(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2010-0130430
(43) 공개일자 2010년12월13일

(51) Int. Cl.
C12N 15/29 (2006.01) *C07K 14/415* (2006.01)
A01H 1/00 (2006.01)
(21) 출원번호 10-2009-0049111
(22) 출원일자 2009년06월03일
심사청구일자 2009년06월03일

(71) 출원인
경희대학교 산학협력단
경기도 용인시 기흥구 서천동 1 경희대학교 국제캠퍼스내
(72) 발명자
박영두
경기도 용인시 수지구 성복동 성동마을 수지자이 1차아파트 108동 702호
이기호
부산광역시 동래구 온천1동 153-8 SK허브스카이 102동 2206호
유재경
경상북도 안동시 노하동 164-1
(74) 대리인
특허법인다울

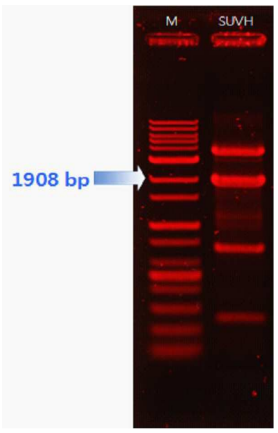
전체 청구항 수 : 총 10 항

(54) 배추 유래의 ＳＵ（ＶＡＲ）３－９ 호몰로그 ２ 단백질 및 이를 코딩하는 유전자

(57) 요 약

본 발명은 배추(*Brassica rapa* ssp. *pekinensis*) 유래의 SU(VAR)3-9 호몰로그(SU(VAR)3-9 homologue) 2 단백질에 관한 것이다. 또한, SU(VAR)3-9 호몰로그 2 단백질 검출용 프라이머를 제작하고 이를 사용하여 SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 유전자를 분리하는 방법을 제공한다.

대 표 도 - 도1



이 발명을 지원한 국가연구개발사업
과제고유번호 200901FHT020711394
부처명 농촌진흥청
연구관리전문기관
연구사업명 바이오그린21사업
연구과제명 유전자 발현 제어시스템 확립 및 이용기술 개발
기여율
주관기관 국립농업과학원
연구기간 2009.01.01~2011.12.31

특허청구의 범위

청구항 1

서열번호 6의 아미노산 서열을 가지는, 배추(*Brassica rapa* ssp. *pekinensis*) 유래의 SU(VAR)3-9 호몰로그(SU(VAR)3-9 homologue) 2 단백질.

청구항 2

제 1항의 SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 핵산 서열로 이루어지는 유전자.

청구항 3

제 2항에 있어서, 상기 핵산 서열은 서열번호 5의 핵산 서열을 가지는 것인 유전자.

청구항 4

제 2항의 유전자를 포함하는 재조합 벡터.

청구항 5

제 4항에 있어서, 상기 재조합 벡터는 도 7에 도시된 pSUVH-RNAi 벡터인 것을 특징으로 하는 재조합 벡터.

청구항 6

제 4항의 재조합 벡터로 형질전환된 형질전환체.

청구항 7

제 6항의 형질전환체를 이용한 형질전환성 배추.

청구항 8

제7항에 있어서, 상기 형질전환성 배추는 왜성 현상이 나타나는 미니 배추인 것을 특징으로 하는 형질전환성 배추.

청구항 9

(i) 서열번호 1 내지 2에 기재된 프라이머를 제작하는 단계; (ii) 배추의 RNA로부터 합성한 cDNA를, 상기 제작된 서열번호 1 내지 2로 이루어지는 프라이머 쌍을 사용하여 RT-PCR을 수행하여 SU(VAR)3-9 호몰로그 2 단백질에 대한 유전자를 분리하는 단계; 및 (iii) 상기 분리된 유전자를 이용하여 SU(VAR)3-9 호몰로그 2의 전체 서열을 확인하는 단계를 포함하는, 배추로부터 SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 유전자를 분리하는 방법.

청구항 10

서열번호 1 내지 2에서 선택되는 SU(VAR)3-9 호몰로그 2 단백질 검출용 프라이머.

명 세 서

발명의 상세한 설명

기 술 분 야

[0001] 본 발명은 배추(*Brassica rapa* ssp. *pekinensis*)의 DNA 및 히스톤의 메틸화 과정에 관여하는 SU(VAR)3-9 호몰로그(SU(VAR)3-9 homologue) 2의 단백질을 배추로부터 분리하는 방법과, 분리된 SU(VAR)3-9 호몰로그(SU(VAR)3-9 homologue) 2 및 이를 코딩하는 유전자에 관한 것이다.

배 경 기 술

[0002] 식물체의 유전자 발현을 조절하기 위해서는 알맞은 메틸화 패턴의 유지가 필요하다. 이를 위해 초기에 메틸화 되어 있지 않은 DNA에 메틸화를 일으키는 과정과 함께 고정된 패턴을 유지하는 형태의 상보적인 시스템이 필요하다. 식물에서 DNA 메틸화는 게놈 DNA의 CpG 또는 CpNpG 염기의 사이토신에 메틸기가 결합하는 것이다. 이 메틸화 과정에 의해서 식물의 발달을 조절하고 바이러스로부터 게놈을 보호하고, 전체 게놈의 순수한 유전 정보를 유지하는 등의 유전자 조절이 가능하게 된다.

[0003] DNA 염기서열 상의 메틸화 외에도 식물체의 발달 과정에 따른 유전자의 전사 여부의 조절에 크로마틴 도멘인, 유크로마틴(euchromatin)과 헤테로크로마틴(heterochromatin)도 영향을 준다. 유크로마틴은 유전자 발현(genes activation)에 관여하고, 헤테로크로마틴은 유전자의 발현 억제(gene silencing)에 관여하는 것으로서, 히스톤의 전사후 변형(post-transcriptional modifications)에 의해 상호 전환이 가능함이 보고되었다(Bannister, A.J., P. Zegerman, J.F. Partridge, E.A. Miska, J.O. Thomas et al., 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410: 120.240.).

[0004] 현재까지의 보고에 따르면, 이와 같은 히스톤 변형에 의한 유전자의 발현 조절은 “히스톤 코드 가설”을 바탕으로 크로마틴의 상태조절에 기인한 것으로 설명되고 있다(Strahl, B. D., and Allis, C. D., 2000. The language of covalent histone modifications, Nature 403, 41-5.; Jenuwein, T. and Allis, C.D. 2001. Translating the histone code. *Science*. 294(5551): 2477). 즉, 히스톤의 말단(N-terminal)에 일시적으로 변형 정보를 저장한다는 것이다. 이러한 히스톤 변형에 관여하는 것이 SET 도메인 단백질이다. SET 도메인은 *Drosophila*에서 발견된 유전자인 *Su(var)3-9*, *E(z)*, *Trithorax*에 공통적으로 나타나는 C-터미널 쪽의 130 아미노산 이하의 모티프(motif)이다(Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., Reuter, G. 1994. The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J.* 13(16): 3822-3831). 이러한 SET 도메인을 가지고 있는 단백질은 크로마틴과 관련하여 유전자의 발현을 촉매 및 억제하는 것으로 보고되었다(Francis N.J. and Kingston R.E. 2001. Mechanisms of transcriptional memory. *Nat. Rev. Mol. Cell. Biol.* 2: 409-421).

[0005] *Arabidopsis*에서는 현재까지 10개의 SET 도메인을 가지는 유전자가 확인되었으며, 이를 *Su(var)3-9* homologue (SUVH)로 명명되었다(Baumbusch L.O., Thorstensen T., Krauss V., Fischer A., Naumann K., Assalkhou R., Schulz I., Reuter G., Aalen R.B., 2001. The Arabidopsis thaliana genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. *Nucleic. Acids. Res.* 29: 4319-4333). 이 SUVH homologous는 담배, 벼 및 옥수수에서도 보고되었다(Baumbusch L.O., Thorstensen T., Krauss V., Fischer A., Naumann K., Assalkhou R., Schulz I., Reuter G., Aalen R.B., 2001. The Arabidopsis thaliana genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. *Nucleic. Acids. Res.* 29: 4319-4333; Shen W.H., 2001, NtSET1, a member of a newly identified subgroup of plant SET-domain-containing proteins, is chromatin-associated and its ectopic overexpression inhibits tobacco plant growth. *The Plant Journal* 28:371-383; Springer M.S., Murphy W.J., Eizirik E., O'Brien S.J., 2003. Placental mammal diversification and the Cretaceous-Tertiary boundary. *Proc Natl Acad Sci USA.* 100(3): 1056-61; the

Plant Chromatin Database, http://www.chromdb.org).

[0006] 최근 보고에 따르면, SUVH2와 SUVH4를 과발현시킨 *Arabidopsis*에서는 직접적으로 후생유전적 변형(epigenetic modification)에 영향을 주어 헤테로크로마틴이 감소하였다(Yu Y., Dong A., Shen W.H., 2004. Molecular characterization of the tobacco SET domain protein NtSET1 unravels its role in histone methylation, chromatin binding, and segregation. *The Plant Journal* 40: 699-711; Naumann K., Fischer A., Hofmann I., 2005. Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in *Arabidopsis*. *EMBO Journal* 24: 1418-1429; Fischer A., Hofman I., Naumann K., Reuter G., 2006. Heterochromatin proteins and the control of heterochromatic gene silencing in *Arabidopsis*. *Journal of Plant Physiology* 163:358-368). 즉, 식물에서의 SUVH는 히스톤 메틸트랜스퍼라아제 능력을 가지고 있어, 헤테로크로마틴을 조절하는 중요한 요소라 할 수 있다.

[0007] 또한 최근에는 SUVH는 TGS(transcriptional gene silencing)와도 직접적인 연관이 있음이 보고되었다. 이것은 4개의 루시페라아제를 도입하여 상호 억제작용(co-suppression)을 의도적으로 발생시킨 형질전환 애기장대에서 SUVH를 과발현 시킬 경우, 도입 유전자의 사일런싱이 증가되어 발현이 잘 되지 못하였고, 전반적인 생육이 늦어져 미니 식물체(mini-plant) 형상을 보인 반면에, SUVH를 억제한 결과 도입 유전자의 상호 억제작용 현상이 없어져서 SUVH가 도입되지 않은 형질전환체에 비하여 루시페라아제의 발현이 2배 정도 더 증가되는 것으로 확인되었다. 이는 LUC 도입 유전자의 메틸화에 SUVH2가 영향을 주었기 때문이다. 또한 SUVH2가 도입되지 않는 LUC 유전자의 경우 메틸화 패턴이 변화되었으나, SUVH2의 과발현 및 억제 상태의 LUC 유전자의 경우, 메틸화 패턴이 후대로도 그대로 전이됨이 확인되었다. 즉, SUVH2는 후생 유전학적으로 세대 진전시, DNA 메틸화의 안전성을 제공하며 유전자의 사일런싱을 조절한다는 것이 확인 되었다 (Naumann K., Fischer A., Hofmann I., Krauss V., Phalke S., Irmler K., Hause G., Aurich A.C., Dorn R., Jenuwein T., Reuter G., 2005. Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in *Arabidopsis*. *EMBO. J.* 24(7): 1418-29).

[0008] 현대 분자 생물학의 발달 및 유용 유전자의 형질전환은 우수한 작물 육종에 필수적이다. 그러나 이러한 형질전환에서 *Agrobacterium tumefaciens*을 이용한 식물 형질전환은, T-DNA는 무작위로 여러 카피의 T-DNA가 대상 작물에 도입되어 1 카피가 아닌 멀티 카피 형질전환체는 사일런싱이 일어나 해당 유전자의 발현이 억제된다. 그러므로 도입 유전자의 안전적인 발현은 식물학 및 원예생명공학에서 주요한 요소로서 확인되고 있다(Hobbs SL, Kpodar P, DeLong CM. 1990. The effect of T-DNA copy number, position and methylation on reporter gene expression in tobacco transformants. *Plant Molecular Biology* 15, 851-864; Holtorf S, Apel K, Bohlmann H. 1995. Comparison of different constitutive and inducible promoters for the overexpression of transgenes in *Arabidopsis thaliana*. *Plant Molecular Biology* 29, 637-646). 그러나 SUVH를 사일런싱시키는 것은, 도입 유전자의 유전자 발현 제어의 내재적인 요인(유전자)를 발굴하는 데 적합하며, 외부요인 투여가 없어 결과해석에 용이할 것으로 판단된다. 즉, 유전자의 효과적인 발현을 확인할 수 있는 형질전환성 식물(competent plant)을 육성할 수 있다. 이러한 형질전환성 식물에 관한 보고는 아직 없으며, 앞으로의 식물 생명공학의 발전에 발맞추어 꼭 필요한 것이라고 본다. 또한 SUVH를 과발현에 의하여 나타나는 식물의 왜화 현상은 배추를 단순히 식용을 위한 김치의 재료 뿐만이 아니라, 디자인적인 가치를 제공하여 하나의 장식용 식물로 만들 수도 있다.

[0009] 그러나 아직 배추에서의 SU(VAR)3-9 호몰로그 (SU(VAR)3-9 homologue; SUVH) 2를 코딩(coding)하는 유전자에 관해서는 보고된 바가 없다.

[0010] 본 발명자들은 배추 유래의 SU(VAR)3-9 호몰로그 2를 코딩하는 유전자를 규명하고자 노력하였다. 또한 상기 유전자를 이용하여 배추에서의 히스톤 전사후 변형(histone post-transcriptional modifications)의 역할 및 이것에 의한 TGS(transcriptional gene silencing)을 규명하고자 하였다. 또한, 상호 억제작용 현상과 상관없이 안정적으로 발현 및 후대로 유전되는 형질전환성(competent) 배추 품종을 육성하여, 식물학 및 육종학적으로 유용한 재료를 공급하고자 한다.

발명의 내용

해결 하고자하는 과제

[0011] 본 발명은 배추 유래의 SU(VAR)3-9 호몰로그 2 단백질 및 이를 코딩하는 유전자를 제공하기 위한 것이다.

[0012] 또한 본 발명은 상기 SU(VAR)3-9 호몰로그 2 단백질 검출용 프라이머 및 이를 이용하여 배추로부터 SU(VAR)3-9 호몰로그 2를 코딩하는 유전자를 분리하는 방법을 제공하기 위한 것이다.

[0013] 또한 본 발명은 배추 유래의 SU(VAR)3-9 호몰로그 2 단백질 코딩 유전자를 이용한 형질전환성 배추를 제공하기 위한 것이다.

과제 해결수단

[0014] 본 발명은 서열번호 6의 아미노산 서열을 가지는, 배추 유래의 SU(VAR)3-9 호몰로그 2 단백질에 관한 것이다.

[0015] 본 발명의 "SU(VAR)3-9 호몰로그 2 단백질" 분리를 위해 프라이머쌍을 제작하여 사용하였다. 제작된 프라이머쌍을 사용하여 배추로부터 SU(VAR)3-9 호몰로그 2 단백질에 해당하는 유전자를 분리하기 위해 RT-PCR을 수행하여 SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 유전자의 전체 서열을 확보하였다.

[0016] 본 발명의 SU(VAR)3-9 호몰로그 2 단백질은 서열번호 6로 기재된 아미노산 서열을 가진다. 또한, 하나 이상의 위치에서 일부 아미노산 서열이 결실, 삽입, 비보전적 또는 보전적 치환 또는 이들의 조합에 의한 변이된 변이체도 본 발명의 범주에 포함된다.

[0017] 또한, 본 발명은 서열번호 6의 아미노산 서열을 가지는 배추 유래의SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 핵산 서열을 포함하는 유전자에 관한 것이다.

[0018] 본 발명의 서열번호 6의 아미노산 서열을 가지는 SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 유전자는, 바람직하게는 서열번호 5의 핵산서열을 가지는 유전자이다.

[0019] SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 핵산 서열은 하나 이상의 염기가 치환, 결실, 삽입 또는 이들의 조합에 의해 변이될 수 있다. 상기 핵산 서열 변이체 또한, 본 발명의 범주에 포함된다.

[0020] 또한, 본 발명은 SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 핵산 서열을 포함하는 유전자를 포함하는 재조합 벡터에 관한 것이다.

[0021] 본 발명에서 "재조합 벡터"란 유전자 삽입물이 발현되도록 작동가능하게 연결된 필수적인 조절 요소를 포함하는 유전자 작제물로, 플라스미드 벡터, 코스미드 벡터, 박테리오파지 벡터 및 바이러스 벡터 등을 포함한 통상의 모든 벡터를 포함한다.

[0022] 본 발명의 벡터는 SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 핵산 서열이 발현조절 서열과 기능적으로 연결되어 있다. 예를 들어, 벡터는 프로모터, 오퍼레이터, 개시코돈, 종결코돈, 폴리아데닐화 시그널, 인핸서 같은 발현 조절 요소 외에도 막 표적화 또는 분비를 위한 신호 서열 또는 리더 서열을 포함하며 목적에 따라 다양하게 제조될 수 있다. 벡터는 선택성 마커를 포함할 수 있다. 또한, 벡터는 자가 복제하거나 숙주 DNA에 통합될 수 있다. 본 발명의 벡터는 당해 기술 분야에서 잘 알려진 유전자 재조합 기술을 이용하여 제조할 수 있으며, 부위-특이적 DNA 절단 및 연결은 당해 기술 분야에서 일반적으로 알려진 효소 등을 사용한다.

[0023] 또한, 본 발명은 SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 핵산 서열을 포함하는 유전자를 포함하는 재조합 벡터로 형질전환된 형질전환체에 관한 것이다. 구체적으로, 본 발명에 의해 SU(VAR)3-9 호몰로그 2 단백질과 이에 대한 유전자의 서열을 이용하여, 히스톤 전사후 변형의 역할 및 이것에 의한 TGS(transcriptional gene silencing) 규명이 가능하고, 이로 인하여, 상호 억제작용 현상과 상관없이 안전적으로 발현되고 및 후대로 유전되는 형질전환성 배추 품종을 육성할 수 있다.

[0024] 숙주세포로의 벡터 도입은 숙주 세포에 적합한 방법을 선택하여 수행할 수 있다. 예를 들어, 전기충격유전자전달법(electroporation), 인산칼슘($CaPO_4$) 침전, 염화칼슘($CaCl_2$) 침전, 실리콘 카바이드 섬유 이용한 교반, 아그로박테리아 매개된 형질전환, 입자총법, 원형질 세포법, 텍스트란 설페이트, 리포펙타민 등의 공지된 방법으로 수행할 수 있다.

[0025] 숙주세포로는 에스케리치아 콜라이(*Escherichia coli*), 바실러스 서브틸리스(*Bacillus subtilis*), 스트렙토마이세스(*Streptomyces*), 슈도모나스(*Pseudomonas*), 프로테우스 미라빌리스(*Proteus mirabilis*) 또는 스타필로코쿠스(*Staphylococcus*)와 같은 원핵 숙주 세포가 있으나, 이로 제한되는 것은 아니다. 또한, 진균(예를 들어, 아스

페르길러스(*Aspergillus*), 효모(예를 들어, 피치아 파스토리스(*Pichia pastoris*), 사카로마이세스 세르비시애(*Saccharomyces cerevisiae*), 쉬조사카로마세스(*Schizosaccharomyces*), 뉴로스포라 크라사(*Neurospora crassa*), 곤충 세포, 식물 세포, 포유동물 등의 진핵 생물 유래의 세포가 있으나, 이로 제한되는 것은 아니다.

[0026] 또한, 본 발명의 SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 유전자는 식물 개체 내로 형질전환 될 수 있다. 식물로의 형질전환법은 바람직하게 아그로박테리움법, 원형질세포법, 입자총법 등이 있다. 형질전환되는 식물은 바람직하게는 배추이다.

[0027] 또한, 본 발명은 서열번호 6의 아미노산 서열을 가지는 SU(VAR)3-9 호몰로그 2 단백질의 세포내 발현 수준을 증가 또는 감소시켜 식물, 바람직하게는 배추의 히스톤 전사후 변형의 역할 및 이것에 의한 TGS(transcriptional gene silencing) 규명이 가능하다. 또한 상호 억제작용 현상과 상관없이 안전적으로 발현되고 후대로 유전되는 형질전환성 배추 품종을 육성할 수 있다.

[0028] 예를 들어, SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 재조합 벡터로 식물, 보다 구체적으로 배추를 형질전환시킴으로써 SU(VAR)3-9 호몰로그 2 단백질의 세포내 발현 수준을 증가시킬 수 있다. SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 재조합 벡터는 상술한 바와 같다. 또한, SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 게놈 유전자의 엑손으로 하나 이상의 외생 핵산 서열을 도입시켜 유전자를 불활성화시켜 SU(VAR)3-9 호몰로그 2 단백질의 세포내 발현 수준을 억제 및 차단시킬 수 있다. 본 발명에서 용어, "외생(Exogenous) 핵산 서열"이란 숙주세포에서 정상적으로 발견되지 않는 핵산 서열을 의미하고, 외생 핵산서열을 유전자 게놈의 엑손으로 삽입함으로써 유전자의 불활성을 유도할 수 있다.

[0029] 상기에서 예시한 방법으로, 목적에 맞게 식물, 바람직하게는 배추의 SU(VAR)3-9 호몰로그 2 단백질의 발현을 효과적으로 조절할 수 있다.

[0030] 또한, 본 발명은 (i) 서열번호 1 내지 2에 기재된 프라이머를 제작하는 단계; (ii) 배추의 RNA로부터 합성한 cDNA를, 상기 제작된 서열번호 1 내지 2로 이루어지는 프라이머 쌍을 사용하여 PCR을 수행하여 SU(VAR)3-9 호몰로그 2 단백질에 대한 유전자를 분리하는 단계; 및 (iii) 분리된 유전자를 이용하여 SU(VAR)3-9 호몰로그 2 유전자의 전체 서열을 확인하는 단계를 포함하는, 배추로부터 SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 유전자를 분리하는 방법에 관한 것이다.

[0031] 각 단계를 각각 살펴보면 다음과 같다.

[0032] 상기 단계 (i) 전에, 식물에 대한 공지된 SU(VAR)3-9 호몰로그 2 단백질을 코딩하는 유전자 단백질의 유전자에 대한 서열을 수집하고 유전자 코딩 서열을 정렬하는 단계와 상기 정렬된 서열 중 상동성이 높은 영역을 선택하는 단계를 먼저 수행할 수 있다. 이 때 식물에 대한 공지된 SU(VAR)3-9 호몰로그 2 단백질의 유전자에 대한 서열은 NCBI GenBank 등으로부터 획득할 수 있으며, 상동성 비교는 육안으로나 비교 프로그램을 이용하여 수행할 수 있다.

[0033] 그 다음에 상기 (i) 단계의 프라이머를 제작하게 되는데, 상기 상동성이 높은 영역에 특이적인 프라이머쌍을 제작하게 된다. 본 발명에서, 주형 가닥 복사를 위한 시작 지점으로 기능을 하는 프라이머의 길이는 특별히 제한되지 않으며, 10 내지 50, 보다 바람직하게는 10 내지 40개의 염기서열을 가진다. 구체적으로, 본 발명에서 사용된 프라이머의 서열은 서열번호 1 내지 2로 기재되었다.

[0034] 단계 (ii)에서, 시료로부터 RNA 분리와 PCR은 일반적인 방법으로 수행할 수 있다. 상기 단계 (iii)에서는 분리된 유전자 단편을 이용하여 RT-PCR을 수행하는 방법을 이용할 수 있는데, 상기 RT-PCR은 공지된 일반적인 방법으로 수행할 수 있다.

[0035] 본 발명의 방법으로 SU(VAR)3-9 호몰로그 2 단백질 분리에 사용될 수 있는 식물의 종류는 특별히 제한되지 않는다. 예를 들어, 식량작물류, 채소작물류, 특용작물류, 과수류, 화훼류 및 사료작물류에 속하는 다양한 식물체를 시료로 사용할 있다.

[0036] 구체적으로, 본 발명에서 SU(VAR)3-9 호몰로그 2 단백질의 분리에 사용된 배추(*Brassica rapa* ssp. *pekinensis*)는 기내의 무균 상태에서 키운 것에서 샘플을 채취하여 RNA를 분리하였다.

[0037] 또한, 본 발명은 서열번호 1 내지 2에서 선택되는 SU(VAR)3-9 호몰로그 2 단백질 검출용 프라이머를 제공한다.

[0038] 본 발명에서 제공하는 프라이머들은 식물의 SU(VAR)3-9 호몰로그 2 단백질 간에 상동성이 높은 영역을 선택적으로 인지하므로, 다양한 식물에서 SU(VAR)3-9 호몰로그 2 단백질의 검출에 효율적으로 사용될 수 있는 프라이머

들이다.

[0039] 서열번호 1 내지 2의 정방향 프라이머 및 역방향 프라이머 쌍은 적절한 완충용액 및 온도에서 중합반응을 위한 효소 및 상이한 4가지 뉴클레오사이드 트리포스페이트의 존재 하에서 cDNA 합성을 개시하게 된다.

[0040] 본 발명의 프라이머 서열은 분광학적, 광화학적, 생화학적, 면역화학적 또는 화학적 수단에 의해 직접적으로 또는 간접적으로 검출 가능한 표지를 포함할 수 있다. 표지로는 호스래디쉬 퍼옥시다제, 알칼린 포스파타아제 등과 같은 효소; ^{32}P과 같은 방사성 동위원소; cy3, cy5 등의 형광성 분자; 바이오틴과 같은 화학 그룹을 예로 들 수 있다.

[0041]

효 과

[0042] 본 발명의 방법으로 SU(VAR)3-9 호몰로그 2 단백질 유전자를 식물체로부터 효과적으로 분리할 수 있다. 또한, 본 발명에 따른 SU(VAR)3-9 호몰로그 2 단백질은 히스톤 전사후 변형의 역할 및 이것에 의한 TGS(transcriptional gene silencing)을 규명할 수 있다. 또한 이를 이용하여 발현억제 벡터(down-regulation vector)를 만들어 식물체 형질전환 시, 상호 억제현상과 상관없이 안전적으로 발현되고 후대로 유전되는 형질전환성 배추를 품종 육성하여, 식물학 및 육종학적으로 유용한 재료를 제공하는 효과를 얻을 수 있다.

발명의 실시를 위한 구체적인 내용

[0043] 이하에서는, 본 발명의 구성을 실시예를 들어 더욱 상세히 설명하지만 본 발명의 권리범위가 하기 실시예로만 한정되는 것은 아니다.

[0044] **실시예 1: SU(VAR)3-9 호몰로그(SU(VAR)3-9 homologue) 2 단백질 유전자를 분리하기 위한 프라이머의 작성**

[0045] SU(VAR)3-9 호몰로그 2 단백질 유전자의 공통된 부분을 찾기 위하여 기존에 발표되어 있는 작물별 SU(VAR)3-9 호몰로그 2 단백질 유전자를 "National Center for Biotechnology Information(NCBI) Genbank"에서 검색하였다. 유전자의 전체 염기서열을 수집하기 위해 Entrez 프로그램을 사용하였다. 검색방법으로는 유전자명으로 SU(VAR)3-9 호몰로그(SU(VAR)3-9 homologue) 2를 검색어로 선정하였다.

[0046] 수집된 SU(VAR)3-9 호몰로그 2 단백질의 유전자 가운데, 모델 식물인 배추과의 애기장대(*Arabidopsis thaliana*)의 염기서열을 선발, 염기서열의 상동성을 많이 보이는 부분을 BLAST 프로그램으로 찾은 뒤 그 부분을 중심으로 프라이머를 제작하였다. 제작된 프라이머들의 서열은 표 1에 기재된 바와 같다.

표 1

[0047]

프라이머	염기서열
SUVH-F (서열번호 1)	5'-ATG AAT ACT CTA TTC GCA TTT CCA-3'
SUVH-R (서열번호 2)	5'-TCA ATT ACA GAT GGC GAG C-3'

[0048] **실시예 2: SU(VAR)3-9 호몰로그 2 단백질의 유전자 분리**

[0049] 배추(*Brassica rapa* ssp. *pekinensis*)로부터 RNA의 분리는 GibcoBRL회사에서 제공된 트리졸(Trizol) 시약을 사용하여 수행하였다. 배추의 잎을 액체질소를 이용하여 파쇄한 다음, 1ml 트리졸(Trizol) 용액을 첨가하고 잘 섞은 후 5분 동안 상온에 두었다. 다음, 200 ㎕의 클로로포름(chloroform)을 첨가하여 섞은 다음 다시 4℃에서 13,000 rpm으로 10분 동안 원심분리한 뒤 그 혼합액의 상등액을 새 튜브로 옮겼다. 이소프로페놀(isopropanol)을 500 ㎕를 넣고 13,000 rpm으로 15분간 원심분리하여 침전시켰다. 다음, 이소프로페놀을 제거하고 침전물을 75% 에탄올로 세척한 다음, 에탄올이 없도록 잘 건조시킨 후 DEPC 처리된 멸균 증류수에 녹여 사용하였다.

[0050] SU(VAR)3-9 호몰로그 2 단백질 유전자의 분리는 확보된 RNA를 주형으로 올리고 dT 프라이머(primer)와 역전사효

소(reverse transcriptase)를 이용, cDNA를 합성한 뒤 작성한 프라이머로 PCR하여 분리하였다. cDNA합성 프로그램 조건은 1 ㎍의 RNA를 올리고 dT 프라이머와 70℃에서 두어 RNA 이차구조를 제거한 뒤 45℃에서 30분간 역전사 시키고 그 사용된 역전사효소를 불활성화 시키기 위해 94℃에서 5분간 두었다. 그 후 합성된 cDNA 1 마이크로그램을 사용, 작성한 프라이머와 태크 중합효소(tag polymerase)로 PCR하였는데 그 프로그램은 94℃에서 2분간 초기 변성 시간을 둔 뒤 94℃에서 30초, 61℃에서 30초, 72℃에서 2분으로 40 반복(cycle) 실행하고 마지막으로 72℃에서 10분간 최종 연장시킨 후 4℃를 유지하여 전체 반응을 종료하였다.

[0051] 각각의 프라이머들로 증폭된 단편들은 그 각각의 염기서열을 분석하기 위해 PCR 산물용 벡터 pGem-T easy 벡터에 클로닝하고, 올바른 유전자가 삽입되었는지 확인하기 위하여 제한효소 인식 서열인 *Eco*RI을 이용하였다. 그 결과는 도 2와 같다. 이후 상기 유전자가 삽입된 자리에 근접한 T7 프로모터와 SP6 프로모터에 결합하는 T7과 SP6 프라이머를 이용하여 염기서열을 분석하였다.

[0052] **실시예 3: SU(VAR)3-9 호몰로그 2 단백질 유전자의 전체 염기서열 확인**

[0053] 염기서열 분석용 pGem-T easy 벡터 내부에 클로닝 된 PCR 단편의 염기서열 분석은 GenBank BLAST 프로그램을 사용하여 이루어졌다. 그 결과 SUVH-F(서열번호 1)와 SUVH-R(서열번호 2) 프라이머로 증폭된 1908bp 산물의 염기서열(도 1)과 아미노산 서열을 분석한 결과, 애기장대(*Arabidopsis thaliana*)의 SU(VAR)3-9 호몰로그 2 유전자와 75%의 상동성을 보였다. 즉, RNA를 대상으로 SUVH-F(서열번호 1)와 SUVH-R(서열번호 2) 프라이머를 이용해 RT-PCR을 수행한 결과 배추의 SU(VAR)3-9 호몰로그 2 유전자의 전체 유전자가 증폭된다는 것을 확인할 수 있었다.

[0054] 또한 NCBI Conserved Domain 프로그램을 이용하여(도 4 참고), 분리한 SU(VAR)3-9 호몰로그 2 유전자에 *Drosophila*에서 발견된 유전자인 *Su(var)3-9*, *E(z)*, *Trithorax*에 공통적으로 나타나는 C-터미널 쪽의 130 아미노산 이하의 모티프인 pre_SET 수퍼패밀리 도메인과 SET 수퍼패밀리 도메인이 각각 1e-20, 2e-17로 정확하게 포함되는 것을 통해, 그 정확도가 높음을 확인하였다. 확인된 SU(VAR)3-9 호몰로그 2 유전자 염기서열을 바탕으로 시작 코돈(codon)과 종료 코돈을 결정한 후, 이를 배추로부터 분리된 전체 SU(VAR)3-9 호몰로그 2 유전자로 결정하고, SU(VAR)3-9 호몰로그 2 단백질의 아미노산 서열은 서열번호 6에 나타내었다. 그리고 상기 단백질에 대한 유전자 서열은 서열번호 5에 나타내었다.

[0055] **실시예 4: SU(VAR)3-9 호몰로그 2 단백질 유전자를 이용한 재조합 벡터 제작**

[0056] 식물발현용 벡터 중 유전자 발현 조절 벡터를 만들기 위하여, 발현 억제 RNA*i* 벡터로는 Gateway® system(invitrogen, USA)을 포함하고 있는 pB7GWIWG2(II) 벡터 (VIB-Ghent University, USA)를 이용하였다. 발현 억제 RNA*i* 벡터로는 효과적인 발현 억제를 위하여, NCBI Conserved Domain 프로그램을 이용한 분석 결과를 바탕으로(도 4 참고) 분리한 SU(VAR)3-9 호몰로그 2 유전자에 *Drosophila*에서 발견된 유전자인 *Su(var)3-9*, *E(z)*, *Trithorax*에 공통적으로 나타나는 C-터미널 쪽의 130 아미노산 이하의 모티프인 pre_SET 수퍼패밀리 도메인과 SET 수퍼패밀리 도메인 부분을 효과적으로 억제할 수 있는 프라이머를 제작하였다. 제작된 프라이머들의 서열은 표 2에 기재된 바와 같다.

표 2

[0057]

프라이머	염기서열
SUVH-i-F (서열번호 3)	5'-CTG TGG TGT TTG AGT GTG GGG TTT-3'
SUVH-i-R (서열번호 4)	5'-CAG CTT ATC GTC CAA CCC ATA GTC C-3'

[0058] 실시예 2에서 염기서열 분석용 pGem-T easy 벡터 내부에 클로닝 된 SU(VAR)3-9 호몰로그 2 단백질 유전자의 염기서열을 작성한 프라이머로 PCR하여 분리하였다. 합성 프로그램 조건은 플라스미드 1 마이크로그램을 사용, 작

성한 프라이머와 태크 중합효소(tag polymerase)로 PCR하였는데, 그 프로그램은 표 2에서 제시한 프라이머를 이용한 발현 억제 RNA*i* 벡터 단편의 경우 94℃에서 2분간 초기 변성 시간을 둔 뒤 94℃에서 30초, 58℃에서 30초, 72℃에서 1분으로 40 반복(cycle) 실행하고 마지막으로 72℃에서 10분간 최종 연장시킨 후 4℃를 유지하여 전체 반응을 종료하였다.

[0059] 그 결과는 도 5와 같이 발현 억제 RNA*i* 벡터 단편의 경우 524 bp를 획득하였다. 이후, 각각의 프라이머들로 증폭된 단편들은 그 각각을 Gateway® system(invitrogen, USA)을 적용하기 위하여, pCR 8/GW/TOPO TA 벡터에 클로닝하고, 올바른 유전자가 삽입되었는지 확인하기 위하여 제한효소 인식 서열인 *Eco*RI을 이용하였으며, 이는 도 6과 같다. 이후 상기 유전자가 삽입된 자리에 근접한 T7 프로모터와 T3 프로모터에 결합하는 T7과 T3 프라이머를 이용하여 염기서열을 분석하였다. 이후 Gateway® LR Clonase™ II enzyme mix (invitrogen, USA)를 이용하여, 발현 억제 RNA*i* vector는 pB7GWIWG2(II) 벡터에 센스 및 안티센스 방향으로 삽입되도록 하였다. 이 재조합 벡터는 pSUVH-RNAi라 명명하였으며, 도 7에 나타내었다. 상기 도 7의 벡터에서, Sm/SpR은 streptomycin/spectinomycin 저항성 Sm/SpR 유전자 이고, LB와 RB는 T-DNA의 좌측 보더와 우측 보더를 나타내며, BAR 는 제초제 저항성 BAR 유전자이고, p35S는 콜리프라워 모자이크 바이러스(CaMV) 35S 프로모터를 나타내는 p35S이며, T35S는 CaMV 35S의 종결자를 나타내는 T35S이고, CmR은 콜로람페니콜(chloramphenicol) 저항성 CmR 유전자이다.

도면의 간단한 설명

[0060] 도 1은 각 SUVH-F와 SUVH-R을 사용한 RT-PCR에 의해 배추로부터 증폭된 SU(VAR)3-9 호몰로그 (SU(VAR)3-9 homologue) 2를 코딩하는 유전자의 단편을 전기영동한 사진이다.

[0061] 도 2는 SUVH-F(서열번호 1)와 SUVH-R(서열번호 2)을 사용한 RT-PCR에 의해 배추로부터 증폭된 SU(VAR)3-9 호몰로그 (SU(VAR)3-9 homologue) 2를 코딩하는 유전자를 pGem-T easy 벡터에 넣은 후 정확한 삽입을 확인하기 위해 *Eco*RI을 처리하여 확인한 사진이다.

[0062] 도 3은 RT-PCR에 의해 분리된 배추의 SU(VAR)3-9 호몰로그 (SU(VAR)3-9 homologue) 2를 코딩하는 유전자의 시퀀스를 GenBank BlastX 프로그램을 사용하여 애기장대(*Arabidopsis thaliana*) 시퀀스 간에 상동성 비교 결과이다.

[0063] 도 4는 RT-PCR에 의해 분리된 배추의 SU(VAR)3-9 호몰로그 (SU(VAR)3-9 homologue) 2를 코딩하는 유전자를 NCBI Conserved Domain 프로그램을 이용하여 분석한 결과이다.

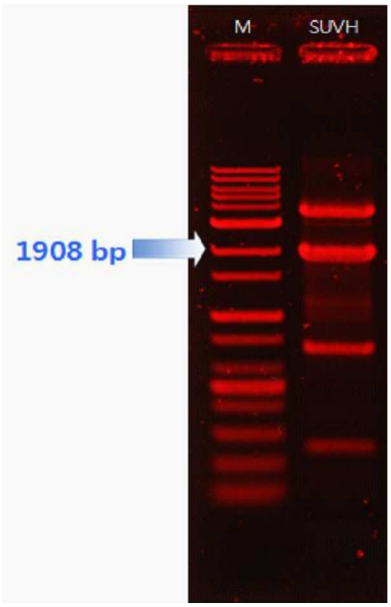
[0064] 도 5은 각 SUVH-i-F(서열번호 3)와 SUVH-i-R(서열번호 4)을 사용한 PCR에 의해 유채로부터 증폭된 SU(VAR)3-9 호몰로그 (SU(VAR)3-9 homologue) 2를 코딩하는 유전자의 RNA*i* 벡터 단편을 전기영동한 사진이다.

[0065] 도 6는 SUVH-i-F(서열번호 3)와 SUVH-i-R(서열번호 4)을 사용한 PCR에 의해 배추의 SU(VAR)3-9 호몰로그 (SU(VAR)3-9 homologue) 2를 코딩하는 유전자를 pCR 8/GW/TOPO TA 벡터에 넣은 후 정확한 삽입을 확인하기 위해 *Eco*RI을 처리하여 확인한 사진이다.

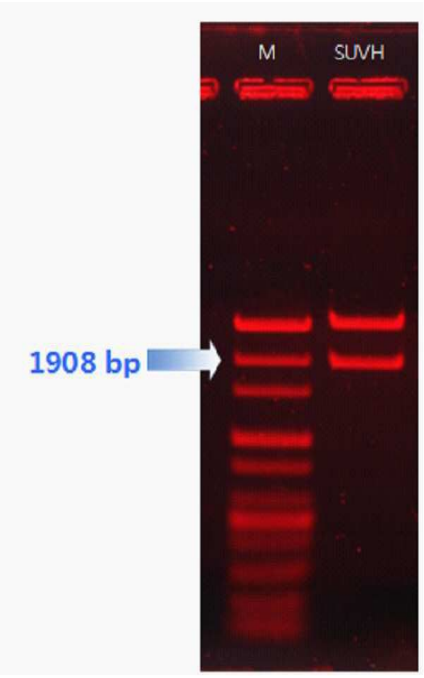
[0066] 도 7은 SU(VAR)3-9 호몰로그 (SU(VAR)3-9 homologue) 2를 코딩하는 유전자를 조절하는 식물발현 억제용 RNA*i* 벡터인 pSUVH-RNA*i* 를 나타내는 모식도이다.

도면

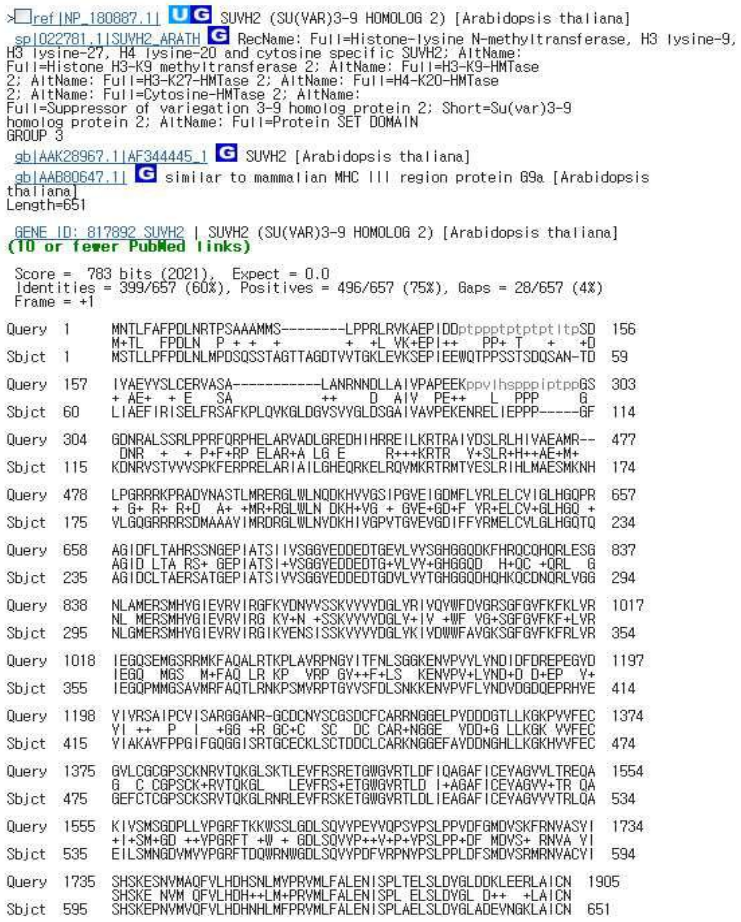
도면1



도면2



도면3



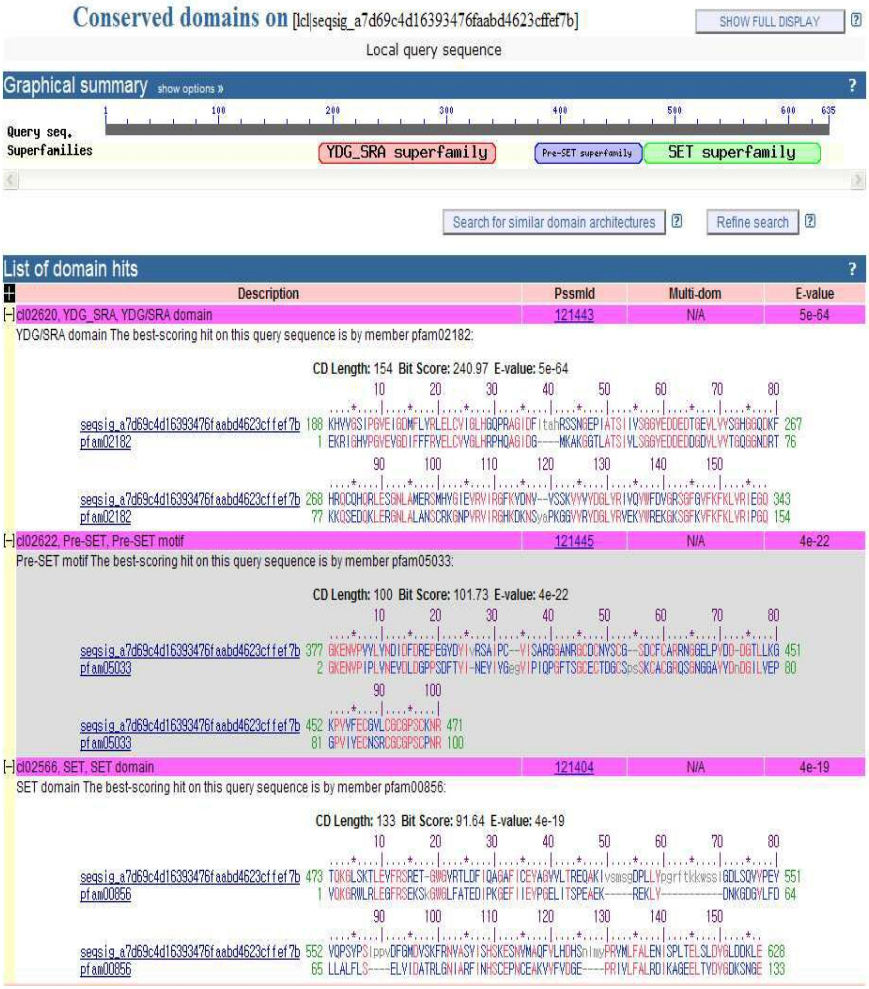
>ref|NP_180887.1| SUVH2 (SU(VAR)3-9 HOMOLOG 2) [Arabidopsis thaliana]
sp|Q22781.1|SUVH2_ARATH RecName: Full=Histone-lysine N-methyltransferase, H3 lysine-9, H3 lysine-27, H4 lysine-20 and cytosine specific SUVH2; AltName: Full=Histone H3-K9 methyltransferase 2; AltName: Full=H3-K9-HMTase 2; AltName: Full=H3-K27-HMTase 2; AltName: Full=H4-K20-HMTase 2; AltName: Full=Cytosine-HMTase 2; AltName: Full=Suppressor of variegation 3-9 homolog protein 2; Short=Su(var)3-9 homolog protein 2; AltName: Full=Protein SET DOMAIN GROUP 3
gb|AAK28967.1|AF344445_1 SUVH2 [Arabidopsis thaliana]
gb|AAB80647.1| similar to mammalian MHC III region protein G9a [Arabidopsis thaliana]
Length=651

GENE_ID: 817892 SUVH2 | SUVH2 (SU(VAR)3-9 HOMOLOG 2) [Arabidopsis thaliana]
(10 or fewer PubMed links)

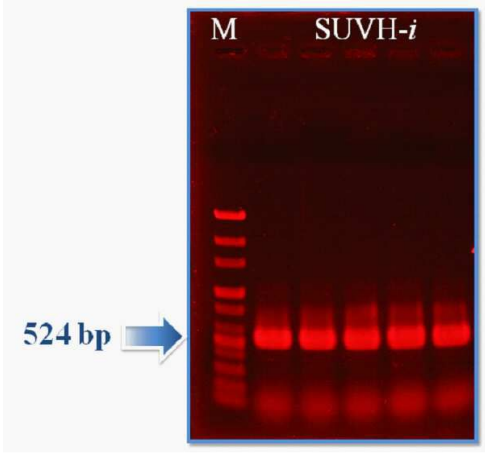
Score = 783 bits (2021), Expect = 0.0
Identities = 399/657 (60%), Positives = 496/657 (75%), Gaps = 28/657 (4%)
Frame = +1

```
Query  1     MNTLFAFPDLNRTPSAAAMMS--------LPPRLRVKAEPIDDptppptptptptitpSD  156
             M+TL  FPDLN  P + +  +          +  +L VK+EPI++   PP+ T   +   +D
Sbjct  1     MSTLLPFPDLNLMPDSQSSTAGTTAGDTVVTGKLEVKSEPIEEWQTPPSSTSDQSAN-TD  59

Query  157   IYAEVYSLCERVASA----------LANRNNDLLAIVPAPEEKppvihspppiptppGS  303
             + AE+  + E   SA              ++   D  AIV  PE++   L  PPP     G
Sbjct  60    LIAEFIRISELFRSAFKPLQVKGLDGVSVVGLDSGAIVAVPEKENRELIEPPP-----GF  114

Query  304   GDNRALSSRLPPRFQRPHELARVADLGREDHIHRREILKRTRAIYDSLRLHIVAEAMR--  477
              DNR  +  + P+F+RP ELAR+A LG E      R+++KRTR  Y+SLR+H++AE+M+
Sbjct  115   KDNRVSTVVVSPKFERPRELARIAILGHEQRKELRQVMKRTRMTYESLRIHLMAESMKNH  174

Query  478   LPGRRRKPRADYNASTLMRERGLWLNQDKHVVGSIPGVEIGDMFLYRLELCVIGLHGQPR  657
             + G+ R+ R+D  A+ +MR+RGLWLN DKH+VG + GVE+GD+F YR+ELCV+GLHGQ +
Sbjct  175   VLGQGRRRRSDMAAAYIMRDRGLWLNYDKHIVGPVTGVEVGDIFFYRMELCVLGLHGQTQ  234

Query  658   AGIDFLTAHRSSNGEPIATSIIVSGGYEDDEDTGEVLVYSGHGGQDKFHRQCQHQRLESG  837
             AGID LTA RS+ GEPIATSI+VSGGYEDDEDTG+VLVY+GHGGQD  H+QC +QRL  G
Sbjct  235   AGIDCLTAERSATGEPIATSIVVSGGYEDDEDTGDVLVYTGHGGQDHQHKQCDNQRLVGG  294

Query  838   NLAMERSMHYGIEVRVIRGFKYDNVYSSKVVYVDGLYRIVQYWFDVGRSGFGVFKFKLVR  1017
             NL MERSMHYGIEVRVIRG KY+N +SSKVVYVDGLY+IV +WF VG+SGFGVFKF+LVR
Sbjct  295   NLGMERSMHYGIEVRVIRGIKYENSISSKVVYVDGLYKIVDWWFAVGKSGFGVFKFRLVR  354

Query  1018  IEGQSEMGSRRMKFAQALRTKPLAVRPNGYITFNLSGGKENVPVYLYNQIDFDREPEGYD  1197
             IEGQ  MGS  M+FAQ LR KP  VRP GY++F+LS  KENVPV+LYNQ+D D+EP  Y+
Sbjct  355   IEGQPMMGSAVMRFAQTLRNKPSMVRPTGYVSFDLSNKKENVPVFLYNDVDGDQEPRHYE  414

Query  1198  YIVRSAIPCVISARGGANR-GCDCNVSCGSDCFCARRNGGELPYDDDGTLLKGKPVVFEC  1374
             YI ++  P  I  +GG +R GC+C  SC  DC CAR+NGGE  YDD+G LLKGK VVFEC
Sbjct  415   YIAKAVFPPGIFGQGGISRTGCECKLSCTDDCLCARKNGGEFAYDDNGHLLKGKHVVFEC  474

Query  1375  GVLCGCGPSCKNRVTQKGLSKTLEVFRSRETGWGVRTLDFIQAGAFICEYAGVVLTREQA  1554
             G  C CGPSCK+RVTQKGL   LEVFRS+ETGWGVRTLD I+AGAFICEYAGVV+TR QA
Sbjct  475   GEFCTCGPSCKSRVTQKGLRNRLEVFRSKETGWGVRTLDLIEAGAFICEYAGVVVTRLQA  534

Query  1555  KIVSMSGDPLLVPGRFTKKWSSLGDLSQVVPEYVQPSVPSLPPVDFGMDVSKFRNVASYI  1734
             +I+SM+GD ++YPGRFT +W + GDLSQVVP++V+P+VPSLPP+DF MDVS+ RNVA YI
Sbjct  535   EILSMNGDVMVYPGRFTDQWRNWGDLSQVVPDFVRPNVPSLPPLDFSMDVSRMRNVACYI  594

Query  1735  SHSKESNVMAQFVLHDHSNLMVPRVMLFALENISPLTELSLDYGLDDKLEERLAICN  1905
             SHSKE NVM QFVLHDH++LM+PRVMLFALENISPL ELSLDYGL D++  +LAICN
Sbjct  595   SHSKEPNVMVQFVLHDHNHLMFPRVMLFALENISPLAELSLDYGLADEVNGKLAICN  651
```

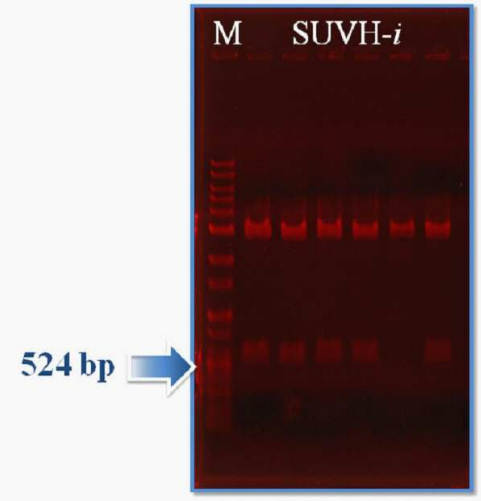
도면4


Conserved domains on [lcl|seqsig_a7d69c4d16393476faabd4623cffef7b]
SHOW FULL DISPLAY
Local query sequence
Graphical summary show options
Query seq.
Superfamilies
YDG_SRA superfamily
Pre-SET superfamily
SET superfamily
Search for similar domain architectures
Refine search
List of domain hits
Description
Pssmid
Multi-dom
E-value
cl02620, YDG_SRA, YDG/SRA domain
121443
N/A
5e-64
YDG/SRA domain The best-scoring hit on this query sequence is by member pfam02182:
CD Length: 154 Bit Score: 240.97 E-value: 5e-64
cl02622, Pre-SET, Pre-SET motif
121445
N/A
4e-22
Pre-SET motif The best-scoring hit on this query sequence is by member pfam05033:
CD Length: 100 Bit Score: 101.73 E-value: 4e-22
cl02566, SET, SET domain
121404
N/A
4e-19
SET domain The best-scoring hit on this query sequence is by member pfam00856:
CD Length: 133 Bit Score: 91.64 E-value: 4e-19

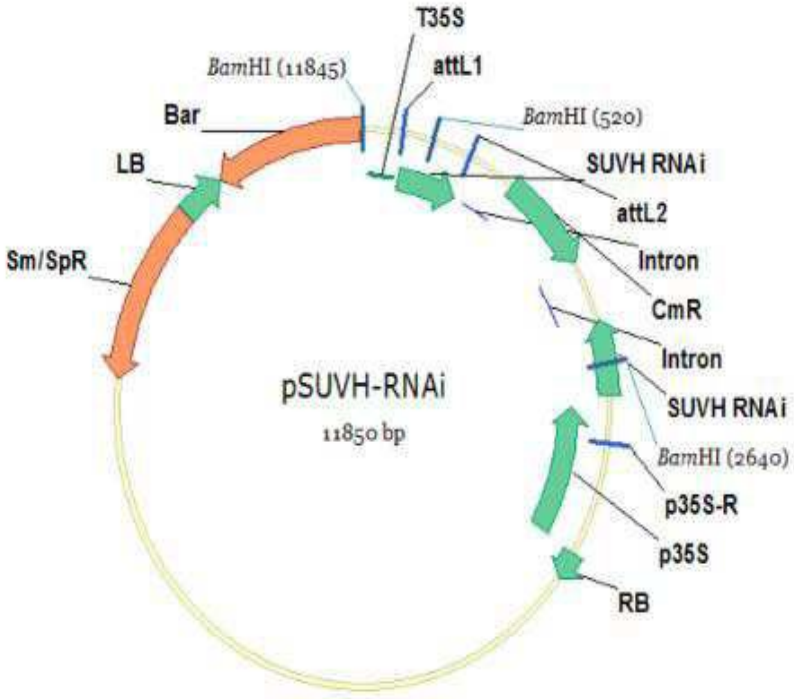
도면5


M
SUVH-i
524 bp

도면6



도면7



서 열 목 록

<110> UNIVERSITY-INDUSTRY COOPERATION GROUP OF KYUNG HEE UNIVERSITY

<120> SU(VAR)3-9 HOMOLOGUE 2 PROTEIN OF Brassica rapa ssp. pekinensis AND A GENE ENCODING THE SAME

<130> P09-140

<160> 6

<170> KopatentIn 1.71

<210> 1
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> SUVH-F Primer

<400> 1
atgaatactc tattcgcatt tcca 24

<210> 2
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> SUVH-R Primer

<400> 2
tcaattacag atggcgagc 19

<210> 3
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> SUVH-i-F Primer

<400> 3
ctgtggtgtt tgagtgtggg gttt 24

<210> 4
<211> 25
<212> DNA

<213> Artificial Sequence

<220>
<223> SUVH-i-R Primer

<400> 4
cagcttatcg tccaacccat agtcc 25

<210> 5
<211> 1908
<212> DNA
<213> Brassica rapa ssp. pekinensis

<220>
<221> CDS
<222> (1)..(1905)

<400> 5
atg aat act cta ttc gca ttt cca gac ctt aac cgc acg ccc tcc gcc 48
Met Asn Thr Leu Phe Ala Phe Pro Asp Leu Asn Arg Thr Pro Ser Ala
1 5 10 15

gcc gcg atg atg tca ctg cct ccg aga cta cgc gtc aag gcc gaa cca 96
Ala Ala Met Met Ser Leu Pro Pro Arg Leu Arg Val Lys Ala Glu Pro
20 25 30

att gac gat cct acc cct cct cct act cct act cca act ccc acc ctc 144
Ile Asp Asp Pro Thr Pro Pro Pro Thr Pro Thr Pro Thr Pro Thr Leu
35 40 45

acc ccc tcc gat atc tac gcc gag tac tac agc ctc tgc gaa cgt gtc 192
Thr Pro Ser Asp Ile Tyr Ala Glu Tyr Tyr Ser Leu Cys Glu Arg Val
50 55 60

gct tcc gca ttg gct aac cgt aac aac gac ctc ctc gcg atc gtc ccc 240
Ala Ser Ala Leu Ala Asn Arg Asn Asn Asp Leu Leu Ala Ile Val Pro
65 70 75 80

gca ccg gaa gaa aaa cct ccc gtc ctc cac tct cct cca ccg atc cca 288
Ala Pro Glu Glu Lys Pro Pro Val Leu His Ser Pro Pro Pro Ile Pro
85 90 95

acc ccg ccg gga tca gga gac aac cgc gct ctc tcc tcc cgc ctc ccg 336

```
Thr Pro Pro Gly Ser Gly Asp Asn Arg Ala Leu Ser Ser Arg Leu Pro
            100                 105                 110

ccg agg ttc cag aga ccg cat gag ctc gct cgc gtc gcc gac cta ggg          384
Pro Arg Phe Gln Arg Pro His Glu Leu Ala Arg Val Ala Asp Leu Gly
        115                 120                 125

cgc gaa gac cat ata cac cgc aga gag atc ctg aaa cgg act cga gcg          432
Arg Glu Asp His Ile His Arg Arg Glu Ile Leu Lys Arg Thr Arg Ala
    130                 135                 140

atc tac gac tcc ctc cgc ctc cat ata gtc gcg gag gcg atg cgg ctc          480
Ile Tyr Asp Ser Leu Arg Leu His Ile Val Ala Glu Ala Met Arg Leu
145                 150                 155                 160

ccc ggt cgg cgg cgt aaa cct aga gcc gat tac aac gcc tcg act ctg          528
Pro Gly Arg Arg Arg Lys Pro Arg Ala Asp Tyr Asn Ala Ser Thr Leu
                165                 170                 175

atg agg gag cgc gga ctg tgg ctg aac cag gac aaa cac gtc gtg ggt          576
Met Arg Glu Arg Gly Leu Trp Leu Asn Gln Asp Lys His Val Val Gly
            180                 185                 190

tcg att cct ggg gtt gag att gga gac atg ttt ctt tac cgt ttg gag          624
Ser Ile Pro Gly Val Glu Ile Gly Asp Met Phe Leu Tyr Arg Leu Glu
        195                 200                 205

ctc tgt gtg att ggc tta cac ggc cag ccg cga gct gga att gat ttt          672
Leu Cys Val Ile Gly Leu His Gly Gln Pro Arg Ala Gly Ile Asp Phe
    210                 215                 220

tta acg gct cac cgg agc tca aac ggt gag ccg ata gcc act agc att          720
Leu Thr Ala His Arg Ser Ser Asn Gly Glu Pro Ile Ala Thr Ser Ile
225                 230                 235                 240

atc gtc tcg ggt ggt tac gag gat gat gag gat aca ggg gag gtt ctt          768
Ile Val Ser Gly Gly Tyr Glu Asp Asp Glu Asp Thr Gly Glu Val Leu
                245                 250                 255

gtt tac tca ggg cac ggt ggg cag gac aaa ttc cac agg cag tgt cag          816
Val Tyr Ser Gly His Gly Gly Gln Asp Lys Phe His Arg Gln Cys Gln
            260                 265                 270

cac cag agg ctg gaa agt ggg aac ctt gcg atg gag agg agt atg cat          864
His Gln Arg Leu Glu Ser Gly Asn Leu Ala Met Glu Arg Ser Met His
```

```
        275                 280                 285

tac ggt atc gaa gtt agg gtg att cgt ggg ttt aag tat gat aat gtg          912
Tyr Gly Ile Glu Val Arg Val Ile Arg Gly Phe Lys Tyr Asp Asn Val
    290                 295                 300

gtt agc tcg aag gtg tat gtg tac gat ggg ttg tat agg att gta cag          960
Val Ser Ser Lys Val Tyr Val Tyr Asp Gly Leu Tyr Arg Ile Val Gln
305                 310                 315                 320

tat tgg ttt gat gtt ggc agg tca ggg ttt ggt gtg ttt aag ttt aag         1008
Tyr Trp Phe Asp Val Gly Arg Ser Gly Phe Gly Val Phe Lys Phe Lys
                325                 330                 335

tta gtt aga atc gag ggg cag agt gag atg ggg agt agg agg atg aag         1056
Leu Val Arg Ile Glu Gly Gln Ser Glu Met Gly Ser Arg Arg Met Lys
            340                 345                 350

ttt gct caa gcg ctt agg act aag ccc ttg gcg gtt agg cct aat ggg         1104
Phe Ala Gln Ala Leu Arg Thr Lys Pro Leu Ala Val Arg Pro Asn Gly
        355                 360                 365

tat ata acg ttt aac ctt tct ggt ggt aag gag aat gtt cct gtt tat         1152
Tyr Ile Thr Phe Asn Leu Ser Gly Gly Lys Glu Asn Val Pro Val Tyr
    370                 375                 380

ctg tat aac gac att gat ttt gat cgt gaa ccg gag ggg tat gac tac         1200
Leu Tyr Asn Asp Ile Asp Phe Asp Arg Glu Pro Glu Gly Tyr Asp Tyr
385                 390                 395                 400

att gtg aga tct gcg att ccc tgt gtg att tct gct cgc gga ggg gct         1248
Ile Val Arg Ser Ala Ile Pro Cys Val Ile Ser Ala Arg Gly Gly Ala
                405                 410                 415

aac agg ggc tgt gat tgt aac tat tcg tgt ggt agt gat tgt ttt tgc         1296
Asn Arg Gly Cys Asp Cys Asn Tyr Ser Cys Gly Ser Asp Cys Phe Cys
            420                 425                 430

gcg agg agg aat ggt ggt gag ctt cct tat gat gac gac ggg act ctg         1344
Ala Arg Arg Asn Gly Gly Glu Leu Pro Tyr Asp Asp Asp Gly Thr Leu
        435                 440                 445

ttg aag ggt aag cct gtg gtg ttt gag tgt ggg gtt ttg tgt ggg tgt         1392
Leu Lys Gly Lys Pro Val Val Phe Glu Cys Gly Val Leu Cys Gly Cys
```

```
    450                 455                 460

ggt ccg agc tgt aag aac cgt gtg act cag aag ggg ttg agt aag acg        1440
Gly Pro Ser Cys Lys Asn Arg Val Thr Gln Lys Gly Leu Ser Lys Thr
465                 470                 475                 480

ctg gag gtt ttc agg tcc agg gaa aca ggt tgg ggt gtt agg aca ttg        1488
Leu Glu Val Phe Arg Ser Arg Glu Thr Gly Trp Gly Val Arg Thr Leu
                485                 490                 495

gat ttc att caa gcc ggg gcc ttc atc tgt gag tat gct ggg gtg gtt        1536
Asp Phe Ile Gln Ala Gly Ala Phe Ile Cys Glu Tyr Ala Gly Val Val
            500                 505                 510

ctc aca agg gag caa gcc aag att gta tca atg agt ggg gat cct ctt        1584
Leu Thr Arg Glu Gln Ala Lys Ile Val Ser Met Ser Gly Asp Pro Leu
        515                 520                 525

ctc tat ccg ggt cgg ttc aca aag aaa tgg agt agt ttg ggt gat tta        1632
Leu Tyr Pro Gly Arg Phe Thr Lys Lys Trp Ser Ser Leu Gly Asp Leu
    530                 535                 540

tct caa gtg tac ccg gag tat gtt cag ccg agt tat cct tct ctg ccg        1680
Ser Gln Val Tyr Pro Glu Tyr Val Gln Pro Ser Tyr Pro Ser Leu Pro
545                 550                 555                 560

ccg gtt gat ttt ggg atg gat gtg tcg aag ttt aga aac gtg gct tct        1728
Pro Val Asp Phe Gly Met Asp Val Ser Lys Phe Arg Asn Val Ala Ser
                565                 570                 575

tac att agt cac agc aaa gaa tca aat gtg atg gcg cag ttt gta ttg        1776
Tyr Ile Ser His Ser Lys Glu Ser Asn Val Met Ala Gln Phe Val Leu
            580                 585                 590

cat gac cat agt aac ttg atg tac cct cgg gtc atg ctt ttt gcc ctg        1824
His Asp His Ser Asn Leu Met Tyr Pro Arg Val Met Leu Phe Ala Leu
        595                 600                 605

gag aat atc tct ccg ttg acc gag cta agc ctg gac tat ggg ttg gac        1872
Glu Asn Ile Ser Pro Leu Thr Glu Leu Ser Leu Asp Tyr Gly Leu Asp
    610                 615                 620

gat aag ctg gaa gag cgg ctc gcc atc tgt aat      tga                   1908
Asp Lys Leu Glu Glu Arg Leu Ala Ile Cys Asn
```

625 630 635

<210> 6
<211> 635
<212> PRT
<213> Brassica rapa ssp. pekinensis

<400> 6
Met Asn Thr Leu Phe Ala Phe Pro Asp Leu Asn Arg Thr Pro Ser Ala
1 5 10 15

Ala Ala Met Met Ser Leu Pro Pro Arg Leu Arg Val Lys Ala Glu Pro
20 25 30

Ile Asp Asp Pro Thr Pro Pro Pro Thr Pro Thr Pro Thr Pro Thr Leu
35 40 45

Thr Pro Ser Asp Ile Tyr Ala Glu Tyr Tyr Ser Leu Cys Glu Arg Val
50 55 60

Ala Ser Ala Leu Ala Asn Arg Asn Asn Asp Leu Leu Ala Ile Val Pro
65 70 75 80

Ala Pro Glu Glu Lys Pro Pro Val Leu His Ser Pro Pro Pro Ile Pro
85 90 95

Thr Pro Pro Gly Ser Gly Asp Asn Arg Ala Leu Ser Ser Arg Leu Pro
100 105 110

Pro Arg Phe Gln Arg Pro His Glu Leu Ala Arg Val Ala Asp Leu Gly
115 120 125

Arg Glu Asp His Ile His Arg Arg Glu Ile Leu Lys Arg Thr Arg Ala
130 135 140

Ile Tyr Asp Ser Leu Arg Leu His Ile Val Ala Glu Ala Met Arg Leu
145 150 155 160

Pro Gly Arg Arg Arg Lys Pro Arg Ala Asp Tyr Asn Ala Ser Thr Leu
165 170 175

Met Arg Glu Arg Gly Leu Trp Leu Asn Gln Asp Lys His Val Val Gly
180 185 190

```
Ser Ile Pro Gly Val Glu Ile Gly Asp Met Phe Leu Tyr Arg Leu Glu
        195                 200                 205

Leu Cys Val Ile Gly Leu His Gly Gln Pro Arg Ala Gly Ile Asp Phe
    210                 215                 220

Leu Thr Ala His Arg Ser Ser Asn Gly Glu Pro Ile Ala Thr Ser Ile
225                 230                 235                 240

Ile Val Ser Gly Gly Tyr Glu Asp Asp Glu Asp Thr Gly Glu Val Leu
                245                 250                 255

Val Tyr Ser Gly His Gly Gly Gln Asp Lys Phe His Arg Gln Cys Gln
            260                 265                 270

His Gln Arg Leu Glu Ser Gly Asn Leu Ala Met Glu Arg Ser Met His
        275                 280                 285

Tyr Gly Ile Glu Val Arg Val Ile Arg Gly Phe Lys Tyr Asp Asn Val
    290                 295                 300

Val Ser Ser Lys Val Tyr Val Tyr Asp Gly Leu Tyr Arg Ile Val Gln
305                 310                 315                 320

Tyr Trp Phe Asp Val Gly Arg Ser Gly Phe Gly Val Phe Lys Phe Lys
                325                 330                 335

Leu Val Arg Ile Glu Gly Gln Ser Glu Met Gly Ser Arg Arg Met Lys
            340                 345                 350

Phe Ala Gln Ala Leu Arg Thr Lys Pro Leu Ala Val Arg Pro Asn Gly
        355                 360                 365

Tyr Ile Thr Phe Asn Leu Ser Gly Gly Lys Glu Asn Val Pro Val Tyr
    370                 375                 380

Leu Tyr Asn Asp Ile Asp Phe Asp Arg Glu Pro Glu Gly Tyr Asp Tyr
385                 390                 395                 400

Ile Val Arg Ser Ala Ile Pro Cys Val Ile Ser Ala Arg Gly Gly Ala
                405                 410                 415

Asn Arg Gly Cys Asp Cys Asn Tyr Ser Cys Gly Ser Asp Cys Phe Cys
```

420 425 430

Ala Arg Arg Asn Gly Gly Glu Leu Pro Tyr Asp Asp Asp Gly Thr Leu
435 440 445

Leu Lys Gly Lys Pro Val Val Phe Glu Cys Gly Val Leu Cys Gly Cys
450 455 460

Gly Pro Ser Cys Lys Asn Arg Val Thr Gln Lys Gly Leu Ser Lys Thr
465 470 475 480

Leu Glu Val Phe Arg Ser Arg Glu Thr Gly Trp Gly Val Arg Thr Leu
485 490 495

Asp Phe Ile Gln Ala Gly Ala Phe Ile Cys Glu Tyr Ala Gly Val Val
500 505 510

Leu Thr Arg Glu Gln Ala Lys Ile Val Ser Met Ser Gly Asp Pro Leu
515 520 525

Leu Tyr Pro Gly Arg Phe Thr Lys Lys Trp Ser Ser Leu Gly Asp Leu
530 535 540

Ser Gln Val Tyr Pro Glu Tyr Val Gln Pro Ser Tyr Pro Ser Leu Pro
545 550 555 560

Pro Val Asp Phe Gly Met Asp Val Ser Lys Phe Arg Asn Val Ala Ser
565 570 575

Tyr Ile Ser His Ser Lys Glu Ser Asn Val Met Ala Gln Phe Val Leu
580 585 590

His Asp His Ser Asn Leu Met Tyr Pro Arg Val Met Leu Phe Ala Leu
595 600 605

Glu Asn Ile Ser Pro Leu Thr Glu Leu Ser Leu Asp Tyr Gly Leu Asp
610 615 620

Asp Lys Leu Glu Glu Arg Leu Ala Ile Cys Asn
625 630 635