

## (19) 대한민국특허청(KR)

## (12) 공개특허공보(A)

(51) Int. Cl.<sup>6</sup>  
C12N 15/82(11) 공개번호 특2000-0048782  
(43) 공개일자 2000년07월25일

(21) 출원번호	10-1999-7002770
(22) 출원일자	1999년03월30일
번역문제출일자	1999년03월30일
(86) 국제출원번호	PCT/US1997/17555
(86) 국제출원출원일자	1997년09월30일
(81) 지정국	AP ARIPO특허 : 캐나 레소토 말라위 수단 스와질랜드 우간다 EA 유라시아특허 : 아르메니아 아제르바이잔 벨라루스 키르기즈 카자흐스탄 몰도바 러시아 타지키스탄 투르크메니스탄 EP 유럽특허 : 오스트리아 벨기에 스위스 리히텐슈타인 독일 덴마크 스페인 프랑스 영국 그리스 아일랜드 이탈리아 룩셈부르크 모나코 네덜란드 포르투칼 스웨덴 핀란드 OA OAPI특허 : 부르키나파소 베냉 중앙아프리카 콩고 코트디브와르 카메룬 가봉 기네 말리 모리타니 니제르 세네갈 차드 토고 국내특허 : 알바니아 아르메니아 오스트리아 오스트레일리아 아제르바이잔 보스니아-헤르체고비나 바베이도스 불가리아 브라질 벨라루스 캐나다 스위스 중국 쿠바 체코 독일 덴마크 에스토니아 스페인 핀란드 영국 그루지야 헝가리 이스라엘 아이슬란드 일본 캐나 키르기즈 북한 대한민국 카자흐스탄 세인트루시아 스리랑카 라이베리아 레소토 리투아니아 룩셈부르크 라트비아 몰도바 마다가스카르 마케도니아 몽고 말라위 멕시코 노르웨이 뉴질랜드 슬로베니아 슬로바키아 타지키스탄 투르크메니스탄 터키 트리니다드토바고 우크라이나 우간다 우즈베키스탄 베트남 폴란드 포르투칼 루마니아 러시아 수단 스웨덴 싱가포르
(30) 우선권주장	60/026,855 1996년09월30일 미국(US)
(71) 출원인	액씨드 제네틱스 엘.엘.씨. 미합중국, 아이오와 50011-1061, 에임, 아이오와 스테이트 유니버시티, 후드 싸이언스빌딩 1573
(72) 발명자	킬링파터 미합중국, 아이오와50014, 에임, 오크랜드스트리트3409 구안한평 미합중국, 아이오와50010, 에임, 크레스트우드써클1608
(74) 대리인	박세걸, 조성욱, 최홍순

심사청구 : 없음(54) 전분 기질내에 폴리펩티드의 캡슐화요약

본 발명은 동화작용 단백질을 암호화하는 유전자로부터의 캡슐화 영역을 가지는 잡종 폴리펩티드를 제공한다. 특히, 본 발명은 기질내에 부착된 단백질을 캡슐화하는 유전자를 암호화하는 재조합 핵산 분자에 관한 것이며, 이러한 유전자는 전분내에, 특히 전분 입자 기질내에 원하는(페이로드)폴리펩티드를 캡슐화한다. 이러한 재조합 핵산 분자를 포함하는 발현 벡터, 및 이것에 대한 숙주, 특히 이러한 벡터에 의해 형질변환된 상기 숙주의 전분-형유 부분이 또한 제공된다. 바람직하게, 전분내에 캡슐화된 외부 단백질을 함유하는 곡물이 제공되는데, 이것은 포유류, 어류 및 조류의 먹이를 제조하는데 유용하다. 또한, 본 발명은 전분, 특히 전분 입자로부터 정제된 단백질을 제조하는 방법 및 이러한 단백질의 공업적 사용을 포함한다.

대표도도1a색인어

풀리펩티드, 전분, 단백질

## 영세서

### 기술분야

본 발명은 동화작용 단백질을 암호화하는 유전자로부터의 캡슐화 영역을 가지는 잡종 풀리펩티드에 관한 것이다.

### 배경기술

다당류 효소.

원핵 세포 및 진핵 세포 둘 모두는 저장물(storage reserve)로서 다당류 효소를 이용한다. 원핵 세포에 있어서 1차 저장 다당류는 글리코겐이다. 글리코겐은 대부분의 유관속식물에서 발견되는 전분과 유사하긴 하지만 상이한 사슬 길이 및 종합도를 나타낸다. 많은 식물에 있어서, 전분이 1차 저장 다당류로서 사용된다. 전분은 전분 함유 식물의 여러 가지 조직내에 저장된다. 전분은 대부분 두 성분으로 이루어지는데, 하나는 아밀로스이고, 다른 하나는 아밀로펩틴이다. 아밀로스는 선형 글루칸으로서 형성되고 아밀로펩틴은 글루칸의 분지 사슬로서 형성된다. 대표적인 전분은 25% 아밀로스 및 75% 아밀로펩틴으로 이루어진다. 그밖에, 흔히 상이한 식물로부터 얻어지는 전분은 상이한 특성을 가진다. 옥수수 전분과 감자 전분은 인산기가 존재하거나 또는 존재하지 않기 때문에 상이한 것으로 판단된다. 특정의 식물의 전분 특성들은 식물 계통내로 도입되어진 돌연변이 때문에 상이하다. 돌연변이 전분은 옥수수, 쌀 및 완두콩 등에서 잘 확인된다.

전분 분지 또는 전분 성분을 변화시키면 전분 특성이 상이해진다. 전분의 한 가지 특성은, 특히 잎, 뿌리, 괴경 및 씨에 형성되는 전분 입자의 형성이다. 이러한 입자는 전분 합성 과정동안에 형성된다. 전분의 특정의 신타아제, 특히 입자-결합 전분 신타아제, 가용성 전분 신타아제 및 분지 효소는 전분 입자를 형성할 때 전분 입자내에 캡슐화되는 단백질이다.

동물 및 세균 글리코겐 신타아제의 cDNA 클론의 사용이 국제특허출원 공개번호 GB92/01881호에 기술되어 있다. 글리코겐 신타아제의 누클레오티드 및 아미노산 서열이 상기 문헌으로부터 확인된다. 예를들어, 글리코겐 신타아제를 암호화하는 *E. coli* *gIgA* 유전자에 대한 누클레오티드 서열은 GenBank/EMBL(SWISSPROT) 데이터베이스의 접근번호 J02616호로부터 구해질 수 있다(Kumar 등, 1986, *J. Biol. Chem.*, 261:16256-16259 참조). *E. coli* 글리코겐 생합성 효소 구조 유전자가 Okita 등에 의해 또한 클론화되었다(1981, *J. Biol. Chem.*, 256(13):6944-6952 참조). 글리코겐 신타아제의 *gIgA* 구조 유전자는 Leung 등에 의해 *Salmonella typhimurium* LT2로부터 클론화되었다(1987, *J. Bacteriol.*, 169(9):4349-4354 참조). 토끼 골격근(Zhang 등, 1989, *FASEB J.*, 3:2532-2536 참조) 및 사람의 근육(Browner 등, 1989, *Proc. Natl. Acad. Sci.*, 86:1443-1447)으로부터 얻어지는 글리코겐 신타아제의 서열이 또한 알려져 있다.

식물의 가용성 전분 신타아제의 cDNA 클론의 사용이 보고되었다. 완두콩의 가용성 전분 신타아제 이소포름 I 및 II의 아미노산 서열이 다음문헌(Dry 등, 1991, *Plant Journal*, 2:193202)에 발표되었다. 벼의 가용성 전분 신타아제의 아미노산 서열이 다음문헌[Baba 등, 1993, *Plant Physiology*]에 기재되었다. 이러한 벼의 가용성 전분 신타아제의 아미노산 서열은 N-말단 서열을 부정확하게 열거하고 있기 때문에 오해를 일으킬 수 있다. 아마 이것은 추출된 효소에서 프로테아제 감성 또는 다른 고유한 불안정성을 수반하는 일부의 추출 실수때문인 것으로 판단된다. 정확한 N-말단 서열(AECSR을 이용하여 출발)은 벼의 가용성 전분 신타아제의 트랜싯(transit) 펩티드 서열을 나타내는 것으로 보인다.

옥수수 분지 효소 I의 서열이 Baba 등에 의해 연구되었다(1991, *BBRC*, 181:8794 참조). 옥수수의 배젖으로부터 얻은 전분 분지 효소 II가 Fisher 및 Shrable에 의해 연구되었다(1993, *Plant Physiol.*, 102:1045-1046 참조). 식물, 세균 및 동물의 분지 효소의 cDNA 클론의 사용이 보고되었다. 세균의 분지 효소(BE)에 대한 누클레오티드 및 아미노산 서열이 문헌으로부터 확인된다. 예를들어, Kiel 등은

*Cyanobacterium synechococcus* sp. PCC7942(1989, *Gene*(Amst), 78(1):918 참조)로부터 그리고 *Bacillus stearothermophilus*(Kiel 등, 1991, *Mol. Gen. Genet.* 230(12):136-144 참조)로부터 분지 효소 유전자 *gIgB*를 클론화했다. *S. cerevisiae*의 유전자 *glc3* 및 *ghal*은 대립유전자이고 글리코겐 분지 효소를 암호화한다(Rowen 등, 1992, *Mol. Cell Biol.*, 12(1):22-29 참조). Matsumomo 등은 *Neurospora crassa*로부터 글리코겐 분지 효소를 연구했다(1990, *J. Biochem.*, 107:118-122 참조). 또한, Genbank/EMBL 데이터베이스는 분지 효소를 암호화하는 *E. coli* *gIgB* 유전자에 대한 서열을 포함하고 있다.

전분 효소(EC 2.4.1. II)는 전분 분자를 연장시키고 아밀로스 및 아밀로펩틴 둘 모두에 대하여 작용하는 것으로 판단된다. 전분 신타아제(SYS)의 활성도는 입자 및 플라스티드의 기질과 관련이 있는 것으로 밝혀질 수 있다. 결합된 전분 신타아제 효소의 전분 회합 능력이 잘 알려져 있다. 전분 생합성과 관련이 있는 여러 가지의 효소들은 상이한 결합성을 가지는 것으로 확인된다(Mu-Forster 등, 1996, *Plant Phys.* 111:821-829 참조). 입자-결합된 전분 신타아제(GBSTS)의 활성은 *waxy* 유전자의 생성물과 아주 상관이 있다(Shure 등, 1983, *Cell* 35:225-233 참조). 옥수수, 쌀 및 감자와 같은 다수의 종의 아밀로스의 합성은 이러한 유전자의 발현에 의존하는 것으로 밝혀졌다(Tasi, 1974, *Biochem Gen* 11: 83-96; Hovenkamp-Hermelink 등, 1987, *Theor. Appl. Gen.* 75: 217-221 참조). Vassar 등은 감자로부터 얻은 입자-결합 전분 신타아제에 대한 유전자의 분자 클론화 및 부분적 특성화를 발표했다(1989, *Plant Sci.* 64(2):185-192).

참조). 또한, Vasser등은 감자의 유전자-결합 전분 신타아제에 대한 유전자의 발현을 안티센스 구성물(antisense construct)을 이용하여 억제하는 것을 발표했다(1991, *Mol. Gen. Genet.* 225(2):289-296 참조).

다른 STS 효소가 Frydman 과 Cardini의 최초 연구후 가용성 전분 신타아제인 것으로 확인되었다(Frydman 과 Cardini, 1964, *Biochem. Biophys. Res. Communication* 17:407-411 참조). 최근에, 용어 '가용성'의 적절함이, 상기 효소가 입자와 회합될 뿐만 아니라 가용성 상으로 존재한다는 발견에 비추어 의문시되었다(Denyer등, 1993, *Plant J.* 4: 191-198; Denyer등, 1995, *Planta* 97: 57-62; Mu-Forster등, 1996, *Plant Physiol.* 111: 821-829 참조). 일반적으로, 아밀로펙틴의 생합성은 가용성 전분 신타아제와 전분 분지 효소의 상호작용을 수반하는 것으로 믿어진다. 가용성 전분 신타아제의 상이한 이소포름들이 확인되었고 완두콩(Denyer 및 Smith, 1992, *Planta* 186: 609-617; Dry등, 1992, *Plant Journal*, 2:193-202 참조), 감자(Edwards 등, 1995, *Plant Physiol.* 112: 89-87; Marshall 등, *Plant Cell* 8: 1121-1135 참조) 및 쌀(Bata 등, 1993, *Plant Physiol.* 103: 565-573 참조)에서 클론화되었지만, 일부가 전분 분지 효소와 회합되는 다중 이소포름은 거의 함유하지 않는 것으로 보인다(Tyynela 및 Schulman, 1994, *Physiol. Plantarum* 89: 835-841 참조). STS 클론들의 공통의 특징은 효소의 ADP-Glc 결합 부위인 것으로 판단되는 KXGGLGDV 교감 서열이 존재한다는 것이다(Furukawa 등, 1990, *J. Biol. Chem.* 265: 2086-2090; Furukawa 등, 1993, *J. Biol. Chem.* 268: 23837-23842 참조).

옥수수에 있어서, 이소포름 I 및 II로서 알려져있는 STS의 두 가지 가용성 형태가 확인되었지만(Macdonald 및 Preiss, 1983, *Plant Physiol.* 73: 175-178; Boyer 및 Preiss, 1978, *Carb. Res.* 61: 321-334; Pollock 및 Preiss, 1980, *Arch Biochem. Biophys.* 204: 578-588; Macdonald 및 Preiss, 1985 *Plant Physiol.* 78: 849-852; Dang 및 Boyer, 1988, *Phytochemistry* 27: 1255-1259; Mu등, 1994, *Plant J.* 6: 151-159 참조), 이들중 어느것도 클론화되지 않았다. 최근에, 옥수수 내배젖의 STS I 활성이 가용성 및 입자-회합된 단편들에서 발견되는 76-kDa 폴리펩티드와 상관되었다(Mu 등, 1994, *Plant J.* 6: 151-159 참조). STS II의 정체는 아직 알려지지 않고 있다. STS I 및 STS II는 상이한 효소학적 특성을 나타낸다. STS I은 프라이머-비의존 활성을 나타내는 반면에 STS II는 글리코실 전이에 촉매작용하기 위한 글리코겐 프라이머를 필요로 한다. 가용성 전분 신타아제는 전분 퇴적을 위하여 높은 대사속도 조절 계수를 가지며(Jenner 등, 1993, *Aust. Plant Physiol.* 22: 703-709; Keeling 등, 1993, *Planta* 191: 342-348 참조), 승온에서 보통과는 다른 반응속도 특성을 가지는 것으로 보고되었다(Keeling 등, 1995, *Aust. J. Plant Physiol.* 21: 807-827 참조). 옥수수에 있어서 이소포름들은 서로 상당히 상이한 온도 최적성 및 안정성을 나타낸다.

식물의 전분 신타아제( 및 *E. coli* 글리코겐 신타아제)서열은 ADPG 결합 영역인 것으로 알려져 있는 서열 KTGG을 포함한다. 이러한 전분 신타아제 단백질중 어떠한 것에 대한 유전자가 본 발명에 따른 구성물에서 사용될 수 있다.

때때로 Q-효소로서 명명되는 분지 효소[ $\alpha$  1,40글루칸6D( $\alpha$  1,40글루카노)트렌스퍼라아제(E.C.2.4.1.18)]는 아밀로스를 아밀로펙틴으로 전환시킨다.  $\alpha$  1,40글루칸 사슬의 분절이 유사한 글루칸 사슬의 1차 히드록시 기에 전달된다.

세균의 분지 효소 유전자 및 식물 서열이 보고되었다(쌀 내배젖: Nakamura등, 1992, *Physiologia Plantarum*, 84:329-335 및 Yamanouchi, 1992, *Plant Physiol.*, 99:1265-1266; 완두콩: Smith, 1988, *Planta*, 175:270-279 및 Bhattacharyya등, 1989, *J. Cell Biochem.*, Suppl. 13D:331; 옥수수 내배젖: Singh 및 Preiss, 1985, *Plant Physiology*, 79:34-40; VosScherperkeuter 등, *Plant Physiology*, 90:75-84; 감자: Kossmann 등, 1991, *Mol. Gen. Genet.*, 230(12):39-44; 카사바: Salehuzzaman 및 Visser, 1992, *Plant Mol Biol*, 20:809-819 참조).

다당류 효소의 분야에 있어서, 여러 가지의 식물종에서 다수의 전분 생합성 유전자를 사용하여 식물의 전분 경로를 공학적으로 변경하기 위한 벡터가 보고되었다. 이러한 다당류 효소들중 일부가 셀룰로오스, 전분 또는 글리코겐에 결합한다는 것이 잘 알려져있다. 다당류 효소의 사용이 있는 한 가지 특별한 특허문현은 식물 전분을 변형시키기 위하여 글리코겐 생합성 효소를 사용하는 것을 보여준다. Shewmaker에게 하여된 미합중국 특허 제 5,349,123호에 있어서, 식물 세포내에서 글리코겐 생합성 효소를 형성하기 위한 DNA를 함유하는 벡터가 개시되어 있다. 구체적으로, 상기 특허는 상기 효소를 도입시킴으로써 감자 전분이 변화한다고 나타내고 있다. 또한, 다른 전분 생합성 유전자들 및 이들의 사용이 보고되었다.

#### 잡종(융합)펩티드.

잡종 단백질(또한, '융합 단백질'이라고도 명명함)은 하나의 폴리펩티드로 일제히 융합된 두 개 이상의 단백질로 이루어지는 폴리펩티드 사슬이다. 흔히 단백질들중 하나는 특정의 수용 세포에 결합하는 리간드이다. 융합 펩티드를 암호화하는 벡터가 미생물의 발효를 통해 이질 단백질을 생성하는데 주로 사용된다. 이어서, 생성된 융합 단백질은 잡종 폴리펩티드를 친화성 기질에 부착하기 위해 사용된다. 예를 들어, 융합 단백질은 칼럼에 결합될 수 있는 베타 갈락토시다제를 이용하여 형성할 수 있다. 이러한 방법은 비루스성 항원을 형성하는데 사용되었다.

또다른 사용은 잡종 폴리펩티드의 폴리펩티드들중 하나의 것을 회수하는 것이다. 융합된 펩티드를 분열시키기 위한 화학적 및 생물학적 방법이 알려져 있다. 산-불안정성 아스파르til-프롤린 결합이 펩티드들 사이에서 이용되고 상기 펩티드들이 산에 영향을 받지 않는 경우 펩티드들을 분열시키는데 낮은 pH가 사용될 수 있다. 호르몬은 시아노브로마이드에 의해 분열되어왔다. 그밖에, 부위-특이 단백질 가수분해에 의한 분열이 보고되었다. 이온 크로마토그래피와 같은 다른 단백질 정제 방법은, 단백질의 전체 염기성을 증가시켜서 이온 교환 칼럼에 대한 결합을 향상시키는 폴리아르기닌 테일(tail)을 사용함으로써 향상되었다. 다수의 특허들은 잡종 펩티드 또는 특정의 사용을 위한 특별한 잡종 펩티드를 제조하는 방법의 개선을 설명했다. Pastan등에게 하여된 미합중국 특허 제 5,635,599호는 잡종 단백질의 개선을 설명하고 있다. 상기 특허는 잡종 펩티드의 일부로서 원형적으로 변경된 리간드를 보고하고 있다. 이러한 리간드는 특이성 및 우수한 결합 친화성을 가진다. 잡종 단백질의 또다른 개선이 Kuliopulos에게 하여된

미합중국 특허 제 5,648,244호에 보고되어 있다. 이러한 특허는 캐리어 펩티드를 가지는 잡종 펩티드를 제조하기 위한 방법을 설명하고 있다. 핵산중간분해효소에 의해 인식되는 때 상기의 핵산 영역은 비현문구조의 3-베이스 오버행(nonpaliindromic 3-base overhang)을 발생시킨다. 따라서 벡터의 분열이 가능해진다.

구체적인 목적을 가지는 잡종 단백질의 예가 미합중국 특허 제 5,643,756호에 보고되어 있다. 이러한 특허는 세포내에서 글리코실화 단백질을 발현시키기 위한 벡터를 보고하고 있다. 이러한 잡종 단백질은 HIV gp 120의 적당한 면역 반응성에서 사용하기에 알맞다. 크게 글리코실화되는 gp 120 도메인의 분리는 이러한 보고된 벡터에 의해 향상된다.

미합중국 특허 제 5,202,247호 및 5,137,819호는 다당류 결합 도메인을 갖는 잡종 단백질 및, 다당류 기질에 결합될 수 있는 잡종 단백질의 제조 방법 및 조성물을 논하고 있다. 구체적으로, 미합중국 특허 제 5,202,247호는 관심있는 펩티드에 셀룰라아제 결합 영역을 연결시키는 잡종 단백질을 가르치고 있다. 이러한 특허는 잡종 단백질이 셀룰로오스상에서의 친화성 크로마토그래피에 의해 세균 숙주에서 발현된 후 정제될 수 있다고 기재하고 있다.

유전자 공학 기술의 발달에 따라, 여러 가지의 유기체 및 식물로부터 다른 유기체 또는 식물내로 유전자를 전달할 수 있게 되었다. 과거에 전분은 형질전환 또는 돌연변이 생성에 의해 변경되어 왔지만 더욱 더 변경시킬 필요성이 여전히 존재한다. 따라서, 전분내에서 구체적으로는 전분 입자내에서 원하는 아미노산 또는 펩티드의 캡슐화(encapsulation)를 제공하는 벡터들이 요구된다. 얻어지는 전분은 변경되고 상기 벡터를 운반하는 식물로부터 얻은 조직이 변형된다.

### 발명의 상세한 설명

본 발명은 전분-캡슐화 영역(starch-encapsulating region; SER)에 내생하지 않는, 즉 상기 전분-캡슐화 영역에 결합된 채로 자연 발생하지 않는 페이로드폴리펩티드(payload polypeptide)에 융합된 전분-결합 효소(starch-binding)로부터의 전분-캡슐화 영역을 포함하는 잡종 폴리펩티드를 제공한다. 이러한 잡종 폴리펩티드는 상기 페이로드 폴리펩티드를 포함하는 변경된 전분을 제조하는데 유용하다. 이러한 변경된 전분은 특정의 아미노산이 풍부한 곡물을 제공하는데 사용될 수 있다. 또한, 이러한 변경된 전분은 위산에 의한 분해에 내성이 있도록 전분-캡슐화 형태를 가지는 호르몬 및 다른 약제, 예를 들어 인슐린과 같은 펩티드를 제공하는데에도 유용하다. 또한, 상기 잡종 폴리펩티드는 쉽게 정제되는 형태의 페이로드 폴리펩티드를 제공하는데에도 유용하다. 예를 들어, 세균 발효에 의해, 또는 곡물 또는 동물에 의해 제조되는 상기의 잡종 폴리펩티드는 당업계에 공지된 방법에 의해 회합되는 변경된 전분으로부터 분리 및 정제될 수 있다.

본원에서 사용되는 용어 '폴리펩티드'는 같거나 또는 상이한 복수의 아미노산을 의미하고, 단백질을 또한 포함한다.

용어 '잡종 폴리펩티드'는 호르몬과 같은 또 다른 폴리펩티드에 융합된 전분-결합 효소의 전분 캡슐화-영역과 같이, 최소한 2개의 상이한 소오스로 부터의 펩티드 또는 폴리펩티드로 이루어진 폴리펩티드를 의미하며, 이러한 잡종 폴리펩티드의 2개 이상의 성분 부분들은 사실상 일제히 융합된 채로 발생하지 않는다.

용어 '페이로드 폴리펩티드'는 이러한 페이로드 폴리펩티드를 함유하는 변경된 전분을 발현하기 위한 전분-캡슐화 영역과의 회합상태로 발현이 요구되는 것으로, 전분-캡슐화 영역에 내생(endogenous)하지 않는 폴리펩티드를 의미한다.

이러한 페이로드 폴리펩티드가 변경된 전분내의 특정의 아미노산의 아미노산 함량을 향상시키기 위해 사용되는 경우, 이러한 페이로드 폴리펩티드는 Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, 및 Val로 이루어진 그룹에서 선택된 세 가지이하의 상이한 타입들의 아미노산으로 이루어지는 것이 바람직하다.

페이로드 폴리펩티드가 숙주 유기체 또는 또 다른 유기체에 생물학적 활성 폴리펩티드를 공급하기 위해 사용되는 경우, 상기 페이로드 폴리펩티드는 호르몬, 예를 들어 인슐린, 성장 인자, 예를 들어 소마토트로핀, 항체, 효소, 면역글로불린, 또는 색소와 같은 생물학적 활성 폴리펩티드일 수 있거나 또는 당해 기술분야에 알려져 있는 바와 같이 상기 생물학적 활성 폴리펩티드의 활성 단편일 수 있다. 이러한 폴리펩티드는 생물학적 활성을 가지는 한에 있어서는, 자연 발생 폴리펩티드일 필요는 없지만, 돌연변이 생성되거나, 절단되거나(truncated), 또는 그렇지 않으면 변경될 수 있다. 이러한 생물학적 활성 폴리펩티드는 생물학적 활성 폴리펩티드의 생물학적 활성 부문만을 함유하는 변경된 폴리펩티드일 수 있다. 또한, 이러한 생물학적 활성 폴리펩티드는 생물학적 활성을 유지하는 자연 발생 아미노산에 상동(바람직하게는 약 75% 이상의 상동)인 아미노산 서열일 수 있다.

잡종 폴리펩티드의 전분-캡슐화 영역은 당업계에 알려져 있는 어떠한 전분-결합 효소, 예를 들어 가용성 전분 신타아제 I, 가용성 전분 신타아제 II, 가용성 전분 신타아제 III, 입자-결합된 전분 신타아제, 분자 효소 I, 분자 효소 IIa, 분자 효소 IIb 및 글루코아밀라제 폴리펩티드의 전분-캡슐화 영역일 수 있다.

잡종 폴리펩티드가 순수한 또는 부분적으로 정제된 형태의 페이로드 폴리펩티드를 제조하기 위해 사용되는 경우, 이러한 잡종 펩티드는 전분-캡슐화 영역과 페이로드 폴리펩티드 사이에서 분열 부위를 포함하는 것이 바람직하다. 정제된 페이로드 폴리펩티드를 분리하는 방법은 잡종 폴리펩티드를 그 분열 부위에 대하여 특이성이 있는 분열제와 접촉시키는 것을 포함한다.

또한, 본 발명은 잡종 폴리펩티드를 암호화하는 재조합 핵산(RNA 또는 DNA)분자를 제공한다. 이러한 재조합 핵산 분자는 선택된 숙주내에서 잡종 폴리펩티드를 발현시키는데 적합한 조절 서열을 포함하는 것이 바람직하다. 용어 '조절 서열'은 프로모터, 인트론, 특정의 숙주 유기체에 대하여 바람직한 코돈 서열, 및 특정의 숙주에서 DNA 또는 RNA의 발현에 영향을 미치는 것으로 당업계에 알려져 있는 다른 서열을 포함한다. 전분-캡슐화 영역 및 페이로드 폴리펩티드를 암호화하는 핵산 서열은 자연 발생 핵산 서열,

또는 이것의 생물학적 단편일 수 있거나, 또는 이러한 서열에 대하여 상동, 바람직하게는 이러한 서열에 대하여 최소한 75% 상동인 생물학적 활성 서열일 수 있다.

숙주 유기체로는 세균, 식물 및 동물이 있다. 바람직한 숙주는 식물이다. 외떡잎식물 및 쌍떡잎식물 둘 모두가 본 발명의 잡종 폴리펩티드를 발현하기 위한 유용한 숙주이다.

또한, 본 발명은 본 발명의 잡종 단백질을 암호화하는 핵산을 포함하는 발현 벡터를 제공한다. 이러한 발현 벡터는 핵산을 숙주 유기체내에 형질변환하기 위해 사용되며 숙주 유기체에서 핵산의 발현을 돋는 서열을 또한 포함한다. 이러한 발현 벡터는 플라스미드, 변경된 바이러스, 또는 DNA 또는 RNA 분자, 또는 당업계에 알려져있는 형질변환계에서 유용한 다른 벡터일 수 있다.

본 발명에 따라, 본 발명의 잡종 펩티드를 발현할 수 있는 재조합 핵산 분자를 포함하는 형질변환된 세포가 생성된다. 이러한 세포는 단세포 유기체, 식물, 또는 동물로 부터의 원핵 또는 진핵 세포일 수 있다. 이러한 세포는 잡종 폴리펩티드가 얻어질 수 있는 세균성 세포일 수 있다. 또는 이러한 세포는 잡종 폴리펩티드가 얻어질 수 있는 식물로 재생될 수 있는 식물 세포일 수 있거나, 또는 이러한 식물 세포는 잡종 폴리펩티드를 암호화하는 핵산을 함유하는 종자에 의해 번식력이 있는 식물로 재생될 수 있다. 바람직한 실시예에 있어서, 이러한 종자는 페이로드 폴리펩티드를 포함하는 변경된 전분을 함유한다.

용어 '변경된 전분'은 페이로드 폴리펩티드를 포함하도록 변경되어진 자연 발생 전분을 의미한다.

소화 과정에서 특정의 단계로 페이로드 폴리펩티드의 소화를 목표로하는 방법, 예를들어 동물의 위장에서 페이로드 폴리펩티드의 분해를 방지하는 방법이 또한 제공되는데, 이 방법은 페이로드 폴리펩티드를 포함하는 본 발명의 변경된 전분을 동물에게 공급함으로써, 폴리펩티드가 동물의 위장에서 분해되는 것을 전분에 의해 방지한다. 선택적으로, 전분은 위장에서 소화되어 폴리펩티드를 방출하는 것으로 알려져있는 것일 수 있다.

본 발명의 바람직한 재조합 핵산 분자는 본 명세서의 표에서 기재된 전분 합성 유전자 서열로부터 선택된 전분-캡슐화 영역을 암호화하는 DNA를 포함한다.

본 발명의 바람직한 플라스미드는 특정의 숙주와 함께 사용하기에 적합하다. 프로모터, 플라스미드-목표 서열, 전분-캡슐화 영역을 암호화하는 핵산 서열, 및 종결 서열을 포함하는 플라스미드가 제공된다. 이러한 플라스미드는 선택된 숙주에서 발현하기 위한 전분-캡슐화 영역 및 페이로드 폴리펩티드를 암호화하는 DNA 서열의 삽입을 위해 적당한 것이다.

본 발명의 플라스미드는 SER을 암호화하는 핵산과 페이로드 폴리펩티드를 암호화하는 핵산 사이의 융합 부위에 근접한 스페이서(spacer) 또는 링커 유니트(linker unit)를 임의로 포함할 수 있다. 본 발명은 원핵 또는 진핵 숙주에 적합한 프로모터를 포함하는 플라스미드를 포함한다. 또한, 이러한 프로모터는 외떡잎식물 및 쌍떡잎식물에서의 발현에 특히 적합하다.

본 발명의 펩티드-변경된 전분을 형성하는 방법은 전분-캡슐화 영역을 암호화하는 핵산과 회합된 프로모터를 갖는 플라스미드를 공급하는 단계와, 상기 플라스미드를 이용하여 숙주를 형질변환하여 상기 숙주가 펩티드-변경된 전분을 발현하도록하는 단계를 포함하는데, 상기 전분-캡슐화 영역을 암호화하는 핵산 서열은 페이로드 폴리펩티드를 암호화하는 핵산 영역에 연결된다.

또한, 본 발명은 배자; 영양 조직; 및 변경된 전분 입자를 포함하는 전분-함유 입자를 포함하는 전분-함유 곡물을 포함하는데, 상기 변경된 전분 입자는 변경되지 않은 전분 입자에 내생하지 않는 캡슐화된 단백질 입자를 갖는다. 이러한 전분 함유 곡물은 배자가 옥수수 배자, 쌀 배자, 또는 밀 배자인 곡물일 수 있다.

본원에 나타낸 모든 간행물들은 본 발명과 불일치하지않는 정도까지 본원에 참조로 인용된다.

본 발명은 잡종 폴리펩티드, 잡종 폴리펩티드를 제조하는 방법, 및 이러한 잡종 폴리펩티드를 암호화하는 핵산을 제공한다. 잡종 폴리펩티드는 하나의 펩티드 사슬로 일제히 융합된 두 개이상의 부분(subpart)들로 이루어진다. 이러한 부분은 아미노산 또는 펩티드 또는 폴리펩티드일 수 있다. 이러한 부분들중 하나는 전분-캡슐화 영역이다. 따라서, 잡종 폴리펩티드는 이러한 잡종 폴리펩티드를 발현하는 유기체에 의해 생성된 전분 입자내로 도입(target)될 수 있다.

세포내에서 잡종 폴리펩티드를 제조하는 방법은, 중요한 폴리펩티드(페이로드 폴리펩티드)를 암호화하는 DNA 서열에 연결되고, 부착 DNA의 발현 생성물을 전분 입자상에 결합하도록 작용하는 서열을 암호화하는 DNA의 최소한 하나의 단편을 포함하는 DNA 구성물의 제조를 포함한다. 이러한 구성물은 원핵 또는 진핵 세포내에서 발현된다. 잡종 폴리펩티드는 정제된 단백질을 생성하거나, 중요한(페이로드) 단백질을 전분 입자내에서 고정화하거나, 또는 외부 아미노산 또는 펩티드를 함유하는 입자를 생성하는데 사용될 수 있다.

본 발명에 따른 잡종 폴리펩티드는 하기와같은 세 개의 영역을 가진다.

페이로드 펩티드(X)	중심 부위(CS)*	전분-캡슐화 영역(SER)
-------------	------------	----------------

여기서, X는 중요한 관심있는 어떠한 아미노산 또는 펩티드임.

\* 임의 성분.

X에 대한 유전자는 아래에 기술된 DNA 구성물내에서 5' 또는 3' 위치에 있을 수 있다.

CS는 당업계에 알려져있는 바와같은 이탈 부위(leaving site), 분열 부위, 또는 스페이서일 수 있다. 분열 부위는 분열 효소에 의해 인식된다. 분열 효소는 특별한 부위에서 펩티드를 분열시키는 효소이다. 펩티드를 분열시키기위해 이용되어온 효소의 예로는 트롬빈, 트립신, 시아노브로마이드, 포름산,

하드록실 아민, 콜라게나제, 및 알라서브틸리신(alasubtilisin)이 있다. 스페이서는 잡종 폴리펩티드를 포함하는 펩티드들을 결합시키는 펩티드이다. 보통, 스페이서는 펩티드들을 결합시키거나, 어느정도의 최소 거리를 유지하거나, 또는 단백질의 중첩(folding), 전하, 및 물의 수용도(water acceptance)에 영향을 주는 것 이외에 어떠한 다른 활성을 갖지는 않는다. 스페이서는 잡종 폴리펩티드의 생물학적 활성을 방해하지 않는 어떠한 펩티드 서열일 수 있다.

전분-캡슐화 영역(SER)은 전분에 대하여 결합 친화성을 가지는 폴리펩티드의 영역이다. 보통, SER은 식물의 분지 효소 및 전분 신타아제의 전분-결합 영역을 포함하는 펩티드들로 이루어진 그룹에서 선택되지만, 글루코아밀라제 등과같은 다른 소오스로부터의 전분-결합 영역을 포함할 수도 있다. 본 발명의 바람직한 실시예에 있어서, SER은 전분 합성 경로중에 자연 발생하는 유전자의 펩티드 생성물을 포함한다. 이러한 바람직한 SER들은 전분-형성 캡슐화 영역(SFER)으로 정의된다. 이밖에, 본 발명에서 바람직한 SER은 특정의 효소 전분 신타아제(STS), 입자-결합된 전분 신타아제(GBSTS) 및 전분 함유 식물의 분지 효소(BE)로부터의 특정의 전분-캡슐화 영역(SSER)이다. 가장 바람직한 유전자 생성물은 GBSTS이다. 그밖에, 전분 신타아제 I 및 분지 효소 II가 유용한 유전자 생성물이다. 바람직하게, SER(및 위에서 열거한 모든 전분 캡슐화 영역)은, 절단부가 전분-캡슐화 영역을 포함하도록 하는 완전한 길이의 전분 합성 효소 유전자의 절단된 버전(truncated version)이다.

숙주내에서 잡종 폴리펩티드를 발현하기위한 DNA 구성물은 아래와같다.

프로모터	인트론*	트렌싯 펩티드 암호화 영역*	X	SER	터미네이터
------	------	--------------------	---	-----	-------

여기서,\*는 임의 성분이다. 다른 임의 성분들이 또한 사용될 수 있다.

이 기술분야에 알려져있는 바와같이, 프로모터는 전사 제어 DNA의 영역이다. 상이한 유형의 프로모터들이 상이한 숙주들에 대하여 선택된다. Lac 및 T7 프로모터들은 원핵생물에서 잘 작동하고, 35S CaMV 프로모터는 쌍떡잎식물에서 잘 작동하고, 폴리유비퀴틴 프로모터는 많은 외떡잎식물에서 잘 작동한다. 어떠한 수의 상이한 프로모터들이 이 기술분야에 알려져있고 본 발명의 범위내에서 사용될 수 있다.

또한 이 기술분야에 알려져있는 바와같이, 인트론은 유전자 생성물을 암호화하지않는 유전자내부의 누클레오티드 서열이다. 외떡잎식물에서의 발현을 흔히 증가시키는 인트론의 한 가지 예는 Adh1이다. 상기 구성물에서 이러한 성분은 임의적인 것이다.

트렌싯(transit) 펩티드 암호화 영역은 색소체와 같은 세포기관내로 단백질의 전좌를 암호화하는 누클레오티드 서열이다. 트렌싯 펩티드를 이용하는 숙주에 의해 인식되고 또 그것과 양립할 수 있는 트렌싯 펩티드를 선택하는 것이 바람직하다. 본 발명에 있어서 선택되는 플라스티드는 전분체이다.

잡종 폴리펩티드는 전분체에서 전분을 합성하고 저장하는 식물세포와 같은 세포내의 전분체내에 위치하는 것이 바람직하다. 숙주가 전분체를 함유하지않는 세균 또는 다른 세포인 경우, 트렌싯 펩티드 암호화 영역이 있을 필요는 없다.

터미네이터는 전사를 종결하는 DNA 서열이다.

X는 페이로드 폴리펩티드에 대한 암호화 영역으로서, 중요한 어떠한 폴리펩티드 또는 아미노산의 사슬일 수 있다. 이것은 공지된 폴리펩티드의 전체 서열을 가질 수 있거나 또는 그것의 유용한 단편을 포함할 수 있다. 상기 페이로드 폴리펩티드는 폴리펩티드, 또는 이것의 단편, 또는 효소, 호르몬, 성장 인자, 면역글로불린, 색소 등과같은 생물학적 활성 단백질일 수 있다. 본 발명에서 이용될 수 있는 페이로드 폴리펩티드의 일부의 예로는 프로락틴(PRL), 혈청 알부민, 성장 인자 및 성장 호르몬 즉, 소메토트로핀이 있다. 혈청 알부민으로는 소, 양, 말, 닭, 및 사람의 혈청 알부민이 있다. 성장 인자로는 표피 성장 인자(EGF), 인슐린-유사 성장 인자 I (IGI-I), 인슐린-유사 성장 인자 II (IGI-II), 섬유아세포성장인자(FGF), 형질 변환 성장 인자 알파(TGF-alpha), 형질변환 성장인자 베타(TGF-beta), 신경 성장 인자(NGF), 혈소판-유도된 성장 인자(PDGF), 및 재조합된 사람의 인슐린-유사 성장 인자 I (rHuIGF-I) 및 II (rHuIGF-II)가 있다. 본 발명을 실시하기위해 이용될 수 있는 성장 호르몬으로는 소, 돼지, 양, 말, 닭, 및 사람의 성장 호르몬이 있다. 돼지의 성장 호르몬으로는 유럽 특허 출원 공보제 104,920호(Biogen)에서 설명되고 청구된 바와같은 델타-7 재조합 돼지 성장 호르몬이 있다. 바람직한 페이로드 폴리펩티드는 성장 호르몬, 인슐린 A 및 B 사슬, 칼시토닌, 베타 엔도르핀, 우로가스트론, 베타 글로빈, 마이오글로빈, 사람의 성장 호르몬, 안지오텐신, 프롤린, 프로테아제, 베타-갈락토시다제, 및 셀룰라리아제가 있다.

또한, 잡종 폴리펩티드, SER 영역 및 페이로드 폴리펩티드는 글리코실화, 아실화, 및 펩티드의 원하는 활성을 방해하지않는 다른 변경과 같은 이 기술분야에 알려져있는 후번역 변경을 포함할 수도 있다.

잡종 폴리펩티드의 발생.

SER은 생합성에 관련된 유전자내에 존재한다. 이러한 유전자를 분리하기 위한 방법은 게놈 DNA 라이브러리 및 cDNA 라이브러리로부터 스크리닝하는 것을 포함한다. 이러한 유전자는 결찰, 돌연변이생성제, 소화, 제한 및 다음문헌[Maniatis 등, Molecular Cloning, Cold Spring Harbor Labs, Cold Spring Harbor, N.Y]에 설명된 다른 이와같은 과정에 따라 절단 및 변화될 수 있다. SER 영역에 접근하기위한 우수한 출발물질의 예로는 전분 신타아제 I, II, III, IV, 분지효소 I, II A 및 B 및 입자-결합된 전분 신타아제(GBSTS)가 있다. 이러한 유전자는 벼, 옥수수, 완두콩, 감자, 밀등과 같은 전분 함유 식물에 존재한다. DNA 또는 cDNA 또는 mRNA 또는 SER에 대하여 발생된 형체로 만든 SER 프로브를 사용하면, 클로닝에 유용한 유전자의 분리 및 확인이 가능해진다. 이러한 전분 효소-암호화 서열은 회합된 폴리펩티드를 캡슐화하기위한 SER 영역의 능력을 방해하지않는다는 조건으로 변경될 수 있다.

전분 입자내에 캡슐화되는 단백질을 암호화하는 유전자가 위치하는때, 이 기술분야에 잘 알려져있는 바와같은 몇 가지의 방법이 SER을 분리하기위해 이용될 수 있다. 한 가지 방법은 유전자를 제한 효

소를 이용하여 여러 가지의 부위에서 절단하여 N-말단으로부터 부위들을 제거하여 얻어지는 단백질을 발현시키는 것이다. 이어서, 발현된 절단된 단백질은 전분 겔상에 놓여져서 잔류 단백질의 회합 및 해리 상수가 구해진다. 이 기술분야에 알려져있는 표지 유전자, 예를들어 그린 형광 단백질 유전자가 절단된 단백질에 부착되어 전분 입자내에서 표지 유전자의 존재를 결정하기 위하여 사용될 수 있다.

SER 유전자 서열 영역이 분리되면 이것은 전분내에 캡슐화된 페이로드 폴리펩티드를 발현시키는 유전자 단편 서열을 제조하는데 사용될 수 있다. SER 유전자 서열 및 페이로드 폴리펩티드를 암호화하는 유전자 서열은 일제히 결찰될 수 있다. 이어서, 얻어지는 융합된 DNA는 여러 가지의 숙주에서 발현하기 위한 다수의 벡터 구성물에 위치할 수 있다. 바람직한 숙주는 색소체내에서 전분 입자를 형성하지만, SER의 테스트는 *E. coli*와 같은 세균 숙주내에서 쉽게 수행될 수 있다.

페이로드 폴리펩티드를 암호화하는 핵산 서열은 DNA, RNA, 게놈 DNA, cDNA, mRNA로부터 유도될 수 있거나, 또는 전체적으로 또는 부분적으로 합성될 수 있다. 페이로드 폴리펩티드의 서열은, 돌연변이를 함유하여 생성된 단백질이 생물학적 기능을 유지한다는 조건으로 새로운 돌연변이 단백질이 되도록 조작될 수 있다.

페이로드 폴리펩티드-암호화 핵산 서열이 SER-암호화 서열에 결찰되는 때, 이러한 페이로드 폴리펩티드에 대한 유전자 서열은 N-말단을 암호화하는 SER 서열의 말단에서 부착되는 것이 바람직하다. N-말단이 바람직하긴 하지만, 페이로드 폴리펩티드가 SER의 N-말단에 결찰되는지 또는 C-말단에 결찰되는지는 본 발명에 있어서 중요한 것으로 여겨지지 않는다. 명백히, 합성에 의해서든지 또는 클로닝 또는 결찰에 의해서든지 상관없이 본 발명의 재조합 핵산 분자를 형성하는 방법은 본 발명에 있어서 중요한 것이다.

잡종 폴리펩티드의 중심 영역은 임의적인 것이다. 본 발명의 일부의 사용에 있어서 이러한 영역에서 적절한 프로테아제 분열 부위를 암호화하는 DNA를, 잡종 폴리펩티드를 발현하기 위해 사용되는 재조합 분자내로 도입하는 것은 매우 유용할 수 있다. 선택적으로, 중심 영역을 형성하는데 pH-민감성인 아미노산 서열을 암호화하는 DNA를 도입하는 것이 유용할 수도 있다. 본 발명이 프로테아제등에 의해 전분 입자로부터 추출 및 방출될 수 있는 순수한 단백질을 발달시키기 사용되는 경우에는, 프로테아제 분열 부위가 유용하다. 그밖에, 단백질이 동물에서 소화되어야하는 경우에는 프로토아제 분열 부위는 동물의 소화관에서 효소가 전분으로부터 단백질을 방출시키는 것을 돋는데에 유용할 수 있다.

중심 부위는 스페이서를 포함할 수 있다. 스페이서는 잡종 폴리펩티드를 포함하는 단백질을 결합시키는 펩티드를 의미한다. 보통 스페이서는 단백질들을 결합시키거나, 어느정도의 최소한의 거리를 유지하거나, 또는 잡종 단백질의 겹침, 전하, 또는 소수성 또는 친수성에 영향을 미치는 것 이외에 다른 어떠한 특별한 활성을 갖지는 않는다.

구성물 발생.

잡종 폴리펩티드를 암호화하는 결찰된 DNA가 형성되면, DNA를 잡종 폴리펩티드를 발현시키기 위한 숙주에 전달할 수 있는 클로닝 벡터 또는 플라스미드가 제조된다. 본 발명의 재조합 핵산 서열이 적절한 클로닝 벡터 또는 플라스미드내로 삽입된다. 본 발명에 있어서 바람직한 숙주는 전분 입자-생성 숙주이다. 그러나, 세균 숙주가 이용될 수도 있다. 특히 유용한 것은 식물의 일부 또는 전부의 전분-합성 유전자를 함유하도록 형질변환된 세균 숙주이다. 당업자는 플라스미드가 숙주에 알맞는 것으로 이해하고 있다. 예를들어, 세균 숙주에 있어서 전사 조정 프로모터로는 lac, TAC, trp등이 있다. 그밖에, 트랜싯 펩티드(transit peptide)를 암호화하는 DNA는 거의 사용될 수 없었고 이러한 구조적 유전자보다 앞선 비밀스런 선두물질(secretory leader)이 폴리펩티드를 매질내로 도입하기 위해 사용될 수 있다. 선택적으로, 이러한 생성물은 숙주내에서 유지되고 숙주는 분해되고 생성물이 전분 추출법에 의해서 또는 물질을 전분 기질(또는 아밀로스 또는 아밀로펙틴, 글리코겐 등)에 결합시켜서 생성물을 추출함으로써 분리 및 정제될 수 있다.

바람직한 숙주가 식물이므로 바람직한 플라스미드는 식물에서 사용하기에 적합하다. 플라스미드는 프로모터, 바람직하게는 식물의 전분-함유 조직에서 단백질의 발현을 목표로하도록 적합화된 프로모터를 함유한다. 이러한 프로모터는 씨, 뿌리, 괴경등과 같은 여러 가지의 조직에 대하여 특이성이 있을 수 있거나, 또는 식물의 조직도처에서 유전자를 발현하기 위한 구성 프로모터일 수 있다. 잘 알려진 프로모터로는 10 kD 제인(옥수수) 프로모터, CAB 프로모터, 파타스틴(patastatin), 35S 및 19S 콜리플라워 모자이크 바이러스 프로모터(쌍떡잎식물에서 매우 유용함), 폴리유비퀴틴 프로모터(외떡잎식물에서 유용함) 및 이 기술분야에 알려져있는 상기 것들의 개량물 또는 변형물이 있다.

클로닝 벡터는 플라스미드를 정확한 위치로 항하게하기 위한 트랜싯 펩티드 암호화 서열을 함유할 수 있다. 트랜싯 펩티드 암호화 서열의 예들이 서열표에서 나타나있다. 다른 트랜싯 펩티드에 대한 암호화 서열이 사용될 수 있다. 사용되는 숙주에서 자연 발생하는 트랜싯 펩티드가 바람직하다. 옥수수에 대한 바람직한 트랜싯 펩티드 암호화 서열이 표 및 도면에서 확인된다. 트랜싯 펩티드의 목적은 벡터를 세포내의 정확한 영역에 조준하는데 있다.

트랜싯 펩티드 암호화 서열에 부착된 것은 페이로드 폴리펩티드의 N-말단을 암호화하는 DNA 서열이다. 이러한 페이로드 폴리펩티드를 암호화하는 서열의 방향은 감지 전사(sense transcription)가 요구되는가 또는 항감지(antisense) 전사가 요구되는가에 따라 변화한다. 본원에서 상세하게 설명된 DNA 구성물은 N-말단에서 페이로드 폴리펩티드를 암호화하는 서열을 가지지만 SER 암호화 영역이 N-말단 및 페이로드 폴리펩티드 서열 이후에 있을 수도 있다. DNA 구성물의 말단에는 터미네이터 서열이 있다. 이러한 서열은 이 기술분야에 잘 알려져있다.

클로닝 벡터는 숙주내로 도입되어 형질변환된다. 클로닝 벡터, 바람직하게 플라스미드를 숙주내로 도입하는 것은 이 기술분야에 알려져있는 다수의 형질변환 방법에 의해 수행될 수 있다. 이러한 방법은 숙주에 따라 변화할 수 있지만 예로는 미립자 충격법, 미세 주입법, *Agrobacterium* 형질변환법, 'whiskers' 기술(미국특허 제 5,302,523호 및 5,464,765호), 전기천공법등이 있다. 숙주가 식물인 경우, 세포는 재생되어 식물을 형성할 수 있다. 숙주가 형질변환되고 단백질이 숙주안에서 발현되면, 숙주내에

서 페이로드 폴리펩티드를 암호화하는 DNA의 존재가 확인될 수 있다. 발현된 단백질의 존재는 Western Blot 또는 ELISA에 의해서 또는 식물 또는 세포내의 변화의 결과에 따라 확인될 수 있다.

#### 캡슐화된 단백질의 사용.

본 발명은 다수의 용도를 가진다. 잡종 폴리펩티드는 전분으로부터 순수한 형태로 분열될 수 있고(분열 부위가 포함될 수 있음) 순수한 단백질이 회수될 수 있다. 선택적으로, 전분내에 캡슐화된 페이로드 폴리펩티드는 소비 동물(동물은 포유류, 조류 및 어류를 포함함)의 소화관의 여러 가지 부위에 단백질을 운반하기 위해 미가공의 형태로 사용될 수 있다. 예를들어, 물질이 캡슐화되어있는 전분이 소화에 내성이 있는 경우, 단백질은 동물의 장내로 서서히 방출되므로 값어치있는 단백질이 위장에서 분해되는 것이 방지된다. 메티오닌 및 리신과같은 아미노산이 동물에게 공급되는 입자내로 직접 흡입되도록 캡슐화됨으로써, 이러한 아미노산을 갖는 식이를 다른 형태로 보충하기위한 필요성이 제거된다.

본 발명은 호르몬, 효소, 단백질, 단백질성 영양소 및 단백질성 의약품을 동물의 소화관내의 특정의 소화 부위로 조준하는 것을 가능하게한다. 전분내에 캡슐화되고 상부 소화관에서 일반적으로 소화되는 단백질은 비소화 방식으로 위를 통과할 수 있고 장에 부분적으로 또는 완전히 흡수될 수 있다. 장벽을 통과할 수 있는 경우, 페이로드 폴리펩티드는 동물을 약으로 치료하거나, 또는 동물의 예방접종을 위해서 또는 동물에게 이용될 수 있는 영양소를 강화하기 위해서 성장 호르몬과 같은 성장 인자를 제공하기위해 사용될 수 있다.

사용된 전분이 위에서의 소화에 대하여 내성이 없는 경우, 추가된 단백질이 동물의 상부 소화관에서 흡수되도록 목표될 수 있다. 이를위하여, 변경된 전분을 생성하기위해 사용된 숙주는 슈가리 2 타입(sugary 2 type) 전분이 만들어지도록 돌연변이생성 또는 형질변환될 필요가있다. 본 발명은 변경된 전분을 형성하는 돌연변이생성 유기체를 숙주로서 사용하는 것을 포함한다. 이러한 돌연변이생성 숙주의 일부의 예로는 슈가리 1 돌연변이, 슈가리 2 돌연변이, 취성(brITTLE) 돌연변이, 수축된(shrunken) 돌연변이, 밀같은(waxy) 돌연변이, 아밀로스 익스텐더 돌연변이, 흐린(dull) 돌연변이, 불투명(opaque) 돌연변이, 가루가많은(floury) 돌연변이 등이있다. 이러한 돌연변이 전분 및 상이한 식물원으로부터의 전분은 상이한 수준의 소화성을 가진다. 따라서, DNA의 발현을 위한 숙주 및 변경된 전분이 공급되는 동물을 선택함으로써, 잡종 단백질이 목표되는 곳에서 소화될 수 있다. 상이한 단백질들이 상이한 신체 부분들에 가장 효율적으로 흡수된다. 선택된 소화성을 갖는 전분으로 단백질을 캡슐화시킴으로써, 단백질이 소화관도처의 어디에서든지 그리고 소화 과정 동안의 특정의 시기에 공급될 수 있다.

본 발명의 또다른 이점은 원하는 폴리펩티드의 글리코실화를 상이한 수준으로 발현하거나 또는 억제하려는 능력이다. 캡슐화 과정은 단백질이 다른 DNA 분자에 의해 발현되는 경우와 상이한 글리코실화 상태로 입자내에서 발현되는 것을 가능하게한다. 글리코실화는 캡슐화의 양, 이용되는 숙주 및 폴리펩티드의 서열에 의존할 수 있다.

위에서 설명한 특징을 갖는 개량된 곡물이, 다른 바람직한 특징을 갖는 것으로 알려져있는 식물의 유전자 조작에 의해서 생성될 수 있다. 전분-합성 효소 유전자의 누클레오티드 서열을 조작함으로써, 식물에서 생성되는 핵심 아미노산, 단백질 또는 펩티드의 양을 변경시킬 수 있다. 식물, 진균, 세균 또는 동물의 기원을 갖는 한 가지 이상의 유전자 공학적 유전자 구조물이 성적 교차 또는 형질변화에 의해서 식물 개놈내로 도입될 수 있다. 이러한 공학적 유전자는 야생형 유전자의 추가의 복제를 포함할 수 있거나 또는 새로운 성질을 갖는 변경 또는 대립 또는 교번 효소를 암호화할 수 있다. 이러한 유전자 구성을 (들)을 도입하면, 감지 또는 향감자 배열로 도입된 유전자(들)의 양 및 유형에 따라 여러 가지의 효과가 얻어진다. 이것은 특이 단백질 또는 펩티드를 생성하려는 식물의 능력을 증가시키거나 또는 개선된 아미노산 균형을 제공하려는 식물의 능력을 증가시킬수 있다.

#### 전분 생합성에 포함된 효소의 클로닝.

본 발명의 DNA 구성을 제공하기위해 공자의 클로닝법이 사용될 수 있다. 특별한 형태의 SSTS, GBSTS, BE, 글리코겐 신타아제(GS), 아밀로펙틴, 또는 본원에서 사용된 다른 유전자의 소오스는 전분 또는 글리코겐을 만들 수 있는 어떠한 유기체일 수 있다. 가능한 공여 유기체가 스크리닝 및 확인된다. 이어서, 아래와같은 두 가지의 방법이 수행될 수 있다. (a)본원에 설명된 프로토콜후 효소 정제 및 형체/서열 발생을 사용하며; (b)관련된 유기체로부터의 라이브러리에서 SSTS, GBSTS, BE, GS, 아밀로펙틴, 또는 다른 전분-캡슐화 효소에 대한 게놈 DNAs를 확인하기위한 이종 프로브로서 SSTS, GBSTS, BE, GS, 아밀로펙틴 또는 다른 cDNA를 사용한다. 유전자 형질변환, 식물 재생, 및 테스트 프로토콜은 이 기술분야에 알려져있다. 이러한 경우에 있어서, 전분 형성동안 발현을 확실히하는 조절 서열을 함유하는 형질변환을 위한 유전자 구성을 만들 필요가 있다. 이러한 조절 서열은 많은 소립자와 괴경 및 뿌리에 존재한다. 예를들어 이러한 조절 서열은 입자 결합 전분 신타아제 합성(GBSTS), 가용성 전분 신타아제 (SSTS), 또는 분지 효소(BE)에서 옥수수 배젖으로서 또는 전분 합성 경로 효소에서 다른 옥수수 배젖으로서 쉽게 입수 할 수 있다. 상기 배젖으로부터의 조절 서열은 정확한 발생시기에 단백질의 발현을 확실하게한다(예를들어, ADGP 피로포스포릴라제).

이 방법에 있어서 본 발명의 발명자들은 글리코겐 또는 아밀로펙틴과같은 적당한 농도의 탄수화물의 존재하에 자생 단백질 전기영동법을 이용하여 전분-결합 단백질의 전분-결합 함량을 측정했다. 전분-캡슐화 영역은 지정부위돌연변이유발법 또는 당업자에게 알려져있는 다른 유전자 공학 방법을 이용하여 밝혀질 수 있다. 새로운 펩티드 또는 아미노산 조합을 가지는 새로운 유전자 공학적 단백질은 본원에 설명된 방법을 사용하여 평가될 수 있다.

#### 실시예

##### 실시예 1:

###### 전분-캡슐화 단백질의 확인 방법.

###### 전분-입자 단백질 분리:

12.5g 입자를 25ml 추출 완충액(50mM 트리스 아세테이트, pH 7.5, 1mM EDTA, 1mM DTT)에서 3x20초동안 Waring 블렌더를 이용하여 1분의 배합 간격으로 균질화한다. 시료를 얼음위에서 유지한다. 미라 천(mira cloth)을 통해 여과하고 6,000rpm으로 30분간 원심분리한다. 상징액을 버리고 백색 전분 펠릿위에 있는 변색된 고체를 버린다. 펠릿을 25ml 완충액에 재현탁하고 재원심분리한다. 세척을 2회 이상 반복한다. 재현탁 및 세척된 펠릿은 -20°C에 두는 것이 가능하다. 전분을 공기 흐름중에서 건조시킨다. -20°C에서 저장한다.

단백질 추출:

전분 50mg을 에펜도르프(eppendorf)에서 2%SDS와 혼합한다. 4°C에서 5분간 18,000rpm으로 와동 회전시킨다. 1ml 시료를 완충액(4ml 종류수, 1ml M Tris-HCl, pH 6.8, 0.8 ml 글리세롤, 1.6ml 10% SDS, 0.4ml B-메르캡토에탄올, 0.2ml 0.5% 브롬페놀 블루)에 가한다. 엔도르프를 덮개의 구멍을 이용하여 10분간 끓인다. 냉각하고, 10,000rpm으로 10분간 원심분리한다. 상징액을 새로운 엔도르프내로 경사분리한다. 위에서 설명한 바와같이 4분간 끓인다. 냉각한다.

SDS-PAGE 걸: (변성되지않음)

	10%재용해	4%스택(stack)
아실/비스(Bis)40%스톡(stock)	2.5 ml	1.0ml
1.5M Tris pH 8.8	2.5ml	-
0.5M Tris pH 8.8	-	2.5ml
10% SDS	100 $\mu$ l	100 $\mu$ l
물	4.845ml	6.34ml

15분 탈기후 새로이 추가

10%과황산암모늄	50 $\mu$ l	50 $\mu$ l
TEMED	5 $\mu$ l	10 $\mu$ l

미니-프로테안 II 이중 슬라브 세포(Mini-Protean II Dual Slab Cell); 걸당 3.5ml의 재용해 완충액. 4% 스택을 상부(top)에 놓는다. 걸은 200V 일정 전압에서 런닝(running)한다. 10 x 런닝 완충액(250 mM Tris, 1.92 M 글리신, 1%SDS,pH8.3).

전분-캡슐화 영역의 측정 방법:

용액:

추출 완충액: 50 mM Tris-아세테이트 pH 7.5, 10 mM

EDTA, 10% 수크로스, 2.5 mM

DTT-프레시

스택킹 완충액: 0.5M Tris-HCl, pH 6.8

재용해 완충액: 1.5M Tris-HCl, pH 8.8

10X 하부 전극 완충액: 30.3g Tris+144g 글리신qs 1L.(pH는 ~8.3, 조절없음). 사용을 위해 희석.

상부 전극 완충액: 하부 전극과 동일함.

수크로스 용액: 18.66g 수크로스+100 ml dH<sub>2</sub>O

30%아실/비스 스톡(2.67%): 146g 아크릴아미드+4 g비스+25ml dH<sub>2</sub>O.  
500ml 수준에 도달시킴. 여과하고 어둠 속에서 1달 동안 4°C에서 저장함. 500ml 수준에 도달시킴.

15%아실/비스 스톡(20%): 6g 아크릴아미드+1.5g 비스+25ml dH<sub>2</sub>O. 여과하고 어둠 속에서 1달 동안 4°C에서 저장함. 500ml 수준에 도달시킴.

리보플라빈 용액: 1.4g 리보플라빈+100 dH<sub>2</sub>O. 어둠 속에서 1 달간 저장함.

SS 평가분석 믹스( mix): 25mM 시트르산나트륨, 25mM Bicine-NaOH(pH 8.0), 2mM EDTA, 1mM DTT-프레시, 1 mM 아데노신 5' 디포스포글루코스-프

레시, 10mg/ml 토끼간 글리코겐 타입 III-프레시.

요오드 용액: 2g 요오드+20g KI, 0.1 N HCl 1 리터 이하.

추출:

4ml 추출 완충액+12g 배젖. 균질화한다.

미라천 또는 4층 치즈천을 통해 여과한다. 20,000g(14,500 rpm, SM-24

로우터)회전시킨다. 20분. 4°C.

글래스 피펫을 사용하여 상징액을 제거한다.

0.85ml 추출물+0.1ml 글리세롤+0.05ml 0.5% 브로모페놀 블루.

미세구축물(microfuge)을 충분한 속도로 5분간 와동 회전시킨다. 직접사

용하거나 또는 액체 질소에서 동결시키고 -80°C에서 2주이하동안 저장한다.

캐스트 겔(cast gel):

양면 스카치 테이프의 친수성 면를 사용하여 Gel Bond PAG 필름(FMC

Industris, Rockland, ME)을 외부 유리판에 부착한다. 상기 테이프 및 필름을 상기 판의 바닥부와 가능하면 접근하게하고 평평하게 정렬한다. 상기 필름은 상기 판보다 약간 더 작다. 필름과 판의 사이로 물을 분출시켜서 필름을 부착한다. 과량의 물을 배출하기위한 조직을 사용한다. 판들을 보통상태로 세우고 판들의 바닥부를 점착성 점착제로 밀봉한다. 회색 러버가 캐스팅 스탠드(casting stand)로부터 제거되지않는 경우 그 캐스팅 스탠드내로 카셋트를 맞추어넣는다. 겔을 필름과함께 종합하고 차후의 모든 조작동안 부착된 채로 유지한다.

캐스트 4.5% T 재용해 미니-겔(0.75g):

2.25ml dH<sub>2</sub>O

+3.75ml 수크로스 용액

+2.5ml 재용해 완충액

+1.5ml 30%아실/비스 스톡

+각각의 겔에 대하여 여러 가지 양의 글리코겐(즉, 0-1.0%)

15분간 탈기

+50μl 10% APS

+5μl TEMED

30분 또는 밤새 종합

캐스트 3.125% T 스택:

1.59ml dH<sub>2</sub>O

+3.75ml 수크로스 용액

+2.5ml 스택 완충액

+2.083ml 15% 아크릴/비스 스톡

탈기하지않음.

15μl 10% APS

+35μl 리보플라빈 용액

+30μl TEMED

라이트 별브(light bulb)에 근접하여 2.5시간동안 종합.

콤(comb)을 잡아뽑기전에 4°C에서 냉각한다. 또한 콤은 사용할 수 없고 1

센티미터의 스태커를 캐스트할 수 있다.

상기의 과정:

상이한 온도에서 진행할 수 있음: 겔 및 용액을 예비배양함.

200V에서 15분동안 예비진행함.

겔을 적재시킴: 웰(well)당  $7\mu\text{l}$ , 또는 콤이없는 경우에는  $115\mu\text{l}$ .

색소 프론트( front)가 바닥부에 접근할 때 까지 140V에서 진행함. 여러

가지의 진행온도는 전체의 겔 장비를 물중탕내에 위치시킴으로써 달성됨.

효소 평가분석: 색소 프론트에서 겔을 절단한다. SS에서 배양한다. 평가분

석 믹스를 온건한 교반과함께 실온에서 밤새 유지함. 겔을 물로 세정한다.

I2K/KI 용액을 이용하여 플러딩(flooding)시킨다. 라이트 박스(light box)에

서 겔의 사진을 찍고 그 사진을 평가한다.  $R_m$ =겔의 상부로부터 활성 밴드까

지의  $\text{mm}$ /겔의 상부로부터 겔의 바닥부까지의  $\text{mm}$ .  $1/R_m$ 에 대한 글리코겐%를 플

로트한다. 선이  $x$  축을 교차하는 점은  $-K(y=0)$ 이다.

SER 영역 길이에 대한 테스트 및 평가 프로토콜:

SER의 선택을 위한 상기의 과정후, 4개의 기본 단계가 필요하다. 우선 전분-암호화 영역을 갖는 단백질을 암호화하는 DNA가 선택되어야한다. 이것은 예를들어 아밀라제에 대한 유전자와같은 공지의 전분-결합 유전자 또는 전분-합성 유전자로부터 선택될 수 있다. 단백질이 추출되어야한다. 다수의 단백질 추출 방법이 이 기술분야에 잘 알려져있다. 이러한 단백질은 프로테아제에 의해 처리되어 상이한 길이의 단백질 단편이 형성될 수 있다. 바람직한 단편은 단백질의 N-말단으로부터 1차적으로 삭제된 부분을 가진다. 이러한 SER 영역은 N-말단보다는 C-말단에 더욱 가까이 위치한다. 이러한 단백질은 상기에 기술된 바와같이 겔상에 두어지고 겔 기질에 대한 친화성이 평가된다. 더욱높은 친화성은 단백질의 그 영역이 기질을 더욱 선호한다는 것을 나타낸다. 이러한 방법은 상이한 단백질들의 비교를 통해 전분 캡슐화 영역이 천연 단백질인지 합성 단백질인지 확인하는 것을 가능하게한다.

실시예 2:

SER 융합 벡터:

하기의 융합 벡터들이 *E. coli*에서의 사용을 위해 적용된다. 이러한 벡터에서 예상되는 SER에 부착된 융합 유전자는 유전자 형광 단백질(GFP)를 암호화했다. 단백질 및 폴리펩티드를 암호화하는 어떠한 수의 상이한 유전자들이 상기의 벡터에 결합될 수 있다. 이러한 경우의 GFP에 있어서, 제 2 유전자 또는 유전자 단편에 융합된 waxy 옥수수의 SER을 갖는 융합 벡터를 구성했다.

pEXS114(도 1a참조): 합성GFP를 프라이머 EXS74(5'-GACTAGTCATATG GTG AGC AAG GGC GAG GAG-3')[서열 번호 1] 및 EXS74(5'-CTAGATCTTCATATG CTT GTA CAG CTC GTC CAT GCC-3')[서열번호 2]를 사용하여 플라스미드 HBT-SGFP(Jen Sheen으로부터 입수; Dept. of Molecular; Wellman 11, MGH; Boston, MA 02114)로부터 PCR-증폭했다. 이러한 PCR 생성물의 말단을 T DNA 폴리머라제를 이용하여 연마하여 무딘 말단을 발생시킨다음, 상기 PCR 생성물을 Spe I를 이용하여 소화시켰다. 이러한 SGFP 단편을 pBSK(Stratagene, 11011North Torrey Pines Rd. La Jolla, Ca. 소재)의 EcoRV-Spe I 부위내에 서브클론화하여 pEXS114를 발생시켰다. pEXS(도 1b 참조): 합성 GFP(SGFP)를 프라이머 EXS73 및 EXS75(5'-CTAGATCTTGGC CTT GTA CAG CTC GTC CAT GCC-3')[서열번호 3]을 사용하여 플라스미드 HBT-SGFP(Jen Sheen으로부터 입수)로부터 PCR-증폭했다. 이러한 PCR 생성물의 말단을 T DNA 폴리머라제를 이용하여 연마하여 무딘 말단을 발생시킨다음, 상기 PCR 생성물을 Spe I를 이용하여 소화시켰다. 이러한 SGFP 단편을 pBSK(Stratagene)의 EcoRV-Spe I 부위내에 서브클론화하여 pEXS115를 발생시켰다.

pEXSWX(도 2a 참조):pET-21a의 NdeI-NotI 부위내에 서브클론화된 옥수수의 WX(도 2b 참조). 개농 DNA 서열 및, mRNA 서열이 발생되는 회합된 아미노산이 다음 표 1a 및 1b에서 확인되고 선택적으로 하기의 표에서 열거된 DNA가 이용될 수 있다.

표 1a

옥수수에서 waxy 유전자의 DNA 서열 및 추론된 아미노산 서열

[서열 번호 4 및 서열 번호 5]

유전자자리 ZMWAXY 4800bp DNA PLN

정의 UDP-글루코스 전분 글리코실 트랜스퍼라제에 대한 *Zea mays* waxy ( $wx^+$ ).

접근 X03935 M24258

키워드 글리코실 트랜스퍼라제; 트랜싯 펩티드; UDP-글루코스 전분  
글리코실 트랜스퍼라제; waxy 유전자자리.

소오스 옥수수

유기체 Eukaryota; Plantae; Embryobionta; Magnoliophyta;  
Liliopsida; Commelinidae; Cyperales; Poaceae.

참조문헌 1(영기 1내지 4800)

저자 Kloesgen,R.B., Gierl,A., Schwarz-Sommer, Z., 및 Saedler,H.  
 제목 ze a mays의 waxy 유전자자리의 분자 분석(Molecular  
                   analysis of the waxy locus of ze a mays)  
 저널 Mol. Gen. Genet. 203, 237-244(1986)  
 기준 완전 자동  
 코멘트 NCBI gi: 22509  
 특징 위치/한정률(qualifier)  
 소오스 1..4800  
                   /유기체='Zea mays'  
 반복 영역 283..287  
                   /주(note)='직접 반복 1'  
 반복 영역 288..292  
                   /주='직접 반복 1'  
 반복 영역 293..297  
                   /주='직접 반복'  
 반복 영역 298..302  
                   /주='직접 반복'  
 혼합(misc)특징 372..385  
                   /주='GC 스트레치(stretch)(가능성있는 조절 인자 결  
                   합 부위)'  
 혼합 특징 442..468  
                   /주='GC 스트레치(가능성있는 조절 인자 결합 부위)'  
 혼합 특징 768..782  
                   /주='GC 스트레치(가능성있는 조절 인자 결합 부위)'  
 혼합 특징 810..822  
                   /주='GC 스트레치(가능성있는 조절 인자 결합 부위)'  
 혼합 특징 821..828  
                   /주='목표 복제 부위(Ac7)'  
 CATT 시그날 821..828  
 TATA 시그날 867..873  
 혼합 특징 887..900  
                   /주='GC 스트레치(가능성있는 조절 인자 결합 부위)'  
 혼합 특징 901  
                   /주='전사 출발 부위'  
 엑손(exon) 901..1080  
                   /번호=1  
 인트론(intron) 1081..1219  
                   /번호=1  
 엑손 1220..1553  
                   /번호=2  
 트렌싯 펩티드 1233..1448  
 CDS 결합(1449..1553, 1685..1765, 1860..1958, 2055..2144, 2226..2289, 2413..2513,  
                   2651..2760, 2858..3101, 3212..3394, 3490..3681, 3793..3879, 3977..4105, 4227..4343)  
                   주/='NCBI gi: 22510  
                   /코돈 출발=1

/생성물='글루코실 트렌스퍼라제'

/번역='ASAGMNVVFGAEMAPWSKTGGLGDVLGGLPPAMAANGHRVMVV

SPRYDQYKDAWDTSVVSEIKMGDGYETVRFFHCYKRGVDRVFUDHPLFLERVWGKTEE  
 KIYGPVAGTDYRDNQLRFSLLCQAALAEAPRILSLNNNPYFSGPYGEDVVFCNDWHTG  
 PLSCYLNQSHGIYRDAKTAFCIHNISYQGRFAFSDYPELNLPERFKSSFDFIDGY  
 EKPVTEGRKINWMKAGILEADRVLTVSPYYAEELISGIARGCELDNIMRLTGITGIVNG  
 MDVSEWDPSRDKYIAVKYDVSTAVEAKALNKEALQAEVGLPVDRNIPLVAFIGRLEEQ  
 KGPDVMAAIPOLMEMVEDVQIVLLGTGKKKFERMLMSAEKFPKGKVRAVVKFNAALA  
 HHIMAGADVLAVTSRFEPGGLIQLQGMRYGTPCACASTGGLVDTIIEGKTGFHMGRSL  
 VDCNVVEPADVKKVATTLQRAIKVVGTPAYEEMVRNCMIQDLSWKGPAKNWENVLLSL  
 GVAGGEPGVEGEEIAPLAKENVAAP"

인트론 1554..1684

/번호=2

엑손 1685..1765

/번호=3

인트론 1766..1859

/번호=3

엑손 1860..1958

/번호=4

인트론 1959..2954

/번호=4

엑손 2055..2144

/번호=5

인트론 2145..2225

/번호=5

엑손 2226..2289

/번호=6

인트론 2290..2412

/번호=6

엑손 2413..2513

/번호=7

인트론 2514..2650

/번호=7

엑손 2651..2760

/번호=8

인트론 2761..2857

/번호=8

엑손 2858..3101

/번호=9

인트론 3102..3211

/번호=9

엑손 3212..3394

/번호=10

흔합 특징 3358..3365

/주='목표 복제 부위(Ac9)'

인트론 3395..3489

/번호=10

엑손 3490..3681  
           /번호=11

흔합 특징 3570..3572  
           /주='목표 복제 부위(Spm 18)'

인트론 3682..3792  
           /번호=11

엑손 3793..3879  
           /번호=12

인트론 3880..3976  
           /번호=12

엑손 3977..4105  
           /번호=13

인트론 4106..4226  
           /번호=13

엑손 4277..4595  
           /번호=14

폴리A 시그날 4570..4575

폴리A 시그날 4593..4598

폴리A 부위 4595

폴리A 시그날 4597..4602

폴리A 부위 4618

폴리A 부위 4625

영기 짐 계 935 A 1413C      1447G      1005T

## 기원(ORIGIN)

```

 1 CAGCGACCTA TTACACAGCC CGCTCGGGCC CGCGACGTCG GGACACATCT TCTTCCCCCT
 61 TTTGGTGAAG CTCTGCTCGC AGCTGTCCGG CTCCCTGGAC GTTCGGTGTGG CAGATTCACTC
121 TGGTGTCTCG TCTCCTGTGC TTCCCTGGTA GCTTGTGTAG TGGAGCTGAC ATGGTCTGAG
181 CAGGCTTAAA ATTTGCTCGT AGACGAGGGAG TACCAGCACA GCACGTTGCG GATTTCCTCTG
241 CCTGTGAAGT GCAACGTCTA GGATTGTCAC ACGCCTTGGT CGCGTCGCGT CGCGTCGCGT
301 CGATGCGGTG GTGAGCAGAG CAGCAACAGC TGGGCGGCC AACGTTGGCT TCCGTGTCTT
361 CGTCGTACGT ACGCGCGCGC CGGGGACACG CAGCAGAGAG CGGAGAGCGA GCCGTGCACG
421 GGGAGGTGGT GTGGAAGTGG AGCCCGCGCG CCGGCCGCC CGCCTGGTG GGCAACCCAA
481 AAGTACCCAC GACAAGCGAA GGCGCCAAAG CGATCCAAGC TCCGAAACGC AACAGCATGC
541 GTCGCGTCGG AGAGCCAGCC ACAAGCAGCC GAGAACCGAA CCGGTGGCG ACAGCGTATG
601 GGACGGACGC GGGCGACGCT TCCAAACGGG CCACGTACGC CGCGCTGTGC GTGCGTGCAG
661 ACGACAAGCC AAGGCGAGGC AGCCCCCGAT CGGGAAAGCG TTTTGGCGC GAGCGCTGGC
721 GTGCGGGTCA GTCGCTGGT CGCAGTGCAGGG GGGGGAAACGG GTATCGTGGG GGGCGCGGGC
781 GGAGGAGAGC GTGGCGAGGG CCGAGAGCAG CGCGCGCGCG GGTCAAGCAA CGCGCCCCAC
841 GTACTGCCCT CCCCCCTCCGC CGCGCGCTAGA AATACCGAGG CCTGGACCGG GGGGGGGCCC
901 CGTCACATCC ATCCATCGAC CGATCGATCG CCACAGCCAA CACCAACCGC CGAGGGACAG
961 CGACAGCCGC CAGGAGGAAG GAATAAACTC ACTGCCAGCC AGTGAAGGGG GAGAAAGTGT
1021 CTGCTCCGTC GACCAGTGC CGCACCGCCC GGCAGGGCTG CTCATCTCGT CGACGACCG
1081 GTTCTGTTCC GTTCCGATCC GATCCGATCC TGTCTTGAG TTTCGTCCAG ATCTGGCGC
1141 GTATCTGCGT GTTTGATGAT CCAGGTTCTT CGAACCTAAA TCTGTCGCGT CACAGCTTT
1201 TTCTCTCTCT CCTACGCAGT GGATTAATCG GCATGGCGGC TCTGGCCACG TCGCAGCTCG
1261 TCGCAACCGC CGCCGGCCTG GGGGTCCCCG ACGCGTCCAC GTTCCGCCGC GGGCGCGCGC

```

## 표 1b

벼에서 waxy 유전자의 DNA 서열 및 추론된 아미노산 서열

서열 번호 6 및 확인 번호 7

유전자자리	OSWX	2542 bp	RNA	PLN
정의	O. sativa waxy mRNA.			
접근	X62134	S39554		
키워드	글루코실트렌스퍼라제; 전분 생합성; waxy 유전자.			
소오스	벼			
유기체	Oryza sativa			
	Eukaryota; Plantae; Embryobionta; Magnoliophyta;			
	Liliopsida; Commelinidae; Cyperales; Poaceae.			
참조문헌	1 (영기 1내지 2542)			
저자	Okayaka, R.J.			
제목	직접 제출			
저널	EMBL/GenBank/DDBJ 데이터베이스에 제출되었음(1991.9.12)			
R. J.				
	Okayaka, University of Florida, Dep of Vegetable Crops,			
	1255 Fifield Hall, 514 IFAS, Gainesville, Florida 32611, USA			
기준	완전 자동			
참조문헌	2 (영기 1내지 2542)			
저자	Okayaka, R.J.			
제목	벼의 waxy 유전자로부터 긴 cDNA의 누클레오티드 서열 (Nucleotide sequence of a long cDNA from the rice waxy gene)			
저널	Plant Mol. Biol. 19, 513-516(1992)			
기준	완전 자동			
코멘트	NCBI gi: 20402			
특징	위치/한정물			
소오스	1..2542			
	/유기체='Oryza sativa'			
	/발생 단계='미숙한 씨'			
	/조직 타입='씨'			
CDS	453..2282			
	/유전자='Wx'			
	/표준명='Waxy 유전자'			
	/EC 번호='2.4.1.21'			
	/주='NCBI gi'			
	/코돈 출발=1			
	/작용='전분 생합성'			
	/생성물='전분(세균 글리코겐)신타아제'			

/번역='MSALTTSQLATSTAGFGIADRSAPSSLLRHGFQGLKPRSPAGGD  
 ATSLSVTTSARATPKQQRSVQRGSRRFPSVVVYATGAGMNVVFGAEMAPWSKTGGLG  
 DVLGGLPPAMAANGHRVMVISPRYDQYKDAWTSVVAEIKVADRYERVFFFHCYKRGV  
 DRVFIDHPSFLEKVWGBTGEKIYGPDTGVDYKDQNQMRFSLLCQAALAEAPRILNLNNNP  
 YFKGTYGEDVVFVCNDWHTGPLASYLKNYQPNQIYRNAKVAFCIHNI SYQGRFAFED  
 YPELNLSERFRSSFDIDGYDTPVEGRKINWMKAGILEADRVLTVSPYYAEELISGIA  
 RGCELDNIMRLTGITGIVNGMDVSEWDPSKDKYITAKYDATTIAEAKALNKEALQAEA  
 GLPVDRKIPLIAFIGRLEEQKGPDVMAAAIPELMQEDVQIVLLGTGKKFEKLLKSME  
 EKYPGKVRAVVKFNAPLAHЛИМАГАДВЛАВPSRFEPGCLIQLQGMRYGTPCACASTGG  
 LVDTVIEGKTGFHMGRSLVDCKVVEPSDVKKVAATLKRAIKVVGTPAYEEMVRNCMNQ  
 DLSWKGPAKNWENVLLGLGVAGSAPGIEGDEIAPLAKE  
 3'UTR 2283..2535

플리A 부위 2535  
 염기 집계 610A 665C 693G 574T  
 기원

1 GAATTCACTG TGAAGGAATA GATTCTCTTC AAAACAATTT AATCATTCAAT CTGATCTGCT  
 61 CAAAGCTCTG TGCACTCTCCG GGTGCAACGG CCAGGATATT TATTGTGAG TAAAAAAATG  
 121 TCATATCCCC TAGCCACCCA AGAAACTGCT CCTTAAGTCC TTATAAGCAC ATATGGCATT  
 181 GTAATATATA TGTGAGTT TTAGCGACAA TTTTTTAA AACTTTGGT CCTTTTTATG  
 241 AACGTTTAA GTTTCACTGT CTTTTTTTAA CGAATTTAA ATGTAGCTTC AAATTCTAAT  
 301 CCCCAATCCA AATTGTAATA AACTTCATT CTCCTAATTA ACATCTTAAT TCATTTATTT  
 361 GAAAACCACT TCAAATTCTT TTTAGGCTCA CCAAACCTTA ACAATTCAA TTCAGTGCAG  
 421 AGATCTTCCA CAGCAACAGC TAGACAACCA CCATGCGGC TCTCACCAAG TCCCAGCTCG  
 481 CCACCTCGGC CACCGGCTTC GGCATGCGC ACAGGTGGC GCGGTGCGT CGCTCCGCC  
 541 ACGGGTCCA GGGCTCAAG CCCCGCAGCC CCGCCGGCG CGACGCCAGC TCGCTCAGCG  
 601 TGACGACCAG CGCGCGCG ACGCCAAGC AGCACGGTC GGTGCAGCGT GGCAGCCGGA  
 661 GGTTCCCCCTC CGTCGTCGTG TACGCCACCG GCGCCGGCAT GAACGTCGTG TTCGTCGGCG  
 721 CCGAGATGGC CCCCTGGAGC AAGACCGGGCG GCCTCGGTGA CGTCCTCGGT GGCCTCCCCC  
 781 CTGCCATGGC TGCGAATGGC CACAGGGTCA TGGTGATCTC TCCTCGGTAC GACCACTACA  
 841 AGGACGCTTG GGATACCAGC GTTGTGGCTG AGATCAAGGT TGCAGACAGG TACGAGAGGG  
 901 TGAGGTTTT CCATTGCTAC AAGCGTGGAG TCGACCGTGT GTTCATCGAC CATCCGTAT  
 961 TCCTGGAGAA GGTTGGGGA AAGACCGGTG AGAAGATCTA CGGACCTGAC ACTGGAGTTG  
 1021 ATTACAAAGA CAACCAGATC CGTTTCAGCC TTCTTGCCA GGCAGCACTC GAGGCTCTA  
 1081 GGATCCTAAA CCTCAACAAAC AACCCATACT TCAAAGGAAC TTATGGTGA GATGGTGTG  
 1141 TCGTCTGCAA CGACTGGCAC ACTGGCCAC TGGCGAGCTA CCTGAAGAAC AACTACCAGC  
 1201 CCAATGGCAT CTACAGGAAT GCAAAGGTG CTTCTGCAT CCACACATC TCCTACCAAG  
 1261 GCCGTTTCGC TTTCGAGGAT TACCCCTGAGC TGAACCTCTC CGAGAGGTTG AGGTCACTC  
 1321 TCGATTTCAT CGACGGGTAT GACACGGCGG TGGAGGGCAG GAAGATCAAC TGGATGAAGG  
 1381 CCGGAATCCT GGAAGCCGAC AGGGTGCTCA CCGTGAGGCC GTACTACGCC GAGGAGCTCA  
 1441 TCTCCGGCAT CGCCAGGGGA TCGAGCTCG ACAACATCAT GCGGCTCACC GGCATCACCG  
 1501 GCATCGTCAA CGGCATGGAC GTCAGCGAGT GGGATCCTAG CAAGGACAAG TACATCACCG  
 1561 CCAAGTACGA CGCAACCACG GCAATCGAGG CGAAGGGCCT GAACAAGGAG GCGTTGCAGG  
 1621 CGGAGGCGGG TCTTCCGGTC GACAGGAAAA TCCCACGTGAT CGCGTTCATC GGCAGGCTGG  
 1681 AGGAACAGAA GGGCCCTGAC GTCATGGCCG CCGCCATCCC GGAGCTCATG CAGGAGGACG  
 1741 TCCAGATCCT TCTCTGGGT ACTGGAAGA AGAAGTTCGA GAAGCTGCTC AAGAGCATGG  
 1801 AGGAGAAGTA TCCGGGCAAG GTGAGGGCGG TGGTGAAGTT CAACGGCCG CTTGCTCATC  
 1861 TCATCATGGC CGGAGCCGAC GTGCTCGCCG TCCCCAGGCC CTTCGAGGCC TGTGGACTCA  
 1921 TCCAGCTGGA GGGGATGAGA TACGGAACCG CCTGTGCTTG CGCGTCACCC GGTGGGCTCG  
 1981 TGGACACGGT CATCGAAGGC AAGACTGGTT TCCACATGGG CCGTCTCAGC GTGCACTGCA  
 2041 AGGTGGTGA GCCAAGCGAC GTGAAGAAGG TGGCGGCCAC CCTGAAGCGC GCCATCAAGG  
 2101 TCGTCGGCAC GCGGGCGTAC GAGGAGATGG TCAGGAACCTG CATGAACCAAG GACCTCTCCT  
 2161 GGAAGGGGCC TGCGAAGAAC TGGGAGATG TGCTCCTGGG CCTGGCGTC GCGGGCAGCG  
 2221 CGCCGGGGAT CGAAGGGCAG GAGATCGCCG CGCTCGCCAA GGAGAACGTG GCTGCTCCTT  
 2281 GAAGAGCCTG AGATCTACAT ATGGAGTGTAT TAATTAATAT AGCACTATAT GGATGAGAGA  
 2341 CGAATGAACC AGTGGTTGT TTGTTGTACT GAATTGTAG CTATAGCCAA TTATATAGGC  
 2401 TAATAAGTTT GATGTTGTAC TCTTCTGGGT GTGCTTAAGT ATCTTATCGG ACCCTGAATT  
 2461 TATGTGTGTG GCTTATTGCC AATAATATTA AGTAATAAAG GGTTTATTAT ATTATTATAT  
 2521 ATGTTATATT ATACTAAAAA AA

## 표 2

옥수수에서 가용성 전분 신타아제 IIa 유전자의

DNA 서열 및 추론된 아미노산 서열

[서열 번호 8 및 서열 번호9]

파일 명: MSS2C 서열 : 노르말 2007 BP

코돈 테이블 : UNIV.TCH

서열 영역 : 1-2007

번역 영역 : 1-2007

\*\*\* DNA 번역 \*\*\*

1	GCT	GAG	GCT	GAG	GCC	GGG	GGC	AAG	GAC	GCG	CCG	CCG	GAG	AGG	AGC	GGC	48
1	A	E	A	E	A	G	G	K	D	A	P	P	E	R	S	G	16
49	GAC	GCC	GCC	AGG	TTG	CCC	CGC	GCT	CGG	CGC	AAT	GCG	GTC	TCC	AAA	CGG	96
17	D	A	A	R	L	P	R	A	R	R	N	A	V	S	K	R	32
97	AGG	GAT	CCT	CTT	CAG	CCG	GTC	GGC	CGG	TAC	GGC	TCC	GCG	ACG	GGA	AAC	144
33	R	D	P	L	Q	P	V	G	R	Y	G	S	A	T	G	N	48
145	ACG	GCC	AGG	ACC	GGC	GCC	GCG	TCC	TGC	CAG	AAC	GCC	GCA	TTG	GCG	GAC	192
49	T	A	R	T	G	A	A	S	C	Q	N	A	A	L	A	D	64
193	GTT	GAG	ATC	GTT	GAG	ATC	AAG	TCC	ATC	GTC	GGC	GCG	CCG	CCG	ACG	AGC	240
65	V	E	I	V	E	I	K	S	I	V	A	A	P	P	T	S	80
241	ATA	GTG	AAG	TTC	CCA	GGG	CGC	GGG	CTA	CAG	GAT	GAT	CCT	TCC	CTC	TGG	288
81	I	V	K	F	P	G	R	G	L	Q	D	D	P	S	L	W	96
289	GAC	ATA	GCA	CCG	GAG	ACT	GTC	CTC	CCA	GCC	CCG	AAG	CCA	CTG	CAT	GAA	336
97	D	I	A	P	E	T	V	L	P	A	P	K	P	L	H	E	112
337	TCG	CCT	GGC	GTT	GAC	GGG	GAT	TCA	AAT	GGG	ATT	GCA	CCT	CCT	ACA	GTT	384
113	S	P	A	V	D	G	D	S	N	G	I	A	P	P	T	V	128
385	GAG	CCA	TTA	GTA	CAG	GAG	GCC	ACT	TGG	GAT	TTC	AAG	AAA	TAC	ATC	GGT	432
129	E	P	L	V	Q	E	A	T	W	D	F	K	K	Y	I	G	144
433	TTT	GAC	GAG	CCT	GAC	GAA	GCG	AAG	GAT	GAT	TCC	AGG	GTT	GGT	GCA	GAT	480

145	F	D	E	P	D	E	A	K	D	D	S	R	V	G	A	D	160
481	GAT	GCT	GGT	TCT	TTT	GAA	CAT	TAT	GGG	ACA	ATG	ATT	CTG	GGC	CTT	TGT	528
161	D	A	G	S	F	E	H	Y	G	T	M	I	L	G	L	C	176
529	GGG	GAG	AAT	GTT	ATG	AAC	GTG	ATC	GTG	GTG	GCT	GCT	GAA	TGT	TCT	CCA	576
177	G	E	N	V	M	N	V	I	V	V	A	A	E	C	S	P	192
577	TGG	TGC	AAA	ACA	GGT	GGT	CTT	GGA	GAT	GTT	GTG	GGG	GCT	TTA	CCC	AAG	624
193	W	C	K	T	G	G	L	G	D	V	V	G	A	L	P	K	208
625	GCT	TTA	GCG	AGA	AGA	GGA	CAT	CGT	GTT	ATG	GTT	GTG	GTA	CCA	AGG	TAT	672
209	A	L	A	R	R	G	H	R	V	M	V	V	V	P	R	Y	224
673	GGG	GAC	TAT	GTG	GAA	GCC	TTT	GAT	ATG	GGA	ATC	CGG	AAA	TAC	TAC	AAA	720
225	G	D	Y	V	E	A	F	D	M	G	I	R	K	Y	Y	K	240
721	GCT	GCA	GGA	CAG	GAC	CTA	GAA	GTG	AAC	TAT	TTC	CAT	GCA	TTT	ATT	GAT	768
241	A	A	G	Q	D	L	E	V	N	Y	F	H	A	F	I	D	256
769	GGA	GTC	GAC	TTT	GTG	TTC	ATT	GAT	GCC	TCT	TTC	CGG	CAC	CGT	CAA	GAT	816
257	G	V	D	F	V	F	I	D	A	S	F	R	H	R	Q	D	272
817	GAC	ATA	TAT	GGG	GGA	AGT	AGG	CAG	GAA	ATC	ATG	AAG	CGC	ATG	ATT	TTG	864
273	D	I	Y	G	G	S	R	Q	E	I	M	K	R	M	I	L	288
865	TTT	TGC	AAG	GTT	GCT	GTT	GAG	GTT	CCT	TGG	CAC	GTT	CCA	TGC	GGT	GGT	912
289	F	C	K	V	A	V	E	V	P	W	H	V	P	C	G	G	304
913	GTG	TGC	TAC	GGA	GAT	GGA	AAT	TTG	GTG	TTC	ATT	GCC	ATG	AAT	TGG	CAC	960
305	V	C	Y	G	D	G	N	L	V	F	I	A	M	N	W	H	320
961	ACT	GCA	CTC	CTG	CCT	GTT	TAT	CTG	AAG	GCA	TAT	TAC	AGA	GAC	CAT	GGG	1008
321	T	A	L	L	P	V	Y	L	K	A	Y	Y	R	D	H	G	336
1009	TTA	ATG	CAG	TAC	ACT	CGC	TCC	GTC	CTC	GTC	ATA	CAT	AAC	ATC	GGC	CAC	1056
337	L	M	Q	Y	T	R	S	V	L	V	I	H	N	I	G	H	352
1057	CAG	GGC	CGT	GGT	CCT	GTA	CAT	GAA	TTC	CCG	TAC	ATG	GAC	TTG	CTG	AAC	1104
353	Q	G	R	G	P	V	H	E	F	P	Y	M	D	L	L	N	368
1105	ACT	AAC	CTT	CAA	CAT	TTC	GAG	CTG	TAC	GAT	CCC	GTC	GGT	GGC	GAG	CAC	1152
369	T	N	L	Q	H	F	E	L	Y	D	P	V	G	G	E	H	384
1153	GCC	AAC	ATC	TTT	GCC	GCG	TGT	GTT	CTG	AAG	ATG	GCA	CAC	CGG	GTG	GTG	1200
385	A	N	I	F	A	A	C	V	L	K	M	A	D	R	V	V	400
1201	ACT	GTC	AGC	CGC	GGC	TAC	CTG	TGG	GAG	CTG	AAG	ACA	GTG	GAA	GGC	GGC	1248
401	T	V	S	R	G	Y	L	W	E	L	K	T	V	E	G	G	416
1249	TGG	GGC	CTC	CAC	GAC	ATC	ATC	CGT	TCT	AAC	GAC	TGG	AAG	ATC	AAT	GGC	1296
417	W	G	L	H	D	I	I	R	S	N	D	W	K	I	N	G	432
1297	ATT	CGT	GAA	CGC	ATC	GAC	CAC	CAG	GAG	TGG	AAC	CCC	AAG	GTG	GAC	GTG	1344
433	I	R	E	R	I	D	H	Q	E	W	N	P	K	V	D	V	448
1345	CAC	CTG	CGG	TCG	GAC	GGC	TAC	ACC	AAC	TAC	TCC	CTC	GAG	ACA	CTC	GAC	1392
449	H	L	R	S	D	G	Y	T	N	Y	S	L	E	T	L	D	464
1393	GCT	GGA	AAG	CGG	CAG	TGC	AAG	GGC	GCC	CTG	CAG	CGG	GAC	GTG	GGC	CTG	1440
465	A	G	K	R	Q	C	K	A	A	L	Q	R	D	V	G	L	480
1441	GAA	GTG	CGC	GAC	GAC	CTG	CCG	CTG	CTC	GGC	TTC	ATC	GGG	CGT	CTG	GAT	1488
481	E	V	R	D	D	V	P	L	L	G	F	I	G	R	L	D	496
1489	GGA	CAG	AAG	GGC	GTG	GAC	ATC	ATC	GGG	GAC	GCG	ATG	CCG	TGG	ATC	GCG	1536
497	G	Q	K	G	V	D	I	I	G	D	A	M	P	W	I	A	512
1537	GGG	CAG	GAC	GTG	CGG	CTG	GTG	ATG	CTG	GGC	ACC	GGC	CCA	CCT	GAC	CTG	1584
513	G	Q	D	V	Q	L	V	M	L	G	T	G	P	P	D	L	528
1585	GAA	CGA	ATG	CTG	CAG	CAC	TTG	GAG	CGG	GAG	CAT	CCC	AAC	AAG	GTG	CGC	1632
529	E	R	M	L	Q	H	L	E	R	E	H	P	N	K	V	R	544
1633	GGG	TGG	GTC	GGG	TTC	TCG	GTC	CTA	ATG	GTG	CAT	CGC	ATC	ACG	CCG	GGC	1680
545	G	W	V	G	F	S	V	L	M	V	H	R	I	T	P	G	560
1681	GCC	AGC	GTG	CTG	GTG	ATG	CCC	TCC	CGC	GGC	GGG	CTG	AAC	CAG			1728
561	A	S	V	L	V	M	P	S	R	F	A	G	G	L	N	Q	576
1729	CTC	TAC	GCG	ATG	GCA	TAC	GGC	ACC	GTC	CCT	GTG	GTG	CAC	GCC	GTG	GGC	1776
577	L	Y	A	M	A	Y	G	T	V	P	V	V	H	A	V	G	592
1777	GGG	CTC	AGG	GAC	ACC	GTG	GGC	CCG	TTC	GAC	CCG	TTC	GGC	GAC	GCC	GGG	1824
593	G	L	R	D	T	V	A	P	F	D	P	F	G	D	A	G	608
1825	CTC	GGG	TGG	ACT	TTT	GAC	CGC	GCC	GAG	GCC	AAC	AAG	CTG	ATC	GAG	GTG	1872
609	L	G	W	T	F	D	R	A	E	A	N	K	L	I	E	V	624
1873	CTC	AGC	CAC	TGC	CTC	GAC	ACG	TAC	CGA	AAC	TAC	GAG	GAG	AGC	TGG	AAG	1920
625	L	S	H	C	L	D	T	Y	R	N	Y	E	E	S	W	K	640
1921	AGT	CTC	CAG	GGC	CGC	GGC	ATG	TCG	CAG	AAC	CTC	AGC	TGG	GAC	CAC	GCG	1968
641	S	L	Q	A	R	G	M	S	Q	N	L	S	W	D	H	A	656
1969	GCT	GAG	CTC	TAC	GAG	GAC	GTC	CTT	GTC	AAG	TAC	CAG	CAG	TGG			2007
657	A	E	L	Y	E	D	V	L	V	K	Y	Q	W				669

옥수수에서 가용성 전분 신타아제 IIb 유전자의

DNA 서열 및 추론된 아미노산 서열

[서열 번호 10 및 서열 번호 11]

파일 명: MSS3CFULL. DNA 서열 : 노르말 2097 BP

코돈 테이블 : UNIV.TCH

서열 영역 : 1-2097

번역 영역 : 1-2097

\*\*\* DNA 번역 \*\*\*

1	ATG	CCG	GGG	GCA	ATC	TCT	TCC	TCG	TCG	TCG	GCT	TTT	CTC	CTC	CCC	GTC	48
1	M	P	G	A	I	S	S	S	S	S	A	F	L	L	P	V	16
49	GCG	TCC	TCC	TCG	CCG	CGG	CGC	AGG	CGG	GGC	AGT	GTG	GGT	GCT	GCT	CTG	96
17	A	S	S	S	P	R	R	R	R	G	S	V	G	A	A	L	32
97	CGC	TCG	TAC	GGC	TAC	AGC	GGC	GCG	GAG	CTG	CGG	TTG	CAT	TGG	GCG	CGG	144
33	R	S	Y	G	Y	S	G	A	E	L	R	L	H	W	A	R	48
145	CGG	GGC	CCG	CCT	CAG	GAT	GGA	GCG	GCG	TCG	GTA	CGC	GCC	GCA	GCG	GCA	192
49	R	G	P	P	Q	D	G	A	A	S	V	R	A	A	A	A	64
193	CCG	GCC	GGG	GGC	GAA	AGC	GAG	GAG	GCA	GCG	AAG	AGC	TCC	TCC	TCG	TCC	240
65	P	A	G	G	E	S	E	E	A	A	K	S	S	S	S	S	80
241	CAG	GCG	GGC	GCT	GTT	CAG	GGC	AGC	ACG	GCC	AAG	GCT	GTG	GAT	TCT	GCT	288

81	Q	A	G	A	V	Q	G	S	T	A	K	A	V	D	S	A	96
289	TCA	CCT	CCC	AAT	CCT	TTG	ACA	TCT	GCT	CCG	AAG	CAA	AGT	CAG	AGC	GCT	336
97	S	P	P	N	P	L	T	S	A	P	K	Q	S	Q	S	A	112
337	GCA	ATG	CAA	AAC	GGA	ACG	AGT	GGG	GGC	AGC	AGC	GCG	AGC	ACC	GCC	GCG	384
113	A	M	Q	N	G	T	S	G	G	S	S	A	S	T	A	A	128
385	CCG	GTG	TCC	GGA	CCC	AAA	GCT	GAT	CAT	CCA	TCA	GCT	CCT	GTC	ACC	AAG	432
129	P	V	S	G	P	K	A	D	H	P	S	A	P	V	T	K	144
433	AGA	GAA	ATC	GAT	GCC	AGT	GGG	GTG	AAG	CCA	GAG	CCC	GCA	GGT	GAT	GAT	480
145	R	E	I	D	A	S	A	V	K	P	E	P	A	G	D	D	160
481	GCT	AGA	CCG	GTG	GAA	AGC	ATA	GGC	ATC	GCT	GAA	CCG	GTG	GAT	GCT	AAG	528
161	A	R	P	V	E	S	I	G	I	A	E	P	V	D	A	K	176
529	GCT	GAT	GCA	GCT	CCG	GCT	ACA	GAT	GGG	GGC	GCG	AGT	GCT	CCT	TAT	GAC	576
177	A	D	A	A	P	A	T	D	A	A	A	S	A	P	Y	D	192
577	AGG	GAG	GAT	AAT	GAA	CCT	GGC	CCT	TTG	GCT	GGG	CCT	AAT	GTG	ATG	AAC	624
193	R	E	D	N	E	P	G	P	L	A	G	P	N	V	M	N	208
625	GTC	GTC	GTG	GTG	GCT	TCT	GAA	TGT	GCT	CCT	TTC	TGC	AAG	ACA	GGT	GGC	672
209	V	V	V	V	A	S	E	C	A	P	F	C	K	T	G	G	224
673	CTT	GGA	GAT	GTC	GTG	GGT	GCT	TTG	CCT	AAG	GCT	CTG	GCG	AGG	AGA	GGA	720
225	L	G	D	V	V	G	A	L	P	K	A	L	A	R	R	G	240
721	CAC	CGT	GTT	ATG	GTC	GTG	ATA	CCA	AGA	TAT	GGA	GAG	TAT	GCC	GAA	GCC	768
241	H	R	V	M	V	V	I	P	R	Y	G	E	Y	A	E	A	256
769	CGG	GAT	TTA	GGT	GTA	AGG	AGA	CGT	TAC	AAG	GTA	GCT	GGA	CAG	GAT	TCA	816
257	R	D	L	G	V	R	R	R	Y	K	V	A	G	Q	D	S	272
817	GAA	GTT	ACT	TAT	TTT	CAC	TCT	TAC	ATT	GAT	GGG	GTA	GTT	GAT	TTT	GTA	864
273	E	V	T	Y	F	H	S	Y	I	D	G	V	D	F	V	F	288
865	GTA	GAA	GCC	CCT	CCC	TTC	CGG	CAC	CGG	CAC	AAT	AAT	ATT	ATT	TAT	GGG	912
289	V	E	A	P	P	F	R	H	R	H	N	N	N	I	Y	G	304
913	GAA	AGA	TTG	GAT	ATT	TTG	AAG	CGC	ATG	ATT	TTG	TTC	TGC	AAG	GCC	GCT	960
305	E	R	L	D	I	L	K	R	M	I	L	F	C	K	A	A	320
961	GTT	GAG	GTT	CCA	TGG	TAT	GCT	CCA	TGT	GGC	GGT	ACT	GTC	TAT	GGT	GAT	1008
321	V	E	V	P	W	Y	A	P	C	G	G	T	V	Y	G	D	336
1009	GGC	AAC	TTA	GTT	TTC	ATT	GCT	AAT	GAT	TGG	CAT	ACC	GCA	CTT	CTG	CCT	1056
337	G	N	L	V	F	I	A	N	D	W	H	T	A	L	L	P	352
1057	GTC	TAT	CTA	AAG	GCC	TAT	TAC	CGG	GAC	AAT	GGT	TTG	ATG	CAG	TAT	GCT	1104
353	V	Y	L	K	A	Y	Y	R	D	N	G	L	M	Q	Y	A	368
1105	CGC	TCT	GTG	CTT	GTG	ATA	CAC	AAC	ATT	GCT	CAT	CAG	GGT	CGT	GGC	CCT	1152
369	R	S	V	L	V	I	H	N	I	A	H	Q	G	R	G	P	384
1153	GTA	GAC	GAC	TTC	GTC	AAT	TTT	GAC	TTG	CCT	GAA	CAC	TAC	ATC	GAC	CAC	1200
385	V	D	D	F	V	N	F	D	L	P	E	H	Y	I	D	H	400
1201	TTC	AAA	CTG	TAT	GAC	AAC	ATT	GGT	GGG	GAT	CAC	AGC	AAC	GTT	TTT	GCT	1248
401	F	K	L	Y	D	N	I	G	G	D	H	S	N	V	F	A	416
1249	GCG	GGG	CTG	AAG	ACG	GCA	GAC	CGG	GTG	GTG	ACC	GTT	AGC	AAT	GGC	TAC	1296
417	A	G	L	K	T	A	D	R	V	V	T	V	S	N	G	Y	432
1297	ATG	TGG	GAG	CTG	AAG	ACT	TCG	GAA	GGC	GGG	TGG	GGC	CTC	CAC	GAC	ATC	1344
433	M	W	E	L	K	T	S	E	G	G	W	G	L	H	D	I	448

1345	ATA AAC CAG AAC GAC TGG AAG CTG CAG GCC ATC GTG AAC GGC ATC GAC	1392
449	I N Q N D W K L Q G I V N G I D	464
1393	ATG AGC GAG TGG AAC CCC GCT GTG GAC GTG CAC CTC CAC TCC GAC GAC	1440
465	M S E W N P A V D V H L H S D D	480
1441	TAC ACC AAC TAC ACG TTC GAG ACG CTG GAC ACC GGC AAG CGG CAG TGC	1488
481	Y T N Y T F E T L D T G K R Q C	496
1489	AAG GCC GCC CTG CAG CGG CAG CTG GGC CTG CAG GTC CGC GAC GAC GTG	1536
497	K A A L Q R Q L G L Q V R D D V	512
1537	CCA CTG ATC GGG TTC ATC GGG CGG CTG GAC CAC CAG AAG GGC GTG GAC	1584
513	P L I G F I G R L D H Q K G V D	528
1585	ATC ATC GCC GCG ATC CAC TGG ATC GCG GGG CAG GAC GTG CAG CTC	632
529	I I A D A I H W I A G Q D V Q L	544
1633	GTG ATG CTG GGC ACC GGG CGG GCC GAC CTG GAG GAC ATG CTG CGG CGG	1680
545	V M L G T G R A D L E D M L R R	560
1681	TTC GAG TCG GAG CAC AGC GAC AAG GTG CGC GCG TGG GTG GGG TTC TCG	1728
561	F E S E H S D K V R A W V G F S	576
1729	GTG CCC CTG GCG CAC CGC ATC ACG GCG GCC GCG GAC ATC CTG CTG ATG	1776
577	V P L A H R I T A G A D I L L M	592
1777	CCG TCG CGG TTC GAG CCG TGC GGG CTG AAC CAG CTC TAC GCC ATG GCG	1824
593	P S R F E P C G L N Q L Y A M A	608
1825	TAC GGG ACC GTG CCC GTG GTG CAC GCC GTG GGG GGG CTC CGG GAC ACG	1872
609	Y G T V P V V H A V G G L R D T	624
1873	GTG GCG CCG TTC GAC CCG TTC AAC GAC ACC GGG CTC CGG TGG ACG TTC	1920
625	V A P F D P F N D T G L G W T F	640
1921	GAC CGC GCG GAG GCG AAC CGG ATG ATC GAC GCG CTC TCG CAC TGC CTC	1968
641	D R A E A N R M I D A L S H C L	656
1969	ACC ACG TAC CGG AAC TAC AAG GAG AGC TGG CGC GCC TGC AGG GCG CGC	2016
657	T T Y R N Y K E S W R A C R A R	672
2017	GGC ATG GCC GAG GAC CTC AGC TGG GAC CAC GCC GCC GTG CTG TAT GAG	2064
673	G M A E D L S W D H A A V L Y E	688
2065	GAC GTG CTC GTC AAG GCG AAG TAC CAG TGG TGA	2097
689	D V L V K A K Y Q W *	699

표 4

옥수수에서 가용성 전분 신타아제 I 유전자의  
DNA 서열 및 추론된 아미노산 서열

[서열 번호 12 및 서열 번호 13]

파일 명: MSS1FULL. DNA 서열 : 노르말 1752 BP

코돈 테이블 : UNIV.TCH

서열 영역 : 1-1752

번역 영역 : 1-1752

\*\*\* DNA 번역 \*\*\*

TGC GTC GCG GAG CTG AGC AGG GAG GGG CCC GCG CCG CGC CCG CTG CCA Cys Val Ala Glu Leu Ser Arg Glu Gly Pro Ala Pro Arg Pro Leu Pro 700 705 710 715	48
CCC GCG CTG CTG GCG CCC CCG CTC GTG CCC GGC TTC CTC GCG CCG CCG Pro Ala Leu Leu Ala Pro Pro Leu Val Pro Gly Phe Leu Ala Pro Pro 720 725 730	96
GCC GAG CCC ACG GGT GAG CCG GCA TCG ACG CCG CCC GTG CCC GAC Ala Glu Pro Thr Gly Glu Pro Ala Ser Thr Pro Pro Pro Val Pro Asp 735 740 745	144
GCC GGC CTG GGG GAC CTC GGT CTC GAA CCT GAA GGG ATT GCT GAA GGT Ala Gly Leu Gly Asp Leu Gly Leu Glu Pro Glu Gly Ile Ala Glu Gly 750 755 760	192
TCC ATC GAT AAC ACA GTA GTT GTG GCA AGT GAG CAA GAT TCT GAG ATT Ser Ile Asp Asn Thr Val Val Val Ala Ser Glu Gln Asp Ser Glu Ile 765 770 775	240
GTG GTT GGA AAG GAG CAA GCT CGA GCT AAA GTC ACA CAA AGC ATT GTC Val Val Gly Lys Glu Gln Ala Arg Ala Lys Val Thr Gln Ser Ile Val 780 785 790 795	288
TTT GTA ACC GGC GAA GCT TCT CCT TAT GCA AAG TCT GGG GGT CTA GGA Phe Val Thr Gly Glu Ala Ser Pro Tyr Ala Lys Ser Gly Gly Leu Gly 800 805 810	336
GAT GTT TGT GGT TCA TTG CCA GTT GCT CTT GCT GCT CGT GGT CAC CGT Asp Val Cys Gly Ser Leu Pro Val Ala Leu Ala Ala Arg Gly His Arg 815 820 825	384
GTG ATG GTT GTA ATG CCC AGA TAT TTA AAT GGT ACC TCC GAT AAG AAT Val Met Val Val Met Pro Arg Tyr Leu Asn Gly Thr Ser Asp Lys Asn 830 835 840	432
TAT GCA AAT GCA TTT TAC ACA GAA AAA CAC ATT CGG ATT CCA TGC TTT Tyr Ala Asn Ala Phe Tyr Thr Glu Lys His Ile Arg Ile Pro Cys Phe 845 850 855	480
GGC GGT GAA CAT GAA GTT ACC TTC TTC CAT GAG TAT AGA GAT TCA GTT Gly Gly Glu His Glu Val Thr Phe Phe His Glu Tyr Arg Asp Ser Val 860 865 870 875	528
GAC TGG GTG TTT GTT GAT CAT CCC TCA TAT CAC AGA CCT GGA AAT TTA Asp Trp Val Phe Val Asp His Pro Ser Tyr His Arg Pro Gly Asn Leu 880 885 890	576
TAT GGA GAT AAG TTT GGT GCT TTT GGT GAT AAT CAG TTC AGA TAC ACA Tyr Gly Asp Lys Phe Gly Ala Phe Gly Asp Asn Gln Phe Arg Tyr Thr 895 900 905	624
CTC CTT TGC TAT GCT GCA TGT GAG GCT CCT TTG ATC CTT GAA TTG GGA Leu Leu Cys Tyr Ala Ala Cys Glu Ala Pro Leu Ile Leu Glu Leu Gly 910 915 920	672
GGA TAT ATT TAT GGA CAG AAT TGC ATG TTT GTT GTC AAT GAT TGG CAT Gly Tyr Ile Tyr Gly Gln Asn Cys Met Phe Val Val Asn Asp Trp His 925 930 935	720
GCC AGT CTA GTG CCA GTC CTT CCT GCT GCA AAA TAT AGA CCA TAT GGT Ala Ser Leu Val Pro Val Leu Leu Ala Ala Lys Tyr Arg Pro Tyr Gly 940 945 950 955	768
GTT TAT AAA GAC TCC CGC AGC ATT CTT GTA ATA CAT AAT TTA GCA CAT Val Tyr Lys Asp Ser Arg Ser Ile Leu Val Ile His Asn Leu Ala His 960 965 970	816

CAG GGT GTA GAG CCT GCA AGC ACA TAT CCT GAC CTT GGG TTG CCA CCT Gln Gly Val Glu Pro Ala Ser Thr Tyr Pro Asp Leu Gly Leu Pro Pro 975 980 985	864
GAA TGG TAT GGA GCT CTG GAG TGG GTC TTC CCT GAA TGG GCG AGG AGG Glu Trp Tyr Gly Ala Leu Glu Trp Val Phe Pro Glu Trp Ala Arg Arg 990 995 1000	912
CAT GCC CTT GAC AAG GGT GCA GTT AAT TTT TTG AAA GGT GCA GTT His Ala Leu Asp Lys Gly Glu Ala Val Asn Phe Leu Lys Gly Ala Val 1005 1010 1015	960
GTG ACA GCA GAT CGA ATC GTG ACT GTC AGT AAG GGT TAT TCG TGG GAG Val Thr Ala Asp Arg Ile Val Thr Val Ser Lys Gly Tyr Ser Trp Glu 1020 1025 1030 1035	1008
GTC ACA ACT GCT GAA GGT GGA CAG GCC CTC AAT GAG CTC TTA AGC TCC Val Thr Thr Ala Glu Gly Gln Gly Leu Asn Glu Leu Leu Ser Ser 1040 1045 1050	1056
AGA AAG AGT GTA TTA AAC GGA ATT GTA AAT GGA ATT GAC ATT AAT GAT Arg Lys Ser Val Leu Asn Gly Ile Val Asn Gly Ile Asp Ile Asn Asp 1055 1060 1065	1104
TGG AAC CCT GCC ACA GAC AAA TGT ATC CCC TGT CAT TAT TCT GTT GAT Trp Asn Pro Ala Thr Asp Lys Cys Ile Pro Cys His Tyr Ser Val Asp 1070 1075 1080	1152
GAC CTC TCT GGA AAG GCC AAA TGT AAA GGT GCA TTG CAG AAG GAG CTG Asp Leu Ser Gly Lys Ala Lys Cys Lys Gly Ala Leu Gln Lys Glu Leu 1085 1090 1095	1200
GGT TTA CCT ATA AGG CCT GAT GTT CCT CTG ATT GGC TTT ATT GGA AGG Gly Leu Pro Ile Arg Pro Asp Val Pro Leu Ile Gly Phe Ile Gly Arg 1100 1105 1110 1115	1248
TTG GAT TAT CAG AAA GGC ATT GAT CTC ATT CAA CTT ATC ATA CCA GAT Leu Asp Tyr Gln Lys Gly Ile Asp Leu Ile Gln Leu Ile Ile Pro Asp 1120 1125 1130	1296
CTC ATG CGG GAA GAT GTT CAA TTT GTC ATG CTT GGA TCT GGT GAC CCA Leu Met Arg Glu Asp Val Gln Phe Val Met Leu Gly Ser Gly Asp Pro 1135 1140 1145	1344
GAG CTT GAA GAT TGG ATG AGA TCT ACA GAG TCG ATC TTC AAG GAT AAA Glu Leu Glu Asp Trp Met Arg Ser Thr Glu Ser Ile Phe Lys Asp Lys 1150 1155 1160	1392
TTT CGT GGA TGG GTT GGA TTT AGT GTT CCA GTT TCC CAC CGA ATA ACT Phe Arg Gly Trp Val Gly Phe Ser Val Pro Val Ser His Arg Ile Thr 1165 1170 1175	1440
GCC GGC TGC GAT ATA TTG TTA ATG CCA TCC AGA TTC GAA CCT TGT GGT Ala Gly Cys Asp Ile Leu Leu Met Pro Ser Arg Phe Glu Pro Cys Gly 1180 1185 1190 1195	1488
CTC AAT CAG CTA TAT GCT ATG CAG TAT GGC ACA GTT CCT GTT GTC CAT Leu Asn Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val Pro Val Val His 1200 1205 1210	1536
GCA ACT GGG GGC CTT AGA GAT ACC GTG GAG AAC TTC AAC CCT TTC GGT Ala Thr Gly Leu Arg Asp Thr Val Glu Asn Phe Asn Pro Phe Gly 1215 1220 1225	1584
GAG AAT GGA GAG CAG GGT ACA GGG TGG GCA TTC GCA CCC CTA ACC ACA Glu Asn Gly Glu Gln Gly Thr Gly Trp Ala Phe Ala Pro Leu Thr Thr 1230 1235 1240	1632
GAA AAC ATG TTT GTG GAC ATT GCG AAC TGC AAT ATC TAC ATA CAG GGA Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile Tyr Ile Gln Gly 1245 1250 1255	1680
ACA CAA GTC CTC CTG GGA AGG GCT AAT GAA GCG AGG CAT GTC AAA AGA Thr Gln Val Leu Leu Gly Arg Ala Asn Glu Ala Arg His Val Lys Arg 1260 1265 1270 1275	1728
CTT CAC GTG GGA CCA TGC CGC TGA Leu His Val Gly Pro Cys Arg * 1280	1752
GAA AAC ATG TTT GTG GAC ATT GCG AAC TGC AAT ATC TAC ATA CAG GGA Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile Tyr Ile Gln Gly 1245 1250 1255	1680
ACA CAA GTC CTC CTG GGA AGG GCT AAT GAA GCG AGG CAT GTC AAA AGA Thr Gln Val Leu Leu Gly Arg Ala Asn Glu Ala Arg His Val Lys Arg 1260 1265 1270 1275	1728
CTT CAC GTG GGA CCA TGC CGC TGA Leu His Val Gly Pro Cys Arg * 1280	1752

(2) 서열 확인 번호 130에 대한 정보:

(i) 서열 특징:

(A) 길이: 584 아미노 산

(B) 타입: 아미노 산

(C) 토플로지: 선형

(ii) 문자 유형: 단백질

(xi) 서열 설명: 서열 확인 번호 13:

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

Ala Ser Leu Val Pro Val Leu Leu Ala Ala Lys Tyr Arg Pro Tyr Gly  
 245 250 255  
 Val Tyr Lys Asp Ser Arg Ser Ile Leu Val Ile His Asn Leu Ala His  
 260 265 270  
 Gln Gly Val Glu Pro Ala Ser Thr Tyr Pro Asp Leu Gly Leu Pro Pro  
 275 280 285  
 Glu Trp Tyr Gly Ala Leu Glu Trp Val Phe Pro Glu Trp Ala Arg Arg  
 290 295 300  
 His Ala Leu Asp Lys Gly Glu Ala Val Asn Phe Leu Lys Gly Ala Val  
 305 310 315 320  
 Val Thr Ala Asp Arg Ile Val Thr Val Ser Lys Gly Tyr Ser Trp Glu  
 325 330 335  
 Val Thr Thr Ala Glu Gly Gly Gln Gly Leu Asn Glu Leu Leu Ser Ser  
 340 345 350  
 Arg Lys Ser Val Leu Asn Gly Ile Val Asn Gly Ile Asp Ile Asn Asp  
 355 360 365  
 Trp Asn Pro Ala Thr Asp Lys Cys Ile Pro Cys His Tyr Ser Val Asp  
 370 375 380  
 Asp Leu Ser Gly Lys Ala Lys Cys Lys Gly Ala Leu Gln Lys Glu Leu  
 385 390 395 400  
 Gly Leu Pro Ile Arg Pro Asp Val Pro Leu Ile Gly Phe Ile Gly Arg  
 405 410 415  
 Leu Asp Tyr Gln Gly Ile Asp Leu Ile Gln Leu Ile Ile Pro Asp  
 420 425 430  
 Leu Met Arg Glu Asp Val Gln Phe Val Met Leu Gly Ser Gly Asp Pro  
 435 440 445  
 Glu Leu Glu Asp Trp Met Arg Ser Thr Glu Ser Ile Phe Lys Asp Lys  
 450 455 460  
 Phe Arg Gly Trp Val Gly Phe Ser Val Pro Val Ser His Arg Ile Thr  
 465 470 475 480  
 Ala Gly Cys Asp Ile Leu Leu Met Pro Ser Arg Phe Glu Pro Cys Gly  
 485 490 495  
 Leu Asn Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val Pro Val Val His  
 500 505 510  
 Ala Thr Gly Gly Leu Arg Asp Thr Val Glu Asn Phe Asn Pro Phe Gly  
 515 520 525  
 Glu Asn Gly Glu Gln Gly Thr Gly Trp Ala Phe Ala Pro Leu Thr Thr  
 530 535 540  
 Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile Tyr Ile Gln Gly  
 545 550 555 560  
 Thr Gln Val Leu Leu Gly Arg Ala Asn Glu Ala Arg His Val Lys Arg  
 565 570 575  
 Leu His Val Gly Pro Cys Arg \*  
 580

## 표 5

옥수수 분지 효소 II 유전자 및 트렌싯 펩티드의

mRNA서열 및 추론된 아미노산 서열

[서열 번호 14 및 서열 번호 15]

유전자자리	MZEGLUCTRN	2725 bp	ss-mRNA	PLN
정의	옥수수 전분 분지 효소 II	mRNA	완전한 cds.	
접근	L08065			
키워드	1,4-알파-글루칸 분지 효소; 아밀로-글리코실라아제; 글루카노 트렌스퍼라아제; 전분 분지 효소 II.			
소오스	Zea mays cDNA에서 mRNA			
유기체	Zea mays			
	Eukaryota; Plantae; Embryobionta; Magnoliophyta; Liliopsida; Commelinidae; Cyperales; Poaceae.			

참조문헌 1 (영기 1내지 2542)

저자 Fisher, D. K., Boyer, C. D., 및 Hannah, L. C.

제목 옥수수 배젖으로부터의 전분 분자 효소(Starch branching enzyme II from maize endosperm)

저널 Plant Physiol. 102, 1045-1046(1993)

기준 완전 자동

코멘트 NCBI gi: 168482

특징 위치/한정률

소오스 1..2725

/품종='W64Ax182E'

/발생 단계='꽃가루 발생후 29일'

/조직 타입='배젖'

유기체='Zea mays'

시그날 펩티드 91..264

CDS 91..2490

/EC 번호='2. 4. 1. 18'

/주='NCBI gi: 168483'

/코돈 출발=1

/생성물='전분 분자 효소 II'

번역='MAFRVSGAVLGGAVRAPRLLTGGGEGSLVFRHTGLFLTRGARVGC  
SGTHGAMAAAARKAVMVPEGENDGLASRADSAQFQSDELEVVDISEETTCGAGVAD  
AQALNRVRVPPPSDGQKIFQIDPMLQGYKYHLEYRYSLYRRIRSDIDEHEGGLEAFS  
RSYEKFGFNASAEGITYREWAPGAFSAALVGDVNNWDPNADRMSKNEFGVWEIFLPNN  
ADGTSPIPHGSRVKVRMDTPSGIKDSIPAWIKYSVQAPCEIPYDGIVYDPPEEVKYVF  
RHAQPKRPKSLRIYETHVGMSSPEPKINTYVNFRDEVLPRIKKLGYNQIMAIQEHS  
YYGSFGYHVTNFFAPSSRGTPEDLKSOLIDRAHELGLLVLMMDVVSASSNTLDGLNG  
FDGTDTHYFHSGPRGHWMWDSRLFNYGNWEVRLFLLSNARWWLEYKFDGFRFDGVT  
SMMYTHHGLQVTFTGNFNEYFGFATDVDAVYMLVNDLIHGLYPEAUTIGEDVSGMP  
TFALPVHDGGVGFYRHMMAVADKWDLLKQSDETWMGDIVHTLTNRRWLEKCVTYA  
ESHDQALVGDKTIAFWLMDKDMYDFMALDRPSTPTIDRGIALHKMIRLITMGLGEGY  
LNFMGNFEGHPEWIDPFRGPQRLPSGKFIPGNNNNSYDKCRRRFDLGDADYLRYHGMQE  
FDQAMQHLEQKYEFMTSDHQYISRKHEEDKVIVFEKGDLVVFVNFHCNNSYFDYRIGC  
RKPGVYKVVLDSDAGLFGGFSRIHAAEHFTADCSDHNRPYSFSVYTPSRTCVVYAPV  
mat 펩티드 265..2487

/코돈 출발=1

/생성물='전분 분자 효소 II'

영기 집계 727A 534C 715G 749T

```

1  GGCCCAGAGC AGACCCGGAT TTCGCTCTTG CGGTGGCTGG GGTTTTAGCA TTGGCTGATC
61 AGTCGATCC GATCCGGCTG CGAAGCGAG ATGGCGTCC GGGTTCTGG GGCGGTCTC
121 GGTGGGGCCG TAAGGGCTCC CCGACTCACC GGCGCGGGGG AGGGTAGTCT AGTCTTCCGG
181 CACACCGGGC TCTCTTAACT CGGGGTGCT CGAGTTGGAT GTTCCGGGAC GCACGGGGCC
241 ATGCCCGCGG CGGGCGGGC CAGGAAGGGC GTCATGGTTC CTGAGGGCGA GAATGATGCC
301 CTCGCATCAA GGGCTGACTC GGCTCAATTG CAGTCGGATG AACTGGAGGT ACCAGACATT
361 CTGAAGAGA CAACGTGGG TGCTGGTGTG GCTGATGTC AAGCCTTGAA CAGAGTTCGA
421 GTGGTCCCCC CACCAAGCGA TGGACAAAAA ATATCCAGA TTGACCCCAT GTTGCAGGCA
481 TATAAGTACC ATCTTGAGTA TCGGTACAGC CTCTATAGAA GAATCCGTT AGACATTGAT
541 AACATGAAGA GAGGCTTGGG AGCCTTCTCC CGTAGTTGG AGAAGTTGG ATTTAAATGCC
601 AGCCGGAAAG GTATCACATA TCGAGAATG GCTCTGGAG CATTTCCTGC AGCATTGGTG
661 GGTGACGTCA ACAACTGGGA TCCAAATGCA GATCGTATGA GCAAAATATGA GTTTGGTGT
721 TGGGAAATT TCTTGCTTAA CAATCGCAGAT GTTACATCAC CTATTCTCA TGGATCTCGT
781 GTAAAGGTGA GAATGGATAC TCCCATCAGGG ATAAAGGGT CAATCCAGC CTGGATCAAG
841 TACTCAGTGC AGGCCCCAGG AGAAATACCA TATGATGGGA TTTATTATGA TCCTCCTGAA
901 GAGGTAAAGT ATGTGTTCAAG GCATCGCAGA CCTAAACGAC CAAATCATT GCGGATATAT
961 GAAACACATG TCGGAATGAG TAGCCGGAA CGGAAGATAA ACACATATGT AAACCTTAA
1021 GATGAAGTCC TCCCAAGAAAT AAAAAGACTT GGATACAATG CAGTCAAAT AATGGCAATC
1081 CAAGAGCACT CATAATTATGG AAGCTTTGGG TACCATGTA CTAATTTTTT TGCGCCAAGT
1141 AGTCGTTTG GTACCCAGA AGATTTGAAG TCTTGATGG ATAGAGCACA TGAGCTTGGT
1201 TTGCTAGTTC TCATGGATGT GGTTCATAGT CATCGCTCAA GTAAACTCT GGATGGTTG
1261 ATATGGTTTG ATGTCACAGA TACACATTAC TTTCACAGTG GTCCACCGTGG CCATCACTGG
1321 ATGTGTTGATT CTCGGCTTCT TAATATGGG AACTGGGAG TTTTAAGATT TCTTCTCTCC
1381 ATATCTAGAT GGTGGCTCGA GGAATATAAG TTTGATGGTTC TCCGTTTTGA TGGGTGTGACC
1441 TCCATGATGT ACACTCACCA CGGATTACAA GTAACATTAA CGGGGAACTT CAATGAGTAT
1501 TTTGGCTTGC CCACCGATGT AGATGCGATG TTGTTACTTGA TGCTGGTAAA TGATCTAATT
1561 CATGGACTTT ATCTTGAGGC TGTAACCAATT GGTAAAGATG TTAGTGGAAAT GCCTACATT
1621 GCCCTTCCTG TTCACGATGG TGGGGTAGGT TTTGACTATC GGATGCATAT GGCTGTGGCT
1681 GACAATGGG TTGACCTTCT CAAGCAAATG GATGAAACTT GGAAGATGGG TGATATTGTG
1741 CACACACTGA CAAATAGGAG GTGGTTAGAG AAGTGTGAA CTTATGCTGA AAGTCATGAT
1801 CAAGCATTAG TCGGGACACAA GACTATTGGG TTTGGTGTGA TGGACAAGGA TATGTATGAT
1861 TTGATAGGAC TTCAACTCC ACCATTGATC GTGGGATAGC ATTACATAAG
1921 ATGATTAGAC TTATCACAAAT GGGTTAGGA GGAGAGGGAT ATCTTAATT CATGGGAAT
1981 GAGTTGGAC ATCTGAATG GATAGATTTC CCAAGAGGTC CGCAAGACT TCCAAGTGGT
2041 AAGTTTATTG CAGGGATAAA CAACAGTTT GACAATGTC GTCGAAGATT TGACCTGGGT
2101 GATGCAGACT ATCTTAGGTA TCATGGTATG CAAGAGTTTG ATCAGGCAAT GCAACATCTT
2161 GAGCAAAAT ATGAATTCAT GACACTGTAT CACCACTGATA TTTCCCGAA ACATGAGGAG
2221 GATAAGGTGA TTGTTCTCGA AAAGGGAGAT TTGGTATTG TGTTCAACTT CCACGTGAAAC
2281 AACAGCTATT TTGACTACCC TATTGTTGT CGAAAGCTC GGGTGTATAA GGTGGTCTTG
2341 GACTCCGAGC CTGGACTATT TTGGTGGATT AGCAGGATCC ATCACCAGC CGAGCACTTC
2401 ACCGGCGACT GTTCGGCATGA TAATAGGCCA TATTGATCT CGGTTTATAC ACCAAGCAGA
2461 ACATGTTGCG TCTATGCTCC AGTGGAGTGA TAGCGGGGTA CTGGTGTGCTG CGCGGCATGT
2521 GTGGGGCTGT CGATGTGAGG AAAAACCTTC TTCCAAAAGG GGCAGATGCA TGCACTGATG
2581 CTACATAAG GTTCTGATAC TTAAATCGAT GCTGGAAAGC CCATGCATCT CGCTGCGTTG
2641 TCCCTCTCAT ATATATAAGA CCTTCAAGGT GTCAATTAAA CATAGAGTTT TCGTTTTTCG
2701 CTTTCCCTAAA AAAA AAAAAA AAAAAA

```

표 6

옥수수 분지 효소 I 유전자 및 트렌싯 펩티드의  
mRNA서열 및 추론된 아미노산 서열

[서열 번호 16 및 서열 번호 17]

유전자자리 MZEBET 2725 bp ss-mRNA PLN  
정의 옥수수 전분 분지 효소 I (BE- I ).  
접근 D11081  
키워드 전분 분지 효소 I .  
소오스 Zea mays L. (inbred dh43), cDNA에서 mRNA  
유기체 Zea mays  
Eukaryota; Plantae; Embryobionta; Magnoliophyta;  
Liliopsida; Commelinidae;

참조문헌 1 (영기 1내지 2763)

저자 Baba, T., Kimura, K., Mizuno, K., Etoh, H., Ishida, Y., Shida, O. 및 Arai, Y.

제목 옥수수 분지 효소 I에서 전분 가수분해효소의 촉매 영역의 서열전환(Sequence conservation of the catalytic region of Amylolytic enzyme in maize branching enzyme I )

저널 Biochem. Biophys. Res. Commun. 181, 87-94(1991)

기준 완전 자동

코멘트 Tadashi Baba에 의해 DDBJ에 제출되었음(1992년 4월 30일)

일본국, 이바라키 305, 츄쿠바, 츄쿠바대학, 응용 생화학 연구소.

전화: 0298-53-6632

팩스: 0298-53-6632

NCBI gi: 217959

특징 위치/한정물

소오스 1..2763  
/유기체='Zea mays'

CDS <1..2470  
/주='NCBI gi: 217960'  
/코돈 개시=2  
/생성물='분자 효소-1 전구체'  
/번역='LCLVSPSSPTPLPPRRSRSHADRAAPPGIAGGGNVRSLSVLSV  
QCKARRSGVRKVSKFATAATVQEDKTMATAKGVDHLPYDLDPKLEIFKDHFRYRM  
KRFLEQKGSIEENEGSLESFSKGYLGINTNEDGTVYREWAPAAQEAEELIGDFNDWN  
GANHKMEKDGFVWSIKIDHVKGKPAIPHNSKVKFRFLHGGVVWDRIPALIRYATVDA  
SKFGAPYDGVHWDPPASERYTFKHPRPSKPAAPRIYEAHVGMGSGEKPAVSTYREFADN  
VLPRIRANNYNTVQLMAVMEMHSYYASFGYHVTNFFAVSSRSGTPEDLKYLVDKAHSLG  
LRVLMDVVHSHASNNTDGLNGYDVGQSTQESYFHAGDRGYHKLWDSRLFNYANWEVL  
RFLLSNLRYWLDEFMDGFRFDGVTSMLYHHHGINVGFTGNYQEYFSLDTAVDAVVY  
MLANHLMHKLLPEATVVAEDVSGMPVLCRPVDEGGVGFYRLAMAIPDRWIDYLNKD  
DSEWSMGEIAHTLTNRRYTEKCIAYAESHDSIVGDKTIAFLMDKEMYTGMSDLQPA  
SPTIDRGIALQKMIHFITMALGGDGYLNFMGNEFGHPEWIDFPREGNNWSYDKCRRQW  
SLVDTDHLRYKYMNAFDQAMNALDERFSFLSSSKQIVSDMNDEEKVIVFERGDLVVF  
NFHPKKTYEGYKVGCDLPGKYRVALDSALVFGGHGRVGHVDHFSTSPGVPGVPETN  
FNNRPNSFKVLSPPRTCVAYYRVDEAGAGRRLHAKAETGKTPAESIDVKASRASSKE  
DKEATAGGKKGWKFARQPSDQDTK"

트랜싯 펩티드 2..190

성숙한 펩티드 191..2467  
/EC 번호='2. 4. 1. 18'  
코돈 출발=1  
생성물='분자 효소-1 전구체'

폴리A 시그날 2734..2739

영기 집계 719A 585C 737G 722T

## 기원

```

1 GCTGTGCGCTC GTGTCGCCCT CTTCCCGCCC GACTCCGCTT CCGCCGCCGC GGCCTCTCG
61 CTCGCATGCT GATCGGGCGG CACCGCCGG GATCGGGGT GCGGCCAATG TGCCTCTGAG
121 TGTGTTGCTC GTCCAGTGCAG AGGCTCGCCG GTCAGGGGTG CGGAAGGTCA AGACAAATT
181 CGCCACTGCA GCTACTGTC AAGAAGATAA AACTATGGCA ACTGCCAAG GCGATGTCGA
241 CCATCTCCCC ATATACGACC TGGACCCCAA GCTGGAGATA TTCAAGGACC ATTTCAGGTA
301 CGGGATGAAA AGATTCCTAG AGCAGAAAGG ATCAATTGAA GAAAATGAGG GAAGTCTTGA
361 ATCTTTTCTT AAAGGCTATT TGAAATTG GATTAATACA AATGAGGATG GAACGTATA
421 TCGTGAATGG GCACCTGCTC CGCAGGAGGC AGAGCTTATT GTGACTTCATGACTGGAA
481 TGGTGCACAA CATAAGATGG AGAAGGATAA ATTTGGTGTG TGTCGATCA AAATTGACCA
541 TGTCAAAGGG AACCTGCGA TCCCTCACAA TTCCAAGGTT AAATTTCGCT TTCTACATGG
601 TGGAGTATGG GTTGATCGTA TTCCAGCATT GATTCTGGT GCGACTGTT ATGCCCTCTAA
661 ATTTGGAGCT CCGTATGATG GTGTTCATG GGATCTCTC GCTTCGAAA GGTACACATT
721 TAAGGCATCCT CGGCCTTCAGA AGCCTGCTGC TCCACGATTC TATGAAGGCC ATGTAAGGTAT
781 GAGTGGTCAA AAGCCAGCAG TAAGCACATA TAGGGAATTG GCACACATG TGTGCCCCAG
841 CATAACGAGCA AATAACTACA ACACAGTCA GTTGATGGCA GTTATGGAGC ATTCGTA
901 TGCTTCTTCC GGGTACCATG TGACAAATTG CTTTGCCTT AGCAGCAGAT CAGGCACACC
961 AGAGGACCTC AAATATCTTG TTGATAAGGC ACACAGTTG GGTGCGGAG TTCTGATGGA
1021 TGTGTCATG AGCCATGCGA GTAATAATGT CACAGATGGT TAAATGGCT ATGATGTTGG
1081 ACAAAAGCACC CAAGAGTCATC ATTTTCATGC GGGAGATAGA GTTATCATA AACTTTGGGA
1141 TAGTCGGCTG TCAACTATG CTAACGTGGC GGTTATTAGG TTCTCTTT CTACACCTGAG
1201 ATATTGGTTG GATGAATTCA TGTTTGATGG CTTCCGATTG GATGGAGTTA CATCAATGCT
1261 GTATCATCAC GATGGATATC ATGTTGGTTT TACTGGAAAC TACCAAGGAAT ATTTCAGTTT
1321 GGACACAGCT GTGGATCGAG TTGATTTACAT GATGCTTGC AACCATTAA TGCAACAAACT
1381 CTTGCCAGAA GCAACTGTTG TTGCTGAAGA TGTTCAAGG ATGCCGGTCC TTGCGGGCC
1441 AGTTGATGAA GGTGGGGTTG GGTGTTGACTA TCGCCTGGCA ATGGCTATCC CTGATAGATG
1501 GATTGACTAC CGAAGAATAA AGATGACTC TGAGTGGTCG ATGGTGAAA TAGGGCATA
1561 TTTGACTAAC AGGAGATATA CTGAAAATG CATCGCATAT GCTGAGAGCC ATGATCAGTC
1621 TATTGTTGGC GACAAAATCA TTGCAATTCT CCGTGGACG AAGGAAATGT ACACGGCAT
1681 GTCAGACTTG CAGCCTGCG CACCTACAT TGATCGAGGG ATTGCACTCC AAAAGATGAT
1741 TCACCTTCATC AACAATGCCCT TTGGAGGTGA TGCGTACTTG AATTATATGG GAAATGAGTT
1801 TGGTCACCCA GAATGGATTG ACTTTCCAAG AGAAGGGAAAC AACTGGAGCT ATGATAAATG
1861 CAGACAGCAG TGGAGCTTG TGAGACACTGA TCACTGGCGG TACAAGTACA TGATGCGTT
1921 TGACCAAGCG ATGAATGGC CCGATGAGAG ATTTCCTTC CTTTCGCTGTT CAACGAGAT
1981 CGTCAGCGAC ATGAACGATG AGGAAAAGGT TATTGCTTTT GAACTGGAG ATTTCAGTTT
2041 TGTTTTCAAT TTCCCATCAGA AGAAAACCTA CGAGGGCTAC AAAGTGGGAT GCGATTGCG
2101 TGGGAAATAC AGAGTAGCCC TGGAACCTGA TGCTCTGTC TTGGTGGAC ATGGAAGAGT
2161 TGGCCACGAC GTGGATCACT TCACGTGCC TGAAAGGGTG CCAGGGGTG CCGAACGAA
2221 CTTCAACAAAC CGGCGAACAT CGTTCAAAGT CCTTCTCGG CCCCCGACCT GTGTCGCTTA
2281 TTACCGTGTG GACGAACGAG GGGCTGGACG ACGTCTTCAC GCGAACAGCAG AGACAGGAAA
2341 GACGTCTCCA CGAGAGAGCA TCGACGTCAA AGCTTCCAGA GCTAGTAGCA AAGAAGACAA
2401 GGAGGCACAG GCTGGTGGCA AGAAGGGATG GAAGTTGCG CGCGAGCCAT CGCATCAAGA
2461 TACCAAATGA AGCCACAGT CCGTGGTGGAG GACTGGACTG GCTGCCGGCG CCCTGTTAGT
2521 AGTCTGCTC TACTGGACTA GCGCCCGCTG CGCCCGCTTGG AACGGTCCCT TCCGTAGCT
2581 TGCAGGGGAC TGGTGTCTCA TCACCGAGCA GGCAGGGACT GCTTGATAG CTTTCTAGA
2641 ATAATAATCA GGGATGGATG GATGGTGTGT ATTGGCTATC TGGCTAGACG TGCAATGTGCC
2701 CAGTTGTAT GTACAGGAGC AGTCCCCGTC CAGAATAAAAA AAAAACTTGT TGGGGGGTTT
2761 TTC

```

표 7

가용성 전분 신타아제 I 옥수수 유전자(153bp)의 트렌싯  
펩티드 영역의 암호화 서열 및 추론된 아미노산 서열

[서열 번호 18 및 서열 번호 19]

파일 명: MSS1TRPT. DNA 서열 : 노르말 153 BP

코돈 테이블 : UNIV.TCH

서열 영역 : 1-153

번역 영역 : 1-153

\*\*\* DNA 번역 \*\*\*

1 ATG GCG ACG CCC TCG GCC GTG GGC GCG TGC CTC CTC CTC GCG CGG	48
1 M A T P S A V G A A C L L L A R	16
49 GCC GCC TGG CCG GCC GCC GTC GGC GAC CGG GCG CGC CCG CGG AGG CTC	96
17 A A W P A A V G D R A R P R R L	32
97 CAG CGC GTG CTG CGC CGC CGG TGC GTC GCG GAG CTG ACC AGG GAG GGG	144
33 Q R V L R R R C V A E L S R E G	48
145 CCC CAT ATG	153
49 P H M	51

GFP 구성물:

1. pET-21a에서만의 GFP:

pEXS115는 Nde I 및 Xho I에 의해 소화되고 SGFP 암호화 서열을 함유하는 740 bp 단편은 pET-21a(Novagen 601 Science Dr. Madison WI)의 Nde I 및 Xho I 부위내에 서브클론화된다. (도 2b GFP-21a 지도 참조).

2. 완전한 길이의 성숙한 WX의 5' 말단에서 프레임내(in-frame)서브클론화된 GFP:

pEXS114로 부터의 SGSP를 함유하는 740bp Nde I 단편은 pEXSWX의 Nde I부위내에 서브클론화된다.

(도 3a GFP-FLWX 지도 참조).

3. N-말단이 절단된 WX의 5'에서 프레임내 서브클론화된 GFP:

N-말단에서 700bp에 의해 절단된 WX.

pEXSWX로부터의 WX의 C-말단을 암호화하는 1 kb BamH I 단편은 pEXS115의 Bgl II 부위내에 서브클론화된다. 이어서 전체의 SGFP-절단된 WX단편은 Nde I-HindIII 단편으로서 pET21내에 서브클론화된다. (도 3b GFP-BamHIWX 지도 참조).

4. 절단된 WX의 5' 말단에서 프레임내 서브클론화된 GFP. N-말단에서 100bp에 위해 절단된 WX.

pEXS115로부터의 SGFP를 함유하는 740bp Nde I-Nco I 단편은 Nde I 및 Nco I 부위에서 pEXSWX내에 서브클론화된다. (도 4 GDP-NcoWX 지도 참조).

실시예 3

세균내로 플라스미드의 형질변환:

Escherichia coli 반응성 세포 제조:

1. 단일 콜로니의 원하는 *E. coli* 균주를 이용하여 2.5ml LB 배지를 접종한다: 선택된 균주는 Stratagene로부터 입수한 XL1BLUE DL21DE3였으며, 적절한 항생물질을 포함했다. 37°C에서 200rpm으로 밤새 성장시킨다.

2. 100ml의 LB 배지를, 적절한 항생물질을 함유하는 1:50의 상기 밤새 배양된 희석액으로 접종했다. OD<sub>600</sub>=0.3-0.5가 될 때까지 37°C에서 200rpm으로 성장시킨다.

3. 배양체를 살균 원심분리병에 옮기고 얼음위에서 15분간 유지하여 차갑게 한다.

4. 3,000xg(4°C)에서 5분간 원심분리한다.

5. 펠릿을 얼음-냉각된 형질변환 완충액 1에 혼탁시킨다.

6. 3,000xg(4°C)에서 5분간 원심분리한다.

7. 펠릿을 얼음-냉각된 형질변환 완충액 2에 혼탁시킨다.

8. 분취량을 액체 질소중에서 속성 동결시키고, -70°C에서 저장한다.

형질변환 완충액 1

형질변환 완충액 2

RbCl	1.2g	MOPS(10mM)	0.209g
------	------	------------	--------

MnCl <sub>2</sub> 4H <sub>2</sub> O	0.99g	RbCl	0.12g
-------------------------------------	-------	------	-------

초산칼륨	0.294g	CaCl <sub>2</sub> 2H <sub>2</sub> O	1.1g
------	--------	-------------------------------------	------

CaCl <sub>2</sub> 2H <sub>2</sub> O	0.15g	글리세롤	15g
-------------------------------------	-------	------	-----

글리세롤	15g	dH <sub>2</sub> O	100ml
------	-----	-------------------	-------

dH <sub>2</sub> O	100ml	NaOH에 의해 pH 6.8	
-------------------	-------	-----------------	--

0.2M 아세트산에 의해 pH 5.8	필터 살균
----------------------	-------

필터 살균

영화루비듐 염총격법에 의한 *Escherichia coli* 형질변환: Hanadan, D.(1985) in DNA cloning: a practical approach(Glover, D.M. 발행), pp.109-135, IRL Press 참조.

1. 얼음위에있는 1-5μl의 DNA를 150μl *E. coli* 반응성 세포와 함께 30분간 배양한다.

2. 42°C에서 45초간 염총격을 가한다.

3. 그후 즉시 2분간 얼음위에 둔다.

4. 600μl LB 배지를 가하고 37°C에서 1시간 배양한다.

5. 적절한 항생물질을 포함하는 LB 아가(agar)상에서 배양한다.

이러한 플라스미드는 그린 형광 단백질을 함유하는 잡종 폴리펩티드를 세균내에서 발현시킬 수 있다.

실시예 4

*E. coli*에서 구성물의 발현:

1. 3ml LB를 관심있는 플라스미드를 함유하는 *E. coli*로 접종한다. 적절한 항생물질을 포함시킨다. 37°C, 250rpm, 밤새.

2. 이와같이 밤새 배양한 것으로 100ml LB를 접종시킨다. 적절한 항생물질을 포함시킨다. 37°C에서 250rpm으로 성장시킨다.

3. OD<sub>600</sub>가 약 0.4-0.5일 때, 실온 및 200rpm으로 한다.

4. OD<sub>600</sub>가 약 0.6-0.8일 때, 100 $\mu$ l 1M 1PTG를 이용하여 유도한다. 최종 1PTG 농도는 1mM이다.
5. 실온에서 200rpm으로 4-5시간동안 성장시킨다.
6. 원심분리에 의해 세포를 수거한다.
7. 액체 질소에서 속성 동결시키고 사용할 때까지 -70°C에서 저장한다.

세포는  $\text{dH}_2\text{O}$ 에 혼탁될 수 있고 UV 광( $\lambda_{\text{max}} = 395\text{nm}$ )하에서 고유 형광성에 대하여 관찰 될 수 있다. 선택적으로, 세포는 초음파처리될 수 있고 세포 추출물의 분취량이 SDS-PAGE에 의해 분리되고 GFP를 검출하기 위해 UV 광하에서 관찰될 수 있다. 이용된 단백질이 그린 형광 단백질인 경우, 용해된 물질내에서 단백질의 존재가 라이트 박스에서 395nm의 UV하에서 평가될 수 있고 그린색의 열 신호가 확인될 수 있다.

## 실시예 5

### 세균으로부터 플라스미드 추출:

하기의 사항은 본 발명을 실시하는데 있어서 유용한 가장 일반적인 알칼리성 용해 플라스미드 제제 프로토콜들 중 하나이다.

1. 상기에 기술된 플라스미드들중 하나의 것을 이용하여 형질변환시킨 단일 콜로니의 *E. coli*로 100-200ml LB를 접종시킨다. 적절한 항생물질을 함유시키킨다. 37°C에서 200rpm으로 밤새 성장시킨다.
2. 5,000x g(4°C)에서 10분간 원심분리한다.
3. 세포를 10ml 물에 혼탁시키고, 15ml 원심분리관에 옮겨서 원심분리를 반복한다.
4. 펠릿을 5ml 0.1 M NaOH, 0.5% SDS에 혼탁시킨다. 얼음위에서 10분간 배양한다.
5. 2.5ml의 3M 초산나트륨(pH 5.2)을 가하고, 온건하게 전화시키고, 얼음위에서 10분간 배양한다.
6. 15,000-20,000x g(4°C)에서 5분간 원심분리한다.
7. 동일 체적의 페놀: 클로로포름: 이소아밀 알콜(25:24:1)을 이용하여 상정액을 추출한다.
8. 6,000-10,000x g(4°C)에서 10분간 원심분리한다.
9. 수성상을 세척관에 옮기고 1배 부피의 이소프로판올을 이용하여 침전시킨다.
10. 12,000x g(4°C)에서 15분간 원심분리한다.
11. 펠릿을 0.5ml TE에 용해시키고, 20μl의 10mg/ml RNase를 부가하고, 37°C에서 1시간 배양한다.
12. 페놀: 클로로포름: 이소아밀 알콜(25:24:1)을 이용하여 두 번 추출한다.
13. 클로로포름을 이용하여 한 번 추출한다.
14. 1배 체적의 이소프로판올 및 0.1배 체적의 3M 초산나트륨을 이용하여 수성상을 침전시킨다.
15. 펠릿을 70% 에탄올을 이용하여 한 번 세척한다.
16. 펠릿을 스피드백(SpeedVac)에서 건조시키고 TE에 혼탁시킨다.

이어서, 이러한 플라스미드를 다른 숙주내로 삽입한다.

五 8

pEXS52로 부터의 전분 신타아제 암호화 영역의

## DNA 서열 및 추론된 아미노산 서열

[서열 번호 20 및 서열 번호 21]

파일 명: MSS1DELN. DNA 서열 : 노르말 1626 BP

코돈 테이블 : UNIV.TCH

서열 영역 : 1-1626

번역 영역 : 1-1626

(xi) 서열 설명: 서열 번호: 20:

TGC GTC GCG GAG CTG AGC AGG GAG GAC CTC GGT CTC GAA CCT GAA GGG Cys Val Ala Glu Leu Ser Arg Glu Asp Leu Gly Leu Glu Pro Glu Gly 55 60 65	48
ATT GCT GAA GGT TCC ATC GAT AAC ACA GTC GTT GTG GCA AGT GAG CAA Ile Ala Glu Gly Ser Ile Asp Asn Thr Val Val Val Ala Ser Glu Gln 70 75 80	96
GAT TCT GAG ATT GTG GTT GGA AAG GAG CAA GCT CGA GCT AAA GTC ACA Asp Ser Glu Ile Val Val Gly Lys Glu Gln Ala Arg Ala Lys Val Thr 85 90 95	144
CAA AGC ATT GTC TTT GTA ACC GGC GAA GCT TCT CCT TAT GCA AAG TCT Gln Ser Ile Val Phe Val Thr Gly Glu Ala Ser Pro Tyr Ala Lys Ser 100 105 110 115	192
GGG GGT CTA GGA GAT GTT TGT GGT TCA TTG CCA GTT GCT CTT GCT GCT Gly Gly Leu Gly Asp Val Cys Gly Ser Leu Pro Val Ala Leu Ala Ala 120 125 130	240
CGT GGT CAC CGT GTG ATG GTT GTA ATG CCC AGA TAT TTA AAT GGT ACC Arg Gly His Arg Val Met Val Met Pro Arg Tyr Leu Asn Gly Thr	288

135	140	145	
TCC GAT AAG AAT TAT GCA AAT GCA TTT TAC ACA GAA AAA CAC ATT CGG Ser Asp Lys Asn Tyr Ala Asn Ala Phe Tyr Thr Glu Lys His Ile Arg 150 155 160			336
ATT CCA TGC TTT GGC GGT GAA CAT GAA GTT ACC TTC TTC CAT GAG TAT Ile Pro Cys Phe Gly Gly His Glu Val Thr Phe Phe His Glu Tyr 165 170 175			384
AGA GAT TCA GTT GAC TGG GTG TTT GTT GAT CAT CCC TCA TAT CAC AGA Arg Asp Ser Val Asp Trp Val Phe Val Asp His Pro Ser Tyr His Arg 180 185 190 195			432
CCT CGA AAT TTA TAT GGA GAT AAG TTT GGT GCT TTT GGT GAT AAT CAG Pro Gly Asn Leu Tyr Gly Asp Lys Phe Gly Ala Phe Gly Asp Asn Gln 200 205 210			480
TTC AGA TAC ACA CTC CTT TGC TAT GCT GCA TGT GAG GCT CCT TTG ATC Phe Arg Tyr Thr Leu Leu Cys Tyr Ala Ala Cys Glu Ala Pro Leu Ile 215 220 225			528
CTT GAA TTG GGA GGA TAT ATT TAT GGA CAG AAT TGC ATG TTT GTT GTC Leu Glu Leu Gly Gly Tyr Ile Tyr Gly Gln Asn Cys Met Phe Val Val 230 235 240			576
AAT GAT TGG CAT GCC AGT CTA GTG CCA GTC CTT CTT GCT GCA AAA TAT Asn Asp Trp His Ala Ser Leu Val Pro Val Leu Leu Ala Ala Lys Tyr 245 250 255			624
AGA CCA TAT GGT TAT AAA GAC TCC CGC AGC ATT CTT GTC ATA CAT Arg Pro Tyr Gly Val Tyr Lys Asp Ser Arg Ser Ile Leu Val Ile His 260 265 270 275			672
AAT TTA GCA CAT CAG GGT GTA GAG CCT GCA AGC ACA TAT CCT GAC CTT Asn Leu Ala His Gln Gly Val Glu Pro Ala Ser Thr Tyr Pro Asp Leu 280 285 290			720
GGG TTG CCA CCT GAA TGG TAT GGA GCT CTG GAG TGG GTC TTC CCT GAA Gly Leu Pro Pro Glu Trp Tyr Gly Ala Leu Glu Trp Val Phe Pro Glu 295 300 305			768
TGG GCG AGG AGG CAT GCC CTT GAC AAG GGT GAG GCA GTT AAT TTT TTG Trp Ala Arg Arg His Ala Leu Asp Lys Gly Glu Ala Val Asn Phe Leu 310 315 320			816
AAA GGT GCA GTT GTG ACA GCA GAT CGA ATC GTG ACT GTC AGT AAG GGT Lys Gly Ala Val Val Thr Ala Asp Arg Ile Val Thr Val Ser Lys Gly 325 330 335			864
TAT TCG TGG GAG GTC ACA ACT GCT GAA GGT GGA CAG GGC CTC AAT GAG Tyr Ser Trp Glu Val Thr Thr Ala Glu Gly Gln Gly Leu Asn Glu 340 345 350 355			912
CTC TTA AGC TCC AGA AAG AGT GTA TTA AAC GGA ATT GTC AAT GGA ATT Leu Leu Ser Ser Arg Lys Ser Val Leu Asn Gly Ile Val Asn Gly Ile 360 365 370			960
GAC ATT AAT GAT TGG AAC CCT GCC ACA GAC AAA TGT ATC CCC TGT CAT Asp Ile Asn Asp Trp Asn Pro Ala Thr Asp Lys Cys Ile Pro Cys His 375 380 385			1008
TAT TCT GTT GAT GAC CTC TCT GGA AAG GCC AAA TGT AAA GGT GCA TTG Tyr Ser Val Asp Asp Leu Ser Gly Lys Ala Lys Cys Lys Gly Ala Leu 390 395 400			1056

CAG AAG GAG CTG GGT TTA CCT ATA AGG CCT GAT GTT CCT CTG ATT GGC Gln Lys Glu Leu Gly Leu Pro Ile Arg Pro Asp Val Pro Leu Ile Gly 405 410 415	1104
TTT ATT GGA AGG TTG GAT TAT CAG AAA GGC ATT GAT CTC ATT CAA CTT Phe Ile Gly Arg Leu Asp Tyr Gln Lys Gly Ile Asp Leu Ile Gln Leu 420 425 430 435	1152
ATC ATA CCA GAT CTC ATG CGG GAA GAT GTT CAA TTT GTC ATG CTT GGA Ile Ile Pro Asp Leu Met Arg Glu Asp Val Gln Phe Val Met Leu Gly 440 445 450	1200
TCT GGT GAC CCA GAG CTT GAA GAT TGG ATG AGA TCT ACA GAG TCG ATC Ser Gly Asp Pro Glu Leu Glu Asp Trp Met Arg Ser Thr Glu Ser Ile 455 460 465	1248
TTC AAG GAT AAA TTT CGT GGA TGG GTT GGA TTT AGT GTT CCA GTT TCC Phe Lys Asp Lys Phe Arg Gly Trp Val Gly Phe Ser Val Pro Val Ser 470 475 480	1296
CAC CGA ATA ACT GCC GGC TGC GAT ATA TTG TTA ATG CCA TCC AGA TTC His Arg Ile Thr Ala Gly Cys Asp Ile Leu Leu Met Pro Ser Arg Phe 485 490 495	1344
GAA CCT TGT GGT CTC AAT CAG CTA TAT GCT ATG CAG TAT GGC ACA GTT Glu Pro Cys Gly Leu Asn Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val 500 505 510 515	1392
CCT CTT GTC CAT GCA ACT GGG GGC CTT AGA GAT ACC GTG GAG AAC TTC Pro Val Val His Ala Thr Gly Gly Leu Arg Asp Thr Val Glu Asn Phe 520 525 530	1440
AAC CCT TTC GGT GAG AAT GGA GAG CAG GGT ACA GGG TGG GCA TTC GCA Asn Pro Phe Gly Glu Asn Gly Glu Gln Gly Thr Gly Trp Ala Phe Ala 535 540 545	1488
CCC CTA ACC ACA GAA AAC ATG TTT GTG GAC ATT GCG AAC TGC AAT ATC Pro Leu Thr Thr Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile 550 555 560	1536
TAC ATA CAG GGA ACA CAA GTC CTC CTG GGA AGG GCT AAT GAA GCG AGG Tyr Ile Gln Gly Thr Gln Val Leu Leu Gly Arg Ala Asn Glu Ala Arg 565 570 575	1584
CAT GTC AAA AGA CTT CAC GTG GGA CCA TGC CGC TGA His Val Lys Arg Leu His Val Gly Pro Cys Arg *	1620
580 585 590	

## (2) 서열번호 21에 대한 정보

## (i) 서열특징

(A)길이: 540 아미노산

(B)타입: 아미노산

(C)토풀로지: 선형

## (ii) 문자 유형: 단백질

## (iii) 서열 설명: 서열 번호 21

Cys Val Ala Glu Leu Ser Arg Glu Asp Leu Gly Leu Glu Pro Glu Gly  
 1 5 10 15  
 Ile Ala Glu Gly Ser Ile Asp Asn Thr Val Val Val Ala Ser Glu Gln  
 20 25 30  
 Asp Ser Glu Ile Val Val Gly Lys Glu Gln Ala Arg Ala Lys Val Thr  
 35 40 45  
  
 Gln Ser Ile Val Phe Val Thr Gly Glu Ala Ser Pro Tyr Ala Lys Ser  
 50 55 60  
 Gly Gly Leu Gly Asp Val Cys Gly Ser Leu Pro Val Ala Leu Ala Ala  
 65 70 75 80  
 Arg Gly His Arg Val Met Val Val Met Pro Arg Tyr Leu Asn Gly Thr  
 85 90 95  
 Ser Asp Lys Asn Tyr Ala Asn Ala Phe Tyr Thr Glu Lys His Ile Arg  
 100 105 110  
 Ile Pro Cys Phe Gly Gly Glu His Glu Val Thr Phe Phe His Glu Tyr  
 115 120 125  
 Arg Asp Ser Val Asp Trp Val Phe Val Asp His Pro Ser Tyr His Arg  
 130 135 140  
 Pro Gly Asn Leu Tyr Gly Asp Lys Phe Gly Ala Phe Gly Asp Asn Gln  
 145 150 155 160  
 Phe Arg Tyr Thr Leu Leu Cys Tyr Ala Ala Cys Glu Ala Pro Leu Ile  
 165 170 175  
 Leu Glu Leu Gly Gly Tyr Ile Tyr Gly Gln Asn Cys Met Phe Val Val  
 180 185 190  
 Asn Asp Trp His Ala Ser Leu Val Pro Val Leu Leu Ala Ala Lys Tyr  
 195 200 205  
 Arg Pro Tyr Gly Val Tyr Lys Asp Ser Arg Ser Ile Leu Val Ile His  
 210 215 220  
 Asn Leu Ala His Gln Gly Val Glu Pro Ala Ser Thr Tyr Pro Asp Leu  
 225 230 235 240  
 Gly Leu Pro Pro Glu Trp Tyr Gly Ala Leu Glu Trp Val Phe Pro Glu  
 245 250 255  
 Trp Ala Arg Arg His Ala Leu Asp Lys Gly Glu Ala Val Asn Phe Leu  
 260 265 270  
 Lys Gly Ala Val Val Thr Ala Asp Arg Ile Val Thr Val Ser Lys Gly  
 275 280 285  
 Tyr Ser Trp Glu Val Thr Thr Ala Glu Gly Gln Gly Leu Asn Glu  
 290 295 300  
 Leu Leu Ser Ser Arg Lys Ser Val Leu Asn Gly Ile Val Asn Gly Ile  
 305 310 315 320  
 Asp Ile Asn Asp Trp Asn Pro Ala Thr Asp Lys Cys Ile Pro Cys His  
 325 330 335  
 Tyr Ser Val Asp Asp Leu Ser Gly Lys Ala Lys Cys Lys Gly Ala Leu  
 340 345 350  
 Gln Lys Glu Leu Gly Leu Pro Ile Arg Pro Asp Val Pro Leu Ile Gly  
 355 360 365  
 Phe Ile Gly Arg Leu Asp Tyr Gln Lys Gly Ile Asp Leu Ile Gln Leu  
 370 375 380  
 Ile Ile Pro Asp Leu Met Arg Glu Asp Val Gln Phe Val Met Leu Gly  
 385 390 395 400  
 Ser Gly Asp Pro Glu Leu Glu Asp Trp Met Arg Ser Thr Glu Ser Ile  
 405 410 415  
  
 Phe Lys Asp Lys Phe Arg Gly Trp Val Gly Phe Ser Val Pro Val Ser  
 420 425 430  
 His Arg Ile Thr Ala Gly Cys Asp Ile Leu Leu Met Pro Ser Arg Phe  
 435 440 445  
 Glu Pro Cys Gly Leu Asn Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val  
 450 455 460  
 Pro Val Val His Ala Thr Gly Gly Leu Arg Asp Thr Val Glu Asn Phe  
 465 470 475 480  
 Asn Pro Phe Gly Glu Asn Gly Glu Gln Gly Thr Gly Trp Ala Phe Ala  
 485 490 495  
 Pro Leu Thr Thr Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile  
 500 505 510  
 Tyr Ile Gln Gly Thr Gln Val Leu Leu Gly Arg Ala Asn Glu Ala Arg  
 515 520 525  
 His Val Lys Arg Leu His Val Gly Pro Cys Arg \* 540  
 530 535 540  
 121-37

## 실시예 6:

이 실험은 옥수수 프로모터, 옥수수 트렌싯 펩티드, 전분 신타아제 I 유전자로 부터의 전분-캡슐화 영역 및 부착된 결찰 유전자 단편을 가지는 플라스미드를 이용한다. 도 6에서 도시된 플라스미드는 도 8에서 열거된 DNA 서열을 함유한다.

플라스미드 pEXS52를 하기의 프로토콜에 따라 구성하였다.

트랜스게닉 플라스미드를 구성하기위해 사용된 물질은 다음과같다:

플라스미드 pBluescript SK-

플라스미드 pMF6(nos3' 터미네이터를 함유함)

플라스미드 pHHK1(옥수수 adhI인트론을 함유함)

플라스미드 MstsI(6-4)(옥수수 stsi 트렌싯 펩티드를 함유함, PCT stsi 트렌싯 펩티드에 대한 주형으로서 사용함)

pBluescript SK-내의 플라스미드 MstsIII

프라이머 EXS29 (GTGGATCCATGGCGACGCCCTGGCCGTGG) [서열번호: 22]

EXS35(CTGAATTCCATATGGGCCCTCCCTGCTCAGCTC) [서열번호:23]. 둘 모두 PCTstsi 트렌싯 펩티드에 대하여 사용됨.

프라이머 EXS31(CTCTGAGCTAAGCTTGCCTTCTTCTTAATG) [서열 번호:24]

EXS32(GTCTCCGGGGTGTGCTTCCTGCTTCAG) [서열 번호 25]

둘 모두 PCR 옥수수 10KD 제인 프로모터에 대하여 사용됨(Journal: Gene 71:359-370[1988]).

옥수수 A62 게놈 DNA(PCR 옥수수 10KD 제인 프로모터로서 사용됨).

단계 1: pBluescript SK-에서 10KD 제인 프로모터를 클론화한다(pEXS10zp로서 명명함).

1. PCR 1.1Kb 옥수수 10KD 제인 프로모터

프라이머: EXS31, EXS32

주형: 옥수수 A632 게놈 DNA

2. 1.1Kb 옥수수, 10KD 제인 프로모터 PCR 생성물을 SacI 및 SacII 부위에서 pBluescript SK-LH에 클론화한다(도 7 참조).

단계 2: pEXS10zp에서 NdeI를 제거한다(pEXS10zp-NdeI로 명명함).

NdeI를 pBluescript SK에서의 옥수수 10KD 제인 프로모터로부터 무단(blunt)말단 결찰에 의하여 제거한다.

단계 3: pBluescript SK-에서 옥수수 adhI 인트론을 클론화한다(pEXSadhl로 명명함).

옥수수 adhI 인트론을 XbaI 및 BamHI 부위에서 플라스미드 pHHK1로부터 방출한다. 옥수수 adhI 인트론(XbaI/BamHI 단편)를 XbaI 및 BamHI 부위에서 pBluescript SK-내에 클론화한다(도 7참조).

단계 4: 옥수수 10KD 제인 프로모터 및 옥수수 adhI 인트론을 pBluescript SK-내에 클론화한다(pEXS10zp-adhI로 명명함).

옥수수 10KD 제인 프로모터를 SacI 및 SacII 부위에서 플라스미드 pEXS10zp-NdeI로부터 방출한다. 옥수수 10KD 제인 프로모터(SacI /Sac II 단편)를 SacI 및 SacII 부위에서 플라스미드 pEXSadhl(옥수수 adhI 인트론을 함유함)내에 클론화한다(도 7참조).

단계 5: 옥수수 nos3' 터미네이터를 플라스미드 pEXSadhl내에 클론화한다(pEXSadhl-nos3'로 명명함).

옥수수 nos3' 터미네이터를 EcoRI 및 HindIII 부위에서 플라스미드 pMF6로부터 방출한다.

옥수수 nos3' 터미네이터(EcoRI 및 HindIII 단편)를 EcoRI 및 HindIII 부위에서 플라스미드 pEXSadhl내에 클론화한다(도 7참조).

단계 6: 옥수수 nos3' 터미네이터를 플라스미드 pEXS10zp-adhI내로 클론화한다 (pEXS10zp-adhI-nos3'로 명명함).

옥수수 nos3' 터미네이터를 EcoRI 및 ApaI 부위에서 플라스미드 pEXSadhl-nos3'로부터 방출한다. 옥수수 nos3' 터미네이터(EcoRI 및 ApaI III 단편)를 EcoRI 및 ApaI 부위에서 플라스미드 pEXS10zp-adhIp내에 클론화한다(도 7참조).

단계 7: 옥수수 STSI 트렌싯 펩티드를 플라스미드 pEXS10zp-adhI-nos3'내에 클론화한다(pEXS33으로 명명함).

1. PCR 150bp 옥수수 STSI 트렌싯 펩티드

프라이머: EXS29, EXS35

주형: MSTS1(6-4)플라스미드

2. 150bp 옥수수 STSI 트렌싯 펩티드 PCR 생성물을 EcoRI 및 BamHI 부위에서 pEXS10zp-adhI-

nos3'내에 클론화한다(도 7참조).

단계 8: pEXS33의 옥수수 STS1 트렌싯 펩티드에 관한 부위-목표 돌연변이유발(site-directed mutagensis)(pEXS33(m)로 명명함).

플라스미드 pEXS33의 옥수수 STS1 트렌싯 펩티드에는 돌연변이(정지 코돈)이 있다. 정지 코돈을 비정지 코돈으로 변화시키기 위하여 부위-목표 돌연변이유발을 수행한다. 새로운 플라스미드(옥수수 10kd 제인 프로모터, 옥수수 STS1 트렌싯 펩티드, 옥수수 adh1 인트론, 옥수수 nos3' 터미네이터를 함유함)를 pEXS33(m)로 명명한다.

단계 9: pEXS33(m)에서 NotI 부위를 제거한다(pEXS50으로 명명함).

NotI 부위를 NotI 대체, 무딘 말단 결찰에 의해 pEXS33로부터 제거하여 pEXS50을 형성한다(도 8참조).

단계 10: pEXS33(m)에서 옥수수 adh1 인트론을 부위를 제거한다(pEXS60으로 명명함).

옥수수 adh1 인트론을 NotI/BamHI 소화, Klenow 단편에 의한 대체 및 무딘 말단 결찰에 의해 제거하여 pEXS60을 형성한다(도 9참조).

단계 11: 옥수수 STSIII를 pEXS50, pEXS60내에 클론화한다.

옥수수 STSIII를 NdeI 및 EcoRI 부위에서 pBluescript SK-의 STSIII FH로부터 방출한다. 옥수수 STSIII (NdeI-EcoRI 단편)를 pEXS50, pEXS60내에 클론화하여 pEXS51, pEXS61로 각각 명명한다(각각 도 8 및 도 9참조).

단계 12: 표8에서의 유전자를 NdeI/NotI 부위에서 pEXS51내에 클론화하여 pEXS52를 형성한다. 다른 유전자 (STS1, II, WX, glgA, glgB, glgC, BEI, BEII, 등)를 NdeI/NotI 부위에서 pEXS51, pEXS61내에 클론화함으로써 다른 유사한 플라스미드가 제조될 수 있다.

플라스미드 pEXS52를 벼내에 형질변환시켰다. pEXS52에 의해 형질변환된 재생된 벼 식물을 표지하여 머젠티 박스(magenta box)내에 두었다.

각각의 라인의 두 개의 자매세포(sibling)를 상기 머젠티 박스로부터 선택하여 토양 혼합물(이탄-질석과 50/50으로 혼합된 상부토양)이 채워진 2.5 인치 포트(pot)에 옮겼다. 그 포트들을 1/2인치의 물이 채워진 유리수조내에 두었다. 유리 수조의 상부를 덮어서 고습도가 유지되게 하였다(열방출을 막기 위해 몇 개의 구멍을 만듬). 온도계로 온도를 측정했다. 상기 유리 수조에 형광을 쪼여주었다. 제 1주간은 상기 식물에 비료를 사용하지 않았다. 빛을 쪼여주는 기간은 오전 6시에서 오후 8시 까지 최소 14시간이었다. 온도는 밤에는 최소 68°F, 그리고 낮에는 80-90°F를 유지하였다. 필요한 경우 뿌리 성장을 막기 위해 상기 유리 수조의 아래에 가열 매트(heating mat)를 사용했다. 이러한 상태에서 1주동안 식물을 유지시켰다. (주: 묘종은 낮은 광세기 때문에 크게 성장하기 시작했다).

제 1 주후, 상기 유리 수조의 상부를 개방하고 벼 형질변환체를 성장 챔버에 옮겨서 고습도 및 높은 광세기로 3주간 유지시켰다.

선택적으로, 고습도를 유지하기 위해 그린하우스의 물 혼합물이 사용될 수 있다. 상기 식물을 3주간 성장시켰다. 이어서, 상기 식물을 토양 혼합물(상부토양 및 이탄-질석, 50/50)이 채워진 6인치 포트(최소 5인치 포트)에 옮겼다. 이러한 포트를 1/2인치의 물이 채워진 트레이에 두었다. 식물에 비료작용을 주기위하여 15-16-17(N-K-P)를 250ppm으로 1주일에 한번 또는 외관에 의해 평가된 식물의 필요성에 따라 사용했다. 이러한 식물을 오전 6시부터 오후 8시까지 최소한 14시간동안 높은 세기의 광을 쪼여주고, 밤에는 최소70°F, 그리고 낮에는 80-90°F의 온도에서 유지시켜 주었다.

상기 식물에는 쌀알이 열렸고 상기 쌀알을 수확했다. 이와같이 수확된 쌀알은 상기 결찰된 아미노산 C, V, A, E, L, S, R, E[서열번호 27]의 존재에 대하여 추출 및 분석된 전분을 가질 수 있다.

실시예 7:

식물에 대한 SER 벡터:

도 6에서 도시된 플라스미드는 외떡입식물, 즉 옥수수에서 사용하기에 적합하다. 플라스미드 pEXS52는 프로모터, 트렌싯 펩티드(옥수수로부터의), 및 아미노산 서열 C, V, A, E, L, S, R, E[서열번호 27]를 암호화하는 결찰된 유전자 단편(TGC GTC GCG GAG CTG AGC AGG GAG)[서열 번호 26]을 가진다.

이러한 유전자 단편은 옥수수 가용성 전분 신타아제(MSTS1) 유전자의 N-말단에 근접하여 자연발생한다. 표 8에서 확인되는 바와같이, 대략 아미노산 292에서 전분 신타아제로부터 SER이 시작된다. 이러한 벡터는 옥수수 숙주내에 형질변환되는 것이 바람직하다. 상기 트렌싯 펩티드는 옥수수에 대하여 적을 되므로 옥수수가 바람직한 숙주이다. 명백히, 트렌싯 펩티드 및 프로모터는, 필요하다면, 원하는 숙주 식물에 대하여 적절하게 되도록 변경될 수 있다. 'whiskers' 기술(미합중국 특허 제 5,302,523호 및 제 5,464,765호)에 의한 형질변환후, 형질변환된 숙주 세포는 이 기술분야에 알려져있는 방법에 의해 재생되고, 형질변환체는 수분되고, 결과의 날알은 수거되어 전분 및 전분 입자내에서 펩티드의 존재에 대하여 분석될 수 있다.

하기의 바람직한 유전자들이 먹이를 개선하기 위하여 옥수수에서 이용될 수 있다: 피타아제, 성장호르몬: AUG AUG AUG AUG AUG AUG AUG AUG[서열 번호 28]; 및/또는 AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG[서열 번호 29]; 및/또는 AAA AAA AAA AAA AAA AAA[서열 번호 30], 또는 사슬내에서 리신 아미노산을 암호화하는 코돈들의 결합물 또는 리신 및 메티오닌 코돈 둘 모두를 암호화하는 코돈들의 결합물 또는 상기 아미노산들중 두 개 또는 세 개의 어떠한 결합물. 사슬의 길이는 과도하게 길지 않아야 하지만 사슬의 길이는 중요한 것으로 보이지 않는다. 따라서 아미노산은 전분 입자내에 캡슐화될 수 있거나 또는 식물 숙주의 전분 함유 영역에 형성된 전분내에 결합될 수 있다.

이러한 플라스미드는 거의 변경되지 않으면서 벼, 밀, 귀리, 수수, 또는 기장과 같은 다른 곡물내에 형질변환될 수 있다. 프로모터는 서열이 발표되어 있는 waxy 유전자이거나, 또는 이 기술분야에 알려져 있는 다른 제인 프로모터일 수 있다.

그밖에 이러한 플라스미드는 과도한 실험이 없이도 감자, 고구마, 타로토란, 망우수, 카사바, 땅콩, 완두콩, 콩, 또는 이집트콩과 같은 쌍떡잎 식물내에 형질변환될 수 있다.

외떡입식물 및 쌍떡입식물을 형질변환하는 여러 가지의 방법들이 산업에 알려져 있으므로 유전자들을 형질변환하는 방법은 본 발명에 대하여 중요한 것이 아니다. 플라스미드는 다음문헌[An 등, (1988) *Bibary Vectors, in Plant Molecular Biology Manual A3*, S.B. Gelvin 및 R. A. Schilperoort, 발행 (Dordrecht, The Netherland: Kluwer Academic Publishers, pp1-19)]의 동결-해동 방법에 의해 *Agrobacterium tumefaciens*내에 도입될 수 있다. 식물 물질의 구성 및 접종, 새싹의 재생, 및 새싹의 뿌리내림을 포함하는 *Agrobacterium* 접종물의 제조가 다음문헌[Edwards 등, 'Biochemical and molecular characterization of a novel starch synthase from phototodes,' *Plant J.* 8, 283-294]에 설명되어 있다. 다수의 캡슐화 영역이 다수의 상이한 유전자내에 존재한다. 단백질은 전분 입자내에 캡슐화되는 것(입자 캡슐화)이 바람직하긴 하지만, 비입자 전분내에서 캡슐화하는 것도 '캡슐화'라는 측면에서 본 발명의 범위에 또한 포함된다. 이러한 목적을 위하여 하기와 같은 유형의 유전자들이 바람직하다.

글리코겐 신타아제의 전분-캡슐화 영역의 사용:

*E. coli* 글리코겐 신타아제는 큰 단백질이 아니다: 이러한 구조적 유전자는 1431개의 염기쌍 길이를 가지며, 구체적으로 49,000의 평가된 분자량을 갖는 477개 아미노산의 단백질이다. 식물 계획내에 삽입된 세균성 유전자의 경우 코돈 사용의 문제가 발생할 수 있지만, 이러한 문제는 *E. coli*의 경우에 있어서는 *Bacillus*로부터의 것들과 같은 다른 세균으로부터의 것들 만큼이나 크지는 않은 것으로 알려져 있다. *E. coli*로부터의 글리코겐 신타아제는 옥수수의 유전자와 공통되는 코돈 사용 프로필을 가지지만, 공자의 방법에 따라 번역 개시점에서 서열을 변경시켜서 식물 콘센서스 서열과 더욱 화합성이 있게하는 것이 바람직하다:

glgA G A T A A T G C A G[서열 번호 31]

cons A A C A A T G G C T [서열 번호 32]

가용성 전분 신타아제의 전분-캡슐화 영역의 사용:

식물-가용성 전분 신타아제의 cDNA 클론이 상기의 발명의 배경 부분에서 설명되어 있고 본 발명에서 사용될 수 있다. 어떠한 이와 같은 SSTS 단백질에 대한 유전자가 본 발명에 따른 구성물에서 사용될 수 있다.

분지 효소의 전분-캡슐화 영역의 사용:

식물, 세균 및 동물 분지 효소의 cDNA 클론이 상기의 발명의 배경 부분에서 설명되어 있고 본 발명에서 사용될 수 있다. 분지 효소[1,4D글루칸: 1,4D글루칸6D(1,4D글루카노) 트렌스퍼라아제 (E.C.2.4.1.18)는 아밀로스를 Q-효소로 때때로 불리워지는 아밀로펙틴으로 전환시킨다.

옥수수 분지 효소 I의 서열이 Baba등에 의해 조사되었다( Baba등(1991) BBRC, 181:87-94 참조). 옥수수 배젖으로부터의 전분 분지 효소 II가 Fisher등에 의해 조사되었다( Fisher등(1993), *Plant Physiol.* 102:1045-1046 참조).

입자-결합 전분 신타아제의 전분-결합 도메인의 사용:

식물 입자-결합 전분 신타아제의 cDNA 클론의 사용이 다음문헌[Shure등, (1983) *Cell* 35:225-233, 및 Visser등, (1989) *Plant Sci.* 64(2):185-192]에서 설명되어 있다. 또한, Visser등은 감자에 있어서 입자-결합 전분 신타아제의 발현을 항감지 구성물(antisense construct)에 의해 억제하는 것을 설명했다(Visser등, (1991), *Mol. Gen. Genetic* 225(2):289-296; (1994) *The Plant Cell* 6:43-52 참조). Shimada등은 벼에 있어서의 항감지를 보여주고 있다(1993, *Theor. Appl. Genet.* 86:665-672 참조). Van der Leij 등은 야생형 waxy 감자 유전자에 의한 형질변환후 저 아밀로스 감자에서 아밀로스 생합성의 복원을 보여주고 있다(1991, *Theor. Appl. Genet.* 82:289-295 참조).

예를 들어, 옥수수, 벼, 밀, 감자, 카사바, 완두콩 또는 보리로부터 유도한 입자 전분 신타아제의 아미노산 서열 및 누클레오티드 서열은 잘 알려져 있다. 어떠한 이와 같은 GBSTS 단백질이 본 발명에 따른 구성물에서 사용될 수 있다.

식물 형질 변환 벡터의 구성:

본 발명의 방법에서 사용하기 위한 식물 형질변환 벡터는 표준 방법을 사용하여 구성할 수 있다.

트렌싯 펩티드 서열의 사용:

일부의 유전자 구성물은 전분체내에서 정확한 국소화를 확실히 하기 위하여 전분체 트렌싯 펩티드의 존재를 필요로 한다. 엽록체 트렌싯 펩티드는 유사한 서열을 가지는 것으로 판단된다. Heijne등은 다음문헌(1991, *Plant Mol. Biol. Reporter*, 9(2):104-126)에서 엽록체 트렌싯 펩티드의 데이터베이스를 설명하고 있다. 본 발명에서 유용한 다른 트렌싯 펩티드로는, ADPG 피로포스포리아제(1991, *Plant Mol. Biol. Reporter*, 9(2):104-162 참조), 작은 서브유니트 RUBISCO, 아세토락테이트 신타아제, 글리세르알데히드3P디히드로게나아제 및 리덕타아제의 것들이 있다.

많은 게노타입으로부터의 작은 서브유니트 RUBISCO의 트렌싯 펩티드의 콘센서스 서열은 다음과 같다:

MASSMLSSAAVATRTNPAQASM VAPFTGLKSAAFPVSRKQNLDI

TSIASNGGRVQC[서열 번호 33].

옥수수의 작은 서브유니트 RUBISCO는 다음과같은 서열을 갖는다:

MAPTVMMASSATATRNPQAQS AVAPFQGLKSTASLPVARRSSR

SLGNVASNGGRIRC[서열 번호 34]

옥수수 잎의 글리세르알데하이드3P디하이드로게나아제의 트렌싯 펩티드는 다음과같은 서열을 갖는다:

MAQILAPSTQWQMRITKTSPCA TPITSKMWSSLVMKQTKVVAHS

AKFRVMAVNSENGT[서열 번호 35]

옥수수 배젖-결합 전분 신타아제의 트렌싯 펩티드 서열은 다음과같다:

MAALATSQLVATRAGHGVPDASTFRRGAAQGLRGARASAAADTLSMRTSASAAPRHQQQARRGGRFPFPSSLVVC[서열 번호 36].

옥수수 배젖의 가용성 전분 신타아제의 트렌싯 펩티드 서열은 아래와같다:

MATPSAVGAACLLLARXAWPAAVGDRARPRRLQRVLRRL[서열 번호 37].

전분-캡슐화 단백질내에 새로운 아미노산 또는 유전자의 훈입:

본 발명에서 사용된 전분 결합 단백질은 새로운 아미노산 결합물이 훈입되도록 이 분야에 통상의 지식을 가진 사람에 의해 변경될 수 있다. 예를들어, 전분-결합 단백질의 서열은 리신, 메티오닌 또는 트립신이 보통 수준보다 더 높은 수준으로 발현되도록 변경될 수 있다. 이러한 수준은 자연 수준이상으로 유용하게 상승될 수 있으며 이러한 단백질은 곡물과 같은 농작물에 영양 증진을 제공한다.

아미노산 균형을 변경시키는 것 외에도, 값어치있는 펩티드가 전분-결합 단백질내로 훈입되도록 전분-결합 단백질을 조작할 수 있다. 단백질의 N-말단에서 페이로드 폴리펩티드를 전분-결합 단백질에 부착시키면, 펩티드 단편을 추가하고 전분-결합 능력을 유지시키는 효과가 얻어진다. 그밖에, 전분-캡슐화 영역에 대한 페이로드 폴리펩티드의 부착 부위내에 특정의 프로테아제 분열 부위를 훈합시킴으로써 개선이 이루어질 수 있다. 프로테아제가 상이한 아미노산 결합에 대하여 바람직한 특이성을 가진다는 것은 당업자에게 잘 알려져있다. 이러한 특이성은 동물 및 사람의 소화관의 상이한 영역에 값어치있는 펩티드를 운반하기위한 매개물을 제공하는데 사용될 수 있다.

또한, 본 발명의 또다른 실시예에 있어서, 페이로드 폴리펩티드는 전분 입자의 정제 및 가공후에 방출될 수 있다. 전분 가수분해 및/또는 젤라틴화 과정을 사용하면, 전분 입자에 결합된 단백질은 방출될 수 있거나 단백질 가수분해에 이용될 수 있음이 알려져있다. 따라서, 전분 입자 기질로부터 상업적인 양의 단백질 및 펩티드의 회수가 가능하게된다.

여전히 본 발명의 또다른 실시예에 있어서, 전분의 소화성을 변경시키기위하여 전분 입자를 여러 가지의 상이한 방식으로 처리하는 것이 가능하다. 이러한 방법을 사용하면 전분 입자내에 갇혀있는 단백질, 펩티드 또는 아미노산의 생물학적이용효능을 변화시킬 수 있다.

전술한 발명은 명료성 및 이해의 목적을 위해 예시 및 실시예에 의해 상세히 설명되었지만, 이 분야에서 통상의 지식을 가진 당업자에게는 본 발명의 가르침에 비추어 동봉한 청구의 범위를 일탈함이없이 특정의 변경 및 변화가 가능함이 쉽게 명백해질 것이다.

## 도면의 간단한 설명

도 1a는 Stratagene로부터의 pBSK내에 서브클론화된 합성 GFP(Green Fluorescent Protein)을 함유하는 플라스미드 pEXS114를 보여준다.

도 1b는 플라스미드 pEXS115를 보여준다.

도 2a는 상업적으로 이용가능한 플라스미드내에 서브클론화된 제한 부위를 갖는 waxy 유전자를 보여준다.

도 2b는 서브클론화된 pEXS115로부터의 GFP 단편을 갖는 Novagen으로부터 상업적으로 이용가능한 pET-21A 플라스미드를 보여준다.

도 3a는 pEXSWX내에 서브클론화된 pEXS114, 및 GFP-FLWX 지도를 보여준다.

도 3b는 GFP-Bam HIWX 플라스미드를 보여준다.

도 4는 pEXSWX내에 서브클론화된 pEXS115의 SGFP 단편, 및 GFP-NcoWX 지도를 보여준다.

도 5는 외떡입식물에서 사용하기에 적합한 플라스미드의 선형도를 보여준다.

도 6은 플라스미드 pEXS52를 보여준다.

도 7은 pEXS51 및 pEXS560을 형성하기위해 사용되는 여섯 개의 도입 플라스미드를 보여주는 것으로서, 도 7a는 pEXS adh1을 보여주고, 도 7b는 pEXS adh1-nos3'을 보여주고, 도 7c는 pEXS33을 보여주고, 도 7d는 pEXS10zp를 보여주고, 도 7e는 pEXS10zp-adh1을 보여주고, 도 7f는 pEXS10zp-adh1-nos3'를 보여준다.

도 8a 및 도 8b는 각각, 전분 가용성 신타아제 유전자인 MS-SIII를 각각 함유하는 pEXS50 및 pEXS51을 보여준다.

도 9a는 pEXS50에서 보여진 인트론이 제외된 플라스미드 pEXS60을 보여주고, 도 9b는 pEXS60에서 보여진 인트론이 제외된 플라스미드 pEXS61을 보여준다.

## 서열 목록

## (1) 일반 정보:

( i )출원인: 퀼링, 페퍼, 구안, 한핑( Keeling, Peper, Guan, Hanping)

( ii )발명의 명칭: 전분 캡슐화

( iii )서열의 수: 37

( iv )서신 주소:

(A)주소: 그린리, 위너 앤드 설리반, 피.씨.(Greenlee, Winner and Sullivan, P. C.)

(B)스트리트: 5370 맨하탄 서클(5370 Manhattan Circle)

(C)시티: 보울더(Boulder)

(D)주:

(E)국명: 미합중국

(F)우편번호: 80303

## ( v )컴퓨터 판독 형태

(A) 매개 유형: 플로피 디스크

(B) 컴퓨터: IBM PC 호환성

(C) 오퍼레이팅 시스템: PC-DOS/MS-DOS

(D) 소프트웨어: 페이턴트인 릴리스(Patent In Release)#1.0, 버전 #1.30

## ( vi )현재 출원 정보:

(A) 출원 번호: 미합중국

(B)출원일: 1997년 9월 30일

(C) 분류:

## ( vii )이전의 출원 정보:

(A) 출원 번호: US60/026,855호

(B) 출원일: 1996년 9월 30일

## (viii)대리인 정보

(A)성명: 위너, 엘렌 피(Winner, Ellen P)

(B)등록 번호: 28,547

(C)참조/다켓(DOCKET)번호: 89-97

## ( ix )통신 정보:

(A)전화: (303) 499-8080

(B)팩스: (303) 499-8089

## (2)서열 번호 1에 대한 정보:

## ( i ) 서열 특징:

(A)길이: 31개 염기쌍

(B)유형: 핵산

(C)가닥수: 단일

(D)토플로지: 선형

## ( ii ) 분자 유형: 다른 핵산

(A)설명: /desc='올리고누클레오티드'

## ( iii ) 가정: 없음

## (xi) 서열 설명: 서열 번호 1:

GACTAGTCAT ATGGTGAGCA AGGGCGAGGA G

## (2)서열 번호 2에 대한 정보:

## ( i ) 서열 특징:

(A)길이: 36개 염기쌍

(B)유형: 핵산

(C)가닥수: 단일

(D)토플로지: 선형

( ii ) 분자 유형: 다른 핵산

(A)설명: /desc='올리고누클레오티드'

( iii ) 가정: 없음

(xi) 서열 설명: 서열 번호 2:

CTAGATCTTC ATATGCTTGT ACAGCTCGTC CATGCC

(2)서열 번호 3에 대한 정보:

( i ) 서열 특징:

(A)길이: 39개 염기쌍

(B)유형: 핵산

(C)가닥수: 단일

(D)토플로지: 선형

( ii ) 분자 유형: 다른 핵산

(A)설명: /desc='올리고누클레오티드'

( iii ) 가정: 없음

(xi) 서열 설명: 서열 번호 3:

CTAGATCTTG GCCATGGCCT TGTACAGCTC GTCCATGCC

(2)서열 번호 4에 대한 정보:

( i ) 서열 특징:

(A)길이: 4800개 염기쌍

(B)유형: 핵산

(C)가닥수: 이중

(D)토플로지: 관련없음

( ii ) 분자 유형: DNA (게놈)

( iii ) 가정: 없음

(vi) 최초 소오스:

(A)유기체: Zea mays

( ix )특징:

(A) 명칭/키: CDS

(B) 위치: 결합(1449..1553, 1685..1765, 1860..1958, 2055..2144..2226

..2289..2413..2513, 2651..2760, 2858..3101, 3212..3394,

3490..3681, 3793..3879, 3977..4105, 4227..4343)

(xi) 서열 설명: 서열 번호 4:

CAGCGACCTA TTACACAGCC CGCTCGGGCC CGCGACGTGCG GGACACATCT TCTTCCCCCT	60
TTTGGTGAAG CTCTGCTCGC AGCTGTCCGG CTCCTGGAC GTTCGTGTGG CAGATTCACTC	120
TGTTGCTCG TCTCCTGTGC TTCTGGTA GCTTGTGTAG TGGAGCTGAC ATGGTCTGAG	180
CAGGCTTAAA ATTTGCTCGT AGACGAGGAG TACCAGCACA GCACGTTGCC GATTCTCTG	240
CCTGTGAAGT GCAACGTCTA GGATTGTAC ACACGCTTGGT CGCGTCGCGT CGCGTCGCGT	300
CGATGCGGTG GTGAGCAGAG CAGCAACAGC TGGGCGGCC AACGTTGGCT TCCGTGTCTT	360
CGTCGTACGT ACACGCGCGC CGGGGACACG CAGCAGAGAG CGGAGAGCGA GCCGTGCACG	420
GGGAGGTGGT GTGGAAGTGG AGCCGCGCGC CCGGCCGCC GCGCCCGGTG GGCAACCCAA	480
AAGTACCCAC GACAAGCGAA GGCGCCAAAG CGATCCAAGC TCCGGAACGC AACAGCATGC	540
GTCGCGTCGG AGAGCCAGCC ACAAGCAGCC GAGAACCGAA CCGGTGGCG ACACGTCATG	600
GGACGGACGC GGGCGACGCT TCCAAACGGG CCACGTACGC CGCGTGTGC GTCGTGCAG	660
ACGACAAGCC AAGGGAGGC AGCCCCCGAT CGGGAAAGCG TTTTGGCGC GAGCGCTGGC	720
GTGCGGGTCA GTCGCTGGT CGCAGTGCCTG GGGGGAACGG GTATCGTGGG GGGCGCGGGC	780
GGAGGAGAGC GTGGCGAGGG CCGAGAGCAG CGCGCGGCCG GGTCACGCAA CGCGCCCCAC	840
GTACTGCCCT CCCCCCTCCGC GCGCGCTAGA AATACCGAGG CCTGGACCGG GGGGGGGCCC	900
CGTCACATCC ATCCATCGAC CGATCGATCG CCACAGCCAA CACCAACCGC CGAGGCGACG	960
CGACAGCCGC CAGGAGGAAG GAATAAAACTC ACTGCCAGCC AGTGAAGGGG GAGAAGTGT	1020
CTGCTCCGTC GACCACTGCG CGCACCGCCC GGCAGGGCTG CTCATCTCGT CGACGACCAG	1080
GTTCTGTTCC GTTCCGATCC GATCCGATCC TGTCTTGAG TTTCGTCCAG ATCCTGGCGC	1140
GTATCTGCGT GTTGATGAT CCAGGTTCTT CGAACCTAAA TCTGTCCGTG CACACGTCTT	1200
TTCTCTCTCT CCTACGCAGT GGATTAATCG GCATGGCGGC TCTGGCCACG TCGCAGCTCG	1260
TCGCAACCGG CGCCGGGCGTGG GGGCTCCCGG ACGCGTCCAC GTTCCGCCGC GGGCCGCCGC	1320
AGGGCCTGAG GGGGGCCCGG GCGTCGGCGG CGGCGGACAC GCTCAGCATG CGGACCAGCG	1380
CGCGCGCGGC GCCCAGGCAC CAGCAGCAGG CGCGCGCGGG GGGCAGGTTTC CGCTCGCTCG	1440
TCGTGTGC GCC AGC GCC GGC ATG AAC GTC GTC TTC GTC GGC GCC GAG ATG Ala Ser Ala Gly Met Asn Val Val Phe Val Gly Ala Glu Met	1490
1 5 10	
GCG CCG TGG ACC AAG ACC GGC GGC CTC GGC GAC GTC CTC GGC GGC CTG	1538

Ala Pro Trp Ser Lys Thr Gly Gly Leu Gly Asp Val Leu Gly Gly Leu			
15	20	25	30
CCG CCG GCC ATG GCC GTAAGCGCGC GCACCGAGAC ATGCATCCGT TGGATCGCGT	1593		
Pro Pro Ala Met Ala			
35			
CTTCTTCGTG CTCTTGCAGC GTGCATGATG CATGTGTTTC CTCCTGGCTT GTGTTCGTGT	1653		
ATGTGACGTG TTTGTTGGG CATGCATGCA G GCG AAC GGG CAC CGT GTC ATG	1705		
Ala Asn Gly His Arg Val Met			
40			
GTC GTC TCT CCC CGC TAC GAC CAG TAC AAG GAC GCC TGG GAC ACC AGC	1753		
Val Val Ser Pro Arg Tyr Asp Gln Tyr Lys Asp Ala Trp Asp Thr Ser			
45	50	55	
GTC GTG TCC GAG GTACGGCCAC CGAGACCAGA TTCAGATCAC AGTCACACAC	1805		
Val Val Ser Glu			
60			
ACCGTCATAT GAACCTTCT CTGCTCTGAT GCCTGCACT GCAAATGCAT GCAG ATC	1862		
Ile			
AAG ATG GGA GAC GGG TAC GAG ACG GTC AGG TTC TTC CAC TGC TAC AAG	1910		
Lys Met Gly Asp Gly Tyr Glu Thr Val Arg Phe Phe His Cys Tyr Lys			
65	70	75	
CGC GGA GTG GAC CGC GTG TTC GTT GAC CAC CCA CTG TTC CTG GAG AGG	1958		
Arg Gly Val Asp Arg Val Phe Val Asp His Pro Leu Phe Leu Glu Arg			
80	85	90	95
GTGAGACGAG ATCTGATCAC TCGATACGCA ATTACCAACC CATTGTAAGC AGTTACAGTG	2018		
AGCTTTTTT CCCCCGGCC TGGTCGCTGG TTTCAG GTT TGG GGA AAG ACC GAG	2072		
Val Trp Gly Lys Thr Glu			
100			
GAG AAG ATC TAC GGG CCT GTC GCT GGA ACG GAC TAC AGG GAC AAC CAG	2120		
Glu Lys Ile Tyr Gly Pro Val Ala Gly Thr Asp Tyr Arg Asp Asn Gln			
105	110	115	
CTG CGG TTC AGC CTG CTA TGC CAG GTCAGGATGG CTTGGTACTA CAACTTCATA	2174		
Leu Arg Phe Ser Leu Leu Cys Gln			

120

125

TCATCTGTAT GCAGCAGTAT ACACTGATGA GAAATGCATG CTGTTCTGCA G GCA GCA	2231
Ala Ala	
CTT GAA GCT CCA AGG ATC CTG AGC CTC AAC AAC AAC CCA TAC TTC TCC	2279
Leu Glu Ala Pro Arg Ile Leu Ser Leu Asn Asn Asn Pro Tyr Phe Ser	
130 135 140	
GGA CCA TAC G GTAAGAGTTG CAGTCTTCGT ATATATATCT GTTGAGCTCG	2329
Gly Pro Tyr	
145	
AGAATCTTCA CAGGAAGCGG CCCATCAGAC GGACTGTCAT TTTACACTGA CTACTGCTGC	2389
TGCTCTTCGT CCATCCATAC AAG GG GAG GAC GTC GTG TTC GTC TGC AAC	2438
Gly Glu Asp Val Val Phe Val Cys Asn	
150 155	
GAC TGG CAC ACC GGC CCT CTC TCG TGC TAC CTC AAG AGC AAC TAC CAG	2486
Asp Trp His Thr Gly Pro Leu Ser Cys Tyr Leu Lys Ser Asn Tyr Gln	
160 165 170	
TCC CAC GGC ATC TAC AGG GAC GCA AAG GTTGCCTTCT CTGAAGTGAA	2533
Ser His Gly Ile Tyr Arg Asp Ala Lys	
175 180	
CAACGCCGTT TTCGTTCTCC ATGCTCGTAT ATACCTCGTC TGGTAGTGGT GGTGCTTCTC	2593
TGAGAAACTA ACTGAAACTG ACTGCATGTC TGTCTGACCA TCTTCACGTA CTACCAAG	2650
ACC GCT TTC TGC ATC CAC AAC ATC TCC TAC CAG GGC CGG TTC GCC TTC	2698
Thr Ala Phe Cys Ile His Asn Ile Ser Tyr Gln Gly Arg Phe Ala Phe	
185 190 195	
TCC GAC TAC CCG GAG CTG AAC CTC CCG GAG AGA TTC AAG TCG TCC TTC	2746
Ser Asp Tyr Pro Glu Leu Asn Leu Pro Glu Arg Phe Lys Ser Ser Phe	
200 205 210	
GAT TTC ATC GAC GG GTCTGTTTC CTGCGTGCAT GTGAACATTC ATGAATGGTA	2800
Asp Phe Ile Asp Gly	
215	
ACCCACAACT GTTCGCGTCC TGCTGGTTCA TTATCTGACC TGATTGCATT ATTGCAG C	2858

TAC GAG AAG CCC GTG GAA GGC CGG AAG ATC AAC TGG ATG AAG GCC GGG Tyr Glu Lys Pro Val Glu Gly Arg Lys Ile Asn Trp Met Lys Ala Gly 220 225 230	2906
ATC CTC GAG GCC GAC AGG GTC CTC ACC GTC AGC CCC TAC TAC GCC GAG Ile Leu Glu Ala Asp Arg Val Leu Thr Val Ser Pro Tyr Tyr Ala Glu 235 240 245	2954
GAG CTC ATC TCC GGC ATC GCC AGG GGC TGC GAG CTC GAC AAC ATC ATG Glu Leu Ile Ser Gly Ile Ala Arg Gly Cys Glu Leu Asp Asn Ile Met 250 255 260 265	3002
CGC CTC ACC GGC ATC ACC GGC ATC GTC AAC GGC ATG GAC GTC AGC GAG Arg Leu Thr Gly Ile Thr Gly Ile Val Asn Gly Met Asp Val Ser Glu 270 275 280	3050
TGG GAC CCC AGC AGG GAC AAG TAC ATC GCC GTG AAG TAC GAC GTG TCG Trp Asp Pro Ser Arg Asp Lys Tyr Ile Ala Val Lys Tyr Asp Val Ser 285 290 295	3098
ACG GTGAGCTGGC TAGCTCTGAT TCTGCTGCCT GGTCCCTCCTG CTCATCATGC Thr	3151
TGGTTCGGTA CTGACGCGGC AAGTGTACGT ACGTGCGTGC GACGGTGGTG TCCGGTTCAAG	3211
GCC GTG GAG GCC AAG GCG CTG AAC AAG GAG GCG CTG CAG GCG GAG GTC Ala Val Glu Ala Lys Ala Leu Asn Lys Glu Ala Leu Gln Ala Glu Val 300 305 310	3259
GGG CTC CCG GTG GAC CGG AAC ATC CCG CTG GTG GCG TTC ATC GGC AGG Gly Leu Pro Val Asp Arg Asn Ile Pro Leu Val Ala Phe Ile Gly Arg 315 320 325 330	3307
CTG GAA GAG CAG AAG GGC CCC GAC GTC ATG GCG GCC GCG ATC CCG CAG Leu Glu Glu Gln Lys Gly Pro Asp Val Met Ala Ala Ala Ile Pro Gln 335 340 345	3355
CTC ATG GAG ATG GTG GAG GAC GTG CAG ATC GTT CTG CTG GTACGTGTGC Leu Met Glu Met Val Glu Asp Val Gln Ile Val Leu Leu 350 355	3404
GCCGGCCGCC ACCCGGCTAC TACATGCGTG TATCGTTCGT TCTACTGGAA CATGGTGTG	3464
AGCAACGCGA TGGATAATGC TGCAG GGC ACG GGC AAG AAG AAG TTC GAG CGC	3516

Gly Thr Gly Lys Lys Lys Phe Glu Arg		
360	365	
ATG CTC ATG AGC GCC GAG GAG AAG TTC CCA GGC AAG GTG CGC GCC GTG	3564	
Met Leu Met Ser Ala Glu Glu Lys Phe Pro Gly Lys Val Arg Ala Val		
370	375	380
GTC AAG TTC AAC GCG GCG CTG GCG CAC CAC ATC ATG GCC GGC GCC GAC	3612	
Val Lys Phe Asn Ala Ala Leu Ala His His Ile Met Ala Gly Ala Asp		
385	390	395
GTG CTC GCC GTC ACC AGC CGC TTC GAG CCC TGC GGC CTC ATC CAG CTG	3660	
Val Leu Ala Val Thr Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln Leu		
405	410	415
CAG GGG ATG CGA TAC CGA ACG GTACGAGAGA AAAAAAAAT CCTGAATCCT	3711	
Gln Gly Met Arg Tyr Gly Thr		
420		
GACGGAGGG ACAGAGACAG ATTATGAATG CTTCATCGAT TTGAATTGAT TGATCGATGT	3771	
CTCCCGCTGC GACTCTTGCA G CCC TGC GCC TGC GCG TCC ACC GGT GGA CTC	3822	
Pro Cys Ala Cys Ala Ser Thr Gly Gly Leu		
425	430	
GTC GAC ACC ATC ATC GAA GGC AAG ACC GGG TTC CAC ATG GGC CGC CTC	3870	
Val Asp Thr Ile Ile Glu Gly Lys Thr Gly Phe His Met Gly Arg Leu		
435	440	445
AGC GTC GAC GTAAGCCTAG CTCTGCCATG TTCTTTCTTC TTTCTTCCTG	3919	
Ser Val Asp		
450		
TATGTATGTA TGAATCAGCA CCGCCGTTCT TGTGTCGTCG TCGTCCTCTC TTCCCAG	3976	
TGT AAC GTC GTG GAG CCG GCG GAC GTC AAG AAG GTG GCC ACC ACA TTG	4024	
Cys Asn Val Val Glu Pro Ala Asp Val Lys Lys Val Ala Thr Thr Leu		
455	460	465
CAG CGC GCC ATC AAG GTG GTC GGC ACG CCG GCG TAC GAG GAG ATG GTG	4072	
Gln Arg Ala Ile Lys Val Val Gly Thr Pro Ala Tyr Glu Glu Met Val		
470	475	480
AGG AAC TGC ATG ATC CAG GAT CTC TCC TGG AAG GTACGTACGC CCGCCCCGCC	4125	
Arg Asn Cys Met Ile Gln Asp Leu Ser Trp Lys		

485	490	495		
CCGGCCCCGCC	AGAGCAGAGC	GCCAAGATCG	ACCGATCGAC CGACCACACG TACGCGCCTC	4185
GCTCCTGTCG	CTGACCGTGG	TTTAATTGTC	GAAATGCGCA G GGC CCT GCC AAG	4238
			Gly Pro Ala Lys	
AAC TGG GAG AAC GTG CTG CTC AGC CTC GGG GTC GCC GGC GGC GAG CCA				4286
Asn Trp Glu Asn Val Leu Leu Ser Leu Gly Val Ala Gly Gly Glu Pro				
500	505	510	515	
GGG GTC GAA GGC GAG GAG ATC GCG CCG CTC GCC AAG GAG AAC GTG GCC				4334
Gly Val Glu Gly Glu Ile Ala Pro Leu Ala Lys Glu Asn Val Ala				
520	525	530		
GCG CCC TGA AGAGTTCGGC CTGCAGGGCC CCTGATCTCG CGCGTGGTGC				4383
Ala Pro *				
AAAGATGTTG GGACATCTTC TTATATATGC TGTTTCGTTT ATGTGATATG GACAAGTATG				4443
TGTAGCTGCT TGCTTGTGCT AGTGTAAATGT AGTGTAGTGG TGGCCAGTGG CACAACCTAA				4503
TAAGCCCATG AACTAATTGC TTGCGTGTGT AGTTAAGTAC CGATCGGTAA TTTTATATTG				4563
CGAGTAAATA AATGGACCTG TAGTGGTGGA GTAAATAATC CCTGCTGTTG GGTGTTCTTA				4623
TCGCTCCTCG TATAGATATT ATATAGAGTA CATTTCCTC TCTCTGAATC CTACGTTTGT				4683
GAAATTCTA TATCATTACT GTAAAATTTC TGCGTTCCAA AAGAGACCCT AGCCTATCTT				4743
TGGCCCTGTT TGTTTCGGCT TCTGGCAGCT TCTGGCCACC AAAAGCTGCT GCGGACT				4800

(2) 서열 번호 5에 대한 정보:

- (i) 서열 특징:
  - (A)길이: 534개 아미노산
  - (B)유형: 아미노
  - (D)토플로지: 선형
- (ii) 문자 유형: 단백질

(xi) 서열 설명: 서열 번호 5:

Ala Ser Ala Gly Met Asn Val Val Phe Val Gly Ala Glu Met Ala Pro  
 1 5 10 15

Trp Ser Lys Thr Gly Gly Asp Val Leu Gly Gly Leu Pro Pro  
 20 25 30

Ala Met Ala Ala Asn Gly His Arg Val Met Val Val Ser Pro Arg Tyr  
 35 40 45

Asp Gln Tyr Lys Asp Ala Trp Asp Thr Ser Val Val Ser Glu Ile Lys  
 50 55 60

Met Gly Asp Gly Tyr Glu Thr Val Arg Phe Phe His Cys Tyr Lys Arg  
 65 70 75 80

Gly Val Asp Arg Val Phe Val Asp His Pro Leu Phe Leu Glu Arg Val  
 85 90 95

Trp Gly Lys Thr Glu Glu Lys Ile Tyr Gly Pro Val Ala Gly Thr Asp  
 100 105 110

Tyr Arg Asp Asn Gln Leu Arg Phe Ser Leu Leu Cys Gln Ala Ala Leu  
 115 120 125

Glu Ala Pro Arg Ile Leu Ser Leu Asn Asn Asn Pro Tyr Phe Ser Gly  
 130 135 140

Pro Tyr Gly Glu Asp Val Val Phe Val Cys Asn Asp Trp His Thr Gly  
 145 150 155 160

Pro Leu Ser Cys Tyr Leu Lys Ser Asn Tyr Gln Ser His Gly Ile Tyr  
 165 170 175

Arg Asp Ala Lys Thr Ala Phe Cys Ile His Asn Ile Ser Tyr Gln Gly  
 180 185 190

Arg Phe Ala Phe Ser Asp Tyr Pro Glu Leu Asn Leu Pro Glu Arg Phe  
 195 200 205

Lys Ser Ser Phe Asp Phe Ile Asp Gly Tyr Glu Lys Pro Val Glu Gly  
 210 215 220

Arg Lys Ile Asn Trp Met Lys Ala Gly Ile Leu Glu Ala Asp Arg Val  
 225 230 235 240

Leu Thr Val Ser Pro Tyr Tyr Ala Glu Glu Leu Ile Ser Gly Ile Ala  
 245 250 255

Arg Gly Cys Glu Leu Asp Asn Ile Met Arg Leu Thr Gly Ile Thr Gly  
 260 265 270

Ile Val Asn Gly Met Asp Val Ser Glu Trp Asp Pro Ser Arg Asp Lys  
 275 280 285

Tyr Ile Ala Val Lys Tyr Asp Val Ser Thr Ala Val Glu Ala Lys Ala  
 290 295 300

Leu Asn Lys Glu Ala Leu Gln Ala Glu Val Gly Leu Pro Val Asp Arg  
 305 310 315 320

Asn Ile Pro Leu Val Ala Phe Ile Gly Arg Leu Glu Glu Gln Lys Gly  
 325 330 335

Pro Asp Val Met Ala Ala Ala Ile Pro Gln Leu Met Glu Met Val Glu  
 340 345 350

Asp Val Gln Ile Val Leu Leu Gly Thr Gly Lys Lys Lys Phe Glu Arg  
 355 360 365

Met Leu Met Ser Ala Glu Glu Lys Phe Pro Gly Lys Val Arg Ala Val  
 370 375 380

Val Lys Phe Asn Ala Ala Leu Ala His His Ile Met Ala Gly Ala Asp  
 385 390 395 400

Val Leu Ala Val Thr Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln Leu  
 405 410 415

Gln Gly Met Arg Tyr Gly Thr Pro Cys Ala Cys Ala Ser Thr Gly Gly  
 420 425 430

Leu Val Asp Thr Ile Ile Glu Gly Lys Thr Gly Phe His Met Gly Arg  
 435 440 445

Leu Ser Val Asp Cys Asn Val Val Glu Pro Ala Asp Val Lys Lys Val  
 450 455 460

Ala Thr Thr Leu Gln Arg Ala Ile Lys Val Val Gly Thr Pro Ala Tyr  
 465 470 475 480

Glu Glu Met Val Arg Asn Cys Met Ile Gln Asp Leu Ser Trp Lys Gly  
 485 490 495

Pro Ala Lys Asn Trp Glu Asn Val Leu Leu Ser Leu Gly Val Ala Gly  
 500 505 510

Gly Glu Pro Gly Val Glu Gly Glu Ile Ala Pro Leu Ala Lys Glu  
 515 520 525

Asn Val Ala Ala Pro \*  
 530

(2) 서열 번호 6에 대한 정보:

(i) 서열 특징:

- (A)길이: 2542개 염기쌍
- (B)유형: 핵산
- (C)가닥수: 이중
- (D)토플로지: 관련없음

(ii) 분자 유형: cDNA에서 mRNA

( iii ) 가정: 없음

( vi ) 최초 소오스:

(A)유기체: Oryza sativa

( ix )특징:

(A) 명칭/키: CDS

(B) 위치: 453..2282

(xi) 서열 설명: 서열 번호 6:

GAATTCAAGTG TGAAGGAATA GATTCTCTTC AAAACAATTT AATCATTCACTTGATCTGCT	60
CAAAGCTCTG TGCATCTCCG GGTGCAACGG CCAGGATATT TATTGTGCAG TAAAAAAATG	120
TCATATCCCC TAGCCACCCA AGAAACTGCT CCTTAAGTCC TTATAAGCAC ATATGGCATT	180
GTAATATATA TGTTTGAGTT TTAGCGACAA TTTTTTTAAA AACTTTTGGT CCTTTTTATG	240
AACGTTTTAA GTTTCAGTGT CTTTTTTTTT CGAATTTAA ATGTAGCTTC AAATTCTAAT	300
CCCCAATCCA AATTGTAATA AACTTCATT CTCCTAATTA ACATCTTAAT TCATTTATT	360

GAAAACCACT G TCAAATTCTT TTTAGGCTCA CCAAACCTTA AACAAATTCAA TTCAGTGCAG	420
AGATCTTCCA CAGCAACAGC TAGACAACCA CC ATG TCG GCT CTC ACC ACC ACG TCC Met Ser Ala Leu Thr Thr Ser	473
535 540	
CAG CTC GCC ACC TCG GCC ACC GGC TTC GGC ATC GCC GAC AGG TCG GCG Gln Leu Ala Thr Ser Ala Thr Gly Ile Ala Asp Arg Ser Ala	521
545 550 555	
CCG TCG TCG CTG CTC CGC CAC GGG TTC CAG GGC CTC AAG CCC CGC AGC Pro Ser Ser Leu Leu Arg His Gly Phe Gln Gly Leu Lys Pro Arg Ser	569
560 565 570	
CCC GCC GGC GGC GAC GCG ACG TCG CTC AGC GTG ACG ACC AGC GCG CGC Pro Ala Gly Gly Asp Ala Thr Ser Leu Ser Val Thr Thr Ser Ala Arg	617
575 580 585	
GCG ACG CCC AAG CAG CAG CGG TCG GTG CAG CGT GGC AGC CGG AGG TTC Ala Thr Pro Lys Gln Gln Arg Ser Val Gln Arg Gly Ser Arg Arg Phe	665
590 595 600 605	
CCC TCC GTC GTC GTG TAC GCC ACC GGC GCC GGC ATG AAC GTC GTG TTC Pro Ser Val Val Val Tyr Ala Thr Gly Ala Gly Met Asn Val Val Phe	713
610 615 620	
GTC GGC GCC GAG ATG GCC CCC TGG AGC AAG ACC GGC GGC CTC GGT GAC Val Gly Ala Glu Met Ala Pro Trp Ser Lys Thr Gly Gly Leu Gly Asp	761
625 630 635	
GTC CTC GGT GGC CTC CCC CCT GCC ATG GCT GCG AAT GGC CAC AGG GTC Val Leu Gly Leu Pro Pro Ala Met Ala Ala Asn Gly His Arg Val	809
640 645 650	
ATG GTG ATC TCT CCT CGG TAC GAC CAG TAC AAG GAC GCT TGG GAT ACC Met Val Ile Ser Pro Arg Tyr Asp Gln Tyr Lys Asp Ala Trp Asp Thr	857
655 660 665	
AGC GTT GTG GCT GAG ATC AAG GTT GCA GAC AGG TAC GAG AGG GTG AGG Ser Val Val Ala Glu Ile Lys Val Ala Asp Arg Tyr Glu Arg Val Arg	905
670 675 680 685	
TTT TTC CAT TGC TAC AAG CGT GGA GTC GAC CGT GTG TTC ATC GAC CAT Phe Phe His Cys Tyr Lys Arg Gly Val Asp Arg Val Phe Ile Asp His	953
690 695 700	

CCG TCA TTC CTG GAG AAG GTT TGG GGA AAG ACC GGT GAG AAG ATC TAC Pro Ser Phe Leu Glu Lys Val Trp Gly Lys Thr Gly Glu Lys Ile Tyr 705 710 715	1001
GGA CCT GAC ACT GGA GTT GAT TAC AAA GAC AAC CAG ATG CGT TTC AGC Gly Pro Asp Thr Gly Val Asp Tyr Lys Asp Asn Gln Met Arg Phe Ser 720 725 730	1049
CTT CTT TGC CAG GCA GCA CTC GAG CCT AGG ATC CTA AAC CTC AAC Leu Leu Cys Gln Ala Ala Leu Glu Ala Pro Arg Ile Leu Asn Leu Asn 735 740 745	1097
AAC AAC CCA TAC TTC AAA GGA ACT TAT GGT GAG GAT GTT GTG TTC GTC Asn Asn Pro Tyr Phe Lys Gly Thr Tyr Gly Glu Asp Val Val Phe Val 750 755 760 765	1145
TGC AAC GAC TGG CAC ACT GGC CCA CTG GCG AGC TAC CTG AAG AAC AAC Cys Asn Asp Trp His Thr Gly Pro Leu Ala Ser Tyr Leu Lys Asn Asn 770 775 780	1193
TAC CAG CCC AAT GGC ATC TAC AGG AAT GCA AAG GTT GCT TTC TGC ATC Tyr Gln Pro Asn Gly Ile Tyr Arg Asn Ala Lys Val Ala Phe Cys Ile 785 790 795	1241
CAC AAC ATC TCC TAC CAG GGC CGT TTC GCT TTC GAG GAT TAC CCT GAG His Asn Ile Ser Tyr Gln Gly Arg Phe Ala Phe Glu Asp Tyr Pro Glu 800 805 810	1289
CTG AAC CTC TCC GAG AGG TTC AGG TCA TCC TTC GAT TTC ATC GAC GGG Leu Asn Leu Ser Glu Arg Phe Arg Ser Ser Phe Asp Phe Ile Asp Gly 815 820 825	1337
TAT GAC ACG CCG GTG GAG GGC AGG AAG ATC AAC TGG ATG AAG GCC GGA Tyr Asp Thr Pro Val Glu Gly Arg Lys Ile Asn Trp Met Lys Ala Gly 830 835 840 845	1385
ATC CTG GAA GCC GAC AGG GTG CTC ACC GTG AGC CCG TAC TAC GCC GAG Ile Leu Glu Ala Asp Arg Val Leu Thr Val Ser Pro Tyr Tyr Ala Glu 850 855 860	1433
GAG CTC ATC TCC GGC ATC GCC AGG GGA TGC GAG CTC GAC AAC ATC ATG Glu Leu Ile Ser Gly Ile Ala Arg Gly Cys Glu Leu Asp Asn Ile Met 865 870 875	1481
CGG CTC ACC GGC ATC ACC GGC ATC GTC AAC GGC ATG GAC GTC AGC GAG	1529

Arg Leu Thr Gly Ile Thr Gly Ile Val Asn Gly Met Asp Val Ser Glu			
880	885	890	
TGG GAT CCT AGC AAG GAC AAG TAC ATC ACC GCC AAG TAC GAC GCA ACC			1577
Trp Asp Pro Ser Lys Asp Lys Tyr Ile Thr Ala Lys Tyr Asp Ala Thr			
895	900	905	
ACG GCA ATC GAG GCG AAG GCG CTG AAC AAG GAG GCG TTG CAG GCG GAG			1625
Thr Ala Ile Glu Ala Lys Ala Leu Asn Lys Glu Ala Leu Gln Ala Glu			
910	915	920	925
GCG GGT CTT CCG GTC GAC AGG AAA ATC CCA CTG ATC GCG TTC ATC GGC			1673
Ala Gly Leu Pro Val Asp Arg Lys Ile Pro Leu Ile Ala Phe Ile Gly			
930	935	940	
AGG CTG GAG GAA CAG AAG GGC CCT GAC GTC ATG GCC GCC GCC ATC CCG			1721
Arg Leu Glu Glu Gln Lys Gly Pro Asp Val Met Ala Ala Ile Pro			
945	950	955	
GAG CTC ATG CAG GAG GAC GTC CAG ATC GTT CTT CTG GGT ACT GGA AAG			1769
Glu Leu Met Gln Glu Asp Val Gln Ile Val Leu Leu Gly Thr Gly Lys			
960	965	970	
AAG AAG TTC GAG AAG CTG CTC AAG AGC ATG GAG GAG AAG TAT CCG GGC			1817
Lys Lys Phe Glu Lys Leu Leu Lys Ser Met Glu Glu Lys Tyr Pro Gly			
975	980	985	
AAG GTG AGG GCG GTG GTG AAG TTC AAC GCG CCG CTT GCT CAT CTC ATC			1865
Lys Val Arg Ala Val Val Lys Phe Asn Ala Pro Leu Ala His Leu Ile			
990	995	1000	1005
ATG GCC GGA GCC GAC GTG CTC GCC GTC CCC AGC CGC TTC GAG CCC TGT			1913
Met Ala Gly Ala Asp Val Leu Ala Val Pro Ser Arg Phe Glu Pro Cys			
1010	1015	1020	
GGA CTC ATC CAG CTG CAG GGG ATG AGA TAC GGA ACG CCC TGT GCT TGC			1961
Gly Leu Ile Gln Leu Gln Gly Met Arg Tyr Gly Thr Pro Cys Ala Cys			
1025	1030	1035	
GCG TCC ACC GGT GGG CTC GTG GAC ACG GTC ATC GAA GGC AAG ACT GGT			2009
Ala Ser Thr Gly Gly Leu Val Asp Thr Val Ile Glu Gly Lys Thr Gly			
1040	1045	1050	
TTC CAC ATG GCC CGT CTC AGC GTC GAC TGC AAG GTG GTG GAG CCA AGC			2057
Phe His Met Gly Arg Leu Ser Val Asp Cys Lys Val Val Glu Pro Ser			

1055	1060	1065	
GAC GTG AAG AAG GTG GCG GCC ACC CTG AAG CGC GCC ATC AAG GTC GTC			2105
Asp Val Lys Lys Val Ala Ala Thr Leu Lys Arg Ala Ile Lys Val Val			
1070	1075	1080	1085
GGC ACG CCG GCG TAC GAG GAG ATG GTC AGG AAC TGC ATG AAC CAG GAC			2153
Gly Thr Pro Ala Tyr Glu Glu Met Val Arg Asn Cys Met Asn Gln Asp			
1090	1095	1100	
CTC TCC TGG AAG GGG CCT GCG AAG AAC TGG GAG AAT GTG CTC CTG GGC			2201
Leu Ser Trp Lys Gly Pro Ala Lys Asn Trp Glu Asn Val Leu Leu Gly			
1105	1110	1115	
CTG GGC GTC GCC GGC AGC GCG CCG GGG ATC GAA GGC GAC GAG ATC GCG			2249
Leu Gly Val Ala Gly Ser Ala Pro Gly Ile Glu Gly Asp Glu Ile Ala			
1120	1125	1130	
CCG CTC GCC AAG GAG AAC GTG GCT GCT CCT TGA AGAGCCTGAG ATCTACATAT			2302
Pro Leu Ala Lys Glu Asn Val Ala Ala Pro *			
1135	1140		
GGAGTGATTA ATTAATATAG CAGTATATGG ATGAGAGACG AATGAACCAAG TGGTTTGT			2362
GTTGTAGTGA ATTTGTAGCT ATAGCCAATT ATATAGGCTA ATAAGTTGA TGTTGTACTC			2422
TTCTGGGTGT GCTTAAGTAT CTTATCGGAC CCTGAATTAA TGTGTGTGGC TTATTGCCAA			2482
TAATATTAAG TAATAAAGGG TTTATTATAT TATTATATAT GTTATATTAT ACTAAAAAAA			2542

(2) 서열 번호 7에 대한 정보:

( i ) 서열 특징:

- (A)길이: 610개 아미노산
- (B)유형: 아미노산
- (D)토플로지: 선형

( ii ) 문자 유형: 단백질

(xi) 서열 설명: 서열 번호 7:

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

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

Tyr Gly Thr Pro Cys Ala Cys Ala Ser Thr Gly Gly Leu Val Asp Thr  
 500 505 510  
 Val Ile Glu Gly Lys Thr Gly Phe His Met Gly Arg Leu Ser Val Asp  
 515 520 525  
 Cys Lys Val Val Glu Pro Ser Asp Val Lys Lys Val Ala Ala Thr Leu  
 530 535 540  
 Lys Arg Ala Ile Lys Val Val Gly Thr Pro Ala Tyr Glu Glu Met Val  
 545 550 555 560  
 Arg Asn Cys Met Asn Gln Asp Leu Ser Trp Lys Gly Pro Ala Lys Asn  
 565 570 575  
 Trp Glu Asn Val Leu Leu Gly Leu Gly Val Ala Gly Ser Ala Pro Gly  
 580 585 590  
 Ile Glu Gly Asp Glu Ile Ala Pro Leu Ala Lys Glu Asn Val Ala Ala  
 595 600 605  
 Pro \*

610

(2) 서열 번호 8에 대한 정보:

(i) 서열 특징:

- (A) 길이: 2007개 염기쌍
- (B) 유형: 핵산
- (C) 가닥수: 이중
- (D) 토플로지: 관련없음

(ii) 분자 유형: cDNA에서 mRNA

(iii) 가정: 없음

(vi) 최초 소오스:

(A) 유기체: Zea mays

(ix) 특징:

- (A) 명칭/키: CDS
- (B) 위치: 1..2007

(xi) 서열 설명: 서열 번호 8:

GCT GAG GCT GAG GCC GGG GGC AAG GAC GCG CCG CCG GAG AGG AGC GGC Ala Glu Ala Glu Ala Gly Gly Lys Asp Ala Pro Pro Glu Arg Ser Gly 615 620 625	48
GAC GCC GCC AGG TTG CCC CGC GCT CGG CGC AAT GCG GTC TCC AAA CGG Asp Ala Ala Arg Leu Pro Arg Ala Arg Arg Asn Ala Val Ser Lys Arg 630 635 640	96
AGG GAT CCT CTT CAG CCG GTC GGC CGG TAC GGC TCC GCG ACG GGA AAC Arg Asp Pro Leu Gln Pro Val Gly Arg Tyr Gly Ser Ala Thr Gly Asn 645 650 655	144
ACG GCC AGG ACC GGC GCC GCG TCC TGC CAG AAC GCC GCA TTG GCG GAC Thr Ala Arg Thr Gly Ala Ala Ser Cys Gln Asn Ala Ala Leu Ala Asp 660 665 670	192
GTT GAG ATC GTT GAG ATC AAG TCC ATC GTC GCC GCG CCG CCG ACG AGC Val Glu Ile Val Glu Ile Lys Ser Ile Val Ala Ala Pro Pro Thr Ser 675 680 685 690	240
ATA GTG AAG TTC CCA GGG CGC GGG CTA CAG GAT GAT CCT TCC CTC TGG Ile Val Lys Phe Pro Gly Arg Gly Leu Gln Asp Asp Pro Ser Leu Trp 695 700 705	288
GAC ATA GCA CCG GAG ACT GTC CTC CCA GCC CCG AAG CCA CTG CAT GAA Asp Ile Ala Pro Glu Thr Val Leu Pro Ala Pro Lys Pro Leu His Glu 710 715 720	336
TCG CCT GCG GTT GAC GGA GAT TCA AAT GGA ATT GCA CCT CCT ACA GTT Ser Pro Ala Val Asp Gly Asp Ser Asn Gly Ile Ala Pro Pro Thr Val 725 730 735	384
GAG CCA TTA GTA CAG GAG GCC ACT TGG GAT TTC AAG AAA TAC ATC GGT Glu Pro Leu Val Gln Glu Ala Thr Trp Asp Phe Lys Lys Tyr Ile Gly 740 745 750	432
TTT GAC GAG CCT GAC GAA GCG AAG GAT GAT TCC AGG GTT GGT GCA GAT Phe Asp Glu Pro Asp Glu Ala Lys Asp Asp Ser Arg Val Gly Ala Asp 755 760 765 770	480
GAT GCT GGT TCT TTT GAA CAT TAT GGG ACA ATG ATT CTG GGC CTT TGT Asp Ala Gly Ser Phe Glu His Tyr Gly Thr Met Ile Leu Gly Leu Cys 775 780 785	528
GGG GAG AAT GTT ATG AAC GTG ATC GTG GTG GCT GCT GAA TGT TCT CCA	576

Gly Glu Asn Val Met Asn Val Ile Val Val Ala Ala Glu Cys Ser Pro			
790	795	800	
TGG TGC AAA ACA CGT GGT CTT GGA GAT GTT GTG GGA GCT TTA CCC AAG			624
Trp Cys Lys Thr Gly Gly Leu Gly Asp Val Val Gly Ala Leu Pro Lys			
805	810	815	
GCT TTA GCG AGA AGA GGA CAT CGT GTT ATG GTT GTG GTA CCA AGG TAT			672
Ala Leu Ala Arg Arg Gly His Arg Val Met Val Val Val Pro Arg Tyr			
820	825	830	
GGG GAC TAT GTG GAA GCC TTT GAT ATG GGA ATC CGG AAA TAC TAC AAA			720
Gly Asp Tyr Val Glu Ala Phe Asp Met Gly Ile Arg Lys Tyr Tyr Lys			
835	840	845	850
GCT GCA GGA CAG GAC CTA GAA GTG AAC TAT TTC CAT GCA TTT ATT GAT			768
Ala Ala Gly Gln Asp Leu Glu Val Asn Tyr Phe His Ala Phe Ile Asp			
855	860	865	
GGA GTC GAC TTT GTG TTC ATT GAT GCC TCT TTC CCG CAC CGT CAA GAT			816
Gly Val Asp Phe Val Phe Ile Asp Ala Ser Phe Arg His Arg Gln Asp			
870	875	880	
GAC ATA TAT GGG GGA AGT AGG CAG GAA ATC ATG AAG CGC ATG ATT TTG			864
Asp Ile Tyr Gly Gly Ser Arg Gln Glu Ile Met Lys Arg Met Ile Leu			
885	890	895	
TTT TGC AAG GTT GCT GTT GAG GTT CCT TGG CAC GTT CCA TGC GGT GGT			912
Phe Cys Lys Val Ala Val Glu Val Pro Trp His Val Pro Cys Gly Gly			
900	905	910	
G TG TGC TAC GGA GAT GGA AAT TTG GTG TTC ATT GCC ATG AAT TGG CAC			960
Val Cys Tyr Gly Asp Gly Asn Leu Val Phe Ile Ala Met Asn Trp His			
915	920	925	930
ACT GCA CTC CTG CCT GTT TAT CTG AAG GCA TAT TAC AGA GAC CAT GGG			1008
Thr Ala Leu Leu Pro Val Tyr Leu Lys Ala Tyr Tyr Arg Asp His Gly			
935	940	945	
TTA ATG CAG TAC ACT CGC TCC GTC CTC GTC ATA CAT AAC ATC GGC CAC			1056
Leu Met Gln Tyr Thr Arg Ser Val Leu Val Ile His Asn Ile Gly His			
950	955	960	
CAG GGC CGT GGT CCT GTA CAT GAA TTC CCG TAC ATG GAC TTG CTG AAC			1104
Gln Gly Arg Gly Pro Val His Glu Phe Pro Tyr Met Asp Leu Leu Asn			

965	970	975	
ACT AAC CTT CAA CAT TTC GAG CTG TAC GAT CCC GTC GGT GGC GAG CAC Thr Asn Leu Gln His Phe Glu Leu Tyr Asp Pro Val Gly Gly Glu His			1152
980	985	990	
GCC AAC ATC TTT GCC GCG TGT GTT CTG AAG ATG GCA GAC CGG GTG GTG Ala Asn Ile Phe Ala Ala Cys Val Leu Lys Met Ala Asp Arg Val Val			1200
995	1000	1005	1010
ACT GTC AGC CGC GGC TAC CTG TGG GAG CTG AAG ACA GTG GAA GGC GGC Thr Val Ser Arg Gly Tyr Leu Trp Glu Leu Lys Thr Val Glu Gly Gly			1248
1015	1020	1025	
TGG GGC CTC CAC GAC ATC ATC CGT TCT AAC GAC TGG AAG ATC AAT GGC Trp Gly Leu His Asp Ile Ile Arg Ser Asn Asp Trp Lys Ile Asn Gly			1296
1030	1035	1040	
ATT CGT GAA CGC ATC GAC CAC CAG GAG TGG AAC CCC AAG GTG GAC GTG Ile Arg Glu Arg Ile Asp His Gln Glu Trp Asn Pro Lys Val Asp Val			1344
1045	1050	1055	
CAC CTG CGG TCG GAC GGC TAC ACC AAC TAC TCC CTC GAG ACA CTC GAC His Leu Arg Ser Asp Gly Tyr Thr Asn Tyr Ser Leu Glu Thr Leu Asp			1392
1060	1065	1070	
GCT GGA AAG CGG CAG TGC AAG GCG GCC CTG CAG CGG GAC GTG GGC CTG Ala Gly Lys Arg Gln Cys Lys Ala Ala Leu Gln Arg Asp Val Gly Leu			1440
1075	1080	1085	1090
GAA GTG CGC GAC GAC GTG CCG CTG CTC GGC TTC ATC GGG CGT CTG GAT Glu Val Arg Asp Asp Val Pro Leu Leu Gly Phe Ile Gly Arg Leu Asp			1488
1095	1100	1105	
GGA CAG AAG GGC GTG GAC ATC ATC GGG GAC GCG ATG CCG TGG ATC GCG Gly Gln Lys Gly Val Asp Ile Ile Gly Asp Ala Met Pro Trp Ile Ala			1536
1110	1115	1120	
GGG CAG GAC GTG CAG CTG GTG ATG CTG GGC ACC GGC CCA CCT GAC CTG Gly Gln Asp Val Gln Leu Val Met Leu Gly Thr Gly Pro Pro Asp Leu			1584
1125	1130	1135	
GAA CGA ATG CTG CAG CAC TTG GAG CGG GAG CAT CCC AAC AAG GTG CGC Glu Arg Met Leu Gln His Leu Glu Arg Glu His Pro Asn Lys Val Arg			1632
1140	1145	1150	

GGG TGG GTC GGG TTC TCG GTC CTA ATG GTG CAT CGC ATC ACG CCG GGC Gly Trp Val Gly Phe Ser Val Leu Met Val His Arg Ile Thr Pro Gly	1155	1160	1165	1170	1680
GCC AGC GTG CTG GTG ATG CCC TCC CGC TTC GCC GGC GGG CTG AAC CAG Ala Ser Val Leu Val Met Pro Ser Arg Phe Ala Gly Leu Asn Gln	1175	1180	1185		1728
CTC TAC GCG ATG GCA TAC GGC ACC GTC CCT GTG GTG CAC GCC GTG GGC Leu Tyr Ala Met Ala Tyr Gly Thr Val Pro Val Val His Ala Val Gly	1190	1195	1200		1776
GGG CTC AGG GAC ACC GTG GCG CCG TTC GAC CCG TTC GGC GAC GCC GGG Gly Leu Arg Asp Thr Val Ala Pro Phe Asp Pro Phe Gly Asp Ala Gly	1205	1210	1215		1824
CTC GGG TGG ACT TTT GAC CGC GCC GAG GCC AAC AAG CTG ATC GAG GTG Leu Gly Trp Thr Phe Asp Arg Ala Glu Ala Asn Lys Leu Ile Glu Val	1220	1225	1230		1872
CTC AGC CAC TGC CTC GAC ACG TAC CGA AAC TAC GAG GAG AGC TGG AAG Leu Ser His Cys Leu Asp Thr Tyr Arg Asn Tyr Glu Glu Ser Trp Lys	1235	1240	1245	1250	1920
AGT CTC CAG GCG CGC GGC ATG TCG CAG AAC CTC AGC TGG GAC CAC GCG Ser Leu Gln Ala Arg Gly Met Ser Gln Asn Leu Ser Trp Asp His Ala	1255	1260	1265		1968
GCT GAG CTC TAC GAG GAC GTC CTT GTC AAG TAC CAG TGG Ala Glu Leu Tyr Glu Asp Val Leu Val Lys Tyr Gln Trp	1270	1275			2007

(2) 서열 번호 9에 대한 정보:

( i ) 서열 특징:

- (A)길이: 669개 아미노산
- (B)유형: 아미노산
- (D)토플로지: 선형

( ii ) 분자 유형: 단백질

(xi) 서열 설명: 서열 번호 9:

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

245

250

255

Gly Val Asp Phe Val Phe Ile Asp Ala Ser Phe Arg His Arg Gln Asp  
 260 265 270

Asp Ile Tyr Gly Gly Ser Arg Gln Glu Ile Met Lys Arg Met Ile Leu  
 275 280 285

Phe Cys Lys Val Ala Val Glu Val Pro Trp His Val Pro Cys Gly Gly  
 290 295 300

Val Cys Tyr Gly Asp Gly Asn Leu Val Phe Ile Ala Met Asn Trp His  
 305 310 315 320

Thr Ala Leu Leu Pro Val Tyr Leu Lys Ala Tyr Tyr Arg Asp His Gly  
 325 330 335

Leu Met Gln Tyr Thr Arg Ser Val Leu Val Ile His Asn Ile Gly His  
 340 345 350

Gln Gly Arg Gly Pro Val His Glu Phe Pro Tyr Met Asp Leu Leu Asn  
 355 360 365

Thr Asn Leu Gln His Phe Glu Leu Tyr Asp Pro Val Gly Gly Glu His  
 370 375 380

Ala Asn Ile Phe Ala Ala Cys Val Leu Lys Met Ala Asp Arg Val Val  
 385 390 395 400

Thr Val Ser Arg Gly Tyr Leu Trp Glu Leu Lys Thr Val Glu Gly Gly  
 405 410 415

Trp Gly Leu His Asp Ile Ile Arg Ser Asn Asp Trp Lys Ile Asn Gly  
 420 425 430

Ile Arg Glu Arg Ile Asp His Gln Glu Trp Asn Pro Lys Val Asp Val  
 435 440 445

His Leu Arg Ser Asp Gly Tyr Thr Asn Tyr Ser Leu Glu Thr Leu Asp  
 450 455 460

Ala Gly Lys Arg Gln Cys Lys Ala Ala Leu Gln Arg Asp Val Gly Leu  
 465 470 475 480

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

485

490

495

Gly Gln Lys Gly Val Asp Ile Ile Gly Asp Ala Met Pro Trp Ile Ala  
 500 505 510

Gly Gln Asp Val Gln Leu Val Met Leu Gly Thr Gly Pro Pro Asp Leu  
 515 520 525

Glu Arg Met Leu Gln His Leu Glu Arg Glu His Pro Asn Lys Val Arg  
 530 535 540

Gly Trp Val Gly Phe Ser Val Leu Met Val His Arg Ile Thr Pro Gly  
 545 550 555 560

Ala Ser Val Leu Val Met Pro Ser Arg Phe Ala Gly Leu Asn Gln  
 565 570 575

Leu Tyr Ala Met Ala Tyr Gly Thr Val Pro Val Val His Ala Val Gly  
 580 585 590

Gly Leu Arg Asp Thr Val Ala Pro Phe Asp Pro Phe Gly Asp Ala Gly  
 595 600 605

Leu Gly Trp Thr Phe Asp Arg Ala Glu Ala Asn Lys Leu Ile Glu Val  
 610 615 620

Leu Ser His Cys Leu Asp Thr Tyr Arg Asn Tyr Glu Glu Ser Trp Lys  
 625 630 635 640

Ser Leu Gln Ala Arg Gly Met Ser Gln Asn Leu Ser Trp Asp His Ala  
 645 650 655

Ala Glu Leu Tyr Glu Asp Val Leu Val Lys Tyr Gln Trp  
 660 665

(2) 서열 번호 100에 대한 정보:

(i) 서열 특징:

- (A) 길이: 2097개 염기쌍
- (B) 유형: 핵산
- (C) 가닥수: 이중
- (D) 토플로지: 관련없음

(ii) 분자 유형: cDNA에서 mRNA

(iii) 가정: 없음

(vi) 최초 소오스:

- (A) 유기체: Zea mays

(ix) 특징:

- (A) 명칭/URI: CDS
- (B) 위치: 1..2097

(xi) 서열 설명: 서열 번호 10:

ATG CCG GGG GCA ATC TCT TCC TCG TCG GCT TTT CTC CTC CCC GTC	48
Met Pro Gly Ala Ile Ser Ser Ser Ser Ala Phe Leu Leu Pro Val	
670 675 680 685	
GCG TCC TCC TCG CCG CGG CGC AGG CGG GGC AGT GTG GGT GCT GCT CTG	96
Ala Ser Ser Ser Pro Arg Arg Arg Gly Ser Val Gly Ala Ala Leu	
690 695 700	
CGC TCG TAC GGC TAC AGC GGC GCG GAG CTG CGG TTG CAT TGG GCG CGG	144
Arg Ser Tyr Gly Tyr Ser Gly Ala Glu Leu Arg Leu His Trp Ala Arg	
705 710 715	
CGG GGC CCG CCT CAG GAT GGA GCG GCG TCG GTA CGC GCC GCA GCG GCA	192
Arg Gly Pro Pro Gln Asp Gly Ala Ala Ser Val Arg Ala Ala Ala Ala	
720 725 730	
CCG GCC GGG GGC GAA AGC GAG GAG GCA GCG AAG AGC TCC TCC TCG TCC	240
Pro Ala Gly Gly Glu Ser Glu Glu Ala Ala Lys Ser Ser Ser Ser Ser	
735 740 745	
CAG GCG GGC GCT GTT CAG GGC AGC ACG GCC AAG GCT GTG GAT TCT GCT	288
Gln Ala Gly Ala Val Gln Gly Ser Thr Ala Lys Ala Val Asp Ser Ala	
750 755 760 765	
TCA CCT CCC AAT CCT TTG ACA TCT GCT CCG AAG CAA AGT CAG AGC GCT	336
Ser Pro Pro Asn Pro Leu Thr Ser Ala Pro Lys Gln Ser Gln Ser Ala	
770 775 780	
GCA ATG CAA AAC GGA ACG AGT GGG GGC AGC AGC GCG AGC ACC GCC GCG	384
Ala Met Gln Asn Gly Thr Ser Gly Gly Ser Ser Ala Ser Thr Ala Ala	
785 790 795	
CCG GTG TCC GGA CCC AAA GCT GAT CAT CCA TCA GCT CCT GTC ACC AAG	432

Pro Val Ser Gly Pro Lys Ala Asp His Pro Ser Ala Pro Val Thr Lys		
800	805	810
AGA GAA ATC GAT GCC AGT GCG GTG AAG CCA GAG CCC GCA GGT GAT GAT		480
Arg Glu Ile Asp Ala Ser Ala Val Lys Pro Glu Pro Ala Gly Asp Asp		
815	820	825
GCT AGA CCG GTG GAA AGC ATA GGC ATC GCT GAA CCG GTG GAT GCT AAG		528
Ala Arg Pro Val Glu Ser Ile Gly Ile Ala Glu Pro Val Asp Ala Lys		
830	835	840
GCT GAT GCA GCT CCG GCT ACA GAT GCG GCG GCG AGT GCT CCT TAT GAC		576
Ala Asp Ala Ala Pro Ala Thr Asp Ala Ala Ser Ala Pro Tyr Asp		
850	855	860
AGG GAG GAT AAT GAA CCT GGC CCT TTG GCT GGG CCT AAT GTG ATG AAC		624
Arg Glu Asp Asn Glu Pro Gly Pro Leu Ala Gly Pro Asn Val Met Asn		
865	870	875
GTC GTC GTG GTG GCT TCT GAA TGT GCT CCT TTC TGC AAG ACA GGT GGC		672
Val Val Val Val Ala Ser Glu Cys Ala Pro Phe Cys Lys Thr Gly Gly		
880	885	890
CTT GGA GAT GTC GTG GGT GCT TTG CCT AAG GCT CTG GCG ACG AGA GGA		720
Leu Gly Asp Val Val Gly Ala Leu Pro Lys Ala Leu Ala Arg Arg Gly		
895	900	905
CAC CGT GTT ATG GTC GTG ATA CCA AGA TAT GGA GAG TAT GCC GAA GCC		768
His Arg Val Met Val Val Ile Pro Arg Tyr Gly Glu Tyr Ala Glu Ala		
910	915	920
925		
CGG GAT TTA GGT GTA AGG AGA CGT TAC AAG GTA GCT GGA CAG GAT TCA		816
Arg Asp Leu Gly Val Arg Arg Tyr Lys Val Ala Gly Gln Asp Ser		
930	935	940
GAA GTT ACT TAT TTT CAC TCT TAC ATT GAT GGA GTT GAT TTT GTA TTC		864
Glu Val Thr Tyr Phe His Ser Tyr Ile Asp Gly Val Asp Phe Val Phe		
945	950	955
GTA GAA GCC CCT CCC TTC CGG CAC CGG CAC AAT AAT ATT TAT GGG GGA		912
Val Glu Ala Pro Pro Phe Arg His Arg His Asn Asn Ile Tyr Gly Gly		
960	965	970
GAA AGA TTG GAT ATT TTG AAG CGC ATG ATT TTG TTC TGC AAG GGC GCT		960
Glu Arg Leu Asp Ile Leu Lys Arg Met Ile Leu Phe Cys Lys Ala Ala		

975	980	985	
GTT GAG GTT CCA TGG TAT GCT CCA TGT GGC GGT ACT GTC TAT GGT GAT Val Glu Val Pro Trp Tyr Ala Pro Cys Gly Gly Thr Val Tyr Gly Asp			1008
990	995	1000	1005
GGC AAC TTA GTT TTC ATT GCT AAT GAT TGG CAT ACC GCA CTT CTG CCT Gly Asn Leu Val Phe Ile Ala Asn Asp Trp His Thr Ala Leu Leu Pro			1056
1010	1015	1020	
GTC TAT CTA AAG GCC TAT TAC CGG GAC AAT GGT TTG ATG CAG TAT GCT Val Tyr Leu Lys Ala Tyr Tyr Arg Asp Asn Gly Leu Met Gln Tyr Ala			1104
1025	1030	1035	
CGC TCT GTG CTT GTG ATA CAC AAC ATT GCT CAT CAG GGT CGT GGC CCT Arg Ser Val Leu Val Ile His Asn Ile Ala His Gln Gly Arg Gly Pro			1152
1040	1045	1050	
GTA GAC GAC TTC GTC AAT TTT GAC TTG CCT GAA CAC TAC ATC GAC CAC Val Asp Asp Phe Val Asn Phe Asp Leu Pro Glu His Tyr Ile Asp His			1200
1055	1060	1065	
TTC AAA CTG TAT GAC AAC ATT GGT GGG GAT CAC AGC AAC GTT TTT GCT Phe Lys Leu Tyr Asp Asn Ile Gly Gly Asp His Ser Asn Val Phe Ala			1248
1070	1075	1080	1085
GCG GGG CTG AAG ACG GCA GAC CGG GTG GTG ACC GTT AGC AAT GGC TAC Ala Gly Leu Lys Thr Ala Asp Arg Val Val Thr Val Ser Asn Gly Tyr			1296
1090	1095	1100	
ATG TGG GAG CTG AAG ACT TCG GAA GGC GGG TGG GGC CTC CAC GAC ATC Met Trp Glu Leu Lys Thr Ser Glu Gly Gly Trp Gly Leu His Asp Ile			1344
1105	1110	1115	
ATA AAC CAG AAC GAC TGG AAG CTG CAG GGC ATC GTG AAC GGC ATC GAC Ile Asn Gln Asn Asp Trp Lys Leu Gln Gly Ile Val Asn Gly Ile Asp			1392
1120	1125	1130	
ATG AGC GAG TGG AAC CCC GCT GTG GAC GTG CAC CTC CAC TCC GAC GAC Met Ser Glu Trp Asn Pro Ala Val Asp Val His Leu His Ser Asp Asp			1440
1135	1140	1145	
TAC ACC AAC TAC ACG TTC GAG ACG CTG GAC ACC GGC AAG CGG CAG TGC Tyr Thr Asn Tyr Thr Phe Glu Thr Leu Asp Thr Gly Lys Arg Gln Cys			1488
1150	1155	1160	1165

AAG GCC GCC CTG CAG CGG CAG CTG GGC CTG CAG GTC CGC GAC GAC GTG Lys Ala Ala Leu Gln Arg Gln Leu Gly Leu Gln Val Arg Asp Asp Val 1170 1175 1180	1536
CCA CTG ATC GGG TTC ATC GGG CGG CTG GAC CAC CAG AAG GGC GTG GAC Pro Leu Ile Gly Phe Ile Gly Arg Leu Asp His Gln Lys Gly Val Asp 1185 1190 1195	1584
ATC ATC GCC GAC GCG ATC CAC TGG ATC GCG GGG CAG GAC GTG CTC Ile Ile Ala Asp Ala Ile His Trp Ile Ala Gly Gln Asp Val Gln Leu 1200 1205 1210	1632
GTG ATG CTG GGC ACC GGG CGG GCC GAC CTG GAG GAC ATG CTG CGG CGG Val Met Leu Gly Thr Gly Arg Ala Asp Leu Glu Asp Met Leu Arg Arg 1215 1220 1225	1680
TTC GAG TCG GAG CAC AGC GAC AAG GTG CGC GCG TGG GTG GGG TTC TCG Phe Glu Ser Glu His Ser Asp Lys Val Arg Ala Trp Val Gly Phe Ser 1230 1235 1240 1245	1728
GTG CCC CTG GCG CAC CGC ATC ACG GCG GGC GCG GAC ATC CTG CTG ATG Val Pro Leu Ala His Arg Ile Thr Ala Gly Ala Asp Ile Leu Leu Met 1250 1255 1260	1776
CCG TCG CGG TTC GAG CCG TGC GGG CTG AAC CAG CTC TAC GCC ATG GCG Pro Ser Arg Phe Glu Pro Cys Gly Leu Asn Gln Leu Tyr Ala Met Ala 1265 1270 1275	1824
TAC GGG ACC GTG CCC GTG GTG CAC GCC GTG GGG GGG CTC CGG GAC ACG Tyr Gly Thr Val Pro Val Val His Ala Val Gly Gly Leu Arg Asp Thr 1280 1285 1290	1872
GTG GCG CCG TTC GAC CCG TTC AAC GAC ACC GGG CTC GGG TGG ACG TTC Val Ala Pro Phe Asp Pro Phe Asn Asp Thr Gly Leu Gly Trp Thr Phe 1295 1300 1305	1920
GAC CGC GCG GAG GCG AAC CGG ATG ATC GAC GCG CTC TCG CAC TGC CTC Asp Arg Ala Glu Ala Asn Arg Met Ile Asp Ala Leu Ser His Cys Leu 1310 1315 1320 1325	1968
ACC ACG TAC CGG AAC TAC AAG GAG AGC TGG CGC GCC TGC AGG GCG CGC Thr Thr Tyr Arg Asn Tyr Lys Glu Ser Trp Arg Ala Cys Arg Ala Arg 1330 1335 1340	2016
GGC ATG GCC GAG GAC CTC AGC TGG GAC CAC GCC GCC GTG CTG TAT GAG	2064
Gly Met Ala Glu Asp Leu Ser Trp Asp His Ala Ala Val Leu Tyr Glu 1345 1350 1355	
GAC GTG CTC GTC AAG GCG AAG TAC CAG TGG TGA Asp Val Leu Val Lys Ala Lys Tyr Gln Trp *	2097
1360 1365	

(2) 서열 번호 110에 대한 정보:

( i ) 서열 특징:

- (A)길이: 699개 아미노산
- (B)유형: 아미노산
- (D)토플로지: 선형

( ii ) 분자 유형: 단백질

(xi) 서열 설명: 서열 번호 11:

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

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

Val Asp Asp Phe Val Asn Phe Asp Leu Pro Glu His Tyr Ile Asp His  
 385 390 395 400

Phe Lys Leu Tyr Asp Asn Ile Gly Gly Asp His Ser Asn Val Phe Ala  
 405 410 415

Ala Gly Leu Lys Thr Ala Asp Arg Val Val Thr Val Ser Asn Gly Tyr  
 420 425 430

Met Trp Glu Leu Lys Thr Ser Glu Gly Gly Trp Gly Leu His Asp Ile  
 435 440 445

Ile Asn Gln Asn Asp Trp Lys Leu Gln Gly Ile Val Asn Gly Ile Asp  
 450 455 460

Met Ser Glu Trp Asn Pro Ala Val Asp Val His Leu His Ser Asp Asp  
 465 470 475 480

Tyr Thr Asn Tyr Thr Phe Glu Thr Leu Asp Thr Gly Lys Arg Gln Cys  
 485 490 495

Lys Ala Ala Leu Gln Arg Gln Leu Gly Leu Gln Val Arg Asp Asp Val  
 500 505 510

Pro Leu Ile Gly Phe Ile Gly Arg Leu Asp His Gln Lys Gly Val Asp  
 515 520 525

Ile Ile Ala Asp Ala Ile His Trp Ile Ala Gly Gln Asp Val Gln Leu  
 530 535 540

Val Met Leu Gly Thr Gly Arg Ala Asp Leu Glu Asp Met Leu Arg Arg  
 545 550 555 560

Phe Glu Ser Glu His Ser Asp Lys Val Arg Ala Trp Val Gly Phe Ser  
 565 570 575

Val Pro Leu Ala His Arg Ile Thr Ala Gly Ala Asp Ile Leu Leu Met  
 580 585 590

Pro Ser Arg Phe Glu Pro Cys Gly Leu Asn Gln Leu Tyr Ala Met Ala  
 595 600 605

Tyr Gly Thr Val Pro Val Val His Ala Val Gly Gly Leu Arg Asp Thr  
 610 615 620

Val Ala Pro Phe Asp Pro Phe Asn Asp Thr Gly Leu Gly Trp Thr Phe  
 625 630 635 640

Asp Arg Ala Glu Ala Asn Arg Met Ile Asp Ala Leu Ser His Cys Leu  
 645 650 655

Thr Thr Tyr Arg Asn Tyr Lys Glu Ser Trp Arg Ala Cys Arg Ala Arg  
 660 665 670

Gly Met Ala Glu Asp Leu Ser Trp Asp His Ala Ala Val Leu Tyr Glu  
 675 680 685

Asp Val Leu Val Lys Ala Lys Tyr Gln Trp \*  
 690 695

(2) 서열 번호 120에 대한 정보:

(i) 서열 특징:

(A)길이: 1752개 염기쌍

(B)유형: 핵산

(C)가닥수: 이중

(D) 토플로지: 관련없음

(ii) 분자 유형: cDNA 에서 mRNA

(iii) 가정: 없음

(vi) 최초 소오스:

(A) 유기체: Zea mays

(ix) 특징:

(A) 명칭/URI: CDS

(B) 위치: 1..1752

(xi) 서열 설명: 서열 번호 12:

TGC GTC GCG GAG CTG AGC AGG GAG GGG CCC GCG CCG CGC CCG CTG CCA	48
Cys Val Ala Glu Leu Ser Arg Glu Gly Pro Ala Pro Arg Pro Leu Pro	
700	705
	710
	715

CCC GCG CTG CTG GCG CCC CCG CTC GTG CCC GGC TTC CTC GCG CCG CCG	96
Pro Ala Leu Leu Ala Pro Pro Leu Val Pro Gly Phe Leu Ala Pro Pro	
720	725
	730

GCC GAG CCC ACG GGT GAG CCG GCA TCG ACG CCG CCG CCC GTG CCC GAC Ala Glu Pro Thr Gly Glu Pro Ala Ser Thr Pro Pro Pro Val Pro Asp 735 740 745	144
GCC GGC CTG GGG GAC CTC GGT CTC GAA CCT GAA GGG ATT GCT GAA GGT Ala Gly Leu Gly Asp Leu Gly Leu Glu Pro Glu Gly Ile Ala Glu Gly 750 755 760	192
TCC ATC GAT AAC ACA GTA GTT GTG GCA AGT GAG CAA GAT TCT GAG ATT Ser Ile Asp Asn Thr Val Val Val Ala Ser Glu Gln Asp Ser Glu Ile 765 770 775	240
GTG GTT GGA AAG GAG CAA GCT CGA GCT AAA GTA ACA CAA AGC ATT GTC Val Val Gly Lys Glu Gln Ala Arg Ala Lys Val Thr Gln Ser Ile Val 780 785 790 795	288
TTT GTA ACC GGC GAA GCT TCT CCT TAT GCA AAG TCT GGG GGT CTA GGA Phe Val Thr Gly Glu Ala Ser Pro Tyr Ala Lys Ser Gly Gly Leu Gly 800 805 810	336
GAT GTT TGT GGT TCA TTG CCA GTT GCT CTT GCT GCT CGT GGT CAC CGT Asp Val Cys Gly Ser Leu Pro Val Ala Leu Ala Ala Arg Gly His Arg 815 820 825	384
GTG ATG GTT GTA ATG CCC AGA TAT TTA AAT GGT ACC TCC GAT AAG AAT Val Met Val Val Met Pro Arg Tyr Leu Asn Gly Thr Ser Asp Lys Asn 830 835 840	432
TAT GCA AAT GCA TTT TAC ACA GAA AAA CAC ATT CGG ATT CCA TGC TTT Tyr Ala Asn Ala Phe Tyr Thr Glu Lys His Ile Arg Ile Pro Cys Phe 845 850 855	480
GGC GGT GAA CAT GAA GTT ACC TTC TTC CAT GAG TAT AGA GAT TCA GTT Gly Gly Glu His Glu Val Thr Phe Phe His Glu Tyr Arg Asp Ser Val 860 865 870 875	528
GAC TGG GTG TTT GTT GAT CAT CCC TCA TAT CAC AGA CCT GGA AAT TTA Asp Trp Val Phe Val Asp His Pro Ser Tyr His Arg Pro Gly Asn Leu 880 885 890	576
TAT GGA GAT AAG TTT GGT GCT TTT GGT GAT AAT CAG TTC AGA TAC ACA Tyr Gly Asp Lys Phe Gly Ala Phe Gly Asp Asn Gln Phe Arg Tyr Thr 895 900 905	624
CTC CTT TGC TAT GCT GCA TGT GAG GCT CCT TTG ATC CTT GAA TTG GGA	672

Leu Leu Cys Tyr Ala Ala Cys Glu Ala Pro Leu Ile Leu Glu Leu Gly			
910	915	920	
GGA TAT ATT TAT GGA CAG AAT TGC ATG TTT GTT GTC AAT GAT TGG CAT			720
Gly Tyr Ile Tyr Gly Gln Asn Cys Met Phe Val Val Asn Asp Trp His			
925	930	935	
GCC AGT CTA GTG CCA GTC CTT CTT GCT GCA AAA TAT AGA CCA TAT GGT			768
Ala Ser Leu Val Pro Val Leu Leu Ala Ala Lys Tyr Arg Pro Tyr Gly			
940	945	950	955
GTT TAT AAA GAC TCC CGC AGC ATT CTT GTA ATA CAT AAT TTA GCA CAT			816
Val Tyr Lys Asp Ser Arg Ser Ile Leu Val Ile His Asn Leu Ala His			
960	965	970	
CAG GGT GTA GAG CCT GCA AGC ACA TAT CCT GAC CTT GGG TTG CCA CCT			864
Gln Gly Val Glu Pro Ala Ser Thr Tyr Pro Asp Leu Gly Leu Pro Pro			
975	980	985	
GAA TGG TAT GGA GCT CTG GAG TGG GTA TTC CCT GAA TGG GCG AGG AGG			912
Glu Trp Tyr Gly Ala Leu Glu Trp Val Phe Pro Glu Trp Ala Arg Arg			
990	995	1000	
CAT GCC CTT GAC AAG GGT GAG GCA GTT AAT TTT TTG AAA GGT GCA GTT			960
His Ala Leu Asp Lys Gly Glu Ala Val Asn Phe Leu Lys Gly Ala Val			
1005	1010	1015	
GTC ACA GCA GAT CGA ATC GTG ACT GTC AGT AAG GGT TAT TCG TGG GAG			1008
Val Thr Ala Asp Arg Ile Val Thr Val Ser Lys Gly Tyr Ser Trp Glu			
1020	1025	1030	1035
GTC ACA ACT GCT GAA GGT GGA CAG GCC CTC AAT GAG CTC TTA AGC TCC			1056
Val Thr Thr Ala Glu Gly Gln Gly Leu Asn Glu Leu Leu Ser Ser			
1040	1045	1050	
AGA AAG AGT GTA TTA AAC GGA ATT GTA AAT GGA ATT GAC ATT AAT GAT			1104
Arg Lys Ser Val Leu Asn Gly Ile Val Asn Gly Ile Asp Ile Asn Asp			
1055	1060	1065	
TGG AAC CCT GCC ACA GAC AAA TGT ATC CCC TGT CAT TAT TCT GTT GAT			1152
Trp Asn Pro Ala Thr Asp Lys Cys Ile Pro Cys His Tyr Ser Val Asp			
1070	1075	1080	
GAC CTC TCT GGA AAG GCC AAA TGT AAA GGT GCA TTG CAG AAG GAG CTG			1200
Asp Leu Ser Gly Lys Ala Lys Cys Lys Gly Ala Leu Gln Lys Glu Leu			

1085	1090	1095	
GGT TTA CCT ATA AGG CCT GAT GTT CCT CTG ATT GGC TTT ATT GGA AGG Gly Leu Pro Ile Arg Pro Asp Val Pro Leu Ile Gly Phe Ile Gly Arg 1100 1105 1110 1115			1248
TTG GAT TAT CAG AAA GGC ATT GAT CTC ATT CAA CTT ATC ATA CCA GAT Leu Asp Tyr Gln Lys Gly Ile Asp Leu Ile Gln Leu Ile Ile Pro Asp 1120 1125 1130			1296
CTC ATG CGG GAA GAT GTT CAA TTT GTC ATG CTT GGA TCT GGT GAC CCA Leu Met Arg Glu Asp Val Gln Phe Val Met Leu Gly Ser Gly Asp Pro 1135 1140 1145			1344
GAG CTT GAA GAT TGG ATG AGA TCT ACA GAG TCG ATC TTC AAG GAT AAA Glu Leu Glu Asp Trp Met Arg Ser Thr Glu Ser Ile Phe Lys Asp Lys 1150 1155 1160			1392
TTT CGT GGA TGG GTT GGA TTT AGT GTT CCA GTT TCC CAC CGA ATA ACT Phe Arg Gly Trp Val Gly Ser Val Pro Val Ser His Arg Ile Thr 1165 1170 1175			1440
GCC GGC TGC GAT ATA TTG TTA ATG CCA TCC AGA TTC GAA CCT TGT GGT Ala Gly Cys Asp Ile Leu Leu Met Pro Ser Arg Phe Glu Pro Cys Gly 1180 1185 1190 1195			1488
CTC AAT CAG CTA TAT GCT ATG CAG TAT GGC ACA GTT CCT GTC CAT Leu Asn Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val Pro Val Val His 1200 1205 1210			1536
GCA ACT GGG GGC CTT AGA GAT ACC GTG GAG AAC TTC AAC CCT TTC GGT Ala Thr Gly Gly Leu Arg Asp Thr Val Glu Asn Phe Asn Pro Phe Gly 1215 1220 1225			1584
GAG AAT GGA GAG CAG GGT ACA GGG TGG GCA TTC GCA CCC CTA ACC ACA Glu Asn Gly Glu Gln Gly Thr Gly Trp Ala Phe Ala Pro Leu Thr Thr 1230 1235 1240			1632
GAA AAC ATG TTT GTG GAC ATT GCG AAC TGC AAT ATC TAC ATA CAG GGA Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile Tyr Ile Gln Gly 1245 1250 1255			1680
ACA CAA GTC CTC CTG GGA AGG GCT AAT GAA GCG AGG CAT GTC AAA AGA Thr Gln Val Leu Leu Gly Arg Ala Asn Glu Ala Arg His Val Lys Arg 1260 1265 1270 1275			1728
CTT CAC GTG GGA CCA TGC CGC TGA Leu His Val Gly Pro Cys Arg *			1752
1280			

(2) 서열 번호 130에 대한 정보:

(i) 서열 특징:

- (A)길이: 584개 아미노산
- (B)유형: 아미노산
- (D)토플로지: 선형

(ii) 문자 유형: 단백질

(xi) 서열 설명: 서열 번호 13:

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

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

Gly Leu Pro Ile Arg Pro Asp Val Pro Leu Ile Gly Phe Ile Gly Arg  
 405 410 415  
  
 Leu Asp Tyr Gln Lys Gly Ile Asp Leu Ile Gln Leu Ile Ile Pro Asp  
 420 425 430  
  
 Leu Met Arg Glu Asp Val Gln Phe Val Met Leu Gly Ser Gly Asp Pro  
 435 440 445  
  
 Glu Leu Glu Asp Trp Met Arg Ser Thr Glu Ser Ile Phe Lys Asp Lys  
 450 455 460  
  
 Phe Arg Gly Trp Val Gly Phe Ser Val Pro Val Ser His Arg Ile Thr  
 465 470 475 480  
  
 Ala Gly Cys Asp Ile Leu Leu Met Pro Ser Arg Phe Glu Pro Cys Gly  
 485 490 495  
  
 Leu Asn Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val Pro Val Val His  
 500 505 510  
  
 Ala Thr Gly Gly Leu Arg Asp Thr Val Glu Asn Phe Asn Pro Phe Gly  
 515 520 525  
  
 Glu Asn Gly Glu Gln Gly Thr Gly Trp Ala Phe Ala Pro Leu Thr Thr  
 530 535 540  
  
 Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile Tyr Ile Gln Gly  
 545 550 555 560  
  
 Thr Gln Val Leu Leu Gly Arg Ala Asn Glu Ala Arg His Val Lys Arg  
 565 570 575  
  
 Leu His Val Gly Pro Cys Arg \*  
 580

(2) 서열 번호 140에 대한 정보:

(i) 서열 특징:

- (A) 길이: 2725개 염기쌍
- (B) 유형: 핵산
- (C) 가닥수: 단일
- (D) 토플로지: 관련없음

(ii) 분자 유형: mRNA

(iii) 가정: 없음

(vi) 최초 소오스:

- (A) 유기체: Zea mays

(ix) 특징:

- (A) 명칭/키: sig. 펩티드
- (B) 위치: 91..264

(ix) 특징:

- (A) 명칭/키: mat. 펩티드
- (B) 위치: 265..2487

(ix) 특징:

- (A) 명칭/키: CDS
- (B) 위치: 91..2490

(xi) 서열 설명: 서열 번호 14:

GGCCCAGAGC AGACCCGGAT TTTCGCTCTTG CGGTCGCTGG GGTTTTAGCA TTGGCTGATC	60
AGTTCGATCC GATCCGGCTG CGAAGGGCGAG ATG GCG TTC CCG GTT TCT GGG GCG	114
Met Ala Phe Arg Val Ser Gly Ala	
-58 -55	
GTG CTC GGT GGG GCC GTC AGG GCT CCC CGA CTC ACC GGC GGC GGG GAG	162
Val Leu Gly Gly Ala Val Arg Ala Pro Arg Leu Thr Gly Gly Gly Glu	
-50 -45 -40 -35	
GGT AGT CTA GTC TTC CGG CAC ACC GGC CTC TTC TTA ACT CGG GGT GCT	210
Gly Ser Leu Val Phe Arg His Thr Gly Leu Phe Leu Thr Arg Gly Ala	
-30 -25 -20	
CGA GTT GGA TGT TCG GGG ACG CAC GGG GCC ATG CGC GCG GCG GCC GCG	258
Arg Val Gly Cys Ser Gly Thr His Gly Ala Met Arg Ala Ala Ala Ala	
-15 -10 -5	
GCC AGG AAG GCG GTC ATG GTT CCT GAG GGC GAG AAT GAT GGC CTC GCA	306
Ala Arg Lys Ala Val Met Val Pro Glu Gly Glu Asn Asp Gly Leu Ala	
1 5 10	
TCA AGG GCT GAC TCG GCT CAA TTC CAG TCG GAT GAA CTG GAG GTA CCA	354
Ser Arg Ala Asp Ser Ala Gln Phe Gln Ser Asp Glu Leu Glu Val Pro	
15 20 25 30	

GAC ATT TCT GAA GAG ACA ACG TGC GGT GCT GGT GTG GCT GAT GCT CAA Asp Ile Ser Glu Glu Thr Thr Cys Gly Ala Gly Val Ala Asp Ala Gln	35	40	45	402
GCC TTG AAC AGA GTT CGA GTG GTC CCC CCA CCA AGC GAT GGA CAA AAA Ala Leu Asn Arg Val Arg Val Val Pro Pro Ser Asp Gly Gln Lys	50	55	60	450
ATA TTC CAG ATT GAC CCC ATG TTG CAA GGC TAT AAG TAC CAT CTT GAG Ile Phe Gln Ile Asp Pro Met Leu Gln Gly Tyr Lys Tyr His Leu Glu	65	70	75	498
TAT CGG TAC AGC CTC TAT AGA AGA ATC CGT TCA GAC ATT GAT GAA CAT Tyr Arg Tyr Ser Leu Tyr Arg Arg Ile Arg Ser Asp Ile Asp Glu His	80	85	90	546
GAA GGA GGC TTG GAA GCC TTC TCC CGT AGT TAT GAG AAG TTT GGA TTT Glu Gly Gly Leu Glu Ala Phe Ser Arg Ser Tyr Glu Lys Phe Gly Phe	95	100	105	594
AAT GCC AGC GCG GAA GGT ATC ACA TAT CGA GAA TGG GCT CCT GGA GCA Asn Ala Ser Ala Glu Gly Ile Thr Tyr Arg Glu Trp Ala Pro Gly Ala	115	120	125	642
TTT TCT GCA GCA TTG GTG GGT GAC GTC AAC AAC TGG GAT CCA AAT GCA Phe Ser Ala Ala Leu Val Gly Asp Val Asn Asn Trp Asp Pro Asn Ala	130	135	140	690
GAT CGT ATG AGC AAA AAT GAG TTT GGT GTT TGG GAA ATT TTT CTG CCT Asp Arg Met Ser Lys Asn Glu Phe Gly Val Trp Glu Ile Phe Leu Pro	145	150	155	738
AAC AAT GCA GAT GGT ACA TCA CCT ATT CCT CAT GGA TCT CGT GTA AAG Asn Asn Ala Asp Gly Thr Ser Pro Ile Pro His Gly Ser Arg Val Lys	160	165	170	786
GTG AGA ATG GAT ACT CCA TCA GGG ATA AAG GAT TCA ATT CCA GCC TGG Val Arg Met Asp Thr Pro Ser Gly Ile Lys Asp Ser Ile Pro Ala Trp	175	180	185	834
ATC AAG TAC TCA GTG CAG GCC CCA GGA GAA ATA CCA TAT GAT GGG ATT Ile Lys Tyr Ser Val Gln Ala Pro Gly Glu Ile Pro Tyr Asp Gly Ile	195	200	205	882
TAT TAT GAT CCT CCT GAA GAG GTA AAG TAT GTG TTC AGG CAT GCG CAA				930

Tyr Tyr Asp Pro Pro Glu Glu Val Lys Tyr Val Phe Arg His Ala Gln		
210	215	220
CCT AAA CGA CCA AAA TCA TTG CGG ATA TAT GAA ACA CAT GTC GGA ATG	978	
Pro Lys Arg Pro Lys Ser Leu Arg Ile Tyr Glu Thr His Val Gly Met		
225	230	235
AGT AGC CCG GAA CCG AAG ATA AAC ACA TAT GTA AAC TTT AGG GAT GAA	1026	
Ser Ser Pro Glu Pro Lys Ile Asn Thr Tyr Val Asn Phe Arg Asp Glu		
240	245	250
GTC CTC CCA AGA ATA AAA CTT GGA TAC AAT GCA GTG CAA ATA ATG	1074	
Val Leu Pro Arg Ile Lys Lys Leu Gly Tyr Asn Ala Val Gln Ile Met		
255	260	265
270		
GCA ATC CAA GAG CAC TCA TAT TAT GGA AGC TTT GGA TAC CAT GTA ACT	1122	
Ala Ile Gln Glu His Ser Tyr Tyr Gly Ser Phe Gly Tyr His Val Thr		
275	280	285
AAT TTT TTT GCG CCA AGT AGT CGT TTT GGT ACC CCA GAA GAT TTG AAG	1170	
Asn Phe Phe Ala Pro Ser Ser Arg Phe Gly Thr Pro Glu Asp Leu Lys		
290	295	300
TCT TTG ATT GAT AGA GCA CAT GAG CTT GGT TTG CTA GTT CTC ATG GAT	1218	
Ser Leu Ile Asp Arg Ala His Glu Leu Gly Leu Leu Val Leu Met Asp		
305	310	315
GTG GTT CAT AGT CAT GCG TCA AGT AAT ACT CTG GAT GGG TTG AAT GGT	1266	
Val Val His Ser His Ala Ser Ser Asn Thr Leu Asp Gly Leu Asn Gly		
320	325	330
TTT GAT GGT ACA GAT ACA CAT TAC TTT CAC AGT GGT CCA CGT GGC CAT	1314	
Phe Asp Gly Thr Asp Thr His Tyr Phe His Ser Gly Pro Arg Gly His		
335	340	345
350		
CAC TGG ATG TGG GAT TCT CGC CTA TTT AAC TAT GGG AAC TGG GAA GTT	1362	
His Trp Met Trp Asp Ser Arg Leu Phe Asn Tyr Gly Asn Trp Glu Val		
355	360	365
TTA AGA TTT CTT CTC TCC AAT GCT AGA TGG TGG CTC GAG GAA TAT AAG	1410	
Leu Arg Phe Leu Leu Ser Asn Ala Arg Trp Trp Leu Glu Glu Tyr Lys		
370	375	380
TTT GAT GGT TTC CGT TTT GAT GGT GTG ACC TCC ATG ATG TAC ACT CAC	1458	
Phe Asp Gly Phe Arg Phe Asp Gly Val Thr Ser Met Met Tyr Thr His		

385

390

395

CAC GGA TTA CAA GTA ACA TTT ACG GGG AAC TTC AAT GAG TAT TTT GGC			1506
His Gly Leu Gln Val Thr Phe Thr Gly Asn Phe Asn Glu Tyr Phe Gly			
400	405	410	
TTT GCC ACC GAT GTA GAT GCA GTG GTT TAC TTG ATG CTG GTA AAT GAT			1554
Phe Ala Thr Asp Val Asp Ala Val Val Tyr Leu Met Leu Val Asn Asp			
415	420	425	430
CTA ATT CAT GGA CTT TAT CCT GAG GCT GTA ACC ATT GGT GAA GAT GTT			1602
Leu Ile His Gly Leu Tyr Pro Glu Ala Val Thr Ile Gly Glu Asp Val			
435	440	445	
AGT GGA ATG CCT ACA TTT GCC CTT CCT GTT CAC GAT GGT GGG GTA GGT			1650
Ser Gly Met Pro Thr Phe Ala Leu Pro Val His Asp Gly Gly Val Gly			
450	455	460	
TTT GAC TAT CGG ATG CAT ATG GCT GTG GCT GAC AAA TGG ATT GAC CTT			1698
Phe Asp Tyr Arg Met His Met Ala Val Ala Asp Lys Trp Ile Asp Leu			
465	470	475	
CTC AAG CAA AGT GAT GAA ACT TGG AAG ATG GGT GAT ATT GTG CAC ACA			1746
Leu Lys Gln Ser Asp Glu Thr Trp Lys Met Gly Asp Ile Val His Thr			
480	485	490	
CTG ACA AAT AGG AGG TGG TTA GAG AAG TGT GTA ACT TAT GCT GAA ACT			1794
Leu Thr Asn Arg Arg Trp Leu Glu Lys Cys Val Thr Tyr Ala Glu Ser			
495	500	505	510
CAT GAT CAA GCA TTA GTC GGC GAC AAG ACT ATT GCG TTT TGG TTG ATG			1842
His Asp Gln Ala Leu Val Gly Asp Lys Thr Ile Ala Phe Trp Leu Met			
515	520	525	
GAC AAG GAT ATG TAT GAT TTC ATG GCC CTC GAT AGA CCT TCA ACT CCT			1890
Asp Lys Asp Met Tyr Asp Phe Met Ala Leu Asp Arg Pro Ser Thr Pro			
530	535	540	
ACC ATT GAT CGT GGG ATA GCA TTA CAT AAG ATG ATT AGA CTT ATC ACA			1938
Thr Ile Asp Arg Gly Ile Ala Leu His Lys Met Ile Arg Leu Ile Thr			
545	550	555	
ATG GGT TTA GGA GGA GAG GGC TAT CTT AAT TTC ATG GGA AAT GAG TTT			1986
Met Gly Leu Gly Glu Gly Tyr Leu Asn Phe Met Gly Asn Glu Phe			
560	565	570	

GGA CAT CCT GAA TGG ATA GAT TTT CCA AGA GGT CCG CAA AGA CTT CCA Gly His Pro Glu Trp Ile Asp Phe Pro Arg Gly Pro Gln Arg Leu Pro 575 580 585 590	2034
AGT GGT AAG TTT ATT CCA GGG AAT AAC AAC AGT TAT GAC AAA TGT CGT Ser Gly Lys Phe Ile Pro Gly Asn Asn Ser Tyr Asp Lys Cys Arg 595 600 605	2082
CGA AGA TTT GAC CTG GGT GAT GCA GAC TAT CTT AGG TAT CAT GGT ATG Arg Arg Phe Asp Leu Gly Asp Ala Tyr Leu Arg Tyr His Gly Met 610 615 620	2130
CAA GAG TTT GAT CAG GCA ATG CAA CAT CTT GAG CAA AAA TAT GAA TTC Gln Glu Phe Asp Gln Ala Met Gln His Leu Glu Gln Lys Tyr Glu Phe 625 630 635	2178
ATG ACA TCT GAT CAC CAG TAT ATT TCC CCG AAA CAT GAG GAG GAT AAG Met Thr Ser Asp His Gln Tyr Ile Ser Arg Lys His Glu Glu Asp Lys 640 645 650	2226
GTG ATT GTG TTC GAA AAG GGA GAT TTG GTA TTT GTG TTC AAC TTC CAC Val Ile Val Phe Glu Lys Gly Asp Leu Val Phe Val Phe Asn Phe His 655 660 665 670	2274
TGC AAC AAC AGC TAT TTT GAC TAC CGT ATT GGT TGT CGA AAG CCT GGG Cys Asn Asn Ser Tyr Phe Asp Tyr Arg Ile Gly Cys Arg Lys Pro Gly 675 680 685	2322
GTG TAT AAG GTG GTC TTG GAC TCC GAC GCT GGA CTA TTT GGT GGA TTT Val Tyr Lys Val Val Leu Asp Ser Asp Ala Gly Leu Phe Gly Gly Phe 690 695 700	2370
AGC AGG ATC CAT CAC GCA GCC GAG CAC TTC ACC GCC GAC TGT TCG CAT Ser Arg Ile His His Ala Ala Glu His Phe Thr Ala Asp Cys Ser His 705 710 715	2418
GAT AAT AGG CCA TAT TCA TTC TCG GTT TAT ACA CCA AGC AGA ACA TGT Asp Asn Arg Pro Tyr Ser Phe Ser Val Tyr Thr Pro Ser Arg Thr Cys 720 725 730	2466
GTC GTC TAT GCT CCA GTG GAG TGA TAGCGGGTA CTCGTTGCTG CGCGGCATGT Val Val Tyr Ala Pro Val Glu *735 740	2520
GTGGGGCTGT CGATGTGAGG AAAAACCTTC TTCCAAAACC GGCAGATGCA TGCATGCATG	2580
CTACAATAAG GTTCTGATAC TTTAATCGAT GCTGGAAAGC CCATGCATCT CGCTGCGTTG	2640
TCCTCTCTAT ATATATAAGA CCTTCAGGT GTCAATTAAA CATAGAGTTT TCGTTTTCG	2700
CTTTCCTAAA AAAAAAAA AAAAA	2725

(2) 서열 번호 15에 대한 정보:

(i) 서열 특징:

- (A)길이: 800개 아미노산
- (B)유형: 아미노산
- (D)토풀로지: 선형

(ii) 문자 유형: 단백질

(xi) 서열 설명: 서열 번호 15:

Met Ala Phe Arg Val Ser Gly Ala Val Leu Gly Gly Ala Val Arg Ala			
-58	-55	-50	-45
Pro Arg Leu Thr Gly Gly Glu Gly Ser Leu Val Phe Arg His Thr			
-40	-35	-30	
Gly Leu Phe Leu Thr Arg Gly Ala Arg Val Gly Cys Ser Gly Thr His			
-25	-20	-15	
Gly Ala Met Arg Ala Ala Ala Ala Arg Lys Ala Val Met Val Pro			
-10	-5	1	5
Glu Gly Glu Asn Asp Gly Leu Ala Ser Arg Ala Asp Ser Ala Gln Phe			
10	15	20	
Gln Ser Asp Glu Leu Glu Val Pro Asp Ile Ser Glu Glu Thr Thr Cys			
25	30	35	
Gly Ala Gly Val Ala Asp Ala Gln Ala Leu Asn Arg Val Arg Val Val			
40	45	50	
Pro Pro Pro Ser Asp Gly Gln Lys Ile Phe Gln Ile Asp Pro Met Leu			
55	60	65	70
Gln Gly Tyr Lys Tyr His Leu Glu Tyr Arg Tyr Ser Leu Tyr Arg Arg			
75	80	85	

Ile Arg Ser Asp Ile Asp Glu His Glu Gly Gly Leu Glu Ala Phe Ser  
 90 95 100  
  
 Arg Ser Tyr Glu Lys Phe Gly Phe Asn Ala Ser Ala Glu Gly Ile Thr  
 105 110 115  
  
 Tyr Arg Glu Trp Ala Pro Gly Ala Phe Ser Ala Ala Leu Val Gly Asp  
 120 125 130  
  
 Val Asn Asn Trp Asp Pro Asn Ala Asp Arg Met Ser Lys Asn Glu Phe  
 135 140 145 150  
  
 Gly Val Trp Glu Ile Phe Leu Pro Asn Asn Ala Asp Gly Thr Ser Pro  
 155 160 165  
  
 Ile Pro His Gly Ser Arg Val Lys Val Arg Met Asp Thr Pro Ser Gly  
 170 175 180  
  
 Ile Lys Asp Ser Ile Pro Ala Trp Ile Lys Tyr Ser Val Gln Ala Pro  
 185 190 195  
  
 Gly Glu Ile Pro Tyr Asp Gly Ile Tyr Tyr Asp Pro Pro Glu Glu Val  
 200 205 210  
  
 Lys Tyr Val Phe Arg His Ala Gln Pro Lys Arg Pro Lys Ser Leu Arg  
 215 220 225 230  
  
 Ile Tyr Glu Thr His Val Gly Met Ser Ser Pro Glu Pro Lys Ile Asn  
 235 240 245  
  
 Thr Tyr Val Asn Phe Arg Asp Glu Val Leu Pro Arg Ile Lys Lys Leu  
 250 255 260  
  
 Gly Tyr Asn Ala Val Gln Ile Met Ala Ile Gln Glu His Ser Tyr Tyr  
 265 270 275  
  
 Gly Ser Phe Gly Tyr His Val Thr Asn Phe Phe Ala Pro Ser Ser Arg  
 280 285 290  
  
 Phe Gly Thr Pro Glu Asp Leu Lys Ser Leu Ile Asp Arg Ala His Glu  
 295 300 305 310  
  
 Leu Gly Leu Leu Val Leu Met Asp Val Val His Ser His Ala Ser Ser  
 315 320 325

Asn Thr Leu Asp Gly Leu Asn Gly Phe Asp Gly Thr Asp Thr His Tyr  
 330 335 340  
  
 Phe His Ser Gly Pro Arg Gly His His Trp Met Trp Asp Ser Arg Leu  
 345 350 355  
  
 Phe Asn Tyr Gly Asn Trp Glu Val Leu Arg Phe Leu Leu Ser Asn Ala  
 360 365 370  
  
 Arg Trp Trp Leu Glu Glu Tyr Lys Phe Asp Gly Phe Arg Phe Asp Gly  
 375 380 385 390  
  
 Val Thr Ser Met Met Tyr Thr His His Gly Leu Gln Val Thr Phe Thr  
 395 400 405  
  
 Gly Asn Phe Asn Glu Tyr Phe Gly Phe Ala Thr Asp Val Asp Ala Val  
 410 415 420  
  
 Val Tyr Leu Met Leu Val Asn Asp Leu Ile His Gly Leu Tyr Pro Glu  
 425 430 435  
  
 Ala Val Thr Ile Gly Glu Asp Val Ser Gly Met Pro Thr Phe Ala Leu  
 440 445 450  
  
 Pro Val His Asp Gly Gly Val Gly Phe Asp Tyr Arg Met His Met Ala  
 455 460 465 470  
  
 Val Ala Asp Lys Trp Ile Asp Leu Leu Lys Gln Ser Asp Glu Thr Trp  
 475 480 485  
  
 Lys Met Gly Asp Ile Val His Thr Leu Thr Asn Arg Arg Trp Leu Glu  
 490 495 500  
  
 Lys Cys Val Thr Tyr Ala Glu Ser His Asp Gln Ala Leu Val Gly Asp  
 505 510 515  
  
 Lys Thr Ile Ala Phe Trp Leu Met Asp Lys Asp Met Tyr Asp Phe Met  
 520 525 530  
  
 Ala Leu Asp Arg Pro Ser Thr Pro Thr Ile Asp Arg Gly Ile Ala Leu  
 535 540 545 550  
  
 His Lys Met Ile Arg Leu Ile Thr Met Gly Leu Gly Glu Gly Tyr  
 555 560 565

Leu Asn Phe Met Gly Asn Glu Phe Gly His Pro Glu Trp Ile Asp Phe  
 570 575 580  
  
 Pro Arg Gly Pro Gln Arg Leu Pro Ser Gly Lys Phe Ile Pro Gly Asn  
 585 590 595  
  
 Asn Asn Ser Tyr Asp Lys Cys Arg Arg Arg Phe Asp Leu Gly Asp Ala  
 600 605 610  
  
 Asp Tyr Leu Arg Tyr His Gly Met Gln Glu Phe Asp Gln Ala Met Gln  
 615 620 625 630  
  
 His Leu Glu Gln Lys Tyr Glu Phe Met Thr Ser Asp His Gln Tyr Ile  
 635 640 645  
  
 Ser Arg Lys His Glu Glu Asp Lys Val Ile Val Phe Glu Lys Gly Asp  
 650 655 660  
  
 Leu Val Phe Val Phe Asn Phe His Cys Asn Asn Ser Tyr Phe Asp Tyr  
 665 670 675  
  
 Arg Ile Gly Cys Arg Lys Pro Gly Val Tyr Lys Val Val Leu Asp Ser  
 680 685 690  
  
 Asp Ala Gly Leu Phe Gly Gly Phe Ser Arg Ile His His Ala Ala Glu  
 695 700 705 710  
  
 His Phe Thr Ala Asp Cys Ser His Asp Asn Arg Pro Tyr Ser Phe Ser  
 715 720 725  
  
 Val Tyr Thr Pro Ser Arg Thr Cys Val Val Tyr Ala Pro Val Glu \*  
 730 735 740

(2) 서열 번호 160에 대한 정보:

(i) 서열 특징:

- (A) 길이: 2763개 엔기쌍
- (B) 유형: 핵산
- (C) 가닥수: 단일
- (D) 토플로지: 관련없음

(ii) 분자 유형: mRNA

(iii) 가정: 없음

(vi) 최초 소오스:

- (A) 유기체: Zea mays

(ix) 특징:

- (A) 명칭/키: 트렌싯 펩티드
- (B) 위치: 2..190

(ix) 특징:

- (A) 명칭/키: mat. 펩티드
- (B) 위치: 195..2467

(ix) 특징:

- (A) 명칭/키: CDS
- (B) 위치: 2..2470

(xi) 서열 설명: 서열 번호 16:

G CTG TGC CTC GTG TCG CCC TCT TCC TCG CCG ACT CCG CTT CCG CCG	46
Leu Cys Leu Val Ser Pro Ser Ser Pro Thr Pro Leu Pro Pro	
-63 -60 -55 -50	
CCG CGG CGC TCT CGC TCG CAT GCT GAT CGG GCG GCA CCG CCG GGG ATC	94
Pro Arg Arg Ser Arg Ser His Ala Asp Arg Ala Ala Pro Pro Gly Ile	
-45 -40 -35	
GCG GGT GGC GGC AAT GTG CGC CTG AGT GTG TTG TCT GTC CAG TGC AAG	142
Ala Gly Gly Asn Val Arg Leu Ser Val Leu Ser Val Gln Cys Lys	
-30 -25 -20	
GCT CGC CGG TCA GGG GTG CGG AAG GTC AAG AGC AAA TTC GCC ACT GCA	190
Ala Arg Arg Ser Gly Val Arg Lys Val Lys Ser Lys Phe Ala Thr Ala	
-15 -10 -5	
GCT ACT GTG CAA GAA GAT AAA ACT ATG GCA ACT GCC AAA GGC GAT GTC	238
Ala Thr Val Gln Glu Asp Lys Thr Met Ala Thr Ala Lys Gly Asp Val	
1 5 10 15	
GAC CAT CTC CCC ATA TAC GAC CTG GAC CCC AAG CTG GAG ATA TTC AAG	286
Asp His Leu Pro Ile Tyr Asp Leu Asp Pro Lys Leu Glu Ile Phe Lys	
20 25 30	
GAC CAT TTC AGG TAC CGG ATG AAA AGA TTC CTA GAG CAG AAA GGA TCA	334
Asp His Phe Arg Tyr Arg Met Lys Arg Phe Leu Glu Gln Lys Gly Ser	
35 40 45	

ATT GAA GAA AAT GAG GGA AGT CTT GAA TCT TTT TCT AAA GGC TAT TTG Ile Glu Glu Asn Glu Gly Ser Leu Glu Ser Phe Ser Lys Gly Tyr Leu	50	55	60	382	
AAA TTT GGG ATT AAT ACA AAT GAG GAT GGA ACT GTA TAT CGT GAA TGG Lys Phe Gly Ile Asn Thr Asn Glu Asp Gly Thr Val Tyr Arg Glu Trp	65	70	75	80	430
GCA CCT GCT GCG CAG GAG GCA GAG CTT ATT GGT GAC TTC AAT GAC TGG Ala Pro Ala Ala Gln Glu Ala Glu Leu Ile Gly Asp Phe Asn Asp Trp	85	90	95	478	
AAT GGT GCA AAC CAT AAG ATG GAG AAG GAT AAA TTT GGT GTT TGG TCG Asn Gly Ala Asn His Lys Met Glu Lys Asp Lys Phe Gly Val Trp Ser	100	105	110	526	
ATC AAA ATT GAC CAT GTC AAA GGG AAA CCT GCC ATC CCT CAC AAT TCC Ile Lys Ile Asp His Val Lys Gly Lys Pro Ala Ile Pro His Asn Ser	115	120	125	574	
AAG GTT AAA TTT CGC TTT CTA CAT GGT GGA GTA TGG GTT GAT CGT ATT Lys Val Lys Phe Arg Phe Leu His Gly Gly Val Trp Val Asp Arg Ile	130	135	140	622	
CCA GCA TTG ATT CGT TAT GCG ACT GTT GAT GCC TCT AAA TTT GGA GCT Pro Ala Leu Ile Arg Tyr Ala Thr Val Asp Ala Ser Lys Phe Gly Ala	145	150	155	160	670
CCC TAT GAT GGT GTT CAT TGG GAT CCT CCT GCT TCT GAA AGG TAC ACA Pro Tyr Asp Gly Val His Trp Asp Pro Pro Ala Ser Glu Arg Tyr Thr	165	170	175	718	
TTT AAG CAT CCT CGG CCT TCA AAG CCT GCT GCT CCA CGT ATC TAT GAA Phe Lys His Pro Arg Pro Ser Lys Pro Ala Ala Pro Arg Ile Tyr Glu	180	185	190	766	
GCC CAT GTA GGT ATG AGT GGT GAA AAG CCA GCA GTA AGC ACA TAT AGG Ala His Val Gly Met Ser Gly Glu Lys Pro Ala Val Ser Thr Tyr Arg	195	200	205	814	
GAA TTT GCA GAC AAT GTG TTG CCA CGC ATA CGA GCA AAT AAC TAC AAC Glu Phe Ala Asp Asn Val Leu Pro Arg Ile Arg Ala Asn Asn Tyr Asn	210	215	220	862	
ACA GTT CAG TTG ATG GCA GTT ATG GAG CAT TCG TAC TAT GCT TCT TTC				910	

Thr Val Gln Leu Met Ala Val Met Glu His Ser Tyr Tyr Ala Ser Phe			
225	230	235	240
GGG TAC CAT GTG ACA AAT TTC TTT GCG GTT AGC AGC AGA TCA GGC ACA			958
Gly Tyr His Val Thr Asn Phe Phe Ala Val Ser Ser Arg Ser Gly Thr			
245	250	255	
CCA GAG GAC CTC AAA TAT CTT GTT GAT AAG GCA CAC AGT TTG GGT TTG			1006
Pro Glu Asp Leu Lys Tyr Leu Val Asp Lys Ala His Ser Leu Gly Leu			
260	265	270	
CGA GTT CTG ATG GAT GTT GTC CAT AGC CAT GCA AGT AAT AAT GTC ACA			1054
Arg Val Leu Met Asp Val Val His Ser His Ala Ser Asn Asn Val Thr			
275	280	285	
GAT GGT TTA AAT GGC TAT GAT GTT GGA CAA AGC ACC CAA GAG TCC TAT			1102
Asp Gly Leu Asn Gly Tyr Asp Val Gly Gln Ser Thr Gln Glu Ser Tyr			
290	295	300	
TTT CAT GCG GGA GAT AGA GGT TAT CAT AAA CTT TGG GAT AGT CGG CTG			1150
Phe His Ala Gly Asp Arg Gly Tyr His Lys Leu Trp Asp Ser Arg Leu			
305	310	315	320
TTC AAC TAT GCT AAC TGG GAG GTA TTA AGG TTT CTT CTT TCT AAC CTG			1198
Phe Asn Tyr Ala Asn Trp Glu Val Leu Arg Phe Leu Leu Ser Asn Leu			
325	330	335	
AGA TAT TGG TTG GAT GAA TTC ATG TTT GAT GGC TTC CGA TTT GAT GGA			1246
Arg Tyr Trp Leu Asp Glu Phe Met Phe Asp Gly Phe Arg Phe Asp Gly			
340	345	350	
GTT ACA TCA ATG CTG TAT CAT CAC CAT GGT ATC AAT GTG GGG TTT ACT			1294
Val Thr Ser Met Leu Tyr His His Gly Ile Asn Val Gly Phe Thr			
355	360	365	
GGA AAC TAC CAG GAA TAT TTC AGT TTG GAC ACA GCT GTG GAT GCA GTT			1342
Gly Asn Tyr Gln Glu Tyr Phe Ser Leu Asp Thr Ala Val Asp Ala Val			
370	375	380	
GTT TAC ATG ATG CTT GCA AAC CAT TTA ATG CAC AAA CTC TTG CCA GAA			1390
Val Tyr Met Met Leu Ala Asn His Leu Met His Lys Leu Leu Pro Glu			
385	390	395	400
GCA ACT GTT GTT GCT GAA GAT GTT TCA GGC ATG CCG GTC CTT TGC CGG			1438
Ala Thr Val Val Ala Glu Asp Val Ser Gly Met Pro Val Leu Cys Arg			

405

410

415

CCA GTT GAT GAA GGT GGG GTT GGG TTT GAC TAT CGC CTG GCA ATG GCT			1486
Pro Val Asp Glu Gly Gly Val Gly Phe Asp Tyr Arg Leu Ala Met Ala			
420	425	430	
ATC CCT GAT AGA TGG ATT GAC TAC CTG AAG AAT AAA GAT GAC TCT GAG			1534
Ile Pro Asp Arg Trp Ile Asp Tyr Leu Lys Asn Lys Asp Asp Ser Glu			
435	440	445	
TGG TCG ATG GGT GAA ATA GCG CAT ACT TTG ACT AAC AGG AGA TAT ACT			1582
Trp Ser Met Gly Glu Ile Ala His Thr Leu Thr Asn Arg Arg Tyr Thr			
450	455	460	
GAA AAA TGC ATC GCA TAT GCT GAG AGC CAT GAT CAG TCT ATT GTT GGC			1630
Glu Lys Cys Ile Ala Tyr Ala Glu Ser His Asp Gln Ser Ile Val Gly			
465	470	475	480
GAC AAA ACT ATT GCA TTT CTC CTG ATG GAC AAG GAA ATG TAC ACT GGC			1678
Asp Lys Thr Ile Ala Phe Leu Leu Met Asp Lys Glu Met Tyr Thr Gly			
485	490	495	
ATG TCA GAC TTG CAG CCT GCT TCA CCT ACA ATT GAT CGA GGG ATT GCA			1726
Met Ser Asp Leu Gln Pro Ala Ser Pro Thr Ile Asp Arg Gly Ile Ala			
500	505	510	
CTC CAA AAG ATG ATT CAC TTC ATC ACA ATG GCC CTT GGA GGT GAT GGC			1774
Leu Gln Lys Met Ile His Phe Ile Thr Met Ala Leu Gly Gly Asp Gly			
515	520	525	
TAC TTG AAT TTT ATG GGA AAT GAG TTT GGT CAC CCA GAA TGG ATT GAC			1822
Tyr Leu Asn Phe Met Gly Asn Glu Phe Gly His Pro Glu Trp Ile Asp			
530	535	540	
TTT CCA AGA GAA GGG AAC AAC TGG AGC TAT GAT AAA TGC AGA CGA CAG			1870
Phe Pro Arg Glu Gly Asn Asn Trp Ser Tyr Asp Lys Cys Arg Arg Gln			
545	550	555	560
TGG AGC CTT GTG GAC ACT GAT CAC TTG CGG TAC AAG TAC ATG AAT GCG			1918
Trp Ser Leu Val Asp Thr Asp His Leu Arg Tyr Lys Tyr Met Asn Ala			
565	570	575	
TTT GAC CAA GCG ATG AAT GCG CTC GAT GAG AGA TTT TCC TTC CTT TCG			1966
Phe Asp Gln Ala Met Asn Ala Leu Asp Glu Arg Phe Ser Phe Leu Ser			
580	585	590	

TCG TCA AAG CAG ATC GTC AGC GAC ATG AAC GAT GAG GAA AAG GTT ATT Ser Ser Lys Gln Ile Val Ser Asp Met Asn Asp Glu Glu Lys Val Ile 595 600 605	2014
GTC TTT GAA CGT GGA GAT TTA GTT TTT GTT TTC AAT TTC CAT CCC AAG Val Phe Glu Arg Gly Asp Leu Val Phe Val Phe Asn Phe His Pro Lys 610 615 620	2062
AAA ACT TAC GAG GGC TAC AAA GTG GGA TGC GAT TTG CCT GGG AAA TAC Lys Thr Tyr Glu Gly Tyr Lys Val Gly Cys Asp Leu Pro Gly Lys Tyr 625 630 635 640	2110
AGA GTA GCC CTG GAC TCT GAT GCT CTG GTC TTC GGT GGA CAT GGA AGA Arg Val Ala Leu Asp Ser Asp Ala Leu Val Phe Gly Gly His Gly Arg 645 650 655	2158
GTT GGC CAC GAC GTG GAT CAC TTC ACG TCG CCT GAA GGG GTG CCA GGG Val Gly His Asp Val Asp His Phe Thr Ser Pro Glu Gly Val Pro Gly 660 665 670	2206
GTG CCC GAA ACG AAC TTC AAC AAC CGG CCG AAC TCG TTC AAA GTC CTT Val Pro Glu Thr Asn Phe Asn Asn Arg Pro Asn Ser Phe Lys Val Leu 675 680 685	2254
TCT CCG CCC CGC ACC TGT GTG GCT TAT TAC CGT GTA GAC GAA GCA GGG Ser Pro Pro Arg Thr Cys Val Ala Tyr Tyr Arg Val Asp Glu Ala Gly 690 695 700	2302
GCT GGA CGA CGT CTT CAC GCG AAA GCA GAG ACA GGA AAG ACG TCT CCA Ala Gly Arg Arg Leu His Ala Lys Ala Glu Thr Gly Lys Thr Ser Pro 705 710 715 720	2350
GCA GAG AGC ATC GAC GTC AAA GCT TCC AGA GCT AGT AGC AAA GAA GAC Ala Glu Ser Ile Asp Val Lys Ala Ser Arg Ala Ser Ser Lys Glu Asp 725 730 735	2398
AAG GAG GCA ACG GCT GGT GGC AAG AAG GGA TGG AAG TTT GCG CGG CAG Lys Glu Ala Thr Ala Gly Lys Lys Gly Trp Lys Phe Ala Arg Gln 740 745 750	2446
CCA TCC GAT CAA GAT ACC AAA TGA AGCCACGAGT CCTTGGTGAG GACTGGACTG Pro Ser Asp Gln Asp Thr Lys *	2500
755 760	
GCTGCCGGCG CCCTGTTAGT AGTCCTGCTC TACTGGACTA GCCGCCGCTG GCGCCCTTGG	2560
AACGCTCCCTT CCCTGTAGCT TGCAGGCGAC TGGTGTCTCA TCACCGAGCA GGCAGGCAC	2620
GCTTGTATAG CTTTCTAGA ATAATAATCA GGGATGGATG GATGGTGTGT ATTGGCTATC	2680
TGGCTAGACG TGCATGTGCC CAGTTGTAT GTACAGGAGC AGTTCCCGTC CAGAATAAAA	2740
AAAAACTTGT TGGGGGGTTT TTC	2763

(2) 서열 번호 170에 대한 정보:

( i ) 서열 특징:

(A)길이: 823개 아미노산

(B)유형: 아미노산

(D)토플로지: 선형

( ii ) 문자 유형: 단백질

(xi) 서열 설명: 서열 번호 17:

Leu Cys Leu Val Ser Pro Ser Ser Ser Pro Thr Pro Leu Pro Pro Pro  
 -63 -60 -55 -50

Arg Arg Ser Arg Ser His Ala Asp Arg Ala Ala Pro Pro Gly Ile Ala  
 -45 -40 -35

Gly Gly Gly Asn Val Arg Leu Ser Val Leu Ser Val Gln Cys Lys Ala  
 -30 -25 -20

Arg Arg Ser Gly Val Arg Lys Val Lys Ser Lys Phe Ala Thr Ala Ala  
 -15 -10 -5 1

Thr Val Gln Glu Asp Lys Thr Met Ala Thr Ala Lys Gly Asp Val Asp  
 5 10 15

His Leu Pro Ile Tyr Asp Leu Asp Pro Lys Leu Glu Ile Phe Lys Asp  
 20 25 30

His Phe Arg Tyr Arg Met Lys Arg Phe Leu Glu Gln Lys Gly Ser Ile  
 35 40 45

Glu Glu Asn Glu Gly Ser Leu Glu Ser Phe Ser Lys Gly Tyr Leu Lys  
 50 55 60 65

Phe Gly Ile Asn Thr Asn Glu Asp Gly Thr Val Tyr Arg Glu Trp Ala

70

75

80

Pro Ala Ala Gln Glu Ala Glu Leu Ile Gly Asp Phe Asn Asp Trp Asn  
 85 90 95

Gly Ala Asn His Lys Met Glu Lys Asp Lys Phe Gly Val Trp Ser Ile  
 100 105 110

Lys Ile Asp His Val Lys Gly Lys Pro Ala Ile Pro His Asn Ser Lys  
 115 120 125

Val Lys Phe Arg Phe Leu His Gly Gly Val Trp Val Asp Arg Ile Pro  
 130 135 140 145

Ala Leu Ile Arg Tyr Ala Thr Val Asp Ala Ser Lys Phe Gly Ala Pro  
 150 155 160

Tyr Asp Gly Val His Trp Asp Pro Pro Ala Ser Glu Arg Tyr Thr Phe  
 165 170 175

Lys His Pro Arg Pro Ser Lys Pro Ala Ala Pro Arg Ile Tyr Glu Ala  
 180 185 190

His Val Gly Met Ser Gly Glu Lys Pro Ala Val Ser Thr Tyr Arg Glu  
 195 200 205

Phe Ala Asp Asn Val Leu Pro Arg Ile Arg Ala Asn Asn Tyr Asn Thr  
 210 215 220 225

Val Gln Leu Met Ala Val Met Glu His Ser Tyr Tyr Ala Ser Phe Gly  
 230 235 240

Tyr His Val Thr Asn Phe Phe Ala Val Ser Ser Arg Ser Gly Thr Pro  
 245 250 255

Glu Asp Leu Lys Tyr Leu Val Asp Lys Ala His Ser Leu Gly Leu Arg  
 260 265 270

Val Leu Met Asp Val Val His Ser His Ala Ser Asn Asn Val Thr Asp  
 275 280 285

Gly Leu Asn Gly Tyr Asp Val Gly Gln Ser Thr Gln Glu Ser Tyr Phe  
 290 295 300 305

His Ala Gly Asp Arg Gly Tyr His Lys Leu Trp Asp Ser Arg Leu Phe

310

315

320

Asn Tyr Ala Asn Trp Glu Val Leu Arg Phe Leu Leu Ser Asn Leu Arg  
 325 330 335

Tyr Trp Leu Asp Glu Phe Met Phe Asp Gly Phe Arg Phe Asp Gly Val  
 340 345 350

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

Asn Tyr Gln Glu Tyr Phe Ser Leu Asp Thr Ala Val Asp Ala Val Val  
 370 375 380 385

Tyr Met Met Leu Ala Asn His Leu Met His Lys Leu Leu Pro Glu Ala  
 390 395 400

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

Val Asp Glu Gly Val Gly Phe Asp Tyr Arg Leu Ala Met Ala Ile  
 420 425 430

Pro Asp Arg Trp Ile Asp Tyr Leu Lys Asn Lys Asp Asp Ser Glu Trp  
 435 440 445

Ser Met Gly Glu Ile Ala His Thr Leu Thr Asn Arg Arg Tyr Thr Glu  
 450 455 460 465

Lys Cys Ile Ala Tyr Ala Glu Ser His Asp Gln Ser Ile Val Gly Asp  
 470 475 480

Lys Thr Ile Ala Phe Leu Leu Met Asp Lys Glu Met Tyr Thr Gly Met  
 485 490 495

Ser Asp Leu Gln Pro Ala Ser Pro Thr Ile Asp Arg Gly Ile Ala Leu  
 500 505 510

Gln Lys Met Ile His Phe Ile Thr Met Ala Leu Gly Gly Asp Gly Tyr  
 515 520 525

Leu Asn Phe Met Gly Asn Glu Phe Gly His Pro Glu Trp Ile Asp Phe  
 530 535 540 545

Pro Arg Glu Gly Asn Asn Trp Ser Tyr Asp Lys Cys Arg Arg Gln Trp

550	555	560
<b>Ser Leu Val Asp Thr Asp His Leu Arg Tyr Lys Tyr Met Asn Ala Phe</b>		
565	570	575
<b>Asp Gln Ala Met Asn Ala Leu Asp Glu Arg Phe Ser Phe Leu Ser Ser</b>		
580	585	590
<b>Ser Lys Gln Ile Val Ser Asp Met Asn Asp Glu Glu Lys Val Ile Val</b>		
595	600	605
<b>Phe Glu Arg Gly Asp Leu Val Phe Val Phe Asn Phe His Pro Lys Lys</b>		
610	615	620
625		
<b>Thr Tyr Glu Gly Tyr Lys Val Gly Cys Asp Leu Pro Gly Lys Tyr Arg</b>		
630	635	640
<b>Val Ala Leu Asp Ser Asp Ala Leu Val Phe Gly Gly His Gly Arg Val</b>		
645	650	655
<b>Gly His Asp Val Asp His Phe Thr Ser Pro Glu Gly Val Pro Gly Val</b>		
660	665	670
<b>Pro Glu Thr Asn Phe Asn Asn Arg Pro Asn Ser Phe Lys Val Leu Ser</b>		
675	680	685
<b>Pro Pro Arg Thr Cys Val Ala Tyr Tyr Arg Val Asp Glu Ala Gly Ala</b>		
690	695	700
705		
<b>Gly Arg Arg Leu His Ala Lys Ala Glu Thr Gly Lys Thr Ser Pro Ala</b>		
710	715	720
<b>Glu Ser Ile Asp Val Lys Ala Ser Arg Ala Ser Ser Lys Glu Asp Lys</b>		
725	730	735
<b>Glu Ala Thr Ala Gly Gly Lys Lys Gly Trp Lys Phe Ala Arg Gln Pro</b>		
740	745	750
<b>Ser Asp Gln Asp Thr Lys *</b>		
755	760	

(2) 서열 번호 180에 대한 정보:

(i) 서열 특징:

- (A)길이: 153개 염기쌍
- (B)유형: 핵산
- (C)가닥수: 단일
- (D)토플로지: 관련없음

(ii) 분자 유형: cDNA 에서 mRNA

(iii) 가정: 없음

(vi) 최초 소오스:

(A)유기체: Zea mays

(ix)특징:

- (A) 명칭/URI: CDS
- (B) 위치: 1..153

## (xi) 서열 설명: 서열 번호 18:

ATG GCG ACG CCC TCG GCC GTC GGC GCG TGC CTC CTC GCG CGG	48
Met Ala Thr Pro Ser Ala Val Gly Ala Ala Cys Leu Leu Leu Ala Arg	
765 770 775	
GCC GCC TGG CCG GCC GCC GTC GGC GAC CGG GCG CGC CCG CGG AGG CTC	96
Ala Ala Trp Pro Ala Ala Val Gly Asp Arg Ala Arg Pro Arg Arg Leu	
780 785 790	
CAG CGC GTG CTG CGC CGC CGG TGC GTC GCG GAG CTG AGC AGG GAG GGG	144
Gln Arg Val Leu Arg Arg Cys Val Ala Glu Leu Ser Arg Glu Gly	
795 800 805	
CCC CAT ATG	153
Pro His Met	
810	

## (2) 서열 번호 19에 대한 정보:

## (i) 서열 특징:

(A) 길이: 51개 아미노산

(B) 유형: 아미노산

(D) 토플로지: 선형

## (ii) 분자 유형: 단백질

## (xi) 서열 설명: 서열 번호 19:

Met Ala Thr Pro Ser Ala Val Gly Ala Ala Cys Leu Leu Leu Ala Arg	
1 5 10 15	
Ala Ala Trp Pro Ala Ala Val Gly Asp Arg Ala Arg Pro Arg Arg Leu	
20 25 30	
Gln Arg Val Leu Arg Arg Cys Val Ala Glu Leu Ser Arg Glu Gly	
35 40 45	
Pro His Met	
50	

## (2) 서열 번호 20에 대한 정보:

## (i) 서열 특징:

(A) 길이: 1620개 염기쌍

(B) 유형: 핵산

(C) 가닥수: 이중

(D) 토플로지: 관련없음

## (ii) 분자 유형: cDNA에서 mRNA

## (iii) 가정: 없음

## (ix) 특징:

(A) 명칭/키: CDS

(B) 위치: 1..1620

(xi) 서열 설명: 서열 번호 20:

TGC GTC GCG GAG CTG AGC AGG GAG GAC CTC GGT CTC GAA CCT GAA GGG	48	
Cys Val Ala Glu Leu Ser Arg Glu Asp Leu Gly Leu Glu Pro Glu Gly		
55	60	65
ATT GCT GAA GGT TCC ATC GAT AAC ACA GTA GTT GTG GCA AGT GAG CAA	96	
Ile Ala Glu Gly Ser Ile Asp Asn Thr Val Val Val Ala Ser Glu Gln		
70	75	80
GAT TCT GAG ATT GTG GTT GGA AAG GAG CAA GCT CGA GCT AAA GTA ACA	144	
Asp Ser Glu Ile Val Val Gly Lys Glu Gln Ala Arg Ala Lys Val Thr		
85	90	95
CAA AGC ATT GTC TTT GTA ACC GGC GAA GCT TCT CCT TAT GCA AAG TCT	192	

Gln Ser Ile Val Phe Val Thr Gly Glu Ala Ser Pro Tyr Ala Lys Ser			
100	105	110	115
GGG GGT CTA GGA GAT GTT TGT GGT TCA TTG CCA GTT GCT CTT GCT GCT			240
Gly Gly Leu Gly Asp Val Cys Gly Ser Leu Pro Val Ala Leu Ala Ala			
120	125	130	
CGT GGT CAC CGT GTG ATG GTT GTA ATG CCC AGA TAT TTA AAT GGT ACC			288
Arg Gly His Arg Val Met Val Val Met Pro Arg Tyr Leu Asn Gly Thr			
135	140	145	
TCC GAT AAG AAT TAT GCA AAT GCA TTT TAC ACA GAA AAA CAC ATT CGG			336
Ser Asp Lys Asn Tyr Ala Asn Ala Phe Tyr Thr Glu Lys His Ile Arg			
150	155	160	
ATT CCA TGC TTT GGC GGT GAA CAT GAA GTT ACC TTC TTC CAT GAG TAT			384
Ile Pro Cys Phe Gly Gly Glu His Glu Val Thr Phe Phe His Glu Tyr			
165	170	175	
AGA GAT TCA GTT GAC TGG GTG TTT GTT GAT CAT CCC TCA TAT CAC AGA			432
Arg Asp Ser Val Asp Trp Val Phe Val Asp His Pro Ser Tyr His Arg			
180	185	190	195
CCT GGA AAT TTA TAT GGA GAT AAG TTT GGT GCT TTT GGT GAT AAT CAG			480
Pro Gly Asn Leu Tyr Gly Asp Lys Phe Gly Ala Phe Gly Asp Asn Gln			
200	205	210	
TTC AGA TAC ACA CTC CTT TGC TAT GCT GCA TGT GAG GCT CCT TTG ATC			528
Phe Arg Tyr Thr Leu Leu Cys Tyr Ala Ala Cys Glu Ala Pro Leu Ile			
215	220	225	
CTT GAA TTG GGA GGA TAT ATT TAT GGA CAG AAT TGC ATG TTT GTT GTC			576
Leu Glu Leu Gly Gly Tyr Ile Tyr Gly Gln Asn Cys Met Phe Val Val			
230	235	240	
AAT GAT TGG CAT GCC AGT CTA GTG CCA GTC CTT CTT GCT GCA AAA TAT			624
Asn Asp Trp His Ala Ser Leu Val Pro Val Leu Leu Ala Ala Lys Tyr			
245	250	255	
AGA CCA TAT GGT GTT TAT AAA GAC TCC CGC AGC ATT CTT GTA ATA CAT			672
Arg Pro Tyr Gly Val Tyr Lys Asp Ser Arg Ser Ile Leu Val Ile His			
260	265	270	275
AAT TTA GCA CAT CAG GGT GTA GAG CCT GCA AGC ACA TAT CCT GAC CTT			720
Asn Leu Ala His Gln Gly Val Glu Pro Ala Ser Thr Tyr Pro Asp Leu			

280

285

290

GGG TTG CCA CCT GAA TGG TAT GGA GCT CTG GAG TGG GTA TTC CCT GAA Gly Leu Pro Pro Glu Trp Tyr Gly Ala Leu Glu Trp Val Phe Pro Glu 295	300	305	768	
TGG GCG AGG AGG CAT GCC CTT GAC AAG GGT GAG GCA GTT AAT TTT TTG Trp Ala Arg Arg His Ala Leu Asp Lys Gly Glu Ala Val Asn Phe Leu 310	315	320	816	
AAA GGT GCA GTT GTG ACA GCA GAT CGA ATC GTG ACT GTC AGT AAG GGT Lys Gly Ala Val Val Thr Ala Asp Arg Ile Val Thr Val Ser Lys Gly 325	330	335	864	
TAT TCG TGG GAG GTC ACA ACT GCT GAA GGT GGA CAG GGC CTC AAT GAG Tyr Ser Trp Glu Val Thr Ala Glu Gly Gly Gln Gly Leu Asn Glu 340	345	350	355	912
CTC TTA AGC TCC AGA AAG AGT GTA TTA AAC GGA ATT GTA AAT GGA ATT Leu Leu Ser Ser Arg Lys Ser Val Leu Asn Gly Ile Val Asn Gly Ile 360	365	370	960	
GAC ATT AAT GAT TGG AAC CCT GCC ACA GAC AAA TGT ATC CCC TGT CAT Asp Ile Asn Asp Trp Asn Pro Ala Thr Asp Lys Cys Ile Pro Cys His 375	380	385	1008	
TAT TCT GTT GAT GAC CTC TCT GGA AAG GCC AAA TGT AAA GGT GCA TTG Tyr Ser Val Asp Asp Leu Ser Gly Lys Ala Lys Cys Lys Gly Ala Leu 390	395	400	1056	
CAG AAG GAG CTG GGT TTA CCT ATA AGG CCT GAT GTT CCT CTG ATT GGC Gln Lys Glu Leu Gly Leu Pro Ile Arg Pro Asp Val Pro Leu Ile Gly 405	410	415	1104	
TTT ATT GGA AGG TTG GAT TAT CAG AAA GGC ATT GAT CTC ATT CAA CTT Phe Ile Gly Arg Leu Asp Tyr Gln Lys Gly Ile Asp Leu Ile Gln Leu 420	425	430	435	1152
ATC ATA CCA GAT CTC ATG CGG GAA GAT GTT CAA TTT GTC ATG CTT GGA Ile Ile Pro Asp Leu Met Arg Glu Asp Val Gln Phe Val Met Leu Gly 440	445	450	1200	
TCT GGT GAC CCA GAG CTT GAA GAT TGG ATG AGA TCT ACA GAG TCG ATC Ser Gly Asp Pro Glu Leu Glu Asp Trp Met Arg Ser Thr Glu Ser Ile 455	460	465	1248	

TTC AAG GAT AAA TTT CGT GGA TGG GTT GGA TTT AGT GTT CCA GTT TCC Phe Lys Asp Lys Phe Arg Gly Trp Val Gly Phe Ser Val Val Pro Val Ser 470 475 480	1296
CAC CGA ATA ACT GCC GGC TGC GAT ATA TTG TTA ATG CCA TCC AGA TTC His Arg Ile Thr Ala Gly Cys Asp Ile Leu Leu Met Pro Ser Arg Phe 485 490 495	1344
GAA CCT TGT GGT CTC AAT CAG CTA TAT GCT ATG CAG TAT GGC ACA GTT Glu Pro Cys Gly Leu Asn Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val 500 505 510 515	1392
CCT GTT GTC CAT GCA ACT GGG GGC CTT AGA GAT ACC GTG GAG AAC TTC Pro Val Val His Ala Thr Gly Leu Arg Asp Thr Val Glu Asn Phe 520 525 530	1440
AAC CCT TTC GGT GAG AAT GGA GAG CAG GGT ACA GGG TGG GCA TTC GCA Asn Pro Phe Gly Glu Asn Gly Glu Gln Gly Thr Gly Trp Ala Phe Ala 535 540 545	1488
CCC CTA ACC ACA GAA AAC ATG TTT GTG GAC ATT GCG AAC TGC AAT ATC Pro Leu Thr Thr Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile 550 555 560	1536
TAC ATA CAG GGA ACA CAA GTC CTC CTG GGA AGG GCT AAT GAA GCG AGG Tyr Ile Gln Gly Thr Gln Val Leu Leu Gly Arg Ala Asn Glu Ala Arg 565 570 575	1584
CAT GTC AAA AGA CTT CAC GTG GGA CCA TGC CGC TGA His Val Lys Arg Leu His Val Gly Pro Cys Arg *	1620
580 585 590	

(2) 서열 번호 21에 대한 정보:

(i) 서열 특징:

- (A)길이: 540개 아미노산
- (B)유형: 아미노산
- (D)토플로지: 선형

(ii) 문자 유형: 단백질

(xi) 서열 설명: 서열 번호 21:

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

245

250

255

Trp Ala Arg Arg His Ala Leu Asp Lys Gly Glu Ala Val Asn Phe Leu  
 260 265 270

Lys Gly Ala Val Val Thr Ala Asp Arg Ile Val Thr Val Ser Lys Gly  
 275 280 285

Tyr Ser Trp Glu Val Thr Thr Ala Glu Gly Gly Gln Gly Leu Asn Glu  
 290 295 300

Leu Leu Ser Ser Arg Lys Ser Val Leu Asn Gly Ile Val Asn Gly Ile  
 305 310 315 320

Asp Ile Asn Asp Trp Asn Pro Ala Thr Asp Lys Cys Ile Pro Cys His  
 325 330 335

Tyr Ser Val Asp Asp Leu Ser Gly Lys Ala Lys Cys Lys Gly Ala Leu  
 340 345 350

Gln Lys Glu Leu Gly Leu Pro Ile Arg Pro Asp Val Pro Leu Ile Gly  
 355 360 365

Phe Ile Gly Arg Leu Asp Tyr Gln Lys Gly Ile Asp Leu Ile Gln Leu  
 370 375 380

Ile Ile Pro Asp Leu Met Arg Glu Asp Val Gln Phe Val Met Leu Gly  
 385 390 395 400

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

Phe Lys Asp Lys Phe Arg Gly Trp Val Gly Phe Ser Val Pro Val Ser  
 420 425 430

His Arg Ile Thr Ala Gly Cys Asp Ile Leu Leu Met Pro Ser Arg Phe  
 435 440 445

Glu Pro Cys Gly Leu Asn Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val  
 450 455 460

Pro Val Val His Ala Thr Gly Leu Arg Asp Thr Val Glu Asn Phe  
 465 470 475 480

Asn Pro Phe Gly Glu Asn Gly Gln Gly Thr Gly Trp Ala Phe Ala

485

490

495

Pro Leu Thr Thr Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile  
 500 505 510

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

His Val Lys Arg Leu His Val Gly Pro Cys Arg \*

530 535 540

(2) 서열 번호 22에 대한 정보:

(i) 서열 특징:

(A)길이: 30개 엔티티

(B)유형: 헥사

(C)가닥수: 단일

(D)토플로지: 선형

(ii) 문자 유형: 다른 헥사

(A)설명: /desc='올리고누클레오티드'

(iii) 가정: 없음

(xi) 서열 설명: 서열 번호 22:

GTGGATCCAT GGCAGCGCCC TCGGCCGTGG

(2)서열 번호 23에 대한 정보:

( i ) 서열 특징:

(A)길이: 35개 염기쌍

(B)유형: 핵산

(C)가닥수: 단일

(D)토플로지: 선형

( ii ) 분자 유형: 다른 핵산

(A)설명: /desc='올리고누클레오티드'

(xi) 서열 설명: 서열 번호 23:

CTGAATTCCA TATGGGGCCC CTCCCTGCTC AGCTC

(2)서열 번호 24에 대한 정보:

( i ) 서열 특징:

(A)길이: 36개 염기쌍

(B)유형: 핵산

(C)가닥수: 단일

(D)토플로지: 선형

( ii ) 분자 유형: 다른 핵산

(A)설명: /desc='올리고누클레오티드'

(xi) 서열 설명: 서열 번호 24:

CTCTGAGCTC AAGCTTGCTA CTTCTTTCC TTAATG

(2)서열 번호 25에 대한 정보:

( i ) 서열 특징:

(A)길이: 29개 염기쌍

(B)유형: 핵산

(C)가닥수: 단일

(D)토플로지: 선형

( ii ) 분자 유형: 다른 핵산

(A)설명: /desc='올리고누클레오티드'

(xi) 서열 설명: 서열 번호 25:

GTCTCCCGGG TGGTGTCCCTT GCTTCCTAG

(2)서열 번호 26에 대한 정보:

( i ) 서열 특징:

(A)길이: 53개 염기쌍

(B)유형: 핵산

(C)가닥수: 이중

(D)토플로지: 관련없음

( ii ) 분자 유형: cDNA 에서 mRNA

(iii) 가정: 없음

(xi) 서열 설명: 서열 번호 26:

TGCGTCCGG AGCTGAGCAG GGAGGTCTCC GCGGTGGTGT CCTTGCTTCC TAG

(2)서열 번호 27에 대한 정보:

## ( i ) 서열 특징:

- (A)길이: 8개 아미노산
- (B)유형: 아미노산
- (C)가닥수: 단일
- (D)토플로지: 관련없음

## ( ii ) 분자 유형: 펩티드

## (xi) 서열 설명: 서열 번호 27:

Cys Val Ala Glu Leu Ser Arg Glu

1 5

## (2) 서열 번호 28에 대한 정보:

## ( i ) 서열 특징:

- (A)길이: 16개 염기쌍
- (B)유형: 핵산
- (C)가닥수: 이중
- (D)토플로지: 관련없음

## ( ii ) 분자 유형: cDNA 에서 mRNA

## (xi) 서열 설명: 서열 번호 28:

AGAGAGAGAG AGAGAG

## (2) 서열 번호 29에 대한 정보:

## ( i ) 서열 특징:

- (A)길이: 36개 염기쌍
- (B)유형: 핵산
- (C)가닥수: 이중
- (D)토플로지: 관련없음

## ( ii ) 분자 유형: cDNA 에서 mRNA

## (iii) 가정: 없음

## (xi) 서열 설명: 서열 번호 29:

AAGAAGAAGA AGAAGAAGAA GAAGAAGAAG AAGAAG

## (2) 서열 번호 30에 대한 정보:

## ( i ) 서열 특징:

- (A)길이: 18개 염기쌍
- (B)유형: 핵산
- (C)가닥수: 이중
- (D)토플로지: 관련없음

## ( ii ) 분자 유형: cDNA 에서 mRNA

## (iii) 가정: 없음

## (xi) 서열 설명: 서열 번호 30:

AAAAAAAAAA AAAAAAAA

## (2) 서열 번호 31에 대한 정보:

## ( i ) 서열 특징:

- (A)길이: 11개 염기쌍
- (B)유형: 핵산
- (C)가닥수: 단일
- (D)토플로지: 관련없음

## ( ii ) 분자 유형: 다른 핵산

(A)설명: /desc='올리고누클레오티드'

(iii) 가정: 없음

(xi) 서열 설명: 서열 번호 31:

AGATAATGCA G

(2) 서열 번호 32에 대한 정보:

(i) 서열 특징:

(A)길이: 10개 염기쌍

(B)유형: 핵산

(C)가닥수: 단일

(D)토플로지: 관련없음

(ii) 분자 유형: 다른 핵산

(A)설명: /desc='올리고누클레오티드'

(iii) 가정: 없음

(xi) 서열 설명: 서열 번호 32:

AACAAATGGCT

(2) 서열 번호 33에 대한 정보:

(i) 서열 특징:

(A)길이: 56개 아미노산

(B)유형: 아미노산

(C)가닥수: 단일

(D)토플로지: 관련없음

(ii) 분자 유형: 펩티드

(iii) 가정: 없음

(xi) 서열 설명: 서열 번호 33:

Met Ala Ser Ser Met Leu Ser Ser Ala Ala Val Ala Thr Arg Thr Asn  
1 5 10 15

Pro Ala Gln Ala Ser Met Val Ala Pro Phe Thr Gly Leu Lys Ser Ala  
20 25 30

Ala Phe Pro Val Ser Arg Lys Gln Asn Leu Asp Ile Thr Ser Ile Ala  
35 40 45

Ser Asn Gly Gly Arg Val Gln Cys  
50 55

(2) 서열 번호 34에 대한 정보:

(i) 서열 특징:

(A)길이: 58개 아미노산

(B)유형: 아미노산

(C)가닥수: 단일

(D)토플로지: 관련없음

(ii) 분자 유형: 펩티드

(iii) 가정: 없음

(xi) 서열 설명: 서열 번호 34:

```

Met Ala Pro Thr Val Met Met Ala Ser Ser Ala Thr Ala Thr Arg Thr
1          5          10          15

Asn Pro Ala Gln Ala Ser Ala Val Ala Pro Phe Gln Gly Leu Lys Ser
20         25          30

Thr Ala Ser Leu Pro Val Ala Arg Arg Ser Ser Arg Ser Leu Gly Asn
35         40          45

Val Ala Ser Asn Gly Gly Arg Ile Arg Cys
50         55

```

(2) 서열 번호 35에 대한 정보:

(i) 서열 특징:

- (A)길이: 58개 아미노산
- (B)유형: 아미노산
- (C)가닥수: 단일
- (D)토플로지: 관련없음

(ii) 분자 유형: 펩티드

(iii) 가정: 없음

(xi) 서열 설명: 서열 번호 35:

```

Met Ala Gln Ile Leu Ala Pro Ser Thr Gln Trp Gln Met Arg Ile Thr
1          5          10          15

Lys Thr Ser Pro Cys Ala Thr Pro Ile Thr Ser Lys Met Trp Ser Ser
20         25          30

Leu Val Met Lys Gln Thr Lys Lys Val Ala His Ser Ala Lys Phe Arg
35         40          45

Val Met Ala Val Asn Ser Glu Asn Gly Thr
50         55

```

(2) 서열 번호 36에 대한 정보:

(i) 서열 특징:

- (A)길이: 74개 아미노산
- (B)유형: 아미노산
- (C)가닥수: 단일
- (D)토플로지: 관련없음

(ii) 분자 유형: 펩티드

(iii) 가정: 없음

(xi) 서열 설명: 서열 번호 36:

```

Met Ala Ala Leu Ala Thr Ser Gln Leu Val Ala Thr Arg Ala Gly His
1          5          10          15

Gly Val Pro Asp Ala Ser Thr Phe Arg Arg Gly Ala Ala Gln Gly Leu
20         25          30

Arg Gly Ala Arg Ala Ser Ala Ala Ala Asp Thr Leu Ser Met Arg Thr
35         40          45

Ser Ala Arg Ala Ala Pro Arg His Gln Gln Ala Arg Arg Gly Gly
50         55          60

Arg Phe Pro Phe Pro Ser Leu Val Val Cys
65         70

```

(2) 서열 번호 37에 대한 정보:

(i) 서열 특징:

- (A)길이: 39개 아미노산
- (B)유형: 아미노산
- (C)가닥수: 단일
- (D)토플로지: 관련없음

(ii) 분자 유형: 펩티드

(iii) 가정: 없음

(xi) 서열 설명: 서열 번호 37:

```

Met Ala Thr Pro Ser Ala Val Gly Ala Ala Cys Leu Leu Leu Ala Arg
1           5           10          15

Xaa Ala Trp Pro Ala Ala Val Gly Asp Arg Ala Arg Pro Arg Arg Leu
20          25           30

Gln Arg Val Leu Arg Arg Arg
35

```

### (57) 청구의 범위

#### 청구항 1

(a) 전분-캡슐화 영역(starch-encapsulating region)과;

(b) 상기 전분-캡슐화 영역에 융합된 페이로드 폴리펩티드(payload polypeptide)를 포함하는 잡종 폴리펩티드.

#### 청구항 2

제 1 항에 있어서, 상기 페이로드 폴리펩티드는 Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, 및 Val로 이루어진 그룹에서 선택된 세 타입이하의 아미노산들로 이루어짐을 특징으로하는 잡종 폴리펩티드.

#### 청구항 3

제 1 항에 있어서, 상기 페이로드 폴리펩티드는 생물학적 활성 폴리펩티드인 잡종 폴리펩티드.

#### 청구항 4

제 3 항에 있어서, 상기 페이로드 폴리펩티드는 호르몬, 성장 인자, 항체, 펩티드, 폴리펩티드, 효소 면역글로불린, 삭소 및 이들의 생물학적 활성 단편으로 이루어진 그룹에서 선택되는 잡종 폴리펩티드.

#### 청구항 5

제 1 항에 있어서, 상기 전분-캡슐화 영역은 가용성 전분 신타아제 I, 가용성 전분 신타아제 II, 가용성 전분 신타아제 III, 입자-결합 전분 신타아제, 분지 효소 I, 분지 효소 IIa, 분지 효소 IIb, 및 글루코아밀라아제 폴리펩티드로 이루어진 그룹에서 선택되는 효소의 전분-캡슐화 영역인 잡종 폴리펩티드.

#### 청구항 6

제 1 항에 있어서, 상기 전분-캡슐화 영역과 상기 페이로드 폴리펩티드의 사이에서 분열 부위를 포함하는 잡종 폴리펩티드.

#### 청구항 7

제 1 항의 잡종 폴리펩티드를 암호화하는 재조합 핵산 분자.

#### 청구항 8

제 7항에 있어서, 세균 숙주내에서 상기 전분-캡슐화 영역 및 상기 페이로드 폴리펩티드의 발현에 적합한 조절 서열을 포함하는 DNA 분자인 재조합 핵산 분자.

#### 청구항 9

제 7항에 있어서, 식물 숙주내에서 상기 전분-캡슐화 영역 및 상기 페이로드 폴리펩티드의 발현에 적합한 조절 서열을 포함하는 DNA 분자인 재조합 핵산 분자.

**청구항 10**

제 9 항에 있어서, 상기 조절 서열은 외떡입식물에서 상기 전분-캡슐화 영역 및 상기 페이로드 폴리펩티드의 발현에 적합한 것인 재조합 분자.

**청구항 11**

제 9 항에 있어서, 상기 조절 서열은 쌍떡입식물에서 상기 전분-캡슐화 영역 및 상기 페이로드 폴리펩티드의 발현에 적합한 것인 재조합 분자.

**청구항 12**

제 9 항에 있어서, 상기 조절 서열은 동물 숙주에서 상기 전분-캡슐화 영역 및 상기 페이로드 폴리펩티드의 발현에 적합한 것인 재조합 분자.

**청구항 13**

제 7항의 재조합 분자를 포함하는 발현 벡터.

**청구항 14**

DNA 분자를 발현할 수 있으며, 제 7항의 재조합 분자를 포함하도록 형질변환된 세포.

**청구항 15**

제 14 항에 있어서, 상기 세포는 식물 세포인 세포.

**청구항 16**

제 15항의 세포로부터 재생된 식물.

**청구항 17**

재조합 분자를 발현할 수 있는 제 16 항의 식물로부터 얻은 씨앗.

**청구항 18**

페이로드 폴리펩티드를 포함하는, 제 14 항의 세포로부터 유도한 변경된 전분.

**청구항 19**

동물의 소화계통의 선택된 부위에서 페이로드 폴리펩티드가 소화되도록하는 방법으로서, 소화관의 선택된 부위에서 소화되도록 선택되는 전분 기질내에 상기 페이로드 폴리펩티드가 포함되어있는 제 18 항의 변경된 전분을 상기 동물에게 공급하는 방법.

**청구항 20**

제 1 항의 잡종 폴리펩티드로부터 순수한 페이로드 폴리펩티드를 제조하는 방법으로서,

(a)상기 잡종 폴리펩티드를 암호화하는 DNA를 이용하여 숙주 유기체를 형질변환하고;

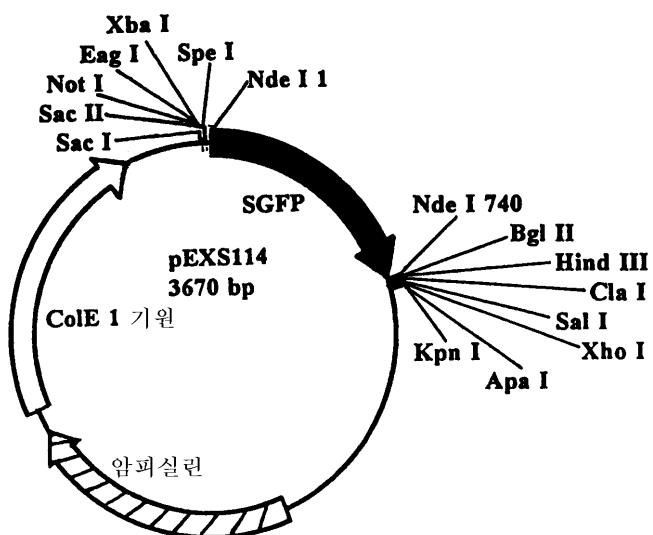
(b)상기 잡종 폴리펩티드를 상기 숙주내에서 발현시키고;

(c)상기 숙주로부터 상기 잡종 폴리펩티드를 분리하고;

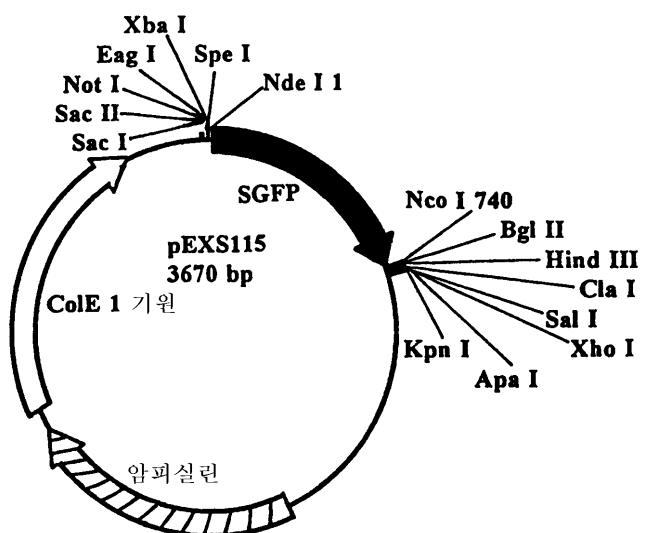
(d)상기 잡종 폴리펩티드로부터 상기 페이로드 폴리펩티드를 정제하는 것을 포함하는 방법.

**도면**

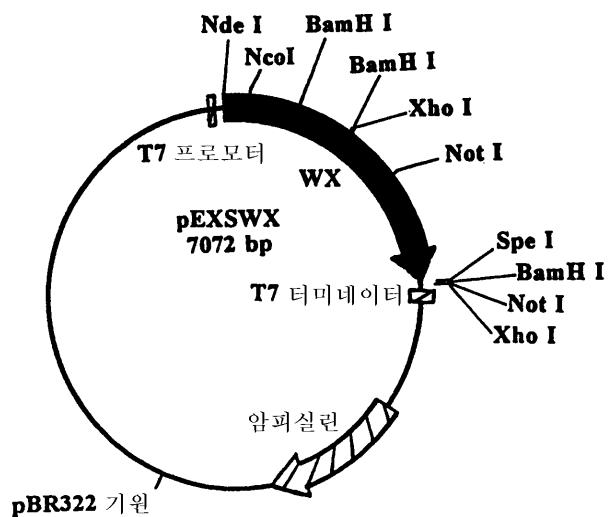
## 도면 1a



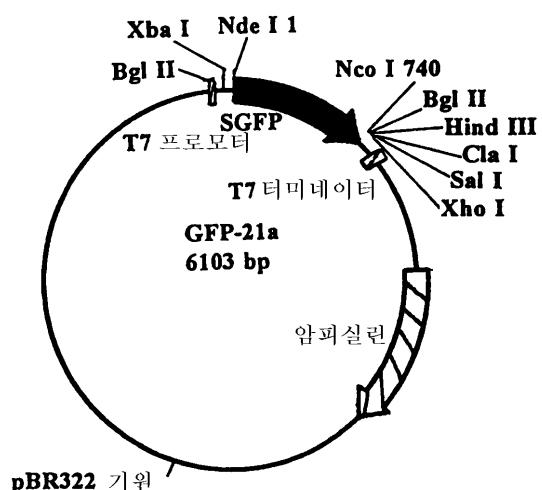
## 도면 1b



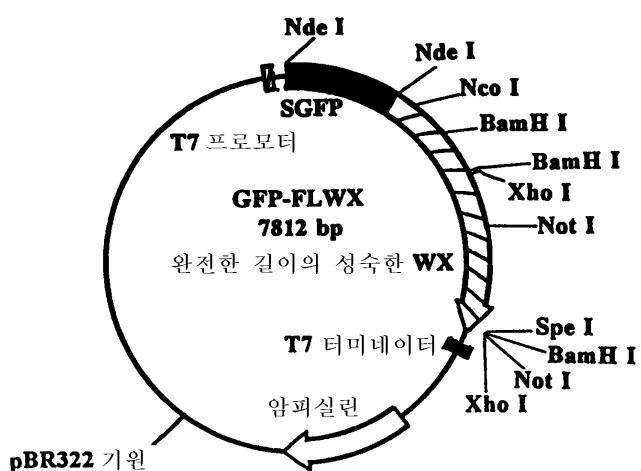
도면2a



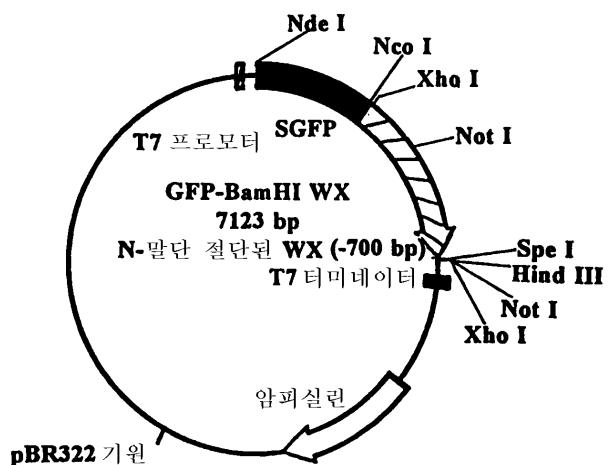
도면2b



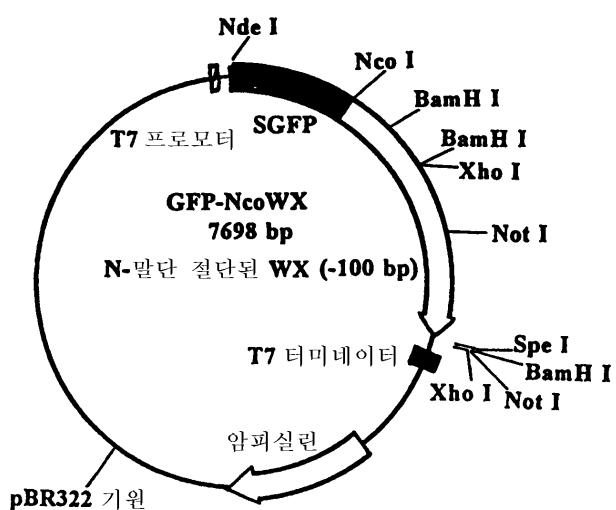
도면3a



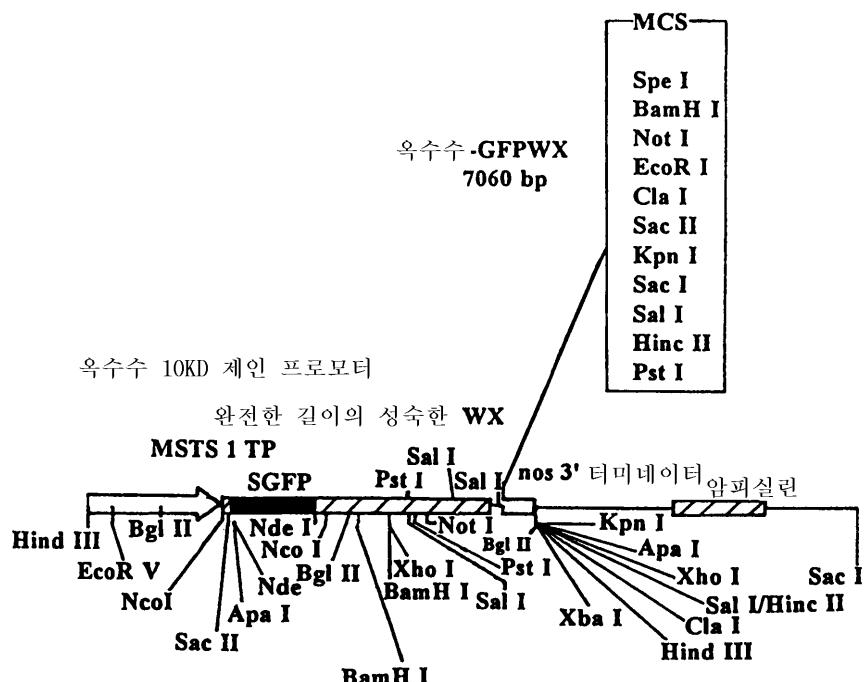
## 도면3b



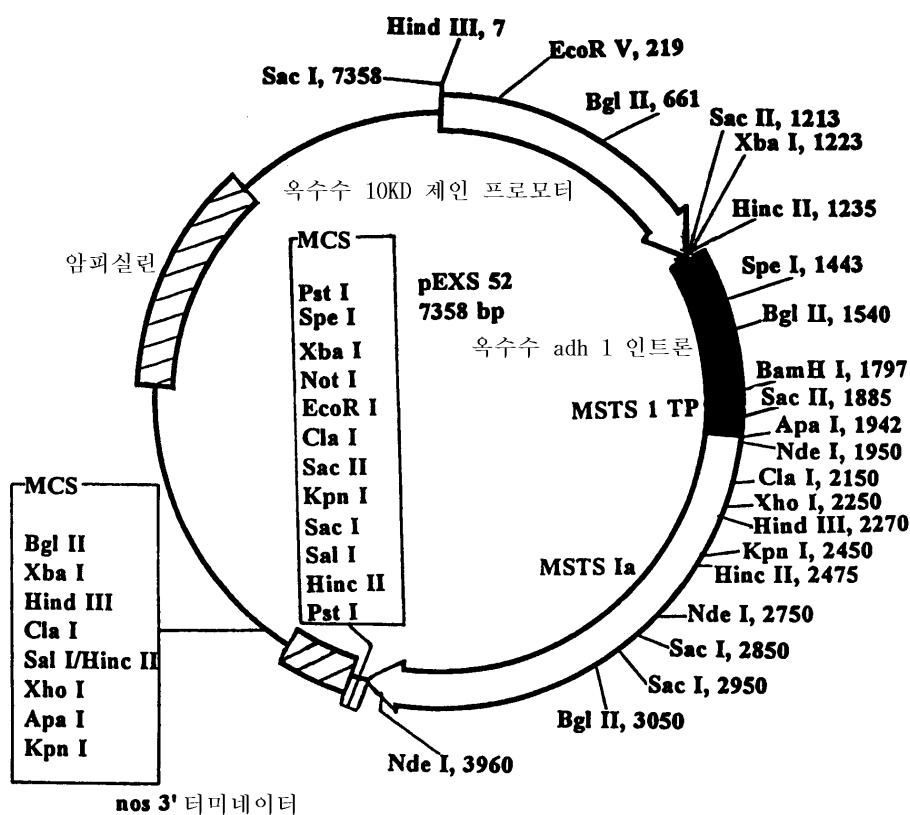
## 도면4



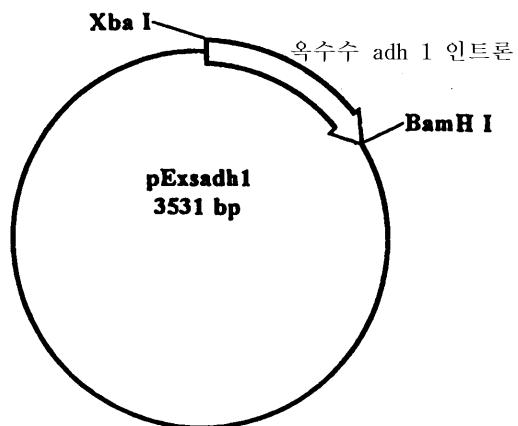
도면5



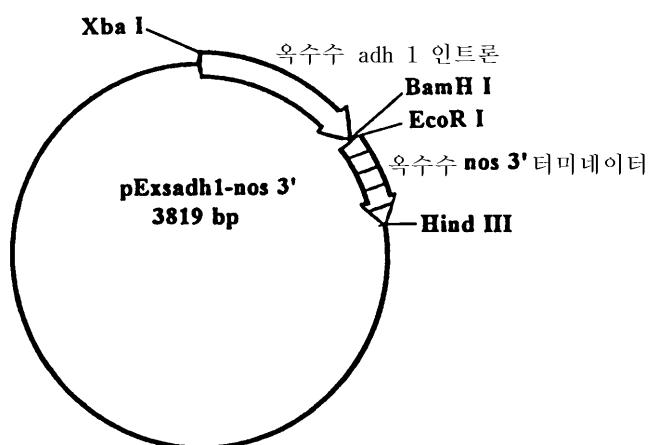
도면6



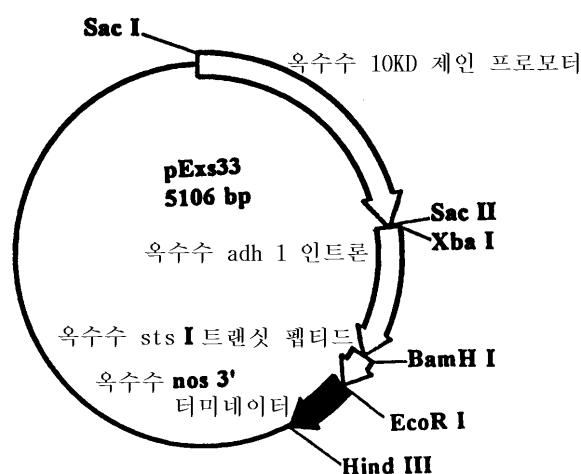
도면7a



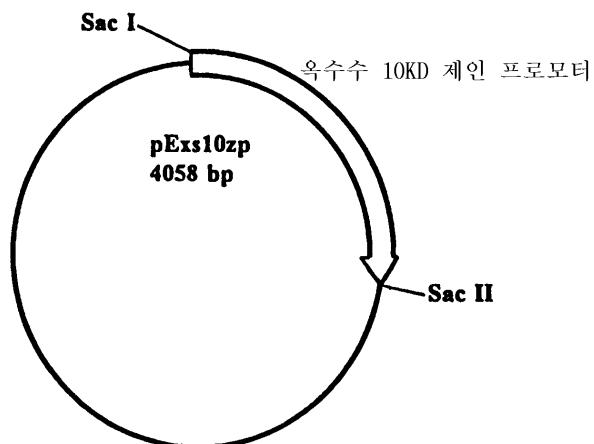
도면7b



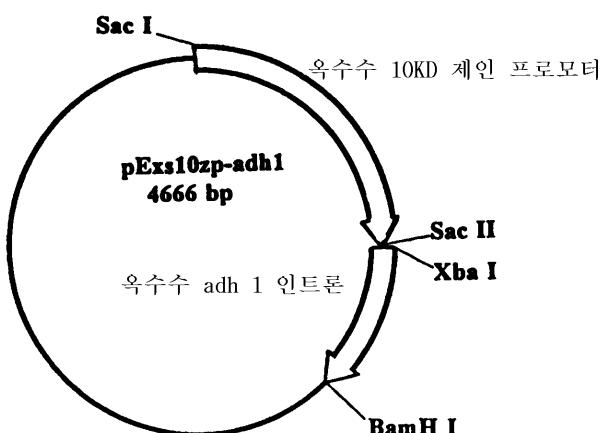
도면7c



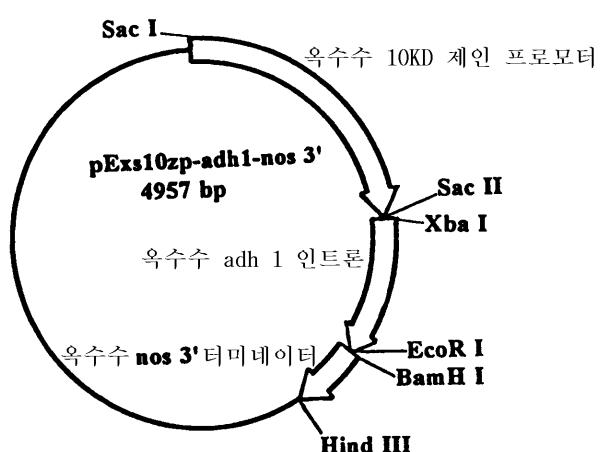
## 도면7d



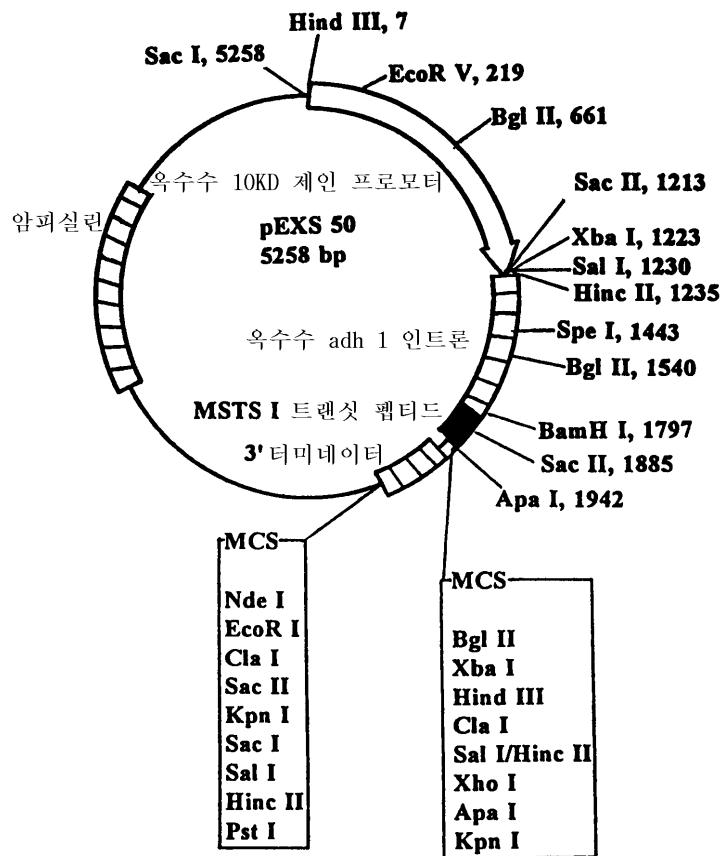
## 도면7e



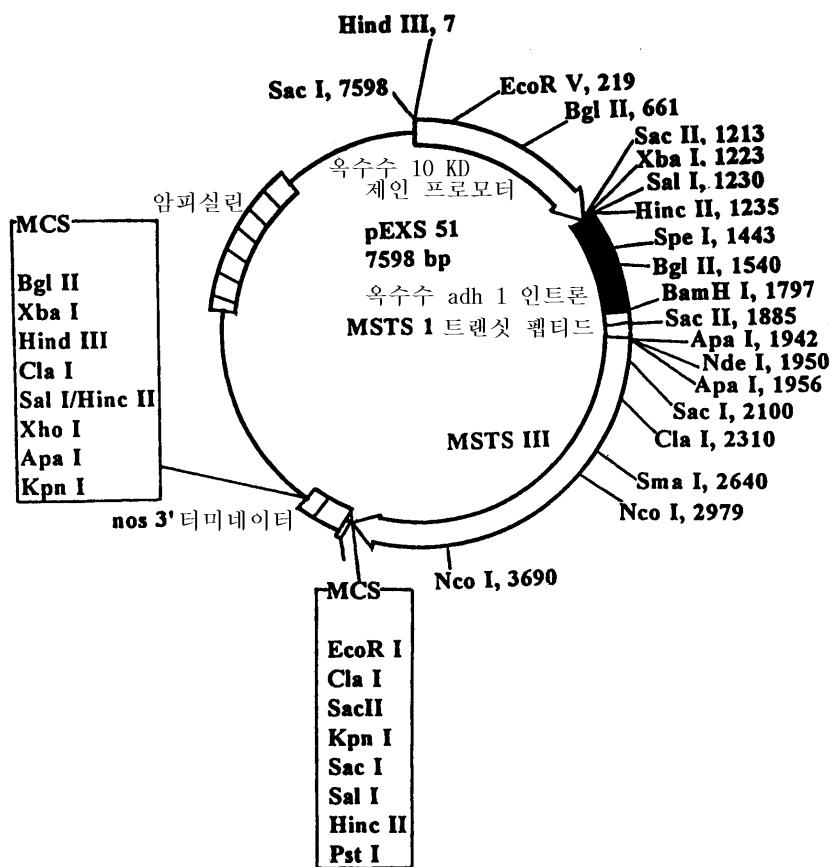
## 도면7f



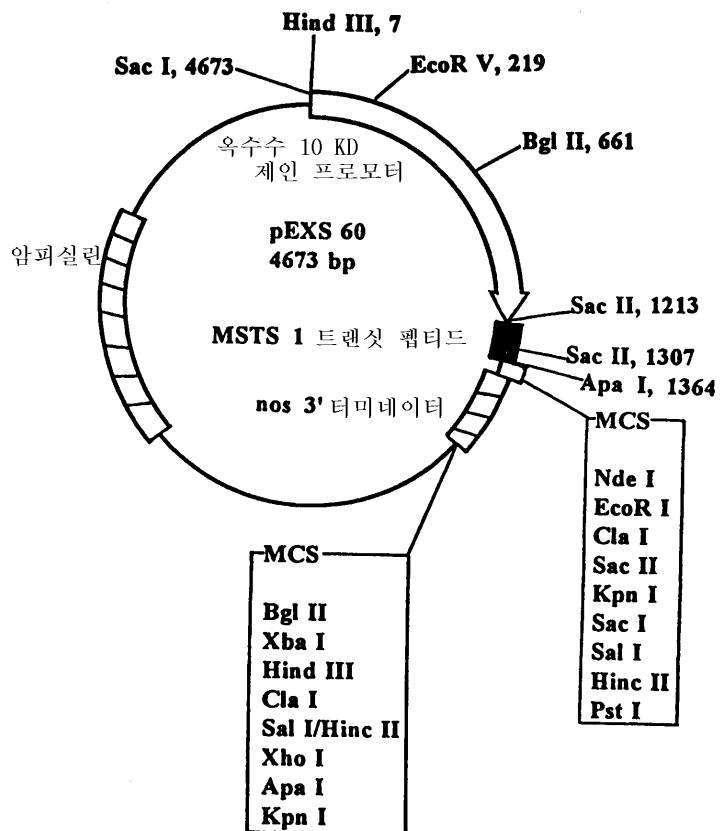
## 도면8a



## 도면8b



도면9a



도면9b

