



## (19) 대한민국특허청(KR)

## (12) 등록특허공보(B1)

(45) 공고일자 2019년05월29일

(11) 등록번호 10-1983115

(24) 등록일자 2019년05월22일

(51) 국제특허분류(Int. Cl.)

C12N 15/82 (2006.01) A01H 5/02 (2018.01)

C12N 9/04 (2006.01) C12N 9/10 (2006.01)

C12P 19/46 (2006.01)

(52) CPC특허분류

C12N 15/8201 (2013.01)

A01H 5/02 (2018.05)

(21) 출원번호 10-2017-7037475(분할)

(22) 출원일자(국제) 2012년08월07일

심사청구일자 2018년01월25일

(85) 번역문제출일자 2017년12월27일

(65) 공개번호 10-2018-0004306

(43) 공개일자 2018년01월10일

(62) 원출원 특허 10-2014-7006130

원출원일자(국제) 2012년08월07일

심사청구일자 2015년09월22일

(86) 국제출원번호 PCT/IB2012/001513

(87) 국제공개번호 WO 2013/021261

국제공개일자 2013년02월14일

(30) 우선권주장

61/521,171 2011년08월08일 미국(US)

(뒷면에 계속)

(56) 선행기술조사문헌

Scientia Horticulturae. Vol. 112, No. 2, 페이  
지 172-175 (2007.)GenBank Accession Number AY262037  
(2003.11.01.)\*

\*는 심사관에 의하여 인용된 문헌

전체 청구항 수 : 총 10 항

심사관 : 최준호

(54) 발명의 명칭 사프란 화합물의 재조합 생성을 위한 방법 및 물질

**(57) 요약**

제아잔틴 분해 디옥시게나제를 단독으로 또는 UDP-글리코실트랜스페라제(UGT)를 엔코딩하는 재조합 유전자와 조합하여 발현하도록 조작된 재조합 미생물, 식물, 및 식물 세포가 개시된다. 상기 미생물, 식물, 또는 식물 세포는 크로세틴, 크로세틴 디알데하이드, 크로신, 또는 피크로크로신과 같은 사프란으로부터의 화합물을 생성시킬 수 있다.

(52) CPC특허분류

*C12N 9/0006* (2013.01)  
*C12N 9/1048* (2013.01)  
*C12N 9/1051* (2013.01)  
*C12P 19/46* (2013.01)  
*C12Y 114/99042* (2013.01)

(72) 발명자

**손카르, 샤일렌드라**

인도 600 041 첸나이 티리반미유르 메튜 스트리트  
23/3 쥬-2 바이바브 아파트먼트

**쿠마르, 사티쉬**

인도 603 106 칸치푸람 디스트릭트 우타라메를 아  
럼빌리유르 빌리지 포스트 73

**쿠마르, 칼리안, 케이.**

인도 600 096 라지브 나가르 폐룬구디 아넥스 첸나  
이 프라샨티 아파트먼츠 플랫 넘버 디1

**판차파게사, 무랄리**

인도 600 113 첸나이 타라마니 타라마니 로드 티셀  
바이오 파크 리미티드 4 플로어 401-405 에볼바 바  
이오텍 프라이빗 리미티드

**한센, 에스벤, 할크야례르**

덴마크 1972 프레테릭스버그 3. 티브이 보르사에스  
베이 30

**한센, 클라브스, 리이쉐데**

덴마크 2300 코펜하겐 그라에肯랜드스베이 88비 예  
스티

(30) 우선권주장

61/576,460 2011년12월16일 미국(US)  
61/595,450 2012년02월06일 미국(US)

## 명세서

### 청구범위

#### 청구항 1

하이드록실- $\beta$ -사이클로시트랄 (HBC)을 아글리콘 O-글리코실 유리딘 5'-디포스포(UDP) 글리코실 트랜스페라제 (O-글리코실 UGT) 및 UDP-글루코스와 접촉시키는 단계를 포함하고, 상기 아글리콘 O-글리코실 UGT는 SEQ ID NO:12에 기재된 아미노산 서열을 갖는 UGT73-EV12인, 피크로크로신을 생산하는 방법.

#### 청구항 2

SEQ ID NO:12에 기재된 아미노산 서열을 갖는 UGT73 폴리펩티드를 엔코딩하는 분리된 핵산.

#### 청구항 3

제2항의 핵산에 작동 가능하게 연결된 조절 영역을 포함하는 핵산 작제물.

#### 청구항 4

SEQ ID NO:12에 기재된 아미노산 서열을 갖는 분리된 폴리펩티드.

#### 청구항 5

작제

#### 청구항 6

작제

#### 청구항 7

제2항에 따른 분리된 핵산을 포함하는 DNA 발현 카세트로서, 분리된 핵산은 프로모터에 작동 가능하게 연결된, DNA 발현 카세트.

#### 청구항 8

제7항의 DNA 발현 카세트를 포함하는 재조합 벡터.

#### 청구항 9

제7항에 따른 DNA 발현 카세트를 포함하는 재조합 숙주 세포.

#### 청구항 10

제9항에 있어서, 상기 세포는 효모 세포, 에스케리키아 콜리 (*Escherichia coli*) 세포, 식물 세포, 포유동물 세포 및 곤충 세포로 구성되는 군으로부터 선택된 재조합 숙주 세포.

#### 청구항 11

제10항에 있어서, 상기 효모 세포는 사카로마이세스 세레비지애 (*Saccharomyces cerevisiae*)인, 재조합 숙주 세포.

#### 청구항 12

크로세틴 디알데하이드를 알데하이드 데하이드로게나제와 접촉시켜 제10항의 세포에서 크로세틴을 생산하는 단계를 포함하는, 크로세틴을 생산하는 방법.

#### 청구항 13

작제

**청구항 14**

삭제

**청구항 15**

삭제

**청구항 16**

삭제

**청구항 17**

삭제

**청구항 18**

삭제

**청구항 19**

삭제

**청구항 20**

삭제

**청구항 21**

삭제

**청구항 22**

삭제

### **발명의 설명**

#### **기술 분야**

[0001] 본 출원은 2011년 8월 8일에 출원된 미국 가출원 번호 61/521171호, 2011년 12월 16일에 출원된 미국 가출원 번호 61/576460호, 및 2012년 2월 6일에 출원된 미국 가출원 번호 61/595450호를 우선권으로 주장한다.

#### **기술 분야**

[0003] 본 발명은 사프란 식물인 크로쿠스 사티부스(*Crocus sativus*)로부터 화합물을 재조합적으로 생성시키기 위한 방법 및 물질, 및 더욱 특히 재조합 숙주에서 사프란 식물로부터의 플라보란트(*flavorant*), 아로마탄트(*aromatant*), 및 착색제 화합물을 재조합적으로 생성시키기 위한 방법 및 물질에 관한 것이다.

### **배경 기술**

#### **배경**

[0005] 사프란은 크로쿠스 사티부스 L. 꽃의 암술머리로부터의 추출에 의해 제조된 건조 향신료이며, 이는 3500년 이상 사용된 것으로 생각된다. 이러한 향신료는 다수의 의학적 목적을 위해 역사적으로 사용되어 왔으나, 최근에는 이의 착색제 특성에 대해 주로 이용된다. 사프란의 주요 성분 중 하나인 크로세틴은 관련된 카로티노이드-타입 분자와 유사한 항산화제 특성을 가지며, 또한 착색제이다. 사프란의 주요 색소는 황적색의 색을 부여하는 글리코시드의 혼합물인 크로신이다. 크로신의 주요 성분은 색에 있어서 황색인  $\alpha$ -크로신이다. 사프라날은 건조 과정의 생성물인 것으로 생각되며, 식품 제조에 사용될 수 있는 후각자극제 특성을 또한 갖는다. 사프라날은 무색인 사프란 추출물인 피크로크로신의 쓴 부분의 아글리콘 형태이다. 따라서, 사프란 추출물은 착색제 또는 플라보란트로서 또는 이의 후각자극제 특성에 대해 많은 목적에 사용된다.

[0006]

사프란 식물은 이탈리아, 프랑스, 인도, 스페인, 그리스, 모로코, 터키, 스위스, 이스라엘, 파키스탄, 아제르바이잔, 중국, 이집트, 아랍에미리트 연합, 일본, 호주, 및 이란을 포함하는 많은 국가에서 상업적으로 재배된다. 이란은 전세계의 매년 사프란 생산량(200톤 바로 위로 추정됨)의 약 80%를 생산한다. 1 kg의 생성에 150,000개 초과의 꽃이 필요한 것으로 보고되었다. 생산량을 증가시키기 위한 식물 품종개량 노력은 열매를 맺지 않는 식물을 발생시키는 식물의 유전체의 세배수체에 의해 복잡해진다. 또한, 상기 식물은 10월 중순 또는 하순에 시작하여 약 15일 동안만 개화 상태로 존재한다. 통상적으로, 생산은 또한 비효율적인 과정인 꽃으로부터 암술머리의 수작업 제거를 포함한다. 사프란 kg 당 \$1000을 초과하는 판매 가격이 통상적이다. 매력적 대안은 사프란 성분의 생물전환 또는 새로운 생합성이다.

## 발명의 내용

개요

본 발명의 개시는 재조합 숙주에서 사프란 식물로부터의 화합물의 생성을 개선시키기 위한 방법 및 물질, 뿐만 아니라 피크로크로신, 사프라날, 크로신, 크로세틴, 또는 크로세틴 에스테르와 같은 화합물의 생성을 위한 재조합 경로를 확립하는데 유용한 누클레오티드 및 폴리펩티드의 발견을 기초로 한다. 본 발명의 개시는 또한 크로세틴 및 크로세틴 에스테르를 함유하는 조성물에 관한 것이다. 생성물은 푸드 시스템 또는 의약 보충을 위한 최적 특징을 위해 단독으로 및 조합되어 생성될 수 있다. 다른 구체예에서, 화합물은 혼합물로 생성될 수 있다. 일부 구체예에서, 숙주 균주는 재조합 효모이다. 다른 구체예에서, 본원에 기재된 누클레오티드는 식물 유전학에서 사용될 수 있고, 식물 품종개량 방법에서 마커로서 보조하기 위해 사용될 수 있다.

한 양태에서, 본 문헌은 제아잔틴 분해 디옥시제나제(ZCD)를 엔코딩하는 외인성 핵산을 포함하는 재조합 카로티노이드 생성 숙주(예를 들어, 미생물)를 특징으로 한다. 숙주는 검출가능한 양의 크로세틴 및/또는 크로세틴 디알데하이드 및/또는 하이드록실- $\beta$ -사이클로시트랄(HBC)을 생성할 수 있다. ZCD는 크로쿠스 사티부스 ZCD일 수 있다.

숙주는 제라닐제라닐 디포스페이트 신타제(GGPPS), 피토엔 신타제, 피토엔 데하이드로제나제, 및  $\beta$ -카로틴 신타제를 엔코딩하는 내인성 유전자를 포함할 수 있다.

숙주는 GGPPS, 피토엔 신타제, 피토엔 데하이드로제나제, 및  $\beta$ -카로틴 신타제를 엔코딩하는 적어도 하나의 외인성 핵산을 추가로 포함할 수 있다.

본 문헌은 또한 GGPPS, 피토엔 신타제, 피토엔 데하이드로제나제,  $\beta$ -카로틴 신타제,  $\beta$ -카로틴 하이드록실라제, 및 제아잔틴 분해 디옥시제나제(zeaxanthin cleavage dioxygenase, ZCD)(예를 들어, 크로쿠스 사티부스 ZCD)를 엔코딩하는 적어도 하나의 외인성 핵산을 포함하는 재조합 숙주를 특징으로 한다. 적어도 하나의 외인성 핵산의 발현은 숙주에서 검출가능한 양의 크로세틴 및/또한 크로세틴 디알데하이드를 생성할 수 있다.

본원에 기재된 숙주 중 임의의 숙주는 알데하이드 데하이드로제나제를 엔코딩하는 내인성 유전자 또는 알데하이드 데하이드로제나제(ALD)를 엔코딩하는 외인성 핵산을 추가로 포함할 수 있다. 알데하이드 데하이드로제나제는 사카로마이세스 세레비지애 알데하이드 데하이드로제나제(예를 들어, ALD2-ALD6 또는 HFD1)일 수 있다.

본원에 기재된 숙주 중 임의의 숙주는  $\beta$ -카로틴 하이드록실라제를 엔코딩하는 내인성 유전자 또는  $\beta$ -카로틴 하이드록실라제를 엔코딩하는 외인성 핵산을 추가로 포함할 수 있다.  $\beta$ -카로틴 하이드록실라제는 찬토필로마이세스 덴드로로우스(*Xanthophyllomyces dendrorhous*)  $\beta$ -카로틴 하이드록실라제일 수 있다.

본원에 기재된 숙주 중 임의의 숙주는 아글리콘 O-글리코실 유리딘 5'-디포스포(UDP) 글리코실 트랜스페라제(O-글리코실 UGT)를 엔코딩하는 외인성 핵산을 추가로 포함할 수 있다. 이러한 숙주는 검출가능한 양의 피크로크로신 또는 크로신을 생성할 수 있다. 아글리콘 O-글리코실 UGT는 UGT85C2, UGT73-EV12, 또는 UGT71 하이브리드 효소일 수 있다. 아글리콘 O-글리코실 UGT는 또한 크로쿠스 사티부스로부터의 Cs VrUGT2일 수 있다.

본원에 기재된 숙주 중 임의의 숙주는 O-글리코실 UGT를 엔코딩하는 외인성 핵산을 추가로 포함할 수 있다. 이러한 숙주는 검출가능한 양의 크로세틴 모노 및 디 글루코실 에스테르를 생성할 수 있다. 아글리콘 O-글리코실 UGT는 UGT76G1, 또는 UGT71 하이브리드 효소(예를 들어, 71C125571C2 및/또는 71C125571E1)일 수 있다.

본원에 기재된 숙주 중 임의의 숙주는 2개의 글루코스 모이어티 사이의  $\beta$  글루코실 결합(예를 들어,  $\beta$  1,6 결합)을 촉매하는 UGT를 엔코딩하는 외인성 핵산을 추가로 포함할 수 있다. 이러한 숙주는 검출가능한 양의 크로

세틴 젠티비오실 에스테르를 생성할 수 있다. 2개의 글루코스 모이어티 사이의  $\beta$  글루코실 결합을 촉매하는 UGT는 UGT71 하이브리드 효소, 예를 들어, 71C125571C2 또는 71C125571E1일 수 있다.

[0018] 본원에 기재된 숙주 중 임의의 숙주는 유리딘-5'-디포스포글루코스(UDP-글루코스)-크로세틴 8,8'-글루코실트랜스페라제를 엔코딩하는 외인성 핵산을 추가로 포함할 수 있다. 이러한 숙주는 검출가능한 양의 크로세틴 모노글루코시드를 생성할 수 있다. UDP-글루코스-크로세틴 8,8'-글루코실트랜스페라제는 크로쿠스 UDP-글루코스-크로세틴 8,8'-글루코실트랜스페라제일 수 있다.

[0019] 본원에 기재된 숙주 중 임의의 숙주는 2개의 글루코스 모이어티 사이의  $\beta$  글루코실 결합(예를 들어,  $\beta$  1,6 결합)을 촉매하는 UGT를 엔코딩하는 외인성 핵산을 추가로 포함할 수 있다. 이러한 숙주는 검출가능한 양의 크로신을 생성할 수 있다. 2개의 글루코스 모이어티 사이의  $\beta$  글루코실 결합을 촉매하는 UGT는 UGT76G1, UN4522, 또는 U1671일 수 있다.

[0020] 본원에 기재된 숙주 중 임의의 숙주는 미생물, 식물, 또는 식물 세포일 수 있다. 미생물은 유성 효모, 효모균과(Saccharomycete), 예를 들어, 사카로마이세스 세레비지애(*Saccharomyces cerevisiae*), 또는 에스cherichia coli일 수 있다. 식물 또는 식물 세포는 크로쿠스 사티부스일 수 있다.

[0021] 본원에 기재된 숙주 중 임의의 숙주는 데옥시자일룰로스 5-포스페이트 신타제(DXS), D-1-데옥시자일룰로스 5-포스페이트 리덕토이소머라제(DXR), 4-디포스포시티딜-2-C-메틸-D-에리트리톨 신타제(CMS), 4-디포스포시티딜-2-C-메틸-D-에리트리톨 키나제(CMK), 4-디포스포시티딜-2-C-메틸-D-에리트리톨 2,4-사이클로디포스페이트 신타제(MCS), 1-하이드록시-2-메틸-2(E)-부테닐 4-디포스페이트 신타제(HDS), 및 1-하이드록시-2-메틸-2(E)-부테닐 4-디포스페이트 환원효소(HDR) 중 하나 이상을 엔코딩하는 외인성 핵산을 추가로 포함할 수 있다.

[0022] 본원에 기재된 숙주 중 임의의 숙주는 트렁케이션된 3-하이드록시-3-메틸-글루타릴(HMG)-CoA 환원효소(tHMG), 메발로네이트 키나제(MK), 포스포메발로네이트 키나제(PMK), 및 메발로네이트 피로포스페이트 데카르복실라제(MPPD) 중 하나 이상을 엔코딩하는 외인성 핵산을 추가로 포함할 수 있다.

[0023] 또 다른 양태에서, 본 문헌은 피크로크로신을 생성하는 방법을 특징으로 한다. 상기 방법은 피크로크로신을 생성시키기 위해 HBC와 아글리콘 O-글리코실 UGT 및 UDP-글루코스를 접촉시키는 것을 포함하며, 상기 아글리콘 O-글리코실 UGT는 UGT85C2, UGT73-EV12, 또는 UGT71 하이브리드 효소로 이루어진 군으로부터 선택된다. UGT는 또한 Cs VrUGT2일 수 있다.

[0024] 또 다른 양태에서, 본 문헌은 UGT73 폴리펩티드를 엔코딩하는 분리된 핵산을 특징으로 한다. UGT73 폴리펩티드는 도 3에 기재된 UGT73 아미노산 서열과 적어도 80%의 서열 동일성을 가질 수 있다. 본 문헌은 또한 상기 핵산에 작동 가능하게 연결된 조절 영역을 포함하는 핵산 작제물 뿐만 아니라 상기 핵산 또는 핵산 작제물을 포함하는 제조합 숙주를 특징으로 한다.

[0025] 또 다른 양태에서, 본 문헌은 도 3에 도시된 UGT73 아미노산 서열과 적어도 80%의 서열 동일성을 갖는 분리된 폴리펩티드를 특징으로 한다. 폴리펩티드는 도 3에 도시된 UGT73 아미노산 서열과 적어도 90%의 서열 동일성을 가질 수 있다. 폴리펩티드는 도 3에 도시된 UGT73 아미노산 서열과 적어도 95%의 서열 동일성을 가질 수 있다. 폴리펩티드는 도 3에 도시된 UGT73 아미노산 서열을 가질 수 있다.

[0026] 또 다른 양태에서, 본 문헌은 도 9에 도시된 아미노산 서열을 갖는 분리된 폴리펩티드 및 이러한 폴리펩티드를 엔코딩하는 핵산을 특징으로 한다.

[0027] 본 문헌은 또한 크로세틴을 생성시키는 방법을 특징으로 한다. 상기 방법은 크로세틴을 생성시키기 위해 크로세틴 디알데하이드와 알데하이드 데하이드로게나제를 접촉시키는 것을 포함한다.

[0028] 본 발명의 또 다른 양태는 SEQ ID NO:57에 기재된 아미노산 서열을 엔코딩하는 SEQ ID NO:58에 기재된 합성 DNA 서열을 제공하는 것이다.

[0029] 또 다른 양태에서, 본 발명은 SEQ ID NO:66에 기재된 아미노산 서열을 엔코딩하는 SEQ ID NO:65에 기재된 합성 DNA 서열을 특징으로 한다.

[0030] 달리 정의되지 않는 한, 본원에서 사용되는 모든 기술 및 과학 용어는 본 발명이 속하는 분야의 당업자에 의해 일반적으로 이해되는 것과 동일한 의미를 갖는다. 본원에 기재된 것과 유사하거나 동등한 방법 및 물질이 본 발명을 실시하는데 사용될 수 있으나, 적합한 방법 및 물질은 하기에 기재된다. 본원에 언급된 모든 간행물, 특히 출원, 특히, 및 다른 참고문헌은 이들의 전체내용이 참조로서 포함된다. 상충되는 경우, 정의를 포함하는

본 명세서가 통제할 것이다. 또한, 상기 물질, 방법, 및 예는 단지 예시이고, 제한하고자 하는 것이 아니다. 본 발명의 다른 특징 및 장점은 하기 상세한 설명으로부터 명백해질 것이다. 본 출원의 출원인은 특허법 하에서 표준 실시에 따라 과도적 관용구 "포함하는", "-로 본질적으로 포함하는" 또는 "-로 구성되는"을 이용하여 임의의 개시된 발명을 대안적으로 주장할 권리를 보유한다.

#### [0031] 상세한 설명

[0032] 다양한 크로세틴 에스테르는 사프란 추출물의 착색제 특성을 담당한다. 크로세틴은 어느 한쪽 말단에 2개의 카르복실산 기를 갖는 C18 백본으로부터 형성된 디터트렌이다. 크로세틴은  $\beta$ -카로틴 및 제아잔틴을 함유하는 카로티노이드 경로로부터 유래된다(도 2 참조). 사프란의 주요 색소는 2개의 젠티오비오스 모이어티(디젠티오비오시드)를 갖는 크로세틴 디에스테르인 크로신이다. 크로신은 크로세틴의 에스테르의 우세한 형태이다. 크로세틴의 다른 글루코시드 형태( $\alpha$ -크로세틴 또는 크로세틴-1으로도 언급됨)는 젠티오비오시드, 글루코시드, 젠티오글루코시드, 및 디글루코시드를 포함한다. 모노- 또는 디-메틸에스테르 형태에서의  $\gamma$ -크로세틴은 또한 13-시스-크로세틴, 및 트랜스 크로세틴 아이소머와 함께 사프란 내에 존재한다.

[0033] 무색인 피크로크로신은 사프란의 쓴 맛을 담당한다. 이는 HBC를 통해 제아잔틴으로부터 생성된 모노터트렌 알데하이드이다. 피크로크로신의 데글루코실레이션(deglycosylation)은 사프란 향신료의 주요 방향 성분인 사프라날(4-하이드록시-2,4,4-트리메틸 1-사이클로헥센-1-카르복스알데하이드, 또는 페하이드로- $\beta$ -사이클로시트랄)을 발생시킨다.

[0034] 사프란 추출물은 또한 왁스 및 지방, 단백질, 정유(essential oil), 안토시아닌, 플라보노이드, 비타민(리보플라빈 및 티아민), 아미노산, 전분, 미네랄, 겸을 함유한다. 모노터트렌 알데하이드 및 이소포론-관련 화합물은 사프라날과 함께 사프란의 휘발성 성분이다.

[0035] 본 문헌은 크로세틴, 크로세틴 디알데하이드, 피크로크로신, 크로신, 또는 사프라날과 같은 사프란으로부터의 화합물의 생합성에 유용한 폴리펩티드를 발현하는 재조합 숙주, 예를 들어, 식물 세포, 식물, 또는 미생물이 개발될 수 있다는 발견을 기초로 한다. 상기 숙주는 제아잔틴 분해 디옥시게나제(ZCD)(제아잔틴 분해 옥시게나제(ZCO)(예를 들어, 크로쿠스 사티부스로부터 유래됨)로도 언급됨), 및 일부 구체예에서, 하나 이상의 유리딘 5'-디포스포(UDP) 글리코실 트랜스페라제를 발현시킬 수 있다. 다양한 미생물 본체에서의 상기 생합성 폴리펩티드의 발현은 에너지 및 탄소원, 예를 들어, 당, 글리세롤, CO<sub>2</sub>, H<sub>2</sub>, 및 일광으로부터 일관되게 재현가능한 방식으로 크로세틴, 크로세틴 디알데하이드, 피크로크로신, 크로신, 또는 사프라날과 같은 사프란으로부터의 화합물이 생성되는 것을 가능케 한다. 재조합 숙주에 의해 생성되는 각각의 화합물의 비율은 숙주로 미리 선택된 생합성 효소를 혼입시키고, 이들을 적절한 수준으로 발현시킴으로써 조정될 수 있다.

[0036] 유전자 중 적어도 하나는 재조합 유전자이며, 특정 재조합 유전자(들)은 사용을 위해 선택된 종 또는 균주에 좌우된다. 화합물 수율을 증가시키고/시키거나, 에너지 및 탄소원이 사프란 화합물로 전환되는 효율을 개선시키고/시키거나, 세포 배양물 또는 식물로부터의 생산성을 향상시키기 위해 추가 유전자 또는 생합성 모듈이 포함될 수 있다. 상기 추가 생합성 모듈은 터르페노이드 전구체, 이소펜테닐 디포스페이트 및 디메틸알릴 디포스페이트의 합성과 관련된 유전자를 포함한다. 추가 생합성 모듈은 터르펜 신타제 및 터르펜 사이클라제 유전자, 예를 들어, 제라닐제라닐 디포스페이트 신타제를 엔코딩하는 유전자, 및 카로테이노이드 합성과 관련된 효소를 엔코딩하는 유전자를 포함하며; 이를 유전자는 내인성 유전자 또는 재조합 유전자(예를 들어, 외인성 핵산)일 수 있다.

#### [0037] 글루코스로부터의 IPP

[0038] 일부 구체예에서, 본원에 기재된 재조합 숙주는 디터트렌 생합성 또는 터르페노이드 전구체의 생성과 관련된 재조합 유전자, 예를 들어, 메틸에리트리톨 4-포스페이트(MEP) 또는 메발로네이트(MEV) 경로 내의 유전자를 발현시킨다. 예를 들어, 재조합 숙주는 이소프레노이드 생합성을 위한 MEP 경로와 관련된 효소를 엔코딩하는 하나 이상의 유전자를 포함할 수 있다. MEP 경로 내의 효소는 데옥시자일룰로스 5-포스페이트 신타제(DXS), D-1-데옥시자일룰로스 5-포스페이트 리덕토이소머라제(DXR), 4-디포스포시티딜-2-C-메틸-D-에리트리톨 신타제(CMS), 4-디포스포시티딜-2-C-메틸-D-에리트리톨 키나제(CMK), 4-디포스포시티딜-2-C-메틸-D-에리트리톨 2,4-사이클로디포스페이트 신타제(MCS), 1-하이드록시-2-메틸-2(E)-부테닐 4-디포스페이트 신타제(HDS) 및 1-하이드록시-2-메틸-2(E)-부테닐 4-디포스페이트 환원효소(HDR)를 포함한다. 하나 이상의 DXS 유전자, DXR 유전자, CMS 유전자, CMK 유전자, MCS 유전자, HDS 유전자 및/또는 HDR 유전자가 재조합 미생물로 혼입될 수 있다. 문헌 [Rodriguez-Conception and Boronat, Plant Phys. 130: 1079-1089 (2002)]을 참조하라.

- [0039] DXS, DXR, CMS, CMK, MCS, HDS 및/또는 HDR 폴리펩티드를 엔코딩하는 적합한 유전자는 E. 콜리, 아라비돕시스 탈리아나 및 시네코코쿠스 레오폴리엔시스(*Synechococcus leopoliensis*)에 의해 만들어진 것을 포함한다. DXR 폴리펩티드를 엔코딩하는 누클레오티드 서열은, 예를 들어, 미국 특허 번호 7,335,815호에 기재되어 있다.
- [0040] 일부 구체예에서, 재조합 숙주는 이소프레노이드 생합성을 위한 메발로네이트 경로와 관련된 효소를 엔코딩하는 하나 이상의 유전자를 함유한다. 숙주로의 형질전환에 적합한 유전자는 메발로네이트 경로 내의 효소, 예를 들어, 트렁케이션된 3-하이드록시-3-메틸-글루타릴(HMG)-CoA 환원효소(tHMG)를 엔코딩하고/하거나, 유전자는 메발로네이트 키나제(MK)를 엔코딩하고/거나, 유전자는 포스포메발로네이트 키나제(PMK)를 엔코딩하고/거나, 유전자는 메발로네이트 피로포스페이트 데카르복실라제(MPPD)를 엔코딩한다. 따라서, 하나 이상의 HMG-CoA 환원효소 유전자, MK 유전자, PMK 유전자, 및/또는 MPPD 유전자가 재조합 숙주, 예를 들어, 미생물로 혼입될 수 있다.
- [0041] 메발로네이트 경로 폴리펩티드를 엔코딩하는 적합한 유전자는 공지되어 있다. 예를 들어, 적합한 폴리펩티드는 E. 콜리, 파라코쿠스 데니트리플리칸스(*Paracoccus denitriflcans*), 사카로마이세스 세레비지애, 아라비돕시스 탈리아나(*Arabidopsis thaliana*), 키타사토스포라 그리세올라(*Kitasatospora griseola*), 호모 사피엔스(*Homo sapiens*), 드로소필라 멜란고가스터(*Drosophila melanogaster*), 갈루스 갈루스(*Gallus gallus*), 스트렙토마이세스 종 K0-3988, 니코티아나 아테누아타(*Nicotiana attenuata*), 키타사토스포라 그리세올라(*Kitasatospora griseola*), 헤베아 브라질리엔시스(*Hevea brasiliensis*), 엔테로코쿠스 파에시움(*Enterococcus faecium*), 및 헤마토코쿠스 플루비알리스(*Haematococcus pluvialis*)에 의해 제조된 것을 포함한다. 예를 들어, 미국 특허 번호 7,183,089호, 5,460,949호, 및 5,306,862호를 참조하라.
- IPP로부터 β-카로틴**
- [0043] 일부 구체예에서, 본원에 기재된 재조합 숙주는 IPP로부터 β-카로틴으로의 생합성 경로와 관련된 유전자를 발현시킨다(도 1). 유전자는 숙주에 대해 내인성(즉, 숙주는 자연적으로 카로티노이드를 생성함)일 수 있거나, 외인성, 예를 들어, 재조합 유전자(즉, 숙주는 카로티노이드를 자연 생성하지 않음)일 수 있다. IPP로부터 β-카로틴으로의 생합성 경로 내의 첫번째 단계는 EC 2.5.1.29로 분류된 제라닐제라닐 디포스페이트 신타제(GGPPS 또는 GGDPS)로도 공지됨, GGDP 신타제, 제라닐제라닐 피로포스페이트 신세타제 또는 CrtE)에 의해 촉매된다. EC 2.5.1.29에 의해 촉매되는 반응에서, 트랜스, 트랜스-파르네실 디포스페이트 및 이소펜테닐 디포스페이트는 디포스페이트 및 제라닐제라닐 디포스페이트로 전환된다. 따라서, 일부 구체예에서, 재조합 숙주는 GGPPS를 엔코딩하는 핵산을 포함한다. 적합한 GGPPS 폴리펩티드는 공지되어 있다. 예를 들어, 비제한적인 적합한 GGPPS 효소는 스테비아 래바우디아나(*Stevia rebaudiana*), 지베렐라 푸지쿠로이(*Gibberella fujikuroi*), 무스 무스쿨루스(*Mus musculus*), 탈라시오시라 슈도나나(*Thalassiosira pseudonana*), 잔토필로마이세스 텐드로로우스(*Xanthophyllomyces dendrorhous*), 스트렙토마이세스 클라불리게루스(*Streptomyces clavuligerus*), 술풀로부스 아시디칼다리우스(*Sulfulobus acidocaldarius*), 시네코코쿠스 종(*Synechococcus sp.*) 및 아라비돕시스 탈리아나에 의해 제조된 것을 포함한다. GenBank Accession Nos. ABD92926; CAA75568; AAH69913; XP\_002288339; ZP\_05004570; BAA43200; ABC98596; 및 NP\_195399를 참조하라.
- [0044] 도 1의 경로에서의 다음 단계는 EC 2.5.1.32로 분류된 피토엔 신타제 또는 CrtB에 의해 촉매된다. EC 2.5.1.32에 의해 촉매된 상기 반응에서, 2개의 제라닐제라닐 디포스페이트 분자는 반응하여 2 피로포스페이트 분자 및 피토엔을 형성한다. 이러한 단계는 또한 피토엔-β-카로틴 신타제 또는 CrtYB로 공지된 효소에 의해 촉매될 수 있다. 따라서, 일부 구체예에서, 재조합 숙주는 피토엔 신타제를 엔코딩하는 핵산을 포함한다. 적합한 피토엔 신타제의 비제한적인 예는 X. 텐드로로우스 피토엔-β-카로틴 신타제를 포함한다.
- [0045] β-카로틴의 생합성에서의 다음 단계는 피토엔 디새츄라제(phytoene desaturase) 또는 CrtI로도 공지된 피토엔 데하이드로게나제에 의해 촉매된다. 이러한 효소는 피토엔을 리코펜으로 전환시킨다. 따라서, 일부 구체예에서, 재조합 숙주는 피토엔 데하이드로게나제를 엔코딩하는 핵산을 포함한다. 적합한 피토엔 데하이드로게나제의 비제한적인 예는 뉴로스포라 크라사(*Neurospora crassa*) 피토엔 디새츄라제(GenBank Accession no. XP 964713)를 포함한다. 이를 효소는 또한 식물 및 시아노박테리움에서 풍부하게 발견된다.
- [0046] β-카로틴은 CrtY 또는 CrtL-b로도 언급되는 효소 β-카로틴 신타제를 이용하여 리코펜으로부터 형성된다. 이러한 단계는 또한 다기능성 CrtYB에 의해 촉매될 수 있다. 따라서, 일부 구체예에서, 재조합 숙주는 β-카로틴 신타제를 엔코딩하는 핵산을 포함한다.
- β-카로틴으로부터의 제아잔틴 및 사프란 화합물**

- [0048] 도 2는  $\beta$ -카로틴으로부터 다양한 사프란 화합물로의 경로를 예시한다.
- [0049] 최초 단계에서,  $\beta$ -카로틴은 제아잔틴으로 전환된다. 이러한 전환은  $\beta$ -카로틴을  $\beta$ -크립토잔틴으로 전환시키는  $\beta$ -카로틴 하이드록실라제(BCII)에 의해 촉매되어, 상기  $\beta$ -크립토잔틴은 추가로 반응하여 제아잔틴을 형성한다. 이러한 효소는 CrtZ로도 공지되어 있다. 적합한  $\beta$ -카로틴 하이드록실라제는 잔토필로마이세스 텐드로우스, 아라비돕시스 탈리아나, 아도니스 아에스티발리스(*Adonis aestivalis*), 뿐만 아니라 다른 카로티노이드 생성 미생물로부터 이용가능하다.
- [0050] 제아잔틴은 효소 제아잔틴 분해 디옥시게나제(ZCD)(제아잔틴 분해 옥시게나제(ZCO)로도 공지됨)를 통해 하이드록실- $\beta$ -사이클로시트랄(HBC) 및 크로세틴 디알데하이드로 전환된다. 적합한 ZCD는 크로쿠스 사티바 식물로부터 이용가능하다. 실시예 6을 참조하라. 도 8은 적합한 ZCD를 엔코딩하는 코돈 최적화된 유전자 서열을 함유한다.
- [0051] HBC는 글루코스 공여체로서 UDP-글루코스를 이용하는 아글리콘 O-글리코실 UGT 효소를 이용하여 피크로크로신으로 전환된다. 적합한 UGT는 스테비아 레바우디아나(*Stevia rebaudiana*)로부터의 UGT85C2, 스테비아 73-동족체, 및 2개의 UGT 패밀리 71 하이브리드 UGT를 포함한다. 상기 UGT의 누클레오티드 및 아미노산 서열에 대해서는 도 3을 참조하라(SEQ ID NO:1-20). 변이체 Cs UGT2가 또한 사용될 수 있다(도 9 및 10 참조). 이들 효소는 도 2에서 UGTb로 언급된다. 역 반응은 공지되지 않은 글루코시다제에 의해 촉매된다. 피크로크로신의 생성을 위한 수율 및 역가를 개선시키기 위해, 선택 숙주 유기체 내에서  $\beta$ -글루코시다제 기능성을 높아웃시키는 것이 요망될 수 있다.
- [0052] 사프라날은 사프란의 공정 동안 자연스럽게 형성되며, 이것이 물리적 전환으로 인한 것인지, 또는 효소 또는 효소들에 의한 촉매작용을 필요로 하는 것인지는 공지되어 있지 않다. HBC가 탈수를 통해 사프라날로 직접 전환될 수 있는지, 또는 피크로크로신이 중간체인지 여부는 공지되어 있지 않다.
- [0053] 크로세틴 디알데하이드는 알데하이드 산화환원효소로도 공지된 알데하이드 데하이드로게나제(ADH)에 의해 사프란 식물에서 크로세틴으로 전환될 수 있다. 실시예 9에 기재된 바와 같이, S. 세레비지에는 이종성 유전자의 도입 없이 카르복실레이트 형태로 디알데하이드를 전환시키는데 사용될 수 있는 다수의 내인성 알데하이드 데하이드로게나제를 갖는다. 실시예 9를 참조하라.
- [0054] 크로신 형성에서의 두번째 단계는 크로세틴 분자의 카르복실산 말단으로의 글루코스 모이어티의 첨가이다. 크로쿠스 사티부스 UGT2(CsUGT2)는 크로세틴을 크로세틴의 모노글루코시드(크로세틴 모노글루코실 에스테르 또는 크로세틴 디글리오실 에스테르)로 전환시키는 것으로 밝혀졌다. 이러한 효소는 EC 2.4.1, 유리딘-5'-디포스포글루코스(UDP-글루코스)-크로세틴 8,8'-글루코실트랜스페라제로 분류된다. 이와 같이, 재조합 숙주는 UGT2를 엔코딩하는 핵산을 포함할 수 있다. 크로쿠스 사티부스 UGT2의 핵산 및 아미노산 서열, 및 코돈-최적화된 핵산 서열에 대해서는 도 7을 참조하라. CsUGT2에 대한 GenBank Accession Number는 A Y262037.1이다.
- [0055] 재조합 숙주는 또한 크로세틴으로부터 글루코스 에스테르(크로세틴 모노글루코실 에스테르 또는 크로세틴 디글리오실 에스테르)의 형성을 촉매하는 크로쿠스 사티부스 UGT(Cs VrUGT2)를 포함할 수 있다. 실시예 8을 참조하라. Cs VrUGT2의 아미노산 서열은 도 9에 제공된다. Cs VrUGT2 및 Cs UGT2의 정렬에 대해서는 또한 도 10을 참조하라.
- [0056] 재조합 숙주는 또한 크로신이 크로세틴 디알데하이드로부터 형성될 수 있도록 2개의 글루코스 모이어티 사이의  $\beta$  글루코실 결합(예를 들어,  $\beta$ -1,6 글루코실 결합)을 촉매하는 UGT를 포함할 수 있다. 이러한 UGT는 도 2에서 UGTa로 언급된다. 이와 같이, 재조합 숙주는 UGT2를 엔코딩하는 핵산을 포함할 수 있다. 스테비아 레바우디아나 UGT, UGT76G1은 4개의 글루코스 모이어티와 함께 크로세틴 에스테르를 형성할 수 있는 것으로 밝혀졌다. 실시예 4를 참조하라. 생성물이 크로신 또는 크로신 유사체인 경우 아이소머 특성규명이 결정될 것이다.
- [0057] 스테비아 레바우디아나로부터의 UGT76G1인 3개의 UGT 및 크로쿠스로부터의 2개의 UN1761 및 UN4522는 4개의 글루코스 모이어티를 이용하여 크로세틴 에스테르를 형성할 수 있는 것으로 밝혀졌다. 실시예 4를 참조하라. 스테비아 UGT76G1에 대해, 생성물이 크로신 또는 크로신 유사체인 경우 아이소머 특성규명이 결정될 것이다. UN1761 및 UN4522 각각의 아미노산 서열은 도 5에 기재되어 있다.
- [0058] 재조합 숙주는 또한 어느 하나 또는 둘 모두의 말단 카르복실 말단에서 아글리콘 크로세틴을 촉매하는 UGT를 포함할 수 있다. 3개의 UGT UGT76G1, 또는 UGT71 하이브리드 효소(71C125571C2 및 71C125571E1)는 크로세틴으로부터 모노 및 디글루코실 에스테르의 형성을 나타내었다. 실시예 7을 참조하라.

- [0059] 재조합 숙주는 또한 크로세틴으로부터 직접 젠티비오실 에스테르의 형성을 촉매하는 UGT를 포함할 수 있다. 2개의 UGT UGT71 하이브리드 효소(71C125571C2 및 71C125571E1)는 크로세틴으로부터 젠티비오실 에스테르의 형성을 나타내었다. 실시예 7을 참조하라.
- [0060] 기능적 동족체
- [0061] 상기 기재된 폴리펩티드의 기능적 동족체가 또한 재조합 숙주에서 사프란 화합물을 생성시키는데 사용하기에 적합하다. 기능적 동족체는 참조 폴리펩티드와 서열 유사성을 갖고, 참조 폴리펩티드의 생화학적 또는 생리학적 기능(들) 중 하나 이상을 수행하는 폴리펩티드이다. 기능적 동족체 및 참조 폴리펩티드는 자연 발생 폴리펩티드일 수 있고, 서열 유사성은 수렴 또는 분지 진화 사건으로 인한 것일 수 있다. 이와 같이, 기능적 동족체는 때때로 동족체, 또는 오솔로그(ortholog), 또는 파라로그(paralog)로 문헌에서 표시된다. 자연 발생 기능성 동족체의 변이체, 예를 들어, 야생형 코딩 서열의 돌연변이에 의해 엔코딩된 폴리펩티드는 그 자체로 기능적 동족체일 수 있다. 기능적 동족체는 또한 폴리펩티드에 대한 코딩 서열의 부위-특이적 돌연변이유발을 통하여나, 다양한 자연 발생 폴리펩티드에 대한 코딩 서열로부터의 도메인을 조합("도메인 스와핑(domain swapping)")시킴으로써 생성될 수 있다. 본원에 기재된 기능적 UGT 폴리펩티드를 엔코딩하는 유전자를 변형시키기 위한 기술은 공지되어 있고, 이는 특히 특이적 진화 기술, 부위-특이적 돌연변이유발 기술 및 무작위 돌연변이유발 기술을 포함하며, 이는 폴리펩티드의 특정 활성을 증가시키거나, 기질 특이성을 변경시키거나, 발현 수준을 변경시키거나, 세포하 위치를 변경시키거나, 요망되는 방식으로 폴리펩티드:폴리펩티드 상호작용을 변형시키는데 유용할 수 있다. 이러한 변형된 폴리펩티드는 기능적 동족체로 간주된다. 용어 "기능적 동족체"는 때때로 기능적으로 상동성인 폴리펩티드를 엔코딩하는 핵산에 적용된다.
- [0062] 기능적 동족체는 누클레오티드 및 폴리펩티드 서열 정렬의 분석에 의해 확인될 수 있다. 예를 들어, 누클레오티드 또는 폴리펩티드 서열의 데이터베이스에 대한 질의를 수행하는 것은 본원에 기재된 폴리펩티드의 동족체를 확인할 수 있다. 서열 분석은 참조 서열로서 관심 아미노산 서열을 이용하는 비여분 데이터베이스(nonredundant database)의 BLAST, 상호 BLAST, 또는 PSI-BLAST 분석을 포함할 수 있다. 아미노산 서열은 일부 경우에 누클레오티드 서열로부터 추론된다. 40%를 초과하는 서열 동일성을 갖는 데이터베이스 내의 폴리펩티드는 사프란으로부터의 화합물의 합성에 유용한 폴리펩티드로서의 적합성에 대한 추가 평가를 위한 후보이다. 아미노산 서열 유사성은 보존성 아미노산 치환, 예를 들어, 한 소수성 잔기의 또 다른 소수성 잔기로의 치환 또는 한 극성 잔기의 또 다른 극성 잔기로의 치환을 가능케 한다. 요망시, 상기 후보의 수작업 검사는 추가로 평가되는 후보의 수를 협소화시키기 위해 수행될 수 있다. 수작업 검사는 보존성 기능적 도메인을 갖는 것으로 보이는 후보를 선택함으로써 수행될 수 있다.
- [0063] 보존된 영역은 반복된 서열이거나, 일부 이차 구조(예를 들어, 헬릭스 및 베타 시트)를 형성하거나, 양성 또는 음성으로 하전된 도메인을 확립하거나, 단백질 모티프 또는 도메인을 나타내는 본원에 기재된 폴리펩티드의 일차 아미노산 서열 내의 영역의 위치를 결정함에 의해 확인될 수 있다. 예를 들어, 월드 와이드 웹 sanger.ac.uk/Software/Pfam/ 및 pfam.janelia.org/에서 다양한 단백질 모티프 및 도메인에 대한 컨센서스 서열을 기재하는 Pfam 웹 사이트를 참조하라. Pfam 데이터베이스에 포함된 정보는 문헌[Sonnhammer *et al.*, *Nucl. Acids Res.*, 26:320-322 (1998); Sonnhammer *et al.*, *Proteins*, 28:405-420 (1997); 및 Bateman *et al.*, *Nucl. Acids Res.*, 27:260-262 (1999)]에 기재되어 있다. 보존된 영역은 또한 밀접하게 관련된 종으로부터의 동일하거나 관련된 폴리펩티드의 서열을 정렬시킴으로써 결정될 수 있다. 밀접하게 관련된 종은 바람직하게는 동일 패밀리로부터 유래된다. 일부 구체예에서, 2개의 상이한 종으로부터의 서열의 정렬이 적당하다.
- [0064] 통상적으로, 적어도 약 40%의 아미노산 서열 동일성을 나타내는 폴리펩티드가 보존된 영역을 확인하는데 유용하다. 관련된 폴리펩티드의 보존된 영역은 적어도 45%의 아미노산 서열 동일성(예를 들어, 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 또는 적어도 90%의 아미노산 서열 동일성)을 나타낸다. 일부 구체예에서, 보존된 영역은 적어도 92%, 94%, 96%, 98%, 또는 99%의 아미노산 서열 동일성을 나타낸다.
- [0065] 참조 핵산 또는 폴리펩티드에 비한 임의의 후보 핵산 또는 폴리펩티드에 대한 동일성 퍼센트가 하기와 같이 결정될 수 있다. 참조 서열(예를 들어, 핵산 서열 또는 아미노산 서열)은 전장에 걸쳐 핵산 또는 폴리펩티드 서열의 정렬(전역 정렬)이 수행되는 것을 가능케 하는 컴퓨터 프로그램 ClustalW(version 1.83, default parameters)를 이용하여 하나 이상의 후보 서열과 정렬된다(Chenna *et al.*, *Nucleic Acids Res.*, 31(13):3497-500 (2003)).
- [0066] ClustalW는 참조 서열과 하나 이상의 후보 서열 사이의 최적 매치를 계산하고, 동일성, 유사성 및 차이가 결정될 수 있도록 이들을 정렬시킨다. 서열 정렬을 최대화시키기 위해 참조 서열, 후보 서열, 또는 둘 모두에 하나

이상의 잔기의 캡이 삽입될 수 있다. 핵산 서열의 신속한 쌍을 이룬 정렬을 위해, 다음과 같은 디폴트 파라미터가 사용된다: 워드(word) 크기: 2; 윈도우(window) 크기: 4; 스코어링(scoring) 방법: 백분율; 상부 대각선의 수: 4; 및 캡 패널티(gap penalty): 5. 핵산 서열의 다수의 정렬을 위해, 다음과 같은 파라미터가 사용된다: 캡 오프닝 패널티(gap opening penalty): 10.0; 캡 신장 패널티(gap extension penalty): 5.0; 및 중량 전이(weight transitions): 있음. 단백질 서열의 신속한 쌍을 이룬 정렬을 위해, 다음과 같은 파라미터가 사용된다: 워드 크기: 1; 윈도우 크기: 5; 스코어링 방법: 백분율; 상부 대각선의 수: 5; 캡 패널티: 3. 단백질 서열의 다수의 정렬을 위해, 다음과 같은 파라미터가 사용된다: 중량 매트릭스(weight matrix): 블로섬(blosum); 캡 오프닝 패널티: 10.0; 캡 신장 패널티: 0.05; 친수성 캡: on; 친수성 잔기: Gly, Pro, Ser, Asn, Asp, Gln, Glu, Arg, 및 Lys; 잔기-특이적 캡 패널티: on. ClustalW 출력은 서열 사이의 관계를 반영하는 서열 정렬이다. ClustalW는, 예를 들어, 월드 와이드 웹(searchlauncher.bcm.tmc.edu/multi-align/multi-align.html)의 Baylor College of Medicine Search Launcher 사이트 및 월드 와이드 웹(ebi.ac.uk/clustalw)의 European Bioinformatics Institute 사이트에서 수행될 수 있다.

[0067] 참조 서열에 대한 후보 핵산 또는 아미노산 서열의 동일성 퍼센트를 결정하기 위해, 서열은 ClustalW를 이용하여 정렬되고, 정렬 내의 동일한 매치의 수가 참조 서열의 길이에 의해 나누어지고, 결과에 100이 곱해진다. 동일성 퍼센트 값은 소수점 첫째 자리까지 반올림 및 반내림될 수 있는 것이 인지된다. 예를 들어, 78.11, 78.12, 78.13, 및 78.14는 78.1로 반내림되는 반면, 78.15, 78.16, 78.17, 78.18, 및 78.19는 78.2로 반올림된다.

[0068] 본원에 기재된 폴리펩티드가 글루코실화 또는 효소에 의해 수행되는 다른 효소 활성과 관련되지 않은 추가의 아미노산을 포함할 수 있고, 이에 따라 상기 폴리펩티드는 그렇지 않은 경우보다 더 길 수 있음이 인지될 것이다. 예를 들어, 폴리펩티드는 정제 태그(예를 들어, HIS 태그 또는 GST 태그), 염록체 전이 웨프티드, 미토콘드리아 전이 웨프티드, 아밀로플라스트 웨프티드, 신호 웨프티드, 또는 아미노 또는 카르복시 말단에 첨가되는 분비 태그를 포함할 수 있다. 일부 구체예에서, 폴리펩티드는 리포터, 예를 들어, 녹색 형광 단백질 또는 황색 형광 단백질로서 기능하는 아미노산 서열을 포함한다.

#### 핵산

[0069] 본원에 기재된 폴리펩티드를 엔코딩하는 재조합 유전자는 폴리펩티드를 발현시키는데 적합한 하나 이상의 조절 영역에 센스 배향으로 작동 가능하게 연결된, 상기 폴리펩티드에 대한 코딩 서열을 포함한다. 많은 미생물이 폴리시스템트론 mRNA로부터 다수의 유전자 생성물을 발현시킬 수 있으므로, 다수의 폴리펩티드가 요망시 상기 미생물에 대한 단일 조절 영역의 조절하에서 발현될 수 있다. 코딩 서열 및 조절 영역은, 조절 영역이 서열의 전사 또는 번역을 조절하는데 효과적이 되도록 조절 영역과 코딩 서열이 위치되는 경우에 작동 가능하게 연결된 것으로 간주된다. 통상적으로, 코딩 서열의 번역 판독틀의 번역 판독틀의 번역 개시 부위는 모노시스템트론 유전자에 대한 조절 영역의 다운스트림에 1 내지 약 50개의 누클레오티드 사이에 위치된다.

[0070] 많은 경우에, 본원에 기재된 폴리펩티드에 대한 코딩 서열은 재조합 숙주가 아닌 종에서 확인되고, 즉, 이종성 핵산이다. 따라서, 재조합 숙주가 미생물인 경우, 코딩 서열은 다른 원핵생물 또는 진핵생물 미생물, 식물 또는 동물로부터 유래될 수 있다. 그러나, 일부 경우에, 코딩 서열은 숙주에 대해 자연이고, 상기 유기체로 재조합되는 서열이다. 자연 서열은 종종 외인성 핵산에 연결된 비-자연 서열, 예를 들어, 재조합 핵산 자체물 내의 자연 서열의 측면에 존재하는 비-자연 조절 서열의 존재에 의해 자연 발생 서열과 구별될 수 있다. 또한, 안정적으로 형질전환된 외인성 핵산은 통상적으로 자연 서열이 발견되는 위치가 아닌 위치에 통합된다.

[0071] "조절 영역"은 전사 또는 번역 생성물의 전사 또는 번역 개시 및 속도, 및 안정성 및/또는 이동성에 영향을 미치는 누클레오티드 서열을 갖는 핵산을 나타낸다. 조절 영역은 프로모터 서열, 인핸서 서열, 반응 요소, 단백질 인지 부위, 유도성 요소, 단백질 결합 서열, 5' 및 3' 비번역 영역(UTR), 전사 개시 부위, 종료 서열, 아데닐중합체형성 서열, 인트론, 및 이들의 조합물을 포함하나, 이에 제한되지는 않는다. 조절 영역은 통상적으로 적어도 코어(기본) 프로모터를 포함한다. 조절 영역은 또한 적어도 하나의 조절 요소, 예를 들어, 인핸서 서열, 업스트림 요소, 또는 업스트림 활성화 영역(UAR)을 포함할 수 있다. 조절 영역은 조절 영역이 서열의 전사 또는 번역을 조절하는데 효과적이 되도록 조절 영역 및 코딩 서열을 위치시킴으로써 코딩 서열에 작동 가능하게 연결된다. 예를 들어, 코딩 서열 및 프로모터 서열을 작동 가능하게 연결시키기 위해, 코딩 서열의 번역 판독틀의 번역 개시 부위는 통상적으로 프로모터의 다운스트림에 1 내지 약 50개의 누클레오티드 사이에 위치된다. 그러나, 조절 영역은 번역 개시 부위의 업스트림의 약 5,000개 만큼 많은 누클레오티드, 또는 전사 시작 부위의 업스트림의 약 2,000개 만큼 많은 누클레오티드에 위치될 수 있다.

- [0073] 포함되는 조절 영역의 선택은 효율, 선택성, 유도성, 요망되는 발현 수준, 및 특정 배양 단계 동안의 우선적 발현을 포함하나, 이에 제한되지는 않는 여러 요인에 좌우된다. 코딩 서열에 대해 조절 영역을 적절히 선택하고 위치시킴으로써 코딩 서열의 발현을 조절하는 것은 당업자에게 통상적인 일이다. 하나 이상의 조절 영역, 예를 들어, 인트론, 인핸서, 업스트림 활성화 영역, 전사 중료자, 및 유도성 요소가 존재할 수 있는 것이 이해될 것이다.
- [0074] 하나 이상의 유전자는 사프란으로부터의 화합물의 생성의 별개의 양태에 유용한 "모듈"에서 재조합 핵산 작제물 내에서 조합될 수 있다. 복수의 유전자를 모듈, 특히 폴리시스트론 모듈에서 조합시키는 것은 다양한 종에서의 모듈의 사용을 촉진한다. 예를 들어, 제아잔틴 분해 디옥시게나제, 또는 UGT 유전자 클러스터는 폴리시스트론 모듈 내에서 조합될 수 있어서, 적합한 조절 영역의 삽입 후, 모듈은 매우 다양한 종으로 도입될 수 있다. 또 다른 예로서, UGT 유전자 클러스터는 각각의 UGT 코딩 서열이 별개의 조절 영역에 작동 가능하게 연결되어 UGT 모듈을 형성하도록 조합될 수 있다. 이러한 모듈은 모노시스트론 발현이 필요하거나 요망되는 종에서 사용될 수 있다. 사프란으로부터의 화합물의 생성에 유용한 유전자에 더하여, 재조합 작제물은 통상적으로 또한 복제 기점, 및 적절한 종에서의 작제물의 유지를 위한 하나 이상의 선택가능한 마커를 함유한다.
- [0075] 본 발명의 한 구체예는 SEQ ID NO:57에 기재된 아미노산 서열을 엔코딩하는 SEQ ID NO:58에 기재된 합성 DNA 서열을 제공한다.
- [0076] 본 발명의 또 다른 구체예는 SEQ ID NO:66에 기재된 아미노산 서열을 엔코딩하는 SEQ ID NO:65에 기재된 합성 DNA 서열을 제공한다.
- [0077] 본 발명의 또 다른 구체예는 도 3에 기재된 UGT73 아미노산 서열과 적어도 80%의 서열 동일성을 갖는 UGT73 폴리펩티드를 엔코딩하는 분리된 핵산을 포함하는 DNA 발현 카세트 또는 상기 핵산에 작동 가능하게 연결된 조절 영역을 포함하는 핵산 작제물을 제공한다.
- [0078] 본 발명의 또 다른 구체예는 SEQ ID NO:57에 기재된 아미노산 서열을 엔코딩하는 SEQ ID NO:58에 기재된 합성 DNA 서열을 포함하는 DNA 발현 카세트를 제공하며, 상기 분리된 핵산 또는 합성 DNA 서열은 프로모터에 작동 가능하게 연결된다.
- [0079] 본 발명의 또 다른 구체예는 SEQ ID NO:66에 기재된 아미노산 서열을 엔코딩하는 SEQ ID NO:65에 기재된 합성 DNA 서열을 포함하는 DNA 발현 카세트를 제공하며, 상기 분리된 핵산 또는 합성 DNA 서열은 프로모터에 작동 가능하게 연결된다.
- [0080] 본 발명의 또 다른 구체예는 도 3에 도시된 UGT73 아미노산 서열과 적어도 80%의 서열 동일성을 갖는 UGT73 폴리펩티드를 엔코딩하는 분리된 핵산을 포함하는 DNA 발현 카세트 또는 상기 핵산에 작동 가능하게 연결된 조절 영역을 포함하는 핵산 작제물을 포함하는 재조합 벡터를 제공한다.
- [0081] 본 발명의 또 다른 구체예는 SEQ ID NO:57에 기재된 아미노산 서열을 엔코딩하는 SEQ ID NO:58에 기재된 합성 DNA 서열을 포함하는 DNA 발현 카세트를 포함하는 재조합 벡터를 제공하며, 상기 분리된 핵산 또는 합성 DNA 서열은 프로모터에 작동 가능하게 연결된다.
- [0082] 본 발명의 또 다른 구체예는 SEQ ID NO:66에 기재된 아미노산 서열을 엔코딩하는 SEQ ID NO:65에 기재된 합성 DNA 서열을 포함하는 DNA 발현 카세트를 포함하는 재조합 벡터를 제공하며, 상기 분리된 핵산 또는 합성 DNA는 프로모터에 작동 가능하게 연결된다.
- [0083] 본 발명의 또 다른 구체예는 본 발명에 개시된 바와 같은 DNA 발현 카세트 또는 재조합 벡터를 포함하는 재조합 세포를 제공한다.
- [0084] 본 발명의 또 다른 구체예는 효모, E. 콜리, 식물 세포, 포유동물 세포 및 곤충 세포로 구성되는 군으로부터 선택된 재조합 세포에 관한 것이다.
- [0085] 본 발명의 또 다른 구체예는 사카로마이세스 세레비지애인 재조합 세포에 관한 것이다.
- [0086] 유전 부호의 축퇴성으로 인해, 다수의 핵산이 특정 폴리펩티드를 엔코딩할 수 있고, 즉, 많은 아미노산에 대해, 아미노산에 대한 코돈으로 작용하는 하나 이상의 누클레오티드 트리플렛(triplet)이 존재하는 것이 인지될 것이다. 따라서, 제공된 폴리펩티드에 대한 코딩 서열 내의 코돈은 숙주(예를 들어, 미생물)에 대한 적절한 코돈 편향(bias) 표를 이용하여 특정 숙주에서의 최적 발현이 수득되도록 변형될 수 있다. 분리된 핵산과 같이, 상기 변형된 서열은 정제된 문자로 존재할 수 있고, 재조합 핵산 작제물을 위한 모듈을 작제하는데 사용하기 위한

벡터 또는 바이러스로 통합될 수 있다.

#### [0087] 재조합 속주

다수의 원핵생물 및 진핵생물은 본원에 기재된 재조합 미생물, 예를 들어, 그람-음성 박테리아, 효모 및 진균을 작제하는데 사용하기에 적합하다. 사프란 화합물의 생성을 위한 균주로 사용하기 위해 선택된 종 및 균주는 생성 유전자가 균주에 대해 내인성이고, 유전자가 존재하지 않는(예를 들어, 카로티노이드 유전자) 것을 결정하기 위해 먼저 분석된다. 내인성 대응부가 균주에 존재하지 않는 유전자는 하나 이상의 재조합 작제물에 어셈블리되고, 이는 이후에 손실된 기능(들)을 공급하기 위해 균주로 형질전환된다.

예시적 원핵생물 및 진핵생물 종은 하기에 더욱 상세히 기재된다. 그러나, 다른 종이 적합할 수 있는 것이 인지될 것이다. 예를 들어, 적합한 종은 아가리쿠스(*Agaricus*), 아스페질루스(*Aspergillus*), 바실루스(*Bacillus*), 칸디다(*Candida*), 코리네박테리움(*Corynebacterium*), 에스케리키아(*Escherichia*), 푸사리움(*Fusarium*)/지베렐라(*Gibberella*), 클루이베로마이세스(*Kluyveromyces*), 라에티포루스(*Laetiporus*), 렌티누스(*Lentinus*), 파플라(*Phaffia*), 파네로카에테(*Phanerochaete*), 피키아(*Pichia*), 피스코미트렐라(*Physcomitrella*), 로도투룰라(*Rhodoturula*), 사카로마이세스, 스키조사카로마이세스(*Schizosaccharomyces*), 스파셀로마(*Sphaeloma*), 잔토필로마이세스(*Xanthophyllomyces*) 및 야로위아(*Yarrowia*)로 구성되는 군으로부터 선택된 속 내에 속할 수 있다. 상기 속으로부터의 예시적 종은 렌티누스 티그리누스(*Lentinus tigrinus*), 라에티포루스 술푸레우스(*Laetiporus sulphureus*), 파네로카에테 크리소스포리움(*Phanerochaete chrysosporium*), 피키아 파스토리스(*Pichia pastoris*), 피스코미트렐라 패텐스(*Physcomitrella patens*), 로도투룰라 글루티니스 32(*Rhodoturula glutinis* 32), 로도투룰라 무실라기노사(*Rhodoturula mucilaginosa*), 파피아 로도지마 UVB-AX(*Phaffia rhodozyma* UVB-AX), 잔토필로마이세스 덴드로로우스(*Xanthophyllomyces dendrorhous*), 푸사리움 푸지쿠로이(*Fusarium fujikuroi*)/지베렐라 푸지쿠로이(*Gibberella fujikuroi*), 칸디다 우틸리스(*Candida utilis*) 및 야로위아 리포리티카(*Yarrowia lipolytica*)를 포함한다. 일부 구체예에서, 미생물은 아스코마이세테(Ascomycete), 예를 들어, 지베렐라 푸지쿠로이, 클루이베로마이세스 락티스(*Kluyveromyces lactis*), 스키조사카로마이세스 폼비(*Schizosaccharomyces pombe*), 아스페질루스 니거(*Aspergillus niger*), 또는 사카로마이세스 세레비지애일 수 있다. 일부 구체예에서, 미생물은 원핵생물, 예를 들어, 에스케리키아 콜리, 로도박터 스파에로이데스(*Rhodobacter sphaeroides*), 또는 로도박터 캡슐라투스(*Rhodobacter capsulatus*)일 수 있다. 특정 미생물이 높은 처리량의 방식으로 관심 유전자를 스크리닝하고 시험하는데 사용될 수 있으나, 요망되는 생산성 또는 성장 특성을 갖는 다른 미생물이 사프란으로부터 화합물의 대규모 생성에 사용될 수 있음이 인지될 것이다.

#### [0090] 사카로마이세스 세레비지애

사카로마이세스 세레비지애는 합성 생물학에서 널리 사용되는 본체(chassis) 유기체이며, 이는 재조합 미생물 플랫폼으로 사용될 수 있다. 생성 수율을 향상시키기 위해 다양한 모듈의 합리적인 설계를 가능하게 하는, S. 세레비지애에 대한 돌연변이, 플라스미드, 대사의 상세한 컴퓨터 모델 및 다른 정보의 라이브러리가 존재한다. 재조합 미생물을 만드는 방법이 공지되어 있다.

본원에 기재된 유전자는 다수의 공지된 프로모터 중 임의의 프로모터를 이용하여 효모에서 발현될 수 있다. 터르펜을 과생성하는 균주가 공지되어 있고, 사프란 화합물의 생성에 이용가능한 제라닐제라닐 디포스페이트의 양을 증가시키는데 사용될 수 있다.

S. 세레비지애의 적합한 균주는 저장 지질의 증가된 축적 및/또는 아세틸-CoA와 같은 이용가능한 전구체 분자의 증가된 양을 가능케 하도록 변형될 수 있다. 예를 들어, S. 세레비지애에서의 30%까지의 트리아실글리세롤(TAG)의 축적은 전사 인자 SNF2의 봉괴, 식물-유래 디아실 글리세롤 아실트랜스페라제 1(DGA1)의 과발현, 및 효모 LEU2의 과발현에 의해 아미사카 등(amisaka et al.)(Biochem. J. (2007) 408, 61-68)에 의해 입증되었다. 또한, 프로이사드 등(Froissard et al.)(FEMS Yeast Res 9 (2009) 428-438)은 식물 오일 본체-형성 단백질인 AtClol의 효모에서의 발현이 오일 본체 형성을 촉진하고, 저장 지질의 과축적을 발생시키는 것을 밝혀내었다. 상기 축적된 TAG 또는 지방산은, 예를 들어, TAG(POX 유전자)(예를 들어, 야로위아 리폴리티카 POX 유전자)로부터 아세틸-CoA를 형성할 수 있는 것으로 공지된 효소를 추가로 발현시킴으로써 아세틸-CoA 생합성으로 전환될 수 있다.

#### [0094] 아스페질루스 종

아스페질루스 종, 예를 들어, A. 오리자에(*A. oryzae*), A. 니거(*A. niger*) 및 A. 소자에(*A. sojae*)가 식품 생성에 널리 사용되는 미생물이며, 이는 또한 재조합 미생물 플랫폼으로 사용될 수 있다. 누클레오티드 서열은

유출을 향상시키고, 생성 수율을 증가시키는 내인성 경로의 합리적 설계 및 변형을 가능케 하는, A. 니둘란스 (*A. nidulans*), A. 푸미가투스(*A. fumigatus*), A. 오리자에, A. 클라바투스(*A. clavatus*), A. 플라부스(*A. flavus*), A. 니거, 및 A. 테레우스(*A. terreus*)의 유전체에 대해 이용가능하다. 아스퍼질루스에 대한 대사 모델 뿐만 아니라 전사체학 연구 및 단백질체학 연구가 개발되었다. A. 니거는 다수의 식품 성분, 예를 들어, 시트르산 및 글루콘산의 산업적 생산을 위해 배양되며, 따라서, A. 니거와 같은 종이 사프란으로부터의 화합물의 생성에 일반적으로 적합하다.

#### [0096] 에스캐리키아 콜리

합성 생물학에서 또 다른 널리 사용되는 플랫폼 유기체인 에스캐리키아 콜리가 또한 재조합 미생물 플랫폼으로 사용될 수 있다. 사카로마이세스와 유사하게, 생성 수율을 향상시키기 위해 다양한 모듈의 합리적 설계를 가능케 하는, E. 콜리에 대해 이용가능한 돌연변이, 플라스미드, 대사의 상세한 컴퓨터 모델 및 다른 정보의 라이브러리가 존재한다. 사카로마이세스에 대해 상기 기재된 것과 유사한 방법이 재조합 E. 콜리 미생물을 제조하는데 사용될 수 있다.

#### [0098] 아가리쿠스, 지베렐라, 및 파네로카에테 종

아가리쿠스, 지베렐라, 및 파네로카에테 종이 유용할 수 있는데, 이는 이들이 배양 중에 많은 양의 지베렐린을 생성시키는 것으로 공지되어 있기 때문이다. 따라서, 사프란으로부터 많은 양의 화합물을 생성시키기 위한 터르펜 전구체는 이미 내인성 유전자에 의해 생성된다. 따라서, 사프란으로부터의 화합물의 생합성을 위한 재조합 유전자를 함유하는 모듈은 메탈로네이트 또는 MEP 경로 유전자를 도입시킬 필요 없이 상기 속으로부터의 종에 도입될 수 있다.

#### [0100] 로도박터 종

로도박터는 재조합 미생물 플랫폼으로 사용될 수 있다. E. 콜리와 유사하게, 생성 수율을 향상시키기 위해 다양한 모듈의 합리적 설계를 가능케 하는, 이용가능한 돌연변이 뿐만 아니라 적합한 플라스미드 벡터의 라이브러리가 존재한다. 카로티노이드 및 CoQ10의 증가된 생성을 위한 로도박터의 막 박테리아 종에서 이소프레노이드 경로가 조작되었다. 미국 특히 공개 번호 20050003474호 및 20040078846호를 참조하라. E. 콜리에 대해 상기 기재된 것과 유사한 방법이 재조합 로도박터 미생물을 제조하는데 사용될 수 있다.

#### [0102] 피스코미트렐라 종

현탁액 배양물에서 성장하는 경우 피스코미트렐라 모세스(*Physcomitrella mosses*)는 효모 또는 다른 진균 배양물과 유사한 특징을 갖는다. 이러한 속은 다른 세포 유형을 생성시키는 것이 어려울 수 있는 식물 이차 대사물의 생성을 위한 중요한 세포 유형이 되고 있다.

#### [0104] 식물 및 식물 세포

일부 구체예에서, 본원에 기재된 핵산 및 폴리펩티드는 사프란으로부터 화합물을 생성시키기 위해 식물 또는 식물 세포로 도입된다. 따라서, 숙주는 본원에 기재된 적어도 하나의 재조합 유전자를 포함하는 식물 또는 식물 세포일 수 있다. 식물 또는 식물 세포는 이의 유전체로 통합된 재조합 유전자를 가짐으로써 형질전환될 수 있고, 즉, 안정적으로 형질전환될 수 있다. 안정적으로 형질전환된 세포는 통상적으로 각각의 세포 분열과 함께 도입된 핵산을 보유한다. 식물 또는 식물 세포는 또한, 재조합 유전자가 이의 유전체로 통합되지 않도록 일시적으로 형질전환될 수 있다. 일시적으로 형질전환된 세포는 통상적으로 각각의 세포 분열과 함께 도입된 핵산의 전부 또는 일부를 상실하여, 도입된 핵산은 충분한 수의 세포 분열 후에 딸세포에서 검출될 수 없다. 일시적으로 형질전환된 트랜스제닉 식물 및 안정적으로 형질전환된 트랜스제닉 식물 및 식물 세포 모두는 본원에 기재된 방법에서 유용할 수 있다.

본원에 기재된 방법에서 사용되는 트랜스제닉 식물 세포는 전체 식물의 일부 또는 전부를 구성할 수 있다. 이러한 식물은 성장 챔버, 온실, 또는 논밭에서 고려 중인 종에 적합한 방식으로 성장될 수 있다. 트랜스제닉 식물은 특정 목적을 위해 요망되는 바에 따라, 예를 들어, 다른 균주로 재조합 핵산을 도입시키거나, 다른 종으로 재조합 핵산을 전달하거나, 다른 요망되는 특성의 추가 선택을 위해 품종개량될 수 있다. 대안적으로, 트랜스제닉 식물은 상기 기술에 순종적인 종에 대해 식물적으로 증식될 수 있다. 본원에서 사용되는 바와 같은, 트랜스제닉 식물은 또한 프로제니가 트랜스진(transgene)을 유전하는 한 최초 트랜스제닉 식물의 프로제니를 나타낸다. 트랜스제نك 식물에 의해 생성된 종자는 성장될 수 있고, 이후 핵산 작제물에 대해 동형접합인 종자를 수득하기 위해 자가수분(또는 이종교배 및 자가수분)될 수 있다.

- [0107] 트랜스제닉 식물은 혼탁 배양, 또는 조직 또는 기관 배양으로 성장될 수 있다. 본 발명의 목적상, 고체 및/또는 액체 조직 배양 기술이 사용될 수 있다. 고체 배지를 이용하는 경우, 트랜스제닉 식물 세포는 배지 상에 직접 배치될 수 있거나, 이후에 배지와 접촉하여 배치되는 필터에 배치될 수 있다. 액체 배지를 이용하는 경우, 트랜스제닉 식물 세포는 부유 장치, 예를 들어, 액체 배지와 접촉하는 다공성 막에 배치될 수 있다.
- [0108] 일시적으로 형질전환된 식물 세포가 사용되는 경우, 리포터 활성을 갖는 리포터 폴리펩티드를 엔코딩하는 리포터 서열이 형질전환 절차에 포함될 수 있고, 리포터 활성 또는 발현을 위한 검정이 형질전환 후 적합한 시간에 수행될 수 있다. 검정을 수행하는데 적합한 시간은 통상적으로 형질전환 후 약 1-21일, 예를 들어, 약 1-14일, 약 1-7일, 또는 약 1-3일이다. 일시적 검정의 사용은 다양한 종에서의 신속한 분석, 또는 특정 수용체 세포에서 발현이 이전에 확인되지 않은 이종성 폴리펩티드의 발현을 확인하는데 특히 편리하다.
- [0109] 단자엽 및 쌍자엽 식물로 핵산을 도입시키기 위한 기술은 당분야에 공지되어 있고, 이는 아그로박테리움-매개 형질전환, 바이러스 벡터-매개 형질전환, 전기천공 및 입자 건(particle gun) 형질전환, 미국 특허 번호 5,538,880호; 5,204,253호; 6,329,571호; 및 6,013,863호를 포함하나, 이에 제한되지는 않는다. 세포 또는 배양된 조직이 형질전환을 위한 수용체 조직으로 사용되는 경우, 식물은 요망시 당업자에게 공지된 기술에 의해 형질전환된 배양물로부터 재생될 수 있다.
- [0110] 트랜스제닉 식물의 집단은 트랜스진의 발현에 의해 부여되는 특성 또는 표현형을 갖는 집단의 일원에 대해 스크리닝되고/되거나 선택될 수 있다. 예를 들어, 단일 형질전환 사건의 프로제니의 집단은 ZCD 또는 UGT 폴리펩티드 또는 핵산의 요망되는 수준의 발현을 갖는 식물에 대해 스크리닝될 수 있다. 물리적 및 생화학적 방법이 발현 수준을 확인하는데 사용될 수 있다. 상기는 폴리누클레오티드의 검출을 위한 서던 분석 또는 PCR 증폭; RNA 전사체를 검출하기 위한 노던 블롯, SI RNase 보호, 프라이머-신장, 또는 RT-PCR 증폭; 폴리펩티드 및 폴리누클레오티드의 효소 또는 리보자임 활성을 검출하기 위한 효소 검정; 및 폴리펩티드를 검출하기 위한 단백질 겔 전기영동, 웨스턴 블롯, 면역침전, 및 효소-결합 면역검정을 포함한다. 다른 기술, 예를 들어, 제자리부합법(*in situ hybridization*), 효소 염색, 및 면역염색이 또한 폴리펩티드 및/또는 핵산의 존재 또는 발현을 검출하기 위해 사용될 수 있다. 언급된 기술 모두를 수행하는 방법은 공지되어 있다. 대안으로서, 독립적 형질전환 사건을 포함하는 식물의 집단은 사프란으로부터의 화합물의 생성과 같은 요망되는 특성을 갖는 식물에 대해 스크리닝될 수 있다. 선택 및/또는 스크리닝은 하나 이상의 세대에 걸쳐, 및/또는 하나 이상의 지리적 위치에서 수행될 수 있다. 일부 경우에, 트랜스제닉 식물은 요망되는 표현형을 유도하거나, 트랜스제닉 식물에서 요망되는 표현형을 생성시키는데 달리 요구되는 조건하에서 성장되고 선택될 수 있다. 또한, 선택 및/또는 스크리닝은 표현형이 식물에 의해 나타날 것으로 예상되는 특정 발달 단계 동안 적용될 수 있다. 선택 및/또는 스크리닝은 트랜스진이 결핍된 대조 식물에 비해 사프란 화합물의 수준에서 통계적으로 유의한 차이를 갖는 트랜스제닉 식물을 선택하기 위해 수행될 수 있다.
- [0111] 본원에 기재된 핵산, 재조합 유전자, 및 작제물은 다수의 단자엽 및 쌍자엽 식물 및 식물 세포 시스템을 형질전환시키기 위해 사용될 수 있다. 적합한 단자엽의 비제한적인 예는, 예를 들어, 곡물 작물, 예를 들어, 벼, 호밀, 수수, 기장, 밀, 옥수수, 및 보리를 포함한다. 식물은 또한 쌍자엽, 예를 들어, 콩, 목화, 해바라기, 완두, 제라늄, 시금치, 또는 담배일 수 있다. 일부 경우에, 식물은 세포질 및 미토콘드리아에서 통상적으로 발견되는 페닐 포스페이트 생성을 위한 전구체 경로, 예를 들어, 메발로네이트 경로를 함유할 수 있다. 비-메발로네이트 경로가 식물 색소체에서 더욱 흔히 발견된다[Dubey, et al., 2003 *J. Biosci.* 28 637-646]. 당업자는 선도 서열의 사용을 통해 적절한 세포소기관으로 생합성 폴리펩티드의 발현을 표적화할 수 있어, 생합성이 식물 세포의 요망되는 위치에서 발생한다. 당업자는 직접 합성을 위해, 예를 들어, 요망되는 경우 식물의 잎에 대해 적절한 프로모터를 사용할 것이다. 발현은 또한 조직 배양물, 예를 들어, 요망되는 경우 칼루스(callus) 배양물 또는 모상근(hairy root) 배양물에서 발생할 수 있다.
- [0112] 본 발명은 청구항에 기재된 본 발명의 범위를 제한하지 않는 하기 실시예에 추가로 기재될 것이다.
- ### 도면의 간단한 설명
- [0113] 도면의 설명
- 도 1은 IPP로부터  $\beta$ -카로틴으로의 생합성 경로의 개략도이다.
- 도 2는 사프란 내에서의 생합성 경로의 개략도이다.
- 도 3은 스테비아 레바우디아나(*Stevia rebaudiana*) UGT88B1(SEQ ID NO:1 및 2), UGT76G1(SEQ ID NO:3 및 4),

UGT74G1(SEQ ID NO:5 및 6), UGT91D2e(SEQ ID NO:7 및 8), UGT85C2(SEQ ID NO:9 및 10), 및 UGT73(SEQ ID NO:11 및 12), 카타란투스 로세우스(*Catharanthus roseus*) UGT2(SEQ ID NO:13 및 14), 아라비돕시스 탈리아나(*Arabidopsis thaliana*) UGT75B1(SEQ ID NO:15 및 16), 및 2개의 A. 탈리아나 하이브리드 UGT(UGT71 하이브리드 효소 1: 71C125571C2, SEQ ID NO:17 및 18) 및 UGT71 하이브리드 효소 2: 71C125571E1, SEQ ID NO:19 및 20)의 누클레오티드 및 아미노산 서열을 함유한다.

도 4는 공지된 UGT91 서열을 갖는 UN1671, UN3356, UN4522, UN4666, UN6460, 및 UN2281 UGT 클러스터의 아미노산 서열을 도시하는 개략도이다.

도 5는 실시예 4에서 확인된 UGT의 서열(UN6338, SEQ ID NO:21; UN4666, SEQ ID NO:22(DNA) 및 23(아미노산); UN3356, SEQ ID NO:24(DNA) 및 25(아미노산); UN6428, SEQ ID NO:26; UN3131, SEQ ID NO:27; UN1671, SEQ ID NO:28(DNA) 및 29(아미노산); UN4522, SEQ ID NO:30(DNA) 및 31(아미노산); UN6460, SEQ ID NO:32(DNA) 및 33(아미노산); UN2281, SEQ ID NO:34(DNA) 및 35(아미노산); 및 UN2644, SEQ ID NO:36)을 함유한다.

도 6은 사카로마이세스 세레비지애(소스: DNA 2.0>) 내의 EUGT1-EUGT19의 발현을 위한 코돈 최적화된 누클레오티드 서열 SEQ ID NO:37-55의 서열을 함유한다.

도 7은 크로쿠스 사티부스 글루코실트랜스페라제 2(UGT2)(GenBank Accession No. AY262037.1)의 누클레오티드(SEQ ID NO:56) 및 아미노산 서열(SEQ ID NO:57), 뿐만 아니라 코돈-최적화된 핵산 서열(SEQ ID NO:58)을 함유한다.

도 8은 실시예 6에서 사용된 코돈 최적화된 유전자 서열(SEQ ID NO:59-64)을 함유한다. 소문자 서열은 코딩 영역과 관계가 없고, 클로닝 목적에 사용된다.

도 9는 실시예 8에서 사용된 변이체 크로쿠스 UGT(Cs VrUGT2)의 코돈 최적화된 누클레오티드 서열(소스: GenScript)(SEQ ID NO:65) 및 아미노산 서열(SEQ ID NO:66)을 함유한다.

도 10은 크로쿠스 사티부스로부터의 CsUGT2(GenBank Accession Number: AY262037.1) 및 변이체 Cs VrUGT2의 정렬, 뿐만 아니라 각각의 폴리펩ти드의 아미노산 서열(SEQ ID NO:57 및 66)을 함유한다.

도 11은 알데하이드 데하이드로게나제(ALD) 2, ALD3, ALD4, ALD5, ALD6, 및 HFD1(알데하이드 데하이드로게나제인 것으로 또한 예측됨)을 엔코딩하는 누클레오티드 서열(SEQ ID NO:67-72)을 함유한다.

다양한 도면 내의 유사한 참조 기호는 유사한 요소를 나타낸다.

## 발명을 실시하기 위한 구체적인 내용

[0114]

실시예

[0115]

실시예 1: 효모에서의  $\beta$ -카로틴의 생성

[0116]

$\beta$ -카로틴 생성 효모 리포터 균주를 사프란 생합성 유전자의 최적 조합을 찾기 위해 설계된 eYAC 실험을 위해 작제하였다. 뉴로스포라 크라사 피토엔 디세츄라제(피토엔 데하이드로게나제로도 공지됨)(등록 번호 XP\_964713) 및 제라닐제라닐 피로포스페이트 신세타제 또는 CrtE(등록 번호 DQ0 12943)로도 공지된 잔토필로마이세스 텐드로로우스 GGDP 신타제 및 X. 텐드로로우스 피토엔- $\beta$ -카로틴 신타제 CrtYB(등록 번호 AY177204) 유전자를 모두 발현 카세트에 삽입하였고, 이들 발현 카세트를 실험실 효모 균주 사카로마이세스 세레비지애 CEN.PK 113-11의 유전체로 통합시켰다. 피토엔 디세츄라제 및 CrtYB를 강한 항시성 GPD1 프로모터의 조절하에서 과발현시킨 한편, 강한 항시성 TPI1 프로모터를 이용하여 CrtE의 과발현이 가능하였다. X. 텐드로로우스 CrtE 및 뉴로스포라 크라사 피토엔 디세츄라제 발현 카세트의 염색체 통합을 S. 세레비지애 ECM3-YOR093C 유전자간 영역에서 수행한 한편, CrtYB 발현 카세트의 통합을 S. 세레비지애 KIN1-INO2 유전자간 영역에서 수행하였다.

[0117]

SC 드롭아웃(dropout) 플레이트에서 성장된 콜로니는  $\beta$ -카로틴이 생성되는 경우 오렌지 색 형성을 나타낸다.  $\beta$ -카로틴의 존재는 메탄올로의 추출 및 LC/MS 분석에 의해 정량된다.

[0118]

실시예 2: HBC 및 크로세틴 디알데하이드의 최적화된 효모 생성

[0119]

크로세틴이 크로세틴 디알데하이드로부터 형성되고, 크로세틴 디알데하이드 및 하이드록실-베타-사이클로시트랄(HBC)이 효소 제아잔틴 분해 디옥시게나제(ZCD)를 이용한 제아잔틴 분해 시에 생성되는 것이 공지되어 있다. 유전자 수집물을 실시예 1에 기재된 eYAC 및  $\beta$ -카로틴 생성 효모 균주를 이용하여 크로세틴 디알데하이드 및

HBC의 생합성을 위한 최적 경로를 확립시키기 위해 eYAC에서 어셈블리시켰다.

[0120]  $\beta$ -카로틴을 크로세틴 디알데하이드로 전환시키는 효소에 대한 유전자 유사체의 수집물은 효모 코돈 최적화된 합성(DNA 2.0)에 의해 공급되었고, 다양한 메티오닌 억제성 유전자 프로모터 하에서 eYAC Entry Vector 내에 삽입되었다. eYAC 기술의 사용은 문헌[Naesby *et al.*, *Microb Cell Fact.* 8:45 (2009)]에 기재되었다. 표 1에 제시된 37개의 사프란 생합성 유전자에 대한 발현 카세트를 연결시키고(UGT 유전자를 갖거나 갖지 않음), eYAC에 라이제이션시켰다. 둘 모두의 유형의 eYAC를  $\beta$ -카로티노이드 생성 효모 균주 EFSC301으로 형질전환시켰다. 이러한 균주는 효모 ECM3 및 KIN1 3'UTR 영역 내의 GPD/TPI 프로모터-기반 CrtYB/CrtE/Nc-AI-1 유전자 발현 카세트의 통합에 의해 제조된 안정적인 카로티노이드 생산자이다.

[0121] 약 800개의 콜로니/플레이트의 효모 형질전환 효율을 단일 영양요구 선택 플레이트를 이용하여 수득하였다. 이후, 형질전환체를 이중 영양요구 선택 플레이트(류신-, 트립토판-) 상에 재-스트리킹(re-streaking)시켰다. 양성 형질전환체를 SC 드롭아웃 배지(-류신, -트립토판, 및 -메티오닌)에서 성장시켰다. 세포를 진탕 플라스크에서 30°C에서 24-72시간 동안 성장시키고, 세포-비함유 브로쓰(broth) 뿐만 아니라 세포 추출물을 유기 용매로 추출하고, HBC, 크로세틴 디알데하이드, 및 크로세틴의 존재에 대해 분석하였다.

[0122] 형질전환된 효모에서 생합성된 크로세틴 디알데하이드, 크로세틴 및 HBC의 함량을 기초로 하여, 높은 생산자, 중간 생산자 및 낮은 생산자를 확인하였다. 이들 형질전환체를 PCR에 의해 스크리닝하여, 높은 생산자, 중간 생산자 및 낮은 생산자의 유전자 조성을 결정하였다. PCR 결과를 기초로 하여, 크로세틴 디알데하이드, 크로세틴 및 HBC 생성에 필수적 및 비-필수적인 유전자를 확인하고, 새로운 조합 및 새로운 eYAC 작제물에서 유전자를 추가하거나 삭제함으로써 작제물을 추가로 개선시킬 수 있다.

표 1: eYAC 작제를 위한 유전자 공급원

S.No.	코드	등록 번호	유전자명	소스	크기(bp)
1	CH-1	D90087	$\beta$ -카로틴 하이드록실라제	판노에야 아나나티스(박테리아)	567
2	CH-2	DQ201828	$\beta$ -카로틴 3-하이드록실라제 (crtS)	판노셀로마이세스 엔드로로우스	1713
3	CH-3	NM_124636	$\beta$ -고리 하이드록실라제 (CHY2)	아라비누시스 달리아나(식물)	951
4	CH-4	AF125576	$\beta$ -카로틴 하이드록실라제	아라비누시스 달리아나	972
5	ZCO-1	AJ489276	제아잔틴 분해 디옥시케나제 (CsZCO )	크로쿠스 사티부스	1149
6	ZCO-2	AJ132927	카로티노이드 9,10(9',10')-분해 디옥시케나제(CsCCD)	크로쿠스 사티부스	1680
7	ZCO-3	AJ489277	리코펜 분해 디옥시케나제	빅사 오렐라나(식물)	1149
8	ZCO-4	AB247160	카로티노이드 분해 디옥시케나제 (CmCCD4a)	[크리산체움 x 보리풀리움]	1824
9	ZCO-5	AB120111	카로티노이드 분해 디옥시케나제1 (Ls CCD1)	락투카 사티파	1818
10	ZCO-6	EU334343	카로티노이드 분해 디옥시케나제4	オス반부스 프라그린스	1869
11	ZCO-7	AY662342	9-시스-에폭시-카로티노이드 디옥시케나제 1	솔라눔 퓨페로솔	1851
12	EUGT-1	AY262037	글루코실트랜스페라제 2, UGTCs2	크로쿠스 사티부스	1422
13	EUGT-2	AP003270	추정 UDP-글루코실트랜스페라제	오리자 사티파	1461
14	EUGT-3	AP005171	추정 UDP-글루코실트랜스페라제	오리자 사티파	1563
15	EUGT-4	AP005643	추정 UDP-글루코실트랜스페라제	오리자 사티파	1335
16	EUGT-5	AY290820	글루코실트랜스페라제 , UGTCs3	크로쿠스 사티부스	1467
17	CH-5	U58919	베타-카로틴 하이드록실라제	아라비누시스 달리아나	887
18	CH-6	EF120636	$\beta$ - 카로틴 하이드록실라제	아도니스 아에스티발리스	930
19	CH-7	Y14810	베타-카로틴 하이드록실라제	솔라눔 티코페르시뮴	945
20	CH-8	NM_001036638	카로틴 베타-고리 하이드록실라제 (BETA-OHASE 1)	아라비누시스 달리아나	675
21	CH-9	NC_010475	베타-카로틴 옥시게나제 CrtR	시네코쿠스 종 PCC 7002	888
22	CH-10	NC_008819	베타-카로틴 하이드록실라제	프로플로로코쿠스 마리누스	1032
23	CH-11	NC_010296	베타-카로틴 하이드록실라제 (crtR)	미크로시스터스 아에루기노사	894
24	EUGT-6	AP005259	추정 UDP-글루코실트랜스페라제	오리자 사티파	1539
25	EUGT-7	AP005171	추정 UDP-글루코실트랜스페라제	오리자 사티파	1524
26	EUGT-8	XN_470006	추정 UDP-글루코로노신 및 UDP-글루코신	오리자 사티파	1452
27	EUGT-9	AP005643	추정 UDP-글루코실트랜스페라제	오리자 사티파	1296
28	EUGT-10	AC133334	추정 UDP-글루코로노신 및 UDP-글루코신 트랜스페라제	오리자 사티파	1419
29	EUGT-11	AC133334	추정 UDP-글루코로노신 및 UDP-글루코신 트랜스페라제	오리자 사티파	2319
30	EUGT-12	AP004741	추정 UDP-글루코실트랜스페라제	오리자 사티파	1338
31	EUGT-13	AB012241	글루코실트랜스페라제-유사 단백질	아라비누시스 달리아나	1056
32	EUGT-14	AL133314	글루코실트랜스페라제-유사 단백질	아라비누시스 달리아나	1317
33	EUGT-15	Z25802	UDP-蹂노스; 암토시아니딘-3-글루코사이드 람노실트랜스페라제	페류나 X 허브리다	1416
34	EUGT-16	AC004786	추정 펄리보논 3-O-글루코실트랜스페라제	아라비누시스 달리아나	1329
35	EUGT-17	AB294391	글루코실트랜스페라제	디안투스 코리오펄루스	1386
36	EUGT-18	AB192314	글루코실트랜스페라제	이포모에야 날	1380
37	EUGT-19	NM_001074394	가상 단백질	오리자 사티파	1413

[0124]

### 실시예 3: 피크로크로신-형성 UGT의 발견

[0125] 글루코실트랜스페라제 효소는 하이드록실-베타-사이클로시트랄(HBC)로부터 피크로크로신을 형성시키는 평 필요하다. 이러한 반응은 글루코스-글루코스 결합-형성 반응에 반대되는 아글리콘 글루코실화이며, 상기 유형의 활성을 스크리닝하기 위해 UDP-글루코스를 이용하는 글리코실트랜스페라제의 많은 패밀리가 존재한다.

- [0127] *HBC 기질의 소싱(sourcing)*
- [0128] HBC를 합성하고, 요망되는 화합물을 키랄 컬럼 크로마토그래피(GVK, Hyderabad)에 의해 정제하였다.
- [0129] *UGT 효소의 스크리닝*
- [0130] 다음과 같은 UGT를 피크로크로신 형성에 대해 검정하였다: 스테비아 레바우디아나 88B1, 76G1, 74G1, 91D2e, 85C2, 73EV12; 카타란투스 로세우스 UGT2; 및 아라비돕시스 탈리아나 UGT 75B1, 및 아라비돕시스 하이브리드 효소 UGT 353 및 UGT354(도 3에 서열이 제공됨).
- [0131] 상기 UGT를 엔코딩하는 유전자를 T7 프로모터를 이용하여 플라스미드로 클로닝하고, 발현 연구를 위해 E. 콜리 BL21 세포로 형질전환시켰다. 상기 UGT를 갖는 균주를 0.1 mM IPTG를 이용하여 유도하고, 유도된 배양물을 밤새 20°C에서 성장시켰다. 이후, 유도된 세포를 BugBuster 시약(Novagen)을 이용하여 용해시키고, 정화된 용해질을 UGT 검정에 사용하였다.
- [0132] UGT 검정을 글루코스-수용체 기질( $10 \mu\text{M}$  최종 농도)로서 HBC 및 공여체로서 UDP-글루코스( $1 \text{mM}$  최종 농도)에 첨가된  $98 \mu\text{L}$ 의 유도된 정화 용해질을 이용한  $100 \mu\text{L}$  반응으로 수행하였다. 반응을 3시간 동안 30°C에서 수행하였고,  $300 \mu\text{L}$ 의 물 포화 1-부탄올의 첨가에 의해 종료시켰다. 샘플을  $300 \mu\text{L}$ 의 물-포화 1-부탄올을 이용하여 3회 추출하였다. 푸울링된 부탄올 분획을 Speed-vac에서 완전 건조시키고, 하기 방법을 이용하여 LC/MS에 의해 분석하였다. Luna-SL C18 컬럼( $5 \mu\text{m}$ , 100 옹스트롬) 모델 G1316B(4.6 mm ID)을 LC 분리에 사용하였고, 440 nm에서 모니터하였다. 0.25% 포름산(FA)인 다른 용매와 함께 20-80% 아세토나트릴로부터의 구배를 이용하여  $0.8 \text{ mL}/\text{분}$ 으로 20분의 분리를 수행하였다. MS 분석을 위해 LC를 Q-TOF에 커플링시켰다.
- [0133] 상기 UGT 중, 스테비아로부터의 UGT85C2 및 UGT73EV12 및 2개의 하이브리드 아라비돕시스 효소는 검정된 조건하에서 HBC로부터 피크로크로신의 형성을 나타내었다. 예비 분석은 스테비아 UGT85C2와의 반응이 HBC를 피크로크로신 표준과 유사한 보유 시간 및 질량을 갖는 화합물로 부분적으로 전환시킨 것을 나타내었다. HBC 피크 영역을 표준의 보유 시간으로 모니터하였다.
- [0134] 스테비아 UGT85C2는 HBC를 생성시키는 것으로 밝혀진 효모 균주에서 공동 발현된다(실시예 2 & 6 참조). 상기 효소는 시험판내에서 밝혀진 것과 동일 반응을 수행하여 피크로크로신을 형성할 수 있는 3개의 추가 UGT를 확인하였다: 스테비아 UGT73, 및 2개의 아라비돕시스 UGT71 하이브리드 효소(하이브리드 효소와 관련하여 Hansen, et al., *Phytochemistry* 70 (2009) 473-482 참조). 도 3은 UGT73 및 UGT71 하이브리드 효소의 누클레오티드 및 아미노산 서열을 제공한다.
- [0135] *스크리닝 UGT 수집물*
- [0136] 광범위한 특이성을 갖는 170개 초과의 UGT 효소의 수집물을 E. coli에서 발현시키고, 상기 기재된 것과 유사한 방식으로 검정하였다. HBC와 함께 당화 반응을 수행하여 피크로크로신을 형성할 수 있는 3개의 추가 UGT를 확인하였다: 스테비아 UGT73, 및 2개의 아라비돕시스 UGT71 하이브리드 효소(하이브리드 효소와 관련하여 Hansen, et al., *Phytochemistry* 70 (2009) 473-482 참조). 도 3은 UGT73 및 UGT71 하이브리드 효소의 누클레오티드 및 아미노산 서열을 제공한다.
- [0137] **실시예 4: 크로신-형성 글리코실트랜스페라제 효소의 발견**
- [0138] 크로신은 연속적인 반응으로 크로세틴에 첨가되는 4개의 글루코스 모이어티를 갖는 크로세틴의 유도체이다. 최종 2개의 글루코스 분자는 하나의 글리코실트랜스페라제의 작용에 의한 것임이 매우 유력한  $\beta$ -1,6-결합에 의해 2개의 제 1의 글루코스 분자에 부착된다. 제 2의 글루코스의 첨가를 촉매하는 UGT 효소는 아글리콘 글리코실라제 트랜스페라제보다 덜 흔하고, UGT 서브페밀리 91 또는 79에 속할 것이다. 이를 2개의 서브페밀리는 1,2 또는 1,6 글루코스-글루코스 결합의 형성을 촉매하는 것으로 현재 공지된 오직 2개의 서브페밀리이다.
- [0139] 크로쿠스로부터의 유전자를 확인하기 위한 노력으로, 크로쿠스 스티그마로부터의 서브페밀리 79 및 91 UGT 뿐만 아니라 다른 서브페밀리 91 UGT를 확인하고, 분리하였다.
- [0140] *크로쿠스 피로시퀀싱(pyrosequencing)*
- [0141] 크로쿠스 스티그마 cDNA에 대한 피로시퀀싱을 MOgene LC(St. Louis, MO, USA)로부터 받았다. 전체 전사체 시퀀싱을 2개의 FLX 티타늄 플레이트를 이용하여 수행하고, 전체 약 1100 MB의 미가공 시퀀싱 데이터를 발생시키고, 새로운 어셈블리를 수행하였다.
- [0142] 피로시퀀싱된 데이터의 66,000개의 독특한 콘티그(contig)를 분석한 후, 약 10개의 UGT-유사 서열(서브페밀리 91)을 공지된 UGT에 대한 블라스트 분석에 의해 확인하였다. 이를 기초로 하여, 유전자/대립유전자 특이적 역

PCR 프라이머를 크로쿠스 cDNA 라이브러로부터 전장 유전자를 분리시키도록 설계하였다.

[0143] 유전자-특이적 및 벡터-특이적 프라이머를 괴로시퀀싱 데이터를 기초로 하여 설계하고, UGT 유전자의 5'- 및 3'-말단을 얻기 위해 사용하였다. 프루프 리딩 폴리머라제(proof reading polymerase)(예를 들어, Advantage 2 및 KOD 폴리머라제)를 이용하여 유전자 및 벡터 특이적 프라이머의 조합을 이용한 UGT 서열의 5' 및 3' 말단의 성공적인 증폭 후, 증폭된 PCR 단편을 다운스트림 가공을 위해 젤 추출하였다. PCR 증폭된 단편을 PCR 정제 키트를 이용하여 정제한 후, 계속해서 TA 클로닝 벡터(InstaTA cloning kit, Fermentas)로 클로닝하고, E. 콜리 균주(NEB 10- $\beta$  Competent cells, New England Biolabs, UK)로 형질전환시켰다. 유전자 특이적 콜로니-PCR 을 이용한 PCR 단편의 정량 분석 후, 플라스미드 DNA 샘플을 시퀀싱하였다.

[0144] 서브페밀리 91로부터의 6개의 전장 UGT 크로쿠스 cDNA 서열을 상기 방식으로 확인하였다. 6개 모두 UGT(UN1671, UN3356, UN4522, UN4666, UN6460 및 UN2281)의 아미노산 서열은 공지된 UGT91 서열과 함께 군집을 이룬다(도 4 참조; 도 5는 UN1671, 3356, 4522, 4666, 6460 및 2281의 서열을 함유함). 상기 6개 중에서, UN1671 전사체 및 U4522 전사체가 전사체에서의 이의 풍부함을 기초로 하여 발견된 91개의 동족체 중 가장 고도로 발현되었다.

[0145] UN1671, UN4522, UN4666, UN6460, UN3356 및 UN2281의 6개의 전장 서열을 유전자 특이적 프라이머를 이용하여 추가로 증폭시키고, E. 콜리 발현 및 시험관내 발현을 위해 플라스미드 벡터에 삽입하였다.

[0146] SMART PCR cDNA 합성 방법을 추가 7개의 UGT의 완전 서열의 증폭에 이용하였다. 이러한 방법은 나노그램 양의 전체 RNA로부터 고품질의 cDNA를 생성시키는 능력을 갖는다. mRNA의 아데닐중합체형성된 영역을 포획하는 친화성 방법을 기초로 하여 정제된 크로쿠스 mRNA로부터 RACE cDNA를 제조하였다. 유전자 특이적 및 대립유전자 특이적 프라이머를 전장 UGT 코딩 영역을 수득하기 위해 이용하였다. 코딩 영역을 각각의 UGT를 갖는 E. 콜리 T7 Express lysY/I<sup>q</sup> 적격 E. 콜리(New England Biolabs, UK) 균주에서 형질전환시키고, 항생제를 함유하는 Luria Broth 배지에서 성장시키고, 16시간 동안 37°C에서 인큐베이션(250 rpm에서 진탕)하였다. 세포를 새로운 LB 중에 0.01의 OD600으로 접종시키고, 0.4 내지 0.5의 OD600에 도달할 때까지 30°C에서 성장시켰다. 온도를 20°C로 낮추고, 세포를 0.1 mM IPTG로 유도시키고, 24시간 동안 인큐베이션하였다. 세포를 실온에서 1분 동안 12,000 rpm에서 펠렛화시키고, 제조업체의 프로토콜에 따라 Bug buster 시약(Novagen)에서 용해시켰다. 정화된 상층액을 3시간 동안 30°C에서 인큐베이션된 반응물 중의 10 mM UDP-글루코스(최종 농도) 및 1 mM 디-글루코실 에스테르(최종 농도)를 이용하는 UGT 검정에 사용하였다.

#### 시험관내 번역된 효소의 스크리닝

[0147] 전체 19개의 UGT 유전자(표 2 참조)를 다른 서브페밀리 79 또는 91 UGT 서열과의 이들의 상동성으로 인해 부분적 모노-당화된 크로세틴 에스테르의 크로신으로의 전환을 위한 후보로 선택하였다. 모든 유전자를 효모 코돈 사용(도 6 내의 누클레오티드 서열)에 대한 최적화를 이용하여 합성하였다.

[0149]

표 2

코드	등록 번호	유전자명	크기 (bps)
EUGT-1	AY262037	글루코실트랜스페라제 2, UGTCs2	1383
EUGT-2	AP003270	추정 UDP-글루코실트랜스페라제	1422
EUGT-3	AP005171	추정 UDP-글루코실트랜스페라제	1524
EUGT-4	AP005643	추정 UDP-글루코실트랜스페라제	1296
EUGT-5	AY290820	글루코실트랜스페라제 , UGTCs3	1428
EUGT-6	AP005259	추정 UDP-글루코실트랜스페라제	1539
EUGT-7	AP005171	추정 UDP-글루코실트랜스페라제	1524
EUGT-8	XM_470006	추정 UDP-글루코로노실 및 UDP-글루코실	1452
EUGT-9	AP005643	추정 UDP-글루코실트랜스페라제	1296
EUGT-10	AC133334	추정 UDP-글루코로노실 및 UDP-글루코실 트랜스페라제	1419
EUGT-11	AC133334	추정 UDP-글루코로노실 및 UDP-글루코실 트랜스페라제	1389
EUGT-12	AP004741	추정 UDP-글루코실트랜스페라제	1338
EUGT-13	AB012241	글루코실트랜스페라제-유사 단백질	1056
EUGT-14	AL133314	글루코실트랜스페라제-유사 단백질	1317
EUGT-15	Z25802	UDP 람노스: 안토시아닌-3-글루코시드 람노실트랜스페라제	1416
EUGT-16	AC004786	추정 플라보놀 3-O-글루코실트랜스페라제	1329
EUGT-17	AB294391	글루코실트랜스페라제	1386
EUGT-18	AB192314	글루코실트랜스페라제	1380
EUGT-19	NM_0010743 94	가상 단백질	1413

[0150]

\* 시험관내 번역을 위한 T7 프로모터 서열을 이용하여 PCR 증폭될 수 없음/EUGT # 2, 8 및 11에 대해 표현됨.

[0152]

시험관내 번역은 16개의 UGT에 대해 성공적이었고; 다른 3개의 UGT를 E. 콜리 기반 발현 시스템으로 클로닝하였다. 16개의 시험관내 번역된 UGT를 글루코스-수용체 기질로서 크로세틴 젠티오비오실글루코실 에스테르(크로세틴-3G, GVK, India) 및 글루코스 공여체로서 UDP-글루코스를 이용하여 크로신 형성에 대해 스크리닝하였다. 40 μL의 시험관내 번역된 단백질을 3 mM의 최종 농도의 MgCl<sub>2</sub>, 10 μg/mL의 BSA, 50 μM의 기질, 및 1 mM UDP-글루코스를 함유하는 100 μL의 반응에서 사용하였다. 50 mM 포타슘 포스페이트 완충액 pH 7.2 중에서 3시간 동안 30°C에서 반응을 수행하고, 300 μL의 물 포화 1-부탄올을 첨가하여 종료시켰다. 샘플을 300 μL의 물-포화 1-부탄올로 3회 추출하였다. 푸울링된 부탄올 분획을 Speed-Vac에서 완전히 건조시키고, 메탄올 중에 재현탁시키고, Agilent 1200 HPLC & Q-TOF LC/MS 6520에 의해 분석하였다. 시험된 샘플 중 어느 것도 검정된 반응 조건하에서 크로신을 생성시키는 것으로 보이지 않았다.

[0153]

#### 식물 UGT 효소의 스크리닝

[0154]

스테비아로부터의 5개의 UGT(88B1, 76G1, 74G1, 912D2e, 및 85C2) 뿐만 아니라 카타란투스 로세우스 UGT2 및 아라비돕시스 탈리아나 UGT 75B1(실시예 3 참조)을 또한 크로신 생성에 대해 검정하였다.

[0155]

상기 UGT 중에서, 크로쿠스 UGT UN1671 및 UN4522 및 스테비아 UGT76G1이 크로세틴-3G를 당화시키는 능력을 나타내었다. LC-MS에 의한 예비 분석은 크로신과 동일한 분자 질량을 갖는 생성물 분자의 출현을 나타내었다. 서브페밀리 76의 UGT가 통상적으로 2개의 글루코스 모이어티 사이에 1,3 결합을 만들에 따라, 글루코스-글루코스 결합의 유형은 크로신 또는 크로신 유사체가 생성되었는지 결정하기 위한 NMR에 의해 확인된다.

[0156]

#### 실시예 5: 크로세틴 글루코실 에스테르 형성을 위한 크로쿠스 UGT2의 클로닝

[0157]

크로쿠스 UGT2(CsUGT2, GenBank Accession Number: A Y262037.1)는 카르복실레이트 위치에서 크로세틴의 2개의 일차 글루코실화를 촉매하여, 크로세틴 모노- 및 디-글루코실 에스테르를 발생시키는 것으로 생각된다. CsUGT2는 T7 프로모터를 이용하여 폴리-히스티딘 태그 융합과 함께 및 폴리-히스티딘 태그 융합 없이 박테리아 발현 벡터로 클로닝되었다. 유전자를 또한 강한 항시성 GPD1 프로모터를 이용하여 효모 발현 작제물로 클로닝하였다. 최적화된 효모 발현을 위한 유전자를 클로닝에 사용하였다. 도 7은 CsUGT2의 누클레오티드 및

아미노산 서열, 뿐만 아니라 코돈-최적화된 누클레오티드 서열을 제공한다.

[0158] 형질전환된 XJa(DE3) 자가용해 E. 콜리 K 균주를 제조업체의 프로토콜(Zymo research, CA 92614, U.S.A)에 따라 IPTG로 유도시켰다. 형질전환된 사카로마이세스 세레비지애 세포(Strain DSY5, Dualsystems Biotech, Switzerland)를 2% 글루코스, pH 5.8을 함유하는 SC 드롭아웃 배지에서 성장시켰다. CsUGT2 유전자를 갖는 DSY5 균주의 단일 콜로니를 SC 글루코스 배지에 접종시키고, 밤새 250 rpm에서 30°C에서 인큐베이션시켰다. 효모 세포를 신선한 SC 브로쓰에서 1.0의 OD600과 동등하게 신선한 배지에 재접종시키고, 추가 72시간 동안 인큐베이션시켰다. 이후, 세포를 펠렛화시키고, YeastBuster™ Protein Extraction Reagent(Merck, India)를 이용하여 용해시켰다. 세포-비함유 추출물을 10 mM UDP-글루코스(최종 농도), Chromadex(US)로부터 구입한 1 mM 크로세틴(최종 농도)을 이용하여 크로세틴 당화 활성에 대해 검정하고, 3시간 동안 30°C에서 인큐베이션하였다. 분석을 미정제 반응 혼합물에 대해 수행하였고, 모노 및 디-글루코실 에스테르의 존재가 참고문헌[J. Mass. Spectrom. 2009, 44, 1661-1667]에 따른 질량분광법을 이용하여 이들의 질량을 기초로 하여 관찰되었다.

#### [0159] 실시예 6: 크로세틴을 생성하는 효모

[0160] 크로세틴의 생성을 위한 기능적 생합성 경로를 하기와 같이 개발하였다. β-카로틴을 생성하는 실시예 1에 기재된 조작된 효모 균주(EYS886)를 사프란 생합성 경로를 조작하는데 이용하였다. C. 사티부스 제아잔틴 분해 옥시게나제(ZCO, 제아잔틴 분해 디옥시게나제 또는 ZCD로도 공지됨) 및 잔토필로마이세스 텐드로로우스 카로틴 하이드록실라제(CH) CH-2 유전자의 공동 발현은 LC 및 MS 분석에 의해 나타나는 바와 같이 크로세틴의 생성을 발생시켰다. 크로세틴 디알데하이드의 크로세틴으로의 전환을 위해 이종성 유전자는 제공되지 않았으며; 이러한 활성을 S. 세레비지애 세포에서 자연적으로 발생해야 한다.

[0161] 표 3(도 8은 최적화된 DNA 서열을 함유함)에 기재된 바와 같은 소스의 카로틴 하이드록실라제("CH") 및 제아잔틴 분해 옥시게나제("ZCO")의 유전자 유사체의 여러 조합의 발현에 높은 카페수의 pRS416 E. 콜리/효모 셔틀 벡터를 이용하였다. ZCO 유전자를 TEF 프로모터의 조절 하에서 발현시키고; CH 유전자를 GPD 프로모터를 이용하여 발현시켰다. 다음과 같은 조합을 시험하였다: CH2/ZC01, CH3/ZC02, 및 CH6/ZC04.

표 3: CH 및 ZCO 유전자의 소스

CH2	잔토필로마이세스 텐드로로우스 (진균)	β-카로틴 3-하이드록실라제 (crtS)
CH3	아라비돕시스 탈리아나(식물)	β-고리 하이드록실라제 (CHY2)
CH6	아도니스 아에스티발리스	β-카로틴 하이드록실라제
ZCO1	크로쿠스 사티부스	제아잔틴 분해 디옥시게나제 (CsZCO)
ZCO2	크로쿠스 사티부스	카로티노이드 9,10(9',10')-분해 디옥시게나제 (CsCCD)
ZCO4	크리산데룸x 모리폴리움	카로티노이드 분해 디옥시게나제 (CmCCD4a)

[0163]

[0164] ZCO/CH6 조합을 함유하는 플라스미드를 제조업체의 프로토콜(Frozen-EZ Yeast Transformation II Kit,™ Zymo research, Switzerland)에 따라 β-카로틴 생성 균주로 형질전환시켰다. 형질전환체를 2% 글루코스를 함유하는 SC Ura-플레이트(pH 5.8)에 플레이팅시키고, 3일 동안 30°C에서 인큐베이션하였다.

[0165]

양성 효모 클론을 30°C에서 글루코스를 함유하는 액체 SC Ura-배지에서 성장시키고, 밤새 진탕 인큐베이터에서 200 rpm에서 통기시켰다.

[0166]

배양물을 원심분리에 의해 놓축시키고, 신선한 SC Ura-배지에서 1.2와 동등한 OD로 재현탁시켰다. 세포를 추가 72시간 동안 200 rpm에서 30°C에서 추가로 인큐베이션시켰다. 이후, 세포를 펠렛화시키고, 분석을 위해 추출물을 제조하였다. 펠렛을 저온 PBS 완충액(10mM; pH7.2)으로 2회 세척하고, 2ml의 메탄올:PBS 완충액(3:1) 중에 혼탁시키고, 밤새 -18°C에서 저장하였다. 이러한 혼합물을 해동시키고, 3분 동안 10,000 rpm에서 원심분리시키고, 펠렛을 3 ml의 클로로포름:메탄올(1:2)과 함께 볼텍스 혼합기를 이용하여 재추출하였다. 이러한 혼합물을 2분 동안 10,000 rpm에서 원심분리시키고, 상층액을 HPLC에 의한 분석을 위해 주입하였다. 유사한 방식으로, 상층액을 제공된 순서로 클로로포름, 메탄올, 및 물로 추출하고, HPLC에 의해 분석하였다.

[0167]

#### 분석

[0168]

세포 추출물을 1 ml/분으로 40분의 기간에 걸쳐 1% 아세트산 및 물 중 60% 내지 100%의 선형 메탄올 구배와 함께 C18 Discovery HS HPLC 컬럼을 이용하여 분석하였다. 440 nm 흡광도에서의 일차 분석과 함께 Shimadzu SPD M20A Photo Diode Array 검출기를 지닌 Shimadzu 분취용 LC 8A 시스템을 이용하였다.

- [0169] C. 사티부스 ZCO1(GenBank accession number AJ489276, GenBank protein ID CAD33262.1) 및 X. 텐드로로우스 CH-2를 함유하는 재조합 균주 중 하나의 분석은 새로운 화합물의 생성이 크로세틴 및 크로세틴 디알데하이드의 표준과 동등한 시점에서 용리되는 것을 나타내었다. 상기 효모 균주에 의해 생성된 세포내 대사물을 GC-MS 분석에 추가로 적용시키고, 크로세틴 및 크로세틴 디알데하이드의 질량을 확인하였다.
- [0170] ZCO 및 CH의 다른 조합이 또한 용해성 단백질 발현에 적절한 조건하에서 가능적일 것이 예상된다.
- [0171] 이들 데이터는 효모가 글루코스로부터 크로세틴 디알데하이드를 제조할 수 있고, 효모가 적어도 일부의 크로세틴 디알데하이드를 크로세틴으로 산화시킬 수 있는 효소 활성을 갖는 것을 입증한다. 추가로, HBC는 ZCO 반응의 부산물이므로, 효모는 또한 HBC를 생성시킬 수 있다. 상기 기재된 UGT 및 CsUGT2의 첨가와 함께, 효모가 또한 피크로크로신 및 크로신을 생성할 것으로 예상된다.
- [0172] **실시예 7: 크로세틴 에스테르를 형성시키는 글리코실트랜스페라제 효소의 발견**
- [0173] 크로세틴이 2개의 별개의 UGT를 포함하는 다단계 경로에 의해 효소적으로 글루코실화되는 것이 제안되었다. 하나의 UGT는 모노글루코실- 및 디글루코실-에스테르의 형성과 함께 크로세틴의 말단 카르복실 말단으로의 글루코스 모이어티의 첨가를 촉매한다. 다른 UGT는 글루코스 모이어티를 글루코실기로 전달하여 크로세틴 모노겐티오비오실- 및 디젠티오비오실에스테르를 형성시킨다.
- [0174] 다음과 같은 UGT를 크로세틴으로부터의 모노, 디 또는 젠티오비오실 분자와 같은 크로세틴 에스테르의 형성에 대해 스크리닝하였다: 스테비아 래바우디아나(88B1, 76G1, 74G1, 912D2e, 및 85C2, UGT73) 및 2개의 아라비돕시스 UGT71 하이브리드 효소(71C125571C2 및 71C125571E1).
- [0175] 상기 UGT를 엔코딩하는 유전자를 T7 프로모터 하에서 플라스미드로 클로닝하고, 발현 연구를 위해 E. 콜리 BL21(Autolysis: XJb(DE3), Zymoresearch) 세포로 형질전환시켰다. 상기 UGT를 갖는 균주를 0.1 mM IPTG로 유도하고, 유도된 배양물을 밤새 20°C에서 성장시켰다. 이후, 유도된 세포를 동결 및 해동 방법에 의해 용해시켰다.
- [0176] UGT 검정을 글루코스-수용체 기질( $10 \mu\text{M}$  최종 농도)로서 크로세틴 및 공여체로서 UDP-글루코스( $1 \text{mM}$  최종 농도)와 함께 인큐베이션된  $98 \mu\text{L}$  유도된 정화 용해질을 갖는  $100 \mu\text{L}$  반응으로 수행하였다. 반응을 3시간 동안 30°C에서 수행하고,  $300 \mu\text{L}$ 의 물 포화 1-부탄올의 첨가로 종료시켰다. 샘플을  $300 \mu\text{L}$ 의 물-포화 1-부탄올로 3회 추출하였다. 푸울링된 부탄올 분획을 Speed-vac에서 완전히 건조시키고, 다음과 같은 방법을 이용하여 LC/MS에 의해 분석하였다. 기계: Agilent 1200 HPLC & Q-TOF LC/MS 6520, 컬럼: c18 reverse Luna,  $4 \mu\text{m}$ ,  $4.6 \times 150 \text{mm}$ , 주입 부피:  $20 \mu\text{l}$ , 이동상: 2성분의 아세토니트릴(B):물(A)(0.1% HCOOH), 유량:  $0.8 \text{ mL}/\text{분}$ , 작업 시간: 20분, 검출: 440nm, 구배: 5분 동안 20% B, 15분 동안 80% B, 20분 동안 80% B, 이온 소스-Dual ESI, 획득 모드-MS, 질량 범위-100-1500, 모드-음성 모드.
- [0177] 이들 중, 3개의 UGT(스테비아로부터 76G1, 및 2개의 아라비돕시스 UGT71 하이브리드 효소)는 크로세틴의 글루코실화를 촉매하여 모노 및 디 글루코실 에스테르를 형성시킨다. 2개의 아라비돕시스 UGT71 하이브리드 효소(71C125571C2 및 71C125571E1)가 또한 크로세틴 젠티비오실 에스테르를 형성하는 능력을 나타내었다. LC-MS에 의한 예비 분석은 모노, 디 및 젠티비오실 에스테르와 동일한 분자 질량을 갖는 생성물 분자의 출현을 나타내었다.
- [0178] **실시예 8: 크로쿠스 사티부스로부터 글리코실트랜스페라제를 형성시키는 크로세틴 모노 및 디 글루코실 에스테르의 발견**
- [0179] 실시예 4의 피로시퀀싱 데이터는 또한 변이체 크로쿠스 UGT, Cs VrUGT2를 나타내었다. 도 9는 Cs VrUGT2의 아미노산 서열을 함유한다. 변이체 UGT의 서열을 BLAST를 이용하여 크로쿠스 UGT2(CsUGT2, GenBank Accession No.: AY262037.1)와 비교하였다. 도 10은 크로쿠스 사티부스로부터의 CsUGT2 및 변이체 Cs VrUGT2의 정렬 뿐만 아니라 각각의 폴리펩ти드의 아미노산 서열을 함유한다. BLAST 분석을 기초로 하여, 유전자/대립유전자 특이적 역 PCR 프라이머를 크로쿠스 cDNA 라이브러리로부터 전장 유전자를 분리시키도록 설계하였다.
- [0180] Cs VrUGT2를 엔코딩하는 코돈 최적화된 누클레오티드 서열을 T7 프로모터 하에서 플라스미드로 클로닝하고, 발현 연구를 위해 E. 콜리 BL21(Autolysis: XJb(DE3), Zymoresearch) 세포로 형질전환시켰다. Cs VrUGT2를 갖는 균주를 0.1 mM IPTG로 유도시키고, 유도된 배양물을 밤새 20°C에서 성장시켰다. 이후, 유도된 세포를 동결 및 해동에 의해 용해시켰다.
- [0181] UGT 검정을 글루코스-수용체 기질( $10 \mu\text{M}$  최종 농도)로서 크로세틴 및 공여체로서 UDP-글루코스( $1 \text{mM}$  최종

농도)와 함께 인큐베이션된 유도된 배양물로부터의 98  $\mu\text{L}$ 의 정화 용해질을 이용한 100  $\mu\text{L}$  반응으로 수행하였다. 반응을 3시간 동안 30°C에서 수행하고, 300  $\mu\text{L}$ 의 물 포화 1-부탄올의 첨가에 의해 종료시켰다. 샘플을 300  $\mu\text{L}$ 의 물-포화 1-부탄올로 3회 추출하고, 분획을 푸울링시켰다. 푸울링된 부탄올 분획을 Speed-vac에서 완전히 건조시키고, 다음과 같은 방법을 이용하여 LC/MS에 의해 분석하였다. 기계: Agilent 1200 HPLC & Q-TOF LC/MS 6520, 컬럼: c18 reverse Luna, 4  $\mu\text{m}$ , 4.6x 150mm, 주입 부피: 20  $\mu\text{l}$ , 이동상: 2성분의 아세토니트릴(B):물(A)(0.1% HCOOH), 유량: 0.8  $\text{ml}/\text{분}$ , 작업 시간: 20분, 검출: 440nm, 구배: 5분 동안 20% B, 15분 동안 80% B, 20분 동안 80% B, 이온 소스-Dual ESI, 획득 모드-MS, 질량 범위-100-1500, 모드-음성 모드.

[0182] LC-MS에 의한 예비 분석은 모노 및 디 글루코실 에스테르와 동일한 문자 질량을 갖는 생성물 문자의 출현을 나타내었다.

[0183] 실시예 9: 크로세틴 디알데하이드를 크로세틴으로 전환시킬 수 있는 내인성 효모 알데하이드 데하이드로제나제의 발견

[0184] 사프란의 색은 주로 크로세틴의 연속적 당화로부터 유래되는 카로티노이드 글리코시드로 인한 것이다. 사프란 생합성 경로에서 중요 단계 중 하나는 크로세틴 디알데하이드의 크로세틴으로의 산화이다. 상기 전환을 수행하는 사카로마이세스 세레비지애에서의 내인성 알데하이드 데하이드로제나제의 능력을 시험하였다. 효모 유전체는 5개의 공지된 알데하이드 데하이드로제나제 코딩 유전자(ALD2 내지 ALD6) 뿐만 아니라 알데하이드 데하이드로제나제인 것으로 예측되는 추가 유전자, HFD1을 갖는다. 참조 균주 S288C로부터의 ALD2, ALD3, ALD4, ALD5, ALD6, 및 HFD1을 엔코딩하는 누클레오티드 서열(SEQ ID NO:67-72)에 대해 도 11을 참조하라. 서열은 참조 균주 S288C에 대한 것이다. 사용된 균주 내의 유전자 서열에서 약간의 변화가 존재할 수 있다. 밤새 성장된 효모 배양물로부터 세포 비함유 추출물을 제조한 후, 기계적 용해에 의해 분쇄시켰다. 용해질을 정화시키고, 표 4에 기재된 바와 같이 수행된 시험관내 반응에서 크로세틴 디알데하이드를 크로세틴으로 전환시키는 이들의 능력에 대해 시험하였다. 임의의 전체 세포 추출물이 없는 음성 대조군을 또한 포함시켰다. 반응을 60분 동안 25°C에서 수행한 후, 3 부피(1500  $\text{ml}$ )의 물 포화 부탄올을 첨가하여 중지시켰다.

[0185] 표 4

성분	최종 농도	반응 당 양
1M Tris-HCl pH7.5	100mM	50 $\mu\text{l}$
1M KCl	100mM	50 $\mu\text{l}$
0.5M MgCl <sub>2</sub>	3.75mM	3.75 $\mu\text{l}$
1M 2-머캅토에탄올	10mM	5 $\mu\text{l}$
10 mM 크로세틴 디알데하이드	200 $\mu\text{M}$	10 $\mu\text{l}$
20 mM $\beta$ -NAD	0.67mM	16.7 $\mu\text{l}$
세포 비함유 추출물		50 $\mu\text{l}$
물		314.55 $\mu\text{l}$
전체		500 $\mu\text{l}$

[0186]

[0187] 유기상을 원심분리에 의해 분리시키고, 진공 건조에 적용시킨 후, 이들을 질량분석법과 커플링된 고성능 액체 크로마토그래피(LC-MS)에 의해 분석하였다. 5 마이크론 가드(guard) 컬럼이 장비된 Luna C18 5  $\mu\text{m}$  컬럼(4.6 x 150  $\text{m}$ )과 함께 Agilent 1200 HPLC & Q-TOF LC/MS 6520을 사용하였다. 이동상은 0.8  $\text{ml}/\text{분}$ 의 유량의 아세토니트릴(B)(0.1% 포름산(HCOOH)): $\text{H}_2\text{O}$ (A)(0.1% HCOOH)였다. 작업 시간은 통상적으로 1분의 후속 작업과 함께 15분이었다.

시간	용매 비 B
4	70
10	80
12	90
15	90

[0188]

[0189] MS 파라미터는 다음을 포함하였다: 이온 소스로서 ESI, 이중 ESI 획득 모드; 100–450 Da 질량 범위; +/- ve (신속 극성 스위칭) 모드.

[0190]

효모 내인성 알데하이드 데하이드로게나제(들)은 LC-MS 결과에 의해 입증되는 바와 같이 크로세틴 디알데하이드를 크로세틴으로 전환시킬 수 있었다.

[0191]

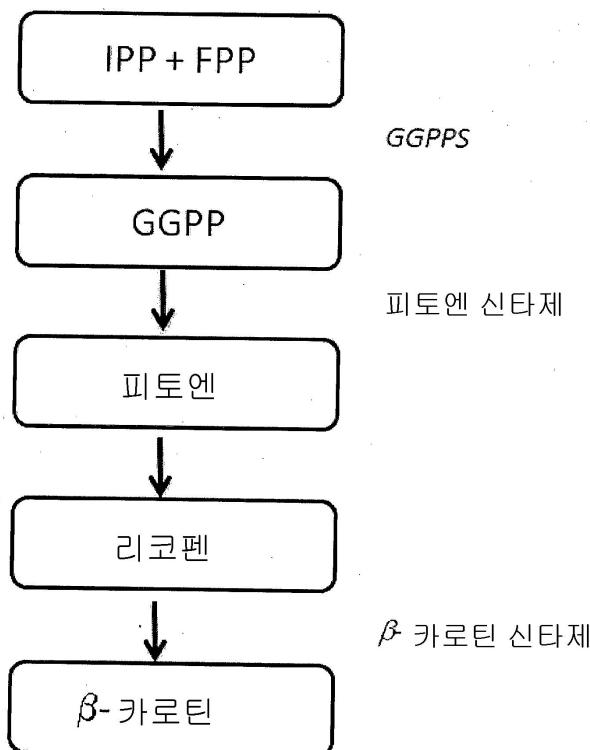
### 다른 구체예

[0192]

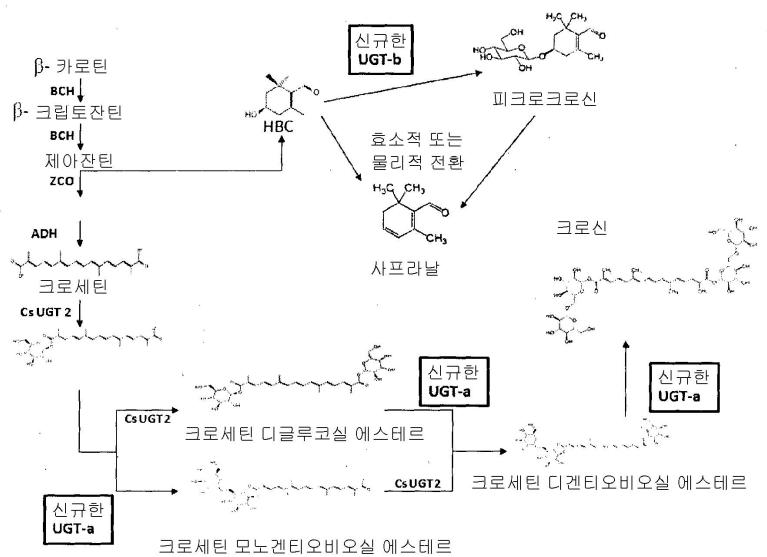
본 발명은 이의 상세한 설명과 관련하여 기재되었으나, 상기 설명은 예시를 위한 것으로, 첨부된 청구항의 범위에 의해 규정되는 본 발명의 범위를 제한하는 것이 아님이 이해되어야 한다. 다른 양태, 장점, 및 변형이 하기 청구항의 범위 내이다.

## 도면

### 도면1

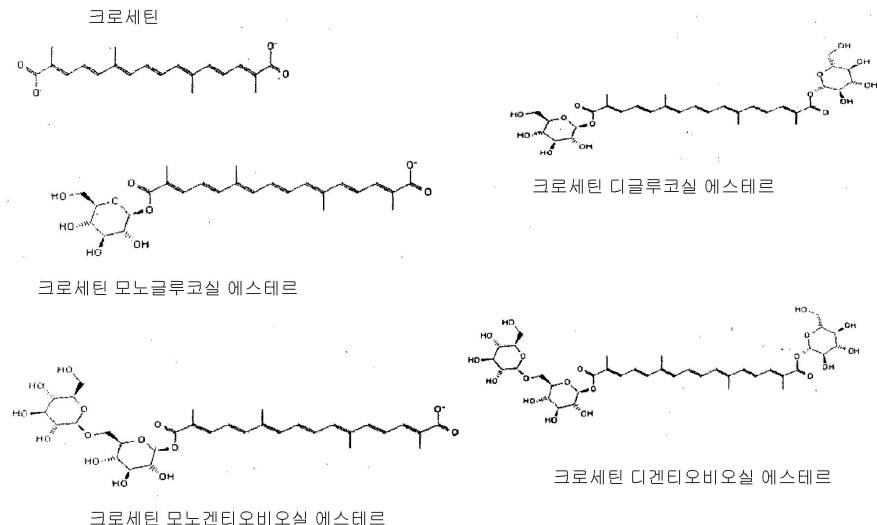


## 도면2a

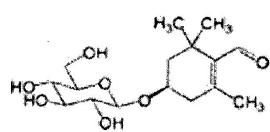
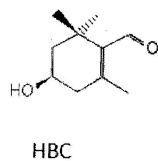


## 도면2b

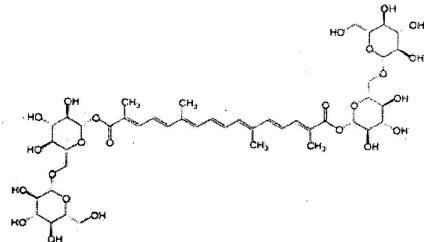
## 구조



도면2c



크로신



## 도면3a

스티비아 레비우디아나 88B1, 76G1, 74G1, 91D2e, 85C2

카타린투스 로세우스 UGT12 및

아란버드시스 틸리아나 UGT75B1

스티비아 레비우디아나 UGT88B1 누클레오티드 (SEQ ID NO: 01)

```

ATGGAGTCCTTAAGGTGATCCTGTATACCTTCTCCGGAAATCGGCCATCTGGTTCCATGGGGAGCTGGAACACTCATCCACACCCACCCCTCACTCTCGT
TATCATCTCTGTAACCTGGCTPACATATAAGAACCGGGCCACCACTACATCACACCGCTCCACCCCTTGATCATCTTCACCTCCGCTTAACCTCCGCTCGT
TCCCTCTCACCAGAGATCATCTTGAAATCATGGCTCATGACCTTGCCCTGATATCCCTCAACTTACAACACCTGGAAACTTAAAGATCTAGATCTTCAATTAATTCCTGGGTACCTCCATTCATTCT
TCTCTGIGGGTCTCATCTCCAAACCATCTACAAACATATTCGGAAACTTAAAGATCTAGATCTTCAATTAATTCCTGGGTACCTCCATTCATTCT
CCGATAIGCCCACAGTTTGTGATAGAACCTACATCGGAGGTTAATCTACAAAACACTCTGTAACAAACCTCAATACATGGCAAATCTCCGGAGTCATGCAACAGCTC
TTGGAGTTGAGGAAAGCTGGTGTGAGGAGACTGCTCACTGGCAAGAGTAACTCTCCGAGAGTGTAACTCATCACGGACGCTCCCTCCACCTATTCTAATCGGCCTTATGCCATCGGGAA
TCAAGTTGATCATACGAAACAGGTGTTAAATGGCTGACACACACCTAGTAAGTGTAGTGTGTTGTTGCTTGGGGGGGCTGACGGGCTCAACAGCAGCTGCTG
AATGAGAGAAATAGCGGTGTTGAGGAGACTGCTCACTGGCAAGAGTGTGTTGGGGGGGCTGACGGGCTCAACAGCAGCTGCTG
CTTCCTGTAGGGTTGAGGAGACTGCTCACTGGCAAGAGTGTGTTGGGGGGGCTGACGGGCTCAACAGCAGCTGCTG
TCATGCGGGTGGACTTGGTGTGAGAATGTCAGGGGGTTGTCAGGGGGTAGAGGAGACGGGTGAGAAGCTGAGTTGTTGAGG
AAATAAGGTGGACTTGGTGTGAGAATGTCAGGGGGTTGTCAGGGGGTAGAGGAGACGGGTGAGAAGCTGAGTTGTTGAGG
CGGATTGAGATGAGTACAAAGCCAAAGGTGCGGGGAGGAGCGGGGTTCTCGAGTTGATTTCTCAATTAATCTGAGTCATGGACCCACAGTGA

```

스티비아 레비우디아나 UGT88B1 단백질 (SEQ ID NO: 02)

```

MESSKVILYPSPGTGHLYSMVLEGLKLTHPRSLSVLVIPATYETGSTTYINTVSTTPFTFHILPVTLPLPDSSSEFTIDAFDIPOLNPVYNTVALSET
STIKAVTILDFVNAAFQFLPSKSLDPLPTYFFSGASGICAFHLPLTYKTYSGNPKDLDFINIPGVPPVHSIMPTVIFDKRSNSYRNFKTSNNMAKSSGVIANSF
LQBERRAQOTLRIGKSTDGPSPPIYLIGPLIAGSNQVDNECNEIWKWLTOPSKSVWLFMCFGSOGVKKEOLRAGLERSGORFLWVRKPSPDGKEFGFLDV
LPEGFVARTKEKLGIVKINWAPLSPALIGHESYGFVSHGMNSLIAVVEPMATPLAEQKNRUVLVEELKVAWLMADGEVSAAVEETVRLQMDGRYRE
RILEMSTRAKAAYEDGSSSRVDFEFLTESWTHK*

```

## 도면3b

스테비아 세미우드이나 UGT76G1 ₐ ॥ 회오리드 (SEQ ID NO. 3)

ATGGAAATAAACGGAGACCACCGTTGCCGGGAGAATAATTTCACACCAACTTCACAAACCAAACATCTAATACCTC  
 TCGCTAGCCAATGTGTACTCCAAAGGATTCACTTACGATTCACCATTCACCTTCACACCAACTTCACAAACCAAACATCTAATACCTC  
 ACTTCACTTCAGATTCACTTCACCTTGACAAACGAGCTGAGGACTGAGAAATTAGACGCGAACTGGAACTGTTGATGTTGCTGCGG  
 ATTCCTGATTATCAAGAACACGGAGCTGAGGACTGAGAAATTAGACGCGAACTGGAACTGGAACTGTTGCTGCGG  
 GTGTTAATCGCCGATCAGATTGGTACTTCAGTCACTTCAGTCACTTGAGGTTACCTGAGTGTGAGCTTCAACCTCCAGGGCTGTTGGTACAGCAGCT  
 TGTTAATTTCTGAGTACACCACTCCCAGCATTGACAGCATCTCCAGCAGCTTACAGACATTCAGCAGCTTACAGACATTCAGCAGCT  
 GCGAGTGGTTCTGAGTACAGTAAAGGATATCAAGTGTGAGCTTCAAGTGTGAGCTTCAAGGAGTCTGAGCTTGGAGAGGATAC  
 GAAACAAACAAAGCATCTCAGGAGTCAAGACTCAAGAACCCGGTTGGAGAGAACAA  
 CGGCCTCCAAGTTCTGAGTACACCACTCCCAGCATTGACAGCATCTCCAGCAGCTTACAGACATTCAGCAGCTTACAGACATTCAGCAGCT  
 TTAGACCAACAACCGTACGTTGGTACTGTGAGTTGGAGTACTGAAGTACTGGATGAGAAAGGATTCTGGAAATAG  
 TCGTGGTTGGTCTGAGTACAGCAGTCGTTTATGGGTTGGTGCACCTGGGTTGCAAGGGTTGACGGTGGTGGACCCGGTTGCCAG  
 ATGGGGTCTGGGTGAAGAGGAGCTATGTCGAATGGGTTCCGGAGCAGAAGAGTGTCTAGCTAGCAGCTTCCGGAGATAGC  
 CATAGCGGATGGAACHTACGTGGAAAGGTTGAGGTTGGGGTGTATTGGAAAGGTTGGGCTGAGGAGTCTGGACT  
 TAGATACATGAGGATGTTGAGGTTGGGGTGTATTGGAAAGGTTGGGCTGAGGAGTCTGGACT  
 TGGTGGGATGAGGAGGAAATAGACAGAAATGCAAGGAGTGTGAAACAAGGAGATAGCAATAGAGAGGTTGAGGTTGATC  
 TACGAATCATAGAGCTCTAGTTCTTACATTCTCATCGTTGAA

스테비아 세미우드이나 UGT76G1 ₐ ॥ 회오리드 (SEQ ID NO. 4)

MENKTETTVRRRRRILIFPVFPQGHINPMQLANLVLYSKFSITIFHTNFNPKTSNYPHTFRFLNDPQDWRIISNLPTHGLAVMR  
 ILINEHGADELREIELLMLASEEDGEVSCSLIADQIMWYFTOSVADSLNRLRVLTSSLNFHABVSLPQDFELGYLDPPDKTRLEQ  
 ASGFPMILKVDIKSFSMWKRYKEYFENTRKTKASSGVIVNSKELESELETFIREITAPSFLIPLPKHLTASSSSLLHDRTVFPW  
 LDOQPSRSVLYVSGSGTEVLDKFLETARGLVDKRSFLWVVRPGFVGSTWVEPLPDGFLIGERGRIVKWMVPOQEVLAHGAIWF  
 HSGWNSTLESVCEGYPMIFSDFGLDQPLNARYMSDVLKVGVYLENGWERGEIANAIRVMVDEEGEYIRQNARVLKOKADVSIMKGSS  
 YESLEISVYISSI \*

스테비아 세미우드이나 UGT74G1 ₐ ॥ 회오리드 (SEQ ID NO. 5)

ATGGCGGAACAAACAAAGATCAAGAAATCACACACGTTCTACTCATCCATTCCATACAGGCCATATAACCCCTTCATCCAGTT  
 TGGCAACGATTAATCCAAAGGTGTCACAAACACACTGTGTTACCACTTCACCTTAACCTAACCTAACACAGTACACCA  
 CCACCACTCCATCGAAATCCAAAGCAATTCCGATGGTTGTGATGAGGCGGTTATGAGTCCAGGAGATCATATTGGAAACATTC  
 AAACAAAGTGGGTCTAATCACTAGCTGACTTAATCAAGAAGCTCAGTCAAGTGAAGGAAACAAATGTGCAATCATTTATGATTCTAT

### 도면3c

스테비아 레바우디아나 UGT74G1 단백질 (SEQ ID NO: 06)

MAEQQKIKSPHILLIPPLQHINTIFQGKRILISKGVTTIHLINSTLNHSNTTTSIOAISDGDCDEGGFMSAGESILETF  
KQVGSKSLADLIKQSEGTTIDAIYDSMEEWVLWDVAIEFGJLGGSFFTOACUVNSLYVWHKGLLISLPLGETWSPGFPLQRWETP  
LTLQNHEQIQSPWSOMLFGKFANIDQARWFTNSFYKLEEVIEWENTRKWINLVIGPTFSMLDURDDDKNGDNLYKANHECMNW  
LDDKPXSVVYAFQSOLVKHGPEQEETIRALIDSVDNFMWIKHEKEGLPLNSEVITKGGLIIVAWCKQLDVIAHESVGCFVTHC  
ENSTLEAISLGWVYAMPQFSDQTTNAKULDEILIGVGRVKADENGIVRRGNLASCIRMMEEERGVUIRKNAVWKEDLAKAVHBEQQ  
SDNDIVEFVSSELKA\*

스테비아 레바우디아나 UGT91D2e 풀레오티드 (SEQ ID NO: 07)

ATGGCTACAGTGCCTAGTGACGCCGTAAGCAGCTTCATGTGCGACGTCCCATGCGTGTGTTCGGTCAACATCCTCCCTA  
CCTTCAGCTTCTGAAATTGATAGTGCTGAAAGGGTCACAAGTCTCGTTCTTCTACCAACAGAACATTCAACGCTCTCTCTCATA  
TCTCGCCACTCTCAAATGTTGTTCAACTCACACTTCCACGTGTCCAAAGAGCTGCCGGAGATGCGAGGCCAACGTGAGTCACCCCT  
GAAGATTCATATCTCAAGAAGGCTCTGATGCTCACCGGGAGTACCCGGTTCTAGAACACACTCTCGGACTGGATTAT  
TATGATTAATCTACTACTGTGTCATGCCATCACGGGTAGGGTACGGTCTCTGGTACCTCCGTACAGGCC  
TTGCTTATATGGACGCCAGTGACGCCATGATAAATGGTCAGATGGTCAAGACCGGGTACTCTCCGTACAGGCCATGGCCA  
TTGCTTATATGGACGCCAGTGACGCCATGATAAATGGTCAGATGGTCAAGACCGGGTACTCTCCGTACAGGCCAAGCTGG  
TTTCCCTTTCGACCAAAGTAGATGTGCTGGCGAGCATGATGACTGTGCGGACTGTGCGCTTACAAAGCTCCGGGATATCTGATGGATACCG  
TTATGGGGCTGGTCTTAAGGGATCTGTGTTCCAAATAGTTACCATGAGTTGGAACTCAATGGCTACCTCTTGGAGACAC  
TACACCAAGTACCGGGTGTCCGGGGGATTACTGCCCCGGAAATCCCGGAGACGAGAAGATGAAACATGGGTGTCATCAAGAAA

도면3d

TRVEYVSYOFVDYLJERNARAVADHES\*  
스티비아 비보우드나 UGT85C2 누클레오티드 (SEQ ID NO: 09)  
ATGGATGCAATGGCTACAACTGAGAGAACACACAGTCATCTTCATACCATTTCCAGCACAAAGCCACATAAGCCATGCTCAA  
ATGCCACACTTCTCCACACAAAGGACTCCAGAYAACCTTCTGTCACACCGACTTCATCCACACCGACTTCTGGAATCA  
ATTGCTTAGAGCGGTGACGGGTTCCGGTCCGGACATTCCGGATGGTTCTGGATCGTCAACAGTCGGAGGAGGAGGAG  
TCACTTGAGCGGTGACGGGTTCCGGTCCGGACATTCCGGATGGATCGTCAACAGTCGGAGGAGGAGGAGGAG  
AGATGGGTCTGTGGGTTTCACATGAGCTGCACAAAGCTGGGATTCGGTCACTGGGATTCGGTCACTGGGATTC  
AGAAGGAG  
TCATGGGTTTTACCATATTCTCATGAGAAAGGATTTCGACCACTTAAGATGCAAGTTACCTGACAAATTGGGATTGGAC  
ACCGTCATGATTGGGTCGGGAATTGGAGGCATTCGGCTCTCAAGGATTCCCGCTGGACTGGAGGACTGACCTCAT  
GATGTTCACTACGGAAGCTCTCAAAGGTCAACAAAGGTTCAACATCATATTTCACACGTTGATGAGTTGGAGG  
AAACTTTGTCATTGAGGATAATACATTCACATGGCCCACTGCAATTACTCTTGTCAATACCGGAAGGAAAGCA  
AAACTTGTCACTACGGAAGCTCTCAAAGGTCAACAAAGGTTCAACATCATATTTCACACGTTGATGAGTTGGAGG  
GGAAATACGAGTCCTCAAGTACAGTTAGTAAAGAGAACCCAGAGTTCCAGTGCTTAAGACCAATTCCGTC  
TTATGTAATTTGGAACTACAGTTAGTCAATGCTTGAAGAACATGACGGATTTCAGTGCTTAAGAACATT  
TTTGGATCATCCGATCAACTTGGTGTAGGGAAATGGCAGTTTGGCCCTGTGACTTGAGGACATATAAGAAA  
GCTAGCTGTGGTGTCAACAGAAAGGTCTTGGAGGACCCCTTCGGTTGGAGGGTTCTGACTCATGTTGGGG  
CTTGTCTGCTGGGGTGGCAATGATATGCTGGCTTATCGTGGACAGGTGACCAACTGTAGGTTATATGCA  
GAAAGAATGGGAGGTTG

TGGCTCGATGGTAAACAAAAGGCAGTGGGTACGTTGCATTAGAACGGAGGTTGGTGACCAACCGAGGTTGTTGAGTTAGC  
ATTGGGTCTCGAGCTTCTGGGTGCCATTGTTGGCTTATAGAACCCAAAAGGTCCCGGAAGTCAGACTCGGIGGAGTTGCCAG  
ACGGGTTCTGGAAACGACTCGTGACCGCTGGGTTGGCTGGAGCTGGAGTTGGCACCTCAGTTAGGAACTGAGCTAGTGGTTGTTG  
GGTTCTGTTGACTCATCTGGGTTCTGATCAATTGGAGGGCTAATGTTGGTACCCCTCAACATGCTACCGATTTGGGACCA  
ACCTCTGATGCTCGATTAACAGGGGGATCGAGATAACAAAGAATGAGGAGATGTTGCTGACCAGGGAGCTGAGTAACATTAACGAC  
TTGCTAGATCACTGAGGCCGCGTGTGGAAAGAAGGGGAGATCTACAGGCAGCCGCGAGGGAGCTGAGTAACATTAACGAC  
ACTAAGGTGAAAAGAATGTAAGCCAATTCTGAGACTATTGGAAAGAATGCGCTGGCTGCCATCGATCATGAGAGTTA

## 도면3e

KAKDWKKERATAPNSSSINTKMKYETIVLARN\*  
 ASWSUVERVLTESVGGUHGWGMLLESLISAGVMILWFOIWLDLQINCRICREWEVGHEMIRVKRUEVKLVVLEUMGGDNTR  
 캐리풀스로세우스 CsUGT2 뉴클레오티드 (SEQ ID NO: 13)  
 ATGGTTAATCGCTCATATTTCACHTCCCATTCAGGGACAGGCCATATGTTACCCGCCTAGACATGGCCAATCTATTCACHTCTGGGAGTCAGTAAC  
 ATTAAATCAACCCATCACATCCTTCCATGTACAAACCATAGAAAGGAGGAATTCCTGGAATTCATCAATTCGCCAGCTCAG  
 AAGTGTGTTAACCTGAGGAGATGGAAGTCAGTCACTGAGGAGGAATGCTCTGATGAAATTCCTGAGGAGGAATTCCTGAGGAGGAATTCCTGAGGAG  
 CAACATTCAGAGATCTGTCCTATGTCCTCTGTCCTGAGGAGGAATTCCTGAGGAGGAATTCCTGAGGAGGAATTCCTGAGGAGGAATTCCTGAGGAG  
 GTCTGCTGCTTGCCTCTCTGAGGAGGAATTCCTGAGGAGGAATTCCTGAGGAGGAATTCCTGAGGAGGAATTCCTGAGGAGGAATTCCTGAGGAG  
 AAATTAATTAACCGAGAACAAATTTCACACATACAGAAAGGGAATATGAGTCAGATTTACCAAATGCTGAGGAGATGAGAATCACATCTAC  
 GGAGTGTGAGTCATAATGTTCTATGAACTTGACCATGATGCCATTATACATCACAGTTTGGAGAAAGCATGGATATTAGGGCTTTTGCTTGTAA  
 CAAATACAGCTGAGATAACCCAAAGGGGAAAGAAATTCAGCAATTGCAATTGAGTCAGGAGATGAGAATGAGAACAAGAAAGGGTAATTATAAGGGAT  
 TCTGTTCTGGAGATGGCCAAATTAAATTCTGCCAAATTACAGAAATGCAAGCCTGTAATCTCCGCAAATTCTCGTATTAC  
 GAGCAGAGAACACTTCAAATGTTCCAGAAGGAGATTGAGTCAGGAGATGAGAACAAGAAAGGGTAATTATAAGGGATGGCACAACCTAATTCTGACA  
 CGAATCAGTAGGAGCATTGTTACCATGTTGGTGGAAATTCACTCTGAGGAGATCTGCAAGCCTGTAATCTGGGTCTGGCTCGCAATGAGTAGGTT  
 TCAATGAGAAGATTGATGAGGTACTGAAACGGGATAGGGAGTTGGAGTGGGCTGAGTGGGAGATGAGTAGGTTTCAACAGAGATATTAAGGGAAAG  
 GCTTAAATGAGGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATG  
 TCGTGTACCTTACTGCTCTTATGAGGAGATGGGGATATGTTCTCACTGAGAAGAAAGGGTAATTATAAGGGATGGCACAACCTAATTCTGACA  
 CGAATCAGTAGGAGCATTGTTACCATGTTGGTGGAAATTCACTCTGAGGAGATCTGCAAGCCTGTAATCTGGGTCTGGCTCGCAATGAGTAGGTT  
 TCAATGAGAAGATTGATGAGGTACTGAAACGGGATAGGGAGTTGGAGTGGGCTGAGTGGGAGATGAGTAGGTTTCAACAGAGATATTAAGGGAAAG  
 GCTTAAATGAGGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATG  
 TCGTGTACCTTACTGCTCTTATGAGGAGATGGGGATATGTTCTCACTGAGAAGAAAGGGTAATTATAAGGGATGGCACAACCTAATTCTGACA

카터린트루스로서우스 CSUGT2 누클레오티드 (SEQ ID NO: 13) TAKDMKEKARATAPNGSSNINTDKMVKELITVLARN\*

MDAMATTEKKPHVIFIFPPAQSHIKAMKILQQLHHKGLOITFVNNTDFIHNOFLESGPFCILDGAPGFRFETIPDGVHSPEASIPIRE  
UGT85C2 (SEQ ID NO: 10)  
SSBIAI-SMBI-SSBIAI  
GGCTTGAGATGGGACCAAAGTGAACAGGAGAAGTCAGGGGCTGACAAATGGTCAGGAAT  
CACCGTGTAGCAAGAACATAG  
AAGGCTTAAGATTTGGAAAGAAGGCTTGGCATAGCTCTTACGGTICATCTCTTGAACATAGACAAATGGTCAGGAAT

도면3f

카티린투스-로세우스 CsUGT2 단백질 (SEQ ID NO: 14)  
 MYNQDHLIFNFPFMAQGHMLPAIDMANLFTSRGVKVLITTHQHVPMFHKSTERSRNNSGFDISIQSIFPASEVGLPEGIESLJQVSGDD  
 EMI.PKFMRGVNLQQPLEOLIQESRPHCLLS.DMFPWTTESAQAFGIPLIFHGSCS.FALSAE.SVRNKP.FENVSTDTEEFVVPDLPH  
 QTKLRTQISTYRENTESDFTKMLKKVRDSESTSYGVVNSFYLEPDYADYYINVLGRKAWHIGPFLCNKLQAEQDAQRGKSAID  
 ADECLNWLDSKQPNNSVIYICFGSMANLNSAQLHEIATALESSGQNFTWPKVDEENSSKWPEGEEFRTEKGLLIKGMAPQTLIE  
 HESTYGA.FVTHCGWNSTLEGICAGVPLVTWPF.FAEOFNEKLITEVLKTYGYGARQWSRVSTEIKGEAIANA.INRVMVGDEAVEMRN  
 AKDLRKARRKALEEDGSSYRDITLALIEELGAYRSOVERQQD\*

아라비들시스 블리아나 UGT75B1 누클레오틴 (SEQ ID NO: 15)

ATGGGGGCCACCGCATTTCTACTTGGTAACGTTCCGGCCAAAGGTACCGTGACGGTCACTCCTCCGGCTCATCAAAG  
 AACCGGGCCACGGTGTCACTTGGTCACTTGGTCTCTCCACACTCATGATCGCAACACACAAGTCGAAACAAACTCTCT  
 TCCTTACTTCTCCGACGGTTGACATGGGACTCTACGAGCTACTAAGATGGTGAATCTCAGGTTACGAGATCTCCACACTCAT  
 GATAAGGCACTATGGTAGATTTCAACTCCCTCCGCTCTCTCTGGATCCACCGGCTTGGTTCAACATCTTACACTCAT  
 TCCAAAGTAGCAGTAGCTGGTTGGTTACCTATCTGCTTCTCTGGAAATCAGAGATCTCCACACTTCCACACAAC  
 TGGAACACAGTCGCGTTGGTTACCTATCTGCTTCTCTGGAAATCAGAGATCTCCACACTTCCACACAAC  
 AAAGCGCATACGATGGGTTCAGAAATGATGGAGTTCTCATAAAGAACACAAACCGAAATTCATCACACTTCGATTCGCT  
 GGAAACAGAGGCCCTAACGGTTCCGGATATCGATATGGGGGGTTGGTTACTCTCCACGGAGATTCTCAGGAGGACCA  
 ACAAAATCAGTTAACAGTCAAAAGTAGTAGTAGTACACTGGCTAGACTCGAAAACAGAACGACCCGTTGGGTATAGTAA  
 ATGGTTGAGTTGCTCCAAGAACAGAGGAGAACAGAACGAGGATGGAGATAGCTGGATTCAAGACGAGGATGGCTGAAGAGGTGG  
 ATCCAAACAGAACGAAACAGAGGAGAACAGAACGAGGATGGAGATAGCTGGATTCAAGACGAGGATGGCTGAAGAGGTGG  
 TGATGTTGCTGGTTGGAGATAGGTTTAAGTCACCGAGCCGGTAGGTGTTGTGACTCAATGGGGAGCTGACGCTG  
 GAGAGTTGGTTGGCTTGGCTTCGGTTGGCTGTTGGATCACCGACGACCGGAAGCTACTGGAAAGAACGAGTGG  
 GACTGGTGGAGGTAAGAGAGAACAGGATGGTTGGAGAGGAGGAGATCAGGAGGTGTTGGAAAGCCGGTGTGGAGGAGAAGTGG  
 CGGTGGAGGTGAGGAAACGCAAGAACGAAATGGAAAGCTTAGCGATGGAGAACGGGTAGAGAACGAGGGAGATCTGGATAAGACATGG  
 GCTTTGTTGGAGGATATTGTTGGAGAATCTCTTATCAAACACTTGTGTAAGCAGAGGAGGATCTGGATAAGACATGG  
 AAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA

아라비들시스 블리아나 UGT75B1 단백질 (SEQ ID NO: 16)  
 MAPPHLLVTFPAQGVNFSLRFARRLKRIGARVTFTCVSVFHNMSLANKVKENLSELTFSDFDDGGIISTYEDOKRSVNLKVGDKAUSDFLEATKNGDSPV  
 TCLIYIILWWAPKRYARRFOLPSALMIPOLAVENIYTHEMNKSFELPNLSSLETRDULPSFLPSNINKGAYDAFOEMMEFLIKEKRPKLINTFDSLPEALI  
 APPNIDOMAVGPLPTEIFSGSTNKSVDQSSYTMILDSTKEESSVIVYSEGMVELSKOIEELARALIEGKRPFLWVITDKSNRETKTEGEETEIEKLAGFRHE

## 도면3g

LAMEAGREGGSSDRIMEAVFEDICGESIQLNCEAEVVKV+  
LEEVGMIVNSUQSJELSHKAVGFVHUGWSSTLESILVUGPVVAEPMWSDQPTNAKLUESWKGVRUKNUGLVERGEIRCRULEAVMEEKSVELRENAKKKR

스테비아 UGT73 누클레오틴 (SEQ IB No: 11)

스비아이 UGT73 단백질 (SEQ ID NO: 12)  
MARVDATNHLHFVPLPFLPFGHMLPMVLLAERGSTVITTPLMANEFKPVITARAIIKDRKIQLELKIPSTEGLPECCENFDMIESAQFFHKMFEATYKLAEPDVNVAYORLTHPPSCTIADNLLPWNTNDIAQKFKIPRIVFHGPFCFTILCITHAMNSNVYDIGSDSERILIPGLPDRIELTKGOALSWRGKDTKEASFWNRWÖRDIDFANGVWNSFHAPLEYPYVYELAKVKGKRWVIGPVLSICNKSFDIAERGNGKADEDEHCKWLDSMSRSVIFVCLGSLVYRGVGTQNIDIGLGLEASKPFWCLRHTTEEFERMWSEQGYEERVKDRGLITRWPAGFELSHRAGGELTHCGNINSTLEGITAGPMVWTMPQTDFQFINERFIVDVLKIUVKGGMVEPVVVGQDGDKFGVLVHTKKEETRSFEDIMGEEGEEGTRRRSRLRDMAKSAMEDGSSHRDMTSMIQDIVECKNR\*

## 도면3h

A  
CTGAAGGAGATTCGGGAGGGGGAAAGAGGCTGTGATTGGACGGTGGATCTTCGTTTGTTGCGGTTAAAGATTCAATAGTGGGTTGGGTTGG  
아리아시스 UGT71 하이브리드 흡수 1:7C125571C2 단백질 (SEQ ID NO: 18)  
MGKQEDAEILVILPFPFSGHILATEILAKRLISQNPRINTITLTYWGLPFPTQADTAIFRLSVKNEPRLVLTPEVQDPPPMELFVEE  
FAESYIILEYVKMVPIIREALSTLSSRSDEGSVRYAGIVLDFFCVPMIDGNEFNLPYIIFTCSAGFLGMKYLPERHREKSEFNR  
SENFEINLIPGYVNNSVPTKVLPSGLFMKETYEPWELAERFPEAKGLVNNTYALEPNGFKYFDRCPDNYPTIYPIGPLIICSNDRPNLID  
LSEDRRIKWLDDQPESSVVFLCFFGSIKSLAASQIKEIAQALLVELVGIRFLMSIRTDPKEYASPNELIIPDGFMNRVMGLVLCGWAQVEE  
ILAHKAIGGFVSHCGMNSILESLRFGVPIATWMPYAEQQLNRAFTIVEKELGALEMRLDYVSEYGETVRADELAGAVRSLMDGEDVPRRK  
LKEIAEAGKEAVMDGSSSFVAVKRFDGL\*

아세틸시스 UGT71 총이온화도 1: 71C125571C2 누출오염 (SEQ ID NO. 17)

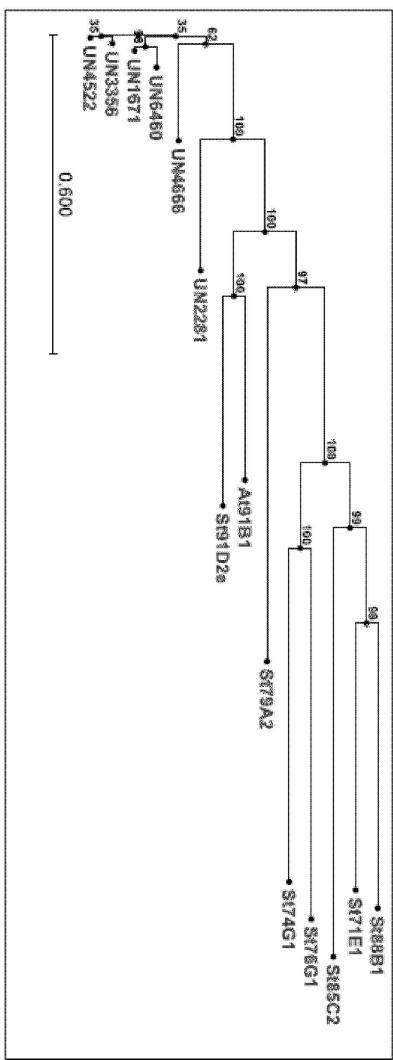
## 도면3i

TCCGATCCCTAGTCAAAATGAGCTCGTATCCGCTCGTTACGTTGCCGGAAAGCTCCACCAATGGAACTCTTGTGGAA  
 TTTGGCGAATCTTACATTCTGATAAGCTCAAGAAATGGTCCCACATCAGAGAAGCTCTCCACTCTTGCTCTCCGGATGA  
 ATCGGGTTCACTGAGGAGGTGAACTCTCATTCCTGTTATGTCACAGTAGGTTTGCCTCTGATGATGATGAGGTTAACCTCCCTCT  
 ACATTTCTTGACGTGTTAGGGCTGGGTTGGTACTGTCAGGAGACACCGGAAATCGGAATCGGAATCAGCTCCGGGGATGA  
 AGCTTCACGAGGAGGTGAACTCTCATTCCTGCTTGGGTTGGTACTGTCAGGAGACAGGTTTCCCTGAGGTTAACCTCCCTCT  
 CTACGGAGCCTGGGTGCGAACTAGAGAGGTTCCCTACTAAGGTTTGGTAAATTCAACAGCTCTCGAGCCAAACGGTT  
 TTAATATTCTGATCGTGTGTCGGGATAACTACCCAAACATTACCCAACTCGGGCCATTGGAACTTGAACCTTGAACAAACAAAGAGGATGCT  
 AAAACGAGGAGATATGAGGTTGGTGGTTATGGTGGAGGATAGATTATGGTCGCTCGTGGTGGAGGTTGGAGGATGCT  
 GAAACGAGTGAAGGAGATGGGGTGGCATGGAGCATGGAGGATGGTGGAGGATGGTGGAGGATGGTGGAGGATGGTGGAGGATGCT  
 TAGAGTTTCCGAAGAATATGAAACTGGAGAAGTCTCCAGTGGGGTTGGTGGAGGATGGTGGAGGATGGTGGAGGATGGTGGAGGATGCT  
 TGGGCCAACAAATGGGGTTGGTGGTGGAGGATGGTGGAGGATGGTGGAGGATGGTGGAGGATGGTGGAGGATGGTGGAGGATGGTGGAGGATGCT  
 GTGTGGGGTCCGATGGCAGCTGGCATTATATGCTGACACAACGTTGAAATGCTTCTACTGTGGTGGAACTGGGATGGGATTAGG  
 AGATTAGGATGGATATCGGAGGATAGGAAGGGGTATGACGGTGGGGATGGGATGGGATGGGATGGGATTAGGATGGGATTAGG  
 AAGTGTGATGAGTGTGATGGTGGAGATAGGAATAGGTAAGGAGATGGTGGAGGAGCTGGGTTGGTGGAGGATGGGATTAGG  
 CGCATCCATGGAAATTCTCGAGCATGATCGAATGTTACGATTAA

이전페이지     UGT71     하이브리드     2. 71C12557IE1     단백질 (SEQ ID NO. 20)

MGKOEAEELVITIPFPFSGHILATELAKRILISODNPRHHTITLYWGLPFLIQADTAFLRSVLNEPRIRLVTLPEVQDPPEMLFVE  
 FAESYTLIEYVRKMVPITREALSTLSSRDESGSVRVAGILVLDFFCVPMDVGEMFLNLPYIFLTCSAGTFLGMKYLPERHRETRSEFNR  
 SFNEELILIPGYVNVSUPTKIVLPSGLFMKETEPWELAERFPEAKGLVNNTALEPNFGFKYFDRCPDNYPTIPIGPILNIENKKDDA  
 KTDDEMRMLNEQPESSVVELCFEGSMSFNEKOVKELAVAILERSHRFLMSLRPTPKIEPKVEYENLEEVLEPEGFLKRISIGKVIG  
 WAPOMAVLSPHSVGFVSHCGINISTLESMWCGVPMMAWPLIYAEQTINAFLUVELGLAEIRMVDYRTDKAGYDGGMEVTVEEIEDGIR  
 KLMSDGEIRNKVKDVKEKSRAAVEGGSSYASIGKFIEHWNSVNI\*

도면4



### 도면5a

**UN6338** (부분) 누클레오티드 (SEQ ID NO: 21)  
 ATGGAATTCGATCCGGATTGGTGCAGCTCATCAGGAAGCTGTATTGAAACCGGGTGGCGCAGCGGAATGCTGCCAC  
 AACTAGTCGGGAAAGAAGAAGGGAAAGAAGGGAAAGTGGTAAAGAGHTGGAGTGCTTGAATGGGATCTACAAGAAC  
 AAATCCGCGTTATGCCGCGTCTGGAAAGGAGGGAAAGCTCACCGTCGAACAAGICGGAGAGTAGCATTAGGGCTC  
 GAGTCGGGGCTGAATTATGAGACTCTGAGGGGCGACGGCTCGCTGGGTTGAGGAGGACAAAGGAC  
 CGGGAAATGGTGGATGGGACTAGTCTGGCTACTGGTGGCTACTGGTGGCTCTGGTGGGGCTCTGGTGGCT  
 GGGGTGAGCTGATTGCGAGGGCTCTCGTGGCTGGGCTCTGGTGGCTACTGGTGGCTACTGGTGGCT  
 CAATGCTAGGTACGGGGACAAAGAAATCGGGTGGAGGAGGAGCAAGAAGAGGATGGGACTTTACTGGTAG  
 AGGAATTGCGAGATAGTTGAAGTTAGCTATGTTGGAAAGAAGAAGGTAAGGATTTCAG

**UN4666** (전체) 누클레오티드 (SEQ ID NO: 22)  
 ATGGAAGCTGGTGGTGTAAACTCCACATAGTAGTATTCTCATGGCTTCGGCCACATGGCTCTCTTCTAGAGCTT  
 CTCAAATCTCTCGAAGAGGGCATCTCATCCATCTGTATCCACCCAAAGAACATCCAGAGATCCCACAAATCT  
 CCTCCACAAATCTCTCCTAAATTCATCCTTATCCTACTCCAAAGTGAAGGATGGATGGCGGCTGGTGA  
 CCACACCGCTCCGGCGCAAACCTCCAGTACCTCAAAAGGCGCTGACGGCTCGAGCAGCTTCCGGAGCT  
 CCTCCGGAGGAAGCTTCCCCAACCCGATTTGATAATCCAAAGACCTTCTGCAGACTGGTACCCAAATAGGGCGGA  
 GCTCCACGTGCCGTCGATGTTACTTCGGCACCGTGCAGCGTGTGACTTCTCTGGCCACCCGTCGGAGTCTCG  
 AAGCGTAAGAAAGGATCGAGGACTGGCTGGTTCTCCGGCGTGGTCCCTTCAGGTGGCTACCCCTCC  
 ACCGAGATGATGGCGAAAGACACGGGGTCCCTCACTGGCGTGACCGACGTCGGCGCATGGGG  
 CCATCTGGGGTGTCTGGCGATAGGCACCTGGGGAGCTGGTGCGGAGTGGCTGGCGATTCCTGAAGAGA  
 TTACGGGAAGCCGGTGTCTGGGTTAGGCTACTGCTGCCTACTGGCGACGAAGACCGATGGCAATAGTATTCTGA  
 TTGGCTGGCACCGCGAAGCCAGGAATCTGGGGTGTACATCGCGTGTGGGAGGATCTGGGCTCTGGG  
 ACACGAGCTGGCGCTGGGCTCGAGCTCGGGGTTGCTTCTGGGCTCTGGGAGGCGAGGGTGTGAGG  
 ACACGAGCTGGCGCTGGGCTCGAGCTCGGGGTTGCTTCTGGGCTCTGGGAGGCGAGGGTGTGAGG  
 GGGTGGAGACCCGCTGTTGTTGCGATATTGGGGGACCAAGGGGCTCAACGGCGAGGCGTGTGG  
 GGGTGGAGATTGGAGAGGAAGGGGGACGGCTTTCAGCGGAAATGAGGTTGGCGAAGGGTGTGAGG  
 AAGGGGATGGATCAGGGAGGTCGTATAGGAAGAAAGCCAAAGGAGATGA

## 도면5b

**UN4666** (전정) 단백질 (SEQ ID NO: 23)

MEAGGKQLHIVFPMIAFGMILPFLIESLKSIAKRGHITISFVSTPKNIQREFPKSPSTNISS  
HKFHFEITPOSGRMFQGDVEATDIPPAULQYKKALDGLQPRFSFIREASPKPDWID  
LIQHMPPTAAELHVPSMFGTVAAALTFFGIPSESKRKIGTEWLVSPWPWFFPSKV  
AYRLHEMIVKADITAGLHSQVTDRMEEAVGCAVALRCRELESEWEPLILELYK  
FVTPYQGLIPTADESTDGNTIDWTRSOESTVYTAIGSEVSISGYELTHIALGIELAG  
LRFIMARTRYGLSSDELLPGGEERTRGIGRVMGIVPQMRVLADGSVGGFTVHCGWS  
SVVESIHFQHPLVLLPTEGDOGINARLLEEKGSGSNRGRTRLRRMWRRLS1WSWWK  
GMDQYRIGKRR

**UN3356** (전정) 누클레오티드 (SEQ ID NO: 24)

ATGGAAGCTGGTGGTGTATAACTCCACATAGTAGTTCCATGGCTAGCCTCGGCCACATGCTTCCTCTAGAGCT  
CTCAAAATCTCTCGCAAAGAGAGGCATCTCATATCCTCTGATCCACCCCAAAGAACATCCAGAGATCCCAAATCTC  
CCTCCACAATATCTCTCTCATAATTICATCCCTTATCACTCCCAAAGTGGAAAGGCATGCCCGTGACGTCGAGG  
CCACCAACCGAACCTCCGGGGCAAACCTCAGTACCTCAGTACCCAAACCCGAAATGGATAATCCAAGACCTCTCAGCAGC  
CCTCGAGAAGCTTCCCCAAACCCGAAATGGATAATCCAAGACCTCTCAGCAGC  
GCTCACGTGCCGTGATGTACTCTGGCACCGTGCGCCGGCGCAGCGTGTGACTTTCTTCAGCCACCCGTGGAGTCTCG  
AAGCTTAAGAAAGGATGGACGACGGGCTTCTCCCGCTGGTCTCCAAAGGGTGGTACCCGCTCC  
ACAGGATGATGTGATGGCGAAAGACAGCCGGTCCCTCCACTCGGCCGTGACGGACGTCGGAGGCTCCAGGGCGATGGCGG  
CCATCGTGGGTGCTGCCGCGTGCAGTACGCACCTCCACTCGGCCGTGACGGACGTCGGAGGCTCCAGGGCGATGGCGG  
TTAACGGGAAGCCCGTGTCCGGTAGGCCATCTGGTGTACATCGGTTGGGAGCGAGGGGATGGCAATAGTATTCGA  
TTGGCTCGGCACCGGAAGCAGCAGGAATCTGTTGGTGTACATCGGTTGGGAGCGAGGGGATGGCAATAGTATTCGA  
ACACGAGCTGGCTCGGCCCTGAGCTCGGGGGTGCCTTCTTGGGCTCTCAGGAGCCGTACGGGGTGTGAGCTG  
TTACGGAGATCTGCCGGGGCTTGAGGAGGGACAGGGGTACGGGAAGGTGGTGTGGGGTGTGAGCTG  
TTGGACACCCGCTTGTGTCGGTGGAGGATCTGACGCACTGGGTGGAGTCTGGTGGAGCTG  
GGGTGCGAAGTGGAGAGGAAGGGGGAGGGCTTACCGCGGAATGAGTGCGGAAGGCTCAATCTGTCATGGTGG  
AAGGGATGGATCAGGGAGTCTGCTAGTAGGAAGAACCCAAAGGAGATGAGAAGAAGTTTGGTGAACAAGAGGTGCA  
GGAGAAGTATGTGGATGAGTTATCAGTCTGCTAGTAATGGAACACGAAAGGGTAG

**UN3356** (전정) 단백질 (SEQ ID NO: 25)

MEAGGDKLHIVFPMIAFGMILPFLIESLKSIAKRGHITISFVSTPKNIQREFPKSPSTNISS  
INFIPSLPLRVEGMPGDVVEATDIPPAULQYKKALDGLQPRFSFIREASPKPDWID

### 도면5c

**UN6248** (부분) 누클레오티드 (SEQ ID NO: 26)  
GCACGGAAATCAGTGGCGATTGCTACAAGTCATGGGAAGCTGTACACGAAGGCCGGTGCTACTGTGGATTGGCC  
ACCCGAAGAAGTCTGACATGGAGCTGGCAGAGTGTCTCAATGGCTGGATCTACAGCAGCAAAATCCGTGT  
TATGTCGCTTCTGGAAAGCAGGGGAAGCTCACCGTGCAGAACAGTGGAGAGTAGCATTAAGGCTGGAGCTGAGCG  
GCTGAAGGGCTGGATTCACAGACTAGGTCTGGCTCATCTGTCGGTGGTTCTGACACATGGCCGGTGCAG  
TCGATAGTCGAGGGGCTCTGGCTGGTGGATGGCTGCCACTGTGTTGGGGCAAGGGCTCAATGCTAGGTT  
ATTGGTGGACAAGAAAGTGGGGTCTGGAGGTGCCGAGAAATGAGGAG  
  
**UN3131** (부분) 누클레오티드 (SEQ ID NO: 27)  
ACAAGGAGCGCCTGAAACCACCTCGTCGGAGTTATCGGACTGCGACGTGGTCGCCGTCTGGGATGCAACGAGATC  
GAGCGGGAGTGGATCGAGCTGTGGGGAGCTCTACAAAGAACCCGTTGGCCACTTCCCTCCAATGGCTCC  
AGCAAACACTACAGATGGCGATGGAGATAGGGAGTATAACGGAATCTTCCAATGGCTGGACGATCAAGAACCTGGCTCC  
GTGGCTCTACGTGGCTTCGGCAGCGAGGTAGACTGGAGAACAGGGTTCGGAGATGCGACTCGGGCTGAGCTG  
TCGGGACTTCCCCTCTGGGGCCCTGGAGGTGGCCCGGCCCTGGACGACGGTGTGGGGTGGGGTGGGGTGGGGTGGGGT  
ACGGCCGACGGGGATCTGGCTCACGGGGGGGCCACAGGGAGAACGGGGATGGGGTGGGGTGGGGTGGGGTGGGGTGGGGT  
CGTACGCACTGGGGTGGAACTGGTCGGTGGGGATGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGT  
GACGGGGCTCAACGGAGGCTCTCGAGGACAAGGGAGTGGCAAGGGAGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGT  
GGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGT  
AGTCGATGGGGAGGAATGGGAAGAGCAGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGT  
AGTAGAGGGATGAAGAGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGT  
GAGCATGGCCAACCTTAATGAGGAATTCTTACCGCAGACCGAGTCTACGGCCACGGGGTGGATTGAACGGCTAATAATT  
TTCCGGAAATTATGGCTTGTAAATGC

LJQHWPPIAELHVPSMFVPAALTFEGHSEFSKRKGTEDRPGSPWPFSSKV  
AYRLHEMTVAKDAGPLISGVTYVRMEEATVGCCATAIRCRELSEWEULPTELEYKG  
PVTPVGLLITADESTGTLISIDLGTSQSSEWVYIAIGSEVIGVLEHLAULGLIG  
LPFLWALRRGYLGSSTDTELLPGGEERTRGKGVMWPMQMVLAWSGGVFTWHGWS  
SVVESELHFGPLVLWIPFGOGIMARLBIEKGIGVEVERKGDAFSTRNEVAKVLNSWMK  
GMGSSSSYRKAKEMKKIFGDKEQEYKVDEFIQFLSNGTAKG-

## 도면5d

ATGGAAAGCTGGTGGTGATAAACTCCACATAGTAGTTCCATGGCTAGGCCATGCTTCCTTCTAGAGCT  
 CTCAAAATCTCCGAAGAGAACGCCATTCATATCCTTCGTATTCACCCAAAGAACATCCAGGATTCCTCAAATCTC  
 CCTCCCAAATATCTCTCTCATTAATTCATCCCTACTCCCAAGTGGAAAGGATGCCGGGACGCTGAGG  
 CCACCACCGACCTCCCGGGGAAACCTCCAGTACCTCAAAAGGCCTCGACGGCTCGAGCAGCTTCCGGAGCT  
 CCTCCGAGAAGCTTCCCCCAACCCGATTGGATAATCCAAGACCTTCTCAGGACTCTGGATACCCAAATAGGGCCGG  
 CTCCACGTGCCGTGATGTACTTCGGACGGTGCCGGCCGGCAGCGGTGACTTCTCGGCCACCCGTCAGTTGCTGA  
 GCCGGGTAAGGGCTCGAGGGCTGGCTGGCTTCCTCAGCCGTGGCTCCCTTCAGGTTGGTACCGCCTCCA  
 CGAGTTGATGTGATGGCGAAAGACGGGGGGTCCCTCCACTCGGGCATGACGCCGGCCATGGAGGG  
 CATCGTGGTGTGCTGCCGTGCGA-TACGCACCTGCCGGAGCTGGAGTGGAGTGGGTGGCGATTCTCGAAGAGATT  
 TACGGGAAGGCCGTGATTCGGTAGGCCCTACTGCTGTGCTACTGCGCACGAAAGCACCGAATGGCAATAGTATTCGATT  
 GGCTCGGCCACGGAAAGCCAGAACCTGIGGTGATCACTCGGCGTGGGAGGCGAGGTGTCATCGGGTGGAGGCTGATCG  
 ACAGGCTGGCGCTGGGCTCGAGCTCGGGGGCTGGCTGGAGGGGACGGGAGGGGGTACGGGAAGGTTGGTATGGGGGGTCCCAA  
 TACCGAGATCTCCGGGGGGCTGGAGGGGGAGGGGGTACGGGAAGGTTGGTATGGGGGGTCCCAA  
 TGAGGGTGTGGCCGATAGGTGCTGGGGAGGAGGATCTGTGACCGCACTGGGGTTGGAGGTCTGGTGGAGGAGCTTGCA  
 TGGACACCCGCTGTTTGTGGCGATATTCGGGGACAGGGGCTCAACCGCAGGGCTGTTTACGGAGGAATGGAGGTGGCGAAGGGGATCA  
 GGTGGAGGTGGAGAGGAAGGGGGACGGGGCTTACGGAGGAATGGAGGTGGCGAAGGGGATCAATCTGATCATGGTGG  
 AAGGGGATGGATCGGGTAGTCGTATAGGAAGAAAGGGAGGAGTGAAGAAGGAGATTTCGAGACAAGAATGGCAG  
 GAGAAGTATGTGGATGAGTTGTCAGTCTGCTCGAGTAATGGAACAGCAGAACAGGAG

## 도면5e

ATGGAAAGCTGGTGTGATAACTCCACATAGTAGTTTCATGGCTAGCCTTCGGCCACATGCCCTTCCTAGAGCT  
 CCTCAAATCTCTCGCAAGAGGCCATCTCATATCCTCTGACTCACCCCAAGAACATCCAGAGATTCCAATCTC  
 CCTCCCAAATATCTCTCTCATAAATTCTACCCATATCACTCCCAAGTGAAGGCATGCCGGTGAAGCTGAGG  
 CCACACCGACCTCCCGCCGCAACCTCCAGTACCTCAAAAGGCCCTCGACGCCCTCGAGCAGCCTTCGGAGCTT  
 CCTCCGAGAAGCTTCCCCCAACCCGATTTGATAATCCAAAGACCTTCTGCAGCACTGGATACCCAAATAGGCCCGA  
 GCTCCACGTGCCGTGATGTACTTCGGCACGGTGCGCCGCAGGGTGAECTTCTCGGCCACCCGTCGGAGTTCTCG  
 AAGCGTAAGAAAGGGATCGAGGACTGGCTGGTTCTCCGGCGTGGTCCCTTCAGGTGCGTACCCGCTCC  
 ACGAGATGATTGATGGCAGAACACGGGGGTCCTCAACTGGCGTGACCGACGTCGCCGACATGGGGCG  
 CCATCGTGGGTTGTCGCCGTGGATACCGACCTGGGGAGGGTGGAGTGGCTGGAGTGGCTGCGGATTCCTGAAGAGA  
 TTACCGGGAAAGGCCGTGATTGGTAGGGCTACTGTGTGCCCTACTGGCAGAACGACCCGATGGCAATAGTATCGA  
 TTGGCTCGGCACGGAAAGCCAGGAATCTGGGACATCGCGTIGGGAGGCGAGGTGTCATGGGGGGCTGCGT  
 ACACGAGCTGGCGCTGGCGTGGAGCTGGGGGGGCTCTGGCTCTAGGGGGTACGGGGTACGGGGTACGGGG  
 GATACCGAGATCTCGCCGGGGCTCGAGGAGGGTACGGGGTACGGGGTACGGGGTACGGGG  
 AATGAGGGTGTGGCGATGGGTGGTGGAGGATTCCTGJACCGCACTGGGGTGGAGTTCGGGGTGGAGGAGCTGCA  
 TTGGACACCCGCTGGTGTGGCGATTCGGGGGACCGGGCTCAACGGGAGGCTGTTGGAGGAGAAAGGGATCC  
 GGGTCTCGAAGTGGAGAGGAAGGGGGACCGGGCTTACGGGGATGAGGGTGGCGAAGGGCTGTCACCTGGTCATGGT  
 GAAGGGGATGGATCAGGGAGTCGTATAGGAAGAAAGCCAAAGGAGATGAAGAAGATTGGTACAAGAGGGTAG  
 GGAGAAGTATGGATGAGTTTACGTTGCTCACTAATGGAACAGCAAAGGGTAG

**UN4522** (전장) 단백질 (SEQ ID NO: 31)  
MEAGGDKIHTWEPWLAFGHMIPETELISKSTAKRGHITSS

MEAGGKUHTWEPWMAFGMIPPEFLSKSALKGHLSFVSTTRNLTORPNPUDQPL  
INFIPLSLKVEKGMDNEATDPLPANTQYKLKGUDLQFESFLWSPWPWTSPKV  
LIOUWPPRRAELHPSMYGTVPAULFEGHSFSRSKKGGDMLWSPWPWTSPKV  
AYRLHEMIWAKDTAGPLHSGVTDVRMEAIVGCAVARTERLESEBWLPIELEYK  
PVTPVGYLLPTADESTDGNSITDWIGTRSSESVYTAIGSEVSIGVELHELAIGLELAG  
LPPFLWAIARPPGGSITTELFGGFERTTRYKGVKGWMWQMRLADASVGGETWCCWS  
LPPFLWAIARPPGGSITTELFGGFERTTRYKGVKGWMWQMRLADASVGGETWCCWS  
SVFSEHGFPLVLPFGOQLWNLARVEKGGDAS-TRNEVAKAVNLWVE  
GDGSSESYRKAKEMKJFEDKECOKKEDEFIOPFLNGTAK-  
GDGSSESYRKAKEMKJFEDKECOKKEDEFIOPFLNGTAK-

## 도면5f

UN6460

(부록)

(SEQ ID NO: 32)

ATGGAAGCTGGGTGATAAACTCACATAGTAGTTCCATGGCTAGCCTCGGCCACATGCTCCTTCAGAGCT  
 CTACAATCTCGCAAGAGAGGCCATCTCATATCCTCTGCATTCCTTATCCTGATTCCTACCTCCAAAGCATCCAAATCTC  
 CCTCCACAATATCTCTCTCATAAATTCTCCTTATCCTGAGCTTCCAGTACTCTGGCACGGTGCAGGGCAGCGTGTACTTCTGGCC  
 GCTCCACGTGCCGTGATGTACTCTGGCACGGTGCAGGGCAGCGTGTACTTCTGGCCACCCGTCGCAAGTGGCAGCTGCGAG  
 AGCCGGTAAGGGCTGGGGCTGGCTGGCTTCCTCCACTCGGGCATGACCAGCCGCCATGGAGGGCG  
 CCATCGTGGGTGCTGCGCGTCGGATACGCACCTGGGGAGCTGGAGTGCGGAGTTCGAGGAG  
 TTACGGGAAGCCCCGTGATTCGGTAGGCCTACTGTGTGCCTACTGCGGTGGGTGATACCGCAAGAACGACCTCTGCAAG  
 TGGCTGGCACGGCAAGCCAGGATCTGTGGTGTACATCGGGTGGGTGACCGGAGGTTGGCATCGGTGGAGCTGAT  
 ACAGAGCTGGCTGGGGCTCGAGCTGGGGCTGGGGAGGGTGGGTGACCGGAGGTTGGCATCGGTGGAGCTGAT  
 GATACCGAGATCTGCCCGGGGCTCGAGCTGGGGCTGGGGAGGGTGGGTGACCGGAGGTTGGCATCGGTGGAGCTGAGC  
 AATGGAGGGTGGGGCTCGAGCTGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGG  
 TTTGGACACCGGCTGTTGGCGATGGTGGGGAGGATTCGTGACGCACTGGGTTGGAGGTGGGTGGAGGAGCTGCA  
 GGGTCGAAGTGGAGAGGAAGGGGACGGCTTITACGGGACCGGCTCAACGGCAGGGCITCAATCTGGTCATGG  
 GAAGGGATGGATCAGGGAGTCGTATAGGAAGAAGCAAGGAGATGAAGAAGAAGATTTGGTGGACAACAGCAAAGGGTAG  
 GGAGAAGTATGGATGAGTTATCAGTTCTGTCAGTAATGGAACACAGCAAAGGGTAG

UN6460

(설명)

단백질

(SEQ ID NO: 33)

MEAGGDKHLIVVFFMIAFGHMLPFLIEYKISRKERSHTRIHPNQEPNLPPQISPL  
 INFPLSLPKVCEMPFVPAALTFDLPPLPANLOXIKKALDGLRPFNSFLRASPKPDWIDQ  
 LLQHWIIPFTAAEIRVPSMFPTGVPAALTFDLPPLPANLOXIKKALDGLRPFNSFLRASPKPDWIDQ  
 AYRHETLVMKAODAAGLHSGMTDARNEEAATIVGCCAATRCBLESWMLPLEETIGK  
 PVTFVGULIPTADESDGNSIDLWLGRSRSQESVUVTALGSEVSIGEVLTIELAIGLEJAG  
 LPFWALIURPYGSISDPEIILUGGFEBRTGKVNWVOMRYVLADGSVGGFVTHCWS  
 SVVESLHGHPLVLLPFGDQGINARLJEEKGIGEVERKGDASTRNEVAKAVNLWVE  
 GDGSGSYSRKAKEMKIKFGDKECQEKYVDEFIQFLISNGTAKG-

도면5g

**UN2281** (전종) 누클레오티드 (SEQ ID NO: 34)

```

ATGGAAGCTCAAGCTGGTAGTGATCAGAAGCTCACGTAGTAGTATCCATGGCTAGTTGGCACATGATCCCT
TCCTCGAACTATCAAATCTAGCAAGAAGAGGCATCTCATATCCTCATCCTCCACCCGAAGAACATCGATAGACT
CGTCCCAAACCTCCATCCGGAAATAATTTCATCCCTTCGTTACCTCAAGAAAGCCTTAACGGGCTCGAAAGGCCATGGGAGTCCGAG
GCCACGACCGACCTCCACCCGAAGACCTCCGTACCTCAAGAAAGCCTTAACGGGCTCGAAAGGCCATGGGAGTCCGAG
TCATCCGAGAAGCTCTCCCAACCCGATGGATAATCCAGCATTGGGACCTCCACCTCCACACTGGATGACCCGGAACTT
CGCTAATTCTCTCAAGACTTCCCTGCCGCCACCACTCTCTCATGGGCCACCTCCACACTGGATGACCCGGAACTT
GACAGCTGGTAGTTCTCTCCGGGCCCTTCCCTCCAAGTGTGCTATCGCTCACGAGATGTTCCATGGC
CGCAAGCCACCAACAACAGAATGCCAACACCGACACGGGCCGACCTCACCGAGATGTGCTCAATGGCTCAACGC
CCTACGGGTACGGAGCTGCAATGGAGCTGGAGCGCCGACCTGGCGCTCTCGAGAAGATCTACAAATGGGTTATT
TCCGGTAGGCCTACTTCTCCGGGCCAGCAAGTICGGGGCAAGCCAGCAATITGGATATTATGGATGGCTCGGTACA
CAGAGTCAAAACACCGTCTCTATATCGCTGGGGAGCGAAGTGGACCGTCTCGTAACCGAGCTGATACACGGAGCTGGCT
CTGGGCTAGAGCTCGGGGGCTTCAGGAGCTGGGGAGCTGGGGAGCTGGGGAGCTGGGGAGCTGGGGAGCTGGGGATGT
TGCCCGAGGCTCGAAGAGCGAACGAGGAGTATGGGAAGGTTGGGGATGGGGAGCTGGGGAGCTGGGGAGCTGGGGATGT
GCCGAAGAGTCGGGGAGGATCTGACGCACTGGGGAGCTGGGGAGCTGGGGAGCTGGGGAGCTGGGGAGCTGGGGAGCTGGGGAG
TCGIGATGTTGCCCGCTTACGAGGGATGACGCACTGGGGAGCTGGGGAGCTGGGGAGCTGGGGAGCTGGGGAGCTGGGGAG
GAAGGAGGAGGATGGGCTTACGAGGGATGACGCACTGGGGAGCTGGGGAGCTGGGGAGCTGGGGAGCTGGGGAGCTGGGGAG
GTCGGGATAGGAGAAAGCTCGGGAGATGAAGACGGGTCTGGGGGAAAGGAAGTGTACATGGGGAGCTGGGGAGCTGGGGAG
TTTGTTCAGTATCTGCTCAATCATCGTGCCTCCATACAGAACTGA

```

**UN2281** (전종) 단백질 (SEQ ID NO: 35)

```

MEQAGDQDKIHVWFLAHCMTPEELSKSLARRHILSTSPKNDLDRVPLNHLRE
INTPPSLPUKVEGMASEVEANTDIPEDPELYKKRNLGEKPNFRTEASPKDWLIOH
WVPIIAESGASLIFFFTPAKTFMFGHSWHRLELWLSWPFWWPPFSKVSYRHE
MSMAMAHQOKNSNTDINGPDTETIUNGCHALAVRSMELEPDIPPLEKTYMPVPG
LIPPAQYGDASNCIDMWIGTQSOKTLYALGSEVTRNLTHELAUGLEGAEFL
WAURRPQSGAGVGMIEFEEERTKEYKTYAMEWPMEVYLAEBSVGFRFLICGWGSVE
SLHFGHPLMLVFGIQQFLNARMLEKKGIGIVEVEREEDGSFRDDVAKVLTMEEGG
RYRKAREMKTVLGKRCQDKYADEFVQYLINHRASION-

```

도면5h

UN2644 (부분) 누클레오티드 (SEQ ID NO: 36)

GGGGAGCGAGGTGACCGATGAGCGGGAGCTGACGAAGAGCTGGCTCTGGGCTGGAGCTGCCGGCTGCCCTTCCTG  
GGGCCTCGGAGGCCTTGGCTCGCTCCATCGCGATGTCGAGATTTGGGGAGGGTCCGAAGAGCGGACGA  
GGGATTGGCAAAAGTGGGAGGGTGGTGGCTCCGCAACTCGAGGTCTGGCCACTCGCGTGGAGGATTCTGA  
CGCACTGGGGTTGGATCGATTATCGAGGGCTAACATTGGACGCCATCTGATGCCGGTGGAGGATTCTGA  
AGGGCTCACCGGAGGCTGGAGGAGAAGGGATCGGAGCTGGAGGAGAGGGACGGGTGGTACCG  
AGGAATGAGGTGGCAAGGGCGTGAAGTTAGTCATGGTGAAGAGGATGGAGGGTGGTATAGGGAGAACGCG

### 도면6a

**EUGT-1** 누출로트드 (SEQ ID NO: 37)

ATGCTAAACGCCAATTAGTGCCACATTCCTACTTTAACATGCCAGCTAGGGCAGCATCAATTCACTTGGAAGAGGTGCACTCATACUTTG  
GACAACGTGCAACAGGGTCTGCTAACAGCTACTAAGTAGAACAGGTCAGTTAACATGTATCTCAGAGGTTGACCCAGGGCTGAAAG  
GCTGAGCACCTAGAGAGGCTACTCTGCGACGACCCAAAGTCAGATGTTAACATGTAGGAGGAGCCTGCTAACAT  
TGTTGGGATTAAGACAGGCTGCAATTGTTGCGACTGCTGAAATCAATGGCCAAACGGTCCCTCAACTTGAGGATTCAGGAGGTTGCT  
ACAGACATGTTGGGAAGGCGAGGATTAAGTGCTCAGATGTTGCGACTGCTGAAATCAATGGCCAAACGGTCCCTCAACTTGAGGATTCAG  
GATTGGTAGAGTCAACCCAGACCTACTACCTTGCGTTAACACACAAAATTAGATAAAGCTGATATGATGGCAGAAAATAGTACTAGGAATTG  
AAGCAGACCTCTGAGATGGTCCCGTTGCGACTGCTGAAATCAATGGCCAAACGGTCCCTCAACTTGAGGATTCAGGCTTGTAGGATTCAG  
GGGTAAACCTAAACCTCGAGATACACTCCAGATACACTCCCATATCTGGATTGGCTGAGCTCAAAGCACTAAAGCACTAACTTGCTGAGCTG  
CCTGATCAACAAATGAGATTGCTCTGAGTATGATAGTCAACAAATCTTATCTGGTAGTCAAGAACATGCTGAGCTGTTGATGCTGAGCT  
AGAGAAGGGTAGCCCTGGTGTGTTCCAATGTTGGGGTGCCTCAATGGCTGAGCTGACAGGCAATGTAAGCTGAGGAGGAGGTTGCAAAAG  
TATGGAGGGTAGCCCTGGTGTGTTCCAATGTTGGGGTGCCTCAATGGCTGAGCTGACAGGCAATGTAAGCTGAGGAGGAGGTTGCAAAAG  
AGAGAAGGGTAGCCCTGGTGTGTTCCAATGTTGGGGTGCCTCAATGGCTGAGCTGACAGGCAATGTAAGCTGAGGAGGAGGTTGCAAAAG  
GAGCTAAACATATGGTAAAGACTCTGTTAGAGGAGAGAGGTTGCAAAAGATGGCTGAGCTGAGGAGGAGGTTGCAAAAG  
CCGGCAGATGGTAAACTGGCCAAGAATCTGTAAGTGAAAGGGCGAGTAGTGTAAAGTGCTATAAGGAGTCTCATACATGTCATGTC  
TCATGGTATAA

- 45 -

도면6b

GATGCCCTGAGAAGGGTTATGGGGGTGAAGAGGGCGATGAGTTCGGTTAAAGTAAAGAGTTGGCAAAGTGTTGGCATGAGATGAAAGTTAATGACAGAT  
ATGTCGGTATTCCCTAAAGTCGTTGCTCAGATACAAATGCAAGGCAGGTAA

**EUGT-3** 누클레오티드 (SEQ ID NO: 39)

ATGGCTGCAACTTCTGATTCACACCTGCTGCGCCCTGCAGCAGCTGCCTCTAGTAGTTCATCCCCCTACACATTGTAGTATTTCCTGGCATTTGGACA  
CATGATTCATTTCTGAGTIAATCTAAAGACTCTGGCTAGTCGTTACTCTGGCTACTTCTGCTACTTCTGCTACTCTGGCATAGAACATGCAAGCTGGGCTACACCTCA  
GCCCACTGTCATCTCCAGACTAAGAGTAGTCCTAGACTACACGGCTTGCTGAGGCGCTGAATCACAGCAGACGTGCCACCAGAA  
AAGGTGGGTCTATGAAAAGGCGAGATGGTATGGTTGCTACTTCTGCTACTCTGGCATAGTCGAGGAGCTGAGGAGCTGAGGAGCTGGCAG  
TGCCGCTTCTAAGGAAGCCAGATGGTATGGTTGCTACTTCTGCTACTCTGGCATAGTCGAGGAGCTGAGGAGCTGAGGAGCTGGCAG  
GTCCTAGCTGCCTTAAGGAGACCTCTGGCATAGTCGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGGCAG  
TCTAACATTCGATACAGGAGACGTATGAGAAGCCAGATGGTATGGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGGCAG  
ACAACATCTAATGTTAGTATGATCATACAGAACTTGTCCAGAATGACCAAAATGTTCCATGCTGACAGAGTTACACAAAGCCAGCTATCCAT  
GGCTTGTAGTCTCAGCATGGAGAGATAATGATATAGGCGCTACATGCTGAGCAGATCATCTGGCTGCTGAGAACAAAGCCAAACA  
AATCCGGATCTAGTGTAGTCTGGTACTGAAGSACTTACAGCGGATCATATGATGAACACTAGCTTGGATTAGGATAGTCTGGTGTAGTCTGG  
ATTAGGAGAGACCTCTGGTATCACTGCTGAGTATGGTATGGCTGAGACAGGAGCTGGGTTCTACCCACTGGGTTGGGTGAGAAGTGGTCC  
ACAGGGAGAATGTCATGGGAGCTGGGTTGGGTGAGAAGTGGCTACAGGAGCTGGGTTGGGTGAGAAGTGGTCC  
CCTTICATCGCTGATCAAGGACTGATGAGGAAAGCTGGGAGCTGGGTTGGGTGAGAAGTGGCTACAGGAGCTGGGTTGGGTGAGAAGTGGTCC  
CGTGGCTGCTGCTATCAAGAGCTGATGGTGGAGGAAAGAGTGTAGCTCATAAAGCTATCGACTTGGGTTGAGAGAGCTACAGAACAAAG  
AGATGTTACTTATAGGAACTATGGGATACTTGGCAATGCTACAGATA

**EUGT-4** 누클레오티드 (SEQ ID NO: 40)

ATGGATGCCTCACCATGCTGTTAGTGTCCCTGGTTAGCTTGGACATCTTCTACCAAGCTGTTGGAGTTGGCTCTAGAGGGCTAAGAG  
TCTCTTGTCTCTACTCCAAGGAACATGCTAGCTGAGTGGCTATGGAAAGCTGACCCATGCCCACATGCTGTTGGAAAGCTGAGGCTATGCTGAG  
GCTGAAGCCACACAGACGCTCCAGATCATGTCAGTGGCTATGGAAAGCTGACCCATGCCCACATGCTGTTGGAAAGCTGAGGCTATGCTGAG  
CGGAACAAAGCTGGATGGTGTGATGGCTGAGCTGGAGCTGAGTGGCTGAGCTGGAGAGCTGAGTGGCTGAGCTGGAGCTGAGTGGCTGAG  
CTACTGCTGCTGCTACTTGGAGCTGAGTGGCTGAGCTGGAGCTGAGTGGCTGAGCTGGAGCTGAGTGGCTGAGCTGGAGCTGAGTGGCTGAG  
TACTTGGCTGAGCTGAGTGGCTGAGCTGGAGCTGAGTGGCTGAGCTGGAGCTGAGTGGCTGAGCTGGAGCTGAGTGGCTGAGCTGGAGCTGAG  
TACTTGGCTGAGCTGAGTGGCTGAGCTGGAGCTGAGTGGCTGAGCTGGAGCTGAGTGGCTGAGCTGGAGCTGAGTGGCTGAGCTGGAGCTGAG  
TACTTGGCTGAGCTGAGTGGCTGAGCTGGAGCTGAGTGGCTGAGCTGGAGCTGAGTGGCTGAGCTGGAGCTGAGTGGCTGAGCTGGAGCTGAG  
ATTCGACCTGGGTTGAGAGCGTACCCAGAGGAGTAGGATGAGGAGCTGGTAAAGACAGAAATGGGTCACCAACTTAAGGAGCCACCGTGGGATGACGGGCT  
TACACATGGGGACACTCTCTGTCATGAGGCTGAGATTGGTACCCCTCTAGTATGTTACCTCTGTTGGGATCAGTTCACCTGGGCTGAGCTGGGCT  
GCGCTGGGGTAGGTTGAGAGCTGGTACAGAGGAGCTGGTACAGAGGAGCTGGTACAGAGGAGCTGGTACAGAGGAGCTGGTACAGAGGAGCTGGTACAG

### 도면6c

AGGAATCCAAAAGCTTGGCTAAATGCCGTAAAGATGGGAGGGTTGCTGACACCGAATGTCATGAAAGATGTTGCTCATAAACATTGAGATCTTACACAATCACGGTACAGGATATTA

**EUGT-5** 누클레오티드 (SEQ ID NO: 41)

EUGT-6. 누클레오티드 (SEQ ID NO: 42)

ATGGAAGCTACTAGAGAGCTGCCCAATGGCTCATCAGCATCTACTAGTCATT  
CATTAAGTGTACCTGTTCCATGGCTGGCTTGGCTATGATCCCATATCTTGAACTTGCTAAGAGACTGGCCAGAAGAGGG  
GTTGTCACACTGGACCTTCAGCGTGTGATGGTCTACTCTGAGGAGCTGATCATCAGCAGATGCTCACCAAGAAAGGTA  
GGTTTAACTAAAGAACGCTTGTAGTTGTTAGCAGCTCTTGCCTGGTTGAGCTTGGCTATATTAGATTGCTCAGCAAAATGGGCTGGCT  
GAATCACACCAACTGCTGCTGGATTTCAGAACGCTGATGGTATTAGATTGATTCAGCTGAGCTTGGCTAGTGGAGCTTGGCAATGGT  
ATTGCGGAAGAGCATAAATTCCTGTGCAATGTCTCTATCTTTCAGCTGAGCTTGGTAGCATTTGAGGACCAAGGCAA  
GAGAACCTGACATCTAGAACATAAACAGAACATTATGGTGCACCTCTGGATACCATTCTCATTAATGTAGCA  
TACAGAGGGCTCACGGGGCTGAGTGGATTCGCTGAGCTTGCAGCTTGGCTAGCTTGGCTAGTGTAGCATGGCAGAGATT  
TGGAAATGGAGCACGCTGCTGAGTGGATTCGCTGAGCTTGGCTAGTGTAGCATGGCAGCTTGGCTAGTGTAGCATGGCAGAGATT  
ACCGAATTGTTGCCAACACCATCGTCTGGCGGTTGTTAATGCCACCTCCACCTCAGCTGCGAGGGGTAGATGACGAC  
GATGATGAGCTATCTATGGATGACAACACATCGCTATGGCAATGGCTGGCTGGATGAAACCTGAGAGGCTGTGATA

도면6d

TATGTGGCATGGGTTAGAAGACTTCACAGACAGTAGGTACCGTAAGAGAAATTGGCTCTGGCTTGGAGTTGGCAGGGTT  
 AGATTCCTATGGGACTTAAAGGCCCTCATCGCCCATCTGICACAGAGACAATGTCAGCGCGCATCTATA  
 TTGCCAGATGGCTTAAAGATCAAGAGTGTCTGCTGCCAGAGGAGGTCGGCTGGCTGGCAAGATGGGTTCCACGTTACGTT  
 TTAGCCATAGAGCTACTGGTGGATTCTAACACATGTTGGCTGCTCCATCTTGAAATCACTGAGAATGAGATTGCTTACCA  
 TGGTGTGATGGTGCCTATTGCCGACCAAGGCTTGGGGTCAAGCTTACAGAAGAGAAATCGAGTCGAAGTGGCC  
 TGTAAATGATGACGGCTTAAAGAGAATGAGAGATGCCATCGCAGCTGCCGTAGACAGATTTGTTGAAGAGAAAGGTAAGCA  
 CTTAGTAGAAGAGGAGAGGAGTTGAGGGATGTGCTGGGTGACGAGGGTAGACAGAAATGTACCTAGATGAATAGTCGGC  
 TATCTACAAAGATACAAAtaa

**EUGT-7** 뉴클레오티드 (SEQ ID NO: 43)

ATGGCTGCTACCTCAGATTCTACCCAGCTGCTGCCAGCGCGGAGCATCTCTCA  
 TCTAGTCTCTCATATAGTGGTATTCCCATGGCTTGGTTATGATTCATGTCATTCTGGAATGTCCTAAAGACTG  
 GCTTCTAGAGGACATGGCTTACTTGTCTACAGAACTCTAGAAATGCAAGAATGAGATTGGAGGTACACTCTGCCCATIG  
 TCTAGTCTCATCTAGATTAAGATGAGTTGGTTCATAGTTGCTGAGTGTATGGATTAACCTGAAGGTGGAGTCACAGCT  
 GATGTTCCACCTGAAAGGTCGGACTGTGAAGAAGGCTTITGATGCTTGGCGCTCATTTGCTGATGATTCGTTGAGAA  
 GCATGTGCAAGCGGGGATGGGGAGCTGTAACCTGAGAGCAGCAGGGCTCTGAGAAACCTGATGATGATTCCTGATTIC  
 GCACACTCTGGATCTGCCAACAGAAACAAAATCCATAGCTACATTGCTAGTCAGTGTG  
 GCCATTITGGGTCACCTGAGAAATTAACTCTCACGTACTACAGCGGAAGATATCAGTGGCAACCACATGGATA  
 CCATTTCATCCAATATGCAATATGGAGAAGGCCACCGAGCGCAATGGATGTTAGCTGATTCGGCTAATGCTAATGGC  
 GTGTCGACATGGATGATTGGAAAGTGGAGCACACCTCTAACCTGATGAGCTATTAATACAGAACCTGCCAGAAC  
 GAACCTAGACTTCTCACTTGTACTGAATTGTACACAAAGGCCAGAAATCCAGGGCTTGGTCCACCGCTTA  
 GACGATTAATGATGATGTTGTTAACAGATCAGACAGATCTTGTGCTTATGCAATGTTGGACAAACACCAAAT  
 AAGTCGTCATCTATGTCCTGGGACTTGTGAGCTTGTGAGGACATCACTGAGGACATGTGTTACCTCTGG  
 TTAGCAGGTTGAGGACTTGTGCTGGGCTCTGTAGACCTGAGGACATCACTGAGGACATGTGTTACCTCTGG  
 TTGAGACTAGAGTCAACACTGCGGGTTAACCGCTGAATCTTCACTACGGTCAACCTGTAGCTAGTCATGTGCTTC  
 GGTGTTCTAACACACTGCGGGTTAACCGCTGAATCTTCACTACGGTCAACCTGTAGCTAGTCATGTGCTTC  
 ATAGCTGATCAGGGACTAATGCAAGACGGGTGCGCAACAGAGTTATGGTCGAAGAGGGAAAAGAGCTGCTCATAACT  
 TCACTACGGGACTGTCGCGCTGCTATCCAAAGAGTTATGGTCGAAGAGGGAAAAGAGCTGCTCATAACT  
 ATCGAATTATGTTGGGACTTAGGAGATGAGTACAACAGGAATGTACCTTACGAATTGATGTTACCTCAATGCTAT  
 AAAtaa

**EUGT-8** 뉴클레오티드 (SEQ ID NO: 44)

### 도면6e

**EUGT-9** 노출오리드 (SEQ ID NO: 45)  
 ATGGACGCATCCACCTCACTGTTGATTCCTGGCTGCTTGGGACTCTGCA  
 CCAGGGCAATGTTGGCTGTAGATGGCATCCAGAGGTTGAGAGTGCTTGTATC  
 TTGAAAGGGCAATGTCATCTGTCGAATTACCTAACAAAGAGTGTATGGGTTACCGATGGGCCAAGGA  
 ACAACTGATGTTCCAGATCACATGTCAGTGTCTTATGGAAAGGCCATGATGGACTGACAGACCACTCTGCAATTCTG  
 GACGCTGCACTGGCGCAGGAATAAAGTGTATGGCTTAACCTGCAACAGCTGTCACATTGGCGTICCTGCTCAGCGCT  
 GACCGTAAATGTCATGCGTGTGATGTCATTCACCTGCAACAGCTGTCACATTGGCGTICCTGCTGATGAGCAAGA  
 GATGCTGACAGATTCATGCGCATGCTGTAGATGCTGTTCTGCAAGAAGTAGTGTAGTTAGCTGTTAGATA  
 TTGTTAGAAATTGAACTGAAAGTGTCCATTATCTTAATTCATTTGGCAACCCAGTAGTACCAATGCTTACTACCT  
 CCACCTGTTGATGGAGATGGAGACAGACACATGTCATCTGTTAGACAGAACCCAGACTAAGTCGTG  
 GTCTATGTTGCCCTTGGATCAGAGAGGCCCTCTGACTGCGAGCAAGGAGAGAAATTAGCACTTGGATTTGAGAAAG  
 GCTCTTATGGCATTGAGAAAACCTACGGGGTGTATGCACTGATGGGGTCTGCTGCTCCAGGCTTGGAGAAAG  
 GTAGGAGCTGCGTITAGGAGAAGTGTGGATGAGGAGATGAGGACATGAGGAGATTTGGAGAGCT  
 GCAAGAACATGGCAGAAGGGTTGGCTGTGGAGTGTCTAGAGGACATGAGGAGCT  
 GTAGGAGCTGCGTITAGGAGAAGTGTGGATGAGGAGATGAGGACATGAGGAGCT  
 GTAGGAGCTGCGTITAGGAGAAGTGTGGATGAGGAGATGAGGACATGAGGAGCT  
 GAAAGTACTGtaa

**EUGT-9** 누클레오티드 (SEQ ID NO: 45)

AAGGAGGUAGAUCCUACUCAAGGGTGTGATGTTCTTACCCCAAGGACATGCCAGA  
CACCGGCTGAACTTGGCTGAGTGGCAAGGTTGAGTGCTTGTGATCAGGCTTACCCCAAGGACATGCCAGA  
TTGAGAAGGCCATGTCATCTGTGGAATTACCTTACCAAGGTTGAGTGGGTTACCATGATGTCGCCAGAGCA  
ACAATCTGATGTTCCAGATCACATGTCAGGCTTATGGAGGCTCAGATGGACTGAGACGACCATCTGCAATTCTG  
GACGCTGCAGCTGCCAGGAATAAAGTGTGGTATCTGGATGGGATGTGTCATGGGCTGCTGCTTCAGCCGCT  
GACCGTAAGAATGCCATGCGTGTGATGATGCTCCTAACATGCAACAGCTGTCGACATTTGGTCTCTGATGAAAGA  
GATGCTGACAGATTCATCTGCCATGTCCTGAGATCTGCTTCAGAGTAGTGTGAAATTAGCTGTTAGATCA  
TGTGAGAATTGAACTGTGAAAGTGTCCCATAATCTAATTCTGGCAACACCTGATGACCATGTTTACTACCT  
CCACCTCAGGTGATGGAGGAGGAGGAGGAGAATGTCATCTGGTATGACCAACCCATACTGCTT  
GTCATGTTGGCTTGGATCTAGAAGGCCCTCTGACTGCGAGCAAGGAGAGAAATTAGCAGCTGGATTGAAATTCTGG  
GTCCTCAATTCTATGGGCTTGGGATGAGAAAACCTCAGGGGGTGTGAGGATGGTGGTCTGCTGCTCCAGGCTTGGAGAAAG  
GTCCTCAATTCTATGGGCTTGGGATGAGAAAACCTCAGGGGGTGTGAGGATGGTGGTCTGCTGCTCCAGGCTTGGAGAAAG

## 도면6f

ACTCTGGGGAGGATGTTAAAGACTGAATGGGTACCTCAACTGAAGAATTGGTCATCTGCAGTGGTCATTCTCA  
ACACATTGGGCCATTCTCGTCATAGAGGATGAGATCTGGGCACCCATTAGTATGCTTCATTGTCCTAGATCAA  
TTACAATGCTCTCATCTTGAGAGGAGCAGAGGTTGGGTGTCACAGAGTGGCTAGAGAATGGTGAACATGGGAGCTT  
GACAGATGGCGTAGGCCCTGCTGTGAGAGCTGGCTGTGAGAGATGTCATGAGAGATGTATAGTCATTCTAAC  
AAAATGGTGGAGCTGCTGAGACAGAGTGGCCATGAGAGATGTATAGTCATTCTAACAGCTAGTATGTCATT  
ACAAGTGAACGGCTAA  
  
**EUGT-10** 누클레오티드 (SEQ ID NO: 46)  
ATGGAGGATGACACTCTCACTCTCCCTGCATGTCGGTATTCCTGATGGTGGCA  
TTGGCACACCTATGCCCTGCTGGATTGCTGAAGAGTTGGCTGAAAGGCCACGGTCAGTTTGTATCAACACCT  
AGAAACTTGCTAGATTACACCCAGTGTAGCCAGGAATTAGCTGAATTGGTTGATCTAGTCATTACACTCTCTAGAGTA  
GAATGGTTGCGACAGCAGAGCAGAGCTACATCAGAGTCGCAATTGACAAGTCTGAGTACACAGAAAGGCTTGTGAGGG  
CTAGCGCCATTCTGAGCTTCTGACAGCAGGCTGAGGGTAAAGGGAGATGTCAGGAGTGGCTAGTACACTCTGAGAGTA  
CACCCTGGTAGCTTCTGACAGAAGAGGAGTGTGTCAGGAGGAGTGTGTCAGGAGTGGCTAGTACACTCTGAGAGTA  
AGITCAGCACTCCACCGAACATCTCTGGCAGCAAAAGGGAGAACATTGTTAGATCAATTGGGAATCTGGCCTCATCATC  
GAAGCTAAAGAGAGACTCTGAGAGGTTGTCAGGAGGAGTGTGTCAGGAGGAGCTTAAAGAATCTTAACCTAGATTTCAGGG  
CAAAGATCCAAATTGGTGCATGAGGAGGTTGTCAGGAGGAGTGTGTCAGGAGGAGCTTAAAGAATCTTAACCTAGATTTCAGGG  
AAGCCTGTTGCCCCCTTGGGCTACTACTCTCCAGTCAGATGGTGCAGAGGGAGGAGTGTCTCCAAAATGGGAAGCATGATCC  
ATCATGCAATGTTGGATGCTCAACCGAAATCTGTAGTTTATGTTGCTCTAGTTGGCTGAGGCCAAATGCTCTGAGAT  
TATGAGAGAAATTAGCCCATGGATTGGATAATAGCTGTTACAGATCTCTTGGCTATGAGAAAACCTGCTGGTGTAGAC  
GCTGATCTCTGCTGCTGGTTGGGAGAACACTGGTGAACKGTTGGCTAGTAACTACAGAGTGGCTCACAGGTT  
GCTCTTAGACAGTGGCTAGCAGCTGTTGCTGAGGAAACTCTGAGGAAATGTCAGTGGCTAGTGGGAGGAGAG  
CATCCATTGATAATGCTACTTATGGTGTCAAGGTCAAATGCTAGAACTTGGGAGGAGTGTGAGGAGGAGGAG  
GTTGTAAGACCTTTTGTCAACGCCAGAAACTCTGAGGAAATGTCAGTGGCTAGTGGGAGGAGAG  
GTGTCAGTGGCTAGTGGGAGGAGAG  
TTGTGCAACATTGACATCCTCTGGAATGACTGAAATAACTCAGATGGTCAATTACCAAA  
TTGTGCAACATTGACATCCTCTGGAATGACTGAAATAACTCAGATGGTCAATTACCAAA

ATGGAGATGTTGGCTACTCTCTAATCTTATGCTGTCGCCGTGGTATGCACGTGTTGATCTGC  
CCTGGTGTGGCTTGGTCACTCTGTACCGTCTGGATTAGCCAAAAGACTGGCCATAGAGGCCATAGATATCATT  
GTGTCTACTCTAGAAATATCTCTCGTTACCCACCGTCAGACCTGCTCTAGCTCCCTAGTGCATCTGTTGCTCTCCA  
CTTCAAAGAGTAGAAGGATTTGCCAGACGGGCTGAATCTACTAATGAGTACCATGATAGACCTGACATGGTCAATG

도면6g

CATAGAAGAGCCTTGTATGGATTGGCAGCTCCATTTCTGAGTTCCTGGCACAGCATGTGCAGACTGGGTATAGTCGAT  
 GTATTCATCACTGGGCGCTGCAGGCCATGGAAACATAAGGTGCTTGTGCTATGATGTTAGTGAAGCAGCAGCACACATG  
 ATCGCATCCATAGCTGATAGAAGATGGAAAGAGCTGAACAGAACTCCCAAGCGCAGCAGGACAAGGTAGGCCAGCTGCC  
 GCCCAACCTTGAAAGTGGCTAGAATGATTGCTACTAAAGTAGTCAAGGATGAGTCCTGCTGAAAGGTTCTGCTGAAAGGTTCT  
 CTGACATTCTAGATCATCATCTAGTTGCTGTTGCTTAATGCTCCATTACATGAGGAAGGAGAGATGGTGAAGATGCTACT  
 TTGAGAGGCAAACCTTACTCTTCTGTTGCTTAATGCTCCATTACATGAGGAAGGAGAGATGGTGAAGATGCTACT  
 GTTAAAGGGTTAGATGCCAACCTGGCTGAGCTGGGTTGGAGAGAAAGAACCCGGTGTCTGAGAAGAAAACCACCGGTGTTCTGAC  
 GCGACTTGCTACCAGCTGGGTGAAGAGAACAGAGGCGTGGTGTGTTGACTAGATGGTCCACAAATGAG  
 ATTCTAGCTCATGCAGCTGTTAGGGCCCTTCAACCACTGGGTGGAACTCAACATAGAGGACTGATGTTGCTAT  
 CCACTTATTATGTTACCAACTTGGGATCAGGGACCTAACGAAAGGCTGCAAGTGAAGGAGTGGTCAAGGAGCTGCGTGGAGGTT  
 GCACGTAATGTGGTGAATGGTCTTGTATAGAGAAGGGCTGCACTAGAGCAGTCGCCATCAGGAGCTGCGTGGAGAAGGCTA  
 TCTAAAGTTCCAAGCTAAGGCCAAAAATTACAAGAGATTGGGCTGACATGGCTGTACCGAAAGATACATCGATGGT  
 TTCCATCCACACAATTGAGAAGTTATAAAGActaa

**EUGT-12** 누클레오티드 (SEQ ID NO: 48)

ATGGACACTGGTAGATCAAATCTTACCCGTTGCACTTGTGATCTTCTTCTGGCT  
 GCCTTGGGCAATTGTTGCCATCTGGAAGACTAGCTGAAAGAGTGGCTCAAGGGACATAGGTATCTTCTGTTAACC  
 CCTAGAAACCTAGCCAGACTTCCAGATGGCACAGAATGACGCCAAAGTGAATTGTAATTGTTGAAAG  
 GCTTTGATGCCTIAGCTGCACTTGGCAGTTCTAGGTGCAAGGATGATGCTGCCGGAAAGACCCAGACTGGATT  
 ATCGCTGATACATCATGGCACCACTAGTGTCTTGCACAAATAAGTGCCTGTCATGTGTTGCTCTGCT  
 TCTATGATGGCAGGGGGCTACTAGATCATCTGAACTGGCTCTATCTTCAAGTTAGGAGTCTGGTCAG  
 AAAGCAAGTTCTGTCATGATCTGGATCTGGCTCAGTGGAGTGTGTCAGCAAGAGGGTGTCTTACGAAAGGAGTACT  
 CTAGCTGCCATGAGATCTGCCAGAGTGGGAACCTGATGCTTCAACAGCAAGGTGGAGCTGGCTTGAAAGAACCTTA  
 ATACCATGGACTCTCCATCCAGGAGGAGAAGAGGAGCTGGAGCTGGAGATAATTCAACTATGGTTGG  
 TGGATGTTCAAGCTAAGTCGTTGTTATAGTCGCTCTGGGTAGAAAGTCTTGGCTTAGAAAGCTGCACTGAG  
 TTAGCTCTGGTTGGAAATTAGCTGGTACAAGGTCTATGGGCCCTTGAAGAACACAGCGGTAGATTTATCTGATGTT  
 TTGCCCCCTGGTACCAAGAGAGAACAAATCTACGGGCACTGTCGAATGGGTGGGTCTCAATCACTATCTGAC  
 CATGCACTGTCGGAGCTTCTACACATGGCGAGAAATAGTTAGTAGAAGGTTATTATTGGCAATCCACTGATC  
 ATGCTCCAATATTCGCTGACCAAGGGCAAACGCTAGATGATGGGGTATAAAGTGGCTCACAACTGAGTGGGTGAC  
 GATACTGACGGGATCTCTGATAGACAGCAGGTTGAGCTGCGTAAAGAGGCCGTTAGGTAGAAGGAAACTAGAGAGTT  
 TTGGCGCAATGCTTATAGCTACAGGAAITGGTGGCCGACAAGGAATIGCACGAGAGATACTTGTGAATTCTCAA

도면6h

CAACTGGTTTCTCACGGCGCAGACGGTAGTGTAAACACCGCCGACAGTGCCATCAAGTaa

**EUGT-13** 누클레오티드 (SEQ ID NO: 49)

ATGATGCAATTGGGGGTGCCTACTCCCTTAAGGTTAGTAACTGACAAACA  
 AAATTAAATTACCTGACCCATCCAAGGACCTAGCAGATTCCAAATTAAACATCCCAGAATCTTACCGCCTCAGAT  
 TTGAAACACTTAGGCCATCTGTTTATCATCAAGTTAAACAAAGACTGTGAATCTTAAAGAGTGTGTTGGTCAA  
 TTCTACTTCAGCAGAAGAGGAANTTGCTTGTTGTTCTATGATGAGTTCTATGATGAGTTCTATGAGTTCTGAG  
 TTAACCTACCAAAAGTTATTTCTACTGAAAATGCTACCGCTTITGCTTGAAGAGTGTATGTTAAGTTAGCT  
 AAAGAGGAATGCTCCACTACAGAAGGCTGGTAGGGAGGGACTAGTCCAGAAUTGCACTTCTGAGATACAAA  
 GATTGGCACACATGCGATTGCCCTGTTGAAGGCTCTGGGAGGTGTCAAATCTGAGTGGAAAGGGCACAGCATCT  
 TCTATGATTATCACACACTGCTCATGCTTGAAGAAATTACTTCTATGGAATGGTIGCAACAGAATTGAAAGATGCT  
 CCTATGGTCCACTGTAACATGGTAAGTCTGCCCCAACCACTTCACTTATGAAATGAAATGATGAAATGATCTGCT  
 AATAAACAAAGCCATCATGTCATACATTCCTGGTCATTACCTTACCTTATGGAAACTAAGGAAGTACTGAAATG  
 GCTTCGGATGGTATCAAGTAATCAATTTCTTGGCTTATTAGACCTGGTCTATCTGGTCTGAACTATCTTAAT  
 GAGGAATGTTTCTATGAGGATGTCGATGAGGGCTACATGTTAAATGGGCTACACAGAAACAGATTAGCACAT  
 GCCGCACTCGGAGCTCTGGTACACTGGAACCTCCTGGAATTCATGGGAAGGTATACCAATTGTTGG  
 TTGTCGCTTAATTAAGATaa

**EUGT-14** 누클레오티드 (SEQ ID NO: 50)

ATGGAAAAAAAAATGGAAAGCTAAACGGTAGATCGTTAGTACCAATTCTGCACAGGG  
 CAGTGACTCCATGTCATGGAAAGTTCTAACTCAAGGGATTTCATACAGTGTAGAAGGTCTATTCAAC  
 AAGCTGGGGGCATGGAATCTATGATCACACTTAAATAGACTCTGGGGCATTTAACAGATGAA  
 TTGCAACAAGGTAACGATAAGCTTACCAACCTGCTGCTAATTAGCTGAGTACATGTTCTCCATCCAGATA  
 ATTCCAAGTGTAAATATCTCTACCAACCTGCTGCTAATTAGCTGAGTACATGTTCTCCATCCAGATA  
 TTGATCTTAAAGATAACAAAGATTCAACTACCTGGCATGGTCACTAGCAGATCTCTGAGAGTC  
 GCAAATAGAGAGACTGCTGACATGAGCTCATCCATGAGAGACAA  
 AAGGTTGGAAATTCTCTGTTATCCCTTGACCTTACACATGACAGACTCATCACCCTCACTTATAGGAGAGATA  
 TCCTGTTGAAATGGCTGAATAACAGAAACCAAAACAGCTTATATTAGTATAGTACTCTGAGACAGATGGAAACC  
 AAGGAAGTTAGAAATGCTCTGGGCTGTTAAATAGTAAACCAACCTTCTATGGGTTATAGAGGGGTTATTTA  
 GGCACATAATGGGATCTTGGCGAAGATGTTATAAAATGGTCTGAGAGAGGGTACATGTTAAAGAGGGCACC  
 CAAATTGAAGTATGGGCCACCCAGCTGTTGGGGATCTGGTACATGGGTGGAACTCTCTGGAATCAATCGGT

도면6i

GAGGGAGTGCCAATGATCTGTAAGGCCCTTCACGGGGAACAGAAACTGAAACGCCATGTACATGAGTCAGTTGGAGAATC  
GGGTTCCAAGTAGAAGGAAAGGTCGACCGTGGGGAGTAGAGAGAGCTGTCAAAAGATTGATAGTCGATGACGAGGTGCA  
GGTATGAGAGAAAGGCCCTGGTCAAAGAGAAATCAAGGCATCTGTTCTAGTGGGGTCATCCATCATGCTTG  
GAAGATATGTGAACTACCTAAAGAACAAAGATGctaa

**EUGT-15** 누클레오티드 (SEQ ID NO: 51)

ATGAGATGAAACATCTAATGATGCCTTCACGTGTCATGTTCCAAATTGGCTT  
GGTCATATTCTCCATTGTCACACTGCAATAATGTCATCACAGCATGAGC  
AATGCACTAGAGTGAATCTATGCTAAATGCACTACAACACATTGTCACACTGACTTTGCTCATGAGAAGGT  
TTGCACCCAGGAGCTGAGGAGTACTGTCGAATTAAACCTCTTCACTAAACCTTCTGCTCTATGATTCGCTGCTCTGACCTTATGCAA  
CCTCAAAATAAGACTCTGTTCTCACTAAACCTTCTGCTCTATGATTCGCTGCTGAGGATGTTACCTAAATG  
GCCAACGGGGGGTAACTGCTACTACTGAGTGTGAGCATATCTACTGCTTCTACTGCTTCCATGCTCCAGCAAG  
GTTTGGAGCCTTAAAGTACCCCTCATGGAGATATGAAAAAACACCTTGGCTTCTCAACACATCAGTTACATC  
GTGAGAACCTTGAAGCTAGAGTATTCTATGATTCAACTCTTCCACACGGACTTACCCCTTACGACAGAA  
TCTGGATGAGGAGGTGTAGTGAATCTAGCCAAACATGCTCAAGAACATGGAAAGGTCATACAAATATGIGGAAGCT  
CAGTCAACAAACCCAGTGTGTTCTGACAGGAGCTGCTGACCCACATGGTAACTGAGGAGAAATGGCTTAC  
TGGTAATAAATTGAAAGGGTACTGTTAAATATTGCTTCTTGGCTGAAACTTCTTACCGAACAGGTTGAAG  
GAATTGGCACTAGGGCTAGAACAGAACGGTCTGCCATTCTCTAGTTAAATTCTCACTAACGGGATGTCAGTGCC  
GAATTGGAACAGAGCTTACCAAGGGTTCTGGAGAGGATAAAGGTATATCCACTCCGGGGGGTCAACAA  
CAACATATAATTGGCTCATTCTCATCAGGGGTCTATGCTCATGCCGGCTCTCTCTGTCATGAAAGCATTTGTTAAT  
GACTGTCAACTGAGTGTGCTGCCACAAAAGGGAGATCAAATTGATGCTAAATTGGTTCTGGGTGATATGGAAGCTGGT  
GTTGAAATTACAGAGGTGAGGAAGTGGCTACTTGGTAAGGAAGATAATTAGGGAGCACTGAGGAGCT  
GTGAAAGAGGCCAGGGAAACTTACAGGGAAATCAAAAAAGGGAAAGAGTTCTACTAAACAAAGGATATCCAACT  
AAGTATATCGTAACTTAGTGAATGACGCCATGGCTAAAGGTTCAACTACAtaa

**EUGT-16** 누클레오티드 (SEQ ID NO: 52)

ATGGAACCTACATTCAGCTTITATGTTTCATGGTTCGGACATATGATACAA  
TCCCTCATTAGCAACACAAGTGGAGAAAAGGGGCAAAATCACATTCTGTCCTTAAAGACACAAAGCAATT  
GAACACACATAATTATTCCAGATTCTCATAGTATTCTCATCCTTCAGTCAAGCATCCCTACAGTGAATGGGGTGCAGCAGGGTGC  
GAAACTACTAGTGTGATCTCAATCTCTATGGATAATTCTATGTCAGAAGGCTGACTTGGACAGAGATCAAGTGGAGCT  
GCCCTTAAAGGCTTGCTGCTGATTTGATATTCTGACTTGGACAGAGATCAAGTGGAGCT  
AAGTCGGTATCATACATGATTTGTTCCGCACTACAACTGCTTATACCTCGCTCTGGAGGAGTTGGGTGTACCA  
ACCA

도면6j

CCTGGCTACCCCTCAAGTAAGGTTTGTATAGAGAAATGATGCTCATGCTTGGCACACTTCTATTTACAAGA  
 STACACCAGATAACAATGGATTCAAATCTGGCACATCATGCTTAAGGACTTGCAACGAATCGAAGGCAAGT  
 TGTGACTACATCTCTCAATACATAAAAAGTATTACTGACGGTCATGTTACAGACAAGAATCTAACAAA  
 CTAGAAGGCAATTGCTCATTCTGTCAGAATTTCACCAAGGTTGGTTAGGACTTACAGGTTGCAATTCTGCT  
 CTGGAAAGGATCAATTCAAGAGGTATGTTAGGATGGACTTACAGGTTGCAATTCTGCT  
 AGAGGAGTCCACTGTGAGGAGGACTTCCAGAAGGTTCCAAGAAGAGTAAAGGTTAGAGGGTCTGGGGAGGG  
 TGCTTAATGACTGACTGTCAAATGGTTACTACCTTCTTGGTATGCAAGTACTATTCACAGTATGTTAGGAG  
 TTAAAGTCTCGTIGAGTCAGAGAAAAGACTGGCGGTTCTAAAGAGTACTGTCAGGACTCTGAGGAA  
 ATGGATAAGTAGTGACCTGGAAATGGTATGATCACCCTGCCAACACTTAAAGGAACACTTGGTCTACGGTTA  
 CTGACAGGTACGTGATAATTGTCGAAGAGCTACAGGAATATTGATIaa

**EUGr-17** 누름해오드 (SEQ ID NO.: 53)

ATGTCACAAATAACGGGACACCATATTGCAATGTACCCCTGGCTAGCAATGGC  
 CACATCACCTCCCTCTAGAATGACAAGCTGGCCAAAGAGGACACAGAAATTCTCTGCCACCTAAACA  
 CAATTAAGGATTAACCTGTTACCTGAAACCATTAACCAAGAATTAACCTTCTGATTCAGGTT  
 CCTAAAGGTAACCAATCAAGGGGCCACCTAATGTCCTCTGGTCAGAGCGCTTAACCTCTTCATTATCTGTAATAACA  
 CCAAGCTGAAGCAGAGCAACTAATGATATCTCAGCCATGCCATGTGTTACCATCTCCACATGACCCATAGAGCC  
 AACATGGAAATTAAACCTCTACTATATGTCCTCTGGTCAGAGCGCTTAACCTCTTCATTATCTGTAATAACA  
 CCAAGGCTGAACCTGTTACCTGCTGGCCATGGCATCAAACCTGAGAGAAATCTACATGTCATGGTAAATCT  
 TCAAGAGAATGIGGAGCCATGGCATCAAACCTGAGAGAAATCTACATGTCATGGTAAATCTACATGTCATGGTAAATCT  
 GGTAAAGCCTGTTAACCTGCTGGCCATGGCATCAAACCTGAGAGAAATCTACATGTCATGGTAAATCTACATGTCATGGT  
 GCATCTGGGGTCAACGGCTACTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCT  
 GTTGGGTTGGAAATTGACTGGCTTACCTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCT  
 TTGCGAGAAGGATTCCTGAAAGGACAAAGGATAGGGCATCGTTATGGGGTGGTGCACACAACTTAAATTGAG  
 CATCCATCGTAGGTTCTGTCATGTTACCTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCT  
 ATGATCCCACATGCTGTTACCTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCT  
 CGTAGGAGGAGATGGGAGATGTGAGAAGGAGCTAGTAAGGAGAATCTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCT  
 AAGGAAGTATGGCAATCAGCTAAATGGAGAGAATTACTAAGAGGAATTTGAGGAAATCTGTCAGGCTTACCTGTCAGGCTTACCTGTCAGGCTT  
 ATTGACAAACTATGACTTGTGAGAaa

도면6k

**EUGT-18** 누클레오티드 (SEQ ID NO: 54)

ATGGGGTCACAAGCACACAACTTATCAGGCAATGTATCCATGGTTGGCGTGGGACAC  
 TIGACTGGCTCTCTCAGACTGTCTAACAACACTGGCGGGAAAGGTCATAGAATTCTAATACCTAACAGAACACAA  
 TCAAATGGAAAGTTTAATTACACCCACACTTATCATTTGCGCAATCGTAGTACCTCTATACTGGTTGGCT  
 CCAGGGAGCTGAAACACCTCCGAGCTCTTCCACACTCTACTCTTTAATGGAAAGGCATGGACAAGACACAAATGAT  
 ATGGAATTATCTGAAAGGATTGAAAGTTGATGTGTTAGCTTACCCATTGGTACTCTAGTTAGCTAGAAAG  
 ATTGGTATAAAATCTGATTCTACTCTACAACTTACCCATTGATGCGATGGCTACGGCTTATCTCCAGAGAGAGAGTAGT  
 GGGAAACAACTTACAGAAGCTGACATGATGAAAGCAGCAGCTTATCCAGACCCCTATTAAGCTACATGCTATGAA  
 GCACGTGGATTACTCTGAGCTTGTGATGAGTTGGGTGATATAACCTTGTGATCTCACAGCAGTTCT  
 GAATCAGATGGCTGGCATACTCAACTTGTAGAGAAATTGAGGTCATCTGACTACATGAAAGAACACAAATCCAAAG  
 CCACTCTTATTAGCAGGCCCTGCACTTGCGCTTGGTCTGAATGTACTTGAGAAAGCAGCTTATGCAAGAGCTGTTATGG  
 TTAAAGAAGGGTCTGGTAATCTATGCGCTTGGTCTGAATGTACTTGAGAAAGCAGCTTATGCAAGAGCTGTTATGG  
 GGATTTGAACTGACTGTGATGCCATTTCGCTGCTTGAAGGCTCCCTTGAAGCTGAAGTGGTGAATGGTCAACAGCAACTGTT  
 GAAGAGTTAAAGGAAAGATCCAGGGGAGGGTATGTCACGGTGAATGGTCAACAGCAACTGTTACACATCCA  
 TCACTGGCTGCTTGTCTCAGTGGTGGCATCCCTTAGTGAGGCTTGGTAAAGGCTGAGTGTCAAATCGTCTATA  
 CCACAGGGAGATCAAATCAAGCAACTTGTGAGGCTTGGTAAAGGCTGAGTGTCAAATCGTCTATA  
 GAGGAGGAGTCTCAAGAGATACTGTTGAGGTTGGTAAAGGCTGAGTGTCAAATCGTCTATA  
 GTCAAGGGAAACACGATAATTAGGGTTCTGCTTAATGCTGATCTAGATCCAAATACATGGATTCCTCAACAA  
 AAGCTACAGGATCTCTAGGAtaa

**EUGT-19** 누클레오티드 (SEQ ID NO: 55)

ATGGCCGCCGTGTTGAAGCTGATGAGCAGGCTATGCGTGTGCTTGTCCCCTT  
 CTGGCATTCGGACATATATCACCATTGCCAGTTAGCAAGAGATCCTGGGCTGTAGGGGTTGAGGTTGCTAGAGTTACTTCT  
 TCTGCTGAGCCAACTGCTGCGACTGCGGAATCAACAGCTGAAGTTGATGAGATGGGCTGAACTGTTAAATAGCTTA  
 AGAGTGCCTGGTTACCACTGAGTGGTGGCGGAATCAACAGCTGAAGTTGATGAGATGGGCTGAACTGTTAAATAGCTTA  
 GACGGAAACAGACCAACTGAGTGGCGGAAGGGCTGAGGCTTGGGCTGAGGCTTGGGCTGAACTGTTAAATAGCTTA  
 GTAGCTGATGTTAGGCGAACACTAGTGCAGAGGGCTACTTCTGTTACTGCTACTTCTGCAACTGTTAGGCTTGGGCTGA  
 GTGCCCTGACGTAGGAGACTACATCTGGGCTGCACTCTGCTTACCTTCCAGCTGCAACTGTTAGGCTTGGGCTGA  
 CCACCTGATCTCCCTAGCAACTTCAAGCTGAGACTTACTTGTGTTACCTTCCAGCTGCAACTGTTAGGCTTGGGCTGA  
 CCTATATTGAAATACGTGCTACACAAATACGATAAGCCAATCTGAGACTGGACCTCTGTTCCAGAACCTCACATGGG  
 GAACCTGAAGAGATGGGAAACATGGTGTCTCTTCCAGATAATGCACTGGGTTGGCTGCAAGTTGGCTGAAACC

도면61

TTCTGCCAACAGCCGCTGCTACAGAGCTGTTACTTGGCTTAGAAGCAACAGGTCAACCATTGTCGCAAGTACTTAATTC  
CCTAGGCTCTGTCGATGCTGAAGCCGAAGTAAAAAATGATGCCCCAGGTTTGAGGAAGAGTGAAGAGTGAAGGTAGAGGTT  
GTACACAGGGTGGTCAAACAGCAACATATCTTAAGACACAGGTCITGTTGCTAGTTAACAGCTGGATTAGT  
TCTGTTGCGAAGGACTTGTGCTGGGTAGACTAGTATGCTACCAATGAAGTCAGATCAATTCAATGCTGGCTTGG  
CTGGCCAGAGAAATIAAGAGTCGGACTGAAGTAGCTGGACGAGTGGGACGGTTGGCTGACGCTGTTAGAGAT  
GCTGTAACGCTGCACTGCTGATGCTGGTGCCTGGAGATGATGAGAGAGAATGGAGAGAATGGAGAGAATTCACAGATGATGCT  
GTTCAACGTAGATTGTCGAGGGAGTTAGAGAGTTGAGAAGATGTTGGTGTaa

### 도면7a

누클레오티드 (SEQ ID NO\_ 56)  
 ATGTGACCGGAAACAAATGCCACATCTCTCCCTGTCAGCACAGGCCATATCAACCCATT  
 TCCATTGGCAGGACTGGCCTCTAACAACTCTCACCAACTCTGTCAAACCCGGTTCTCTCAA  
 CTCGACCAAAATCGGACCCGGTCCGGTCAACATCCAGTGAATCGACGGGATCTGCATCCGGGGAT  
 ATATGGCAGCACCTACTGTCGGGTCAGGGGAGGAGCTTAAGGTTCCGGTTGCCGAGCCGGTCCGGTTGCCGGAT  
 ATCTACCGCCACGGTGGGAGGGACAAATTAAAGGTTCCGGTTGCCGAGCCGGTCCGGTTGCCGGAT  
 CCCCACTCGAGCCTAGTGATCTGCCGCGTTGTAACGGATTGGACAGGAGTCGTAAACCCGGATCTCT  
 GCCACTCTGGAGTGAATCAGCAAAATTGGACAAAGCTGACATGATGGGGAGGAACTCCATATACGAG  
 CTAGAGGGGGATTGTTGGATGGGTCCGATTACCATTAAGGGGAATGTCGATGGCCGACCGTGGCGT  
 CTACTTATCTCGACAACCGAATCCCGTGGATTACACATACTGGATTCAATCATACACCCCTGACACTAC  
 TCCCTACTTGGACTTGGCTGACTCCAGGGCCCGAACTCTGTCATCTATGTCATCTGTCATTGGTAGTTGHC  
 AGTCTGAGCCCTGATCAAACCAATGGAGTGGCTGGGTCTGATGCCACCAACAAAGCTTATTTGG  
 TGGTAGCAGCTTGGAGTTAGCCAAACTCCCTGCAACTTCACCCAAAGAGAATCGAGCCGGTGGGGTGG  
 TGTGACCTGTGGACCAACTCGATCTCTGCAACAGTGGACCCGGTTGCTGTGACTTCACCTGGGA  
 TGGAACTCGACTATGGAGGGTGTGGACTTGGTGTGGAGGGTGGGGAGTGGCCGAGTGGTCCGGACAGC  
 CTATGAATGCCAAGTATGTAGAGGTGTGGAGGTGGGGAGACTTATGGGAAGGATT  
 TGTGAGAGGGAGGAGGTCAAAGATGTGTGGAGGAGGTGGATGGAGAGGAGGAGTCTCTGAGAAGTGA  
 GAGATGCGCAAGGGTGTAGTTGCCAAGGGACTCTGTGTAGTGGGGGGAGTTCTGACAAGTGA  
 TCAAGGAGTTCATACATCAGTGTGCAATGACTCCAAATTCCTAGTTA

공개된 단백질 서열 (SEQ ID NO: 57)

### 도면7b

MUNGNICKHTILLIPCAQHNNPILQFGKRLASHNLITLTVNTRFLNSNKSEPGVNNICISDGFDPGGWNNAESRRAYFDRFQSRSGKHGVLIESLRSGRGAC  
FGLRPPVLMAMNVAERSGLSVAFTTOPCAVADTYRHMEGRIKVVAEPVRPLGPLPLEPSDLPCVNRGFRVNNPDILPIRVNKHNLKADMGRNISTYLEAD  
LIDGSRLPLVKPSIGTPTSYLNRPSSDHGFNLYTPDTPTPYLWDLSDKAPNSVIVSFGLSSLPDQEINASGLITANKSFIWVRTSELAKIPANFTQEN  
ASRGIVTWCBQDLDIAYATGCFVTHGMWNSTMEGVALGVEMGVPOWSQPMIAKYEDWVKVGRKACTYKDFVRGEFFKRCVEEMDGRSGKIRENAWRCK  
LAKDSVEGSSSDKCIKEFHQCQNDKSLV

합성물 Cs<sup>+</sup> UGt2 투플레오티드 서열

( 소스 : DNA 2.0) 누클레오티드 (SEQ ID NO: 58)

도면8a

**CH2: (Accession №: DQ201828) ㅋ螺旋形 (SEQ ID NO: 59)**

cccggggaattcaagcttaaatgtttcatgtttggcgttgtaaactgggacatgggtggctgcatttccgtggcaaggatattgc  
 TTTCTTCAGTTGTGACCTAGCACCTAGAAGAGATCATCCTTATAAATGGCAAGGGCCAATACAGAGAGAGTGGTCAACTCTTAGATTGCTGGAAATT  
 GCTGGGCTCCAGGTGACTCAACTGACCAGGGAAAGGCTTACAGCTGAGGGATGGTTACAGCTGAGGGAGCTACAAGAACATAGAGGATCATGA  
 TACCATTTATGCCAAGCTGTTAAGTGTATGGTACCTATCTCTAGAAAGGTATGGAATTAGTAGACAAGATGATGGAA  
 GATGCAGCTGAGAAGATATGGCTGGGTGAATCTGAGGGAGAAAAAGGCTACCGATGGAAACTGAGGGGTGGACGTTAA  
 AGATTTGGTGGGGAGGCCACATTAGATGTTATGGCTATGCTTACAGCTGAGGGATGGCTTGCCCTTACATTAGATTGCTTACAAACAGACTAATG  
 AACATTATGTTGCTTGTGGCCTGCTTACAGCTTACATTAGATTGCTTAAAGCCATTACTGAGCTTCTAGAAGGGATTAGGGACTTGTACCT  
 TACITTAAGAACATGAAGAGAAGGCATGAGATCCATTAACTCAAGGACTAGCAGTTCTAGAAGGGTAGGTATTGATGAGTGA  
 GGAGAAAACAGCTGTTAGGTCTGCTGAGTAAAGGCTAGGGTGGGACAAGAAAGATGTTCAAGGGAGATAACTTCCTTG  
 TTGTTAGAGCCATTATCGCTGCTAACCTGCCAGAATCACAGAAATTGAGTGTAGAGGACTAGCACAGGTTCTAGACTG  
 TTGCGAGGTTACGAAACATCCACAGCTTAACTGGTGTAGACTGCTTACAGCTTAAAGCCATTAGGATGACTGTTG  
 TGAGGAATCTGTCATTGCAACAGACATGCCACTTGTAGAGGATTTGAGGATTTGAGGATTTGAGGATTTGAGGATTTGAGGATTTGAGG  
 CATTAGCTAGACCTCTTCACTTATGCAAAATAGAGAATGCTTAAGGATGAGTGTAGGTTACAGCTTACCTTACCTT  
 GGAAGAGAGACGGTAGTGTATCAATGAAGTCAGAATTAACAAAGGCACAATGGTTAGGTTAGGACCAAGAGATGGTTAGG  
 GTTTATCTACGGTAGGACGCTGAAGAGTTAGGACCAAGAGATGGTTAGGAGTTGCGCATTTGAGGATTTGAGGATTTG  
 CATAAGGACATCAGGCTTCATCTTCTGGACCTAGAGCTTGGGAGATGGGAGTTGCGTTGAGGATTTGCGTTG  
 TTGTAACCTAGAAGGGTCAATTGACCAATCATCTCACACCCAGAATACGAAACATATCACCTTGTATCATATCTAGAC  
 TATAGTTGGTAGAGAGAAGGAGTTATCAAATGAGACTACAAGTGGAAACCGTAGAGTaaacgcggctcgagccgg  
 ggg

**CH3: (Accession №: NM\_124636) ㅋ螺旋形 (SEQ ID NO: 60)**

cccggggaattcaagcttaaatggcgcgggtctgcacaattggccgtactttgaagccactttatcggttcctcattttctgc  
 TAATCCTATATCTACTGCGCTTCCACCTTCAATTAGATTCAGGCTTCAAGGGAAATCTTACAGTATGCTTC  
 TAGTCGAGGAACGTAACAGTCCCTCCAAATGGACGATGATAACAAACCTGAATCTACTACAAGTTCATCTGAATTC  
 ATGACA

도면8b

ZC01: (Accession No: AJ489276) 누를메오티드 (SEQ ID NO: 62)

ccgggggaattcaagaacctaaATGCAAGTGGACCCCTACAAAGGAATTCTTACAGGACCA  
ATTACATGCATGTAAGTACTGCTTACTGCTGTGAGATAGTTCCTCAGGAGATTTGGTAGATAG  
AGAAATACCGTCAACTAAGTCTCACACAGCCAAAGACCAGTCCTACTGGGAAACTTCTCTTCTACGGTCA  
ATACAACCCATAGTACTTACTCAAGATAATGTTGTGATGGCAAAGCTGGCCAGGATGTCCTCATTCAGGAACCC  
AAGTTTGTACATGTTGCAATATACAGAGCACTATGCTGTTTCCAGATAACAGATCGTGATGAAACCCAGCTGAATTGTTA  
GAGGGAGAGGTATGATGGTCCAGACCTGGAAAGGTCTCTAGACTTAGGCTTTACCAAGATATGCTACATCCGATAGTGAATG

## 도면8c

도면8d

ZC04: (Accession No: AB247160)

(SEQ ID NO: 64)

## 도면9

크로쿠스 씨리부스 Cs VrUGT2  
 합성된 누클레오티드 서열 누클레오티드 (SEQ ID NO: 65)  
 ATGGAACAAAGGAGTCACGGTAACAGTGTCACATCTTATTATTGCGCATGCCCTGCACAGGTACATCACAC  
 CCTATCTAACATTGGTAAAGATTTGGCTCACATAATTGTTGACTACATGGTCATAATCAGGATCTTGCC  
 AACAGTACTAAGTCTGACCAGCAGTCTGTTACATAGAATGGAACTCGTTGGTAGTGAACATTGTCGTTGATGAA  
 GCTGCACCATCTGAGAGCATATTGATAGATGGAACTCGTTGGTAGTGAACATTGTCGTTGATGAA  
 TCTTAAGGATCAAGGGTAGACCGTGCCTGGCCATGTTAGTCTAGTCTAGCCATTGCGCATGCCATTGGCTATGAA  
 GAAAGATCAGGTTGAGATCAGTTGCTTACACACCAAGGCGCAGTCAGTACCCATTGCGCATGCCATTGGCTATGAA  
 GAAGGTAGAATTAAGTCCAGTACCGAACCTGTTAGATTGCCAGGGTTACCACTTACGATCCATTGCGCATGCCATTGGCTATGAA  
 CCAAGTTTGTACTGATCTGACCTTGTTAGATTGCCAGGGTTACCACTTACGATCCATTGCGCATGCCATTGGCTATGAA  
 TTAGATAAGGCCGACATGATGTGATCAACTCAATCTATGAATTAGAACAGAACAGAATTGATGGATGGAATCC  
 AGATTGCCATTACCTGTTAACCTATGCGTCCACAGTCCCTCACCTACTGGATAATAGAACATCCATTGAC  
 TCACATTACGGTTAACCTATACACTCCAGATAACCTACCTCCATTGCGTCAATTGCGATCCAGGCAACATAAGAAC  
 AGTGTAAATATAGCTTCCTCGTAGTTACCTGCTCTGATGCAACTAATGAATAGCATCCGGTT  
 ATCGCCACAAACAAAGTTATCTGGTAGATACTGCAATTGCGAAGTCCAGGCTAATTCACTAA  
 GAAACCGCTTCAGAGGTTAGCTGTAACATGGTTGATGCAATTGCGAAGTCCAGGCTAATTCACTAA  
 TTGTGTACACACTGGGGTTGAACTCTACATGGAGGTATGGCATGGTGTGACCTGCA  
 TGGTC CGATCACCAATGACGCCAAATATGTCGAAGAGCTATGGGTGTGACCTGCA  
 AAAGATTCTGTTAGAGGTGAAGGTTAAAGATGTGTGAAGAAGTCTATGGACGGTGTGAAAGAAGTGGTAATA  
 AGAGAAAATGCCGCTAGATGGTCAATAGGATCTGTTCTCAGAAGGTTGTTCTCAGACAAGTGTATC  
 AAAGAATTATCCATCAATGCTGTTAGTAA

## 크로쿠스 씨리부스 Cs VrUGT2 단백질 (SEQ ID NO: 66)

MEQKDVGNGKCHILLPCPAQHINPILQFGKRLASHNLITLVNTRFLSNSTKSEPGPV  
 NIECISDGFSGGMMAAPSRAYRFLRSRGRPRPAHLWYDPPEFLPN  
 AMNVAERSGIRSVAFFTOPCAVDTIYRVWEGRIRKVPTEPYRUPGLPPLPLEPSDLRPSFVT  
 DSDPVVNPDLPILLVNVQHNLKDADMMLNSYELEHEEDWMESRSLPFLPVRSIGPTVPS  
 TYLDNRIPSISHYGFNLYTPTDTPYLWDLSDKAPSIVYWSFGSLLSISPDQTNELASGL  
 IATNKSFIMWVRTSELAKLPANFTQENASRGGLVWTWCQDOLYLAHVATGCFVTHCGINST  
 MEGIALGVPMYGVFQWSDQPMNAKYVEDVMKVGVRAKTYGKDFVRGEFFKRCVEEVMDGE  
 RSGKIRENAARMCKLARDSVSEGSSDCRCIREFIHQCK.\*

## 도면10a

**Cs UGT2 (460)** 아데노신. 전장 (SEQ ID NO: 57)

MLNGNICKHILLPCPAQGHINPILQFGKRLASHNLTLTUVNTRFLSNSTKSEPGPVNIQCIISDGFDPGGMNAAPSRRAYFDRQO  
SRSGQQKHVGIESLRSRGGRPGACFGGLRPVPLWAMNVAAERSGLRSVAFFTQPCAVDTIYRHVWEGRIKPVVAEPVRLPGLPPL  
EPSDLPCVRNGFGRAVVNPDLLPLRVNQHKNLKDADMGRNSIYELEADLDSRSLPLPVKSIGPTVPSIYLDNRIPSDSHYGF  
NLYTPDITPYLDWLDSKAPNSVIVVSEGSSLSPDQTNEIASGLATNKSFIWVVRTSELAKLPANFTQENASRGLVNTWCD  
QLDLLAHVATGCFVTHCGWNSTMMEGVALGVPMVGPQWSQDPMNAKYVEDVWKVGVRAKTYGKDFVRGEFFKRCVEE  
VMDGERSGKIRENAARWCKLAKDSVSEGSSDKCIKEFIHQCCNDISKLV#

**Cs Vr UGT2 (459)** 아데노신. 전장 (SEQ ID NO: 66)

MEQKDVGNGNICKHILLPCPAQGHINPILQFGKRLASHNLTLTUVNTRFLSNSTKSEPGPVNIQCIISDGFDSSGMNAAPSRRAY  
FDRLESVGSETLFRGLIESLRSRGRAHVLVYDPFLPWAMNVAAERSGLRSVAFFTQPCAVDTIYRHVWEGRIKPVVTEPVRLPG  
LPPLESIDLPSFVTDSDPVPVNPDLLPLLVNQHKNLKDADMMLINSIYELEHEFFDWMMESRRLPLPVKSIGPTVPSIYLDNRIPSD  
SHYGFNLYTIPDITPYLDWLDSKAPNSVIVVSEGSSLSPDQTNEIASGLATNKSFIWVVRTSELAKLPANFTQENASRGLVV  
TWCQDQLDVLAHVATGCFVTHCGWNSTMMEGIALGVPMVGPQWSQDPMNAKYVEDVWKVGVRAKTYGKDFVRGEFFKRC  
CVEEVMDGERSGKIRENAARWCKLAKDSVSEGSSDKCIKEFIHQCC#

Score = 835 bits (2158), Expect = 0.0, Method: Compositional matrix adjust.  
Identities = 410/453 (91%), Positives = 417/453 (92%), Gaps = 2/453 (0%)  
Source: NCBI blast

**Query - Cs UGT2**  
**Subject - Cs UGT2**

Query	2	LNIGNICKHILLPCPAQGHINPILQFGKRLASHNLTLTUVNTRFLSNSTKSEPGPVNIQCI	61
Subject	6	+NGNICKHILLPCPAQGHINPILQFGKRLASHNLTLTUVNTRFLSNSTKSEPGPVNI+CI VNGNICKHILLPCPAQGHINPILQFGKRLASHNLTLTUVNTRFLSNSTKSEPGPVNI	65

## 도면10b

Query	62	SDGFDPCGMNAAPSRAYFDRPQSRSGOKHVGILIESLRSRGPGACFGLRVPLWAMNV SDGFD GGMNIAAPSIRAYFDR + S + GILESIERSGRP Sbjct	66	P WAMNV P WAMNV
Query	122	ERSGLRSVAFTOPCAVDIYRHMEGRIKVPAEPYRPLGIPPLPSDLPCVRNGFRV ERSGLRSVAFTOPCAVDIYRHMEGRIKVPAEPYRPLGIPPLPSDLPCVRNGFRV Sbjct	126	V 125
Query	182	VNPDLIPLRVNQHKLKADMIGMGNISIELEADID--GSRLPLPKSIGTVPTSYLDN VNPDLIPLRVNQHKLKADM NSELE + D SRPLPKSIGTVPTSYLDN Sbjct	186	V 185
Query	240	RIPSDSHYGENLYTDTPTYLWLDISKAPNSVIYVSFGSLISSLSPDNTETASGLATNK RIPSDSHYGENLYTDTPTYLWLDISKAPNSVIYVSFGSLISSLSPDNTETASGLATNK Sbjct	246	239
Query	300	SFIWVVRTSELAKLPAFTQENASRGLIVWTWCIDOLILAHVATGCWVTHCGWNSTMGEVA SFIWVVRTSELAKLPAFTQENASRGLIVWTWCIDOLILAHVATGCWVTHCGWNSTMGEVA Sbjct	306	305
Query	360	LGVPVMYGVPOwSDPMNAKVEDWVKVGRAKTYKGDFVRGEEFKRCVEEMDGERSGKI LGVPVMYGVPOwSDPMNAKVEDWVKVGRAKTYKGDFVRGEEFKRCVEEMDGERSGKI Sbjct	366	365
Query	420	RENAWMCKLAKDSSEGSSDKCIKEFTHQCC RENAWMCKLAKDSSEGSSDKCIKEFTHQCC Sbjct	426	452
				458

## 도면11a

## &gt;ALD2 누클레오티드 (SEQ ID NO: 67)

ATGCCCTACCTTGATACTGATATCGAAATCCCACAATTGAAAATCTTTAAAGCAACCG  
CTAGGGTTTTATCAACAATGAGTTTGTCATCATCAGATGGAAAGACCATCGAAACT  
GTGAACCCAGCTACTGGCGAACCGATAACATCCTCAAGCAGCTAACGAAAAGGATGTA  
GACAAGCTGTGAAAGCTGCCAGGGCTGCTTTGATAACGTTGGTCGAAGACATTTCT  
GAGAACGTGGTATTATCTTCAACTTATTAAACATTATTGAGGGAGGAAGACACA  
CTTGGCATTAGGAGACTTAGCCTGGAAAGCCTTACCATTAACATGCAAAGGTGAT  
TTGGCACAAATTITACAGCTTACCATGATTTGCTGGTCCGCTGATAATTGGACAAA  
GGTGCACACCATAACCATGACTTTAACAGTTGCATATACTCTAAAAGTCCTTTGGC  
GTTGTTGCTCAAATCGTCCCATGGAATTATCCTCTAGCTATGGCTGTTGGAAATGGCAA  
GGTGCCTTAGCAGCGGTATCATCAACACGTTGCTGAGAATAACCTCTCTATCT  
CTACTTATTGCTACTTTAAATTTAAAGCAGGTTTCCACCTGGTGTGTAATATC  
GTTCTGGTTATGGATCACTGTAGGCCAACGCTAGCATCTCACATGGATATGACAAA  
ATACTTAAACGCTTACACTAGAATGCCGTGTAAGCTCTCGCTCTCGTATTTGAAGAT  
GCAGACCTTGATAAGGCTATCGATTGGATAGCAGCTGGCATTTTACAATTAGGACAG  
AATTGTAACGCCAACCTCAAGAGTTATGTTCAAAGTTCGATCTACGACAAGTTGTTGAA  
AAGTTAAAGAAAAGGAGTTGGATGGTGGAGAAAAATTGATCCGTTTGT  
GAGAAATCAAGTGTGGTCAGTTATCAAGTACACAGTATGCCGATCAAAGTTAC  
ATAGAACGCTGGAAAAGGGAGAAAAGTTGGACATGTTCCAGACCTCTGAATTCTTATT  
GGTAGGAGCTAAAGCTACTTCACTCCCCAACCATCTCACTGATGTCGGCAAACATCG  
AAACTGTTACAGGAGATATTGGCCCGGTTGGTTGTTAGCAAGGTTCAAATTAT  
GATGACGCTCTGAAGCTGGCTAATGATACTTGCTACGGGCTGCCCTGCGGTCTTCACA  
AAAGATGTCAGAAAGCGCACATGTTGCTCGCGATATTAAAGCAGGAACGTGTTGGATC  
AACTCATCTAACGATGAAGATGTTACCGTTCTTTGGCGGGTTAAAGTGGTATT  
GGTAGAGAACTGGGCAAAGTGGTGTGATACTTCAAACAAAAGCAGTTACATA  
AATCTCTTTGGACACTAA

## &gt;ALD3 누클레오티드 (SEQ ID NO: 68)

ATGCCCTACCTTGATACTGATATCGAAATCCCACAATTGAAAATCTTTAAAGCAACCG  
CTAGGGTTTTATCAACAATGAGTTTGTCATCATCAGATGGAAAGACCATCGAAACT  
GTGAACCCAGCTACTGGCGAACCGATAACATCCTCAAGCAGCTAACGAAAAGGATGTA  
GACAAGCTGTGAAAGCTGCCAGGGCTGCTTTGATAACGTTGGTCGAAGACATCTCT

**도면11b**

GAGCAACGTGGTATTCTTCAAACTTAATTAACGAGGAGGAAGACACA  
 CTTGCCGATTAGAGACTTAGACGCTGGTAAGCCTTCATTCAAATGCTAAACAAGAC  
 TTAGCCCAGATTAGAACACTAACAGATACTATGCAGGGCGGTGACAAGTCAATATG  
 GGTGAAACCATTCATTGACTTTAACAGTTGCACTATACTCTAAAGTCCCTTGGC  
 GTTGGTCTCAATCGTCCATGGAAATTATCCTCTAGCTATGGCTTGAGAAAATGCAA  
 GGTGCTTAGCGCCCGTAACCGTTATCATCAAACCTGCTGAAAATACCTCTATCT  
 CTAACTTTATTTGCTACTTTAACAGTTGCACTATACTCTAAAGTCCCTTGGC  
 ATTCCTGGTATGGTCCGGTGGGGAAAGCTTAGGAACCCACATGGATATCGACAAA  
 ATATCTTACGGGAAAGTACTAACGGTTGGCGCTAGTATGGAACTTCGGCCATCG  
 AACCTTAAGGATATCACACTAGAATGCGGTGGTAAGTCTCTGCTTGTATTGAGAT  
 GCAGACCTGATAAGGCTAGAATGGTACCAAATGGTATTTCCTTAAATTCCGGGACAG  
 ATCTGCACTGCAAACACTAACAGTTATGTCAAAGTTCGATCTACGACAAGTTGTTGAA  
 AAAGTTAAAGAAACTGCAAAGAAGGAGTGGGATGTTGAGGAAATTGATCCGTTGAT  
 GAGAAATGCACTGTTGCTCCAGTTATCAAGTACACAGTATGACCGCATCAAAGTTAC  
 ATAGAAACGTTGAAAGGAGGAAAGTGGACATGTTGAGACAGCTCTGAATTTCTTAAIT  
 GGTGGAGCTAAAGCTACTTCTTCCCCCAACCATCTTCACTGATGTACAGAACATCT  
 AAGTGTGCGTGTGAAATATTGGCCCGGTTGTTGAGCAAGTTCACAAATTAT  
 GATGACGCTGAAAGCTGGCAATGATACTTGTACGGGCTCGCCTCTGCGGTCTTCACC  
 AAAGATGCAAGAACGGCACATGTTGCTCGGATATTAAAGCAGGAACGTGTTGGATC  
 AATCAAACCAATCAAGAAGAACCTAAAGTCTTGGCGGATTAAAGATGAGTGGTATT  
 GGTAGAGAATCAGGGCACACGGCGTTGATAACTATTACAATAAAACTGTCATGTG  
 GATCTTCATTGGATAATAA

&gt;ALD4 누클레오티드 (SEQ ID NO: 69)

ATGTCAGTAGATCTACGCTCTGCTTAAAGACGCTGCTGCATCCTCATTGGAGACTTCAA  
 TTGAGATATTCTCACACCTCCTATGACAGTGCCTATCAAAGCTGCCAATGGGTGGAA  
 TATGAGAACAAACAGGGTTGTTCATCAACAAACAAGTTGTTCTCTAAACAGAACAAAG  
 ACCTCGAAGTCTTCAACCCCTCCACGGAAGAAGAAATATGTCATATTATGAAGGTAGA  
 GAGGACGATGTGGAAGAGGCGTGCAGGCCGCCGACCGTGCCTTCTCAATGGGCTTGG  
 AACGGTATGACCCATTGACAGGGTAAGGCTTGTACAGGTTAGCGAATTAAATTGAA  
 CAGGACAAAGGATGTCATTGCTCATCGAGACTTTGATAACGGTAAAGCTATCTTCC  
 TCGAGAGGATGTTGATTAGTCATCAACTATTGAAATCTCTGCTGGCTTGTGAT  
 AAAATTGATGGTAGAAATGATTGATACGGTAAAGCCATTTCCTTACACTAACAGAACAG  
 CCTTGGGTGTTGGGCAGATTTCCTTGGAAATTCCCACTTGTGATGTGGCCTGG  
 AAGATTGCCCCTGTGTTGCTAACACCCGCTGACGGTAAAGCAGACATGAGAAATTGTC  
 CCATTGTCGCTTGTGTTGCTAACATCCACAGGGGTATTCCACCTGGTGTG  
 ATCAACATTGATCCGGGTTGGTAAGATTGTTGGCAGGCAATTACAAACATCCAAAAA

**도면11c**

ATCAAAAAGGTTGCCTTACAGGGTCCACGGCTACGGGTAGACACATTACAGTCGC  
 GCCGAGGCTTGGAAAAAGTGAATTGGAGCTGGGTTGAAATCACCACATTGCTTC  
 GCGGAGCCGAGTGGAAAAAGCCGTGCAAAACATTATCCTTGGTATCTACTAACATTCT  
 GGTGAGGCTGTGCTGGGTTCAAGGGTGTATGTTGAGAAATCTTACGACAATTC  
 ATTGAAAGAGTCTAACGGCTTCTGAAATCCATCAAGTGGCAGCCATTGATGAATCT  
 ACTTCCAAGGTGACAAACCTCTAAATGCAACTAAACAAATCTTGAATACGTTGAC  
 ATTGTTGAAAGAATGAGGTGCTACTTGTGATTACCGGTTGGTGAAGGATTAGGTAGCAAGG  
 TACTTCAATTAGCAACTGTTGCTGGTCACTGTAACCAAATTCTGGCAGAACGTCATTAAAC  
 ATGGCGAACGATTGCAATACGGGTTGGCTGCTGGTATTACACCTCTAAATTAAACC  
 GCCTTAAAGTGGCTGATAGAGTTATGCGGGTACGGCTGGATAAAACACTTAAACGAT  
 TCCACACCGAGTCTTTCGGGGTTCAATGCACTGGTTGGCAGGGAAATGTCT  
 GTTGATGTTTACAAAACACTTGCAGTTAAAGCGGTGCGTGCACATTGGACAGTAA

&gt;ALD5 누클레오티드 (SEQ ID NO: 70)

ATGCTTTCTCGCACAAGAGCTGCAGCTCCGAATTCCAGAATATTCACTAGAACGCTGTTA  
 CGCTCTTATTCTCAAGCACCATTACCGCTTCAATTACTCTTCAAATGGTTCACCTAC  
 GAAACGCAAACAGGGTATTGATCAATGGTAATTGTTGCTGAGCAAAGAAAACG  
 TTTGACGTGATCAATCCATCTAACCGAAGAAAAGATAACAACTGTATACAGGCTATGGAA  
 GATGATGTTGATGAGGCTGGCGGCTTAAAGCTTTGAAACAGAACGAGTGGCTTATT  
 GTAGAGGCGGGAGGTTCCGGCTAAAGCTTTATTCAATCTCGCTGACTTGGTGGAGAAACAC  
 CAAGAAAACACTGGCTGCCATTGAGTCATGGATAATGGTAAGTCATTGTTTGTGCGCG  
 GGTGACGCTGCTTGTAGTATCTAAATACTTGGGTTCTGCGGTGGTGGCAGATAAAATC  
 TACGGTAACGTTGATGACACAGGTAAGGAAACCCATTACACTAACATTAGGAACCATTA  
 GGGGTTTGGCCCAAATAATCCTGGAACTTCCCTTATTGATGTTGCTGATGGAAATT  
 GGGCCTGCTCTGGTACAGGTAACACCGCTGATTGAAACCCGCTGAAACACACCTTA  
 TCTGCCCTTTTGGCTTCCAGTTGTCAGGAAGGAGGCTACCCGCTGTTGAGTCAT  
 ATCTTGGGGGGTCCGGTAGAGTTGTTGGAGAAAGATTGAGTGACACACCCAGACGTC  
 AGAGATTGTTTACAGGCTACTGGCACCGGCCCATATTAGAGGTGCGTGCCT  
 ACTGTCAGAACAGTACTTGGAGCTGGGAGGTAATACACCAATTATTGTTGCTGAC  
 GCTGATCTAGATAAGCCGTCAAGAACATTGCTTGGTATTTCACAACTCTGGTAA  
 GTTGGTCTGCGCTGGTCCAGAAATATCATTCAAGATACAGTATACGAGGGGTGTTGCAA  
 AAACATAAGGATTACACCGAGTCACAAAGGTCGGTGACCCATTGATGAGGAAATTTC  
 CAAGGTGCTCAAACATCTGACAAACAGCTGCATAAAATTAGACTATGTCGATGAG  
 AAATCAGAGGGGGCTGCTTGTGACTGGAGGGGGCAGACATGGCAGTAAGGTTATT  
 GTCAAGCCAACAGTGTGCTGATGTCAGGAAAGATATGAGAAATTGTTAAGGAGGAAGT  
 TTTGGTCCATTGTAACTGATCAAAGTTTCTACTGTTGATGAGTGGTATGGCA

**도면11d**

AATGATTCTCAATATGGTTAGCCGCAGGTATTCACACTAACGATATAACAAGGCTGTT  
GATGTGTCAAAAGAGTGAACAGCTGGTACTGTTGGATAAATACCTATAACAACCTCCAC  
CAAATGTTCTTCGGTGGCTCGGCAGTCAGGTATTGCCGTGAAATGGGTGAGGCT  
GCTTAAGTAACACTCAAACAAATCTGTCAGAATTGCCATTGACAAGCCAATTGCGT  
TGA

## &gt;ALD6 누클레오티드 (SEQ ID NO: 71)

ATGACTAACGCTACACTTGACACTGCTGAAACCAGTCAGATCACACTTCAAATGGTTG  
ACATACAGAGCACAACCGGTCTATTCAACAAACAGTAAAGCTCAAGACGGT  
AAGACCTATCCCGTCAAGATCCTCCACTGAAAACACCGTTGAGGTCTCTCGCC  
ACCACTGAAGATGTGAATATGCTATCGAATGTGCCGACCGTCTTCCACGACACTGAA  
TGGCTACCCAAGACCCAAGAGAAAAGAGGCGTCACTAACGTAAGTGGCTGACGAAATTG  
GAAAGCCAATTGACTTCCATTGACATTGGACAATGGTAAAAGCTTGGCC  
TTAGCCGTGGGATGTTACCATTGAACACTGCTAACAGATGCTGCTGCTATGCC  
GACAAAGTCACCGTAGAACAACTAACACCGGTGACGGCTACATGAACTTACCCCTTA  
GAGGCAATCGGTCTGTGGTCAAATTTCATGGAAACTTCCAAATAATGATGTTGGCT  
TGGAGATGCGTGGCCATGGGTAACGCTGTATCTGAAACACCGCTGCTGTC  
ACACCTTAAATGCCCTATACTTGTCTCTTATGAAAGAGGTTGTTACCGTGGT  
GTCGTCACATCGTCCAGGTCTGGTAGAACACTGTTGGTCTGCTTGGACAACGACCCA  
AGAGCTAACAGCTGGTTTACCGGTCTACAGAACGTAAGAGTGTGCTGTCGAC  
TCTCTGAATCTAACATTGAGAAAATCATTGGAAACTAGGTGTAAGTCCGCCATTG  
GTCCTTGACGATGCTAACATTAAGAAGACTTACCAATCTAGTAAACGGTATTTCAG  
AACGCTGGTCAAATTGTTCTCTGGTCTAGAATTACGTTCAAGAAGGTTACGAC  
GAACATTGGCTGCTTCAAGGCTACTTGGAAACCGAAATCAAAGTTGGTAACTCCATT  
GACAGGCTAACACGGTCTATCACTAACCGTCAACAAATTGGACACAATTATGAAAC  
TACATCGATATCGTAAGAAAAGGGCCAAAGATCTAACGTTGGCAGAAAAGTTGGT  
GACAAGGGTACTTCATCAGACCAACCGTTTCTACGATGTTAAGAGACATGAGAATT  
GTTAGGAAAGAAAATTGGACGAGTGTCAAGTGGCAAGGTTAGGTTAGAAGGAA  
GGTCTCGAAATGCTAACAGCTGTAATTGCTGAGGTTCTGGTATCGAAACAGAAATCT  
TTGAGCACAGGTTGAAGGTGGCCAAGATGTTGAAGGCGTACCGTCTGGATCAACACA  
TACAACGATTGACTCCAGAGTCCATTGCGTGGTTAAGCAATCTGGTACCGTAGA  
GAAATGGTGAAGAAGTCTACCATGCTACACTGAAGTAAAAGCTGTCAGAATTAGTTG  
TAA

## &gt;HFD1 누클레오티드 (SEQ ID NO: 72)

ATGCTAAACGACGGCTAAAAATATTGAAATTACCCCAGTGTCTAAATAGATGAAATA  
GTTGAAATCTCAAGAAATTCTCTTGAGAACAAATTGAAATTGTCACGAAAATAAC

**도면11e**

CCAAGGAAAAAGATCTAGAATTCAAGGCAGTTGCAAGTTGAAAAAAACTCTATTATGCCGT  
AAAGATCATGAGGAGAACTGATCGATGCTATGACAAGGACTTTCATCGAACAAATT  
GAATCGGTTCTGAATGAAACGACCAAACTTATGAACGATATACTTCACCTAATTGAGATT  
TTACCAAATATTGATCAAACCTCGAGAGTATCTGATTCTCTCCCAATTATGTTGGT  
AAACAAATCGTGGAGAAAATATCAAGGGCAGTGTCTGATTATGCTCCTTCAATT  
CCCCACTTTGACATTGCCCCATTGGCAGCAGCTTGTGAGGTAAACACCAATTGTT  
CTGAAGCCAAGTGAACAAACACACACTGCTGTAGTTGGAAAATTGTTAACAC  
GCTGGTTCTCTGATGGATTGATGCAAGTGTCAAGGAGCTATAGTAAAGACTACAAGA  
CTACTAGATTGGAAAATTGACTAATATTCTACACAGGTTCTCCCGTGTGGATCA  
ATAGTTGCTGAGAAAAGCAGCAAAGACTCTAACACCTTGTGTACTTGAACCTGGTAAA  
TCACCTACCTTATTACAGAAAATTCAAAGCAAGTAACATAAAATTGCTTGGAAAAGG  
ATTTTTTGGTCTTGGAAATTCTGGCCAGATTGTTGTTACCAAGATTATTGTTA  
GTACATAATCTCTATCTACAAAGCTTAAAGAGTGTGAATCAGTACTAAATGAATT  
TATCCAAGCTTGTGACACAAACAGATTCACTCGTATGATTGAGCTGCTTACAAA  
AAGGCCGTTGCAAGTAAACTCAACTAACGGCTCAAGATTGTCCTCAAAATTCT  
ATCAATTGAGATGAGGATCTATGCTTGTACACCAACCATAGTTATAACATTGGT  
TGGGATGATCTTGTGAAACAGGAAAACCTTGTCTGTATTGCCCATATTGAGTAC  
GAGGATCTTGTGAGGACCTTAAACAGATAATAGAAGAACATGACACTCCATTGTC  
TACATATTCTCTGATAGCCAACACTGAATAATCGTATCTGACCGCTTAAGATCTGGT  
GACTGTGTTGCGTGTACAGTGTAGGAAATTACCGAGGCTCCATTGGAGGG  
ATCGGTAATTCTGAGGTTAGGAACTATGGTGGATATTGTTGATTCAATACCTTGT  
GAAAGAACAAATTGAAACACCATATTGGAAATGATTACCCCTTTATGAGATACCC  
CCAATAGCGCACAAGGAAAAGCTCGTCCGTTTGCAGTGGAAAGAAAACCTGGTT  
GACGAAATGGCAATAACAGTGGGGGTTACGCCAATATTTCATTATGCCCCCGTT  
ATTTCATTAGTACCATTCACGCTCATGTTCTCTG

**서 열 목록**

## SEQUENCE LISTING

&lt;110&gt; Raghavan, Shriram

Hansen, Jorgen

Sonkar, Shailendra

Kumar, Sathish

Panchapagesa, Murali

Halkjaer Hansen, Esben

Riishede Hansen, Klavs

<120> METHODS AND MATERIALS FOR RECOMBINANT PRODUCTION OF SAFFRON

COMPOUNDS

<130> MBHB 13-1499-WO

<150> US 61/521,171

<151> 2011-08-08

<150> US 61/576,460

<151> 2011-12-16

<150> US 61/595,450

<151> 2012-02-06

<150> PCT/IB2012/001513

<151> 2012-08-07

<160> 72

<170> PatentIn version 3.5

<210> 1

<211> 1386

<212> DNA

<213> Stevia rebaudiana

<400> 1

atggaggct ctaaggtgat cctgtatcct tctccggaa tcggccatct tgttccatg	60
gtggagctt gaaaactcat ccacacccac caccctcac tctccgttat catcctcgta	120
cttccggcta catatgaaac cgggtccacc actacataca tcaacaccgt ctccaccacc	180
accccccattca tcacattcca ccacccccc gttatccctc ttccaccaga ctcatcttct	240
gaattcatag accttgcctt cgatatccct caacttaca acccggtcgt ctacaacacc	300

ctcgtagcca tctccgaaac ctcaccatc aaagctgtca tccttgattt ctgtaaac	360
gcagctttc agatctctaa aagtctcgat cttccactt actacttctt taccagtggc	420
gcctctggtc tctgtgcgtt cttacatctt ccaaccatct aaaaaacata ttccggaaac	480
ttaaagatc tagatacttt cattaatatt cctgggtac ctccattca ttctccgt	540
atgccccacag ttttgttga taaggaaagt aattctaca aaaacttcgt aaaaacctca	600
aataacatgg caaaatctt cggagtcatt gcaaacagct tcttgagtt ggaggaaga	660

gctgctcaaa ctctccgaga tggtaaatcc atcacggacg gtcgcctctcc acctattat 720

ctaatacgccc cttaatcgc tagcggcaat caagttgatc ataacgaaaa cgagtgttta 780  
 aaatggctga acacacaacc tagtaaaaagt gtatgtttt tggctttgg gagccagggt 840  
 gtgttaaga aagaacaatt gaaggaaata gcgggtgggt tagagagaag tggcaaaga 900  
 ttttgtgg tggtgcaaa gccgccatca gatggtggt aagagttcgg tcttgcgtat 960  
 gttcttcctg aagggtttgt agccaggact aaagaaaagg gtctggtggt gaagaactgg 1020  
 ggcctcaac cagcgattct tggcatgaa tcgggtggag gatttgcgtat tcattgcggg 1080  
 tggaactcgt cacttgaagc ggttgggg ggtgtgccga tggtgccatg gcccgtgtac 1140

gcagagcaga agatgaacag agtgtatgg gttgaggaaa taaaggtggc actttgggg 1200  
 agaatgtcg cagatgggtt tgtgagtgcg gaggcggt tagagacggt gagacagtta 1260  
 atggatggga gaagagttag agaacggatt ttggagatga gtacaaaagc caaggctcg 1320  
 gtggaggacg gcggttcctc tcgagttgtat ttcttcaaat taactgagtc atggaccac 1380  
 aagtga 1386

<210> 2

<211> 461

<212> PRT

<213> Stevia rebaudiana

<400> 2

Met Glu Ser Ser Lys Val Ile Leu Tyr Pro Ser Pro Gly Ile Gly His

1 5 10 15

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

20 25 30

Ser Leu Ser Val Ile Ile Leu Val Leu Pro Ala Thr Tyr Glu Thr Gly

35 40 45

Ser Thr Thr Thr Tyr Ile Asn Thr Val Ser Thr Thr Pro Phe Ile

50 55 60

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

65 70 75 80

Glu Phe Ile Asp Leu Ala Phe Asp Ile Pro Gln Leu Tyr Asn Pro Val

85 90 95

Val Tyr Asn Thr Leu Val Ala Ile Ser Glu Thr Ser Thr Ile Lys Ala

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

Gly Gly Phe Val Ser His Cys Gly Trp Asn Ser Ser Leu Glu Ala Val

355 360 365

Val Phe Gly Val Pro Met Val Ala Trp Pro Leu Tyr Ala Glu Gln Lys

370 375 380

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

385 390 395 400

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

405 410 415

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

420 425 430

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

435 440 445

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

450 455 460

<210> 3

<211> 1380

<212> DNA

<213> Stevia rebaudiana

<400> 3

atggaaaata aaacggagac caccgttcgc cgccgcggga gaataatatt attcccgta	60
ccatttcaag gtcacataaa cccaatgctt cagcttagcca atgtgttgta ctccaaagga	120
ttcagtatca ccatcttca caccaacttc aacaaaccca aaacatctaa ttaccctcac	180
ttcactttca gattcatcct cgacaacgac ccacaagacg tacgcatttc caatctaccg	240
actcatggtc cgctcgctgt tatgcggatt ctgattatca acgaacacgg agctgaccaa	300

ttacgacgca aactggaact gttgatgtta gcttctgaag aagatggaga ggtatcgtgt	360
ttaatcgccg atcagatttg gtacttcacg caatctgtt ctgacagtct taacctccga	420
cggcttgcgtt tggtgacaag cagcttgcgtt aatttcatg cacatgttc acttcctcag	480
tttgcgttgcg ttggttaccc cgatcctgat gacaaaaccc gtttggaga acaagcgtgt	540
gggtttccata tgctgaaagt gaaagatatc aagtgttagtt tttcgatgtg gaaaaaaatac	600
aaagagtatt tcgagaacat tacgaaacaa acaaaagcat cttcaggagt catctggAAC	660
tcatttaagg aactcgaaga gtctgagctc gaaactgtta tccgtgagat cccgctcca	720

agtttcttga taccactccc caagcattg acagcatctt ccagcagctt actagaccac	780
gatcgAACG ttttccatg gtagaccaa caaccgtcac gttcggtact gtatgttagt	840
tttggtagtg gtactgaagt actggatgag aaagatttct tgaaatagc tcgtggttg	900
gttgatagca agcagtgcgtt ttatgggtg gttcgacctg gtttgtcaa gggttcgacg	960
tgggtcgaac cttgccaga tgggttctt ggtgaaagag gacgtattgt gaaatgggtt	1020
ccgcagcaag aagtgcgtc tcatggagca ataggcgcat tctggactca tagcggatgg	1080
aactctacgt tgaaagcgt ttgtgaaggt gttcctatga tttctcgga tttgggctc	1140

gatcaaccgt tgaatgctag atacatgagt gatgtttga agtaggggt gtattggaa	1200
aatgggtgg aaagaggaga gatagcaa gcaataagaa gagttatggt ggatgaagaa	1260
ggagaataca ttagacagaa tgcaagagt ttgaaacaaa aggcatgtt ttcttgatg	1320
aagggtggtt catttacga atcatttagag tctctagtt cttacatttc atcgtttaaa	1380

&lt;210&gt; 4

&lt;211&gt; 459

&lt;212&gt; PRT

&lt;213&gt; Stevia rebaudiana

&lt;400&gt; 4

Met Glu Asn Lys Thr Glu Thr Thr Val Arg Arg Arg Arg Arg Ile Ile

1 5 10 15

Leu Phe Pro Val Pro Phe Gln Gly His Ile Asn Pro Met Leu Gln Leu

20 25 30

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

35 40 45

Asn Phe Asn Lys Pro Lys Thr Ser Asn Tyr Pro His Phe Thr Phe Arg

50 55 60

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

65 70 75 80

Thr His Gly Pro Leu Ala Val Met Arg Ile Leu Ile Asn Glu His

85 90 95

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

100 105 110

Glu Glu Asp Gly Glu Val Ser Cys Leu Ile Ala Asp Gln Ile Trp Tyr

115 120 125

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

130 135 140

Val Thr Ser Ser Leu Phe Asn Phe His Ala His Val Ser Leu Pro Gln

145 150 155 160

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

165 170 175

Glu Gln Ala Ser Gly Phe Pro Met Leu Lys Val Lys Asp Ile Lys Cys

180 185 190

Ser Phe Ser Met Trp Lys Lys Tyr Lys Glu Tyr Phe Glu Asn Ile Thr

195 200 205

Lys Gln Thr Lys Ala Ser Ser Gly Val Ile Trp Asn Ser Phe Lys Glu

210 215 220

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

225 230 235 240

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

245 250 255

Leu Leu Asp His Asp Arg Thr Val Phe Pro Trp Leu Asp Gln Gln Pro

260 265 270

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

275 280 285

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

290 295 300

Gln Ser Phe Leu Trp Val Val Arg Pro Gly Phe Val Lys Gly Ser Thr

305 310 315 320

Trp Val Glu Pro Leu Pro Asp Gly Phe Leu Gly Glu Arg Gly Arg Ile

325 330 335

Val Lys Trp Val Pro Gln Gln Glu Val Leu Ala His Gly Ala Ile Gly

340 345 350

Ala Phe Trp Thr His Ser Gly Trp Asn Ser Thr Leu Glu Ser Val Cys

355 360 365

Glu Gly Val Pro Met Ile Phe Ser Asp Phe Gly Leu Asp Gln Pro Leu

370	375	380
Asn Ala Arg Tyr Met Ser Asp Val Leu Lys Val Gly Val Tyr Leu Glu		
385	390	395
Asn Gly Trp Glu Arg Gly Glu Ile Ala Asn Ala Ile Arg Arg Val Met		
405	410	415
Val Asp Glu Glu Gly Glu Tyr Ile Arg Gln Asn Ala Arg Val Leu Lys		
420	425	430
Gln Lys Ala Asp Val Ser Leu Met Lys Gly Gly Ser Ser Tyr Glu Ser		
435	440	445
Leu Glu Ser Leu Val Ser Tyr Ile Ser Ser Leu		
450	455	
<210> 5		
<211> 1383		

<212> DNA

<213> Stevia rebaudiana

<400> 5

atggcggAAC aacaaaAGAT caagaaATCA ccacacGTTC tactcatccc attccCTTA	60
caaggCCATA taaaccCTTT catccAGTTT ggcaaACGAT taatctccAA aggtgtCAA	120
acaacACTTG ttaccACCAT ccacacCTTA aactcaACCC taaaccACAG taacaccACC	180
accacCTCCA tgeaaATCCA agcaattCCG gatggTTGtG atgaaggCgg ttttatGAGt	240
gcaggagaAT catatTTGGA aacattCAAA caagttGGGT ctaaatCAct agctgactTA	300
atcaagaAGC ttcaaAGTGA aggaaccACA attgatGCA A tcatttatGA ttctatGACT	360

gaatgggtt tagatgttgc aattgagttt ggaatcgatG gtggTTcgTT tttcactCAA	420
gcttggttg taaacagCTT atattatcat gttcataagg gtttgatttc tttGCCATTG	480
ggtaaaACTG tticggTTCC tggatttcca gtgctcaac ggtgggagac accgttaatt	540
ttgcagaATC atgagcaaAT acagagCCt tggtctcaga tggTTttgg tcagtTTGCT	600
aatattGATC aagcacgttG ggtcttcaca aatagtTTTt acaagctcGA ggaagaggtA	660
atagagtGGA cgagaaAGAT atggaacttG aaggtaatcG ggccaacact tccatccatG	720
taccttgaca aacgacttGA tgatgataAA gataacggat ttaatctcta caaagcaaAC	780

catcatgAGT gcatgaACTG gtttagacGAT aagccaaAGG aatcagtTGT ttacgtAGCA	840
tttggtagCC tggtaaACA tggacCCGAA caagtGGAAG aaatcacACG ggcttAATA	900

gatagtgtatc tcaacttctt gtgggttatac aaacataaag aagagggaaa gctcccgaa	960
aatcttcgg aagtaataaa aaccggaaag ggtttgattt tagcatggtg caaacatgg	1020
gatgtgttag cacacgaatc agtaggatgc ttgttacac attgtgggtt caactcaact	1080
cttgaagcaa taagtcttgg agtccccgtt gttcaatgc cttcaatttc ggatcaaact	1140
acaatgcca agcttctaga taaaatttg ggtgtggag ttagagttaa ggctgatgag	1200

aatggatag tgagaagagg aaatcttgcg tcatgtatta agatgattt ggaggaggaa	1260
agaggagtaa taatccgaaa gaatgcggta aaatggaagg atttgctaa agtagccgtt	1320
catgaaggtg gtatgtcaga caatgatatt gtcgaatttg taagtggct aattaaggct	1380
taa	1383

&lt;210&gt; 6

&lt;211&gt; 460

&lt;212&gt; PRT

&lt;213&gt; Stevia rebaudiana

&lt;400&gt; 6

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

1 5 10 15

Pro Phe Pro Leu Gln Gly His Ile Asn Pro Phe Ile Gln Phe Gly Lys

20 25 30

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

35 40 45

Thr Leu Asn Ser Thr Leu Asn His Ser Asn Thr Thr Thr Ser Ile

50 55 60

Glu Ile Gln Ala Ile Ser Asp Gly Cys Asp Glu Gly Gly Phe Met Ser

65 70 75 80

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

85 90 95

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

100 105 110

Ala Ile Ile Tyr Asp Ser Met Thr Glu Trp Val Leu Asp Val Ala Ile

115 120 125

Glu Phe Gly Ile Asp Gly Ser Phe Phe Thr Gln Ala Cys Val Val

130	135	140
Asn Ser Leu Tyr Tyr His Val His Lys Gly Leu Ile Ser Leu Pro Leu		
145	150	155
Gly Glu Thr Val Ser Val Pro Gly Phe Pro Val Leu Gln Arg Trp Glu		
165	170	175
Thr Pro Leu Ile Leu Gln Asn His Glu Gln Ile Gln Ser Pro Trp Ser		
180	185	190
Gln Met Leu Phe Gly Gln Phe Ala Asn Ile Asp Gln Ala Arg Trp Val		
195	200	205
Phe Thr Asn Ser Phe Tyr Lys Leu Glu Glu Glu Val Ile Glu Trp Thr		
210	215	220
Arg Lys Ile Trp Asn Leu Lys Val Ile Gly Pro Thr Leu Pro Ser Met		
225	230	235
Tyr Leu Asp Lys Arg Leu Asp Asp Asp Lys Asp Asn Gly Phe Asn Leu		
245	250	255
Tyr Lys Ala Asn His His Glu Cys Met Asn Trp Leu Asp Asp Lys Pro		
260	265	270
Lys Glu Ser Val Val Tyr Val Ala Phe Gly Ser Leu Val Lys His Gly		
275	280	285
Pro Glu Gln Val Glu Glu Ile Thr Arg Ala Leu Ile Asp Ser Asp Val		
290	295	300
Asn Phe Leu Trp Val Ile Lys His Lys Glu Glu Gly Lys Leu Pro Glu		
305	310	315
Asn Leu Ser Glu Val Ile Lys Thr Gly Lys Gly Leu Ile Val Ala Trp		
325	330	335
Cys Lys Gln Leu Asp Val Leu Ala His Glu Ser Val Gly Cys Phe Val		
340	345	350
Thr His Cys Gly Phe Asn Ser Thr Leu Glu Ala Ile Ser Leu Gly Val		
355	360	365
Pro Val Val Ala Met Pro Gln Phe Ser Asp Gln Thr Thr Asn Ala Lys		
370	375	380

Leu Leu Asp Glu Ile Leu Gly Val Gly Val Arg Val Lys Ala Asp Glu  
 385                    390                    395                    400

Asn Gly Ile Val Arg Arg Gly Asn Leu Ala Ser Cys Ile Lys Met Ile  
 405                    410                    415

Met Glu Glu Glu Arg Gly Val Ile Ile Arg Lys Asn Ala Val Lys Trp  
 420                    425                    430

Lys Asp Leu Ala Lys Val Ala Val His Glu Gly Gly Ser Ser Asp Asn  
 435                    440                    445

Asp Ile Val Glu Phe Val Ser Glu Leu Ile Lys Ala  
 450                    455                    460

&lt;210&gt; 7

&lt;211

&gt; 1422

&lt;212&gt; DNA

&lt;213&gt; Stevia rebaudiana

&lt;400&gt; 7

atggctacca gtgactccat agttgacgac cgtaaggcgc ttcatgttgc gacgttccca	60
tggcttgctt tcggtcacat cctcccttac cttcagcttt cgaaattgtat agctgaaaag	120
ggtcacaaag tctcgtttct ttctaccacc agaaacatcc aacgtctctc ttctcatatc	180
tgcgccactca taaaatgttgt tcaactcaca cttccacgtg tccaagagct gccggaggat	240
gcagaggcga ccactgacgt ccacctgaa gatattccat atctcaagaa ggcttctgat	300
ggtcttcaac cggaggtcac ccggtttcta gaacaacact ctccggactg gattattat	360

gattatactc actactgggtt gccatccatc gcggctagcc tcggtatctc acgagcccac	420
ttctccgtca ccactccatg ggccattgtat tatatggac cttcagctga cgccatgata	480
aatggttcag atggtcgaac cacgggttag gatctcacga caccgc当地 gtgggttccc	540
tttccgacca aagtatgctg gcggaaagcat gatcttgccc gactggtgcc ttacaaagct	600
ccggggatat ctgatggata ccgtatgggg ctgggttcta agggatctga ttgtttgctt	660
tccaaatgtt accatgagtt tggaactcaa tggctacctc tttggagac actacaccaa	720
gtaccgggtgg ttccgggtggg attactgcca cggaaatac cggagacga gaaagatgaa	780

acatgggtgt caatcaagaa atggctcgat ggttaaacaaa aaggcagtgt ggtgtacgtt	840
gcattagaa gcgagggtttt ggtgagccaa accgagggttg ttgagtttagc attgggtctc	900
gagcttctg gttgccatt tggtggct tatagaaaac caaaagggtcc cggaaagtca	960

gactcggtgg agttgccaga cgggttcgtg gaacgaactc gtgaccgtgg gttggctgg	1020
acgagttggg caccttagtt acgaatactg agccatgagt cggtttgtgg tttcttgact	1080
cattgtggtt ctggatcaat tgtgaaaggg ctaatgttg gtcaccctct aatcatgcta	1140
ccgatTTTG gggaccaacc tctgaatgtc cgattactgg aggacaaaca ggtggaaatc	1200

gagataccaa gaaatgagga agatggttgc ttgaccaagg agtcggttgc tagatcactg	1260
aggtccgtt tggtggaaaa agaaggggag atctacaagg cgaacgcgag ggagctgagt	1320
aaaatctata acgacactaa ggttgaaaaa gaatatgtaa gccattcgt agactatTTG	1380
gaaaagaatg cgctgcggt tgccatcgat catgagagtt aa	1422

&lt;210&gt; 8

&lt;211&gt; 473

&lt;212&gt; PRT

&lt;213&gt; Stevia rebaudiana

&lt;400&gt; 8

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

1 5 10 15

Ala Thr Phe Pro Trp Leu Ala Phe Gly His Ile Leu Pro Tyr Leu Gln

20 25 30

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

35 40 45

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

50 55 60

Asn Val Val Gln Leu Thr Leu Pro Arg Val Gln Glu Leu Pro Glu Asp

65 70 75 80

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

85 90 95

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

100 105 110

His Ser Pro Asp Trp Ile Ile Tyr Asp Tyr Thr His Tyr Trp Leu Pro

115 120 125

Ser Ile Ala Ala Ser Leu Gly Ile Ser Arg Ala His Phe Ser Val Thr

130 135 140

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

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

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

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

385	390	395	400
-----	-----	-----	-----

Glu Ile Pro Arg Asn Glu Glu Asp Gly Cys Leu Thr Lys Glu Ser Val

405	410	415
-----	-----	-----

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

420	425	430
-----	-----	-----

Lys Ala Asn Ala Arg Glu Leu Ser Lys Ile Tyr Asn Asp Thr Lys Val

435	440	445
-----	-----	-----

Glu Lys Glu Tyr Val Ser Gln Phe Val Asp Tyr Leu Glu Lys Asn Ala

450	455	460
-----	-----	-----

Arg Ala Val Ala Ile Asp His Glu Ser

465	470
-----	-----

<210> 9

<211> 1446

<212> DNA

<213> Stevia rebaudiana

<400> 9

atggatgcaa tggctacaac tgagaagaaa ccacacgtca ttttcatacc atttccagca	60
caaaggccaca tttaaggccat gctcaaacta gcacaacttc tccaccacaa aggactccag	120
ataaaccttcg tcaacaccga cttcatccac aaccagttc ttgaatcatc gggcccat	180
tgtcttagacg gtgcaccggg ttccgggttc gaaaccatc eggatggtgt ttctcacagt	240
ccggaagcga gcatccaat cagagaatca ctcttgagat ccattgaaac caacttcttg	300

gatcgttca ttgatcttgt aaccaaactt ccggatcctc cgacttgtat tatctcagat	360
gggttcttgt cggtttcac aattgacgct gaaaaaaagc ttgaaattcc ggtcatgtat	420
tattggac ac ttgctgcctg tggttcatg ggttttacc atattcattc tctcattgag	480
aaaggattt caccactaa agatgcaat tacttgacaa atgggtatgg ggacaccgtc	540
atgattggg ttccggaaat ggaaggcatc cgtctcaagg attcccgct ggactggagc	600
actgaccta atgacaaatg ttgatgttc actacggaag ctctcaaag gtcacacaag	660
gtttcacatc atatccca cacgttcgat gagttggagc ctgttattaa aaaaactttg	720

tcattgaggt ataatcacat ttacaccatc ggcccactgc aattacttct tcatcaaata	780
cccgaaagaga aaaagcaaac tggaaattacg agtctccatg gatacagttt agtaaaagaa	840
gaaccagagt gttccagtg gttcagtct aaagaaccaa attccgtcgt ttatgtaaat	900

tttggaga	ctacagtaat	gtcttagaa	gacatgacgg	aattgggtt	gggacttgct	960
aatagcaacc	attatttcct	ttggatcatc	cgtcaact	tggtgatagg	ggaaaatgca	1020
gtttggccc	ctgaacttga	ggaacatata	aagaaaagag	gcttattgc	tagctgggt	1080
tcacaagaaa	aggcttigaa	gcaccctcg	gttggaggt	tcttgactca	ttgtgggtgg	1140

ggatcgacca	tcgagagctt	gtctgctggg	gtgccaatga	tatgctggcc	ttattcgtgg	1200
gaccagctga	ccaaactgttag	gtatatatgc	aaagaatggg	aggttgggct	cgagatggga	1260
accaaagtga	aacgagatga	agtcaagagg	cttgtacaag	agttgatggg	agaaggaggt	1320
cacaaaatga	ggaacaaggc	taaagattgg	aaagaaaagg	ctcgattgc	aatagctcct	1380
aacggttcat	ctctttgaa	catagacaaa	atggtcaagg	aaatcaccgt	gctagcaaga	1440
aactag						1446

&lt;210&gt; 10

&lt;211&gt; 481

&lt;212&gt; PRT

&lt;213&gt; Stevia rebaudiana

&lt;400&gt; 10

Met Asp Ala Met Ala Thr Thr Glu Lys Lys Pro His Val Ile Phe Ile

1 5 10 15

Pro Phe Pro Ala Gln Ser His Ile Lys Ala Met Leu Lys Leu Ala Gln

20 25 30

Leu Leu His His Lys Gly Leu Gln Ile Thr Phe Val Asn Thr Asp Phe

35 40 45

Ile His Asn Gln Phe Leu Glu Ser Ser Gly Pro His Cys Leu Asp Gly

50 55 60

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

65 70 75 80

Pro Glu Ala Ser Ile Pro Ile Arg Glu Ser Leu Leu Arg Ser Ile Glu

85 90 95

Thr Asn Phe Leu Asp Arg Phe Ile Asp Leu Val Thr Lys Leu Pro Asp

100 105 110

Pro Pro Thr Cys Ile Ile Ser Asp Gly Phe Leu Ser Val Phe Thr Ile

115 120 125

Asp Ala Ala Lys Lys Leu Gly Ile Pro Val Met Met Tyr Trp Thr Leu  
 130 135 140  
 Ala Ala Cys Gly Phe Met Gly Phe Tyr His Ile His Ser Leu Ile Glu  
 145 150 155 160  
 Lys Gly Phe Ala Pro Leu Lys Asp Ala Ser Tyr Leu Thr Asn Gly Tyr  
 165 170 175  
 Leu Asp Thr Val Ile Asp Trp Val Pro Gly Met Glu Gly Ile Arg Leu  
 180 185 190  
  
 Lys Asp Phe Pro Leu Asp Trp Ser Thr Asp Leu Asn Asp Lys Val Leu  
 195 200 205  
 Met Phe Thr Thr Glu Ala Pro Gln Arg Ser His Lys Val Ser His His  
 210 215 220  
 Ile Phe His Thr Phe Asp Glu Leu Glu Pro Ser Ile Ile Lys Thr Leu  
 225 230 235 240  
 Ser Leu Arg Tyr Asn His Ile Tyr Thr Ile Gly Pro Leu Gln Leu Leu  
 245 250 255  
  
 Leu Asp Gln Ile Pro Glu Glu Lys Lys Gln Thr Gly Ile Thr Ser Leu  
 260 265 270  
 His Gly Tyr Ser Leu Val Lys Glu Glu Pro Glu Cys Phe Gln Trp Leu  
 275 280 285  
 Gln Ser Lys Glu Pro Asn Ser Val Val Tyr Val Asn Phe Gly Ser Thr  
 290 295 300  
 Thr Val Met Ser Leu Glu Asp Met Thr Glu Phe Gly Trp Gly Leu Ala  
 305 310 315 320  
  
 Asn Ser Asn His Tyr Phe Leu Trp Ile Ile Arg Ser Asn Leu Val Ile  
 325 330 335  
 Gly Glu Asn Ala Val Leu Pro Pro Glu Leu Glu Glu His Ile Lys Lys  
 340 345 350  
 Arg Gly Phe Ile Ala Ser Trp Cys Ser Gln Glu Lys Val Leu Lys His  
 355 360 365  
 Pro Ser Val Gly Gly Phe Leu Thr His Cys Gly Trp Gly Ser Thr Ile

370	375	380	
Glu Ser Leu Ser Ala Gly Val Pro Met Ile Cys Trp Pro Tyr Ser Trp			
385	390	395	400
Asp Gln Leu Thr Asn Cys Arg Tyr Ile Cys Lys Glu Trp Glu Val Gly			
405	410	415	
Leu Glu Met Gly Thr Lys Val Lys Arg Asp Glu Val Lys Arg Leu Val			
420	425	430	
Gln Glu Leu Met Gly Glu Gly His Lys Met Arg Asn Lys Ala Lys			
435	440	445	
Asp Trp Lys Glu Lys Ala Arg Ile Ala Ile Ala Pro Asn Gly Ser Ser			
450	455	460	
Ser Leu Asn Ile Asp Lys Met Val Lys Glu Ile Thr Val Leu Ala Arg			
465	470	475	480
Asn			
<210> 11			
<211> 1470			
<212> DNA			
<213> Stevia rebaudiana			
<400> 11			
atggctagag tcgatagagc cacaacatt cactcgct tgttccgct actgactcca			60
ggtcataatga tacccatggc cgacatagcc cggttactag ccgaacgcgg ttcaacggta			120
accataatca ccacaccact gaacgcgaac cgttcaaac cggtcattgc tcggccatc			180
aaagacccgc tcaagatcca agttcttcaa ctcaaactcc cctcaaccga aggtttaccc			240
gaaggatgctc agaattttga catgatcgaa tcggctcagt ttttcataa aatgttcgag			300
gcaacatata agttagccga acccgccgtt aacgcggtcc agagactaac tccaccacca			360
agttgcata ttgctgataa tctttacct tggacaaatg atttagccca aaagttaaa			420
attccaagaa ttgttttca tggcccgga tgctcacaa tcttatgcat acatattgca			480
atgaatagta acgtgttata tgacatcggt tccgattcggt agcgatctt gctaccgggt			540
ttaccggacc gtattgagct aaccaaagga caagcttga gttggggag gaaagacaca			600
aaggaagccg cgagtttttg gaaccgcgtg caacgagacg aagatttcgc aaatggatc			660

gtggtaata gtttcacgc gttggacct tactatgtt aagagcttc aaaggtaaa	720
gttaagaag ttgggttat tggccggtt tcgttatgtt acaaaagttt cgaagatata	780
gccgagagag gaacaagg agcgattgtt gaacatgaat gtttgaatgt gtttagattcg	840
atggagtcac ggtcagtgtt attcgttgtt ttggggagtc tggttcgtgt tgggaccgag	900
caaaaattt acctcggtt agggttggag gcatcgaaga aaccgtttt gtggtccta	960
cgcacatacaa ccgaagaatt cgaaagatgg ttgtcgagc aagggtatga agaaagggtg	1020

aaagatagag ggctaataat ccgtgggtgg gccccacaag tttttatccc gtcgcaccga	1080
gccattggtg gtttttaac acattgtggg tggaactcga ctcttgaagg gattacagct	1140
ggagtccta tggttacatg gcctcagttt acggaccagt ttataaacga aagattttt	1200
gtagatgttt tgaagatcggtt ggtgaacaaa gaagagatca cgcgatcgat cgaagatcta	1260
atggacgaag gtgaggaagg tgaacaaga agaaggagaa gtagagaact acgcgatatg	1320
gcaaaaagcg cgatggagga tggaggttca tcgcacatcgat atatgacatc aatgattcag	1380
gatattgtcg agttgtcaa aaatcgtaa	1440

gatattgtcg agttgtcaa aaatcgtaa	1470
--------------------------------	------

<210> 12

<211> 489

<212> PRT

<213> Stevia rebaudiana

<400> 12

Met Ala Arg Val Asp Arg Ala Thr Asn Leu His Phe Val Leu Phe Pro

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

Leu Leu Thr Pro Gly His Met Ile Pro Met Val Asp Ile Ala Arg Leu

20	25	30
----	----	----

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

35	40	45
----	----	----

Ala Asn Arg Phe Lys Pro Val Ile Ala Arg Ala Ile Lys Asp Arg Leu

50	55	60
----	----	----

Lys Ile Gln Val Leu Glu Leu Lys Leu Pro Ser Thr Glu Gly Leu Pro

65	70	75	80
----	----	----	----

Glu Gly Cys Glu Asn Phe Asp Met Ile Glu Ser Ala Gln Phe Phe His

85	90	95
----	----	----

Lys Met Phe Glu Ala Thr Tyr Lys Leu Ala Glu Pro Ala Val Asn Ala

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

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

355 360 365

Cys Gly Trp Asn Ser Thr Leu Glu Gly Ile Thr Ala Gly Val Pro Met

370 375 380

Val Thr Trp Pro Gln Phe Thr Asp Gln Phe Ile Asn Glu Arg Phe Ile

385 390 395 400

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

405 410 415

Val Val Gly Asp Gln Asp Lys Phe Gly Val Leu Val Asn Lys Glu Glu

420 425 430

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

435 440 445

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

450 455 460

Met Glu Asp Gly Gly Ser Ser His Arg Asp Met Thr Ser Met Ile Gln

465 470 475 480

Asp Ile Val Glu Leu Cys Lys Asn Arg

485

<210> 13

<211> 1464

<212> DNA

<213> Catharanthus roseus

<400> 13

atggtaatc agtccatat tttcaactc ccattcatgg cacagggcca tatgttaccc	60
gccttagaca tggccaatct attcacttct cgtggagtca aagtaacatt aatcacaacc	120
catcaacatg ttcccatgtt tacaaaatcc atagaaagga gcagaaattc tggatttgat	180
atatccattc aatccatcaa attcccatgt tcagaagttg gtttacctga aggaatcgaa	240
agtcttagatc aagtttcagg ggacgacgaa atgcttccta agttcatgag aggagttat	300
ttactccaac aacctctcgaa acaactatttgc caagaatctc gtctcatttgc tcttcttct	360
gatatgttct tcccttggac tactgaatct gctgtaaat ttggatttcc cagattgttt	420

tttcatgggt cctgttcctt tgccctctct gcagctgaaa gtgtgagaag aaataaacct	480
---	-----

ttcgagaatg ttccacaga cacagaggaa ttgttgtgc ctgatcttc ccaccaaatt	540
aaattaacca gaacacaat tcaacatac gaaaggaaa atattgagtc agatttacc	600
aaaatgtga agaaagttag ggattcagaa tccacatctt acggagttgt agtcaatagt	660
ttctaigaac ttgaaccaga ttatgccat tattacatca acgttttggg aagaaaagca	720
tggcatatag ggccttttt gcttgtaac aaattacaag ctgaagataa agccaaagg	780
gggaagaaat cagcaattga tgcagacgaa tgttaaatt ggcttgattc gaaacaacca	840

aattccgtaa ttatctctg ttccgaagt atggccaatt taaattctgc ccaattacac	900
gaaattgcaa cagcccttga atcctccggc caaaatttca tctgggttgt tagaaaatgt	960
gtggacgaag aaaacagttc aaaatggtt ccagaaggat tcgaagaaag aacaaaagaa	1020
aaaggctaa ttataaaggg atgggcacca caaaccctaa ttcttgaaca cgaatcagta	1080
ggagcattt tgaccattt tggttggat tcaactctg aaggaatctg cgccagggtt	1140
cctctggta ctggccctt ctttgcttag caattttca atgagaattt gattacagag	1200
gtactgaaaa cgggatacgg agttgggct cggcaatgga gttagatccc aacagagatt	1260

ataaaaggag aagccatagc taatgctatt aatcgagtaa tggtggtga tgaagctgtt	1320
gagatgagaa acagagcaaa agatttgaag gaaaaggcaa gaaaagctt ggaagaagat	1380
ggatcttctt atcgatct tactgctttt attgaagaat tggggcata tcgttctcaa	1440
gttgaagaa agcaacaaga ctag	1464

&lt;210&gt; 14

&lt;211&gt; 487

&lt;212&gt; PRT

&lt;213&gt; Catharanthus roseus

&lt;400&gt; 14

Met Val Asn Gln Leu His Ile Phe Asn Phe Pro Phe Met Ala Gln Gly

1 5 10 15

His Met Leu Pro Ala Leu Asp Met Ala Asn Leu Phe Thr Ser Arg Gly

20 25 30

Val Lys Val Thr Leu Ile Thr Thr His Gln His Val Pro Met Phe Thr

35 40 45

Lys Ser Ile Glu Arg Ser Arg Asn Ser Gly Phe Asp Ile Ser Ile Gln

50 55 60

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

65	70	75	80
----	----	----	----

Ser Leu Asp Gln Val Ser Gly Asp Asp Glu Met Leu Pro Lys Phe Met

85	90	95	
----	----	----	--

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

100	105	110	
-----	-----	-----	--

Ser Arg Pro His Cys Leu Leu Ser Asp Met Phe Phe Pro Trp Thr Thr

115	120	125	
-----	-----	-----	--

Glu Ser Ala Ala Lys Phe Gly Ile Pro Arg Leu Leu Phe His Gly Ser

130	135	140	
-----	-----	-----	--

Cys Ser Phe Ala Leu Ser Ala Ala Glu Ser Val Arg Arg Asn Lys Pro

145	150	155	160
-----	-----	-----	-----

Phe Glu Asn Val Ser Thr Asp Thr Glu Glu Phe Val Val Pro Asp Leu

165	170	175	
-----	-----	-----	--

Pro His Gln Ile Lys Leu Thr Arg Thr Gln Ile Ser Thr Tyr Glu Arg

180	185	190	
-----	-----	-----	--

Glu Asn Ile Glu Ser Asp Phe Thr Lys Met Leu Lys Lys Val Arg Asp

195	200	205	
-----	-----	-----	--

Ser Glu Ser Thr Ser Tyr Gly Val Val Val Asn Ser Phe Tyr Glu Leu

210	215	220	
-----	-----	-----	--

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

225	230	235	240
-----	-----	-----	-----

Trp His Ile Gly Pro Phe Leu Leu Cys Asn Lys Leu Gln Ala Glu Asp

245	250	255	
-----	-----	-----	--

Lys Ala Gln Arg Gly Lys Lys Ser Ala Ile Asp Ala Asp Glu Cys Leu

260	265	270	
-----	-----	-----	--

Asn Trp Leu Asp Ser Lys Gln Pro Asn Ser Val Ile Tyr Leu Cys Phe

275	280	285	
-----	-----	-----	--

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

290	295	300	
-----	-----	-----	--

Ala Leu Glu Ser Ser Gly Gln Asn Phe Ile Trp Val Val Arg Lys Cys

305	310	315	320
-----	-----	-----	-----

Val Asp Glu Glu Asn Ser Ser Lys Trp Phe Pro Glu Gly Phe Glu Glu

325 330 335

Arg Thr Lys Glu Lys Gly Leu Ile Ile Lys Gly Trp Ala Pro Gln Thr

340 345 350

Leu Ile Leu Glu His Glu Ser Val Gly Ala Phe Val Thr His Cys Gly

355 360 365

Trp Asn Ser Thr Leu Glu Gly Ile Cys Ala Gly Val Pro Leu Val Thr

370 375 380

Trp Pro Phe Phe Ala Glu Gln Phe Phe Asn Glu Lys Leu Ile Thr Glu

385 390 395 400

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

405 410 415

Ser Thr Glu Ile Ile Lys Gly Glu Ala Ile Ala Asn Ala Ile Asn Arg

420 425 430

Val Met Val Gly Asp Glu Ala Val Glu Met Arg Asn Arg Ala Lys Asp

435 440 445

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

450 455 460

Arg Asp Leu Thr Ala Leu Ile Glu Glu Leu Gly Ala Tyr Arg Ser Gln

465 470 475 480

Val Glu Arg Lys Gln Gln Asp

485

<210> 15

<211> 1410

<212> DNA

<213> Arabidopsis thaliana

<400> 15

atggcgccac	cgcattttct	actggtaacg	tttccggcgc	aaggtcacgt	gaaccatct	60
ctccgttttg	ctcgctggct	catcaaaaga	accggcgac	gtgtcacttt	cgtcacttgt	120
gtctccgtct	tccacaactc	catgatcgca	aaccacaaca	aagtgcaaaa	tctctttc	180
cttactttct	ccgacggttt	cgacgatgga	ggcatttcca	cctacgaaga	ccgtcagaaa	240

aggtcggta atctcaaggt taacggcgat aaggcactat cgattcat cgaagctact	300
aagaatggtg actctccgt gacttgctt atctacacga ttcttctcaa ttgggctcca	360
aaagtagcac gttagattca acttccctcc gctttctct ggatccaacc ggcttggtt	420
ttcaacatct atiacactca tttcatggaa aacaagtccg tttcgagtt acctaatctg	480
tcttctctgg aaatcagaga tcttccatct ttccctcacac ctccaaacac aaacaaaggc	540
gcatacgtgcgttcaaga aatgatggag tttctataa aagaaaccaa accgaaaatt	600
ctcatcaaca ctgcgattc gctggAACCA gaggccttaa cgcttccc gaatatcgat	660

atggtggcg ttggccctt acttcccacg gagattttct caggaagcac caacaaatca	720
gttaaagatc aaagtagtag ttatacactt tggctagact cgaaaacaga gtcctctgtt	780
atttacgtt ccgttggAACAC aatgggttag ttgtccaaga aacagataga ggaactagcg	840
agagcactca tagaaggaa acgaccgtt ttgtgggtta taactgataa atccaaacaga	900
gaaacgaaaa cagaaggaga agaagagaca gagattgaga agatagctgg attcagacac	960
gagcttgaag agttggat gattgtgtcg tgggtttcgc agatagaggt tttaaatcac	1020
cgagccgttag gtgttttgt gactcatgtt gggtggagct cgacgctggaa gagtttggtt	1080

cttggcgttc cggttgtggc gttcccgatg tggtcgatc aaccgacgaa cgcaagcta	1140
ctggaaagaaa gtggaaagac tgggtgtgagg gtaagagaga acaaggatgg tttgggtggag	1200
agaggagaga tcaggaggtg tttggaaagcc gtgtggagg agaagtcgg ggagttgagg	1260
gaaaacgcaa agaaatggaa gcgttagcg atggaaagcgg gttagagaagg aggatctcg	1320
gataagaaca tggaggctt tgtggaggat atttggagg aatctttat tcaaaacttg	1380
tgtgaagcag aggaggtaaa agtaaagtaa	1410

&lt;210&gt; 16

&lt;211&gt; 469

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 16

Met Ala Pro Pro His Phe Leu Leu Val Thr Phe Pro Ala Gln Gly His

1 5 10 15

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

20 25 30

Ala Arg Val Thr Phe Val Thr Cys Val Ser Val Phe His Asn Ser Met

35 40 45

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

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

Glu Gly Glu Glu Glu Thr Glu Ile Glu Lys Ile Ala Gly Phe Arg His  
 305 310 315 320

Glu Leu Glu Glu Val Gly Met Ile Val Ser Trp Cys Ser Gln Ile Glu  
 325 330 335

Val Leu Ser His Arg Ala Val Gly Cys Phe Val Thr His Cys Gly Trp  
 340 345 350

Ser Ser Thr Leu Glu Ser Leu Val Leu Gly Val Pro Val Val Ala Phe  
 355 360 365

Pro Met Trp Ser Asp Gln Pro Thr Asn Ala Lys Leu Leu Glu Glu Ser  
 370 375 380

Trp Lys Thr Gly Val Arg Val Arg Glu Asn Lys Asp Gly Leu Val Glu  
 385 390 395 400

Arg Gly Glu Ile Arg Arg Cys Leu Glu Ala Val Met Glu Glu Lys Ser  
 405 410 415

Val Glu Leu Arg Glu Asn Ala Lys Lys Trp Lys Arg Leu Ala Met Glu  
 420 425 430

Ala Gly Arg Glu Gly Ser Ser Asp Lys Asn Met Glu Ala Phe Val  
 435 440 445

Glu Asp Ile Cys Gly Glu Ser Leu Ile Gln Asn Leu Cys Glu Ala Glu  
 450 455 460

Glu Val Lys Val Lys

465

<210> 17

<211> 1425

<212> DNA

<213> Arabidopsis thaliana

<400> 17

atgggaaagc aagaagatgc agagctcgac atcacatcc tccctttctc cggacacatt	60
ctcgcaacaa tcgaactcgc caaacgtctc ataagtcaag acaatcctcg gatccacacc	120
atcaccatcc tctattgggg attaccttt attcctcaag ctgacacaat cgcttcctc	180
cgtcccttag tcaaaaatga gcctcgatc cgtctcgta cgttgccgaa agtccaagac	240



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

Ile Ala Gln Ala Leu Glu Leu Val Gly Ile Arg Phe Leu Trp Ser Ile  
 305 310 315 320

Arg Thr Asp Pro Lys Glu Tyr Ala Ser Pro Asn Glu Ile Leu Pro Asp  
 325 330 335

Gly Phe Met Asn Arg Val Met Gly Leu Gly Leu Val Cys Gly Trp Ala  
 340 345 350

Pro Gln Val Glu Ile Leu Ala His Lys Ala Ile Gly Gly Phe Val Ser  
 355 360 365

His Cys Gly Trp Asn Ser Ile Leu Glu Ser Leu Arg Phe Gly Val Pro  
 370 375 380

Ile Ala Thr Trp Pro Met Tyr Ala Glu Gln Gln Leu Asn Ala Phe Thr  
 385 390 395 400

Ile Val Lys Glu Leu Gly Leu Ala Leu Glu Met Arg Leu Asp Tyr Val  
 405 410 415

Ser Glu Tyr Gly Glu Ile Val Lys Ala Asp Glu Ile Ala Gly Ala Val  
 420 425 430

Arg Ser Leu Met Asp Gly Glu Asp Val Pro Arg Arg Lys Leu Lys Glu  
 435 440 445

Ile Ala Glu Ala Gly Lys Glu Ala Val Met Asp Gly Gly Ser Ser Phe  
 450 455 460

Val Ala Val Lys Arg Phe Ile Asp Gly Leu  
 465 470

<210> 19

<211> 1473

<212> DNA

<213> Arabidopsis thaliana

<400> 19

atgggaaagc aagaagatgc agagctcgta atcacatcc tccctttctc cggacacatt	60
ctcgcaacaa tcgaactcgc caaacgtctc ataagtcaag acaatcctcg gatccacacc	120
atcaccatcc tctattgggg attaccttt attcctcaag ctgacacaat cgcttcctc	180

cgatccctag tcaaaaatga gcctcgatc cgtctcgta cggtgcccgaa agtccaagac	240
--	-----

cctccaccaa	tggaacttct	tgtggattt	gccgaatctt	acattcttga	atacgtcaag	300
aaaatggttc	ccatcatcag	agaagctctc	tccactctct	tgtctcccg	cgtatgaatcg	360
ggttcagttc	gtgtggcigg	attggttctt	gacttcttct	gcgtccctat	gatcgatgta	420
ggaaacgagt	ttaatctccc	ttcttacatt	ttcttgacgt	gtagcgcagg	gttcttggtt	480
atgatgaagt	atcttcaga	gagacaccgc	gaaatcaaat	cggaaatcaa	ccggagcttc	540
aacgaggagt	tgaatctcat	tcctggttat	gtcaactctg	ttcctactaa	ggttttgcgg	600

ttcatcgagc atgtatcgaa tgttacgatt taa 1473

<210> 20

<211> 490

<212> PRT

<213> *Arabidopsis thaliana*

<400> 20

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

1                    5                    10                    15

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

20 25 30

35	40	45
----	----	----

Pro Phe Ile Pro Gln Ala Asp Thr Ile Ala Phe Leu Arg Ser Leu Val

50	55	60
----	----	----

Lys Asn Glu Pro Arg Ile Arg Leu Val Thr Leu Pro Glu Val Gln Asp

65	70	75	80
----	----	----	----

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

85	90	95
----	----	----

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

100	105	110
-----	-----	-----

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

115	120	125
-----	-----	-----

Val Leu Asp Phe Phe Cys Val Pro Met Ile Asp Val Gly Asn Glu Phe

130	135	140
-----	-----	-----

Asn Leu Pro Ser Tyr Ile Phe Leu Thr Cys Ser Ala Gly Phe Leu Gly

145	150	155	160
-----	-----	-----	-----

Met Met Lys Tyr Leu Pro Glu Arg His Arg Glu Ile Lys Ser Glu Phe

165	170	175
-----	-----	-----

Asn Arg Ser Phe Asn Glu Glu Leu Asn Leu Ile Pro Gly Tyr Val Asn

180	185	190
-----	-----	-----

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

195	200	205
-----	-----	-----

Tyr Glu Pro Trp Val Glu Leu Ala Glu Arg Phe Pro Glu Ala Lys Gly

210	215	220
-----	-----	-----

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

225	230	235	240
-----	-----	-----	-----

Phe Asp Arg Cys Pro Asp Asn Tyr Pro Thr Ile Tyr Pro Ile Gly Pro

245	250	255
-----	-----	-----

Ile Leu Asn Leu Glu Asn Lys Lys Asp Asp Ala Lys Thr Asp Glu Ile

260	265	270
-----	-----	-----

Met Arg Trp Leu Asn Glu Gln Pro Glu Ser Ser Val Val Phe Leu Cys

275	280	285
-----	-----	-----

Phe Gly Ser Met Gly Ser Phe Asn Glu Lys Gln Val Lys Glu Ile Ala

290 295 300

Val Ala Ile Glu Arg Ser Gly His Arg Phe Leu Trp Ser Leu Arg Arg

305 310 315 320

Pro Thr Pro Lys Glu Lys Ile Glu Phe Pro Lys Glu Tyr Glu Asn Leu

325 330 335

Glu Glu Val Leu Pro Glu Gly Phe Leu Lys Arg Thr Ser Ser Ile Gly

340 345 350

Lys Val Ile Gly Trp Ala Pro Gln Met Ala Val Leu Ser His Pro Ser

355 360 365

Val Gly Gly Phe Val Ser His Cys Gly Trp Asn Ser Thr Leu Glu Ser

370 375 380

Met Trp Cys Gly Val Pro Met Ala Ala Trp Pro Leu Tyr Ala Glu Gln

385 390 395 400

Thr Leu Asn Ala Phe Leu Leu Val Val Glu Leu Gly Leu Ala Ala Glu

405 410 415

Ile Arg Met Asp Tyr Arg Thr Asp Thr Lys Ala Gly Tyr Asp Gly Gly

420 425 430

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

435 440 445

Ser Asp Gly Glu Ile Arg Asn Lys Val Lys Asp Val Lys Glu Lys Ser

450 455 460

Arg Ala Ala Val Val Glu Gly Gly Ser Ser Tyr Ala Ser Ile Gly Lys

465 470 475 480

Phe Ile Glu His Val Ser Asn Val Thr Ile

485 490

<210> 21

<211> 605

<212> DNA

<213> Crocus sativus

<400> 21

atggaattcg atcccgattg gttgcagtc atcaggaagc tgtattcgaa accgggtggc	60
gcagccggaa tgctgccaca actagtccccggg gaagaagaag aagcgaaag tggtaagagt	120
tggagtgccct taaaatggtt ggatctacaa gaagcaaaat ccgtcggttgcgcgttc	180
gaaagcgagg cgaagctcac cgtcgaacaa gtcggagaga tagcattagg gctcgagtgc	240
agcgggctga aatttatatg gactctgagg gccgacggct gcctcgaaa tttcgaggag	300
aggacaaagg accgggaaat gtttgatg ggatggatc cgagactaa gttctggct	360
cactcgtcgg tcgggtggtt ctgacgcat tgcgggtcga gctcgattgt cgagggctc	420
tcgttcggc tggtgatggt gcccctgcca atgggttca ccaaggcttcaatgctagg	480
tacgtgggg acaagaaaat cgggtggag gtggcggaaa atgaagagga tggacttt	540
actggtagag gaattgcaga tagttgaag ttagctatgt tggaaagaaga aggttaaggat	600
ttcag	605
<210> 22	
<211> 1305	
<212> DNA	
<213> Crocus sativus	
<400> 22	
atggaagctg gtggataa actccacata gtagtatttc catggcttgc cttcgccac	60
atgcttcctt tccttagct ctcaaaatct ctgcggaa gaggccatct cataccttc	120
gtatccaccc caaagaacat ccagagattt ccaaaatctc cttccacaaa tatctcttct	180
cataaaatcc atccctttat cactccccaa agtggaaagga tgccgggtga cgtcgaggcc	240
accacccgacc tcccgccggc aaacctccag tacctcaaaa aagccctcga cggcctcgag	300
cagccttcc ggagcttcct ccgagaagct tcccccaaac ccgattggat aatccaagac	360
cttctgcgc actggatacc accaatagcg gccgagctcc acgtgcgtc gatgtacttc	420
ggcacgggtgc cggccgcagc gttgacttgc ttccggccacc cgtcgagtt ctgcggcgt	480
aagaaaggaa tcgaggactg gctggttct ccggcgtggg tcccttccc ttccaaagggt	540
gcgtaccgcc tccacgagat gattgtgatg gcgaaagaca cggcggtcc cttccactcg	600
ggcggtgaccg acgtccgcgc catggaggcg gccatcggtgg gttgctgcgc cgtcgatcata	660
cgcacccgtcc gggagctgga gtcggagtggtgg ctggcgttgc tcaagagat ttacggaaag	720
cccggttcc cggtaggcct actgctgcct actgcccacg aaagcaccga tggcaatagt	780
attatcgatt ggctcgac gcaagccag gaatctgtgg tgtacatcgc gttggggagc	840
gagggtgtcca tcgggtggaa gctgatacac gagctggcgc tccggcgttgc gctcgccgggg	900

ttgccttcc ttgggctct caggaggccg tacgggttgt cgagcgatac cgagatcctg 960

cccgaaaaact tcgaggagcg gacgaggggg tacgggaagg tggtatggg gtgggtcca 1020  
 caaatgaggg ttttgccga tggtcggtg ggaggattcg tgacgcactg cggttgagg 1080  
 tcgggtgtgg agagcttgca tttggacac ccgctgttt tggccat attcgggac 1140  
 cagggctca acgcgaggct gttggaggag aaggatcg ggtcgaattg gagaggaagg 1200  
 gggacgcgtc tttacgcgg aatgaggtgg cgaaggctgt caatctggc atggtggaa 1260  
 gggatggatc agggatcg tataggaaga aagccaagga gatga 1305

<210> 23

<211> 434

<212> PRT

<213> Crocus sativus

<400> 23

Met Glu Ala Gly Gly Asp Lys Leu His Ile Val Val Phe Pro Trp Leu

1 5 10 15

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

20 25 30

Lys Arg Gly His Leu Ile Ser Phe Val Ser Thr Pro Lys Asn Ile Gln

35 40 45

Arg Phe Pro Lys Ser Pro Ser Thr Asn Ile Ser Ser His Lys Phe His

50 55 60

Pro Phe Ile Thr Pro Gln Ser Gly Arg Met Pro Gly Asp Val Glu Ala

65 70 75 80

Thr Thr Asp Leu Pro Pro Ala Asn Leu Gln Tyr Leu Lys Lys Ala Leu

85 90 95

Asp Gly Leu Glu Gln Pro Phe Arg Ser Phe Leu Arg Glu Ala Ser Pro

100 105 110

Lys Pro Asp Trp Ile Ile Gln Asp Leu Leu Gln His Trp Ile Pro Pro

115 120 125

Ile Ala Ala Glu Leu His Val Pro Ser Met Tyr Phe Gly Thr Val Pro

130 135 140

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

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

Gly Thr Arg Leu Leu Arg Gly Met Arg Trp Arg Arg Leu Ser Ile Trp  
 405 410 415  
 Ser Trp Trp Lys Gly Met Asp Gln Gly Val Arg Ile Gly Arg Lys Pro  
 420 425 430  
 Arg Arg

&lt;210&gt; 24

&lt;211&gt; 1395

&lt;212&gt; DNA

&lt;213&gt; Crocus sativus

&lt;400&gt; 24

atggaagctg gtggtgataa actccacata gtagtatttc catggctagc cttcgccac 60

atgcttcctt tccttagagct ctcaaaatct ctgc当地 aaga gaggccatct catatcattc  
 gtatccaccc caaagaacat ccagagattc ccaaattctcc ctccacaat atctcctctc  
 ataaatttca tccctttatc actccccaaa gtggaggca tgcccgtga cgtcgaggcc  
 accaccgacc tcccgccgac aaacctccag tacctaaaaa aagccctcga cggcctcgag  
 cagcctttcc ggagcttcct ccgagaagct tccccaaac ccgattggat aatccaagac  
 cttctgcagc actggatacc accaatagcg gccgagctcc acgtgccgtc gatgtacttc  
 ggcacgggtgc cgccgcagc gttgacttcc ttggccacc cgtcgaggat ctcgaagcgt 480

aagaaaggga tcgaggaccg gccgggttct cc当地 tgggg tcccttccc ttccaagggtg  
 gcgttccccc tccacgagat gattgtgatg gc当地 agaca cggcgggtcc cctccactcg  
 ggcgtgaccg acgtccgccc catggaggcg gccatcgatgg gttgctgcgc cgtcgcgata  
 cgcacctgcc gggagctgga gtc当地 gggatgg ctgcccattc tc当地 agagat ttacggaaag  
 cccgtgattc cggtaggcct actgctgcct actgcccacg aaagcaccga tggcaatagt  
 attatcgatt ggctcggcac gc当地 agccag gaatctgtgg tgtacatcgc gttggggagc  
 gaggtgtcca tc当地 ggtgtgg gctgatacac gagctggcgcc tc当地 ctcgaagcgt 720

ttgc当地 tttggctct caggaggccg tacgggttgt cgagcgatac cgagatcctg  
 cccggggct tc当地 gggccg gacgaggggg tacgggagg tggatggg gtgggtccca  
 caaatgaggg tggatggccga tggatggccg ggaggattcg tgacgcactg cggatggag  
 tc当地 ggtgg agagcttgca tttggacac cc当地 ttgtgttt tgatggccat attcggggac  
 caggggctca acgc当地 gggct gttggaggaa agggatcg gggatcg ggatggaaag  
 gggacgcgt ct当地 tacgcg gaatgaggatg gc当地 aggctg tcaatctgtc atggatggaaag 1260

gggatggat cagggagttc gtataggaag aaagccaagg agatgaagaa gatttttgt 1320  
 gacaaagagt gccaggagaa gtatgtggat gagtttattc agttcttgct cagtaatgga 1380  
 acagcaaaag ggtag 1395  
 <210> 25  
 <211> 464  
 <212> PRT  
 <213> Crocus sativus  
 <400> 25  
 Met Glu Ala Gly Gly Asp Lys Leu His Ile Val Val Phe Pro Trp Leu  
 1 5 10 15  
 Ala Phe Gly His Met Leu Pro Phe Leu Glu Leu Ser Lys Ser Leu Ala  
 20 25 30  
 Lys Arg Gly His Leu Ile Ser Phe Val Ser Thr Pro Lys Asn Ile Gln  
 35 40 45  
 Arg Phe Pro Asn Leu Pro Pro Gln Ile Ser Pro Leu Ile Asn Phe Ile  
 50 55 60  
 Pro Leu Ser Leu Pro Lys Val Glu Gly Met Pro Gly Asp Val Glu Ala  
 65 70 75 80  
 Thr Thr Asp Leu Pro Pro Ala Asn Leu Gln Tyr Leu Lys Lys Ala Leu  
 85 90 95  
 Asp Gly Leu Glu Gln Pro Phe Arg Ser Phe Leu Arg Glu Ala Ser Pro  
 100 105 110  
 Lys Pro Asp Trp Ile Ile Gln Asp Leu Leu Gln His Trp Ile Pro Pro  
 115 120 125  
 Ile Ala Ala Glu Leu His Val Pro Ser Met Tyr Phe Gly Thr Val Pro  
 130 135 140  
 Ala Ala Ala Leu Thr Phe Phe Gly His Pro Ser Glu Phe Ser Lys Arg  
 145 150 155 160  
 Lys Lys Gly Ile Glu Asp Arg Pro Gly Ser Pro Pro Trp Val Pro Phe  
 165 170 175  
 Pro Ser Lys Val Ala Tyr Arg Leu His Glu Met Ile Val Met Ala Lys

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

Lys Glu Met Lys Lys Ile Phe Gly Asp Lys Glu Cys Gln Glu Lys Tyr

435 440 445

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

450 455 460

<210> 26

<211> 518

<212> DNA

<213> Crocus sativus

<400> 26

gcacggaatt ccatgccat tggctacaag tcatacgaa gctgtacacg aagccggcg	60
tcactgtcg attttgcca cccagaaga ttcaaattga catgagctgg gcgagtgc	120

tcaaatggct ggatctacaa gcagcaaaat ccgttgttta tgtcgcttc ggaagcgagg	180
cgaagctcac cgtcgaacaa gtcggagaga tagcattagg gctcgagtcg agcgggctga	240
aatttatatg gactctgagg gccgacggct tgcctcgaaa tttcgaggag aggacaaagg	300
accggggaaat gtttgcaag ggctggattc cacagactag gttctggct cattcgtcgg	360
tccgggtt ctgtacat ggcgggtcga gctcgatgt cgaggggctc tcgttcggc	420
tggtgatggt ggtgtgcca ctgttgaaa gccaaggct caatgctagg tatttgggg	480
acaagaaat tgggtcgag gtgccgagaa atgaggag	518

<210> 27

<211> 882

<212> DNA

<213> Crocus sativus

<400> 27

acaaggacgc gctgaaccac ttctcgaa gttattcgaa actgcgacgt ggtcgccgt	60
cgcggatgca acgagttcga gcgggagtgg atcgagctgt tggtggagct ctacaagaag	120
accgttttgc cccctcgccc acttccctcc caattggagc aacaactacg atggcgatgg	180
agatagggag tataacggaa tcttccatg gctggacgtt caagaacgtg gctccgtcgt	240
ctacgtggct ttccggcagcg aggtgagact ggagaaggaa caggttcgca agatcgact	300
cgggctcgag ctgtcgaaat tcccttctt gtggggccctg aggtggcccg ccgcctgga	360

cgcacgtgctg ccggagggt tcgagacccg gacggccgaa cggggatcg tgcgtacgg	420
gtgggtccca caggcgagaa tcctggcca cccggcggtc gggggattcg tgacgcactg	480
cgggtggaaat tcggtcgtgg agggatgcg gtggggccctt ggcgtgattc tggtgcccgt	540

gattttcgac caggcgctca acgcgaggct cctcgaggac aaggagatcg gaaaggaggt	600
gccgaggggac gaggaggatg gaaagttcga tggggaggga attgcagaaga cgatgaggat	660
agtgtatggtg gaagaggagg gagacagtat tagggttgcg agtagagcga tgaagagtgt	720
gcttggggac aaggaggtgc aagattccta catigtcaag ttgtggagt acttgaaga	780
gcataggcca acttaatgag aaattcttac gcagaccgag tctacggcca cgggggtggat	840
tgaacggcta ataatttttc cgaaattat gccttgtaat gc	882
<210> 28	
<211> 1395	
<212> DNA	
<213> Crocus sativus	
<400> 28	
atggaagctg gtgggtataa actccacata gtagtatttc catggctagc ctccggcac	60
atgccttcctt tccttagagct ctcaaaatct ctgc当地 gaggccatct catatccttc	120
gtatccaccc caaagaacat ccagagattc ccaaacttcc ctccacaat atctcctctc	180
ataaatttca tcccttatac actccccaaa gtggaggca tgccggcga cgtcgaggcc	240
accacccgacc tcccgccggc aaacctccag tacctcaaaa aagccctcgaa cggcctcgag	300
cagcctttcc ggagcttcctt cc当地 gagatggat aatccaagac	360
cttccatcacttcc actggataacc accaatagcg gccc当地 gctccactcg gatgtacttc	420
ggcacgggtgc cggccgc当地 gttgacttcc ttccggccacc cgtcgagtt gtc当地 gccgc	480
ggtaaagggc tecagggtctg gtc当地 gtttcc ttccaaagggtg	540
gcgttccaccc tccacgagtt gattgtgatg gcgaaagacg cggcgggtcc cctccactcg	600
ggcatgaccg acgccc当地 cc当地 catggaggcg gccatcgatgg gttgctgc当地 cgtcgagata	660
cgcacccgtcc gggagctgga gtc当地 gatggctt ctggccatc tcaagagat ttacgggaa	720
cccgatcccgatcc cgatggccctt actgctgcctt actgccc当地 acgaaaccgaa tggcaatagt	780
attatcgattt ggctccggccac gcaagccag gaatctgtgg tgtacatcgc gttggggagc	840
gagggtgtcca tc当地 ggttggaa gctgatacc gagctggccg tccggccatc gctcgccggg	900
ttgccttcctt tttggctctt caggaggccg tacgggttgtt cgagcgatcc cgatccctg	960
cccgccggctt tc当地 gggagccg gacgaggggg tacgggaaagg tggatggg gtc当地 cccca	1020
caaatgaggg tggatggccgat taggtcgatgg ggaggattcg tgc当地 cactcg cgttggagt	1080
tc当地 ggttggaa agagcttgca tttggacac cc当地 ttgtttt tggccat attcggggac	1140
caggggctca acgcgaggct gttggaggag aaaggatcg gggc当地 gaggttggagg ggagagga	1200

ggggacgggt ctttacgag gaatgaggtg gcgaaggcga tcaatctgat catggtgaa 1260  
 ggggatggat cggtagttc gtataggaag aaagcgaagg agatgaagaa gatttcgca 1320  
 gacaaagaat gccaggagaa gtatgtggat gagttgttc agtcttgct cagtaatgga 1380  
 acagcaaaag ggtag 1395

&lt;210&gt; 29

&lt;211&gt; 464

&lt;212&gt; PRT

&lt;213&gt; Crocus sativus

&lt;400&gt; 29

Met Glu Ala Gly Gly Asp Lys Leu His Ile Val Val Phe Pro Trp Leu

1 5 10 15

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

20 25 30

Lys Arg Gly His Leu Ile Ser Phe Val Ser Thr Pro Lys Asn Ile Gln

35 40 45

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

50 55 60

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

65 70 75 80

Thr Thr Asp Leu Pro Pro Ala Asn Leu Gln Tyr Leu Lys Lys Ala Leu

85 90 95

Asp Gly Leu Glu Gln Pro Phe Arg Ser Phe Leu Arg Glu Ala Ser Pro

100 105 110

Lys Pro Asp Trp Ile Ile Gln Asp Leu Leu Gln His Trp Ile Pro Pro

115 120 125

Ile Ala Ala Glu Leu His Val Pro Ser Met Tyr Phe Gly Thr Val Pro

130 135 140

Ala Ala Ala Leu Thr Phe Phe Gly His Pro Ser Gln Leu Ser Ser Arg

145 150 155 160

Gly Lys Gly Leu Glu Gly Trp Leu Ala Ser Pro Pro Trp Val Pro Phe

165 170 175

Pro Ser Lys Val Ala Tyr Arg Leu His Glu Leu Ile Val Met Ala Lys

180                    185                    190

Asp Ala Ala Gly Pro Leu His Ser Gly Met Thr Asp Ala Arg Arg Met

195                    200                    205

Glu Ala Ala Ile Val Gly Cys Cys Ala Val Ala Ile Arg Thr Cys Arg

210                    215                    220

Glu Leu Glu Ser Glu Trp Leu Pro Ile Leu Glu Glu Ile Tyr Gly Lys

225                    230                    235                    240

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

245                    250                    255

Asp Gly Asn Ser Ile Ile Asp Trp Leu Gly Thr Arg Ser Gln Glu Ser

260                    265                    270

Val Val Tyr Ile Ala Leu Gly Ser Glu Val Ser Ile Gly Val Glu Leu

275                    280                    285

Ile His Glu Leu Ala Leu Gly Leu Glu Leu Ala Gly Leu Pro Phe Leu

290                    295                    300

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

305                    310                    315                    320

Pro Gly Gly Phe Glu Glu Arg Thr Arg Gly Tyr Gly Lys Val Val Met

325                    330                    335

Gly Trp Val Pro Gln Met Arg Val Leu Ala Asp Arg Ser Val Gly Gly

340                    345                    350

Phe Val Thr His Cys Gly Trp Ser Ser Val Val Glu Ser Leu His Phe

355                    360                    365

Gly His Pro Leu Val Leu Leu Pro Ile Phe Gly Asp Gln Gly Leu Asn

370                    375                    380

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

385                    390                    395                    400

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

405                    410                    415

Ile Met Val Glu Gly Asp Gly Ser Gly Ser Ser Tyr Arg Lys Lys Ala

420                    425                    430

Lys Glu Met Lys Lys Ile Phe Ala Asp Lys Glu Cys Gln Glu Lys Tyr

435 440 445

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

450 455 460

<210> 30

<211> 1395

<212> DNA

<213> Crocus sativus

<400> 30

atggaagctg gtggtgataa actccacata gtagtattc catggctagc cttcgccac	60
atgcttcctt tccttagagct ctcaaaatct ctcgcaaaga gaggccatct catatcttc	120
gtatccaccc caaagaacat ccagagattc ccaaatactcc ctccacaat atctcctctc	180
ataaatttca tcccttatac actccccaaa gtggaaggca tgcccggtga cgtcgaggcc	240

accaccgacc tcccgccggc aaacctccag tacctaataa aagccctcgaa cgccctcgag	300
cagccttcc ggagcttctt ccgagaagct tccccaaac ccgattggat aatccaagac	360
cttctgcage actggataacc accaatagcg gccgagctcc acgtgccgtc gatgtacttc	420
ggcacgggtgc cggccgcagc gttgacttcc ttccggccacc cgtcgaggat ctcgaagcgt	480
aagaaaggaa tcgaggactg gctggttctt ccggcgtgg tcccttccc ttccaaggat	540
gcgttaccggcc tccacgagat gattgtatg gcgaaagaca cggcgggtcc cctccactcg	600
ggcgtgaccg acgtccggccg catggaggcg gccatcgatgg gttgctgcgc cgtcgcgata	660

cgcacctgcc gggagctgga gtcggagtgg ctggcgatcc tcgaagagat ttacggaaag	720
cccggttattt cggtaggcctt actgctgcctt actggccacg aaaggcaccga tggcaatagt	780
attatcgattt ggctcgacgc gcgaagccag gaatctgtgg tgtacatcgc gttggggagc	840
gagggtgtcca tcgggtgtgg agtgtatcac gagctggcgc tccggctcgaa gctcgccggg	900
ttgcctttcc ttgggtctt caggaggccg tacgggttgtt cgagcgatac cgagatcctg	960
cccgaaaaatc tcgaggagcg gacgaggggg tacgggaagg tgggtatggg gtgggtccca	1020
caaataatgggg tgggtggccga tgggtcggtt ggaggattcg tgacgcactg cggttggat	1080

tcgggtgtgg agagcttgca tttggacac ccgttggttt tgggtccgtt attcgccggac	1140
caggggctca acgcgaggctt gttggaggaa aaggatcg gggtcgaatgtt ggaggaa	1200
ggggacgcgtt cttaacgcgtt gaatggatgtt gcaaggctgtt taaatctgtt catgggtggaa	1260
ggggatggat caggagttt gatggatggaa aagccaagg agatgaagaa gatgtttggat	1320

gacaaaaggt gccaggagaa gtatgtggat gagtttattc agttcttgct cagtaatgga 1380  
 acagcaaaag ggtag 1395  
 <210> 31  
 <211> 464  
 <212> PRT  
 <213> Crocus sativus

<400> 31  
 Met Glu Ala Gly Gly Asp Lys Leu His Ile Val Val Phe Pro Trp Leu  
 1 5 10 15  
 Ala Phe Gly His Met Leu Pro Phe Leu Glu Leu Ser Lys Ser Leu Ala  
 20 25 30  
 Lys Arg Gly His Leu Ile Ser Phe Val Ser Thr Pro Lys Asn Ile Gln  
 35 40 45  
 Arg Phe Pro Asn Leu Pro Pro Gln Ile Ser Pro Leu Ile Asn Phe Ile  
 50 55 60

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

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

Asp Thr Ala Gly Pro Leu His Ser Gly Val Thr Asp Val Arg Arg Met

195 200 205

Glu Ala Ala Ile Val Gly Cys Cys Ala Val Ala Ile Arg Thr Cys Arg

210 215 220

Glu Leu Glu Ser Glu Trp Leu Pro Ile Leu Glu Glu Ile Tyr Gly Lys

225 230 235 240

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

245 250 255

Asp Gly Asn Ser Ile Ile Asp Trp Leu Gly Thr Arg Ser Gln Glu Ser

260 265 270

Val Val Tyr Ile Ala Leu Gly Ser Glu Val Ser Ile Gly Val Glu Leu

275 280 285

Ile His Glu Leu Ala Leu Gly Leu Glu Leu Ala Gly Leu Pro Phe Leu

290 295 300

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

305 310 315 320

Pro Gly Gly Phe Glu Glu Arg Thr Arg Gly Tyr Gly Lys Val Val Met

325 330 335

Gly Trp Val Pro Gln Met Arg Val Leu Ala Asp Gly Ser Val Gly Gly

340 345 350

Phe Val Thr His Cys Gly Trp Ser Ser Val Val Glu Ser Leu His Phe

355 360 365

Gly His Pro Leu Val Leu Leu Pro Ile Phe Gly Asp Gln Gly Leu Asn

370 375 380

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

385 390 395 400

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

405 410 415

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

420 425 430

Lys Glu Met Lys Lys Ile Phe Gly Asp Lys Glu Cys Gln Glu Lys Tyr

435

440

445

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

450

455

460

&lt;210&gt; 32

&lt;211&gt; 1395

&lt;212&gt; DNA

&lt;213&gt; Crocus sativus

&lt;400&gt; 32

atgaaagctg gtggtgataa actccacata gtagtatttc caaggcttagc cttcgccac	60
atgccttcctt tccttagagct ctacaaaatc tctcgaaag agaggccatc tcatatcctt	120
cgtatccacc caagaacat ccagagattc ccaaattctcc ctccacaat atctccttc	180
ataaatttca tcccttatac actccccaaa gtggaaggca tgcccggtga cgtcgaggcc	240

accacccgacc tcccggcggc aaacctccag tacctaaaaa aagccctcga cggcctcgag	300
cagcctttcc ggagcttctt ccgagaagct tccccaaac ccgattggat aatccaagac	360
cttctgcagc actggatacc accaatagcg gccgagctcc acgtgccgtc gatgtacttc	420
ggcacgggtgc cggccgcgc gttgactttc ttccggcacc cgtcgagttt gtcgagccgc	480
ggtaaaggc tcgagggtct gctggcttcc ccggcgtggg tcccttccc ttccaagggtg	540
gcgtaccgcc tccacgagtt gattgtgtatg gcgaaagacg cggcgggtcc cctccactcg	600
ggcatgaccg acgccccccg catggaggcg gccatcgtgg gttgctgcgc cgtcgcgata	660

cgcacctgcc gggagctgga gtcggagtgg ctggccattc tcgaagagat ttacgggaag	720
cccggttcc cggtggctt actgctgcctt actggccacg aaaggccatc tggcaatagt	780
attatcgatt ggctcgccac gcgaagccag gaatctgtgg tgtacatcgc gttggggagc	840
gagggtgtcca tcgggtgtgg gctgatacac gagctggcgc tccggccatc gctcgccggg	900
ttgccttcc ttgggtctt caggaggccg tacgggttgtt cgagcgatac cgagatctg	960
cccgccccct tcgaggagcg gacggggggg tacgggaagg tggtgatggg gtgggtccca	1020
caaatacgagg tggtgccga tgggtcggtt ggaggattcg tgacgcactg cggttggagt	1080

tcgggtgtgg agagcttgca tttggacac ccgttgcattt tggccat attcgccac	1140
caggggctca acgcgaggctt gttggaggaa aaggatcg gggtcaagt ggagaggaa	1200
ggggacgcgtt cttaacgcgtt gaatggatgtt gcgaaggctt tcaatctgtt catgggtggaa	1260
ggggatggat caggagttc gtataggaag aaagccaggat agatgaagaa gattttggat	1320
gacaaagatg gccaggagaa gtatgtggat gagtttattc agttctgtt cagtaatggaa	1380

acagcaaaag ggtag 1395

<210> 33

<211> 464

<212> PRT

<213> Crocus sativus

<400> 33

Met Glu Ala Gly Gly Asp Lys Leu His Ile Val Val Phe Pro Trp Leu

1 5 10 15

Ala Phe Gly His Met Leu Pro Phe Leu Glu Leu Tyr Lys Ile Ser Arg

20 25 30

Lys Glu Arg Pro Ser His Ile Leu Arg Ile His Pro Lys Asn Ile Gln

35 40 45

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

50 55 60

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

65 70 75 80

Thr Thr Asp Leu Pro Pro Ala Asn Leu Gln Tyr Leu Lys Lys Ala Leu

85 90 95

Asp Gly Leu Glu Gln Pro Phe Arg Ser Phe Leu Arg Glu Ala Ser Pro

100 105 110

Lys Pro Asp Trp Ile Ile Gln Asp Leu Leu Gln His Trp Ile Pro Pro

115 120 125

Ile Ala Ala Glu Leu His Val Pro Ser Met Tyr Phe Gly Thr Val Pro

130 135 140

Ala Ala Ala Leu Thr Phe Phe Gly His Pro Ser Gln Leu Ser Ser Arg

145 150 155 160

Gly Lys Gly Leu Glu Gly Trp Leu Ala Ser Pro Pro Trp Val Pro Phe

165 170 175

Pro Ser Lys Val Ala Tyr Arg Leu His Glu Leu Ile Val Met Ala Lys

180 185 190

Asp Ala Ala Gly Pro Leu His Ser Gly Met Thr Asp Ala Arg Arg Met

195	200	205
Glu Ala Ala Ile Val Gly Cys Cys Ala Val Ala Ile Arg Thr Cys Arg		
210	215	220
Glu Leu Glu Ser Glu Trp Leu Pro Ile Leu Glu Glu Ile Tyr Gly Lys		
225	230	235
Pro Val Ile Pro Val Gly Leu Leu Leu Pro Thr Ala Asp Glu Ser Thr		
245	250	255
Asp Gly Asn Ser Ile Ile Asp Trp Leu Gly Thr Arg Ser Gln Glu Ser		
260	265	270
Val Val Tyr Ile Ala Leu Gly Ser Glu Val Ser Ile Gly Val Glu Leu		
275	280	285
Ile His Glu Leu Ala Leu Gly Leu Glu Leu Ala Gly Leu Pro Phe Leu		
290	295	300
Trp Ala Leu Arg Arg Pro Tyr Gly Leu Ser Ser Asp Thr Glu Ile Leu		
305	310	315
320		
Pro Gly Gly Phe Glu Glu Arg Thr Arg Gly Tyr Gly Lys Val Val Met		
325	330	335
Gly Trp Val Pro Gln Met Arg Val Leu Ala Asp Gly Ser Val Gly Gly		
340	345	350
Phe Val Thr His Cys Gly Trp Ser Ser Val Val Glu Ser Leu His Phe		
355	360	365
Gly His Pro Leu Val Leu Leu Pro Ile Phe Gly Asp Gln Gly Leu Asn		
370	375	380
Ala Arg Leu Leu Glu Glu Lys Gly Ile Gly Val Glu Val Glu Arg Lys		
385	390	395
400		
Gly Asp Ala Ser Phe Thr Arg Asn Glu Val Ala Lys Ala Val Asn Leu		
405	410	415
Val Met Val Glu Gly Asp Gly Ser Gly Ser Ser Tyr Arg Lys Lys Ala		
420	425	430
Lys Glu Met Lys Lys Ile Phe Gly Asp Lys Glu Cys Gln Glu Lys Tyr		
435	440	445

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

450 455 460

<210> 34

<211> 1380

<212> DNA

<213> Crocus sativus

<400> 34

atggaagctc aagctggtag tgcatacgaa cttcacgtat tagtattccc atggctatgt	60
tgcgtaca tgcataccctt ctcgaacta tcaaaaatctc tagcaagaag aggccatctc	120
atatccctca tctccaccccc gaagaacatc gatagactcg tccaaacctt ccattccgaa	180
ataaatttca tcccttctc gtcctccaaa gtagaaggca tggccgagtc cgtcgaggcc	240

accaccgacc tcccacccga agaccttccg tacctaaga aagcccttaa cgggctcgaa	300
aagcccttca accgggtcat ccgagaagct tctccaaac ccgattggat aatccagcat	360
tgggtccac caatagcggc cgagtccggt gcgtcgctaa tttcttcaa gactttccct	420
gccgccacca tctccatcc gggccaccc tccacttggc tgacccgcgaa acttgacagc	480
tggctagttt ctccctcggt ggtcccttc cttccaaag tgcgtatcg cctccacgag	540
atgccttcca tggccgcaag ccaccaacaa cagaagtcca acaccgacac cggcccccac	600
ctcaccgaga tcgtcgtaa tggctgcaac gccctagcgg tacggagctg catggagctg	660

gagccggacc acctgtccgct ctcgagaag atctacaaaa tgcggattt tccgttaggc	720
ctacttcctc cggcccgacca agtccggcgc gatgccagca attgcgatat tatggattgg	780
ctcggtacac agagtcaaaa aaccgtccctc tatatcggt tggtggagcga agtgcggc	840
cgttaacgacg tgcatacacga gctggctctg ggtctagagc tgcgggggtt gccttcctt	900
tgggctctca ggaggccctt cggctccgcc gggaaatgtcg ggatgttgcc cgaggcgttc	960
gaagagcga cgaaggagta tggaaagggtg gcgatggagt gggtcccaca aatggaggtg	1020
ctggccgacg agtccgtggg aggattcttgc acgcactgctg gttgggttc ggtcgtag	1080

agtctgcatt ttggacatcc gtcgtgtat tgcggctct tcggcgatca gttccat	1140
gcgcgtatgt tggaggagaa agggatcgga gtcgaggatgg agaaggagga ggtgggtct	1200
tttacgaggg atgacgtggc gaaggcagtgc aagtgtatca tgggtggaaaga ggaggaggt	1260
cggatagga agaaagctcg ggagatgaag acgggtctcg gggtaaaaaa atgtcaagac	1320
aagtatgcgg atgagttgt tcgtatctg ctcaatcatc gtgcgtccat acagaactga	1380

<210> 35

<211> 459

&lt;212&gt; PRT

&lt;213&gt; Crocus sativus

&lt;400&gt; 35

Met Glu Ala Gln Ala Gly Ser Asp Gln Lys Leu His Val Val Val Phe

1 5 10 15

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

20 25 30

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

35 40 45

Asn Ile Asp Arg Leu Val Pro Asn Leu His Pro Glu Ile Asn Phe Ile

50 55 60

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

65 70 75 80

Thr Thr Asp Leu Pro Pro Glu Asp Leu Pro Tyr Leu Lys Lys Ala Leu

85 90 95

Asn Gly Leu Glu Lys Pro Phe Asn Arg Phe Ile Arg Glu Ala Ser Pro

100 105 110

Lys Pro Asp Trp Ile Ile Gln His Trp Val Pro Pro Ile Ala Ala Glu

115 120 125

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

130 135 140

Ser Phe Met Gly His Pro Ser His Trp Met Thr Arg Glu Leu Asp Ser

145 150 155 160

Trp Leu Val Ser Pro Pro Trp Val Pro Phe Pro Ser Lys Val Ser Tyr

165 170 175

Arg Leu His Glu Met Leu Ser Met Ala Ala Ser His Gln Gln Gln Lys

180 185 190

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

195 200 205

Cys Asn Ala Leu Ala Val Arg Ser Cys Met Glu Leu Glu Pro Asp His

210 215 220

Leu Pro Leu Leu Glu Lys Ile Tyr Lys Met Pro Val Phe Pro Val Gly  
 225 230 235 240  
 Leu Leu Pro Pro Ala Gln Gln Val Gly Gly Asp Ala Ser Asn Cys Asp  
 245 250 255  
 Ile Met Asp Trp Leu Gly Thr Gln Ser Gln Lys Thr Val Leu Tyr Ile  
 260 265 270  
 Ala Leu Gly Ser Glu Val Thr Val Arg Asn Glu Leu Ile His Glu Leu  
 275 280 285  
 Ala Leu Gly Leu Glu Leu Ala Gly Leu Pro Phe Leu Trp Ala Leu Arg  
 290 295 300  
 Arg Pro Phe Gly Ser Ala Gly Asn Val Gly Met Leu Pro Glu Gly Phe  
 305 310 315 320  
 Glu Glu Arg Thr Lys Glu Tyr Gly Lys Val Ala Met Glu Trp Val Pro  
 325 330 335  
 Gln Met Glu Val Leu Ala Asp Glu Ser Val Gly Gly Phe Leu Thr His  
 340 345 350  
 Cys Gly Trp Gly Ser Val Val Glu Ser Leu His Phe Gly His Pro Leu  
 355 360 365  
 Val Met Leu Pro Val Phe Gly Asp Gln Phe Leu Asn Ala Arg Met Leu  
 370 375 380  
 Glu Glu Lys Gly Ile Gly Val Glu Val Glu Lys Glu Glu Asp Gly Ser  
 385 390 395 400  
 Phe Thr Arg Asp Asp Val Ala Lys Ala Val Lys Leu Ile Met Val Glu  
 405 410 415  
 Glu Glu Gly Gly Arg Tyr Arg Lys Lys Ala Arg Glu Met Lys Thr Val  
 420 425 430  
 Leu Gly Gly Lys Lys Cys Gln Asp Lys Tyr Ala Asp Glu Phe Val Gln  
 435 440 445  
 Tyr Leu Leu Asn His Arg Ala Ser Ile Gln Asn  
 450 455

&lt;211&gt; 464

&lt;212&gt; DNA

&lt;213&gt; Crocus sativus

&lt;400&gt; 36

ggggagcgag gtgacgatga gcgcggagct gacgaagagc tggctctggg gctggagctc	60
gccgggctgc cttccctgtg ggccctcggg aggccttcg gtcgtctcc tcatacgat	120
gtcgagattt tgccggaggg gttcgaagag cggacgaggg gattcggcaa agtggcgagg	180
gttgggtcc cgcaactcga ggtttggcc cactccggg tgggaggatt cttgacgcac	240
tgcgggttg gatcgattat cgagggccta catttcggac gtccgctgat cctgatgccg	300
gtgagggggg accagggct caacgcgagg ctgttgagg agaaaggat cggagtcgag	360

gtggagagga aggaggacgg gtcgtttagc aggaatgagg tggcgaaggc cgtgaagtta	420
gtcatggtgg aagaggatgg agggtcgtat agggagaaag cgcg	464

&lt;210&gt; 37

&lt;211&gt; 1383

&lt;212&gt; DNA

&lt;213&gt; Saccharomyces cerevisiae

&lt;400&gt; 37

atgctaaacg gcaataagtgc cacatccta ctttaccat gcccagctca gggcacatc	60
aatccaatttca tacaattcgg gaagagactt gcatctata acttgttgc aacacttgc	120
aacaccaggt ttctgtctaa cagtaactaag tcagaaccag gtccagttaa cattcaatgt	180
atctcagatg gttcgaccc aggaggcatg aacgctgcac catctagaag agcttacttc	240

gacagaccac agtcaagatc aggtcaaaaag catgttagtc taattgaatc tttaagatcc	300
agaggttagac caggtgcttg tttggatta agaccagtgc cattgtggc aatgaatgtt	360
gctgaaaggt caggtcttag gtctgttagca tttttactc aaccttgc agtcgatacc	420
atatacagac atgtgtggaa aggaggatt aagggtccag ttgctgaacc tgtcagatta	480
cctggcttgc ctccattaga accaagtgc ttaccttgc tttagaaatgg atttggtaga	540
gtagtcaacc cagacctact acctttgcgt gttaatcaac aaaaaattt agataaagct	600
gatatgatgg gcagaaatag tatctacgaa ttggaaggcag accttgcgt tggtcccgt	660

ttgccactgc ctgtgaaatc aattggccca accgtccctt caacttaccc ggataatcgt	720
atcccttcag actcacacta tgggttaac ctatacactc cagatactac cccatatctg	780
gattggctgg actctaaagc acctaattct gtcatctacg tctcatttgc ttcaatcc	840
tctttgtctc ctgatcaaac aaatgagatt gcttgcgt tgatagctac aaacaatcc	900

tttatctggg tagtcagaac atccgagttt gcaaaggta ctgcattt cactcaagag	960
aatgcctcca gaggggttgt agtaacttgg tgtgatcaat tggatcttt agcacatgt	1020
gctacaggct gtttgttac acattgttgt tggaactcta ctatggaagg ggttagccct	1080
ggtgttccaa tggggcgt gcctcaatgg tctgaccagg caatgaatgc caaatacg	1140
gaagatgtgt ggaaaggatgg agttagagct aaaacatatg gtaaagactt cgtaggaga	1200
gaagagttaa aaagatgcgt tgaagaggtt atggatgggg aaagatccgg aaagattaga	1260
gaaaatgccg ccagatgggt taaactggcc aaagattctg taagtgaagg cggcagtagt	1320
gataagtaca taaaggagtt catacatcaa tggcaatg attccaaaat ctcattggta	1380
taa	1383
<210> 38	
<211> 1422	
<212> DNA	
<213> <i>Saccharomyces cerevisiae</i>	
<400> 38	
atggatggcg atgggtggg tggatcaatg catgttgtt tggatccatg gcttgcattt	60
ggtcacattt tgcatttcgc tgaatttgc aagaggtag cacgtcaagg gcatagagta	120
acactattct caacccttag aaatactaga agattgtatcg acgtaccacc aagtcttagct	180
ggcagaatac gtgttgaga cattcctcta ccaagagtgc aacacttgcc tgaacacgct	240
gaagcaacaa tcgatcttcc atccaacgt cttagaccat acttagctg agcatatgac	300
gaggcctttt cttagagagtt atcaagattt ttgcaggaaa ctggccatc aagaccagac	360
tgggtgttgg ctgactatgc cgcatactgg gctccagctg ctgttcaag gcacgggtgc	420
ccatgtgcct tcctgtctt gttggcgct gccgcattgt gctttttgg tccagccaa	480
accttacaag gcagaggccc atacgcacaa acagagccgt ctcatttac tgcagtgct	540
gaataacgtgc ctcccaac aactgtcgct tttagaggaa acgaagcaag agaactattc	600
aaaccatccc ttataccaga cgaaagtggc gtctctgaat cctacagatt ttctcaatca	660
atcgaaggat gtcaatttgc tgcagtcagg tctaatcagg aatttgaacc agaatggta	720
gaactgttag gtgaatttgc ccaaaaggct gttataccat tggcatgtt tccacccct	780
ccacctcaag acgttgccgg tcatgaggaa acttaagat ggctggatag acaggaacct	840
aattcagttt tctacgctgc tttgggtct gaagtttgc taacagccga gcaattacag	900
agaatttgc tgggtctggc agcttagtggaa ttaccattca tctggccctt tagagcacct	960
cctgacgctg gagatgggaa tggctgcca ggtgggttca aggagagat aaacggtaga	1020

gggttagttt gtagaggttg gttcctcaa gtgaagtcc tagcacatgc ttctgttagga	1080
ggattcctaa cacatgcagg ttggaactct attgccgaag gtctagctaa tggagtgaga	1140
cttgcattc tgccacttat gttgaacaa ggattgaacg caagacaact ggcagagaaa	1200
aaagtcgtcg tcgaggtcgc tagagatgaa gatgtggct ctttgcgc aaatgtatt	1260
gtggatgcct tgagaagggt tatggtgggt gaagagggcg atgagtcgg tgttaaagta	1320
aaagagttgg ccaaagtgtt tggcgatgt gaagttaatg acagatatgt tcgtgatttc	1380
ctaaagtgtc tgtccgagta caaaatgca aggcaaggaa aa	1422
<210> 39	
<211> 1524	
<212> DNA	
<213> <i>Saccharomyces cerevisiae</i>	
<400> 39	
atggctgcaa ctctgattc aacacctgct gctgccgtc cagcagtcgc ctctagtagt	60
tcatcccctc tacacattgt agtatttcct tggatggcat ttggacacat gattccattt	120
ctggagttt ctaaaagact ggctagtcgt ggccatgccg ttacttcgt tactaccct	180
agaaatgcag cttagactgg cgctacacct ccageccccac tgtcatcatc ttccagacta	240
agagtagtcc cattagactt accagccgtt gacggcttgc ctgaaggcgc tgaatcaaca	300
gcagacgtgc caccagaaaa ggtgggtcta ttgaaaaaag cctttatgg tttggcagca	360
ccattcgcca gattcgtac cgaaggctgt gctgcaggtg atgggaagc ttttaccgca	420
gctgccggct ttcttaaggaa gccagattgg atcttccag atttgctca ctctggata	480
tggccaatag ctgaagagca caagattcca tatgcaactt ttctaatcgt cccagctgcc	540
tttagtagcca tcttggacc acgttagagaa aatcttacac accctagaac tactgctgag	600
gactatatgg ttcaaccacc ttggattctt ttccctcta acattgcata caggagacgt	660
catgaagccg aatggatggt cgctgtttt agagctaatg cctccgggt gtcatatgt	720
gataggtttt gggatcaga acaacatctt aattgtatgt tgcataata cagaacttgt	780
ccagaaattt aaccaagatt gttccattt ctgacagagt tatacacaaa gccagctatc	840
ccatctggct ttttagttcc tccagcattt gacgataatg atataggcgt ctacaatcgt	900
tctgacagat cattcggttc cgtcatgca tggctggaca aacagccaaa caaatccgt	960
atctacgtta gtcttggta tgaagcacctt attacagccg atcatatgca tgaacttagct	1020
tttgattt aattttttttt tttttttttt ttggatggatcc ctttggcat tgaggagacc atctggat	1080
aactgtcatg atgatatgtt attgccttcc gggttcgaga caagagtgc agcaagaggt	1140

ttggttgt a cagaatgggt tccacaggtg agaatgttg ctcatgggc agttgggtt	1200
ttcctaacc actcggttg gggttctaca gtcgaatctt ttcatatgg tcagccatt	1260
gtcatgttc ctcatcgc tgatcaagga ctgatagcac aagcagtggc cgcaaccggg	1320
gtgggttag aagtgcgt aactatgtat gatggaaatt tctacagaga tgacgtggct	1380
gctgttattc aaagagtcat gggttggaa gagggaaaag agttagctca taaagctatc	1440
gaactttcg gtatcttagg agacagagta caacaagaga tgtacttata cgaacttatac	1500
ggatacttgc aatgctacaa gtaa	1524
<210> 40	
<211> 1296	
<212> DNA	
<213> <i>Saccharomyces cerevisiae</i>	
<400> 40	
atggatgcct caccattgca tgtggtagtg ttcccttgg tagttttgg acatcttcta	60
ccagcttgg agttagccgg tagattggct tctagaggc taagagtctc tttgtctct	120
actccaagga acattgttagt actgagaagg ccatgcccattt ctgttgaatt tgtggattt	180
cctctgccta gagttgatgg ttacctgtat ggcgtgaag ccacaacaga cgttccagat	240
catatgtcca gtgccttatg gaaagcatca gacggtctaa ctgcaccatt ttccgcattc	300
cttgatgctg ccgcgtgcagc cgaaaaacaaa gttgatttgg ttagtcttggc cgcatgttg	360
tcttggcag cagctgtgc agcagataga aaggtgccat gcttctaat gatgccttat	420
actgctactg catgtgtca ctggagtc cctgtatgaag caagagatgc agacagattt	480
ccatcagcca tagccagaag atttgtaatg gcttccagaa gttcagaatt acttgcgtt	540
agatcatgtg tcgagtttga gccagaatca gtgccttctc tatctaacat ctgcgttcaa	600
cctgttgttc caatcggtt gttaccttca ccacaggttag atggcgacgg ttagggat	660
actgcttga tgccttttg gctggaccgt caaccaccaa aatctgttagt ctacgtcgcc	720
ttggggtcag aagcaccatt aactgcccggaa caaaggagag aatttttttttggatgg	780
ttatctggcg ctccatttctt atggcttta aggaagccac acgggtggcga tgacgtggc	840
ggcttattgc cacctgggtt tgaagagcgtt accagaggtt gaggtatggta taagacagaa	900
tgggttccac aacttaatgtt ttggcccat gcagctgttagt ggcctttctt tacacattgc	960
ggacacttctt ctgtcatcgaa aggtctgaga ttgggttacc ctcttagttt gttaccttgc	1020
tttctggatc agttcaccaa tgcttcatac cttgaaggcg ctgcgtgggtt aggtgtacaa	1080
gtggcttagag acggtaaca tggaggtgtt tttgacagag atggagtggc tgccgcgtt	1140

agagcagcag ttgtcgatga ggaatccaaa aaagcttgg ctgctaattgc cggttaagatg	1200
ggagaggttg ttgtgcacac cgaatgtcat gaaagatgtt ttgatgcctt catacaacaa	1260
ttgagatcct acacaactac acgtacagga tattaa	1296
<210> 41	
<211> 1428	
<212> DNA	
<213> <i>Saccharomyces cerevisiae</i>	
<400> 41	
atggcaaagg aacatatcgt gctatttcca ttcatgtcac aaggccacat aatcccatc	60
ctaagtctag ccaaacttat ctcagagaga catccaacat ataccattac tctttigaat	120
acaccattga acattctgaa tctacaatct acattgcctc caaatagtaa cattcactta	180
aagagttac cttagatc ctctgatttt gggttgccac ctgatgtga aaacacagat	240
agtttaccat tcccattagt tttgagttt tatcaatctg gtgaatctt agccacacat	300
ttcaactact ttgttctga tttgaccaga caaaatcatg atacaccacc attgcttac	360
gtagccgatg tctttttgg ctggacagct gaaatgcaaa agaggctgaa taccatgt	420
tcatttcta cttgtggtgc ttacggact gcagcttact tttcagttt gttacatcta	480
ccacatgctg aaactgactt acctgactt acagcaccag gttcccaaaa aacattcaaa	540
ctgcaacgta atcaattgtc aacttacttg aaaaaagctg atgggtctga caggtggcc	600
aaattcttcc aaagacatg atcactatct ttgaccagtg acgccatgtat ctgcaataca	660
gttggggaaa ttggggctgtc aggtttggaa ctgttaagaa aaaacactgg attgagat	720
tggcttattt gaccattact tccatcaact cctcctaact catccttggg tagatcaggc	780
agaaaaatcag gaatggaaatg atcttacattt atgaaatggt tggattcaca tcctccagg	840
tccgtcgctc acgtatcctt tgggtctata cacgacactg ccgtcaaat gacctctgt	900
gcagtaggtt tggctgtcga attggccact agatcctgac gacattcagg tcgttagattt	960
ggaggtata gaaatagaaa ttcaaatctt aatgggtgtcc cagatgagtt tgaggctagg	1020
atgagagggt ctggcagagg aatcctaattt cacggctggg cacctcaattt agaaattcta	1080
gaacatgaaa gtacaggtgc atttgggtt cactgtgggtt ggaacagtttac tttagaatct	1140
ttgtcttagat gtgtgtgtat gatcgggtgg cctcttgctg cagaacagttt ttacaactcc	1200
aaaatggttt gagaagattt ggaatgggtt ggtacttgc agggatctgg aggtggcg	1260
agatccgaag aggtggaaag acttgtgaga ttagttacag aagatgaaaa gggttcagac	1320
gaagagaacg aacaatacga tggaaatgtt ggtggctatg aggaaaaggcggtgaggc	1380

tccctgtctg gccaacttat caagtttata ggaatggaat ctcatgtaa	1428
<210> 42	
<211> 1539	
<212> DNA	
<213> <i>Saccharomyces cerevisiae</i>	
<400> 42	
atggaagcta ctagagaagc tgcccaatg gctgcattcg catcatctag tagtccattg	60
catatagtta tggccatg gctggcttt ggccatatga tcccatctt tgaacttgc	120
aagagactgg ccagaagagg gttggcgttg accttcgtt ctactcctag aaatgccgt	180
agattaggag ctattccacc agccctgtct gcacacttga gagttgtccc acttgacctt	240
ccagccgtcg atggcttacc tgaaggagct gaatctacag cagatgtcc accagaaaag	300
gttggtttac taaagaaagc ctttgcgtt ttgcgttgc ctttcgtgg ttttgcgttgc	360
gaggcgttgt cagctggca tggtaatca acaccaactg ctgcgtggatt ttccagaaag	420
cctgatttgc ttatatttgc ttgcacaa aattgggtct ggcttattgc cgaagagcat	480
aaaatccctt gtgcattttt ctctatctt ccagctgtt tggtagcatt tggtaggacca	540
aggcaagaga accttgacca tcctagaact aaaacagaac atttatggt gcaacctcct	600
tggataccat ttccatctaa tggatcacac agaaggcgac acggcgctt gttggattgc	660
gcagtttca ggcctaacgc ttctgggttc agtgcgttccg acagattttggag	720
cacgcttgct gtagacttat tatccatgt tcatgcctg aagcagaacc aagatgttt	780
ccattgttgc cggaaattttt cgccaaacca tccgttccgtt ccgggttgc aatggccacct	840
ccaccccttccag ctgcagggtt agatgacgac gatgtacgtt tatctatggat tggatcaacac	900
atcgctatgg caatgcgtt gctggatgaa caacctgaga ggtctgtat atatgtggca	960
ttgggttcag aagctccact gacagtaggtt cacgtaaag aattggctt tggcttggag	1020
ttggcagggtt ttagatttgc atggccactt agagccccctc catccgcctc atctgtcaac	1080
agagacaat gtgcagccga cggcgatcta ttattgccag atggctttag atcaagatgtt	1140
gctgctgcca gaggagggtct ggtctgcgcgca agatgggttc cacagttacg tattttggcc	1200
catagagctt ctgggtggatt cctaacacat tgggttgcgtt cttccatctt tgaatcaactg	1260
agatttgcattt taccattgtt gatgttgccat ttttgcgtt accaaggctt aggggttgc	1320
gcttaccatg caagagaaat cggagtcgaa gtggccgtt atgtacgg ctcttttgc	1380
agagatgcca tcgcagctgc cgttagacaa gttatgggtt aagagaaagg taaagcactt	1440
agttagaaagg cagaagatgtt gagggatgtt ctgggtgcgcg aggtagaca gaaatgtac	1500

ctagatgaat tagtcggcta tctacaaaga tacaataa	1539
<210> 43	
<211> 1524	
<212> DNA	
<213> <i>Saccharomyces cerevisiae</i>	
<400> 43	
atggctgcta cctcagattc taccccgact gctgccgcag ccgccgcagc atcttcctca	60
tcttagtcctc ttcatatagt ggtattccca tggcttgctt ttggtcataat gattccattt	120
ctggaaattgt ctAAAAGACT ggcttctaga ggacatgcgg ttacttcgt cacaactcct	180
agaaatgcag caagattagg agctacacct cctgccccat tgtctagttc atctagatta	240
agagtggttc cattagattt gcctgcagtt gatggattac ctgaagggtgc cgagtccaca	300
gctgatgttc cacctgaaaa ggtcgactg ttgaagaagg ctttgatgg cttggccgt	360
ccatttgcta gattcggtgc agaagcatgt gcagccggcg atgggaagc tgtaactgca	420
gcagcaggct tcctgagaaa acctgattgg attatccctg atttcgcaca ctcctggatc	480
tggccaatcg ccgaggaaca caaaattcca tacgctacat ttctaatgt tccagctgct	540
ctagtggcca tttgggtcc acgttagagaa aatttaactc atccacgtac tacagccaa	600
gattacatgg tcaaccacc atggatacca tttccatcca atatagcata taggagaagg	660
cacgaagccg aatggatggt agctgcattc cgtgctaatg ctagtggcgt gtccgacatg	720
gatagattt gggaaagtga gcaacatcct aactgtagac taataattta cagaacctgc	780
ccagaaatcg aaccttagact attccactt ttgactgaat tgtacacaaa gccagcaata	840
ccatcagggc tttgggtcc accagctta gacgataatg atattgggtt ttacaacaga	900
tcagacagat catttgtc tggtatgcaa tggtggaca aacaaccaaa taagtctgtc	960
atctatgtt ctgggtac tgaggccca atcacagccg accatatgca cgagtttagca	1020
ttcggttag aattagcagg tgtcagattt ctgtggctc ttctgtagacc ttcaggcatc	1080
aactgtcatg acgacatgtt gttaccttct ggttcgaga ctagatgtc agctagaggc	1140
ttggttgtta cagaatgggt tccacaggtt agaatgctag ctcatggtgc cgtaggtgt	1200
tttctaacaac actgcgggtg gggttcaacc gtagaatctt ttcattacgg tcaacctcta	1260
gtcatgttgc cttagtgc ttagtgcggta ctaattgcac aagccgttgc cgcaacagga	1320
gttagtgtcg aagttgctag aaactacgac gatggctcat ttataggga cgtatgtcgcc	1380
gctgctatcc aaagagttat ggtcgaagag gagggaaaag agcttgctca taaagctatc	1440
gaattatgtg ggtatcttagg agatagagta caacaggaaa tgtacctta cgaatttgatt	1500

ggttaccttc aatgctataa ataa	1524
<210> 44	
<211> 1452	
<212> DNA	
<213> Saccharomyces cerevisiae	
<400> 44	
atgggtgatg gtggaggagg cggttggac gttgtcgat tccatgggt agcattcgga	60
cacatgatac catabtctgga actgtctaaa agactggctg cccgtggcca ttagttaaca	120
ttagtctcca ctccatagaaa tttgtcaaga ttacccatg tcccagccgg attatcagcc	180
agacttagat tcgtgtctt accaatgcct cctgttgatg gtctaccaga aggtgccaa	240
agtactgcag atgtgcctcc aggttaacgtt gaactgatca aaaaaggctg cgacggattt	300
gcagctccat tcgctgcctt catggccat ttagtagctg ctggaggtt aaaaaccat	360
tggataatca ttgatttcgc ttaccatgg ttgccaccaa ttgcagccga acataacgct	420
gctgcaatcg ctittctagg cccttagatgg gctaattgtc ctcaccctag ggctccctta	480
gactttaccc caccaccccg ttgggttcca ccaccatctg ctatggccta tcgttagaaat	540
gaggccagat ggggtggcg tgcctttaga ccaaaccat caggtgtttc tgacatttag	600
agaatgtgga gaaccatcgat atctttaga tttactatct acagatcctg tgacgaagta	660
gaaccagggtg tttttggactt acttattgtt ttgttccgtt gaccagctgtt accaggcagg	720
atcttattgtt cacccctcc agatttggcc gccgcagatg atgacgacgt tgacgggttgc	780
tcatcagccg atagagctga aactttgcgtt tggttggacg aacaaccaac caagagtgtt	840
atttacgtgg cacttgggtt cgaaggcacct gttacagcaa agaacttgca agagtttgt	900
ttgggattttt agcttagctgg cgttagattt ctatggcct tgagaaagcc tgctgcagga	960
actttgtcac acgcctctgc agctgtatgtt gatgaactac ttccagatgg ttttgaagag	1020
agaacaagag gaagagggtgt tttttggaca ggctgggtcc cacaggttggat agttttggca	1080
catgctgcgg ttgggtcatt ctttactcat tgcggctggg gtttacaat agaatctctg	1140
gtattttggc atcccttagt gatgcttcca ttctgttggat accaaggctt tggttgcaga	1200
gcaatggcag aaagagggtt tggcgtggag gtttgcgttggat aggacgttga tgaaggagtt	1260
tttgggagac atgatgttgc agctggcggtt aggaggtca tgggttggat tggaaaggaaa	1320
gtgtttgggtt aaaaatgttgcgtt gaaaatgttgcgtt gaaatgttgcgtt gaaatgttgcgtt	1380
caataactttt atgatgttgcgtt gaaatgttgcgtt gaaatgttgcgtt gaaatgttgcgtt	1440
aagtactgtt aa	1452

<210> 45

<211> 1296

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 45

atggacgcat ctccacttca tgtggttgtatcccttggc ttgcgg tcaacttgcta 60  
ccagcccttg aattggctgg tagattggca tccagagggt tgagagtgtc ttttgtatct 120

accccaagga acatcgccag attgagaagg ccatgtccat ctgtcgaatt tgtcgaatta	180
cctttaccaa gagttgtatgg gttaccagat ggtgccgaag caacaactga tggccagat	240
cacatgtcaa gtgcttatg gaaggcctca gatggactga cagcaccatt ctctgcattt	300
ctggacgctg cagctgccgc aggaaataaa gttgattggc taatcttggta tggatgttg	360
tcatgggctg ctgcttcagc cgctgaccgt aaagtccat gcgtgttgc gatgccttac	420
actgcaacag ctgtgcaca tttggcggtt cctgatgaag caagagatgc tgacagattt	480
ccatctgcac ttgtctgttag attcgttct gcttcagaa gtagtgcattt gtttagcttt	540

agatcatgtg tagaatttga acctgaaagt gtccattat tatctaata tttcgcaaa	600
ccagtagtac caatcggtt actacccca cctcagggtt atggagatgg cgatggagac	660
acagcactaa tgtcatcaig gtttagacaga caaccaccta agtctgttgt ctatgtgcc	720
cttggatcag aagccccctt gactgccgag caaaggagag aattagcact tggattggaa	780
ttatctgggg ctccatttct atggcattt agaaaacctc acggcggtga tgacgatgg	840
ggtctgctgc ctccaggctt tgagggaaagg actcgtggga gaggtatgg aaagactgaa	900
tgggtacctc aactgaaat tttggctcat gctgcagggtt gtgcatttct aacacattgc	960

ggcattcct ccgtcataga gggattgaga ttccggcacc cattagttat gcttccattg	1020
ttcctagatc aatttacaaa tgcttcata tttgaaggag ccagaggtgt gggtgtccaa	1080
gtggctagag atggtaaca tggtgagct tttgacagag atggcgtagc cgctgcgtg	1140
agagctgccg tcgttgacga ggaatctaag aaagcattgg ccgcaaacgc tggtaaaatg	1200
ggtaggtcg ttgcagacac agagtccat gagatgtta tagatgcatt tatccaacag	1260
cttagatcct acaccacaac tagaaccggc tactaa	1296

<210> 46

<211> 1419

<212> DNA

<213> *Saccharomyces cerevisiae*

&lt;400&gt; 46

atggacatg cacacttcc tcaatctct ctgcattcg tgatattcc atggttggca	60
ttcggacacc tattgccttg cttggatttg gctgaaagat tggctgcaag aggccaccgt	120
gtcagtttg tatcaacacc tagaaacttg gcttagattac caccaggtag accagaaita	180
gctgaattgg ttgatctgt agcattacca cttcctagag tagatggttt gccagacgga	240
gcagaggcta catcagacgt gccatttgac aagtccgagt tacacagaaa ggctttgtat	300
gggctagccg ccccatttc agcttcctt gacacagcat gtgcaggcgg taaaaggcca	360
gattgggttt tggccgacct tatgcaccat tggtagctt tagcctcaca agagagagga	420

gtgccatgct ctagatact tccatgttgc gcccgcgttg tggcaagttc agcaccccca	480
accgaatctt ctgcccacca aaggaaagcc attgttagat caatggaaac tgccgttcca	540
tcatcgaaat ctaagagagc tactgaagag tttgcacag aaggcgcttc tggagttct	600
atcatgacta gatacgttct tactttacaa agatccaaat tggttgcaat gaggagttgt	660
cctgagttgg aaccaggagc tttacaatc ttaacttagat tttacggaa gcctgttgc	720
ccttcggcc tactacctcc acgtccagat ggtgccagag gagtctccaa aaatggaaag	780
catgatgcca tcatgcaatg gttggatgttca acccagccaa aatctgttagt ttatgttgc	840

ctaggttccg aagccccaaat gtctgcagat ttattgagag aattagccca tggattggat	900
ttagctggta caagattcct ttgggctatg agaaaacctg ctgggtttaga cgctgattct	960
gttctgcctg ctgggtttctt gggcagaact ggtgaacgtg ggcttagtaac taccagatgg	1020
gctccacagg ttccattctt agcacatgca gctgtttgtg cttccatttac acattgtggc	1080
tggggctctg tcgttgaagg gtcgcgttc ggtcatccat tgataatgtt acctatTTA	1140
ggtgatcaag gtccaaatgc tagaatcttta gaaggcagaa aattgggtgt cgccgttcc	1200
agaaatgtt aagatggtag ttgcacaga ggtggatgtt caggcgcgtt gagagccgtt	1260

gtcgtggagg aagagggtaa gacctttttt gctaacgcca gaaaacttca ggaaatcgta	1320
gctgatagag aacgtgagga aagatgcattt gatgaatttg tgcaacattt gacatctgg	1380
aatgaactga aaaataactc agatggtcaa taccctaa	1419

&lt;210&gt; 47

&lt;211&gt; 1389

&lt;212&gt; DNA

<213> *Saccharomyces cerevisiae*

&lt;400&gt; 47

atggatagtg gctactcctc atcttatgtt gctgcccgtg gtatgcacgt tgtgtatctgc	60
ccttggttgg cctttgtca cctgttacca tgtctggatt tagccaaag actggcctca	120

agaggccata gagtatcatt tgtgtctact cctagaata tctctcgttt accaccagtc	180
agacctgctc tagtcctct agttgcattc gttgctctc cactccaag agtagaagga	240
ttgccagacg gcgcgaatc tactaatgac gtaccacatg atagacctga catggcga	300
ttgcatagaa gagccttga tggattggca gctccattt ctgagttcct gggcacagca	360
tgtcagact gggttatagt cgatgttattt catcaactgg ctgctgcage cgccatggaa	420
cataaggcgc ctttgctat gatgttgtt gggtcagcac acatgatcgc atccatagct	480
gatagaagat tggaaagagc taaaacagaa tccccagccg cagcaggaca aggttaggcca	540
gctgccccc caaccttga agtggctaga atgaaattga ttctactaa agtagttca	600
gggatgagtc ttgtgaaag gtttctctg acattatcta gatcatcatt agtttaggt	660
agatcctgcg tcgagttcga acctgaaaca gtaccttac tatctacttt gagaggcaaa	720
cctattactt tccttggctt aatgcctcca ttacatgaag gaaggagaga agatggtaa	780
gatgctactg ttaggtggtt agatgcccac cctgctaagt ctgttggta cggtgcattt	840
ggttctgagg taccactagg ggtggaaaag gtgcataat tagcattagg acttgagctg	900
gcccggaaaca gattccttgc ggcttgaga aaaccaaccg gtgttctga cgccgacttgc	960
ctaccagctg gttcgaaga gagaacaaga ggccgtggc tcgttgcata tagatggtc	1020
ccacaaatga gtattcttagc tcatgcagct gtagggccct ttctaacccca ttgcggttgg	1080
aactcaacaa tagaaggact gatgttggt catccactta ttatgttacc aatcttggc	1140
gatcagggac ctaacgcaag attgatttag gcaaagaacg caggtctgca gttgcacgt	1200
aatgatggtg atggccctt tgatagagaa ggcgttgcag ctgcacatcg agcagtcgc	1260
gttgaggaaag agtcatctaa agtttccaa gctaaggcca aaaaattaca agagatttg	1320
gctgacatgg ctgtcacga aagataacatc gatggttca tccaaacatt gagaagttat	1380
aaagactaa	1389

<210> 48  
<211> 1338  
<212> DNA  
<213> *Saccharomyces cerevisiae*  
<400> 48

atggacactg gtagatcaaa atttcatcc cgttgcatt ttgtgatctt tcctggctg	60
gccttggc attgttgcc atatctggaa ctagctgaaa gagtggctc aagaggacat	120
agagtatctt ttgtatcaac ccctagaaac ctagccagac ttccagatgg cacagaatgt	180
acaaatgacg tcccaagtgg taaattgaa ttattgtgga aggctttga tgccttagct	240

gcaccatttgcgagttcctt aggtgcagca tgtgatgctt ccggcgaaagg accagactgg	300
attatcgctg atacatttca tcattggca ccactagttt ctttgcaca taagggtgcctt	360
tgtctatgt ttttccttc tgcttctatgt atggcaggctt gggctactatgt atcatctgaa	420
ccagccgggtt ctcttatctt tcaagttta ggagtctggg cttagaaagc aagtttcttc	480
gtcgatcatg gtgcctccgg aatgtctgtt gcaaagaggtt gtccttagc aatggaaaga	540
tgtactctatgt ctgccatgatgatc atcctgcctt gagggtggaaac ctgtatgcatt tcaacaagtc	600
gcagctggcc tgaagaacaa accattaata ccattggac ttgttcctcc atcccctgat	660
ggcgggagga gaagagcagg tatgacagat aattcaacta tgcgttggtt ggatgttcag	720
ccagctaagt ctgttgtata cgtcgctctg gtticagaag ttccttgcc tttagaacaa	780
gtgcatgatgt tagctttgg tttggattttt gctggtacaa ggttcctatg gcccctttaga	840
aaaccacacgtt gtgttagattt atctgatgtt ttgccacctt gttaccaaga gagaacaaaa	900
tctcacgggc acgttgcattt ggggtgggtt cctcaaatca ctattctatgc ccatgcagct	960
gtcggagctt tccttacaca ttgcggcaga aatagtttag tagaaggattt attatttggc	1020
aatccactga tcatgcttcc aatattcggtt gaccaaggcgc caaacgttagt attgtatggag	1080
ggttaataaag ttggctcaca agtgaggcgtt gacgatatgg acggatctt cgtatagacac	1140
gggtttgcag ctgcccgtttagt gtagaagaggaaactagaag agtttcgtt	1200
gcaaatgcata tttagactaca ggaattgggtt gccgacaagg aattgcacga gagatacattt	1260
gatgaatttttta ttcaacaactt ggtttctcac ggccgacacgtt gtagttgtttaa caccggccca	1320
ccagtgccat caagttaa	1338
<210> 49	
<211> 1056	
<212> DNA	
<213> <i>Saccharomyces cerevisiae</i>	
<400> 49	
atgatgcattt tggggcgtgc tcactccctt aaaggtttttta gtattacagt tgcacaaaca	60
aaatttaattt acctgaaccc atccaaggac cttagcaggattt tccaattttt aacaatccca	120
gaatctttac cagcctcaga ttggaaaactt ctggccctt tctggttttt catcaagtta	180
aacaaagagt gtgaaatctc attttgggtt aatttcctact tcaacagcaa	240
gagggaaattt ctgtgtttat ctatgtatgtt ttcatgttattt ttgcagaagc agccgccaaa	300
gagtttaactt taccaaaatgtt tttttttttt actgaaaatgtt ctaccgtttt tgcttgtaga	360
agtgtatgtt gtaagttata cgctaaagac ggaatcgctc cacttacaga aggctgcgggtt	420

agggaagagg agctagttcc agaattgcat cctctgagat acaaagattt gccaacatct	480
gcatttgcgc ctgttgaagc ctctgtggag gtgttcaaat ctgttgcgaa aaagggcaca	540
gcacatctcta tgattatcaa cactgtctca tgcttagaaa tatcttcatt ggaatggttg	600
caacaggaat tgaagattcc tatctaccct atcggccac tgtacatggt aagttctgcc	660
ccaccaactt cactattaga tgaaaatgaa tcctgcattg actggctgaa taaacaaaag	720
ccatcatcg tcatacatat ttctcttggt tcattcacct tattggaaac taaggaagta	780
cttggaaatgg ctccggatt ggtatcaagt aatcaatatt tccttggc tattagacct	840
gggtctatct tgggttctga actatctaatt gaggaattgt tttctatgtat ggagattcct	900
gatagaggct acattgttaa atgggctaca cagaaacaag tattagcaca tgccgcagtc	960
ggagctttct ggtcacattt tggatggaac tccactctgg aatccatcgg cgaaggtata	1020
ccaattgtgg gtttgttgcgtttaattaaa agataaa	1056
<210> 50	
<211> 1317	
<212> DNA	
<213> <i>Saccharomyces cerevisiae</i>	
<400> 50	
atggaaaaaa aaatggaagc taaacgtaga atcggtctag taccaattcc tgcacaagg	60
catgtgactc cattgtatgc attggtaaa gttctaaact caaaggattt ttcaatcaca	120
gtcgttagaag gtcatttcaa ccagggtgtca tcatttttc aacatttccc aggatttcaa	180
tttgcataaca tcaaagagtc ttacatgaa tcagaatttgg aaaagctggg cgccatcgaa	240
tctatgtatca cacttaataa gacttctgag gcatctttt aagattgtat atcacaacta	300
cattttgcac aaggtaacgatatacgttattttacg atgagtgatgtatgttctgt	360
ggccgcgtt ctaaagatgtt tagtattccaa agtgtaatat tctctaccca atctgtgt	420
aattacgttt cccatccaga tatgcaagac aaagttgtgg aaaattgtatcc ttctttaaga	480
tacaaagatt taccaaccc tcggcatgggt ccactagaca gattttgtatgcaga	540
gaagtcgcaa ataagagaac tgcttagtgcc gtcataatca atacagttttagtgcataatca	600
tcatcttcct tgcataatca ggaacaaaag gtcggatatt ctgttatcc ttctggacat	660
ttacacatga cagactcatc accatctca cttagagg aagatagatc ctgttatcgaa	720
tggctgaata aacagaaacc aaaatcgttcc atttatatta gtataggtac tcttggacag	780
atggaaacca aggaagttt agaaatgtcc tgggtctgt gtaatagtaa ccaacctttt	840
ctatgggtta ttagagccgg ttctatTTTtta ggcactaatg ggatcgaaatc tttggccagaa	900

gatgttaata aaatggtgc tgagagagg tacattgtaa agagggcacc tcaaattgaa	960
gtattggcc acccagctgt tggcgattc tggcacatt gcggttggaa ctctatcctg	1020
gaatcaatcg gtgagggagt gccaatgatc tgtaagcctt ttacacgtga acagaaaactg	1080
aacccatgt acatcgagtc agtttgaga atcggttcc aagtagaagg aaaggtcac	1140
cgtgggaag tagagagac tgtcaaaaga ttgatagtcg atgacgaagg tgcaggtatg	1200
agagaaaagg ctttgttct aaaagagaaa atcaaggcat ctgttctag tggcggtgca	1260
tcctacaatg ctttggaga tattgtgaac tacctaaaga caaagcaaag atgctaa	1317
<210> 51	
<211> 1415	
<212> DNA	
<213> <i>Saccharomyces cerevisiae</i>	
<400> 51	
atgagatgaa acattctaat gatgctttc acgttgtcat gttccattt ttcgccttg	60
gtcatatttc tccatttcta caacttgcaa ataaattgtc atcatacgtt gtaaaagtct	120
ccttcttac agcatcaggc aatgcatactg gagtgaatc tatgctaaat agtgcaccta	180
caacacatat tgtaccactg actttgcctc atgtagaagg tttgccacca ggagctgaga	240
gtactgctga attaacacca gcatccgccc agttattgaa ggtcgctt gaccttatgc	300
aacctcaaat aaagactctg ttatctact taaaacctca tttcgctcta tttgattcg	360
ctcaggaaatg gttacctaaa atggccaacg ggttggat caaaactgtc tactactcag	420
ttgtttagc attatctact gccttcctaa cttgtccagc aagagtttg gagcctaaaa	480
agtacccttc attggaagat atgaaaaaac caccttggg ctttcctcaa acatcgttta	540
catctgtgag aaccttggaa gctagagatt ttctatacgtt attcaaatcc ttccacaacg	600
gacctaccct ttacgacaga atccaatctg gattgagagg atgttgtca atcttagcca	660
aaacatgctc acaaataggaa ggtccatata tc当地atgtt ggaagctcag ttcaacaaac	720
cagtgttct gataggacca gttgttcctg acccaccatc cgtaagttt gaggaaaagt	780
gggctacatg gttaaataaa ttgaaggcg gtactgtat atattgtct tttggctctg	840
aaactttctt taccgacgac caggtgaagg aattggcaact aggcttagaa cagacaggtc	900
tgccattttt tctatgttta aattttccag ctaacgtgga tgcgttgcc gaattgaaca	960
gagcttacc agaaggttt ctggagagag tgaaagataa ggttattatc cactccgggt	1020
gggttcaaca acaaataa ttggctcatt catcagtgg gtgctatgtc tgtcatgccc	1080
gcttctttc tgtcattgaa gcattggta atgactgtca agtagtcatg ctgccacaaa	1140

agggagatca aatttgaat gctaaattgg tttccggta tatggaagct ggtgtgaaa	1200
ttaacagacg tgacgaagat ggctactttg gtaaggaga tattaaggag gcagtgcaga	1260
aagtaatggt tcatgttcaa aaggagccag ggaaacttat cagggaaaat caaaaaaagt	1320
gaaagagtt tcactaaac aaggataatcc aatctaagta ttcgttaac ttatgtaaatg	1380
aatgaccgc catggctaa gttcaacta cataa	1415

&lt;210&gt; 52

&lt;211&gt; 1329

&lt;212&gt; DNA

<213> *Saccharomyces cerevisiae*

&lt;400&gt; 52

atggAACCTA cattcacgc tttatgttt ccatggtttgc cttcgacata tatacaccat	60
ttccttcatt tagcaaacaa gttggcagaa aaggggcatc aaatcacatt tctgttgcc	120
aaaaaagcac aaaagcaatt agaacaccat aatTTTCC cagattctat agtatttcat	180
ccttgacaa tccctcacgt gaatgggttgc ccagcagggtc ccgaaactac tagtgatatc	240
tcaatctcta tggataatct attgtcagaa gctcttgact tgaccagaga tcaagtggaa	300
gctgccgtta gagccttgcg tcctgatttgc atatttttgc acattgcaca ctggattcca	360

gagatcgcaa aggaacacat gatcaagtcc gtatcataca tgattgttgc cgccactaca	420
attgttata ctttcgtcc tggaggagtt ttgggtgtac caccacccgg ctacccttca	480
agtaaggttt tgtatagaga aatgtatgtt catgtttgg caacacttcc tatattttac	540
aaaagactat accaccagat aacaacttggaa ttcaaatctt ggcacatcat tgctctaagg	600
acttgcacg aaatcgaagg caagttttgt gactacatct cttctcaata ccataaaaaaa	660
gtattactga ccggccttat gttaccagaa caagatactt caaaaccactt agaagagcaa	720
ttgtctcatt tcctgtcaag attccacca aggtctgtgg tgttttgtgc tctaggctcc	780

cagattgtgc tggaaaagga tcaatttcaa gagttatgtt taggtatggaa acttacaggt	840
ttggccatttc tgattgcagt caaaccacctt agagggagtt ccactgttgc ggaaggactt	900
ccagaagggtt tccaagagag agtaaaaaggat agaggcgtcg tctggggagg gtgggttcaa	960
cagccattaa tttggatca tccatcaatc ggctgttttgc ttaatcatttgc tggtccaggc	1020
accatttggg agtgcttaat gactgactgtt caaatgggtt tactacctt cttgggtgat	1080
caagtactat tcactagatt aatgacagag gaattaaag ttcgttgc agtctcaaga	1140
gaaaagactg gctggttctc taaagagtca ctgtctgacg ctatcaaattc tgttatggat	1200

aaagatagtg acctggaaa attggtaga tccaaaccatg ccaaacttaa ggaaacactt	1260
--	------



&lt;210&gt; 54

&lt;211&gt; 1380

&lt;212&gt; DNA

<213> *Saccharomyces cerevisiae*

&lt;400&gt; 54

atgggtcac aagcaacaac ttatcacatg gcaatgtatc catggtcgg cgtggcacac	60
ttgactggct tcttcagact tgctaacaaa ctggccggga aaggcatag aatttcattt	120
ctaataccta agaacacaca atcaaaattt gaaagttta atttacaccc acatcttata	180
tcatttgctc caatcgtagt accttctata cctgggttgc ctccaggagc tgaacaacc	240
tccgacgttc cttdccatc tactcattt ttaatggaa ccatggacaa gacacaaaat	300
gatatcgaaa ttatcttcaa agatttgaaa gttgatgtag tggatgttttacga cttaaccat	360

tggttaccta gtttagctag aaagattggg ataaaaatctg tattctactc tacaattca	420
ccattgtatgc atggctacgc tctatctcca gagagaagag tagttggaa acaacttaca	480
gaagctgaca tgaatggaaagc accagcttca tttccagacc cttctattaa gctacatgct	540
catgaagcac gtggatttac tgcttaggact gttatgaat ttgggttgtga tataaccttt	600
tttgatcgta tcttcacagc agtttctgaa tcagatggct tggcataactc aacttgtaga	660
gaaatttgaag gtcaattctg tgactacatt gaaacacaat tccaaaagcc agtcttatta	720
gcaggccctg cactgccagt tccatctaaa tccacaatgg agcagaaaatg gtctgattgg	780

ttggcaagt taaaagaggg gtccgtaatc tattgccc tttttctga atgtaccc	840
agaaaagaca aattccaaga gctgttatgg ggatttggaaac tgactggat gccattttc	900
gctgcttga agccctctt tgaagctgaa agtattgaag ccgcattcc agaagagttt	960
aaggaaaaga tccagggggag aggtatcgat cacggtaat gggtgcaaca gcaactgttt	1020
ctacaacatc catcagtcgg ctgtttgtt tctcattgcg gttggccatc ctttagtgag	1080
gcttggta acgattgtca aatcgatctt ttaccacaag tggagatca aatcatcaat	1140
gccagaatta tgagtgtctc ttgaaagta ggcgtcgaag tcgaaaagg tgaagaggac	1200

ggagtgttct caagagaatc agttgtaaa gcagtaaagg ccgttatgga tgagaaaatg	1260
gagataggtt gagaagtcag gggaaaccac gataaattaa gaggtttct gcttaatgct	1320
gatctagatt ccaaatacat ggattccttc aatcaaagc tacaggatct tctaggataa	1380

&lt;210&gt; 55

&lt;211&gt; 1413

&lt;212&gt; DNA

<213> *Saccharomyces cerevisiae*

&lt;400&gt; 55

atggccgccc ctgttgtga agctgatgac gaggctatgc acgttgcttt gttcccttt	60
ctggcattcg gacatatatac accattgcc cagttagcaa gatccttggg cgctgttagga	120
ggtgtcagag ttacttcct ttctgtcgca gccaacgttg cacgtgtaga agcaatgtta	180

cctggcgcacg gtactgctgt tgcgcagcc ttgcattgc ctagagtgc tggtttacca	240
gttaggtgccg aatcaacagc tgaagttgat gcagatggcg ctgaactgtt aaaattagct	300
ttagacggaa caagaccaca ggtcgaggct ctgcttgccc gttgaggcc agacgttg	360
ttgtttgatt ttgcaacacc atggtagct gatgtagcca gacaactagg tgcaagagcc	420
gctcactttt ctgtgtttac tgcagttact tccgcctact tgacagtgc tgcacgtagg	480
agactacatc atggcgctgc atctgtcca actgttgacg acttagcaac agccctgtc	540
ggtttccac cttagtcatac cctagcaact gttccaacct atcaagctgc agactttact	600

tacgttta cctttcca tggatgcca tcagcttacg atagagtgcg cgcctgtat	660
aaagcatcag acgtatttgt gtttaagacc tgccgaaaaa tggaaaggcc ttatattgaa	720
tacgtcgcta cacaatacga taagccaaatc ctagtaactg gacctttgt tccagaacct	780
ccacatgggg aacttgaaga gagatggaa acatggttgt cctttccc agataatgca	840
gtgggttcg caagttttgg ctctgaaacc ttctgccaa cagccgctgc tacagagctg	900
ttacttggct tagaagcaac aggtcaacca ttctgtcgag tacttaattt ccctaggct	960
gtcgatgctg aagccgaagt aaaaaaatgt atggccccag gttttgagga aagagtgaaa	1020

ggtagaggtg ttgtacacag tgggggtt caacagcaac atatcttaag acacaggct	1080
gttggttgct acgttaatca cgctggatt agttctgttgc tgcaggact tggtgtgg	1140
tgttagactatg tattgttacc aatgaagtca gatcaattct tcaatgtgc ttgtgtggcc	1200
agagaattaa gagtcgggac tgaagttgatc agaagagatg gggacgggtt gttggccat	1260
gacgcgttta gagatgtgtt aaacgctgca gttgtatgc ctgggtggcg agatgtat	1320
gagagaaaat ggagagaatt tctaacagat gatgtgttc aacgttagatt cgtcgaggag	1380
tttgttagag agttgagaaa gttgggtttt gta	1413

&lt;210&gt; 56

&lt;211&gt; 1383

&lt;212&gt; DNA

&lt;213&gt; Crocus sativus

&lt;400&gt; 56

atgttgaacg gcaacaaatg ccacatcctc ctcccccgttccagcaca aggccatatc	60
--	----

aaccccatc tccaattcg	caagcgactg	gcctctaca	acctcctcac	cacactcg	tc	120	
aacacccgg	tcc	tccaa	ctcgaccaa	tccgaaccg	gtccggtaa	catccagtgc	180
atatccgac	gattcgatcc	cggtggatg	aatgcagcac	ctagtctcg	ggcctattc	240	
gaccgac	tc	agagtcgg	gggtcagaaa	catgttgcc	tcatcgagtc	actccgtcc	300
agaggccgac	ccggcgc	atg	tttttgtcta	cgacc	ccctctggc	aatgaacgta	360

gccgagcg	gtcggtcg	ca	ttcttactc	agccgtgt	gc	agtggacaca	420
atctacc	gccc	acgtgtgg	gggacgaatt	aagg	ttccgag	ggtccgg	480
cctggatt	ccca	ctcg	gcctagt	gtat	ctggcgtcg	ttcgtaacgg	540
gtcgtaacc	cggat	ctt	gccact	cg	gtgaatc	acgg	600
gacatgat	ggagaact	c	atac	tagggc	atttgtt	gggtccg	660
ttaccatt	cgtgaag	tc	gatggc	accgtgc	ctacttat	ctcgaca	720
atccgtcg	attcac	acta	tggatt	taat	ctatac	acccttc	780

gactggct	actccaagg	cc	cgaa	ct	gtcatctat	gtcatttgg	tagttgt	840
agtctgag	cc	ctgat	caaac	caat	gagatt	gcgtcggt	tgtcgcc	900
tttatttgg	tt	gg	tttgc	ac	ttcg	gagttt	gccaact	960
aatgcgag	cc	gtgg	ctgg	gt	tgac	ctgg	tgac	1020
gcgaccgg	tt	gtt	cgt	gac	tca	ctgc	gcaact	1080
ggtgtgcc	ga	tgg	tggt	gg	tttgc	tttgc	tttgc	1140
gaggatgt	gt	gaa	agg	tcg	cgt	gagg	ggc	1200

gaggagtt	ca	aaagatgt	gt	ggaggagg	tg	atggatgg	agaggagt	gg	aaagat	taga	1260			
gagaatg	ccg	caagg	tggt	taag	ttgg	cc	aaggact	ctg	tttagt	gagg	tggagg	ttct	1320	
gacaagt	gt	taa	aggag	ttt	cata	cat	tgc	aat	ttc	tttgc	aat	tttgc	tttgc	1380
taa														1383

&lt;210&gt; 57

&lt;211&gt; 460

&lt;212&gt; PRT

&lt;213&gt; Crocus sativus

&lt;400&gt; 57

Met Leu Asn Gly Asn Lys Cys His Ile Leu Leu Leu Pro Cys Pro Ala

1 5 10 15

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

20	25	30
His Asn Leu Leu Thr Thr Leu Val Asn Thr Arg Phe Leu Ser Asn Ser		
35	40	45
Thr Lys Ser Glu Pro Gly Pro Val Asn Ile Gln Cys Ile Ser Asp Gly		
50	55	60
Phe Asp Pro Gly Gly Met Asn Ala Ala Pro Ser Arg Arg Ala Tyr Phe		
65	70	75
Asp Arg Pro Gln Ser Arg Ser Gly Gln Lys His Val Gly Leu Ile Glu		
85	90	95
Ser Leu Arg Ser Arg Gly Arg Pro Gly Ala Cys Phe Gly Leu Arg Pro		
100	105	110
Val Pro Leu Trp Ala Met Asn Val Ala Glu Arg Ser Gly Leu Arg Ser		
115	120	125
Val Ala Phe Phe Thr Gln Pro Cys Ala Val Asp Thr Ile Tyr Arg His		
130	135	140
Val Trp Glu Gly Arg Ile Lys Val Pro Val Ala Glu Pro Val Arg Leu		
145	150	155
Pro Gly Leu Pro Pro Leu Glu Pro Ser Asp Leu Pro Cys Val Arg Asn		
165	170	175
Gly Phe Gly Arg Val Val Asn Pro Asp Leu Leu Pro Leu Arg Val Asn		
180	185	190
Gln His Lys Asn Leu Asp Lys Ala Asp Met Met Gly Arg Asn Ser Ile		
195	200	205
Tyr Glu Leu Glu Ala Asp Leu Leu Asp Gly Ser Arg Leu Pro Leu Pro		
210	215	220
Val Lys Ser Ile Gly Pro Thr Val Pro Ser Thr Tyr Leu Asp Asn Arg		
225	230	235
Ile Pro Ser Asp Ser His Tyr Gly Phe Asn Leu Tyr Thr Pro Asp Thr		
245	250	255
Thr Pro Tyr Leu Asp Trp Leu Asp Ser Lys Ala Pro Asn Ser Val Ile		
260	265	270

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

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

Glu Glu Phe Lys Arg Cys Val Glu Glu Val Met Asp Gly Glu Arg Ser  
 405 410 415  
 Gly Lys Ile Arg Glu Asn Ala Ala Arg Trp Cys Lys Leu Ala Lys Asp  
 420 425 430  
 Ser Val Ser Glu Gly Ser Ser Asp Lys Cys Ile Lys Glu Phe Ile  
 435 440 445  
 His Gln Cys Cys Asn Asp Ser Lys Ile Ser Leu Val  
 450 455 460  
 <210> 58

&lt;211&gt;

&gt; 1383

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

<220><223> Synthesized Crocus sativus UGT2 nucleotide sequence codon  
optimized for expression in *Saccharomyces cerevisiae*

&lt;400&gt; 58

atgctaaacg gcaataagtg ccacatccta cttttaccat gcccagctca gggtcacatc 60

aatccaattc tacaattcgg gaagagactt gcatctata acttgtgac aacactgtc	120
aacaccagg ttcgtctaa cagtactaag tcagaaccag gtccagttaa cattcaatgt	180
atctcagatg gttcgaccc aggaggcatg aacgctgcac catctagaag agcttactc	240
gacagaccac agtcaagatc aggtcaaaaag catgtaggc taattgaatc tttaagatcc	300
agaggttagac caggtgcttg tttggatta agaccagtgc cattgtggc aatgaatgtt	360
gctgaaaggt caggtcttag gtctgttagca tttttactc aaccttgc agtcgatacc	420
atatacagac atgtgtggg aggaggatt aagggtccag ttgctgaacc tgtcagatta	480
cctggcttc ctccattaga accaagtgc ttaccttgc tttagaaatgg atttggtaga	540
gtagtcacc cagacctact acctttgcgt gttaatcaac aaaaaatit agataaagct	600
gatatgatgg gcagaaatag tatctacgaa ttggaagcag accttctgga tggtcccg	660
ttgcactgc ctgtgaaatc aattggcca accgtccctt caacttacct ggataatgt	720
atcccttcag actcacacta tgggttaac ctataactc cagatactac cccatatctg	780
gattggctgg actctaaagc acctaattct gtcatactcgt tctcattgg ttcactatcc	840
tctttgtctc ctgtcaaac aatgagatt gcttctggat tgatagctac aaacaatcc	900
tttatctggg tagtcagaac atccgagttt gcaaaagtac ctgccaattt cactcaagag	960
aatgcctcca gagggttggt agtaacttgg tgtgatcaat tggatcttt agcacatgt	1020
gctacaggct gtttggttac acattgtggt tggaactcta ctatgaaagg ggtagccctt	1080
ggtgttccaa tggtgccgt gcctcaatgg tctgaccagc caatgaatgc caaatacgtt	1140
gaagatgtgt ggaagttgg agttagagct aaaacatatg gtaaagactt cgtaggaa	1200
gaagagttca aaagatgcgt tgaagaggat atggatgggg aaagatccgg aaagattaga	1260
gaaaatgccg ccagatggtg taaactggcc aaagattctg taagtgaagg cggcagtgt	1320
gataagtgc taaaggagtt catacatcaa tggcaatg attccaaaat ctcattggta	1380
taa	1383
<210> 59	
<211> 1713	
<212> DNA	
<213> Xanthophyllomyces dendrorhous	
<400> 59	
cccgggaaat tcaagcttaa aatgttcatt ttgggtctgt taactggagc attgggttg	60
gctgcatttt cctggcaag tattgtttc ttcatgttgc acctagcacc tagaagatca	120
tccttataca atttgcagg tccaaatcac accaattact ttacaggcaa ttccctagac	180

attctgtctg ctagaaccgg cgaagagcac gccaaataca gagagaagta tggttcaact	240
cttagatttgc ctggaatcgc tggcgctcca gtattgaact caactgaccc aaaagtcttt	300
aaccatgtga taaaagaggc ctatgattac ccaaaggctg gtatggcagc cagagtctta	360
cgtatagcaa ctggggatgg ttttgtaca gctgaaggag aagctcacaa aagacataga	420
aggatcatga taccatctt atctgccaa gctgttaagt ctatggtacc tatttccta	480
gaaaaggta tgaaattagt agacaagatg atggaagatg cagctgagaa agatatggct	540
gtgggtgaat ctgcaggcga gaaaaaggct accagattgg aaactgaagg ggtggacgtt	600
aaagattggg tcggagagc cacattagat gttatggcct tggctggctt tgattacaaa	660
tctgattctc tacaacaaacta gactaatgaa cttttagttt ctttttcgg cctgactgat	720
ggatttgcctt ctacattaga ttcccttaag gccattatgt gggactttgt accttacttt	780
agaacaatga agagaaggca tgagatccaa ttaactcaag gactagcagt ttctagaagg	840
gttagtattt aattgtatgga gcagaaaaaa caagctttt tagttctgc atctgatcaa	900
gccgtggaca aaaaagatgt tcaagggaga gatatacttt ctttgcttgt tagagccat	960
atcgctgcta acctgccaga atcacagaaa ttgagtgtatg aagaggtact agcacagatt	1020
tcaaattctac ttttgcagg ttacgaaaca tcatccacag tcttaacttg gatgttccat	1080
agactgtccg aagataaagc cgtcaagat aaactacgtg agggaaatctg tcaaattgac	1140
acagacatgc caactcttga tgaacttaat gctttgcctt acttagaagc attcgtaag	1200
gaatcattac gtctagaccc tccttcaccc tatgcaaaata gagaatgctt aaaggatgaa	1260
gatttcatac ctttagcaga gccagtcatt ggaagagacg gtgtgttat caatgaagtc	1320
agaattacaa aaggcacaat gtttatgtt ccattgttta acattaacag atcaaagttt	1380
atctacggtg aggacgctga agatgtttaga ccagaaagat gtttggaaaga tggtacagat	1440
tctttaaca gtatcgaaac tccatacgga catcaggctt catttttc tggacctaga	1500
gcatgttttgc tggagatt tgccgttgct gaaatgaaag cattttttt cgtaacactt	1560
agaagggtgc aatttgaacc aatcatctca cacccagaat acgaacatata cacccgtatc	1620
atatctagac cacgtatagt tggtagagag aaggaaggaa atcaaattgag actacaagtg	1680
aaaccagtag agtaaccgcg gctcgagccc ggg	1713

&lt;210&gt; 60

&lt;211&gt; 951

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 60

ccggggaat tcaagcttaa aatggcagca ggtctgtcca caattgccgt aacttgaag	60
ccacttaatc gttcctcatt ttctgctaat catcctataat ctactgccgt ctttccacct	120
tcatattaat tcaacggcctt cagaaggagg aaaatcctaa cagtagtgc tt ctagtcgag	180
gaacgtaaac agtccctc aatggacgt gataacaac ctgaatctac tacaagtca	240
tctgaaattc taatgacatc tagactgcta aagaaagctg agaaaaaaaaa atctgaaaga	300
ttcacttaatc tgattgctgc agttatgagt tccttcggta tcacatcaat ggcaatcatg	360
gctgtgtatt acagatttc atggcaaatg aaggcggtg aagttcagt tttggagatg	420
tttggcactt tcgctctatc tgttaggagca gcagtcggaa tggagtttg ggctagatgg	480
gccccacagag cttgtggca cgactctta tggatatgc atgaatccca tcacaaacca	540
agagaaggta ctttgaatt gaatgatgtg tttgcaatta ccaacgcgt tccagccatt	600
ggtttactgt attacggatt ccttaacaaa ggtttagtc caggcttatg ttttgggtct	660
gggttggaa tcacaatgtt tggatggcc tacatgttg tgcgtatgg tctggcac	720
aaaagatttc ctgtcgccc aatagctaat gttccatatac tttagaaaaat tgctggca	780
catcaattgc atcataccga taagttcaaa ggtgttcctt acgggtgtt tctggacca	840
aaggaagtag aggaagttgg aggttaaggaa gagttagaaaa aggaatcag tagaaggata	900
aagtataaca ataagggttc atctaccta taaccgcggc tcgagccgg g	951
<210> 61	
<211> 969	
<212> DNA	
<213> Adonis aestivalis	
<400> 61	
ccggggaat tcaagcttaa aatgcttagct tctatggcag ctgctacctc tataaccta	60
tcttcttagat cttcagatt ccatagaggg ttattccta atacaagcc taatatcaga	120
aaccacccat gcttattgtt ttccccactg ctaatgcgtt acagaaatgg agcagggct	180
ttgacaattt gttcgtcgc tgagagaaca agaggaagag aaattccaca aatcgaagag	240
gatgagaaga atatggacga agtatttgaa cagatgaata gtgctagtgt aagggttgca	300
gagaaacttg cacgtaaaaa atctgaaaga ttacttatt taattgccgc ttaatgagt	360
tcaatggta ttacttccat ggctataactt tcagtctact acagatttc ctggcaaatg	420
gaggggtggcg atatccctgt tacagaaatg ttggcactt ttgcattgtc ttaggtgt	480
gcagtcggta tggatatttg ggcaaggtgg gctcatagag ccctgtggca cgcctcattt	540
tggcacatgc atgaatcaca tcacaaacct agagaaggac catttgaatt gaacgtatgtt	600

ttcgcataaa tcaacgccgt tcctgctata gccctattga atttcggctt tttccataaa	660
gttttgcattc cagggttatg tttgggtca ggtctggta tcacagtgtt tggaatggct	720
tacatgttcg tgcacatgcgg tttagtgcattt agaagattcc cagtagggcc aattgctaact	780
gtgccttact tttagaaaagt tgccgcagca caccaaatcc accataactga taaatttcaa	840
ggagttccat atggtctatt tctaggccct aaggaactgg aggaagttgg cggaatgag	900
gaatttagaaa aggaatcga acgtagaattt aagagaatga atgccttta accgcggctc	960
gagccccggg	969

&lt;210&gt; 62

&lt;211&gt; 1149

&lt;212&gt; DNA

&lt;213&gt; Crocus sativus

&lt;400&gt; 62

cccgggaaat tcaagcttaa aatgcaagt gaccctacaa aaggaattgg tttagccat	60
acatcccttc aattctcaaa tggcagatta catgcattgt gtgaatatga cttgccttac	120
gtcgtgagac ttctccaga ggatggagac atttccactg ttggtagaat agagaataac	180
gtgtcaacta agtctaccac agcccatcca aagaccgatc ctgtcactgg cgaaacttc	240
tcttcttctt acggtccat acaaccatac gttacttact caagatatga ttgtatggc	300
aaaaagtctg gcccagatgt gccaatctt tcattcaagg aaccaagttt tgtacatgat	360

tttgcattt cagagcacta tgctgtttc ccagatatac agatcgtat gaaaccagct	420
gaaattgtta gagggagacg tatgattggt ccagacctgg aaaagggtcc tagacttaggt	480
cctttaccaa gatatgctac atccgatagt gaaatgagat ggttgtatgt tcctgggtt	540
aacatggttc atgcgttta tgcttggag gaggaggag gcgaagggtt cgttattgtat	600
gctcctaactg tatcaccaat cgaaaatgca atcgacatg ttgatctgt gcatgtctt	660
gttgaatgg cccgtataga attgaaatcc ggatcgtttt ctagaacact attgagtgt	720
gaaaacttag acttgggt gatccataga gtttactctg ggagaaaatc aagatacgct	780

tacttagtg tgggtgaccc tatgcataag atcagaggtg tggtaaagt agatggaa	840
ttggccgta ggggtgaatg cgttgttagca agacgtgagt tcggcgctgg atgtttgg	900
ggagagccat ttctgttacc tgcattttt aaaaaatccg gtggcgaaga ggatgacgga	960
tacgtggta gtacatcaca cgtatgaaggc aaggagaat catctttgt agttatggat	1020
gctagaagtc cagaactaga aatcttggcc gaagtcgtac ttccaaggag gttccatac	1080
gttttcacg gtttgttcgt cactgaagca gaattgttat cacaacaata accgcggctc	1140

gagcccggg 1149

<210> 63

<211> 1680

<212> DNA

<213> *Crocus sativus*

<400> 63

ccggggaaat tcaagcttaa aatggcgaa gtcgcaaagg aagaggttga agagagaaga 60

tccatgttg ccgtcaaccc acaaccttcc aaaggattgg tatcctccgc tgttgatctt 120

atagagaaag ctgtggtta tctttccat gataagagt aaccttgcca ctaccttct 180

ggtaacttg cacctgttgt agacgaaact cctccttgc ctgattgcc tgtagaggc 240

cacttaccag aatgtctgaa tggggagttt gtaagagtgc gtccaaatcc aaagtttatg 300

cctgtggctg ggtaccattg gtttgcgtt gatggtatga ttcatggat gaggattaag 360

gatggaaaag ctacttatgc ttcttagatat gtcaaaaacat ctagactaaa acaggaggaa 420

tactttgaag gtccaaagtt catgaagatc ggagacttga aagggtttt tggcctgtt 480

atggttcaaa tgcaactttt gagggcaaag ttaaaggtaa tcgatgtaa ttacggtgta 540

ggaaccggta atacagccct gatataccat catggtaaac tgttggcact ttcagaagca 600

gataaggctt acgttagtcaa agttttggaa gatggcgatc tacaaacatt gggcttattg 660

gactacgata agagactgtc ccactcattc acagctcacc caaaggtaga tcctttaca 720

gacgaaatgt tcactttgg gtacgctcac acaccacat acgtcactta tagagtgtc 780

tcaaaaagacg gtgttatgag agacccagtt cctattacaa tacctgcctc agtgatgtg 840

catgattcg ctattactga gaattactca atctttatgg acctaccact atactttcaa 900

ccaaaggaaa tggtgaaggg tggaaagtgc atttcagtt tcgacgccac caaaaaagcc 960

agatttggtg tttacctag atacgtaaa gatgactcac ttatcagatg gtttgaatta 1020

ccaaatttgtt tcatttcca caacgccaat gcttggagg aaggtgatga agtagttcta 1080

atcacctgca gattagagaa cccagacttg gatatggta atggagcgtt caaagagaaa 1140

ctagaaaaact ttaaaaatga actgtatgaa atgagattca acatgaaaac tggagcagca 1200

tcacaaaagc aactatctgt gtctgccgta gatTTccac gtatcaatga atcttacaca 1260

accagaaaac aacgttacgt ctatggact atcttgaca atatcactaa agtcaaaggc 1320

ataatcaagt ttgatctaca tgctgaacca gaagcaggaa aaaagaattt ggaagtgggt 1380

ggtaatgttc aaggcatatt cgatttggc ccaggaaggt atggctctga agctgtctt 1440

catgatgaga acacaggaa atctgaggtt aacgttattg atgcaaagac aatgagtgcc	1560
gaaccagtgg ctgttgga acttccaaat cgtgttccat acggttcca tgcattctt	1620
gtaaacagg aacagttaca gtggcagcaa actgatgtt aaccgcccgt cgagcccccgg	1680
<210> 64	
<211> 1824	
<212> DNA	
<213> Chrysanthemum x morifolium	
<400> 64	
cccgaaaaat tcaagcttaa aatggattgt ttatcatcct cttccgttc cacttctca	60
ccaacaaatt catactcatc tagtccacca ttgcctacct ctaaaccttc atctagttt	120
agagtgtttt ccgttagaac agaagataaa ctacagactg ttacaactac aaccaaaagg	180
ccttccgatg aacaggtcaa aaaacacact acaccatctt tcaatatcga aaagagaagt	240
agatccgtcg ttgtcgacca atctctgcca tctactttcc ttaatgctt cgacaatatc	300
ataaacaatt tcatcgatcc acctttgaga gttttagtgc acccaaagca tgtcttatca	360
gacaactttt ctccagtgga tgagttaccc cctacagaat gccaagtaat tgaaggtaca	420
ctaccatctt gcttagatgg tgcatacttc agaaatggtc ctaaccacaa atttctgcca	480
agagggccat atcatttgtt ttagggat ggtatgtgc acgcaattag aatttccaaat	540
ggcaagcta ctttctgttc aagatatgtt aagacttaca aatacaatat cgagaaagat	600
gcaggcttcc caatcattcc aatgttttg gctggttca atggatgac agcatcagcc	660
gccagaatgg ctgtttggc cggaaatggc cttggccgtc aatacgaccc tacaaagggt	720
atcgccctag ccaatacata tttggcttac tttggtaaca agctatacgc acttggtaaa	780
tctgatcttc ctacgcgtt gaaattggca cctaatggtg acatcattac aactggcagg	840
catgattttg atggcaact gtttatgtct atgacagctc atccaaagat agacccagaa	900
actaaggaag cattttccctt cagatacgtt ccaatgccac catttctaac cttttttaga	960
tttgaccaaa acggtgagaa acaaccagac gttccaaatct tctcaatgac ttctccaagt	1020
tttctgcattt acttagctat aaccaaaaac tacgcattat tcccaagat ccaatcgcc	1080
atgtcaccaa tgaaatgtt agggatggg tctcctgtt ctgctgattc cggcaagata	1140
ccaagattgg ggctgattcc aaggtacgtt aaggacaaaa gtgaaatgaa atggtttgg	1200
gttccaggct tcaatgtat ccactgcattt aatgcattgg aggaagatgg tggagacaca	1260
gttgggttgg tagtcctaa cattttgtca gttgaacaca cccttgagag aatggatctg	1320
attcatgctt ctattgaaaa ggtcaccatc aacttaaaga caggaatggt gtcaagacat	1380

cctttgtcta ctagaaacctt agatttgtc gtattgaatc ctgccttcata tgccgtcaaa	1440
aatagataca tctactgtgg agttgggtat cctatgccaag agatttccgg agttgtaaaa	1500
ctggatgtgt cacttagtga agtagataga cgtgagtgtat ttgttgctat tagaatgttt	1560
ggaccagggtt gtttggcgagagccttt tttgtagcta gagaacctga taatcctgaa	1620
gccgatgaag atgatggcta tgtgattca tatgtccata acgaaaacac aggtgaatct	1680
cgttttgtt ttagggatgc aaaatctcca actttggaaa tagtagcagc tgtaaaacta	1740
ccttagacgtt ttccatatgg gttcacggg ttattcgta aggaaagtga catcaacaaa	1800
ctttaaccgc ggctcgagcc cgccc	1824
<210> 65	
<211> 1380	
<212> DNA	
<213> Crocus sativus	
<400> 65	
atgaaacaaa aggacgtcaa cggttaacaag tgtcacatct tattattgcc atgcctgct	60
caaggtcaca tcaaccctat cttacaattt ggtaaaagat tggcttcaca taatttttg	120
actacattgg tcaataccag attcttgtcc aacagtacta agtctgaacc aggtcctgtt	180
aacatagaat gtagtctga tggtttgac tcaggtggta tgaacgctgc accatctaga	240
agagcatatt ttgatagatt ggaatccgtt ggtatgaaa cattgtcgg tttgatcgaa	300
tcttaagat caagaggttag acctgccccat gtttagtct acgatccatt ctgccttgg	360
gctatgaacg tagcagaag atcagggttg agatcgttg cttctttac acaaccatgc	420
gcagtcgata ccatctatag acacgttatgg gaaggtagaa ttaaagttcc agttaccgaa	480
cctgttagat tgccagggtt accaccttg gaaccttcgtt attgccaaat ttttgttact	540
gattctgacc ctgttgtcaa tccagacttg ttaccttgtt tggttaacca acataagaac	600
ttagataagg ccgacatgat gttgatcaac tcaatctatg aattagaaca cgaagaattt	660
gattggatgg aatccagatt gccattacct gttaaatcta tcgggtccaaat agtccctca	720
acctacttgg ataatagaat accatctgac tcacattacg gtttaactt atacactcca	780
gataccactc ctatattgga ttgggtggac tccaaaggcac caaatagtgtaatatacgat	840
tccttcggta gtttatcttc attgtctctt gatcaaacta atgaaatagc atccggtttgc	900
atcgccacaa acaaaagttt tatctggta gtttagaacct ctgaatttgcg caagttgcac	960
gctaatttca ctcaagaaaaa cgcttcaaga ggttttagtgc taacatggta tgatcaatta	1020
gacgtattgg cacatgttgc cactgggtgtt ttgttacac actgcgggttgc gaactctaca	1080

atggaaggta ttgcattggg tttccaatg gtcgggtac ctcaatggc cgatcaacca	1140
atgaacgcca aatatgtcg aagacgtatgg aagggtggc tcagagctaa gacttatggt	1200
aaagatttcg ttagaggtga agagttaaa agatgtgtg aagaagtcat ggacggtgaa	1260
agaagtgta aaataagaga aaatgccgt agatggtgca aattagctaa ggattctgtt	1320

tcagaaggta gttcctcaga caagtgtatc aaagaattt tccatcaatg ctgtaagtaa	1380
--	------

&lt;210&gt; 66

&lt;211&gt; 459

&lt;212&gt; PRT

&lt;213&gt; Crocus sativus

&lt;400&gt; 66

Met Glu Gln Lys Asp Val Asn Gly Asn Lys Cys His Ile Leu Leu Leu	
---	--

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

Pro Cys Pro Ala Gln Gly His Ile Asn Pro Ile Leu Gln Phe Gly Lys	
---	--

20	25	30
----	----	----

Arg Leu Ala Ser His Asn Leu Leu Thr Thr Leu Val Asn Thr Arg Phe	
---	--

35	40	45
----	----	----

Leu Ser Asn Ser Thr Lys Ser Glu Pro Gly Pro Val Asn Ile Glu Cys	
---	--

50	55	60
----	----	----

Ile Ser Asp Gly Phe Asp Ser Gly Gly Met Asn Ala Ala Pro Ser Arg	
---	--

65	70	75	80
----	----	----	----

Arg Ala Tyr Phe Asp Arg Leu Glu Ser Val Gly Ser Glu Thr Leu Phe	
---	--

85	90	95
----	----	----

Gly Leu Ile Glu Ser Leu Arg Ser Arg Gly Arg Pro Ala His Val Leu	
---	--

100	105	110
-----	-----	-----

Val Tyr Asp Pro Phe Leu Pro Trp Ala Met Asn Val Ala Glu Arg Ser	
---	--

115	120	125
-----	-----	-----

Gly Leu Arg Ser Val Ala Phe Phe Thr Gln Pro Cys Ala Val Asp Thr	
---	--

130	135	140
-----	-----	-----

Ile Tyr Arg His Val Trp Glu Gly Arg Ile Lys Val Pro Val Thr Glu	
---	--

145	150	155	160
-----	-----	-----	-----

Pro Val Arg Leu Pro Pro Leu Glu Pro Ser Asp Leu Pro	
---	--

165	170	175
-----	-----	-----

Ser Phe Val Thr Asp Ser Asp Pro Val Val Asn Pro Asp Leu Leu Pro

180	185	190
-----	-----	-----

Leu Leu Val Asn Gln His Lys Asn Leu Asp Lys Ala Asp Met Met Leu

195	200	205
-----	-----	-----

Ile Asn Ser Ile Tyr Glu Leu Glu His Glu Glu Phe Asp Trp Met Glu

210	215	220
-----	-----	-----

Ser Arg Leu Pro Leu Pro Val Lys Ser Ile Gly Pro Thr Val Pro Ser

225	230	235
-----	-----	-----

240

Thr Tyr Leu Asp Asn Arg Ile Pro Ser Asp Ser His Tyr Gly Phe Asn

245	250	255
-----	-----	-----

Leu Tyr Thr Pro Asp Thr Pro Tyr Leu Asp Trp Leu Asp Ser Lys

260	265	270
-----	-----	-----

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

275	280	285
-----	-----	-----

Ser Pro Asp Gln Thr Asn Glu Ile Ala Ser Gly Leu Ile Ala Thr Asn

290	295	300
-----	-----	-----

Lys Ser Phe Ile Trp Val Val Arg Thr Ser Glu Leu Ala Lys Leu Pro

305	310	315
-----	-----	-----

320

Ala Asn Phe Thr Gln Glu Asn Ala Ser Arg Gly Leu Val Val Thr Trp

325	330	335
-----	-----	-----

Cys Asp Gln Leu Asp Val Leu Ala His Val Ala Thr Gly Cys Phe Val

340	345	350
-----	-----	-----

Thr His Cys Gly Trp Asn Ser Thr Met Glu Gly Ile Ala Leu Gly Val

355	360	365
-----	-----	-----

Pro Met Val Gly Val Pro Gln Trp Ser Asp Gln Pro Met Asn Ala Lys

370	375	380
-----	-----	-----

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

385	390	395
-----	-----	-----

400

Lys Asp Phe Val Arg Gly Glu Glu Phe Lys Arg Cys Val Glu Glu Val

405	410	415
-----	-----	-----

Met Asp Gly Glu Arg Ser Gly Lys Ile Arg Glu Asn Ala Ala Arg Trp

420 425 430

Cys Lys Leu Ala Lys Asp Ser Val Ser Glu Gly Gly Ser Ser Asp Lys

435 440 445

Cys Ile Lys Glu Phe Ile His Gln Cys Cys Lys

450 455

<210> 67

<211> 1521

<212> DNA

<213> Saccharomyces cerevisiae

<400> 67

atgcctacct tgtatactga tatcgaaatc ccacaattga aaatctctt aaagcaaccg	60
ctagggttgt ttatcaacaa ttagtttgt ccatcatcag atggaaagac catgaaaact	120
gtgaacccag ctactggcga accgataaca tccttccaag cagctaacga aaaggatgta	180

gacaaagctg taaaagctgc cagggctgct tttgataacg tttggcgaa gacatcttct	240
gagcaacgtg gtatttatct ttcaaactta taaaactta ttgaggagga gcaagacaca	300
cttgcgcatt tagagacttt agacgctgga aagccttacc attcaaatgc caaagggtat	360
ttggcacaaa tttagactt taccagatat ttggctgggt ccgctgataa gtttgacaaa	420
ggtgcaacca taccattgac ttttacaag tttgcatata ctctaaaagt tcctttggc	480
gttggcgttc aaatcggtcc atgaaattat cctctagcta tggcttggg gaaattgcaa	540
ggtgcccttag cagccggtaa cacggttatc atcaaaccctg ctgagaatac ctctctatct	600

ctactttatt ttgctacttt aattaaaaaa gcagggtttc cacctgggtgt tgtcaatatc	660
gttccgggtt atggatcaact tggggccaa gcccggatcat ctcacatggat tatggacaaa	720
atatcttta cgggaagcac caaggctgggt ggattttgtgt tggaagcttc cggccaatcg	780
aaccttaaag acgttacact agaatcggtt ggtaagtctc ctgtctcgat atttgaagat	840
gcagacccggata ataaggctat cgattggata gcagctggca ttttctacaa ttcaggacag	900
aattgttacgg caaaactcaag agtttatgtt caaagttcga tctacgacaa gtttggtaa	960
aagttaaag aaactgcaaa gaaggagtggtt gatgttgcag gaaaatttga tccgttgc	1020

gagaaatgca tcgttggtcc agtttatca agtacacagt atgaccgcatttcaaaatgtac	1080
atagaacgtg gtaaaaggaa ggaaaagttt gacatgttcc agacctctga atttcctatt	1140
ggggggacta aaggctactt cattccccca accatcttca ctgtatgtccc gcaaacatcg	1200

aaactgttac aggatgagat attggcccg gttgtggtg ttagcaagtt cacaattat	1260
gatgacgctc tgaagctggc taatgatact tgctacggc tcgcctctgc ggtttcaca	1320
aaagatgtca agaaagcgca catgtttgt cgcgatatta aagcaggaac tgttggatc	1380
aactcatcta acgatgaaga ttttaccgtt cctttggcg gtttaaat gagtggtatt	1440
ggttagagaac tggggcaaag tgggtgtat acctatctc aaacaaaagc agttcacata	1500
aatctcttt tggacaacta a	1521
<210> 68	
<211> 1521	
<212> DNA	
<213> <i>Saccharomyces cerevisiae</i>	
<400> 68	
atgcctacct tgtatactga tatcgaaatc ccacaattga aaatctctt aaagcaaccg	60
ctagggttgt ttatcaacaa ttagtttgtt ccatcatcag atggaaagac catgaaact	120
gtgaacccag ctactggcga accgataaca tccttccaag cagctaacga aaaggatgt	180
gacaaagctg taaaagctgc caggctgtct ttgataacg tttggcgaa gacatctct	240
gagcaacgtg gtatttatct ttcaaactta taaaactta ttgaggagga gcaagacaca	300
cttgcgcatt tagagacttt agacgctggt aagccttcc attccatgc taaacaagac	360
ttagcccaga ttatagaact tacaagatac tatgcgggg cggtcgacaa gttcaatatg	420
ggtgaaacca ttccattgac ttttacaag tttgcatata ctctaaaagt tcctttggc	480
gttgggtgtc aaatcggtcc atgaaattat cctctagcta tggcttgtag aaaaatgcaa	540
ggtgccttag cggccggtaa cacggttatac atcaaaccctg ctgaaaatac ctcttatct	600
ctactttatt ttgctacttt aattaaaaaa gcaggttttc cacctgggtgt tgtcaatgtc	660
attccctgggtt atggttccgt tggggaaa gcttttagaa cccacatgga tatcgacaaa	720
atatcttta cggaaagtac taagggtggc ggctcagtat tggaaagcttc cggccatcg	780
aaccttaagg atatcacact agaatgcggt ggtaagtctc ctgctttgtt atttgaagat	840
gcagaccttg ataaggctat agaatggta gcaaatggta tttttttaa ttcgggacag	900
atctgcactg caaaactcaag agtttatgtt caaagttcga tctacgacaa gtttggtaa	960
aagttaaag aaactgcaaa gaaggagtgg gatgttgcag gaaaatttga tccgtttgat	1020
gagaaatgca tcgttggtcc agtttatatca agtacacagt atgaccgcat caaaagtac	1080
atagaacgtg gtaaaaagga ggaaaagttt gacatgttcc agacctctga atttcctatt	1140
ggtggagcta aaggctactt cattccccca accatttca ctgatgtacc agaaacatct	1200

aagtgtgc gtgatgaaat attggcccg gttgtggtg ttagcaagtt cacaattat	1260
gatgacgctc tgaagctggc taatgatact tgctacggc tcgcctctgc ggtttcacc	1320
aaagatgtca agaaagcgca catgttgcg cgcgatatta aagcaggaac tggttggatc	1380
aatcaaacc aatcaagaaga agctaaagtt cctttggcg gatctaagat gagtggtatt	1440
ggtagagaat caggcgacac cggcggttat aactattac aaataaaatc agtccatgtg	1500
gatcttcat tggataaata a	1521
<210> 69	
<211> 1560	
<212> DNA	
<213> <i>Saccharomyces cerevisiae</i>	
<400> 69	
atgttcagta gatctacgct ctgcttaaag acgtctgcat cctccattgg gagacttcaa	60
ttgagatatt ttcacacacct tcctatgaca gtgcctatca agctgccaa tgggttggaa	120
tatgagcaac caacggggtt gttcatcaac aacaagtttgc ttccctctaa acagaacaag	180
accttcgaag tcattaaaccc ttccacggaa gaagaaatat gtcatattta tgaaggtaga	240
gaggacgatg tggaaagaggc cgtgcaggcc gccgaccgtg cttctctaa tgggtcttgg	300
aacggtatcg accctattga cagggtaag gctttgtaca gtttagccga attaattgaa	360
caggacaagg atgtcattgc ttccatcgag actttggata acggtaaagc tatctttcc	420
tcgagaggag atgttgattt agtcatcaac tatttgcattt cttctgtgg ctttgctgat	480
aaaatttgatg gtagaatgat tgatactggt agaaccatt tttcttacac taagagacag	540
ccttgggtt ttgtggca gattattcct tggaaatttcc cactgttgat gtggcctgg	600
aagattgccc ctgcttggt cacggtaac accgtcggt tgaagactgc cgaatccacc	660
ccattgtccg cttgtatgt gtctaaatac atcccacagg cgggtattcc acctgggttg	720
atcaacatttgc tatccgggtt tggtaagatt gtgggtgagg ccattacaaa ccatccaaaa	780
atcaaaaagg ttgccttcac agggtccacg gctacggta gacacattta ccagtccca	840
gccgcaggct tggaaaaagt gactttggag ctgggtggta aatcaccaaa cattgttttc	900
gcggacgccc agttgaaaaa agccgtgcaa aacattatcc ttggatctta ctacaattct	960
ggtgaggctt gttgtgcggg ttcaagggtt tatgttgaaat aatctattta cgacaaattc	1020
attgaagagt tcaaagccgc ttctgaatcc atcaaggtgg gcgaccatt cgtgaatct	1080
actttccaag gtgcacaaac ctctcaaatg caactaaaca aaatctgaa atacgttgac	1140
attggtaaga atgaaggtgc tactttgatt accgggtggtaaagatttgg tagcaagggt	1200

tacttcatta agccaactgt ctttggtgac gttaaggaag acatgagaat tgtcaaagag	1260
gaardtttg gcccttgtt cactgtaacc aaattcaaat ctgccgacga agtcattaac	1320
atggcgaacg attctgaata cgggttggct gctggtattc acacctctaa tattaatacc	1380
gccttaaaag tggctgatag agttaatgcg ggtacggtct ggataaacac ttataacgat	1440
ttccaccacg cagttccctt cggtggttc aatgcatctg gttggcag ggaaatgtct	1500
gttgatgctt tacaaaacta cttgcaagtt aaagcggtcc gtgccaattt ggacgagtaa	1560

&lt;210&gt; 70

&lt;211&gt; 1563

&lt;212&gt; DNA

<213> *Saccharomyces cerevisiae*

&lt;400&gt; 70

atgcttctc gcacaagagc tgcaagtcgg aattccagaa tattcactag aagcttgttta	60
cgtcttatt ctcaaggcacc attacgcgtt ccaattactc ttccaaatgg tttcacctac	120
gaacagccaa cagggttatt catcaatggt gaatttgtt cctcgaagca aaagaaaaacg	180
tttgcgtga tcaatccatc taacgaagaa aagataacaa ctgtatacaa ggctatggaa	240
gatgatgtt atgaagccgt tgcaagcggct aaaaaagctt ttgaaacgaa gtggcttatt	300
gttagagccgg aggttcgcgc taaagctta ttcaatctcg ctgacttggt tgagaaacac	360

caagaaaacac tggctgccat tgagtcaatg gataatggta agtcatgtt ttgtgcgcgc	420
gggtgacgtcg cttagtatac taaatacttg cgttttgcg gtgggtggc agataaaaatc	480
tacggtaacg ttattgacac aggtaaaaac cattttacct actcaattaa ggaaccattt	540
ggcggttgcg gccaaataat cccttggAAC ttcccttat tggatgtggc atggaaaattt	600
gggcctgctc tggctacagg taacaccgtc gtattgaaac ccgctgaaac aacacctta	660
tctgccctt tcgcttccc gttgtgtcag gaagcaggca taccgcgtt gtttagtcaat	720
atccctccgg gtccggtag agttgttggaa gaaagattga gtgcacaccc agacgtgaag	780

aagattgctt ttacaggctc tactgccacc ggccgcata ttatgaaggt cgctgccat	840
actgtcaaga aagtcaactt ggagctggga ggttaatcac caaatattgt gtttgcgtac	900
gctgatctag ataaagccgt caagaacatt gccttcggta tttttacaa ctctggtaa	960
gtttgcgtcg ctgggtccag aatatacatt caagatacag tatacgagga ggtgttgcac	1020
aaactaaagg attacaccga gtcactaaag gtcggtgacc catttgcgtga ggaagtttc	1080
caaggtgctc aaacatctga caaacagctg cataaaattt tagactatgt cgatgtac	1140
aaatcagagg gggctcgctc tgtgacttggaa gggccagac atggcagtaa agtttattt	1200

gtcaagccaa cagtgtttgc tcatgtcaaa gaagatataa ggaggaaatg	1260
tttggtccca ttgttaactgt atccaagttt tctactgttg atgaagtgtat tgctatggca	1320
aatgattctc aatatgggtt agccgcaggattcacacta acgatattaa caaggcttt	1380
gatgtgtcca aaagagtgaa agctggtaact gtttgataaa atacctataa caacttccac	1440
caaaaatgttc cttcgggtgg cttcggccag tcaggtattt ggcgtgaaat gggtgaggct	1500
gctttaagta actacactca aacaaaatct gtcagaattt ccattgacaa gccaattcgat	1560
tga	1563

&lt;210&gt; 71

&lt;211&gt; 1503

&lt;212&gt; DNA

<213> *Saccharomyces cerevisiae*

&lt;400&gt; 71

atgactaagg tacactttga cactgctgaa ccagtcaaga tcacacttcc aaatggtttgg	60
acatacggc aaccAACCGG tctattcatt aacaacaagt ttatgaaagc tcaagacggat	120
aagacctatc cccgtcgaaga tccttccact gaaaacacccg tttgtgaggt ctcttctgcc	180
accactgaag atgttgaata tgctatcgaa tggccgacc gtgtttccca cgacactgaa	240
tgggctaccc aagacccaag agaaagaggc cgtctactaa gtaagttggc tgacgaaattt	300
gaaagccaaa ttgacttggt ttcttccattt gaagcttgg acaatggtaa aactttggcc	360

tttagccgtg gggatgttac cattgcaatc aactgtctaa gagatgctgc tgcctatgcc	420
gacaaagtca acggtagaac aatcaacacc ggtgacggct acatgaactt caccaccta	480
gagccaatcg gtgtctgtgg tcaaattttt ccatgaaact ttccaaataat gatgttgct	540
tggaaatcg ccccagcatt ggccatgggt aacgtctgtt tcttggaaacc cgctgctgtc	600
acaccttaa atgcctata ctgtttttt ttatgtaaga aggttggat tccagcttgt	660
gtcgtaaca tcgttccagg tcctggtaga actgttgggt ctgttttgc caacgaccca	720
agaatcagaa agctggctttt taccggttt acagaagtcg gtaagagtgt tgctgtcgac	780

tcttctgaat ctaacttggaa gaaaatcact ttggaaacttag gtggtaagtc cgcccatgg	840
gtctttgacg atgctaacat taagaagact ttaccaaaatc tagtaaacgg tattttcaag	900
aacgtggtc aaatttggcc ctctggttt agaatttacg ttcaagaagg tattttacgac	960
gaactattgg ctgtttcaa ggcttacttg gaaacccaaa tcaaagggttgg taatccattt	1020
gacaaggcta acttccaaagg tgctatcact aaccgtcaac aattcgacac aattatgaac	1080
tacatcgata tcggtaagaa agaaggcgcc aagatctttaa ctgggtggcga aaaagttgg	1140

gacaagggtt acttcatcag accaaccgtt ttctacgatg ttaatgaaga catgagaatt	1200
gttaaggaag aaattttgg accagttgtc actgtcgcaa agttcaagac tttagaagaa	1260
ggtgtcgaaa tggctaacag ctctgaattc ggtcttaggtt ctggtatcga aacagaatct	1320
ttgagcacag gtttgaaggt gccaagatg ttgaaggccg gtaccgtctg gatcaacaca	1380
tacaacgatt ttgactccag agttccattc ggtgggtta agcaatctgg ttacggtaga	1440
gaaatgggtg aagaagtcta ccatgcatac actgaagtaa aagctgtcag aattaagttg	1500
taa	1503
<210> 72	
<211> 1599	
<212> DNA	
<213> <i>Saccharomyces cerevisiae</i>	
<400> 72	
atgtcaaacg acggctcaaa aatattgaat tataccccag tgtctaaaat agatgaaata	60
gttgaatct caagaaattt cttcttttag aacaattga aattgtccc cgaaaataac	120
ccaaggaaaa aagatctaga attcaggcag ttgcagttga aaaaactcta ttatgccgtc	180
aaagatcatg aggaagaact gatcgatgt atgtacaagg actttcatcg gaacaaaatt	240
gaatcggttc tgaatgaaac gaccaaactt atgaacgata tactcacct aattgagatt	300
ttaccaaaaat tgaatcaaacc tcggagagta tctgattttt ctccctcatt tatgtttgg	360
aaaacaatcg tggagaaaat atcaaggggc agtgtttga ttattgctcc tttcaatttt	420
ccctacttt tagcattgc cccattggca gcagctttg ctgcaggtaa caccattgtt	480
ctgaagccaa gtgaactaac accacacact gctgtagttt tgaaaaattt gttaccaca	540
gctggttcc ctgatggatt gattcaagta gttcagggag ctatagatga aactacaaga	600
ctactagatt gtggaaaatt tgacctaata ttctacacag gttctccccg tgtcggtca	660
atagggtctg agaaagcgc aaaaagtcta acaccttgta tacttgaact tggtggtaaa	720
tcacctacct ttattacaga aaatttcaaa gcaagtaaca taaaattgc tttgaaaagg	780
attttttttg gtgtttcgg aaattctggc cagatttgta ttccaccaga ttattttgttta	840
gtacataaaat ctatctatcc aaaagtctttaa aagagtgtg aatcagtact aaatgaattt	900
tatccaagct ttgatgaaca aacagattc actcgatgtt ttcattgagcc tgcttacaaa	960
aaggccgttcaagtataaa ctcaactaac ggctccaaga ttgtgccttc aaaaatttct	1020
atcaattcag atactgagga tctatgcctt gtaccaccaa ccatgttta taacattgg	1080
tggatgatc ctttgatgaa acaggaaaac ttgtccctg tattgcccattt cattgagttac	1140

gaggatctt atgagaccat taacaagata atagaagaac atgacactcc attggtgcaa	1200
tacatattct ctgatagcca aactgaaata aatcgatct tgacgcgctt aagatctgg	1260
gactgtgtt tcgggtatac agtgattcat gtaggaatta ccgacgctcc atttggaggg	1320
atcggtactt caggttatgg taactatggt ggatattatg gattcaatac cttagtcat	1380
gaaagaacaa ttttaaaca accatattgg aatgattttt cccttttat gagataccct	1440
ccaaatagcg cacaaaagga aaagctcgtc cgtttgcga tggaaagaaa accttggtt	1500
gacagaaatg gcaataacaa gtgggggtta cgccaatatt tttcattatc tgccgccgtt	1560
attttaatta gtaccattta cgctcatgt tcttcctga	1599