(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) 。Int. Cl. 7
C12N 15/29

(45) 공고일자 2002년10월04일
(11) 등록번호 10-0354633
(24) 등록일자 2002년09월16일

(21) 출원번호	10-2001-7001201	(65) 공개번호	특0000-0000000
(22) 출원일자	2001년01월29일	(43) 공개일자	0000년00월00일
(62) 원출원	특허특1998-0710210		
	원출원일자 : 1998년12월12일	심사청구일자	1998년12월12일
번역문 제출일자	2001년01월29일		
(86) 국제출원번호	PCT/US1997/09743	(87) 국제공개번호	WO 1997/47731
(86) 국제출원출원일자	1997년06월10일	(87) 국제공개일자	1997년12월18일

(81) 지정국 국내특허 : 아일랜드, 알바니아, 오스트레일리아, 보스니아-헤르체고비나, 바베이도스, 불가리아, 브라질, 캐나다, 중국, 쿠바, 체코, 에스토니아, 그루지야, 헝가리, 이스라엘, 아이슬란드, 일본, 북한, 대한민국, 세인트루시아, 스리랑카, 라이베리아,
AP ARIPO특허: 케냐, 레소토, 말라위, 수단, 스와질랜드, 우간다,
EA 유라시아특허: 아르메니아, 아제르바이잔, 벨라루스, 키르기즈, 카자흐스탄,
EP 유럽특허: 오스트리아, 벨기에, 스위스, 독일, 덴마크, 스페인, 프랑스, 영국, 그리스, 이탈리아, 룩셈부르크, 모나코, 네덜란드, 포르투칼, 스웨덴,

(30) 우선권주장 60/019,940 1996년06월14일 미국(US)

(73) 특허권자 이.아이,듀우판드네모아앤드캄파니
미합중국19898델라웨어주윌밍톤마킷트스트리트1007

(72) 발명자 킨니,안토니,제이.
미국19809델라웨어주윌밍톤로어애비뉴609
페이더,게리,마이클
미국19350펜실바니아주랜던버그우즈레인100

(74) 대리인 주성민
장수길

심사관 : 최규환

(54) 대두 종자 단백질 유전자 특정 계열의 억제 방법

요약

본 발명은 종자 저장 단백질을 코딩하는 유전자의 발현을 조절하여 유전자전환 식물의 종자 저장 단백질 프로파일을 변화시키는 유전자전환 대두의 제조에 관한 것이다. 종자 저장 단백질 프로파일 변화 결과, 독특하고 유용한 기능성을 갖는 신규의 대두 단백질 제품을 생산할 수 있다.

대표도

도 1

색인어

대두 종자 저장 단백질, 유전자 전환, 대두 단백질

명세서

도면의 간단한 설명

도 1은 본 발명의 키메라 유전자의 제작에서 중간 클로닝 비히클로서 사용되는 플라스미드 pML70의 제한 지도(restriction map)이다.

도 2는 본 발명의 키메라 유전자의 제작에서 중간 클로닝 비히클로서 사용되는 플라스미드 pCW109의 제한 지도이다.

도 3은 본 발명의 키메라 유전자의 제작에서 중간 클로닝 비히클로서 사용되는 플라스미드 pKS18HH의 제한 지도이다.

도 4는 플라스미드 pJo1의 제한 지도이다. 이 플라스미드는 식물 전사 단위 KTi 프로모터/β-콘글리시닌의 절단된 α 서브유닛/KTi 3′ 말단을 pKS18HH의 BamH I 위치에 클로닝함으로써 유도되었다.

도 5는 pJo1로 형질전환시킨 체세포 배로부터 추출된 단백질의 SDS-PAGE이다.

도 6은 플라스미드 pBS43의 제한 지도이다. 이 플라스미드는 대두 β-콘글리시닌 프로모터의 전사 조절하에 글리신 맥스 미크로조말 델타-12 데세츄라제(Glycine max microsomal delta-12 desaturase)를 코딩하는 핵산 서열을 포함한다.

도 7은 pBS43으로 형질전환시킨 식물로부터 얻어진 대두 종자로부터 추출된 단백질의 SDS-PAGE이다.

도 8은 플라스미드 pJo3의 제한 지도이다. 이 플라스미드는 식물 전사 단위 KTi 프로모터/β-콘글리시닌의 α 서브유닛의 완전 길이 cDNA/KTi 3′ 말단을 pKS18HH의 HindIII 위치에 클로닝함으로써 유도되었다.

도 9는 플라스미드 pRB20의 제한 지도이다. 이 플라스미드는 전사 단위 β-콘글리시닌 프로모터/파세올린 3′ 말단을 pKS18HH의 HindIII 위치에 클로닝함으로써 유도되었다. 이는 본 발명의 키메라 유전자의 제작에서 중간 클로닝 비히클로서 사용된다.

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 종자 저장 단백질을 코딩하는 유전자의 발현을 조절하여 형질전환 식물의 종자 저장 단백질 프로파일을 변화시키는 형질전환 대두 계열의 제조에 관한 것이다. 이러한 변형된 형질전환 대두 계열은 독특하고 유용한 기능성을 갖는 신규 대두 단백질 제품의 생산에 사용된다.

대두 종자는 건조 중량 기준으로 35 내지 55 % 단백질을 함유한다. 이 단백질의 대부분은 저장 단백질이고, 이는 발아하는 동안 가수분해되어, 성장하는 묘목에 필요한 에너지 및 대사 중간 물질을 제공한다. 대두 종자의 저장 단백질은 수확시의 중요한 영양 공급원이고 가축 사료로서 사용된다. 또한, 대두는 통상적으로, 사람이 소비하기에 가장 경제적인 단백질 공급원이라고 알려져 있다. 대두 단백질 또는 단백질 단리물은 이미 세계 여러 다른 지역에서 식품으로서 널리 사용되고 있다. 대두 종자내 저장 단백질의 양과 질을 개선하는데 많은 노력을 기울이고 있다.

대부분의 식물류 종자는 당업계 종자 저장 단백질로서 공지된 것을 함유한다. 식물류 종자는 단백질의 크기 및 용해도를 기준으로 분류된다(히긴스, 티. 제이(Higgins)의 문헌[Ann. Rev. Plant Physiol. 35: 191-221(1984)]). 모든 종류의 단백질이 모든 식물류에서 발견되는 것은 아니지만, 대부분의 식물류 종자는 한 종류 이상의 단백질을 함유한다. 특정 용해도 또는 크기를 갖는 단백질 계열은 통상적으로, 동일한 식물류 범위내 다른 종류의 단백질보다 상이한 식물류 범위내 동일한 종류의 단백질에 구조적으로 좀더 연관되어 있다. 많은 식물류에 있어서, 특정 계열의 종자 단백질은 서열 상동성을 근거로 하여 서브클래스로 나누어질 수 있는 아주 복잡한 복수 유전자(multigene)군에 의해서 코딩된다.

두 종의 주요 대두 종자 저장 단백질, 즉, 글리시닌(이는 또한 11S 글로불린으로서 공지됨) 및 β-콘글리시닌(이는 또한 7S 글로불린으로서 공지됨)이 있다. 또한, 이들은 종자의 총 단백질의 70 내지 80 % 또는 종자 건조 중량의 25 내지 35 %를 이룬다. 글리시닌은 약 360 kDa 분자량을 갖는 큰 단백질이다. 이는 G1, G2, G3, G4 및 G5의 5 개 주요 이소형(통상, 서브유닛이라고 부름)의 다양한 조합으로 이루어진 육량체이다. 각 서브유닛은 이황화 결합에 의해서 결합된 하나의 산성 및 하나의 염기성 폴리펩티드로 이루어진다. 하나의 서브유닛의 산성 및 염기성 폴리펩티드 모두는 단일 유전자에 의해서 코딩된다. 그러므로, 5 종의 글리시닌 서브유닛을 코딩하는 5 개의 비-대립 유전자가 있다. 이들 유전자는 각 서브유닛 G1, G2, G3, G4 및 G5에 상응하여 Gy1, Gy2, Gy3, Gy4 및 Gy5로서 명명된다(닐센, 엔. 씨(Nielsen, N. C.)의 문헌[Plant Cell 1: 313-328(1989)]).

글리시닌 서브유닛 유전자에 대한 게놈 클론 및 cDNA는 서열화되었고, 뉴클레오티드 및 아미노산 서열 유사성을 근거로 하여 2 개의 군으로 나누어진다. 군 I은 Gy1, Gy2 및 Gy3을 포함하고, 군 II는 Gy4 및 Gy5를 포함한다. 하나의 군 범위내 유전자 간에는 85 %를 넘는 유사성이 있지만, 군 I 및 군 II의 유전자 간에는 단지 42 내지 46 %의 유사성이 있다.

β-콘글리시닌(7S 글로불린)은 150 내지 240 kDa의 분자량을 갖는 이종 당단백질이다. 이는 α, α' 및 β, 즉, 3 종의 음으로 크게 하전된 서브유닛의 다양한 조합으로 이루어진다. α 및 α' 서브유닛을 코딩하는 유전자의 코딩 부위를 나타내는 cDNA 클론은 서열화되어있고, 유사한 크기이며, 서열 상동성은 85 %로 제한된다. 그러나, β 서브유닛의 코딩 부위를 나타내는 cDNA의 서열은 α 및 α' cDNA보다 약 0.5 kb 정도 더 작다. 이러한 차이를 배제하고 α 및 α' 서브유닛에 대한 서열 상동성은 75 내지 80 %이다. 3 종의 β-콘글리시닌 서브유닛은 게놈 대두내 몇몇 부위에 밀집된 총 15 개 서브유닛 유전자에 의해서 코딩된다(하라다, 제이. 제이.(Harada, J. J.) 등의 문헌[Plant Cell 1 : 415-425(1989)]).

전통적 동양식 대두 기재 식품이 필수적이지 않은 나라에서도 단백질 농축물, 단리물, 및 텍스쳐드(textured) 단백질 제품과 같은 신규 대두 기재 제품을 점차적으로 많이 이용하고 있다. 그러나, 식품 분야에서 이러한 신규 제품의 용도는 지역적인 기호 및 조리시의 요구 조건에 대한 단백질 제품의 기능적 특성에 따라서 다르다. 지난 10년 동안, 대두 단백질의 기능적 특성을 이해하기 위한 많은 노력을 기울여 왔다. 기능적 특성의 예로는 물 흡착 파라미터, 습윤성, 팽창성, 수분 보유력, 용해도, 점착성, 점도, 응고성, 겔화 특성 및 유화성이다. 연구의 주된 부분은 β-콘글리시닌 및 글리시닌 단백질 각각의 연구 뿐 아니라 이들 각 단백질이 대두 단백질계에 전체적으로 어떻게 영향을 주는가 하는 연구에 중점

을 두고 있다(킨셀라, 제이. 이.(Kinsella, J. E.) 등의 문헌[New Protein Foods 5: 107-179(1985)]; 모르, 씨. 브이.(Morr. C. V.)의 문헌[JAOCS 67: 265-271(1987)]; 펭, 엘. 씨.(Peng, L. C.) 등의 문헌[Cereal Chem 61:480-489(1984)]). 기능적 특성은 단백질의 물리화학적 성질에 직접 연관되어 있기 때문에, β-콘글리시닌 및 글리시닌의 구조적 차이는 이들 두 단백질로 하여금 매우 상이한 기능적 특성을 갖게 한다. 열 응집(aggregation), 유화성, 및 수분 보유력에서의 차이가 보고되고 있다. 또한, 글리시닌을 가하면 겔화성이 변하여 보다 큰 신장 응력, 스트레스 및 전단력을 갖는 겔을 생성하고, 용매 보유력이 보다 좋아지고 탁도가 보다 낮아진다. 그러나, 오늘날 제조된 대두 단백질 제품은 글리시닌 및 β-콘글리시닌 두 종류의 블렌드이고, 그러므로 글리시닌 및 β-콘글리시닌의 블렌드 각각의 특성에 따라서 다른 기능적 특성을 갖는다. 예를 들어, 글리시닌이 100 ℃로 가열되면, 단백질의 약 50 %가 가용성 응집물로 신속하게 전환된다. 또한, 가열 결과, 응집물이 증가하고 침전이 발생한다. 이 침전물에는 글리시닌의 염기성 폴리펩티드가 포함되어 있고, 산성 폴리펩티드는 용해된 상태로 남아 있다. β-콘글리시닌의 존재는 염기성 폴리펩티드와 가용성 복합물을 형성시킴으로써 염기성 폴리펩티드의 침전을 억제한다. 열 변성이 바람직한지 아닌지는 목적하는 용도에 따라서 다르다. 단지 한 종류 또는 나머지 한 종류의 저장 단백질만을 함유하는 대두 단백질 제품을 제조할 수 있다면, 특정 대두 단백질로부터 유도된 특정 물리적 특성을 필요로 하는 제품을 제조하는데 유용하거나 보다 경제적일 것이다.

지난 20년에 동안, 여러 가지 저장 단백질 서브유닛 중 1 종 이상이 결핍된 대두 계열(널(null) 돌연변이)이 대두 생식세포질에서 확인되었거나 돌연변이적 육종 기술을 사용하여 생산되었다. 돌연변이 사건을 결합하려는 육종 연구 결과, β-콘글리시닌 표준량의 약 반을 함유하는 대두 종자 계열을 생산하였다(다까사시, 케이.(Takashashi, K.) 등의 문헌[Breeding Science 44: 65-66(1994); 기따무라, 제이.(Kitamura, J.) 등의 문헌[JARQ 29:1-8(1995)]. 또한, 3 종의 독립 열성 돌연변이에 의해서 β-콘글리시닌의 감소가 조절된다. 글리시닌 서브유닛 널 돌연변이를 재조합함으로써 글리시닌의 양을 크게 감소시킨 계열의 종자를 생산하였다(기따무라, 제이.(Kitamura, J.) 등의 문헌[JARQ 29:1-8(1995)]). 또한, 3 종의 독립 열성 돌연변이에 의해서 β-콘글리시닌의 감소가 조절된다. 글리시닌만을 함유하거나 β-콘글리시닌만을 함유하는 상기 계열의 종자로부터 농업적으로 존속가능한 대두 변종을 개발하는 것은 시간이 많이 걸리고 비용도 많이 들 것이다. 각각의 이종교배는 3 종의 돌연변이 사건을 독립적으로 분리시킬 것이다. 또한, 각각의 돌연변이 사건은 동형접합 상태이어야 할 것이다. 농업적으로 우수한 대두 계열을 고수율로 얻으려는 개발은 많은 세대 동안 수많은 자손의 스크리닝 및 분석을 필요로 할 것이다.

종자내 특정 저장 단백질을 감소시키기 위해서 안티센스 기술을 사용하여 왔다. 브라시카 나푸스(Bracicca napus) 중의 나핀(2S 알부민) 및 크루시페린(11S 글로불린)은 두 개의 주요 저장 단백질이고, 이들은 총 종자 단백질의 각각 약 25 % 및 60 %를 이룬다. 나핀 단백질은 16 개 이하 유전자의 복수 유전자군에 의해서 코딩된다. 몇몇 cDNA 및 게놈 클론이 서열화되었다(조세프손, 엘.쥐.(Joseffson, L.G.) 등의 문헌[J. Biol Chem 262: 12196-12201(1987); 쇼필드, 에스.(Schofield, S.) 및 크로우취, 엠. 엘. (Crouch, M. L.(1987)) 등의 문헌[J. Biol. Chem. 262: 12202-12208(1987)]. 이 유전자는 이들의 코딩 및 플랭킹(franking) 부위 모두에서 90 %가 넘는 서열 상동성을 나타낸다. 크루시페린 유전자군도 동일하게 복잡하며, 이는 총 8 개의 유전자를 갖는 세 서브패밀리를 포함한다(로빈. 제이(Robin. J.) 등의 문헌[Plant Mol. Biol. 20: 559-563(1992)]). 코노-무라즈(Kohno.-Murase) 등의 문헌[Plant Mol. Biol. 26: 1115-1124(1994)])에는 나핀을 소량 함유하거나 함유하지 않는 형질전환 식물 종자를 제작하기 위해서, 나프A 프로모터에 의해서 유도된 나프A 유전자를 사용하는 나핀 안티센스 유전자를 사용할 수 있다고 기재하고 있다.

동일한 그룹(코노-무라즈 등의 문헌[Theoret. Applied Genetics 91: 627-631(1995)]은 나프A 프로모터의 조절하에 크루시페린 유전자(cruA, 알파 2/3 이소형 코딩)의 안티센스형을 발현시킴으로써 브라시카 나푸스내 크루시페린(11S 글로불린) 발현을 감소시키는 시도를 하였다. 이 경우에서의 결과는 보다 복잡하였다. 크루시페린은 서열 상동성(알파1, 2/3 및 4)을 근거로 하여 3 개의 서브클래스로 나누어지는데 즉, 각 종류는 서로에 대하여 60 내지 75 %의 서열 상동성을 갖는다(로빈. 제이 등의 문헌[Plant Mol. Biol. 20: 559-563(1992)]). 알파 2/3 이소형을 코딩하는 안티센스 유전자의 발현은 보다 낮은 수준의 알파 1 및 2/3 이소형을 생성하였다. 그러나, 알파 4 종류의 발현은 감소되

지 않았다.

안티센스 기술을 사용하여 쌀 중의 종자 저장 단백질, 즉, 글루텔린의 수준을 감소시켰다. 완전한 길이의 안티센스 글루텔린 코딩 부위에 작동가능하게 결합된 종자 특이 글루텔린 프로모터의 발현은 글루텔린 단백질 수준에 있어서 약 25 %의 감소를 나타내었다(미국 특허 제5,516,668호).

발명이 이루고자 하는 기술적 과제

본 발명은 종자 저장 단백질을 코딩하는 유전자의 발현을 조절하여 유전자전환 식물의 종자 저장 단백질 프로파일을 변화시키는 유전자전환 대두의 제조에 관한 것이다.

발명의 구성 및 작용

본 발명은 대두 중의 글리시닌 또는 β - 콘글리시닌(각각 11S 또는 7S 글로불린) 종자 저장 단백질의 양을 감소시키는 방법을 제공한다. 하나의 실시태양에서, 공동억제(cosuppression) 기술을 사용하여 종자 단백질 유전자의 7S 글로불린류를 코딩하는 유전자의 발현을 억제하였다. β - 콘글리시닌의 하나의 서브클래스 (α)를 코딩하는 유전자의 발현에 의해서 7S - 글로불린의 두 개 (α 및 α ') 또는 세 개 (α, α ' 및 β) 모두의 서브클래스를 코딩하는 유전자가 억제되었고, 결과적으로 변화된 종자 저장 단백질 프로파일을 갖는 대두 계열을 생산하였다. 또다른 실시태양에서, 완전하게 다른 두 종의 유전자(이들 중 한 종은 종자 단백질 유전자임)를 억제시키는 방법은 종자 조성을 크게 변화시키는 것으로 나타났다. 놀랍게도, 형질전환 대두 종자의 지방산 프로파일을 변화시키도록 발현하는 대두 유전자의 코딩 부위와 작동가능하게 결합된 대두 종자 저장 단백질의 프로모터 부위를 함유하는 키메라 유전자의 발현 결과, 두 개의 뚜렷한 표현 형질, 즉 종자 저장 단백질 프로파일 및 종자 오일 프로파일을 동시에 변화시켰다.

대두 종자 저장 단백질의 양을 감소시키는 방법은

(a) (i) 대두 종자 세포 내에서 기능적인 프로모터를 코딩하는 핵산 단편,

(ii) (i)의 프로모터에 대하여 센스 또는 안티센스 배향으로 놓인, 대두 종자 저장 단백질의 전부 또는 일부를 코딩하는 핵산 단편, 및

(iii) 전사 종결 부위

를 함유하는 키메라 유전자를 제작하는 단계;

(b) (a)의 키메라 유전자를 대두 세포 중에 도입하여 형질전환 대두 세포를 생성하는 단계; 및

(c) 단계 (a)의 키메라 유전자를 발현시키는 조건하에 단계 (b)의 형질전환 대두 세포를 성장시키는 단계

로 이루어지고, 단계 (a)의 키메라 유전자를 함유하지 않는 대두와 비교하여 대두 종자 저장 단백질 서브유닛 계열 중 1 종 이상의 양이 감소된다고 본 명세서에 교시되어 있다.

< 서열의 간단한 설명

본 발명은 하기 상세한 설명 및 본 출원의 일부를 이루는 서열 설명으로부터 보다 충분히 이해할 수 있을 것이다. 본 명세서에 참고로 인용된 37 C.F.R.1.822에 정의된 바와 같이, 서열 설명은 아미노산을 위한 3 개의 문자 코드를 포함한다.

서열 번호: 1은 β-콘글리시닌 대두 종자 저장 단백질의 α 서브유닛을 코딩하는 뉴클레오티드 서열을 5′에서 3′방향으로 나타낸다.

서열 번호: 2는 β-콘글리시닌 대두 종자 저장 단백질의 α′ 서브유닛을 코딩하는 뉴클레오티드 서열을 5′에서 3′방향으로 나타낸다.

서열 번호: 3은 β-콘글리시닌 대두 종자 저장 단백질의 β 서브유닛을 코딩하는 뉴클레오티드 서열을 5′에서 3′방향으로 나타낸다.

서열 번호: 4 및 5는 β-콘글리시닌 대두 종자 저장 단백질의 α 및 α′ 서브유닛을 코딩하는 핵산 단편을 단리하는데 사용되는 PCR 프라이머 ConS 및 Con1.4a의 뉴클레오티드 서열(각각)을 나타낸다.

서열 번호: 6 및 7은 β-콘글리시닌 대두 종자 저장 단백질의 α 및 α′ 서브유닛을 코딩하는 핵산 단편을 구별하는데 사용되는 PCR 프라이머 Con.09 및 Con.8의 뉴클레오티드 서열(각각)을 나타낸다.

서열 번호: 8 및 9는 β-콘글리시닌 대두 종자 저장 단백질의 α 및 α′ 서브유닛을 코딩하는 완전한 길이의 cDNA를 단리하는데 사용되는 PCR 프라이머 Con.Sa 및 Con1.9a의 뉴클레오티드 서열(각각)을 나타낸다.

서열 번호: 10은 β-콘글리시닌 대두 종자 저장 단백질의 α 및 α′서브유닛을 코딩하는 완전한 길이의 cDNA를 확인하는데 사용되는 PCR 프라이머 Con.1.0의 뉴클레오티드 서열을 나타낸다.

서열 번호: 11, 12 및 13은 군 I 글리시닌 대두 종자 저장 단백질의 Gy1, Gy2 및 Gy3 서브유닛을 코딩하는 뉴클레오티드 서열(각각)을 5′에서 3′방향으로 나타낸다.

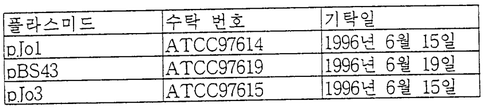
서열 번호: 14 및 15는 군 II 글리시닌 대두 종자 저장 단백질의 Gy4 및 Gy5 서브유닛을 코딩하는 뉴클레오티드 서열(각각)을 5′에서 3′방향으로 나타낸다.

서열 번호: 16, 17 및 18은 군 I 글리시닌 대두 종자 저장 단백질의 서브유닛을 코딩하는 cDNA를 단리하는데 사용되는 PCR 프라이머 G1-1, G1-1039 및 G1-1475의 뉴클레오티드 서열(각각)을 나타낸다.

서열 번호: 19, 20 및 21은 군 I 글리시닌 대두 종자 저장 단백질의 서브유닛을 코딩하는 cDNA를 단리하는데 사용되는 PCR 프라이머 G4-7, G4-1251 및 G4-1670의 뉴클레오티드 서열(각각)을 나타낸다.

< 생물학적 기탁

하기 플라스미드는 부다페스트 조약에 의해 미국 20852 매릴랜드주 록크빌 파크로운 드라이브 12031 소재의 아메리칸 타입 컬쳐 콜렉션(American Type Culture Collection(ATCC))에 기탁하고, 하기 수탁 번호를 얻었다.



플라스미드	수탁 번호	기탁일
pJo1	ATCC97614	1996년 6월 15일
pBS43	ATCC97619	1996년 6월 19일
pJo3	ATCC97615	1996년 6월 15일

< 정의

본 명세서의 관련하여 많은 용어가 사용된다. 용어 '핵산'은 당, 포스페이트 및 퓨린 또는 피리미딘을 함유하는 단량체(뉴클레오티드)로 이루어지는, 단일-가닥 또는 이중-가닥일 수 있는 거대 분자를 말한다. '핵산 단편'은 주어진 핵산

분자의 단편이다. 고등 식물에서 디옥시리보핵산(DNA)은 DNA내 정보를 단백질에 전달하는데 리보핵산(RNA)이 관련되는 유전 물질이다. '게놈'은 유기체의 각 세포에 함유된 유전 물질의 완전체이다. '뉴클레오티드 서열'이라는 용어는 DNA 또는 RNA 중합체의 서열을 말하며, 이는 단일 또는 이중-가닥일 수 있으며, 임의로, DNA 또는 RNA 중합체 중에 삽입할 수 있는 합성, 비-천연 또는 변형된 뉴클레오티드 염기를 포함한다.

본 명세서에 사용된 바와 같이, '상동'이라는 용어는 두 개의 핵산 분자의 뉴클레오티드 서열간 또는 두 개의 단백질 분자의 아미노산 서열간의 밀접한 관련성을 말한다. 당업자에 의해서 잘 이해되는 엄격한 조건하에 DNA-DNA 또는 DNA-RNA 혼성화에 의해서(햄스(Hames) 및 히긴스(Higgins)의 문헌[Nucleic Acid Hybridisation, IRL Press, Oxford, U.K.(1985)]), 또는 니들맨(Needleman) 등의 문헌[J. Mol. Biol. 48:443-453(1970)]의 방법과 같은 두 핵산 또는 단백질 사이의 서열 유사성의 비교에 의해서 이러한 상동성의 평가를 제공한다.

본 명세서에 사용된 바와 같은, '본질상 유사한'이라는 말은 염기 변화와 관련하여 코딩된 아미노산에서 변화를 일으키지 않는 DNA 서열을 말하거나, 염기 변화와 관련하여 하나 이상의 아미노산을 변화시킬 수 있지만 DNA 서열에 의해서 코딩된 단백질의 기능적 특성에는 영향을 주지 않는 DNA를 말한다. 그러므로, 본 발명은 예시한 특정 서열 이상의 것을 함유하는 것으로 이해된다. 생성되는 단백질 분자의 기능적 특성에 실질적으로 영향을 주지않는 침묵(silent) 변화를 일으키는 서열의 삭제, 삽입 또는 치환과 같은 서열에 대한 변형이 또한 예상된다. 이는 예를 들어, 유전 암호의 퇴화를 반영하거나, 주어진 위치에 화학적으로 등가의 아미노산을 생성하는 유전자 서열의 변형이 예상된다. 따라서, 아미노산 알라닌, 즉, 소수성 아미노산에 대한 코돈은 글리신, 발린, 류신 또는 이소류신과 같은 또다른 소수성 아미노산 잔기를 코딩하는 코돈에 의해서 치환될 수 있다. 유사하게, 글루탐산에 대한 아스파르트산과 같이 또다른 잔기에 대하여 하나의 음으로 하전된 잔기로 치환시키는 변화, 또는 아르기닌에 대한 리신과 같이 또다른 잔기에 대하여 하나의 양으로 하전된 잔기로 치환시키는 변화도 또한 생물학적 등가의 생성물을 제작할 수 있을 것으로 예상할 수 있다. 또한, 단백질 분자의 N-말단 및 C-말단 부분의 변화를 일으키는 뉴클레오티드 변화도 단백질 활성을 변화시키지 않을 것으로 예상된다. 단백질의 생물 활성에 대한 변화의 효과를 연구하기 위해서, 어떤 경우에는 서열의 돌연변이체를 제작하는 것이 실제 필요할 수도 있다. 제시된 각 변형들은 코딩된 생성물의 생물학적 활성의 보유 측정과 같이 당업계의 통상적 기술 범위 내에 잘 공지되어 있다. 더우기, 당업자들은 본 발명에 포함된 '본질상 유사한' 서열 또한 본 명세서에 예시된 서열과 함께 엄격한 조건(0.1 X SSC. 0.1 % SDS, 65 ℃)하에서 혼성화하는 그들의 능력에 의해서 정의할 수 있다고 알고 있다.

'유전자'는 코딩 부위 앞(5'비-코딩) 조절 서열 및 뒤(3'비-코딩) 조절 서열를 포함하는 특정 단백질을 발현하는 핵산 단편을 말한다. '고유' 유전자는 천연에서 발견되는 것과 같은 자체 조절 서열을 갖는 단리된 유전자를 말한다. '키메라 유전자'는 자연에서 발견되지 않는 이종의 조절 및 코딩 서열을 포함하는 유전자를 말한다. '내생' 유전자는 통상적으로, 게놈내 고유한 위치에서 발견되는 천연 유전자를 말한다. '외래' 유전자는 숙주 유기체 중에서 정상적으로 발견되지 않지만 유전자 전달에 의해서 도입되는 유전자를 말한다.

'코딩 서열' 또는 '코딩 부위'는 특정 단백질을 코딩하고, 비-코딩 서열을 제외하는 DNA서열을 말한다. 이는 '비중단된(uninterrupted) 코딩 서열'을 구성할 수 있는데 즉, 인트론이 없거나 적당한 스플라이싱 접합부(splice junction)에 의해서 결합된 하나 이상의 인트론을 포함할 수도 있다. '인트론'은 제1 전사체에서 전사되는 뉴클레오티드 서열이지만, 이는 세포내 RNA의 분해 및 재-결합을 통해서 제거되어 단백질로 번역될 수 있는 성숙 mRNA를 생성한다.

'개시 코돈' 및 '종결 코돈'은 단백질 합성(mRNA 번역)의 개시 및 사슬 종결 각각을 구체화하는 코딩 서열 중 3 개의 인접한 뉴클레오티드 단위를 말한다. '오픈 리딩 프레임(open reading frame)'은 아미노산 서열을 코딩하는 개시 및 종결 코돈 사이의 인트론에 의해서 중단되지 않는 코딩 서열을 말한다.

'RNA 전사체'는 DNA서열의 RNA 중합효소-촉매된 전사로부터 생성되는 생성물을 말한다. RNA 전사체는 DNA 서열의 완전한 상보적 카피이며, 이는 제1 전사체로서 불리거나, 제1 전사체의 전사후 반응으로부터 유도된 RNA 서열일 수 있고, 성숙 RNA로 불린다. '메신저 RNA(mRNA)'는 인트론이 없는 것으로, 세포에 의해서 단백질로 번역될 수 있는 RNA를 말한다. 'cDNA'는 mRNA로부터 유도되고 이에 상보적인 이중-가닥 DNA를 말한다. '센스' RNA는 mRNA를 포함하는 RNA 전사체를 말한다. '안티센스' RNA는 목적하는 제1 전사체 또는 mRNA의 모두 또는 일부에 상보적이고, 목적 유전자의 발현을 차단하는 RNA 전사체를 말한다. 특정 유전자 전사체의 임의의 부분, 즉, 5′ 비-코딩 서열, 3′ 비-코딩 서열, 인트론 또는 코딩 서열에 안티센스 RNA의 상보성이 존재할 수 있다.

본 명세서에 사용된 바와 같이, '적합한 조절 서열'은 본 발명의 핵산 단편에 대하여 상부(5′), 중간, 하부(3′)에 위치되는 천연의 뉴클레오티드 서열 또는 키메라 유전자를 말하며, 이는 본 발명의 핵산 단편의 발현을 조절한다. 본 명세서에 사용된 '발현'이라는 용어는 세포의 단백질 조직과 함께 결합하여 표현 형질을 변화시키는, 본 발명의 핵산 단편으로부터 유도된 센스(mRNA) 또는 안티센스 RNA의 안정한 축적 및 전사를 말한다. 유전자의 발현은 유전자의 전사 및 mRNA의 전구체 또는 성숙 단백질로의 번역을 포함한다. '안티센스 억제'는 목적 단백질의 발현을 억제할 수 있는 안티센스 RNA 전사체의 생성을 말한다. '과다발현'은 정상 또는 비-형질전환된 유기체 중에 생산 수준을 초과하는 형질전환 유기체내 유전자 생성물의 생산을 말한다. '공동억제'는 내생 유전자와 실질적인 상동성을 갖는 외래 유전자의 발현으로, 외래 및 내생 유전자 모두의 발현을 억제하는 것을 말한다. '변화 수준'이라는 말은 형질전환된 유기체 내에 정상 또는 비-형질전환된 유기체와는 다른 양이나 비율로 유전자 생성물을 생산하는 것을 말한다. 당업자들은 본 발명에 의해서 추측되는 표현형 효과는 정상 또는 비-형질전환된 유기체와 비교하여 형질전환 유기체 내에 생산되는 유전자 생성물의 수준 변화, 즉, 안티센스 억제 또는 공동억제에 의해서 전달되는 유전자 발현의 감소에 의해서 달성될 수 있음을 알 수 있을 것이다.

'프로모터'는 유전자내 DNA 서열, 통상적으로 코딩 서열의 상부(5′)를 말하고, 이는 RNA 중합효소및 적당한 전사에 요구되는 다른 인자를 인식함으로써 코딩 서열의 발현을 조절한다. 또한, 프로모터는 생리학적 또는 발생적 조건에 반응하여 전사 개시의 효율성을 조절하는 단백질 인자 결합에 관련되는 DNA 서열을 포함할 수 있다. 또한, 인헨서(enhancer) 성분을 포함할 수도 있다. '인헨서'는 프로모터 활성을 촉진시킬 수 있는 DNA 서열이다. 이는 프로모터의 고유한 성분이거나, 프로모터의 수준 또는 조직-특이성을 증가시키기 위해 삽입되는 이종의 성분일 수 있다. '구성(constitutive) 프로모터'는 모든 조직 및 모든 시간에 유전자 발현을 지시하는 것을 말한다. 본 명세서에 언급한 바와 같은 '조직-특이적' 또는 '발생-특이적' 프로모터는 잎 또는 종자와 같은 특정 조직 내에 또는 초기 또는 말기 배형성과 같은 조직내 특정 발생 단계에서 거의 독점적으로 유전자 발현을 나타내는 것이다.

'3′ 비-코딩 서열'은 mRNA 프로세싱 또는 유전자 발현에 영향을 줄 수 있는 폴리아데닐화 신호 및 임의의 다른 조절성 신호를 함유하는 유전자의 DNA 서열 부분을 말한다. 폴리아데닐화 신호는 통상적으로, mRNA 전구체의 3'말단에 폴리아데닐산 트랙(tract)의 첨가에 영향을 주는 것을 특징으로 한다.

'작동가능하게 결합된'이라는 말은 하나의 기능이 나머지 하나에 의해서 영향을 받도록 결합되어 있는 단일 핵산 분자상의 핵산 서열을 말한다. 예를 들어, 프로모터는 임의의 구조적 유전자와 작동가능하게 결합하여 구조적 유전자의 발현에 영향을 줄 수 있다(즉, 구조적 유전자는 프로모터의 전사 조절하에 있다).

'형질전환'은 핵산 단편을 숙주 유기체의 게놈 내에 전달하여 유전적으로 안정한 고유 특성을 나타내는 것을 말한다. 형질전환시킨 핵산 단편을 함유하는 숙주 유기체는 '형질전환' 유기체라고 부른다.

본 발명은 형질전환 대두 계통의 제작에 관한 것이고, 여기서 종자 저장 단백질을 코딩하는 유전자의 발현을 조절하여 형질전환 식물의 종자 저장 단백질 프로파일을 변화시킨다. 종자 저장 단백질 프로파일의 변화는 독특하고 유용한 기능성을 갖는 신규 대두 단백질 제품을 생산할 수 있다.

식물내 유전자 발현은 그러한 식물내에서 기능적인 조절 서열을 사용한다. 식물 내에서 외래 유전자의 발현은 잘 알려져 있다(드 블라에르(De Blaere) 등의 문헌[Meth. Enzymol. 153:277-291(1987)]). 본 발명 단편의 발현을 유도하기 위해서 선택된 프로모터 공급원은 목적 종자 저장 단백질 유전자의 발현을 감소시킴으로써 본 발명을 달성하는 충분한 전사 활성을 갖는 한, 중요하지 않다. 바람직한 프로모터는 콜리플라워 모자이크 바이러스(cauliflower mosaic virus)내 19S 및 35S 전사체를 나타내는 것과 같은, 강한 구성 식물 프로모터이다(오델. 제이. 티.(Odell, J. T.) 등의 문헌[Nature 313:810-812(1985)]; 헐(Hull) 등의 문헌[Virology 86:482-493)]. 특히 바람직한 프로모터는 종자-특이적 발현을 가능하게 하는 것들이다. 종자-특이적 프로모터의 예로는, 이에 제한되는 것을 아니지만, 많은 식물내 총 종자 단백질이 90 % 이하인 종자 저장 단백질의 프로모터가 있다. 종자 저장 단백질은 엄격하게 조절되고 높은 조직-특이적 및 단계-특이적 방식으로 종자 내에서 거의 독점적으로 발현된다(히긴스 등의 문헌[Ann. Rev. Plant Physiol. 35: 191-221(1984)]; 골드버그(Goldberg) 등의 문헌[Cell 56: 149-160 (1989)]). 더우기, 다른 종자 저장 단백질은 종자 발달의 다른 단계에서 발현될 수 있다.

종자-특이적 유전자의 발현이 보다 세부적으로 연구되었다(골드버그 등의 문헌[Cell 56: 149-160 (1989)] 및 히긴스 등의 문헌[Ann. Rev. Plant Physiol. 35: 191-221(1984)]에 의한 검토 참고). 현재, 형질전환 쌍떡잎 식물에는 종자 저장 단백질 유전자(천연 또는 키메라 유전자)의 종자-특이적 발현의 수많은 예가 있고, 통상적으로 시간적 및 공간적 발현 패턴이 유지된다. 이러한 예에서 사용되는 프로모터를 사용하여 본 발명에 영향을 줄 수도 있다. 이들은 콩 β-파세올린(센굽타-고팔란(Sengupta-Gopalan) 등의 문헌[Proc. Natl. Acad. Sci. USA 82:3320-3324(1986)]); 호프만(Hoffman) 등의 문헌[Plant Mol. Biol. 11:717-729(1988)]), 콩 렉틴(보엘커(Voelker) 등의 문헌[EMBO J. 6:3571-3577(1987)]), 대두 렉틴(오까무로(Okamuro) 등의 문헌[Proc. Natl. Acad. Sci. USA 83:8240-8244(1986)]), 대두 쿤니츠(Kunitz) 트립신 억제제(페레즈-그로(Perez-Grau) 등의 문헌[Plant Cell 1:095-1109(1989)]), 대두 β-콘글리시닌 (비카이(Beachy) 등의 문헌[EMBO J. 4:3047-3053(1985)]), 완두 빅실린(히긴스 등의 문헌[Plant Mol. Biol. 11:683-695(1988)], 완두 컨빅실린(뉴비긴(Newbigin) 등의 문헌[Planta 180: 461-470(1990)], 완두 레구민(쉬르샛(Shirsat) 등의 문헌[Mol. Gen. Genetics 215:326-331(1989)]), 평지씨 나핀(래드케(Radke) 등의 문헌[Theor. Appl. Genet. 75:685-694(1988)]) 및 아라비돕시스 탈리아나 2S 알부민(반데커크호브(Vandekerckhove) 등의 문헌[Bio/Technology 7:929-932(1989)])의 쌍떡잎 식물 유래의 유전자이다

본 발명의 핵산 단편의 발현에서 특히 유용한 것은 쿤니츠 트립신 억제제(KTi; 조푸쿠(Jofuku) 등의 문헌[Plant Cell 1: 1079-1093(1989)]), 글리신(닐슨(Nielson) 등의 문헌[Plant Cell 1:313-328(1989)]) 및 β-콘글리시닌 (하라다(Harada) 등의 문헌[Plant Cell 1: 415-425(1989)])과 같은 몇몇 대두 종자 저장 단백질 유전자 유래의 이종 프로모터일 것이다. 당업자는 다른 종자 프로모터에 의해 부여되는 시간적 조절의 차이에 주목해야 한다는 것을 인식할 수 있을 것이다. 예를 들어, α-서브유닛 유전자에 대한 프로모터는 β-서브유닛 유전자에 대한 프로모터보다 몇일 전에 발현되기 때문에(비카이 등의 문헌[EMBO J. 4:3047-3053(1985)]), α-서브유닛 발현을 억제하기 위해서 β-서브유닛 유전자를 사용하는 것은 보다 덜 유용할 것이다.

또한, 지질 또는 탄수화물 생합성과 같은 종자 대사의 다른 측면과 관련된 유전자의 프로모터도 사용할 수 있지만 덜 바람직할 것이다. 요약하면, 당업자는 충분한 강도 및 적당한 시간적 발현 패턴을 갖는 임의의 프로모터를 사용하여 본 발명을 달성할 수 있음을 인식하는데 어려움이 없을 것이다. 유사하게, 인헨서 또는 인헨서-유사 성분을 본 발명의 천연 또는 키메라 핵산 단편의 프로모터 부위에 도입함으로써 보다 발현이 증가되므로 본 발명을 달성할 수 있을 것이다. 이는 35S 프로모터에서 발견되는 것과 같은 바이러스성 인헨서(오델 등의 문헌[Plant Mol. Biol. 10:263-272]), 오핀(opine) 유전자 유래의 인헨서(프롬(Fromm) 등의 문헌[Plant Cell 1:977-984]), 또는 임의의 다른 공급원 유래의 인헨서를 포함하며, 본 발명의 핵산 단편에 작동가능하게 결합된 프로모터 내에 위치시켰을 때 전사를 증가시킨다.

구성 프로모터에 종자 특이성을 40 배 증가시킬 수 있는 β-콘글리시닌의 α-서브유닛을 코딩하는 유전자로부터 단리된 DNA서열 성분이 특히 중요하다(첸(Chen) 등의 문헌[Dev. Genet. 10:112-122(1989)]). 형질전환 식물 중의 프로모터와 함께 종자-특이성이 증가된 발현을 달성하기 위해서, 당업자는 이러한 성분을 쉽게 분리하고 그것을 임의의 유전자 프로모터 부위 범위 내에 삽입할 수 있다. 통상적으로, β-콘글리시닌 유전자와는 다른 시간에 발현되는 임의의 종자-특이적 유전자 내에 이러한 성분의 삽입 결과, 종자가 발달하는 동안 보다 장시간 형질전환 식물에 이러한 유전자를 발현시킬 것이다.

본 발명 핵산 단편의 적당한 발현을 위해서 요구되는 폴리아데닐화 신호 및 다른 조절 서열을 제공할 수 있는 임의의 3′비-코딩 부위를 사용하여 본 발명을 달성할 수 있다. 이는 천연 지방산 데세츄라제의 3′말단, 35S 또는 19S 콜리플라워 모자이크 바이러스 전사와 같은 바이러스 유전자, 오핀 합성 유전자, 리불로스 1,5-비스포스페이트 카르복실레이트 또는 클로로필 a/b 결합 단백질을 포함한다. 당업계에는 다른 3'비-코딩 부위의 유용성을 교시하는 수많은 예가 있다.

당업자는 본 발명에 따른 고등 식물 세포를 형질전환시키는 여러 가지 방법을 이용할 수 있다(유럽 특허 공개 EP-A-295,959 및 EP-A-318,341). 이러한 방법들은 아그로박테리움종(Agrobacterium spp.)의 Ti 및 Ri 플라스미드를 이용하는 형질전환 벡터를 기재로 하는 것이다. 이러한 2 성분형(binary type)의 벡터를 사용하는 것이 특히 바람직하다. Ti-유도된 벡터는 떡잎 및 쌍떡잎 식물을 포함하여 넓은 범위의 고등 식물을 형질전환시킨다(숙하핀다(Sukhapinda) 등의 문헌[Plant Mol. Biol. 8:209-216(1987)]; 포트리쿠스(Potrykus) 등의 문헌[Mol. Gen. Genet. 199:183(1988)]). 당업자들은 외래 DNA 작제물의 직접 도입(유럽 특허 공개 EP-A-295,959 참조), 일렉트로포레이션(electroporation) 기술(프롬 등의 문헌[Nature(London) 319:791(1986)]) 또는 핵산 작제물로 코팅된 금속 입자를 갖는 고속 탄도 포격(클레인(Klein) 등의 문헌[Nature(London) 327:70(1987)])과 같은 다른 형질전환 방법도 이용할 수 있다. 일단 형질전환되면 당업자에 의해서 세포를 재생시킬 수 있다. 맥카브(MaCabe) 등의 문헌[Bio/Technology 6:923-926(1988), 피너(Finer)등의 문헌[In Vitro Cell. Dev. Biol. 27: 175-182)(1991)] 및 힌치, 엠. 에이. 더블유(Hinchee, M.A.W.)의 문헌[Bio/Technology 6:915-922(1988)]에는 특히, 최근에 설명된 대두 형질전환 방법이 기재되어 있다.

일단 형질전환 식물이 상기 기재된 방법 중 하나에 의해서 얻어지면, 바람직한 표현형을 가장 효과적으로 나타내는 개개의 형질전환체를 스크리닝하는 것이 필요하다. 동일한 작제물을 갖는 개개의 형질전환 식물은 발현 수준이 다를 수 있다는 것은 당업자들에게 널리 알려져 있으며, 이러한 현상은 통상적으로 '위치 효과(position effect)'라고 불린다. 따라서, 본 발명에 있어서, 다른 개개의 형질전환체는 목적 종자 단백질의 억제 효율성에 있어서 다양할 수 있다. 당업자는 특정 유전자의 발현을 감소시키기 위해서 안티센스 또는 공동억제 기술의 사용과 관련되는 특별한 고려 사항을 잘 알고 있을 것이다. 미국 특허 제5,190,931호, 제5,107,065호 및 제5,283,323호에는 이러한 기술이 실시가능하지만, 이들의 효율성은 예상할 수 없다고 공지되어 있다. 따라서, 당업계는 특정 유전자에 대하여 가장 효과적이라고 추측되는 방법을 교시하고 있지 않으므로, 당업자는 억제시켜야 할 유전자 중 하나 이상의 다른 부분을 함유하는 복수 유전자 작제물을 제작할 것이다. 또한, 가장 효과적인 작제물도 단리된 개개의 형질전환 계통 중 단지 일부에서 효과적인 억제 표현형을 얻을 것이다. 예를 들어, 국제 특허 공개 WO 93/11245 및 WO 94/11516에는 카놀라 내에 지방산 데세츄라제 유전자의 발현 억제를 시도할 때, 실제적인 억제는 시험한 계통의 1% 미만이 얻어졌다고 교시하고 있다. 다른 종류에서의 백분율은 다소 높지만, 100%인 경우는 없다. 이를 본 발명상의 한계로 보아서는 안되고, 대신 실제적인 문제로서 당업자에게 인식되고 기대되어야 한다. 따라서, 기술자들은 수많은 형질전환체를 스크리닝하는 방법을 발전시킬 것이다. 통상적으로, 이들 특징적인 스크린(screen)은 실제적인 근거에서 선택될 것이고, 본 발명의 본질적인 부분은 아니다. 다수의 시료가 무효할 것으로 예상되기 때문에, 바람직한 방법은 많은 수의 시료를 쉽게 처리할 수 있는 방법일 것이다.

공동 억제 기작은 불분명하고(임의의 검토 및 고찰을 위해서 플라벨, 알.(Flavell, R)의 문헌[Proc. Natl. Acad. Sci. USA 91:3490-3496(1994)] 참고), 따라서, 필요할 때 이를 유도해내기 위한 필요 조건 또한 불분명하다. 문헌에서 발견되는 대부분의 예는 공동 억제시켜 바람직한 반응을 유도해내는 유전자의 전사된 부위의 전부 또는 큰 부분을 사용하는 것이다. 그러나, 하나 이상의 경우(브루슬란(Brusslan) 등의 문헌[Plant Cell 5:667-677(1993)]; 브루슬란 및 토빈(Tobin)의 문헌[Plant Mol. Biol. 27:809-813(1995)])에서, 아라비돕시스의 cab140 유전자의 사용, 프로모터(1.3 kb 단편)의 사용 및 전혀 관련없는 유전자에 융합된 전사 부위 중 단지 14 bp가 내생 cab140 유전자 뿐 아니라 도입된 키메라 유전자를 공동억제시키기에 충분하였다. 이 결과는 수많은 다른 프로모터-리더(5′비번역된 리더는 전사의 시작과 번역 개시 코돈 사이의 부위로서 정의됨) 단위가 키메라 유전자를 성공적으로 유도하기 위하여 사용되어 왔기 때문에 독특하고 뚜렷하게 예상할 수는 없다. 플라벨은 몇몇 또는 많은 유전자(종자 단백질을 코딩하는 유전자와 같은 복수 유전자군을 포함함)가 공동억제 기작을 피하도록 진화될 수 있지만, 다른 유전자는 그렇지 않으며, 게놈이 진화함에 따라서 조절 수준을 한층 더 높일 수 있을 것이라고 서술한다. 따라서, 본 발명에 있어서, 콘글리시닌 유전자의 프로모터 및 리더를 사용하여 내생 콘글리신의 발현을 억제할 수 있지만, 변형유전자(transgene)의 다른 부분(개시 코돈 이상)을 사용하여 관련없는 유전자를 완전히 억제할 수 있다는 것이 특이하다.

< 실시예

본 발명은 하기 실시예에 의해서 추가 정의된다. 실시예는 단지 설명을 위해서 주어진 것이며, 본 발명을 실시예에 기재된 용도에 제한하지 않는 것으로 이해해야 할 것이다. 본 발명은 다양한 종자 저장 단백질의 변화된 수준을 갖는 형질전환 대두 식물을 생산하는데 사용될 수 있다. 상기 논의 및 하기 실시예로부터, 당업자들은 본 발명의 범위와 취지에서 벗어나지 않고 확실하게 실시할 수 있고, 본 발명의 다양한 변화 및 변형을 여러 용도 및 조건에 적용시킬 수 있다. 이러한 모든 변형은 청구항의 범위 내에 속하도록 실시된다.

제한 효소, 다른 변형 효소, 아가로스 겔 전기영동, 핵산 혼성화와 같은 DNA 조작의 상세한 방법 및 플라스미드 DNA로써 E. coli의 형질전환이 삼브룩(Sambrook) 등의 문헌[Molecular Cloning, A Laboratory manual, 2nd ed, Cold Spring Harbor Laboratory Press(이 후 'Maniatis')]에 기재되어 있다. 모든 제한 효소 및 다른 변형 효소는 지브코(Gibco) BRL (미국 매릴랜드주 개더스버그 소재)로부터 입수하였다.

< 실시예1

발아하는 대두 떡잎내 β-콘글리시닌의 발현이 공동억제의 표적일 수 있는지를 결정하기 위해서, 중합효소 사슬 반응 키트(GeneampTM RNA PCR Kit; 퍼킨 엘머 세투스(Perkin Elmer Cetus))에서 역전사효소를 사용하여 β-콘글리시닌의 α 및 α′서브유닛의 절단된 cDNA 단편을 제조하였다. 상부 프라이머, ConS는 EMBL/Genbank/DDBJ 데이터베이스로부터 얻어진 α 및 α′서브유닛 cDNA 서열의 뉴클레오티드 5-19와 상동이다. 클로닝을 돕기 위해서 부가적 뉴클레오티드를 5′말단에 가하여 Nco I 제한 위치를 코딩하였다. 하부 프라이머, Con 1.4a는 α 및 α′서브유닛 cDNA의 각 서열을 나타내는 서열 번호:1의 뉴클레오티드 1370-1354 및 서열 번호:2의 뉴클레오티드 1472-1456에 상보적이었다. 클로닝을 돕기 위해서, 부가적 뉴클레오티드를 5'말단에 가하여 Kpn I 제한 위치를 코딩하였다. PCR 프라이머 ConS 및 Con 1.4a의 뉴클레오티드 서열을 하기에 나타내었다.

ConS 5'-CGTACCATGGTGAGAGCGCGGTTCC-3' (서열 번호: 4)
Nco I
Con1.4a 5'-CGGTACCGAATTGAAGTGTGGTAG-3' (서열 번호: 5)
Kpn I

제조업체의 프로토콜에 따라서 키트가 공급된 랜덤 육량체 또는 Con1.4a 중 하나를 사용하여, 발아하는 대두 종자로부터 분리된 RNA를 역전사시켰다. 생성되는 cDNA 단편을 ConS 및 Con1.4a의 혼합물을 사용하는 PCR(중합효소 사슬 반응) 반응에서 증폭시켰다. 반응물 농도는 제조업체의 프로토콜에 기재된 대로 하였다. 하기 프로그램이 사용되었다: a) 95 ℃에서 2 분, 1 사이클; b) 50 ℃에서 1.5 분(어닐링), 70 ℃에서 5 분(확장), 95 ℃에서 1.5 분(변성), 35 사이클; 및 c) 50 ℃에서 2 분, 이어 68 ℃에서 10 분, 1 사이클. 각 PCR 반응 혼합물 15 마이크로리터를 아가로스 겔 전기영동에 의해서 분석하였다. 반응은 예상된 크기의 PCR 생성물, 즉, α ′ 에 대하여 1.47 kb 및 α에 대하여 1.37 kb를 얻었다. 반응 혼합물 잔류의 절단된 cDNA 단편을 위자아드(Wizard)TM PCR 프렙 DNA 정제 시스템 키트 (프로메가(Promega)사제) 사용하여 정제하였다.

사용되는 프라이머로 인하여 5 ′ 말단에서 Nco I 제한 위치 및 3 ′ 말단에서 Kpn I 제한 위치를 포함하는, α 및 α ′ 단편을 함유하는 정제된 반응 혼합물을 Kpn I 및 Nco I 제한 효소로 분해하였다. 겔 전기영동을 수행하여 α cDNA 단편을 회수하고, 단편 F8로 명명하였으며, 이를 pCW109(도 1) 및 pML70(도 2) 중에 두 플라스미드 내에 존재하는 Nco I 및 Kpn I 위치를 사용하여 방향적으로 클로닝(센스 배향)하였다. ConS 및 Con1.4a내 프라이머(Con.09 및 Con.8)의 한벌 세트를 사용하는 PCR에 의해서 F8은 α로서 확인되었고, Hind III, Nco I, Kpn I 및 Pst I을 갖는 pCW 109/F8 플라스미드를 분해시킴으로써 α ′ 와 구별하였다(α는 Pst I 위치를 포함하지 않는 반면, Pst I 위치는 α ′ 포함함).

Con.09 5'-TCGTCCATGGAGCGCGGTTCCCATTAC-3' (서열 번호: 6)
Con.8 5'-TCTCGGTCGTCGTTGTT-3' (서열 번호: 7)

전사 단위 KTi 프로모터/절단된 α/KTi 3 ′ 말단을 BamHI로 제한 분해시킴으로써 플라스미드 pML70/F8로부터 유리시키고, 겔 분리하여 F11로 표지하였다. 그 후, F11을 BamH I 위치에서 제한 pKS18HH(도 3) 중으로 클로닝하였다. pKS18HH는 하기 유전 요소 : (i) T7 프로모터/하이그로마이신 B 포스포트랜스퍼라제 (Hygromycin B phosphotransferase)(HPT)/T7 터미네이터 서열, (ii) 콜리플라워 모자이크 바이러스(CaMV)/하이그로마이신 B 포스포트랜스퍼라제 (HPT)/아그로박테리움 투메파시엔스(tumefaciens) T - DNA로부터의 노팔린 신타제(NOS) 3 ′ 으로부터의 35S 프로모터; 및 (iii) 제거된 베타 - 락타마제 코딩 부위를 갖는 pSP72 플라스미드 벡터(프로메가)를 함유하는 플라스미드 작제물이다. 분자 생물학 분야의 숙련자는 널리 공지된 프로토콜(Manistis)을 사용하여 상기 세 성분을 하나의 플라스미드 벡터 중에 결합시킬 수 있다.

클레브시엘라(Klebsiella) - 유도 플라스미드 pJR225를 함유하는 E. coli 균주 W677로부터 PCR 증폭에 의해서 하이그로마이신 B 포스포트랜스퍼라제 (HPT) 유전자를 단리하였다(그리츠 엘.(Gritz L.), 및 다비즈 제이.(Davies J.)의 문헌[Gene 25:179 - 188(1983)]). pKS18HH는 대두와 같은 식물 내에 HPT 효소의 성분 발현을 위한 CaMV 35S/HPT/NOS 카세트를 함유한다. 또한, pKS18HH 플라스미드는 특정 E. coli 균주내 HPT 효소의 발현을 위한 T7 프로모터/HPT/T7 터미네이터 카세트, 예를 들어, 노바블루(NovaBlue)TM(DE3)(노바겐(Novagen))를 함유하며, 이는 (lacUV5 콘트롤 하에 T7 RNA 중합효소 유전자를 갖는) 람다 DE3에 대하여 용원성이다. pKS18HH는 또한, 유전자를 이러한 벡터 중에 클로닝하기에 적합한 특이 제한 효소의 3 위치를 포함한다. 따라서, pKS18HH 플라스미드 벡터는 E. coli 및 식물 모두에서 선택을 위해서 하이그로마이신 B의 사용을 가능하게 한다. Hind III로 분해함으로써 F11 단편의 삽입 및 배향 확인을 달성하였다. 시계방향의 배향으로 F11 단편을 갖는 클론을 선택하고 pJo1로 표지하였다(도 4).

체세포 배 배양균의 형질전환

대두 체세포 배의 형질전환 및 증식을 위하여 하기 스톡 용액 및 배지를 사용하였다.

스톡 용액			배지
MS 술페이트 100x 스톡		(g/L)	SB55(리터 당)
$MgSO_4$	$7H_2O$	37.0	각 MS 스톡 10 mL
$MgSO_4$	H_2O	1.69	B5 비타민 스톡 1 mL
$ZnSO_4$	$7H_2O$	0.86	NH_4NO_3 0.8 g
$CuSO_4$	$5H_2O$	0.0025	KNO_3 3.033 g, 2,4-D (10 mg/mL 스톡) 1 mL
MS 할라이드 100x 스톡			
$CaCl_2$	$2H_2O$	44.0	아스파라긴 0.667 g
KI		0.083	pH 5.7
$CoCl_2$	$6H_2O$	0.00125	
KH_2PO_4		17.0	SB103(리터 당)
H_3BO_3		0.62	무라쉬게 & 스쿠그(Muashige & Skoog) 염 혼합물(기브코 BRL) 1 pk.
Na_2MoO_4	$2H_2O$	0.025	말토스 60 g
Na_2EDTA		3.724	겔리트(gelite) 2 g
$FeSO_4$	$7H_2O$	2.784	pH 5.7 (SB53 플러스 목탄에 대하여, 목탄 5 g 첨가)
B5 비타민 스톡			
미오-이노시톨		100.0	무라쉬게 & 스쿠그 염 혼합물(기브코 BRL) 1 pk.
니코틴산		1.0	B5 비타민 스톡 1 mL
피리독신 HCl		1.0	아가로스 7 g
티아민		10.0	pH 5.7

형광 및 백열광을 동시에 가하면서 28 ℃의 회전 진동기상 액상 매질 (SB55) 35 mL중에서 대두 배 현탁 배양액을 유지하고 16/8 시 낮/밤 스케줄을 제공하였다. 약 35 mg의 조직을 액상 배지 35 mL 중에 접종함으로써 배양균을 매 2 내지 3 주 마다 2차 배양하였다.

입자 발사기에 의한 포격(particle gun barmbardment) 방법에 의해서 대두 배 현탁 배양균을 pJol로 형질전환시켰다(클레인 등의 문헌[Nature 327:70(1987)] 참조). 이러한 형질전환을 위해서 듀폰 바이오리드틱(DuPont Biolistic)™ PDS1000/He 기구를 사용하였다.

pJol 플라스미드 DNA(1 ㎍/㎕) 5 ㎕, $CaCl_2$ (2.5 M) 50 ㎕ 및 스퍼미딘 (0.1 M) 20 ㎕을 60 mg/mL 1 mm 금 입자 현탁액 50 ㎕에 가하였다. 입제를 3 분 동안 젓고, 10 초 동안 마이크로퓨즈에서 회전시키고 상징액을 제거하였다. 그 후, DNA-코팅된 입자를 일단 70 % 에탄올 400 ㎕로 세척하고, 무수 에탄올 40 ㎕ 중에 현탁시켰다. DNA/입자 현탁액에 1 초 동안의 초음파 처리를 3 회 실시하였다. DNA-코팅된 금 입자 5 ㎕를 각각 매크로 캐리어 디스크(macro carrier disk)에 적하하였다.

2 주 묵힌 현탁 배양액 약 300 내지 400 mg를 60 mm x 15 mm 빈 페트리 디쉬에 넣고, 피펫으로 조직로부터 잔류 액체를 제거하였다. 보유 스크린으로부터 약 3.5 인치 떨어진 곳에 조직을 놓고 2 회 포격하였다. 멤브레인 파열 압력을 1000 psi에 고정시키고, 챔버를 비워 Hg -28 인치가 되게 하였다. 두 개의 플레이트를 각 실험 작제물 마다 포격하였다. 포격 후, 조직은 반으로 나뉘어졌고 다시 액상 배지 중에 넣어 상기 설명한 대로 배양하였다.

포격 15 일 후, 액상 배지를 하이그로마이신 50 mg/mL을 함유하는 신선한 SB55로 교환하였다. 선택 배지를 주마다 신선하게 갈았다. 포격 6 주 후, 녹색의 형질전환된 조직을 단리하고, 플라스크 중에 접종하여 새로운 형질전환된 배 현탁 배양균을 생산하였다.

형질전환된 배 클러스터를 액상 배양 배지로부터 옮겨 고체 아가 배지, SB103 + 0.5 % 목탄상에 놓고 성숙시키기 시작하였다. 1 주 후, 배를 SB103 배지 - 목탄으로 옮겼다. SB103 상에서 3 주 후, 성숙중인 배를 분리하여 SB148 배지 위에 놓았다. 배를 성숙시키는 동안의 조건은 26 ℃, 형광 및 백열광을 동시에 가하고 16/8 시 낮/밤 스케줄을 제공하였다. SB148 배지상에서 6 주 후, β-콘글리시닌 서브유닛 단백질의 발현에 대하여 배를 분석하였다. 각 배 클러스터는 5 내지 20 개의 체세포 배를 발생하였다.

형질전환된 체세포 배의 분석

체세포 배가 성숙하는 동안 β-콘글리시닌의 α, α' 및 β 서브유닛을 시각화할 수 있는 때를 결정하기 위해서 SDS-PAGE 겔 전기영동에 의해서 초기 실험을 수행하였다. 형질전환되지 않은 배의 떡잎(포격을 겪지 않았다는 것을 제외하고는 상기와 같이 발아한 것임)을 성숙하기 시작한 후 6, 8, 10 및 12 주에 배로부터 잘라내어 분석할 때까지 -80 ℃에서 냉동보관하였다. 떡잎 조직을 칭량하고, 추출 완충액 10 ㎕/조직 mg을 가하고, 조직을 펠렛 페스틀 디스포저블 믹서(Pellet Pestle Disposable Mixer(Kimble/Kontes))에서 분쇄하였다. 추출 완충액은 50 mM Tris-HCl(pH 7.5), 10 mM β-메르캅토에탄올(BME) 및 0.1 % SDS로 이루어졌다. 그 후, 시료를 12,000 rpm에서 10 분 동안 마이크로퓨징시키고, 상등액을 새로운 마이크로퓨즈 튜브로 옮겼다. 추출물을 사용할 때까지 -20 ℃에서 냉동보관하였다.

SDS-PAGE 분석 동안, 로딩(loading) 완충제(2x) 8 ㎕를 시료 추출액 8 ㎕에 가하였다. 로딩 완충제(2x)은 100 mM Tris-HCl(pH 7.5), 4 % SDS 및 0.2 % 브로모페놀 블루, 15 % 글리세롤 및 200 mM βME로 이루어졌다. 혼합물을 95 ℃에서 4 분 동안 가열하였다. 그 후, 시료 혼합물을 마이크로퓨즈(12,000 rpm에서 20 초 동안)시키고, 미니-프로테인 II 일렉트로포레시스 셀(Mini-Protein II Electrophoresis Cell)(바이오-라드(Bio-Rad))로 조립된, 10 %의 미리 제조한 레디 겔 (Ready Gel)TM (바이오-라드)상에 로딩하였다. 바이오-라드 Tris/글리신/SDS 완충액을 유동 완충액으로서 사용하였고, 전압은 125 V로 일정하였다. 시료 추출물 이외에, 각 겔은 표준 분자량(바이오-라드 SDS-PAGE 표준, 좁은 범위)를 갖는 하나의 레인 및 시판 탈지 대두 분말로부터 추출된 총 대두 종자 단백질을 갖는 하나의 레인을 포함하였다. 완료시, 겔을 쿠마씨 브릴리언트 블루(Coomassie Brilliant Blue)로 염색하고 단백질을 시각화하기 위해서 탈색시켰다(Maniatis). 겔을 사진 촬영하고, 물이 들어 있는 밀폐된 봉투에 넣어 냉장고에 보관하였다. 결과, 성숙 개시 후 8 내지 10 주 사이에 체세포 배의 떡잎에서 β-콘글리시닌의 α, α' 및 β 서브유닛을 검출할 수 있었다.

상기 설명한 방법을 사용하여 성숙 개시 10 주 후에 형질전환된 배 발현 분석을 수행하였다. 처음에, 클론 당 두 개의 배를 분석하였다. 두 개의 배에서 β-콘글리시닌 서브유닛의 억제가 관찰된 경우, 추가의 배를 분석하였다. 이러한 분석의 결과를 표 1에 나타내었고, 여기서 각 β-콘글리시닌 서브유닛의 존재 또는 부재는 각각 (+) 또는 (-)로 표시하였다.

클론	배	α	α'	β
Jol-1	1	-	-	+
	2	-	-	+
	3	+	+	+
	4	-	-	+
	5	+	+	-
Jol-2	1	+	+	+
	2	+	+	+
Jol-3	1	+	+	+
	2	+	+	-
Jol-4	1	-	-	-
	2	-	-	-
	3	-	-	-
	4	+	+	+
	5	-	-	-
Jol-5b	1	+	+	+
	2	+	+	+
Jol-5c	1	-	-	+
	2	-	-	+
Jol-5d	1	+	+	+
	2	+	+	+
Jol-6a	1	-	-	+
	2	-	-	+
	3	-	-	+
	4	-	-	+
	5	+	+	+
Jol-6b	1	+	+	+
	2	+	+	+
Jol-6c	1	+	+	+
Jol-6d	1	+	+	+
	2	+	+	+
Jol-6d	1	+	+	+
	2	+	+	+
Jol-6e	1	+	+	-
	2	+	+	+
Jol-7a	1	-	-	-
	2	+	+	+
Jol-7b	1	-	-	-
Jol-8a	1	+	+	+
Jol-8b	1	+	+	+
	2	-	-	-
Jol-9a	1	-	-	-
	2	-	-	-
Jol-9b	1	-	-	-
	2	-	-	-
Jol-9c	1	-	-	-
Jol-10	1	-	-	-
	2	-	-	-

7 개의 유전자전환 클론이 α 및 α ′ 발현이 억제된 배를 발생하였다. 또한, 하나의 클론(Jol-4)이 3 개의 β-콘글리

시닌 서브유닛 모두가 억제된 배를 발생하였다. 절단된 α변형 유전자 서열이 총 1.32 kb β 서브유닛 cDNA 중 단지 0.75 kb 부분과 겹친다는 결과가 놀라웠다. 따라서, 절단된 α변형 유전자 및 β 서브유닛 cDNA 사이에 단지 52 %의 유사성이 있었다. 이러한 정보로써, 절단된 α변형 유전자가 β-콘글리시닌 유전자의 β 서브유닛을 '공동억제'하는 구조와 충분한 유사성을 갖지 않는 것으로 생각될 것이다.

SDS-PAGE의 실시예를 도 5에 나타내었다. 레인 1 내지 3은 클론 pJol-1로부터 생성된 배로부터 절단된 3 개의 떡잎 추출물이었다. 레인 4 및 5는 각각 단백질 분자량 표준 및 종자로부터 유래된 대두 단백질 표준이었다. 레인 6 및 7은 클론 pJol-4로부터 생성된 배로부터 절단된 떡잎 추출물이었다. 레인 2의 단백질 패턴은 α 및 α′ 모두 공동-억제된 배의 예이다. 레인 6 및 8의 단백질 패턴은 β-콘글리시닌을 이루는 모든 서브유닛이 억제되는 배의 예이다.

< 실시예 2

β-콘글리시닌 프로모터 부위를 사용하여 공동억제함으로써 발아하는 떡잎에서 β-콘글리시닌의 발현을 억제시킬 수 있는지 결정하기 위해서, 대두 β-콘글리시닌 프로모터의 조절하에(비카이 등의 문헌[EMBO J. 4:3047-3053(1985)]) 글리신 맥스 미크로조말 델타-12 데세츄라제 cDNA(GmFad 2-1) 서열을 함유하는 pBS43으로 명명된 플라스미드(헤퍼드(Heppard) 등의 문헌[Plant Physiol 110:311-319 (1996)]; 유전자 은행 승인 번호 L43920)를 제작하였다. 이 벡터의 제작은 pMH40, pCST2 및 pBS13 플라스미드를 사용함으로써 용이하였다. 하기 상술된 플라스미드 작제물은 미국 특허 출원 번호 USSN 08/262,401 및 국제 특허 공개 번호 WO94/11516의 일부에 기재되어 있고, 둘 모두 본 명세서에 참고문헌으로 인용되어 있다.

pMH40 벡터는 시판 클로닝 벡터(프로메가 바이오테크사) 플라스미드 pGEM9z를 E. coli의 β-글루쿠로니다제 유전자에 결합된 CaMV(오델 등의 문헌[Nature 303:810-812(1985)]; 하프스터(Harpster) 등의 문헌[Mol. Gen. Genet. 212:182-190(1988)])로부터의 1.4 kb 35S 프로모터 부위에 삽입함으로써 유도되었다. 이것은 효소 β-글루쿠로니다제를 코딩하는 1.85 kb 단편(제퍼슨(Jefferson) 등의 문헌[PNAS USA 80:4803-4807(1983)]) 및 아가로박테리움 투메파시엔스의 Ti-플라스미드의 노팔린 신타제 유전자로부터의 전사 터미네이터를 함유하는 0.3 kb DNA 단편(프랄리(Fraley) 등의 문헌[PNAS USA 80:4803-4807(1983)])이었다.

벡터 pCST2는 벡터 pML18 및 pCW109A로부터 유도되었다. 플라스미드 pCW109A는 대두 β-콘글리시닌 프로모터 서열 및 파세올린 3′ 비번역 부위를 포함하며, 시판 플라스미드 pUC18(지브코-BRL)로부터 유도된 벡터 pCW109의 변형물이다. 벡터 pCW109는 β-콘글리시닌 유전자의 555 bp 5′ 비-코딩 부위(프로모터 부위 포함)를 클로닝 벡터 pUC18의 Hind III 위치에 삽입하고, 이어 Nco I, Sma I, Kpn I 및 Xba I에 대한 제한 효소 위치를 포함하는 다중 클로닝 서열을 삽입한 후, 통상적인 콩 파세올린 3′ 비번역 부위 1174 bp를 Hind III 위치 중에 삽입함으로써 제조하였다. 사용되는 β-콘글리시닌 프로모터 부위는 27 개의 뉴클레오티드 부분에서 차이가 나는, 공개된 β-콘글리시닌 유전자의 대립 유전자이다(도일(Doyle) 등의 문헌[J. Biol. Chem. 261: 9228-9238(1986)]). 이 유전자의 추가적 서열 설명은 국제 출원 공개 WO91/13993에서 발견할 수 있다.

안티센스 작제물에서 용이하게 사용하기 위해서, Nco I 위치 및 플라스미드 pCW109에서 가능한 번역 출발 위치를 Nco I로 분해함으로써 소실시키고, 멍(mung) 콩 엑소뉴클레아제 분해 및 블런트 위치의 재결합에 의해서 변형된 플라스미드 pCW109A를 얻었다.

벡터 pML18은 비-조직 특이 및 구성 콜리플라워 모자이크 바이러스(35S) 프로모터로 이루어지고(오델 등의 문헌[Nature 313:810-812(1985)]; 헐 등의 문헌[Virology 86:482-493(1987)]로 이루어지고, 이는 네오마이신 포스포

트랜스퍼라제 유전자(벡(Beck) 등의 문헌[Gene 19:327-336(1982)])의 발현을 유도하고, 뉴클레오티드 848 내지 1550를 포함하는 노팔린 신타제 유전자의 3'말단(데픽커(Depicker) 등의 문헌[J. Appl. Genet. 1:561- 574(1982)])을 유도한다. 이 전사 단위는 시판 클로닝 벡터 pGEM9z(지브코-BRL) 중에 삽입되었고, Sal I, Xba I, BAm HI 및 Sma I 제한 위치에 의해서 이러한 순서로 35S 프로모터의 5'말단에 플랭킹되었다. 부가적인 Sal I 위치가 NOS 3′서열의 3′말단에 존재하고, Xba I, Bam HI 및 Sal I 위치는 유일하다. 플라스미드 pML18은 Xba I로 분해되었고, DNA 중합효소 I의 클레나우 (Klenow) 단편을 사용하여 단일 가닥의 말단을 채우고, Xba I 위치를 제거하기 위해서 생성물을 결합시켰다. 생성되는 플라스미드를 pBS16으로 명명하였다.

플라스미드 pCW109A를 Hind III로 분해하고, β-콘글리시닌/안티센스 델타-12 데세츄라제 cDNA/파세올린 3′비번역 부위를 함유하는, 생성되는 1.84 kb 단편을 겔 분리하였다. 이 1.84 kb 단편을 pBS16의 Hind III 위치에 결합시켰다. 바람직한 배향으로 삽입물을 함유하는 플라스미드는 Kpn I로 분해시 3.53 kb 및 4.41 kb 단편을 생산하였고, 이 플라스미드를 pCST2로 명명하였다.

GmFad2-1 cDNA의 공급원으로서 벡터 pBS13을 사용하였고, 이는 대두 미크로조말 델타 12-데세츄라제를 코딩하고, 유전자은행 기탁 번호 L43920으로 기재된 바와 같은 서열을 갖는다. 벡터 pBS13은 벡터 pML70으로부터 유도되고(도 1), 이는 KTi3 프로모터 및 KTi3 3′비번역 부위를 포함하고, 중간 플라스미드 pML51, pML55, pML64 및 pML65를 경유하여 시판 벡터 pTZ18R(Pharmacia)로부터 유도되었다. 5′비번역 부위의 2039 개 뉴클레오티드, 및 조푸쿠의 문헌에서 기술된 서열 중 755 내지 761 염기에 상응하는 Eco RI 위치에서 끝나는 KTi3 유전자의 코딩 서열의 390 염기 모두를 포함하는, 완전한 대두 KTi3 유전자(조푸쿠(Jofuku) 및 골드버그의 문헌[Plant Cell 1: 1079-1093(1989)])의 2.4 kb Bst BI/Eco RI 단편을 pTZ18R의 Acc I/Eco RI 위치에 결합시켜 플라스미드 pML51을 제작하였다. KTi3 삽입물의 5′비번역 부위의 중간의 Nco I 위치를 소멸시키기 위해서, 플라스미드 pML51을 Nco I로 자르고, DNA 중합효소 I의 클레나우 단편을 사용하여 단일 가닥의 말단을 채우고, 생성물을 결합시켜 플라스미드 pML55를 얻었다. 플라스미드 pML55를 Xmm I/Eco RI로 부분적으로 분해하여, 상기 인용된 서열 중 732 내지 755 염기에 해당하는 0.42 kb 단편을 유리시키고 폐기시켰다. Xmn I 위치(5′-TCTTCC-3′) 및 Nco I 위치(5′-CCATGGG-3′)에 대한 코딩 서열로 이루어지는 상보적인 합성 올리고뉴클레오티드 이량체를 제조하고, 이어 Eco RI 위치(5′-GAAGG-3′)의 일부에 의해서 직접, Nco I 위치를 포함하는 합성 Xmm I/Eco RI 링커(linker)를 제작하였다. Xmn I 및 Nco I/Eco RI 위치를 짧은 개재 서열(5′-ATAGCCCCCCAA-3′)에 의해서 결합시켰다. 이러한 합성 링커를 4.94 kb 단편의 Xmm I/Eco RI 위치에 결합시켜 플라스미드 pML64를 제작하였다. 프라이머 ML51 및 ML52를 사용하여 표준 PCR 프로토콜(퍼킨 엘머 세투스, GeneAmp PCR 키트)에 의해서 조푸쿠의 문헌에 기재된 서열로부터 KTi3 유전자의 3′비번역 부위를 증폭시켰다. 프라이머 ML51은 프라이머의 5′말단에서 Eco RV(5′-GATATC-3′), Nco I (5′-CCATGG-3′), Xba I (5′-TCTAGA-3′), Sma I (5′-CCCGGG-3′) 및 Kpn I (5′-GGTACC-3′) 위치에 상응하는 뉴클레오티드의 첨가와 함께 상기 인용된 서열의 염기 1072 내지 1091에 상응하는 20 개의 뉴클레오티드를 포함하였다. 프라이머 ML52는 프라이머의 5′말단에서 Sma I (5′-CCCGGG-3′), Eco RI(5′-GAATTC-3′), Bam HI (5′-GGATCC-3′) 및 Sal I (5′-GTCGAC-3′) 위치에 상응하는 뉴클레오티드의 첨가와 함께 상기 인용된 서열의 염기 1242 내지 1259에 상응하는 뉴클레오티드를 정확하게 보충 함유하였다. KTi3 유전자의 PCR-증폭된 3'말단을 pML64의 Nco I/Eco RI 위치에 결합시켜 플라스미드 pML65를 제작하였다. 합성 다중 클로닝 위치 링커는 Pst I (5′-CTGCA-3′), Sal I (5′-GTCGAC-3′), Bam HI (5′-GGATCC-3′) 및 Pst I (5′-CTGCA-3′) 위치에 대한 코딩 서열로 이루어지는 상보적인 합성 올리고뉴클레오티드 이량체를 제조함으로써 제작되었다. 링커를 pML65의 Pst I 위치(5′에서 직접 KTi3 프로모터 부위로)에 결합시켜 플라스미드 pML70을 제작하였다.

대두 델타 12-데세츄라제 cDNA, GmFad2-1(유전자 은행 기탁 번호 L43920)의 1.46 kb Sma I/Kpn I 단편을 pML70의 상응하는 위치에 결합시켜 플라스미드 pBS10을 제작하였다. 데세츄라제 cDNA 단편은 pBS10내 KTi3 프로모터에 대하여 역(안티센스) 배향이었다. 플라스미드 pBS10는 Bam HI로 분해되었고, KTi3 프로모터/안티센스 데세츄라제 cDNA/KTi3 3′말단 전사 단위를 나타내는 3.47 kb 단편을 아가로스 겔 전기 영동에 의해서 단리하였다. 벡터 pML18은 비-조직 특이 및 구성 콜리플라워 모자이크 바이러스(35S) 프로모터로 이루어지고(오델 등의 문헌[Natur e 313:810 - 812(1985)]; 헐 등의 문헌[Virology 86:482 - 493(1987)], 네오마이신 포스포트랜스퍼라제 유전자의 발현을 유도하고(벡(Beck) 등의 문헌[Gene 19:327 - 336(1982)]), 이어 뉴클레오티드 848 내지 1550를 포함하는 노팔린 신타제 유전자의 3'말단(데픽커(Dedicker) 등의 문헌[J. Appl. Genet. 1:561 - 574(1982)])을 유도한다. 이러한 전사 단위를 시판 클로닝 벡터 pGEM9z(지브코 - BRL)에 삽입하고, Sal I, Xba I, Bam HI 및 Sma I 제한 위치에 의해서 이러한 순서로 35S 프로모터의 5' 말단에서 플랭킹하였다. 부가적인 Sal I 위치가 NOS 3 ′ 서열의 3 ′ 말단에 존재하고, Xba I, Bam HI 및 Sma I 위치는 유일하다. pBS10으로부터 유리된 3.47 kb 전사 단위를 벡터 pML18의 Bam HI 위치에 결합시켰다. 생성되는 플라스미드를 Sma I 및 Kpn I로 분해시킬 경우, 바람직한 배향의 삽입물을 함유하는 플라스미드는 5.74, 2.69 및 1.46 kb의 3 개의 단편을 생산하였다. 정확한 배향의 전사 단위를 갖는 플라스미드를 선택하여 pBS13으로 명명하였다.

pBS13(상기 설명됨)으로부터 1.46 kb XbaI/EcoRV 단편을 벡터 pCST2의 SmaI/XbaI 위치(상기 설명됨)에 직접적으로 클로닝하여 pBS39로 명명된 플라스미드를 생산하였다. 플라스미드 pBS39의 3.3 kb HindIII 단편을 플라스미드 pMV40(상기 설명됨)의 HindIII 위치에 클로닝하여 식물 발현 벡터 pBS43 (도 6)을 얻었다.

벡터 pBS43으로 대두의 형질전환 및 유전자전환 '트랜스위치(Transwich)' 계통의 확인

대두 분열 조직의 입자 포격 방법을 사용하여 벡터 pBS43을 대두 분열 조직에 클로닝하여 형질전환시켰다(크리스토우(Christou) 등의 문헌[Trends Biotechnol. 8:145 - 151(1990)]). 형질전환된 식물 종자(즉, GUS 활성에 대하여 양성으로 확인된 식물 유래)을 지방산 조성에 대해서 스크리닝하였다. 소량(약 10 mg)의 종자 칩(chip)의 헥산 추출물로부터 지방산 메틸 에스테르를 제조하였다(브로우즈(Browse) 등의 문헌[Anal. Biochem. 152:141 - 145(1986)]). 10 개의 다른 유전자전환 계통으로부터 종자 칩을 분석하고, 이들 계통 중 하나로부터 몇몇 R1 종자(260 - 05)는 대조 종자에서의 약 20 %와 비교하여 80 내지 85 %의 총 올레산 함량을 가졌다. 이 표현형은 내생 Fad 2 - 1 유전자의 공동억제에 의해서 발생되고, 이는 pBS43의 2 개 카피를 '트랜스위치'라고 명명된 대두 게놈 위치에 삽입한 결과이다(킨니, 에이. 제이(Kinney, A. J.) 등의 문헌['Induced Mutations and Molecular Techniques for Crop Improvement', International Atomic Energy Agency, Vienna(1995)]). 트랜스위치 위치를 포함하는, 계통 260 - 05로부터의 올레산 고함량 R1 종자를 자가수분하여 트랜스위치에 대하여 동형접합인 R2 종자를 선택하였다. 추가 분석을 위하여, 이들 R2 동형접합 종자들 중 2 개(G94 - 1 및 G94 - 19), 및 G94 - 1 및 G94 - 19의 추가 세대로부터 유도된 종자(R3, R4, R5)를 선택하였다.

아이오와(Iowa) 및 푸에르토 리코(Puerto Rico) 모두에서 재배한 G94 - 1 및 G94 - 19 식물의 R5 종자를 분말로 분쇄하고, 약 1 g을 헥산 5 ml로 추출하였다. 원심분리 후, 헥산을 부어 버리고 플레이크를 공기 건조시켰다. 약 10 mg의 탈지 분말을 상기 설명한 바와 같이 추출하고, SDS - PAGE에 의해서 분석하였다. 두 지역에서 얻어진 유전자전환 계열 모두에 있어서, 대조 대두 계열 및 표준 대두 분말과 비교하여 β - 콘글리시닌의 α ′ 및 α 서브유닛의 발현이 억제되었다(도 7).

< 실시예 3

안티센스 기술을 사용하여 β - 콘글리시닌 발현을 공동억제시킬 수 있는지를 시험하기 위해서, 상기 설명한 바와 같이 중합효소 사슬 반응에서 역전사효소를 사용하여 완전한 길이의 α 및 α ′ cDNA를 제조하였다. 상부 프라이머, ConSa는 Kpn I 제한 위치에 대하여 코딩하는 5 ′ 말단에 가해지는 부가적 뉴클레오티드와 함께 α 및 α ′ 모두의 cDNA 서열 4 - 19 부위에 상동이다. 사용되는 하부 프라이머, Con1.9a는 α 이소형을 나타내는 서열 번호:1의 1818 - 1801 부위 및 α ′ 이소형을 나타내는 서열 번호:2의 1920 - 1903과 상동이다. 후속되는 클로닝 단계를 돕기 위해서, 부가적 뉴클레오티드를 5 ′ 말단에 가하여 Nco I 제한 위치에 대하여 코딩하였다.

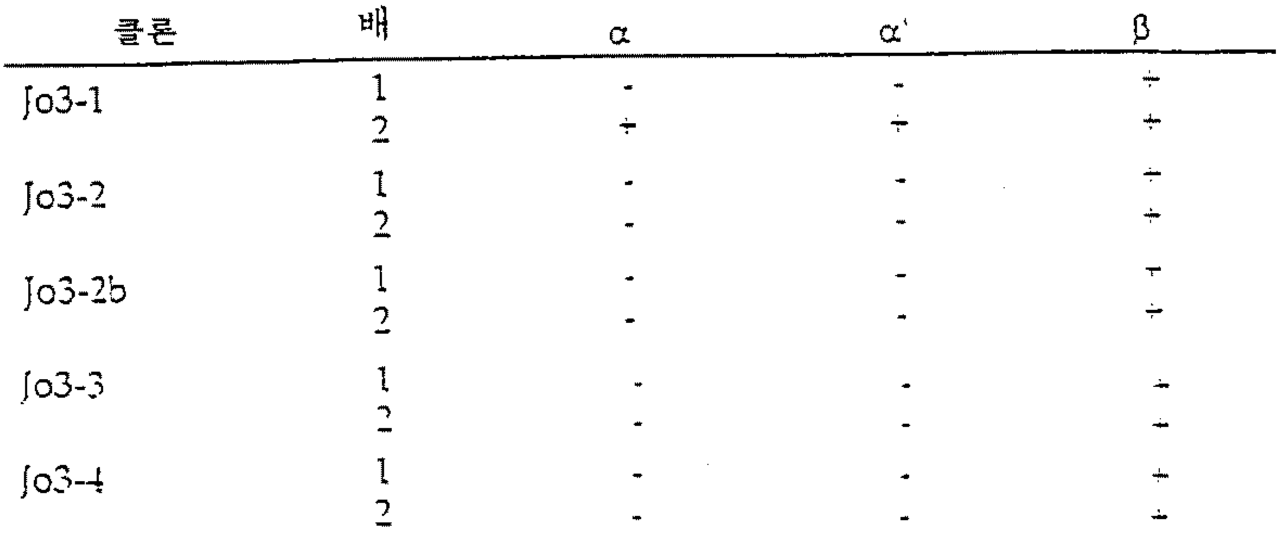
Con1.0 5'-CGGGTATGGCGAGTGTT-3' (서열 번호: 10)

역전사 및 후속 PCR 반응을 상기 설명한 바와 같이 수행하였다. 랜덤 육량체 또는 Con1.9a 중 하나를 사용하여(상기 상세히 설명된 방법), 발아하는 대두 종자로부터 분리된 RNA를 역전사시켰다. 상기 상술된 프로토콜을 사용하고 Con Sa 및 Con 1.9a를 사용하는 PCR 반응으로 cDNA를 증폭시켰다. 각 PCR 반응 혼합물 15 마이크로리터를 아가로스 겔 전기영동에 의해서 분석하였다 α에 대하여 예상된 분자량, 즉, 1.8 kb 밴드를 관찰하였다. 위자아드(Wizard)TM PCR 프렙 DNA 정제 시스템 키트 (프로메가(Promega)사) 사용하여 남아있는 반응 혼합물은 정제하였다. 사용되는 프라이머로 인하여 5′말단에 Kpn I 위치 및 3′말단에 Nco I 위치를 포함하는 α cDNA 단편을 Kpn I 및 Nco I 제한 효소로 분해하였다. 생성되는 α cDNA를 겔 분리하여 F10으로 표지하였고, 이를 플라스미드내 존재하는 Nco I 및 Kpn I 위치를 사용하여 pCW109 중에 직접 클로닝(안티센스 배향)하였다. 네스티드(nested) 프라이머(상부: Con.09 (서열번호:6); 하부 Con1.4a(서열번호:5); 및 Con1.0 (서열번호:10))를 사용하는 PCR에 의해서 F10은 α로서 확인되었다.

Con1.0 5'-CGGGTATGGCGAGTGTT-3' (서열 번호: 10)

전사 단위 β-콘글리시닌 프로모터/α cDNA 안티센스/파세올린 3′말단은 Hind III로 부분 분해함으로써 pCW109/F10으로부터 유리되었다. 부분 분해의 상태는 6 개의 단편 (5.1 kb, 3.8 kb, 3.6 kb, 2.6 kb, 2.4 kb 및 2.6 kb)으로로 생산되었다. 전사 단위를 포함하는 3.6 kb 단편을 겔 분리하여 F14로 표지하였다. 그 후, F14를 pKS18HH의 Hind III 위치 중에 클로닝하였다. Hind III으로 형질전환시킨 세포로부터 제조된 플라스미드 DNA 제작물을 분해시킴으로써 삽입을 확인한 후, 양성 배양액으로부터 플라스미드 DNA를 Kpn I으로 분해하고, 그들이 3.6 kb F14 단편을 함유하고 pCW109/F10을 Hind III로 부분 분해물 유래의 3.8 kb 단편을 함유하지 않는다는 것을 확실하게 하였다. F14는 Kpn I 위치를 포함하는 반면, 3.8 kb 단편은 포함하지 않는다. 확인 후 pKS18HH/F14를 pJo3으로 표지하였다(도 8). 대두 배 현탁 배양균을 상기 상술한 바와 같이 pJo3으로 형질전환하였다. 형질전환 결과, 5 개의 형질전환된 클론을 생성하였고, 성숙 후, 각 클론은 4 내지 8 개의 체세포 배를 발생하였다.

형질전환된 체세포 배의 단백질 추출물을 상기 상술한 바와 같이 SDS-PAGE에 의해서 분석하였다. 결과를 표 2에 나타내었다. 유전자전환된 클론은 α 및 α′모두의 발현이 억제되는 적어도 하나 이상의 체세포 배를 생성하였다.

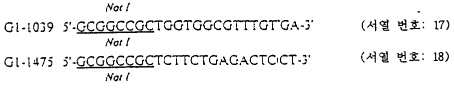


클론	배	α	α'	β
Jo3-1	1	-	-	+
	2	+	+	+
Jo3-2	1	-	-	+
	2	-	-	+
Jo3-2b	1	-	-	+
	2	-	-	+
Jo3-3	1	-	-	+
	2	-	-	+
Jo3-4	1	-	-	+
	2	-	-	+

< 실시예 4

글리시닌 서브유닛을 코딩하는 5 개의 비대립 유전자가 있었다. 게놈 클론 및 cDNA를 서열화함으로써 서브유닛 유전자를 서열 유사성을 근거로 하는 2 개 군으로 나누어졌다 군 I은 Gy1(서열 번호:11), Gy2(서열 번호:12) 및 Gy3(서열 번호:13)으로 이루어지고, 군 II는 Gy4(서열 번호:14) 및 Gy5(서열 번호:15)로 이루어졌다. 하나의 군 범위내 유전자 사이에는 85 %를 넘는 유사성이 있었지만, 다른 군의 유전자 사이에는 42% 내지 46 %의 유사성이 있었다. 공동억제 기술을 사용함으로써 발아하는 떡잎에서 글리시닌의 발현을 억제시킬 수 있는지 결정하기 위해서, 상기 설명한 바와 같은 중합 효소 사슬 반응에서 역전사효소를 사용하여 군 I 및 군 II의 cDNA를 제조하였다.

군 I 반응(G1 - 1)에 사용되는 상부 프라이머는 모든 군 I cDNA의 1 - 19 부위에 상동이었다. 두 개의 하부 프라이머, Gy1의 1038 - 1022 부위, Gy2의 1008 - 992 부위 및 Gy3의 996 - 980 부위에 상동인 G1 - 1039, 또는 Gy1의 1475 - 1460 부위, Gy2의 1445 - 1430 부위 및 Gy3의 1433 - 1418 부위에 상동인 G1 - 1475를 사용하였다. 이후의 클로닝을 돕기 위해서, 모든 프라이머는 5 ' 말단에서 Not I 제한 위치를 코딩하는 부가적 뉴클레오티드를 함유하였다.



G1-1 5'-GCGGCCGCATGGCCAAGCTAGTTTTTTT-3' (서열 번호: 16)
Not I
G1-1039 5'-GCGGCCGCTGGTGGCGTTTGTGA-3' (서열 번호: 17)
Not I
G1-1475 5'-GCGGCCGCTCTTCTGAGACTCCT-3' (서열 번호: 18)
Not I

반응 중 하부 프라이머로서 랜덤 육량체, 또는 G1 - 1475 또는 G1 - 1039 중 하나를 사용하여, 발아하는 대두 종자로부터 분리된 RNA를 역전사시켰다. G1 - 1475 또는 G1 - 1039 중 하나와 G1 - 1의 혼합물을 사용하여 cDNA 단편을 증폭시켰다. PCR 반응 혼합물 15 마이크로리터를 아가로스 겔 전기영동에 의해서 분석하였다 PCR 반응 결과, G1 - 1/G1 - 1475 및 G1 - 1/G1 - 1039 각각의 프라이머 세트에 대하여 예상된 분자량, 약 1 kb 및 1.4 내지 1.5 kb의 생성물을 얻었다. 위자아드(Wizard)TM PCR 프렙 DNA 정제 시스템 키트 (프로메가사) 반응 혼합물 잔사로부터의 cDNA 단편을 정제하였다. 그 후, 정제된 cDNA를 Not I로 분해하고, 아가로스 겔 정제에 의해서 단리하였다.

군 II의 RT-PDR 반응에서 사용되는 상부 프라이머(G4-7)는 군 II의 모든 cDNA 7-22 부위에 상동이었다. 두 개의 하부 프라이머, 즉, Gy4의 1251-1234 부위 및 Gy5의 1153-1235 부위에 상동인 G4-1251; 또는 Gy4의 1668-1653 부위에 상동인 G4-1670을 사용하였다. Gy5에는 유사한 부위가 없었다. 이후의 클로닝을 돕기 위해서, 모든 프라이머는 5' 말단에서 Not I 제한 위치를 코딩하는 부가적 뉴클레오티드를 포함하였다.

G4-7 5'-GCGGCCGCATGCCCTTCACTCTCT-3' (서열 번호: 19)
Not I
G4-1251 5'-GCGGCCGCTGGGAGGGTGAGGCTGTT-3' (서열 번호: 20)
Not I
G4-1670 5'-GCGGCCGCTGAGCCTTGTTGAGAC-3' (서열 번호: 21)
Not I

반응에서 하부 프라이머로서 랜덤 육량체, 또는 G4-1251 또는 G4-1670 중 하나를 사용하여, 발아하는 대두 종자로부터 분리된 RNA를 역전사하였다. G4-1251 또는 G4-1670 중 하나와 G4-7의 혼합물을 사용하여 cDNA 단편을 증폭시켰다. PCR 반응 혼합물 15 마이크로리터를 아가로스 겔 전기영동에 의해서 분석하였다 PCR 반응 결과, G4-7/G4-1251 및 G4-7/G4-16.70 프라이머 세트에 대하여 예상된 분자량, 약 1.25 및 1.7 kb의 생성물을 얻었다. 위자아드(Wizard)TM PCR 프렙 DNA 정제 시스템 키트 (프로메가(Promega)사제)사용하여 반응 혼합물 잔사로부터의 cDNA 단편을 정제하였다. 그 후, 정제된 cDNA를 Not I로 분해하고, 아가로스 겔로부터 단리하였다.

단리된 군 I cDNA를 Not I 위치에서 pRB20(도 9) 중에 클로닝(센스 배향)하였다. Not I로 부분적인 제한 분해 및 한번 절단된 pRB20/군 I 선상 단편의 분리 후, 군 II cDNA를 가하여, 최종 전사 단위 β-콘글리시닌 프로모터/군 I cDNA (센스 배향)/파세올린 3' 말단 및 β-콘글리시닌 프로모터/군 II cDNA (센스 배향)/파세올린 3' 말단을 제작하였다. 그 후, 상기 상술한 방법에 따라서, 생성되는 플라스미드를 사용하여 체세포 배 현탁 배양균을 형질전환시켰다.

미생물 기탁에 관한 표시

(PCT 규칙 13bis)

A. 하기 표시는 명세서 7페이지 1행에서 언급된 미생물에 관한 것이다.	
B. 기탁의 확인	
기탁 기관의 명칭 아메리칸 타입 컬쳐 콜렉션 (ATCC)	
기탁 기관의 주소 (우편번호 및 국명 포함) 미국 20852 메릴랜드주 록크빌 파크로운 드라이브 12301	
기탁일: 1996년 6월 15일	수탁번호 97614
C. 추가 공시 사항	
유럽 특허를 청구한 상기 명명된 미생물에 대하여 기탁된 미생물의 시료는 유럽 특허 공고가 발행되거나 또는 출원이 거절되거나 포기되거나 포기 간주될 때까지 시료의 분양을 신청한 사람에 의해 지명된 숙련가에게 분양함으로써만 이용될 수 있다 (EPC 규칙 28(4)).	
D. 공시 사항이 적용되는 지정국	
E. 공시 사항의 별도 공급	
이 서류는 국제 출원과 함께 접수되었음을 확인함 권한있는 관리	

Form PCT/RO/134(1992년 7월)

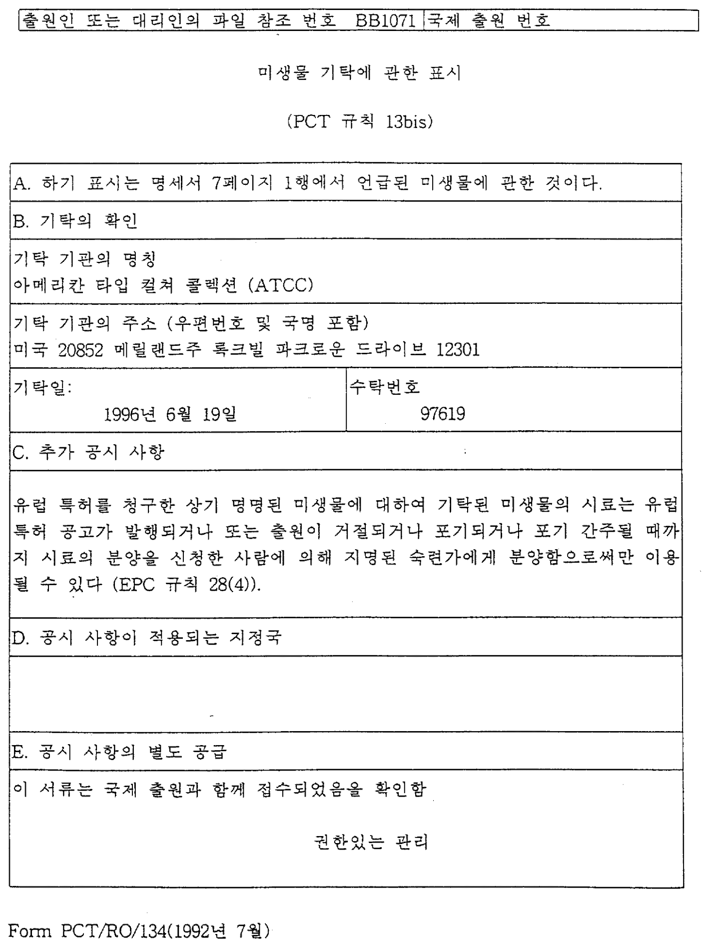
출원인 또는 대리인의 파일 참조 번호 BB1071	국제 출원 번호

미생물 기탁에 관한 표시

(PCT 규칙 13bis)

A. 하기 표시는 명세서 7페이지 1행에서 언급된 미생물에 관한 것이다.	
B. 기탁의 확인	
기탁 기관의 명칭 아메리칸 타입 컬쳐 콜렉션 (ATCC)	
기탁 기관의 주소 (우편번호 및 국명 포함) 미국 20852 메릴랜드주 록크빌 파크로운 드라이브 12301	
기탁일: 1996년 6월 15일	수탁번호 97615
C. 추가 공시 사항	
유럽 특허를 청구한 상기 명명된 미생물에 대하여 기탁된 미생물의 시료는 유럽 특허 공고가 발행되거나 또는 출원이 거절되거나 포기되거나 포기 간주될 때까지 시료의 분양을 신청한 사람에 의해 지명된 숙련가에게 분양함으로써만 이용될 수 있다 (EPC 규칙 28(4)).	
D. 공시 사항이 적용되는 지정국	
E. 공시 사항의 별도 공급	
이 서류는 국제 출원과 함께 접수되었음을 확인함 권한있는 관리	

Form PCT/RO/134(1992년 7월)



출원인 또는 대리인의 파일 참조 번호 BB1071	국제 출원 번호

미생물 기탁에 관한 표시

(PCT 규칙 13bis)

A. 하기 표시는 명세서 7페이지 1행에서 언급된 미생물에 관한 것이다.	
B. 기탁의 확인	
기탁 기관의 명칭 아메리칸 타입 컬쳐 콜렉션 (ATCC)	
기탁 기관의 주소 (우편번호 및 국명 포함) 미국 20852 메릴랜드주 록크빌 파크로운 드라이브 12301	
기탁일: 1996년 6월 19일	수탁번호 97619
C. 추가 공시 사항	
유럽 특허를 청구한 상기 명명된 미생물에 대하여 기탁된 미생물의 시료는 유럽 특허 공고가 발행되거나 또는 출원이 거절되거나 포기되거나 포기 간주될 때까지 시료의 분양을 신청한 사람에 의해 지명된 숙련가에게 분양함으로써만 이용될 수 있다 (EPC 규칙 28(4)).	
D. 공시 사항이 적용되는 지정국	
E. 공시 사항의 별도 공급	
이 서류는 국제 출원과 함께 접수되었음을 확인함 권한있는 관리	

Form PCT/RO/134(1992년 7월)

발명의 효과

본 발명에 의해 종자 저장 단백질 프로파일 변화 결과, 독특하고 유용한 기능성을 갖는 신규의 대두 단백질 제품을 생산할 수 있었다.

(57) 청구의 범위

청구항 1.

(a) (i) 대두 종자 저장 단백질 유전자로부터의 프로모터 부위를 포함하는 핵산 단편,

(ii) (i)의 프로모터에 대하여 센스 또는 안티센스 배향으로 위치한, (i)의 대두 종자 저장 단백질이 아닌 대두 단백질 유전자의 전부 또는 일부를 포함하는 핵산 단편, 및

(iii) 전사 종결 부위

를 포함하는 키메라 유전자를 제작하는 단계;

(b) (a)의 키메라 유전자를 대두 종자 중에 도입하여 유전자전환 대두 종자를 생성하는 단계; 및

(c) 단계 (a)의 키메라 유전자를 발현시키면서 단계 (b)의 유전자전환 대두 종자를 성장시키는 단계

를 포함하는, 단계 (a)의 키메라 유전자를 함유하지 않는 대두와 비교하여 대두 종자 저장 단백질 서브유닛 계열 중 1 종 이상의 양 및 (a) (ii)의 핵산 단편에 의해 코딩되는 단백질의 양이 감소된, 2 종의 대두 유전자의 발현을 동시에 감소시키는 방법.

청구항 2.

제1항에 있어서, (a) (i)의 대두 종자 저장 단백질이 아닌 대두 단백질 유전자의 전부 또는 일부를 포함하는 핵산 단편을 프로모터 부위에 대하여 센스 배향으로 위치시키는 방법.

청구항 3.

제1항에 있어서, (a) (i)의 대두 종자 저장 단백질이 아닌 대두 단백질 유전자의 전부 또는 일부를 코딩하는 핵산 단편을 프로모터 부위에 대하여 안티센스 배향으로 위치시키는 방법.

청구항 4.

제1항에 있어서, 프로모터가 β-콘글리시닌 대두 종자 저장 단백질의 알파 서브유닛을 코딩하는 유전자로부터 유도되는 방법.

청구항 5.

제1항에 있어서, (a) (i)의 대두 종자 저장 단백질이 아닌 대두 단백질 유전자의 전부 또는 일부를 코딩하는 핵산 단편이 지방산 생합성에 관련되는 유전자를 코딩하는 방법.

청구항 6.

제1항에 있어서, 단계 (a)의 키메라 유전자를 함유하지 않는 대두와 비교하여 대두 종자 저장 단백질 서브유닛 계열 중 1 종 이상의 양 및 (a) (ii)의 핵산 단편에 의해서 코딩되는 단백질의 양이 감소되는 방법.

청구항 7.

제6항에 있어서, 단계 (a)의 키메라 유전자를 함유하지 않는 대두 종자와 비교하여 대두 종자 저장 단백질 서브유닛 계열 중 2 종 이상이 감소되고, 단계 (a)의 키메라 유전자를 함유하지 않는 대두 종자와 비교하여 단계 (a)의 키메라 유전자를 함유하는 대두 종자의 지방산 프로파일이 변화되는 방법.

청구항 8.

제 1 항의 방법으로 얻은 대두 종자로부터 제조되는 대두 기재 제품.

청구항 9.

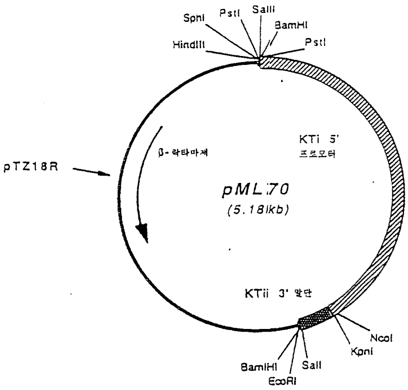
제 1 항의 방법으로 얻은 대두 종자로부터 제조되는 사료.

청구항 10.

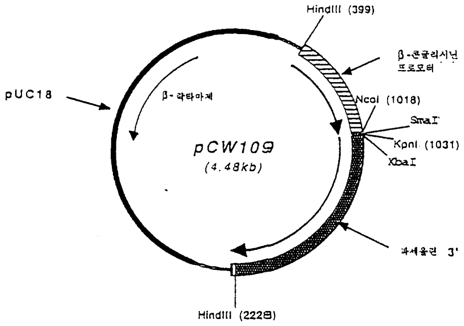
보통의 정상 대두로부터 제조되는 대두 기재 제품에 대하여 대두 저장 단백질의 서브유닛 계열중 2종류 이상이 감소된 대두 기재 제품.

도면

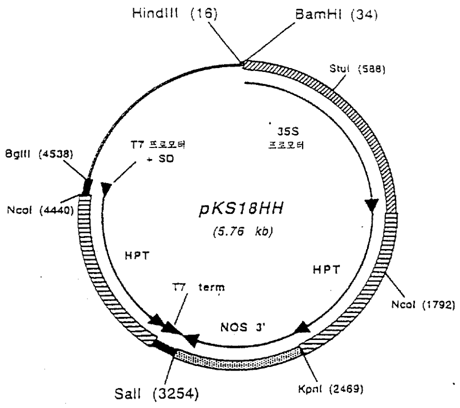
도면 1



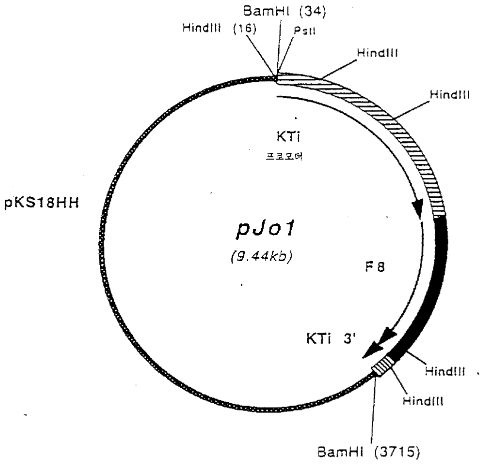
도면 2



도면 3


HindIII (16)
BamHI (34)
StuI (588)
35S
프로모터
T7 프로모터
+ SD
BglII (4538)
NcoI (4440)
pKS18HH
(5.76 kb)
HPT
HPT
T7 term
NOS 3'
NcoI (1792)
SalI (3254)
KpnI (2469)

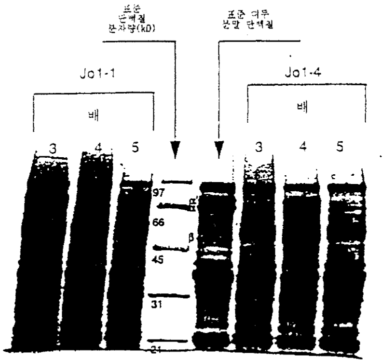
도면 4


BamHI (34)
HindIII (16)
PstI
HindIII
HindIII
KTi
프로모터
pKS18HH
pJo1
(9.44kb)
F8
KTi 3'
HindIII
HindIII
BamHI (3715)

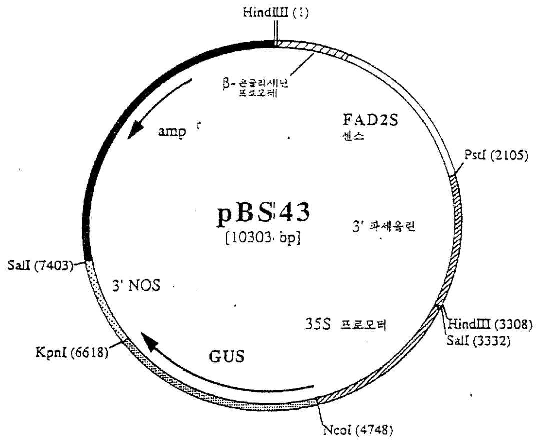
도면 5

형질전환 체세포 배 pJol-1 및 pJol-4로부터 추출된 단백질의 SDS-PAGE

(10 % 겔, 쿠마씨 염색)

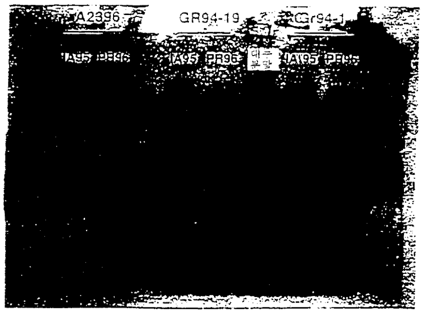


도면 6

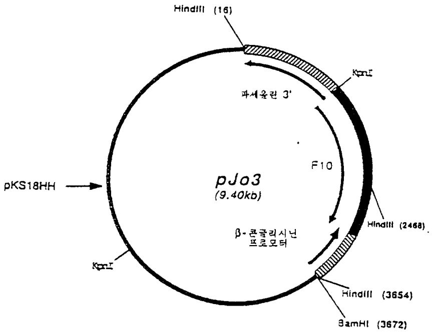


도면 7

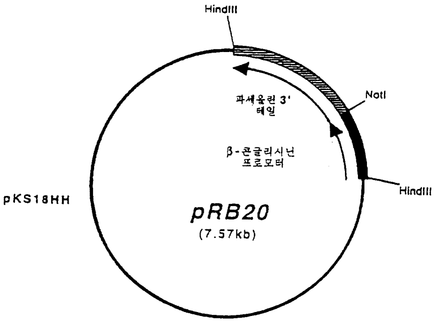
형질전환 대두 종자 G94-1 및 G94-19로부터 추출된 단백질의 SDS-PAGE (10 % 겔, 쿠마씨 염색)



도면 8



도면 9



```
SEQUENCE LISTING (1)GENERAL INFORMATION: (i)APPLICANT:ANTHONY JOHN KINNEY GARY MICH
AEL FADER (ii)TITLE OF INVENTION:SUPPRESSION OF SPECIFIC CLASSES OF SOYBEAN SEED PROTE
IN GENES (iii)NUMBER OF SEQUENCES:  21 (iv)CORRESPONDENCE ADDRESS: (A)NAME:  E. I. DU PO
NT DE NEMOURS AND COMPANY (B)STREET:  1007 MARKET STREET (C)CITY:  WILMINGTON (D)ST
ATE:  DELAWARE (E)COUNTRY:  UNITED STATES OF AMERICA (F)ZIP:  19898 (v)COMPUTER READA
BLE FORM: (A)MEDIUM TYPE:  DISKETTE, 3.50 INCH (B)COMPUTER:  IBM PC COMPATIBLE (C)OPER
ATING SYSTEM:  MICROSOFT WINDOWS 95 (D)SOFTWARE:  MICROFSOFT WORD 7.0A (vi)CURRENT
APPLICATION DATA: (A)APPLICATION NUMBER: (B)FILING DATE: (C)CLASSIFICATION: (vii) PRIOR
APPLICATION DATA: (A)APPLICATION NUMBER:  60/019,940 (B)FILING DATE:  JUNE 14,1996
(viii)A
TTORNEY/AGENT INFORMATION: (A)NAME:  LYNNE M. CHRISTENBURY (B)REGISTRATION NUMBER
:  30,971 (C)REFERENCE/DOCKET NUMBER:  BB-1071-A (ix)TELECOMMUNICATION INFORMATION:
(A)TELEPHONE:  302-992-5481 (B)TELEFAX:  302-773-0164 (2)INFORMATION FOR SEQ ID NO:1:
(i)SEQUENCE CHARACTERISTICS: (A)LENGTH:  1818 base pairs (B)TYPE:  nucleic acid
(C)STRANDED
NESS:  single (D)TOPOLOGY:  linear (ii)MOLECULE TYPE:  cDNA (xi)SEQUENCE DESCRIPTION:
SEQ I
D NO:1: ATGATGAGAG CACGGTTCCC ATTACTGTTG CTGGGACTTG TTTTCCTGGC TTCAGTTTCT    60
GTCTCATTTG GCATTGCTTA CTGGGAAAAA GAGAACCCCA AACACAACAA GTGTCTCCAG   120 AGTTG
CAATA GCGAGAGAGA CTCGTACAGG AACCAAGCAT GCCACGCTCG TTGCAACCTC   180 CTTAAGGTGG
AGAAAGAAGA ATGTGAAGAA GGTGAAATTC CACGACCACG ACCACGACCA   240 CAACACCCGG AGAGG
GAACC TCAGCAACCC GGTGAGAAGG AGGAAGACGA AGATGAGCAA   300 CCACGTCCAA TCCCATTCCC
ACGCCCACAA CCTCGTCAAG AAGAAGAGCA CGAGCAGAGA   360 GAGGAACAGG AATGGCCTCG CAAGGA
GGAA AAACGCGGAG AAAAGGGAAG TGAAGAGGAA   420 GATGAGGATG AGGATGAGGA ACAAGATGAA
CGTCAATTCC CATTCCCACG CCCACCTCAT   480 CAGAAGGAAG AGCGAAACGA AGAGGAAGAT GAGGAT
GAGG AGCAGCAGCG AGAGAGCGAA   540 GAAAGTGAAG ATTCTGAGTT ACGAAGACAT AAGAATAAGA
ACCCTTTTCT CTTCGGCTCT   600 AACAGGTTCG AAACTCTCTT CAAAAACCAA TATGGTCGCA TTCGCG
TCCT CCAGAGGTTC   660 AACCAACGCT CCCCACAACT TCAGAATCTC CGAGACTACC GCATTTTGGA GT
TCAACTCC   720 AAACCCAACA CCCTCCTTCT CCCCAACCAT GCTGACGCTG ATTACCTCAT CGTTATCCT
T   780 AACGGGACTG CCATTCTTTC CTTGGTGAAC AACGACGACA GAGACTCCTA CAGACTTCAA   840
TCTGGTGATG CCCTGAGAGT CCCCTCAGGA ACCACATACT ATGTGGTCAA CCCTGACAAC   900 AACGAA
AATC TCAGATTAAT AACACTCGCC ATACCCGTTA ACAAGCCTGG TAGATTTGAG   960 AGTTTCTTCC T
ATCTAGCAC TGAAGCTCAA CAATCCTACT TGCAAGGATT CAGCAGGAAC  1020 ATTTTAGAGG CCTCCT
ACGA TACCAAATTC GAGGAGATAA ACAAGGTTCT GTTTAGTAGA  1080 GAGGAAGGGC AGCAGCAAGG
GGAGCAGAGG CTGCAAGAGA GCGTGATTGT GGAAATCTCG  1140 AAGGAACAGA TTCGGGCACT GAGCA
AACGT GCCAAATCTA GTTCAAGGAA AACCATTTCT  1200 TCTGAAGATA AACCTTTTAA CTTGAGAAG
C CGCGACCCCA TCTACTCCAA CAAGCTTGGC  1260 AAGTTCTTTG AGATCACCCC AGAGAAAAAC CCCCA
GCTTC GGGACTTGGA TATCTTCCTC  1320 AGTATTGTGG ATATGAACGA GGGAGCTCTT CTTCTACCAC
ACTTCAATTC AAAGGCGATA  1380 GTGATACTGG TAATTAATGA AGGAGATGCA AACATTGAAC TTGT
TGGCCT AAAAGAACAA  1440 CAACAGGAGC AGCAACAGGA AGAGCAACCT TTGGAAGTGC GGAAATATA
G AGCCGAATTG  1500 TCTGAACAAG ATATATTTGT AATCCCAGCA GGTTATCCAG TTGTGGTCAA CGC
TACCTCA  1560 AATCTGAATT TCTTTGCTAT TGGTATTAAT GCCGAGAACA ACCAGAGGAA CTTCCTCG
CA  1620 GGTTCGCAAG ACAATGTGAT AAGCCAGATA CCTAGTCAAG TGCAGGAGCT TGCATTCCCT  16
80 GGGTCTGCAC AAGCTGTTGA GAAGCTATTA AAGAACCAAA GAGAATCCTA CTTTGTGGAT  1740 GC
TCAGCCTA AGAAGAAAGA GGAGGGGAAT AAGGGAAGAA AGGGTCCTTT GTCTTCAATT  1800 TTGAGG
GCTT TTTACTGA                                                  1818 (2)INFORMATION FOR SEQ
ID NO:2: (i)SEQ
UENCE CHARACTERISTICS: (A)LENGTH:  1920 base pairs (B)TYPE:  nucleic acid
(C)STRANDEDNESS:
single (D)TOPOLOGY:  linear (ii)MOLECULE TYPE:  cDNA (xi)SEQUENCE DESCRIPTION:  SEQ ID
NO:2
```

: ATGATGAGAG CGCGGTTCCC ATTACTGTTG CTGGGAGTTG TTTTCCTAGC ATCAGTTTCT 60 GTCT
CATTTG GCATTGCGTA TTGGGAAAAG CAGAACCCCA GTCACAACAA GTGCCTCCGA 120 AGTTGCAAT
A GCGAGAAAGA CTCCTACAGG AACCAAGCAT GCCACGCTCG TTGCAACCTC 180 CTTAAGGTGG AGGA
AGAAGA AGAATGCGAA GAAGGTCAAA TTCCACGACC ACGACCACAA 240 CACCCGGAGA GGGAACGTCA
GCAACACGGT GAGAAGGAGG AAGACGAAGG TGAGCAGCCA 300 CGTCCATTCC CATTCCCACG CCCACG
CCAA CCTCATCAAG AGGAAGAGCA CGAGCAGAAG 360 GAGGAACACG AATGGCATCG CAAGGAGGAA A
AACACGGAG GAAAGGGAAG TGAAGAGGAA 420 CAAGATGAAC GTGAACACCC ACGCCCACAC CAACCTC
ATC AAAAGGAAGA GGAAAAGCAC 480 GAATGGCAAC ACAAGCAGGA AAAGCACCAA GGAAAGGAAA G
TGAAGAAGA AGAAGAAGAC 540 CAAGACGAGG ATGAGGAGCA AGACAAAGAG AGCCAAGAAA GTGAA
GGTTC TGAGTCTCAA 600 AGAGAACCAC GAAGACATAA GAATAAGAAC CCTTTTCACT TCAACTCTAA
AAGGTTCCAA 660 ACTCTCTTCA AAAACCAATA TGGCCACGTT CGCGTCCTCC AGAGGTTCAA CAAACG
CTCC 720 CAACAGCTTC AGAATCTCCG AGACTACCGC ATTTTGGAGT TCAACTCCAA ACCCAACACC 7
80 CTTCTTCTCC CCCACCATGC TGACGCTGAT TACCTCATCG TTATCCTTAA CGGGACTGCC 840 ATTC
TTACCT TGGTGAACAA CGACGACCGA GACTCTTACA ACCTTCAATC TGGCGATGCC 900 CTAAGAGTCC
CTGCAGGAAC CACATTCTAT GTGGTTAACC CTGACAACGA CGAGAATCTC 960 AGAATGATAG CAGGA
ACCAC ATTCTATGTG GTTAACCCTG ACAACGACGA GAATCTCAGA 1020 ATGATAACAC TCGCCATACC
CGTTAACAAA CCCGGTAGAT TTGAGAGTTT CTTCCTATCT 1080 AGCACTCAAG CTCAACAGTC CTACT
TGCAA GGGTTCAGCA AGAATATTCT AGAGGCCTCA 1140 TACGACACCA AATTCGAGGA GATAAACAAG
GTTCTGTTTG GTAGAGAGGA GGGGCAGCAA 1200 CAAGGGGAGG AGAGGCTGCA AGAGAGTGTG ATTG
TGGAAA TCTCAAAGAA ACAAATTCGG 1260 GAACTGAGCA AACATGCCAA ATCTAGTTCA AGGAAAAC
CA TTTCTTCTGA AGATAAACCT 1320 TTCAACTTGG GAAGCCGCGA CCCCATCTAT TCCAACAAGC TTG
GCAAGTT GTTTGAGATT 1380 ACCCAGAGAA ACCCTCAGCT TCGGGACTTG GATGTCTTCC TCAGTGTT
GT GGATATGAAC 1440 GAGGGAGCTC TTTTTCTACC ACACTTCAAT TCAAAGGCCA TAGTGGTACT AG
TGATTAAT 1500 GAAGGAGAAG CAAACATTGA ACTTGTTGGC ATTAAAGAAC AACAACAGAG GCAGCA
ACAG 1560 GAAGAGCAAC CTTTGGAAGT GCGGAAATAT AGAGCTGAAT TGTCTGAACA AGATATATTT
1620 GTAATCCCAG CAGGTTATCC AGTTATGGTC AACGCTACCT CAGATCTGAA TTTCTTTGCT 1680 T
TTGGTATCA ATGCCGAGAA CAACCAGAGG AACTTCCTTG CAGGTTCGAA AGACAATGTG 1740 ATAAGC
CAGA TACCTAGTCA AGTGCAGGAG CTTGCGTTCC CTAGGTCTGC AAAAGATATT 1800 GAGAACCTAA
TAAAGAGCCA AAGTGAGTCC TACTTTGTGG ATGCTCAGCC TCAGCAGAAA 1860 GAGGAGGGGA ACAAG
GGAAG AAAGGGTCCT TTGTCTTCAA TTTTGAGGGC TTTTTACTGA 1920 (2)INFORMATION FOR SEQ
ID NO:3: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 1320 base pairs (B)TYPE: nucleic acid
(C)S
TRANDEDNESS: single (D)TOPOLOGY: linear (ii)MOLECULE TYPE: cDNA (xi)SEQUENCE DESCRIPTI
ON: SEQ ID NO:3: ATGATGAGAG TGCGGTTTCC TTTGTTGGTG TTGCTGGGAA CTGTTTTCCT GGCATC
AGTT 60 TGTGTCTCAT TAAAGGTGAG AGAGGATGAG AATAACCCTT TCTACTTTAG AAGCTCTAAC
120 AGCTTCCAAA CTCTCTTTGA GAACCAAAAC GTTCGCATTC GTCTCCTCCA GAGATTCAAC 180 AAA
CGCTCCC CACAACTTGA GAACCTTCGA GACTACCGGA TTGTCCAGTT TCAGTCAAAA 240 CCCAACACA
A TCCTTCTCCC CCACCATGCT GACGCCGATT TCCTCCTCTT TGTCCTTAGC 300 GGGAGAGCCA TACTT
ACCTT GGTGAACAAC GACGACAGAG ACTCCTACAA CCTTCACCCT 360 GGCGATGCCC AGAGAATCCC A
GCTGGAACC ACTTACTATT TGGTTAACCC TCACGACCAC 420 CAGAATCTCA AAATAATCAA ACTTGC
CATA CCCGTCAACA AACCTGGCAG ATATGATGAT 480 TTCTTCTTAT CTAGCACTCA AGCCCAACAG T
CCTACTTGC AAGGCTTCAG CCATAATATT 540 CTAGAGACCT CCTTCCATAG CGAATTCGAG GAGATA
AACA GGGTTTTGTT TGGAGAGGAA 600 GAGGAGCAGA GGCAGCAAGA GGGAGTGATC GTGGAACTCT
CAAAGGAACA AATTCGGCAA 660 CTGAGCAGAC GTGCCAAATC TAGTTCAAGG AAAACCATTT CCTCC
GAAGA TGAACCATTC 720 AACTTGAGAA GCCGCAACCC CATCTATTCC AACAACTTTG GAAAGTTCTT
TGAGATCACC 780 CCTGAGAAAA ACCCACAGCT TCGGGACTTG GATATCTTCC TCAGTTCTGT GGATA
TCAAC 840 GAAGGAGCTC TTCTTCTACC ACACTTCAAT TCAAAGGCCA TAGTGATACT AGTGATTAAT
900 GAAGGAGATG CAAACATTGA ACTTGTTGGC ATTAAAGAAC AACAACAGAA GCAGAAACAG 960 G
AAGAGGAAC CTTTGGAAGT GCAAAGGTAC AGAGCTGAAT TGTCTGAAGA CGATGTATTT 1020 GTAAT
TCCAG CAGCTTATCC ATTTGTCGTC AACGCTACCT CAAACCTCAA TTTCCTTGCT 1080 TTTGGTATCA

ATGCTGAGAA CAACCAGAGG AACTTCCTTG CAGGCGAGAA AGACAATGTG 1140 GTAAGGCAGA TAGAA
AGACA AGTGCAGGAG CTTGCGTTCC CTGGGTCTGC ACAAGATGTT 1200 GAGAGGCTAT TAAAGAAGCA
GAGGGAATCC TACTTTGTTG ATGCTCAGCC TCAGCAGAAG 1260 GAGGAGGGGA GTAAGGGAAG AAAGG
GTCCT TTTCCTTCAA TCTTAGGTGC TCTCTACTGA 1320 (2)INFORMATION FOR SEQ ID NO:4: (i)SEQ
UENCE CHARACTERISTICS: (A)LENGTH: 25 base pairs (B)TYPE: nucleic acid (C)STRANDEDNESS:
s
ingle (D)TOPOLOGY: linear (ii)MOLECULE TYPE: DNA (genomic) (xi)SEQUENCE DESCRIPTION:
SEQ
ID NO:4: CGTACCATGG TGAGAGCGCG GTTCC   25
(2)INFORMATION FOR SE
Q ID NO:5: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 24 base pairs (B)TYPE: nucleic acid
(C)S
TRANDEDNESS: single (D)TOPOLOGY: linear (ii)MOLECULE TYPE: DNA (genomic) (xi)SEQUENCE
ESCRIPTION: SEQ ID NO:5: CGGTACCGAA TTGAAGTGTG GTAG
24 (2)INF
ORMATION FOR SEQ ID NO:6: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 27 base pairs (B)TYPE
: nucleic acid (C)STRANDEDNESS: single (D)TOPOLOGY: linear (ii)MOLECULE TYPE: DNA
(genomic)
(xi)SEQUENCE DESCRIPTION: SEQ ID NO:6: TCGTCCATGG AGCGCGGTTC CCATTAC

27 (2)INFORMATION FOR SEQ ID NO:7: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 17 base
pairs
(B)TYPE: nucleic acid (C)STRANDEDNESS: single (D)TOPOLOGY: linear (ii)MOLECULE TYPE:
DNA
(genomic) (xi)SEQUENCE DESCRIPTION: SEQ ID NO:7: TCTCGGTCGT CGTTGTT

17 (2)INFORMATION FOR SEQ ID NO:8: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 25 base
pairs
(B)TYPE: nucleic acid (C)STRANDEDNESS: single (D)TOPOLOGY: linear (ii)MOLECULE TYPE:
DNA
(genomic) (xi)SEQUENCE DESCRIPTION: SEQ ID NO:8: ACGGTACCGA TGAGAGCGCG GTTCC

25 (2)INFORMATION FOR SEQ ID NO:9: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 27 base
pairs
(B)TYPE: nucleic acid (C)STRANDEDNESS: single (D)TOPOLOGY: linear (ii)MOLECULE TYPE:
DNA
(genomic) (xi)SEQUENCE DESCRIPTION: SEQ ID NO:9: AACCCATGGT CAGTAAAAAG CCCTCAA

27 (2)INFORMATION FOR SEQ ID NO:10: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 17 base pai
rs (B)TYPE: nucleic acid (C)STRANDEDNESS: single (D)TOPOLOGY: linear (ii)MOLECULE
TYPE: DN
A (genomic) (xi)SEQUENCE DESCRIPTION: SEQ ID NO:10: CGGGTATGGC GAGTGTT

17 (2)INFORMATION FOR SEQ ID NO:11: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 1488 base p
airs (B)TYPE: nucleic acid (C)STRANDEDNESS: single (D)TOPOLOGY: linear (ii)MOLECULE
TYPE: c
DNA (xi)SEQUENCE DESCRIPTION: SEQ ID NO:11: ATGGCCAAGC TAGTTTTTTC CCTTTGTTTT CTGC
TTTTCA GTGGCTGCTG CTTCGCTTTC   60 AGTTCCAGAG AGCAGCCTCA GCAAAACGAG TGCCAGATCC
AAAAACTCAA TGCCCTCAAA   120 CCGGATAACC GTATAGAGTC AGAAGGAGGG CTCATTGAGA CATGG
AACCC TAACAACAAG   180 CCATTCCAGT GTGCCGGTGT TGCCCTCTCT CGCTGCACCC TCAACCGCAA C

GCCCTTCGT   240 AGACCTTCCT ACACCAACGG TCCCCAGGAA ATCTACATCC AACAAGGTAA GGGTATT
TTT   300 GGCATGATAT ACCCGGGTTG TCCTAGCACA TTTGAAGAGC CTCAACAACC TCAACAAAGA   3
60 GGACAAAGCA GCAGACCACA AGACCGTCAC CAGAAGATCT ATAACTTCAG AGAGGGTGAT   420 TTG
ATCGCAG TGCCTACTGG TGTTGCATGG TGGATGTACA ACAATGAAGA CACTCCTGTT   480 GTTGCCGT
TT CTATTATTGA CACCAACAGC TTGGAGAACC AGCTCGACCA GATGCCTAGG   540 AGATTCTATC TTG
CTGGGAA CCAAGAGCAA GAGTTTCTAA AATATCAGCA AGAGCAAGGA   600 GGTCATCAAA GCCAGAAA
GG AAAGCATCAG CAAGAAGAAG AAAACGAAGG AGGCAGCATA   660 TTGAGTGGCT TCACCCTGGA AT
TCTTGGAA CATGCATTCA GCGTGGACAA GCAGATAGCG   720 AAAAACCTAC AAGGAGAGAA CGAAGGG
GAA GACAAGGGAG CCATTGTGAC AGTGAAAGGA   780 GGTCTGAGCG TGATAAAACC ACCCACGGAC GA
GCAGCAAC AAAGACCCCA GGAAGAGGAA   840 GAAGAAGAAG AGGATGAGAA GCCACAGTGC AAGGGTA
AAG ACAAACACTG CCAACGCCCC   900 CGAGGAAGCC AAAGCAAAAG CAGAAGAAAT GGCATTGACG AG
ACCATATG CACCATGAGA   960 CTTCGCCACA ACATTGGCCA GACTTCATCA CCTGACATCT ACAACCCT
CA AGCCGGTAGC 1020 GTCACAACCG CCACCAGCCT TGACTTCCCA GCCCTCTCGT GGCTCAGACT CAGT
GCTGAG 1080 TTTGGATCTC TCCGCAAGAA TGCAATGTTC GTGCCACACT ACAACCTGAA CGCGAACAG
C 1140 ATAATATACG CATTGAATGG ACGGGCATTG ATACAAGTGG TGAATTGCAA CGGTGAGAGA 12
00 GTGTTTGATG GAGAGCTGCA AGAGGGACGG GTGCTGATCG TGCCACAAAA CTTTGTGGTG 1260 GCT
GCAAGAT CACAGAGTGA CAACTTCGAG TATGTGTCAT TCAAGACCAA TGATACACCC 1320 ATGATCGG
CA CTCTTGCAGG GGCAAACTCA TTGTTGAACG CATTACCAGA GGAAGTGATT 1380 CAGCACACTT TCA
ACCTAAA AAGCCAGCAG GCCAGGCAGA TAAAGAACAA CAACCCTTTC 1440 AAGTTCCTGG TTCCACCT
CA GGAGTCTCAG AAGAGAGCTG TGGCTTAG   1488 (2)INFORMATION FOR SEQ ID NO:12: (i)
SEQUENCE CHARACTERISTICS: (A)LENGTH: 1458 base pairs (B)TYPE: nucleic acid
(C)STRANDEDNE
SS: single (D)TOPOLOGY: linear (ii)MOLECULE TYPE: cDNA (xi)SEQUENCE DESCRIPTION: SEQ
ID
NO:12: ATGGCCAAGC TTGTTCTTTC CCTTTGTTTC CTTCTTTTCA GTGGCTGCTT CGCTCTGAGA   60
GAGCAGGCAC AGCAAAATGA GTGCCAGATC CAAAAGCTGA ATGCCCTCAA ACCGGATAAC   120 CGTAT
AGAGT CGGAAGGTGG GTTCATTGAG ACATGGAACC CTAACAACAA GCCATTCCAG   180 TGTGCCGGTG
TTGCCCTCTC TCGCTGCACC CTTAACCGCA ATGCCCTTCG TAGACCTTCC   240 TACACCAACG GTCCCCA
GGA AATCTACATA CAACAAGGTA ATGGTATTTT TGGCATGATA   300 TTCCCGGGTT GTCCTAGCAC T
TATCAAGAG CCGCAAGAAT CTCAGCAACG AGGACGAAGC   360 CAGAGGCCCC AAGACCGTCA CCAAAAG
GTA CATCGCTTCA GAGAGGGTGA TTTGATCGCA   420 GTGCCTACTG GTGTTGCATG GTGGATGTAC A
ACAATGAAG ACACTCCTGT TGTTGCCGTT   480 TCTATTATTG ACACCAACAG CTTGGAGAAC CAGCTC
GACC AGATGCCTAG GAGATTCTAT   540 CTTGCTGGGA ACCAAGAGCA AGAGTTTCTA AAATATCAGC
AGCAGCAGCA AGGAGGTTCC   600 CAAAGCCAGA AAGGAAAGCA ACAAGAAGAA GAAAACGAAG GAAGC
AACAT ATTGAGTGGC   660 TTCGCCCCTG AATTCTTGAA AGAAGCGTTC GGCGTGAACA TGCAGATAGT
GAGAAACCTA   720 CAAGGTGAGA ACGAAGAGGA GGATAGTGGA GCCATTGTGA CAGTGAAAGG AGGTC
TAAGA   780 GTCACAGCTC CAGCCATGAG GAAGCCACAG CAAGAAGAAG ATGATGATGA TGAGGAAGAG
840 CAGCCACAGT GCGTGGAGAC AGACAAAGGT TGCCAACGCC AAAGCAAAAG GAGCAGAAAT   900 GG
CATTGATG AGACCATTTG CACAATGAGA CTTCGCCAAA ACATTGGTCA GAATTCATCA   960 CCTGACA
TCT ACAACCCTCA AGCTGGTAGC ATCACAACCG CCACCAGCCT TGACTTCCCA 1020 GCCCTCTGGC TTC
TCAAACT CAGTGCCCAG TATGGATCAC TCCGCAAGAA TGCTATGTTC 1080 GTGCCACACT ACACCCTGA
A CGCGAACAGC ATAATATACG CATTGAATGG GCGGGCATTG 1140 GTACAAGTGG TGAATTGCAA TGG
TGAGAGA GTGTTTGATG GAGAGCTGCA AGAGGGAGGG 1200 GTGCTGATCG TTCCACAAAA CTTTGCGG
TG GCTGCAAAAT CCCAGAGCGA TAACTTTGAG 1260 TATGTGTCAT TCAAGACCAA TGATAGACCC TC
GATCGGAA ACCTTGCAGG GGCAAACTCA 1320 TTGTTGAACG CATTGCCAGA GGAAGTGATT CAGCACA
CTT TTAACCTAAA GAGCCAGCAG 1380 GCCAGGCAGG TGAAGAACAA CAACCCTTTC AGCTTCCTTG TT
CCACCTCA GGAGTCTCAG 1440 AGGAGAGCTG TGGCTTAG
1458 (2)
INFORMATION FOR SEQ ID NO:13: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 1446 base pairs
(B)

TYPE: nucleic acid (C)STRANDEDNESS: single (D)TOPOLOGY: linear (ii)MOLECULE TYPE:
cDNA (x
i)SEQUENCE DESCRIPTION: SEQ ID NO:13: ATGGCTAAGC TTGTTCTTTC CCTTTGTTTT CTGCTTTTC
A GTGGCTGCTG CTTCGCTTTC   60 AGTTTCAGAG AGCAGCCACA GCAAAACGAG TGCCAGATCC AACGC
CTCAA TGCCCTAAAA   120 CCGGATAACC GTATAGAGTC AGAAGGTGGC TTCATTGAGA CATGGAACCC
TAACAACAAG   180 CCATTCCAGT GTGCCGGTGT TGCCCTCTCT CGCTGCACCC TCAACCGCAA CGCCCTT
CGC   240 AGACCTTCCT ACACCAACGC TCCCCAGGAG ATCTACATCC AACAAGGTAG TGGTATTTTT   3
00 GGCATGATAT TCCCGGGTTG TCCTAGCACA TTTGAAGAGC CTCAACAAAA AGGACAAAGC   360 AGC
AGGCCCC AAGACCGTCA CCAGAAGATC TATCACTTCA GAGAGGGTGA TTTGATTGCA   420 GTGCCAACC
G GTTTTGCATA CTGGATGTAC AACAATGAAG ACACTCCTGT TGTTGCCGTT   480 TCTCTTATTG ACA
CCAACAG CTTCCAGAAC CAGCTCGACC AGATGCCTAG GAGATTCTAT   540 CTTGCTGGGA ACCAAGAGC
A AGAGTTTCTA CAGTATCAGC CACAGAAGCA GCAAGGAGGT   600 ACTCAAAGCC AGAAAGGAAA GCGT
CAGCAA GAAGAAGAAA ACGAAGGAGG CAGCATATTG   660 AGTGGCTTCG CCCCGGAATT CTTGGAACA
T GCGTTCGTCG TGGACAGGCA GATAGTGAGA   720 AAGCTACAAG GTGAGAACGA AGAGGAAGAG AAG
GGTGCCA TTGTGACAGT GAAAGGAGGT   780 CTCAGCGTGA TAAGCCCACC CACGGAAGAG CAGCAACAA
A GACCCGAGGA AGAGGAGAAG   840 CCAGATTGTG ACGAGAAAGA CAAACATTGC CAAAGCCAAA GCAG
AAATGG CATTGACGAG   900 ACCATTTGCA CAATGAGACT TCGCCACAAC ATTGGCCAGA CTTCATCAC
C TGACATCTTC   960 AACCCTCAAG CTGGTAGCAT CACAACCGCT ACCAGCCTCG ACTTCCCAGC CCTCT
CGTGG 1020 CTCAAACTCA GTGCCCAGTT TGGATCACTC CGCAAGAATG CTATGTTCGT GCCACACTAC
1080 AACCTGAACG CAAACAGCAT AATATACGCA TTGAATGGAC GGGCATTGGT ACAAGTGGTG 1140 A
ATTGCAATG GTGAGAGAGT GTTTGATGGA GAGCTGCAAG AGGGACAGGT GTTAATTGTG 1200 CCACA
AAACT TTGCGGTGGC TGCAAGATCA CAGAGCGACA ACTTCGAGTA TGTTTCATTC 1260 AAGACCAATG
ATAGACCCTC GATCGGCAAC CTTGCAGGTG CAAACTCATT GTTGAACGCA 1320 TTGCCGGAGG AAGTG
ATTCA GCAAACTTTT AACCTAAGGA GGCAGCAGGC CAGGCAGGTC 1380 AAGAACAACA ACCCTTTCAG
CTTCCTGGTT CCACCTAAGG AGTCTCAGAG GAGAGTTGTG 1440 GCTTAG

1446 (2)INFORMATION FOR SEQ ID NO:14: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 1689 bas
e pairs (B)TYPE: nucleic acid (C)STRANDEDNESS: single (D)TOPOLOGY: linear
(ii)MOLECULE TYPE
: cDNA (xi)SEQUENCE DESCRIPTION: SEQ ID NO:14: ATGGGGAAGC CCTTCACTCT CTCTCTTTCT TC
CCTTTGCT TGCTACTCTT GTCGAGTGCA   60 TGCTTTGCTA TTAGCTCCAG CAAGCTCAAC GAGTGCCA
AC TCAACAACCT CAACGCGTTG   120 GAACCCGACC ACCGCGTTGA GTTCGAAGGT GGTTTGATTC AAA
CATGGAA CTCTCAACAC   180 CCTGAGCTGA AATGCGCCGG TGTCACTGTT TCCAAACTCA CCCTCAACC
G CAATGGCCTC   240 CACTTGCCAT CTTACTCACC TTATCCCCGG ATGATCATCA TCGCCCAAGG GAAAG
GAGCA   300 CTGCAGTGCA AGCCAGGATG TCCTGAGACG TTTGAGGAGC CACAAGAACA ATCAAACAGA
360 AGAGGCTCAA GGTCGCAGAA GCAGCAGCTA CAGGACAGTC ACCAGAAGAT TCGTCACTTC   420 AAT
GAAGGAG ACGTACTCGT GATTCCTCCT GGTGTTCCTT ACTGGACCTA TAACACTGGC   480 GATGAACCA
G TTGTTGCCAT CAGTCTTCTT GACACCTCTA ACTTCAATAA CCAGCTTGAT   540 CAAACCCCTA GGGT
ATTTTA CCTTGCTGGG AACCCAGATA TAGAGTACCC AGAGACCATG   600 CAACAACAAC AACAGCAGA
A AAGTCATGGT GGACGCAAGC AGGGGCAACA CCAGCAGGAG   660 GAAGAGGAAG AAGGTGGCAG CGTG
CTCAGT GGCTTCAGCA AACACTTCTT GGCACAATCC   720 TTCAACACCA ACGAGGACAT AGCTGAGAAA
CTTCAGTCTC CAGACGACGA AAGGAAGCAG   780 ATCGTGACAG TGGAAGGAGG TCTCAGCGTT ATCAG
CCCCA AGTGGCAAGA ACAACAAGAT   840 GAAGATGAAG ATGAAGACGA AGATGATGAA GATGAACAA
A TTCCCTCTCA CCCTCCTCGC   900 CGACCAAGCC ATGGAAAGCG TGAACAAGAC GAGGACGAGG ACGAA
GATGA AGATAAACCT   960 CGTCCTAGTC GACCAAGCCA AGGAAAGCGT GAACAAGACC AGGACCAGGA
CGAGGACGAA 1020 GATGAAGATG AAGATCAACC TCGCAAGAGC CGCGAATGGA GATCGAAAAA GACAC
AACCC 1080 AGAAGACCTA GACAAGAAGA ACCACGTGAA AGAGGATGCG AGACAAGAAA CGGGGTTGAG
1140 GAAAATATCT GCACCTTGAA GCTTCACGAG AACATTGCTC GCCCTTCACG CGCTGACTTC 1200 TA
CAACCCTA AAGCTGGTCG CATTAGTACC CTCAACAGCC TCACCCTCCC AGCCCTCCGC 1260 CAATTCCAA
C TCAGTGCCCA ATATGTTGTC CTCTACAAGA ATGGAATTTA CTCTCCACAT 1320 TGGAATCTGA ATG

CAAACAG TGTGATCTAT GTGACTCGAG GACAAGGAAA GGTTAGAGTT 1380 GTGAACTGCC AAGGGAA
TGC AGTGTTCGAC GGTGAGCTTA GGAGGGGACA ATTGCTGGTG 1440 GTACCACAGA ACTTCGTGGT G
GCGGAGCAA GCCGGAGAAC AAGGATTCGA ATACATAGTA 1500 TTCAAGACAC ACCACAACGC AGTCAC
TAGC TACTTGAAGG ATGTGTTTAG GGCAATTCCC 1560 TCAGAGGTTC TTGCCCATTC TTACAACCTT
CGACAGAGTC AAGTGTCTGA GCTTAAGTAT 1620 GAAGGAAATT GGGGTCCTTT GGTCAACCCT GAGTC
TCAAC AAGGCTCACC CCGTGTTAAA 1680 GTCGCATAA
168
9 (2)INFORMATION FOR SEQ ID NO:15: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 1551 base pa
irs (B)TYPE: nucleic acid (C)STRANDEDNESS: single (D)TOPOLOGY: linear (ii)MOLECULE
TYPE: cD
NA (xi)SEQUENCE DESCRIPTION: SEQ ID NO:15: ATGGGGAAGC CCTTCTTCAC TCTCTCTCTT TCTTCC
CTTT GCTTGCTACT CTTGTCGAGT   60 GCATGCTTTG CTATTACCTC CAGCAAGTTC AACGAGTGCC A
ACTCAACAA CCTCAACGCG   120 TTGGAACCCG ACCACCGCGT TGAGTCCGAA GGTGGTCTTA TTGAAAC
ATG GAACTCTCAA   180 CACCCTGAGC TGCAATGCGC CGGTGTCACT GTTTCCAAAC GCACCCTCAA CCG
CAACGGC   240 TCCCACTTGC CATCTTACTT ACCTTATCCC CAAATGATCA TTGTCGTTCA AGGGAAGGG
A   300 GCAATTGGAT TTGCATTTCC GGGATGTCCC GAGACGTTTG AGAAGCCACA ACAACAATCA   360
AGCAGAAGAG GCTCAAGGTC ACAGCAGCAA CTACAAGACA GTCACCAGAA GATTCGTCAC   420 TTCAA
TGAAG GAGACGTACT AGTGATTCCT CTTGGTGTTC CTTACTGGAC CTATAACACT   480 GGCGATGAAC
CAGTTGTTGC CATCAGTCCT CTTGACACCT CCAACTTCAA CAATCAGCTT   540 GATCAAAACC CCAGAG
TATT TTACCTTGCT GGGAACCCAG ATATAGAGCA TCCCGAGACC   600 ATGCAACAAC AGCAGCAGCA G
AAGAGTCAT GGTGGACGCA AGCAGGGGCA ACACCGACAG   660 CAGGAGGAAG AAGGTGGCAG TGTGCT
CAGT GGCTTCAGCA AACATTTCTT AGCACAATCC   720 TTCAACACCA ACGAGGACAC AGCTGAGAAA C
TTCGGTCTC CAGATGACGA AAGGAAGCAG   780 ATCGTGACAG TGGAGGGAGG CCTCAGCGTT ATCAGC
CCCA AGTGGCAAGA ACAAGAAGAC   840 GAAGACGAAG ACGAAGACGA AGAATATGGA CGGACGCCCT C
TTATCCTCC ACGACGACCA   900 AGCCATGGAA AGCATGAAGA TGACGAGGAC GAGGACGAAG AAGAAG
ATCA ACCTCGTCCT   960 GATCACCCTC CACAGCGACC AAGCAGGCCC GAACAACAAG AACCACGTGG AA
GAGGATGT 1020 CAGACTAGAA ATGGGGTTGA GGAAAATATT TGCACCATGA AGCTTCACGA GAACAT
TGCT 1080 CGCCCTTCAC GTGCTGACTT CTACAACCCA AAAGCTGGTC GCATTAGCAC CCTCAACAGT 1
140 CTCACCCTCC CAGCCCTCCG CCAATTCGGA CTCAGTGCCC AATATGTTGT CCTCTACAGG 1200 AAT
GGAATTT ACTCTCCAGA TTGGAACTTG AACGCGAACA GTGTGACGAT GACTCGAGGG 1260 AAAGGAAG
AG TTGAGTGGT GAACTGCCAA GGGAATGCAG TGTTCGACGG TGAGCTAAGG 1320 AGGGGACAAT TGC
TAGTGGT GCCGCAGAAC CCCGCGGTGG CTGAGCAAGG GGGAGAACAA 1380 GGATTGGAAT ATGTAGTG
TT CAAGACACAC CACAACGCCG TGAGCAGCTA CATTAAGGAT 1440 GTGTTTAGGG TAATCCCTTC GGA
GGTTCTT TCCAATTCTT ACAACCTTGG CCAGAGTCAA 1500 GTGCGTCAGC TCAAGTATCA AGGAAACT
CC GGCCCTTTGG TCAACCCATA A   1551 (2)INFORMATION FOR SEQ ID NO:16: (i)SEQUENCE CH
ARACTERISTICS: (A)LENGTH: 27 base pairs (B)TYPE: nucleic acid (C)STRANDEDNESS: single
(D)T
OPOLOGY: linear (ii)MOLECULE TYPE: DNA (genomic) (xi)SEQUENCE DESCRIPTION: SEQ ID
NO:16:
GCGGCCGCAT GGCCAAGCTA GTTTTTT   27 (2)INFORMATION FOR
SEQ ID NO
:17: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 23 base pairs (B)TYPE: nucleic acid
(C)STRAND
EDNESS: single (D)TOPOLOGY: linear (ii)MOLECULE TYPE: DNA (genomic) (xi)SEQUENCE
DESCRIP
TION: SEQ ID NO:17: GCGGCCGCTG GTGGCGTTTG TGA   23
(2)INFORMAT
ION FOR SEQ ID NO:18: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 23 base pairs (B)TYPE:
nucl
eic acid (C)STRANDEDNESS: single (D)TOPOLOGY: linear (ii)MOLECULE TYPE: DNA (genomic)
(xi)

SEQUENCE DESCRIPTION: SEQ ID NO:18: GCGGCCGCTC TTCTGAGACT CCT

23 (2)INFORMATION FOR SEQ ID NO:19: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 24 base pairs (B)TYPE: nucleic acid (C)STRANDEDNESS: single (D)TOPOLOGY: linear (ii)MOLECULE TYPE: DN
A (genomic) (xi)SEQUENCE DESCRIPTION: SEQ ID NO:19: GCGGCCGCAT GCCCTTCACT CTCT

24 (2)INFORMATION FOR SEQ ID NO:20: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 26 base pairs (B)TYPE: nucleic acid (C)STRANDEDNESS: single (D)TOPOLOGY: linear (ii)MOLECULE TYPE: DN
A (genomic) (xi)SEQUENCE DESCRIPTION: SEQ ID NO:20: GCGGCCGCTG GGAGGGTGAG GCTGTT

26 (2)INFORMATION FOR SEQ ID NO:21: (i)SEQUENCE CHARACTERISTICS: (A)LENGTH: 24 base pairs (B)TYPE: nucleic acid (C)STRANDEDNESS: single (D)TOPOLOGY: linear (ii)MOLECULE TYPE: DN
A (genomic) (xi)SEQUENCE DESCRIPTION: SEQ ID NO:21: GCGGCCGCTG AGCCTTGTTG AGAC

24