

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(51) Int. Cl.⁶

C12N 15/82

C12N 9/02

C12N 15/53

(11) 공개번호 특1999-0087356

(43) 공개일자 1999년12월27일

(21) 출원번호	10-1998-0706766	(87) 국제공개번호	WO 1997/32028
(22) 출원일자	1998년08월28일	(87) 국제공개일자	1997년09월04일
번역문제출일자	1998년08월28일		
(86) 국제출원번호	PCT/US1997/03343		
(86) 국제출원출원일자	1997년02월27일		
(81) 지정국	AP ARIPO특허 : 말라위		
	EA 유라시아특허 : 벨라루스 키르기즈 카자흐스탄 몰도바 타지키스탄		
	EP 유럽특허 : 오스트리아 벨기에 스위스 독일 덴마크 스페인 프랑스		
	영국 그리스 이탈리아 룩셈부르크 모나코 네덜란드 포르투갈 스웨덴		
	핀란드		
	국내특허 : 아일랜드 오스트레일리아 보스니아-헤르체고비나 바베이도스		
	불가리아 브라질 캐나다 중국 쿠바 체코 그루지야 헝가리 일본 대		
	한민국 세인트루시아 스리랑카 라트비아 마다가스카르 몽고 노르웨이		
	뉴질랜드 슬로바키아 우크라이나 미국 폴란드 루마니아		
(30) 우선권주장	60/012,705 1996년02월28일 미국(US)		
	60/013,612 1996년02월28일 미국(US)		
	60/020,003 1996년06월21일 미국(US)		
(71) 출원인	노바티스 아게 더블류. 하링, 지. 보이롤		
	스위스 4058 바젤 슈바르츠발달레 215		
(72) 발명자	존슨 마리 에이		
	미국 노오쓰 캐롤라이나주 27606 할리 히씨 드라이브 408		
	볼라쓰 산드라 엘		
	미국 노오쓰 캐롤라이나주 27707 듀럼 파인 오크 드라이브 4225		
	워드 에릭 알		
	미국 노오쓰 캐롤라이나주 27707 듀럼 벤틀리 드라이브 3761		
(74) 대리인	이병호		

심사청구 : 없음(54) 식물 프로토포르피리노겐 옥시다제 유전자로부터의 프로모터**요약**

본 발명은 식물 프로토포르피리노겐 옥시다제(프로톡스) 암호화 서열과 자연적으로 결합된 프로모터 및 이의 유도체가 제공된다. 이들 프로모터를 사용하여 식물 세포에서 작동가능하게 결합된 이중성 암호화 서열의 발현을 조절할 수 있다. 특히, 이들 프로모터는 상응하는 비변형된 효소를 억제하는 제초제에 대한 내성을 제공하기 위해 변형된 형태의 제초제 표적 효소, 특히 변형된 형태의 프로톡스를 발현시키는데 유용하다. 이들 프로모터를 포함하는 재조합체 DNA 분자 및 키메라 유전자가 제공될 뿐만 아니라, 이와 같은 키메라 유전자를 함유하는 식물 조직 및 식물이 제공된다.

명세서**기술분야**

본 발명은 식물에서 관련 DNA 서열의 전사 프로모터로서 작용하는 신규한 DNA 서열에 관한 것이다. 보다 상세하게는, 본 발명은 식물 프로토포르피리노겐 옥시다제(프로톡스) 암호화 서열과 자연적으로 결합되는 신규한 프로모터에 관한 것이다.

배경기술

1. 프로톡스 효소 및 이의 엽록소/헴 생합성 경로에서의 역할

엽록소 및 헴을 생성시키는 생합성 경로는 단수의 공통 단계로 이루어진다. 엽록소는 모든 녹색 광합성

생물체에 존재하는 광 수집 색소이다. 헴은 헤모글로빈, 시토크롬, P450 혼합-작용 옥시게나제, 퍼옥시다제 및 카탈라제의 보조인자[참조: Lehninger, Biochemistry. Worth Publishers, New York(1975)]이며, 모든 호기성 생물의 필수 성분이다.

엽록소 및 헴 생합성에서 마지막 공통 단계는 프로토포르피리노겐 IX가 프로토포르피린 IX로 산화하는 것이다. 프로토포르피리노겐 옥시다제(이하 "프로톡스(protox)"라 한다)는 이러한 마지막 산화 단계를 촉매하는 효소이다[참조: Matringe et al., Biochem. J. 260: 231(1989)].

프로톡스 효소는 효모 사카로마이세스 세레비지애(*Saccharomyces cerevisiae*)[참조: Labbe-Bois and Labbe, In Biosynthesis of Heme and Chlorophyll, E. H. Dailey, ed. McGraw Hill: New York, pp. 235-285(1990)], 보리 에티오플라스트(etio plast)[참조: Jacobs and Jacobs, Biochem. J. 244: 219(1987)], 및 마우스 간[참조: Dailey and Karr, Biochem. 26: 2697(1987)]을 포함하는 다수의 생물체로부터 부분적으로 또는 완전히 정제된다. 프로톡스를 암호화하는 유전자가 2개의 원핵 생물체, 즉 에스케리키아 콜라이(*Escherichia coli*)[참조: Sasarman et al., Can. J. Microbiol. 39: 1155(1993)] 및 바실루스 서브틸리스(*Bacillus subtilis*)[참조: Dailey et al., J. Biol. Chem. 269: 813(1994)]로부터 분리되었다. 이들 유전자는 서열 상동성이 없고; 이들의 예상된 단백질 생성물은 어떠한 아미노산 서열 동일성도 없다. 에스케리키아 콜라이 단백질은 약 21kDa이며, 세포 구성원과 관련된다. 바실루스 서브틸리스 단백질은 51kDa이고, 가용성의 세포질 활성을 가진다.

또한, cDNA를 암호화하는 프로톡스가 사람[참조: Nishimura et al., J. Biol. Chem. 270(14): 8076-8080(1995)] 및 식물[참조: 1995년 6월 8일자로 출원된 국제 특허원 제PCT/IB95/00452호, 1995년 12월 21일자로 제W095/34659호로 공개됨]로부터 분리되었다.

11. 제초제 표적으로서의 프로톡스 유전자

작물중의 잡초 또는 식물과 같은 목적하지 않는 초목을 제어하기 위해 제초제를 사용하는 것은 통상의 관행이다. 관련 시장은 매년 10억불을 초과한다. 이들의 광범위한 사용에도 불구하고, 잡초 제어는 경작자에게 심각하고 고비용의 문제를 남긴다.

제초제의 효과적인 사용에는 확실한 관리가 요구되고 있다. 예를 들면, 적용 시간 및 방법 및 잡초 식물의 생육 단계가 제초제로 잡초를 효과적으로 제어하는데 중요하다. 여러가지 잡초 종은 제초제에 내성이 있기 때문에, 효과적인 제초제의 생산이 점차 중요해지게 되었다.

불운하게도, 효능이 보다 우수하고 잡초 범위가 보다 넓으며 토양내에서의 분해가 보다 신속한 제초제는 또한 작물 유해성이 보다 클 수 있다. 이러한 문제에 부합되는 특정 액제는 제초제에 저항성 및 내성을 갖는 작물을 발생시키게 되었다. 제초제에 내성이 있는 작물 하이브리드 또는 변종에는 작물에 대한 부수적인 손상의 위험없이 제초제가 사용된다.

내성의 발달은 제초제에 대한 작물의 민감성으로 인해 제초제의 사용이 이미 배제되거나 제한된 작물에 제초제를 적용할 수 있게 한다. 예를 들면, 안데슨 등(Anderson et al.)의 미국 특허 제4,761,373호는 각종 이미다졸리논 또는 설폰아미드 제초제에 내성이 있는 식물에 관한 것이다. 변화된 아세트하이드록시아시드 신타제(AHAS) 효소가 이러한 내성을 부여한다. 굤만 등(Goodman et al.)의 미국 특허 제4,975,374호는 글루타민 신세타제(GS), 예를 들어 포스포노트리신 및 메티오닌 설포시민을 억제하는 것으로 알려진 제초제에 의한 억제에 내성이 있는 돌연변이체 글루타민 신세타제(GS)를 암호화하는 유전자를 함유하는 식물 세포 및 식물에 관한 것이다. 베드브룩 등(Bedbrook et al.)의 미국 특허 제5,013,659호는 설포닐우레아 제초제에 의한 억제에 대해 식물이 내성을 갖도록 하는 돌연변이체 아세트락테이트 신타제를 발현하는 식물에 관한 것이다. 소머스 등(Somers et al.)의 미국 특허 제5,162,602호는 사이클로헥산디온 및 아릴옥시페녹시프로파노산 제초제에 의한 억제에 내성이 있는 식물을 기술한다. 이러한 내성은 변화된 아세틸 조효소 A 카복실라제(ACCCase)에 의해 부여된다.

프로톡스 효소는 각종 제초제 화합물을 위한 표적으로 사용된다. 프로톡스를 억제하는 제초제는 구조가 매우 상이한 분자 부류를 포함한다[참조: Duke et al., Weed Sci. 39: 465(1991); Nandihalli et al., Pesticide Biochem. Physiol. 43: 193(1992); Matringe et al., FEBS Lett. 245: 35(1989); Yanase and Andoh, Pesticide Biochem. Physiol. 35: 70(1989)]. 이러한 제초제 화합물에는 디페닐에테르{예: 아시플루오르펜, 5-[2-클로로-4-(트리플루오로메틸)페녹시]-2-니트로벤조산; 이의 메틸 에스테르; 또는 옥시플루오르펜, 2-클로로-1-(3-에톡시-4-니트로페녹시)-4-(트리플루오로벤젠)}, 옥시디아졸(예: 옥시디아존, 3-[2,4-디클로로-5-(1-메틸에톡시)페닐]-5-(1,1-디메틸에틸)-1,3,4-옥사디아졸-2-(3H)-온), 사이클릭 이미드(예: S-23142, N-(4-클로로-2-플루오로-5-프로파릴옥시페닐)-3,4,5,6-테트라하이드로프탈이미드; 클로로프탈림, N-(4-클로로페닐)-3,4,5,6-테트라하이드로프탈이미드), 페닐 피라졸(예: TNPP-에틸, 에틸 2-[1-(2,3,4-트리클로로페닐)-4-니트로피라졸릴-5-옥시]프로피오네이트; M&B 39279), 피리딘 유도체(예: LS 82-556), 및 페노피레이트 및 이의 0-페닐피롤리디노- 및 피페리디노카바메이트 동족체가 포함된다. 이들 화합물중 다수는 기질 동족체로서 작용하는 것이 명백한 효소에 의해 촉매된 통상의 반응을 경쟁적으로 억제한다.

전형적으로, 프로톡스에 대한 억제 효과는 약 395 내지 410nm에서 여기시킨 후 약 622 내지 635nm에서 형광을 측정함으로써 결정된다[참조: Jacobs and Jacobs, Enzyme 28: 206(1982); Sherman et al., Plant Physiol. 97: 280(1991)]. 이러한 분석은 프로토포르피린 IX가 형광 색소이고 프로토포르피리노겐 IX가 비형광성이라는 사실에 기초하는 것이다.

프로톡스-억제 제초제의 예상된 작용 방식은 엽록체에 프로토포르피리노겐 IX를 축적시키는 것이다. 이러한 축적은 퍼옥시다제 활성에 의해 프로토포르피린 IX로 산화되는 경우 프로토포르피리노겐 IX를 시토크롬에 연결시키는 것으로 생각된다. 광에 노출되는 경우, 프로토포르피린 IX는 시토크롬내에 단일선 산소를 형성시킬 수 있다. 이러한 단일선 산소는 다른 반응성 산소 부류를 형성시킴으로써, 지질 과산화 및 막 분열을 차례로 유발하여 세포의 급속한 처사를 유도할 수 있다[참조: Lee et al., Plant Physiol. 102: 881(1993)].

모든 프로톡스 효소가 식물 프로톡스 효소를 억제하는 제초제에 민감한 것은 아니다. 에스케리키아 콜라 이[참조: Sasarman et al., Can. J. Microbiol. 39: 1155(1993)] 및 바실루스 서브틸리스[참조: Dailey et al., J. Biol. Chem. 269: 813(1994)]로부터 분리된 유전자에 의해 암호화된 2가지의 프로톡스 효소는 이들 제초성 억제제에 내성이 있다. 또한, 페닐이미드 제초제 S-23142에 내성이 있는 단세포 조류 클라 마이도모나스 레인하르티(*Chlamydomonas reinhardtii*)의 돌연변이체가 보고되어 있다[참조: Kataoka et al., J. Pesticide Sci. 15: 449(1990); Shibata et al., In Research in Photosynthesis, Vol. III, N. Murata, ed. Kluwer: Netherlands, pp. 567-570(1992)]. 이들 돌연변이체중 하나 이상은 돌연변이체가 선택된 제초제 억제제뿐만 아니라 다른 부류의 프로톡스 억제제에 대해 내성이 있는 변형된 프로톡스 활성을 갖는 것으로 여겨진다[참조: Oshio et al., Z. Naturforsch. 48c: 339(1993); Sato et al., In ACS Symposium on Porphyrin Pesticides, S. Duke, ed. ACS Press: Washington, D.C.(1994)]. 또한, 돌연변이체 담배 세포주는 억제제 S-21432에 대해 내성이 있는 것으로 보고되었다[참조: Che et al., Z. Naturforsch. 48c: 350(1993)]. 또한, 변형된, 억제제-내성 형태의 식물 프로톡스 암호화 서열이 1995년 6월 8일자로 출원되어 1995년 12월 21일자로 W095/34659로서 공개된 국제 특허원 제PCT/IB95/00452호에 기술되어 있다.

III. 프로톡스 유전자 발현의 조절

프로톡스 유전자와 관련된 다수의 연구는 프로톡스 억제제에 대해 내성을 가질 수 있는 암호화 서열 및 이러한 효소의 변형에 초점을 맞추어 수행되어 왔다. 식물에서 프로톡스 암호화 서열의 발현을 조절 및 촉진시키는 조절 요소에 대해서는 선행 기술에서 어떠한 정보도 이용할 수 없다.

발명의 요약

본 발명은, 본원에서 일반적으로 "프로톡스 프로모터"로 언급되는 식물의 프로토포르피리노겐 옥시다제(프로톡스) 암호화 서열과 자연적으로 결합된 프로모터 영역이 식물에서 이중 암호화 서열의 발현을 촉진하는데 유용하다는 발견에 기초한 것이다.

식물의 프로토포르피리노겐 옥시다제(프로톡스) 암호화 서열과 자연적으로 결합된 프로모터 영역이 식물에서 이중 암호화 서열의 발현을 촉진하는데 유용하다는 발견에 따라, 본 발명은 식물 프로톡스 프로모터 또는 이와 기능상 동등물을 포함하는 분리된 DNA 분자를 제공한다. 추가로 본 발명은, 이중 암호화 서열과 작동가능하게 연결된 식물 프로톡스 프로모터를 포함하는 키메라(chimera) 유전자를 제공한다. 또한, 상기한 키메라 유전자를 포함하는 식물의 조직 및 식물을 제공한다.

본 발명의 한가지 양태에서, 프로톡스 프로모터를 사용하여 제초제 내성 형태의 표적 단백질을 발현시킴으로써 제초제에 내성을 제공한다. 이러한 양태에서, 프로톡스 프로모터는 비변형된 프로톡스 단백질의 억제제에 내성이 있는 제초제-내성 식물 프로톡스 단백질에 대한 암호화 서열과 작동가능하게 연결될 수 있다.

수 탁

하기의 벡터 분자는 하기에 지시된 날짜에 미국 일리노이주 61604 페오리아 노오쓰 유니버시티 스트리트 1815의 농업 연구 기관(Agricultural Research Service)의 북지역 연구 센터의 특허 배양물 수집 기관(NRRL)에 기탁되어 있다:

아라비돕시스(*Arbidopsis*) 프로톡스-1 프로모터를 포함하는 AraPT1Pro는 pWDC-11(NRRL #-21515)로서 1995년 12월 15일에 기탁되었다.

옥수수 프로톡스-1 암호화 서열의 나머지 및 융합된 옥수수 프로톡스-1 프로모터를 포함한 플라스미드가 pWDC-14(NRRL #B-21546)로서 1996년 3월 19일 기탁되었다.

사탕무 프로톡스-1 프로모터를 포함한 플라스미드가 pWDC-20(NRRL #B-21650)으로서 1996년 12월 6일에 기탁되었다.

서열 목록의 설명

서열 1: 아라비돕시스 탈리아나(*Arabidopsis thaliana*) 프로톡스-1 단백질에 대한 DNA 암호화 서열.

서열 2: 서열 1에 의해 암호화되는 아라비돕시스 프로톡스-1 아미노산 서열.

서열 3: 아라비돕시스 탈리아나 프로톡스-2 단백질에 대한 DNA 암호화 서열.

서열 4: 서열 3에 의해 암호화되는 아라비돕시스 프로톡스-2 아미노산 서열.

서열 5: 옥수수 프로톡스-1 단백질에 대한 DNA 암호화 서열.

서열 6: 서열 5에 의해 암호화되는 옥수수 프로톡스-1 아미노산 서열.

서열 7: 옥수수 프로톡스-2 단백질에 대한 DNA 암호화 서열.

서열 8: 서열 7에 의해 암호화되는 옥수수 프로톡스-2 아미노산 서열.

서열 9: 밀 프로톡스-1 단백질에 대한 DNA 암호화 서열.

서열 10: 서열 9에 의해 암호화되는 밀 프로톡스-1 아미노산 서열.

서열 11: 대두 프로톡스-1 단백질에 대한 DNA 암호화 서열.

서열 12: 서열 11에 의해 암호화되는 대두 프로톡스-1 단백질.

서열 13: 아라비돕시스 탈리아나 프로톡스-1 유전자로부터의 프로모터 서열.

서열 14: 옥수수 프로톡스-1 유전자로부터의 프로모터 서열.

서열 15: 목화 프로토스-1 단백질에 대한 DNA 암호화 서열.

서열 16: 서열 15에 의해 암호화되는 목화 프로토스-1 아미노산 서열.

서열 17: 사탕무 프로토스-1 단백질에 대한 DNA 암호화 서열.

서열 18: 서열 17에 의해 암호화되는 사탕무 프로토스-1 아미노산 서열.

서열 19: 평지 프로토스-1 단백질에 대한 DNA 암호화 서열.

서열 20: 서열 19에 의해 암호화되는 평지 프로토스-1 아미노산 서열.

서열 21: 벼 프로토스-1 단백질에 대한 DNA 암호화 서열.

서열 22: 서열 21에 의해 암호화되는 벼 프로토스-1 아미노산 서열.

서열 23: 수수 프로토스-1 단백질에 대한 DNA 암호화 서열.

서열 24: 서열 23에 의해 암호화되는 수수 프로토스-1 아미노산 서열.

서열 25: 옥수수 프로토스-1 인트론 서열.

서열 26: 사탕무 프로토스-1 유전자로부터의 프로모터 서열.

정 의

본원에서 사용된 바와 같이, "식물 프로토스 프로모터"는 식물에서 프로토포르피리노겐 옥시다제(프로토스) 암호화 서열의 바로 상부에서 자연적으로 발생하고, 본래 발생한 상태에서 관련된 프로토스 암호화 서열의 전사 조절에 관여하는 조절 영역을 언급하는데 사용된다. 식물 프로토스 프로모터에는 전사를 개시하는 RNA 폴리머라제의 결합에 직접적으로 관련되는 DNA 영역 및 추가로 작동가능하게 연결된 암호화 서열의 전사에 영향을 미치는 상부 조절-시스 요소가 포함된다.

본원에서 사용된 바와 같이 "유전자"는 (1) 암호화 서열 및 (2) 적합한 숙주 세포에서 암호화 서열의 전사를 촉진하고 조절하는 관련된 조절 영역을 포함하는 DNA 분자를 언급하는데 사용된다. 암호화 서열은 유용한 전사체(예를 들어 안티센스 RNA) 또는 암호화된 전사체의 해독에 의해 생성되는 폴리펩타이드를 암호화할 수 있다. 유전자는 5' 내지 3' 방향으로 최소한 프로모터 영역, 암호화 서열 및 전사 터미네이터를 포함한다. 또한, 유전자는 최소 요소(예: 암호화 서열내의 리더 또는 시그널 펩타이드)의 부분 또는 구별된 요소(예: 인트론)의 일부로서 발생할 수 있는 추가의 조절 영역을 포함할 수 있다.

본원에서 사용된 바와 같이 "키메라 유전자"는 하나 이상의 성분 부분이 다른 성분 부분과 비교하여 이종성인 자연적으로 발생하지 않는 유전자를 언급하는데 사용된다. 본 발명을 기술하기 위해 본원에서 사용된 바와 같이, "키메라 유전자"는 이종 암호화 서열과 작동가능하게 연결된 본 발명의 프로모터를 포함하는 유전자로 언급된다.

프로모터 및 암호화 서열 사이의 관계에 관해 본원에서 사용된 바와 같이, 용어 "이종"은 자연적으로 발생하지 않은 관계를 언급하는데 사용된다. 예를 들어, 암호화 서열이 프로모터 서열과 연관되어 자연적으로 발생하는 암호화 서열과 상이한 경우 프로모터 서열에 대하여 이종성인 것으로 간주한다. 이는 목적하는 프로모터와 자연적으로 결합된 암호화 서열의 변형된 형태를 포함한다. 따라서, 변형된, 억제제-내성 프로토스 암호화 서열은 비변형된, 억제제-민감성 형태의 암호화 서열과 자연적으로 결합된 프로모터에 대해 이종성인 것으로 간주된다. 이는 추가로 상이한 식물 또는 비-식물 종으로부터의 암호화 서열과 작동가능하게 연결된 본 발명의 프로모터를 포함한다.

본원에서 사용된 바와 같이, 용어 "실질적인 서열 상동성"은 기타의 뉴클레오타이드 또는 아미노산 서열을 갖는 실질적인 구조적 등과 및 기능적 등가를 나타내는 뉴클레오타이드 서열(DNA 또는 RNA의 경우) 또는 아미노산 서열(단백질 또는 폴리펩타이드의 경우)을 지시하는데 사용된다. 임의의 실질적인 서열 상동성을 갖는 서열 사이의 기능적 또는 구조적 차이는 최소일 것이다; 즉, 이들이 본 발명에서 지시되는 바와 같이 기능하는 서열의 능력에 영향을 미치지 않는다는 것이다. 예를 들어, 식물 프로토스 프로모터로 밝혀진 DNA 서열과 실질적으로 서열 상동성을 갖는 서열은 식물 프로토스 프로모터로서 관련된 DNA 서열이 동일한 수준 및 형태로 발현되도록 지시할 수 있다. 본원에서 밝혀진 서열과 실질적인 서열 상동성을 갖는 서열은 일반적으로 돌연변이와 같이 밝혀진 서열의 변이체이지만 또한 합성 서열일 수도 있다. 2개 이상의 서열 사이에 상당량의 서열이 중복되거나 유사한 경우, 또는 상이한 서열이 유사한 물리적 특성을 나타낼 경우, 구조적 차이는 최소인 것으로 간주된다. 상기한 특성에는 예를 들어 면역학적 반응성, 효소 활성 및 구조적 단백질 무결성등이 포함된다.

2개의 뉴클레오타이드 서열이 70%이상, 보다 바람직하게는 80%이상 및 가장 바람직하게는 90%이상의 서열 유사성을 갖는다면, 2개의 뉴클레오타이드 서열은 실질적인 서열 상동성을 가질 수 있다. 2개의 아미노산 서열이 폴리펩타이드의 활성화 부분 사이에 50%이상, 바람직하게는 70%이상 및 가장 바람직하게는 90%이상의 서열 유사성을 갖는다면, 2개의 아미노산 서열은 실질적인 서열 상동성을 가질 수 있다. 프로모터 DNA 서열의 경우, "실질적인 서열 상동성"은 또한 관련된 DNA 서열의 발현을 촉진하도록 작동될 수 있는 프로모터 DNA 서열 단편을 언급한다. 프로모터 DNA 서열의 작동가능한 상기의 단편은, 예를 들어 제한 효소를 사용하여 프로모터 DNA 서열을 절단하고, 프로모터 DNA 서열의 서열에 따라 합성함으로써 프로모터 DNA 서열로부터 유도하거나 PCR 기술을 사용하여 수득할 수 있다[참조: Mullis et al., Meth. Enzymol., 155: 335-350 (1987); Erlich (ed.), PCR Technology, Stockton Press (New York 1989)].

프로모터 DNA 서열이 제2 DNA 서열의 전사 및 해독에 영향을 미치도록 2개의 서열이 붙어있을 경우, 프로모터 DNA 서열은 제2 DNA 서열과 "작동가능하게 연결"되었다고 한다. 예를 들어, 제2 DNA 서열이 단백질 생성을 암호화하는 경우, 프로모터 DNA 서열이 제2 DNA 서열의 단백질 생성물을 발현시키는데 영향을 미친다면 프로모터 DNA 서열은 제2 DNA 서열에 작동가능하게 연결된 것이다. 예를 들어, 동일한 키메라 구조내의 암호화 DNA 서열에 물리적으로 부착된 프로모터 DNA 서열을 포함하는 DNA 서열에서 2개의 서열은 작

동가능하게 연결된 것과 같다.

발명의 상세한 설명

본 발명은 식물 프로토포르피리노겐 옥시다제(본원에서 "프로텍스"로 언급)에 대한 암호화 서열과 자연적으로 결합된 프로모터 DNA 서열에 관한 것이다[참조 문헌: 전문이 참조로서 인용된 WO 95/34659로서 1995년 12월 21일 공개된 1995년 6월 8일자 국제 출원 번호 PCT/IB95/00452; 동일자로 출원되고 본원에서 전문이 참조로서 인용된 "DNA Molecules Encoding Plant Protoporphyrinogen Oxidase and Inhibitor Resistant Mutants Thereof"란 명칭의 공동-계류중인 국제 출원 번호 (참고 번호 PH/520757/P1/CGC1847)]. 이들 프로텍스 프로모터 서열은 식물에서 이중 암호화 서열의 발현에 유용한 것으로 밝혀졌다.

아라비도시스 탈리아나 프로텍스-1 암호화 서열(서열 1)에 대한 프로모터 서열은 서열 13으로 제공된다. 실시예 1에서 기술한 프로브와 같이 관련된 암호화 서열을 사용하여 게놈 라이브러리로부터 상기 프로모터를 분리한다. 옥수수 프로텍스-1 암호화 서열(서열 5)에 대한 프로모터 서열은 서열 14로 제공된다. 실시예 4에서 기술한 프로브와 같이 관련된 암호화 서열을 사용하여 게놈 라이브러리로부터 상기 프로모터를 분리한다. 사탕무 프로텍스-1 암호화 서열(서열 17)에 대한 프로모터 서열은 서열 26으로 제공된다. 실시예 11에서 기술한 프로브와 같이 관련된 암호화 서열을 사용하여 게놈 라이브러리로부터 상기 프로모터를 분리한다.

본원에서 제공된 정보를 기초로, 아라비도시스 및 옥수수 프로텍스-1 프로모터를 분리하는데 사용된 방법을 모든 식물의 프로텍스 유전자로부터 프로모터 서열을 분리하는데 사용할 수 있다. 흥미로운 프로모터와 관련된 프로텍스 암호화 서열에 하이브리드화되는 충분한 상동성을 공유하는 어떤 프로텍스 암호화 서열도 상기 방법에서 프로브로 사용할 수 있다. 모든 식물로부터의 개개의 프로텍스-1 및 프로텍스-2 암호화 서열이 필수적인 상동성 정도를 공유하도록 계획되기 때문에, 프로브로 사용되는 프로텍스 암호화 서열을 선택하는 것이 그리 중요하게 고려되지 않는다. 그러나, 최적의 하이브리드화 결과를 위해, 가장 밀접하게 관련된 프로텍스 암호화 서열을 사용하는 것이 바람직하다. 가장 바람직하게는, 프로브로 사용되는 암호화 서열은 흥미로운 프로텍스 프로모터로서 동종 식물로부터 유래되고, 프로모터와 자연적으로 결합된 암호화 서열이다.

따라서, 본 발명은 식물 프로토포르피리노겐 옥시다제에 대한 암호화 서열과 자연적으로 결합된 분리된 프로모터 DNA 분자에 관한 것이다. 바람직하게는, 아라비도시스, 사탕수수, 대두, 보리, 목화, 담배, 사탕무, 팥지, 옥수수, 밀, 수수, 호밀, 귀리, 잔디 및 마초, 기장 및 벼로 이루어진 그룹으로부터 선택된 식물의 식물 프로토포르피리노겐 옥시다제에 대한 암호화 서열과 자연적으로 결합된 분리된 프로모터 DNA 분자이다. 보다 바람직하게는, 아라비도시스, 대두, 목화, 담배, 사탕무, 팥지, 옥수수, 밀, 수수, 호밀, 귀리, 잔디 및 벼로 이루어진 그룹으로부터 선택된 식물의 식물 프로토포르피리노겐 옥시다제에 대한 암호화 서열과 자연적으로 결합된 분리된 프로모터 DNA 분자이다. 가장 바람직하게는, 아라비도시스의 식물 프로토포르피리노겐 옥시다제에 대한 암호화 서열과 자연적으로 결합된 분리된 프로모터 DNA 분자이다. 가장 바람직하게는, 옥수수의 식물 프로토포르피리노겐 옥시다제에 대한 암호화 서열과 자연적으로 결합된 분리된 프로모터 DNA 분자이다. 가장 바람직하게는, 사탕무의 식물 프로토포르피리노겐 옥시다제에 대한 암호화 서열과 자연적으로 결합된 분리된 프로모터 DNA 분자이다.

본 발명에는, 본원에서 정의한 바와 같이 본 발명에 따른 DNA 분자, 바람직하게는 약간 엄격한 조건하에서 10 뉴클레오타이드 이상인 당해 프로텍스 프로모터 서열의 인접 부분을 포함하는 당해 DNA 분자로부터 수득할 수 있는 올리고뉴클레오타이드 프로브와 하이브리드화되는 DNA 분자가 포함된다. 가장 바람직하게는, 다음의 조건하에서, 서열 13(아라비도시스 프로텍스-1 프로모터), 서열 14(옥수수 프로텍스-1 프로모터) 또는 서열 26(사탕무 프로텍스-1 프로모터)의 뉴클레오타이드 서열과 하이브리드화되는 DNA 분자이다:

(a) 50°C에서 7% 나트륨 도데실 설페이드(SDS), 0.5M NaPO₄ pH 7.0, 1mM EDTA중에 하이브리드화하고,

(b) 2X SSC, 1% SDS로 50°C에서 세척한다.

하이브리드의 안정성에 영향을 미치는 인자는 하이브리드화의 엄격함을 결정한다. 이러한 인자로서 DNA PROBE[George H. Keller and Mark M. Manak, Macmillan Publishers Ltd, 1993, Section one: Molecular Hybridization Technology; Page 8ff]에서 제공되는 식에 따라 용이하게 계산할 수 있는 융점 T_m이 있다. 바람직한 하이브리드화 온도는 계산된 융점 T_m이 약 25°C 이하, 바람직하게는 약 12 내지 15°C 이하이고, 올리고뉴클레오타이드의 경우에는 계산된 융점 T_m이 약 5 내지 10°C 이하이다.

본 발명의 추가 양태는, (a) 식물 프로텍스 유전자 또는 mRAN와 특이적으로 하이브리드화될 수 있는, 프로텍스 단백질 또는 식물 프로텍스 프로모터 서열에 대한 10 뉴클레오타이드 이상의 서열의 인접 부분을 포함하는 뉴클레오타이드 프로브를 제조하는 단계, (b) 클로닝된 게놈 DNA 단편 또는 선택된 생물의 cDNA 단편의 모집단에서 단계 (a)에 따라 제조된 뉴클레오타이드 프로브를 사용하여 다른 프로텍스 암호화 서열에 대하여 프로빙하는 단계 및 (c) 프로텍스 프로모터 서열을 함유하는 DNA 부분 및 프로텍스 단백질을 암호화하는 DNA 부분을 포함하는 DNA 분자를 분리하고 증폭시키는 단계를 포함하여, 프로텍스 프로모터 서열을 포함하는 DNA 부분 및 프로텍스 단백질을 암호화하는 DNA 부분을 포함하는 DNA 분자를 생성하는 방법을 제공하는 것이다.

본 발명의 추가 양태는, (a) 식물 프로텍스 유전자 또는 mRAN와 특이적으로 하이브리드화될 수 있는, 식물 프로텍스 단백질에 대한 10 뉴클레오타이드 이상의 암호화 서열의 인접 부분을 포함한 뉴클레오타이드 프로브를 제조하는 단계, (b) 클로닝된 게놈 DNA 단편 또는 선택된 생물의 cDNA 단편의 모집단에서 단계 (a)에 따라 제조된 뉴클레오타이드 프로브를 사용하여 다른 프로텍스 암호화 서열 또는 프로텍스 프로모터 서열에 대하여 프로빙하는 단계 및 (c) 프로텍스 프로모터 서열을 함유하는 DNA 부분을 포함하는 DNA

분자를 분리하고 증폭시키는 단계를 포함하여, 프로텍스 프로모터 서열을 포함하는 DNA 부분을 포함하는 DNA 분자를 생성하는 방법을 제공하는 것이다.

본 발명의 추가 양태는 (a) 식물 프로텍스 유전자 또는 mRNA와 특이적으로 하이브리드화될 수 있는, 식물 프로텍스 단백질 또는 프로텍스 프로모터 서열에 대한 10 뉴클레오타이드 이상의 암호화 서열의 인접 부분을 포함한 뉴클레오타이드 프로브를 제조하는 단계, (b) 클로닝된 게놈 DNA 단편 또는 선택된 생물의 cDNA 단편의 모집단에서, 단계 (a)에 따라 제조된 뉴클레오타이드 프로브를 사용하여 다른 프로텍스 암호화 서열 또는 프로텍스 프로모터 서열에 대하여 프로빙하는 단계 및 (c) 프로텍스 프로모터 서열을 포함하는 DNA 부분을 포함하는 DNA 분자를 분리하고 증폭시키는 단계를 포함하여, 어떤 식물의 프로텍스 유전자로부터 프로텍스 프로모터 서열을 포함하는 DNA 부분을 포함하는 DNA 분자를 분리하는 방법을 제공하는 것이다.

본 발명은 추가로, 프로텍스 유전자의 암호화 영역 또는 프로모터 영역으로부터 수득할 수 있는, 폴리머라제 연쇄 반응(PCR)에서 10 뉴클레오타이드 이상의 식물 프로텍스 유전자 또는 mRNA와 특이적으로 하이브리드화될 수 있는 뉴클레오타이드 프로브의 용도를 포함한다.

본 발명은 추가로, 게놈 프로텍스 서열에 대한 프로브의 선택적인 하이브리드화에 기초한 표준 기술을 사용하여, 식물 프로텍스 유전자 또는 선택된 식물 게놈에서 프로텍스 유전자의 위치 지도에 대해 특이적으로 하이브리드화할 수 있는 뉴클레오타이드 프로브의 용도를 포함한다.

본 발명은 프로브로서 흥미로운 프로모터와 결합된 프로텍스 암호화 서열과 하이브리드화되기에 충분한 상동성을 공유하는 프로텍스 암호화 서열의 용도를 포함한다. 바람직하게는, 프로브로 사용된 암호화 서열이 흥미로운 프로텍스 프로모터와 동종 식물의 것이고 프로모터와 자연적으로 결합된 암호화 서열인 프로텍스 암호화 서열의 용도를 포함한다.

본 발명의 식물 프로텍스 프로모터에는 서열 13에 기술된 아라비도시스 프로텍스-1 프로모터 서열, 서열 14에 기술된 제아 메이즈(Zea mays, 옥수수) 프로텍스-1 프로모터 서열, 서열 26에 기술된 사탕무 프로텍스-1 프로모터 서열뿐만 아니라 상기한 모든 식물종으로부터 수득할 수 있는 상응하는 프로텍스-1 프로모터 서열이 포함된다. 또한, 본 발명은 예시된 프로텍스 프로모터 서열과 동일한 방법으로 작동가능하게 연결된 암호화 서열의 발현을 조절하는 능력을 보유한 DNA 서열의 기능적 단편들을 포함한다. 이런 기능적 단편은 당해 기술 분야에서 프로텍스 프로모터 활성을 확인하는데 사용되는 결실 분석 또는 기타 표준 기술로 확인될 수 있다[참조: Lewin의 Gene V(Oxford Univ. Press) page 546-549, 1990]. 또한, 본 발명은 작동가능하게 연결된 서열상에 동일한 발현 수준 및 발현 패턴을 부여하는, 식물 유전자로부터 수득할 수 있는 프로텍스 프로모터와 실질적인 서열 상동성을 갖는 DNA 서열을 포함한다. 상기 프로모터 서열은 식물 유전자로부터 분리된 프로텍스 프로모터의 변형을 통해 수득할 수 있고, 식물 프로텍스 프로모터의 기능상 동등한 유도체로 생각된다.

하기의 실시예에서 설명한 바와 같이, 본 발명의 DNA 서열, 벡터 및 유전자이식(transgenic) 식물은 식물 프로텍스 유전자로부터 유도된 프로모터 서열을 포함한다. 프로텍스 프로모터 DNA 서열은 바람직하게는 암호화 DNA 서열, 예를 들어 안티센스 전사체와 같이 유용한 RNA 전사체로 전사될 수 있는 DNA 서열 또는 유용한 단백질 생성물의 생산시 최종적으로 발현되는 암호화 서열에 작동가능하게 연결된다.

바람직한 양태에서, 프로텍스 프로모터는 상응하는 비변형된 형태의 효소를 억제하는 일정 농도의 제초제에 대해 내성이 있는 변형된 제초제 표적 효소의 발현을 유도하는데 사용된다. 따라서, 본 발명은 제초제에 대한 내성을 부여하는 제초제-내성 형태의 식물의 제초제 표적 단백질을 발현시키기 위한 프로텍스 프로모터의 용도에 관한 것이다. 이와 같이 변형된 제초제-내성 효소에는 이미다졸글리세롤 포스페이트 데하이드라타제(IGPD; 1994년 11월 24일 공개된 WO 9426909 참조), EPSP 신타제(미국 특허 제4,535,060호; 제4,769,061호, 제4,940,835호 및 EP 550,633 참조), 글루타민 신세타제(GS; 미국 특허 제4,975,374호 참조), 아세틸 조효소 A 카복실라제(ACCCase; 미국 특허 제5,162,602호 참조) 및 아세토락테이트 신타제(미국 특허 제4,761,373호, 제5,304,732호, 제5,331,107호, 제5,013,659호, 제5,141,870호 및 제5,378,824호 참조)가 포함된다. 가장 바람직한 양태에서, 프로텍스 프로모터는 실시예 2 내지 3에서 설명하는 바와 같이 프로텍스 억제제에 대해 내성이 있는 변형된 프로텍스 효소의 발현을 유도하는데 사용된다[참조: 본원에서 관계된 부분이 참조로서 인용된, 1995년 6월 8일 출원되고 W095/34659로서 1995년 12월 21일 공개된 국제 출원 번호 PCT/IB95/00452; 본 출원과 동일자로 제출된 "DNA Molecules Encoding Plant Protoporphyrinogen Oxidase and Inhibitor Resistant Mutants Thereof"이란 명칭으로 공동 계류중임].

본 발명은 이중 DNA 암호화 서열과 작동가능하게 연결된 식물 프로텍스 프로모터를 포함하는 발현 카세트를 포함하는 키메라 유전자에 관한 것이다. 바람직하게는, 키메라 유전자에서 식물 프로텍스 프로모터가 프로텍스-1 유전자 또는 프로텍스-2 유전자의 것이다. 특히 바람직하게는, 키메라 유전자에서 식물 프로텍스 프로모터는 프로텍스-1 유전자의 것이다. 특히 바람직하게는, 키메라 유전자에서 식물 프로텍스 프로모터는 프로텍스-2 유전자의 것이다.

식물 프로텍스 프로모터가 아라비도시스, 사탕수수, 대두, 보리, 목화, 담배, 사탕무, 평지, 옥수수, 밀, 수수, 호밀, 귀리, 잔디 및 마초, 기장 및 벼로 이루어진 그룹으로부터 선택된 식물로부터 유래하는 키메라 유전자가 바람직하다. 식물 프로텍스 프로모터가 아라비도시스, 대두, 목화, 담배, 사탕무, 평지, 옥수수, 밀, 수수, 호밀, 귀리, 잔디 및 벼로 이루어진 그룹으로부터 선택된 식물로부터 유래하는 키메라 유전자가 바람직하다. 식물 프로텍스 프로모터가 아라비도시스, 옥수수 및 사탕무로 이루어진 그룹으로부터 선택된 식물로부터 유래하는 키메라 유전자가 특히 바람직하다. 식물 프로텍스 프로모터가 아라비도시스 및 옥수수로 이루어진 그룹으로부터 선택된 식물로부터 유래하는 키메라 유전자가 보다 바람직하다. 식물 프로텍스 프로모터가 서열 13에 기술된 서열을 갖는 키메라 유전자가 가장 바람직하다. 식물 프로텍스 프로모터가 서열 14에 기술된 서열을 갖는 키메라 유전자가 가장 바람직하다. 식물 프로텍스 프로모터가 서열 26에 기술된 서열을 갖는 키메라 유전자가 가장 바람직하다. 바람직하게는, 키메라 유전자에서 식물 프로텍스 프로모터가 500 뉴클레오타이드 이상, 보다 바람직하게는 300 뉴클레오타이드 이상이다.

DNA 분자가 프로토क्स-1 활성 또는 프로토क्स-2 활성을 갖는 아라비돕시스 종의 단백질을 암호화하고, 단백질이 바람직하게는 서열 2 또는 서열 4에 기술된 아미노산 서열을 포함하는 키메라 유전자가 바람직하다. 또한, DNA 분자가 프로토क्स-1 활성 또는 프로토क्स-2 활성을 갖는 옥수수의 단백질을 암호화하고, 단백질이 바람직하게는 서열 6 또는 서열 8에 기술된 아미노산 서열을 포함하는 키메라 유전자가 바람직하다. 또한, DNA 분자가 프로토क्स-1 활성을 갖는 밀의 단백질을 암호화하고, 단백질이 바람직하게는 서열 10에 기술된 아미노산 서열을 포함하는 키메라 유전자가 바람직하다. 또한, DNA 분자가 프로토क्स-1 활성을 갖는 대두 단백질을 암호화하고, 단백질이 바람직하게는 서열 12에 기술된 아미노산 서열을 포함하는 키메라 유전자가 바람직하다. 또한, DNA 분자가 프로토क्स-1 활성을 갖는 목화 단백질 암호화하고, 단백질이 바람직하게는 서열 16에 기술된 아미노산 서열을 포함하는 키메라 유전자가 바람직하다. 또한, DNA 분자가 프로토क्स-1 활성을 갖는 사탕무의 단백질을 암호화하고, 단백질이 바람직하게는 서열 18에 기술된 아미노산 서열을 포함하는 키메라 유전자가 바람직하다. 또한, DNA 분자가 프로토क्स-1 활성을 갖는 평지의 단백질을 암호화하고, 단백질이 바람직하게는 서열 20에 기술된 아미노산 서열을 포함하는 키메라 유전자가 바람직하다. 또한, DNA 분자가 프로토क्स-1 활성을 갖는 벼의 단백질을 암호화하고, 단백질이 바람직하게는 서열 22에 기술된 아미노산 서열을 포함하는 키메라 유전자가 바람직하다. 또한, DNA 분자가 프로토क्स-1 활성을 갖는 수수의 단백질을 암호화하고, 단백질이 바람직하게는 서열 24에 기술된 아미노산 서열을 포함하는 키메라 유전자가 바람직하다.

추가로 본 발명은, 상응하는 비변형된 형태의 효소를 억제하는 수준의 제초제에 대해 내성이 있는, 식물의 단백질을 암호화하는 DNA 분자와 작동가능하게 연결된 식물 프로토क्स 프로모터를 포함하는 발현 카세트 포함하는 키메라 유전자에 관한 것이다.

바람직하게는, 키메라 유전자에서 이중 암호화 서열은 식물 효소의 변형된 제초제-내성 형태를 암호화한다. 특히 바람직하게는, 키메라 유전자에서 식물 효소는 이미다졸글리세롤 포스페이트 데하이드라타제 (IGPD), 5-에놀피루빌시킴에이트-3-포스페이트 신타제(EPSP), 글루타민 신세타제(GS), 아세틸 조효소 A 카복실라제, 아세토락테이트 신타제, 히스티딘 데하이드로게나제 및 프로토포르피리노겐 옥시다제(프로토क्स)로 이루어진 그룹으로부터 선택된다. 보다 바람직하게는, 키메라 유전자에서 식물 효소는 이미다졸글리세롤 포스페이트 데하이드라타제(IGPD), 5-에놀피루빌시킴에이트-3-포스페이트 신타제(EPSP), 글루타민 신세타제(GS), 아세틸 조효소 A 카복실라제, 아세토락테이트 신타제 및 프로토포르피리노겐 옥시다제(프로토क्स)로 이루어진 그룹으로부터 선택된다.

특히 바람직하게는, 키메라 유전자에서 식물 효소는 진핵성 프로토क्स이다. 보다 바람직하게는, 키메라 유전자에서 식물 효소는 아미노산 치환체를 갖는 진핵성 프로토क्स이고, 아미노산 치환체는 프로토क्स 억제제에 대해 내성을 부여하는 성질을 갖는다. 가장 바람직하게는, 키메라 유전자에서 식물 효소는 프로토क्स 억제제에 내성을 부여하는 성질을 가지며, "DNA Molecules Encoding Plant Protoporphyrinogen Oxidase"이란 명칭의 공동계류중인 국제 출원에 따른 진핵성 프로토क्स이다.

바람직하게는, 키메라 유전자에서 DNA 분자는 아라비돕시스, 사탕수수, 대두, 보리, 목화, 담배, 사탕무, 평지, 옥수수, 밀, 수수, 호밀, 귀리, 잔디 및 마초, 기장, 사료 작물 및 벼로 이루어진 그룹으로부터 선택된 식물의 단백질을 암호화한다. 보다 바람직하게는, 키메라 유전자에서 DNA 분자는 아라비돕시스, 대두, 목화, 사탕무, 평지, 옥수수, 밀 및 수수로 이루어진 그룹으로부터 선택된 식물의 단백질을 암호화한다. 특히 바람직하게는, 키메라 유전자에서 DNA 분자는 아라비돕시스, 밀, 대두 및 옥수수로 이루어진 그룹으로부터 선택된 식물의 단백질을 암호화한다. 가장 바람직하게는, 키메라 유전자에서 DNA 분자는 대두 및 밀로 이루어진 그룹으로부터 선택된 식물의 단백질을 암호화한다.

추가로, 본 발명은 비변형된 식물 프로토क्स 단백질 억제제에 내성이 있는 제초제-내성 식물 프로토क्स 단백질을 발현시키는 본 발명의 키메라 유전자의 용도에 관한 것이다. 추가로, 본 발명은 당해 키메라 유전자를 숙주 계통내로 안정하게 통합시키는 것에 관한 것이다. 본 발명은 식물 프로토포르피리노겐 옥시다제(프로토क्स) 프로모터 또는 이와 기능상 동등한 유도체를 포함하는 재조합 DNA 분자에 관한 것이다. 추가로 본 발명은 상기한 재조합 DNA 분자를 포함하는 재조합 DNA 벡터에 관한 것이다.

본 발명의 추가 목적은, 당해 키메라 유전자를 포함하는, 식물, 식물 종자, 식물 조직 또는 식물 세포내로 안정하게 형질전환될 수 있는 재조합 벡터이다. 식물 및 이의 자손, 식물 종자, 식물 조직 또는 식물 세포는 비-식물원 또는 식물원, 바람직하게는 식물로부터 유래할 수 있는, 목적 단백질을 암호화하는 DNA 분자를 발현시킬 수 있는 벡터로서 안정하게 형질전환된다. 식물 및 이의 자손, 식물 종자, 식물 조직 또는 식물 세포는, 비-식물원 또는 식물원, 바람직하게는 상응하는 비변형된 형태의 효소를 억제시키는 수준의 제초제에 내성이 있는 식물로부터 유래할 수 있는, 목적 단백질을 암호화하는 DNA 분자를 발현시킬 수 있는 재조합 벡터로서 안정하게 형질전환되는 것이 바람직하다.

추가로 본 발명은, 본 발명에 따른 하나 이상의 키메라 유전자로 안정하게 형질전환된 식물 및 이의 자손, 종자 및 배양된 조직을 포함하는 유전자이식 식물 조직에 관한 것이다. 비-식물원 또는 식물원, 바람직하게는 식물 조직에서 상응하는 비변형된 형태의 효소를 억제시키는 농도의 제초제에 내성이 있는 식물로부터 유래할 수 있는, 단백질을 발현시킬 수 있는 DNA 암호화 서열과 작동가능하게 연결된 식물 프로토क्स 프로모터를 포함한 발현 카세트를 포함하는 하나 이상의 키메라 유전자로 안정하게 형질전환되는 식물, 종자 및 배양된 조직을 포함한 유전자이식 식물 조직이 바람직하다.

또한 본 발명은 본 발명에 따른 벡터로 안정하게 형질전환되고, 상기 DNA 분자를 발현시킬 수 있는 숙주를 포함한다. 숙주 세포는 식물 세포, 세균 세포, 효모 세포 및 곤충 세포로 이루어진 그룹으로부터 선택하는 것이 바람직하다.

추가로 본 발명은, 본원에서 교시한 바와 같이 변형된 억제제-내성 프로토क्स 효소를 발현시키는 유전자에 의해 내성이 부여되는, 상기한 식물에서 자연적으로 발생하는 프로토क्स 활성을 억제하는 제초제에 대해 내성이 있는 식물 및 이의 자손, 식물 조직 및 식물 종자에 관한 것이다. 대표적인 식물에는 제초제가 일반적으로 의도된 목적으로 적용될 수 있는 모든 식물이 포함된다. 농경학적으로 중요한 작물, 예를 들어 아라비돕시스, 대두, 사탕수수, 보리, 목화, 담배, 사탕무, 평지, 옥수수, 밀, 수수, 호밀, 귀리, 잔디 및 마초, 기장 및 벼 등과 같은 피자식물 및 나자식물이 바람직하다. 농경학적으로 중요한 작물, 예

를 들어 아라비도시스, 목화, 대두, 평지, 사탕무, 담배, 옥수수, 벼, 밀, 수수, 호밀, 귀리, 잔디 및 벼 등과 같은 피자식물 및 나자식물이 보다 바람직하다. 농경학적으로 중요한 작물, 예를 들어 아라비도시스, 대두, 목화, 사탕무, 평지, 옥수수, 밀, 수수 및 벼 등과 같은 피자식물 및 나자식물이 보다 바람직하다.

본 발명의 유전자이식 식물은 당해 기술 분야에서 공지된 어떤 형질전환 방법에 의해서도 형질전환될 수 있다. 이러한 방법에는, 예를 들어 직접 감염 또는 아그로박테리움 투메파시엔스(*Agrobacterium tumefaciens*)과 식물, 식물 조직 또는 세포의 공동-배양[참조: Horsch et al., *Science*, 225: 1229 (1985); Marton, "Cell Culture and Somatic Cell Genetic of plants", vol. 1, pp. 514-521 (1984)]; 원형질체로 직접 유전자 전달[참조: Paszkowski et al., *EMBO J.* 12: 2717 (1984); Loerz et al., *Mol. Gen. & Genet.* 1199:178 (1985); Fromm et al., *Nature* 319:719 (1986)]; 미세추진 충격(microprojectile bombardment)[참조: Klein et al., *Bio/Technology*, 6:559-563 (1988)]; 원형질체 배양된 세포 또는 조직에의 주입[참조: Reich et al., *Bio/technology*, 4:1001-1004 (1986)]; 또는 문헌[참조: De La Pena et al., *Nature*, 325:274-276 (1987); Hooykaas-Van Slogteren et al., *Nature*, 311:763-764 (1984); Grimsley et al., *Bio/Technology*, 6:185 (1988) 및 Grimsley et al., *Nature*, 325:117 (1988)]에 기술된 바와 같이 종자 또는 식물의 분열 조직내로의 주입에 의한 형질전환이 포함된다.

상기한 바와 같이 유전자이식 종자 및 식물내로 유전자 조작된 유전성은 생식 생장 또는 영양 생장에 의해 전달되어 식물의 자손에서 유지되고 증식된다. 일반적으로 상기한 유지 및 증식은 경작, 씨뿌림 또는 수확과 같은 특정 목적에 맞도록 개발된 공지된 농업 방법을 이용한다. 또한, 수경법 또는 온실 기술과 같은 특별한 방법을 적용할 수 있다. 생장하는 작물은 곤충 또는 감염에 의해 공격과 손상을 받기 쉬울 뿐만 아니라 잡초와 경쟁해야 하기 때문에, 생산을 증진시키기 위해 잡초, 식물 질병, 곤충, 선충 및 기타 불리한 조건을 제어하기 위한 방법이 취해진다. 여기에는 경운, 잡초 또는 감염된 식물의 제거와 같은 기계적인 방법 뿐만 아니라, 제초제, 살진제, 배우자 불임 유기제, 살선충제, 생장 조절제, 숙성제 및 살충제와 같은 농화학품의 적용이 포함된다.

본 발명에 따른 유전자이식 식물 및 종자의 유리한 유전성은, 충해, 제초제 또는 스트레스에 대한 내성, 개선된 영양가, 증가된 수율, 또는 도복 및 흠어짐이 보다 적게 발생하는 개선된 구조와 같이 개선된 성질을 갖는 식물을 개발할 목적으로 식물 육종에 이용할 수 있다. 각종 육종 단계는 교배시킬 계통의 선택, 모 계통의 직접적인 수분 또는 적절한 자손 식물의 선택과 같이 사람에게 의한 공지된 조정에 의해 특성화된다. 관련 기술은 당해 기술 분야에서는 공지된 것이고, 하이브리드화, 동계 교배, 여교배, 다계통 교배, 각종 융합, 중간 하이브리드화 및 이수화 기술 등을 포함하지만 이에 한정되지는 않는다. 또한, 하이브리드화 기술은 기계적, 화학적 또는 생화학적 수단에 의해 웅성 불임 또는 자성 불임 식물을 생성하는 식물의 불임화를 포함한다. 상이한 계통의 꽃가루로 웅성 불임 식물을 교배시키는 수분 작용은 자성 불임 식물을 제외한 웅성 불임 식물의 게놈이 양쪽 부모 계통의 성질은 일정하게 획득하는 것을 보장한다. 따라서, 본 발명에 따른 유전자이식 종자 및 식물은, 제초제 또는 살충제 처리와 같은 통상적인 방법의 유효성을 증가시키거나 이들의 변형된 유전성에 기인하여 상기 방법이 필요없는 개선된 식물 계통의 번식에 사용될 수 있다. 달리는, 최적화된 유전적 "장치"에 기인하여 수득할 수 있는 개선된 스트레스-내성의 신규한 작물은, 비교가능한 반대의 개발 조건에서 내성이 없는 생성물보다 고품질로 수확된 생성물을 생산한다.

경작자에 의해 수확되고 판매된 종자의 발아성 및 균질성은 중요하지 않지만, 종자 생산에 있어서 종자의 발아성 및 균질성은 본질적인 생산 특성이다. 종자 전염병을 억제하고 훌륭한 발아성을 갖는 종자를 생산하기 위해 곡물을 다른 곡물 및 잡초로부터 보호하는 것은 어렵지만, 정당한 가격의 확실한 종자 생산 실례가 순수한 종자의 생장, 상태 조절 및 판매 분야에 경험있는 종자 생산자에 의해 개발되고 있다. 따라서, 경작자 자신의 작물로부터 수확된 종자를 사용하는 대신 특별한 품질 기준에 부합되는 증명된 종자를 구입하는 것이 일반적이다. 종자로서 사용된 증식 물질은, 통상적으로 제초제, 살충제, 살진제, 살균제, 살선충제, 의연체 동물 억제제 또는 이의 혼합물을 포함한 보호 피복제로 처리한다. 통상적으로 사용되는 보호 피복제에는, 캡탄, 카복신, 티람(TMD^R), 메탈락실(Apron^R) 및 피리미포스-메틸(Actellic^R)과 같은 화합물이 포함된다. 필요하다면, 이들 화합물은, 세균, 곰팡이 또는 동물 기생충에 의해 야기되는 피해에 대한 보호를 제공하는 제형화 분야에서 통상적으로 사용되는 추가의 담체, 계면활성제 또는 적용-촉진 보조제와 함께 제형화된다. 보호 피복제는, 액체 제형을 증식 물질에 침지시키거나 배합된 습윤 또는 무수 제형으로 피복시킴으로써 적용할 수 있다. 또한, 싹 또는 열매에 직접적인 처리와 같은 다른 적용 방법이 가능하다.

본 발명의 추가 양태는, 본 발명에 따른 유전자이식 식물, 유전자이식 식물 물질 또는 유전자이식 종자의 사용으로 특징화되는, 상기 예시한 방법과 같은 신규한 농업적 방법을 제공하는 것이다. 본 발명은, 본 발명의 키메라 유전자가 제초제에 대한 내성을 부여하는 식물의 제초제 표적 단백질의 제초제-내성 형태를 발현시키기에 충분한 양으로 포함되는 유전자이식 식물 또는 이의 자손을 사용하는 농업적 방법에 관한 것이다.

본 발명의 방법에 따라 형질전환된 식물의 자손을 번식시키는데 있어서, 다음과 같은 방법을 사용할 수 있다: 하기에 기술할 실시예에 기술되어 있는 바와 같이 생산된 옥수수 식물을 당해 기술 분야에서 공지된 바와 같이 온실 또는 토양중의 포트에서 성장시키고 개화시킨다; 성숙한 수술로부터 꽃가루를 수득하고, 이를 동일한 식물, 동기 식물 또는 다른 바람직한 옥수수의 이삭을 수분하는데 사용한다. 유사하게, 형질전환된 식물에서 개발된 이삭은 동일한 식물, 동기 식물 또는 다른 바람직한 옥수수 식물로부터 수득할 수 있는 꽃가루로 수분할 수 있다. 상기 방법으로 수득한 형질전환된 자손은 도입된 유전자(들) 및/또는 관련된 DNA(유전형) 또는 부여된 표현형에 의해 비-형질전환된 자손과 구별할 수 있다. 유사하게 형질전환된 자손은, 바람직한 특성을 갖는 임의의 식물이 정상적으로 수행하는 것처럼, 자가 수정하거나 다른 식물과 교배될 수 있다. 유사하게, 상기 방법에 의해 생산된 담배 또는 다른 형질전환된 식물은 바람직한 특성을 갖는 자손을 생산하기 위해 당해 기술 분야에서 공지된 바와 같이 자가 수정하거나 교배될 수 있다. 유사하게, 당해 기술 분야에서 공지된 방법 및 본 발명의 방법의 조합에 의해 생산된 다른 유전자이식 생물은 바람직한 특성을 갖는 자손을 생산하기 위해 당해 기술 분야에서 공지된 방법으로서

번식시킬 수 있다.

본 발명은, 하기에 기술될 실시예에 어떤 제한적 의미를 주지 않으면서 하기 실시예에서 보다 자세하게 설명된다.

실시예

실시예 1

아라비돕시스 탈리아나 프로토क्स-1 프로모터 서열의 분리

아라비돕시스 탈리아나(콜롬비아, 식물 전체)로부터 제조된 람다 캡(Lamda Zap) II 게놈 DNA 라이브러리를 스트라타진(Stratagene)사로부터 구입한다. 약 125,000개의 파지를 15cm의 페트리 디쉬당 25,000pfu (플라크 형성 단위)의 밀도로 플레이팅하고, 콜로니/플라크 스크린 막(NEN Dupont사)상에 복제물을 올린다. 플라크 복제물은 랜덤 프라이밍 방법(random priming method, Life Technologies사)에 의해 ³²P-dCTP로 표지된 아라비돕시스 프로토क्स-1 cDNA(서열 1)로 프로빙한다. 문헌[참조: Church and Gilbert, Proc. Natl. Acad. Sci. USA 81: 1991-1995 (1984)]에 기술된 바와 같이 하이브리드화 및 세척 조건을 65°C로 한다. 양성으로 하이브리드화된 플라크를 정제하고, 생체내에서 pBluescript 플라스미드내로 잘라낸다. 게놈 DNA 삽입체로부터의 서열은 형광 염료(Applied Biosystems, Inc.)로 표지된 디데옥시 터미네이터를 사용한 연쇄 종결 방법에 의해 측정한다. 한가지 클론인 AraPT1Pro는 프로토क्स-1 단백질 암호화 서열의 개시 메티오닌(ATG)으로부터 상부의 아라비돕시스 서열 580bp를 포함하는 것으로 측정되었다. 또한, 이 클론은 프로토क्स-1 cDNA의 1241bp로 확장된 암호화 서열 및 인트론을 포함한다. 5'의 비암호화 단편 580bp는 아라비돕시스 프로토क्स-1 프로모터로 추정되고, 서열은 서열 13에 기술되어 있다.

AraPT1Pro는 1995년 12월 14일 pWDC-11로서 기탁되었다(NRRL #B-21515).

실시예 2

천연 아라비돕시스 프로토क्स-1 프로모터하에서 변화된 프로토क्स-1 유전자를 발현시키는 식물 형질전환 벡터의 작제

변화된 적절한 아라비돕시스 프로토क्स-1 cDNA의 전체 길이의 cDNA를 EcoRI-XhoI의 부분적 절단 단편으로서 분리하고, 식물 발현 벡터 pCGN1761ENX(1995년 6월 8일 출원되고 WO 95/34659로서 1995년 12월 21일 공개된 PCT/IB 95/00452의 실시예 9를 참조)로 클로닝한다. 이 플라스미드를 NcoI 및 BamHI으로 절단하여 전체 프로토क्स-1 cDNA와 아그로박테리움 투메파시엔스의 tml 유전자의 3' 비해독된 서열로부터의 전사 터미네이터를 포함한 단편을 생산한다. 상기한 AraPT1Pro 플라스미드를 NcoI 및 BamHI으로 절단하여 pBluescript 및 추정된 아라비돕시스 프로토क्स-1 프로모터 580bp를 포함하는 단편을 생산한다. 이들 두 단편을 접합하여 천연 프로토क्स 프로모터에 융합된 변화된 프로토क्स cDNA를 생산한다. 프로토क्स-1 프로모터/프로토क्स-1 cDNA/tml 터미네이터 융합체를 포함한 발현 카세트를 KpnI으로 절단하고, 한 쌍의 pCIB200 벡터내로 클로닝한다. 한 쌍의 플라스미드를 전기천공에 의해 아그로박테리움내로 형질전환시킨 후 진공 침투 방법(Bechtold et al. C. R. Acad. Sci. Paris 316: 1194-1199 (1993))을 사용하여 아라비돕시스내로 형질전환시킨다. 변화된 프로토क्स 유전자를 발현하는 형질전환체를 카나마이신상에서 선택하거나, 다양한 농도의 프로토क्स 억제 제초제상에서 선택한다.

실시예 3

천연 프로토क्स-1 프로모터/변화된 프로토क्स-1 융합체의 발현에 의한 제초제-내성 식물의 생산

상기한 공정을 사용하여, 프로토क्स-1 서열(서열 1)에서 뉴클레오타이드 1306 내지 1308이 TAC에서 ATG(티로신에서 메티오닌)로의 변화를 포함한 아라비돕시스 프로토क्स-1 cDNA를 천연 프로토क्स-1 프로모터 단편과 융합시키고, 아라비돕시스 탈리아나내로 형질전환시킨다. 상기 변화된 프로토क्स-1 효소(AraC-2Met)는 세균 발현 시스템에서 시험할 경우, 자연적으로 발생하는 효소보다 각종 프로토क्स-억제 제초제에 대해 10배 이상의 내성을 보여준다[본 출원과 동일자로 제출된 "DNA Molecules Encoding Plant Protoporphyrinogen Oxidase and Inhibitors Resistant Mutant Thereof"란 명칭의 공동 계류중인 국제 출원(참조 번호 PH/5-20757/P1/CGC1847)을 참조하라]. 진공 침투된 식물의 종자를 수집하고, 화학식 XVII의 프로토क्स 억제 아릴우라실 제초제의 농도 범위(10.0nM 내지 1.0uM)에서 플레이팅한다. 야생형 아라비돕시스를 사용한 다중 실험은 정상적인 종자 발아를 억제하는데 상기 화합물 10.0nM의 농도면 충분하다는 것을 보여준다. 천연 프로토क्स-1 프로모터에 융합된 AraC-2Met 변화된 효소를 발현시키는 유전자식 종자가 500nM 농도 이하의 제초제에서 정상적인 아라비돕시스 발아를 생성한다는 것은 야생형 아라비돕시스와 비교하여 50배 이상의 제초제 내성을 보여준다. 따라서, 상기한 프로모터/변화된 프로토क्स 효소 융합체는 식물 형질전환에 대한 효과적인 선별 마커로서 기능한다. 프로토क्स-억제 제초제 100.0nM에서 발아되는 몇몇 식물들을 토양에 옮겨심고, 2 내지 3주 성장시키고, 프로토क्स-억제 제초제의 다양한 농도에서 분무 분석을 시험한다. 벡터가 없는 대조군 형질전환체와 비교하여, AraPT1Pro/AraC-2Met 유전자식체는 제초제 분무에 10배 이상의 내성을 갖는다.

실시예 4

옥수수 프로토क्स-1 프로모터 서열의 분리

람다 FIX II 벡터내의 제아 메이즈(Zea Mays, 미주리에서 17번 근친교배, 황변된 종자) 게놈 DNA 라이브러리를 스트라타진사로부터 구입한다. 약 250,000pfu의 라이브러리를 15cm 플레이트당 50,000파지의 밀도로 플레이팅하고, 콜로니/플라크 스크린 막(NEN Dupont사)상에 복제물을 올린다. 플라크 복제물은 랜덤 프라이밍 방법(Life Technologies사)에 의해 ³²P-dCTP로 표지된 옥수수 프로토क्स-1 cDNA(서열 1)로 프로빙한다. 문헌[참조: Church and Gilbert, Proc. Natl. Acad. Sci. USA 81: 1991-1995 (1984)]에 기술된 바와 같이 하이브리드화 및 세척 조건을 65°C로 한다. 람다 파지 DNA를 위저드 람다 프렙 DNA 정제 시스템(Wizard Lamda Preps DNA purification System, Promega사)을 사용하여 3개의 양성 하이브리드화된

파지로부터 분리한다. 제한 절단에 의한 분석, 하이브리드화 형태 및 DNA 서열 분석으로, cDNA 클론으로 먼저 분리된 옥수수 프로토스-1 암호화 서열 5'에 위치한 옥수수 게놈 DNA 약 3.5kb를 포함한 람다 클론을 확인한다. 이 단편은 옥수수 프로토스-1 프로모터를 포함하는 것으로 생각된다. 이 단편의 서열을 서열 14에 기술한다. 뉴클레오타이드 1 내지 3532의 서열은 5' 비암호화 서열에 포함된다. 뉴클레오타이드 3533 내지 3848의 서열은 옥수수 프로토스-1 단백질의 5' 말단을 암호화한다.

나머지 옥수수 프로토스-1 암호화 서열에 융합된 서열 14의 서열을 포함한 플라스미드는 1996년 3월 19일 pWDC-14로서 기탁되었다(NRRL #B-21546).

실시예 5

식물 형질전환 벡터의 작제

식물 형질전환에 수많은 형질전환 벡터를 이용할 수 있고, 본 발명의 프로모터 및 키메라 유전자를 어떤 벡터와도 결합하여 사용할 수 있다. 이에 사용하기 위한 벡터의 선택은 바람직한 형질전환 기술 및 형질전환시킬 표적 생물종에 의존적이다. 특정 표적 생물종에 있어서는 상이한 항생제 또는 제초제 선택 마커가 바람직하다. 형질전환에 일상적으로 사용되는 선택 마커에는 카나마이신 및 관련 항생제에 대하여 내성을 부여하는 nptII 유전자[참조: Messing & Vierra, Gene 19: 259-268 (1982); Bevan et al., Nature 304: 184-187 (1983)], 제초제 포스포노트리신에 대하여 내성을 부여하는 bar 유전자[참조: White et al., Nucl Acids Res 18: 1062 (1990), Spencer et al., Theor Appl Genet 79: 625-631 (1990)], 항생제 하이그로마이신에 대하여 내성을 부여하는 hph 유전자[참조: Blochinger & Diggelmann, Mol Cell Biol 4:2929-2931] 및 메토크세이트에 대하여 내성을 부여하는 dhfr 유전자[참조: Bourouis et al., EMBO J. 2(7): 1099-1104 (1983)]가 포함된다.

I. 아그로박테리움 형질전환에 적합한 벡터의 작제

아그로박테리움 투메파시엔스를 사용한 형질전환에 많은 벡터를 이용할 수 있다. 이는 전형적으로 1개 이상의 T-DNA 경계 서열을 운반하고, pBIN 19[참조: Bevan, Nucl. Acids Res. (1984)] 및 pXYZ와 같은 벡터를 포함한다. 2가지의 전형적인 벡터의 작제가 하기에 기술되어 있다.

pCIB200 및 pCIB2001의 작제: 한 쌍의 벡터 pCIB200 및 pCIB2001를 아그로박테리움에 사용하기 위한 재조합 벡터를 작제하는데 사용하고, 다음과 같은 방법으로 작제한다. 테트라사이클린-내성 유전자를 절단하고, 이어서 NPTII를 운반하는 pUC4K[참조: Messing & Vierra, Gene 19: 259-268 (1982); Bevan et al., Nature 304: 184-187 (1983); McBride et al., Plant Molecular Biology 14: 266-276 (1990)]의 AccI 단편을 삽입시킨 pTJS75[참조: Schmidhauser & Helinski, J Bacteriol. 164: 446-455 (1985)]를 NarI으로 절단하여 pTJS75kan을 제조한다. T-DNA 경계의 좌우, 식물의 선택적 nos/nptII 키메라 유전자 및 pUC 폴리링커[참조: Rothstein et al., Gene 53: 153-161 (1987)]를 포함하는 pCIB7의 EcoRI 단편을 XhoI 링커와 접합시키고, XhoI-절단 단편을 SalI-절단된 pTJS75kan으로 클로닝시켜 pCIB200을 제조한다[참조: EP 0 332 104, 실시예 19 (1338)]. pCIB200은 다음과 같이 독특한 폴리링커 제한 부위를 포함한다: EcoRI, SstI, KpnI, BglII, XbaI 및 SalI. pCIB2001은 추가의 제한 부위로 폴리링커를 삽입하여 제조한 pCIB200의 유도체이다. pCIB2001의 폴리링커에서 독특한 제한 부위는 EcoRI, SstI, KpnI, BglII, XbaI, SalI, MluI, BclI, AvrII, ApaI, HpaI 및 StuI이다. 이들 독특한 제한 부위를 포함하는 것 이외에, pCIB2001은 또한 식물 및 세균 카나마이신 선택, 아그로박테리움-매개된 형질전환을 위한 좌우 T-DNA 경계, E. Coli 및 기타 숙주 사이의 이동을 위한 RK2-유도된 trfA 기능, 및 또한 RK2로부터의 OriT 및 OriV 기능을 포함한다. pCIB2001 폴리링커는 자체의 조절 시그널을 포함하는 식물 발현 카세트의 클로닝에 적합하다.

pCIB10 및 이의 하이그로마이신 선택 유도체의 작제는 다음과 같다: 한 쌍의 벡터 pCIB10에는 식물에서 선택을 위해 카나마이신 내성을 암호화하는 유전자, 및 T-DNA 좌우 경계 서열이 포함되고, E. Coli 및 아그로박테리움 모두에서 복제를 가능하게 하는 광범위한 숙주-범위의 플라스미드 pRK252 서열이 혼입되어 있다. 이의 작제는 문헌[참조: Rothstein et al., Gene 53: 153-161 (1987)]에 기술되어 있다. 문헌[참조: Gritz et al., Gene 25: 179-188 (1983)]에 기술한 하이그로마이신 B 포스포트랜스퍼라제에 대한 유전자를 혼입시켜 pCIB10의 다양한 유도체를 작제한다. 이들 유도체는 하이그로마이신(pCIB743)상에서 또는 하이그로마이신 및 카나마이신(pCIB715, pCIB717)상에서 유전자이식 식물 세포의 선택을 가능하게 한다.

II. 비-아그로박테리움 형질전환에 적합한 벡터의 작제

아그로박테리움 투메파시엔스를 사용하지 않은 형질전환은 형질전환 벡터를 선택하는데 있어서 T-DNA 서열에 대한 필요성을 피하여 결과적으로 T-DNA 서열을 포함한 상기 종류의 벡터 이외에 이들 서열이 결핍된 벡터가 사용될 수 있다. 아그로박테리움에 의존하지 않는 형질전환 기술에는 입자 충격(particle bombardment), 원형질 흡수(예를 들어 PEG 및 전기천공) 및 미세주입(microinjection)에 의한 형질전환이 포함된다. 벡터의 선택은 형질전환될 바람직한 생물종의 선택에 크게 의존한다. 하기에, 몇몇 전형적인 벡터의 작제를 기술한다.

pCIB3064의 작제: pCIB3064는 제초제 바스타(또는 포스포리노트리신)에 의한 선택과 조합되어 직접적인 유전자 전달 기술에 적합한 pUC-유도된 벡터이다. 플라스미드 pCIB246은 E. Coli GUS 유전자에 작동가능하게 융합된 CaMV 35S 프로모터 및 CaMV 35S 전사 터미네이터를 포함하고, 이는 PCT 공개 특허권 WO 93/07278에 기술되어 있다. 이 벡터의 35S 프로모터는 시작 부위의 2개의 5' ATG 서열을 포함한다. 이들 부위를 표준 PCR 기술을 사용하여 ATG를 제거하는 방식으로 돌연변이시키고, 제한 부위 SapI 및 PvuII를 생성한다. 신규한 제한 부위는 독특한 SalI 부위로부터 96 및 36bp 떨어져 있고, 실질적인 시작 부위로부터 101 및 42bp 떨어져 있다. 생성된 pCIB246의 유도체는 pCIB3025로 지정한다. GUS 유전자를 SalI 및 SacI으로 절단하여 pCIB3025로부터 잘라낸 후, 말단을 불런트로 만들고, 재접합하여 pCIB3060을 생성한다. 플라스미드 pJIT82를 존 인스 센터(John Innes Center, Norwich)로부터 수득하고, 스트렙토마이세스 비리도크로모겐으로부터의 bar 유전자를 포함한 400bp의 SmaI 단편을 잘라내어 pCIB3060[참조: Thompson et al., EMBO J 6: 2519-2523 (1987)]의 HpaI 부위에 삽입한다. 이로써 CaMV 53S 프로모

터 및 제초제 선택에 대한 터미네이터의 조절하의 bar 유전자, 암피실린 내성에 대한 유전자(E. Coli에서의 선택을 위해) 및 독특한 부위 Sph I, Pst I, HindIII 및 BamHI을 갖는 폴리링커를 포함하는 pCIB3064를 생성한다. 이 벡터는 자체적인 조절 시그널을 포함하는 식물 발현 카세트의 클로닝에 적합하다

pSOG19 및 pSOG35의 작제: pSOG35는 메토틱세이트에 대한 내성을 부여하는 선택적인 마커로서 E. Coli 유전자 디하이드로폴레이트 리덕타제(DHFR)를 이용하는 형질전환 벡터이다. 35S 프로모터(약 800bp), 옥수수 Adh 1 유전자로부터의 인트론 6(약 550bp) 및 pSOG10로부터의 GUS 비해독된 리더 서열 18bp를 증폭시키는데 PCR을 사용한다. 또한, E. Coli 디하이드로폴레이트 리덕타제 유형 II 유전자를 암호화하는 250bp의 단편을 PCR로 증폭시키고, 이들 두 PCR 단편을 pUC19 벡터 골격 및 노팔린 신타제 터미네이터를 포함하는 pBI221(클론테크, Clontech)로부터의 SacI-PstI 단편으로 조립한다. 이들 단편들의 조립은 인트로 6 서열과 융합된 35S 프로모터, GUS 리더, DHFR 유전자 및 노팔린 신타제 터미네이터를 포함하는 pSOG19를 생성한다. pSOG19에서 GUS 리더를 옥수수 클로로틱 모틀 바이러스(Maize Chlorotic Mottle Virus, MCMV)의 리더 서열로서 대체하여 벡터 pSOG35를 생성한다. pSOG19 및 pSOG35는 암피실린 내성에 대한 pUC 유전자를 포함하고, 식물 프로토스 프로모터를 포함한 키메라 유전자 서열과 같은 외부 서열의 클로닝에 사용할 수 있는 HindIII, Sph I, Pst I 및 EcoRI 부위를 갖는다.

실시예 6

키메라 유전자/식물 발현 카세트의 작제

식물 프로토스 프로모터의 조절하에 유전자이식 식물에서의 발현이 의도된 암호화 서열은 적합한 프로토스 프로모터 및 상부의 적합한 전사 터미네이터하에서 발현 카세트로 조립할 수 있다. 생성된 키메라 유전자는 실시예 5에서 상기한 식물 형질전환 벡터로 용이하게 전달될 수 있다.

I. 프로토스 프로모터 선택

본 발명에 따라, 키메라 유전자는 식물 프로토스 프로모터를 포함한다. 키메라 유전자에서 사용되는 특정 프로토스 프로모터를 선택하는 것은, 일반적으로 밀접하게 관련된 식물종, 가장 바람직하게는 생성된 키메라 유전자를 포함시키고자 하는 식물종과 동일한 프로토스 프로모터를 사용하는 것이 바람직하지만, 주로 개별적인 연구자에 달려 있다. 예를 들어, 키메라 유전자가 옥수수에 포함될 경우, 외떡잎 식물을 사용하는 것이 바람직하고, 옥수수 프로토스 프로모터를 사용하는 것이 가장 바람직하다.

II. 전사 터미네이터

각종 전사 터미네이터를 발현 카세트에서 사용할 수 있다. 이들은 전달유전자에 있어서의 전사의 종결 및 이의 정확한 폴리아데닐화를 유발한다. 적절한 전사 터미네이터는 식물에서 기능이 공지된 것으로, CaMV 35S 터미네이터, tml 터미네이터, 노팔린 신타제 터미네이터, 완두 rbcS E9 터미네이터 뿐만 아니라 자연적으로 식물 프로토스 유전자와 관련된 터미네이터(즉, "프로토스 터미네이터")가 있다. 이들은 외떡잎 식물 및 쌍떡잎 식물 모두에서 사용될 수 있다.

III. 발현 향상 또는 조절을 위한 서열

전사 단위내에서 이로부터의 유전자 발현을 향상시키는 무수한 서열이 밝혀졌고, 이들 서열은 유전자이식 식물에서 본 발명의 유전자와 결합하여 이의 발현을 증가시키는데 사용할 수 있다.

각종 인트론 서열이 특히 외떡잎 식물 세포에서 발현을 향상시키는 것으로 나타났다. 예를 들어, 옥수수 Adh1 유전자의 인트론은, 옥수수 세포로 도입될 경우, 이의 동족 프로모터하에서 야생형 유전자의 발현을 크게 향상시키는 것으로 밝혀졌다. 특히, 인트론 1은 클로람페니콜 아세틸트랜스퍼라제 유전자를 갖는 작제물과 융합시 특히 효과적이고 향상되게 발현시키는 것으로 밝혀졌다[참조: Callis et al., Genes Develop. 1: 1183-1200 (1987)]. 동일한 실험 시스템에서, 옥수수 브롬즈 1 유전자의 인트론이 발현을 향상시키는데 유사한 효과를 갖는다[참조: Callis et al. 상기 참조]. 인트론 서열은 일상적으로 식물 형질전환 벡터, 특히 비-해독된 리더로 혼입된다.

바이러스로부터 유도된 무수한 비-해독된 리더 서열이 또한 발현을 향상시키는 것으로 공지되었고, 이들은 특히 쌍떡잎 식물에서 효과적이다. 특히, 담배의 모자이크 바이러스(Tobacco Mosaic Virus, TMV, "W-서열"), 옥수수 클로로틱 모틀 바이러스(MCMV) 및 알파파 모자이크 바이러스(Alfafa Mosaic Virus)로부터의 리더 서열이 발현을 향상시키는데 효과적인 것으로 나타났다[참조: Gallie et al. Nucl. Acids Res. 15: 8693-8711 (1987); Skuzeski et al. Plant Molec. Biol. 15: 65-79 (1990)].

IV. 세포내에서 유전자 생성물의 표적화

유전자 생성물을 표적화하는 각종 기작이 식물에 존재하는 것으로 공지되어 있으며, 이런 기작의 기능을 조절하는 서열이 다소 상세히 특성화되어 있다. 예를 들어, 엽록체에 대한 유전자 생성물의 표적화는 각종 단백질의 아미노 말단에서 발견된 시그널 서열에 의해 조절되고, 이는 엽록체가 생산한 성숙 단백질을 이입하는 동안 절단된다[참조: Comai et al. J. Biol. Chem. 263:15104-15109 (1988)]. 이 시그널 서열은 이중 유전자 생성물과 융합되어 이중 생성물을 엽록체내로 이입시킬 수 있다[참조: van den Broeck et al, Nature 313: 358-363 (1985)]. 적절한 시그널 서열을 암호화하는 DNA는 RUBISCO 단백질, CAB 단백질, EPSP 신타제 효소, GS2 단백질 및 엽록체에 집중되는 것으로 공지된 기타의 많은 단백질을 암호화하는 cDNA의 5' 말단으로부터 분리될 수 있다.

기타 유전자 생성물은 미토콘드리아 및 퍼옥시좀과 같은 다른 세포 기관에 집중된다[참조: Unger et al. Plant Molec. Biol. 13: 411-418 (1989)]. 또한, 상기한 기관으로 이중 유전자 생성물의 표적화를 실행하도록 이들 생성물을 암호화하는 cDNA를 조작할 수 있다. 상기한 서열의 예로는 핵-암호화 ATP아제 및 미토콘드리아에 특이적인 아스파르테이트 아미노 트랜스퍼라제 이소형이 있다[참조: Rogers et al., Proc. Natl. Acad. Sci. USA 82: 6512-6516 (1985)].

또한, 기타 세포 구획에서 유전자 생성물을 표적화하는 서열이 특성화되어 있다. 아미노 말단 서열은 ER, 아포플라스트 및 호분 세포로부터의 세포외 분비에 대한 표적화에 관여한다. 또한, 카복시 말단 서

열과 결합된 아미노 말단 서열은 유전자 생성물의 엑포 표적화에 관여한다[참조: Shinshi et al., *Plant Molec. Biol.* 14: 357-368 (1990)].

상기한 적절한 표적화 서열을 흥미로운 전달유전자 서열과 융합시킴으로써 임의의 세포 기관 또는 세포 구획에 전달유전자 생성물을 유도할 수 있다. 엽록체 표적화를 위해, 예를 들어 RUBISCO 유전자, CAB 유전자, EPSP 신타제 유전자 또는 GS2 유전자로부터 선택된 엽록체 시그널 서열을 전달유전자의 아미노 말단 ATG의 구조에 융합시킨다. 선택된 시그널 서열은 공지된 절단 부위를 반드시 포함해야하며, 작제된 융합물은 절단에 필요한 절단 부위 다음의 아미노산이 어떤 것인지 고려해야만 한다. 몇몇 경우에서, 상기 필요성은 절단 부위 및 전달유전자 ATG 사이에 소수의 아미노산을 부가하거나 선택적으로 전달유전자 서열내에 몇몇 아미노산을 대체함으로써 충족시킬 수 있다. 엽록체 이입을 위해 작제된 융합물은 시험관 내에서 전사된 작제물을 시험관내에서 해독한 후 문헌[참조: Bartlett et al. In: Edelmann et al. (Eds.) *Methods in Chloroplast Molecular Biology*, Elsevier, pp. 1081-1091 (1982); Wasmann et al. *Mol. Gen. Genet.* 205: 446-453 (1986)]에 기술된 방법을 사용하여 시험관내 엽록체 흡수의 효능을 시험할 수 있다. 이런 작제 기술은 당해 기술 분야에서 공지된 것이고, 미토콘드리아 및 퍼옥시좀에서 동일하게 적용할 수 있다. 전달유전자의 발현에 필요한 표적 선택은, 주어진 경로에 대한 출발점으로서 요구된 전구체의 세포내 국지성에 의존적이다. 이는 몇몇 경우에서는 미토콘드리아 또는 퍼옥시좀이지만, 일반적으로는 시토플 또는 엽록체일 수 있다. 전달유전자 발현 생성물은 일반적으로 ER, 아포플라스트 또는 액포에 대한 표적화를 필요로 하지 않는다.

세포내 표적화에 대해 상기한 기작은 식물 프로텍스 프로모터와의 결합에 사용되어 프로모터로부터 유도된 표적화 시그널에 대해 상이한 발현 형태를 갖는 프로모터의 전사 조절하에서 특이적인 세포 표적화 목적을 수행할 수 있다.

실시예 7

쌍떡잎 식물의 형질전환

쌍떡잎 식물에 대한 형질전환 기술은 당해 기술 분야에서 공지된 것이고, 아그로박테리움-기본된 기술 및 아그로박테리움을 요구하지 않는 기술이 포함된다. 비-아그로박테리움 기술은 원형질체 또는 세포에 의한 직접적인 외인성 유전 물질의 흡수를 수반한다. 이는 PEG 또는 전기천공-매개된 흡수, 입자 충격-매개된 운반 또는 미세주입에 의해 수행될 수 있다. 이들 기술의 예는 문헌[참조: Paszkowaski et al., *EMBO J* 3: 2717-2722 (1984), Potrykus et al., *Mol. Gen. Genet.* 199: 169-177 (1985), Reich et al., *Biotechnology* 4: 1001-1004 (1986), and Klein et al., *Nature* 327: 70-73 (1987)]에 기술되어 있다. 각 경우에서, 형질전환된 세포는 당해 기술 분야에서 공지된 표준 기술을 사용하여 식물 전체로 재생된다.

아그로박테리움-매개된 형질전환은, 이의 형질전환 효율이 높고 많은 상이한 종을 사용하는 이의 광범위한 용도 때문에 쌍떡잎 식물의 형질전환을 위해 바람직한 기술이다. 아그로박테리움에 의해 일상적으로 형질전환되는 많은 작물 종에는 담배, 토마토, 해바라기, 목화, 평지, 감자, 대두, 알파파 및 포플러가 포함된다[참조: EP 0 317 511(목화), EP 0 249 432(토마토, Calgene), WO 87/07299(브래시카, Calgene), US 4,795,855(포플러)] .

재조합 아그로박테리움에 의한 표적 식물종의 형질전환은 일반적으로 식물의 이식편과 아그로박테리움의 공동배양을 수반하고, 당해 기술 분야에서 익히 공지된 프로토콜을 따른다. 형질전환된 조직은 한 쌍의 플라스미드 T-DNA 경계사이에 존재하는 항생제-내성 또는 제초제-내성 마커를 포함하는 선택적 배지상에서 재생된다.

실시예 8

외떡잎 식물의 형질전환

대부분의 외떡잎 식물종의 형질전환은 현재 일반적인 것이다. 바람직한 기술에는 PEG 또는 전기천공 기술을 사용한 원형질체로의 직접적인 유전자 전달 및 융합 조직내로의 입자 충격이 포함된다. 형질전환은 단일종의 DNA 또는 다종의 DNA(즉, 공동-형질전환)로 수행될 수 있고, 이들 두 기술은 본 발명에 사용하기에 적합하다. 공동-형질전환은 복잡한 벡터 작제를 피할 수 있고, 흥미로운 유전자 및 선택적인 마커에 연결되지 않은 유전적 위치를 갖는 유전자이식 식물을 생성하고, 후속의 증식에서 바람직한 것으로 여겨지는 선택 마커의 제거가 가능하다는 잇점이 있다. 그러나, 공동-형질전환을 사용할 경우의 단점은 개개의 DNA 종류가 계능에 통합되는 빈도가 100% 미만이라는 것이다[참조: Schocher et al. *Biotechnology* 4: 1093-1096(1986)].

특허원 EP 0 292 453(Ciba-Geigy), EP 0 392 225(Ciba-Geigy), WO 93/07278(Ciba-Geigy) 및 미국 특허 제 5,350,689(Ciba-Geigy)는 옥수수의 선택된 근친교배 계통으로부터 융합 조직 및 원형질체의 제조, PEG 및 전기천공을 사용한 원형질체의 형질전환 및 형질전환된 원형질체로부터 옥수수를 증식시키는 기술을 기술하고 있다. Gordon-Kamm 등 및 Fromm 등[참조: *Plant Cell* 2: 603-618 (1990) 및 *Biotechnology* 8: 833-839 (1990)]은 입자 충격을 사용한 A188-유도된 옥수수 계통의 형질전환 기술을 발표하였다. 또한, 특허원 WO 93/07278(Ciba-Geigy) 및 문헌[참조: Koziel et al., *Biotechnology* 11: 194-200 (1993)]에는 입자 충격을 사용하여 옥수수의 선택된 근친교배 계통의 형질전환 기술을 기술하고 있다. 이 기술은 수분 후 14 내지 15일째 옥수수 이삭으로부터 잘라낸 1.5 내지 2.5mm 길이의 미숙한 옥수수 씨눈 및 충격을 위해 PDS-1000He 바이오리스팅 장치를 사용한다.

벼의 형질전환은 원형질 또는 입자 충격을 사용한 직접적인 유전자 전달 기술에 의해 수행된다. 원형질체-매개된 형질전환은 일본형(Japonica type) 및 인도형(Indica type)에 대해 기술된다[참조: Zhang et al., *Plant Cell Rep* 7: 379-384 (1988); Shimamoto et al., *Nature* 338: 274-277 (1989); Datta et al. *Biotechnology* 8: 736-740 (1990)]. 또한, 양쪽 형태는 입자 충격을 사용하여 일상적으로 형질전환된다[참조: Christou et al. *Biotechnology* 9: 957-962 (1991)].

특허원 EP 0 332 581(Ciba-Geigy)에는 푸이디에(Pooideae) 원형질체 생성, 형질전환 및 증식 기술이 기술되어 있다. 이들 기술은 닥틸리스(Dactylis) 및 밀의 형질전환을 가능하게 한다. 또한, 장기간 증식될

수 있는 유합 세포 C 유형내로 입자 충격을 사용한 밀의 형질전환이 문헌[참조: Vasil et al., Biotechnology 10: 667-674 (1992)]에 기술되어 있고, 미숙한 씨눈 및 미숙한 씨눈-유도된 유합 세포의 입자 충격을 사용한 밀의 형질전환이 또한 Vasil 등 및 Weeks 등의 문헌[Biotechnology 11: 1553-1558 (1993) and Plant Physiol. 102: 1077-1084 (1993)]에 기술되어 있다. 그러나, 밀의 형질전환을 위한 바람직한 기술은 미숙한 씨눈의 입자 충격에 의한 밀의 형질전환을 수반하고, 유전자 전달 단계에 선행하고-수크로즈 또는 고-말토즈 단계를 포함한다. 충격에 선행하여, 암실에서 실시되는 체세포 씨눈의 유도를 위해 특정수의 씨눈(0.75 내지 1mm 길이)을 3% 수크로즈[Murashige & Skoog, Physiologia Plantarum 15: 473-497 (1962)] 및 2,4-D 3mg/l를 포함한 MS 배지상에 플레이팅한다. 충격 선택일에, 씨눈을 유도 배지로부터 제거하고 삼투 배지(즉, 바람직한 농도, 전형적으로 15%로 가해진 수크로즈 또는 말토즈를 갖는 유도 배지)상에 둔다. 씨눈을 2 내지 3시간 동안 원형질분리시킨 후 충격을 가한다. 중요하지는 않지만, 통상적으로 표적 플레이트 당 씨눈은 20개이다. 적절한 유전자-함유 플라스미드(예를 들어 pCIB3064 또는 pSG35)를 표준 공정을 사용하여 마이크로미터 크기의 금 입자상에 침전시킨다. 각각의 씨눈 플레이트를 표준 80 메쉬 스크린을 사용한 약 1000psi의 파열압을 사용하여 듀폰 바이오리시틱 헬름 장치로 발사한다. 충격을 가한 후, 씨눈을 다시 암실에 두고, 약 24시간 동안 회복시킨다(여전히 삼투 배지). 24시간 후, 씨눈을 삼투 배지로부터 제거하고, 유도 배지상에 다시 두어 증식하기 전 약 1개월 동안 방치한다. 약 1개월 후, 씨눈발생 유합 세포를 발달시키는 씨눈 이식편을, 적절한 선택제(pCIB3064의 경우 바스타 10mg/l이고, pSG35의 경우 메토크세이트 2mg/l)를 추가로 함유한 증식 배지(MS + NAA 1mg/l, GA 5mg/l)로 전달한다. 약 1개월 후, 발달된 발아를 1/2 농도의 MS, 2% 수크로즈 및 동일한 농도의 선택제를 함유한, "GA7s"로 공지된 보다 큰 멸균 용기로 옮긴다. WO 94/13822는 밀의 형질전환 방법을 기술하고 있고, 이는 본원에서 참고로서 인용된다.

실시에 9

천연 옥수수 프로토क्स-1 프로모터하에서 변화된 프로토क्स-1 유전자를 발현시키는 식물 형질전환 벡터의 작제

3848bp의 옥수수 게놈 단편(서열 14)을 Sal I-Kpn I으로 부분적으로 절단된 생성물로서 분리된 람다 파지로부터 잘라내고, 아미노산 164에서 알라닌이 루신으로 대체된 변화된 옥수수 프로토क्स-1 cDNA로부터 유도된 Kpn I-Not I 단편(서열 6)과 접합시킨다. 이로써 천연 옥수수 프로토क्स-1 프로모터와 세균 시스템에서 제초제에 대한 내성을 부여하는 것으로 나타난 총 길이의 cDNA와의 융합물을 생성한다["DNA Molecules Encoding Plant Protoporphyrinogen Oxidase and Inhibitor Resistant Mutants Thereof"(참고 번호 PH/5-20757/P1/CGC1847)란 명칭의 공동 계류중인 국제 출원, 실시예 8 내지 13을 참조하라]. 이 융합물은 CaMV 35S 터미네이터 서열을 포함한 pUC18 유도된 벡터로 클로닝되어 프로토क्स 프로모터/변화된 프로토क्स cDNA/터미네이터 카세트를 생성한다. 이 카세트를 포함한 플라스미드를 pWCo-1으로 지정한다.

옥수수 형질전환을 위한 제2 작제물은 옥수수 게놈 클론으로부터의 암호화 서열에서 발견된 제1 인트론을 옥수수 cDNA로 조작하여 생성한다. 인트론(서열 25)은 길이가 93bp이고, 서열 5의, 정확하게는 실시예 4에 기술된 람다 클론에서 자연적인 함유물처럼 보이는 바와 같이 뉴클레오타이드 203 및 204 사이에 삽입되어 있다. 상기 인트론-함유 형태의 발현 카세트는 pWCo-2로 지정한다.

실시에 10

유전자이식 옥수수 식물에서 옥수수 프로토क्स-1 프로모터 활성의 증명

옥수수 프로토क्स 프로모터/변화된 프로토क्स 융합물로 형질전환된 옥수수 식물은 전달유전자에 특이적인 프라이머로 PCR 분석을 사용하여 확인한다. 총 RNA를 PCR 양성 식물로부터 제조하고, 슈퍼스크립트(Superscript) M-MLV(Life Technologies 사)를 사용하여 추천된 조건하에서 역-전사시킨다. 역전사 반응물 2μl를 변화된 프로토क्स 서열에 대해 특이적으로 고안된 PCR 반응에 사용한다. 상기 반응에서 비형질전환된 대조군이 생성물을 생산하지 않는 반면, pWCo-1으로 형질전환된 식물의 약 85%가 전달유전자로부터 유도된 mRNA의 존재를 나타내는 양성 결과를 생성하였다. 이는 옥수수 프로토क्स 프로모터에 대한 어느 정도의 활성 수준을 증명한다. 또한, 유전자이식 옥수수 식물로부터의 RNA는 프로브로서 서열 5로부터의 방사능표지된 옥수수 프로토क्स cDNA 단편을 사용한 표준 노던 블롯팅 분석을 실시한다. 비형질전환된 대조군 농도를 크게 증가하는 프로토क्स-1 mRNA 농도가 몇몇 유전자이식 옥수수 식물에서 검출되었다. 이런 상승된 mRNA 농도는 클로닝된 옥수수 프로토क्स 프로모터로부터의 변화된 프로토क्स-1 mRNA의 발현에 기인한 것으로 추정된다.

실시에 11

사탕무 프로토क्स-1 프로모터 서열의 분리

사탕무 게놈 라이브러리는 람다 FIX II 벡터내에서 스트라타진사에 의해 제조되었다. 라이브러리 약 300,000pfu를 플레이팅하고, 실시예 4에서 옥수수에 대해 기술한 바와 같이 사탕무 프로토क्स-1 cDNA 서열(서열 17)로 프로빙한다. 제한 절단에 의한 분석, 하이브리드화 형태 및 DNA 서열 분석으로, cDNA 클론으로 먼저 분리된 사탕무 암호화 서열에 대해 5' 방향으로 위치한 약 7kb의 사탕무 게놈 DNA를 함유한 람다 클론을 확인하였다. 2606bp의 Pst I-Sal I 단편을 람다 클론으로부터 pBluescript 벡터내로 서브클로닝한다. 상기 단편은 2068bp의 5' 비암호화 서열을 포함하고, 사탕무 프로토क्स-1 프로모터 서열을 포함한다. 또한, 이는 프로토क्स-1 암호화 서열의 처음 453bp 및 암호화 서열내에 포함된 85bp의 제1 인트론을 포함한다. 이 단편의 서열은 서열 26에 기술되어 있다.

서열 서열 26을 포함한 플라스미드를 pWDC-20(NRRL #B-21650)으로서 1996년 2월 6일 기탁하였다.

실시에 12

천연 사탕무 프로토क्स-1 프로모터하에서 변화된 사탕무 프로토क्स-1 유전자를 발현시키는 식물 형질전환 벡터의 작제

사탕무 게놈 단편(서열 26)을 5' 비암호화 서열 2068bp 및 사탕무 프로토क्स-1 암호화 서열의 처음 300bp

를 포함한 Sac I -BsrG I 단편으로서 실시예 11에서 기술한 게놈 서브클론으로부터 잘라낸다. 이 단편을, 아미노산 449에서 티로신이 메티오닌으로 치환된 변화된 사탕무 프로토스-1 cDNA(서열 18)로부터 유도된 BsrG I -Not I 단편과 접합한다. 이로써 천연 사탕무 프로토스-1 프로모터와 세균 시스템에서 제초제 내성을 부여하는 것으로 나타난 cDNA 총 길이와의 융합물을 생성한다(공동 계류중인 국제 출원 NO. (참고 번호 PH/5-20757/P1/CGC1847)). 이 융합물을 CaMV 35S 터미네이터 서열을 포함한 pUC18-유도된 벡터내로 클로닝하여 프로토스 프로모터/변화된 프로토스 cDNA/터미네이터 카세트를 생성한다. 이 카세트를 포함하는 플라스미드를 pWCo-3이라 지정한다.

실시예 13

천연 사탕무 프로토스-1 프로모터/변화된 사탕무 프로토스-1 융합물의 발현에 의해 제초제 내성 식물의 생산

pWCo-3으로부터의 발현 카세트를 쌍떡잎 식물에 적용할 수 있는, 아그로박테리움, 원형질체 및 바이오리 스틱 형질전환 기술을 포함한 임의의 형질전환 방법을 사용하여 사탕무내로 형질전환시킨다. 변화된 프로토스-1 효소를 발현하는 유전자이식 사탕무는 RNA-PCR에 의해 확인하고, 비형질전환된 사탕무에 치명적인 농도에서 프로토스-억제 제초제에 대한 내성을 시험한다.

본 발명은 이의 특별한 양태를 참조로 기술하고 있지만, 수많은 변화, 변형 및 양태가 가능하다는 것을 이해하고, 따라서 모든 상기한 변화, 변형 및 양태는 본 발명의 정신 및 범주내에 있는 것으로 간주한다.

서열 목록

(1) 일반 정보:

- (i) 출원인 : 존슨, 마리에, 볼라쓰, 산드라, 워드, 에릭
- (ii) 발명의 명칭 : 식물 프로토포르피리노겐 옥시다제 유전자로부터의 프로모터
- (iii) 서열 수 : 26
- (iv) 통신 주소 :

- (A) 주소 : 노바르티스 코포레이션
- (B) 거리 : 피 오 박스 2005 화이트 플레인스 로오드 520
- (C) 도시 : 테리타운
- (D) 주 : 뉴욕
- (E) 국가 : 미국
- (F) 우편 번호 : 10591-9005
- (v) 컴퓨터 판독 형태 :
- (A) 매체 형태 : 플로피 디스크
- (B) 컴퓨터 : IBM PC 호환
- (C) 운영 체계 : PC-DOS/MS-DOS
- (D) 소프트웨어 : PatentIn Release #1, Version #1.30

(vi) 현재 출원 정보 :

- (A) 출원 번호 :
- (B) 출원일 :
- (C) 분류 번호 :
- (vii) 선행 출원 정보 :
- (A) 출원 번호 : US 60/012,705
- (B) 출원일 : 1996년 2월 28일

(vii) 선행 출원 정보 :

- (A) 출원 번호 : US 60/013,612
- (B) 출원일 : 1996년 2월 28일
- (vii) 선행 출원 정보 :

- (A) 출원 번호 : US 60/020,003
- (B) 출원일 : 1996년 6월 21일

(viii) 변호인/대리인 정보 :

- (A) 성명 : 메이그스 제이 티모티
- (B) 등록 번호 : 38,241

(C) 참조/도켓 번호 : CGC 1846

(xi) 원거리통신 정보 :

(A) 전화 번호 : (919) 541-8587

(B) 팩스 번호 : (919) 541-8689

(2) 서열 1에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 1719개 염기쌍

(B) 유형 : 핵산

(C) 채 : 일본채

(D) 형태 : 선형

(ii) 분자 유형 : cDNA

(iii) 가설 : 없음

(iv) 안티-센스 : 없음

(vi) 최초 공급원 :

(A) 생물 : 아라비돕시스 탈리아나(Arabidopsis thaliana)

(vii) 중간체 공급원 :

(B) 클론 : pWDC-2(NRRL B-21238)

(ix) 특성 :

(A) 명칭/키 : CDS

(B) 위치 : 31..1644

(D) 기타 정보 : /생성물="아라비돕시스 프로토스-1"

(xi) 서열 기술 :

서열 1

TGACAAAATT CCGAATTCTC TCGATTTC ATG GAG TTA TCT CTT CTC CGT CCG Met Glu Leu Ser Leu Leu Arg Pro 1 5	54
ACG ACT CAA TCG CTT CTT CCG TCG TTT TCG AAG CCC AAT CTC CGA TTA Thr Thr Gln Ser Leu Leu Pro Ser Phe Ser Lys Pro Asn Leu Arg Leu 10 15 20	102

AAT GTT TAT AAG CCT CTT AGA CTC CGT TGT TCA GTG GCC GGT GGA CCA 150
Asn Val Tyr Lys Pro Leu Arg Leu Arg Cys Ser Val Ala Gly Gly Pro
25 30 35 40

ACC GTC GGA TCT TCA AAA ATC GAA GGC GGA GGA GGC ACC ACC ATC ACG 198
Thr Val Gly Ser Ser Lys Ile Glu Gly Gly Gly Gly Thr Thr Ile Thr
45 50 55

ACG GAT TGT GTG ATT GTC GGC GGA GGT ATT AGT GGT CTT TGC ATC GCT 246
Thr Asp Cys Val Ile Val Gly Gly Gly Ile Ser Gly Leu Cys Ile Ala
60 65 70

CAG GCG CTT GCT ACT AAG CAT CCT GAT GCT GCT CCG AAT TTA ATT GTG 294
Gln Ala Leu Ala Thr Lys His Pro Asp Ala Ala Pro Asn Leu Ile Val
75 80 85

ACC GAG GCT AAG GAT CGT GTT GGA GGC AAC ATT ATC ACT CGT GAA GAG 342
Thr Glu Ala Lys Asp Arg Val Gly Gly Asn Ile Ile Thr Arg Glu Glu
90 95 100

AAT GGT TTT CTC TGG GAA GAA GGT CCC AAT AGT TTT CAA CCG TCT GAT 390
Asn Gly Phe Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp
105 110 115 120

CCT ATG CTC ACT ATG GTG GTA GAT AGT GGT TTG AAG GAT GAT TTG GTG 438
Pro Met Leu Thr Met Val Val Asp Ser Gly Leu Lys Asp Asp Leu Val
125 130 135

TTG GGA GAT CCT ACT GCG CCA AGG TTT GTG TTG TGG AAT GGG AAA TTG 486
Leu Gly Asp Pro Thr Ala Pro Arg Phe Val Leu Trp Asn Gly Lys Leu
140 145 150

AGG CCG GTT CCA TCG AAG CTA ACA GAC TTA CCG TTC TTT GAT TTG ATG 534
Arg Pro Val Pro Ser Lys Leu Thr Asp Leu Pro Phe Phe Asp Leu Met
155 160 165

AGT ATT GGT GGG AAG ATT AGA GCT GGT TTT GGT GCA CTT GGC ATT CGA 582
Ser Ile Gly Gly Lys Ile Arg Ala Gly Phe Gly Ala Leu Gly Ile Arg
170 175 180

CCG TCA CCT CCA GGT CGT GAA GAA TCT GTG GAG GAG TTT GTA CGG CGT 630

AAC CTC GGT GAT GAG GTT TTT GAG CGC CTG ATT GAA CCG TTT TGT TCA Asn Leu Gly Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser 205 210 215	678
GGT GTT TAT GCT GGT GAT CCT TCA AAA CTG AGC ATG AAA GCA GCG TTT Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe 220 225 230	726
GGG AAG GTT TGG AAA CTA GAG CAA AAT GGT GGA AGC ATA ATA GGT GGT Gly Lys Val Trp Lys Leu Glu Gln Asn Gly Gly Ser Ile Ile Gly Gly 235 240 245	774
ACT TTT AAG GCA ATT CAG GAG AGG AAA AAC GCT CCC AAG GCA GAA CGA Thr Phe Lys Ala Ile Gln Glu Arg Lys Asn Ala Pro Lys Ala Glu Arg 250 255 260	822
GAC CCG CGC CTG CCA AAA CCA CAG GGC CAA ACA GTT GGT TCT TTC AGG Asp Pro Arg Leu Pro Lys Pro Gln Gly Gln Thr Val Gly Ser Phe Arg 265 270 275 280	870
AAG GGA CTT CGA ATG TTG CCA GAA GCA ATA TCT GCA AGA TTA GGT AGC Lys Gly Leu Arg Met Leu Pro Glu Ala Ile Ser Ala Arg Leu Gly Ser 285 290 295	918
AAA GTT AAG TTG TCT TGG AAG CTC TCA GGT ATC ACT AAG CTG GAG AGC Lys Val Lys Leu Ser Trp Lys Leu Ser Gly Ile Thr Lys Leu Glu Ser 300 305 310	966
GGA GGA TAC AAC TTA ACA TAT GAG ACT CCA GAT GGT TTA GTT TCC GTG Gly Gly Tyr Asn Leu Thr Tyr Glu Thr Pro Asp Gly Leu Val Ser Val 315 320 325	1014
CAG AGC AAA AGT GTT GTA ATG ACG GTG CCA TCT CAT GTT GCA AGT GGT Gln Ser Lys Ser Val Val Met Thr Val Pro Ser His Val Ala Ser Gly 330 335 340	1062
CTC TTG CGC CCT CTT TCT GAA TCT GCT GCA AAT GCA CTC TCA AAA CTA Leu Leu Arg Pro Leu Ser Glu Ser Ala Ala Asn Ala Leu Ser Lys Leu 345 350 355 360	1110
TAT TAC CCA CCA GTT GCA GCA GTA TCT ATC TCG TAC CCG AAA GAA GCA Tyr Tyr Pro Pro Val Ala Ala Val Ser Ile Ser Tyr Pro Lys Glu Ala 365 370 375	1158
ATC CGA ACA GAA TGT TTG ATA GAT GGT GAA CTA AAG GGT TTT GGG CAA	1206

Ile Arg Thr Glu Cys Leu Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln	
380 385 390	
TTG CAT CCA CGC ACG CAA GGA GTT GAA ACA TTA GGA ACT ATC TAC AGC	1254
Leu His Pro Arg Thr Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser	
395 400 405	
TCC TCA CTC TTT CCA AAT CGC GCA CCG CCC GGA AGA ATT TTG CTG TTG	1302
Ser Ser Leu Phe Pro Asn Arg Ala Pro Pro Gly Arg Ile Leu Leu Leu	
410 415 420	
AAC TAC ATT GGC GGG TCT ACA AAC ACC GGA ATT CTG TCC AAG TCT GAA	1350
Asn Tyr Ile Gly Gly Ser Thr Asn Thr Gly Ile Leu Ser Lys Ser Glu	
425 430 435 440	
GGT GAG TTA GTG GAA GCA GTT GAC AGA GAT TTG AGG AAA ATG CTA ATT	1398
Gly Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile	
445 450 455	
AAG CCT AAT TCG ACC GAT CCA CTT AAA TTA GGA GTT AGG GTA TGG CCT	1446
Lys Pro Asn Ser Thr Asp Pro Leu Lys Leu Gly Val Arg Val Trp Pro	
460 465 470	
CAA GCC ATT CCT CAG TTT CTA GTT GGT CAC TTT GAT ATC CTT GAC ACG	1494
Gln Ala Ile Pro Gln Phe Leu Val Gly His Phe Asp Ile Leu Asp Thr	
475 480 485	
GCT AAA TCA TCT CTA ACG TCT TCG GGC TAC GAA GGG CTA TTT TTG GGT	1542
Ala Lys Ser Ser Leu Thr Ser Ser Gly Tyr Glu Gly Leu Phe Leu Gly	
490 495 500	
GGC AAT TAC GTC GCT GGT GTA GCC TTA GGC CGG TGT GTA GAA GGC GCA	1590
Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala	
505 510 515 520	
TAT GAA ACC GCG ATT GAG GTC AAC AAC TTC ATG TCA CGG TAC GCT TAC	1638
Tyr Glu Thr Ala Ile Glu Val Asn Asn Phe Met Ser Arg Tyr Ala Tyr	
525 530 535	
AAG TAAATGTAAA ACATTAAATC TCCCAGCTTG CGTGAGTTTT ATTAAATATT	1691
Lys	
TTGAGATATC CAAAAAAAAA AAAAAAAAAA	1719

(2) 서열 2에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 537개 아미노산

(B) 유형 : 아미노산

(D) 형태 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열 기술 :

서열 2

```

Met Glu Leu Ser Leu Leu Arg Pro Thr Thr Gln Ser Leu Leu Pro Ser
 1          5          10          15

Phe Ser Lys Pro Asn Leu Arg Leu Asn Val Tyr Lys Pro Leu Arg Leu
      20          25          30

Arg Cys Ser Val Ala Gly Gly Pro Thr Val Gly Ser Ser Lys Ile Glu
      35          40          45

Gly Gly Gly Gly Thr Thr Ile Thr Thr Asp Cys Val Ile Val Gly Gly
 50          55          60

Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu Ala Thr Lys His Pro
 65          70          75          80

Asp Ala Ala Pro Asn Leu Ile Val Thr Glu Ala Lys Asp Arg Val Gly
      85          90          95

Gly Asn Ile Ile Thr Arg Glu Glu Asn Gly Phe Leu Trp Glu Glu Gly
 100          105          110

Pro Asn Ser Phe Gln Pro Ser Asp Pro Met Leu Thr Met Val Val Asp
 115          120          125

Ser Gly Leu Lys Asp Asp Leu Val Leu Gly Asp Pro Thr Ala Pro Arg
 130          135          140

Phe Val Leu Trp Asn Gly Lys Leu Arg Pro Val Pro Ser Lys Leu Thr
 145          150          155          160

Asp Leu Pro Phe Phe Asp Leu Met Ser Ile Gly Gly Lys Ile Arg Ala
      165          170          175

```

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

```

Pro Pro Gly Arg Ile Leu Leu Leu Asn Tyr Ile Gly Gly Ser Thr Asn
      420              425              430

Thr Gly Ile Leu Ser Lys Ser Glu Gly Glu Leu Val Glu Ala Val Asp
      435              440              445

Arg Asp Leu Arg Lys Met Leu Ile Lys Pro Asn Ser Thr Asp Pro Leu
      450              455              460

Lys Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val
      465              470              475              480

Gly His Phe Asp Ile Leu Asp Thr Ala Lys Ser Ser Leu Thr Ser Ser
      485              490              495

Gly Tyr Glu Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala
      500              505              510

Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu Thr Ala Ile Glu Val Asn
      515              520              525

Asn Phe Met Ser Arg Tyr Ala Tyr Lys
      530              535

```

(2) 서열 3에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 1738개 염기쌍

(B) 유형 : 핵산

(C) 채 : 일본채

(D) 형태 : 선형

(ii) 분자 유형 : cDNA

(iii) 가설 : 없음

(iv) 안티-센스 : 없음

(vi) 최초 공급원 :

(A) 생물 : 아라비도시스 탈리아나(Arabidopsis thaliana)

(vii) 중간체 공급원 :

(B) 클론 : pWDC-2(NRRL B-21237)

(ix) 특성 :

(A) 명칭/키 : CDS

(B) 위치 : 70..1596

(D) 기타 정보 : /생성물="아라비도시스 프로토스-2"

(xi) 서열 기술 :

서열 3

TTTTTACTT ATTTCCGTCA CTGCTTTCGA CTGGTCAGAG ATTTTGACTC TGAATTGTTG	60
CAGATAGCA ATG GCG TCT GGA GCA GTA GCA GAT CAT CAA ATT GAA GCG	108
Met Ala Ser Gly Ala Val Ala Asp His Gln Ile Glu Ala	
1 5 10	
GTT TCA GGA AAA AGA GTC GCA GTC GTA GGT GCA GGT GTA AGT GGA CTT	156
Val Ser Gly Lys Arg Val Ala Val Val Gly Ala Gly Val Ser Gly Leu	
15 20 25	
GCG GCG GCT TAC AAG TTG AAA TCG AGG GGT TTG AAT GTG ACT GTG TTT	204
Ala Ala Ala Tyr Lys Leu Lys Ser Arg Gly Leu Asn Val Thr Val Phe	
30 35 40 45	
GAA GCT GAT GGA AGA GTA GGT GGG AAG TTG AGA AGT GTT ATG CAA AAT	252
Glu Ala Asp Gly Arg Val Gly Gly Lys Leu Arg Ser Val Met Gln Asn	
50 55 60	
GGT TTG ATT TGG GAT GAA GGA GCA AAC ACC ATG ACT GAG GCT GAG CCA	300
Gly Leu Ile Trp Asp Glu Gly Ala Asn Thr Met Thr Glu Ala Glu Pro	
65 70 75	
GAA GTT GGG AGT TTA CTT GAT GAT CTT GGG CTT CGT GAG AAA CAA CAA	348
Glu Val Gly Ser Leu Leu Asp Asp Leu Gly Leu Arg Glu Lys Gln Gln	
80 85 90	
TTT CCA ATT TCA CAG AAA AAG CGG TAT ATT GTG CGG AAT GGT GTA CCT	396
Phe Pro Ile Ser Gln Lys Lys Arg Tyr Ile Val Arg Asn Gly Val Pro	
95 100 105	
GTG ATG CTA CCT ACC AAT CCC ATA GAG CTG GTC ACA AGT AGT GTG CTC	444
Val Met Leu Pro Thr Asn Pro Ile Glu Leu Val Thr Ser Ser Val Leu	
110 115 120 125	
TCT ACC CAA TCT AAG TTT CAA ATC TTG TTG GAA CCA TTT TTA TGG AAG	492
Ser Thr Gln Ser Lys Phe Gln Ile Leu Leu Glu Pro Phe Leu Trp Lys	
130 135 140	

AAA AAG TCC TCA AAA GTC TCA GAT GCA TCT GCT GAA GAA AGT GTA AGC	540
Lys Lys Ser Ser Lys Val Ser Asp Ala Ser Ala Glu Glu Ser Val Ser	
145 150 155	
GAG TTC TTT CAA CGC CAT TTT GGA CAA GAG GTT GTT GAC TAT CTC ATC	588
Glu Phe Phe Gln Arg His Phe Gly Gln Glu Val Val Asp Tyr Leu Ile	
160 165 170	
GAC CCT TTT GTT GGT GGA ACA AGT GCT GCG GAC CCT GAT TCC CTT TCA	636
Asp Pro Phe Val Gly Gly Thr Ser Ala Ala Asp Pro Asp Ser Leu Ser	
175 180 185	
ATG AAG CAT TCT TTC CCA GAT CTC TGG AAT GTA GAG AAA AGT TTT GGC	684
Met Lys His Ser Phe Pro Asp Leu Trp Asn Val Glu Lys Ser Phe Gly	
190 195 200 205	
TCT ATT ATA GTC GGT GCA ATC AGA ACA AAG TTT GCT GCT AAA GGT GGT	732
Ser Ile Ile Val Gly Ala Ile Arg Thr Lys Phe Ala Ala Lys Gly Gly	
210 215 220	
AAA AGT AGA GAC ACA AAG AGT TCT CCT GGC ACA AAA AAG GGT TCG CGT	780
Lys Ser Arg Asp Thr Lys Ser Ser Pro Gly Thr Lys Lys Gly Ser Arg	
225 230 235	
GGG TCA TTC TCT TTT AAG GGG GGA ATG CAG ATT CTT CCT GAT ACG TTG	828
Gly Ser Phe Ser Phe Lys Gly Gly Met Gln Ile Leu Pro Asp Thr Leu	
240 245 250	
TGC AAA AGT CTC TCA CAT GAT GAG ATC AAT TTA GAC TCC AAG GTA CTC	876
Cys Lys Ser Leu Ser His Asp Glu Ile Asn Leu Asp Ser Lys Val Leu	
255 260 265	
TCT TTG TCT TAC AAT TCT GGA TCA AGA CAG GAG AAC TGG TCA TTA TCT	924
Ser Leu Ser Tyr Asn Ser Gly Ser Arg Gln Glu Asn Trp Ser Leu Ser	
270 275 280 285	
TGT GTT TCG CAT AAT GAA ACG CAG AGA CAA AAC CCC CAT TAT GAT GCT	972
Cys Val Ser His Asn Glu Thr Gln Arg Gln Asn Pro His Tyr Asp Ala	
290 295 300	
GTA ATT ATG ACG GCT CCT CTG TGC AAT GTG AAG GAG ATG AAG GTT ATG	1020
Val Ile Met Thr Ala Pro Leu Cys Asn Val Lys Glu Met Lys Val Met	
305 310 315	

AAA GGA GGA CAA CCC TTT CAG CTA AAC TTT CTC CCC GAG ATT AAT TAC Lys Gly Gly Gln Pro Phe Gln Leu Asn Phe Leu Pro Glu Ile Asn Tyr 320 325 330	1068
ATG CCC CTC TCG GTT TTA ATC ACC ACA TTC ACA AAG GAG AAA GTA AAG Met Pro Leu Ser Val Leu Ile Thr Thr Phe Thr Lys Glu Lys Val Lys 335 340 345	1116
AGA CCT CTT GAA GGC TTT GGG GTA CTC ATT CCA TCT AAG GAG CAA AAG Arg Pro Leu Glu Gly Phe Gly Val Leu Ile Pro Ser Lys Glu Gln Lys 350 355 360 365	1164
CAT GGT TTC AAA ACT CTA GGT ACA CTT TTT TCA TCA ATG ATG TTT CCA His Gly Phe Lys Thr Leu Gly Thr Leu Phe Ser Ser Met Met Phe Pro 370 375 380	1212
GAT CGT TCC CCT AGT GAC GTT CAT CTA TAT ACA ACT TTT ATT GGT GGG Asp Arg Ser Pro Ser Asp Val His Leu Tyr Thr Thr Phe Ile Gly Gly 385 390 395	1260
AGT AGG AAC CAG GAA CTA GCC AAA GCT TCC ACT GAC GAA TTA AAA CAA Ser Arg Asn Gln Glu Leu Ala Lys Ala Ser Thr Asp Glu Leu Lys Gln 400 405 410	1308
GTT GTG ACT TCT GAC CTT CAG CGA CTG TTG GGG GTT GAA GGT GAA CCC Val Val Thr Ser Asp Leu Gln Arg Leu Leu Gly Val Glu Gly Glu Pro 415 420 425	1356
GTG TCT GTC AAC CAT TAC TAT TGG AGG AAA GCA TTC CCG TTG TAT GAC Val Ser Val Asn His Tyr Tyr Trp Arg Lys Ala Phe Pro Leu Tyr Asp 430 435 440 445	1404
AGC AGC TAT GAC TCA GTC ATG GAA GCA ATT GAC AAG ATG GAG AAT GAT Ser Ser Tyr Asp Ser Val Met Glu Ala Ile Asp Lys Met Glu Asn Asp 450 455 460	1452
CTA CCT GGG TTC TTC TAT GCA GGT AAT CAT CGA GGG GGG CTC TCT GTT Leu Pro Gly Phe Phe Tyr Ala Gly Asn His Arg Gly Gly Leu Ser Val 465 470 475	1500
GGG AAA TCA ATA GCA TCA GGT TGC AAA GCA GCT GAC CTT GTG ATC TCA Gly Lys Ser Ile Ala Ser Gly Cys Lys Ala Ala Asp Leu Val Ile Ser 480 485 490	1548
TAC CTG GAG TCT TGC TCA AAT GAC AAG AAA CCA AAT GAC AGC TTA TAACATTGTC 1603 Tyr Leu Glu Ser Cys Ser Asn Asp Lys Lys Pro Asn Asp Ser Leu 495 500 505	
AAGGTTTCGTC CCTTTTATC ACTTACTTTG TAAACTTGTA AAATGCAACA AGCCGCCGTG	1663
CGATTAGCCA ACAACTCAGC AAAACCCAGA TTCTCATAAG GCTCACTAAT TCCAGAATAA	1723
ACTATTTATG TAAAA	1738

(2) 서열 4에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 508개 아미노산

(B) 유형 : 아미노산

(D) 형태 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열 기술 :

서열 4

```

Met Ala Ser Gly Ala Val Ala Asp His Gln Ile Glu Ala Val Ser Gly
 1           5           10           15

Lys Arg Val Ala Val Val Gly Ala Gly Val Ser Gly Leu Ala Ala Ala
          20           25           30

Tyr Lys Leu Lys Ser Arg Gly Leu Asn Val Thr Val Phe Glu Ala Asp
      35           40           45

Gly Arg Val Gly Gly Lys Leu Arg Ser Val Met Gln Asn Gly Leu Ile
      50           55           60

Trp Asp Glu Gly Ala Asn Thr Met Thr Glu Ala Glu Pro Glu Val Gly
      65           70           75           80

Ser Leu Leu Asp Asp Leu Gly Leu Arg Glu Lys Gln Gln Phe Pro Ile
          85           90           95

Ser Gln Lys Lys Arg Tyr Ile Val Arg Asn Gly Val Pro Val Met Leu
      100           105           110

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

```

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

355	360	365
Lys Thr Leu Gly Thr Leu Phe Ser Ser Met Met Phe Pro Asp Arg Ser		
370	375	380
Pro Ser Asp Val His Leu Tyr Thr Thr Phe Ile Gly Gly Ser Arg Asn		
385	390	395 400
Gln Glu Leu Ala Lys Ala Ser Thr Asp Glu Leu Lys Gln Val Val Thr		
405	410	415
Ser Asp Leu Gln Arg Leu Leu Gly Val Glu Gly Glu Pro Val Ser Val		
420	425	430
Asn His Tyr Tyr Trp Arg Lys Ala Phe Pro Leu Tyr Asp Ser Ser Tyr		
435	440	445
Asp Ser Val Met Glu Ala Ile Asp Lys Met Glu Asn Asp Leu Pro Gly		
450	455	460
Phe Phe Tyr Ala Gly Asn His Arg Gly Gly Leu Ser Val Gly Lys Ser		
465	470	475 480
Ile Ala Ser Gly Cys Lys Ala Ala Asp Leu Val Ile Ser Tyr Leu Glu		
485	490	495
Ser Cys Ser Asn Asp Lys Lys Pro Asn Asp Ser Leu		
500	505	

(2) 서열 5에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 1691개 염기쌍

(B) 유형 : 핵산

(C) 채 : 일본채

(D) 형태 : 선형

(ii) 분자 유형 : cDNA

(iii) 가설 : 없음

(iv) 안티-센스 : 없음

(vi) 최초 공급원 :

(A) 생물 : 제아 메이스(Zea mays(옥수수))

(vii) 중간체 공급원 :

(B) 클론 : pWDC-4(NRRL B-21260)

(ix) 특성 :

(A) 명칭/키 : CDS

(B) 위치 : 1..1443

(D) 기타 정보 : /생성물="옥수수 프로텍스-1 cDNA"

(xi) 서열 기술 :

서열 5

GCG GAC TGC GTC GTG GTG GGC GGA GGC ATC AGT GGC CTC TGC ACC GCG	48
Ala Asp Cys Val Val Val Gly Gly Gly Ile Ser Gly Leu Cys Thr Ala	
1 5 10 15	
CAG GCG CTG GCC ACG CGG CAC GGC GTC GGG GAC GTG CTT GTC ACG GAG	96
Gln Ala Leu Ala Thr Arg His Gly Val Gly Asp Val Leu Val Thr Glu	
20 25 30	
GCC CGC GCC CGC CCC GGC GGC AAC ATT ACC ACC GTC GAG CGC CCC GAG	144
Ala Arg Ala Arg Pro Gly Gly Asn Ile Thr Thr Val Glu Arg Pro Glu	
35 40 45	
GAA GGG TAC CTC TGG GAG GAG GGT CCC AAC AGC TTC CAG CCC TCC GAC	192
Glu Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp	
50 55 60	
CCC GTT CTC ACC ATG GCC GTG GAC AGC GGA CTG AAG GAT GAC TTG GTT	240
Pro Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Asp Leu Val	
65 70 75 80	
TTT GGG GAC CCA AAC GCG CCG CGT TTC GTG CTG TGG GAG GGG AAG CTG	288
Phe Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly Lys Leu	
85 90 95	
AGG CCC GTG CCA TCC AAG CCC GCC GAC CTC CCG TTC TTC GAT CTC ATG	336
Arg Pro Val Pro Ser Lys Pro Ala Asp Leu Pro Phe Phe Asp Leu Met	
100 105 110	
AGC ATC CCA GGG AAG CTC AGG GCC GGT CTA GGC GCG CTT GGC ATC CGC	384
Ser Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu Gly Ile Arg	
115 120 125	

CCG CCT CCT CCA GGC CGC GAA GAG TCA GTG GAG GAG TTC GTG CGC CGC	432
Pro Pro Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg	
130 135 140	
AAC CTC GGT GCT GAG GTC TTT GAG CGC CTC ATT GAG CCT TTC TGC TCA	480
Asn Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser	
145 150 155 160	
GGT GTC TAT GCT GGT GAT CCT TCT AAG CTC AGC ATG AAG GCT GCA TTT	528
Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe	
165 170 175	
GGG AAG GTT TGG CGG TTG GAA GAA ACT GGA GGT AGT ATT ATT GGT GGA	576
Gly Lys Val Trp Arg Leu Glu Glu Thr Gly Gly Ser Ile Ile Gly Gly	
180 185 190	
ACC ATC AAG ACA ATT CAG GAG AGG AGC AAG AAT CCA AAA CCA CCG AGG	624
Thr Ile Lys Thr Ile Gln Glu Arg Ser Lys Asn Pro Lys Pro Pro Arg	
195 200 205	
GAT GCC CGC CTT CCG AAG CCA AAA GGG CAG ACA GTT GCA TCT TTC AGG	672
Asp Ala Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Ala Ser Phe Arg	
210 215 220	
AAG GGT CTT GCC ATG CTT CCA AAT GCC ATT ACA TCC AGC TTG GGT AGT	720
Lys Gly Leu Ala Met Leu Pro Asn Ala Ile Thr Ser Ser Leu Gly Ser	
225 230 235 240	
AAA GTC AAA CTA TCA TGG AAA CTC ACG AGC ATT ACA AAA TCA GAT GAC	768
Lys Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys Ser Asp Asp	
245 250 255	
AAG GGA TAT GTT TTG GAG TAT GAA ACG CCA GAA GGG GTT GTT TCG GTG	816
Lys Gly Tyr Val Leu Glu Tyr Glu Thr Pro Glu Gly Val Val Ser Val	
260 265 270	
CAG GCT AAA AGT GTT ATC ATG ACT ATT CCA TCA TAT GTT GCT AGC AAC	864
Gln Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val Ala Ser Asn	
275 280 285	
ATT TTG CGT CCA CTT TCA AGC GAT GCT GCA GAT GCT CTA TCA AGA TTC	912
Ile Leu Arg Pro Leu Ser Ser Asp Ala Ala Asp Ala Leu Ser Arg Phe	
290 295 300	

TAT TAT CCA CCG GTT GCT GCT GTA ACT GTT TCG TAT CCA AAG GAA GCA	960
Tyr Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro Lys Glu Ala	
305 310 315 320	
ATT AGA AAA GAA TGC TTA ATT GAT GGG GAA CTC CAG GGC TTT GGC CAG	1008
Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly Phe Gly Gln	
325 330 335	
TTG CAT CCA CGT AGT CAA GGA GTT GAG ACA TTA GGA ACA ATA TAC AGT	1056
Leu His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser	
340 345 350	
TCC TCA CTC TTT CCA AAT CGT GCT CCT GAC GGT AGG GTG TTA CTT CTA	1104
Ser Ser Leu Phe Pro Asn Arg Ala Pro Asp Gly Arg Val Leu Leu Leu	
355 360 365	
AAC TAC ATA GGA GGT GCT ACA AAC ACA GGA ATT GTT TCC AAG ACT GAA	1152
Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly Ile Val Ser Lys Thr Glu	
370 375 380	
AGT GAG CTG GTC GAA GCA GTT GAC CGT GAC CTC CGA AAA ATG CTT ATA	1200
Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile	
385 390 395 400	
AAT TCT ACA GCA GTG GAC CCT TTA GTC CTT GGT GTT CGA GTT TGG CCA	1248
Asn Ser Thr Ala Val Asp Pro Leu Val Leu Gly Val Arg Val Trp Pro	
405 410 415	
CAA GCC ATA CCT CAG TTC CTG GTA GGA CAT CTT GAT CTT CTG GAA GCC	1296
Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu Glu Ala	
420 425 430	
GCA AAA GCT GCC CTG GAC CGA GGT GGC TAC GAT GGG CTG TTC CTA GGA	1344
Ala Lys Ala Ala Leu Asp Arg Gly Gly Tyr Asp Gly Leu Phe Leu Gly	
435 440 445	
GGG AAC TAT GTT GCA GGA GTT GCC CTG GGC AGA TGC GTT GAG GGC GCG	1392
Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala	
450 455 460	
TAT GAA AGT GCC TCG CAA ATA TCT GAC TTC TTG ACC AAG TAT GCC TAC	1440
Tyr Glu Ser Ala Ser Gln Ile Ser Asp Phe Leu Thr Lys Tyr Ala Tyr	
465 470 475 480	
AAG TGATGAAAGA AGTGGAGCGC TACTTGTTAA TCGTTTATGT TGCATAGATG	1493
Lys	
AGGTGCCTCC GGGGAAAAAA AAGCTTGAAT AGTATTTTTT ATTCTTATTT TGTAAATTGC	1553
ATTTCTGTTC TTTTTTCTAT CAGTAATTAG TTATATTTTA GTTCTGTAGG AGATTGTTCT	1613
GTTCACTGCC CTTCAAAGA AATTTTATTT TTCATTCTTT TATGAGAGCT GTGCTACTTA	1673
AAAAAAAAAAAA AAAAAAAAAA	1691

(2) 서열 6에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 481개 아미노산

(B) 유형 : 아미노산

(D) 형태 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열 기술 :

서열 6

```

Ala Asp Cys Val Val Val Gly Gly Gly Ile Ser Gly Leu Cys Thr Ala
 1           5           10           15

Gln Ala Leu Ala Thr Arg His Gly Val Gly Asp Val Leu Val Thr Glu
      20           25           30

Ala Arg Ala Arg Pro Gly Gly Asn Ile Thr Thr Val Glu Arg Pro Glu
      35           40           45

Glu Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp
      50           55           60

Pro Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Asp Leu Val
      65           70           75           80

Phe Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly Lys Leu
      85           90           95

Arg Pro Val Pro Ser Lys Pro Ala Asp Leu Pro Phe Phe Asp Leu Met
      100           105           110

Ser Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu Gly Ile Arg

```

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

355	360	365
Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly Ile Val Ser Lys Thr Glu		
370	375	380
Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile		
385	390	395 400
Asn Ser Thr Ala Val Asp Pro Leu Val Leu Gly Val Arg Val Trp Pro		
405	410	415
Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu Glu Ala		
420	425	430
Ala Lys Ala Ala Leu Asp Arg Gly Gly Tyr Asp Gly Leu Phe Leu Gly		
435	440	445
Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala		
450	455	460
Tyr Glu Ser Ala Ser Gln Ile Ser Asp Phe Leu Thr Lys Tyr Ala Tyr		
465	470	475 480

Lys

(2) 서열 7에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 2061개 염기쌍

(B) 유형 : 핵산

(C) 색 : 일본색

(D) 형태 : 선형

(ii) 분자 유형 : cDNA

(iii) 가설 : 없음

(iv) 안티-센스 : 없음

(vi) 최초 공급원 :

(A) 생물 : 제아 메이스(Zea mays(옥수수))

(vii) 중간체 공급원 :

(B) 클론 : pWDC-3(NRRL B-21259)

(ix) 특성 :

(A) 명칭/키 : CDS

(B) 위치 : 64..1698

(D) 기타 정보 : /생성물="옥수수 프로텍스-2"

(xi) 서열 기술 :

서열 7

CTCTCCTACC TCCACCTCCA CGACAACAAG CAAATCCCCA TCCAGTTCCA AACCCCTAACT	60
CAA ATG CTC GCT TTG ACT GCC TCA GCC TCA TCC GCT TCG TCC CAT CCT	108
Met Leu Ala Leu Thr Ala Ser Ala Ser Ser Ala Ser Ser His Pro	
1 5 10 15	
TAT CGC CAC GCC TCC GCG CAC ACT CGT CGC CCC CGC CTA CGT GCG GTC	156
Tyr Arg His Ala Ser Ala His Thr Arg Arg Pro Arg Leu Arg Ala Val	
20 25 30	
CTC GCG ATG GCG GGC TCC GAC GAC CCC CGT GCA GCG CCC GCC AGA TCG	204
Leu Ala Met Ala Gly Ser Asp Asp Pro Arg Ala Ala Pro Ala Arg Ser	
35 40 45	
GTC GCC GTC GTC GGC GCC GGG GTC AGC GGG CTC GCG GCG GCG TAC AGG	252
Val Ala Val Val Gly Ala Gly Val Ser Gly Leu Ala Ala Ala Tyr Arg	
50 55 60	
CTC AGA CAG AGC GGC GTG AAC GTA ACG GTG TTC GAA GCG GCC GAC AGG	300
Leu Arg Gln Ser Gly Val Asn Val Thr Val Phe Glu Ala Ala Asp Arg	
65 70 75	
GCG GGA GGA AAG ATA CGG ACC AAT TCC GAG GGC GGG TTT GTC TGG GAT	348
Ala Gly Gly Lys Ile Arg Thr Asn Ser Glu Gly Gly Phe Val Trp Asp	
80 85 90 95	
GAA GGA GCT AAC ACC ATG ACA GAA GGT GAA TGG GAG GCC AGT AGA CTG	396
Glu Gly Ala Asn Thr Met Thr Glu Gly Glu Trp Glu Ala Ser Arg Leu	
100 105 110	
ATT GAT GAT CTT GGT CTA CAA GAC AAA CAG CAG TAT CCT AAC TCC CAA	444
Ile Asp Asp Leu Gly Leu Gln Asp Lys Gln Gln Tyr Pro Asn Ser Gln	
115 120 125	
CAC AAG CGT TAC ATT GTC AAA GAT GGA GCA CCA GCA CTG ATT CCT TCG	492
His Lys Arg Tyr Ile Val Lys Asp Gly Ala Pro Ala Leu Ile Pro Ser	

130	135	140	
GAT CCC ATT TCG CTA ATG AAA AGC AGT GTT CTT TCG ACA AAA TCA AAG			540
Asp Pro Ile Ser Leu Met Lys Ser Ser Val Leu Ser Thr Lys Ser Lys			
145	150	155	
ATT GCG TTA TTT TTT GAA CCA TTT CTC TAC AAG AAA GCT AAC ACA AGA			588
Ile Ala Leu Phe Phe Glu Pro Phe Leu Tyr Lys Lys Ala Asn Thr Arg			
160	165	170	175
AAC TCT GGA AAA GTG TCT GAG GAG CAC TTG AGT GAG AGT GTT GGG AGC			636
Asn Ser Gly Lys Val Ser Glu Glu His Leu Ser Glu Ser Val Gly Ser			
180	185	190	
TTC TGT GAA CGC CAC TTT GGA AGA GAA GTT GTT GAC TAT TTT GTT GAT			684
Phe Cys Glu Arg His Phe Gly Arg Glu Val Val Asp Tyr Phe Val Asp			
195	200	205	
CCA TTT GTA GCT GGA ACA AGT GCA GGA GAT CCA GAG TCA CTA TCT ATT			732
Pro Phe Val Ala Gly Thr Ser Ala Gly Asp Pro Glu Ser Leu Ser Ile			
210	215	220	
CGT CAT GCA TTC CCA GCA TTG TGG AAT TTG GAA AGA AAG TAT GGT TCA			780
Arg His Ala Phe Pro Ala Leu Trp Asn Leu Glu Arg Lys Tyr Gly Ser			
225	230	235	
GTT ATT GTT GGT GCC ATC TTG TCT AAG CTA GCA GCT AAA GGT GAT CCA			828
Val Ile Val Gly Ala Ile Leu Ser Lys Leu Ala Ala Lys Gly Asp Pro			
240	245	250	255
GTA AAG ACA AGA CAT GAT TCA TCA GGG AAA AGA AGG AAT AGA CGA GTG			876
Val Lys Thr Arg His Asp Ser Ser Gly Lys Arg Arg Asn Arg Arg Val			
260	265	270	
TCG TTT TCA TTT CAT GGT GGA ATG CAG TCA CTA ATA AAT GCA CTT CAC			924
Ser Phe Ser Phe His Gly Gly Met Gln Ser Leu Ile Asn Ala Leu His			
275	280	285	
AAT GAA GTT GGA GAT GAT AAT GTG AAG CTT GGT ACA GAA GTG TTG TCA			972
Asn Glu Val Gly Asp Asp Asn Val Lys Leu Gly Thr Glu Val Leu Ser			
290	295	300	



TCT GTT GAT TCG AAG GAT AGC GGT GAC AAG GAC CTT GCT AGT AAC CAA	1068
Ser Val Asp Ser Lys Asp Ser Gly Asp Lys Asp Leu Ala Ser Asn Gln	
320 325 330 335	
ACC TTT GAT GCT GTT ATA ATG ACA GCT CCA TTG TCA AAT GTC CGG AGG	1116
Thr Phe Asp Ala Val Ile Met Thr Ala Pro Leu Ser Asn Val Arg Arg	
340 345 350	
ATG AAG TTC ACC AAA GGT GGA GCT CCG GTT GTT CTT GAC TTT CTT CCT	1164
Met Lys Phe Thr Lys Gly Gly Ala Pro Val Val Leu Asp Phe Leu Pro	
355 360 365	
AAG ATG GAT TAT CTA CCA CTA TCT CTC ATG GTG ACT GCT TTT AAG AAG	1212
Lys Met Asp Tyr Leu Pro Leu Ser Leu Met Val Thr Ala Phe Lys Lys	
370 375 380	
GAT GAT GTC AAG AAA CCT CTG GAA GGA TTT GGG GTC TTA ATA CCT TAC	1260
Asp Asp Val Lys Lys Pro Leu Glu Gly Phe Gly Val Leu Ile Pro Tyr	
385 390 395	
AAG GAA CAG CAA AAA CAT GGT CTG AAA ACC CTT GGG ACT CTC TTT TCC	1308
Lys Glu Gln Gln Lys His Gly Leu Lys Thr Leu Gly Thr Leu Phe Ser	
400 405 410 415	
TCA ATG ATG TTC CCA GAT CGA GCT CCT GAT GAC CAA TAT TTA TAT ACA	1356
Ser Met Met Phe Pro Asp Arg Ala Pro Asp Asp Gln Tyr Leu Tyr Thr	
420 425 430	
ACA TTT GTT GGG GGT AGC CAC AAT AGA GAT CTT GCT GGA GCT CCA ACG	1404
Thr Phe Val Gly Gly Ser His Asn Arg Asp Leu Ala Gly Ala Pro Thr	
435 440 445	
TCT ATT CTG AAA CAA CTT GTG ACC TCT GAC CTT AAA AAA CTC TTG GGC	1452
Ser Ile Leu Lys Gln Leu Val Thr Ser Asp Leu Lys Lys Leu Leu Gly	
450 455 460	
GTA GAG GGG CAA CCA ACT TTT GTC AAG CAT GTA TAC TGG GGA AAT GCT	1500
Val Glu Gly Gln Pro Thr Phe Val Lys His Val Tyr Trp Gly Asn Ala	
465 470 475	
TTT CCT TTG TAT GGC CAT GAT TAT AGT TCT GTA TTG GAA GCT ATA GAA	1548
Phe Pro Leu Tyr Gly His Asp Tyr Ser Ser Val Leu Glu Ala Ile Glu	
480 485 490 495	

AAG ATG GAG AAA AAC CTT CCA GGG TTC TTC TAC GCA GGA AAT AGC AAG Lys Met Glu Lys Asn Leu Pro Gly Phe Phe Tyr Ala Gly Asn Ser Lys	1596
500 505 510	
GAT GGG CTT GCT GTT GGA AGT GTT ATA GCT TCA GGA AGC AAG GCT GCT Asp Gly Leu Ala Val Gly Ser Val Ile Ala Ser Gly Ser Lys Ala Ala	1644
515 520 525	
GAC CTT GCA ATC TCA TAT CTT GAA TCT CAC ACC AAG CAT AAT AAT TCA Asp Leu Ala Ile Ser Tyr Leu Glu Ser His Thr Lys His Asn Asn Ser	1692
530 535 540	
CAT TGAAAGTGTC TGACCTATCC TCTAGCAGTT GTCGACAAAT TTCTCCAGTT His	1745
545	
CATGTACAGT AGAAACCGAT GCGTTGCAGT TTCAGAACAT CTTCACTTCT TCAGATATTA	1805
ACCCCTTCGTT GAACATCCAC CAGAAAGGTA GTCACATGTG TAAGTGGGAA AATGAGGTTA	1865
AAAACCTATTA TGGCGGCCGA AATGTTCCCTT TTTGTTTTCC TCACAAGTGG CCTACGACAC	1925
TTGATGTTGG AAATACATTT AAATTGTTG AATTGTTTGA GAACACATGC GTGACGTGTA	1985
ATATTTGCCT ATTGTGATTT TAGCAGTAGT CTTGGCCAGA TTATGCTTTA CGCCTTTAAA	2045
AAAAAAAAAA AAAAAA	2061

(2) 서열 8에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 544개 아미노산

(B) 유형 : 아미노산

(D) 형태 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열 기술 :

서열 8

Met	Leu	Ala	Leu	Thr	Ala	Ser	Ala	Ser	Ser	Ala	Ser	Ser	His	Pro	Tyr
1				5						10				15	
Arg	His	Ala	Ser	Ala	His	Thr	Arg	Arg	Pro	Arg	Leu	Arg	Ala	Val	Leu
			20						25					30	

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

Phe Ser Phe His Gly Gly Met Gln Ser Leu Ile Asn Ala Leu His Asn
 275 280 285
 Glu Val Gly Asp Asp Asn Val Lys Leu Gly Thr Glu Val Leu Ser Leu
 290 295 300
 Ala Cys Thr Phe Asp Gly Val Pro Ala Leu Gly Arg Trp Ser Ile Ser
 305 310 315 320
 Val Asp Ser Lys Asp Ser Gly Asp Lys Asp Leu Ala Ser Asn Gln Thr
 325 330 335
 Phe Asp Ala Val Ile Met Thr Ala Pro Leu Ser Asn Val Arg Arg Met
 340 345 350
 Lys Phe Thr Lys Gly Gly Ala Pro Val Val Leu Asp Phe Leu Pro Lys
 355 360 365
 Met Asp Tyr Leu Pro Leu Ser Leu Met Val Thr Ala Phe Lys Lys Asp
 370 375 380
 Asp Val Lys Lys Pro Leu Glu Gly Phe Gly Val Leu Ile Pro Tyr Lys
 385 390 395 400
 Glu Gln Gln Lys His Gly Leu Lys Thr Leu Gly Thr Leu Phe Ser Ser
 405 410 415
 Met Met Phe Pro Asp Arg Ala Pro Asp Asp Gln Tyr Leu Tyr Thr Thr
 420 425 430
 Phe Val Gly Gly Ser His Asn Arg Asp Leu Ala Gly Ala Pro Thr Ser
 435 440 445
 Ile Leu Lys Gln Leu Val Thr Ser Asp Leu Lys Lys Leu Leu Gly Val
 450 455 460
 Glu Gly Gln Pro Thr Phe Val Lys His Val Tyr Trp Gly Asn Ala Phe
 465 470 475 480
 Pro Leu Tyr Gly His Asp Tyr Ser Ser Val Leu Glu Ala Ile Glu Lys
 485 490 495
 Met Glu Lys Asn Leu Pro Gly Phe Phe Tyr Ala Gly Asn Ser Lys Asp
 500 505 510
 Gly Leu Ala Val Gly Ser Val Ile Ala Ser Gly Ser Lys Ala Ala Asp
 515 520 525
 Leu Ala Ile Ser Tyr Leu Glu Ser His Thr Lys His Asn Asn Ser His
 530 535 540

(2) 서열 9에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 1811개 염기쌍

(B) 유형 : 핵산

- (C) 소재 : 일본소재
 (D) 형태 : 선형
 (ii) 분자 유형 : cDNA
 (iii) 가설 : 없음
 (vi) 최초 공급원 :
 (A) 생물 : 트리티쿰 아에스티쿰(*Triticum aestivum*)(밀)
 (vii) 중간체 공급원 :
 (B) 클론 : pWDC-13(NRRL B-21545)
 (ix) 특성 :
 (A) 명칭/키 : CDS
 (B) 위치 : 3..1589
 (D) 기타 정보 : /생성물="밀 프로텍스-1"
 (xi) 서열 기술 :

서열 9

GC GCA ACA ATG GCC ACC GCC ACC GTC GCG GCC GCG TCG CCG CTC CGC	47
Ala Thr Met Ala Thr Ala Thr Val Ala Ala Ala Ser Pro Leu Arg	
1 5 10 15	
GGC AGG GTC ACC GGG CGC CCA CAC CGC GTC CGC CCG CGT TGC GCT ACC	95
Gly Arg Val Thr Gly Arg Pro His Arg Val Arg Pro Arg Cys Ala Thr	
20 25 30	
GCG AGC AGC GCG ACC GAG ACT CCG GCG GCG CCC GGC GTG CGG CTG TCC	143
Ala Ser Ser Ala Thr Glu Thr Pro Ala Ala Pro Gly Val Arg Leu Ser	

35	40	45	
GCG GAA TGC GTC ATT GTG GGC GCC GGC ATC AGC GGC CTC TGC ACC GCG			191
Ala Glu Cys Val Ile Val Gly Ala Gly Ile Ser Gly Leu Cys Thr Ala			
50	55	60	
CAG GCG CTG GCC ACC CGA TAC GGC GTC AGC GAC CTG CTC GTC ACG GAG			239
Gln Ala Leu Ala Thr Arg Tyr Gly Val Ser Asp Leu Leu Val Thr Glu			
65	70	75	
GCC CGC GAC CGC CCG GGC GGG AAC ATC ACC ACC GTC GAG CGT CCC GAC			287
Ala Arg Asp Arg Pro Gly Gly Asn Ile Thr Thr Val Glu Arg Pro Asp			
80	85	90	95
GAG GGG TAC CTG TGG GAG GAG GGA CCC AAC AGC TTC CAG CCC TCC GAC			335
Glu Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp			
100	105	110	
CCG GTC CTC ACC ATG GCC GTG GAC AGC GGG CTC AAG GAT GAC TTG GTG			383
Pro Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Asp Leu Val			
115	120	125	
TTC GGG GAC CCC AAC GCG CCC CGG TTC GTG CTG TGG GAG GGG AAG CTG			431
Phe Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly Lys Leu			
130	135	140	
AGG CCG GTG CCG TCG AAG CCA GGC GAC CTG CCT TTC TTC AGC CTC ATG			479
Arg Pro Val Pro Ser Lys Pro Gly Asp Leu Pro Phe Phe Ser Leu Met			
145	150	155	
AGT ATC CCT GGG AAG CTC AGG GCC GGC CTT GGC GCG CTC GGC ATT CGC			527
Ser Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu Gly Ile Arg			
160	165	170	175
CCA CCT CCT CCA GGG CGC GAG GAG TCG GTG GAG GAG TTT GTG CGC CGC			575
Pro Pro Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg			
180	185	190	
AAC CTC GGT GCC GAG GTC TTT GAG CGC CTC ATC GAG CCT TTC TGC TCA			623
Asn Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser			
195	200	205	
GGT GTA TAT GCT GGT GAT CCT TCG AAG CTT AGT ATG AAG GCT GCA TTT			671
Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe			
210	215	220	

GGG AAG GTC TGG AGG TTG GAG GAG ATT GGA GGT AGT ATT ATT GGT GGA Gly Lys Val Trp Arg Leu Glu Glu Ile Gly Gly Ser Ile Ile Gly Gly 225 230 235	719
ACC ATC AAG GCG ATT CAG GAT AAA GGG AAG AAC CCC AAA CCG CCA AGG Thr Ile Lys Ala Ile Gln Asp Lys Gly Lys Asn Pro Lys Pro Pro Arg 240 245 250 255	767
GAT CCC CGA CTT CCG GCA CCA AAG GGA CAG ACG GTG GCA TCT TTC AGG Asp Pro Arg Leu Pro Ala Pro Lys Gly Gln Thr Val Ala Ser Phe Arg 260 265 270	815
AAG GGT CTA GCC ATG CTC CCG AAT GCC ATC GCA TCT AGG CTG GGT AGT Lys Gly Leu Ala Met Leu Pro Asn Ala Ile Ala Ser Arg Leu Gly Ser 275 280 285	863
AAA GTC AAG CTG TCA TGG AAG CTT ACG AGC ATT ACA AAG GCG GAC AAC Lys Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys Ala Asp Asn 290 295 300	911
CAA GGA TAT GTA TTA GGT TAT GAA ACA CCA GAA GGA CTT GTT TCA GTG Gln Gly Tyr Val Leu Gly Tyr Glu Thr Pro Glu Gly Leu Val Ser Val 305 310 315	959
CAG GCT AAA AGT GTT ATC ATG ACC ATC CCG TCA TAT GTT GCT AGT GAT Gln Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val Ala Ser Asp 320 325 330 335	1007
ATC TTG CGC CCA CTT TCA ATT GAT GCA GCA GAT GCA CTC TCA AAA TTC Ile Leu Arg Pro Leu Ser Ile Asp Ala Ala Asp Ala Leu Ser Lys Phe 340 345 350	1055
TAT TAT CCG CCA GTT GCT GCT GTA ACT GTT TCA TAT CCA AAA GAA GCT Tyr Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro Lys Glu Ala 355 360 365	1103
ATT AGA AAA GAA TGC TTA ATT GAT GGG GAG CTC CAG GGT TTC GGC CAG Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly Phe Gly Gln 370 375 380	1151
TTG CAT CCA CGT AGC CAA GGA GTC GAG ACT TTA GGG ACA ATA TAT AGC Leu His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser 385 390 395	1199

서열 10

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

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

```

Asp Leu Val Gly Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile Asn
  435                      440                      445

Pro Arg Ala Ala Asp Pro Leu Ala Leu Gly Val Arg Val Trp Pro Gln
  450                      455                      460

Ala Ile Pro Gln Phe Leu Ile Gly His Leu Asp Arg Leu Ala Ala Ala
  465                      470                      475                      480

Lys Ser Ala Leu Gly Gln Gly Gly Tyr Asp Gly Leu Phe Leu Gly Gly
          485                      490                      495

Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Ile Glu Gly Ala Tyr
      500                      505                      510

Glu Ser Ala Ser Gln Val Ser Asp Phe Leu Thr Lys Tyr Ala Tyr Lys
      515                      520                      525

```

(2) 서열 11에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 1847개 염기쌍

(B) 유형 : 핵산

(C) 채 : 일본채

(D) 형태 : 선형

(ii) 분자 유형 : cDNA

(iii) 가설 : 없음

(vi) 최초 공급원 :

(A) 생물 : 대두

(vii) 중간체 공급원 :

(B) 클론 : pWDC-12(NRRL B-21516)

(ix) 특성 :

(A) 명칭/키 : CDS

(B) 위치 : 55..1683

(D) 기타 정보 : /생성물="대두 프로텍스-1"

(xi) 서열 기술 :

서열 11

CTTTAGCACA GTGTTGAAGA TAACGAACGA ATAGTGCCAT TACTGTAACC AACC ATG	57
Met	
1	
GTT TCC GTC TTC AAC GAG ATC CTA TTC CCG CCG AAC CAA ACC CTT CTT	105
Val Ser Val Phe Asn Glu Ile Leu Phe Pro Pro Asn Gln Thr Leu Leu	
5 10 15	
CGC CCC TCC CTC CAT TCC CCA ACC TCT TTC TTC ACC TCT CCC ACT CGA	153
Arg Pro Ser Leu His Ser Pro Thr Ser Phe Phe Thr Ser Pro Thr Arg	
20 25 30	
AAA TTC CCT CGC TCT CGC CCT AAC CCT ATT CTA CGC TGC TCC ATT GCG	201
Lys Phe Pro Arg Ser Arg Pro Asn Pro Ile Leu Arg Cys Ser Ile Ala	
35 40 45	
GAG GAA TCC ACC GCG TCT CCG CCC AAA ACC AGA GAC TCC GCC CCC GTG	249
Glu Glu Ser Thr Ala Ser Pro Pro Lys Thr Arg Asp Ser Ala Pro Val	
50 55 60 65	
GAC TGC GTC GTC GTC GGC GGA GGC GTC AGC GGC CTC TGC ATC GCC CAG	297
Asp Cys Val Val Val Gly Gly Gly Val Ser Gly Leu Cys Ile Ala Gln	
70 75 80	
GCC CTC GCC ACC AAA CAC GCC AAT GCC AAC GTC GTC GTC ACG GAG GCC	345
Ala Leu Ala Thr Lys His Ala Asn Ala Asn Val Val Val Thr Glu Ala	
85 90 95	
CGA GAC CGC GTC GGC GGC AAC ATC ACC ACG ATG GAG AGG GAC GGA TAC	393
Arg Asp Arg Val Gly Gly Asn Ile Thr Thr Met Glu Arg Asp Gly Tyr	
100 105 110	
CTC TGG GAA GAA GGC CCC AAC AGC TTC CAG CCT TCT GAT CCA ATG CTC	441
Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Pro Met Leu	
115 120 125	
ACC ATG GTG GTG GAC AGT GGT TTA AAG GAT GAG CTT GTT TTG GGG GAT	489
Thr Met Val Val Asp Ser Gly Leu Lys Asp Glu Leu Val Leu Gly Asp	
130 135 140 145	
CCT GAT GCA CCT CGG TTT GTG TTG TGG AAC AGG AAG TTG AGG CCG GTG	537
Pro Asp Ala Pro Arg Phe Val Leu Trp Asn Arg Lys Leu Arg Pro Val	
150 155 160	
CCC GGG AAG CTG ACT GAT TTG CCT TTC TTT GAC TTG ATG AGC ATT GGT	585

Pro Gly Lys Leu Thr Asp Leu Pro Phe Phe Asp Leu Met Ser Ile Gly	
165 170 175	
GGC AAA ATC AGG GCT GGC TTT GGT GCG CTT GGA ATT CGG CCT CCT CCT	633
Gly Lys Ile Arg Ala Gly Phe Gly Ala Leu Gly Ile Arg Pro Pro Pro	
180 185 190	
CCA GGT CAT GAG GAA TCG GTT GAA GAG TTT GTT CGT CGG AAC CTT GGT	681
Pro Gly His Glu Glu Ser Val Glu Glu Phe Val Arg Arg Asn Leu Gly	
195 200 205	
GAT GAG GTT TTT GAA CGG TTG ATA GAG CCT TTT TGT TCA GGG GTC TAT	729
Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr	
210 215 220 225	
GCA GGC GAT CCT TCA AAA TTA AGT ATG AAA GCA GCA TTC GGG AAA GTT	777
Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe Gly Lys Val	
230 235 240	
TGG AAG CTG GAA AAA AAT GGT GGT AGC ATT ATT GGT GGA ACT TTC AAA	825
Trp Lys Leu Glu Lys Asn Gly Gly Ser Ile Ile Gly Gly Thr Phe Lys	
245 250 255	
GCA ATA CAA GAG AGA AAT GGA GCT TCA AAA CCA CCT CGA GAT CCG CGT	873
Ala Ile Gln Glu Arg Asn Gly Ala Ser Lys Pro Pro Arg Asp Pro Arg	
260 265 270	
CTG CCA AAA CCA AAA GGT CAG ACT GTT GGA TCT TTC CGG AAG GGA CTT	921
Leu Pro Lys Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys Gly Leu	
275 280 285	
ACC ATG TTG CCT GAT GCA ATT TCT GCC AGA CTA GGC AAC AAA GTA AAG	969
Thr Met Leu Pro Asp Ala Ile Ser Ala Arg Leu Gly Asn Lys Val Lys	
290 295 300 305	
TTA TCT TGG AAG CTT TCA AGT ATT AGT AAA CTG GAT AGT GGA GAG TAC	1017
Leu Ser Trp Lys Leu Ser Ser Ile Ser Lys Leu Asp Ser Gly Glu Tyr	
310 315 320	
AGT TTG ACA TAT GAA ACA CCA GAA GGA GTG GTT TCT TTG CAG TGC AAA	1065
Ser Leu Thr Tyr Glu Thr Pro Glu Gly Val Val Ser Leu Gln Cys Lys	
325 330 335	
ACT GTT GTC CTG ACC ATT CCT TCC TAT GTT GCT AGT ACA TTG CTG CGT	1113
Thr Val Val Leu Thr Ile Pro Ser Tyr Val Ala Ser Thr Leu Leu Arg	

340	345	350	
CCT CTG TCT GCT GCT GCT GCA GAT GCA CTT TCA AAG TTT TAT TAC CCT			1161
Pro Leu Ser Ala Ala Ala Ala Asp Ala Leu Ser Lys Phe Tyr Tyr Pro			
355	360	365	
CCA GTT GCT GCA GTT TCC ATA TCC TAT CCA AAA GAA GCT ATT AGA TCA			1209
Pro Val Ala Ala Val Ser Ile Ser Tyr Pro Lys Glu Ala Ile Arg Ser			
370	375	380	385
GAA TGC TTG ATA GAT GGT GAG TTG AAG GGG TTT GGT CAA TTG CAT CCA			1257
Glu Cys Leu Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln Leu His Pro			
390	395	400	
CGT AGC CAA GGA GTG GAA ACA TTA GGA ACT ATA TAC AGC TCA TCA CTA			1305
Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu			
405	410	415	
TTC CCC AAC CGA GCA CCA CCT GGA AGG GTT CTA CTC TTG AAT TAC ATT			1353
Phe Pro Asn Arg Ala Pro Pro Gly Arg Val Leu Leu Leu Asn Tyr Ile			
420	425	430	
GGA GGA GCA ACT AAT ACT GGA ATT TTA TCG AAG ACG GAC AGT GAA CTT			1401
Gly Gly Ala Thr Asn Thr Gly Ile Leu Ser Lys Thr Asp Ser Glu Leu			
435	440	445	
GTG GAA ACA GTT GAT CGA GAT TTG AGG AAA ATC CTT ATA AAC CCA AAT			1449
Val Glu Thr Val Asp Arg Asp Leu Arg Lys Ile Leu Ile Asn Pro Asn			
450	455	460	465
GCC CAG GAT CCA TTT GTA GTG GGG GTG AGA CTG TGG CCT CAA GCT ATT			1497
Ala Gln Asp Pro Phe Val Val Gly Val Arg Leu Trp Pro Gln Ala Ile			
470	475	480	
CCA CAG TTC TTA GTT GGC CAT CTT GAT CTT CTA GAT GTT GCT AAA GCT			1545
Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu Asp Val Ala Lys Ala			
485	490	495	
TCT ATC AGA AAT ACT GGG TTT GAA GGG CTC TTC CTT GGG GGT AAT TAT			1593
Ser Ile Arg Asn Thr Gly Phe Glu Gly Leu Phe Leu Gly Gly Asn Tyr			
500	505	510	
GTG TCT GGT GTT GCC TTG GGA CGA TGC GTT GAG GGA GCC TAT GAG GTA			1641
Val Ser Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu Val			
515	520	525	
GCA GCT GAA GTA AAC GAT TTT CTC ACA AAT AGA GTG TAC AAA			1683
Ala Ala Glu Val Asn Asp Phe Leu Thr Asn Arg Val Tyr Lys			
530	535	540	
TAGTAGCAGT TTTTGTTTT GTGGTGAAT GGGTGATGGG ACTCTCGTGT TCCATTGAAT			1743
TATAATAATG TGAAAGTTTC TCAAATTCGT TCGATAGGTT TTTGGCGGCT TCTATTGCTG			1803
ATAATGTAAA ATCCTCTTTA AGTTTGAAAA AAAAAAAAAA AAAA			1847

(2) 서열 12에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 543개 아미노산

(B) 유형 : 아미노산

(D) 형태 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열 기술 :

서열 12

```

Met Val Ser Val Phe Asn Glu Ile Leu Phe Pro Pro Asn Gln Thr Leu
 1           5           10           15

Leu Arg Pro Ser Leu His Ser Pro Thr Ser Phe Phe Thr Ser Pro Thr
          20           25           30

Arg Lys Phe Pro Arg Ser Arg Pro Asn Pro Ile Leu Arg Cys Ser Ile
          35           40           45

Ala Glu Glu Ser Thr Ala Ser Pro Pro Lys Thr Arg Asp Ser Ala Pro
          50           55           60

Val Asp Cys Val Val Val Gly Gly Gly Val Ser Gly Leu Cys Ile Ala
65           70           75           80

Gln Ala Leu Ala Thr Lys His Ala Asn Ala Asn Val Val Val Thr Glu
          85           90           95

Ala Arg Asp Arg Val Gly Gly Asn Ile Thr Thr Met Glu Arg Asp Gly
          100          105          110

Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Pro Met

```

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

355	360	365
Pro Pro Val Ala Ala Val Ser Ile Ser Tyr Pro Lys Glu Ala Ile Arg		
370	375	380
Ser Glu Cys Leu Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln Leu His		
385	390	395 400
Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser		
405	410	415
Leu Phe Pro Asn Arg Ala Pro Pro Gly Arg Val Leu Leu Leu Asn Tyr		
420	425	430
Ile Gly Gly Ala Thr Asn Thr Gly Ile Leu Ser Lys Thr Asp Ser Glu		
435	440	445
Leu Val Glu Thr Val Asp Arg Asp Leu Arg Lys Ile Leu Ile Asn Pro		
450	455	460
Asn Ala Gln Asp Pro Phe Val Val Gly Val Arg Leu Trp Pro Gln Ala		
465	470	475 480
Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu Asp Val Ala Lys		
485	490	495
Ala Ser Ile Arg Asn Thr Gly Phe Glu Gly Leu Phe Leu Gly Gly Asn		
500	505	510
Tyr Val Ser Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu		
515	520	525
Val Ala Ala Glu Val Asn Asp Phe Leu Thr Asn Arg Val Tyr Lys		
530	535	540

(2) 서열 13에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 583개 염기쌍

(B) 유형 : 핵산

(C) 채 : 일본채

(D) 형태 : 선형

(ii) 분자 유형 : DNA(게놈)

(iii) 가설 : 없음

(ix) 특성 :

(A) 명칭/키 : 프로모터

(B) 위치 : 1..583

(D) 기타 정보 : /기능="아라비도시스 프로토스-1 프로모터"

(xi) 서열 기술 :

서열 13

```

GAATTCGGAT CGAATTATAT AATTATCATA AATTTGAATA AGCATGTTGC CTTTATTAA      60
AGAGGTTTAA TAAAGTTTGG TAATAATGGA CTTTGACTTC AAACCTGATT CTCATGTAAT      120
TAATTAATAT TTACATCAAA ATTTGGTCAC TAATATTACC AAATTAATAT ACTAAAATGT      180
TAATTCGCAA ATAAAACACT AATTCCAAAT AAAGGGTCAT TATGATAAAC ACGTATTGAA      240
CTTGATAAAG CAAAGCAAAA ATAATGGGTT TCAAGGTTTG GGTTATATAT GACAAAAAAA      300
AAAAAAGGTT TGGTTATATA TCTATTGGGC CTATAACCAT GTTATACAAA TTTGGGCCTA      360
ACTAAAATAA TAAAAATAAC GTAATGGTCC TTTTATATTT TGGGTCAAAC CCAACTCTAA      420
ACCCAAACCA AAGAAAAAGT ATACGGTACG GTACACAGAC TTATGGTGTG TGTGATTGCA      480
GGTGAATATT TCTCGTCGTC TTCTCCTTTC TTCTGAAGAA GATTACCCAA TCTGAAAAAA      540
ACCAAGAAGC TGACAAAATT CCGAATTCTC TGCGATTTCC ATG                          583

```

(2) 서열 14에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 3848개 염기쌍

(B) 유형 : 핵산

(C) 채 : 일본채

(D) 형태 : 선형

(ii) 분자 유형 : DNA(게놈)

(iii) 가설 : 없음

(ix) 특성 :

(A) 명칭/키 : 프로모터

(B) 위치 : 1..3848

(D) 기타 정보 : /기능="옥수수 프로텍스-1 프로모터"

(xi) 서열 기술 :

서열 14

TCGATCTTTC TAGGCTGATC CCCAAATCTT CCTCCGAAGC CCCTGGCGCC TCTGCCCCTT	60
GGAGCTGGTG GCCTGAAAGA GCTTTGCTGT TGCCCCGAAG ATTGTGAGGT ATATTGTGAC	120
CTCTGAGACT GACTTCCTTT GTCGTCACTT TGAGTGGAGT TATGGATTGA CCTGACGTGC	180
CTCAGATGGA TTCTTCCTCC GAAGCCCCTG GTCATTTCCG AGAATCTGTA ATCTTATTCC	240
CTTCTTTGGC GAAAACTGT CAGCTTGGAT GTACTCATCC ATCTTCTGAA GCAGCTTCTC	300
CAGAGTTTGT GGAGGCTTCC TGGCGAAATA TTGGGCTGTA GGTCCCTGGAC GAAGACCCTT	360
GATCATGGCC TCAATGACAA TCTCATTGGG CACCGTAGGC GCTTGTGCCC TCAATCGCAA	420
GAACCTTCGT ACATATGCCT GAAGGTATTC TTCGTGATCT TGTGTGCATT GGAACAGAGC	480
CTGAGCTGTG ACCGACTTCG TTTGAAAGCC TTGGAAGCTA GTAACCAACA TGTGCTTAAG	540
CTTCTGCCAC GACGTGATAG TCCCTGGCCG AAGAGAAGAA TACCATGTTT GGGCTACATT	600
CCGGACTGCC ATGACGAAGG ACTTCGCCAT GACTACAGTG TTGACCCCAT ACGAAGATAT	660
AGTTGCTTCG TAGCTCATCA GAAACTGCTT TGGATCTGAG TGCCCATCAT ACATGGGGAG	720
CTGAGGTGGC TTGTATGATG GGGGCCATGG GGTAGCCTGC AGTTCTGCTG CCAAGGGAGA	780
AGCATCATCA AAAGTAAAGG CATCATGATT AAAATCATCA TACCATCCAT CCTCGTTGAA	840
TAAGCCTTCT TGACGAAGCT CCCTGTGTTG GGGCCTTCGA TCTTGTTTAT CTTGAACAAG	900
ATGACGCACT TCTTCAGTGG CTTCGTCGAT CTTTCTTTGG AGATCAGCCA GTCGCACCAT	960
GTTCTCCTTC TTTCTTTGTA CTTGTTGATG GATGATCTCC ATGTCCCTGA TCTCTTGGTC	1020
CAACTCCTCC TCTTGAGTG TCAGACTGGT GGCTTTCCTC TTCTGGCTTC GAGCCTCTCG	1080
AAGAGAAAGA GTTTCTTGAT TTGGGTCCAG CGGCTGCAGT GCAGTGGTCC CTGGTGCTGA	1140

AGCTTTCTTC GGTGGCATGA CAAAGGTCAG TGCTTGCCGA AGGTGGTCGA AAAGGGTTCA	1200
CTAGAGGTGG GAGCCAATGT TGGGGACTTC TCAAGTGCTA TGAGTTAAGA ACAAGGCAAC	1260
ACAAAATGTT AAATATTAAT AGCTTTTCATC TTTCGAAGCA TTATTTCCCT TTGGGTATAA	1320
TGATCTTCAG ACGAAAGAGT CCTTCATCAT TGCATATAT GTTAATAGAA GGAGGAGCAT	1380
ATGAAATGTA AGAGACAACA TGAACAATCG TGTAGCATTG TTAATTCATC ATCATTTTAT	1440
TATTATGGAA AAATAGAAAC AATATTGAAT TACAAATGTA CCTTTGGCTT GACAGAAGAT	1500
AAAAGTACAA GCTTGACGCA CGAGCAAGTA CAAGTCAGTG TGAACAGTAC GGGGGTACTG	1560
TTTCATCTATT TATAGGCACA GGACACAGCC TGTGAGAAAT TACAGTCATG CCCTTTACAT	1620
TTACTATTGA CTTATAGAAA AATCTATGAG GACTGGATAG CCTTTTCCCC TTTAAGTCGG	1680
TGCCTTTTTTC CGCGATTAA GCGAATCTCC CTTGCGCATA GCTTCGGAGC ATCGGCAACC	1740
TTTCGTCACGA TCATGCCCTT CTCATTGTGT ATGCTTTTAA TCCTGAATTC GAAGGTACCT	1800
GTCCATAAAC CATACTTGGA AGACATTGTT AAATTATGTT TTTGAGGACC TTCGGAGGAC	1860
GAAGGCCCCC AACAGTCGTG TTTTGTAGGA CCTTCGGAAG ATGAAGGCCC CCAACAAGAC	1920
CTATCCATAA AACCAACCTA TCCACAAAAC CGACCCCAT CACCCTTCAT TTGCCTCACC	1980
AACAACCCTA ATTAGGTTGT TGGTTTAAAT TTTTATAGGT CAATTGGTC ATCACCATCC	2040
ACTGTCACCT CACAACTCA ATATCAATAA ACAGACTCAA TCACCCAAAC TGACCATACC	2100
CATAAAACCG CCCACCCCTT CTAGCGCCTC GCCAGAAACC AGAAACCCTG ATTCAGAGTT	2160
CAAACCTAAA ACGACCATAA CTTTCACCTT GGAACCTGAA TCAGGTCCAT TTTTTCCTAA	2220
ATCACACAAA ATTAAATTTC GCATCCGATA ATCAAGCCAT CTCTTCACTA TGGTTTAAAG	2280
TGTTGCTCAC ACTAGTGAT TTATGGACTA ATCACCTGTG TATCTCATAC AATAACATAT	2340
CAGTACATCT AAGTTGTTAC TCAATTACCA AAACCGAATT ATAGCCTTCG AAAAAGGTTA	2400
TCGACTAGTC ACTCAATTAC CAAAATAA CTTTAGACTT TCATGTATGA CATCCAACAT	2460
GACACTGTAC TGGACTAAAC CACCTTTCAA GCTACACAAG GAGCAAAAAT AACTAATTTT	2520

CGTAGTTGTA GGAGCTAAAG TATATGTCCA CAACAATAGT TAAGGGAAGC CCCCAAGGAC	2580
TTAAAAGTCC TTTTACCTCT TGAAACTTTT GTCGTGGTCT ACTTTTTCAC TTTAAACTTC	2640
AAAATTTGAC ATTTTATCAC CCCTTAACTC TTAACCACAT TTAAATTACA TTCTTACTAG	2700
ATTATAGATG ATTTTGTGTG GAAAAGTTTT TAAGACATGT TTACACATTG ATTAAAATCA	2760
TTGTTCATAT TTCCTAGAGT TAAATCTAAT CTTATTAAAA CTATTAGAGA TACTTTCACG	2820
AGCTCTAAAT ATTTTATTTT TTTTATTATG GAATTTTGTG AGAATTCTTA TAGACCTTTT	2880
TTGTGGTTT AAAAGCCTTG CCATGTTTTT AACAAAGTTTT TTTTCTATTT TTTGAAATTT	2940
TCTTGGAAC CACTTCTAAC CCGGTAGAAG ATTTATTTTG CTACACTTAT ATCTACAACA	3000
AAATCAACTT ATGAAATTGT CTTGGAAACT ACCTCTAACC CGGTAGAATG AATTGGAATG	3060
AAAATTAAAC CAACTTACGG AATCGCCCAA CATATGTCGA TTAAAGTGGA TATGGATACA	3120
TATGAAGAAG CCCTAGAGAT AATCTAAATG GTTTCAGAAT TGAGGGTTAT TTTTGAAGT	3180
TTGATGGGAA GATAAGACCA TAACGGTAGT TCACAGAGAT AAAAGGGTTA TTTTTTTCAG	3240
AAATATTTGT GCTGCAATTG ATCCTGTGCC TCAAATTCAG CCTGCAACCA AGGCCAGGTT	3300
CTAGAGCGAA CAAGGCCAC GTCACCCGTG GCCCGTCAGG CGAAGCAGGT CTTGTGCAGA	3360
CTTTGAGAGG GATTGGATAT CAACGGAACC AATCACGCAC GGCAATGCGA TTCCCAGCCC	3420
ACCTGTAACG TTCCAGTGGG CCATCCTTAA CTCCAAGCCC AACGGCCCTA CCCCATCTCG	3480
TCGTGTCATC CACTCCGCCG CACAGGCGCT CAGCTCCGCA ACGCCGCCGG AAATGGTCGC	3540
CGCCACAGCC ACCGCCATGG CCACCGCTGC ATCGCCGCTA CTCAACGGGA CCCGAATACC	3600
TGCGCGGCTC CGCCATCGAG GACTCAGCGT GCGCTGCGCT GCTGTGGCGG GCGGCGCGGC	3660
CGAGGCACCG GCATCCACCG GCGCGCGGCT GTCCGCGGAC TGCGTTGTGG TGGGCGGAGG	3720
CATCAGTGGC CTCTGCACCG CGCAGGCGCT GGCCACGCGG CACGGCGTCG GGGACGTGCT	3780
TGTCACGGAG GCCCGCGCCC GCGCGGCGG CAACATTACC ACCGTCGAGC GCGCGGAGGA	3840
AGGGTACC	3848

(2) 서열 15에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 1826개 염기쌍

(B) 유형 : 핵산

(C) 쇠 : 일본쇠

(D) 형태 : 선형

(ii) 분자 유형 : cDNA

(iii) 가설 : 없음

(iv) 안티-센스 : 없음

(vi) 최초 공급원 :

(A) 생물 : 고시피움 히르수툼(*Gossypium hirsutum*)(목화)

(vii) 중간체 공급원 :

(B) 클론 : pWDC-15(NRRL B-21594)

(ix) 특성 :

(A) 명칭/키 : misc_특성

(B) 위치 : 31..1647

(D) 기타 정보 : /생성물="목화 프로토크스-1 암호화 영역"

(xi) 서열 기술 :

서열 15

CCTCTCGCTC GCCTGGCCCC ACCACCAATC ATGACGGCTC TAATCGACCT TTCTCTTCTC	60
CGTTCCCTCGC CCTCCGTTTC CCCTTTCTCC ATACCCCACC ACCAGCATCC GCCCCGCTTT	120
CGTAAACCTT TCAAGCTCCG ATGCTCCCTC GCCGAGGGTC CCACGATTTC CTCATCTAAA	180
ATCGACGGGG GAGAATCATC CATCGCGGAT TGCATCATCG TTGGAGGTGG TATCAGTGGA	240
CTTTGCATTG CTCAAGCTCT CGCCACCAAG CACCGTGACG TCGCTTCCAA TGTGATTGTG	300
ACGGAGGCCA GAGACCGTGT TGGTGGCAAC ATCACTACCG TTGAGAGAGA TGGATATCTG	360
TGGAAGAAG GCCCAACAG TTTTCAGCCC TCCGATCCTA TTCTAACCAT GGCCGTGGAT	420

AGTGGATTGA AGGACGATTT GGTTTTAGGT GACCCTAATG CACCGCGATT TGTACTATGG	480
GAGGGAAAAC TAAGGCCTGT GCCCTCCAAG CCAACCGACT TGCCGTTTTT TGATTTGATG	540
AGCATTGCTG GAAAACCTAG GGCTGGGTTC GGGGCTATTG GCATTGGGCC TCCCCCTCCG	600
GGTTATGAAG AATCGGTGGA GGAGTTTGTG CGCCGTAATC TTGGTGCTGA GGTTTTTGAA	660
CGCTTTATTG AACCATTTTG TTCAGGTGTT TATGCAGGGG ATCCTTCAA ATTAAGCATG	720
AAAGCAGCAT TTGGAAGAGT ATGGAAGCTA GAAGAGATTG GTGGCAGCAT CATTGGTGGC	780
ACTTTCAAGA CAATCCAGGA GAGAAATAAG ACACCTAAGC CACCCAGAGA CCCGCGTCTG	840
CCAAAACCGA AGGGCCAAAC AGTTGGATCT TTTAGGAAGG GACTTACCAT GCTGCCCTGAG	900
GCAATTGCTA ACAGTTTGGG TAGCAATGTA AAATTATCTT GGAAGCTTTC CAGTATTACC	960
AAATTGGGCA ATGGAGGGTA TAACTTGACA TTTGAAACAC CTGAAGGAAT GGTATCTCTT	1020
CAGAGTAGAA GTGTTGTAAT GACCATTCCA TCCCATGTTG CCAGTAACTT GTTGCACTCT	1080
CTCTCGGCTG CTGCTGCAGA TGCATTATCC CAATTTTATT ATCCTCCAGT TGCATCAGTC	1140
ACAGTCTCCT ATCCAAAAGA AGCCATTCTA AAAGAATGTT TGATTGATGG TGAAGTTAAG	1200
GGGTTTGGCC AGTTGCACCC ACGCAGCCAA GGAATTGAAA CTTTAGGGAC GATATACAGT	1260
TCATCACTTT TCCCCAATCG AGCTCCATCT GGCAGGGTGT TGCTCTTGAA CTACATAGGA	1320
GGAGCTACCA ACACTGGAAT TTTGTCCAAG ACTGAAGGGG AACTTGTTAGA AGCAGTTGAT	1380
CGTGATTGTA GAAAAATGCT TATAAATCCT AATGCAAAGG ATCCTCTTGT TTTGGGTGTA	1440
AGAGTATGGC CAAAAGCCAT TCCACAGTTC TTGGTTGGTC ATTTGGATCT CCTTGATAGT	1500
GCAAAAATGG CTCTCAGGGA TTCTGGGTTT CATGGACTGT TTCTGGGGG CAACTATGTA	1560
TCTGGTGTGG CATTAGGACG GTGTGTGGAA GGTGCTTACG AGGTTGCAGC TGAAGTGAAG	1620
GAATTCCTGT CACAATATGC ATACAAATAA TATTGAAATT CTTGTCAGGC TGCAAATGTA	1680
GAAGTCAGTT ATTGGATAGT ATCTCTTTAG CTAAAAAATT GGGTAGGGTT TTTTTGTTA	1740
GTTCCCTTGAC CACTTTTTGG GGTTTTCATT AGAACTTCAT ATTTGTATAT CATGTTGCAA	1800
TATCAAAAAA AAAAAAAAAA AAAAAA	1826

(2) 서열 16에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 539개 아미노산

(B) 유형 : 아미노산

(C) 쇠 : 관련 없음

(D) 형태 : 관련 없음

(ii) 분자 유형 : 단백질

(xi) 서열 기술 :

서열 16

```

Met Thr Ala Leu Ile Asp Leu Ser Leu Leu Arg Ser Ser Pro Ser Val
1           5           10           15

Ser Pro Phe Ser Ile Pro His His Gln His Pro Pro Arg Phe Arg Lys
          20           25           30

Pro Phe Lys Leu Arg Cys Ser Leu Ala Glu Gly Pro Thr Ile Ser Ser
          35           40           45

Ser Lys Ile Asp Gly Gly Glu Ser Ser Ile Ala Asp Cys Val Ile Val
          50           55           60

Gly Gly Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu Ala Thr Lys
65           70           75           80

His Arg Asp Val Ala Ser Asn Val Ile Val Thr Glu Ala Arg Asp Arg
          85           90           95

Val Gly Gly Asn Ile Thr Thr Val Glu Arg Asp Gly Tyr Leu Trp Glu
          100          105          110

Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Pro Ile Leu Thr Met Ala
          115          120          125

Val Asp Ser Gly Leu Lys Asp Asp Leu Val Leu Gly Asp Pro Asn Ala
          130          135          140

```

Pro Arg Phe Val Leu Trp Glu Gly Lys Leu Arg Pro Val Pro Ser Lys
 145 150 155 160

Pro Thr Asp Leu Pro Phe Phe Asp Leu Met Ser Ile Ala Gly Lys Leu
 165 170 175

Arg Ala Gly Phe Gly Ala Ile Gly Ile Arg Pro Pro Pro Gly Tyr
 180 185 190

Glu Glu Ser Val Glu Glu Phe Val Arg Arg Asn Leu Gly Ala Glu Val
 195 200 205

Phe Glu Arg Phe Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp
 210 215 220

Pro Ser Lys Leu Ser Met Lys Ala Ala Phe Gly Arg Val Trp Lys Leu
 225 230 235 240

Glu Glu Ile Gly Gly Ser Ile Ile Gly Gly Thr Phe Lys Thr Ile Gln
 245 250 255

Glu Arg Asn Lys Thr Pro Lys Pro Pro Arg Asp Pro Arg Leu Pro Lys
 260 265 270

Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys Gly Leu Thr Met Leu
 275 280 285

Pro Glu Ala Ile Ala Asn Ser Leu Gly Ser Asn Val Lys Leu Ser Trp
 290 295 300

Lys Leu Ser Ser Ile Thr Lys Leu Gly Asn Gly Gly Tyr Asn Leu Thr
 305 310 315 320

Phe Glu Thr Pro Glu Gly Met Val Ser Leu Gln Ser Arg Ser Val Val
 325 330 335

Met Thr Ile Pro Ser His Val Ala Ser Asn Leu Leu His Pro Leu Ser
 340 345 350

Ala Ala Ala Ala Asp Ala Leu Ser Gln Phe Tyr Tyr Pro Pro Val Ala
 355 360 365

Ser Val Thr Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu
 370 375 380

```

Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln Leu His Pro Arg Ser Gln
385          390          395          400

Gly Ile Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn
          405          410          415

Arg Ala Pro Ser Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ala
          420          425          430

Thr Asn Thr Gly Ile Leu Ser Lys Thr Glu Gly Glu Leu Val Glu Ala
          435          440          445

Val Asp Arg Asp Leu Arg Lys Met Leu Ile Asn Pro Asn Ala Lys Asp
          450          455          460

Pro Leu Val Leu Gly Val Arg Val Trp Pro Lys Ala Ile Pro Gln Phe
465          470          475          480

Leu Val Gly His Leu Asp Leu Leu Asp Ser Ala Lys Met Ala Leu Arg
          485          490          495

Asp Ser Gly Phe His Gly Leu Phe Leu Gly Gly Asn Tyr Val Ser Gly
          500          505          510

Val Ala Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu Val Ala Ala Glu
          515          520          525

Val Lys Glu Phe Leu Ser Gln Tyr Ala Tyr Lys
          530          535

```

(2) 서열 17에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 1910개 염기쌍

(B) 유형 : 핵산

(C) 색 : 일본색

(D) 형태 : 선형

(ii) 분자 유형 : cDNA

(iii) 가설 : 없음

(iv) 안티-센스 : 없음

(vi) 최초 공급원 :

(A) 생물 : 베타 불가리스(Beta vulgaris)(사탕무)

(vii) 중간체 공급원 :

(B) 클론 : pWDC-16(NRRL B-21595N)

(ix) 특성 :

(A) 명칭/키 : misc_특성

(B) 위치 : 1..1680

(D) 기타 정보 : /생성물="사탕무 프로텍스-1 암호화 영역"

(xi) 서열 기술 :

서열 17

ATGAAATCAA TGGCGTTATC AAACGTCATT CCACAGACAC AGTGCATGCC ATTGCGCAGC	60
AGCGGGCATT ACAGGGGTAA TTGTATCATG TTGTCAATTC CATGTAGTTT AATTGGAAGA	120
CGAGGTTATT ATTCACATAA GAAGAGGAGG ATGAGCATGA GTTGCAGCAC AAGCTCAGGC	180
TCAAAGTCAG CGGTTAAAGA AGCAGGATCA GGATCAGGTG CAGGAGGATT GCTAGACTGC	240
GTAATCGTTG GAGGTGGAAT TAGCGGGCTT TGCATCGCGC AGGCTCTTTG TACAAAACAC	300
TCCTCTTCCT CTTTATCCCC AAATTTTATA GTTACAGAGG CCAAAGACAG AGTTGGCGGC	360
AACATCGTCA CTGTGGAGGC CGATGGCTAT ATCTGGGAGG AGGGACCCAA TAGCTTCCAG	420
CCTTCCGACG CGGTGCTCAC CATGGCGGTC GACAGTGGCT TGAAAGATGA GTTGGTGCTC	480
GGAGATCCCA ATGCTCCTCG CTTTGTGCTA TGAATGACA AATTAAGGCC CGTACCTTCC	540
AGTCTCACCG ACCTCCCTTT CTTGACCTC ATGACCATTC CGGGCAAGAT TAGGGCTGCT	600
CTTGGTGCTC TCGGATTTG CCGTCTCCT CCACCTCATG AGGAATCTGT TGAACACTTT	660
GTGCGTCGTA ATCTCGGAGA TGAGGTCTTT GAACGCTTGA TTGAACCCTT TTGTTCAGGT	720
GTGTATGCCG GTGATCCTGC CAAGCTGAGT ATGAAAGCTG CTTTGGGAA GGTCTGGAAG	780
TTGGAGCAAA AGGGTGGCAG CATAATTGGT GGCACCTCTCA AAGCTATACA GGAAAGAGGG	840
AGTAATCCTA AGCCGCCCCG TGACCAGCGC CTCCCTAAAC CAAAGGGTCA GACTGTTGGA	900

TCCTTTAGAA AGGGACTCGT TATGTTGCCT ACCGCCATTT CTGCTCGACT TGGCAGTAGA	960
GTGAAACTAT CTTGGACCCT TTCTAGTATC GTAAAGTCAC TCAATGGAGA ATATAGTCTG	1020
ACTTATGATA CCCCAGATGG CTTGGTTTCT GTAAGAACCA AAAGTGTGT GATGACTGTT	1080
CCATCATATG TTGCAAGTAG GCTTCTTCGT CCACTTTCAG ACTCTGCTGC AGATTCTCTT	1140
TCAAAATTTT ACTATCCACC AGTTGCAGCA GTGTCACTTT CCTATCCTAA AGAAGCGATC	1200
AGATCAGAAT GCTTGATTAA TGGTGAACCT CAAGGTTTCG GGCAACTACA TCCCCGCAGT	1260
CAGGGTGTGG AAACCTTGGG AACAAATTAT AGTTCGTCTC TTTTCCCTGG TCGAGCACCA	1320
CCTGGTAGGA TCTTGATCTT GAGCTACATC GGAGGTGCTA AAAATCCTGG CATATTAAAC	1380
AAGTCGAAAG ATGAACTTGC CAAGACAGTT GACAAGGACC TGAGAAGAAT GCTTATAAAT	1440
CCTGATGCAA AACTTCCTCG TGTACTGGGT GTGAGAGTAT GGCCTCAAGC AATACCCAG	1500
TTTTCTATTG GGCACCTTGA TCTGCTCGAT GCTGCAAAAG CTGCTCTGAC AGATACAGGG	1560
GTCAAAGGAC TGTTTCTTGG TGGCAACTAT GTTTCAGGTG TTGCCTTGGG GCGGTGTATA	1620
GAGGGTGCTT ATGAGTCTGC AGCTGAGGTA GTAGATTTCC TCTCACAGTA CTCAGACAAA	1680
TAGAGCTTCA GCATCCTGTG TAATTCAACA CAGGCCTTTT TGTATCTGTT GTGCGCGCAT	1740
GTAGTCTGGT CGTGGTGCTA GGATTGATTA GTTGCTCTGC TGTGTGATCC ACAAGAATTT	1800
TGATGGAATT TTTCCAGATG TGGGCATTAT ATGTTGCTGT CTTATAAATC CTTAATTTGT	1860
ACGTTTAGTG AATTACACCG CATTTGATGA CTAAAAAAA AAAAAAAAAA	1910

(2) 서열 18에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 560개 아미노산

(B) 유형 : 아미노산

(C) 쇠 : 관련 없음

(D) 형태 : 관련 없음

(ii) 분자 유형 : 단백질

(xi) 서열 기술 :

서열 18

```

Met Lys Ser Met Ala Leu Ser Asn Cys Ile Pro Gln Thr Gln Cys Met
1           5           10           15

Pro Leu Arg Ser Ser Gly His Tyr Arg Gly Asn Cys Ile Met Leu Ser
           20           25           30

Ile Pro Cys Ser Leu Ile Gly Arg Arg Gly Tyr Tyr Ser His Lys Lys
           35           40           45

Arg Arg Met Ser Met Ser Cys Ser Thr Ser Ser Gly Ser Lys Ser Ala
           50           55           60

Val Lys Glu Ala Gly Ser Gly Ser Gly Ala Gly Gly Leu Leu Asp Cys
65           70           75           80

Val Ile Val Gly Gly Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu
           85           90           95

Cys Thr Lys His Ser Ser Ser Ser Leu Ser Pro Asn Phe Ile Val Thr
           100           105           110

Glu Ala Lys Asp Arg Val Gly Gly Asn Ile Val Thr Val Glu Ala Asp
           115           120           125

Gly Tyr Ile Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Ala
           130           135           140

Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Glu Leu Val Leu
145           150           155           160

Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Asn Asp Lys Leu Arg
           165           170           175

Pro Val Pro Ser Ser Leu Thr Asp Leu Pro Phe Phe Asp Leu Met Thr
           180           185           190

Ile Pro Gly Lys Ile Arg Ala Ala Leu Gly Ala Leu Gly Phe Arg Pro
           195           200           205

Ser Pro Pro Pro His Glu Glu Ser Val Glu His Phe Val Arg Arg Asn
210           215           220

```

Leu Gly Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser Gly
 225 230 235 240
 Val Tyr Ala Gly Asp Pro Ala Lys Leu Ser Met Lys Ala Ala Phe Gly
 245 250 255
 Lys Val Trp Lys Leu Glu Gln Lys Gly Gly Ser Ile Ile Gly Gly Thr
 260 265 270
 Leu Lys Ala Ile Gln Glu Arg Gly Ser Asn Pro Lys Pro Pro Arg Asp
 275 280 285
 Gln Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys
 290 295 300
 Gly Leu Val Met Leu Pro Thr Ala Ile Ser Ala Arg Leu Gly Ser Arg
 305 310 315 320
 Val Lys Leu Ser Trp Thr Leu Ser Ser Ile Val Lys Ser Leu Asn Gly
 325 330 335
 Glu Tyr Ser Leu Thr Tyr Asp Thr Pro Asp Gly Leu Val Ser Val Arg
 340 345 350
 Thr Lys Ser Val Val Met Thr Val Pro Ser Tyr Val Ala Ser Arg Leu
 355 360 365
 Leu Arg Pro Leu Ser Asp Ser Ala Ala Asp Ser Leu Ser Lys Phe Tyr
 370 375 380
 Tyr Pro Pro Val Ala Ala Val Ser Leu Ser Tyr Pro Lys Glu Ala Ile
 385 390 395 400
 Arg Ser Glu Cys Leu Ile Asn Gly Glu Leu Gln Gly Phe Gly Gln Leu
 405 410 415
 His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser
 420 425 430
 Ser Leu Phe Pro Gly Arg Ala Pro Pro Gly Arg Ile Leu Ile Leu Ser
 435 440 445
 Tyr Ile Gly Gly Ala Lys Asn Pro Gly Ile Leu Asn Lys Ser Lys Asp
 450 455 460

```

Glu Leu Ala Lys Thr Val Asp Lys Asp Leu Arg Arg Met Leu Ile Asn
465              470              475              480

Pro Asp Ala Lys Leu Pro Arg Val Leu Gly Val Arg Val Trp Pro Gln
              485              490              495

Ala Ile Pro Gln Phe Ser Ile Gly His Phe Asp Leu Leu Asp Ala Ala
              500              505              510

Lys Ala Ala Leu Thr Asp Thr Gly Val Lys Gly Leu Phe Leu Gly Gly
              515              520              525

Asn Tyr Val Ser Gly Val Ala Leu Gly Arg Cys Ile Glu Gly Ala Tyr
              530              535              540

Glu Ser Ala Ala Glu Val Val Asp Phe Leu Ser Gln Tyr Ser Asp Lys
545              550              555              560

```

(2) 서열 19에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 1784개 염기쌍

(B) 유형 : 핵산

(C) 채 : 일본채

(D) 형태 : 선형

(ii) 분자 유형 : cDNA

(iii) 가설 : 없음

(iv) 안티-센스 : 없음

(vi) 최초 공급원 :

(A) 생물 : 브라씨카 나푸스(Brassica napus)(평지)

(vii) 중간체 공급원 :

(B) 클론 : pWDC-17(NRRL B-21615)

(ix) 특성 :

(A) 명칭/키 : misc_특성

(B) 위치 : 47..1654

(D) 기타 정보 : /생성물="평지 프로토스-1 암호화 영역"

(xi) 서열 기술 :

서열 19

GGGCCCCCCC CAAAATTGAG GATTCTCCTT CTCGCGGGCG ATCGCCATGG ATTTATCTCT	60
TCTCCGTCCG CAGCCATTCC TATCGCCATT CTCAAATCCA TTTCCTCGGT CGCGTCCCTA	120
CAAGCCTCTC AACCTCCGTT GCTCCGTATC CGGTGGATCC GTCGTCGGCT CTTCTACAAT	180
CGAAGGCGGA GGAGGAGGTA AAACCGTCAC GCGGGACTGC GTGATCGTCG GCGGAGGAAT	240
CAGCGGCCTG TGCATTGCGC AAGCGCTCGT GACGAAGCAC CCAGACGCTG CAAAGAATGT	300
GATGGTGACG GAGGCGAAGG ACCGTGTGGG AGGGAATATC ATCACGCGAG AGGAGCAAGG	360
GTTTCTATGG GAAGAAGGTC CCAATAGCTT TCAGCCGTCT GATCCTATGC TCACTATGGT	420
GGTAGATAGT GGTTTGAAAG ATGATCTAGT CTTGGGAGAT CCTACTGCTC CGAGGTTTGT	480
GTGTGGAAT GGAAGCTGA GGCCGGTTCC GTCGAAGCTA ACTGACTTGC CTTTCTTTGA	540
CTTGATGAGT ATTGGAGGGA AGATTAGAGC TGGGTTTGGT GCCATTGGTA TTCGACCTTC	600
ACCTCCGGGT CGTGAGGAAT CAGTGAAGA GTTTGTAAGG CGTAATCTTG GTGATGAGGT	660
TTTTGAGCGC TTGATTGAAC CCTTTTGCTC AGGTGTTTAT GCGGGAGATC CTGCGAACT	720
GAGTATGAAA GCAGCTTTTG GGAAGGTTTG GAAGCTAGAG GAGAATGGTG GGAGCATCAT	780
TGGTGGTGCT TTTAAGGCAA TTCAAGCGAA AAATAAGCT CCCAAGACAA CCCGAGATCC	840
GCGTCTGCCA AAGCCAAAGG GCCAACTGT TGTTCTTTC AGGAAAGGAC TCACAATGCT	900
GCCAGAGGCA ATCTCCGCAA GGTGGGTGA CAAGGTGAAA GTTTCTTGGA AGCTCTCAAG	960
TATCACTAAG CTGGCCAGCG GAGAATATAG CTTAACTTAC GAAACTCCGG AGGGTATAGT	1020
CACTGTACAG AGCAAAAGTG TAGTGATGAC TGTGCCATCT CATGTTGCTA GTAGTCTCTT	1080
GCGCCCTCTC TCTGATTCTG CAGCTGAAGC GCTCTCAAAA CTCTACTATC CGCCAGTTGC	1140
AGCCGTATCC ATCTCATACG CGAAGAAGC AATCCGAAGC GAATGCTTAA TAGATGGTGA	1200
ACTAAAAGG TTGCGCCAGT TGCATCCACG CACGCAAAA GTGGAAACTC TTGGAACAAT	1260

```

ATACAGTTCA TCGCTCTTTC CCAACCGAGC ACCGCCTGGA AGAGTATTGC TATTGAACTA 1320
CATCGGTGGA GCTACCAACA CTGGGATCTT ATCAAAGTCG GAAGGTGAGT TAGTGAAGC 1380
AGTAGATAGA GACTTGAGGA AGATGCTGAT AAAGCCAAGC TCGACCGATC CACTTGCTACT 1440
TGGAGTAAAA TTATGGCCTC AAGCCATTCC TCAGTTTCTG ATAGGTCACA TTGATTGGT 1500
AGACGCAGCG AAAGCATCGC TCTCGTCATC TGGTCATGAG GGCTTATTCT TGGGTGGAAA 1560
TTACGTTGCC GGTGTAGCAT TGGGTCGGTG TGTGGAAGGT GCTTATGAAA CTGCAACCCA 1620
AGTGAATGAT TTCATGTCAA GGTATGCTTA CAAGTAATGT AACGCAGCAA CGATTTGATA 1680
CTAAGTAGTA GATTTTGCAG TTTTGACTTT AAGAACACTC TGTTTGTGAA AAATTCAAGT 1740
CTGTGATTGA GTAAATTAT GTATTATTAC TAAAAAAAAA AAAA 1784

```

(2) 서열 20에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 536개 아미노산

(B) 유형 : 아미노산

(C) 쇠 : 관련 없음

(D) 형태 : 관련 없음

(ii) 분자 유형 : 단백질

(xi) 서열 기술 :

서열 20

```

Met Asp Leu Ser Leu Leu Arg Pro Gln Pro Phe Leu Ser Pro Phe Ser
1           5           10           15

Asn Pro Phe Pro Arg Ser Arg Pro Tyr Lys Pro Leu Asn Leu Arg Cys
          20           25           30

Ser Val Ser Gly Gly Ser Val Val Gly Ser Ser Thr Ile Glu Gly Gly
          35           40           45

Gly Gly Gly Lys Thr Val Thr Ala Asp Cys Val Ile Val Gly Gly Gly
          50           55           60

```

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

Ser Ile Thr Lys Leu Ala Ser Gly Glu Tyr Ser Leu Thr Tyr Glu Thr			
305	310	315	320
Pro Glu Gly Ile Val Thr Val Gln Ser Lys Ser Val Val Met Thr Val			
	325	330	335
Pro Ser His Val Ala Ser Ser Leu Leu Arg Pro Leu Ser Asp Ser Ala			
	340	345	350
Ala Glu Ala Leu Ser Lys Leu Tyr Tyr Pro Pro Val Ala Ala Val Ser			
	355	360	365
Ile Ser Tyr Ala Lys Glu Ala Ile Arg Ser Glu Cys Leu Ile Asp Gly			
	370	375	380
Glu Leu Lys Gly Phe Gly Gln Leu His Pro Arg Thr Gln Lys Val Glu			
	385	390	395
Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro			
	405	410	415
Pro Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr			
	420	425	430
Gly Ile Leu Ser Lys Ser Glu Gly Glu Leu Val Glu Ala Val Asp Arg			
	435	440	445
Asp Leu Arg Lys Met Leu Ile Lys Pro Ser Ser Thr Asp Pro Leu Val			
	450	455	460
Leu Gly Val Lys Leu Trp Pro Gln Ala Ile Pro Gln Phe Leu Ile Gly			
	465	470	475
His Ile Asp Leu Val Asp Ala Ala Lys Ala Ser Leu Ser Ser Ser Gly			
	485	490	495
His Glu Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu			
	500	505	510
Gly Arg Cys Val Glu Gly Ala Tyr Glu Thr Ala Thr Gln Val Asn Asp			
	515	520	525
Phe Met Ser Arg Tyr Ala Tyr Lys			
	530	535	

(2) 서열 21에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 1224개 염기쌍

(B) 유형 : 핵산

(C) 소재 : 일본소재

(D) 형태 : 선형

(ii) 분자 유형 : cDNA

(iii) 가설 : 없음

(iv) 안티-센스 : 없음

(vi) 최초 공급원 :

(A) 생물 : 오리자 사티브(Oryza sativa)(벼)

(vii) 중간체 공급원 :

(B) 클론 : pWDC-18(NRRL B-21648)

(ix) 특성 :

(A) 명칭/키 : misc_특성

(B) 위치 : 1..936

(D) 기타 정보 : /생성물="벼 프로토크스-1 부분 암호화 영역"

(xi) 서열 기술 :

서열 21

CGGGCTTTGA AGGCTGCATT TGGGAAGGTG TGGAGGCTGG AGGATACTGG AGGTAGCATT	60
ATTGGTGGAA CCATCAAGAC AATCCAGGAG AGGGGGAAAA ACCCAAACC GCCGAGGGAT	120
CCCCGCCTTC CAACGCCAAA GGGGCAGACA GTTGCATCTT TCAGGAAGGG TCTGACTATG	180
CTCCCGGATG CTATTACATC TAGGTTGGGT AGCAAAGTCA AACTTTCATG GAAGTTGACA	240
AGCATTACAA AGTCAGACAA CAAAGGATAT GCATTAGTGT ATGAAACACC AGAAGGGGTG	300
GTCTCGGTGC AAGCTAAAC TGTGTGCATG ACCATCCCAT CATATGTTGC TAGTGATATC	360
TTGCGGCCAC TTTCAAGTGA TGCAGCAGAT GCTCTGTCAA TATTCTATTA TCCACCAGTT	420
GCTGCTGTAA CTGTTTCATA TCCAAAAGAA GCAATTAGAA AAGAATGCTT AATTGACGGA	480
GAGCTCCAGG GTTTCGGCCA GCTGCATCCG CGTAGTCAGG GAGTTGAGAC TTTAGGAACA	540
ATATATAGCT CATCACTCTT TCCAAATCGT GCTCCAGCTG GAAGGGTGTT ACTTCTGAAC	600
TACATAGGAG GTTCTACAAA TACAGGGATT GTTTCCAAGA CTGAAAGTGA GCTGGTAGAA	660
GCAGTTGACC GTGACCTCAG GAAGATGCTG ATAAATCCTA GAGCAGTGA CCCTTTGGTC	720
CTTGCGTCC GGGTATGGCC ACAAGCCATA CCACAGTTCC TCATTGGCCA TCTTGATCAT	780
CTTGAGGCTG CAAAATCTGC CCTGGGCAAA GGTGGGTATG ATGGATTGTT CCTCGGAGGG	840
AACTATGTTG CAGGAGTTGC CCTGGGCCGA TGCGTTGAAG GTGCATATGA GAGTGCCTCA	900
CAAATATCTG ACTACTTGAC CAAGTACGCC TACAAGTGAT CAAAGTTGGC CTGCTCCTTT	960
TGGCACATAG ATGTGAGGCT TCTAGCAGCA AAAATTTTCAT GGGCATCTTT TTATCCTGAT	1020
TCTAATTAGT TAGAATTTAG AATTGTAGAG GAATGTTCCA TTTGCAGTTC ATAATAGTTG	1080
TTCAGATTTT AGCCATTCAA TTTGTGCAGC CATTTACTAT ATGTAGTATG ATCTTGTAAG	1140
TACTACTAAG AACAAATCAA TTATATTTTC CTGCAAGTGA CATCTTAATC GTCAGCAAAT	1200
CCAGTTACTA GTAAAAA AAAA	1224

(2) 서열 22에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 312개 아미노산

(B) 유형 : 아미노산

(C) 쇠 : 관련 없음

(D) 형태 : 관련 없음

(ii) 분자 유형 : 단백질

(xi) 서열 기술 :

서열 22

Arg	Ala	Leu	Lys	Ala	Ala	Phe	Gly	Lys	Val	Trp	Arg	Leu	Glu	Asp	Thr
1				5					10					15	

Gly Gly Ser Ile Ile Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Gly
 20 25 30
 Lys Asn Pro Lys Pro Pro Arg Asp Pro Arg Leu Pro Thr Pro Lys Gly
 35 40 45
 Gln Thr Val Ala Ser Phe Arg Lys Gly Leu Thr Met Leu Pro Asp Ala
 50 55 60
 Ile Thr Ser Arg Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr
 65 70 75 80
 Ser Ile Thr Lys Ser Asp Asn Lys Gly Tyr Ala Leu Val Tyr Glu Thr
 85 90 95
 Pro Glu Gly Val Val Ser Val Gln Ala Lys Thr Val Val Met Thr Ile
 100 105 110
 Pro Ser Tyr Val Ala Ser Asp Ile Leu Arg Pro Leu Ser Ser Asp Ala
 115 120 125
 Ala Asp Ala Leu Ser Ile Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr
 130 135 140
 Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly
 145 150 155 160
 Glu Leu Gln Gly Phe Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu
 165 170 175
 Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro
 180 185 190
 Ala Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ser Thr Asn Thr
 195 200 205
 Gly Ile Val Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg
 210 215 220
 Asp Leu Arg Lys Met Leu Ile Asn Pro Arg Ala Val Asp Pro Leu Val
 225 230 235 240
 Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Ile Gly
 245 250 255
 His Leu Asp His Leu Glu Ala Ala Lys Ser Ala Leu Gly Lys Gly Gly
 260 265 270
 Tyr Asp Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu
 275 280 285
 Gly Arg Cys Val Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Ser Asp
 290 295 300
 Tyr Leu Thr Lys Tyr Ala Tyr Lys
 305 310

(2) 서열 23에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 1590개 염기쌍

(B) 유형 : 핵산

(C) 출처 : 일본채

(D) 형태 : 선형

(ii) 분자 유형 : cDNA

(iii) 가설 : 없음

(iv) 안티-센스 : 없음

(vi) 최초 공급원 :

(A) 생물 : 소르굼 비콜로(Sorghum bicolor)(수수)

(vii) 중간체 공급원 :

(B) 클론 : pWDC-19(NRRL B-21649)

(ix) 특성 :

(A) 명칭/키 : misc_특성

(B) 위치 : 1..1320

(D) 기타 정보 : /생성물="수수 프로텍스-1 부분 암호화 영역"

(xi) 서열 기술 :

서열 23

TCCACCGTCG AGCGCCCCGA GGAAGGGTAC CTCTGGGAGG AGGGTCCCAA CAGCTTCCAG 60

CCATCCGACC CCGTTCTCTC CATGGCCGTG GACAGCGGGC TGAAGGATGA CCTGGTTTTT 120

GGGGACCCCA ACGCGCCACG GTTCGTGCTG TGGGAGGGGA AGCTGAGGCC CGTGCCATCC	180
AAGCCCGCCG ACCTCCCGTT CTTCGATCTC ATGAGCATCC CTGGCAAGCT CAGGGCCGGT	240
CTCGGCGCGC TTGGCATCCG CCCGCCTGCT CCAGGCCGCG AGGAGTCAGT GGAGGAGTTT	300
GTGCGCCGCA ACCTCGGTGC TGAGGTCTTT GAGCGCCTAA TTGAGCCTTT CTGCTCAGGT	360
GTCTATGCTG GCGATCCTTC CAAGCTCAGT ATGAAGGCTG CATTTGGGAA GGTGTGGCGG	420
TTAGAAGAAG CTGGAGGTAG TATTATGGT GGAACCATCA AGACGATTCA GGAGAGGGGC	480
AAGAATCCAA AACCACCGAG GGATCCCCGC CTTCCGAAGC CAAAAGGGCA GACAGTTGCA	540
TCTTTCAGGA AGGGTCTTGC CATGCTTCCA AATGCCATCA CATCCAGCTT GGGTAGTAAA	600
GTCAAACAT CATGGAACT CACGAGCATG ACAAATCAG ATGGCAAGGG GTATGTTTTG	660
GAGTATGAAA CACCAGAAGG GGTGTTTTG GTGCAGGCTA AAAGTGTAT CATGACCATT	720
CCATCATATG TTGCTAGCGA CATTTTTCGT CCACTTTCAG GTGATGCTGC AGATGTTCTA	780
TCAAGATTCT ATTATCCACC AGTTGCTGCT GTAACGGTTT CGTATCCAAA GGAAGCAATT	840
AGAAAAGAAT GCTTAATTGA TGGGGAATC CAGGGTTTTG GCCAGTTGCA TCCACGTAGT	900
CAAGGAGTTG AGACATTAGG AACAATATAC AGCTCATCAC TCTTTCCAAA TCGTGCTCCT	960
GCTGGTAGGG TGTACTTCT AACTACATA GGAGGTGCTA CAAACACAGG AATTGTTTCC	1020
AAGACTGAAA GTGAGCTGGT AGAAGCAGTT GACCGTGACC TCCGAAAAAT GCTTATAAAT	1080
CCTACAGCAG TGGACCCTTT AGTCCTTGGT GTCCGAGTTT GGCCACAAGC CATACTCAG	1140
TTCTTGGTAG GACATCTTGA TCTTCTGGAG GCCGAAAAAT CTGCCCTGGA CCAAGGTGGC	1200
TATAATGGGC TGTTCCTAGG AGGGAACAT GTTGCAGGAG TTGCCCTGGG CAGATGCATT	1260
GAGGGCGCAT ATGAGAGTGC CGCGCAAATA TATGACTTCT TGACCAAGTA CGCCTACAAG	1320
TGATGGAAGA AGTGGAGCGC TGCTTGTTAA TTGTTATGTT GCATAGATGA GGTGAGACCA	1380
GGAGTAGTAA AAGGCGTCAC GAGTATTTTT CATCTTATT TTGTAAATTG CACTTCTGTT	1440
TTTTTTTCCT GTCAGTAATT AGTTAGATTT TAGTTATGTA GGAGATTGTT GTGTTCACTG	1500
CCCTACAAAA GAATTTTAT TTTGCATTCG TTTATGAGAG CTGTGCAGAC TTATGTAACG	1560
TTTTACTGTA AGTATCAACA AAATCAAATA	1590

(2) 서열 24에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 440개 아미노산

(B) 유형 : 아미노산

(C) 쇠 : 관련 없음

(D) 형태 : 관련 없음

(ii) 분자 유형 : 단백질

(xi) 서열 기술 :

서열 24

```

Ser Thr Val Glu Arg Pro Glu Glu Gly Tyr Leu Trp Glu Glu Gly Pro
1           5           10           15

Asn Ser Phe Gln Pro Ser Asp Pro Val Leu Ser Met Ala Val Asp Ser
          20           25           30

Gly Leu Lys Asp Asp Leu Val Phe Gly Asp Pro Asn Ala Pro Arg Phe
          35           40           45

Val Leu Trp Glu Gly Lys Leu Arg Pro Val Pro Ser Lys Pro Ala Asp
          50           55           60

Leu Pro Phe Phe Asp Leu Met Ser Ile Pro Gly Lys Leu Arg Ala Gly
65           70           75           80

Leu Gly Ala Leu Gly Ile Arg Pro Pro Ala Pro Gly Arg Glu Glu Ser
          85           90           95

Val Glu Glu Phe Val Arg Arg Asn Leu Gly Ala Glu Val Phe Glu Arg
          100          105          110

Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser Lys
          115          120          125

Leu Ser Met Lys Ala Ala Phe Gly Lys Val Trp Arg Leu Glu Glu Ala

```

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

370

375

380

His Leu Asp Leu Leu Glu Ala Ala Lys Ser Ala Leu Asp Gln Gly Gly
 385 390 395 400
 Tyr Asn Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu
 405 410 415
 Gly Arg Cys Ile Glu Gly Ala Tyr Glu Ser Ala Ala Gln Ile Tyr Asp
 420 425 430
 Phe Leu Thr Lys Tyr Ala Tyr Lys
 435 440

(2) 서열 25에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 93개 염기쌍

(B) 유형 : 핵산

(C) 출처 : 일본채

(D) 형태 : 선형

(ii) 분자 유형 : 다른 핵산

(ix) 특성 :

(A) 기술 : /desc="옥수수 프로토스-1 인트론 서열"

(xi) 서열 기술 :

서열 25

GTACGCTCCT CGCTGGCGCC GCAGCGTCTT CTTCTCAGAC TCATGCGCAG CCATGGAATT 60
 GAGATGCTGA ATGGATTTTA TACGCGCGCG CAG 93

(2) 서열 26에 대한 정보:

(i) 서열 특징 :

(A) 길이 : 2606개 염기쌍

(B) 유형 : 핵산

(C) 출처 : 일본채

(D) 형태 : 선형

(ii) 분자 유형 : DNA(게놈)

(iii) 가설 : 없음

(iv) 안티-센스 : 없음

(vi) 최초 공급원 :

(A) 생물 : 베타 불가리스(Beta vulgaris)(사탕무)

(vii) 중간체 공급원 :

(B) 클론 : pWDC-20(NRRL B-21650)

(ix) 특성 :

(A) 명칭/키 : misc_특성

(B) 위치 : 1..6

(D) 기타 정보 : /주목="SalI 부위"

(ix) 특성 :

(A) 명칭/키 : misc_특성

(B) 위치 : 상보성(1..538)

(D) 기타 정보 : /주목="3'에서 5' 방향으로 사탕무 프로토스-1의 부분 cDNA"

(ix) 특성 :

(A) 명칭/키 : misc_특성

(B) 위치 : 539..2606

(D) 기타 정보 : /주목="3'에서 5' 방향으로 나타난 사탕무 프로토스-1 프로모터 영역(pWDC-20으로부터 서브클로닝된 약 3kb PstI-SalI 단편의 부분 서열)"

(xi) 서열 기술 :

서열 26

GTCGACCTAC GCACATGCCA CATTCCACAT TCCACGTTAG GAATTGAATT GAATTGAATT	60
ATGATTATGA ATAATGAAGA GACAGAATTA CCGCCATGGT GAGCACCGCG TCGGAAGGCT	120
GGAAGCTATT GGGTCCCTCC TCCCAGATAT AGCCATCGGC CTCCACAGTG ACGATGTTGC	180
CGCCAACTCT GTCTTTGGCC TCTGTCACTA TAAAATTTGG GGATAAAGAG GACTGTTTTG	240
TACAAAGAGC CTGCGCGATG CAAAGCCCGC TAATTCCACC TCCAACGATT ACGCAGTCTA	300
GCAATCCTCC TGCTCCTGAT CCTGATCCTG ATCCTGCTTC TTTAACCGCT GACTTTGAGC	360
CTGAGCTTGT GCTGCAACTC ATGCTCATCC TCCTCTTCTT ATGTGAATAA TAACCTCGTC	420

TTCCAATTAA ACTACATGGA ATTGACAACA TGATACAATT GCCCCTGTAA TCCCCGCTGC	480
TGTGCAATGG CATGCACTGT GTCTGTGGAA TGCAGTTTGA TAACGCCATT GATTTCATCT	540
CTCTCTCGCT CTCTCGCCTT CCTTATCCTC TATATCCCTT TCTTGCCTGC TCGGGAATTC	600
TAATTAACCT TATATCAAAA TGAACAACCT GTTCTAGTT AAAAGTTTT TTATAAATAG	660
TACTCTAAAT AAACGATTAC ATGTATCTTC TAACCATACT TGTTTGCTGG AGGTGGTGCG	720
TAACCGGTAA CTTACCTTTG TAACTCACCT CAATACCTAC TTATGCTTAA GCATACGGAT	780
TCTTTTAAAC TCTCAGGCAT TGACCTATGT AGCTGGACTG ACTAACATCT GAATTTGTTT	840
CTCTGGTTAT ATATGCAATT TTAAGTGAAT CGAAATTTCT CTGGATGCTA AAAATGTCTT	900
TAACGGGGTT TATCAGGACT AAATTATCTC CTTCATGAG GAGGTTCTTG ATTTGCATGT	960
ATGAGCGTGA AAATGCATTC TTAACGGCTA TAGATTCACT AATAAGTGGT GTTAAAAGTA	1020
AAAAGTACTT GGA AAAATGA TTAAGCGACT TAATTTTTTT TATTTGTTTG AAAGTTGCTT	1080
TTTCTTGGCT ATCTTAACAT GTATTATCA AACACCTTTT TTAATTACAT GGAATCGAA	1140
AAGTTTGAAA AAAAAAATC ATACTCACTA ACCGCTTAA AATATAAGCT GAAGATGTCT	1200
CACTAACAGA GTGCATGTGA AGCACCCCA AAGCAATTAT AACACAACAT CTCCGCTCT	1260
TCAAAATTCC TACAAATACA TCTAATAAAC TTGTTGAAAC AATCAAAGTA ACATGGTGTG	1320
TCAATTGCGG ATGCTTCTCA TTCCAGACTT TATAAGTGA TTTTGTTTAA TCCATAGTCA	1380
ACAACTCACA TAATGGTACC CAAAGAATAC CCAATTTTTT TGCTCAAAAT CCTTAAACAT	1440
TGTAGCTGTG TAAGTTGAC TAACATGTTT CAGCATGCTT GCCATGGGTA AATAAGACTT	1500
AGGGGCAAAT CTCGAATCCA CAACTCATC ATTGGTTTAA GTTGTCTCC AACGTAAAAC	1560
AATGATGTGA AATACACCAC AAAATTCATA CAATCTCOTT ATCTTGGAG CTGAAAGCC	1620
ATAATCTTGT TTGTACTTTC ACTACGTGGA GAAGACAAA TTCAACTAA GAAGAGCTCA	1680
TTGCTCAGTG TCGTGTACTA CTTATCTTTC AACTCATAGA AACAAACAAA CCAATTGTCA	1740

CCTATATACT GTACTTCTCC ATCATATACT TCCAACCTGC CTTAAACTCA ATACTATCAT	1800
AAAAACCACA AAGACATTTT ATAAAAGCAT AATAAAAATG TGTCATCACT CTTCAAAGTT	1860
CCAAAGTGAT TCTAACTACA TTCTAATGAA AATGACATTG GTGTAAACCT AATCCTTGTTG	1920
TTATAAAACA CCTACATACC ACGATTATGT TAGAAATATA TTTATGAATG CAGTACCTAC	1980
ATAAAGCCAT TAAATAACCA GTTTTATGTT ATTTCTGTGAC CAACATAGTT CCTAAAGATT	2040
ACGAAGTAAT TTATAGTCAT TTTGTGGCCA CTTAATTCAT TTAATACCCA GTATATTTAT	2100
AAGTTACCAG CTTAAGTAGT TTTGTGACCA TCTCTACATA CTTCCCTCCG TCCATAATAA	2160
GGGGGCGTTT GGTGCAACG GGGTAAAGGG AATGGAATCA AGAAAGGGAG AGGAGAGGAA	2220
AGGAAAAGAA AACCCCTAGA TTTAGAGTGG TGTTTGGTTA AGATAATGTT AATTCTCTTT	2280
CTTCCTCTTT CTTACCCTTC TTCCACCCTA GCACCACCAC TCCTCCCTCT GTTACTATTC	2340
TCCACGCCGC CTCTCCCTAC CCCAGTAACA CCACCTTGTC GGCCCCCGG TCTTCCCTT	2400
CCCGCGACGG TTCCCCCTC CCCTGCGCCG TCACGTCGTC CCCCTCACCT CCCTGCACCG	2460
TCGAGTTATC CCCCTCCCCT GCGCGTCGCG TTCTCCCCTC CCTCACCATC GCGTTCTCCC	2520
CTCCCTCACC GTCGCGTTCT CCCCTCCCCT ACCGTCGCGG TCTCCCCTCC CTCACCGTCG	2580
CGGTCTCTCT TCCCCTCCCC CTGCAG	2606

(57) 청구의 범위

청구항 1

식물 프로토포르피리노겐 옥시다제(프로톡스) 프로모터 또는 이의 기능상 동등한 유도체를 포함하는 분리된 DNA 분자.

청구항 2

식물 프로토포르피리노겐 옥시다제에 대한 암호화 서열과 자연적으로 결합된 식물 프로톡스 프로모터를 포함하는 분리된 DNA 분자.

청구항 3

제2항에 있어서, 식물이 아라비도시스(Arabidopsis) 종인 분리된 DNA 분자.

청구항 4

제3항에 있어서, DNA 분자가 서열 13에 기술된 뉴클레오타이드 서열 및 적당히 엄격한 조건하에 이와 하이브리드화하는 모든 DNA 분자를 포함하는 분리된 DNA 분자.

청구항 5

제2항에 있어서, 식물이 옥수수인 분리된 DNA 분자.

청구항 6

제5항에 있어서, DNA 분자가 서열 14에 기술된 뉴클레오타이드 서열 및 적당히 엄격한 조건하에 이와 하이브리드화하는 모든 DNA 분자를 포함하는 분리된 DNA 분자.

청구항 7

제2항에 있어서, 식물이 사탕무인 분리된 DNA 분자.

청구항 8

제7항에 있어서, DNA 분자가 서열 26에 기술된 뉴클레오타이드 서열 및 적당히 엄격한 조건하에 이와 하 이브리드화하는 모든 DNA 분자를 포함하는 분리된 DNA 분자.

청구항 9

제1항 내지 제8항 중의 어느 한 항에 기술된 식물 프로토포르피리노겐 옥시다제(프로톡스) 프로모터 또는 이의 기능상 동등한 유도체를 포함하는 재조합체 DNA 분자.

청구항 10

이종성 DNA 암호화 서열과 작동가능하게 결합된 식물 프로톡스 프로모터를 포함하는 키메라 유전자.

청구항 11

제10항에 있어서, 식물 프로톡스 프로모터가 프로톡스-1 유전자로부터 유래하는 키메라 유전자.

청구항 12

제10항에 있어서, 식물 프로톡스 프로모터가 프로톡스-2 유전자로부터 유래하는 키메라 유전자.

청구항 13

제10항에 있어서, 프로톡스 프로모터가 아라비돕시스, 대두, 목화, 담배, 사탕무, 오일시드 평지, 옥수수, 밀, 수수, 호밀, 귀리, 잔디 및 벼로 이루어진 그룹중에서 선택된 식물로부터 유래하는 키메라 유전자.

청구항 14

제10항에 있어서, 프로모터가 아라비돕시스, 사탕무 및 옥수수로 이루어진 그룹중에서 선택된 식물로부터 유래하는 키메라 유전자.

청구항 15

제10항에 있어서, 프로모터가 아라비돕시스 및 옥수수로 이루어진 그룹중에서 선택된 식물로부터 유래하는 키메라 유전자.

청구항 16

제10항에 있어서, 프로모터가 사탕무로부터 유래하는 키메라 유전자.

청구항 17

제10항에 있어서, 프로모터의 뉴클레오타이드 길이가 300개 이상인 키메라 유전자.

청구항 18

제17항에 있어서, 프로모터의 뉴클레오타이드 길이가 500개 이상인 키메라 유전자.

청구항 19

제11항에 있어서, 프로모터가 아라비돕시스로부터 유래하고 서열 13에 기술된 서열을 갖는 키메라 유전자.

청구항 20

제11항에 있어서, 프로모터가 옥수수로부터 유래하고 서열 14에 기술된 서열을 갖는 키메라 유전자.

청구항 21

제11항에 있어서, 프로모터가 사탕무로부터 유래하고 서열 26에 기술된 서열을 갖는 키메라 유전자.

청구항 22

제10항에 있어서, 이종성 암호화 서열이 변형된 제초제-내성 형태의 식물 효소를 암호화하는 키메라 유전자.

청구항 23

제22항에 있어서, 식물 효소가 이미다졸글리세롤 포스페이트 데하이드라타제(IGPD), 5-에놀피루발시킴에이트-3-포스페이트 신타제(EPSP), 글루타민 신세타제(GS), 아세틸 조효소 A 카복실라제, 아세토락테이트 신타제, 히스티딘올 데하이드로게나제 및 프로토포르피리노겐 옥시다제(프로톡스)로 이루어진 그룹중에서 선택되는 키메라 유전자.

청구항 24

제23항에 있어서, 식물 효소가 프로톡스인 키메라 유전자.

청구항 25

제23항에 있어서, 식물 효소가 프로톡스 억제제에 내성을 부여하는 아미노산 치환체를 갖는 진핵성 프로톡스인 키메라 유전자.

청구항 26

제10항에 있어서, 이중성 DNA 분자가 프로톡스-1 활성 또는 프로톡스-2 활성을 갖는 아라비돕시스 종으로부터의 단백질을 암호화하는 키메라 유전자.

청구항 27

제26항에 있어서, 단백질이 서열 2 또는 서열 4에 기술된 아미노산 서열을 포함하는 키메라 유전자.

청구항 28

제10항에 있어서, 이중성 DNA 분자가 프로톡스-1 활성 또는 프로톡스-2 활성을 갖는 옥수수로부터의 단백질을 암호화하는 키메라 유전자.

청구항 29

제28항에 있어서, 단백질이 서열 6 또는 서열 8에 기술된 아미노산 서열을 포함하는 키메라 유전자.

청구항 30

제10항에 있어서, 이중성 DNA 분자가 프로톡스-1 활성을 갖는 밀로부터의 단백질을 암호화하는 키메라 유전자.

청구항 31

제30항에 있어서, 단백질이 서열 10에 기술된 아미노산 서열을 포함하는 키메라 유전자.

청구항 32

제10항에 있어서, 이중성 DNA 분자가 프로톡스-1 활성을 갖는 대두로부터의 단백질을 암호화하는 키메라 유전자.

청구항 33

제32항에 있어서, 단백질이 서열 12에 기술된 아미노산 서열을 포함하는 키메라 유전자.

청구항 34

제10항에 있어서, 이중성 DNA 분자가 프로톡스-1 활성을 갖는 목화로부터의 단백질을 암호화하는 키메라 유전자.

청구항 35

제34항에 있어서, 단백질이 서열 16에 기술된 아미노산 서열을 포함하는 키메라 유전자.

청구항 36

제10항에 있어서, 이중성 DNA 분자가 프로톡스-1 활성을 갖는 사탕무로부터의 단백질을 암호화하는 키메라 유전자.

청구항 37

제36항에 있어서, 단백질이 서열 18에 기술된 아미노산 서열을 포함하는 키메라 유전자.

청구항 38

제10항에 있어서, 이중성 DNA 분자가 프로톡스-1 활성을 갖는 평지로부터의 단백질을 암호화하는 키메라 유전자.

청구항 39

제38항에 있어서, 단백질이 서열 20에 기술된 아미노산 서열을 포함하는 키메라 유전자.

청구항 40

제10항에 있어서, 이중성 DNA 분자가 프로톡스-1 활성을 갖는 벼로부터의 단백질을 암호화하는 키메라 유전자.

청구항 41

제40항에 있어서, 단백질이 서열 22에 기술된 아미노산 서열을 포함하는 키메라 유전자.

청구항 42

제10항에 있어서, 이중성 DNA 분자가 프로톡스-1 활성을 갖는 수수로부터의 단백질을 암호화하는 키메라 유전자.

청구항 43

제42항에 있어서, 단백질이 서열 24에 기술된 아미노산 서열을 포함하는 키메라 유전자.

청구항 44

제9항의 재조합체 DNA 분자를 포함하는 재조합체 DNA 벡터.

청구항 45

식물, 식물 종자, 식물 조직 또는 식물 세포를 안정하게 형질전환시킬 수 있는, 제10항 내지 제43항 중의 어느 한 항에 따른 키메라 유전자를 포함하는 재조합체 벡터.

청구항 46

제10항 내지 제43항 중의 어느 한 항에 따른 키메라 유전자를 포함하는 식물 조직.

청구항 47

제10항 내지 제43항 중의 어느 한 항에 따른 키메라 유전자를 포함하는 식물 및 이의 자손.

청구항 48

제47항에 있어서, 식물이 아라비도시스, 사탕수수, 대두, 보리, 목화, 담배, 사탕무, 오일시드 평지, 옥수수, 밀, 수수, 호밀, 귀리, 잔디, 마초, 기장 및 벼로 이루어진 그룹중에서 선택되는 식물.

청구항 49

제47항에 있어서, 식물이 아라비도시스, 대두, 목화, 담배, 사탕무, 오일시드 평지, 옥수수, 밀, 수수, 호밀, 귀리, 잔디 및 벼로 이루어진 그룹중에서 선택되는 식물.

청구항 50

제초제에 대해 내성을 부여하기 위해 식물에서 제초제 내성 형태의 제초제 표적 단백질을 발현하는 프로토스 프로모터의 용도.

청구항 51

변형되지 않은 식물 프로토스 단백질의 억제제에 대해 내성이 있는 제초제 내성 식물 프로토스 단백질을 발현하는 제25항에 따른 키메라 유전자의 용도.

청구항 52

프로브로서 목적하는 프로모터와 결합된 프로토스 암호화 서열과 하이브리드화하기에 충분한 상동성을 갖는 프로토스 암호화 서열의 용도.

청구항 53

제52항에 있어서, 프로브로서 사용된 암호화 서열이 목적하는 프로토스 프로모터와 동이한 식물 종으로부터 유래하고, 프로모터와 자연적으로 결합된 암호화 서열인 프로토스 암호화 서열의 용도.

청구항 54

(a) 뉴클레오타이드 길이가 10개 이상인 식물로부터의 프로토스 단백질에 대한 암호화 서열의 인접 부분을 포함하고, 식물 프로토스 유전자 또는 mRNA와 특이적으로 하이브리드화할 수 있는 뉴클레오타이드 프로브를 제조하는 단계;

(b) 단계 (a)에 따라 제조된 뉴클레오타이드 프로브를 사용하여 선택된 생물로부터의 클로닝된 게놈 DNA 단편 또는 cDNA 단편의 모집단에서 다른 프로토스 암호화 서열을 프로빙하는 단계;

(c) 프로토스 프로모터 서열을 함유하는 DNA 부분 및 프로토스 단백질을 암호화하는 DNA 부분을 포함하는 DNA 분자를 분리하여 증폭시키는 단계를 포함하는, 프로토스 프로모터 서열을 함유하는 DNA 부분 및 프로토스 단백질을 암호화하는 DNA 부분을 포함하는 DNA 분자를 제조하는 방법.

청구항 55

(a) 뉴클레오타이드 길이가 10개 이상인 식물로부터의 프로토스 단백질에 대한 암호화 서열의 인접 부분을 포함하고, 식물 프로토스 유전자 또는 mRNA와 특이적으로 하이브리드화할 수 있는 뉴클레오타이드 프로브를 제조하는 단계;

(b) 단계 (a)에 따라 제조된 뉴클레오타이드 프로브를 사용하여 선택된 생물로부터의 클로닝된 게놈 DNA 단편 또는 cDNA 단편의 모집단에서 다른 프로토스 암호화 서열을 프로빙하는 단계;

(c) 프로토스 프로모터 서열을 함유하는 DNA 부분을 포함하는 DNA 분자를 분리하여 증폭시키는 단계를 포함하는, 프로토스 프로모터 서열을 함유하는 DNA 부분을 포함하는 DNA 분자를 제조하는 방법.

청구항 56

(a) 뉴클레오타이드 길이가 10개 이상인 식물로부터의 프로토스 단백질에 대한 암호화 서열의 인접 부분을 포함하고, 식물 프로토스 유전자 또는 mRNA와 특이적으로 하이브리드화할 수 있는 뉴클레오타이드 프로브를 제조하는 단계;

(b) 단계 (a)에 따라 제조된 뉴클레오타이드 프로브를 사용하여 선택된 생물로부터의 클로닝된 게놈 DNA 단편 또는 cDNA 단편의 모집단에서 다른 프로토스 암호화 서열을 프로빙하는 단계;

(c) 프로토스 프로모터 서열을 함유하는 DNA 부분을 포함하는 DNA 분자를 분리하는 단계를 포함하는, 식물 프로토스 유전자로부터의 프로토스 프로모터 서열을 함유하는 DNA 부분을 포함하는 DNA 분자를 제조하

는 방법.

청구항 57

제10항 내지 제25항중 어느 한 항에 따른 키메라 유전자를, 제초제에 대한 내성을 부여하기 위해 식물에서 제초제 내성 형태의 제초제 표적 단백질을 발현하기에 충분한 양으로 포함하는 유전자이식 식물 또는 이의 자손이 사용되는 농업적 방법.

청구항 58

제10항에 있어서, DNA 분자에 의해 암호화된 단백질을 염색체에 표적화시킬 수 있는, DNA 분자에 작동가능하게 결합된 시그널 서열을 추가로 포함하는 키메라 유전자.

청구항 59

제10항에 있어서, DNA 분자에 의해 암호화된 단백질을 미토콘드리아에 표적화시킬 수 있는, DNA 분자에 작동가능하게 결합된 시그널 서열을 추가로 포함하는 키메라 유전자.

청구항 60

제22항에 있어서, 식물 효소가 이미다졸글리세롤 포스페이트 데하이드라타제(IGPD), 5-에놀피루빌시키페이트-3-포스페이트 신타제(EPSP), 글루타민 신세타제(GS), 아세틸 조효소 A 카복실라제, 아세토락테이트 신타제 및 프로토포르피리노겐 옥시다제(프로톡스)로 이루어진 그룹중에서 선택되는 키메라 유전자.

청구항 61

제3항에 있어서, DNA 분자가 서열 13에 기술된 뉴클레오타이드 서열, 및 (a) 7% 나트륨 도데실 설페이트(SDS), 0.5M NaPO₄(pH 7.0), 0.1mM EDTA내에서 50℃에서의 하이브리드화 및 (b) 2×SSC, 1% SDS에서 50℃에서의 세척 조건하에 상기 뉴클레오타이드 서열과 하이브리드화하는 모든 DNA 분자를 포함하는 분리된 DNA 분자.

청구항 62

제5항에 있어서, DNA 분자가 서열 14에 기술된 뉴클레오타이드 서열, 및 (a) 7% 나트륨 도데실 설페이트(SDS), 0.5M NaPO₄(pH 7.0), 0.1mM EDTA내에서 50℃에서의 하이브리드화 및 (b) 2×SSC, 1% SDS에서 50℃에서의 세척 조건하에 상기 뉴클레오타이드 서열과 하이브리드화하는 모든 DNA 분자를 포함하는 분리된 DNA 분자.

청구항 63

제7항에 있어서, DNA 분자가 서열 26에 기술된 뉴클레오타이드 서열, 및 (a) 7% 나트륨 도데실 설페이트(SDS), 0.5M NaPO₄(pH 7.0), 0.1mM EDTA내에서 50℃에서의 하이브리드화 및 (b) 2×SSC, 1% SDS에서 50℃에서의 세척 조건하에 상기 뉴클레오타이드 서열과 하이브리드화하는 모든 DNA 분자를 포함하는 분리된 DNA 분자.