

(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) Int. Cl. ⁷ C12N 15/82	(45) 공고일자 2001년01월 15일
	(11) 등록번호 10-0278456
	(24) 등록일자 2000년10월 19일
(21) 출원번호 10-1994-0703487	(65) 공개번호 특1995-0700995
(22) 출원일자 1994년10월04일	(43) 공개일자 1995년02월20일
번역문제출일자 1994년10월04일	
(86) 국제출원번호 PCT/CA 93/00141	(87) 국제공개번호 WO 93/20216
(86) 국제출원일자 1993년04월02일	(87) 국제공개일자 1993년10월 14일
(81) 지정국 AP ARIPO특허 : 말라위 수단 EA 유라시아특허 : 러시아 EP 유럽특허 : 오스트리아 스위스 리히텐슈타인 독일 덴마크 스페인 핀란드 영국 룩셈부르크 네덜란드 스웨덴 OA OAPI특허 : 부르키나파소 베냉 중앙아프리카 콩고 코트디부아르 카 메룬 가봉 기네 말리 모리타니 니제르 세네갈 차드 토고 국내특허 : 오스트레일리아 바베이도스 불가리아 브라질 캐나다 헝가리 일본 북한 대한민국 스리랑카 마다가스카르 몽고 노르웨이 폴란드 루마니아 미국	
(30) 우선권주장 07/862355 1992년04월02일 미국(US) PCT/CA92/00161 1992년04월 15일 캐나다(CA)	
(73) 특허권자 셈비오시스제네텍스인코포레이티드 존호휴알.	
(72) 발명자 캐나다 알베르타 티1와이 5더블류4 켈거리 3605-29 스트리트 엔이 500 올로니, 마우리스, 엠.	
(74) 대리인 캐나다 알베르타 티3에이 2에스4 켈거리 엔.더블유. 에찌힐 플레이스 131 박장원	

심사관 : 한현숙

(54) 조절 신호인 지방체 단백질 시스-요소

요약

후기 구형단계에서부터 배 성숙까지 활성을 가지는 유전자의 5'미번역 서열로 이루어지는 DNA 구조물이 제공된다. 이 구조물은 저장 단백질의 축적에 선행하는 배 형성 단계에서 목적하는 DNA 서열을 발현시키는 데 이용된다.

대표도

도1

명세서

[발명의 명칭]

조절 신호인 지방체(oil-body) 단백질 시스-요소(cis-element)

[도면의 간단한 설명]

제1도는 올레오신 조절 서열, 개시 코돈, 발현시키고자 하는 외부 DNA, 올레오신 종결 서열 및 앰피실린 내성 유전자를 포함하는 벡터 pPAW4의 다이어그램이다.

제2a도는 18KDa인 기름체 단백질을 지정하는 Arabidopsis 게놈 클론의 DNA 서열을 나타낸다. ORE(open reading frame)에는 짧은 인트론이 끼여 있으며, 두 개의 엑손은 번역되어 IUPAC에 의한 1개 문자로 아미노산을 표시하였다.

제2b도는 Arabidopsis의 18KDa 기름체 단백질을 지정하는 서열을 포함하는 Agrobacterium EMBL3 게놈 라이브러리를 제한 효소로 처리하여 얻은 단편들을 나타낸다.

제3도는 총 RNA를 Northern blot으로 분석하여, 발육단계에서 올레오신 mRNA의 발현에 대한 ABA 10 μ m의 영향을 나타낸 것이다.

A) (각 라인 당 70 μ g) ³²PdCTP 50ng으로 표시된 0B990를 프로브로 사용(비 활성 10⁹ dpm/ μ g DNA). 마이크로포어에서 유래된, ABA 10 μ m으로 48시간 동안 처리하거나(+) 처리하지 않은(-), 배의 심장(heart) (H)(13일), 어뢰(torpedo) (T)(17일), 자엽(cotyledonary)(C)(21-25일) 단계이다. 블랏을 Kodak XAR5 필름에 70°C에서 20분간 노출시켰다. 각 라인에서 mRNA 겉보기 크기가 다른 것은 각 mRNA 준비에서의 녹말량이

다르기 때문이다. OD260 측정과 EtBr 염색에 의해 확인되었듯이 모든 라인에는 동일한 양이 로딩되었다.

B) 제3도의 A)를 4.5시간 노출한 결과이다.

C) 스캐닝 밀도측정기로 결정한, mRNA의 축적량의 상대적 강도이다.

제4도는 올레오신의 조직 특이성을 나타낸다. 폴리 (A) + 뿌리(R), 켈러스(ca), 자엽(Co), 잎(L) 및 개화 24일 후의 접합된 배(E)의 RNA에 대해 32 PdCTP 50ng으로 표시된 OB990를 프로브로 사용(비 활성 10^9 dpm/ μ g DNA) 하였다.

제5도는 ABA에 적용된 기름체 단백질 활성의 발육 단계에 대한 민감성을 나타낸다. 라인 A,C,E,G에는 대조구를, 라인 B,D,F,H에는 ABA로 처리한 시료를 약 10,000 dpm 로딩하였다. 각 시료를 ABA로 2일 동안 처리한 후, 1.85MBq/mL [35 S]메티오닌으로 4시간 표지하였다. 심장 단계의 배를 얻기 위해 라인 A와 B (10일 된 배양액)을 62 μ m 스크린으로 걸렀다. 외위 단계의 배를 얻기 위해 라인 E와 F (17일 된 배양액)을 250 μ m 스크린으로 걸렀다. 자엽 단계의 배를 얻기 위해 라인 G와 H (25일 된 배양액)을 500 μ m 스크린으로 걸렀다.

[발명의 상세한 설명]

본 발명은 업스트림(upstream) DNA 서열과, 식물의 종자 발육에서의 유전자 발현 조절에서의 용도에 관한 것이다.

식물 유전자 발명에 대한 연구 결과, 발현을 조절하는 요소에 관한 일반적인 결론을 상당히 얻게 되었다. 원핵 및 진핵과 마찬가지로 식물도 전사 속도를 조절할 수 있는 것으로 보여지는, 유전자의 전사 개시 자리의 5' 위쪽 (upstream)에 보존된 서열 또는 컨센서스(consensus) 서열을 가진다. 진핵생물의 경우, 이러한 서열은 TATAA/TAA/T를, 가지는 전사 개시 자리의 5'쪽의 전형적인 약 25개의 염기를 가진다는 것을 알게 되었으며 이것을 TATA 박스라고 한다. TATA 박스의 기능은 RNA 폴리머레이즈 II의 전사 개시를 정해주는 것으로 보인다. 2번째 업스트림 서열은 CAAT 박스이다. 이 서열은 전형적으로 전사 시작의 업스트림 서열 약 75 번째 염기에서 발견되는데 전사 개시 횟수 조절에 관여한다. 식물에서는 컨센서스 서열이 CCAAT 또는 AGGA이지만, 모든 식물 유전자에서 이들 선택적인 컨센서스 서열이 필요한 것은 아니다. 이들 서열 모티프와 전사 시작 쪽 위쪽의 70-90 사이의 염기 서열은 종종 프로모터라 일컫는다. 일반적으로, 프로모터 부위의 5'과, 그것으로부터 2000개 이내의 염기는 종종 시스적으로 작용하는 요소이며 프로모터의 특성을 다양화시키며, 구조적으로 또는 비구조적으로 전사 활성을 조절할 수 있다. 이와 같이 시스적으로 작용하는 서열을 인핸서 (전사를 증가시키는 경우) 또는 사일렌서 (전사를 억제하거나 감소시키는 경우)라고 한다. 인핸서와 사일렌서는 종종 핵 단백질이 결합하거나 상호 작용하는 자리이다. 이런 경우 그 조절하는 핵 단백질을 트랜스-액팅 요소(trans-acting element)라고 하며, 외부의 자극에 대하여, 또는 특정 기관이나 조직에서의 유전자의 활성을 주로 결정하므로 비 구조적 발현 또는 조절된 발현에서 매우 중요한 역할을 한다. 이 단백질 결합과 인핸서/사일렌서 사이의 관계는 전사 활성을 결정할 것이다. 열, 빛 또는 유전자에 의해 생산되는 호르몬과 같은 화학 물질 등에 의해 활성화되거나 종자, 잎 또는 꽃과 같이 특정 조직에서 활성화되는 유전자를 분리하면 어떻게 발현이 조절되는 결정하는 인자들을 분석할 수 있다. 전부는 아니지만 많은 경우에 주어진 조절되는 유전자로부터 얻은 프로모터와 선택적으로 시스-요소와, 원래는 그 프로모터에 연결되어 있지 않는 리포터 단백질을 지정하는 코딩서열을 포함하는 키메라 유전자는 그 리포터 단백질의 발현이 조절됨을 보여준다. 종자에서의 서열 발현을 위한 종자 특이적 유전자에서 분리한 프로모터를, 보통 종자 특이적인 형태로 발현되지 않거나 또는 변형된 발현 형태를 필요로 하는 유전자에 사용하는 것은 몇몇 경우만 보고되었다. 현재까지 모든 경우에서, 종자 특이적으로 발현되도록 설계된 키메라 유전자는 종자 저장 단백질 조절신호와 프로모터를 이용하였다. 그러나, 저장 단백질 유전자 발현에 대한 연구 결과, 배형성 (embryogenesis)에서 상당히 늦은 단계에서, 즉 쌍떡잎 식물의 경우 배가 고전적인 어뢰 모양이 되었을 때야 발현이 시작되는 것 같다. 따라서 저장 단백질이 높은 정도로 발현되고 그것의 조절이 종종 전사 단계에서 이루어진다 하더라도, 발현의 시간과 정도가 모든 종자 특이적인 응용에 이상적이지는 않다. 따라서, 올레오신(oleosins)의 프로모터와 인핸서와 같은, 종자 저장 단백질의 프로모터와 인핸서와는 다르고 일시적 또는 세포특이성이 있는, 종자 특이적 프로모터와 인핸서를 얻는 데 관심이 집중된다.

아래에서는 형질전환된 식물에서 조직특이적 또는 기관특이적 발현을 가능하게 하는 조직특이적 또는 기관특이적인 조절 서열을 기재한다. 지금까지 비종자 단백질인 클로람페니콜 아세틸 트랜스퍼레이즈에서 조절 유전자 발현에 관한 '고전적인' 예들이, 리포터 단백질의 지정서열이 라이보즈 비스포스페이트 카르복실레이즈를 코딩하는 콩 유전자의 프로모터 및 업스트림 서열과 융합된 트랜스제닉 식물의 경우, 빛 조절에 의해 또는 기관특이적으로 발현될 수 있다는 것이 알려져 있다(Fluhr, Science (1986), 232:1106-1112).

Sengupta-Gopalan et al. Proc. Natl. Acad. Sci. USA, (1985) 82:3320-3324는 강남콩의 주요 저장 단백질인 β -파세올린을 담배에서 발현시킨 것이 보고되었다. 그 유전자는 종자에서 제대로 발현되었으며, 식물의 그밖의 조직에서는 매우 낮은 정도만 발현되었다. 자연상태의 5' 옆쪽의 서열을 포함하는 β -파세올린 전체가 사용되었다. 다른 종에 대한 실험(Radke et al. (1988) Thero. App. Genet. 75:685-694)과, 다른 유전자에 대한 실험 (Perez-Grau, L., Goldgerg, R.B., 1989, Plant Cell, 1:1095-1109)은 내핀 (napin) 유전자 전부와 전사되지 않는 'tag'을 포함하는 'tagged' 유전자를 사용하여 Arabidopsis와 Brassica 모두에서 종자 특이적인 형태로 충실히 발현됨을 보여주었다 (Radke et al. (1988), 위와 같음).

조직 및 기관특이적 발현에서, 전사 시작의 위쪽 서열이 유전공학 기술에 의해 식물에 도입된 유전자에 조직/기관 특이성을 부여하는 데 이용됨을 알게 해주는 몇가지 예가 있다. 그러한 예에는 공학적인 종자 특이적 유전자 조절 (Radke et al (1988), 위와 같음; Bustos set ag. (1989), Plant Cell, 1:839-853)이 있다. 이 두 예에서, 종자 단백질의 지정서열의 위쪽 서열은 리포터 tag(RNA 또는 단백질)에 연결되며, 리포터에 종자 특이성이 부여된다. 이들은 모두 올레오신보다 더 저장단백질 유전자이다. 종자 저장 단백질

질은 올레오신과는 다른 일시적인 발현 형태를 가진다.

종자 특이적 발현을 야기하는 DNA 모티프는 많은 연구의 대상이 되고 있다. Marcotte et al. (Marcotte, W.R., Russel, I.S., Quantrano, R.S., 1989, Plant Cell, 1:969-976)는 밀의 Em 유전자를 연구하였으며, 정자 특이적 발현을 위한 컨센서스 서열일 것으로 추측되는 'Em-박스'라고 하는 두 개의 모티프를 제안하였다. 흥미롭게도, 이들 중 EM-2로 불리는 박스는 외떡잎 식물 (triticin-밀)의 다른 저장 단백질 유전자에서 발견되는 것과 비슷하며, 심지어 쌍떡잎 식물 (β -포르리시닌-완두)에서 발견되는 것과도 비슷하다. Hatzopoulos et al. (1990, Plant Cell, 2:457-467)

DeClercq et al. Plant Physiol., (1990), 94:970-979 에서는 Arabidopsis 2s 알부민의 프로모터와, Arabidopsis와 Brail 너트 2S 알부민의 지정서열을, 거의 보존되지 않은 부위에서 결합시켰다. 담배와 Brassica napus에 형질변환시켜 종자특이적 발현과 변형된 저장 단백질이 제대로 축적되는 것을 확인하였다. 발현 정도는 총 세포 단백질량의 0.05-0.3% 였다.

이와 같은 외래 서열의 종자특이적 발현의 또다른 예는 종자에서의 leu-enkephalins의 발현이다. 종자 특이적인 발현을 위하여, leu-enkephalin(5펩타이드)를 지정하는 짧은 올리고뉴클레오타이드와 2s 알부민을 지정하는 키메릭 DNA 서열을 자연 상태의 알부민 단백질 지정 서열의 6번째와 7번째 시스테인 사이에 넣었다 (Vanderkerhove et al. Bio/Technology, (1989) 7:929-932). 이 유전자는 종자특이적 형태로 발현되었으며, leu-enkephalin을 종자 g당 50nmol까지 축적하였다.

두 종에서 기름체(oil-body) 단백질과 그것과 연관된 위쪽 서열을 지정하는 게놈 클론이 보고되었다; 옥수수 (Zea mays, Bowman-Vance and Huang, (1987) J. Biol. Chem., 265:2238-2243), 당근 (Hatzopoulos et al. (1990) Plant Cell, 2:457-467). 배양된 유지종자인 Brassica napus에 대한 cDNAs와 게놈 클론도 보고되었다 (Murphy, et al. (1991) Biochem. Biophys. Acta, 1088:86-94; Lee and Huang (1991) Plant Physico 96:1395-1397). 발육 중인 종자에서 이들 기름체 단백질 유전자의 발현은 다양하다. Zea mays의 경우, 아이소 형태인 기름체 단백질을 지정하는 유전자의 전사는 종자 발육의 상당히 초기 단계에서 시작되었으며, 수분 후 18일 후에는 쉽게 검출되었다. 쌍떡잎 식물인 Brassica napus(카놀라)와 같이 내배유가 없는 종자에서는 기름체 단백질 유전자의 발현은 종자 발육 중 상당히 늦게 나타나는 것으로 보인다 (Murphy et al. (1989) Biochem. J., 258:285-293)

[발명의 설명]

헤테로 유전자의 종자 특이적 형태의 발현을 위한 번역되지 않는 5'-리더 서열을 선택적으로 동반하는 기름체 단백질의 전사 조절 서열에 대한 연구를 위한 방법 및 조성물을 기재한다. 그러한 방법은 조절 서열과, 그 조절 서열에 원래 있는 ORF (open reading frame) 외의 다른 DNA 서열로 이루어지는 DNA 구조체로 식물세포를 형질전환시키는 단계와, 그 형질 전환된 세포를 발아시키는 단계, 종자가 생성되고 조절 부위의 전자 조절에 의해 조절되는 조건에서 DNA 서열이 발현되는 단계를 포함한다. 또한 종자의 본질적인 가치를 증가시키기 위해 외부 단백질을 함유하는 변형된 종자를 생산하는 데도 이용될 수 있다.

[구체예의 간단한 설명]

본 발명에 따라, 종자에서 식물의 표현형을 조절할 수 있는, 특히 배 형성의 초기 단계에서 조절할 수 있는 DNA 구조물이 제공된다. 이 DNA 구조물은 배의 성숙을 거쳐 후기 구형 단계 (자엽 단계)에서 활성을 나타내는, 유전자의 5'-비번역 서열을 이용하여 종자에서의 전사를 조절한다. 기름체 단백질 유전자의 개시부분의 전사 시작 조절 자리 및 다운스트림은 식물세포의 게놈에 전사 카세트가 끼어들게 해주는 DNA 서열일 것이다. 인테그레이션 구조물이 다양한 DNA 서열에 효율적으로 끼어들 수 있도록 하기 위하여, 종자 특이적 전사 개시 부분의 아래쪽에 멀티 클로닝 자리를 포함할 수 있다.

기름체 단백질 유전자, 특히 쌍떡잎 유지 종자에서 발현되는 기름체 단백질의 유전자에서 얻어지는 조절 서열이 특히 흥미있다. 기름체 단백질은 기름(트리글리세라이드) 또는 저장 단백질보다 상당히 늦게 축적된다고 보고되었다. 이와 같이 늦은 발현은 종자 특이적 발현을 위한 유전자가 배형성의 초기 단계에서는 유전자 발현을 변형시키는데 사용될 수 없으므로 이들 유전자와 관련된 프로모트의 가치를 제한하였다. 그러나 놀랍게도, 쌍떡잎 유지종자에서 이러한 유전자의 발현은 지금까지 생각되어왔던 것보다 훨씬 일찍 일어난다는 것을 알게 되었다. 따라서 이러한 유전자들의 프로모터와 업스트림 요소들은 저장 단백질의 축적에 앞선 배형성 단계 동안 기작을 변형시키는 것과 관련하여 다양하고 가치있는 용도를 가진다.

기름체 단백질은 분류학적으로 광범위한 종에서 동정된다 (Noreau et al. Plant Physiol. (1980), 65:1176-1180; Qu et al. Biochem. J.,(1986) 235:57-65). 이들 단백질은 유지체에 일정하게 분포되어 있으며, 무성 생식 조직의 소기관에서는 발견되지 않는다. Brassica napus (평지씨)에서는 발육 중인 종자의 기름체와 최소한 3개의 폴리펩타이드가 관련이 있다 (Taylor et al. 1990, Planta, 181:18-26). 기름체에 연관된 단백질의 수와 크기는 종에 따라 다양하다. 예를들면, 옥수수에서는 기름체에서 면역학적으로 구분되는 4개의 폴리펩타이드가 발견된다 (Bowman-Vance and Huang, 1988, J. Biol. Chem., 263:1476-1481). 올레오신은 친수성, 소수성 및 친수성 부분이 번갈아 이루어진다 (Bowman-Vance and Huang, 1987, J. Biol. Chem. 262:11275-11279). 옥수수, 평지씨 및 당근의 올레오신의 아미노산 서열이 보고되었다 (Qu and Hunang, 1990, J.Biol.Chem. 265:2238-2243; Hatzopoulos et al. 1990, Plant Cell, 2:457-467). 평지씨와 같은 유지종자에서는 올레오신이 총 종자 단백질의 8% (Taylor et al. 1990, Planta, 181:18-26) 및 20% (Murphy et al. 1989, biochem, J., 258:285-293)을 차지한다. 이러한 비율은 많은 종자 저장 단백질의 비율과 비교될 만하다.

초기 배형성 단계 특히, 저장 단백질의 발현보다 빠른 시기와 관련된 전사 개시 부분은 종자의 초기 발육에서, 관심있는 DNA 서열의 전사를 원하는 수준으로 제공하기 때문에 특히 흥미있다. 일반적인 식물 배형성은 전형적으로 정해진 일련의 단계를 거친다. 쌍떡잎 식물 종자의 배형성은 후형화 단계, 심장 단계, 어뢰 단계 및 자엽 단계를 거친다. 이러한 응용을 위하여, 이와 같은 용어는 식물학에서 Ray, Steves 및 Fults (Saunders College Publishing), 17장 294 페이지에 제공된다. 일반적으로, 전사 개시 부위는 종자의 초기 형성 단계에서 발현되는 유전자로부터 얻을 수 있다. 바람직하게는 이 전사 개시 부위가 후기 구형단계부터 자엽단계까지, 더욱 바람직하게는 구형 단계부터 심장, 어뢰 및 자엽단계까지 그 활성을 유지

하는 것이 좋다. 종자의 초기 형성에서 전사를 제공하기 위한 자연상태의 기름체 단백질 유전자 전사 개시 부위 서열의 염기 서열과 유사한 염기 서열을 가지는 전사 개시 부위를 얻을 수 있다. 그 서열은 자연상태의 것이나, 합성 또는 부분적으로 합성한 것일 수 있다.

기름체 단백질에서 얻은 전사 개시 부위는 일반적으로 전사의 5'-3'방향, 전사 개시 부위, 관심이 있는 DNA 서열 및 전사 종결 부위를 포함하며, 전사 조절 부위가 작동할 수 있도록 연결되고 식물세포에서 기능할 수 있는 카세트에서 제공될 것이다. 하나 이상의 인트론이 존재할 수도 있다. 각각의 조작 후에, 최종 구조물에 사용되는 DNA는 제한 효소로 잘려질 수 있으며, 최종 구조물에 사용되는 다른 DNA에 연결될 수 있다. 바람직한 구체예에서, 양립할 수 있는 제한 효소 자리를 가지는 코딩서열은 기름체 단백질 유전자의 1번 코돈에 해당하는 위치에서 라이게이션될 수 있다. 그러한 치환의 개략적인 다이어그램을 제 1도에 나타내었다. 재조합 코딩 서열은 기름체 단백질 유전자의 코딩 서열을 완전히 치환하고 그것의 3'말단은 기름체 단백질의 터미네이터와 폴리아데닐화 신호에 연결되는 방식에 의해 끼워넣을 수 있다.

선택적으로, PCR (polymerase chain reaction)에 의해, 편리하게 옆에 제한효소 자리를 가지는 전사 개시 부위를 포함하는 DNA 단편을 다량으로 얻을 수 있다. 다량으로 얻는 DNA 단편은, 예를 들면 PCR에 의해 증식된 단편에 있는 전사 개시 부위의 조절하에서 목적 폴리펩타이드의 코딩서열이 전사되는 키메라 유전자를 제조하기 위하여 전사 또는 번역차원에서 목적 폴리펩타이드의 코딩서열에 연결될 수 있다.

이 전사 개시 부위는 자연상태의 것이거나, 숙주세포의 것과 동일하거나, 외부에서 도입된 것이거나 숙주세포와 다를 수 있다. 외부로부터 전사 개시 부위를 포함하는 구조물이 삽입되는 야생형 숙주에서 발견되지 않는 전사 개시 부위를 도입하는 것이 의도된다. 일반적으로, 조절서열은 기름체 단백질 유전자의 번역 개시의 5'쪽 1.5kb 까지의 DNA로 이루어진다. 이 서열은 site-direct 돌연변이에 의해 (Kundel TA, 1985, Proc. Natl. Acad. Sci. USA 82:488-492), 또는 3'-말단 코돈 근처에 적당한 제한효소 자리가 있다면 적당한 링커 올리고뉴클레오타이드를 도입함에 의해 원하는 단백질의 첫번째 코돈에 해당하는 위치에서 변형될 수 있다.

특히 융합 단백질이 기름체 단백질과의 융합 단백질인 경우 목적 DNA 서열을 융합 단백질로 발현시키는 것이 바람직할 것이다. 원하는 DNA 서열은 일반적인 방법에 따라, 제조된 키메라 유전자의 전사에 의해 융합단백질이 생산될 수 있도록 기름체 단백질의 공격에 맞게 기름체 단백질의 코딩 서열로 삽입될 수 있다. 이 경우, 기름체로 이동될 수 있도록, 이 융합 단백질이 제 2a 도에 기재된 바와 같은 Arabidopsis 기름체 단백질의 아미노산 44-122까지의 코딩 부위 또는 Arabidopsis 이외의 종에서 기름체 단백질의 대응되는 부분을 포함하는 것이 바람직할 것이다.

다른 종에서 기름체 단백질 코딩 서열을 분리하기 위하여는 최소한 두 접근이 가능하다. 그 중 하나는, 실시예에 기재된 Arabidopsis 클론을 다른 식물 종의 게놈 라이브러리에 대해 프로브로 사용하는 것이다. 이 클론은 밀접하게 관련된 종, 특히 모든 겨우과 식물의 올레오신과 하이브리드를 잘 이룰 것이다. 진화적으로 Arabidopsis에서 갈려나온 종, 예를 들면, 가지과, 콩과 및 모든 외떡잎 식물의 종에 대해서는, Arabidopsis 클론과 같은 올레오신 클론의 유전자 산물에 대한 항체를 이용하는 방법을 이용할 수 있다. 이러한 항체는 예를 들면 λ gt11를 이용하여 종자에서 유래된 cDNA 발현 라이브러리를 스크리닝하는 데 이용될 수 있다; Huynh et al. 1985, cDNA Cloning, Vol.1. A Practical Approach, Ed, Grover IRL Press, pp.49-78. 이러한 접근에 의해 표준 DNA 하이브리디제이션 기술에 의해 그 종의 게놈 라이브러리로부터 게놈 클론을 분리하는데 이용되는 새로운 종의 올레오신 cDNA 클론을 얻는다.

목적 DNA 서열은 효소와 같은 목적 펩타이드를 지정하는 ORF 이거나, 게놈 서열이 최소한 하나의 ORF, 인트론, 번역되지 않는 리더서열, 또는 전사, 스플라이싱과 같은 mRNA 프로세싱 또는 번역을 억제하는 기타 서열 중에서 선택되는 하나 이상의 서열일 수 있는 게놈 서열에 상보적인 서열일 수 있다. 목적 DNA 서열은 합성되거나, 자연상태에서 유래되거나 이들은 조합한 것일 수 있다. 목적 DNA 서열의 특성에 따라, 식물에서 잘 발현되는 코돈으로 그 서열을 합성하는 것이 바람직할 수 있다. 식물에서 잘 발현되는 코돈은 목적 종의 특정 식물에서 가장 많은 양으로 발현되는 단백질에서 횡수가 가장 높은 코돈으로부터 결정될 수 있다.

목적 DNA 서열은 다양한 재조합 단백질 중 하나를 지정할 수 있다.

이와 같은 방법에 의해 발현될 수 있는 재조합 단백질의 예에는 히루딘과 같은 항응고제, 인터루킨과 같은 림포킨, 고나도트로핀분비 호르몬과 같은 펩타이드 호르몬, 멀티 또는 싱글체인 항체와 같은 면역학적 시약과 단백질 분해효소, 지방분해 효소 및 폴리글루칸 가수분해효소와 같이 상업적으로 가치있는 여러 효소가 있다.

이용되는 종결부위는 그 종결부위가 상대적으로 상호 교환가능하게 보이므로 편리하다. 종결부위는 목적 DNA 서열에 원래 존재하는 것이거나, 다른 기원에서 유래된 것일 수 있다. 편리한 종결 부위는 5' 조절부위를 가지고 온 유전자로부터 얻을 수 있으며, 기름체 단백질 유전자의 터미네이터와 폴리아데닐화 신호를 포함한다. 선택적으로, 동일한 결과를 얻기 위해 Agrobacterium의 노팔린 합성효소 유전자의 터미네이터와 폴리아데닐화 신호와 같이 다른 터미네이터와 폴리아데닐화 신호를 이용할 수도 있다.

발현 카세트는 형질전환된 세포를 동정하기 위한 수단 및/또는 형질 전환된 세포를 선별하기 위한 수단을 부가적으로 포함할 수 있다. 예를 들면, 재조합 유전자는 항생제 내성 또는 제초제 내성 유전자와 같이 구조적으로 발현되는 선별 마커 또는 형질전환된 세포에 생물형광 또는 색을 부여하는 스크리닝 마커에 연결될 수 있다.

5'말단에 기름체 단백질 프로모터를 가지고, 3'말단에 터미네이터를 가지는 목적 DNA 서열을 Agrobacterium Ti 또는 이원(binary) 플라스미드와 같이 적당한 형질변환 벡터, 또는 pUC19, pBr322와 같은 간단한 클로닝 플라스미드에 도입하고, 마이크로인젝션, 일렉트로포레이션, PEG-매개 DNA 도입, biolistic 방법 등에 의해 직접적으로 DNA를 도입한다. 이러한 방법은 식물 형질전환 분야의 전문가에게는 잘 알려져 있다 (Horsch et al. (1985), Science, 227:1229-1231; Newhaus and Spangenberg (1990) Physiol Plant, 79:213-217; Sandford et al. (1990), Physiol Plant, 79:206-209).

형질전환된 식물은 형질전환 방법과 양립할 수 있는 표준 재생 방법을 거쳐, 형질전환된 세포로부터 얻을 수 있다(예를 들면, Moloney et al (1989), Plant Cell Rep., 8:238-242).

위에서 제시된 바와 같이 제조된 발현 카세트는 본질적으로 발육 중인 종자에서 잘 발현된다. 따라서 적절한 융합 펩타이드로 형질 전환된 식물세포는 알려진 방법에 의해 성장되고 종자를 맺는다 (예를 들면, McCormick et al., (1989) Plant Cell Rep. 5:81-84). 2세대 이상 기른 후 형질전환된 동일한 균주 또는 다른 균주와 수분시키고, 원하는 표현형 특성이 안정하게 유지되고 유전되도록 하면서 원하는 표현형 특성을 갖는 하이브리드를 고른 후, 목적 펩타이드를 분리하기 위해 또는 새로운 표현형 특성을 가지는 종자를 제공하기 위해 종자를 수확한다. 재생된 식물을 성장상자, 온실 또는 야외에서 비-재조합 식물과 동일하게 재배한다. mRNA 수준에서, 그리고 종종 폴리펩타이드나 단백질 수준에서 재조합 유전자의 종자 특이적 발현이 나타날 것이다.

폴리펩타이드/단백질은 그 자체로 이용할 수 있을 것이며 원한다면 추출하여 더 정제될 수 있다. 선택적으로 폴리펩타이드/단백질, 심지어는 mRNA 자체도 발육 중인 종자에 새로운 생화학적 표현형을 부여하기 위하여 이용될 수 있다. 새로운 표현형은 종자 단백질 또는 종자 기름의 조성을 바꾸거나, 기존에 존재하던 바람직한 산물 또는 특성을 생산을 증진시키거나, 안티센서, 라이보솜 또는 공동억제(co-suppression) 방법에 의해 바람직하지 않은 유전자 산물을 감소 또는 억제시키는 것과 같은 변형을 포함할 수 있다 (Izant and Weintraub (1984), Cell 36:1007-1015, antisense; Hazelhoff and Gerlach (1988), Nature 334:585-591, ribosome; Napoli, et al., (1990), Plant Cell, 2:279-289, co-suppression).

추출이 필요한 새로운 종자 단백질 또는 펩타이드를 생산하기 위해 형질전환하는 경우, SDS, Triton-X-100, Tween 20, NEGA-8 또는 원하는 단백질을 비가역적으로 불활성시키지 않는다고 알려진 다른 계면활성제를 소량 함유하거나 함유하지 않는 수용액 추출액으로 추출할 수 있다. 단백질 또는 펩타이드를 추출하기 위하여 건조된 종자를 손으로 또는 기계적인 분쇄기로 갈아 수용액상의 슬러리 또는 현탁액을 얻는다. 이것을 50,000g에서 원심분리하여 3 종류의 상 층, 미립자, 수용성 물질 및 소수성 물질로 나눈다. 생성물의 특성에 따라 각 상을 더 정제할 수 있으며, 녹인 후 암모늄 설페이트를 이용하여 선택적으로 침전시키거나 이온 교환, 젤 필터레이션, 친화적 매트릭스 등을 이용한 컬럼 크로마토그래피로 더 정제할 수 있다.

본 명세서에서 보고된 조절 서열에 대한 이상적인 숙주는 겨자와 식물이지만, 이들 프로모터는 올레오신 유전자가 상대적으로 많이 일치하는 식물 종에 광범위하게 사용될 수 있다. 이 프로모터를 이용하는 데 큰 장애가 되는 것은 외떡잎 식물과 쌍떡잎 식물 사이이다. 특정 발현을 위해 외떡잎 식물에 형질전환하기 위해서는, 외떡잎 올레오신 조절 서열이 사용되어야만 한다. 쌍떡잎 식물의 특이적 발현을 위해서는 쌍떡잎 올레오신 조절 서열이 사용되어야만 한다. 일반적으로 평지씨(Brassica) 속과 겨자과에 속하는 모든 식물을 포함하는 많은 쌍떡잎 식물에서 사용될 수 있다고 보고된 서열을 사용할 수 있다. 담배, 토마토와 같은 가지과 식물도 역시 그 서열을 인식하며, 발육 중인 종자에서 발현을 제대로 조절한다.

배의 축과 자엽의 다른 조직에서는 다른 세포성 발현이 검출될 수 있지는 단백질이 모든 배 조직에서 발현될 것으로 기대된다. 본 발명은 변형된 폴리펩타이드 또는 새로운 재조합 펩타이드를 축적하거나, 기작 단계를 더하거나 없애는 의해 식물 종자의 근본적인 가치를 개선하는 것을 비롯하여 많은 용도를 가진다. 간단한 실시예에서 본 발명의 용도는 단백질의 질을 향상시키거나 (예를 들면, 필수 아미노산 또는 희귀 아미노산의 농도를 증가시킨다), 지방산 조성을 변화시켜 액체의 질을 향상시키거나, 탄수화물 조성을 개선 또는 증진시킨다. 황함유 아미노산 잔기가 부족한 종자에 루핀 또는 브라질 너트에서 발견되는 S가 풍부한 단백질을 발현시키는 것도 하나의 예이다. 또, 저장 지방인 트리글리세라이드에서 지방산 비율을 변화시킬 수 있는 지방산 아실 조효소 A(COA) 전이효소를 발현시킬 수 있다. 종자에 재조합 단백질이 축적되는 경우, 그 단백질은 약학적, 상업적 또는 영양학적으로 가치있는 펩타이드일 수 있다. 이 경우, 그 펩타이드는 종자에서 추출될 수 있으며, 용도에 따라, 정제되거나 정제되지 않은 형태로 사용된다. 단백질은 동물 호르몬, 효소, 리포킨, 항응고제 등과 같이 식물에서는 완전히 외래의 것인 단백질일 수 있다. 이러한 헤테로 단백질은 종자에서 추출될 수 있으며, 부분적으로 또는 완전히 정제된 후 실험, 영양 또는 약학 목적으로 사용될 수 있다.

다음의 실시예는 본 발명을 자세히 설명할 목적으로 제공되나, 본 발명의 범위를 제한하지는 않는다.

[실시예 1]

λEMBL 3A의 *Arabidopsis thaliana* V. Columbia 계통 라이브러리로부터 얻은 클론에 있던 15kb 인서트에서 *Arabidopsis*의 기름체 단백질 유전자를 분리하였다. 기름체 단백질 번역 개시의 5'쪽 약 868 염기쌍을 포함하는 약 1.8kb 단편을 플라스미드 벡터에 서브클로닝하였다. *Arabidopsis* 18 KDa 올레오신 유전자를 Nco1 자리와 Kpn1 자리로 싸인 1803 염기쌍 단편 형태로 pPAW4 벡터에 클로닝하였다 (제 1 도). 그 단편을 다양한 외부/교대(alternative) 유전자와 일반적으로 사용할 수 있는 발현 카세트에 전환하기 위하여 두가지의 변형이 가해져야 한다. 첫째, site directed 돌연변이 (Kunkel 앞에서와 동일)로, 미스-매치 올리고뉴클레오타이드를 이용하여 -2, -1, +4 위치에 돌연변이를 일으킨다. 필요한 돌연변이는 A에서 T로 (-2), A에서 C로 (-1), G에서 A로 (+4)이다. 이 돌연변이는 -2에서 +4에 BspHI 자리를 만든다. BspHI 자리는 (T/CATGA)는 ATG 개시 코돈을 포함하며, Nco1으로 절단하면 양립할 수 있는 뒤로 물러난 말단을 제공한다. 두번째 변형은 자연상태의 올레오신의 코딩 서열의 대부분을 포함하는 658개의 염기쌍 단편을 만들기 위해 EcoRV와 Msc1로 처리한다. 이것은 EcoRV-Msc1 단편으로부터 벡터와 부수적인 서열을 분리했을 때 얻어지는 단편이 독특한 말단을 가지는데, 이것은 벡터-프로모터-터미네이터의 조합으로 다시 원형을 이루게 한다. 이 재원형화는 원래의 1803 염기쌍 단편에는 없는 제한효소 자리를 포함하는 올리고뉴클레오타이드 링커의 존재하에서 수행된다.

이 재원형화에서, 올레오신 유전자의 모든 업스트림 서열, 전사 개시자리 및 BspHI 10에 있는 개시 코돈을 포함하는 플라스미드를 얻는다. 이것의 다운스트림의 31개 염기는 하나 이상의 단일 제한효소 자리를 가지는 짧은 폴리링커이다. 이 카세트에 어떤 DNA 서열을 도입하기 위해서는 그 외부 서열이 개시 ATG 자리에 BspHI 또는 Nco1 자리를 가지거나 가지도록 변형된 서열이어야 한다. 서열이 단백질로 발현되기 위해서는 'cap' 자리와 개시코돈 사이에 일정한 거리가 유지되어야 한다.

삽입될 DNA 서열은 플라즈미드에는 없는, 독특한 말단으로 잘려지는 제한효소 자리로 종결되어야 한다. 발현 카세트에 중첩되는 폴리링터는 이러한 자리를 영두에 두고 선택되어질 수 있다. 이 플라즈미드를 BspH1과 그 외부 서열의 3' 말단을 위해 적당한 제한효소로 처리하면 원하는 DNA 단편을 정방향으로 효과적으로 클로닝할 수 있다. 적절한 라이게이션 조건에서, 제 1 도에 기재된 것과 같은 생성물을 얻기 위해, BspH1과 원하는 DNA 단편에 양립하는 자리를 가지는 플라즈미드 발현 카세트와 원하는 DNA를 라이게이션시킨다.

Nco1-Kpn1에서 얻은 완전한 구조물을 잘라, *Agrobacterium* 플라즈미드와 같이 적절한 식물 형질전환 벡터에 도입시킨다. 구조물을 Bin 19와 같은 일반적인 *Agrobacterium* 플라즈미드에 도입하기 위해서는 (Bevan, Nuc1. *Acid Research* (1984) 12:8711-8721) 플라즈미드 pPAW4에 있는 또 다른 제한효소 자리를 사용하여 할 것이다. 하나의 시나리오로써, 그 플라즈미드는 Sma1과 Kpn1으로 잘려질 수 있다. 얻어진 절제된 단편을 Kpn1으로 라이게이션시켜 올리고뉴클레오타이드 링커의 Kpn1 자리에 연결한다. 이것은 Bin19에 도입하기 위한 역방향 Kpn1 단편을 제공한다. 선택적으로, 이 구조물을 Kpn1과 BamH1으로 자르고, 같은 제한 효소로 절단한 pBin19에 정방향으로 연결시킨다. 얻어진 *Agrobacterium* 이원 플라즈미드를, 무장해제된 *Agrobacterium* 균주에 트리파이트 교배로 넣거나 (Ditta, et al. (1980) PANS 77:7347-7351) (An, 91988), 적당한 *Agrobacterium*에 형질변환시킨다 (Plant Mol. Biology Manual, A3 1-19, Kluwer Academic, Dordrecht, Netherlands).

Agrobacterium 내에 있는 재조합 Bin19는 Brassica 균주와 같이 적당한 식물에 표준 explant cocultivation으로 형질전환시키는 데 사용한다 (Horsch et al., (1985), 앞에서와 같음). 플라즈미드 pBin19의 T-DNA 경계에 있는 항생제 내성 유전자 (네오마이신 포스포트랜스퍼레이즈)를 이용하여 카나마이신을 함유하는 배지에서 형질전환된 세포를 선별한다. 표준방법 (예를 들면 평지씨와 같은 유지종자; Moloney et al. *Plant Cell Rep.*, (1989) 8:238-242)에 의해, 이 형질전환된 세포를 완전한 식물로 재생시킨다. 재생된 식물을 개화시키고 자가수분 또는 교차수분시킨다. 구조물에서 외부 DNA가 번역될 수 있는 산물을 지정하는 경우 성숙한 종자를 수용액으로 추출하고, 얻어진 슬러리를 원심 분리하여 (100,000 g에서 30분) 그 생성물을 분리할 수 있다. 원하는 생성물에 따라 얻어진 세가지 상 중 하나를 분배할 수 있다. 원심분리된 시료의 표면에서 지질피막 또는 수용액상 상 또는 팻릿에 존재할 것이다.

선택적으로, 발현되는 산물이 종자내에서 대사에 사용되어 종자의 표현형을 바꾸는 경우 (종자의 크기 또는 색을 바꾸거나, 종자의 지방산 잔기의 비율을 바꾸거나, 종자의 가치를 떨어뜨리는 특정 대사 관계를 막는 것 등)에는 그 생성물을 추출할 필요가 없다. 그러한 대사기작 단계는 종자의 가치를 감소시키는, 영양이 좋지 않은 2차 대사산물을 생산하는 단계를 포함한다. 이런 경우, 그 종자는 그대로 수확하고 보통의 방법에 따라 사용하면 된다.

[실시예 2]

PCR을 이용하여, β -글루쿠코니에이즈 (GUS)의 코딩부위에, 작동가능하게 다양한 양으로 연결된 *Arabidopsis* 올레오신 유전자의 5' 전사 개시 부위 서열을 포함하는 구조물을 제조하였다. 포함된 올레오신 5' 부위의 양에 따라 구조물의 명칭을 정하였다. 예를 들어 구조물 2500은 올레오신 5' 부위의 약 2500 염기쌍을 포함한다. 이 구조물을 Brassica napus와 담배에 도입하고, 아래에 기재된 방법에 따라 GUS 유전자의 발현을 측정하였다. 형질전환된 Brassica napus에서의 5개의 구조물 2500, 1200, 800, 600 및 200의 GUS의 발현을 표 1에 기재하였다. 형질전환되지 않은 대조구에 대한 결과도 표 1에 기재하였다. 형질전환된 담배에서의 구조물 2500과 구조물 800의 발현결과를 표 2에 기재하였다. 표 3에는 트랜스제닉 배에서, 올레오신 프로모터 발현의 발생학적 시기를 나타낸다. 구조물은 제한효소처리, 라이게이션 및 PCR과 같은 일반적인 분자생물학 기술을 이용하여 제조하였다. 이용된 기술의 상세한 사항의 일례로 아래에서 구조물 800의 제조를 기재하였다.

Arabidopsis 올레오신 유전자의 5' 전사개시 부위로부터 약 800 개의 염기쌍을 포함하는 DNA 단편을, GUS 코딩서열에 라이게이션되기에 알맞은 방향으로 얻기 위하여 PCR을 이용하였다. 발현 분석에 필요한 정확한 서열을 다량으로 얻기 위하여 PCR을 이용하였다. 원하는 PCR을 수행하기 위하여, 다음과 같은 서열을 가지는 두 개의 올리고뉴클레오타이드 프라이머를 합성하였다(Milligen-Bioscience, Cyclome DNA synthesizer):

5' 프라이머: 5'CACTGCAGGAACCTCTGGAA3'

(GVR 10)

이탈릭체로 표시된 염기는 제 2a 도의 서열에서 -833에서 -817까지의 서열을 나타낸다. 이 프라이머의 5'에 추가된 뉴클레오타이드는 올레오신 유전자와 일치하지 않으나, PCR로 다량으로 얻은 서열의 5' 말단에 Pst1를 주기 위해 포함시켰다.

3'쪽 프라이머는 다음 서열을 갖도록 합성되었다.

3' 프라이머 (ALP 1):

BamH1 올레오신 서열

5'CTACCCGGGATCTGTTTACTAGAGAGAATG3'

Sma1

이 프라이머는 이탈릭체로 표시된 부분이 제 2a 도의 서열중 -13에서 -30 까지와 완전히 상보적이다. 이 프라이머는 또한 5' 말단에 13개의 염기를 추가로 가지는데 이것은 올레오신 유전자와 상보적이지는 않으나, PCR로 다량으로 얻은 서열의 3' 말단에 두개의 겹치는 제한 효소자리 (Sma1, BamH1)를 주기 위해 포함시켰다.

이들 2개의 프라이머를 이용하여 PCR을 수행하여, 5' 말단에 Pst1 자리를 가지고, 3' 말단에 Sma1, BamH1 자리를 가지는, 올레오신 유전자의 -833에서 -13 서열을 포함하는 DNA 단편을 얻었다. PCR은 Taq 폴리머

레이즈 (Perkin-Elmer-Cetus)를 이용하여 효소 제조 회사에서 추천하는 조건에 따라 92°C(변성)에서 1분, 55°C(어닐링)에서 1분 및 72°C(연장화)에서 1분간 수행하였다. 주형으로는 제 2b 도의 위쪽에 기재된 본래의 λ 라이브러리에서 분리한 *Arabidopsis* DNA 약 15kb를 포함하는 올레오신 게놈 클론을 사용하였다.

다량 생산된 단편 (OLE0 p800)을 0.7% 아가로즈에서 정제한 후 Vogelstein and Gillespie (Preparative and analytical purification of DNA from agarose; Proc. Natl. Acad. Sci. USA 1979 76:615-619)의 글래스비드 방법으로 회수하여 PstI로 잘랐다. 잘려진 단편을 잘로 정제하고, DNA 폴리머레이즈 Klenow 단편으로 말단을 메우고, SmaI로 잘라 뭉툭한 말단을 만들었다. 이것을 pUC19의 SmaI 자리에 클로닝하여 pUCOLEOp800를 얻었다. 제 2b 도에 표시된 올레오신 유전자에서 -649에 해당하는 위치에 해당하는, 인서트의 비균형인 AccI 자리를 이용하여 pUC 벡터에 들어간 인서트의 양쪽 방향 모두를 선별할 수 있다. PCR로 얻은 단편의 5' 말단 (전사방향으로) 대부분이 클로닝 벡터 pUC19에 하나뿐인 HindIII 자리 가까이 에 위치하고, 3'말단 대부분이 pUC19에 하나뿐인 EcoRI 자리 근처에 위치하는 인서트를 가지는 클론을 얻었다.

이렇게 얻은 플라스미드를 BamHI으로 잘라, 옆에, BamHI 자리를 가지는 OLEOp800를 얻었다. BamHI-OLEOp800 단편을, 플라스미드 HspGUS1559를 BamHI으로 자른 것의 BamHI 자리에 클로닝시켰다. HspGUS1559 은 벡터 pCGN1559 (Macbride and Summerfeldt, 1990, Plant Molecular Biology, 14, 269-276); 옆에 BamHI 자리를 가지는 열쇼크 프로모터를 포함하는 인서트; GUS의 ORE(open reading frame)와, pB1221 (Jefferson RA in Cloning Vectors 1988, Eds. Pouwels P., Enger-Valk BE, Brammer WJ., Elsevier Science Pub BV, Amsterdam section VII, A11)에서 유래된 노팔린 합성효소 터미네이터에서 유래된, *Agrobacterium*에서 이원(binary) 벡터로 이용되는 플라스미드이다. HspGUS1559를 BamHI으로 잘라 열쇼크 프로모터를 해리시키고 그 자리에 다른 BamHI 자리를 넣었다. BamHI-OLEOp800 단편을 이 자리에 삽입하여 *Agrobacterium* pOLEOp800GUS1559를 얻었다. 이 플라스미드를 *E. coli* 형질전환에 이용하고, 다량생산된 플라스미드를 일렉트로포레이션방법으로 (Rogers et al., 1988, Plant Molecular Biology Manual, A2: 1-12, Eds. Gelvin S and Schilperoort, R. Kluwer Academic, Dordrecht, Netherland) *Agrobacterium* 균주 EHA101에 도입하였다.

얻어진 *Agrobacterium* (EHA101 × pOLEOp800GUS1559)으로, Moloney et al. (Moloney, M.M., Walker, J.M., Sharma, K.K. (1989) Plant Cell Reports 8:238-242) 방법에 따라 *Brassica napus*를 형질전환시키고, Horsch et al. (Horsch et al. Science (1985) 227:1299-1302) 방법으로 담배를 형질 전환시켰다. 얻어진 트랜스제닉 식물이 종자를 맺도록 한 후 그 발육중인 종자와 대조구인 비-재생성 식물에 대해 GUS 발현을 분석하였다(Jefferson R.A. (1987), Plant Mol. Biol. Rep. 5:387-405). 보고된 GUS 발현은 약 5개의 다른 트랜스제닉 식물에서 각각 얻은 약 5개의 종자로부터 얻은 값의 평균이다.

다른 구조물들 (2500, 1200, 600, 200)도 적절한 프라이머를 사용하여 위에 기재된 것과 동일한 PCR 방법으로 다량 제조하였다. *Brassica napus* 에서 발현된 결과를 GUS효소의 비활성으로 나타내었으며 표 1에 기재하였다. 담배에서의 발현 결과는 표 2에 기재하였다.

이러한 결과는 구조물 800에 사용된 -833에서 -813까지의 올레이신 단편이 트랜스제닉 *Brassica napus* 배에서 심장(heart) 단계부터 리포터 유전자의 종자 특이적 발현을 가능하게 하는 데 충분한 정보를 가지고 있으며, *Arabidopsis* 올레오신 프로모터가 *Arabidopsis* 이외의 식물에서도 전사를 가능하게 한다는 것을 증명한다. 이러한 실험은 이 프로모터에 있는 서열이, 자연상태의 올레오신 프로모터의 특성인, abscisic acid의 첨가에 반응하여 전사를 증가시키는 데 필요한 시스-요소를 포함한다는 것도 증명한다.

본 명세서에서 입증된 종자 특이적 발현은 올레오신 유전자 3' 말단의 자연상태의 터미네이터와의 상호작용에 의존하는 것이 아님에 주목하여야 한다. 이 실시예에서 올레오신 3' 터미네이터는 *Agrobacterium*의 노팔린 합성효소 유전자에서 유래된 터미네이터로 치환되었다. 따라서 구조물 800의 서열은 부가적인 서열의 필요없이 원하는 발현 프로파일을 작동시키는 데 충분하다.

[표 1]

Brassica napus에서의 종자 특이적 발현
GUS 활성 (pmol 생성물/min/mg 단백질에서)

프로모터/GUS 구조물	종자 (어뢰단계)		뿌리	잎	줄기	종자(후기 자엽단계)
	+ABA*	-ABA				
2500	10,185	7,709	444	46.9	88.2	11,607
1200	18,298	1,795				8,980
800	2,250	475	285	277	650	7,130
600	1,506	144				1,365
200	18.1	64.8	260	5.9	26	11
대조구(형질 전환되지 않은 식물)	18.4	13.9	300	6.1	30	14

* ABA는 GUS 활성측정 전에 10^{-5} M abscisic acid를 24시간 처리한 것임.

[표 2]

프로모터/GUS 구조물	종자에서의 GUS활성(pmol 생성물/min/mg 단백질)
2500	11,330
800	10,970

[표 3]

Brassica napus에서의 발생학적 발현
GUS 활성 (pmol 생성물/min/mg 단백질에서)

프로모터/GUS 구조물	심장 단계	어뢰 단계	초기자엽 단계	중기자엽 단계	후기자엽 단계
2500	272	1207	2541	1819	11607
1200	124	262	388	5094	8980
800	149	260	962	2617	7128
600	59	41	29	38	1365
200	30	25	15	20	11
대조구			11	14	14

본 명세서에 언급된 모든 출판물 및 특허출원은 그 내용 전부가 본명세서에 포함된다. 본 발명은 충분히 자세하게 기재되었으며, 이 분야의 통상의 전문가가 본 발명의 특허 청구범위에서 벗어나지 않는 범위에서 본 발명을 변형시킬 수 있을 것이다.

서열 리스트

1) 일반적 사항

(1) 출원인

이름: 유니버시티 케크올로지즈 인터내셔널, Inc.
 거리: 엔.더블류. 2500 유니버시티 드라이브 이에스620
 시: 캘거리
 주: 앨버타
 나라: 캐나다
 우편번호: 티2엔 1엔4
 전화번호: 403-220-5261
 팩스번호: 403-289-9311

이름: 몰로이, 마우리스 엠.
 거리: 엔.더블류. 에드거힐 플라이스 131
 시: 캘거리
 주: 앨버타
 나라: 캐나다
 우편번호: 티2엔 2에스4

(2) 발명의 명칭: 조절인자인 기름체 단백질 시스-요소

(3) 서열의 수 :3

(4) 우편주소

수신인: 스마트 앤드 비거
 거리: 900-55 메트켈프 스트리트
 시: 오타와
 주: 온타리오
 나라: 캐나다
 우편번호: 케이1피 5와이6

(5) 컴퓨터가 인식가능한 형태

중간형태: 3.5 " 디스켓, 1.44 Mb
 컴퓨터: IBM PS/2 모델 50z 또는 55sx
 오페레이팅 시스템: MS-DOS (버전 5.0)
 소프트웨어: Wordperfect (버전 5.1)

(6) 본건 출원 자료

출원번호: PCT/CA/93/00141
 출원일: 02/04/93
 클래스

(7) 우선권 자료

출원번호: 07/862,355
 출원일: 02/04/92
 출원번호: CA92/00161
 출원일: 15/04/92
 출원번호: 07/659,835
 출원일: 22/02/91

(8) 변리사에 관한 사항

이름: 모로우, 조이 디.; 그라벨, 미세린 엘.; 콘, 데이비드
 번호: 75236-1

(9) 통신수단

전화번호: (613) 232-2486
 팩스번호: (613) 232-8440
 텔렉스: 053-3731

2) 서열 1에 관한 정보

(1) 서열특성

길이: 1803
 형태: 핵산
 선형태: 이중선
 토폴로지: 선형

(2) SEQ ID NO 1의 상세한 서열

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```

CCATGGCTAT ACCCAACCTC GGTCTGGTC ACACCAGGAA CTCTCTGGTA AGCTAGCTCC      60
ACTCCCCAGA AACCAACCGC GCCAAATTGC CGGAATTGCT GACCTGAAGA CGGAACATCA      120
TCGTCGGGTC CTTGGGCGAT TGCGGCGGAA GATGGGTCAG CTTGGGCTTG AGGACGAGAC      180
CCGAATCGAG TCTGTTGAAA GGTGTTTCAT TGGGATTGTG ATACGGAGAT TGTCCTCGA      240
GAGGTTTGAG GGAAAGGACA AATGGGTTTG GCTCTGGAGA AAGAGAGTGC GGCTTTAGAG      300
AGAGAATTGA GAGGTTTAGA GAGAGATGCG GCGGCGATGA CGGGAGGAGA GACGACGAGG      360
ACCTGCATTA TCAAAGCAGT GACGTGGTGA AATTTGGAAC TTTTAAGAGG CAGATAGATT      420
TATTATTGTG ATCCATTTC TICATTGTTT TAGAATGTCG CGGAACAAAT TTTAAACTA      480
AATCCTAAAT TTTTCTAATT TTGTTGCCAA TAGTGGATAT GTGGGCCGTA TAGAAGGAAT      540
CTATTGAAGG CCCAAACCCA TACTGACGAG CCCAAAGGTT CGTTTTGCGT TTTATGTTTC      600
GGTTCGATGC CAACGCCACA TTCTGAGCTA GGCAAAAAC AAACGTGTCT TTGAATAGAC      660
TCCTCTCGTT AACACATGCA GCGGCTGCAT GGTGACGCCA TTAACACGTG GCCTACAATT      720
GCATGATGTC TCCATTGACA CGTGACTTCT CGTCTCCTTT CTTAATATAT CTAACAAACA      780
CTCCTACCTC TTCAAAATA TATACACATC TTTTGTATCA ATCTCTCATT CAAAATCTCA      840
TTCTCTCTAG TAAACAAGAA CAAAAAA ATG GCG GAT ACA GCT AGA GGA ACC CAT      894
                               Met Ala Asp Thr Ala Arg Gly Thr His
                               1                               5

CAC GAT ATC ATC GGC AGA GAC CAG TAC CCG ATG ATG GGC CGA GAC CGA.      942
His Asp Ile Ile Gly Arg Asp Gln Tyr Pro Met Met Gly Arg Asp Arg
 10                               15                               20                               25

GAC CAG TAC CAG ATG TCC GGA CGA GGA TCT GAC TAC TCC AAG TCT AGG      990
Asp Gln Tyr Gln Met Ser Gly Arg Gly Ser Asp Tyr Ser Lys Ser Arg
                               30                               35                               40

CAG ATT GCT AAA GCT GCA ACT GCT CTC ACA GCT GGT GGT TCC CTC CTT      1038
Gln Ile Ala Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Lys Lys
                               45                               50                               55

GTT CTC TCC AGC CTT ACC CTT GTT GGA ACT GTC ATA GCT TTG ACT GTT      1086
Val Lys Ser Ser Lys Thr Lys Val Gly Thr Val Ile Ala Lys Thr Val
                               60                               65                               70

GCA ACA CCT CTC CTC GTT ATC TTC AGC CCA ATC CTT GTC CCG GCT CTC      1134
Ala Thr Pro Lys Lys Val Ile Phe Ser Pro Ile Lys Val Pro Ala Lys
                               75                               80                               85

ATC ACA GTT GCA CTC CTC ATC ACC GGT TTT CTT TCC TCT GGA GGG TTT      1182
Ile Thr Val Ala Lys Lys Ile Thr Gly Phe Lys Ser Ser Gly Gly Phe
 90                               95                               100                               105

GGC ATT GCC GCT ATA ACC GTT TTC TCT TGG ATT TAC AAG TAAGCACACA      1231
Gly Ile Ala Ala Ile Thr Val Phe Ser Trp Ile Tyr Lys
                               110                               115

TTTATCATCT TACTTCATAA TTTTGTCGAA TAIGTGATG CATGTGTGA GCCAGTAGCT      1291

```

TTGGATCAAT TTTTTGGTC GAATAACAAA TGTAACAATA AGAAATTGCA AATTCTAGGG	1351
AACATTTGGT TAACTAAATA CGAAATTTGA CCTAGCTAGC TTGAATGTGT CTGTGTATAT	1411
CAICTATATA GGTA AAAATGC TTGGTATGAT ACCTATTGAT TGTGAATAGG TAC GCA ACG Tyr Ala Thr 1	1470
GGA GAG CAC CCA CAG GGA TCA GAC AAG TTG GAC AGT GCA AGG ATG AAG Gly Glu His Pro Gln Gly Ser Asp Lys Leu Asp Ser Ala Arg Met Lys 5 10 15	1518
TTG GGA AGC AAA GCT CAG GAT CTG AAA GAC AGA GCT CAG TAC TAC GGA Leu Gly Ser Lys Ala Gln Asp Leu Lys Asp Arg Ala Gln Tyr Tyr Gly 20 25 30 35	1566
CAG CAA CAT ACT GGT TGG GAA CAT GAC CGT GAC CGT ACT CGT GGT GGC Gln Gln His Tyr Gln Gln Glu His Asp Arg Asp Arg Thr Arg Gly Gly 40 45 50	1614
CAG CAC ACT ACT TAAGTTACCC CACTGATGTC ATCGTCATAG TCCAATAACT Gln His Tyr Tyr 55	1666
CCAATGTCCG GCAGTTAGTT TATCAGCAAT AAAGTGTTA GAATTTGATC AGGGGGAGAT	1726
AATAAAAGCC GAGTTTGAAT CTTTTTGTTA TAAGTAATGT TTATGTGTGT TTCTATATGT	1786
TGTCAAATCG TACCGAG	1803

3) 서열 2에 관한 정보

(1) 서열특성

길이: 22
 형태: 핵산
 선형태: 이중선
 토플로지: 선형

(2) SEQ ID NO 2의 상세한 서열

CACTGCAGGA ACTCTCTGGTAA 22

4) 서열 3에 관한 정보

(1) 서열특성

길이: 31
 형태: 핵산
 선형태: 이중선
 토플로지: 선형

(2) SEQ ID NO 3의 상세한 서열

CTACCCGGGA TCCTGTTTAC TAGAGAGAAT G 31

(57) 청구의 범위**청구항 1**

기름체 단백질 유전자에서 얻을 수 있는 전사 조절 부위, 상기 조절 부위에 이중인 목적 DNA 서열 및 전사 종결 부위로 이루어지며 이들이 5'-3' 방향의 전사에서 작동할 수 있도록 연결된 발현 카세트가 세포 계층에 인테그레이션된 세포를 가지며 종자를 발육시킬 수 있는 식물을, 상기 조절 부위의 전사조절 하에서 상기 DNA 서열이 발현되는 종자가 생산되는 조건하에서 기르는 것으로 이루어지는, 종자세포에서 목적 DNA 서열을 발현시키는 방법.

청구항 2

제1항에 있어서, 상기 유전자가 저장 단백질의 축적전의 배형성 단계에서 발현되는 방법.

청구항 3

제2항에 있어서, 상기 단계가 구형(globular) 배의 형성에서부터 초기 자엽(early cotyledonary) 단계까 지인 방법.

청구항 4

제1항에 있어서, 상기 유전자가 쌍떡잎 식물에서 유래된 것인 방법.

청구항 5

제2항에 있어서, 상기 단계가 구형(globular) 단계, 심장(hart) 단계, 어뢰(torpede) 단계 및 자엽(cotyledonary) 단계로 구성되는 군으로부터 선택되는 방법.

청구항 6

제4항에 있어서, 상기 쌍떡잎 식물이 Brassica napus 또는 Arabidopsis인 방법.

청구항 7

제1항에 있어서, 상기 식물이 Arabidopsis가 아닌 방법.

청구항 8

기름체 단백질을 지정하는 유전자에서 얻을 수 있는 전사 조절 부위와, 상기 조절 부위에 이중인 목적 DNA 서열이 연결된 것으로 이루어지는 DNA 구조물.

청구항 9

제8항에 있어서, 상기 유전자가 겨자과(Brassicaceae)에서 선택되는 식물로부터 얻어지는 DNA 구조물.

청구항 10

제8항에 있어서, 상기 전사 조절 부위가 Brassica napus 또는 Arabidopsis로부터 얻어지는 DNA 구조물.

청구항 11

기름체 단백질 유전자에서 얻을 수 있는 전사조절 부위, 상기 조절 부위에 이중인 목적 DNA 서열 및 전사 종결 부위로 이루어지며 이들이 5'-3'방향의 전사에서 작동할 수 있도록 연결된 발현 카세트.

청구항 12

기름체 단백질을 지정하는 유전자에서 얻을 수 있는 전사 조절 부위로 이루어지며, 상기 조절 부위가 상기 조절 부위에 이중인 목적 DNA 서열에 연결된 키메라 유전자가 세포 계놈에 인테그레이션된 세포를 가지는 식물.

청구항 13

기름체 단백질을 지정하는 유전자에서 얻을 수 있는 전사 조절 부위로 이루어지며, 상기 조절 부위가 상기 조절 부위에 이중인 목적 DNA 서열에 연결된 키메라 유전자가 세포 계놈에 인테그레이션된 세포를 가지는 종자.

청구항 14

제13항에 있어서, 상기 전사 조절 부위가 Brassica napus 또는 Arabidopsis에서 얻을 수 있는 기름체 단백질 유전자에서 유래된 종자.

청구항 15

제13항에 있어서, 상기 종자가 쌍떡잎 식물의 종자인 종자.

청구항 16

제14항에 있어서, 상기 종자가 유지종자인 종자.

청구항 17

기름체 단백질 유전자에서 얻을 수 있는 전사 조절 부위, 상기 조절 부위에 이중인 목적 DNA 서열 및 전사 종결 부위로 이루어지며 이들이 5'-3' 방향의 전사에서 작동할 수 있게 연결된 발현 카세트가 세포 계놈에 인테그레이션된 세포를 가지며 종자를 발육시킬 수 있는 식물을, 상기 조절 부위의 전사조절 하에서 상기 DNA 서열이 발현되는 종자가 생산되는 조건하에서 기르는 것으로 이루어지는, 종자 특이적 대사 기작을 변형시키는 방법.

청구항 18

제17항에 있어서, 상기 변형이 식물 종자에서 발현되는 내생 (endogenous) 유전자의 발현을 감소시키거나 억제하는 것인 방법.

청구항 19

제17항에 있어서, 상기 전사 개시 부위가 사일렌서 (silencer)를 포함하는 방법.

청구항 20

제17항에 있어서, 상기 DNA 서열의 전사된 사슬이 상기 세포의 내생 mRNA에 상보적인 방법.

청구항 21

5'-3'의 전사 방향으로, 기름체 단백질 유전자에서 얻을 수 있는 전사 개시 부위, 식물에 대해 외부 펩타이드를 지정하는 목적 DNA 서열 및 전사 종결 부위로 이루어지는 발현 카세트가 세포 계놈에 인테그레이션된 세포를 가지는 상기 식물을, 상기 개시 부위의 전사조절 하에서 상기 DNA 서열이 발현되는 종자가

생산되는 조건하에서 기르는 것으로 이루어지는, 종자에서 새로운 폴리펩타이드를 생산하는 방법.

청구항 22

제8항에 기재된 DNA 구조물을 가지는 식물의 부분.

청구항 23

제22항에 있어서, 상기 부분이 잎, 줄기, 뿌리, 꽃, 열매 또는 종자인 식물의 부분.

청구항 24

숙주 식물을, 기름체 단백질 유전자에서 얻을 수 있는 전사 조절 부위에 작동가능하도록 연결된 목적 DNA 서열로 이루어지는 구조물로 형질전환시키고; 상기 식물을 상기 조절 부위의 전사조절 하에서 상기 목적 DNA 서열이 발현되는 종자가 생산되는 조건하에서 기르는 것으로 이루어지는, 저장 단백질의 축적적인 배형성의 단계동안 숙주 식물에서 목적 DNA 서열을 발현시키는 방법.

청구항 25

제24항에 있어서, 상기 숙주가 쌍떡잎 유지 종자 세포인 방법.

청구항 26

제25항에 있어서, 상기 유전자가 *Brassica napus*, *Zea mays*, 당근 및 *Arabidopsis*로 구성되는 군으로부터 선택되는 방법.

청구항 27

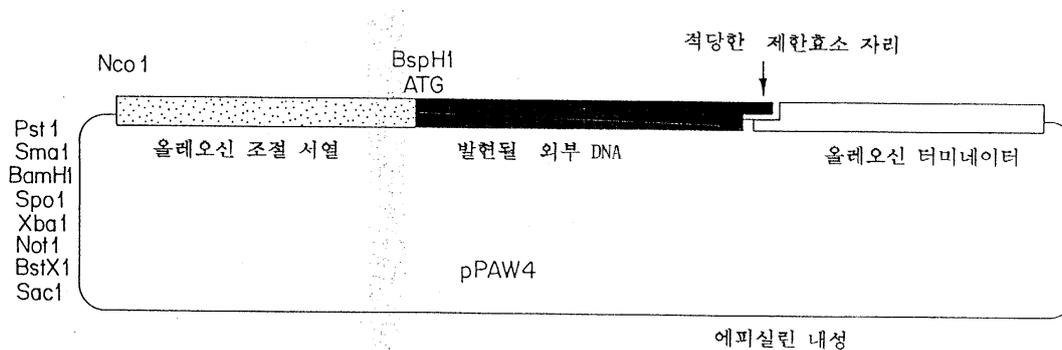
목적하는 폴리펩타이드 유전자를 종자에서 발현시키기 위하여 기름체 단백질 (oil body protein)의 번역 개시 자리쪽으로 충분한 5' 조절 부위를 포함하는 상기 기름체 단백질의 리딩프레임 (reading frame)에, 목적하는 폴리펩타이드를 지정하는 제1 DNA 서열이 종자에서 발현될 수 있고 상기 DNA 서열의 발현이 상기 조절 부위에 의해 조절될 수 있는 자리에 삽입된 것으로 이루어지는 DNA 구조물로, 상기 DNA 구조물이 상기 식물 세포의 게놈에 인테그레이션되도록 게놈 인테그레이션 조건에서 숙주 식물 세포를 형질전환시키고; 상기 식물을, 상기 목적하는 폴리펩타이드가 상기 OBP 유전자의 발현 산물과 융합 단백질로 발현되는 종자 단백질을 생산하도록 성장시키고; 상기 종자의 세포로부터 기름체를 분리하고; 상기 융합 단백질이 해리될 수 있도록 상기 기름체를 분쇄하고; 상기 목적하는 폴리펩타이드를 정제하는 것으로 이루어지는, 정제된 목적하는 폴리펩타이드를 얻는 방법.

청구항 28

종자에서 목적하는 폴리펩타이드를 올레오신의 융합 단백질, 또는 상기 기름체로 표적화하기에 충분한 올레오신의 일부분과의 융합 단백질로 발현시키는 것으로 이루어지는, 상기 기름체에서 목적하는 폴리펩타이드를 얻는 방법.

도면

도면1



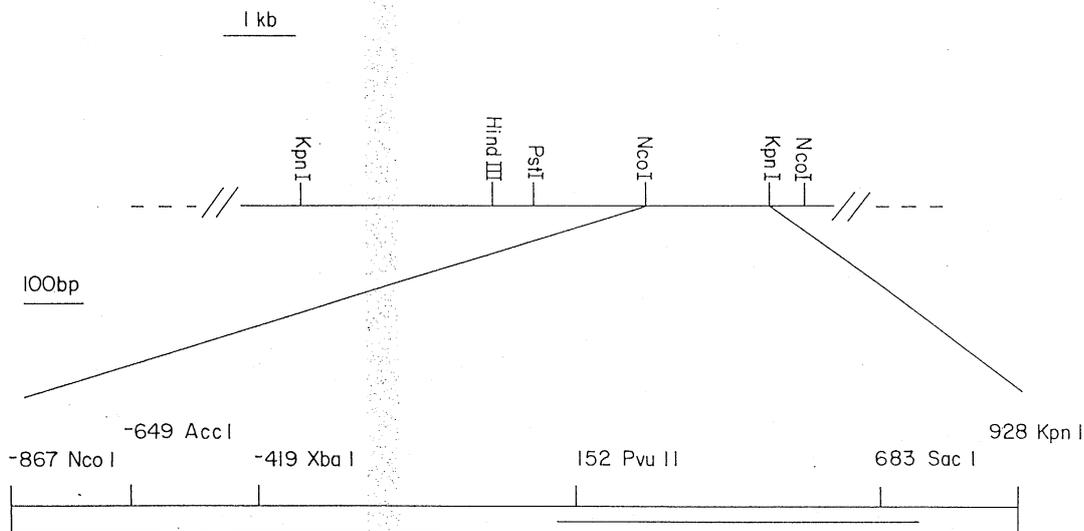
도면2a

-867 ^{NcoI} CCATGGCTATACCCAACCTCGGTCTTGGTCACACCAGGAACCTCTGGTAAGCTAGCTCCACTGCCAGAAAACACGGCGCCAAATTGC
 -777 CGGAATTGCTGACCTGAAGACGGAACATCATCGTCGGGCTCCTTGGGCGATTGCGGC6GAAGATGGGTCAGCTTGGGCTTGGGACGAGAC
 -687 CCGAATCGAGTCTGTGAAAGGTTGTTTCATTGGGATTGTATACGGAGATTGGTCGTCGAGAGGTTTGGGGAAAGGACAAATGGGTTTG
 -97 GCTCTGGAGAAAGAGAGTGGCGCTTTAGAGAGAGAATTGAGAGGTTTAGAGAGAGATGCGGCGGCGATGACGGGAGGAGAGACGACGAGG
 -507 ACCTGCATTATCAAAGCAGTGACGTGGTGAATTTGGAACCTTTAAGAGGCAGATAGATTTATTATTTGTATCCATTTTCTTCATTGTTTC
 -417 TAGAATGTCGCGGAACAAATTTTAAACTAAATCCTAAATTTTCTAATTTTGTGCAATAGTGGATATGTGGGCGTATAGAAGGAAT
 -327 CTATTGAAAGGCCAAACCCTACTGACGAGCCCAAAGGTTCTGTTTTGCGTTTTATGTTTCGGTTCGATGCCAACGCCACATTCTGAGCTA
 -237 GGCAAAAAACAACCGTGTCTTTGAATAGACTCCTCTCGTTAACACATGCGAGCGGCTGCATGGTGACGCCATTAACACGTGGCCTACAAT
 -147 GCATGATGCTCCATTGACACGTGACTTCTCGTCTCCTTTCTTAATATATCTAACAAACACTCCTACCTCTTCCAAAAATATAACACATC
 -57 TTTTGTGATCAATCTCTCATTCAAATCTCATTCTCTCTAGTAAACAAGAACAACAAAAATGGCGGATACAGCTAGAGGAACCCATCACGAT
 I I G R D Q Y P M M G R D R D Q Y Q M S G R G S D Y S K S R
 34 ATCATCGGCAGAGACCAGTACCCGATGATGGGCCGAGACCGAGACCAGTACCAGATGTCGGACGAGGATCTGACTACTCCAAGTCTAGG

도면2a-2

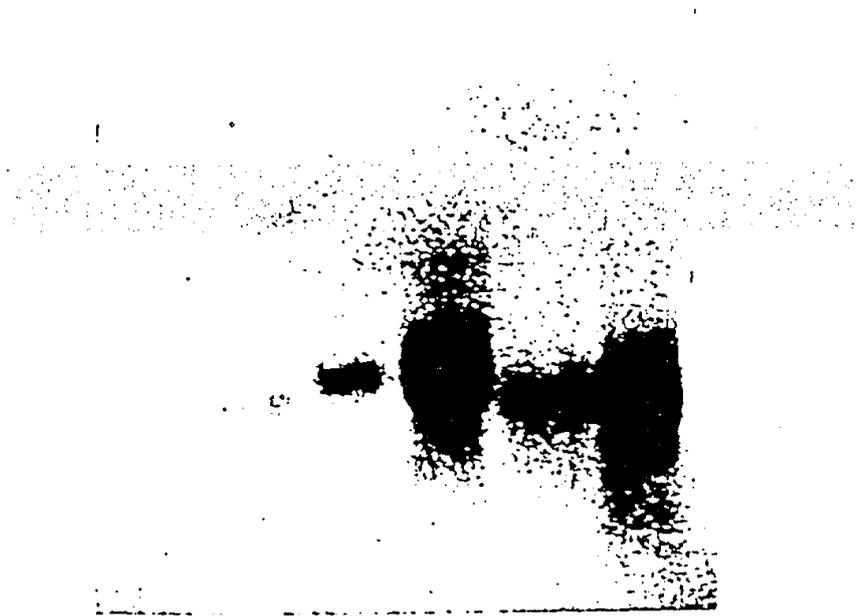
124 Q I A K A A T A V T A G G S L L L V L S S L T L V G T V I A A L
 CAGATTGCTAAAGCTGCAACTGCTGTACAGCTGGTGGTCCCTCCTTGTCTCTCCAGCCTTACCCTTGTGGAACTGTCATAGCTTTG
 214 T V A T P L L L V I F S P I L V P A L I T V A L L I T G F L S
 ACTGTTGCAACACCTCTGCTGTTATCTTCAGCCCAATCCTGTCCCGGCTCTCATCACAGTTGCACCTCATCACCGGTTTCTTTCC
 304 S G G F G I A A I T V F S W I Y K
 TCTGGAGGGTTTGGCATTGCCGTATAACCGTTTTCTCTTGGATTTACAAgtaagcacacatttatcatcttacttcataattttgtgca
 394 atatgtgcatgcatgtgtttgagccagtagctttggatcaatTTTTTggtcgaataacaaaatgtaacaataagaattgcaaatcttagg
 484 gaacatttggttaactaaaacgaaatttgacc tagc tagc ttgaaatgtgtctgtgtatcatctata taggtaaaaatgcttggatga
 574 tacctattgattgtgaaatagTACGCAACGGGAGAGACCCACAGGATCAGACAAGTTGGACAGTGCAAGGATGAAGTTGGGAAGCAAA
 Y A T G E H P Q G S D K L D S A R M K L G S K
 664 A Q D L K D R A Q Y Y G Q Q H T G G E H D R D R T R G G Q H
 GCTCAGGATCTGAAAGACAGAGCTCAGTACTACGGACAGCAACATACTGGTTGGGAACATGACCGTGACCGTACTCGTGGTGGCCAGCAC
 754 T T *
 ACTACTTAAGTTACCCACTGATGTCATCGTCATAGTCCAATAACTCCAATGTCGGGGAGTTAGTTTATGAGGAATAAAGTGTTTAGAAT
 844 TTGATCAGGGGGAGATAATAAAGCCGAGTTTGAATCTTTTTGTTATAAGTAATGTTTATGTGTGTTTCTATATGTTGCAATGGTACC
^{KpnI}

도면2b



도면3a

H T C
- + - + - +

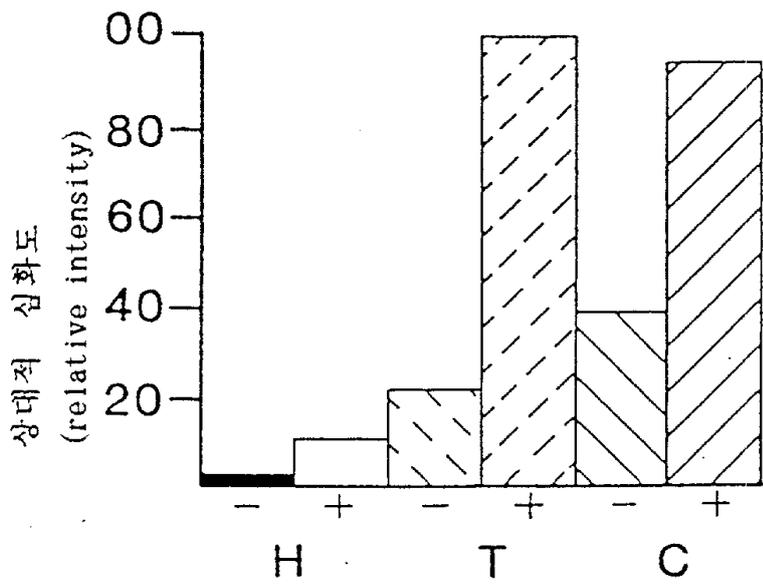


도면3b

H T C
· + · + · +

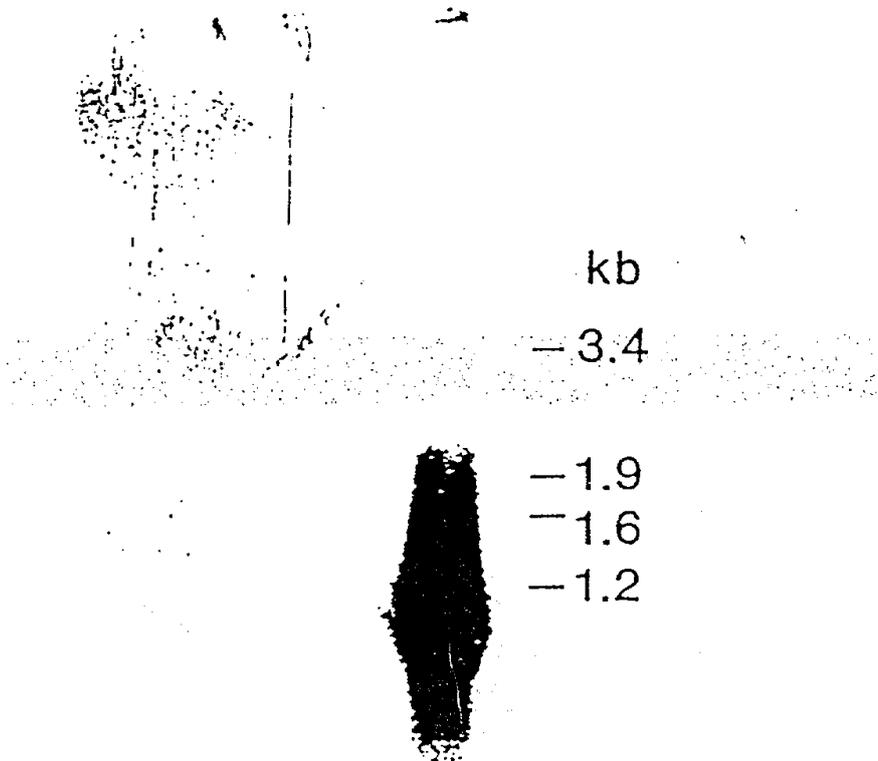


도면3c



도면4

R Ca Co L E



도면5

