70 75 80

Thr Asp Ile Asp Lys Ile Arg Thr Trp Lys Glu Lys Val Ile Asn Gln

85 90 95

Leu Thr Gly Gly Leu Ala Gly Met Ala Lys Gly Arg Lys Val Lys Val

100 105 110

Val Asn Gly Leu Gly Lys Phe Thr Gly Ala Asn Thr Leu Glu Val Glu

115 120 125

Gly Glu Asn Gly Lys Thr Val Ile Asn Phe Asp Asn Ala Ile Ile Ala

130 135 140

Ala Gly Ser Arg Pro Ile Gln Leu Pro Phe Ile Pro His Glu Asp Pro

145 150 155 160

Arg Ile Trp Asp Ser Thr Asp Ala Leu Glu Leu Lys Glu Val Pro Glu

165 170 175

Arg Leu Leu Val Met Gly Gly Gly Ile Ile Gly Leu Glu Met Gly Thr

180 185 190

Val Tyr His Ala Leu Gly Ser Gln Ile Asp Val Val Glu Met Phe Asp
 195 200 205
 Gln Val Ile Pro Ala Ala Asp Lys Asp Ile Val Lys Val Phe Thr Lys
 210 215 220
 Arg Ile Ser Lys Lys Phe Asn Leu Met Leu Glu Thr Lys Val Thr Ala
 225 230 235 240

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

Val Pro His Ile Phe Ala Ile Gly Asp Ile Val Gly Gln Pro Met Leu
 305 310 315 320
 Ala His Lys Gly Val His Glu Gly His Val Ala Ala Glu Val Ile Ala
 325 330 335
 Gly Lys Lys His Tyr Phe Asp Pro Lys Val Ile Pro Ser Ile Ala Tyr
 340 345 350
 Thr Glu Pro Glu Val Ala Trp Val Gly Leu Thr Glu Lys Glu Ala Lys
 355 360 365

Glu Lys Gly Ile Ser Tyr Glu Thr Ala Thr Phe Pro Trp Ala Ala Ser
 370 375 380
 Gly Arg Ala Ile Ala Ser Asp Cys Ala Asp Gly Met Thr Lys Leu Ile
 385 390 395 400
 Phe Asp Lys Glu Ser His Arg Val Ile Gly Gly Ala Ile Val Gly Thr
 405 410 415
 Asn Gly Gly Glu Leu Leu Gly Glu Ile Gly Leu Ala Ile Glu Met Gly
 420 425 430

Cys Asp Ala Glu Asp Ile Ala Leu Thr Ile His Ala His Pro Thr Leu

435	440	445	
His Glu Ser Val Gly Leu Ala Ala Glu Val Phe Glu Gly Ser Ile Thr			
450	455	460	
Asp Leu Pro Asn Pro Lys Ala Lys Lys Lys			
465	470		
<210> 84			
<211> 1422			
<212> DNA			
<213> Artificial			
<220><223> mutated and/or codon-optimized sequence			
<400> 84			
atgtcaactg agattaaaac tcaggtggtc gttctgggcg ccggaccagc tggctattca	60		
gcagcgtttc gctgcgcaga tctgggcctg gagacagtaa tcgttgagcg ttataacacc	120		
cttggcggtg tctgcctgaa cgctcgatgc atcccttcta aggccctgct gcacgtggca	180		
aaagtcatcg aagaagcgaa ggctcttgcc gaacacggca tcgtgtttgg agaaccaaag	240		
actgacatcg acaagatccg cacatggaaa gagaaggtga tcaaccaact caccgttggt	300		
ttggccggtg tggcaaaggg tcgcaaagtt aaggtagtca atggccttgg caaatttacc	360		
ggcgcaaaca ctcttgaagt ggaaggtgag aacggaaaga ctgtgatcaa cttcgacaac	420		
gcaatcatcg cagccggctc gcgtcctatt caactgccgt tcatcccaca cgaagatccg	480		
cgatatctggg atagcaccca tgcgctggag ctgaaagaag taccggaacg attgctcgtc	540		
atgggtggag gtatcatcgg cttggagatg ggaaccgtat accacgcact tggctcccag	600		
atcgacgttg tcgaaatgtt cgatcaggatg attccagcag cagataagga tatcgtgaag	660		
gtgttcacca agcgcatctt aaagaagttc aatctgatgt tggaaaccaa ggttacggca	720		
gtggaagcaa aggaggatgg tatctactgt actatggagg gtaaaaaagc accagcggaa	780		
ccccagcgat acgacgcagt tctcgtcgtc atcgacgcg tcccaaatgg taagaatctg	840		
gacgctggca aagcaggcgt cgaagtggat gaccgtggct tcatccgctg agataaacag	900		
ttgcgcacaa atgttcaca catctttgca attggagaca tcgtaggcca gcccatgctg	960		
gcccaaaaag gtgtccatga aggtcacgtg gcggtgagg ttatcgagg caagaagcac	1020		
tacttcgatc caaaggtgat cccagcatt gcctatacgg agccagaggt agcttgggtt	1080		
ggtttgaccg aaaaagaggg aaaggagaag ggtatctcgt atgaaacagc aacctttcca	1140		
tgggcccgtg ccggacgtgc tatcgcaagc gactgacgag atggcatgac caagctgatc	1200		
ttcgacaaag aatcccaccg cgtgatcggg ggtgctattg ttggcactaa cggcggcgaa	1260		

ctcttgggcg aaatcggtt ggcaattgaa atgggtgctg atgctgagga catcgctttg 1320

accatccacg cacacccgac cctgcatgag tctgtgggac tggcagccga agtggttcgaa 1380

ggctcaatca ccgacctccc taatcccaag gctaagaaga aa 1422

<210> 85

<211> 299

<212> PRT

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 85

Met Asp Phe Pro Gln Gln Leu Glu Ala Cys Val Lys Gln Ala Asn Gln

1 5 10 15

Ala Leu Ser Arg Phe Ile Ala Pro Leu Pro Phe Gln Asn Thr Pro Val

20 25 30

Val Glu Thr Met Gln Tyr Gly Ala Leu Leu Gly Gly Lys Arg Leu Arg

35 40 45

Pro Phe Leu Val Tyr Ala Thr Gly His Met Phe Gly Val Ser Thr Asn

50 55 60

Thr Leu Asp Ala Pro Ala Ala Ala Val Glu Cys Ile His Ala Tyr Phe

65 70 75 80

Leu Ile His Asp Asp Leu Pro Ala Met Asp Asp Asp Asp Leu Arg Arg

85 90 95

Gly Leu Pro Thr Cys His Val Lys Phe Gly Glu Ala Asn Ala Ile Leu

100 105 110

Ala Gly Asp Ala Leu Gln Thr Leu Ala Phe Ser Ile Leu Ser Asp Ala

115 120 125

Asp Met Pro Glu Val Ser Asp Arg Asp Arg Ile Ser Met Ile Ser Glu

130 135 140

Leu Ala Ser Ala Ser Gly Ile Ala Gly Met Cys Gly Gly Gln Ala Leu

145 150 155 160

Asp Leu Asp Ala Glu Gly Lys His Val Pro Leu Asp Ala Leu Glu Arg

165 170 175

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

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

Ala Leu Ala Asp Tyr Ile Ile Gln Arg Asn Lys
 290 295

<210> 86

<211> 897

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 86

atggacttcc cgcaacagct tgaagcgtgt gtgaaacagg ccaaccaagc cctttcacgg	60
ttcattgccc cgcttcggtt ccaaaatacg ccggtttag agacgatgca atacggcgcc	120
ctgcttggtg gtaaacggtt acgccctttc ctgctgtacg caacaggaca tatgtttggt	180
gtcagcaciaa atacactgga cgctccagcg gcggctgtgg aatgtatcca tgcatacttt	240
ttaattcacg atgacttgcc tgcaatggac gatgacgact taagacgtgg gctcccgacg	300
tgccacgtaa aatttggcga agctaacgcc attctggcgg gagacgcgct tcaaacgctc	360
gcctttagca tcctgtccga tgctgatatg ccggaggat cggatcgca tgcattagt	420
atgatctcag aattggccag cgcttcaggg attgccggca tgtgtggcgg gcaggcggtta	480
gacttggacg cggagggcaa acacgtccct ttagatgcgc ttgagagaat acatagacat	540
aagacagggg cgcttatcag agcagctgta agacttggcg cgctgtctgc gggcgataaa	600

ggaagacggg ccctgccgtt tttagataag tatgcagaat cgatcggact ggcgttccaa 660

gtgcaggacg acatccttga tgcgttgggt gataccgca cacttgggaa aagacagggc 720

gcggatcaac agcttggcaa atccacctat ccggttttgc tgggctgga gcaagcgca 780

aaaaagcta gagacttaat tgatgatgcc cgccagagct taaagcaact cgcggagcaa 840

tcgttgata catcagcact ggaagctctg gctgactata ttatccaacg taacaag 897

<210> 87

<211> 897

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 87

atggatttcc cacaacaatt ggaggcctgc gttaagcagg cgaaccaagc tctttcgca 60

ttcatcgcg cactcccggt tcaaaacacg ccagtggctg aaactatgca atacggcgct 120

ttgcttgggt gaaagcggtt gcgccccttc cttgtttacg caaccggtca catgtttggt 180

gtctctacta acacacttga tgcccccgcg gcggccgtgg agtgtatcca cgcttacttc 240

ctcattcacg atgacctccc tgcgatggac gatgatgacc tgcgtcgtgg cctgcccacc 300

tgtaacgtga aatttgggtga agctaagcc atcctggctg gcgacgact gcagaccctt 360

gccttctcca ttctgagcga cgcagacatg cctgaagtct ccgaccgtga tcgcatctcg 420

atgatttccg agctcgcgtc cgccagcggc attgccggta tgtgcggtgg tcaggccctt 480

gacctggatg ctgaaggcaa acacgtgccg cttgatgctc tcgaacgcat ccaccgacac 540

aagaccggcg cgctgattcg tgctgccgtg cgctcggcg ctttgtccgc aggtgacaaa 600

ggacgccgcg ctctcccagt attggataaa tatgccgaat ccattgggtct cgccttccaa 660

gtgcaagacg atatcctcga cgtggtgggc gataccgca cccttgaaa gcgtcaaggt 720

gcggaccagc agctgggaaa gtccacttat ccgcactcc tcggtttgga gcaggcacgc 780

aagaaggccc gcgacctgat tgatgacgcg cgtcaatcac tgaacaact cgcagagcaa 840

tccttgata catctgcctt ggaggccctg gcagactaca tcatccaacg taacaag 897

<210> 88

<211> 897

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 88

atggactttc cacaacagct ggaagcatgc gtgaagcagg ctaaccaggc tctatcaaga	60
ttcatcgcgc ccttaccttt ccaaaatacg cctgttgttg aaactatgca gtatggggca	120
ttattgggtg gtaaaagggt aaggccgttt ttagtgtatg ctacggggca catgttcggt	180
gtcagtacaa atacacttga cgcacctgca gcgctgttg aatgtattca tgcttatttt	240
ttaatccatg atgatttacc tgcaatggat gatgatgatt tgcgtagagg cctaccaaca	300
tgtcatgtga aatttgggtg agcaaatgct attctagctg gggacgcttt gcaaacttta	360

gcgttcagca tcttgtccga tgcagatatg cctgaagtta gcgatagaga tagaattagt	420
atgatctcag aattagcatc cgcttctgga atcgctggca tgtgtggtgg ccaagcgctg	480
gatttggatg ctgaaggaaa acacgtaccg ttggatgcct tggaaagaat tcataggcat	540
aaaacgggtg cattgattcg tgctgctgtg cgtcttggtg ctctgtctgc gggagacaag	600
ggtagaaggg cggtgccagt ttggataaaa tatgctgaaa gcattgggtt agcgtttcaa	660
gttcaagatg acatcttggg tgtggttggg gacacagcca ctttaggtaa aagacaagga	720
gcagaccagc aattaggcaa atctacttac cctgcatgtg tgggtttaga gcaggctcgt	780

aagaaggcta gagacttgat agacgatgct agacagtcac tgaagcaatt ggccgaacag	840
agccttgata cgagcgccct tgaagcctta gctgattaca ttatccaaag aaataaa	897

<210> 89

<211> 897

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 89

atggactttc ctacgaact ggagccctgc gtgaagcagg ctaaccaggc actgagccgg	60
tttatcgccc cccttcctt tcaaaacact cctgtcgtcg agacgatgca atacgtgct	120
cttctcggtg gtaagcgact ccgacccttc ctgctttacg ctacaggcca tatgtttgga	180
gtctccacca ataccctcg cgccccagct gccgccgtcg aatgcatcca tgcttacttc	240

ctgatccacg acgacctgcc cgctatggac gacgatgacc tgcgacgagg actccccact	300
tgtcatgtca aattcggcga ggctaacgcc attctggctg gagacgcct ccagactctc	360
gctttctcta tcctctccga cgctgacatg cctgaagtct ctgatcgaga ccggatttcc	420
atgatttccg agctggcctc ggccctccgc attgccgga tgtgtggtgg tcaggctctc	480
gacctggacg ctgagggcaa gcatgtgccc ctgcagcccc ttgagcggat tcatcggcac	540
aagacgggag ctctcatccg agctgccgtt cggctcggag ctctgtctgc cggagacaag	600

ggtcgacggg ctctccccgt ccttgataag tacgtgagt ccattggcct ggcctttcaa 660

gttcaggatg atattctcga cgctcgtcgt gatacgcca cccttggcaa gcgacaaggt 720

gctgaccagc agctgggcaa gtcgacctac cctgccctgc tcggactcga acaggcccga 780

aagaaggcac gtgatctgat cgacgacgca cgacagagcc tgaacaact ggccgagcag 840

tctcttgaca cctccgccct ggaggtcttc gccgactaca ttattcagcg aaacaag 897

<210> 90

<211> 1056

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 90

atggcatctg agaaggagat cagaagagaa cgctttttaa acgtgtttcc gaaactggta 60

gaagagttaa atgcgagcct gcttgcgtac ggcatgccaa aggaagcatg tgattggtat 120

gcacacagct taaattataa tacaccggga ggtaagctta acagagggtt ttctgtagtg 180

gacacgtatg ctatattaag taataaaaca gtggaacaac ttggccaaga agaatatgaa 240

aaggtggcca tcttgggggt gtgcattgag cttttacaag catattggct tgtggcggat 300

gatatgatgg acaagtcaat tacacgtcgt ggtcagccgt gctggtataa ggtaccggaa 360

gttggggaaa tagctatctg ggacgcattt atgcttgaag ctgctattta taaacttctt 420

aaaagccatt tccgtaatga aaagtattac attgacatta ccgaactttt tcatgaggta 480

acatttcaaa ccgaactggg gcagcttatg gacttgatta cagcacggga ggacaaagtt 540

gatctgtcga aattctcact taaaaaacat tcttttattg taacgttcaa gactgcgtat 600

tattcatttt atctgccgtg cgcgtttgca atgtatgtag ccggcattac agatgaaaag 660

gacttaaaac aggctcggga tgttctgatt ccattggggg aatatttcca aattcaggat 720

gactatctgg actgttttgg taccgccgaa cagattggta aaatcggcac agatatccaa 780

gataataaat gtccttgggt tatcaataaa gcactggaac tggctagcgc tgaacaaaga 840

aagacttttg atgaaaacta cggcaagaaa gatagcgtcg ctgaagctaa gtgcaaaaaa 900

atttttaacg acttaaaaaa cgagcaactg taccacgaat atgaagaaag tattgctaaa 960

gatctcaaag ctaaaataag ccaggtaggt gaatctcgcg ggttcaaagc agatgtattg 1020

acagcgtttc ttaataaagt gtataaacga tccaaa 1056

<210> 91

<211> 1056

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 91

atggctagcg aaaaggagat ccgaagagag agattcctca acgtgttccc taaacttgtg	60
gaggagctca acgcttcgct cctcgtttac ggatgcccga aggaggcctg tgattggtat	120
gcacactccc ttaactacaa cagccttggg ggtaagctca accgaggtct ttctgttgtc	180
gatacctatg caattctttc caacaagacc gtcgagcaac tgggtcagga ggagtacgag	240
aaagtggcaa ttcttggctg gtgcattgaa ctctccaag cttactggct ggtggctgat	300
gacatgatgg acaagagcat caccgacga ggtcagccct gctggtaaa ggtgccagag	360
gttggagaga tcgccatttg ggacgcattt atgctggagg ctgccattta caagtcctt	420
aagtctcact ttcgaaacga gaagtactac atcgacatta ctgagctctt tcacgaagtg	480
acttttcaga ctgagctggg tcagcttatg gatctcatta ccgctcccga ggataaagtg	540
gatctgtcga agttttccct caagaagcat tcgtttattg tcaccttcaa gacagcatac	600
tactcttttt acctgccctg cgccctcgca atgtacgtcg caggaattac cgatgagaaa	660
gacctcaagc aggcacgtga cgtctctatc cccctcggcg agtactttca gattcaggat	720
gactatctcg actgcttcgg taccacagag caaatcggaa agattggaac tgacatccag	780
gacaataagt gctcttgggt tattaacaag gcccttgagc tggcttcgc cgagcaacga	840
aagacccttg acgagaatta cggcaagaag gatagcgtgg ccgaggccaa gtgcaagaag	900
atcttcaacg acctcaagat cgagcagctc taccacgagt acgaggagtc gattgctaag	960
gaccttaagg ctaagattag ccaggtggac gagtcccggg gattcaaggc tgacgtcctg	1020
accgcctttc tcaacaaggt ctacaagcga tccaag	1056

<210> 92

<211> 603

<212> PRT

<213> Perilla setoyensis

<400> 92

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

35 40 45
 Leu Gln Leu Asp Val Lys Pro Val Glu Glu Gly Arg Arg Ser Gly Asn

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

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

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

 245 250 255
 Arg Ile Gln Arg Leu Glu Ala Arg Trp Phe Leu Asp Ala Tyr Ala Thr
 260 265 270
 Arg His Asp Met Asn Pro Ile Ile Leu Glu Leu Ala Lys Leu Asp Phe
 275 280 285

Asn Ile Ile Gln Ala Thr His Gln Glu Glu Leu Lys Asp Val Ser Arg
 290 295 300
 Trp Trp Gln Asn Thr Arg Leu Ala Glu Lys Leu Pro Phe Val Arg Asp

 305 310 315 320
 Arg Leu Val Glu Ser Tyr Phe Trp Ala Ile Ala Leu Phe Glu Pro His
 325 330 335
 Gln Tyr Gly Tyr Gln Arg Arg Val Ala Ala Lys Ile Ile Thr Leu Ala
 340 345 350
 Thr Ser Ile Asp Asp Val Tyr Asp Ile Tyr Gly Thr Leu Asp Glu Leu
 355 360 365
 Gln Leu Phe Thr Asp Asn Phe Arg Arg Trp Asp Thr Glu Ser Leu Gly

 370 375 380
 Gly Leu Pro Tyr Ser Met Gln Leu Phe Tyr Met Val Ile His Asn Phe
 385 390 395 400
 Val Ser Glu Leu Ala Tyr Glu Ile Leu Lys Glu Lys Gly Phe Ile Ala
 405 410 415
 Ile Pro Tyr Leu Gln Arg Ser Trp Val Asp Leu Ala Glu Ser Phe Leu
 420 425 430
 Lys Glu Ala Asn Trp Tyr Tyr Ser Gly Tyr Thr Pro Ser Leu Glu Glu

 435 440 445
 Tyr Ile Asp Asn Gly Ser Ile Ser Ile Gly Ala Val Ala Val Leu Ser
 450 455 460
 Gln Val Tyr Phe Thr Leu Ala Asn Ser Ile Glu Lys Pro Lys Ile Glu
 465 470 475 480
 Ser Met Tyr Lys Tyr His His Ile Leu Arg Leu Ser Gly Leu Leu Val
 485 490 495
 Arg Leu His Asp Asp Leu Gly Thr Ser Leu Phe Glu Lys Lys Arg Gly

 500 505 510
 Asp Val Pro Lys Ala Val Glu Ile Cys Met Lys Glu Arg Asn Asp Thr
 515 520 525
 Glu Glu Glu Ala Glu Glu His Val Lys Tyr Leu Ile Arg Glu Ala Trp

530	535	540	
Lys Glu Met Asn Thr Ala Thr Ala Ala Ala Gly Cys Pro Phe Met Asp			
545	550	555	560
Glu Leu Asn Val Ala Ala Ala Asn Leu Gly Arg Ala Ala Gln Phe Val			
	565	570	575
Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Lys Ile His Gln			
	580	585	590
Gln Met Gly Gly Leu Met Phe Lys Pro Tyr Val			
	595	600	
<210> 93			
<211> 1809			
<212> DNA			
<213> Artificial			
<220><223> mutated and/or codon-optimized sequence			
<400> 93			
atgtgcagca tctctcaaaa ggtcgttatt ggattgaata aggcagcggc caataactgc	60		
ctccaaaacc ttgatcgccg aggtttttaa acacgccgtg tttcatcctc cgaagcggct	120		
tcttgctcc gagcctcaag ctgctgcag cttgatgtca agccggtgga ggaaggccga	180		
cgttccggtg actatcagcc atcaatctgg gattttaatt atgtccagtc cctcaacacc	240		
ccctacaag aggaacgtta cttaccgga cagcggaac tcatcgttca ggtaagccg	300		
cttctcgaaa aaaagatgga ggctactcag cagttggagc tcatcgatga ccttaacaac	360		
ctcggtctga gctacttttt ccaagaccgc atcaagcaga ttctctcatt catctatgat	420		
gaaaatcagt gtttccactc caatatcaat gatcaagctg aaaagcgaga cctgtacttt	480		
accgcccttg gatttcgtct ccttcgccag cacggcttta acgtgtctca ggaggtgttc	540		
gattgcttca aaaacgataa aggtccgat tttaaagcgt ccctgtctgg aaacaccaag	600		
ggcctctcc agctctacga agcctccttc ttggtccgtg agggatgaaga caccctcgaa	660		
ctggcccgtc agttcgcgac caagttctg cgccgcaagc ttgacgaaat cgacgacaac	720		
catctgctct cacgaattca ccactcactc gaaatcccgt tgcactggcg catccagcgt	780		
ctggaggcgc gctggtttct ggacgcgtac gctacagcc acgatatgaa tccattatc	840		
ttggagctcg caaaactgga ctttaatat attcaagcaa cccaccaaga agagctcaag	900		
gacgtgagcc gttggtggca gaatacgcgc ctggccgaaa agctcccgtt tgtgcgcgat	960		

cgacttgtcg agtcatactt ttgggccatc gcactctttg agccacatca atatggatac	1020
cagcgccgtg ttgtgccaa gattatcacc ctgcgacat ccatcgatga cgtgtacgac	1080
atctacggca ctctcgacga gctccaactt ttaccgata acttcggtcg ctgggacacg	1140
gaatccttgg gcggctcttc ttactcgatg cagttgttct acatgggtgat tcacaacttt	1200
gtctctgagt tggcctaiga aatcctgaaa gaaaaagggtt ttatcgccat cccctacctc	1260
caacgctcct gggttgattt ggcagaatcc tttttgaaag aggcaaaactg gtactactct	1320
ggttatacac cctccctcga agaatacatc gacaacggct ccatctcaat tggcgcagta	1380
gctgtgttgt cccaagttaa ttctactttg gcaaattcta tcgaaaagcc aaagattgaa	1440
tctatgtaca agtaccacca tattttgcgt ctgtcgggtt tgttgggtgc cctgcacgat	1500
gacttgggaa ctctcttgtt cgagaagaag cgcggcgatg ttcctaaggc ggtcgaaatt	1560
tgcataaag aacgcaacga tacggaggag gaagcagaag aacacgttaa gtacttgatt	1620
cgcgaagcct ggaaggagat gaacactgct acggcagctg caggatgtcc ttttatggat	1680
gagctcaacg tggctgctgc caatctgggt cgcgcagccc agttcgtgta cctcgacgga	1740
gacggctcatg gcgtgcagca ctccaaaatc catcaacaga tgggtggcct tatgtttaaa	1800
ccgtatgtg	1809
<210> 94	
<211> 1809	
<212> DNA	
<213> Artificial	
<220><223> mutated and/or codon-optimized sequence	
<400> 94	
atgtgtctta tttcccaaaa agtagttatt ggccttaaca aggcggctgc taacaactgc	60
ttacaaaatc ttgacagacg tggtttcaaa actcgtagag tttcatcttc cgaggccgct	120
agctgcctgc gggcctctag ctctttgcag ctggacgtta aaccagtgga agaggggcgc	180
agatcgggga actatcagcc atcgatctgg gactttaatt acgtacaatc actgaatact	240
ccttacaag aagaacggta tctgacacgg cagcagaac tgattgtcca agtgaacca	300
ctgctggaaa agaagatgga agccaccaa caactggaac tcatcgacga cctgaacaac	360
ttgggactgt cctacttttt tcaggacaga attaaacaaa ttctttcatt tatatacgac	420
gaaaatcaat gctttcattc taatatcaac gatcaagccg aaaaacggga cctttacttc	480
acggctttag gttttcgtct gcttagacag catggcttca atgtgagtca agaagtattc	540
gactgcttta aaaacgataa gggatcagac tttaaagctt cattatcagg taacacaaaa	600
ggcttgcttc aattgtatga agcgagtttt ttagtccgag aagggaaga tactcttgaa	660

ttagcccgac agttcgcgac taaattcctg cgtagaaaac tggatgaaat cgacgataac 720

 catcttttaa gccgcattca tcattccttg gaaattccgt tacattggcg cattcagcgg 780
 ttggaagcac gctggttctt tgacgcttat gctacgcgtc atgacatgaa tccaatcatc 840
 ctggaactcg cgaaacttga ctttaataatt attcaagcta cacatcaaga agaacttaaa 900
 gacgtatcta gatgggtggca gaacacacgt ctggccgaga aacttccgtt tgtccgggac 960
 agattagtgg aatcttattt ctgggctatt gctctttttg aacctcatca atatgggtac 1020
 caacgtcgag ttgctgctaa gattattaca ttggcaacaa gcattgatga tgtttacgat 1080
 atttacggga ctttggatga attacaactt ttacggata attttcggag atgggataca 1140

 gagtccctgg gtgggcttcc gtattctaig cagctgtttt atatggttat ccataatttc 1200
 gtgagcgaac ttgcgtatga aattctgaaa gaaaaaggct ttatagccat tccttatctt 1260
 cagagaagtt gggttgacct ggccgagtcc tttctcaaag aggcgaaactg gtactattcc 1320
 ggctatacc ctagccttga agaatacatc gataacgggt ctatcagcat aggcgagtc 1380
 gcagtgtttt cacaggttta ttttacgttg gctaacagta ttgaaaaacc taagatagag 1440
 agcatgtata aatacacca tatcttgccg ttatctgggt tgttagtccg cttacatgat 1500
 gaccttgaa caagtctgtt cgagaagaag cggggcgatg taccgaaggc ggttgagatt 1560

 tgcataaag aacggaatga cacagaagaa gaagcagaag aacatgttaa atacctgatc 1620
 cgcaagctt ggaaagaaat gaacacagct acagcagctg cgggtgttcc gtttatggat 1680
 gaattgaatg tcgcggcgcc aaatctgggt cgggcagccc agtttgtata ttggatggg 1740
 gacggccatg gcgtccagca tagtaaaatt catcaacaga tgggaggctt gatgttcaaa 1800
 ccttacgtc 1809
 <210> 95
 <211> 1809
 <212> DNA
 <213> Artificial
 <220><223> mutated and/or codon-optimized sequence
 <400> 95
 atgtgttcaa tatcacaaaa agtcgtaatt gggttgaaca aggccgtgc caataactgt 60

 ttgcagaatt tggataggag gggtttcaaa acaagaaggg tcagctcttc cgaggcggca 120
 tcatgttaa gagcttcag tagcttgcaa ctatgttta aaccagtcga ggagggtaga 180
 aggtcaggga actaccaacc aagtatctgg gatttcaatt atgtccaatc attgaatact 240
 ccctataagg aggagcgtta cttactaga catgcggaac tgattgtcca agtgaacct 300

ttgttagaaa agaagatgga agctactcaa cagctggaat taattgacga tctaaataat 360
 ttaggtttgt cttatttttt ccaagatagg atcaagcaaa tattgtcatt tatctatgac 420
 gagaatcaat gtttccactc taacattaac gaccaggcag aaaagaggga cctgtatttt 480

 accgcattgg gttttaggct tctacgtcag catggtttta atgtttctca agaagttttt 540
 gattgcttta aaaatgacaa aggtagcgat tttaggcta gcttatctgg caacaccaaa 600
 ggtttgttgc aactatacga agcttctttt ctggtgagag aagggaaga tactctagaa 660
 ctagctagac aattcgctac caaatttctg aggagaaaat tggatgaaat tgatgataac 720
 catctattaa gcagaatata ccattctcta gagattccat tgcattggag aatccaacgt 780
 ttggaagcca gggtgttctt ggatgcttat gcaacacgtc atgatatgaa cccaattatt 840
 ttggagcttg ccaaacttga ttttaacata attcaggcta cccatcaaga agaacttaag 900

 gatgtttcca gatggtggca gaacactcgt ttggctgaaa agcttccatt tgtaagagat 960
 agactagtcg aatcatactt ttgggcaatc gccttatttg aaccacatca gtacggttac 1020
 caacgtagag tcgccgctaa aataataact ttggccactt ctatagatga tgtttatgat 1080
 atttacggga cattagatga attgcaactt ttcaccgaca atttcagaag atgggatacc 1140
 gaaagcctag gcggtctacc ttatagtatg caattgttct atatggttat tcataatttt 1200
 gtttcagaat tggtttacga aatactgaaa gaaaaagggt tcattgccat tccttattta 1260
 caaaggagtt gggtggactt agcagaaaagc tttctgaaag aggccaattg gtattattca 1320

 ggttatactc cgtctctgga ggagtatatt gataacggca gcatctctat tgggtgcagtg 1380
 gcagtattga gccaagtcta ctttacctta gctaacagca ttgaaaaacc aaaaattgag 1440
 tccatgtata aataccatca catcttaaga ctttctggat tgttagtacg ttacacgac 1500
 gatctgggca ccagtttgtt cgaaaaaaaa aggggcgacg tccctaaggc agtagaaatt 1560
 tgtatgaaag aaaggaacga tacagaagaa gaagctgaag agcatgtgaa atatctgatt 1620
 agagaggcgt ggaaggaaat gaacacagct acagctgctg cagggtgtcc attcatggat 1680
 gaattaaacg tagctgcagc caatctgggt cgtgccgctc agttcgttta tttagacggt 1740

 gatggtcattg gggttcagca ttccaagata caccaacaaa tgggggggttt aatgtttaag 1800
 ccttatgta 1809
 <210> 96
 <211> 1809
 <212> DNA
 <213> Artificial
 <220><223> mutated and/or codon-optimized sequence

<400> 96

atgtgttcca tctcgcagaa agtttgtatc ggcctgaaca aggcctgccg taacaactgt	60
ctccagaacc tggatcgacg aggtttcaag acaagacggg tgtcttcttc cgaggctgct	120
tcctgcctgc gagcatctag ctcccttcag cttgacgtca agcctgtcga ggagggtcga	180
cggctctggca actaccaacc ctccatttgg gacttcaatt acgtgcaatc gctgaacact	240
ccttacaagg aggagcgata cctcacccga cagcccgagc tcattgttca ggtcaagccc	300
ctccttgaga agaagatgga ggcaaccag cagctggagc tgattgatga cctgaacaac	360
ctgggcctct cctacttctt tcaggaccgg attaagcaaa ttctgtcctt tatttacgac	420
gagaaccaat gttttattc taacatcaac gaccaggctg aaaagcgaga tctttacttc	480
accgccctcg gattccgtct cctccgacag cacggcttta acgtttcgca agaggctttt	540
gattgcttta agaatgataa aggtccgac ttcaaggctt ctctttctgg taacaccaag	600
ggtctgctgc aactgtacga agcttccctt cttgttagag agggcgagga taccctggag	660
ctggccccgc agtttgctac caagttcctg cgacgaaagc tggatgagat tgacgacaac	720
caccttctgt cccgaatcca ccactcgctt gagatccccc tccactggcg aatccagcga	780
ctggaagcta gatggtttct ggatgcctac gctactcgac acgacatgaa tcctattatc	840
ctggagctgg ccaagcttga cttcaacatc atccaggcca cccaccagga ggaactgaaa	900
gacgtctctc gatgggtgga gaacacccga ctggccgaaa agctccctt tgtgcgagac	960
cgactggtcg agagctactt ttgggctatt gccctctttg agccccacca gtacggttac	1020
cagcgaagag tcgctgctaa aatcatcacc ctggccacta gcatcgacga cgtttacgac	1080
atttacggaa cctcgaatga gctgcaactg ttaccgata actttcggcg ttgggacaca	1140
gaaagccttg gaggtctgcc ctactctatg cagctcttct atatggtgat ccacaacttt	1200
gtgtctgagc tggcttatga gatccttaag gaaaagggtc tcacgcgat cccttacctt	1260
cagcgatcct gggtcgatct tgccgagtct ttctcaagg aggccaaactg gtattactct	1320
ggctacaccc catctctgga ggagtatat gataacggtt ctatttctat cggcgccgtg	1380
gctgtccttt cccaagtcta ttacacactg gctaattcca tcgagaagcc caagattgaa	1440
tctatgtaca aataccatca catcctgcga ctgagcggtc ttctggtgcg gcttcatgac	1500
gatcttggca cttctctgtt cgagaagaaa cgaggagatg tgcctaaggc tgtcgagatt	1560
tgtatgaaag agcgtaacga caccgaggag gaggccgagg agcacgtgaa gtacctgatt	1620
cgagaggcat ggaagaaaat gaacaccgct actgccgccg ctggttgccc ctttatggat	1680
gagctgaacg ttgcagccgc caaccttggc cgagcagccc agttcgtcta cctggacggc	1740
gacggccacg gcgtccagca ttctaaaatt caccagcaga tgggtggcct gatgttcaag	1800

ccctatgtc

1809

<210> 97

<211> 591

<212> PRT

<213> Vitis vinifera (Grape)

<400> 97

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

1 5 10 15

Val Ala Pro Ser Leu Gln Val Ser Ser Ser Pro Cys Ser Trp Arg Thr

20 25 30

Arg Pro Ser Pro Cys Thr Ser Cys His Leu Ser Pro Ser Ser Ser Ser

35 40 45

Lys Pro Leu Leu Gly Ser His Asp Tyr Ser Leu Leu Lys Ser Leu Thr

50 55 60

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

65 70 75 80

Met Lys Glu Val Lys Glu Arg Thr Trp Glu Ala Phe Tyr Arg Ala Trp

85 90 95

Asp Ser Arg Ala Ala Met Glu Met Val Asp Thr Val Glu Arg Leu Gly

100 105 110

Leu Ser Tyr His Phe Glu Asp Glu Ile Asn Ala Leu Leu Gln Arg Phe

115 120 125

Cys Asp Trp Asn Ala Ser Glu Asp Leu Phe Thr Thr Ala Leu Arg Phe

130 135 140

Arg Leu Leu Arg Gln Asn Gly Phe Pro Thr His Ser Asp Val Phe Gly

145 150 155 160

Lys Phe Met Asp Lys Asn Gly Lys Phe Lys Glu Ser Leu Thr Glu Asp

165 170 175

Ile Trp Gly Met Leu Ser Leu His Glu Ala Ser His Leu Gly Ala Lys

180 185 190

Asn Glu Glu Val Leu Ala Glu Ala Lys Glu Phe Thr Arg Ile His Leu

195	200	205	
Ile Gln Ser Met Pro His Met Glu Pro His Phe Ser Ser His Val Gly			
210	215	220	
Arg Ala Leu Glu Leu Pro Arg His Leu Arg Met Val Arg Leu Glu Ala			
225	230	235	240
Arg Asn Tyr Ile Gly Glu Tyr Ser Arg Glu Ser Asn Pro Asn Leu Ala			
245	250	255	
Phe Leu Glu Leu Ala Lys Leu Asp Phe Asp Met Val Gln Ser Leu His			
260	265	270	
Gln Lys Glu Leu Ala Glu Ile Val Arg Trp Trp Lys Gln Leu Gly Leu			
275	280	285	
Val Asp Lys Leu Asp Phe Ala Arg Asp Arg Pro Met Glu Cys Phe Leu			
290	295	300	
Trp Thr Val Gly Ile Phe Pro Asp Pro Arg His Ser Ser Cys Arg Ile			
305	310	315	320
Glu Leu Thr Lys Ala Ile Ala Ile Leu Leu Val Ile Asp Asp Ile Tyr			
325	330	335	
Asp Ser Tyr Gly Ser Leu Asp Glu Leu Ala Leu Phe Thr Asp Ala Val			
340	345	350	
Lys Arg Trp Asp Leu Gly Ala Met Asp Gln Leu Pro Glu Tyr Met Lys			
355	360	365	
Ile Cys Tyr Met Ala Leu Tyr Asn Thr Thr Asn Asp Ile Ala Tyr Arg			
370	375	380	
Ile Leu Lys Glu His Gly Trp Ser Val Ile Glu Asp Leu Lys Arg Thr			
385	390	395	400
Trp Met Asp Ile Phe Gly Ala Phe Leu Ala Glu Ala Tyr Cys Phe Lys			
405	410	415	
Gly Gly His Val Pro Ser Leu Glu Glu Tyr Leu Thr Asn Ala Val Thr			
420	425	430	
Thr Gly Gly Thr Tyr Met Ala Leu Val His Ala Phe Phe Leu Met Gly			
435	440	445	

Gln Gly Val Thr Arg Glu Asn Met Ala Met Leu Lys Pro Tyr Pro Asn
 450 455 460
 Ile Phe Ser Cys Ser Gly Lys Ile Leu Arg Leu Trp Asp Asp Leu Gly
 465 470 475 480
 Thr Ala Arg Glu Glu Gln Glu Arg Gly Asp Asn Ala Ser Ser Ile Glu
 485 490 495
 Cys Tyr Lys Arg Glu Arg Glu Met Asp Thr Val Leu Glu Asp Glu Ala
 500 505 510

Cys Arg Lys His Ile Arg Gln Met Ile Gln Ser Leu Trp Val Glu Leu
 515 520 525
 Asn Gly Glu Leu Val Ala Ser Ser Ala Leu Pro Leu Ser Ile Ile Lys
 530 535 540
 Ala Ala Phe Asn Leu Ser Arg Thr Ala Gln Val Ile Tyr Gln His Gly
 545 550 555 560
 Asp Asp Asn Lys Thr Ser Ser Val Glu Asp His Val Gln Ala Leu Leu
 565 570 575

Phe Arg Pro Val Ser Ser Asn Gly His Ala Gln Ile Thr Met His
 580 585 590

<210> 98

<211> 1773

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 98

atgtctcgct tcgttacaat gccatctcat gtattaccta gttcattcgt tgcgcctagc	60
ctgcaggtgt cgctagccc gtgttcttgg agaactcggc cgtccccttg cacttcatgc	120
catttatcac cgagctcttc atcaaacct ctcttaggat ctcacgatta ttcacttttg	180
aaatccttaa cgttatctcc acacgccgta aactcggaag cagattcacc cacacggcgc	240
atgaaagagg tgaagaaaag aacgtgggaa gccttttacc gagcatggga tagtagagct	300
gccatggaga tggtaggacac agtcgagcgc cttggacttt cctatcactt tgaagatgaa	360
attaatgcgc tgcttcaaag attttgcgac tggaacgcgt cagaggattt gtttacaacc	420
gccttaagat ttagactgct tcggcagaat gggtttccta ctactccga tgtgtttgga	480

aagtttatgg acaagaacgg aaaatttaag gagagtctga cggaagatat ttggggatatg 540
ctttcgcttc atgaagcttc acatcttggg gcaaaaaacg aggaagtcct tgctgaagca 600
aaagagttta ctgcaatata tcttattcag tctatgccgc atatggagcc tcattttctcg 660

tcacacgtag gtcgggcact tgaactccct cgtcatctcc ggatgggtgcg tctcgaagct 720
cgtaattata tcggggaata ttacagtga agcaaccgga acctcgcat ccttgaattg 780
gccaaattgg atttcgatat ggtacaatca ttgcatcaaa aagagctggc ggaaattgtt 840
cgatgggtga aacaattagg attagtcgac aaattagatt ttgctcgcga tcggccgatg 900
gagtgttttc tctggaccgt tggatatctc ccggtatctc ggcatccag ctgtcgtatt 960
gaattgacga aagctattgc gattctgttg gtcattgacg atatttatga ttcatatgga 1020
agcttggatg aactggcttt atttactgat gctgttaaaa gatgggatct ggggtcgatg 1080

gatcaacttc cggagtatat gaaaatctgc tacatggctc tgtacaacac aacaaacgat 1140
atagcatata gaattatata agagcacggc tggagcgtaa ttgaagattt aaaacgcaca 1200
tggatggaca tctttggggc gtttttggca gaagcact gtttcaaagg tggacatgtt 1260
ccgagccttg aggaatactt aacaaacgca gtcacgactg gcggcaccta tatggccctt 1320
gttcacgcgt tctttcttat gggacaaggc gtgactcgtg aaaatatggc catgcttaaa 1380
ccgtatccta atatcttctc atgtagcggg aaaattcttc gactttggga tgacttgggt 1440
acagctcgag aagaacaaga gcgcggggat aatgcgagct ccattgaatg ctacaaaaga 1500

gaaagagaaa tggatacggg actggaagat gaagcatgta gaaaacatat cagacagatg 1560
atccaatcac tgtgggtgga acttaacggc gagctggtgg ccagcagcgc tctgccgttg 1620
tcaattataa aagccgcgtt taacctttcc cgtacagctc aagtatctta tcagcacggg 1680
gatgataaca aaacatcatc agttgaagac catgtccaag cctttttgtt tcgtccggt 1740
tcttcaaattg gtcattgctc aatcacgatg cat 1773

<210> 99
<211> 1773
<212> DNA
<213> Artificial
<220><223> mutated and/or codon-optimized sequence
<400> 99
atgtcccgtt ttgtcactat gccatcgac gtcttgccca gcagcttcgt agcgccatct 60

ctgcaagtgt cttcatcccc atgctcgtgg cgcaccggcc cttccccatg cacctcttgt 120
cattgtcac catcctcttc ttcaaaacca cttttgggca gccacgatta ctccctgctt 180

aagagcctga ccccttctcc acacgccgtg aattccgaag cggattcttc caccgcgcga 240
atgaaggaag tcaaggagcg cacatgggaa gctttctatc gagcctggga ttcacgtgct 300
gctatggaga tggttgacac cgtggaacgt ctgggtcttt cataccactt tgaggacgaa 360
attaatgcgc tcttgcagcg ctttttgat tggaaacgtt ccgaagatct gttcacaacc 420
gccttgcgtt tccgccttct gcgccagaac ggttttccta cccattctga cgtctttggc 480

aaattcatgg ataagaacgg aaaattcaag gaaagcttga ctgaggatat ctggggaatg 540
ctgtcattgc atgaggcatc gcacttgggc gcaaaaaacg aggaggctct cgcggaagca 600
aaggagtcca cccgcatcca tctgatccaa tcaatgccgc atatggaacc tcaattctcc 660
tcccacgtcg gacgcgtctt ggaactccct cgccatcttc gcatgggtgc actcgaagct 720
cgcaattata tcggcgagta ctcccgcgag agcaaccgga atcttgcatc cttggaactc 780
gcgaagttag attttgacat ggtgcagtca ctccaccaa aggaactcgc ggagatcgtt 840
cgctggtgga aacagctggg tctctgggac aagttaggatt tcgcgcgcga ccgacctatg 900

gaatgcttcc tctggaccgt gggcatcttt cctgatcttc gacattcttc ttgtcgcac 960
gaattgacta aagcaatcgc catctctctg gttatcgacg atatttatga ctcatatggc 1020
agcctcgacg aattggcact gttcacagac gcagtcaaac gctgggattt gggagcgatg 1080
gatcagttgc ccgaatatat gaagatctgc tatatggcac tttacaacac cacaaacgat 1140
atcgcgtacc gcatccttaa agaacatggt tggtccttta tcgaagactt gaagcgcacc 1200
tggatggaca tcttcggtcg gttcctcgcg gaggcatact gtttcaaagg cggtcacgta 1260
ccctctctcg aggaatacct caccaatgca gtcaccactg gaggcactta tatggccctg 1320

gtccacgctt ttttctcat gggtcagggc gtgaccgcg aaaatatggc catgcttaag 1380
ccatacccaa acattttctc atgttccgga aaaattctgc gcctttggga cgatcttggt 1440
accgctcgcg aggagcaaga acgcgagac aacgcttctt ccatcgagtg ctataaacga 1500
gagcgcgaaa tggacaccgt gctcgaagac gaagcatgtc gcaagcatat ccgcaaatg 1560
atccagtcac tttgggtgga gcttaacggt gagctcgtgg catcttcgac tcttccactg 1620
tcattatta agcgcggtt caacctgtct cgcacggcac aagtcacta ccaacacgga 1680
gacgataaca agacttcgtc tgtggaggac catgtacagg ccttgctgtt tcgtcccgtt 1740

tcttcaaacg gccacgtca aatcacatg cac 1773

<210> 100

<211> 1773

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 100

atgagtagat ttgttactat gccctcacat gtccctacat cttcttttgt agcccccttc	60
ttgcaagtca gctcaagtcc gtgtagttgg agaactaggc cgagcccatg tacttcatgt	120
catttaagtc cgtcctcaag cagcaaacca ttgctgggat cccatgacta ttctttgcta	180
aaatccttga cgttgtctcc tcacgcagtt aactcagagg cagattcadc tacaaggaga	240
atgaaagaag tgaagaaaag aacttgggaa gcgttctaca gagcatggga ttcaagggca	300
gcaatggaaa tggtcgatac ggtggaaaga ttaggcttgt cataccattt tgaagacgaa	360
ataaatgctt tgcttcaaag gttctgtgac tggaatgcat ctgaggatct ttttacgacc	420
gcactaagat ttagactgtt gagacagaac ggcttcccaa cgcatagcga tgtgtttggt	480
aagttcatgg ataagaatgg gaaatttaag gagtccttaa ccgaagacat ctggggaatg	540
ttatccttgc atgaagcttc ccaccttggg gccaaaaacg aggaagtctt ggccgaagct	600
aaagaattca ccagaattca cctaatacaa tccatgcccc atatggaacc acatttctcc	660
tctcatgtgg ggagagcctt agagttgccg agacacttaa ggatggttag actagaagct	720
aggaactata tcggggaata ttcaagagaa agtaatccga atttggcgtt tttggagctt	780
gctaagttgg attttgatat ggttcaatcc ctgcaccaa aagaactagc cgaaatagtt	840
aggtggtgga agcaattagg tctggttgat aaactggact ttgctagaga tagaccaatg	900
gagtgccttc ttggacggt aggtatcttc cctgaccaa ggcacagctc ttgtagaatt	960
gaacttacta aggcaatcgc catcctattg gtaattgatg atatctacga tagttatggt	1020
agtctggacg aattggcctt atttactgac gcagtaaaga gatgggatct aggtgctatg	1080
gatcagctgc ctgagtatat gaaaatctgt tacatggccc ttataacac cactaatgat	1140
atcgcatata gaatcttgaa ggagcacggt tggtcagtga ttgaagattt aaaaaggact	1200
tggatggata tttttggagc ttttcttgct gaagcttatt gctttaaagg tggcatgtc	1260
ccctcattag aagaatacct gaccaatgcc gtgactactg gtggtaccta catggcttta	1320
gtgcatgcct tttttcttat gggtaaggc gttactcgtg aaaacatggc aatgctgaaa	1380
ccatatccca atatattcag ctgctcaggt aagattttac gtttatggga cgatcttggt	1440
actgcaagag aagaacaaga aaggggggac aatgcatcca gcattgaatg ctataaaaga	1500
gaacgtgaaa tggacacagt gttagaagat gaagcttgca gaaaacatat aagacaaatg	1560
attcaatctt tatgggttga actaaacggg gagctagtgg cgtcaagtgc gcttccctta	1620
agtatcatca aggcagcttt taacctaaagt agaaccgcac aagttattta ccagcacggc	1680
gatgataaca aaacctctc agttgaagac cagttcaag ctctattatt ccgtccagtt	1740

tcaagtaacg gacatgccca aattacaatg cat 1773

<210> 101

<211> 1773

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 101

atgtcgcgat tcgttaccat gccctccac gtgcttccat cgtccttcgt tgctccatct 60

ctccaggttt cgtcgtctcc ctgttcttgg agaactcgac cttccccctg tacttcttgt 120

catctgtccc ccagctcgtc gtccaaacct ctgctgggta gccacgatta cagcctcctg 180

aagtctctca ccttttctcc tcacgtgtt aactccgagg ccgattcttc caccgcacgg 240

atgaaagagg tgaaggagcg aacttgggaa gccttctacc gagcttggga ctcccgggcc 300

gcaatggaga tggtcgacac tgctgaaaga ctgggtctgt cttaccactt tgaggatgag 360

atcaacgctc tcttccagag attctgtgat tggaacgcct cggaggacct gttcacgacc 420

gtcttccgat tccgactcct gagacagaac ggcttcccca cacattcgga cgtcttcggc 480

aagttcatgg acaagaacgg aaagtccaag gagtcgctca cggaagacat ctggggaatg 540

ctttctctcc atgaggcctc gcacctgggc gctaagaatg aggaggtcct cgccgaggcc 600

aaggagttaa cccaatcca cttatccag agcatgcctc acatggagcc tcacttttcc 660

tcccacgtcg gtcgagccct tgagctgccc cgacatcttc gaatggtccg actggaggct 720

cgtaactaca ttggcgagta ctctagagag tccaaccca acctggcctt cctcgaactc 780

gccaagctgg atttcgacat ggtccagtcg ctgcaccaga aggagctcgc cgagattgtc 840

cgatggtgga agcaactcgg ccttgtcgac aaactggatt tcgctcgaga tcgacccatg 900

gagtgcttcc tgtggaccgt tggatatttc cccgaccccc gtcactcttc ctgccgaatc 960

gagctgacca aggccatcgc cattctgctc gttattgatg acatttacga ctcttacgga 1020

agccttgacg agctcgcaat cttaccgac gccgtgaagc gatgggatct gggtgccatg 1080

gatcaactgc ccgagtacat gaagatctgc tacatggccc tttacaacac aaccaacgac 1140

atcgcttacc gaatccttaa ggagcacgga tggctctgta ttgaggacct caagcgaacg 1200

tggatggata tcttcggagc ctttctggct gaggcctact gtttcaaggg aggccatgtc 1260

ccttcgctcg aagagtatct gaccaacgct gtcaccaccg gtggtactta catggctctg 1320

gtccacgcat tcttctgat gggtaagga gtcacaagag agaatatggc tatgcttaag 1380

ccctatccca acatttttctc ttgctctgga aagattctcc ggctctggga tgacctcggc 1440

actgctcgag aggagcagga gcgaggcgat aacgcttcct cgattgagtg ctacaagcga 1500
gagcgagaga tggataccgt ccttgaggac gaggcttgcc gtaagcatat tcgtcaaagt 1560
attcagtcgc tctgggttga gctgaacgga gaactgggtg catcgccgc cctccctctg 1620
tcgatcatca aagccgcttt caacctctcg agaactgctc aggtgattta ccagcacgga 1680
gatgacaaca agaccagctc tgtcgaagac cacgtccagg ctctgctttt cagacctgtg 1740
tcctcgaacg gtcacgctca aattacatg cac 1773

<210> 102

<211> 501

<212> PRT

<213> Phaseolus angularis (Azuki bean) (Vigna angularis)

<400> 102

Met Glu Thr Asn Ser Thr Leu Phe Phe Thr Phe Thr Val Leu Leu Ile

1 5 10 15

Thr Ile Ile Thr Phe Leu Lys Ala Leu Lys Ser Leu Phe Thr Pro Asn

20 25 30

Lys Ser Lys Ser Asn Leu Pro Pro Gly Pro Lys Gly Leu Pro Leu Val

35 40 45

Gly Asn Leu Leu Gln Leu Gly Ala Lys Pro His Gln Thr Leu Ala Thr

50 55 60

Leu Ala Asn Ile His Gly Ser Ile Met Ser Leu Lys Leu Gly Gln Glu

65 70 75 80

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

85 90 95

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

100 105 110

Gly Ser Ser His Asp His Phe Ser Leu Pro Phe Met Pro Val Ser Gln

115 120 125

Gln Trp Arg Glu Leu Arg Lys Leu Cys Asn Glu Leu Leu Phe Ser Asn

130 135 140

Lys Asn Leu Asp Ala Thr Gln Gly Leu Arg Ser Lys Lys Val Arg Glu

145 150 155 160

Leu Tyr Ser Asp Ile His Arg Ser Ser Leu Lys Gly Glu Pro Val Asn

165 170 175
Ile Gly Arg Leu Ala Phe Lys Thr Thr Ile Asn Gln Leu Ser Asn Thr

180 185 190
Ile Tyr Ser Glu Asp Phe Leu Gln Ser Ala Glu Lys Ala Gly Glu Met

195 200 205
Lys Glu Leu Val Thr Asn Ile Met Lys Glu Val Gly Arg Pro Asn Leu

210 215 220
Ala Asp Cys Phe Pro Val Leu Lys Met Ile Asp Pro His Gly Ile Arg

225 230 235 240
Arg Arg Thr Gly Ser Tyr Phe Ser Lys Leu Leu Asn Ile Phe Lys Ser

245 250 255
Leu Ile His Lys Arg Leu Glu Leu Arg Lys Asp Ala Ala Gly Tyr Cys

260 265 270
Thr Lys Lys Asp Met Leu Asp Ala Met Leu Asn Asp Ala Gln His Lys

275 280 285
Met Asp Ile Val Lys Ile Gln Arg Leu Ser Leu Asp Leu Phe Val Ala

290 295 300
Gly Thr Asp Thr Val Thr Ser Thr Val Glu Trp Ala Met Ala Glu Leu

305 310 315 320
Leu His Asn Pro His Val Met Ser Lys Ala Lys Glu Glu Leu Glu Arg

325 330 335
Ile Ile Gly Lys Asp Asn Leu Val Glu Glu Ser Asp Ile Ala Lys Leu

340 345 350
Pro Tyr Leu Gln Ala Ile Val Lys Glu Thr Phe Arg Leu His Pro Ala

355 360 365
Val Pro Leu Leu Leu Pro Arg Lys Ala Glu Val Glu Phe Glu Met His

370 375 380
Gly Tyr Thr Ile Pro Lys Gly Ala Gln Val Leu Ile Asn Val Trp Ala

385 390 395 400
Ile Gly Arg Asp Pro Asn Leu Trp Glu Lys Pro Arg Leu Phe Trp Pro

405 410 415

Glu Arg Phe Leu Glu Ser Glu Ile Asp Phe Lys Gly Arg Ser Phe Glu

420

425

430

Leu Thr Pro Phe Gly Gly Gly Arg Arg Ile Cys Pro Gly Leu Pro Leu

435

440

445

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

450

455

460

Asp Trp Glu Leu Gln Asp Ile Gln Pro Glu Asp Met Asn Met Asp Glu

465

470

475

480

Asn Phe Gly Leu Thr Leu Glu Lys Ala Gln Pro Val Leu Ala Ile Pro

485

490

495

Ile Ile Pro Lys His

500

<210

> 103

<211> 1503

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 103

atggagacta actctaccct ttcttttaca ttcaccgtcc tctgatcac aatcattacg 60

ttctgaagg cgttgaaatc ttgtttacc ctaacaagt caaatctaa cctcccaccc 120

ggccaagg gccttccttt ggttggcaac ctgcttcagt tgggcgcgaa gccgcaccag 180

acctcgcaa cctcgcaaa catccacgga tccatcatgt cctcaaatt gggtaagag 240

acaaccatcg tgatgtcttc cgctgaggcc gcaaaggag ttctgcagat tcatgatcat 300

ttctctcaa accgcaaat ccccgacgca atgcgcggct cctcccacga ccacttctcc 360

cttcttttca tgctgtttc ccaacagtgg cgagaattgc gaaagctttg caacgaactc 420

ctgtttcca acaagaacct ggacgtaca caaggcctgc gtcctcaaaa ggttcgagaa 480

ctgtacagc gatattcaccg cagctcgctg aaggcgagc cgtgaacat cggtcgtctg 540

gcttttaaaa cgaccatcaa ccaactttct aacaccatct actcggaaga cttcttgcaa 600

tccgccgaga aagcgggcga aatgaaggaa ctcgttacca atattatgaa ggaggttggc 660

cgcccgaacc ttgccgactg cttccccgtt ctgaaaatga tcgaccccca cggcatccgc 720

cgtcgaaccg gctcttactt cagcaagctg cttaatatct tcaaatccct gatccacaag 780

cgctctggaac ttcgtaaaga tgctgcgggc tactgtacta aaaaggatat gcttgacgcg 840
atgctgaacg acgccaaca caaaatggac atcgtaaga tccagcgtct ctctcttgat 900
ttgttcgttg caggaacgga taccgtcacc tccactgtcg agtgggctat ggctgaactg 960
cttcacaacc cccacgttat gtcaaaggct aaagaggaac tcgaacgcat tatcggttaag 1020
gataatctgg tggaggagtc agacatcgcc aagttgcctt acttcaggc tatcgtgaag 1080
gagaccttcc gtctccacc agcagtcacc cttttgtctg cacgcaaagc cgaggttgaa 1140
tttgaatgc acggttacac aatcccaaaa ggcgcccagg tccgatcaa tgtctgggcc 1200

attggctcgcg accgaactt gtgggaaaag ccacgcctgt tttggcccga gcgatttctt 1260
gaatccgaga tcgacttcaa gggccgatcg ttcgagctca ccccgttcgg cggcggccgt 1320
cgcatctgcc ctggactccc attggcgatc cgacttgttt ttctgatgct cggcctgttt 1380
attaactctt tcgattggga acttcaggac attcagccgg aagacatgaa tatggacgaa 1440
aacttcggtc tgactcttga aaaggcccag ccagttctcg caattccaat cattccaaaa 1500
cac 1503

<210> 104

<211> 1503

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 104

atggaaacca acagtacttt attctttaca ttacagtcc ttcttatcac tatcataaca 60
tttctcaagg cgctcaaaa tttattcacg ccgaataaat caaaatcaa tcttcgcct 120
ggacaaaaag gactcccgct ggttggaaat cttttgcaac tcggggcaaa acctcatcag 180
acattagcta cgcttgctaa cattcacgga agcattatgt ctctgaagtt aggtcaagaa 240
acaacaattg tcatgtcttc ggcggaggcg gctaaaggcg ttttcagat ccacgatcat 300
tttctgagca accggaaaaa tcttgacgcc atgcgtggta gtcacatga tcatttctca 360

ttaccgttca tgccagtaag ccaacaatgg cgcgagctga gaaaactgtg taatgagtta 420
ttattttcaa ataaaaactt ggatgcaacg caggggctga gaagcaaaaa agttcgtgaa 480
ctgtattcag atatacatcg ctctctcttt aaaggcgaac cggtaacat cgggagactt 540
gcttttaaaa caacaattaa ccaattgagc aatacatct actcggagga cttccttcaa 600
agcgcggaaa aagcaggta aatgaaggaa ctggtcacaa acatcatgaa ggaagtggga 660
cgtccgaatc tggtgactg ttttcgggtg ctgaaaatga ttgatcctca cggaattcgc 720

cgccggacag gttcatatTT ctcaaaactt ctgaatatat tcaaaagctt aattcataaa	780
cgcttagagt tgcgcaaaga tgctgcaggc tattgcacga agaaagatat gctggatgcg	840
atgctgaacg atgctcaaca taaaatggat atcgtgaaaa tccagcggt ctcactcgac	900
ttatttgttg ctggcacaga taccgtcacc tcgaccgtcg agtgggcat ggcagagttg	960
ctgcataatc cacacgtaat gtccaaagct aaagaagaac tggaagaat cattggcaag	1020
gataacctcg ttgaagaaag cgacatcgcg aaacttcctt acttacaggc gattgtaaag	1080
gagacgttta gactgcatcc agcagtacct ctgcttcttc cgcgcaaagc tgaagtagaa	1140
tttgaatgc atggatatac gatcccgaaa ggggcacagg tactcatcaa tgtgtggcg	1200
attggcagag atccaaatTT gtgggaaaaa cctcgtttgt tttggcctga aagattcctg	1260
gagtcagaaa ttgacttcaa gggcagatcc ttcgaactca cgccgtttgg cgggggccgg	1320
cgtatttgcc ctggcctgcc gttagccatt agattggttt tcttgatgtt aggcttattc	1380
atcaatagct tcgattggga gctccaggat attcaaccgg aagacatgaa tatggacgaa	1440
aactttggct taacactgga gaaagcacag ccggtactgg ccattccgat catccgaaa	1500
cat	1503
<210> 105	
<211> 1503	
<212> DNA	
<213> Artificial	
<220><223> mutated and/or codon-optimized sequence	
<400> 105	
atggaaacta acagcacttt attttttaca ttcacagtct tacttattac gataattacg	60
ttctgaaag ctctgaaatc actatttact ccaataaaaa gtaagtctaa ctgccccct	120
ggcccaaagg gtttgccact tgttggaaat ctattgcaat taggtgcaaa acctcatcaa	180
acactggcta ctttggttaa tatacatggg tcattatga gcctaaagt ggccaagag	240
acaactatag tgatgtcttc tgctgaagca gccaagggtg tcttgcaat acatgatcat	300
ttctgagca atagaaagat tctgacgt atgaggggat caagccacga tcatttctca	360
ttgccattca tgctgtatc ccagcaatgg agagaattaa gaaaattatg taatgaactg	420
ttattttcta acaagaatct ggatgcaaca cagggttaa gatcaaagaa agttaggga	480
ttatattcag atatccatcg tagttcttta aaaggcgaac cagtcaacat cggcagactg	540
gcatttaaaa ccaccattaa ccagttgtct aataccatTT atagcgaaga ttctttacaa	600
tctgccgaaa aagctggcga gatgaaagaa ttggttacta atatcatgaa ggaagttggt	660

aggccgaacc ttgcagactg ctttcctgtc ttgaagatga tagatccaca tggcattaga 720
 agaagaacgg gtctttatct ttctaagttg cttaatatct tcaaatcatt gatccacaag 780

aggttggagt tgagaaaaga tgcgcgagga tattgtacga aaaaagacat gctggatgca 840
 atgcttaacg atgctcagca taagatggat atagttaaga tccaaagatt atctttggat 900
 ttgttcgtcg ccggtactga tactgttacc tccactgtag aatgggccat ggcagaattg 960
 ctccataacc cgcatgtcat gtcaaaagct aaagaagaat tggaaagaat tattggcaag 1020
 gataatctgg tagaggaaag tgacatagct aaacttccat acttgcaagc cattgttaag 1080
 gagactttta gattgcatcc agcagtgcct cttcttttgc caagaaaggc tgaagtggaa 1140
 tttgaaatgc acggctacac aattccaaaa ggagctcaag tgttgatcaa tgtatgggcc 1200

attggctcgtg acccaaatct atgggaaaaa ccacgtttat tctggccaga gagatttttg 1260
 gaatcagaaa ttgacttcaa gggcaggtca ttcgagttga cccattcgg cgggtgtaga 1320
 aggatatgtc ctggcttacc acttgccatc cgtttagtat ttttgatgct tgggctattc 1380
 attaactctt ttgactggga gcttcaagat attcagcccg aggatatgaa catggatgag 1440
 aatttcgggc tgacacttga aaaagcgcaa cctgtattag ctattccgat aattccaaaa 1500
 cat 1503

<210> 106
 <211> 1503
 <212> DNA
 <213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 106

atggagacca actcgactct gtcttttact ttactgttc tctgatcac cattatcacc 60
 ttctcaagg cactgaagtc gctgttcacc ccaataagt cgaagagcaa cctcccccc 120
 ggtcctaagg gcctccctct ggtgggaaac ctctccagc tcggcgcaa acctcatcag 180
 actctggcta cctggccaa cattcacggc agcatcatgt cgctcaagct tggtcaggag 240
 actactatcg tcatgtcgtc ggccgaggct gctaaggcg tccttcagat ccacgatcac 300
 ttctgtcta accgaaagat ccccgacgt atgagaggat cttctcacga ccatttctct 360

ctccctttca tgccagtctc gcagcagtgg cgagagcttc gaaagctctg taacgagctt 420
 ctcttctcta ataagaacct tgatgtact cagggtctca gatctaagaa ggttcgagag 480
 ctgtattcgg acattcatcg atcttctctg aagggtgagc ccgtcaacat tggtagactg 540
 gctttcaaga caacaattaa tcagctctct aacacaattt acagcgagga cttctccag 600

tcggctgaga aggccgggtga aatgaaggaa ctctcacca acatcatgaa ggaagtgggc 660
cggcccaacc ttgtgattg cttccccgtt ctcaagatga ttgaccccca tggaatccgg 720
cgacggaccg gtcctatatt ttccaagctc ctgaacattt tcaagtcct catccacaag 780

cgactcgaac tccgaaaaga cgccgccggt tactgtacca aaaaagacat gcttgacgct 840
atgtcaacg acgcacagca caagatggat attgtcaaga ttcagcgact gtcctcgac 900
ctgttcgtgg ctggaaccga caccgttaca tctacagtcg aatgggctat ggccgaactg 960
ctgcataacc ctacgtgat gagcaaggct aaggaggagc ttgagcgaat catcgaaag 1020
gacaacctcg tggaggagtc ggacatcgcc aagctgcctt accttcaagc cattgttaag 1080
gagacctttc gactgcaccc agctgttcca ctgctgctcc caagaaaagc tgaggctcgag 1140
ttcgaatgc acggatatac aatccccaag ggagctcagg ttcttatcaa cgtctgggct 1200

attggacgag accccaacct ctgggagaag cctcgactct tttggcctga gcgattcttc 1260
gagagcgaga ttgactttaa gggtagatct ttcgaactga ccccttttg tggcggacga 1320
cggatctgcc ccggtctgcc cctcgtatct cgactggttt tctcatgct tggcctcttc 1380
atcaatagct ttgactggga gcttcaagac attcagcctg aagacatgaa catggacgag 1440
aacttcggcc tgacctgga gaaggcccag cccgtcctgg caatcccat catccccaag 1500
cac 1503

<210> 107

<211> 495

<212> PRT

<213> Swertia mussotii (Felwort)

<400> 107

Met Asp Phe Asp Phe Leu Thr Ile Ala Ile Gly Phe Leu Phe Thr Ile
1 5 10 15
Thr Leu Tyr Gln Ala Leu Asn Phe Phe Ser Arg Lys Ser Lys Asn Leu
20 25 30
Pro Pro Gly Pro Ser Pro Leu Pro Leu Ile Gly Asn Leu His Leu Leu
35 40 45
Gly Asp Gln Pro His Lys Ser Leu Ala Lys Leu Ala Lys Lys His Gly
50 55 60

Pro Ile Met Gly Leu Gln Leu Gly Gln Val Thr Thr Ile Val Val Thr
65 70 75 80

Ser Ser Gly Met Ala Lys Glu Val Leu Gln Lys Gln Asp Leu Ala Phe
85 90 95

Ser Ser Arg Ser Ile Pro Asn Ala Ile His Ala His Asp Gln Tyr Lys
100 105 110

Tyr Ser Val Ile Trp Leu Pro Val Ala Ser Arg Trp Arg Gly Leu Arg
115 120 125

Lys Ala Leu Asn Ser Asn Met Phe Ser Gly Asn Arg Leu Asp Ala Asn
130 135 140

Gln His Leu Arg Ser Arg Lys Val Gln Glu Leu Ile Ala Tyr Cys Arg
145 150 155 160

Lys Ser Ser Gln Thr Gly Asp Ala Ile Asp Val Gly Arg Ala Ala Phe
165 170 175

Arg Thr Ser Leu Asn Leu Leu Ser Asn Thr Met Phe Ser Lys Asp Leu
180 185 190

Thr Asp Pro Tyr Ser Asp Ser Ala Lys Glu Phe Lys Asp Leu Val Trp
195 200 205

Asn Val Met Val Glu Ala Gly Lys Pro Asn Leu Val Asp Tyr Phe Pro
210 215 220

Leu Leu Asp Lys Val Asp Pro Gln Gly Ile Arg Lys Arg Met Thr Ile
225 230 235 240

His Phe Gly Lys Ile Leu Glu Leu Phe Gly Gly Leu Ile Asp Glu Arg
245 250 255

Leu Gln Gln Lys Lys Ala Lys Gly Val Asn Asp Asp Val Leu Asp Val
260 265 270

Leu Leu Thr Thr Ser Glu Glu Ser Pro Glu Glu Ile Asp Arg Thr His
275 280 285

Ile Gln Arg Met Cys Leu Asp Leu Phe Val Ala Gly Thr Asp Thr Thr
290 295 300

Ser Ser Thr Leu Glu Trp Ala Met Ser Glu Met Leu Lys Asn Pro Glu
305 310 315 320

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

325 330 335
 Lys Ala Val Glu Glu Ala Asp Leu Ala Arg Leu Pro Tyr Leu Arg Cys
 340 345 350
 Ala Ile Lys Glu Thr Leu Arg Ile His Pro Pro Val Pro Leu Leu Ile
 355 360 365
 Pro Arg Arg Thr Glu Gln Glu Val Glu Val Cys Gly Tyr Thr Val Pro
 370 375 380

Lys Asn Ser Gln Val Leu Val Asn Val Trp Ala Ile Ser Arg Asp Asp
 385 390 395 400
 Ala Ile Trp Lys Asp Pro Leu Ser Phe Lys Pro Glu Arg Phe Leu Glu
 405 410 415
 Ser Glu Leu Glu Met Arg Gly Lys Asp Phe Glu Leu Ile Pro Phe Gly
 420 425 430
 Ala Gly Arg Arg Ile Cys Pro Gly Leu Pro Leu Ala Val Arg Met Val
 435 440 445

Pro Val Met Leu Gly Ser Leu Leu Asn Ser Phe Asp Trp Lys Leu Glu
 450 455 460
 Gly Gly Ile Ala Pro Lys Asp Leu Asp Met Glu Glu Lys Phe Gly Ile
 465 470 475 480
 Thr Leu Gln Lys Ala His Pro Leu Arg Ala Val Ala Thr Pro Leu
 485 490 495

<210> 108

<211> 1485

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 108

atggatttcg atttccttac gattgccatt ggattttcttt ttacaattac ctgtgatcag 60
 gctctgaatt ttttcagtcg taaatcgaaa aaccttcgcg caggaccatc accgctgcct 120
 ttgataggga atcttcattt gttaggtgac cagcctcata aatcgttggc caagcttgcg 180
 aagaaacatg gtcctattat gggcttacaa ttgggccaag tgacgactat tgttgttacg 240
 tctagcggca tggccaagga agttcttcaa aagcaggatt tagctttttc atcccgttct 300

attccaaatg cgatacatgc acatgatcag tacaagtact cagtgatttg gcttcctgtg 360
gcatcacggt ggcgtaggct gcgcaaagct ttgaattcca acatgttttc cggaaacaga 420

ctggatgcga accaacaatct gcggtctcgc aaagtgcagg aattaattgc ctattgcaga 480
aaatcaagcc aaactgggta cgcaattgac gtagggcgag ccgcctttcg aacatctctt 540
aacctcctgt ctaacacgat gttttctaaa gacctcactg atccttactc cgattccgcg 600
aaagaattca aggatttggg ctggaacgtg atggttgagg cgggaaaacc taacttgggtg 660
gattattttc cgctccttga taaagtggac ccgaaggaa ttcggaagcg catgacgac 720
cattttggta aaattctcga gctttttggc ggattaattg atgaacgtct tcagcaaaaa 780
aaggcgaagg gagtgaacga cgacgttctt gacgtgcttc tgacgacctc tgaggaatcc 840

ccagaagaga ttgatcggac tcatattcag cgtatgtgcc tggatttggt tgtggccggt 900
acggacacta cctcctccac gctcgaatgg gcaatgtcgg aaatgttaaa aaatcctgag 960
aaaatgaaag cggtcgaagc agaactggcc caagtcatcg gcaaaggcaa agccgtagaa 1020
gaagccgacg ttgctcgctt gccgtattta cgggtgcgcaa ttaaagaaac cctgagaatc 1080
catccgcctg tcccgtttt aatccctcga agaacagaac aagaggtcga agtctgtgga 1140
tataccgtgc ctaaaaattc acaagtgtc gtaaactgtc gggcgattag tcgtgacgac 1200
gcgatatgga aagaccgct tagttttaaa ccggaacggt tcctcgagag cgagttagaa 1260

atgcggggta aggatttcga gctgattcct ttcggggcgg gtagaagaat ctgccctggc 1320
ctgccgtg cggtgcggat ggttccagtt atgcttggtt ctctgcttaa ttcattcgat 1380
tggaactgg aggggtggaat cgcccggaag gacctggata tggaggagaa atttggaatt 1440
actttacaga aggcgcatcc tcttcgcgca gttgccacgc ctctc 1485

<210> 109
<211> 1485
<212> DNA
<213> Artificial
<220><223> mutated and/or codon-optimized sequence
<400> 109

atggattttg attttttaac cattgcaatt ggttttttat tcaccattac gttgtatcaa 60
gcgttaaatt ttttttctag gaagtcctaaa aaccttcgcg caggctcttc tccactgcca 120

ctaactcgga acctacatct tttgggtgat cagcctcaca aatcattggc aaaattggcg 180
aagaaacacg gcccgattat ggggtctaaa ctgggtcaag taacgacat agttgtaaca 240
tctagtggca tggctaaaga agtacttcaa aagcaagatt tagctttttc ctcaagatct 300

ataccaaacg caatacatgc ccatgatcaa tacaagtaca gtgttatatg gctgccagtt 360
 gcttccagat ggcgtgggct acgtaaagcg ctaaattcta acatgttctc tggcaataga 420
 ttggacgcca accagcattt aagatctaga aaagtgaag aattaattgc atactgtaga 480
 aaatctagtc aaacagggga tgccatagat gtagggcgtg ctgcgttttag aaccagcttg 540

 aattttattgt ctaacactat gttttcaaaa gacttgaccg atccctacag cgactccgca 600
 aaggaattta aagatttagt ttggaacgtt atggttagagg ctggaaaacc gaatttggtg 660
 gattattttc ccttggttga taaggtcgac ccccgaggta ttagaaagag aatgacaatt 720
 cacttcgga agatattgga actgttttgt ggattgatag acgaacgttt acaacaaaag 780
 aaagccaaag gtgtgaatga tgatgtgctt gatgttttgc taactacaag cgaggaaagt 840
 ccggaagaaa tagataggac ccatatcaa aggatgtgtt tggacttggt tgttgccgga 900
 actgatacca ctagttcaac attagaatgg gctatgtctg agatgttgaa aaatcctgag 960

 aaaatgaaag ctgcacaggc agagtggcg caagtgatag gtaaaggtaa agctgtagag 1020
 gaagcggatt tgcttagatt accctatttg aggtgtgcta tcaaggaaac tcttaggatt 1080
 catcctcag taccactatt aattccacgt agaaccgaac aggaagtgga agtgtgtgga 1140
 tataccgtac cgaagaactc tcaagtgtg gtaaatgtgt gggccatttc tagagatgat 1200
 gctatatgga aagaccatt atctttttaa cccgagaggt ttctggaaag cgagttggaa 1260
 atgagaggca aagacttga attgatccca ttggggcgg ggagaagaat ctgtcctgga 1320
 ttgccgttgg cagtgagaat ggttccagtt atgctgggtt ctctgttgaa cagtttcgat 1380

 tggaagttag agggagggat agcccctaaa gatctggaca tggaagaaaa gtttggcatt 1440
 actcttcaaa aagcccatcc actaagggcc gtagctacac cgctg 1485

 <210> 110
 <211> 1485
 <212> DNA
 <213> Artificial
 <220><223> mutated and/or codon-optimized sequence
 <400> 110
 atggatttcg atttctgac cattgcaatc ggctttcttt ttaccattac tctttaccag 60
 gctcttaact tcttctctcg aaagtctaag aacctccctc ctggaccttc cccctgccc 120
 ctgatcgga accctcacct cctcggcgac cagccccaca agtccttggc taagcttgct 180
 aaaaagcacg gcccatcat gggctcttcag ctgggacagg ttacaactat tgttgttacc 240

 agctccgga tggttaagga ggtgtccag aaacaggacc ttgccttctc ctcgatgcc 300

atccccaacg ccattccatgc ccacgatcag tacaagtact ctgttatttg gctccctgtg 360
gcctctcggt ggagaggtct ccgaaaagct ctgaattcca acatgttcag cggtaaccgt 420
ctggacgcaa atcagcatct ccggtctaga aaggttcagg agctgattgc ttactgtcga 480
aagtcttcgc agacagggtga cgctatcgat gtgggtagag cagccttccg aacttcgctg 540
aacctcctct cgaataccat gttctctaag gatcttaccg atccatacag cgattctgcc 600
aaggagttaa aggacctggt ttggaacgtc atggtcgagg caggtaagcc taacctggtt 660

gattactttc cctgcttga taaagtggac cctcaggga ttcgaaagag aatgaccatt 720
catttttgta agatcctcga gctgttcggt ggctttatcg acgagcgact ccagcagaaa 780
aaggccaagg gtgttaacga cgacgttctc gacgtgctgc tgacgacatc tgaggagtcc 840
cccagggaat ttgaccgaac ccatatcaa agaattgtgc tggacctctt tgcgctggt 900
actgacacca cttcctccac actggagtgg gcaatgtcgg aaatgtctaa gaacctgag 960
aagatgaagg ctgcccaggc cgagcttgcc caggctcatc gcaagggtaa ggccgtggaa 1020
gaggccgacc tcgctcgact gccctacctt cgatgtgcta tcaaggagac gctgagaatc 1080

cacccccctg tgccccctct catccccctg cgaaccgagc aggaagtga ggtctgtggc 1140
tacaccgtcc caaagaactc ccaggttctc gttaacgtgt gggctatttc cagagatgac 1200
gccatctgga aggacctct gtcttttaag cctgagcgat tcctggagtc cgagctggag 1260
atgcggggaa aggactttga actgatcccc ttggagctg gacgacgaat ctgccccggt 1320
ctccccctgg ccgtccgaat ggtccccgtc atgcttggtat cgctcctcaa ctgctttgac 1380
tggaagcttg aaggcggaat tgctcccaag gaccttgata tggaagagaa gttcggtatc 1440
actctgcaga aagcacaccc cctccgggca gtcgccactc ctctg 1485

<210> 111

<211> 864

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 111

atgacagcag ataacaattc catgcctcac ggtgctgtgt cttcatacgc gaaactggtg 60
cagaacaaaa cactgagga tattctcgag gaattccccg aaatcatccc tctccagcag 120
cgcccgaaca ccgctcctc tgaaacctct aatgacgaat ccggcgagac atgtttctcc 180
ggccacgacg aagagcaaat taagctgatg aacgaaaatt gcacgtgtt ggactgggat 240
gacaacgcca ttggagccgg taccaaaaag gtgtgccatc tgatggagaa cattgaaaag 300

ggactgctcc atcgagcttt ctacagtgttc attttcaatg aacagggtga gctcctgctg 360

caacagcgcg caacggaaaa gattaccttc cccgacctct ggactaacac ctgctgctca 420

cacccattgt gtatcgatga cgaacttggc ttgaaggga agcttgatga taagattaag 480

ggcgcaatca cggccgctgt tcgtaaactc gatcacgaac ttggaattcc tgaagacgaa 540

accaagactc gcggcaaatt tcacttcttg aatcgatatcc actatatggc gccttccaac 600

gagccttggg gcgagcacga gatcgactat attttgttct acaagatcaa tgcaaaggag 660

aacctgactg tgaacccgaa tgtgaatgaa gtgcgtgact tcaagtgggt ctaccgaac 720

gatctgaaga ccatgtttgc cgatccatcc taaaaattca ccccggtgtt caaaattatc 780

tgtgaaaact atctgttcaa ttggtgggaa caactcgacg atttgtccga ggtcgagaac 840

gatcgtaaaa tccaccgtat gctt 864

<210> 112

<211> 1485

<212> DNA

<213> Artificial

<220><223> mutated and/or codon-optimized sequence

<400> 112

atggattttg atttcttgac aatcgcaatt ggtttctgt tcaccattac ctgtaccaaa 60

gcgctcaact tccttagccg caaatccaag aacctccac caggctcgag cccctcccg 120

cttattggca acctgcactt gctgggagat cagccgcaca aatcattggc gaagctcgct 180

aagaagcacg gtccaattat gggctctgcag ttgggccaag ttacaactat cgtcgtgacc 240

tcgtccggaa tggcaaagga ggttttgcaa aaacaagatt tggttttctc gtcccgtcc 300

atccccaacg ctatccacgc acacgaccag tacaagtaca gcgtgatttg getgcccgtt 360

gctagccgct ggcgcggtatt gcgcaaggca ctgaattcca acatgttctc tggcaaccgc 420

ttggacgcca accaacctc tcgatccgc aaagtgaag aactgatcgc ttactgtcgc 480

aaaagctctc agaccggcga tgccattgac gtaggcgcgc cggcatttcg aacctcgctc 540

aaccttctct cgaacacgat gttttcaaaa gaccttaccg acccatattc agattccgcc 600

aaggagtcca aggacctgt ttggaacgtg atggtcgaag ccggaagcc gaatcttgc 660

gattactttc cgttgctcga taagtggtgat cccagggaa tccgcaagcg tatgaccatc 720

catttcggaa aaatcctgga gctgttcggt ggctgatcg acgagcgct tcaacagaag 780

aaggccaagg gagttaatga cgacgtgctg gatgtgttc tcaccacctc tgaagagtec 840

cctgaggaga tcgatcgtac acacatccag cgcattgtc ttgatttgtt cgtggctgga 900

acagatacca caagctcaac cttggagtgg gcaatgtcgg agatgcttaa gaacccggaa	960
aagatgaagg ctgcacaggc ggaactggcc caggtaatcg gcaagggtaa ggcagtagaa	1020
gaggctgacc tcgcacgact cccatatctc cgttgtgcta tcaaggagac cctgcgtatt	1080
cacccgccgg tcctcttctt tattcctcgc cgcaccgagc aggaggttga agtgtgtggt	1140
tacactgtgc caaaaaactc ccaagtgtg gtcaacgttt gggcgatctc gcgcgatgac	1200
gccatctgga aagaccttt gtccttcaag ccagagcgtt ttttgaatc cgagctggag	1260
atgcgtggaa aggattttga attgatcccg ttcggtgccg gtcgtcgcac ctgcccgggt	1320
cttcccccttg cagtacgcat ggtcccgta atgctgggct cgctcctgaa ctcttttgat	1380
tggaaacttg agggcgggat cgctccaaag gacttggata tggaagaaaa gttcggaatt	1440
actttgcaaa aagcacaccc gttgcgcgca gtcgcaaccc ctctg	1485