



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2025년06월24일

(11) 등록번호 10-2825196

(24) 등록일자 2025년06월20일

(51) 국제특허분류(Int. Cl.)
C12N 9/12 (2006.01) *C12N 15/52* (2006.01)
C12N 15/67 (2006.01)

(52) CPC특허분류
C12N 9/1241 (2013.01)
C12N 15/52 (2013.01)

(21) 출원번호 10-2022-0013246

(22) 출원일자 2022년01월28일

심사청구일자 2022년01월28일

(65) 공개번호 10-2023-0116388

(43) 공개일자 2023년08월04일

(56) 선행기술조사문헌

M. Iveth Garcia. et al, "Functionally redundant control of cardiac hypertrophic signaling by inositol 1,4,5-trisphosphate receptors", J Mol Cell Cardiol. 112: 95-103.(2017.11.)*

US20040003420 A1

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

인천대학교 산학협력단

인천광역시 연수구 갯벌로 27, 인천대학교 이노베이션센터(송도동)

(72) 발명자

박준태

인천광역시 연수구 해송로 143, 101동 703호

박지윤

인천광역시 연수구 아카데미로 119 인천대학교, 제3기숙사 B동 417호

(74) 대리인

차준용

전체 청구항 수 : 총 9 항

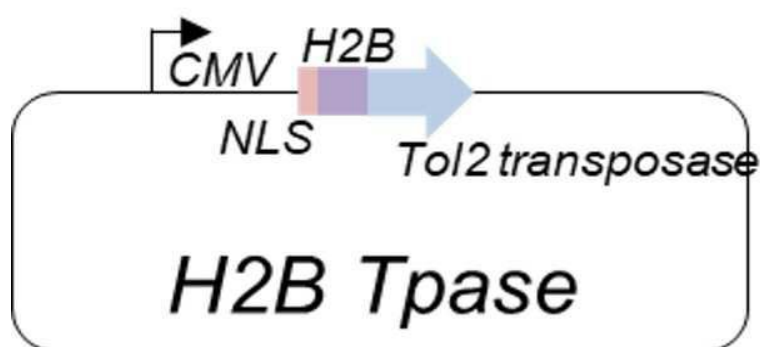
심사관 : 이형곤

(54) 발명의 명칭 신규한 트랜스포사제 및 이를 이용한 트랜스포존 시스템

(57) 요약

본 발명의 일실시예는 히스톤 H2B를 포함하는Tol2 트랜스포사제 발현 벡터 및 Tol2 트랜스포존 시스템을 제공한다. 본 발명의 일실시예에 따른 트랜스포사제 발현 벡터를 포함하는 트랜스포존 시스템은 전이 유전자의 발현량을 향상시킬 수 있다. 또한, 본 발명의 일실시예에 따른 트랜스포존 시스템은 임포틴 알파 KPNA2또는 임포틴 베타 KPNB1를 더 포함할 수 있으며, 상기 임포틴 알파 KPNA2또는 임포틴 베타 KPNB1은 전이 유전자의 발현량을 더욱 증가시킬 수 있다.

대표도 - 도4



(52) CPC특허분류

C12N 15/67 (2013.01)

C12Y 207/07 (2013.01)

C12N 2800/90 (2013.01)

명세서

청구범위

청구항 1

프로모터, Tol2 트랜스포사제(Transposase)를 인코딩하는 염기서열 및 히스톤 H2B를 인코딩하는 염기서열을 포함하고,

상기 프로모터를 기준으로 전사(transcription)되는 방향으로 프로모터, 히스톤 H2B를 인코딩하는 염기서열 및 Tol2 트랜스포사제를 인코딩하는 염기서열이 순서대로 위치하고 있는 것을 특징으로 하는, 트랜스포사제 발현 벡터.

청구항 2

삭제

청구항 3

제 1항에 따른 발현 벡터로부터 발현된 트랜스포사제.

청구항 4

제1항에 따른 트랜스포사제 발현 벡터; 및

Tol2 우측 ITR(Inverted Terminal Repeat), Tol2 좌측 ITR, 프로모터(Promoter) 및 전이 유전자(Transgene)를 포함하는 트랜스포존(Transposon) 벡터;

를 이용하여 세포를 형질전환시키는 단계를 포함하는, 세포에 전이 유전자를 통합하는 방법.

청구항 5

제4항에 있어서,

상기 방법은 임포틴(Importin)을 인코딩하는 염기서열을 포함하는 발현 벡터를 더 포함하는 것을 특징으로 하는 방법.

청구항 6

제5항에 있어서,

상기 임포틴은 임포틴 알파 KPNA2인 것을 특징으로 하는 방법.

청구항 7

제5항에 있어서,

상기 임포틴은 임포틴 베타 KPNB1인 것을 특징으로 하는 방법.

청구항 8

제4항에 있어서,

상기 트랜스포존 벡터 및 트랜스포사제 발현 벡터의 질량비율(트랜스포존 벡터 : 트랜스포사제 발현 벡터)은 1:0.05 내지 1:1인 것을 특징으로 하는 방법.

청구항 9

삭제

청구항 10

제4항에 있어서,

상기 세포는 동물세포인 것을 특징으로 하는 방법.

청구항 11

제4항 내지 제8항 또는 제10항 중 어느 한 항의 방법에 따라서 전이 유전자가 통합된 세포를 배양하는 단계를 포함하는, 세포에서 전이 유전자가 인코딩하는 단백질을 생산하는 방법.

발명의 설명

기술 분야

[0001] 본 발명은 신규한 트랜스포사제 및 이를 이용한 트랜스포존 시스템에 관한 것으로, 더욱 상세하게는 히스톤 H2B를 포함하는 Tol2 트랜스포사제 발현 벡터, 이로부터 발현된 트랜스포사제 및 트랜스포존 시스템에 관한 것이다.

배경 기술

[0002] 유전자 재조합 기술 기반 단백질의 생산기술은 단백질 의약, 식품가공물 등에 널리 이용되고 있다. 특히, 재조합 단백질 생산기술의 경우, 생산하고자 하는 목적 단백질을 인코딩하는 염기 서열을 포함하는 발현 벡터를 대장균, 효모, 곤충 세포, 식물 세포 및 동물 세포 등의 숙주에 도입하고, 이러한 세포주를 적절한 배양 조건으로 배양하여 목적 단백질을 발현시킴으로써 이루어진다. 하지만, 목적 단백질을 효율적으로 생산할 수 있는 숙주를 개발하기 위해서는, 목적으로 하는 단백질마다 생산성이 양호한 숙주 세포를 선정하는 것이 필요하며, 개개의 숙주에서의 목적 단백질 생산 기술에 있어서 한층 더 기술 혁신이 요망되고 있다. 구체적으로 동물 세포 숙주에서는 인간을 비롯한 고등 동물 유래의 단백질에 대하여 인산화, 당쇄 부가 및 폴딩(folding)등의 번역 후 변형(post-translation)을 생체에서 제조되는 것과 보다 동일하게 실시하는 것이 가능하다. 이러한 번역 후 변형은 단백질이 본래 갖는 생리 활성을 재조합 단백질로 재현하기 위해 필수적인 것이며, 그와 같은 생리 활성을 요구하는 의약품 등에는 포유 동물 세포를 숙주로 한 단백질 생산계가 이용되고 있다. 그러나, 동물 세포를 숙주로 한 단백질 발현계의 생산성은 일반적으로 낮고, 도입 유전자의 안정성에도 문제가 있는 경우가 많다. 포유 동물 배양 세포를 숙주로 한 단백질의 생산성 향상은, 치료용 의약품이나 진단약 등의 제조에 있어서 매우 중요할 뿐만 아니라, 이들의 개발 연구에도 많이 기여하고 있다. 그 때문에, 포유 동물 배양 세포, 특히 차이나이즈·헵스타 나소 세포(CHO 세포)를 숙주로 하여 용이하게 고생산주의 획득을 가능하게 하는 유전자 발현계의 개발은 급선무가 되었다.

[0003] 트랜스포존(Transposon)은 염색체의 하나의 유전자좌로부터 별도의 유전자좌로 이동할 수 있는 전위성 유전자이다. 특히, DNA 트랜스포존은 수년간 포유류 세포에서 재조합 단백질 생산을 촉진하기 위해 사용되어 왔다. 이 트랜스포존 시스템은 유전자 치료분야 또는 형질 전환을 위한 비바이러스(non-viral) 전달 도구로 사용되어 왔다. 이 비바이러스성 벡터 시스템은 또한 원래의 DNA를 복제하여 여러 유전자좌에 통합할 수 있는 매우 우수한 능력을 가지고 있다. 트랜스포존 시스템은 "잘라서 붙이기" 메커니즘으로 작동한다. 이 시스템에는 두 가지 구성 요소가 필요하다. 트랜스포사제 부위 및 유전자 전이를 일으키는 트랜스포사제를 운반하는 것은 한 쌍의 역위 말단반복(ITR)이다.

[0004] 주로 사용되는 트랜스포존 시스템은 Sleeping Beauty, PiggyBac, Tol2이다. 특히, Tol2 트랜스포존 시스템은 일본 쌀 물고기(Japanese rice fish)(Oryzias latipes) 또는 메다카로 알려진 물고기에서 유래한 hAT슈퍼 패밀

리에 속하는 Tol2 트랜스포존은 자연에서 발견된 유일한 척추동물 트랜스포존이다. 이 도구는 주로 게놈을 수식하는데 사용되는 도구이며 닭, 마우스 및 인간 세포에서 잘 전치되는 것으로 나타났다. 본 발명자는 신규한 Tol2 트랜스포존 시스템을 발명했다.

발명의 내용

해결하려는 과제

- [0005] 본 발명이 이루고자 하는 기술적 과제는 신규한 트랜스포사제 및 이를 이용한 트랜스포존 시스템을 제공하는 것이고, 구체적으로 목적 단백질의 생산 효율을 높일 수 있는 트랜스포존 시스템을 제공하는 것이다.
- [0006] 본 발명이 이루고자 하는 기술적 과제는 이상에서 언급한 기술적 과제로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재로부터 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0007] 상기 기술적 과제를 달성하기 위하여, 본 발명의 실시예에는 Tol2 트랜스포사제(Transposase)를 인코딩하는 염기서열 및 히스톤 H2B를 인코딩하는 염기서열을 포함하는 트랜스포사제 발현 벡터를 제공한다.
- [0008] 본 발명의 실시예에 있어서, 상기 히스톤 H2B를 인코딩하는 염기서열은 상기 Tol2 트랜스포사제를 인코딩하는 염기서열의 5'방향에 위치할 수 있다.
- [0009] 상기 기술적 과제를 달성하기 위하여, 본 발명의 다른 실시예에는 Tol2 트랜스포사제(Transposase)를 인코딩하는 염기서열 및 히스톤 H2B를 인코딩하는 염기서열을 포함하는 트랜스포사제 발현 벡터로부터 발현된 트랜스포사제를 제공한다.
- [0010] 상기 기술적 과제를 달성하기 위하여, 본 발명의 다른 실시예에는 Tol2 트랜스포사제(Transposase)를 인코딩하는 염기서열 및 히스톤 H2B를 인코딩하는 염기서열을 포함하는 트랜스포사제 발현 벡터; 및
- [0011] Tol2 우측 ITR(Inverted Terminal Repeat), Tol2 좌측 ITR, 프로모터(Promoter) 및 전이 유전자(Transgene)를 포함하는 트랜스포존(Transposon) 벡터;
- [0012] 를 이용하여 세포를 형질전환시키는 단계를 포함하는, 세포에 전이 유전자를 통합하는 방법을 제공한다.
- [0013] 본 발명의 실시예에 있어서, 상기 세포에 전이 유전자를 통합하는 방법은 임포틴(Importin)을 인코딩하는 염기서열을 포함하는 발현 벡터를 더 포함할 수 있다.
- [0014] 본 발명의 실시예에 있어서, 상기 임포틴은 임포틴 알파 KPNA2일 수 있다.
- [0015] 본 발명의 실시예에 있어서, 상기 임포틴은 임포틴 베타 KPNB1일 수 있다.
- [0016] 본 발명의 실시예에 있어서, 상기 트랜스포존 벡터 및 트랜스포사제 발현 벡터의 질량비율(트랜스포존 벡터 : 트랜스포사제 발현 벡터)은 1:0.05 내지 1:1일 수 있다.
- [0017] 본 발명의 실시예에 있어서, 상기 히스톤 H2B를 인코딩하는 염기서열은 상기 Tol2 트랜스포사제를 인코딩하는 염기서열의 5'방향에 위치할 수 있다.
- [0018] 본 발명의 실시예에 있어서, 상기 세포는 동물세포일 수 있다.
- [0019] 상기 기술적 과제를 달성하기 위하여, 본 발명의 또 다른 실시예에는 상기 기술한 세포에 전이 유전자를 통합하는 방법에 따라서 전이 유전자가 통합된 세포를 배양하는 단계를 포함하는, 세포에서 전이 유전자가 인코딩하는 단백질을 생산하는 방법을 제공한다.

발명의 효과

- [0020] 본 발명의 실시예에 따르면, 본 발명의 트랜스포존 시스템을 이용하여 목적 단백질의 생산성을 높일 수 있다.
- [0021] 본 발명의 효과는 상기한 효과로 한정되는 것은 아니며, 본 발명의 상세한 설명 또는 특허청구범위에 기재된 발명의 구성으로부터 추론 가능한 모든 효과를 포함하는 것으로 이해되어야 한다.

도면의 간단한 설명

[0022]

도 1은 CMV프로모터, 전이 유전자로서 루시페라제(Luciferase)를 인코딩하는 염기서열, To2 우측 및 좌측 ITR를 포함하는 트랜스포존 벡터(TP)를 나타낸 그림이다.

도 2는 전이 유전자로서 루시페라제(Luciferase)를 인코딩하는 염기서열, CMV프로모터를 포함하는 비트랜스포존 벡터(non-TP)를 나타낸 그림이다.

도 3은 To12 트랜스포사제를 발현하는 일반적인 To12 트랜스포사제 발현 벡터(Tpase)를 나타낸 그림이다.

도 4는 본 발명에 따른 To12 트랜스포사제 및 히스톤 H2B를 발현하는 트랜스포사제 발현 벡터(H2B Tpase)를 나타낸 그림이다.

도 5는 비트랜스포존(non-TP) 및 트랜스포존(TP)에 대해, 기존 To12 트랜스포사제 벡터(Tpase)와 본 발명에 따른 트랜스포사제 발현 벡터(H2B Tpase)의 양에 따른 형광발현값을 나타낸 그래프이다.

도 6은 단백질 생산성을 측정하기 위해, 본 발명에 따른 트랜스포사제와 일반적인 To12 트랜스포사제를 형광발현값으로 비교하여 나타낸 그래프이다.

도 7은 본 발명에 따른 트랜스포사제 발현 벡터를 포함하는 트랜스포존 시스템(TP+H2B Tpase)의 안정성을 세포 성장률(growth rate)기반으로 나타낸 그래프이다.

도 8은 대조군(Ctrl), 일반적인 To12 트랜스포사제(Tpase) 발현 벡터를 포함하는 트랜스포존 시스템, 본 발명에 따른 트랜스포사제 발현 벡터를 포함하는 트랜스포존 시스템이 세포핵에서 발현되는지 여부를 보여주는 그림이다.

도 9은 대조군(Ctrl), 일반적인 To12 트랜스포사제(Tpase) 발현 벡터를 포함하는 트랜스포존 시스템, 본 발명에 따른 트랜스포사제 발현 벡터를 포함하는 트랜스포존 시스템이 세포핵에서 발현되는지 여부를 수치로 보여주는 그래프이다.

도 10은 본 발명에 따른 트랜스포사제 발현 벡터를 포함하는 트랜스포존 시스템에 임포틴(Importin) 알파 KPN2 및 임포틴 베타 KPNB1을 각각 넣어주었을 때에 대한 형광발현값을 나타낸 그래프이다.

도 11은 임포틴(Importin) 발현에 대한 대조군으로서의 벡터를 나타낸 그림이다.

도 12는 임포틴(Importin) 알파 KPN2를 인코딩하는 염기서열을 포함하는 벡터를 나타낸 그림이다.

도 13은 임포틴(Importin) 베타 KPNB1을 인코딩하는 염기서열을 포함하는 벡터를 나타낸 그림이다.

도 14는 본 발명에 따른 트랜스포사제 발현 벡터를 포함하는 트랜스포존 시스템에 임포틴(Importin) 알파 KPN2 및 임포틴 베타 KPNB1을 각각 넣어주었을 때에, Primer를 이용한 MYC-tag 발현 량 및 GAPDH 발현량을 나타낸 그림이다.

발명을 실시하기 위한 구체적인 내용

[0023]

이하에서는 첨부한 도면을 참조하여 본 발명을 설명하기로 한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 따라서 여기에서 설명하는 실시예로 한정되는 것은 아니다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0024]

명세서 전체에서, 어떤 부분이 다른 부분과 "연결(접속, 접촉, 결합)"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 부재를 사이에 두고 "간접적으로 연결"되어 있는 경우도 포함한다. 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 구비할 수 있다는 것을 의미한다.

[0025]

본 명세서에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

- [0027] 본 발명의 일측면에 따르면, 본 발명의 트랜스포사제(Transposase) 발현 벡터는 Tol2 트랜스포사제를 인코딩하는 염기서열 및 히스톤 H2B를 인코딩하는 염기서열을 포함할 수 있다.
- [0028] 본 발명의 일실시에 따르면, 상기 트랜스포사제 발현 벡터는 트랜스포사제를 인코딩하는 염기서열을 포함하는 벡터일 수 있고, 상기 트랜스포사제 발현 벡터의 염기서열은 서열번호3을 포함할 수 있다. 또한, 상기 트랜스포사제 발현 벡터는 프로모터를 더 포함할 수 있다.
- [0029] 여기서 벡터는 연결된 다른 핵산 분자를 수송할 수 있는 핵산 분자를 지칭할 수 있다. 구체적으로 박테리아, 플라스미드, 파지, 코스미드(cosmid), 에피솜, 바이러스, 및 삽입 가능한 DNA 단편, 즉 동종 재조합에 의해 숙주 세포 계놈에 삽입될 수 있는 단편을 포함할 수 있다. 본 발명의 일실시에 따르면, 상기 벡터는 바람직하게 플라스미드일 수 있다.
- [0030] 본 발명의 일실시에 따르면, 상기 프로모터는 하나 이상의 핵산 서열의 발현을 위해 숙주 세포에 의해 인지되는 핵산 서열이다. 프로모터 서열은 폴리뉴클레오타이드의 발현을 조절하는 전사 제어 서열을 함유한다. 프로모터는 돌연변이체, 절단되고 하이브리드 프로모터를 포함하는 선택된 숙주 세포에서 전사 활성을 보여주는 임의의 핵산 서열일 수 있고 숙주 세포에 상동성 또는 이종성의 세포외 또는 세포내 폴리펩타이드를 암호화하는 유전자로부터 획득될 수 있다. 본 발명에 따른 프로모터는 유도성 및 비-유도성 프로모터를 포함할 수 있다. 핵산 서열은 프로모터의 제어하에 있고 프로모터는 상기 핵산에 대해 이의 기능을 발휘한다. 본 발명의 세포/벡터는 흔히 상기 프로모터를 포함할 수 있다. 본 발명의 일실시에 따르면, 상기 프로모터는 포유동물 세포의 계놈으로부터 유래된 프로모터 또는 포유동물 바이러스로부터 유래된 프로모터일 수 있고, 보다 바람직하게 상기 프로모터는 U6 프로모터, H1 프로모터, CMV(cytomegalo virus) 프로모터, 아데노바이러스 후기 프로모터, 백시니아 바이러스 7.5K 프로모터, SV40 프로모터, HSV의 tk 프로모터, RSV 프로모터, 인간 연장인자 1α(hEF1α) 프로모터, 메탈로티오닌 프로모터, 베타-액틴 프로모터, 인간 IL-2 유전자의 프로모터, 인간 IFN 유전자의 프로모터, 인간 IL-4 유전자의 프로모터, 인간 림포톡신 유전자의 프로모터, 인간 GMCSF 유전자의 프로모터, TERT 프로모터, PSA 프로모터, PSMA 프로모터, CEA 프로모터, E2F 프로모터 AFP 프로모터 또는 알부민 프로모터일 수 있고, 가장 바람직하게는 CMV(cytomegalo virus) 프로모터일 수 있다.
- [0031] 본 발명의 일실시에 따르면, 상기 Tol2 트랜스포사제는 Tol2 트랜스포존(Transposon)시스템에서 전이 유전자(Transgene) 또는 관심 유전자(Gene of Interest)의 양측 말단에 존재하는 역방향 반복 배열(Inverted Terminal Repeat)를 인식 후 결합하여 전이 유전자 또는 관심 유전자를 절단할 수 있다. 본 발명의 일실시에 따르면, 상기 Tol2 트랜스포사제를 인코딩하는 염기서열은 서열번호1을 포함할 수 있다.
- [0032] 본 발명의 일실시에 따르면, 상기 히스톤 H2B는 히스톤(Histone) 단백질 중 하나로서, 상기 히스톤은 염색질(chromatin)을 구성하는 기본단위인 뉴클레오솜(nucleosome)의 중심 단백질이다. 상기 히스톤은 DNA 가닥이 감기는 실패와 같은 역할을 하며 이는 진핵생물의 거대한 유전자가 세포 핵으로 들어가기 위해 필수적인 응축을 할 수 있도록 한다. 또한, 상기 히스톤은 NLSs (nuclear localization sequences)를 포함할 수 있다. 또한, 상기 히스톤 H2B를 인코딩하는 염기서열은 서열번호2를 포함할 수 있다.
- [0033] 도3 및 4를 참고해서 보면, 본 발명의 일실시에 따른 H2B Tol2 트랜스포사제 발현 벡터는 NLS-H2B 서열을 기존의 Tol2 트랜스포사제 발현 벡터(TPase 벡터)의 트랜스포사제 서열 앞에 붙임으로서 발명되었다.
- [0034] 본 발명의 일실시에 따르면, 상기 히스톤 H2B를 인코딩하는 염기서열은 상기 Tol2 트랜스포사제를 인코딩하는 염기서열의 5'방향에 위치할 수 있다.
- [0035] 본 발명의 일측면에 따르면, 본 발명의 트랜스포사제는 본 발명에 따른 발현 벡터로부터 발현될 수 있다. 본 발명의 일실시에 따르면, 본 발명의 트랜스포사제는 Tol2 트랜스포사제에 히스톤 H2B가 융합된 상태일 수 있다. 본 발명의 일실시에 따르면, 상기 트랜스포사제를 인코딩하는 염기서열은 서열번호4를 포함할 수 있다.
- [0036] 도8을 참고해서 보면, 본 발명에 따른 트랜스포사제를 포함하고 있는 세포에서 가장 강한 형광발현이 일어났고 이는 본 발명에 따른 트랜스포사제가 Tol2 트랜스포사제보다 세포핵으로의 신호전달에 있어서 가장 효율적임을 보여주는 것이다. 또한, 도9를 참고해서 보면, 형광발현이 일어난 세포수를 전체 세포수에 대한 백분율로 살펴 보았을 때, 본 발명에 따른 트랜스포사제가 다른 Tol2 트랜스포사제보다 훨씬 많다는 것을 알 수 있으며, 이는 본 발명에 따른 트랜스포사제가 세포내에서 활성을 가지고 작용하고 있다는 것을 보여준다.
- [0037] 본 발명의 다른 일측면에 따르면, 본 발명의 세포에 전이 유전자를 통합하는 방법은 Tol2 트랜스포사제(Transposase)를 인코딩하는 염기서열 및 히스톤 H2B를 인코딩하는 염기서열을 포함하는 트랜스포사제 발현 벡

터; 및 Tol2 우측 ITR(Inverted Terminal Repeat), Tol2 좌측 ITR, 프로모터(Promoter) 및 전이 유전자(Transgene)를 포함하는 트랜스포존(Transposon) 벡터를 이용하여 세포를 형질전환시키는 단계를 포함할 수 있다.

[0038] 본 발명의 일실시에 따르면, 상기 Tol2 우측 ITR(Inverted Terminal Repeat) 및 Tol2 좌측 ITR은 본 발명의 트랜스포사제 발현 벡터에 발현된 트랜스포사제가 인식하는 부위로서, 상기 전이 유전자의 양말단에 위치할 수 있다. 또한, 상기 Tol2 우측 ITR 및 Tol2 좌측 ITR의 염기서열은 각각 서열번호6 및 서열번호7을 포함할 수 있다.

[0039] 본 발명의 일실시에 따르면, 상기 트랜스포존 벡터가 포함하는 프로모터는 포유동물 세포의 계놈으로부터 유래된 프로모터 또는 포유동물 바이러스로부터 유래된 프로모터일 수 있고, 보다 바람직하게 상기 프로모터는 U6 프로모터, H1 프로모터, CMV(cytomegalo virus) 프로모터, 아데노바이러스 후기 프로모터, 벡시니아 바이러스 7.5K 프로모터, SV40 프로모터, HSV의 tk 프로모터, RSV 프로모터, 인간 연장인자 1 α (hEF1 α) 프로모터, 메탈로티오닌 프로모터, 베타-액틴 프로모터, 인간 IL-2 유전자의 프로모터, 인간 IFN 유전자의 프로모터, 인간 IL-4 유전자의 프로모터, 인간 림포톡신 유전자의 프로모터, 인간 GMCSF 유전자의 프로모터, TERT 프로모터, PSA 프로모터, PSMA 프로모터, CEA 프로모터, E2F 프로모터 AFP 프로모터 또는 알부민 프로모터일 수 있고, 가장 바람직하게는 CMV(cytomegalo virus) 프로모터일 수 있다. 본 발명의 일실시에 따르면, 상기 프로모터는 전이 유전자의 5'방향에 위치할 수 있다.

[0040] 상기 전이 유전자(Transgene)는 관심 유전자(Gene of Interest) 또는 목적 유전자로 지칭될 수 있고, 전이 유전자 단백질을 암호화하는 DNA와 관련하여 RNA 전사 개시 신호, 프로모터 또는 인핸서와 같은 전사되지 않는 플랜킹 영역을 포함하지 않는다. 전이 유전자는 세포 내에서 목적 단백질을 생산하기 위하여 도입될 수 있으며 상기 관심 단백질을 인코딩하는 유전자를 의미하고, 기능성 RNA를 암호화하는 DNA 서열을 포함할 수 있다. 따라서, 전이 유전자는 형질감염을 통한 포유동물 숙주 세포와 같은 세포로 도입되고 목적하는 생성물(전이 유전자 발현 생성물, 예를 들어, 이종성 단백질을) 암호화하는 DNA 서열일 수 있다. 관심 유전자는 소포체 및/또는 세포질 막에 걸친 전좌 및/또는 분비를 매개하고/하거나 촉진시키는 신호 펩타이드를 암호화하고 분비 전 또는 분비 동안에 제거되는 신호 펩타이드 암호화 서열에 기능적으로 부착될 수 있다.

[0041] 여기서 사용된 용어, 전이(Transposition)는 하나의 위치에서 트랜스포존의 이동 또는 복제와 다른 위치로의 삽입을 포함하는 복잡한 유전적 재배열(genetic rearrangement) 과정을 지칭하며, 이는 보통 계놈 내, 또는 계놈 간, DNA 구조물(예컨대 플라스미드, 박미드(bacmid), 및 코스미드)간, 또는 계놈과 DNA 구조물 간에서 발생한다. 본 발명의 일실시에 따르면, 전이는 계놈과 DNA 구조물 사이에서 발생할 수 있다.

[0042] 본 발명의 일실시에 따르면, 상기 트랜스포존 벡터의 염기서열은 서열번호5를 포함할 수 있다.

[0043] 본 발명의 일실시에 따르면, 상기 세포에 전이 유전자를 통합하는 방법은 임포틴(Importin)을 인코딩하는 염기서열을 포함하는 발현 벡터를 더 포함할 수 있다. 상기 임포틴은 단백질을 세포질에서 세포핵으로 이동시킬 수 있는 단백질로서, Nuclear Localization Sequences (NLS)라고 불리는 특정 인식부위에 결합할 수 있다. 상기 임포틴은 임포틴 알파(importin α) KPNA2 및 임포틴 베타(importin β KPNB1 라는 2종류의 서브유닛을 포함할 수 있다. 임포틴 알파 KPNA2는 이동시킬 단백질에 존재하는 NLS에 결합하는 어댑터(adaptor) 단백질이고, 임포틴 베타 KPNB1는 이동시킬 단백질에 그 자체가 결합하여 단백질을 이동시키거나, 또는 임포틴 알파 KPNA2와 결합하여 이종이합체(heterodimer)를 구성할 수 있다. 본 발명의 다른 일실시에 따르면, 상기 임포틴은 바람직하게 임포틴 알파 KPNA2 또는 임포틴 베타 KPNB1일 수 있다. 도10을 참고해서 보면, 본 발명의 세포에 전이 유전자를 통합하는 방법에 있어서, 임포틴 베타 KPNB1를 추가로 처리해준 경우가 가장 뛰어난 전이 유전자 발현을 나타냈고 그 다음으로 임포틴 알파 KPNA2로 처리해준 경우였고, 마지막으로 추가 임포틴 처리가 없는 경우였다. 또한, 도14를 참고해서 보면, 본 발명의 세포에 전이 유전자를 통합하는 방법에 있어서 임포틴 알파 KPNA2 또는 임포틴 베타 KPNB1를 처리해 준 경우, 각각 제대로 발현되었기에 트랜스포존 시스템에서 정상적으로 활성을 가지고 있음을 보여주었다.

[0044] 본 발명의 일실시에 따르면, 상기 트랜스포존 벡터 및 트랜스포사제 발현 벡터의 질량비율(트랜스포존 벡터 : 트랜스포사제 발현 벡터)은 1:0.05 내지 1:1일 수 있다. 상기 질량비율이 1:1을 초과하는 경우, 즉, 트랜스포존 벡터에 비해 과도한 트랜스포사제 발현 벡터가 투입되면, 트랜스포사제-트랜스포존 복합체 생성을 방해할 수 있다. 반대로, 상기 질량비율이 1:0.05 미만일 경우, 즉, 트랜스포존 벡터에 비해 트랜스포사제 발현 벡터가 지나치게 덜 투입되면 트랜스포사제-트랜스포존 복합체 생성이 일어나지 않을 수 있다. 본 발명의 다른 일실시에 따르면, 상기 질량비율은 1:0.05 내지 1:0.5일 수 있다. 본 발명의 또다른 일실시에 따르면, 상기 질량비율은 1:0.08 내지 1:0.3일 수 있으며, 1:0.09 내지 1:0.2일 수 있다. 바람직하게는 1:0.1일 수 있다.

- [0045] 도5를 참고해서 보면, 트랜스포존 벡터(TP) 및 Tol2 트랜스포사제 벡터 비율이 10:3인 경우가 비-트랜스포존 값과 비교해서, 루시퍼라제 발현을 증가시켰다. 하지만, 본 발명에 따른 발현 벡터로부터 발현된 Tol2 H2B 트랜스포사제 벡터 및 트랜스포존 벡터가 다른 경우들보다 월등히 루시퍼라제 발현을 증가시켰다. 이는 본 발명에 따른 발현 벡터로부터 발현된 Tol2 H2B 트랜스포사제가 단백질 생산성을 증가시키는 데 상당히 효과적이라는 것을 의미한다.
- [0046] 본 발명의 일실시에 따르면, 상기 세포는 동물세포일 수 있다. 바람직하게는 상기 세포는 포유동물 세포일 수 있으며, 상기 포유동물 세포는 COS 세포, CHO 세포, VERO 세포, MDCK 세포, WI38 세포, V79 세포, B14AF28-G3 세포, BHK 세포, HaK 세포, NSO 세포, SP2/0-Ag14 세포, HeLa 세포, HEK293 세포 및 PER.C6 세포로 이루어진 군으로부터 선택되는 하나일 수 있다. 바람직하게, 상기 세포는 CHO세포일 수 있다.
- [0047] 본 발명의 일실시에 따르면, 상기 트랜스포존 벡터는 항생제 내성 유전자를 더 포함할 수 있다. 항생제 내성 유전자는 마커 유전자로 사용되고, 당업계에서 통상적으로 이용되는 항생제 내성 유전자를 포함하며, 예를 들어 암피실린, 겐타마이신, 카베니실린, 클로람페니콜, 스트렙토마이신, 카나마이신, 제네티신, 네오마이신 및 테트라사이클린에 대한 내성유전자가 있을 수 있다.
- [0048] 본 발명의 일실시에 따르면, 상기 트랜스포존 벡터는 전이 유전자 및 항생제 내성 유전자사이에 추가 프로모터를 더 포함할 수 있다. 상기 추가 프로모터는 포유동물 세포의 게놈으로부터 유래된 프로모터 또는 포유동물 바이러스로부터 유래된 프로모터일 수 있고, 보다 바람직하게 상기 프로모터는 U6 프로모터, H1 프로모터, CMV(cytomegalo virus) 프로모터, 아데노바이러스 후기 프로모터, 벡시니아 바이러스 7.5K 프로모터, SV40 프로모터, HSV의 tk 프로모터, RSV 프로모터, 인간 연장인자 1 α (hEF1 α) 프로모터, 메탈로티오닌 프로모터, 베타-액틴 프로모터, 인간 IL-2 유전자의 프로모터, 인간 IFN 유전자의 프로모터, 인간 IL-4 유전자의 프로모터, 인간 림프톡신 유전자의 프로모터, 인간 GMCSF 유전자의 프로모터, TERT 프로모터, PSA 프로모터, PSMA 프로모터, CEA 프로모터, E2F 프로모터 AFP 프로모터 또는 알부민 프로모터일 수 있고, 가장 바람직하게는 SV40 프로모터일 수 있다.
- [0049] 본 발명의 또다른 일측면에 따르면, 본 발명의 세포에서 전이 유전자가 인코딩하는 단백질을 생산하는 방법은, 상기 세포에 전이 유전자를 통합하는 방법에 따라 전이 유전자가 통합된 세포를 배양하는 단계를 포함할 수 있다.
- [0050] 본 발명의 일실시에 따르면, 상기 단백질은 상기 전이 유전자에 의하여 인코딩되고, 상기 전이 유전자는 세포 내에서 전사 및 번역 단계를 거쳐 단백질을 생산할 수 있다.
- [0051] 본 발명의 일실시에 따르면, 상기 단백질을 생산하는 방법은 형질감염된 세포에 메틸화 억제제(methylationinhibitor) 또는 히스톤 디아세틸레이즈 억제제(histone deacetylase inhibitor)를 처리하는 단계를 더 포함할 수 있다. 상기 메틸화 억제제로의 처리는 풀림이 유지될 때 DNA 축합을 억제하고, 따라서 공동-형질감염된 트랜스포사제 벡터의 발현 수준이 증가하였고, 숙주 게놈 내부의 컷 및 페이스트로 인해 전이유전자가 더욱 안정적으로 통합되고 발현될 수 있다. 상기 메틸화 억제제는 5-아자시티딘(azacytidine), 5-아자-2'-데옥시시티딘, 제볼라린, L-메티오닌, 아피시딘, 히드랄라진 및 프로카인아미드로 구성된 군으로부터 선택된 어느 하나 이상일 수 있다. 또한, 히스톤 디아세틸레이즈 억제제는 TSA(trichostatin A), NaBu(sodium butyrate), VAP(valproic acid) 및 SAHA으로 구성된 군으로부터 선택된 어느 하나 이상일 수 있다.
- [0052] 도6을 참고해서 보면, 본 발명에 따른 트랜스포사제가 다른 Tol2 트랜스포사제보다 적은 양으로도 더욱 높은 단백질 생산성으로 보여준다는 것을 알 수 있다. 또한, 도7을 참고해서 보면, 본 발명에 따른 트랜스포사제는 세포 성장률 기반의 안정성 시험에서 다른 Tol2 트랜스포사제보다 훨씬 높은 성장률을 보여주어 상당히 높은 안정성을 가지고 있다.
- [0054] 이하 첨부된 도면을 참고하여 본 발명의 실시예를 상세히 설명하기로 한다.
- [0055] **실시예1. 세포 배양**
- [0056] CHO (Chinese hamster ovary) DG44 세포들 (A1100001; Thermo Fisher Scientific, Waltham, MA, USA)이 본 연구에서 사용되었다. 세포들은 100 U/ml 페니실린(penicillin) 및 100 μ g/ml 스트렙토마이신(streptomycin) (SV30079.01; Hyclone)과 함께 10% 소태아혈청(fetal bovine serum), 10 mM 소듐 히포산틴(sodium hypoxanthine), 1.6 mM 티미딘(thymidine) (2068642; Gibco, Waltham, MA, USA)이 보충된 25 mM 글루코스(glucose)와 함께 DMEM 배지(Dulbecco's modified Eagle's medium Cells)에서 배양되었다. 세포들은 5% CO₂가

보완된 외기(ambient air)(20% O₂)에서 배양되었다. 세포들은 250 mL 삼각 플라스크(Erlenmeyer flask) (73250; SPL, Pocheon, Korea)에서 40 mL의 작업량(working volume)으로 배양되었고, 150 RPM 교반하에서 인큐베이팅 되었고, 세포 배지를 바꿔줄 때 계대되었다. 세포 밀도(density) 및 생존능력(viability)은 Cedex HiRes Analyzer (05650216001; Roche, Basel, Switzerland)로 측정되었다.

[0058] 실시예2. 플라스미드(Plasmid) 디자인 및 구축

[0059] 제조합 플라스미드를 구축하기 위해서 Infusion cloning 기술을 사용했다. Tol2 트랜스포사제(transposase) 벡터는 Tol2 트랜스포사제를 인코딩하는 cDNA를 pcDNA6 벡터(V22020; Invitrogen, Carlsbad, CA, USA)에 삽입함으로써 만들어졌다. H2B Tol2 트랜스포사제는 이전에 만들어 놓은 Tol2 트랜스포사제 벡터를 PmeI (R0560s; NEB, Ipswich, MA, USA) 효소로 잘라내서 H2B 염기서열을 삽입함으로써 만들었다. TP 벡터는 Tol2 우측 및 좌측 ITRs가 CMV 프로모터 기반 루시퍼라제 유전자 옆에 위치하게끔 디자인되었다. 반면, non-TP벡터는 CMV 프로모터 기반 루시퍼라제 유전자만 포함되도록 디자인되었다. KPNA2 및 KPNB1은 각각 infusion cloning 기술을 통해 pcDNA3.1 벡터에 삽입되었다. EcoRI (R0101s; NEB)가 제한효소로서 사용되었다.

표 1

프라이머	염기서열(5' 에서 3')	프라이머	염기서열(5' 에서 3')
임포틴 알파 Forward	GCTCTACAAAACCAT (서열번호8)	임포틴 알파 Reverse	CCAGATCCTCTTCTG (서열번호9)
임포틴 베타 Forward	CAGCATTTGGGAAGG (서열번호10)	임포틴 베타 Reverse	TGCCAGATCCTCTTC (서열번호11)

[0062] 실시예3. 형질감염(Transfection)

[0063] CHO DG44 세포들은 제조사의 지시에 따라 SG 세포주 4D-Nucleofector™ X Kit L (V4XC-3024; Lonza, Basel, Switzerland)을 사용해서 형질감염되었다. 형질감염된 세포들은 10mg/mL 블라스티시딘(Blasticidin) (ant-bl-05; Invitrogen)을 사용해서 2주동안 선별되었다. 선별과정동안, 배지는 2일마다 바꿔주었다. 선별후, 단일세포들은 96-well 플레이트(353072; Falcon, Franklin Lakes, NJ, USA)에 150 µl 배지로 well당 seeding되었고 15 일동안 인큐베이팅되었다.

[0065] 실시예4. 루시퍼라제(Luciferase) 활성 측정

[0066] 루시퍼라제 활성은 루시퍼라제 검사 시스템(luciferase assay system) (E1500; Promega, Madison, Wisconsin, USA)을 사용해서 분석되었다. 세포들(2×10^6)은 $200 \times g$ 로 2분동안 원심분리되었고 PBS(10010023; Thermo Fisher Scientific)로 2번 세척되었다. 세포들은 100 µL의 1× 세포 배양 용해 시약(cell culture lysis reagent) 및 100 µL의 PBS (10010023; Thermo Fisher Scientific)에서 와류혼합(verotexed)시켜주었다. 그런 다음, 100 µL의 용해된 샘플은 96-well 플레이트(30396; SPL, Pocheon, Kyunggi, Korea)의 각 well에 형질주입되었고, 100 µL의 루시퍼라제 검사 시약II가 각 well에 추가되었다. 루시퍼라제 활성은 VICTOR multi-label plate reader (2030-0050; PerkinElmer, Waltham, MA, USA)를 사용해서 형광세기를 측정함으로써 구해졌다.

[0068] 실시예5. 형광 측정(Immunofluorescence)

[0069] 형광측정을 위해서, 세포들은 세포 배양 슬라이드 4 well (30114; SPL)에 상온에서 15분동안 PBS에 4 % 파라포름알데하이드(paraformaldehyde)로 고정되었고 그런 다음 15분동안 PBS에 0.1 % 트리톤X-100(Triton X-100)으로 투과되도록 처리되었다(permeabilized). 상온에서 1시간동안 PBS에 2.5 % FBS로 차단을 수행하였다. 샘플들은 PBS에 2.5 % FBS 로 희석된 1차 항체와 함께 4 ° C 에서 하룻밤동안(overnight) 인큐베이팅되었다. Myc-Tag (71D10) 토끼 단일클론 항체(Rabbit mAb) (2287s, 1:200 희석배율; Cell signaling Technology, Danvers, Massachusetts, USA)와 함께 인큐베이션한 후, 세포들은 얼음정도 차가운 PBS로 3번 세척되었고, 알렉사 플루오르(Alexa Fluor) 594-결합 항-마우스항체(A11005; 1:200 dilution; Invitrogen)로 1시간동안 상온에서 인큐베이션되었다. 세포핵들은 Hoechst 33342 (H3570; Invitrogen)으로 염색되었고 샘플들은 Dako Fluorescence Mounting 배지 (S3023; Dako)를 사용해서 마운트(mount)되었다. 이미지들은 Carl Zeiss Axio Imager Z1 현미경을 사용해서 캡처되었다.

[0071] 전술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해

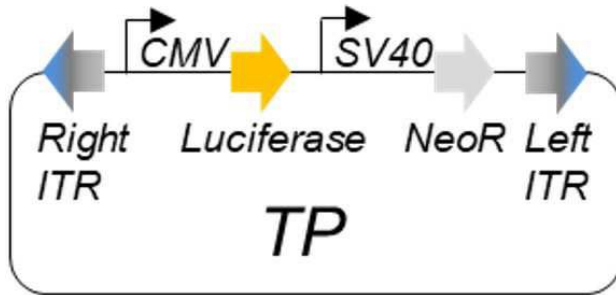
할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.

[0072]

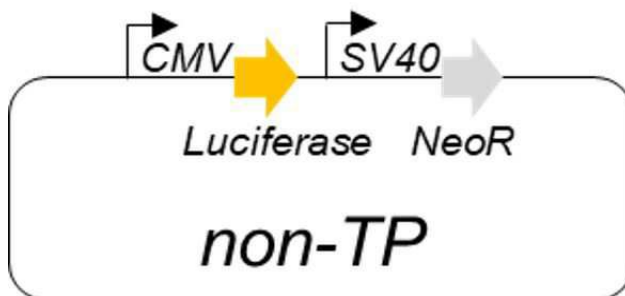
본 발명의 범위는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

도면

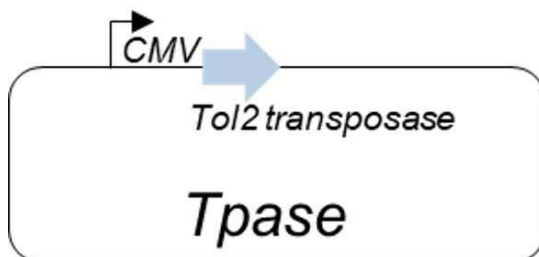
도면1



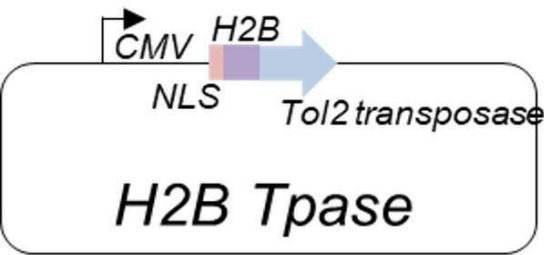
도면2



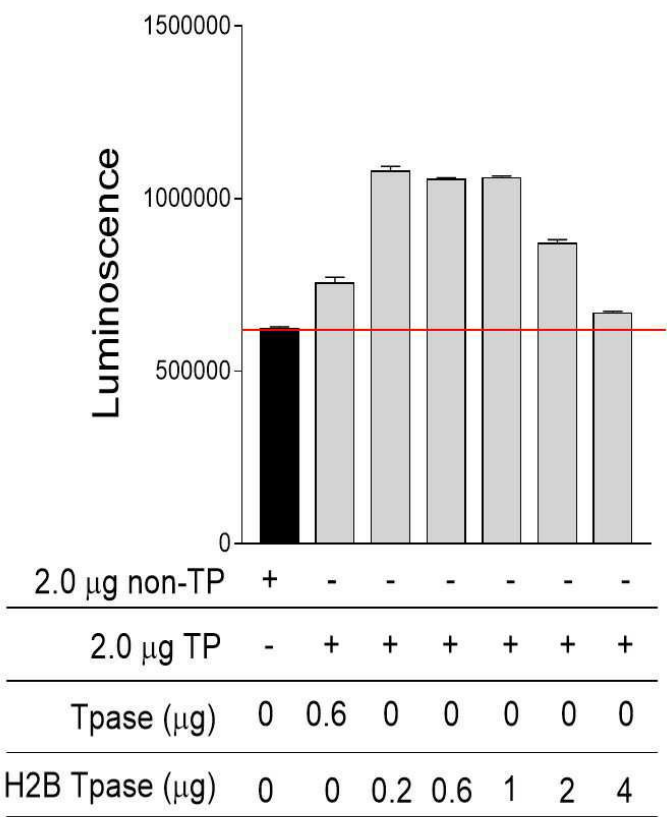
도면3



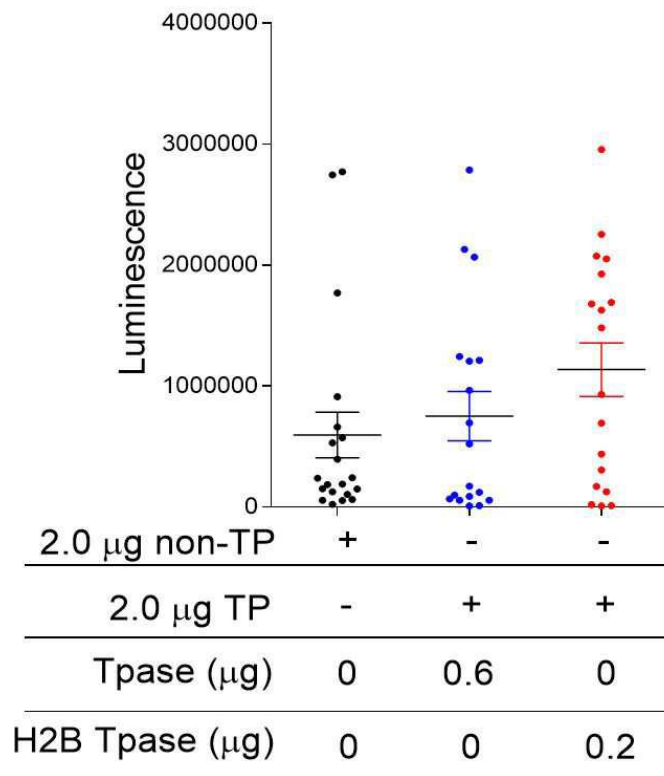
도면4



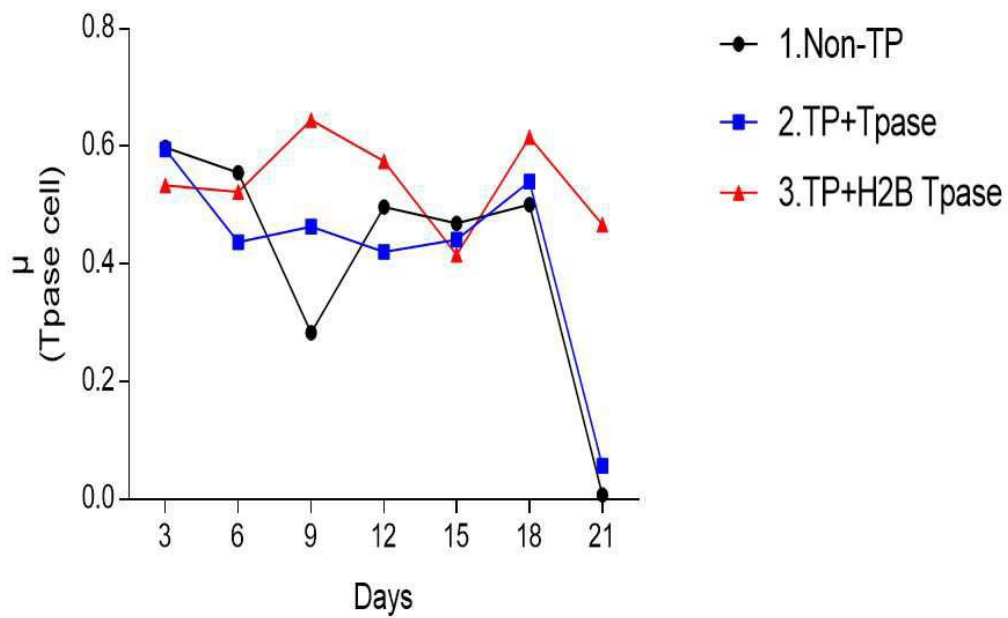
도면5



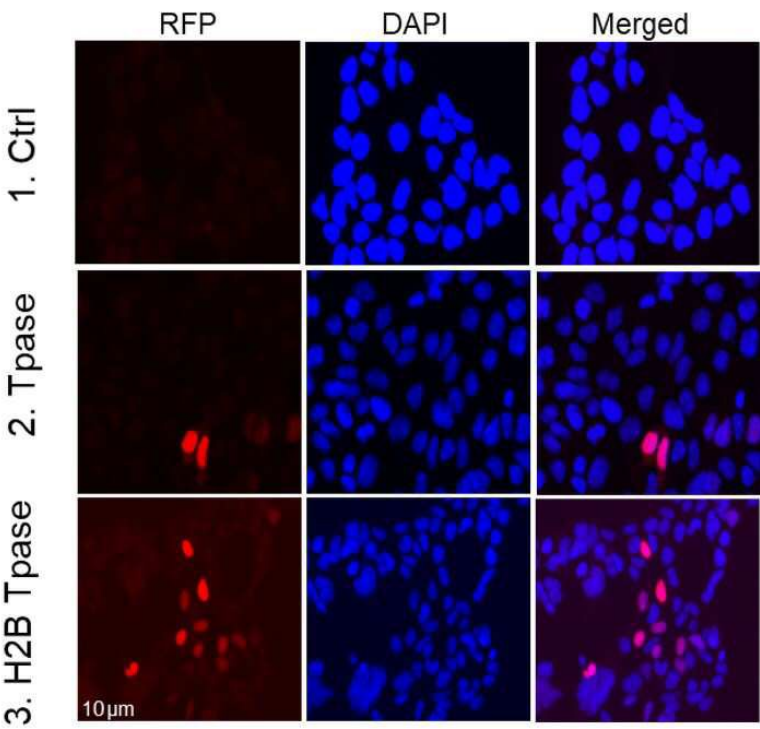
도면6



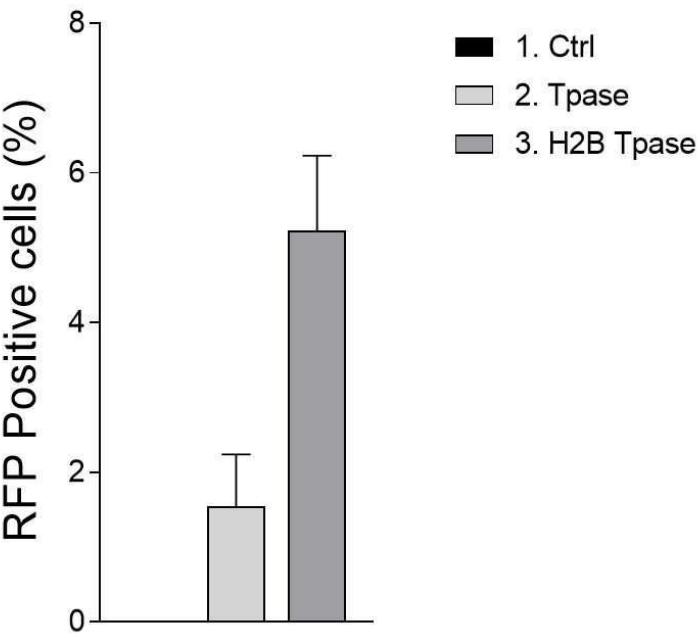
도면7



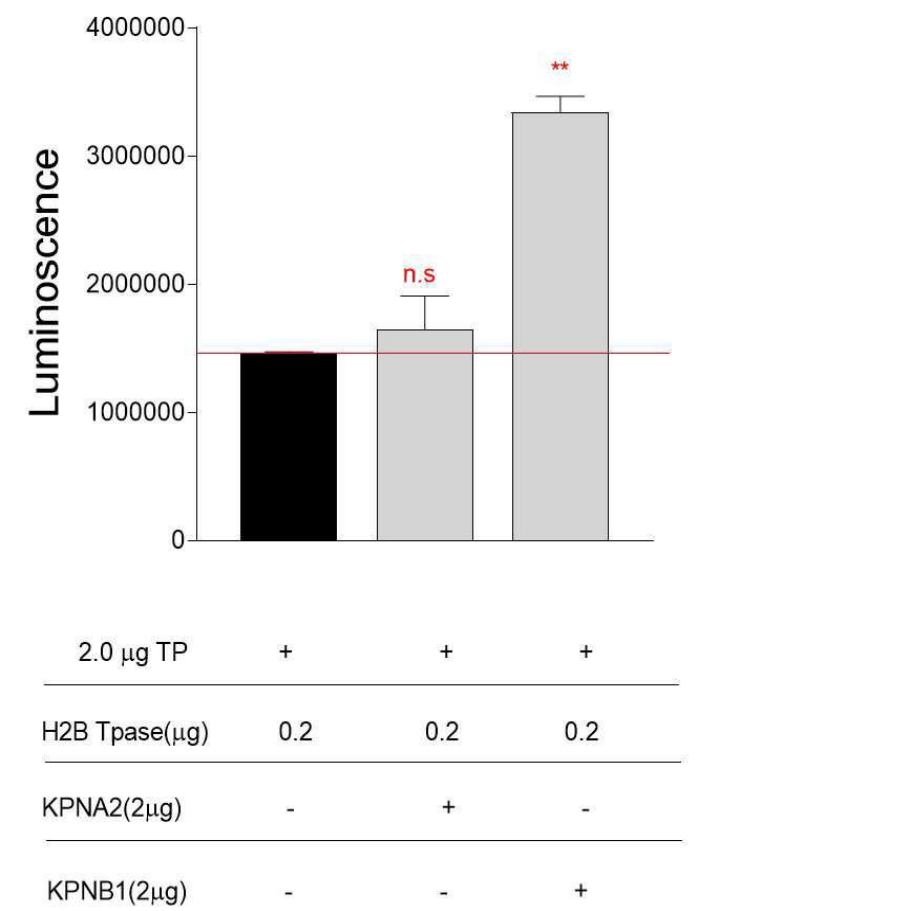
도면8



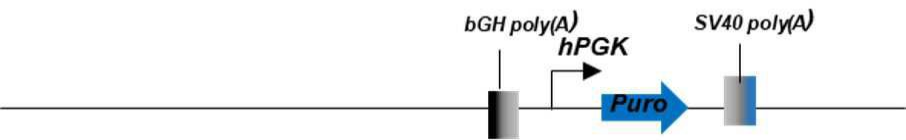
도면9



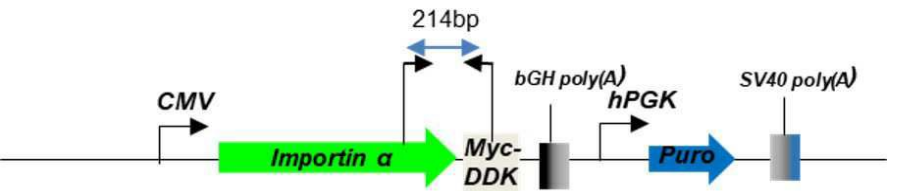
도면10



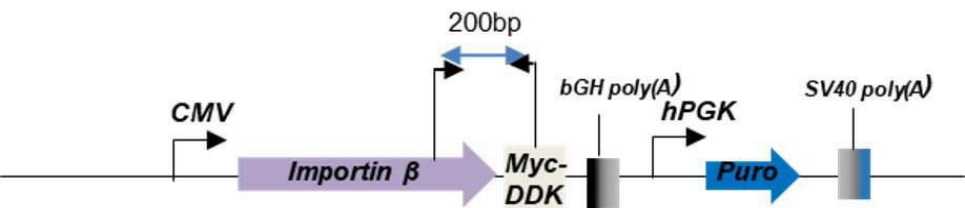
도면11



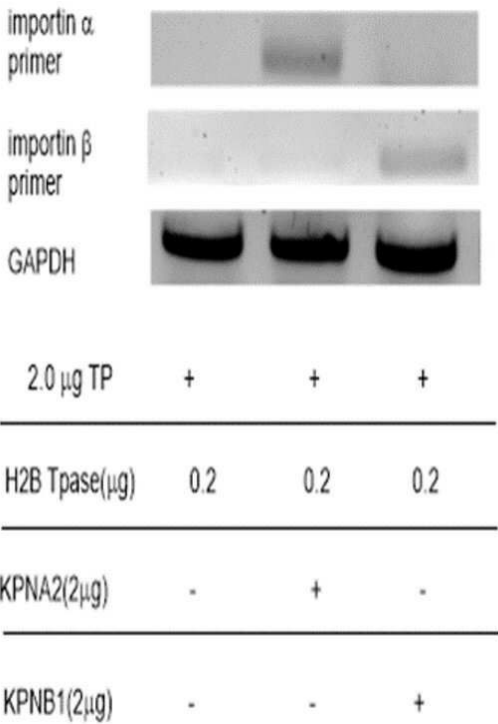
도면12



도면13



도면14



서열 목록

- <110> INCHEON NATIONAL UNIVERSITY RESEARCH AND BUSINESS FOUNDATION
- <120> NOVEL TRANSPOSASE AND TRASNPOSON SYSTEM USING THE SAME
- <130> P22-0004
- <160> 11
- <170> KoPatentIn 3.0
- <210> 1
- <211> 7098
- <212> DNA
- <213> Artificial Sequence
- <220><223> Tol2 transposase sequence
- <400> 1

gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg	60
ccgcatagtt aagccagiat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg	120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc	180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt	240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagtccat agcccatata	300
tggagtccg cgttacataa cttacggtaa atggcccgcg tggctgaccg cccaacgacc	360
cccgcccatg gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc	420
attgacgtca atgggtggag tatttacggt aaactgcca cttggcagta catcaagtg	480
atcatatgcc aagtagcccc cctattgacg tcaatgacgg taaatggccc gcctggcatt	540
atgccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca	600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg	660
actcacgggg atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc	720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc ccattgacg caaatggcg	780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacca	840
ctgcttactg gcttatcgaa attaatacga ctactatag ggagaccaa gctggctagc	900
gtttaaactt aagcttggta ccgagctcgg atccactagt aacggccgcc agtgtgtgg	960
aattcgccct tgatcagcta gcgccacat ggaggaagta tgtgattcat cagcagctgc	1020
gagcagcaca gtccaaaatc agccacagga tcaagagcac ccgtggccgt atcttcgca	1080
attcttttct ttaagtgggtg taaataaaga ttcatccaag atgaaatgtg tcctctgtct	1140
cccgcttaat aaagaaatc cggccttcaa aagttcgcca tcaaacctaa ggaagcatat	1200
tgagagaatg caccctaat acctcaaaaa ctactctaaa ttgacagcac agaagagaaa	1260
gatcgggacc tccacccatg cttccagcag taagcaactg aaagttgact cagttttccc	1320
agtcaaacat gtgtctccag tcaatgtgaa caaagctata ttaaggta tcaattcaagg	1380
acttcactct ttcagcactg ttgatctgcc atcatttaaa gagctgatta gtacactgca	1440
gcctggcatt tcgtcatta caaggcctac ttacgtcc aagatagctg aagctgctct	1500
gatcatgaaa cagaagtga ctgctgcat gagtgaagtt gaatggattg caaccacaac	1560
ggattgttg actgcacgta gaaagtcatt cattgggtga actgtcact ggatcaaccc	1620
tggaagtctt gaaagacatt ccgctgcact tgctgcaa agattaatgg gctctcatac	1680
ttttgaggta ctggccagtg ccatgaatga tatccactca gattatgaaa tacgtgacaa	1740
ggttggttgc acaaccacag acagtgggtc caactttatg aaggttttca gagtttttgg	1800

tgtggaaaac aatgatatcg agactgaggc aagaaggtgt gaaagtgatg acactgattc	1860
tgaaggctgt ggtgaggga gtagatggtgt ggaattccaa gatgcctcac gagtccctgga	1920
ccaagacgat ggcttcgaat tccagctacc aaaacatcaa aagtgtgcct gtcacttact	1980
taacctagtc tcaagcgttg atgccccaaa agctctctca aatgaacact acaagaaact	2040
ctacagatct gtctttggga aatgccaagc tttatggaat aaaagcagcc gatcggtctt	2100
agcagctgaa gctgttgaat cagaaagccg gcttcagctt ttaaggccaa accaaacgcg	2160
gtggaattca acttttatgg ctgttgacag aattcttcaa atttgcaaag aagcaggaga	2220
aggcgcaactt cggaatatat gcacctctct tgaggttcca atgtttaatc cagcagaaat	2280
gctgttcttg acagagtggg ccaacacaaat gcgtccagtt gcaaaagtac tcgacatctt	2340
gcaagcggaa acgaatacac agctgggggtg gctgctgcct agtgtccatc agttaagctt	2400
gaaacttcag cgactccacc attctctcag gtactgtgac ccaacttgtg atgcctaca	2460
acaaggaatc caaacacgat tcaagcatai gtttgaagat cctgagatca tagcagctgc	2520
catccttctc cctaaatttc ggacctcttg gacaaatgat gaaacatca taaaacgagg	2580
catggactac atcagagtgc atctggagcc tttggaccac aagaaggaat tggccaacag	2640
ttcatctgat gatgaagatt ttttcgcttc tttgaaaccg acaacacatg aagccagcaa	2700
agagtgggat ggatatctgg cctgtgtttc agacaccagg gagtctctgc tcacgtttcc	2760
tgctattttg agcctctcta tcaagactaa tacacctctt cccgcatcgg ctgcctgtga	2820
gaggcttttc agcactcgag gattgctttt cagccccaaa agagctaggc ttgacactaa	2880
caattttgag aatcagcttc tactgaagtt aaatctgagg ttttacaact ttgaggggt	2940
cgagtctaga gggcccttcg aaggtaagcc tatccctaac cctctcctcg gtctcgattc	3000
tacgcgtacc ggtcatcatc accatcacca ttgagtttaa acccgctgat cagcctcgac	3060
tgtgccttct agttgccagc catctgttgt ttgccctcc cccgtgcctt ccttgacct	3120
ggaaggtgcc actcccactg tcctttccta ataaaatgag gaaattgcat cgcattgtct	3180
gagtaggtgt cattctattc tgggggggtgg ggtggggcag gacagcaagg gggaggattg	3240
ggaagacaat agcaggcatg ctggggatgc ggtgggctct atggcttctg aggcggaaag	3300
aaccagctgg ggctctaggg ggtatcccca cgcgccctgt agcggcgcac taagcgcggc	3360
gggtgtgggt gttacgcga gcgtgaccgc tacacttgcc agcgccttag cgtccgctcc	3420
tttcgcttcc ttcccttctt ttctcgccac gtctcgccgc tttccccgtc aagctctaaa	3480
tcgggggctc cctttagggt tccgatttag tgctttacgg cacctcgacc ccaaaaaact	3540

tgattagggt gatggttcac gtagtgggcc atcgccctga tagacggttt ttcgcccttt	3600
gacgttggag tccacgttct ttaatagtgg actcttgttc caaactggaa caacactcaa	3660
ccctatctcg gtctattctt ttgatttata agggattttg ccgatttcgg cctattgggt	3720
aaaaaatgag ctgatttaac aaaaatttaa cgcgaattaa ttcgtggaa tgtgtgtcag	3780
ttaggtgtg gaaagtcgcc aggcctccca gcaggcagaa gtatgcaaag catgcatctc	3840
aattagttag caaccagggt tggaaagtcc ccaggctccc cagcaggcag aagtatgcaa	3900
agcatgcatc tcaattagtc agcaaccata gtcccgcccc taactccgcc catcccgccc	3960
ctaaactcgc ccagttccgc ccattctccg ccccatggct gactaatttt ttttatttat	4020
gcagaggccg aggcgcctc tgcctctgag ctattccaga agtagtgagg aggccttttt	4080
ggaggcctag gcttttgcaa aaagctcccg ggagcttgta tatccatttt cggatctgat	4140
cagcagctgt tgacaattaa tcatcgccat agtatatcgg catagtataa tacgacaagg	4200
tgaggaacta aaccatggcc aagcctttgt ctcaagaaga atccaccctc attgaaagag	4260
caacggctac aatcaacagc atccccatct ctgaagacta cagcgtcgcc agcgcagctc	4320
tctctagcga cggccgcctc ttcactgggt tcaatgtata tcattttact gggggacctt	4380
gtgcagaact cgtgggtgctg ggcaactgctg ctgctgcggc agctggcaac ctgacttgta	4440
tcgtcgcgat cggaaatgag aacaggggca tcttgagccc ctgcggacgg tgccgacagg	4500
tgcttctcga tctgcatcct gggatcaaag ccatagttaa ggacagtgat ggacagccga	4560
cggcagttgg gattcgtgaa ttgctgccct ctggttatgt gtgggagggc taagcacttc	4620
gtggccgagg agcaggactg acacgtgcta cgagatttcg attccaccgc cgccttctat	4680
gaaaggttgg gcttcggaat cgttttccgg gacgccggct ggatgatcct ccagcgcggg	4740
gatctcatgc tggagttctt cgcccacccc aacttgttta ttgcagctta taatggttac	4800
aaataaagca atagcatcac aaatttcaca aataaagcat ttttttact gcattctagt	4860
tgtggtttgt ccaaactcat caatgtatct tatcatgtct gtataccgtc gacctctagc	4920
tagagcttgg cgtaatcatg gtcatagtct tttctgtgt gaaattgtta tccgtcaca	4980
attccacaca acatacgagc cggaagcata aagtgtaaag cctggggtgc ctaatgagt	5040
agctaactca cattaattgc gttgcgtca ctgcccgctt tccagtcggg aaacctgtcg	5100
tgccagctgc attaatgaat cggccaacgc gcggggagag gcggtttgcg tattgggcgc	5160
tcttcgctt cctcgctcac tgactcgctg cgtcgggtcg ttcggctgcg gcgagcggt	5220
tcagctcact caaaggcggg aatacggtta tccacagaat caggggataa cgcaggaaag	5280
aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg	5340
ttttccata ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg	5400

tggcgaacc cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg 5460
 cgctctcctg ttccgacct gccgcttacc ggatacctgt ccgcctttct ccttcggga 5520
 agcgtggcgc tttctcatag ctcacgtgt aggtatctca gttcgggtga ggtcgttcgc 5580
 tccaagctgg gcigtgtgca cgaaccccc gttcagcccg accgctgcgc cttatccggt 5640

 aactatcgtc ttgagtccaa cccgtaaga cagcacttat cgccactggc agcagccact 5700
 ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagtctt gaagtgggtg 5760
 cctaactacg gctacactag aagaacagta ttgggtatct gcgctctgct gaagccagtt 5820
 accttcgaa aaagagtgg tagctcttga tccggcaaac aaaccaccgc tggtagcgg 5880
 ggtttttttg ttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct 5940
 ttgatctttt ctacggggtc tgacgtcag tggaacgaaa actcacgtta agggattttg 6000
 gtcattgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt 6060

 aaatcaatct aaagtatata tgagtaaact tggcttgaca gttaccaatg cttaatcagt 6120
 gaggcaccta tctcagcgat ctgtctatct cgttcacca tagttgcctg actccccgtc 6180
 gtgtagataa ctacgatacg ggagggctta ccatctggcc ccagtgtgc aatgataccg 6240
 cgagaccac gtcaccggc tccagattta tcagcaataa accagccagc cggaagggcc 6300
 gagcgagaa gtggctctgc aactttatcc gcctccatcc agtctattaa ttgttgccgg 6360
 gaagctagag taagtagtgc gccagttaat agtttgcgca acgttggtgc cattgctaca 6420
 ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat tcagctccgg tcccaacga 6480

 tcaaggcgag ttacatgac ccccatgttg tgcaaaaaag cggttagctc cttcggctct 6540
 ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg 6600
 cataattctc ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca 6660
 accaagtcat tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata 6720
 cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct 6780
 tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaaccact 6840
 cgtgcacca actgatcttc agcatctttt actttacca gcgtttctgg gtgagcaaaa 6900

 acaggaaggc aaaatgccgc aaaaaaggga ataaggcgca cacggaaatg ttgaatactc 6960
 atactcttc tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga 7020
 tacatatttg aatgtattta gaaaaataa caaatagggg ttccgcgcac atttccccga 7080
 aaagtgccac ctgacgtc 7098

<210> 2

<211> 381

<212> DNA

<213> Artificial Sequence

<220><223> Histone H2B sequence

<400> 2

atggatatgc ctgaaccagc gaagaccgcg cccaagaagg gctccaagaa agccgtcacc 60

aagaccgcgc gtaaaggagg aaagaagcgt aagaggacca ggaaggagag ctacgtatc 120

tacgtgtaca aggtttttag gcaagttcat cccgacaccg gaatctcttc aaaggcgtg 180

ggcatcatga actcgttcgt caacgacatc ttcgagcgca tcgccggtga ggcgtctcgt 240

ctcgtcact acaacaagcg ctccaccatc acttccagag agatccagac cgccgtgcgt 300

ctgtgctac ccggagagtt ggccaaacac gccgtgtctg agggcacaaa ggccgtcacc 360

aagtacacca gctccaagac c 381

<210> 3

<211> 7359

<212> DNA

<213> Artificial Sequence

<220><223> expression vector sequence for H2B Tol2 transposase

<400> 3

gacggatcgg gagatctccc gatccctat ggtgcactct cagtacaatc tgctctgatg 60

ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgt gagtagtgcg 120

cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180

ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240

gattattgac tagttattaa tagtaatcaa ttacggggtc attagtcat agcccatata 300

tggagtccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360

ccgccccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420

attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtg 480

atcatatgcc aagtacgcc cctattgacg tcaatgacgg taaatggccc gcttgccatt 540

atgccagta catgacctta tgggactttc ctacttgca gtacatctac gtattagtea 600

tcgtattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660

actcacgggg atttccaagt ctccaccca ttgacgtcaa tgggagttg ttttggcacc 720

aaaatcaacg ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg 780

gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacca 840

ctgcttactg gcttatcgaa attaatagga ctactatag ggagacccaa gctggctagc	900
gtttggatcc gccaccatgc caaaaaagaa gagaaaggta atggatatgc ctgaaccagc	960
gaagaccgcg cccaagaagg gctccaagaa agccgtcacc aagaccgccc gtaaaggagg	1020
aaagaagcgt aagaggacca ggaaggagag ctacgtctatc tacgtgtaca aggttttgag	1080
gcaagttcat cccgacaccg gaatctcttc aaaggcgatg ggcatcatga actcgttcgt	1140
caacgacatc ttcgagcgca tcgccggtga ggcgctctct ctcgtctact acaacaagcg	1200
ctccaccatc acttccagag agatccagac cgccgtgcgt ctgctgtctac ccggagagtt	1260
ggccaaacac gccgtgtctg agggcacaaa ggccgtcacc aagtacacca gctccaagac	1320
cggatccatg gaggaagtat gtgattcatc agcagctgcg agcagcacag tccaaaatca	1380
gccacaggat caagagcacc cgtggccgta tcttcgcgaa tcttttctt taagtgggtg	1440
aaataaagat tcattcaaga tgaaatgtgt cctctgtctc ccgcttaata aagaaatc	1500
ggccttcaaa agttcgccat caaacctaag gaagcatatt gagagaatgc acccaaatta	1560
cctcaaaaac tactctaat tgacagcaca gaagagaaag atcgggacct ccacccatgc	1620
ttccagcagt aagcaactga aagttgactc agttttccca gtcaaacatg tgtctccagt	1680
cactgtgaac aaagctatat taaggtagat cattcaagga cttcatcctt tcagcactgt	1740
tgatctgcca tcatttaag agctgattag tacactgcag cctggcattt ctgtcattac	1800
aaggcctact ttacgtcca agatagctga agctgctctg atcatgaaac agaaagtgc	1860
tgctgccatg agtgaagtg aatggattgc aaccacaacg gattgttgga ctgcacgtag	1920
aaagtcattc attggtgtaa ctgctcactg gatcaacct ggaagtcttg aaagacattc	1980
cgctgcactt gcctgcaaaa gattaatggg ctctcactt tttgaggtag tggccagtgc	2040
catgaatgat atccactcag agtatgaaat acgtgacaag gttgtttgca caaccacaga	2100
cagtggttcc aactttatga aggctttcag agtttttggt gtggaaaaca atgatatcga	2160
gactgaggca agaagggtgtg aaagtgtatg cactgattct gaaggctgtg gtgagggaag	2220
tgatgggtgtg gaattccaag atgcctcagc agtctggac caagacgatg gcttcgaatt	2280
ccagctacca aaacatcaaa agtgtgcctg tcacttactt aacctagtct caagcgttga	2340
tgcccaaaaa gctctctcaa atgaacacta caagaaactc tacagatctg tctttggcaa	2400
atgccaagct ttatggaata aaagcagccg atcggctcta gcagctgaag ctgttgaatc	2460
agaaagccgg cticagcttt taaggccaaa ccaaaccggt tggaattcaa cttttatggc	2520
tgttgacaga attcttcaaa tttgcaaaga agcaggagaa ggcgcacttc ggaatatatg	2580
cacctctctt gaggttccaa tgtttaatcc agcagaaatg ctgttcttga cagagtgggc	2640
caacacaatg cgtccagttg caaaagtact cgacatcttg caagcggaac cgaatacaca	2700

gctggggtgg ctgctgccta gtgtccatca gttaagcttg aaacttcagc gactccacca	2760
ttctctcagg tactgtgacc cacttttgga tgcctacaa caaggaatcc aaacacgatt	2820
caagcatatg tttgaagatc ctgagatcat agcagctgcc atccttctcc ctaaatttcg	2880
gacctcttgg acaaatgatg aaacatcat aaaacgaggc atggactaca tcagagtga	2940
tctggagcct ttggaccaca agaaggaatt ggccaacagt tcactctgatg atgaagattt	3000
tttcgttct ttgaaaccga caacacatga agccagcaaa gagttggatg gatatctggc	3060
ctgtgtttca gacaccaggg agtctctgct cacgtttcct gctatttgca gcctctctat	3120
caagactaat acacctcttc ccgcatcggc tgcctgtgag aggcttttca gactgcagg	3180
attgcttttc agccccaaaa gagctaggct tgacactaac aattttgaga atcagcttct	3240
actgaagtta aatctgaggt tttacaactt tgagcatcat caccatcacc attgagtta	3300
aaccgctga tcagcctga ctgtgccttc tagttgccag ccactctgtt tttgcccttc	3360
ccccgtgcct tcttgacc ccggaaggtgc cactcccact gtcctttcct aataaaatga	3420
ggaaattgca tcgcatgttc tgagttagtg tcattctatt ctgggggggtg ggggtggggca	3480
ggacagcaag ggggaggatt gggaagacaa tagcagcat gctggggatg cgggtgggctc	3540
tatggcttct gaggcggaaa gaaccagctg gggctctagg gggatcccc acgcgcctg	3600
tagcggcgca ttaagcgcg cggtgtgtgt ggttacgcgc agcgtgaccg ctacacttgc	3660
cagcgcccta gcgcgcctc ctttcgcttt ctcccttcc tttctcgca cgttcgccg	3720
ctttccccgt caagctctaa atcgggggct ccctttaggg ttccgattta gtgctttacg	3780
gcacctgac ccaaaaaaac ttgattaggg tgatggttca cgtagtgggc catcgccctg	3840
atagacggtt tticgccctt tgacgttga gtccacgttc tttaatagt gactcttgtt	3900
ccaaactgga acaacactca accctatctc ggtctattct tttgatttat aagggtttt	3960
gccgatttcg gctatttgt taaaaaatga gctgatttaa caaaaattta acgcaatta	4020
attctgtgga atgtgtgtca gttagggtgt ggaaagtccc caggctccc agcaggcaga	4080
agtatgcaaa gcatgcatct caattagtca gcaaccaggt gtggaaagtc cccaggtcc	4140
ccagcaggca gaagtatgca aagcatgcat ctcaattagt cagcaacct agtccgccc	4200
ctaactccgc ccatcccgcc cctaactccg ccagttccg ccattctcc gcccattggc	4260
tgactaatTT tttttattta tgcagaggcc gaggcgcct ctgcctctga gctattccag	4320
aagtagtgag gaggttttt ttgaggccia ggcttttgca aaaagctccc gggagcttgt	4380
atatccattt tcggatctga tcagcacgtg ttgacaatta atcatcgga tagtatatcg	4440
gcatagtata ataccacaag gtgaggaact aaacatggc caagcctttg tctcaagaag	4500

aatccaccct cattgaaaga gcaacggcta caatcaacag catccccatc tctgaagact	4560
acagcgtcgc cagcgcagct ctctctagcg acggccgcat cttcactggt gtcaatgtat	4620
atcattttac tgggggacct tgtgcagaac tcgtgggtgct gggcactgct gctgctgcgg	4680
cagctggcaa ccigacttgt atcgtcgcga tcggaaatga gaacaggggc atcttgagcc	4740
cctgcggacg gtgccgacag gtgcttctcg atctgcatcc tgggatcaaa gccatagtga	4800
aggacagtga tggacagccg acggcagttg ggattcgtga attgctgccc tctggttatg	4860
tgtgggaggg ctaagcactt cgtggccgag gacgaggact gacacgtgct acgagatttc	4920
gattccaccg ccgccttcta tgaaagggtg ggcttcggaa tcgttttccg ggacgccggc	4980
tggatgatcc tccagcgcgg ggatctcatg ctggagtctt tcgccacccc caactgtttt	5040
attgcagctt ataattggtta caaataaagc aatagcatca caaatttcac aaataaagca	5100
tttttttcac tgcattctag ttgtggtttg tccaaactca tcaatgtatc ttatcatgtc	5160
tgtataccgt cgacctctag ctagagcttg gcgtaatcat ggtcatagct gtttctgtg	5220
tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat aaagtgtaaa	5280
gcctgggggtg cctaagtagt gagctaactc acattaattg cgttgcgtc actgcccgct	5340
ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg cgcggggaga	5400
ggcggtttgc gtattgggcg ctcttccgct tcctcgctca ctgactcgct gcgctcggtc	5460
gttcggctgc ggcgagcggg atcagctcac tcaaaggcgg taatacgggt atccacagaa	5520
tcaggggata acgcaggaaa gaacatgtga gcaaaaaggcc agcaaaaggc caggaaccgt	5580
aaaaaggcgg cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa	5640
aatcgacgct caagtacagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt	5700
ccccctggaa gctccctcgt gcgctctcct gtccgaccc tgccgcttac cggatacctg	5760
tccgcctttc tccttccggg aagcgtggcg ctttctcata gctcacgctg taggtatctc	5820
agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc	5880
gaccgctcgc cttatccgg taactatcgt cttgagtcca acccgtaag acacgactta	5940
tcgccactgg cagcagccac tggtaacagg attagcagag cgaggatatgt aggcggtgct	6000
acagagtctt tgaagtgggt gcctaactac ggctacacta gaagaacagt atttggtatc	6060
tgcgctctgc tgaagccagt taccttcgga aaaagagtig gtagctcttg atccggcaaa	6120
caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa	6180
aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa	6240

aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt 6300
 ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac 6360
 agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc 6420
 atagttgcct gactccccgt cgtgtagata actacgatac gggagggcctt accatctggc 6480
 cccagtgtctg caatgatacc gcgagacca cgtcacccg ctccagattt atcagcaata 6540
 aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc 6600
 cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgcgc 6660

aacgttgttg ccattgctac aggcatcgtg gtgtcacgct cgtcgttttg tatggcttca 6720
 ttcagctccg gttcccaacg atcaaggcga gttacatgat ccccatgtt gtgcaaaaaa 6780
 gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtggtatca 6840
 ctcatggtta tggcagcact gcataattct ctactgtca tgccatccgt aagatgcttt 6900
 tctgtgactg gtgagtactc aaccaagtca tttcgagaat agtgtatgcg gcgaccgagt 6960
 tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac tttaaaagtg 7020
 ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga 7080

tccagttcga tgtaaccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc 7140
 agcgtttctg ggtgagcaaa aacaggaagg caaatgccg caaaaaagg aataaggcg 7200
 acacggaaat gttgaatact catactcttc ctttttcaat attattgaag catttatcag 7260
 ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg 7320
 gttccgcga catttccccg aaaagtgcc cctgacgtc 7359

<210> 4

<211> 2334

<212> DNA

<213> Artificial Sequence

<220><223> H2B Tol2 transposase sequence

<400> 4

atggatatgc ctgaaccagc gaagaccgcg cccaagaagg gctccaagaa agccgtcacc 60
 aagaccgccg gtaaaggagg aaagaagcgt aagaggacca ggaaggagag ctacgtatc 120
 tactgtaca aggttttgag gcaagttcat cccgacaccg gaatctcttc aaaggcgatg 180
 ggcatcatga actcgttcgt caacgacatc ttcgagcga tcgccggtga ggctgtcgt 240
 ctgctcact acaacaagcg ctccaccatc acttcagag agatccagac cgccgtgcgt 300
 ctgctgctac ccggagagtt ggccaacac gccgtgtctg agggcacaaa ggccgtcacc 360

aagtacacca gctccaagac cggatccatg gaggaagtat gtgattcatc agcagctgcg	420
agcagcacag tccaaaatca gccacaggat caagagcacc cgtggccgta tcttcgcgaa	480
ttcttttctt taagtgggtg aaataaagat tcattcaaga tgaaatgtgt cctctgtctc	540
ccgcttaata aagaaatatc ggcccttcaa agttcgccat caaacctaag gaagcatatt	600
gagagaatgc acccaaatta cctcaaaaac tactctaaat tgacagcaca gaagagaaag	660
atcgggacct ccacctatgc ttccagcagt aagcaactga aagttgactc agttttccca	720
gtcaaacatg tgtctccagt cactgtgaac aaagctatat taaggtacat cattcaagga	780
cttcacctt tcagcactgt tgatctgcca tcatttaaag agctgattag tacactgcag	840
cctggcattt ctgtcattac aaggcctact ttacgctcca agatagctga agctgctctg	900
atcatgaaac agaaagtgc tgctgccatg agtgaagttg aatggattgc aaccacaacg	960
gattgttgga ctgcacgtag aaagtcattc attgggtgaa ctgctcactg gatcaaccct	1020
ggaagtcttg aaagacattc cgctgcactt gcctgcaaaa gattaatggg ctctcatact	1080
tttgaggtag tggccagtgc catgaatgat atccactcag agtatgaaat acgtgacaag	1140
gttgtttgca caaccacaga cagtggttcc aactttatga aggctttcag agtttttgg	1200
gtggaaaaca atgatatcga gactgaggca agaaggtgtg aaagtgatga cactgattct	1260
gaaggtgtg gtgaggggaag tgatgggtgtg gaattccaag atgcctcacg agtcttggac	1320
caagacgatg gcttcgaatt ccagctacca aaacatcaaa agtgtgcctg tcacttactt	1380
aacctagtct caagcgttga tgcccaaaaa gctctctcaa atgaacacta caagaaactc	1440
tacagatctg tctttggcaa atgccaagct ttatggaata aaagcagccg atcggtctta	1500
gcagctgaag ctgttgaatc agaaagccgg cttcagcttt taaggccaaa ccaaacgcgg	1560
tggaattcaa cttttatggc tgttgacaga attcttcaaa tttgcaaaga agcaggagaa	1620
ggcgcacttc ggaatatatg cacctctctt gaggttccaa tgtttaatcc agcagaaatg	1680
ctgttcttga cagagtgggc caacacaatg cgtccagtgt caaaagtact cgacatcttg	1740
caagcggaaa cgaatacaca gctggggtgg ctgctgccta gtgtccatca gtttaagcttg	1800
aaacttcagc gactccacca ttctctcagg tactgtgacc cacttgttga tgccctacaa	1860
caaggaatcc aaacacgatt caagcatatg ttgaagatc ctgagatcat agcagctgcc	1920
atccttctcc ctaaatttgc gacctcttgg acaaatgatg aaaccatcat aaaacgaggc	1980
atggactaca tcagagtga tctggagcct ttggaccaca agaaggaatt ggccaacagt	2040
tcactgatg atgaagattt tttcgcttct ttgaaaccga caacacatga agccagcaaa	2100

gagttggatg gatatctggc ctgtgtttca gacaccaggg agtctctgct cacgtttcct 2160
gctatttgca gcctctctat caagactaat acacctcttc ccgcatcggc tgcctgtgag 2220
aggcttttca gcactgcagg attgcttttc agccccaaaa gagctaggct tgacactaac 2280
aattttgaga atcagcttct actgaagtta aatctgaggt tttacaactt tgag 2334
<210> 5
<211> 9272
<212> DNA
<213> Artificial Sequence
<220><223> Tol2 TP (transposon vector) sequence
<400> 5
aaaagtgcc cctgacgtcg acggatcggg agatctcccg atcccctatg gtgcactctc 60

agtacaatct gcctgatgc cgcatagtta agccagtatc tgctccctgc ttgtgtgttg 120
gaggtcgctg agtagtgccg gagcaaaatt taagtacaa caaggcaagg cttgaccgac 180
aattgcata agaactctgt tagggtagg cgttttgccg tgcttcgttt tgcgctgctt 240
cgcatctgc gaagatacgg ccacgggtgc tcttgatcct gtggctgatt ttggactgtg 300
ctgctcgag ctgctgatga atcacatact tcttccattt tcttccactg attgactgtt 360
ataatttccc taatttccag gtcaaggtgc tgtgcatgtt ggtaaatgat gtgacatgac 420
gtcacttcca aaggaccaat gaacatgtct gaccaatttc atataatgtg aaaacgattt 480

tcataggcag aataaataac atttaaatga aactgggcat cagcgcaatt caattggttt 540
ggtaaatgca agggaaaata gaatgaagtg atctccaaaa aataagtact ttttgactgt 600
aaataaaatt gtaaggagta aaaagtactt ttttttctaa aaaaatgtaa ttaagtaaaa 660
gtaaaagtat tgatttttaa ttgtactcaa gtaaagtaaa aatccccaaa aataatactt 720
aagtacagta atcaagtaaa attactcaag tactttacac ctctggttct tgacccccta 780
ccttcattgc ggaacccta tttgtttatt tttctaata cattcaaata tgtatccgct 840
catgagacaa taacctgat aaatgcttca ataatatga aaaaggaaga gtatgagtat 900

tcaacatttc cgtgtcgccc ttattccctt ttttgcggca ttttgccttc ctgtttttgc 960
tcaccagaa acgctggta aagtaaaaga tgctgaagat cagttgggtg cagcagtggtg 1020
ttacatcgaa ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg 1080
ttttccaatg atgagcatt ttaaagttct gctatgtggc gcggtattat cccgtattga 1140
cgccgggcaa gagcaactcg gtcgccgat acactattct cagaatgact tggttgagta 1200
ctcaccagtc acagaaaagc atcttacgga tggcatgaca gtaagagaat tatgcagtgc 1260

tgccataacc atgagtata acactgcggc caacttactt ctgacaacga tcggaggacc	1320
gaaggagcta accgcttttt tgcacaacat ggggatcat gtaactcgcc ttgatcgttg	1380
ggaaccggag ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgcctgtagc	1440
aatggcaaca acgttgcgca aactattaac tggcgaacta cttactctag cttcccggca	1500
acaattaata gactggatgg aggcggataa agttgcagga ccacttctgc gctcgccct	1560
tccggctggc tggtttatgt ctgataaatc tggagccggt gagcgtgggt ctgcggtat	1620
cattgcagca ctggggccag atggttaagc cttccgtatc gtagttatct acacgacggg	1680
gagtcaggca actatggatg aacgaaatag acagatcgct gagatagggt cctcactgat	1740
taagcattgg taactgtcag accaaccat atccactctg ttccacacag gtcagaggtt	1800
tgtccaggag ttcttgacag aggtgtaaaa agtactcaa aattttactc aagtgaagt	1860
acaagtactt agggaaaatt ttactcaatt aaaagtaaaa gtatctggct agaactctac	1920
ttgagtaaaa gtaaaaaagt actccattaa aattgtactt gtagtattaag gaagtaaaag	1980
taaaagcaag aaagaaaact agagattctt gttaagctt ttaatctcaa aaaacattaa	2040
atgaaatgca tacaaggttt tatctgctt tagaactgtt tgtatttaat tatcaaacta	2100
taagacagac aatctaagtc cagtacacgc tactcaaagt tgtaaacct cagatttaac	2160
ttcagtagaa gctgattctc aaaattgtta gtgtcaagcc tagctctttt ggggctgaaa	2220
agcaatcctg cagtgtgaa aagcctctca caggcagccg atgcgggaag aggtgtatta	2280
gtcttgatag agaggctgca aatagcagga aacgtgagca gagactccct ggtgtctgaa	2340
acacaggcca gatgggcctc gcgatgtacg ggccagcgat gtacgggcca gatatacgcg	2400
ttgacattga ttattgacta gttattaata gtaatcaatt acgggggtcat tagttcatag	2460
cccatatatg gaggttccggt ttacataact tacggtaaat ggccgcctg gctgaccgcc	2520
caacgacccc cgccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg	2580
gactttccat tgacgtcaat ggggtggagta ttacggtaa actgccact tggcagtaca	2640
tcaagtgtat catatgccaa gtacgcccc tattgacgtc aatgacggt aatggccgc	2700
ctggcattat gccagtiaca tgacctatg ggactttcct acttggcagt acatctacgt	2760
attagtcac gctattacca tggatgatcg gttttggcag tacatcaatg ggcgtggata	2820
gcggtttgac tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttgtt	2880
ttggcaccaa aatcaacggg actttccaaa atgtcgtaac aactccgcc cattgacgca	2940
aatgggcggt aggcgtgtac ggtgggaggt ctatataagc agagctctct ggctaactag	3000

agaaccact gcttactggc ttatcgaaat taatacgact cactataggg agaccaagc	3060
tggctagcgc caccatggag ttggggctga gctgggtttt cctcgttgct ctttttagag	3120
gtgtccagtg tatggaagac gccaaaaaca taaagaaagg cccggcgcca ttctatccgc	3180
tggaagatgg aaccgctgga gagcaactgc ataaggctat gaagagatac gccctggttc	3240
ctggaacaat tgcttttaca gatgcacata tcgagggtgga catcacttac gctgagtact	3300
tcgaaatgtc cgttcgggtg gcagaagcta tgaaacgata tgggtcgaat acaaatcaca	3360
gaatcgtcgt atgcagtga aactctcttc aattctttat gccgggtgtg ggcgcgttat	3420
ttatcggagt tgcagttgcg cccgcgaacg acatttataa tgaacgtgaa ttgctcaaca	3480
glatgggcat ttgcagcct accgtgggtg tcgtttccaa aaaggggttg caaaaaattt	3540
tgaacgtgca aaaaaagctc ccaatcatcc aaaaaattat tatcatggat tctaaaacgg	3600
attaccaggg atttcagtcg atgtacacgt tcgtcacatc tcactctacct cccggtttta	3660
atgaatacga ttttgtcca gagtccttcg atagggacaa gacaattgca ctgatcatga	3720
actcctctgg atctactggt ctgcctaaag gtgtcgtctt gcctcataga actgcctgcg	3780
tgagattctc gcatgccaga gatcctattt ttggcaatca aatcattccg gatactgcga	3840
ttttaagtgt tgttccattc catcacggtt ttggaatggt tactacactc ggatatttga	3900
tatgtggatt tcgagtcgtc ttaatgtata gatttgaaga agagctgttt ctgaggagcc	3960
ttcaggatta caagattcaa agtgcgctgc tggtgccaac cctattctcc ttcttcgcca	4020
aaagcactct gattgacaaa tacgatttat ctaatttaca cgaaattgct tctggtggcg	4080
ctccctctc taaggaagtc ggggaagcgg ttccaagag gttccatctg ccaggatatca	4140
ggcaaggata tgggctcact gagactacat cagctattct gattacaccc gagggggatg	4200
ataaaccggg cgcggtcggg aaagtgttc catTTTTTga agcgaaggtt gtggatctgg	4260
ataccgggaa aacgctgggc gttaatcaaa gaggcgaact gtgtgtgaga ggtcctatga	4320
ttatgtccgg ttatgtaaac aatccggaag cgaccaacgc cttgattgac aaggatggat	4380
ggctacattc tggagacata gcttactggg acgaagacga acacttcttc atcgttgacc	4440
gcctgaagtc tctgattaag tacaaaggct atcagggtggc tcccgtgaa ttggaatcca	4500
tcttgctcca acacccaac atcttcgacg caggtgtcgc aggtcttccc gacgatgacg	4560
ccggtgaact tcccgccgc gttgttgttt tggagcacgg aaagacgatg acggaaaaag	4620
agatcgtgga ttacgtcgcc agtcaagtaa caaccgcgaa aaagtgtcgc ggaggagtgtg	4680
tgtttgtgga cgaagtaccg aaaggtctta ccggaaaact cgacgcaaga aaaatcagag	4740
agatcctcat aaaggccaag aaggcggaag agatcgccgt gtaactcgag cggccgccac	4800
tgtgtggat atctgcagaa ttccaccaca ctggactagt ggatccgagc tcggtaccaa	4860

gcttaagttt aaaccgctga tcagcctcga ctgtgccttc tagttgccag ccatctgttg	4920
tttgcccttc ccccgctgct tctttgacct tggaaggtgc cactcccact gtcttttct	4980
aataaaatga ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctggggggtg	5040
gggtggggca ggacagcaag ggggaggatt gggaagacaa tagcaggcat gctggggatg	5100
cggtagggctc tatggcttct gaggcggaaa gaaccagctg gggctctagg gggatatccc	5160
acgcgccctg tagcggcgca ttaagcgagg cgggtgtggt ggttacgcgc agcgtgaccg	5220
ctacacttgc cagcgcccta gcgcccgtc ctttcgcttt cttcccttcc tttctcgcca	5280
cgttcgccgg ctttcccggt caagctctaa atcgggggct ccttttaggg ttccgattta	5340
gtgctttacg gcacctgac ccaaaaaaac ttgattaggg tgatggttca cgtagtgggc	5400
catcgccctg atagacggtt tttcgccctt tgacgttga gtccacgttc tttaatagt	5460
gactcttgtt ccaaactgga acaacactca accctatctc ggtctattct tttagattat	5520
aagggatattt gccgatttcg gcctattggt taaaaaatga gctgatttaa caaaaattta	5580
acgcgaatta attctgtgga atgtgtgtca gttagggtgt ggaaagtccc caggctcccc	5640
agcaggcaga agtatgcaaa gcatgcatct caattagtca gcaaccaggt gtggaaagtc	5700
cccaggctcc ccagcaggca gaagtatgca aagcatgcat ctcaattagt cagcaacat	5760
agtcgccccc ctaactccgc ccatcccgcc cctaactccg cccagtcccg cccattctcc	5820
gccccatggc tgactaattt tttttattta tgcagaggcc gaggccgcct ctgcctctga	5880
gctattccag aagtagtgag gaggcctttt tggaggccta ggcttttgca aaaagctccc	5940
gggagcttgt atatccattt tcggatctga tcaagagaca ggatgaggat cgtttcgc	6000
gattgaacaa gatggattgc acgcagggtc tccggccgct tgggtggaga ggctattcgg	6060
ctatgactgg gcacaacaga caatcggtg ctctgatgcc gccgtgttcc ggctgtcagc	6120
gcagggggcg ccggttcttt ttgtcaagac cgacctgtcc ggtgccctga atgaactgca	6180
ggacgaggca gcgcggctat cgtggctggc cagcagggc gttccttgcg cagctgtgct	6240
cgacgtgtc actgaagcgg gaagggactg gctgctattg ggcaagtgc cggggcagga	6300
tctctgtca tctacattg ctctgcccga gaaagtatcc atcatggctg atgcaatgcg	6360
gcggctgcat acgttgatc cggctacctg cccattcgac caccaagcga aacatcgcat	6420
cgagcgagca cgtactcgga tggaagccgg tcttgtgat caggatgatc tggacgaaga	6480
gcatcagggg ctgcgccag ccgaactgtt cgccaggctc aaggcgcgca tccccgacgg	6540
cgaggatctc gtcgtgacct atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg	6600
ccgcttttct ggattcatcg actgtggccg gctgggtgtg gcggaccgct atcaggacat	6660
agcgttggct acccgtgata ttgtgaaga gcttggcggc gaatgggctg accgttctct	6720

cgtagctttac ggtagcgccg ctcccgattc gcagcgcatc gccttctatc gccttcttga	6780
cgagtcttctc tgagcgggac tctgggggttc gaaatgaccg accaagcgac gcccacactg	6840
ccatcacgag atttcgattc caccgccgcc ttctatgaaa ggttgggctt cggaatcgtt	6900
ttccgggacg ccggctggat gatcctccag cgcggggatc tcatgctgga gttcttcgcc	6960
caccccaact tgtttattgc agcttataat ggttacaaat aaagcaatag catcacaat	7020
ttcacaata aagcattttt ttactgcat tctagttgtg gtttgccaa actcatcaat	7080
glatcttate atgtctglat accgtcgacc tctagctaga gcttggcgta atcatggta	7140
tagctgtttc ctgtgtgaaa ttgttatccg ctcaaatc cacacaacat acgagccgga	7200
agcataaagt gtaaagcctg gggtagcctaa tgagttagct aactcacatt aattgcgttg	7260
cgctcactgc ccgttttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc	7320
caacgcgcgg ggagaggcgg ttgctgtatt gggcgtcttt ccgttcttc gctcactgac	7380
tcgtcgcgt cggtcgttcg gctgcggcga gcggtatcag ctactcaaa ggcggttaata	7440
cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa	7500
aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggt ccgccccct	7560
gacgagcatc aaaaaaatcg acgtcaagt cagaggtggc gaaacccgac aggactataa	7620
agataccagg cgtttccccc tggaagctcc ctctgctgct ctctgttcc gacctgccg	7680
cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgcttcc tcatagctca	7740
cgctgtaggt atctcagttc ggtgtaggtc gttcgtcca agctgggctg tgtgcacgaa	7800
cccccgttc agcccgaccg ctgcgcctta tccgtaact atcgtcttga gtccaaccg	7860
gtaagacacg actatcgcc actggcagca gccactggtg acaggattag cagagcgagg	7920
tatgtaggcg gtgtacaga gttcttgaag tggtagccta actacggcta cactagaaga	7980
acagtatttg gtatctgcgc tctgtgaag ccagttacct tcggaaaaag agttggtagc	8040
tcttgatccg gcaaaaaac caccgtggt agcggttttt ttgtttgcaa gcagcagatt	8100
acgcgcagaa aaaaaggatc tcaagaagat cttttgatct ttctacggg gctgacgct	8160
cagtggaacg aaaactcacg ttaagggtt ttggtcatga gattatcaa aaggatcttc	8220
acctagatcc ttttaaatta aaaatgaagt tttaatcaa tctaaagtat atatgagtaa	8280
acttggtctg acagttacca atgttaatc agtgaggcac ctatctcagc gatctgtcta	8340
tttcgttcat ccatagtgc ctgactcccc gtcgtgtaga taactacgat acgggagggc	8400
ttaccatctg gccccagtgc tgcaatgata ccgcgagacc cacgtcacc ggctccagat	8460

ttatcagcaa taaaccagcc agccggaagg gccgagcgca gaagtgggtcc tgcaacttta 8520
 tccgcctcca tccagctctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt 8580
 aatagtttgc gcaacgttgt tgccattgct acaggcatcg tgggtgtcacg ctcgctgttt 8640
 ggtatggctt cattcagctc cggttcccaa cgtcaaggc gagttacatg atcccccatg 8700
 ttgtgcaaaa aagcggtttag ctctctcggt cctccgatcg ttgtcagaag taagtggcc 8760
 gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc 8820
 gtaagatgct ttictgtgac tggtagtagt tcaaccaagt cattctgaga atagtgtatg 8880

cggcgaccga gttgctcttg cccggcgtca atacgggata ataccgcgcc acatagcaga 8940
 actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta 9000
 ccgctgttga gatccagttc gatgtaacc actcgtgcac ccaactgatc ttcagcatct 9060
 ttacttttca ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag 9120
 ggaataaggg cgacacggaa atgttgaata ctcatctct tcctttttca atattattga 9180
 agcatttate agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat 9240
 aaacaaatag gggttccgcg cacatttccc cg 9272

<210> 6

<211> 592

<212> DNA

<213> Artificial Sequence

<220><223> Tol2 Right ITR sequence

<400> 6

acccatattc actctgttcc acacaggtca gaggtttgtc caggagttct tgacagaggt 60
 gtaaaaagta ctcaaaaatt ttactcaagt gaaagtacaa gtacttaggg aaaattttac 120
 tcaattaaaa gtaaaagtat ctggctagaa tcttacttga gtaaaagtaa aaaagtactc 180
 cattaaaatt gtacttgagt attaaggaag taaaagtaaa agcaagaaaag aaaactagag 240
 attcttgttt aagcttttaa tctcaaaaaa cattaaatga aatgcataca aggttttatac 300

ctgctttaga actgttttga tttaattatc aaactataag acagacaatc taatgccagt 360
 acacgtact caaagtttga aaacctcaga tttaacttca gtagaagctg attctcaaaa 420
 ttgttagtgt caagcctagc tcttttgggg ctgaaaagca atcctgcagt gctgaaaagc 480
 ctctcacagg cagccgatgc gggaagaggt gtattagtct tgatagagag gctgcaaata 540
 gcaggaaacg tgagcagaga ctccctggtg tctgaaacac aggccagatg gg 592

<210> 7

<211> 539

<212> DNA

<213> Artificial Sequence

<220><223> Tol2 Left ITR sequence

<400>

7

tgcaagata cgccacggg tgctcttgat cctgtggctg attttgact gtgctgctcg	60
cagctgctga tgaatcacat acttcctcca ttttcttcca ctgattgact gttataattt	120
ccctaatttc caggtcaagg tgctgtgcat tgggtaata gatgtgacat gacgtcactt	180
ccaaaggacc aatgaacatg tctgaccaat ttcataaat gtgaaaacga ttttcatagg	240
cagaataaat aacatttaaa ttaaactggg catcagcgca attcaattgg ttggtaata	300
gcaagggaat atagaatgaa gtgatctcca aaaaataagt actttttgac tgtaataaaa	360
attgtaagga gtaaaaagta cttttttttc taaaaaatg taattaagta aaagtaaaag	420
tattgatttt taattgtact caagtaaagt aaaaatcccc aaaaataata ctttaagtaca	480
gtaatcaagt aaaattactc aagtacttta cacctctggg tcttgacccc ctaccttca	539

<210> 8

<211> 15

<212> DNA

<213> Artificial Sequence

<220><223> forward primer sequence for Importin alpha

<400> 8

gctctacaaa accat	15
------------------	----

<210> 9

<211> 15

<212> DNA

<213> Artificial Sequence

<220><223> reverse primer sequence for Importin alpha

<400> 9

ccagatcctc ttctg	15
------------------	----

<210> 10

<211> 15

<212> DNA

<213> Artificial Sequence

<220><223> forward primer sequence for Importin beta

<400>	10	
cagcatttgg gaagg		15
<210>	11	
<211>	15	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	reverse primer sequence for Importin beta	
<400>	11	
tgccagatcc tcttc		15