



**(19) 대한민국특허청(KR)**  
**(12) 공개특허공보(A)**

(11) 공개번호 10-2012-0131180  
 (43) 공개일자 2012년12월04일

(51) 국제특허분류(Int. Cl.)  
 C12N 5/074 (2010.01) C12N 5/0735 (2010.01)  
 C12N 5/10 (2006.01)  
 (21) 출원번호 10-2012-7023784  
 (22) 출원일자(국제) 2011년02월16일  
 심사청구일자 없음  
 (85) 번역문제출일자 2012년09월11일  
 (86) 국제출원번호 PCT/JP2011/053874  
 (87) 국제공개번호 WO 2011/102531  
 국제공개일자 2011년08월25일  
 (30) 우선권주장  
 61/305,107 2010년02월16일 미국(US)  
 61/379,949 2010년09월03일 미국(US)

(71) 출원인  
**교쿠리츠 다이가쿠 호진 교토 다이가쿠**  
 일본 교토후 교토시 사쿄쿠 요시다혼마치 36반치 1  
**샤단호진 바이오산교조호카 콘소시아무**  
 일본 도쿄도 고토쿠 아오미 2초메 4방 32고  
**도꾸리쯔교세이호진 상교기쥬쯔 소고겐쥬쥬**  
 일본 도쿄도 치요다쿠 카스미가세키 1-3-1  
 (72) 발명자  
**야마나카 신야**  
 일본 교토후 교토시 사쿄쿠 쇼고인 가와하라쥬 53  
 교쿠리츠 다이가쿠 호진 교토 다이가쿠 센터 포  
 아이피에스 셀 리서치 앤드 어플리케이션 나이  
**고시마 나오키**  
 일본 도쿄도 고토쿠 아오미 2초메 3-26 도꾸리쯔  
 교세이호진 상교기쥬쯔 소고겐쥬쥬 나이  
 (뒷면에 계속)  
 (74) 대리인  
**특허법인코리아나**

전체 청구항 수 : 총 31 항

(54) 발명의 명칭 **유도된 다능성 줄기 세포의 효율적 확립 방법**

**(57) 요약**

GLIS 패밀리의 일원 (예, GLIS1) 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질 및 Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질을, 체 세포와 접촉시키는 단계를 포함하는, iPS 세포의 확립 효율 개선 방법, 이 방법으로 수득될 수 있는, GLIS 패밀리의 일원 또는 Klf 패밀리의 일원을 인코딩하는 외생 핵산을 포함하는 iPS 세포, 및 iPS 세포의 분화를 유도함으로써 체세포를 생성하는 방법이 제공된다.

(72) 발명자

**마에카와 모모코**

일본 교토후 교토시 사쿄쿠 쇼고인 가와하라쵸 53  
고쿠리츠 다이가쿠 호진 교토 다이가쿠 센터 포 아  
이피에스 셀 리서치 앤드 어플리케이션 나이

**가와무라 요시후미**

일본 도쿄도 고토쿠 아오미 2쵸메 4-32 샤단호진  
바이오산교쵸호카 콘소시아무 나이

**모치즈키 히로미**

일본 도쿄도 고토쿠 아오미 2쵸메 4-32 샤단호진  
바이오산교쵸호카 콘소시아무 나이

## 특허청구의 범위

### 청구항 1

하기 (1) 및 (2):

- (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 균으로부터 선택되는 1 종 이상의 물질,
- (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 균으로부터 선택되는 1 종 이상의 물질을 체세포와 접촉시키는 것을 포함하는, iPS 세포 확립 효율의 개선 방법.

### 청구항 2

제 1 항에 있어서, 물질 (1) 이 GLIS 패밀리 징크 핑거 (zinc finger) 1 (GLIS1) 또는 GLIS1 을 인코딩하는 핵산을 포함하는 방법.

### 청구항 3

제 1 항 또는 제 2 항에 있어서, 물질 (2) 가 Klf4 또는 Klf4 를 인코딩하는 핵산을 포함하는 방법.

### 청구항 4

하기 (1) 및 (2):

- (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 균으로부터 선택되는 1 종 이상의 물질,
- (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 균으로부터 선택되는 1 종 이상의 물질을 포함하는, iPS 세포 확립 효율 개선제.

### 청구항 5

제 4 항에 있어서, 물질 (1) 이 GLIS1 또는 GLIS1 을 인코딩하는 핵산을 포함하는 개선제.

### 청구항 6

제 4 항 또는 제 5 항에 있어서, 물질 (2) 가 Klf4 또는 Klf4 를 인코딩하는 핵산을 포함하는 개선제.

### 청구항 7

하기 (1), (2) 및 (3):

- (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 균으로부터 선택되는 1 종 이상의 물질,
- (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 균으로부터 선택되는 1 종 이상의 물질,
- (3) 상기 물질 (1) 및 (2) 와 조합됨으로써, 체세포로부터 iPS 세포를 유도할 수 있는 핵 재프로그래밍 물질을 체세포와 접촉시키는 것을 포함하는, iPS 세포의 생성 방법.

### 청구항 8

제 7 항에 있어서, 물질 (1) 이 GLIS1 또는 GLIS1 을 인코딩하는 핵산을 포함하는 방법.

### 청구항 9

제 7 항 또는 제 8 항에 있어서, 물질 (2) 가 Klf4 또는 Klf4 를 인코딩하는 핵산을 포함하는 방법.

### 청구항 10

제 7 항 내지 제 9 항 중 어느 한 항에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct 패밀리의 일원, Sox 패밀리의 일원, Myc 패밀리의 일원, Lin28 패밀리의 일원, Nanog, 및 이들을 인코딩하는 핵산으로 이루어지는 균으로부터 선택되는 방법.

**청구항 11**

제 7 항 내지 제 9 항 중 어느 한 항에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct3/4 또는 이를 인코딩하는 핵산을 포함하는 방법.

**청구항 12**

제 11 항에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct3/4 및 Sox2 또는 이들을 인코딩하는 핵산을 포함하는 방법.

**청구항 13**

제 11 항에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct3/4, Sox2 및 c-Myc 또는 이들을 인코딩하는 핵산을 포함하는 방법.

**청구항 14**

하기 (1), (2) 및 (3):

- (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- (3) 상기 물질 (1) 및 (2) 와 조합됨으로써 체세포로부터 iPS 세포를 유도할 수 있는 핵 재프로그래밍 물질을 포함하는, 체세포로부터의 iPS 세포 유도제.

**청구항 15**

제 14 항에 있어서, 물질 (1) 이 GLIS1 또는 GLIS1 을 인코딩하는 핵산을 포함하는 유도제.

**청구항 16**

제 14 항 또는 제 15 항에 있어서, 물질 (2) 가 Klf4 또는 Klf4 를 인코딩하는 핵산을 포함하는 유도제.

**청구항 17**

제 14 항 내지 제 16 항 중 어느 한 항에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct 패밀리의 일원, Sox 패밀리의 일원, Myc 패밀리의 일원, Lin28 패밀리의 일원, Nanog, 및 이들을 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 유도제.

**청구항 18**

제 14 항 내지 제 16 항 중 어느 한 항에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct3/4 또는 이를 인코딩하는 핵산을 포함하는 유도제.

**청구항 19**

제 18 항에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct3/4 및 Sox2 또는 이들을 인코딩하는 핵산을 포함하는 유도제.

**청구항 20**

제 18 항에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct3/4, Sox2 및 c-Myc 또는 이들을 인코딩하는 핵산을 포함하는 유도제.

**청구항 21**

하기 (1) 및 (2):

- (1) GLIS 패밀리의 일원을 인코딩하는 외생 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 핵산,
- (2) Klf 패밀리의 일원을 인코딩하는 외생 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 핵산

을 포함하는, iPS 세포.

**청구항 22**

제 21 항에 있어서, 외생 핵산이 게놈에 통합되어 있는 iPS 세포.

**청구항 23**

제 21 항 또는 제 22 항에 따른 iPS 세포를 처리하여 체세포로 분화되도록 유도하는 것을 포함하는, 체세포의 생성 방법.

**청구항 24**

하기 단계 (1) 및 (2):

- (1) 제 7 항 내지 제 13 항 중 어느 한 항에 따른 방법에 의해 iPS 세포를 생성하는 단계, 및
- (2) 상기 단계 (1) 을 통해 수득되는 iPS 세포를 처리하여 체세포로 분화되도록 유도하는 단계를 포함하는, 체세포의 생성 방법.

**청구항 25**

하기 (1) 및 (2):

- (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질의, iPS 세포 확립 효율을 개선하기 위한 용도.

**청구항 26**

GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질의, iPS 세포 확립 효율을 개선하기 위한 용도로서, 이 경우 상기 물질이, Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질과 함께, 체세포와 접촉되는 용도.

**청구항 27**

하기 (1), (2) 및 (3):

- (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- (3) 상기 물질 (1) 및 (2) 와 조합됨으로써 체세포로부터 iPS 세포를 유도할 수 있는 핵 재프로그래밍 물질의, iPS 세포를 생성하기 위한 용도.

**청구항 28**

하기 (1) 및 (2):

- (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질의, iPS 세포를 생성하기 위한 용도로서, 이 경우 상기 인자들이, 상기 물질 (1) 및 (2) 와 조합됨으로써 체세포로부터 iPS 세포를 유도할 수 있는 핵 재프로그래밍 물질과 함께, 체세포와 접촉되는 용도.

**청구항 29**

- (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질의, iPS 세포를 생성하기 위한 용도로서, 이 경우 상기 물질이, (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질, 및 상기 물질 (1) 및 (2) 와 조합됨으로써 체세포로부터

iPS 세포를 유도할 수 있는 핵 재프로그래밍 물질과 함께, 체세포와 접촉되는 용도.

**청구항 30**

제 21 항 또는 제 22 항에 따른 iPS 세포의, 체세포 생성에 있어서의 용도.

**청구항 31**

제 21 항 또는 제 22 항에 있어서, 체세포 생성시 세포 공급원으로서의 역할을 하는 iPS 세포.

**명세서**

**기술분야**

[0001] 본 발명은 유도된 다능성 줄기 세포 (이후 iPS 세포로 언급함) 의 확립 효율 개선 방법 및 이를 위한 시약, 더욱 구체적으로는 GLIS 패밀리의 일원 및 Klf 패밀리의 일원을 사용하는 iPS 세포의 확립 효율 개선 방법, 및 이를 위한 시약 등에 관한 것이다.

**배경기술**

[0002] 최근, 마우스 및 인간 iPS 세포가 차례로 확립되었다. Takahashi 및 Yamanaka 는 네오마이신 내성 유전자가 Fbx15 자리에 녹-인 (knock-in) 되어 있는 리포터 마우스로부터의 섬유모세포 내로 Oct3/4, Sox2, Klf4 및 c-Myc 유전자를 전달하고, 세포가 유전자를 발현하도록 강제함으로써 iPS 세포를 유도했다 [Takahashi, K. 및 Yamanaka, S., *Cell*, 126: 663-676 (2006)]. Okita 등은 Fbx15 의 발현보다 다능성 세포에서 국소적으로 발현되는 Nanog 의 자리에 통합된 녹색 형광 단백질 (GFP) 및 푸로마이신 내성 유전자를 갖는 트랜스제닉 마우스를 창조하고, 마우스로부터의 섬유모세포가 앞서 언급된 4 가지 유전자를 발현하도록 강제하고, 푸로마이신-내성 및 GFP-양성 세포를 선별함으로써, 배아 줄기 (ES) 세포와 거의 동일한 유전자 발현 및 후성적 변이 프로파일을 나타내는 iPS 세포 (Nanog iPS 세포) 를 확립하는데 성공했다 [Okita, K. 등, *Nature*, 448: 313-317 (2007)]. 다른 그룹들에 의해 유사한 결과들이 얻어졌다 [Wernig, M. 등, *Nature*, 448: 318-324 (2007); Maherali, N. 등, *Cell Stem Cell*, 1: 55-70 (2007)]. 그 후, iPS 세포가 c-Myc 유전자 이외의 3 가지 인자로도 생성될 수 있음이 밝혀졌다 [Nakagawa, M. 등, *Nat. Biotechnol.*, 26: 101-106 (2008)].

[0003] 게다가, Takahashi 등 [Takahashi, K. 등, *Cell*, 131: 861-872 (2007)] 은 마우스에서 사용된 것과 동일한 4 가지 유전자를 인간 피부 섬유모세포 내로 도입함으로써 iPS 세포를 확립하는데 성공했다. 한편, Yu 등은 Klf4 및 c-Myc 대신 Nanog 및 Lin28 을 사용하여 인간 iPS 세포를 생성했다 [Yu, J. 등, *Science*, 318: 1917-1920 (2007)]. 그러므로, 정의된 인자를 체세포 내로 전달함으로써, 다능성의 면에서 ES 세포와 비슷한 iPS 세포가 인간 및 마우스 둘다에서 생성될 수 있음이 입증되었다.

[0004] 그 후, TERT 및 SV40 라지 T 항원 (인간 세포 불멸화 유전자로서 알려짐) 을, 4 가지 인자 Oct3/4, Sox2, Klf4 및 c-Myc 와 함께, 전달함으로써 확립된 iPS 세포 [Park, I.H. 등, *Nature*, 451: 141-146 (2008)], 앞서 말한 4 가지 인자에 Nanog 및 Lin28 을 추가하여 확립된 iPS 세포 [Liao, J. 등, *Cell Research*, 18: 600-603 (2008)], 및 앞서 말한 4 가지 인자 또는 c-Myc 이외의 3 가지 인자에 UTF1 을 추가하여 확립된 iPS 세포 [Zhao, Y. 등, *Cell Stem Cell*, 3: 475-479 (2008)] 를 포함하는, iPS 세포 확립의 효율을 증가시키려는 다양한 시도가 있어 왔다. 그러나, 현재 만족스러운 개선이 달성되지 않은 상황이다.

**발명의 내용**

[0005] 본 발명자들은, ES 세포와 같은 다능성 세포에서 특이적으로 발현되는 유전자 중에서 뿐만 아니라, 전사 인자의 광범위한 유전자 라이브러리로부터, Klf4 에 대한 대체물로서, iPS 세포를 확립하는데 사용될 수 있는 유전자를 탐색하는 종합적인 연구를 수행했다. 이에 따라 본 발명자들은 GLIS 패밀리에 속하는 유전자 (예, GLIS1), PTX 패밀리에 속하는 유전자 (예, PITX2), 또는 프롤린-풍부 C-말단이 있는 DMRT-유사 패밀리 B 1 유전자 (DMRTB1) 를, 3 가지 유전자 Oct3/4, Sox2 및 c-Myc 와 함께, 마우스 및 인간 진피 섬유모세포에 전달함으로써, iPS 세포를 효율적으로 확립하는데 성공했고, 이들 전사 인자를 Klf4 를 기능적으로 대체할 수 있는 신규한 핵 재프로그래밍 물질로서 동정하였다 (2009 년 2 월 27 일에 출원된 미국 가 출원 제 61/208,853 호, 및 2009 년 9 월 8 일에 출원된 미국 가 출원 제 61/276,123 호).

- [0006] 그 다음에, 본 발명자들은 iPS 세포의 확립시 Klf4 와 조합되어 사용된 이들 Klf4 대체 인자 GLIS1, PITX2 및 DMRTB1 의 효과를 연구했다. 예상 밖의 결과로서, PITX2 및 DMRTB1 은 Klf4 와 조합되었을 때 부가적 효과를 전혀 나타내지 않았지만, GLIS1 과 Klf4 의 조합된 사용은 마우스 및 인간 세포 둘다에서 iPS 세포의 확립에 대해 극적인 시너지 효과를 초래했다. 본 발명자들은 이들 발견에 기초하여 추가 연구를 수행하여, 본 발명을 완성했다.
- [0007] 따라서, 본 발명은 하기를 제공한다:
- [0008] [1] 하기 (1) 및 (2):
- [0009] (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- [0010] (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질
- [0011] 을 체세포와 접촉시키는 것을 포함하는, iPS 세포 확립 효율의 개선 방법.
- [0012] [2] 상기 [1] 에 있어서, 물질 (1) 이 GLIS 패밀리 징크 핑거 (zinc finger) 1 (GLIS1) 또는 GLIS1 을 인코딩하는 핵산을 포함하는 방법.
- [0013] [3] 상기 [1] 또는 [2] 에 있어서, 물질 (2) 가 Klf4 또는 Klf4 를 인코딩하는 핵산을 포함하는 방법.
- [0014] [4] 하기 (1) 및 (2):
- [0015] (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- [0016] (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질을 포함하는, iPS 세포 확립 효율 개선제.
- [0017] [5] 상기 [4] 에 있어서, 물질 (1) 이 GLIS1 또는 GLIS1 을 인코딩하는 핵산을 포함하는 개선제.
- [0018] [6] 상기 [4] 또는 [5] 에 있어서, 물질 (2) 가 Klf4 또는 Klf4 를 인코딩하는 핵산을 포함하는 개선제.
- [0019] [7] 하기 (1), (2) 및 (3):
- [0020] (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- [0021] (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- [0022] (3) 상기 물질 (1) 및 (2) 와 조합됨으로써, 체세포로부터 iPS 세포를 유도할 수 있는 핵 재프로그래밍 물질을 포함하는, iPS 세포의 생성 방법.
- [0023] [8] 상기 [7] 에 있어서, 물질 (1) 이 GLIS1 또는 GLIS1 을 인코딩하는 핵산을 포함하는 방법.
- [0024] [9] 상기 [7] 또는 [8] 에 있어서, 물질 (2) 가 Klf4 또는 Klf4 를 인코딩하는 핵산을 포함하는 방법.
- [0025] [10] 상기 [7] 내지 [9] 중 어느 하나에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct 패밀리의 일원, Sox 패밀리의 일원, Myc 패밀리의 일원, Lin28 패밀리의 일원, Nanog, 및 이들을 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 방법.
- [0026] [11] 상기 [7] 내지 [9] 중 어느 하나에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct3/4 또는 이를 인코딩하는 핵산을 포함하는 방법.
- [0027] [12] 상기 [11] 에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct3/4 및 Sox2 또는 이들을 인코딩하는 핵산을 포함하는 방법.
- [0028] [13] 상기 [11] 에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct3/4, Sox2 및 c-Myc 또는 이들을 인코딩하는 핵산을 포함하는 방법.
- [0029] [14] 하기 (1), (2) 및 (3):
- [0030] (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- [0031] (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- [0032] (3) 상기 물질 (1) 및 (2) 와 조합됨으로써 체세포로부터 iPS 세포를 유도할 수 있는 핵 재프로그래밍 물질

- [0035] 을 포함하는, 체세포로부터의 iPS 세포 유도제.
- [0036] [15] 상기 [14] 에 있어서, 물질 (1) 이 GLIS1 또는 GLIS1 을 인코딩하는 핵산을 포함하는 유도제.
- [0037] [16] 상기 [14] 또는 [15] 에 있어서, 물질 (2) 가 Klf4 또는 Klf4 를 인코딩하는 핵산을 포함하는 유도제.
- [0038] [17] 상기 [14] 내지 [16] 중 어느 하나에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct 패밀리의 일원, Sox 패밀리의 일원, Myc 패밀리의 일원, Lin28 패밀리의 일원, Nanog, 및 이들을 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 유도제.
- [0039] [18] 상기 [14] 내지 [16] 중 어느 하나에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct3/4 또는 이를 인코딩하는 핵산을 포함하는 유도제.
- [0040] [19] 상기 [18] 에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct3/4 및 Sox2 또는 이들을 인코딩하는 핵산을 포함하는 유도제.
- [0041] [20] 상기 [18] 에 있어서, 핵 재프로그래밍 물질 (3) 이 Oct3/4, Sox2 및 c-Myc 또는 이들을 인코딩하는 핵산을 포함하는 유도제.
- [0042] [21] 하기 (1) 및 (2):
- [0043] (1) GLIS 패밀리의 일원을 인코딩하는 외생 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 핵산,
- [0044] (2) Klf 패밀리의 일원을 인코딩하는 외생 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 핵산
- [0045] 을 포함하는, iPS 세포.
- [0046] [22] 상기 [21] 에 있어서, 외생 핵산이 게놈에 통합되어 있는 iPS 세포.
- [0047] [23] 상기 [21] 또는 [22] 에 따른 iPS 세포를 처리하여 체세포로 분화되도록 유도하는 것을 포함하는, 체세포의 생성 방법.
- [0048] [24] 하기 단계 (1) 및 (2):
- [0049] (1) 상기 [7] 내지 [13] 중 어느 하나에 따른 방법에 의해 iPS 세포를 생성하는 단계, 및
- [0050] (2) 상기 단계 (1) 을 통해 수득되는 iPS 세포를 처리하여 체세포로 분화되도록 유도하는 단계
- [0051] 를 포함하는, 체세포의 생성 방법.
- [0052] [25] 하기 (1) 및 (2):
- [0053] (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- [0054] (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질
- [0055] 의, iPS 세포 확립 효율을 개선하기 위한 용도.
- [0056] [26] GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질의, iPS 세포 확립 효율을 개선하기 위한 용도로서, 이 경우 상기 물질이, Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질과 함께, 체세포와 접촉되는 용도.
- [0057] [27] 하기 (1), (2) 및 (3):
- [0058] (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- [0059] (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- [0060] (3) 상기 물질 (1) 및 (2) 와 조합됨으로써 체세포로부터 iPS 세포를 유도할 수 있는 핵 재프로그래밍 물질
- [0061] 의, iPS 세포를 생성하기 위한 용도.
- [0062] [28] 하기 (1) 및 (2):
- [0063] (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질,
- [0064] (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질

- [0065] 의, iPS 세포를 생성하기 위한 용도로서, 이 경우 상기 인자들이, 상기 물질 (1) 및 (2) 와 조합됨으로써 체세포로부터 iPS 세포를 유도할 수 있는 핵 재프로그래밍 물질과 함께, 체세포와 접촉되는 용도.
- [0066] [29] (1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질의, iPS 세포를 생성하기 위한 용도로서, 이 경우 상기 물질이, (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질, 및 상기 물질 (1) 및 (2) 와 조합됨으로써 체세포로부터 iPS 세포를 유도할 수 있는 핵 재프로그래밍 물질과 함께, 체세포와 접촉되는 용도.
- [0067] [30] 상기 [21] 또는 [22] 에 따른 iPS 세포의, 체세포 생성에 있어서의 용도.
- [0068] [31] 상기 [21] 또는 [22] 에 있어서, 체세포 생성시 세포 공급원으로서의 역할을 하는 iPS 세포.
- [0069] 앞서 언급된 바와 같이, 본 발명의 iPS 세포 확립 효율 개선체는 체세포로부터의 iPS 세포 확립 효율을 현저히 개선할 수 있으므로, 예를 들어, 자가 이식에 의하는 인간 이식 의술에 적용시 유용하다.

**도면의 간단한 설명**

- [0070] 도 1 은 인간 Gateway<sup>®</sup> 엔트리 클론으로부터 기능에 의해 엔트리 클론의 범위를 좁히는 (narrowing down) 단계를 보여주는 도해이다 (N. Goshima 등, Nature methods, 2008).
- 도 2 는 전사 인자의 엔트리 클론으로부터 체세포 재프로그래밍 인자를 스크리닝하기 위한 전사 인자 라이브러리를 제조하는데 사용되는 절차의 개요이다.
- 도 3 은 총 4 가지 상이한 유전자, 즉, 3 가지 유전자 (Oct3/4, Sox2, c-Myc) 및 G06 (유전자 코드 명칭: GLIS1), H08 (유전자 코드 명칭: DMRTB1) 또는 H10 (유전자 코드 명칭: PITX2) 을, Nanog-GFP 마우스 진피 섬유모세포 내로 레트로바이러스를 이용하여 전달함으로써 획득된 GFP-양성 콜로니의 형태의 사진 표현이다. "Klf-G6-1" 은 G06 (유전자 코드 명칭: GLIS1) 을 3 가지 유전자와 함께 전달함으로써 획득된 iPS 세포 클론을 나타내고; "Klf-H8-2" 는 H08 (유전자 코드 명칭: DMRTB1) 을 3 가지 유전자와 함께 전달함으로써 획득된 iPS 세포 클론을 나타내고; "Klf-H10-1" 및 "Klf-H10" 은 H10 (유전자 코드 명칭: PITX2) 을 3 가지 유전자와 함께 전달함으로써 획득된 iPS 세포 클론을 나타낸다. P0 은 콜로니 확립시 찍은 사진을 나타내고; P1 은 1번째 세대 (24 웰) 의 사진을 나타내고; P2 는 2번째 세대 (6 웰) 의 사진을 나타낸다. 3 개의 사진으로 이루어진 각각의 세트에서, 왼쪽 패널은 GFP-양성 콜로니의 이미지를 나타내고, 중앙 패널은 위상차 이미지를 나타내고, 오른쪽 패널은 GFP-양성 콜로니 이미지와 위상차 이미지의 겹쳐진 사진을 나타낸다. Klf-H10-1 만 재과종 방법 (Reseed method) 으로 확립되었고, 다른 것들은 MSTO 방법으로 확립되었다.
- 도 4 는 콜로니 확립시, 총 4 가지 상이한 유전자, 즉, 3 가지 유전자 (Oct3/4, Sox2, c-Myc) 및 F09 (유전자 코드 명칭: IRX6), G06 (유전자 코드 명칭: GLIS1), H08 (유전자 코드 명칭: DMRTB1) 또는 H10 (유전자 코드 명칭: PITX2) 을, Nanog-GFP 마우스 진피 섬유모세포 내로 레트로바이러스를 이용하여 전달함으로써 획득된 GFP-양성 콜로니의 형태의 사진 표현이다. "Klf-F9" 는 F09 (유전자 코드 명칭: IRX6) 를 3 가지 유전자와 함께 전달함으로써 획득된 iPS 세포 클론을 나타내고; "Klf-G6-1" 및 "Klf-G6-2" 는 G06 (유전자 코드 명칭: GLIS1) 을 3 가지 유전자와 함께 전달함으로써 획득된 iPS 세포 클론을 나타내고; "Klf-H8-1" 및 "Klf-H8-2" 는 H08 (유전자 코드 명칭: DMRTB1) 을 3 가지 유전자와 함께 전달함으로써 획득된 iPS 세포 클론을 나타내고; "Klf-H10" 은 H10 (유전자 코드 명칭: PITX2) 을 3 가지 유전자와 함께 전달함으로써 획득된 iPS 세포 클론을 나타낸다. "재과종" 은 재과종 방법에 의해 획득된 결과를 나타내고; "MSTO" 는 MSTO 방법에 의해 획득된 결과를 나타낸다.
- 도 5 는 G6-1 (Klf-G6-1), H8-2 (Klf-H8-2) 및 H10 (Klf-H10) iPS 세포 클론에 대한 계놈-PCR 결과의 사진 표현이며, 여기서 "피부" 는 체세포 공급원으로서 사용된 섬유모세포를 나타내고, "플라스미드" 는 pMXs 에 통합된 각각의 유전자를 증폭시킴으로써 제조된 양성 대조군을 나타낸다.
- 도 6 은 도 5 에 나타낸 것 이외의 H10 (Klf-H10) iPS 세포 클론에 대한 계놈-PCR 결과의 사진 표현이다. 도 6 에서, "피부" 는 체세포 공급원으로서 사용된 섬유모세포를 나타내고, "플라스미드" 는 pMXs 에 통합된 각각의 유전자를 증폭시킴으로써 제조된 양성 대조군을 나타낸다.
- 도 7 은 G6-1 (Klf-G6-1), H8-2 (Klf-H8-2) 및 H10 (Klf-H10) iPS 세포 클론에 대한 RT-PCR 결과의 사진 표현이며, 여기서 "피부" 는 체세포 공급원으로서 사용된 섬유모세포를 나타내고; "ES" 및 "iPS" 는 마우스 ES 세포 및 iPS 세포를 나타내고; "Sox2 RT-" 는 음성 대조군이다.

도 8 은 도 7 에 나타낸 것 이외의 H10 (Klf-H10) iPS 세포 클론에 대한 RT-PCR 결과의 사진 표현이다. 도 8 에서, "피부" 는 체세포 공급원으로서 사용된 섬유모세포를 나타내고; "ES" 및 "iPS" 는 마우스 ES 세포 및 iPS 세포를 나타내고; "Sox2 RT-" 는 음성 대조군이다.

도 9 는 2 가지 인자 (Oct3/4, Sox2) 또는 3 가지 인자 (Oct3/4, Sox2, Klf4) 와 G6 (GLIS1), H8 (DMRTB1) 또는 H10 (PITX2) 와의 조합을, Nanog-GFP 마우스 진피 섬유모세포 내로 전달함으로써 확립된 iPS 세포 (GFP-양성 세포) 의 콜로니를 계수한 결과의 그래프 표현이다. 3 회 (4 번째는 오직 대조군에 대한 것임) 의 독립적 실험들의 결과가 요약되어 있다.

도 10 은 표시된 인자로 형질도입된 피부 섬유모세포로부터의, 감염 22 일 후 Nanog-GFP-양성 콜로니의 수를 나타낸다.

도 11 은 표시된 인자로 형질도입된 피부 섬유모세포로부터의, 감염 22 일 후 Nanog-GFP-양성 콜로니의 비율을 나타낸다. 그래프는 3 회의 독립적 실험들의 평균과 표준 편차 (에러 바 (error bar)) 를 나타낸다. \*\*: p<0.01

도 12 는 피부 섬유모세포 (P0; 계대 0) 로부터의 Nanog-GFP-양성 콜로니를 나타낸다. 형광 이미지 (왼쪽); 위상차 이미지 (중앙); 합친 이미지 (오른쪽)

도 13 은 표시된 인자로 형질도입된 MEF 로부터의, 감염 20 일 후 Nanog-GFP-양성 콜로니의 수를 나타낸다. 감염 3 일 후, 섬유모세포를 피더 세포 상에 재파종했다.

도 14 는 표시된 인자로 형질도입된 MEF 로부터의, 감염 20 일 후 Nanog-GFP-양성 콜로니의 비율을 나타낸다. 그래프는 3 회의 독립적 실험들의 평균과 표준 편차 (에러 바) 를 나타낸다. \*\*: p<0.01

도 15 는 MEF (P0; 계대 0) 로부터의 Nanog-GFP-양성 콜로니를 나타낸다. 형광 이미지 (왼쪽); 위상차 이미지 (중앙); 합쳐진 이미지 (오른쪽)

도 16 은 3 가지 인자 (Oct3/4, Sox2, c-Myc) 와 Klf4 및/또는 G6 (GLIS1) 과의 조합을 성인 진피 섬유모세포 (HDF) 내로 전달함으로써 확립된 iPS 세포 (ES-유사 세포) 의 콜로니를 계수한 결과의 그래프 표현이며, 여기서 "104" 및 "105" 는 피더 세포 상으로 재파종된  $5 \times 10^4$  세포/100 mm 디쉬 (dish), 및  $5 \times 10^5$  세포/100 mm 디쉬에 대한 결과를 각각 나타낸다. 3 회의 독립적 실험들의 결과가 요약되어 있다.

도 17 은 3 가지 인자 (Oct3/4, Sox2, c-Myc) 와 Klf4 및/또는 G6 (GLIS1) 과의 조합을 성인 진피 섬유모세포 (HDF) 내로 전달함으로써 확립된 비-iPS 세포 (비-ES-유사 세포) 의 콜로니를 계수한 결과의 그래프 표현이며, 여기서 "104" 및 "105" 는 피더 세포 상으로 재파종된  $5 \times 10^4$  세포/100 mm 디쉬, 및  $5 \times 10^5$  세포/100 mm 디쉬에 대한 결과를 각각 나타낸다. 3 회의 독립적 실험들의 결과가 요약되어 있다.

도 18 은 Oct3/4, Sox2, c-Myc, Klf4 및 G6 으로 확립된 iPS 콜로니 (ES-유사 콜로니) 의 위상차 이미지의 사진 표현이다.

도 19 는 표시된 인자로 형질도입된 인간 진피 섬유모세포 (위쪽:  $5 \times 10^4$  세포, 아래쪽:  $5 \times 10^5$  세포) 로부터의, 감염 약 30 일 후 ESC-유사 콜로니의 수를 나타낸다.

도 20 은 표시된 인자로 형질도입된 인간 진피 섬유모세포 (위쪽:  $5 \times 10^4$  세포, 아래쪽:  $5 \times 10^5$  세포) 의, 감염 약 30 일 후 ESC-유사 콜로니의 비율을 나타낸다. 그래프는 3 회의 독립적 실험들의 평균과 표준 편차 (에러 바) 를 나타낸다. \*: p<0.05; \*\*: p<0.01

도 21 은 OSK + GLIS1 로 생성된 인간 ESC-유사 콜로니를 나타낸다.

도 22 는 확립된 인간 iPS 클론 내의 형질도입된 유전자의 계놈-PCR 분석을 나타낸다. AHDF: 성인 진피 섬유모세포

도 23 은 OSK + GLIS1 로 생성된 인간 iPSC 내의 ESC-마커 유전자의 RT-PCR 분석을 나타낸다. AHDF: 성인 진피 섬유모세포; 201B7: OSKM 으로 생성된 인간 iPS 클론

도 24 는 DNA 마이크로어레이 (microarray) 로 측정된, OSK + GLIS1 로 생성된 iPSC 와 성인 HDF 사이 (위쪽), 및 OSK + GLIS1-형질도입된 iPSC 와 OSKM-형질도입된 iPSC 사이 (아래쪽) 의 전체적 유전자 발현을 비교하는 산

점도를 나타낸다. 상관 계수 ( $R^2$ ) 를 계산했다.

도 25 는 OSK + GLIS1 로 생성된 인간 iPSC 의 테라토마 형성을 나타낸다.

도 26 은 다양한 마우스 조직에서의 GLIS1 의 발현을 나타낸다. 각각의 마우스 조직으로부터 단리된 총 RNA 를 정량적 RT-PCR 로 조사했다. 그래프는 4 회의 독립적 실험들의 평균과 표준 편차 (에러 바) 를 나타낸다.

도 27 은 GLIS1 shRNA 에 노출된 피부 섬유모세포 내의 내생 GLIS1 의 정량적 RT-PCR 분석을 나타낸다. 그래프는 2 회의 독립적 실험들의 평균과 평균 오차 (에러 바) 를 나타낸다.

도 28 은 3 가지 재프로그래밍 인자 (OSK) 에 의한 iPSC 확립 효율에 대한 GLIS1 shRNA 의 효과를 나타낸다. OSK 를 GLIS1 shRNA 와 함께 또는 GLIS1 shRNA 없이 피부 섬유모세포 내로 형질도입한지 4 주 후, Nanog-GFP-양성 콜로니의 수를 계수했다.

**발명을 실시하기 위한 구체적인 내용**

[0071] 본 발명은 체세포의 핵 재프로그래밍 단계에서, (1) GLIS1 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질, 및 (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질 (이후 본 발명의 확립 효율 개선 인자로도 언급됨) 을 체세포와 접촉시키는 것을 포함하는, iPSC 세포의 확립 효율 개선 방법을 제공한다. 체세포 핵 재프로그래밍은 핵 재프로그래밍 물질을 체세포와 접촉시킴으로써 달성되기 때문에, 본 발명은 또한 (3) 상기 물질 (1) 및 (2) 와 조합됨으로써 체세포로부터 iPSC 세포를 유도할 수 있는 핵 재프로그래밍 물질 (이후 단순히 핵 재프로그래밍 물질로도 언급됨) 을, 상기 물질 (1) 및 (2) 와 함께, 체세포와 접촉시키는 것을 포함하는, iPSC 세포의 생성 방법을 제공한다. 여기서, iPSC 세포가 상기 물질 (3) (핵 재프로그래밍 물질) 단독으로는 확립될 수 없지만, 핵 재프로그래밍 물질이, 본 발명의 iPSC 세포 확립 효율 개선 인자와 함께, 체세포와 접촉될 때에는 확립될 수 있는 경우도 "확립 효율의 개선" 으로 여겨진다.

[0072] (a) 체세포의 공급원

[0073] 포유류 기원 (예, 인간, 마우스, 원숭이, 소, 돼지, 랫트, 개 등) 의 생식선 이외의 임의의 세포가 본 발명에서 iPSC 세포의 생성을 위한 출발 물질로서 사용될 수 있다. 예는, 각질화 상피 세포 (예, 각질화 표피 세포), 점막 상피 세포 (예, 혀의 표층의 상피 세포), 외분비샘 상피 세포 (예, 젖샘 세포), 호르몬 분비 세포 (예, 부신수질), 대사 또는 저장을 위한 세포 (예, 간세포), 접면을 구성하는 내막 상피 세포 (예, 타입 I 폐포 세포), 폐쇄관의 내막 상피 세포 (예, 혈관 내피 세포), 수송 능력이 있는 섬모를 갖는 세포 (예, 기도 상피 세포), 세포의 기질 분비를 위한 세포 (예, 섬유아세포), 골약성 세포 (예, 평활근 세포), 혈액 및 면역계의 세포 (예, T 림프구), 감각 관련 세포 (예, 바실러리 세포 (bacillary cells)), 자율신경계 뉴런 (예, 콜린성 뉴런), 감각 기관의 지주세포 및 말초 뉴런 (예, 위성 세포), 중추 신경계의 신경 세포 및 아교 세포 (예를 들어, 별아교 세포), 색소 세포 (예, 망막 색소 상피 세포), 이들의 전구 세포 (조직 전구 세포) 등을 포함한다. 세포 분화 정도, 세포가 수집되는 동물의 연령 등에는 제한이 없으며; 심지어 미분화 전구 세포 (체세포 줄기 세포를 포함함) 와 최종적으로 분화된 성숙 세포가 모두 비슷하게 본 발명에서 체세포의 공급원으로서 사용될 수 있다. 미분화 전구 세포의 예는 조직 줄기 세포 (체세포 줄기 세포), 예컨대 신경 줄기 세포, 조혈 줄기 세포, 간엽 줄기 세포, 및 치수 줄기 세포를 포함한다.

[0074] 체세포의 공급원으로서 포유류 개체의 선택은 특별히 제한되지 않지만; 수득된 iPSC 세포를 인간에서 재생 의료에 사용하려는 경우, 이식편 거부 방지의 관점에서, 환자 또는 그 환자와 동일한 또는 실질적으로 동일한 HLA 타입을 갖는 또다른 사람으로부터 체세포를 수집하는 것이 특히 바람직하다. 본원에서 사용되는 "실질적으로 동일한 HLA 타입" 은 공여자의 체세포에서 유래하는 iPSC 세포의 분화를 유도함으로써 수득된 이식되는 세포가 면역억제제 등을 이용하여 환자에게 이식했을 때 접촉될 수 있을 정도로 공여자의 HLA 타입이 환자의 HLA 타입과 일치함을 의미한다. 예를 들어, 실질적으로 동일한 HLA 타입은, 주 HLA (예를 들어, HLA-A, HLA-B 및 HLA-DR 의 3 개의 주 유전자좌) 가 동일한 HLA 타입 (이하 동일한 의미가 적용됨) 등을 포함한다. 수득되는 iPSC 세포가 인간에게 투여 (이식) 되지는 않으나, 예를 들어 환자의 약물 감수성 또는 역반응을 평가하는 스크리닝을 위한 세포의 공급원으로서 사용되는 경우에도 마찬가지로, 환자 또는 약물 감수성 또는 역반응과 연관되는 동일한 유전적 다형성이 있는 또다른 사람으로부터 체세포를 수집하는 것이 바람직하다.

[0075] 포유류로부터 단리된 체세포는 핵 재프로그래밍 단계에 적용되기 전에 세포의 선택에 따라 그의 배양에 적합한

그 자체로 공지된 배지를 사용하여 예비배양될 수 있다. 상기 배지의 예는, 이에 제한되지는 않으나, 약 5 내지 20% 의 소 태아 혈청 (FCS) 을 함유하는 최소 기본 배지 (MEM), 둘베코 변형 이글 배지 (DMEM), RPMI1640 배지, 199 배지, F12 배지 등을 포함한다. 예를 들어, 양이온성 리포솜과 같은 전달 시약을 사용하여 체세포를 본 발명의 iPS 세포 확립 효율 개선 인자 및 핵 재프로그래밍 물질 (및 필요한 경우 아래 언급되는 또다른 iPS 세포 확립 효율 개선제) 과 접촉시키는 경우, 배지를 무혈청 배지로 교체하여 전달 효율의 감소를 방지하는 것이 때때로 바람직하다.

[0076] (b) 본 발명의 iPS 세포 확립 효율 개선 인자

[0077] 본 발명에서, GLIS 패밀리는 Gli 전사 인자에 대한 유사성에 따라 명명된, 5 개의 C2H2 (Cys<sub>2</sub>-His<sub>2</sub>-타입) 징크 핑거 영역을 갖는 Kruppel-유사 징크 핑거 패밀리아다 [Glis= Gli similar, Kim, Y.S. 등, *J. Biol. Chem.*, 277(34), 30901-30913 (2002)]. GLIS 패밀리는 배아형성 과정에서 다양한 유전자의 발현을 양성 또는 음성적으로 제어하는 전사 인자로 이루어진다. 이러한 유전자 패밀리의 일원의 예는, 이에 제한되지는 않으나, GLIS 패밀리 징크 핑거 1 (GLIS1), GLIS2, GLIS3 등을 포함하며, GLIS1 이 바람직하다. GLIS1 은 마우스 ES 세포에서 발현되지 않음에 유의한다.

[0078] 본 발명에서 사용되는 GLIS 패밀리의 일원은 임의로 선택되는 포유류 (예, 인간, 마우스, 랫트, 원숭이, 소, 말, 돼지, 개 등) 의 세포 또는 조직 [예, 홍선, 골수, 비장, 뇌, 척수, 심장, 골격근, 신장, 폐, 간, 췌장 또는 전립선의 세포 또는 조직, 이들의 상응하는 전구 세포, 줄기 세포 또는 암 세포 등] 에서 유래하는 단백질 또는 이를 인코딩하는 핵산일 수 있으며, 인간 또는 마우스 세포 또는 조직에서 유래하는 것이 바람직하다.

[0079] 인간 및 마우스 기원의 GLIS 패밀리의 일원의 아미노산 서열 및 cDNA 서열에 대한 정보는 표 1 에 제시된 NCBI 접근 번호를 참조하여 획득될 수 있다. 당업자는 cDNA 서열 정보에 기초하여 용이하게 각각의 단백질을 인코딩하는 핵산을 단리하고, 필요한 경우 재조합 단백질을 생성할 수 있다.

[0080] 표 1

유전자 코드 명칭	인간		마우스	
	cDNA	단백질	cDNA	단백질
GLIS1	NM_147193 (SEQ ID NO:1)	NP_671726 (SEQ ID NO:2)	NM_147221 (SEQ ID NO:3)	NP_671754 (SEQ ID NO:4)
GLIS2	NM_032575	NP_115964	NM_031184	NP_112461
GLIS3	NM_001042413	NP_001035878	NM_175459	NP_780668

[0081] 위에 제시된 각각의 아미노산 서열에 대한 동일성이 90% 이상, 바람직하게는 95% 이상, 더욱 바람직하게는 98% 이상, 특히 바람직하게는 99% 이상이고, 야생형 단백질과 동등한 iPS 세포 확립 효율 개선 효과를 보유하는 천연 또는 인공 돌연변이 단백질, 및 이를 인코딩하는 핵산도, 본 발명의 iPS 세포 확립 효율 개선 인자로서 이용될 수 있다. 여기서, iPS 세포 확립 효율 개선 효과는 오직 명시된 재프로그래밍 인자 (예, 2 가지 인자 Oct3/4 및 Sox, 2 가지 인자 및 c-Myc 로 이루어지는 3 가지 인자, 등) 만 체세포에 전달되는 경우와, 재프로그래밍 인자를 전달할 뿐만 아니라 본 발명의 iPS 세포 확립 효율 개선 인자가 체세포와 접촉되는 경우 사이에서, 출현하는 iPS 세포 콜로니의 수를 비교함으로써 입증될 수 있다.

[0083] 본 발명의 GLIS 패밀리의 일원 및 이를 인코딩하는 핵산에 관하여, 그 패밀리에 속하는 인자 중 임의의 1 가지가 단독으로 사용될 수 있고, 2 가지 이상이 조합되어 사용될 수 있다.

[0084] Klf (Kruppel-like factor) 패밀리는 다양한 생물학적 과정 예컨대 증식, 분화, 발생, 및 세포자멸사를 제어하는 전사 인자로 이루어지지만 [McConnell, B.B. 등, *Bioassays*, 29: 549-557 (2007)], 이들의 기능은 여전히 구체적으로 명확하지 않다. 이러한 유전자 패밀리의 일원의 예는, 이에 제한되지는 않으나, Klf1, Klf2, Klf4, Klf5 등을 포함하며, Klf4 가 바람직하다. 앞서 언급된 바와 같이, GLIS 패밀리는 5 개의 C2H2 타입 징크 핑거 영역을 갖지만, Klf 패밀리는 3 개의 C2H2 타입 징크 핑거 영역을 갖는다.

[0085] Yamanaka 등은 동일한 4 가지 유전자 (Oct3/4, Sox2, Klf4 및 c-Myc) 가 동일한 각각의 패밀리에 속하는 다른 유전자로 치환될 수 있다는 가설을 세웠고, Klf4 가 Klf1, Klf2 또는 Klf5 로 대체되는 경우에도 iPS 세포가 확립될 수 있음을 밝혔다 [WO 2007/069666 A1; Nakagawa, M. 등, *Nat. Biotechnol.*, 26: 101-106 (2008)]. ES 세포를 레티노산으로 처리하여 그의 분화를 유도하는 경우, Klf4 뿐만 아니라, Klf2 및 Klf5 도 그의 발현을 감소시켰다. 이러한 사실에 주목하여, Jiang 등의 그룹은 최근에 Klf2, Klf4 및 Klf5 를 동시에 녹 다운

(knocked down) 시켰고, ES 세포에서 분화가 유도되었음을 발견하여, ES 세포에서 Klf 패밀리의 일원 중 적어도 일부, 예컨대 Klf2 및 Klf5 가 Klf4 를 기능적으로 대체할 수 있음을 보였다 [Jiang, J. 등, Nat. Cell Biol., 10: 353-360 (2008)]. 이들은 계속해서 Klf2 또는 Klf5 유전자, 또는 다른 전사 인자 또는 후성적 조절 인자를, 3 가지 유전자 Oct3/4, Sox2 및 c-Myc 와 함께, MEF 내로 전달하여, Klf2 및 Klf5 가 Klf4 를 대체할 수 있음을 확인했고, 에스트로겐 수용체를 닮은 고아 핵 수용체인 Esrrb 도 또한 Klf4 를 대체할 수 있음을 발견했다 [Feng, B. 등, Nat. Cell Biol., 11: 197-203 (2009)]. 이들 발견은 Klf1, Klf2, Klf5, 및 심지어 Esrrb 도 또한 Klf4 의 효과를 보유한다는 생각을 초래했으며 이는 본원에서 제공되는 실시예에서 확인된다 (GLIS 패밀리와 조합되어 사용되는 경우 iPS 세포 확립 효율을 개선함).

[0086] 본 발명에서 사용되는 Klf 패밀리의 일원은 임의로 선택되는 포유류 (예, 인간, 마우스, 랫트, 원숭이, 소, 말, 돼지, 개 등) 의 세포 또는 조직 [예, 홍선, 골수, 비장, 뇌, 척수, 심장, 골격근, 신장, 폐, 간, 췌장 또는 전립선의 세포 또는 조직, 이들의 상응하는 전구 세포, 줄기 세포 또는 암 세포 등] 에서 유래하는 단백질 또는 이를 인코딩하는 핵산일 수 있으며, 인간 또는 마우스 기원의 것이 바람직하다.

[0087] 인간 및 마우스 기원의 Klf 패밀리의 일원의 아미노산 서열 및 cDNA 서열에 대한 정보는 표 2 에 제시된 NCBI 접근 번호를 참조하여 획득될 수 있다. 당업자는 cDNA 서열 정보에 기초하여 각각의 단백질을 인코딩하는 핵산을 용이하게 단리할 수 있고, 필요한 경우 재조합 단백질을 생성할 수 있다.

[0088] 표 2

유전자 코드 명칭	인간		마우스	
	cDNA	단백질	cDNA	단백질
Klf1	NM_006563	NP_006554	NM_010635	NP_034765
Klf2	NM_016270	NP_057354	NM_008452	NP_032478
Klf4	NM_004235 (SEQ ID NO:5)	NP_004226 (SEQ ID NO:6)	NM_010637 (SEQ ID NO:7)	NP_034767 (SEQ ID NO:8)
Klf5	NM_001730	NP_001721	NM_009769	NP_033899

[0089] 위에 제시된 각각의 아미노산 서열에 대한 동일성이 90% 이상, 바람직하게는 95% 이상, 더욱 바람직하게는 98% 이상, 특히 바람직하게는 99% 이상이고, 야생형 단백질과 동등한 iPS 세포 확립 효율 개선 효과를 보여주는 천연 또는 인공 돌연변이 단백질, 및 이를 인코딩하는 핵산도, 본 발명의 iPS 세포 확립 효율 개선 인자로서 이용될 수 있다.

[0091] 본 발명의 Klf 패밀리의 일원 및 이를 인코딩하는 핵산에 관하여, 그 패밀리에 속하는 인자 중 임의의 1 가지가 단독으로 사용될 수 있고, 2 가지 이상이 조합되어 사용될 수 있다.

[0092] 핵 재프로그래밍을 겪는 체세포가, 앞서 기술된 본 발명의 iPS 세포 확립 효율 개선 인자인, GLIS 패밀리의 일원, 또는 Klf 패밀리의 일원 중 임의의 하나의 구성요소 하나 이상을, 확립 효율 개선에 충분한 수준으로, 내생적으로 발현한다고 가정할 때, 내생적으로 발현되는 구성요소를 배제한 나머지 구성요소만의 조합도 또한 본 발명에서 "iPS 세포 확립 효율 개선 인자" 의 범위에 포함될 수 있다.

[0093] 단백질 형태의 본 발명의 iPS 세포 확립 효율 개선 인자를 체세포에 전달하는 것은 세포 내로 단백질을 전달하는 그 자체로 공지된 방법을 사용하여 달성될 수 있다. 상기 방법은 예를 들어 단백질 전달 시약을 사용하는 방법, 단백질 전달 도메인 (PTD) 또는 세포 침투 펩티드 (CPP) 융합 단백질을 사용하는 방법, 미세주입법 등을 포함한다. 단백질 전달 시약은 시판되며, 양이온성 지질 기재의 것, 예컨대 BioPOTER 단백질 운반 시약 (Gene Therapy System), Pro-Ject™ 단백질 트랜스펙션 시약 (PIERCE) 및 ProVectin (IMGENEX); 지질 기재의 것, 예컨대 Profect-1 (Targeting Systems); 막투과 펩티드 기재의 것, 예컨대 Penetrain Peptide (Q biogene) 및 Chariot Kit (Active Motif), HVJ 외피 (일본의 불활성화된 적혈구 응집 바이러스) 를 이용하는 GenomONE (ISHIHARA SANGYO KAISHA, LTD.) 등을 포함한다. 전달은 이들 시약에 첨부된 프로토콜에 따라 달성될 수 있으며, 통상적인 절차는 아래 기재된 바와 같다. 본 발명의 단백질성 iPS 세포 확립 효율 개선 인자가 적당한 용매 (예를 들어, PBS 또는 HEPES 와 같은 완충 용액) 에 희석되고, 전달 시약이 첨가되고, 혼합물이 실온에서 약 5 내지 15 분 동안 인큐베이션되어 복합체를 형성하고, 이러한 복합체가 배지를 무혈청 배지로 교환한

후 세포에 첨가되고, 세포가 37°C 에서 1 시간 내지 수 시간 동안 인큐베이션된다. 그 후, 배지가 제거되고, 혈청을 함유하는 배지로 대체된다.

- [0094] 개발된 PTD 는, 단백질의 세포의 도메인을 이용하는 것 예컨대 초파리-유래 AntP, HIV-유래 TAT (Frankel, A. 등, *Cell* 55, 1189-93 (1988) 또는 Green, M. & Loewenstein, P. M. *Cell* 55, 1179-88 (1988)), Penetratin (Derossi, D. 등, *J. Biol. Chem.* 269, 10444-50 (1994)), Buforin II (Park, C. B. 등 *Proc. Natl Acad. Sci. USA* 97, 8245-50 (2000)), Transportan (Pooga, M. 등 *FASEB J.* 12, 67-77 (1998)), MAP (모델 양친매성 펩티드) (Oehlke, J. 등 *Biochim. Biophys. Acta.* 1414, 127-39 (1998)), K-FGF (Lin, Y. Z. 등 *J. Biol. Chem.* 270, 14255-14258 (1995)), Ku70 (Sawada, M. 등 *Nature Cell Biol.* 5, 352-7 (2003)), 프리온 (Prion) (Lundberg, P. 등 *Biochem. Biophys. Res. Commun.* 299, 85-90 (2002)), pVEC (Elmquist, A. 등 *Exp. Cell Res.* 269, 237-44 (2001)), Pep-1 (Morris, M. C. 등 *Nature Biotechnol.* 19, 1173-6 (2001)), Pep-7 (Gao, C. 등 *Bioorg. Med. Chem.* 10, 4057-65 (2002)), SynB1 (Rousselle, C. 등 *Mol. Pharmacol.* 57, 679-86 (2000)), HN-I (Hong, F. D. & Clayman, G L. *Cancer Res.* 60, 6551-6 (2000)), 및 HSV-유래 VP22 를 포함한다. PTD 에서 유래하는 CPP 는, 폴리아르기닌 예컨대 11R (*Cell Stem Cell*, 4, 381-384 (2009)) 및 9R (*Cell Stem Cell*, 4, 472-476 (2009)) 을 포함한다.
- [0095] 본 발명의 iPS 세포 확립 효율 개선 인자의 cDNA 및 PTD 서열 또는 CPP 서열을 포함하는 융합 단백질 발현 벡터가 제조되고, 이 벡터를 사용하여 재조합 발현이 수행된다. 융합 단백질이 회수되고 전달에 사용된다. 전달은 단백질 전달 시약이 첨가되지 않는 점을 제외하고는 상기와 동일한 방식으로 수행될 수 있다.
- [0096] 팁 직경이 약 1  $\mu$ m 인 유리 바늘에 단백질 용액을 넣고, 이 용액을 세포 내로 주입하는 방법인 마이크로인젝션은 단백질을 세포로 확실히 전달한다.
- [0097] 다른 유용한 단백질 전달 방법은 전기천공법, 반-무손상 (semi-intact) 세포 방법 [Kano, F. 등 *Methods in Molecular Biology*, Vol. 322, 357-365 (2006)], Wr-t 펩티드를 사용하는 전달 [Kondo, E. 등, *Mol. Cancer Ther.* 3(12), 1623-1630 (2004)] 등을 포함한다.
- [0098] 단백질 전달 작업은 하나 이상의 임의로 선택되는 횟수 (예를 들어, 1 회 이상 내지 10 회 이하, 또는 1 회 이상 내지 5 회 이하 등) 로 수행될 수 있다. 바람직하게는, 전달 작업은 2 회 이상 (예를 들어, 3 회 또는 4 회) 반복적으로 수행될 수 있다. 반복되는 전달 작업에 대한 시간 간격은 예를 들어 6 내지 48 시간, 바람직하게는 12 내지 24 시간이다.
- [0099] iPS 세포 확립 효율이 강조되는 경우, 본 발명의 iPS 세포 확립 효율 개선 인자가 단백질로서가 아니라, 이를 인코딩하는 핵산 형태로 사용되는 것이 바람직하다. 핵산은 DNA 또는 RNA 일 수 있고, DNA/RNA 키메라일 수 있고, DNA 가 바람직하다. 핵산은 이중 가닥 또는 단일 가닥일 수 있다. 이중 가닥인 경우에, 이중 가닥 DNA, 이중 가닥 RNA, 또는 DNA/RNA 하이브리드일 수 있다. 바람직하게는, 핵산은 이중 가닥 DNA, 특히 cDNA 이다.
- [0100] 본 발명의 핵산-기재 iPS 세포 확립 효율 개선 인자는, 예를 들어, 인간 또는 다른 포유류 (예, 마우스, 랫트, 원숭이, 돼지, 개 등) 의 세포 또는 조직 [예, 홍선, 골수, 비장, 뇌, 척수, 심장, 골격근, 신장, 폐, 간, 체장 또는 전립선의 세포 또는 조직, 이들의 상응하는 전구 세포, 줄기 세포 또는 암 세포 등] 에서 유래하는 cDNA 로부터 종래의 방법에 따라 클로닝될 수 있다.
- [0101] 본 발명의 iPS 세포 확립 효율 개선 인자를 체세포에 전달하는 것은 세포에 유전자를 전달하는 그 자체로 공지된 방법을 사용하여 달성될 수 있다. 본 발명의 iPS 세포 확립 효율 개선 인자를 인코딩하는 핵산이, 숙주 체세포에서 기능할 수 있는 프로모터를 포함하는 적당한 발현 벡터 내로 삽입된다. 유용한 발현 벡터는, 예를 들어, 바이러스 벡터 예컨대 레트로바이러스, 렌티바이러스, 아데노바이러스, 아데노-관련 바이러스, 헤르페스 바이러스 및 센다이 (Sendai) 바이러스, 동물 세포에서의 발현을 위한 플라스미드 (예를 들어, pA1-11, pXT1, pRc/CMV, pRc/RSV, pcDNA1/Neo) 등을 포함한다.
- [0102] 사용되는 벡터의 유형은 수득되는 iPS 세포의 의도되는 용도에 따라 적당히 선택될 수 있다. 유용한 벡터는 아데노바이러스 벡터, 플라스미드 벡터, 아데노-관련 바이러스 벡터, 레트로바이러스 벡터, 렌티바이러스 벡터, 센다이 바이러스 벡터 등을 포함한다.
- [0103] 발현 벡터에 사용되는 프로모터의 예는, EF1  $\alpha$  프로모터, CAG 프로모터, SR  $\alpha$  프로모터, SV40 프로모터, LTR 프로모터, CMV (사이토메갈로바이러스) 프로모터, RSV (라우스 육종 바이러스) 프로모터, MoMuLV (몰로니 마우스 백혈병 바이러스) LTR, HSV-TK (헤르페스 심플렉스 바이러스 티미딘 키나제) 프로모터 등을 포함하고, EF1  $\alpha$  프

로모터, CAG 프로모터, MoMuLV LTR, CMV 프로모터, SR $\alpha$  프로모터 등이 바람직하다.

- [0104] 발현 벡터는 필요한 경우, 프로모터 이외에, 인헨서, 폴리아데닐화 신호, 선별가능한 마커 유전자, SV40 복제기원 등을 포함할 수 있다. 선별가능한 마커 유전자의 예는, 디히드로폴레이트 리덕타제 유전자, 네오마이신 내성 유전자, 푸로마이신 내성 유전자 등을 포함한다.
- [0105] 본 발명의 iPS 세포 확립 효율 개선 인자를 인코딩하는 핵산에 관하여, 임의의 하나가 발현 벡터에 단독으로 통합될 수 있고, 몇몇이 조합되어 하나의 발현 벡터에 통합될 수 있다. 게다가, 핵산(들)은 하나 이상의 재프로그래밍 유전자와 함께 하나의 발현 벡터에 통합될 수 있다.
- [0106] 앞서 기술된 절차에서, 본 발명의 iPS 세포 확립 효율 개선 인자 및 재프로그래밍 인자의 유전자가 조합되어 하나의 발현 벡터에 통합되는 경우, 이들 유전자는 바람직하게는 폴리시스트론성 발현을 가능케 하는 서열을 통해 발현 벡터에 통합될 수 있다. 폴리시스트론성 발현을 가능케 하는 서열을 사용하면 하나의 발현 벡터에 통합된 복수의 유전자를 더욱 효율적으로 발현하는 것이 가능해진다. 폴리시스트론성 발현을 가능케 하는 유용한 서열은, 예를 들어, 구체역 바이러스의 2A 서열 (SEQ ID NO:9; PLoS ONE 3, e2532, 2008, Stem Cells 25, 1707, 2007), IRES 서열 (미국 특허 제 4,937,190 호) 등을 포함하며, 2A 서열이 바람직하다.
- [0107] 본 발명의 iPS 세포 확립 효율 개선 인자를 인코딩하는 핵산을 포함하는 발현 벡터는 벡터의 선택에 따라 그 자체로 공지된 기술에 의해 세포 내로 도입될 수 있다. 바이러스 벡터의 경우, 예를 들어, 핵산을 함유하는 플라스미드가 적당한 패키징 세포 (예를 들어, Plat-E 세포) 또는 상보 세포주 (예를 들어, 293-세포) 내로 도입되고, 배양 상청액에서 생성된 바이러스 벡터가 회수되고, 상기 벡터가 바이러스 벡터에 적합한 방법에 의해 세포에 감염된다. 예를 들어, 레트로바이러스 벡터를 사용하는 구체적인 수단이 WO2007/69666, *Cell*, 126, 663-676 (2006) 및 *Cell*, 131, 861-872 (2007) 에 개시되어 있다. 렌티바이러스 벡터를 사용하는 구체적인 수단이 *Science*, 318, 1917-1920 (2007) 에 개시되어 있다. iPS 세포가 재생 의료용 세포 공급원으로서 이용되는 경우, 본 발명의 iPS 세포 확립 효율 개선 인자의 발현 (재활성화) 또는 외생 유전자가 통합되어 있는 자리 근처에 존재하는 내생 유전자의 활성화는 iPS 세포 유래의 분화된 세포로부터 재생된 조직에서 발암 위험을 잠재적으로 증가시킨다. 그러므로, 본 발명의 iPS 세포 확립 효율 개선 인자를 인코딩하는 핵산은 바람직하게는 세포의 염색체에 통합되지 않고 일시적으로 발현된다. 이러한 관점에서, 염색체에 통합되는 경우가 드문 아데노바이러스 벡터를 사용하는 것이 바람직하다. 아데노바이러스 벡터를 사용하는 구체적인 수단이 *Science*, 322, 945-949 (2008) 에 개시되어 있다. 아데노-관련 바이러스 벡터도 염색체에 통합되는 빈도가 낮고 세포 독성 및 염증-유발성이 아데노바이러스 벡터보다 낮기 때문에, 또다른 바람직한 벡터로서 언급될 수 있다. 쉰다이 바이러스 벡터는 염색체 외부에서 안정적으로 존재할 수 있고, 필요한 경우 siRNA 를 사용하여 분해 및 제거될 수 있기 때문에, 또한 바람직하게 이용된다. 쉰다이 바이러스 벡터에 관하여, *J. Biol. Chem.*, 282, 27383-27391 (2007), *Proc. Jpn. Acad., Ser. B* 85, 348-362 (2009) 또는 JP-B-3602058 에 기재된 것이 사용될 수 있다.
- [0108] 레트로바이러스 벡터 또는 렌티바이러스 벡터가 사용되는 경우, 심지어 전이유전자 (transgene) 의 침묵이 발생한 경우에도, 재활성화될 수 있다. 그러므로, 바람직하게는, 예를 들어, 본 발명의 iPS 세포 확립 효율 개선 인자를 인코딩하는 핵산이 불필요해졌을 때 Cre-loxP 시스템을 사용하여 잘려 나가는 방법이 사용될 수 있다. 즉, 미리 핵산의 양쪽 말단에 loxP 서열을 배열하여, iPS 세포가 유도된 후, 플라스미드 벡터 또는 아데노바이러스 벡터를 사용하여 Cre 재조합효소가 세포에 작용하도록 하여, loxP 서열 사이에 끼여 있는 영역이 잘려나갈 수 있다. LTR U3 영역의 인헨서-프로모터 서열은 삽입 돌연변이에 의해 인근에 있는 숙주 유전자를 상향조절할 수 있기 때문에, 그 서열을 결실시키거나, 그 서열을 폴리아데닐화 서열 예컨대 SV40 의 폴리아데닐화 서열로 치환시킴으로써 제조되는 3'-자가-불활성화 (SIN) LTR 을 사용하여, 잘려 나가지 않고 게놈에 남아 있는 loxP 서열 바깥쪽에 있는 LTR 에 의한 내인성 유전자의 발현 조절을 회피하는 것이 더욱 바람직하다. Cre-loxP 시스템 및 SIN LTR 을 사용하는 구체적인 수단이 Chang et al., *Stem Cells*, 27: 1042-1049 (2009) 등에 개시되어 있다.
- [0109] 한편, 비-바이러스 벡터인 플라스미드 벡터는 리포펙틴법, 리포솜법, 전기천공법, 인산칼슘 공동침전법, DEAE 텍스트라판법, 미세주입법, 유전자 총 (gene gun) 법 등을 사용하여 세포 내로 전달될 수 있다. 플라스미드를 벡터로서 사용하는 구체적인 수단이, 예를 들어, *Science*, 322, 949-953 (2008) 등에 기재되어 있다.
- [0110] 플라스미드 벡터, 아데노바이러스 벡터 등이 사용되는 경우, 트랜스팩션은 1 회 이상의 임의로 선택되는 횟수 (예를 들어, 1 회 내지 10 회, 1 회 내지 5 회 등) 로 수행될 수 있다. 두 종류 이상의 발현 벡터가 체세포 내로 도입되는 경우, 이들 모든 종류의 발현 벡터가 동시에 체세포 내로 도입되는 것이 바람직하나; 이러한 경

우에도, 트랜스팩션은 1 회 이상의 임의로 선택되는 횟수 (예를 들어, 1 회 내지 10 회, 1 회 내지 5 회 등) 로 수행될 수 있으며, 바람직하게는 트랜스팩션은 2 회 이상 (예를 들어, 3 회 또는 4 회) 반복적으로 수행될 수 있다.

- [0111] 아데노바이러스 또는 플라스미드가 사용되는 경우에도, 전이유전자가 염색체에 통합될 수 있으므로, 서던 블롯 또는 PCR 에 의해, 유전자가 염색체 내로 삽입되지 않았음을 최종적으로 확인할 필요가 있다. 이러한 이유로, 앞서 언급한 Cre-loxP 시스템과 같이, 전이유전자가 염색체에 통합된 후 유전자가 제거되는 수단을 사용하는 것이 유리할 수 있다. 또다른 바람직한 방식의 구현예에서, 트랜스포손을 사용하여 전이유전자가 염색체에 통합된 후, 플라스미드 벡터 또는 아데노바이러스 벡터를 사용하여 트랜스포사제가 세포에 작용하도록 하여, 염색체로부터 전이유전자를 완전히 제거하는 방법이 사용될 수 있다. 바람직한 트랜스포손의 예로서, 인시목 곤충에서 유래하는 트랜스포손인 piggyBac 등이 언급될 수 있다. piggyBac 트랜스포손을 사용하는 구체적인 수단이 Kaji, K. et al., *Nature*, 458: 771-775 (2009), Woltjen et al., *Nature*, 458: 766-770 (2009) 에 개시되어 있다.
- [0112] 또다른 바람직한 비-통합 유형 벡터는 염색체 외부에서 자가-복제할 수 있는 에피솜 벡터이다. 에피솜 벡터를 사용하는 구체적인 수단이 Yu et al., *Science*, 324, 797-801 (2009) 에 개시되어 있다. 필요한 경우, 에피솜 벡터의 복제에 필수적인 벡터 성분의 5' 및 3' 측에 동일한 방향으로 위치하는 loxP 서열을 갖는 에피솜 벡터 내로 본 발명의 iPS 세포 확립 효율 개선 인자를 인코딩하는 핵산을 삽입함으로써 발현 벡터가 구축될 수 있고, 이러한 발현 벡터가 체세포 내로 전달될 수 있다.
- [0113] 에피솜 벡터의 예는 자가-복제에 필수적인 EBV, SV40 등에서 유래하는 서열을 벡터 성분으로서 포함하는 벡터를 포함한다. 자가-복제에 필수적인 벡터 성분은 구체적으로는 복제 기원 및 복제 기원에 결합하여 복제를 조절하는 단백질을 인코딩하는 유전자로 예시되며; 그 예는 복제 기원 oriP 및 EBV 에 대한 EBNA-1 유전자 및 복제 기원 ori 및 SV40 에 대한 SV40 라지 T 항원 유전자를 포함한다.
- [0114] 에피솜 발현 벡터는 본 발명의 iPS 세포 확립 효율 개선 인자를 인코딩하는 핵산의 전사를 제어하는 프로모터를 함유한다. 사용되는 프로모터는 앞서 기술된 바와 같을 수 있다. 에피솜 발현 벡터는, 앞서 기술된 바와 같이, 필요한 경우, 인헨서, 폴리아데닐화 신호, 선별 마커 유전자 등을 추가로 함유할 수 있다. 선별 마커 유전자의 예는 디히드로폴레이트 리덕타제 유전자, 네오마이신 내성 유전자 등을 포함한다.
- [0115] 본 발명에서 유용한 loxP 서열은, 박테리오파지 P1-유래 야생형 loxP 서열 (SEQ ID NO:10) 에 더하여, 전이유전자의 복제에 필수적인 벡터 성분에 인접하는 (flanking) 위치에 동일한 방향으로 위치할 때 재조합에 의해 loxP 서열에 인접하는 서열을 결실시킬 수 있는 임의로 선택되는 돌연변이 loxP 서열을 포함한다. 이러한 돌연변이 loxP 서열의 예는 5' 반복서열 (repeat) 에서 돌연변이화된 lox71 (SEQ ID NO:11), 3' 반복서열에서 돌연변이화된 lox66 (SEQ ID NO:12), 및 스페이서 부위에서 돌연변이화된 lox2272 및 lox511 을 포함한다. 벡터 성분의 5' 및 3' 측에 위치하는 2 개의 loxP 서열은 동일 또는 상이할 수 있지만, 스페이서 부위에서 돌연변이화된 2 개의 돌연변이 loxP 서열은 동일해야 한다 (예, 한 쌍의 lox2272 서열, 한 쌍의 lox511 서열). 5' 반복서열에서 돌연변이화된 돌연변이 loxP 서열 (예, lox71) 및 3' 반복서열에서 돌연변이화된 돌연변이 loxP 서열 (예, lox66) 의 조합이 바람직하다. 이러한 경우, 재조합의 결과로서 염색체에 남은 loxP 서열은 5' 측 및 3' 측의 반복서열에 이중 돌연변이를 가지므로, Cre 재조합효소에 의해 인식되지 않을 것이며, 따라서 원치 않는 재조합으로 인해 염색체에 결실 돌연변이를 야기할 위험이 감소된다. 돌연변이 loxP 서열 lox71 및 lox66 이 조합되어 사용되는 경우, 각각은 앞서 언급된 벡터 성분의 5' 및 3' 측 중 임의의 측에 위치할 수 있으나, 돌연변이화된 자리가 각각의 loxP 서열의 바깥쪽 말단에 위치하도록 하는 방향으로 돌연변이 loxP 서열이 삽입되어야 한다.
- [0116] 2 개의 loxP 서열 각각은 전이유전자의 복제에 필수적인 벡터 성분 (즉, 복제 기점, 또는 복제 기점에 결합하여 복제를 제어하는 단백질을 인코딩하는 유전자 서열) 의 5' 및 3' 측에 동일한 방향으로 위치한다. loxP 서열에 인접하는 벡터 성분은 복제 기점 또는 복제 기점에 결합하여 복제를 제어하는 단백질을 인코딩하는 유전자 서열, 또는 둘다일 수 있다.
- [0117] 에피솜 벡터는, 예를 들어, 리포팩션 방법, 리포솜 방법, 전기천공법, 인산 칼슘 공동침전법, DEAE 텍스트란법, 미세주입법, 유전자 총 방법 등을 사용하여 세포 내로 도입될 수 있다. 구체적으로는, 예를 들어, *Science*, 324: 797-801 (2009) 및 다른 곳에 기재된 방법이 사용될 수 있다.
- [0118] 전이유전자의 복제에 필수적인 벡터 성분이 iPS 세포로부터 제거되었는지 여부는, 프로브 또는 프라이머로서

loxP 서열의 인근에 및/또는 벡터 성분 내에 있는 뉴클레오티드 서열을 포함하는 핵산을, 주형으로서 iPS 세포로부터 단리된 에피솜 분획물과 함께, 사용하여 서던 블롯 분석 또는 PCR 분석을 수행하고, 밴드의 존재 또는 부재 또는 검출된 밴드의 길이를 측정함으로써 확인할 수 있다. 에피솜 분획물은 본 기술분야에서의 명백한 방법에 의해 제조될 수 있다; 예를 들어, *Science*, 324: 797-801 (2009) 등에 기재된 방법이 사용될 수 있다.

- [0119] (c) 핵 재프로그래밍 물질
- [0120] 본 발명에서, "핵 재프로그래밍 물질" 은 체세포에 전달되었을 때, 또는 본 발명의 확립 효율 개선 인자 [(1) GLIS 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질, 및 (2) Klf 패밀리의 일원 및 이를 인코딩하는 핵산으로 이루어지는 군으로부터 선택되는 1 종 이상의 물질] 와 함께 체세포와 접촉되었을 때, 체세포로부터 iPS 세포를 유도할 수 있는 임의의 물질(들)을 언급하며, 단백질성 인자 또는 이를 인코딩하는 핵산 (벡터에 통합된 형태를 포함함), 또는 저분자량 화합물과 같은 임의의 물질로 이루어질 수 있다. 단백질성 인자 또는 이를 인코딩하는 핵산인 공지된 핵 재프로그래밍 물질로서, 예를 들어, 하기 조합이 바람직하다 (이후, 단백질성 인자에 대한 명칭만 제시됨).
- [0121] (1) Oct3/4, Klf4, c-Myc
- [0122] (2) Oct3/4, Klf4, c-Myc, Sox2 (Sox2 는 Sox1, Sox3, Sox15, Sox17 또는 Sox18 로 대체가능함; Klf4 는 Klf1, Klf2 또는 Klf5 로 대체가능함; c-Myc 는 T58A (활성 돌연변이체), 또는 L-Myc 로 대체가능함)
- [0123] (3) Oct3/4, Klf4, c-Myc, Sox2, Fbx15, Nanog, ERas, Tc11
- [0124] (4) Oct3/4, Klf4, c-Myc, Sox2, TERT, SV40 라지 T 항원 (이후 SV40LT)
- [0125] (5) Oct3/4, Klf4, c-Myc, Sox2, TERT, HPV16 E6
- [0126] (6) Oct3/4, Klf4, c-Myc, Sox2, TERT, HPV16 E7
- [0127] (7) Oct3/4, Klf4, c-Myc, Sox2, TERT, HPV16 E6, HPV16 E7
- [0128] (8) Oct3/4, Klf4, c-Myc, Sox2, TERT, Bmi1
- [0129] [상기 제시된 인자에 대한 더 많은 정보에 대해, WO 2007/069666 참조 (상기 조합 (2) 에서 Sox2 를 Sox18 로 대체하고 Klf4 를 Klf1 또는 Klf5 로 대체하는 것에 대한 정보에 대해, *Nature Biotechnology*, 26, 101-106 (2008) 참조); 조합 "Oct3/4, Klf4, c-Myc, Sox2" 에 대해, 또한 *Cell*, 126, 663-676 (2006), *Cell*, 131, 861-872 (2007) 등을 참조; 조합 "Oct3/4, Klf2 (또는 Klf5), c-Myc, Sox2" 에 대해, 또한 *Nat. Cell Biol.*, 11, 197-203 (2009) 참조; 조합 "Oct3/4, Klf4, c-Myc, Sox2, hTERT, SV40 LT" 에 대해, 또한 *Nature*, 451, 141-146 (2008) 참조.]
- [0130] (9) Oct3/4, Klf4, Sox2 (참조, *Nature Biotechnology*, 26, 101-106 (2008))
- [0131] (10) Oct3/4, Sox2, Nanog, Lin28 (참조, *Science*, 318, 1917-1920 (2007))
- [0132] (11) Oct3/4, Sox2, Nanog, Lin28, hTERT, SV40LT (참조, *Stem Cells*, 26, 1998-2005 (2008))
- [0133] (12) Oct3/4, Klf4, c-Myc, Sox2, Nanog, Lin28 (참조, *Cell Research* 18 (2008) 600-603)
- [0134] (13) Oct3/4, Klf4, c-Myc, Sox2, SV40LT (참조, *Stem Cells*, 26, 1998-2005 (2008))
- [0135] (14) Oct3/4, Klf4 (참조, *Nature* 454:646-650 (2008), *Cell Stem Cell*, 2, 525-528 (2008))
- [0136] (15) Oct3/4, c-Myc (참조, *Nature* 454:646-650 (2008))
- [0137] (16) Oct3/4, Sox2 (참조, *Nature*, 451, 141-146 (2008), WO2008/118820)
- [0138] (17) Oct3/4, Sox2, Nanog (참조, WO2008/118820)
- [0139] (18) Oct3/4, Sox2, Lin28 (참조, WO2008/118820)
- [0140] (19) Oct3/4, Sox2, c-Myc, Esrrb (여기서, Esrrb 는 Esrrg 로 치환될 수 있음, *Nat. Cell Biol.*, 11, 197-203 (2009) 참조)
- [0141] (20) Oct3/4, Sox2, Esrrb (참조, *Nat. Cell Biol.*, 11, 197-203 (2009))

- [0142] (21) Oct3/4, Klf4, L-Myc (참조, *Proc. Natl. Acad. Sci. U S A.*, 107(32), 14152-14157 (2010))
- [0143] (22) Oct3/4, Klf4, Sox2, L-Myc, Lin28 (참조, WO2011/016588)
- [0144] (23) Oct3/4, Nanog
- [0145] (24) Oct3/4 (*Cell* 136: 411-419 (2009), *Nature*, 08436, doi:10.1038 온라인상으로 발표됨 (2009))
- [0146] (25) Oct3/4, Klf4, c-Myc, Sox2, Nanog, Lin28, SV40LT (참조, *Science*, 324: 797-801 (2009))
- [0147] 상기 (1)-(25) 에서, Oct3/4 는 Oct 패밀리의 또다른 일원, 예를 들어, Oct1A, Oct6 등으로 대체될 수 있다. Sox2 (또는 Sox1, Sox3, Sox15, Sox17, Sox18) 는 Sox 패밀리의 또다른 일원, 예를 들어, Sox7 등으로 대체될 수 있다. 게다가, 상기 조합 (1)-(25) 에서 핵 재프로그래밍 물질로서 c-Myc 또는 Lin28 이 포함되어 있는 경우, c-Myc 또는 Lin28 대신 각각 L-Myc 또는 Lin28B 가 사용될 수 있다.
- [0148] 상기 인자 (1)-(25) 의 조합이 Klf 패밀리의 일원을 포함하는 경우, 본 발명의 iPS 세포 확립 효율 개선 인자와 조합되어 사용되는 "핵 재프로그래밍 물질" 은 적합하게는 이들 Klf 패밀리의 일원이 아닌 인자를 함유하는 것이다. 상기 조합 (1)-(25) 가 Klf 패밀리의 일원을 포함하지 않는 경우, 본 발명의 iPS 세포 확립 효율 개선 인자와 조합되어 사용되는 핵 재프로그래밍 물질은 인자들의 조합일 수 있다.
- [0149] 전술된 핵 재프로그래밍 물질에 더하여, 또다른 임의로 선택되는 물질을 추가로 포함하는 조합도 본 발명에서 "핵 재프로그래밍 물질" 로서 적합하게 사용된다. 핵 재프로그래밍을 겪는 체세포가 핵 재프로그래밍을 야기하기에 충분한 수준으로 상기 (1) 내지 (25) 중 임의의 하나의 구성요소 하나 이상을 내생적으로 발현하는 경우, 하나 이상의 구성요소를 배제한 나머지 구성요소만의 조합도 본 발명에서 "핵 재프로그래밍 물질" 의 범위에 포함될 수 있다.
- [0150] 이들 조합 중에서, 예를 들어, Oct 패밀리의 일원, Sox 패밀리의 일원, Myc 패밀리의 일원, Lin28 패밀리의 일원 및 Nanog 중에서 선택되는 하나 이상의 물질이 바람직한 핵 재프로그래밍 물질이며, Oct3/4 및 Sox2 의 조합, Oct3/4, Sox2 및 c-Myc 의 조합, Oct3/4, Sox2 및 L-Myc 의 조합, 또는 Oct3/4, Sox2, L-Myc 및 Lin28 의 조합이 특히 바람직하다.
- [0151] c-Myc 는 iPS 세포의 확립을 촉진하지만, 또한 비-iPS 형질전환된 세포 (부분적으로 재프로그래밍된 세포, 무능성 (nullipotent) 형질전환된 세포) 의 생성도 촉진한다. 본 발명자들은 GLIS1 과 Oct3/4, Sox2 및 Klf4 의 동시 발현이 마우스 및 인간 성체 피부 섬유모세포로부터 iPS 세포의 확립을 극적으로 촉진함을 입증했을 뿐만 아니라, c-Myc 와는 달리, GLIS1 이 전술된 비-iPS 형질전환된 세포의 발생을 촉진하지 않음도 밝혔다. 그러므로, c-Myc 를 사용하지 않고, GLIS1 을 사용하는 것이 특히 바람직하다.
- [0152] 전술된 각각의 단백질성 인자의 마우스 및 인간 cDNA 서열에 대한 정보는 WO 2007/069666 (공보에서, Nanog 는 ECAT4 로 기술됨. Lin28, Lin28b, Esrrb, Esrrg 및 L-Myc 에 대한 마우스 및 인간 cDNA 서열 정보는 각각 하기 NCBI 접근 번호를 참조로 하여 획득될 수 있음) 에 언급된 NCBI 접근 번호를 참조하여 입수가 가능하며; 당업자는 이들 cDNA 를 용이하게 단리할 수 있다.
- | 유전자의 명칭 | 마우스          | 인간           |
|---------|--------------|--------------|
| Lin28   | NM_145833    | NM_024674    |
| Lin28b  | NM_001031772 | NM_001004317 |
| Esrrb   | NM_011934    | NM_004452    |
| Esrrg   | NM_011935    | NM_001438    |
| L-Myc   | NM_008506    | NM_001033081 |
- [0153]
- [0154] 핵 재프로그래밍 물질로서 사용하기 위한 단백질성 인자는 수득된 cDNA 를 적당한 발현 벡터에 삽입하고, 벡터를 숙주 세포 내로 도입하고, 배양된 세포 또는 그의 조정 배지로부터 재조합 단백질성 인자를 회수함으로써 제조될 수 있다. 한편, 단백질성 인자를 인코딩하는 핵산이 핵 재프로그래밍 물질로서 사용되는 경우, 수득되는 cDNA 는 상술된 본 발명의 핵산-기재 iPS 세포 확립 효율 개선 인자의 경우와 동일한 방식으로 바이러스 벡터, 에피솜 벡터 또는 플라스미드 벡터 내로 삽입되어 발현 벡터를 구축하고, 이 발현 벡터가 핵 재프로그래밍 단계에 적용된다. 필요한 경우 전술된 Cre-loxP 시스템 또는 piggyBac 트랜스포손 시스템이 또한 이용될 수 있다. 2 가지 이상의 단백질성 인자를 인코딩하는 2 가지 이상의 핵산이 핵 재프로그래밍 물질로서 세포에 전달되는 경우, 상이한 핵산이 별도의 벡터에 의해 운반되거나, 복수의 핵산이 나란히 연결되어 폴리시스트론성

백터를 수득할 수 있다. 후자의 경우, 효율적인 폴리시스트론성 발현을 허용하기 위해, 구제역 바이러스의 2A 자가-절단 펩티드가 핵산 사이에 삽입되는 것이 바람직하다 (참조, 예를 들어, *Science*, 322, 949-953, 2008).

[0155] 핵 재프로그래밍 물질과 체세포와의 접촉은 (a) 물질이 단백질성 인자인 경우 상술된 본 발명의 단백질성 iPS 세포 확립 효율 개선 인자와 동일한 방식으로, 또는 (b) 물질이 단백질성 인자를 인코딩하는 핵산인 경우 상술된 본 발명의 핵산-기재 iPS 세포 확립 효율 개선 인자와 동일한 방식으로 달성될 수 있다. (c) 핵 재프로그래밍 물질이 저분자량 화합물인 경우, 체세포와의 접촉은 저분자량 화합물을 적당한 농도로 수성 또는 비수성 용매에 용해시키고, 그 용액을 인간 또는 그 밖의 포유류로부터 단리된 체세포의 배양에 적합한 배지 [예, 약 5 내지 20%의 소 태아 혈청을 함유하는 최소 기본 배지 (MEM), 돌베코 변형 이글 배지 (DMEM), RPMI1640 배지, 199 배지, F12 배지, 등] 에 첨가함으로써 핵 재프로그래밍 물질의 농도를 체세포에서 핵 재프로그래밍을 야기하기에 충분하고 세포독성을 야기하지 않는 범위 내로 낮추고, 세포를 소정의 기간 동안 배양함으로써 달성될 수 있다. 핵 재프로그래밍 물질의 농도는 사용되는 핵 재프로그래밍 물질의 종류에 따라 다르고, 약 0.1 nM 내지 약 100 nM 의 범위에서 적당히 선택된다. 접촉의 지속시간은 세포의 핵 재프로그래밍을 야기하기에 충분하다면 특별히 제한되지 않으며; 통상적으로는, 양성 콜로니가 출현할 때까지 핵 재프로그래밍 물질이 배지에 동시에 존재하는 것이 허용될 수 있다.

[0156] (d) 다른 iPS 세포 확립 효율 개선제

[0157] 최근에, 전통적으로 낮았던 iPS 세포 확립 효율을 개선하는 다양한 물질들이 잇따라 제안되었다. 다른 iPS 세포 확립 효율 개선제는, 전술된 본 발명의 iPS 세포 확립 효율 개선 인자와 함께 체세포와 접촉되는 경우, iPS 세포 확립 효율을 추가로 증가시킬 것으로 예상된다.

[0158] 다른 iPS 세포 확립 효율 개선제의 예는, 이에 제한되지는 않으나, 히스톤 데아세틸라제 (HDAC) 저해제 (VPA 는 제외함) [예, 저분자량 저해제 예컨대 트리코스타틴 A (TSA), 나트륨 부티레이트, MC 1293, 및 M344, 핵산-기재 발현 저해제 예컨대 HDAC 에 대항하는 siRNA 및 shRNA (예, HDAC1 siRNA Smartpool<sup>®</sup> (Millipore), HDAC1 에 대항하는 HuSH 29mer shRNA 구축물 (OriGene) 등) 등], DNA 메틸트랜스퍼라제 저해제 [예, 5'-아자시티딘 (5'-azaC) [*Nat. Biotechnol.*, 26(7): 795-797 (2008)], G9a 히스톤 메틸트랜스퍼라제 저해제 [예, 저분자량 저해제 예컨대 BIX-01294 (*Cell Stem Cell*, 2: 525-528 (2008)], 핵산-기재 발현 저해제 예컨대 G9a 에 대항하는 siRNA 및 shRNA [예, G9a siRNA (인간) (Santa Cruz Biotechnology) 등], L-채널 칼슘 아고니스트 (예, Bayk8644) [*Cell Stem Cell*, 3, 568-574 (2008)], p53 저해제 [예, p53 에 대항하는 siRNA, shRNA, 우성 음성 돌연변이체 등 (*Cell Stem Cell*, 3, 475-479 (2008)); *Nature* 460, 1132-1135 (2009)], Wnt 신호전달 활성화제 (예, 가용성 Wnt3a) [*Cell Stem Cell*, 3, 132-135 (2008)], 2i/LIF [2i 는 글리코젠 신타제 키나제-3 및 미토겐-활성화 단백질 키나제 신호전달의 저해제임, *PloS Biology*, 6(10), 2237-2247 (2008)], ES 세포-특이적 miRNA [예, miR-302-367 클러스터 (*Mol. Cell. Biol.* doi:10.1128/MCB.00398-08); miR-302 (*RNA* (2008) 14: 1-10); miR-291-3p, miR-294 및 miR-295 (*Nat. Biotechnol.* 27: 459-461 (2009))] 등을 포함한다. 앞서 언급된 바와 같이, 핵산-기재 발현 저해제는 siRNA 또는 shRNA 를 인코딩하는 DNA 를 함유하는 발현 벡터의 형태일 수 있다.

[0159] 전술된 핵 재프로그래밍 물질의 구성요소 중에서, 예를 들어, SV40 라지 T 는 체세포의 핵 재프로그래밍에 비본질적인 보조 인자이기 때문에 iPS 세포 확립 효율 개선제의 범주에 포함될 수도 있다. 핵 재프로그래밍의 메커니즘은 여전히 규명되지 않았지만, 핵 재프로그래밍에 본질적인 인자가 아닌 보조 인자가 핵 재프로그래밍 물질 또는 iPS 세포 확립 효율 개선제로 간주되는지 여부는 중요하지 않다. 이처럼, 체세포 핵 재프로그래밍 과정은 핵 재프로그래밍 물질 및 iPS 세포 확립 효율 개선제와 체세포와의 접촉으로 초래되는 총체적 사건으로 여겨지기 때문에, 당업자가 그 둘 사이를 구별하는 것이 항상 필수적인 것 같지는 않다.

[0160] 개선제가 각각 (a) 단백질성 인자, (b) 단백질성 인자를 인코딩하는 핵산, 또는 (c) 저분자량 화합물인 경우, iPS 세포 확립 효율 개선제와 체세포와의 접촉은 상술된 본 발명의 iPS 세포 확립 효율 개선 인자 및 핵 재프로그래밍 물질과 동일한 방식으로 달성될 수 있다.

[0161] 체세포로부터의 iPS 세포 확립의 효율이 핵 재프로그래밍 물질의 부재시 수득되는 효율에 비해 유의하게 개선되는 한, iPS 세포 확립 효율 개선제, 예컨대 본 발명의 iPS 세포 확립 효율 개선 인자는 핵 재프로그래밍 물질과 동시에 체세포와 접촉될 수 있고, 어느 하나가 먼저 접촉될 수 있다. 하나의 구현예에서, 예를 들어, 핵 재프로그래밍 물질이 단백질성 인자를 인코딩하는 핵산이고 iPS 세포 확립 효율 개선제가 화학적 저해제인 경우, iPS 세포 확립 효율 개선제는 유전자 전달 처리 후 소정의 시간 동안 세포가 배양된 후에 배지에 첨가될 수 있

는데, 그 이유는 핵 재프로그래밍 물질은 유전자 전달 처리로부터 단백질성 인자의 대량-발현까지 소정의 시간 지체를 수반하지만, iPS 세포 확립 효율 개선제는 세포에 신속히 작용할 수 있기 때문이다. 또다른 구현예에서, 핵 재프로그래밍 물질 및 iPS 세포 확립 효율 개선제가 둘다 바이러스 또는 플라스미드 벡터의 형태로 사용되는 경우, 예를 들어, 둘다 세포 내로 동시에 도입될 수 있다.

[0162] (e) 배양 조건에 의한 확립 효율 개선

[0163] 체세포에 대한 핵 재프로그래밍 과정에서 빈산소 조건 하에 세포를 배양함으로써 iPS 세포 확립 효율이 추가로 개선될 수 있다 (참조, *Cell Stem Cell*, 5, p237-241(2009)). 본원에서 언급되는 용어 "빈산소 조건" 은 세포 배양시 주위 산소 농도가 대기 중의 산소 농도보다 유의하게 더 낮음을 의미한다. 구체적으로는, 보통의 세포 배양에 흔히 이용되는 5-10% CO<sub>2</sub>/95-90% 공기 대기 중의 주위 산소 농도보다 더 낮은 산소 농도를 수반하는 조건이 언급될 수 있으며; 예는 18% 이하의 주위 산소 농도를 수반하는 조건을 포함한다. 바람직하게는, 주위 산소 농도는 15% 이하 (예를 들어, 14% 이하, 13% 이하, 12% 이하, 11% 이하 등), 10% 이하 (예를 들어, 9% 이하, 8% 이하, 7% 이하, 6% 이하 등), 또는 5% 이하 (예를 들어, 4% 이하, 3% 이하, 2% 이하 등) 이다. 주위 산소 농도는 바람직하게는 0.1% 이상 (예를 들어, 0.2% 이상, 0.3% 이상, 0.4% 이상 등), 0.5% 이상 (예를 들어, 0.6% 이상, 0.7% 이상, 0.8% 이상, 0.95% 이상 등), 또는 1% 이상 (예를 들어, 1.1% 이상, 1.2% 이상, 1.3% 이상, 1.4% 이상 등) 이다.

[0164] 세포 환경에서 빈산소 상태를 만드는 임의의 방법이 이용될 수 있지만, 산소 농도를 조정할 수 있는 CO<sub>2</sub> 인큐베이터에서 세포를 배양하는 것이 가장 쉬운 방법이며, 적합한 경우에 해당한다. 산소 농도를 조정할 수 있는 CO<sub>2</sub> 인큐베이터는 다양한 제조사에서 시판 중이다 (예를 들어, Thermo scientific, Ikemoto Scientific Technology, Juji Field, Wakenyaku 등에 의해 제조된 빈산소 배양용 CO<sub>2</sub> 인큐베이터).

[0165] iPS 세포 확립 효율이 정상 산소 농도 (20%) 에 비해 개선되는 것을 막지 않는 한, 빈산소 조건 하의 세포 배양을 시작하는 시간은 특별히 제한되지 않는다. 체세포가 본 발명의 iPS 세포 확립 효율 개선 인자 및 핵 재프로그래밍 물질과 접촉되기 전에, 또는 접촉과 동시에, 또는 접촉 후에 배양이 시작될 수 있지만, 예를 들어, 체세포가 본 발명의 iPS 세포 확립 효율 개선 인자 및 핵 재프로그래밍 물질과 접촉된 직후, 또는 접촉 후 소정의 시간 간격 [예를 들어, 1 내지 10 (예를 들어, 2, 3, 4, 5, 6, 7, 8 또는 9) 일] 후에 빈산소 조건 하의 배양이 시작되는 것이 바람직하다.

[0166] iPS 세포 확립 효율이 정상 산소 농도 (20%) 에 비해 개선되는 것을 막지 않는 한, 빈산소 조건 하의 세포 배양 기간은 특별히 제한되지 않으며; 예는, 이에 제한되는 것은 아니나, 3 일 이상, 5 일 이상, 7 일 이상 또는 10 일 이상, 및 50 일 이하, 40 일 이하, 35 일 이하 또는 30 일 이하의 기간 등을 포함한다. 빈산소 조건 하의 바람직한 배양 기간은 주위 산소 농도에 따라 다르며; 당업자는 사용되는 산소 농도에 따라 배양 기간을 적당히 조정할 수 있다. 본 발명의 하나의 구현예에서, iPS 세포 후보 콜로니가 약물 내성을 지표로 이용하여 선별되는 경우, 약물 선별을 시작하기 전에 빈산소 조건에서 정상 산소 농도로 회복되는 것이 바람직하다.

[0167] 게다가, 빈산소 조건 하의 세포 배양에 바람직한 시작 시간 및 바람직한 배양 기간은 또한 사용되는 핵 재프로그래밍 물질의 선택, 정상 산소 농도에서의 iPS 세포 확립 효율 등에 따라 가변적이다.

[0168] (f) iPS 세포의 선별 및 동정

[0169] 세포는 본 발명의 iPS 세포 확립 효율 개선 인자 및 핵 재프로그래밍 물질 (및 또다른 iPS 세포 확립 효율 개선제) 와 접촉된 후, 예를 들어, ES 세포 배양에 적합한 조건 하에 배양될 수 있다. 마우스 세포의 경우, 통상의 배지에 백혈병 저해 인자 (LIF) 를 분화 억제 인자로서 첨가하여 배양이 수행된다. 한편, 인간 세포의 경우, LIF 대신 염기성 섬유모세포 성장 인자 (bFGF) 및/또는 줄기 세포 인자 (SCF) 가 첨가되는 것이 바람직하다. 통상적으로, 방사선 또는 항생제 처리되어 세포 분열이 종결된, 피더 세포로서의 마우스 배아 섬유모세포 (MEF) 의 동시 존재 하에 세포가 배양된다. 통상적으로, MEF 로서 STO 세포 등이 흔히 사용되나, iPS 세포를 유도하는 경우에는, SNL 세포 [McMahon, A.P. & Bradley, A. *Cell* 62, 1073-1085 (1990)] 등이 흔히 사용된다. 이러한 피더 세포와의 동시배양은 본 발명의 iPS 세포 확립 효율 개선 인자 및 핵 재프로그래밍 물질과의 접촉 전, 접촉시, 또는 접촉 후 (예를 들어, 1-10 일 후) 에 시작될 수 있다.

[0170] iPS 세포의 후보 콜로니는 약물 내성 및 리포터 활성을 지표로서 이용하는 방법에 의해 선별될 수 있고, 또한 형태의 육안 검사에 기초하는 방법에 의해서도 선별될 수 있다. 전자의 예로서, 다능성 세포에서 특이적으로 고도로 발현되는 유전자 (예를 들어, Fbx15, Nanog, Oct3/4 등, 바람직하게는 Nanog 또는 Oct3/4) 의 자리

가 약물 내성 유전자 및/또는 리포터 유전자의 표적이 되는 재조합 세포를 사용하여 약물 내성 및/또는 리포터 활성화에 대해 양성인 콜로니가 선별된다. 상기 재조합 세포의 예로서,  $\beta$ geo ( $\beta$ -갈락토시다제 및 네오마이신 포스포트랜스퍼라제의 융합 단백질을 인코딩함) 유전자가 Fbx15 유전자 자리에 녹인 (knocked-in) 되어 있는 마우스-유래 MEF 및 TTF (Takahashi & Yamanaka, *Cell*, 126, 663-676 (2006)), 또는 녹색 형광 단백질 (GFP) 유전자 및 퓨로마이신 내성 유전자가 Nanog 유전자 자리에 통합되어 있는 트랜스제닉 마우스-유래 MEF 및 TTF (Okita 등, *Nature*, 448, 313-317 (2007)) 등이 언급될 수 있다. 한편, 형태의 육안 검사에 의하는 후보 콜로니 선별 방법은, 예를 들어, Takahashi 등에 의해 *Cell*, 131, 861-872 (2007) 에서 기재된 방법을 포함한다. 리포터 세포를 이용하는 방법이 편리하고 효율적이지만, iPS 세포가 인간 치료용으로 제조되는 경우 안정성의 관점에서 육안 검사에 의하는 콜로니 선별이 바람직하다.

[0171] 선별된 콜로니 중 세포의 iPS 세포로서의 정체는 Nanog (또는 Oct3/4) 리포터에 대한 양성 반응 (퓨로마이신 내성, GFP 양성 등), 뿐만 아니라 상기와 같이 가지적 ES 세포성 콜로니의 형성에 의해 확인될 수 있으나; 정확성을 증가시키기 위해, 알칼리성 포스파타제 염색, 다양한 ES-세포-특이적 유전자 발현의 분석, 및 선별된 세포의 마우스로의 이식 및 테라토마 형성의 확인과 같은 시험을 수행하는 것이 가능하다.

[0172] 본 발명의 iPS 세포 확립 효율 개선 인자를 인코딩하는 핵산이 체세포에 전달되는 경우, 수득되는 iPS 세포는 그 안에 외생 핵산이 함유되어 있다는 점에서 종래에 공지된 iPS 세포와 구별되는 신규한 세포이다. 특히, 외생 핵산이 레트로바이러스, 렌티바이러스 등을 사용하여 체세포에 전달되는 경우, 외생 핵산이 통상적으로는 수득되는 iPS 세포의 계능에 통합되어, 외생 핵산을 함유하는 특징이 안정적으로 유지된다.

[0173] (g) iPS 세포의 용도

[0174] 이렇게 확립된 iPS 세포는 다양한 목적으로 사용될 수 있다. 예를 들어, ES 세포와 같은 다능성 줄기 세포에 관하여 보고된 분화 유도 방법 (예, 분화 유도 방법은 JP-A-2002-291469 에 기재된 신경 줄기 세포에 대한 방법, JP-A-2004-121165 에 기재된 채장 줄기-유사 세포에 대한 방법, 및 JP-T-2003-505006 에 기재된 조혈 세포에 대한 방법을 포함하고; 배양체 형성에 의한 분화 유도 방법은 JP-T-2003-523766 에 기재된 방법을 포함함) 을 이용하여, iPS 세포로부터 다양한 세포 (예, 심근 세포, 혈액 세포, 신경 세포, 혈관 내피 세포, 인슐린-분비 세포 등) 로 분화가 유도될 수 있다. 그러므로, 환자 또는 HLA 유형이 동일 또는 실질적으로 동일한 또 다른 사람으로부터 수집한 체세포를 이용하여 iPS 세포를 유도하면, iPS 세포가 원하는 세포 (즉, 환자의 병든 기관의 세포, 질환에 치료 효과가 있는 세포 등) 로 분화되어 환자에게 이식되는 자가 이식에 의하는 줄기 세포 치료가 가능할 것이다. 게다가, iPS 세포로부터 분화된 기능성 세포 (예를 들어, 간세포) 가 상응하는 현존 세포주보다 생체 내에서의 기능성 세포의 실제 상태를 더 잘 반영하는 것으로 여겨지므로, 기능성 세포는 제약 후보 화합물의 유효성 및 독성에 대한 시험관내 스크리닝 등에도 적합하게 사용될 수 있다.

[0175] 본 발명은 이후 하기 실시예에 의해 더욱 상세히 설명되지만, 본 발명이 이에 제한되는 것은 아니다.

[0176] 실시예

[0177] 참조예 1: 신규한 재프로그래밍 인자에 대한 스크리닝

[0178] 도 1 에 제시된 방법에 의해, Goshima 등에 의해 제작된 인간 Gateway® 엔트리 클론 (N. Goshima 등, *Nature methods*, 2008 에 기재된 라이브러리를 사용했음; 데이터는 Y. Maruyama 등, *Nucleic Acid Res.*, 2009 에 개시됨) 에 기초하여 포괄적 인간 유전자의 약 20000 클론을 정리했다. 구체적으로는, 인간 Gateway® 엔트리 클론 중에서, 전장 ORF 를 함유하는 약 50000 클론을, 커버리지 80% 이상 및 아미노산 동일성 95% 이상의 기준을 이용하여, NCBI RefSeq 에 등록된 37900 서열 (24200 유전자) 에 대한 BLASTP 서치에 적용했다. 그의 3' 말단에 정지 코돈이 있는 N-타입, 및 정지 코돈이 없는 F-타입 각각에서 서열 중복이 없는 약 20000 엔트리 클론으로 이루어지는 서브라이브러리를 구축했다. 이들 약 20000 정리된 엔트리 클론을 생물정보학 기술에 의해 단백질 키나제, 단백질 포스파타제, 전사 인자, GPCR, 및 그 밖의 클론으로 분류했으며; 전사 인자의 엔트리 클론으로 이루어지는 서브라이브러리 (모든 인간 전사 인자의 50% 이상을 포함함) 를 구축했다 (도 1). 도 2 에 제시된 바와 같이, pMXs-GW 데스티네이션 (destination) 벡터와의 LR 반응에 의해, 상기 전사 인자의 서브라이브러리로부터의 각각의 엔트리 클론에 대해 발현 클론 DNA 를 제조했다. 이러한 반응 리큐어 (liquor) 를 대장균 (*Escherichia coli*) DH5a 에 전달한 후, 이를 클로닝하여 전사 인자 발현 라이브러리 (재프로그래밍 인자 스크리닝을 위한 전사 인자 발현 라이브러리) 를 구축했다. 또한, 인간 Oct3/4, Sox2, Klf4, c-Myc 유전자 각각을 동일한 pMXs-GW 에 통합시켜 각각의 발현 클론을 구축했다. 이러한 DNA 로부터

재조합 레트로바이러스를 생성하고, 이를 하기 실험에 사용했다.

- [0179] Nanog-GFP 마우스로부터의 진피 섬유모세포를 사용하여 iPS 세포의 유도 실험을 수행했다 [Okita 등, *Nature*, 448, 313-317 (2007)]. 실험은 하기 두 시스템을 사용하여 수행했다: 피더 세포로서 사용되는 MSTO (미토마이신 C 로 처리하여 세포 분열을 종결시킨 SNL 세포) 상에서 레트로바이러스 감염을 수행하는 시스템 [이후 MSTO 방법, *Cell*, 126, 663-676 (2006)] 및 피더 세포를 사용하지 않고 감염시킨 후, 세포 재파종하고 이후 MSTO 상에서 배양을 수행하는 시스템 [이후 재파종 방법, *Nature Biotech.*, 26, pp.101-106 (2008)].
- [0180] 1번째 스크리닝을 위해, 24-웰 플레이트를 사용하여 iPS 세포를 유도했다. Nanog-GFP 마우스 피부 섬유모세포를 젤라틴 (재파종 방법) 또는 MSTO (MSTO 방법) 상으로 파종했다. 다음날, 다양한 플라즈미드로부터 제조한 레트로바이러스로 섬유모세포를 감염시켰다 (제 0 일). 구체적으로는, 섬유모세포를 3 가지 유전자 Oct3/4, Sox2 및 c-Myc, 및 상술된 전사 인자 라이브러리로부터 선택된 1 가지 유전자로 1:1:1:1 비율로 감염시켰다. 음성 대조군의 경우, 섬유모세포를 3 가지 유전자 Oct3/4, Sox2 및 c-Myc 로 1:1:1 비율로 감염시켰다. 양성 대조군의 경우, 섬유모세포를 4 가지 유전자 Oct3/4, Sox2, Klf4 및 c-Myc 로 1:1:1:1 비율로 감염시켰다.
- [0181] 섬유모세포를 감염후 제 2 일까지 10% FBS/DMEM 으로 배양하고, 제 3 일 이후 ES 배지 [*Cell*, 126, 663-676 (2006)] 로 배양했다. 섬유모세포를 초기에 젤라틴 상으로 파종한 경우 (재파종 방법), 그들을 제 3 일에 MSTO 상으로 재파종했다. 그 후, 매 2 일마다 배지를 새로 공급한 동일한 배지로 교체하는 한편, 제 21 일에 푸로마이신 선별을 시작했고, 제 28 일에 세포를 검사했다. 그 결과, 3 가지 유전자와 함께 전달된 각각의 유전자 [샘플 F09 (유전자 코드 명칭: IRX6), 샘플 G06 (유전자 코드 명칭: GLIS1), 샘플 H08 (유전자 코드 명칭: DMRTB1), 및 샘플 H10 (유전자 코드 명칭: PITX2)] 를 포함하는 웰에서 GFP-양성 콜로니가 출현하여, 마우스 iPS 세포의 확립을 확인시켜 주었다. 6-웰 플레이트를 사용하여 iPS 유도를 다시 시도했을 때, 마찬가지로 GFP-양성 콜로니가 출현했다; 재현성이 획득되었다. 콜로니 형성 및 1번째 세대 및 2번째 세대의 시기에 찍은 GFP-양성 iPS 세포 콜로니의 사진 이미지 및 위상차 이미지가 도 3 및 4 에 제시되어 있다.
- [0182] 이들 결과는 이들 4 가지 인자의, Klf4 를 대체할 수 있는 신규한 재프로그래밍 인자로서의 정체성을 입증한다. 성체 마우스 피부 섬유모세포 대신 MEF (마우스 배아 섬유모세포) 또는 HDF (인간 진피 섬유모세포) 를 사용하여 동일한 실험을 수행했을 때, 마찬가지로 iPS 세포 (GFP-양성 콜로니) 가 확립되었다.
- [0183] 참조예 2: 확립된 마우스 iPS 세포의 분석
- [0184] Gentra Puregene Cell Kit (QIAGEN) 를 사용하여 게놈을 추출했고, PCR 효소 (Takara Ex Taq) 및 참조예 1 에서 확립된 iPS 세포를 사용하여 게놈-PCR 을 수행했다. 결과가 도 5 및 6 에 제시되어 있다. 확립된 모든 iPS 세포에서, 게놈 상의 전이유전자만의 존재 및 게놈 상의 다른 유전자의 부재가 확인되었다. G6-1 클론 (유전자 코드 명칭: GLIS1) 의 경우, 전달에 사용된 c-Myc 가 게놈에 삽입되지 않았다 (도 5). 레트로바이러스 벡터는 게놈에 삽입되지 않는 한 안정적으로 발현되지 않기 때문에, 이러한 클론 G6-1 은 오직 3 가지 인자 Oct3/4, Sox2 및 GLIS1 만의 발현으로 확립된 것으로 여겨졌다.
- [0185] 다음으로, Rever Tra Ace kit (Takara) 를 사용하여 RT-PCR 분석을 수행했다. 결과가 도 7 및 8 에 제시되어 있다. 참조예 1 에서 확립된 모든 iPS 세포가 ES 세포-특이적 마커 유전자 Nanog, Oct3/4, Sox2, Rex1 및 ECAT1 을 발현했다. 이들 결과는 신규한 재프로그래밍 인자를 사용하여 확립된 세포의 iPS 세포로서의 정체성을 확인시켜 줬다.
- [0186] 실시예 1: 조합으로 사용된 G6 및 Klf4 에 의한 마우스 iPS 세포의 확립
- [0187] (a) 마우스 iPS 세포의 확립 효율에 대한, 조합으로 사용된 G6 및 Klf4 의 효과
- [0188] 참조예 1 에서 동정된, Klf4 를 대체할 수 있는 신규한 재프로그래밍 인자인, G6 (유전자 코드 명칭: GLIS1), H8 (유전자 코드 명칭: DMRTB1) 및 H10 (유전자 코드 명칭: PITX2) 을, Klf4 와 조합하여 사용하였을 때, iPS 세포가 확립될 수 있었는지 여부를 입증하는 연구를 수행했다. 참조예 1 에서와 같이 Nanog-GFP 마우스 피부 섬유모세포를 사용하여 재파종 방법에 의해 실험을 수행했다. 유전자 전달에 사용된 유전자의 조합은 하기와 같다.
- [0189] (1) Oct3/4, Sox2
- [0190] (2) Oct3/4, Sox2, G6 (유전자 코드 명칭: GLIS1)

- [0191] (3) Oct3/4, Sox2, H8 (유전자 코드 명칭: DMRTB1)
- [0192] (4) Oct3/4, Sox2, H10 (유전자 코드 명칭: PITX2)
- [0193] (5) Oct3/4, Sox2, Klf4
- [0194] (6) Oct3/4, Sox2, Klf4, G6
- [0195] (7) Oct3/4, Sox2, Klf4, H8
- [0196] (8) Oct3/4, Sox2, Klf4, H10
- [0197] 전날 100 mm 배양 디쉬 (Falcon) 1 개 당  $2.5 \times 10^6$  개 세포로 파종된 Plat-E 세포에 레트로바이러스 발현 벡터 (pMXs-Oct3/4, pMXs-Sox2, pMXs-Klf4, pMXs-G6, pMXs-H8, pMXs-H10) 를 별도로 전달함으로써 재프로그래밍에 사용된 레트로바이러스를 제조했다 (Morita, S. 등, *Gene Ther.* 7, 1063-1066). 사용된 배양 브로쓰는 DMEM/10% FCS [10% 소 태아 혈청으로 보충된 DMEM (Nacalai tesque)] 였고, 세포를 37 °C, 5% CO<sub>2</sub> 에서 배양했다.
- [0198] 벡터 전달을 촉진하기 위해, 27  $\mu$ l 의 FuGene6 트랜스펙션 시약 (Roche) 을 300  $\mu$ l 의 Opti-MEM I Reduced-Serum 배지 (Invitrogen) 에 넣고, 혼합물을 실온에서 5 분 동안 놔뒀다. 그 후, 9  $\mu$ g 의 각각의 발현 벡터를 첨가하고, 혼합물을 실온에서 추가로 15 분 동안 놔둔 후, Plat-E 배양 브로쓰에 첨가했다. 제 2 일에, Plat-E 상청액을 새로 공급한 배지로 교체했다. 제 3 일에, 배양 상청액을 회수하고, 0.45  $\mu$ m 멸균 필터 (Whatman) 를 통해 여과하고, 폴리브렌 (Nacalai) 을 4  $\mu$ g/ $\mu$ l 로 첨가하여 바이러스 액체를 수득했다.
- [0199] 사용된 Nanog-GFP 마우스 피부 섬유모세포를, 마우스 등/배 피부로부터 진피를 제거하고, 이를 겔라틴-코팅된 디쉬에서 배양함으로써 수득했다.
- [0200] 사용된 배양 브로쓰는 DMEM/10% FCS 였고, 섬유모세포를 100 mm 디쉬 (Falcon) 에 디쉬 1 개 당  $8.0 \times 10^5$  개 세포로 파종하고, 37 °C, 5% CO<sub>2</sub> 에서 배양했다. 다음 날, 각각의 레트로바이러스 액체 [상기 조합 (1) 내지 (8) 중 임의의 것] 를 첨가하여 오버나이트 감염에 의해 유전자를 전달했다.
- [0201] 바이러스 감염 다음날에, 레트로바이러스 액체를 제거하고 DMEM/10% FCS 로 교체하고, 감염후 제 3 일까지 DMEM/10% FCS 를 사용하여 세포를 배양했다. 감염 후 제 3 일에, 배지를 제거하고, 10 ml 의 PBS 를 첨가하여 세포를 세척했다. PBS 를 제거한 후, 0.25% 트립신/1 mM EDTA (Invitrogen) 을 첨가하고, 반응을 37 °C 에서 약 5 분 동안 진행시켰다. 세포가 떠오른 후, ES 세포 배양 배지 [15% 소 태아 혈청, 2 mM L-글루타민 (Invitrogen), 100  $\mu$ M 비-필수 아미노산 (Invitrogen), 100  $\mu$ M 2-메르캅토에탄올 (Invitrogen), 50 U/ml 페니실린 (Invitrogen) 및 50  $\mu$ g/ml 스트렙토마이신 (Invitrogen) 으로 보충된 DMEM (Nacalai Tesque)] 를 첨가하여 세포를 현탁시키고, 미리 피더 세포를 파종해 둔 100 mm 디쉬에 파종했다. 사용된 피더 세포는 미토마이신 C 로 처리되어 세포 분열이 종결된 SNL 세포였다 [McMahon, A.P. & Bradley, A. *Cell*, 62, 1073-1085 (1990)]. 가시적 콜로니가 출현할 때까지 매 2 일마다 ES 세포 배양 배지를 새로 공급한 동일한 배지로 교체하면서 배양을 계속했다; 감염 26 내지 28 일 후, GFP-양성 콜로니를 계수했다. 3 회의 독립적 실험들의 결과가 표 3 및 도 9 에 제시되어 있다 (도 9 는 표 3 에 제시된 결과의 그래프 표현이다; 4 회의 독립적 실험들의 결과가 대조군에 대해서만 제시되어 있다).

[0202] 표 3

	DsRed	OS+Mock+Mock	OS+Mock+G6	OS+Mock+H8	OS+Mock+H10	OS+Klf+Mock	OS+Klf+G6	OS+Klf+H8	OS+Klf+H10
1 번째	0	0	0	0	0	4	997	21	48
2 번째	0	0	0	0	0	49	1680	6	2
3 번째	0	0	0	0	0	3	1590	6	2
4 번째	0	0	0	6	0	295		223	102

[0203]

[0204]

이들 조건 하에, 심지어 G6, H8 또는 H10 을 Oct3/4 및 Sox2 에 첨가한 경우에도, iPS 세포는 확립될 수 없었거나, 또는 매우 적은 iPS 세포만 확립될 수 있었다. H8 또는 H10 을 Oct3/4, Sox2 및 Klf4 에 첨가한 경우 iPS 콜로니 총수는, Oct3/4, Sox2 및 Klf4 에 첨가하지 않은 경우와 비교할 때 증가하지 않았다. 이와 대조적으로, G6 을 Oct3/4, Sox2 및 Klf4 에 첨가했을 때 iPS 콜로니 총수는, G6 을 Oct3/4 및 Sox2 에 첨가하여 수득된 콜로니 총수 및 Klf4 를 Oct3/4 및 Sox2 에 첨가하여 수득된 콜로니 총수의 합계보다 훨씬 더 높은 수준으로 극적으로 증가했다. Klf4 및 G6 의 조합된 사용은 iPS 세포 확립 효율에 대해 시너지적 효과를 갖는 것으로 나타났다.

[0205]

(b) 3 가지 재프로그래밍 인자 (OSK) 를 사용한, 마우스 iPS 세포의 확립에 대한 GLIS1 및 c-Myc 의 개선 효과의 비교

[0206]

본 발명자들은 그후 GLIS1 및 c-Myc 가 OSK 와 함께 iPSC 생성을 촉진하는 능력을 비교했다. 성체 마우스 피부 섬유모세포에서, 형성된 GFP-양성 콜로니의 수 (도 10) 로 판단할 때, GLIS1 의 효과는 c-Myc 의 효과와

비슷하다. 본 발명자들은 또한 GLIS1 및 c-Myc 를 OSK 와 함께 동시 도입했을 때 GFP-양성 콜로니의 수가 시너지적으로 증가했음을 관찰했다 (도 10).

[0207] 다음에 본 발명자들은 형질도입 후 출현한 GFP 양성 콜로니 대 총 콜로니의 비율을 분석했다. 일방 반복측정 ANOVA 검정 및 사후 Bonferroni 검정을 분석에 사용했다. P-값이 0.05 (\*) 또는 0.01 (\*\*) 미만인 경우 차이가 통계적으로 유의한 것으로 간주되었다. 결과가 도 11 에 제시되어 있다. 중요한 것은, GLIS1 이 GFP-양성 콜로니의 생성은 특이적으로 촉진했으나, GFP-음성 콜로니의 생성은 촉진하지 않은 것인데, 이는 부분적으로 재프로그래밍된 세포 또는 형질전환된 세포를 나타낸다 (도 11). 이와 대조적으로, c-Myc 는 GFP-음성 콜로니의 수를 GFP-양성 콜로니보다 현저히 증가시켰다 (도 11). c-Myc 의 이러한 바람직하지 않은 효과는 GLIS1 이 동시 발현되었을 때 상쇄되었다. 마우스 배아 섬유모세포 (MEF) 로도 유사한 결과가 수득되었다 (도 13 및 도 14). GFP-양성 콜로니가 도 12 및 도 15 에 제시되어 있다.

[0208] 또한 본 발명자들은 MEF 로부터 OSK + GLIS1 로 확립된 iPS 세포가 생식선 적격 (germline-competent) 임을 확인했다.

[0209] 실시예 2: 조합으로 사용된 G6 및 Klf4 에 의한 인간 iPS 세포의 확립

[0210] (a) 인간 iPS 세포의 확립 효율에 대한 조합으로 사용된 G6 및 Klf4 의 효과

[0211] 성인 진피 섬유모세포 (HDF) 를 사용하여 인간 세포에서도 조합으로 사용된 Klf4 및 G6 (GLIS1) 의 시너지적 효과가 관찰되는지 여부를 확인하는 연구를 수행했다. 유전자 전달에 사용된 유전자의 조합이 이하에 제시되어 있다.

[0212] (1) Oct3/4, Sox2, c-Myc

[0213] (2) Oct3/4, Sox2, c-Myc, Klf4

[0214] (3) Oct3/4, Sox2, c-Myc, G6 (유전자 코드 명칭: GLIS1)

[0215] (4) Oct3/4, Sox2, c-Myc, Klf4, G6

[0216] Takahashi, K. 등에 의해 *Ce11*, *131*: 861-872 (2007) 에 기재된 바와 같이 렌티바이러스 (pLenti6/UbC-Slc7a1) 를 사용하여 HDF 가 마우스 동종지향성 바이러스 수용체 Slc7a1 유전자를 발현하도록 강제했다. Takahashi, K. 등에 의해 *Ce11*, *131*: 861-872 (2007) 에 기재된 바와 같이 레트로바이러스를 사용하여 이들 세포 ( $2.6 \times 10^5$  세포/60 mm 디쉬) 를 상기 조합 (1) 내지 (4) 의 유전자로 트랜스펙션시켰다. 바이러스 감염 6 일 후, 세포를 회수하고 피더 세포 상에 재파종했다 ( $5 \times 10^4$  개 세포 또는  $5 \times 10^5$  개 세포/100 mm 디쉬). 사용된 피더 세포는 미토마이신 C 로 처리되어 세포 분열이 종결된 SNL 세포였다 [McMahon, A.P. & Bradley, A. *Ce11*, *62*, 1073-1085 (1990)]. 감염 7 일 후에 시작하여, 4 ng/ml 재조합 인간 bFGF (WAKO) 로 보충된 영양류 ES 세포 배양 배지 (ReproCELL) 에서 세포를 배양했다. 감염 30 내지 35 일 후, ES 세포-유사 콜로니를 계수했다. 3 회의 독립적 실험들의 결과가 도 16 (ES-유사 콜로니) 및 도 17 (비-ES-유사 콜로니) 에 제시되어 있다. Oct3/4, Sox2, c-Myc, Klf4 및 G6 으로 확립된 iPS 콜로니의 위상차 이미지가 도 18 에 제시되어 있다. Klf4 를 Oct3/4, Sox2 및 c-Myc 에 첨가하고, G6 (GLIS1) 을 Oct3/4, Sox2 및 c-Myc 에 첨가한 경우와 비교할 때, Klf4 및 G6 둘다를 Oct3/4, Sox2 및 c-Myc 에 첨가한 경우가 훨씬 더 많은 수의 ES 세포-유사 콜로니의 출현을 초래했다 (도 16). 이들 콜로니는 ES 세포-유사 형태를 나타냈다 (도 18). 요약하면, 인간 세포에서도 또한, Klf4 및 G6 을 조합으로 사용했을 때 iPS 세포 확립 효율에 대한 시너지적 효과가 관찰되었다.

[0217] (b) 3 가지 재프로그래밍 인자 (OSK) 를 사용한, 인간 iPS 세포의 확립에 대한 GLIS1 및 c-Myc 의 개선 효과의 비교

[0218] 본 발명자들은 그후 실시예 1(b) 에 기재된 바와 동일한 방식으로 OSK 와 함께 iPSC 생성을 촉진하는 GLIS1 및 c-Myc 의 능력을 비교했다. 성인 섬유모세포에서, GLIS1 은 c-Myc 와 비슷한 정도로 유사한 효과를 나타냈고, OSK 와 함께 도입되었을 때 ESC-유사 콜로니의 생성을 촉진했다 (도 19). 중요한 것은, GLIS1 이 ESC-유사 콜로니의 생성은 특이적으로 촉진했으나, 비-ESC-유사 콜로니의 생성은 촉진하지 않다는 것이다. 이와 대조적으로, c-Myc 는 비-ESC-유사 콜로니의 수를 ESC-유사 콜로니보다 현저히 증가시켰다 (도 20). OSK + GLIS1 로 생성된 인간 ESC-유사 콜로니가 도 21 에 제시되어 있다.

[0219] 그 후, QIAGEN "Gentra Puregene Cell Kit" 를 사용하여 게놈을 추출했고, PCR 효소 (Takara Ex Taq) 를 사용

하여 게놈-PCR 을 수행했다. 결과가 도 22 에 제시되어 있다. 본 발명자들은 확립된 인간 iPSC 라인에서 전이유전자의 존재를 확인했다 (도 22). Rever Tra Ace kit (Takara) 를 사용하여 RT-PCR 분석을 수행했다. 결과가 도 23 에 제시되어 있다. OSK + GLIS1 로 생성된 세포는 *Oct3/4*, *Sox2*, *Nanog*, 및 *Rex1* 을 포함하는 미분화된 ESC 마커 유전자를 발현했다 (도 23). 다음으로 본 발명자들은 DNA 마이크로어레이 분석을 수행했다. 총 RNA 를 Cy3 으로 라벨링하고, 제조사의 프로토콜에 따라 Whole Human Genome Microarray (Agilent) 에 하이브리드화시켰다. G2505C Microarray Scanner System (Agilent) 을 사용하여 어레이를 스캔했다. Genepring GX11.0.1 소프트웨어 프로그램 (Agilent) 을 사용하여 데이터를 분석했다. 결과가 도 24 에 제시되어 있다. OSK + GLIS1 로 확립된 세포는 OSKM 으로 생성된 iPSC 와 전체적인 유전자 발현의 면에서 유사했다 (도 24). 그 후 본 발명자들은 이전에 기술된 바와 같이 (*Cell*, 131(5), 861-872 (2007)) 테라토마 형성을 수행했다. OSK + GLIS1 로 생성된 세포는 모든 3 개의 배엽의 다양한 조직을 함유하는 테라토마를 생성했다 (도 25). 이들 결과는 GLIS1 이 OSK 에 의한 인간 iPSC 의 생성을 강하고 특이적으로 촉진했음을 입증했다.

[0220] 실시예 3: GLIS1 의 발현 및 기능적 분석

[0221] 본 발명자들은 그후 GLIS1 의 발현 패턴을 조사했다. 마우스 발현 서열 태그 (EST) 데이터베이스의 분석으로 GLIS1 발현이 집합체에, 특히 수정란에, 편중되었음을 예측했다 (<http://www.ncbi.nlm.nih.gov/UniGene/ESTProfileViewer.cgi?uglist=Mm.331757>; 2010 년 4 월 24 일자 현재). 또한, MGI 에 의해 제공된 유전자 발현 데이터 (Gene Expression Data) 는 중기 II 난모세포에서 중도의 GLIS1 발현 및 2-세포 배아에서 약한 발현, 및 8-세포 내지 E4.5 배아에서 발현이 탐지되지 않았음을 나타냈다 (<http://www.informatics.jax.org/searches/expression.cgi?32989>; 2010 년 4 월 24 일자 현재). 이들 웹 기반 분석은 난모세포 및 1-세포 배아에서 GLIS1 의 특이적 발현을 강하게 시사했다. 이들 발견을 실험적으로 확인하기 위해, 본 발명자들은 미수정란, 1-세포 배아, 2-세포 배아, 및 배반포로부터, 뿐만 아니라 신장, 태반, 뇌, 폐, 간, 비장, 및 난소를 포함하는 여러 성체 마우스 조직으로부터 총 RNA 를 단리했다. 또한, 본 발명자들은 마우스 ESC, MEF, 및 성체 피부 섬유모세포로부터 단리된 총 RNA 를 사용했다. 실시간 PCR 분석으로 1-세포 배아 및 미수정란에서 GLIS1 의 최고의 발현을 탐지했다 (도 26). 2-세포 배아 및 태반 조직에서 보통의 발현 수준이 탐지되었다 (도 26). 신장, 난소, ESC, MEF 및 피부 섬유모세포를 포함하는 여러 조직에서 약한 발현이 존재했다 (도 26). 이들 데이터는 미수정란 및 1-세포 배아에 난소 GLIS1 RNA 가 풍부함을 확인시켜 줬다.

[0222] 다음으로 본 발명자들은 섬유모세포 내의 내생 GLIS1 이, 비록 낮은 수준으로 발현됨에도 불구하고, OSK 에 의한 iPSC 생성 동안 역할을 수행하는지 여부를 조사했다. 이를 위하여, 본 발명자들은 GLIS1 shRNA 를 발현하는 여러 레트로바이러스 벡터를 구축했다. 이전에 기술된 바와 같이 (*Nature*, 460(7259), 1132-1135 (2009)) shRNA-매개 녹다운 (knockdown) 을 수행했다. 본 발명자들은 shRNA2 (목표 서열 (SEQ ID NO:3 의 위치 822-842): `ggcctcaccaacctgcacct`; SEQ ID NO:13) 및 shRNA6 (목표 서열 (SEQ ID NO:3 의 위치 1457-1477): `gcccttcaatgccgctacaa`; SEQ ID NO:14) 이 성체 마우스 피부 섬유모세포 내로 트랜스펙션되었을 때 GLIS1 을 효과적으로 억제한 반면에, shRNA4 (목표 서열 (SEQ ID NO:3 의 위치 857-877): `gggcaatgaaccatctcaga`; SEQ ID NO:15) 는 덜 효과적이었음을 발견했다 (도 27, 통계적 분석에 대응 (paired) *t*-검정을 사용했다). 그 후 본 발명자들은 각각의 이들 shRNA 를 OSK 와 함께 Nanog-GFP 리포터를 함유하는 섬유모세포 내로 도입했다. 본 발명자들은 shRNA2 및 shRNA6 이 GFP-양성 콜로니의 수를 상당히 감소시켰음을 발견했다 (도 28). shRNA4 의 경우 더 약한 효과가 관찰되었다. 이러한 결과는 내생 GLIS1 이 OSK 에 의한 iPSC 생성 동안 보조 역할을 수행함을 시사한다.

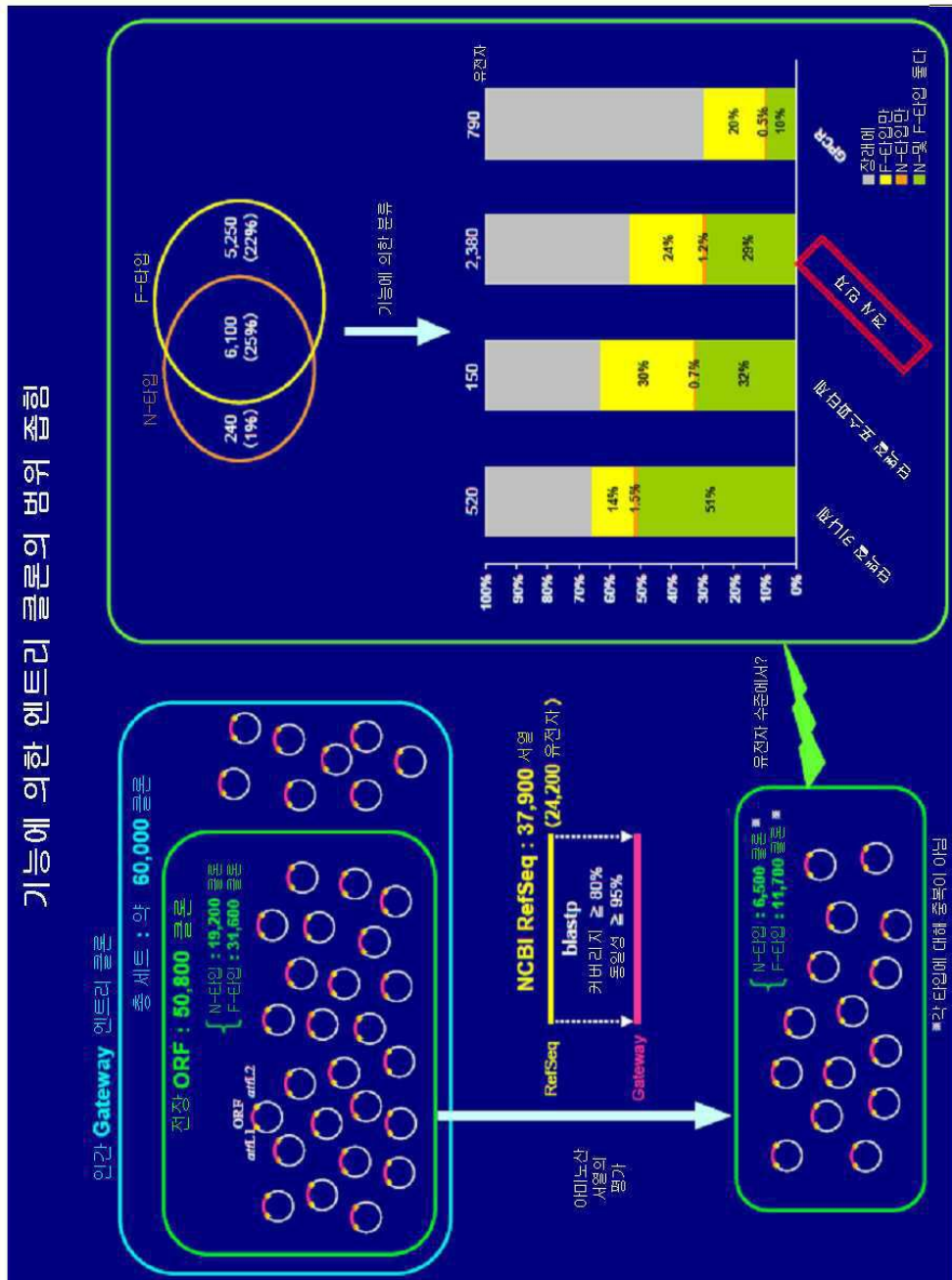
[0223] 본 발명은 바람직한 구현예를 중심으로 기술되었지만, 바람직한 구현예가 변형될 수 있음이 당업자에게 명백하다. 본 발명은, 본 명세서에서 상세히 기술된 방법 이외의 방법으로 본 발명이 구현될 수 있음을 의도한다. 따라서, 본 발명은 첨부된 "청구항" 의 요지 및 범주에 포함되는 모든 변형을 포괄한다.

[0224] 또한, 본원에서 인용된 특허 및 특허 출원을 포함하는 임의의 발행물에 개시된 내용은 본원에 개시된 정도로 그 전문이 본원에서 참고로 포함된다.

[0225] 본 출원은 본원에 그 내용이 참고로 포함되는 미국 가 특허 출원 제 61/305,107 호 및 제 61/379,949 호에 기초한다.

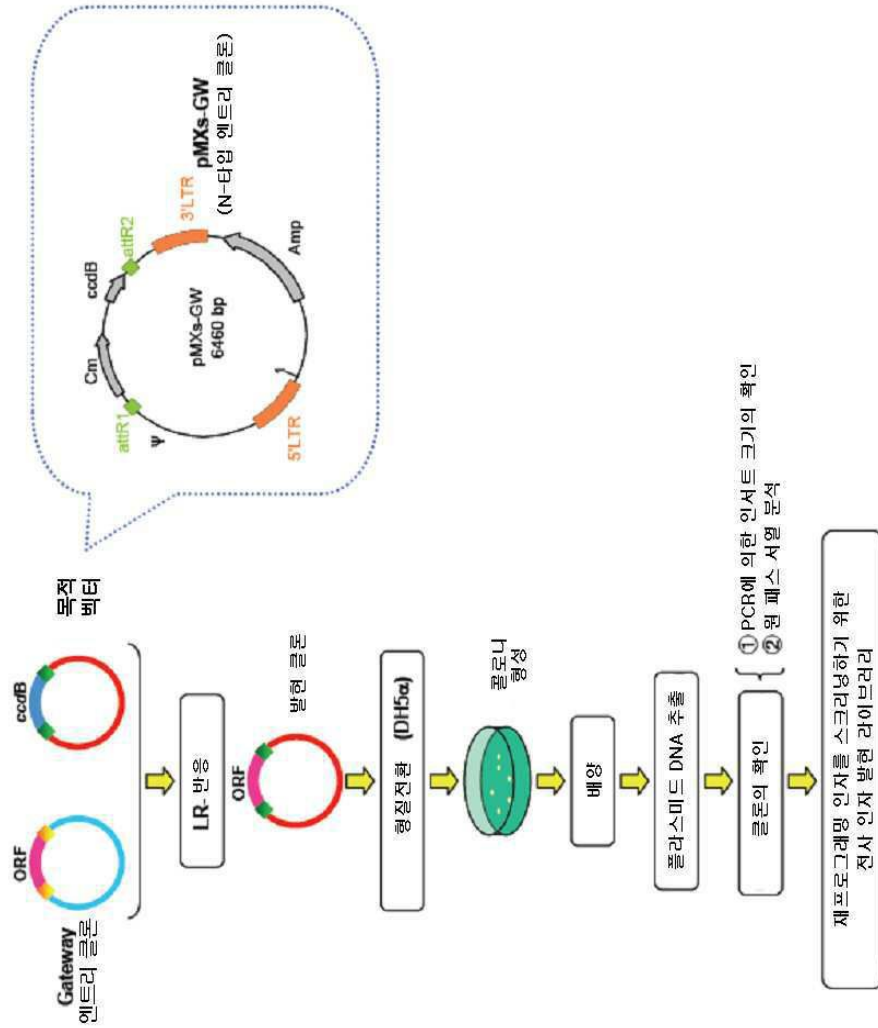
도면

도면1

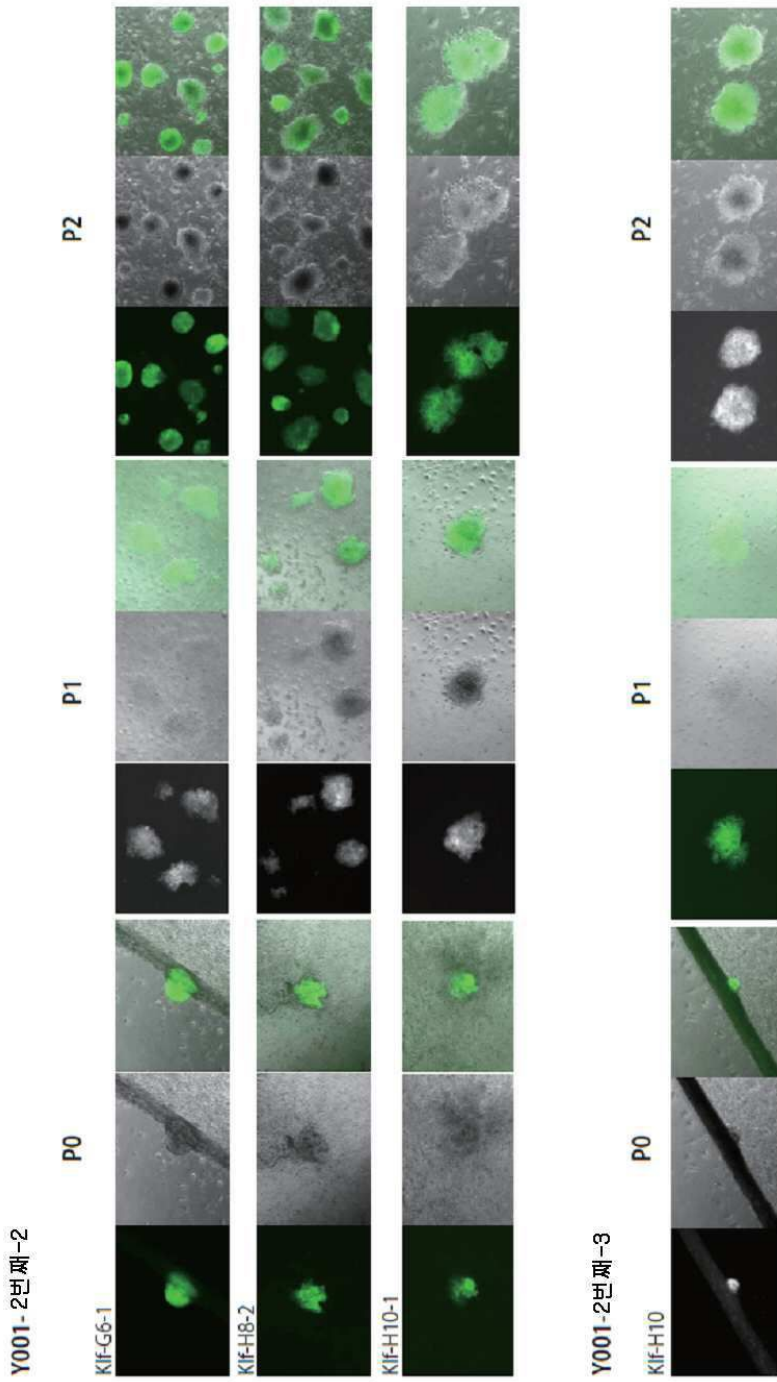


도면2

신규한 체세포 재프로그래밍 인자를 스크리닝하기 위한 발현 클론의 제조 개요



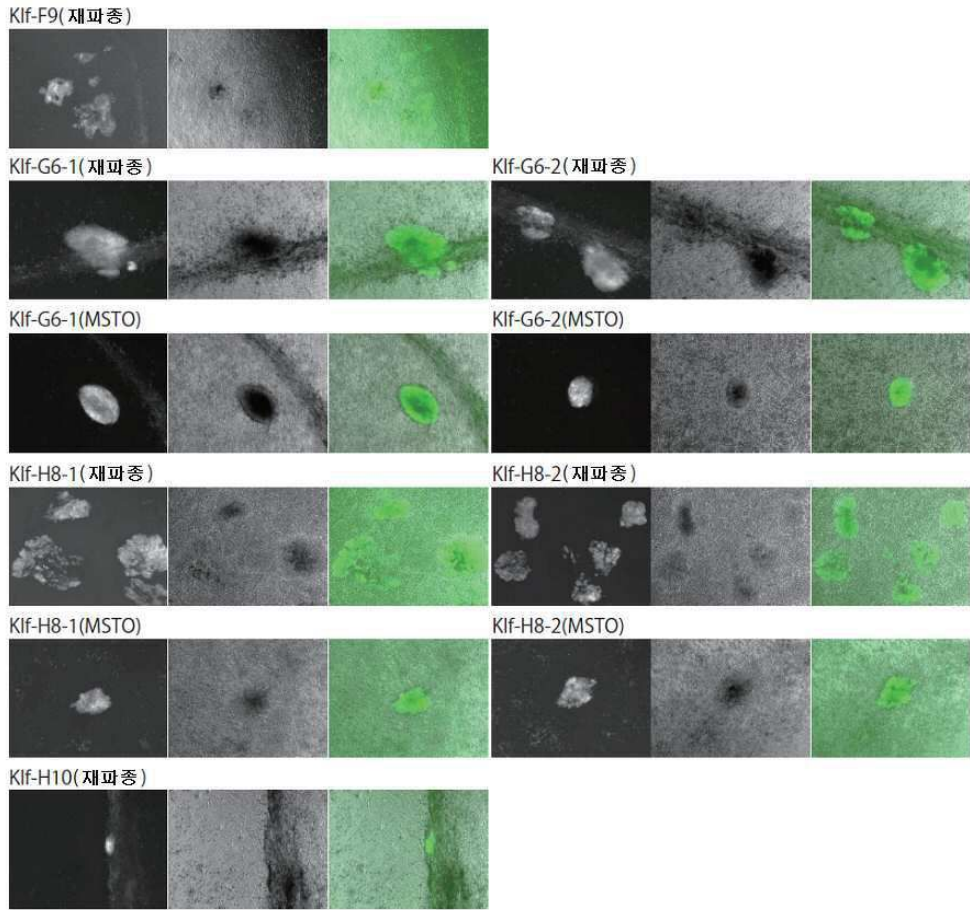
도면3



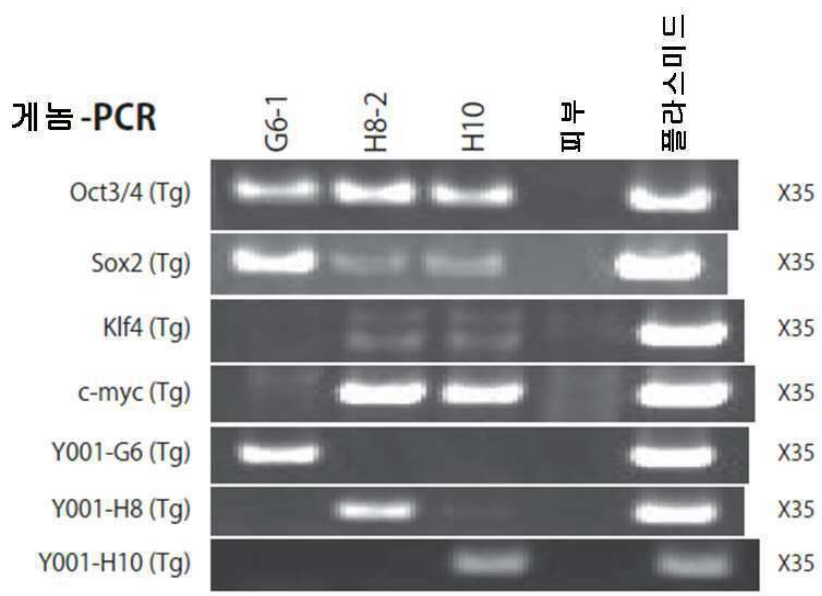
도면4

Y001-2번째-5

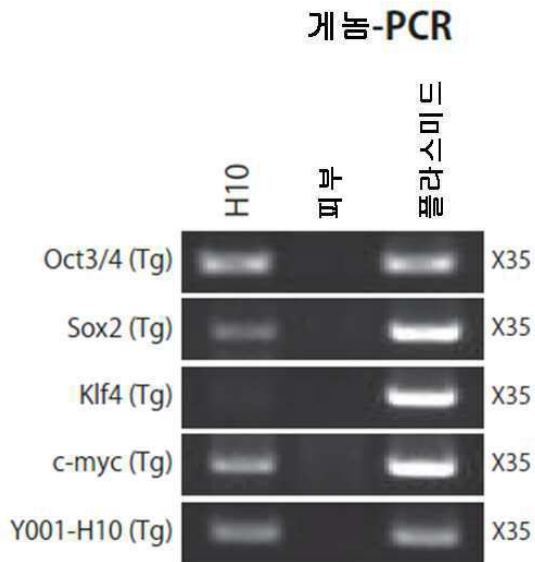
P0



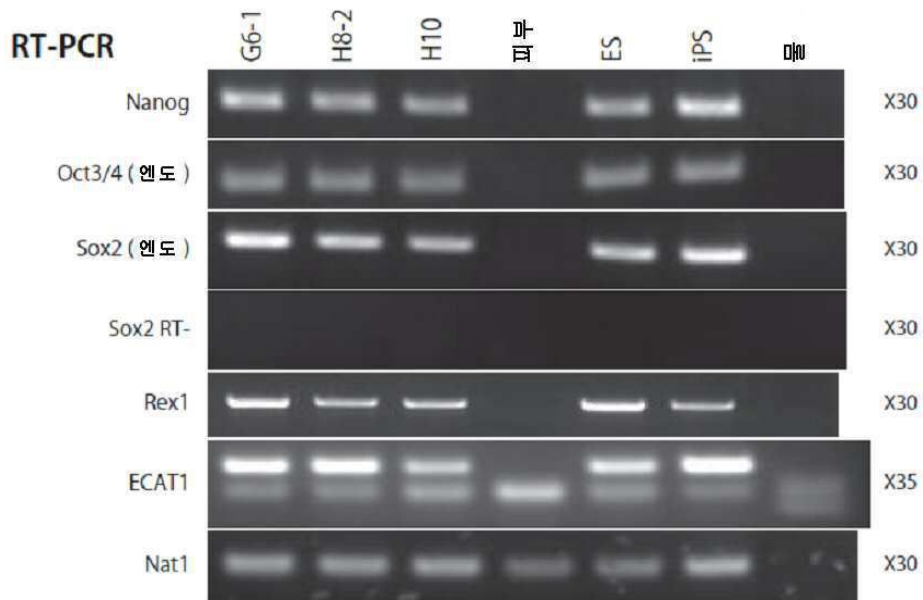
도면5



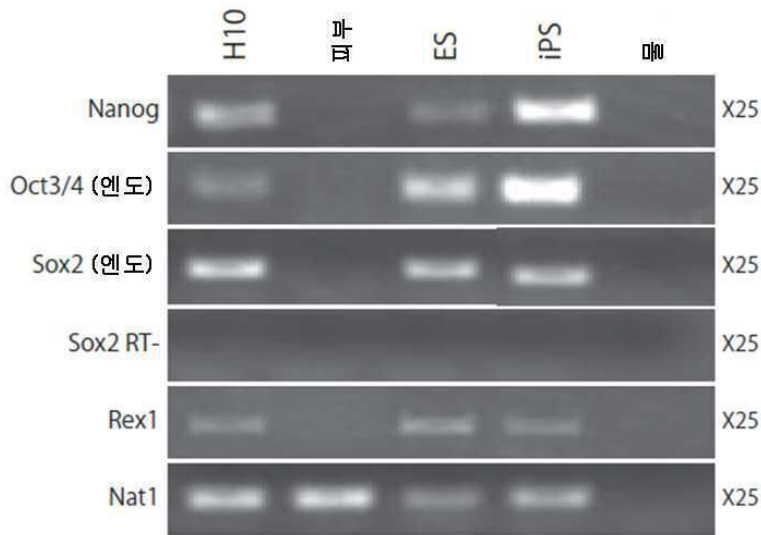
도면6



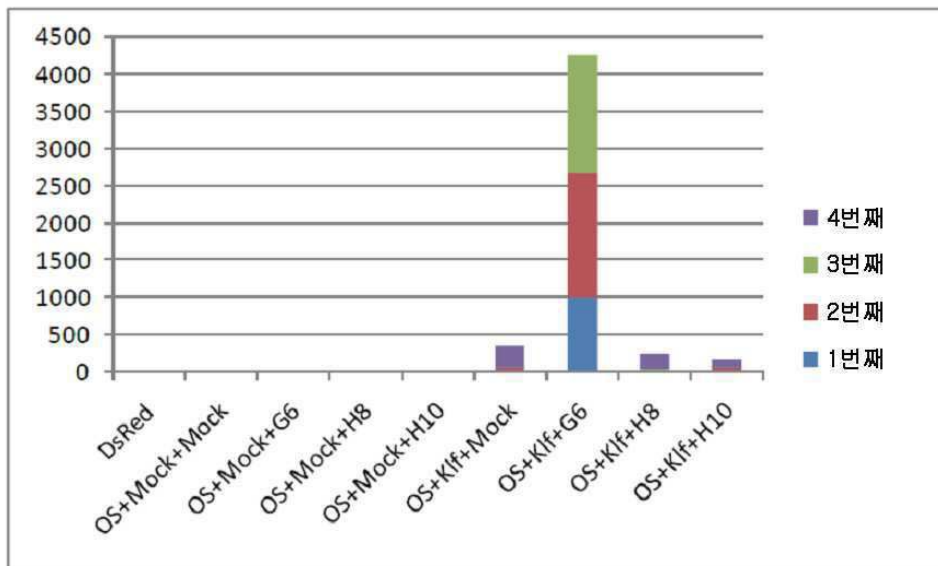
도면7



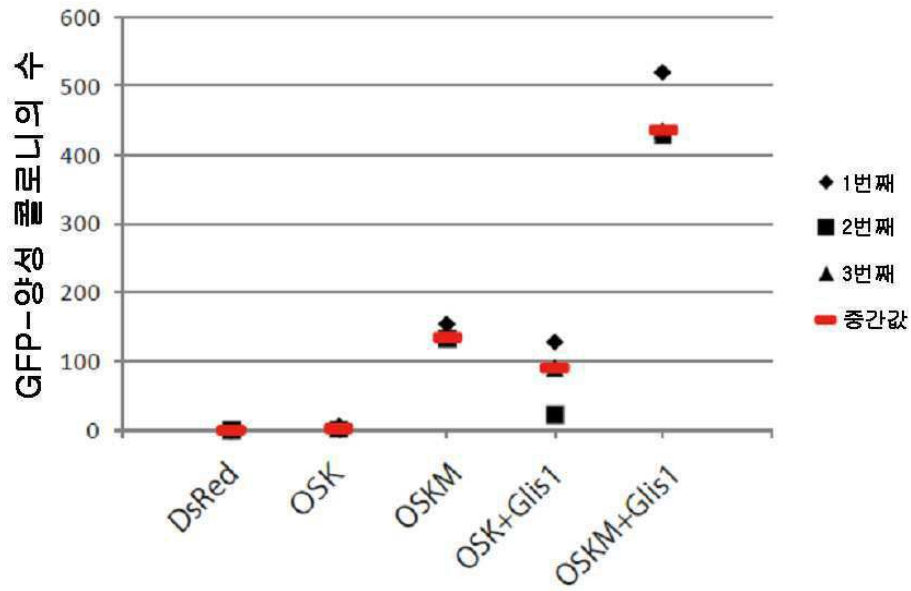
도면8



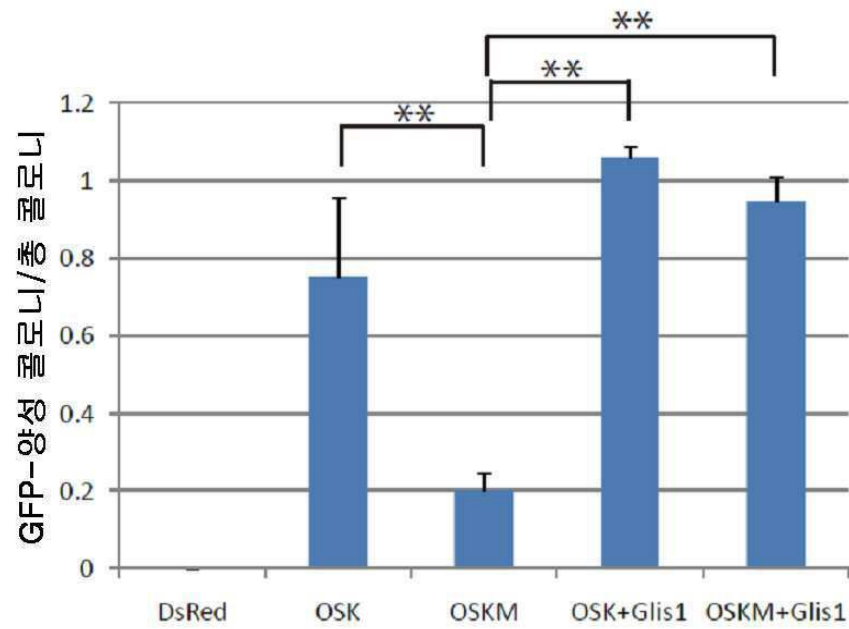
도면9



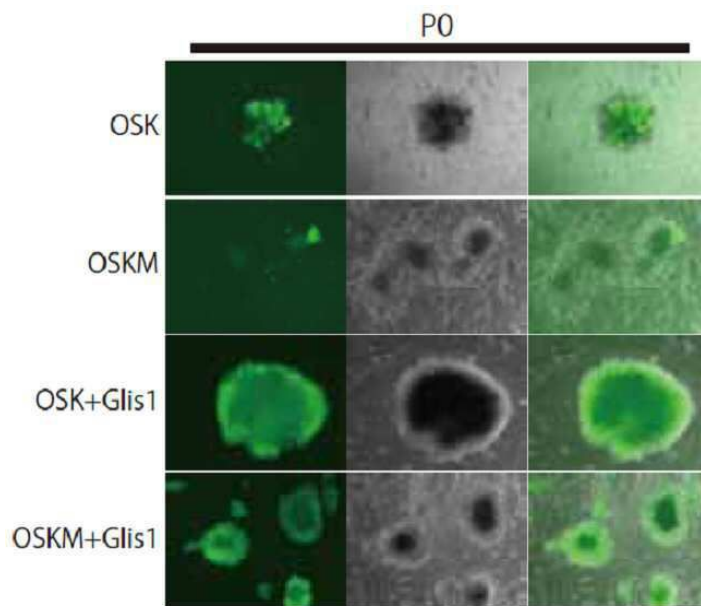
도면10



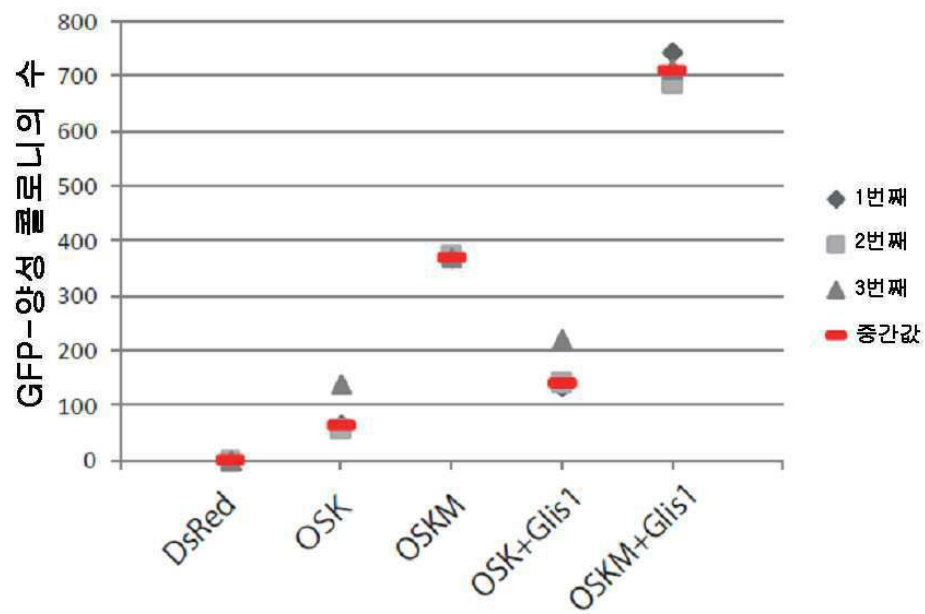
도면11



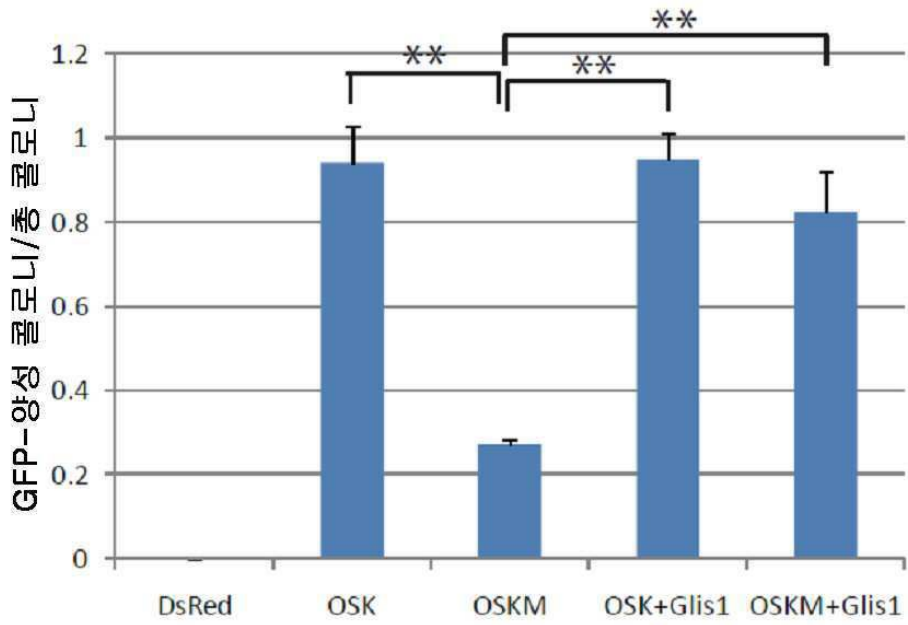
도면12



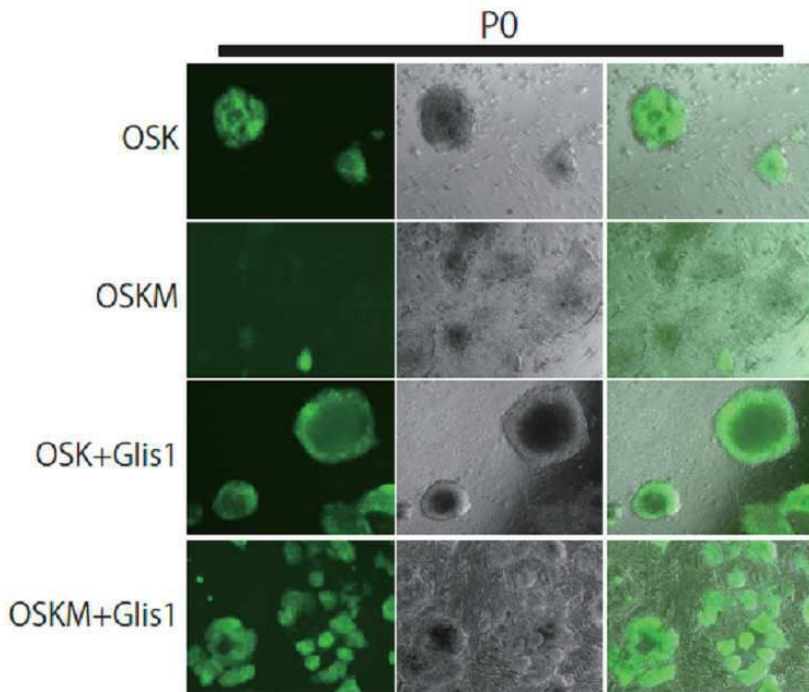
도면13



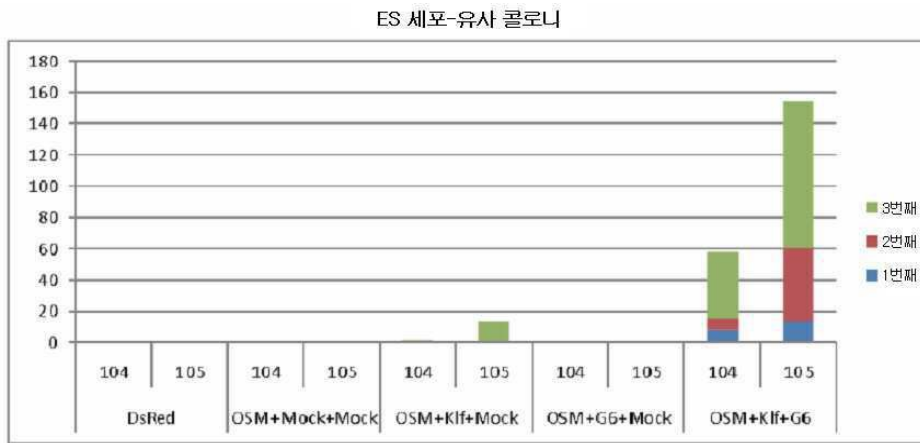
도면14



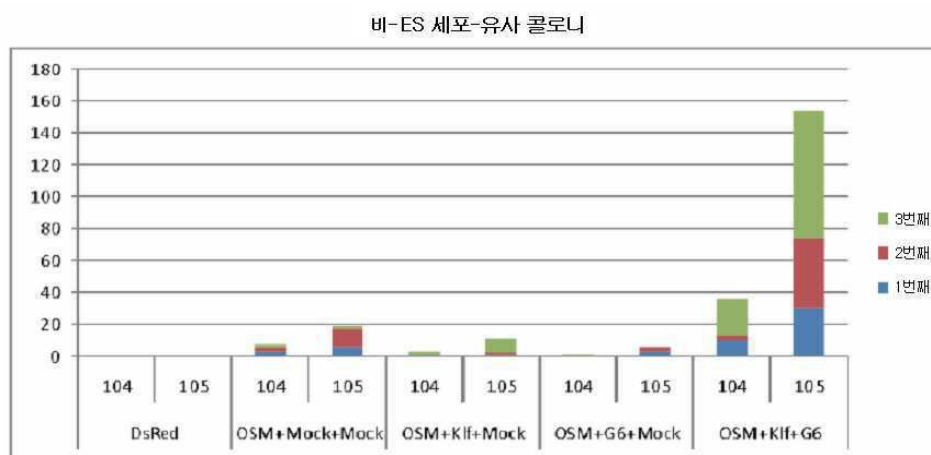
도면15



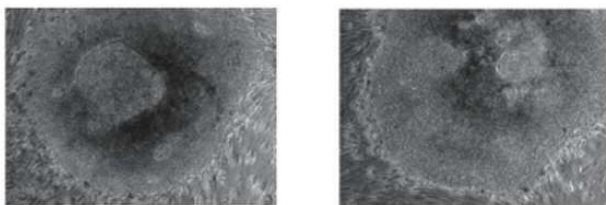
도면16



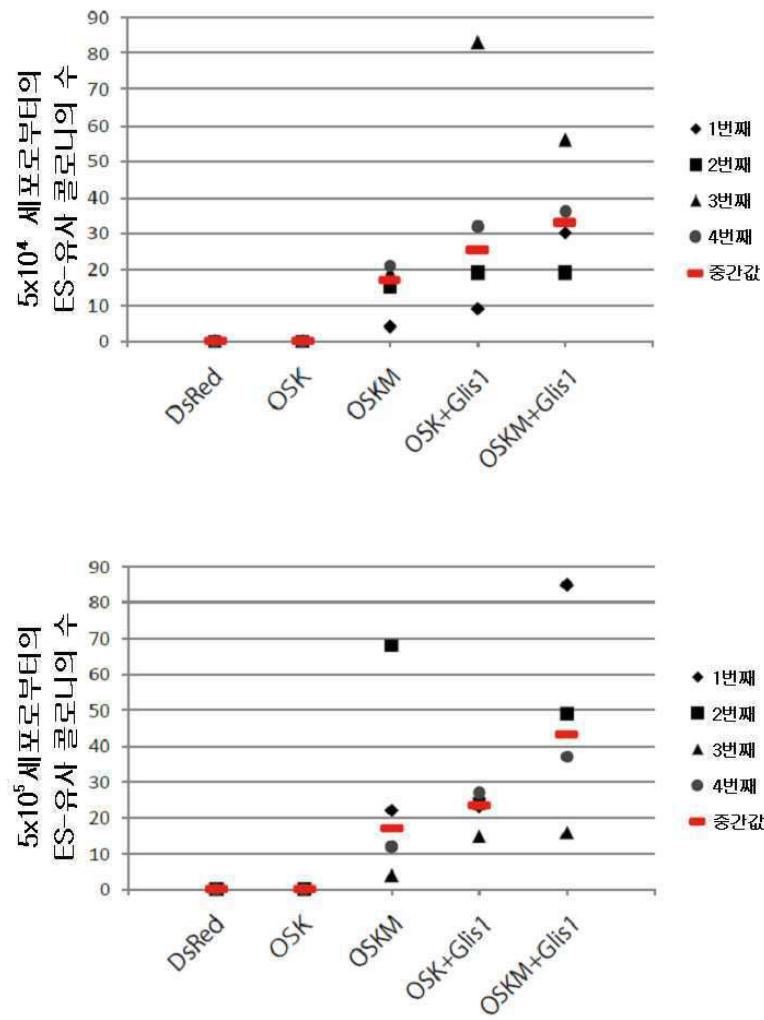
도면17



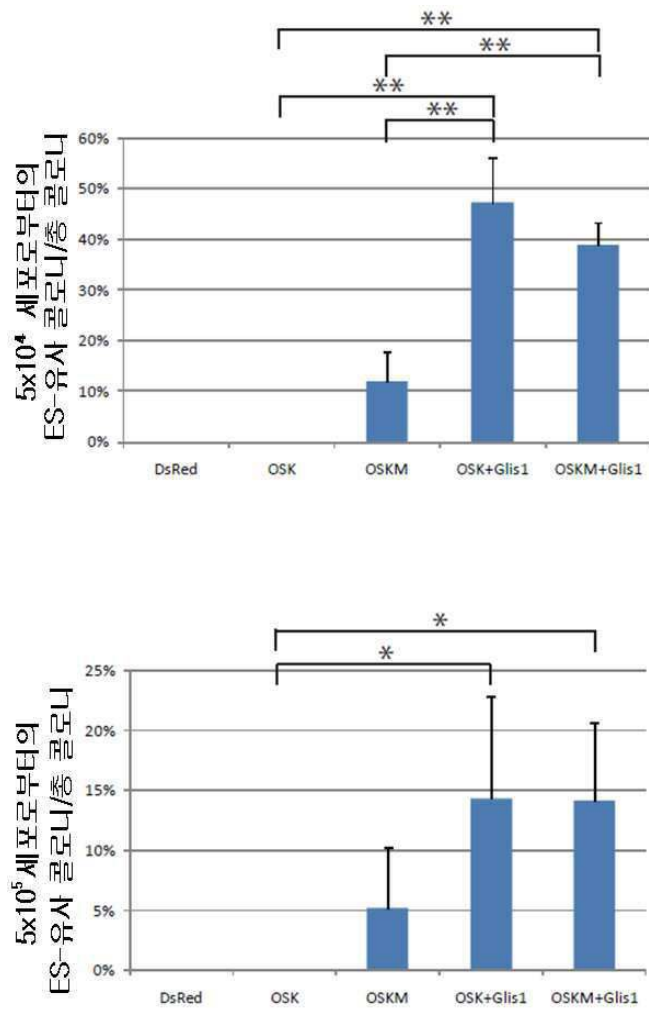
도면18



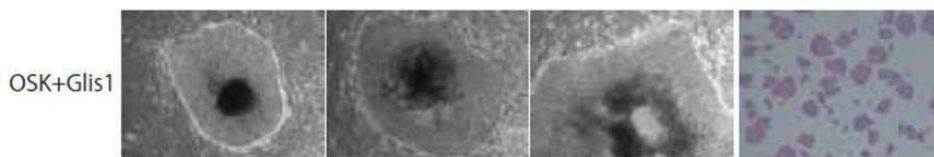
도면19



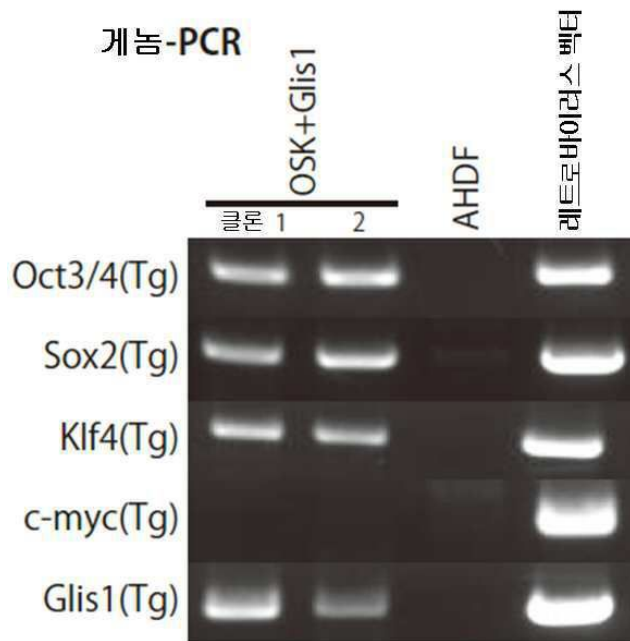
도면20



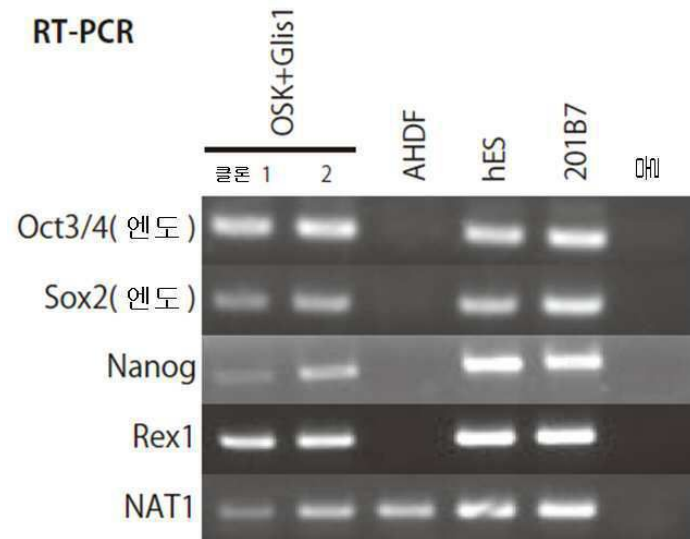
도면21



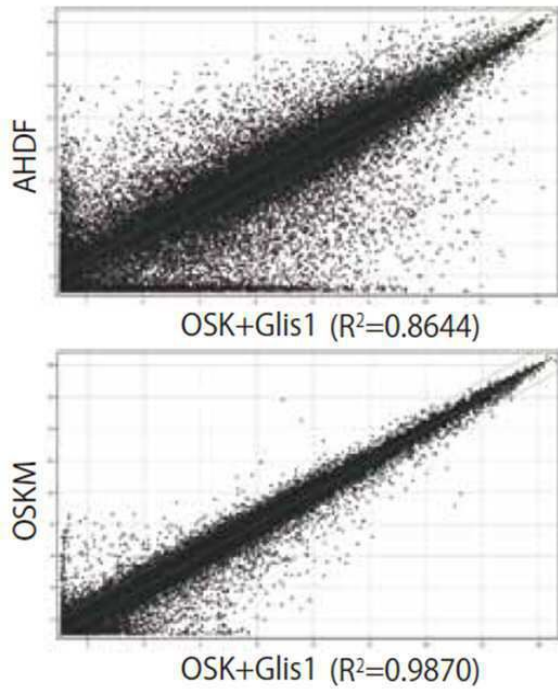
도면22



도면23



도면24

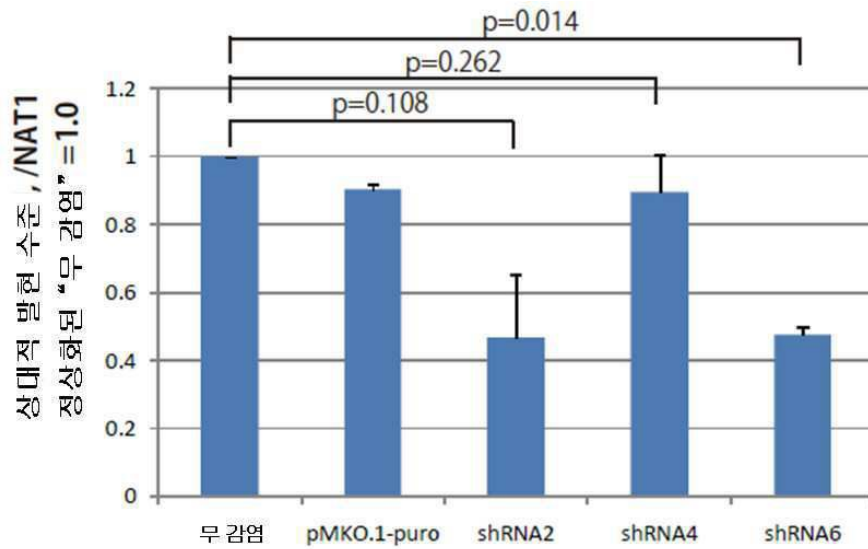


도면25

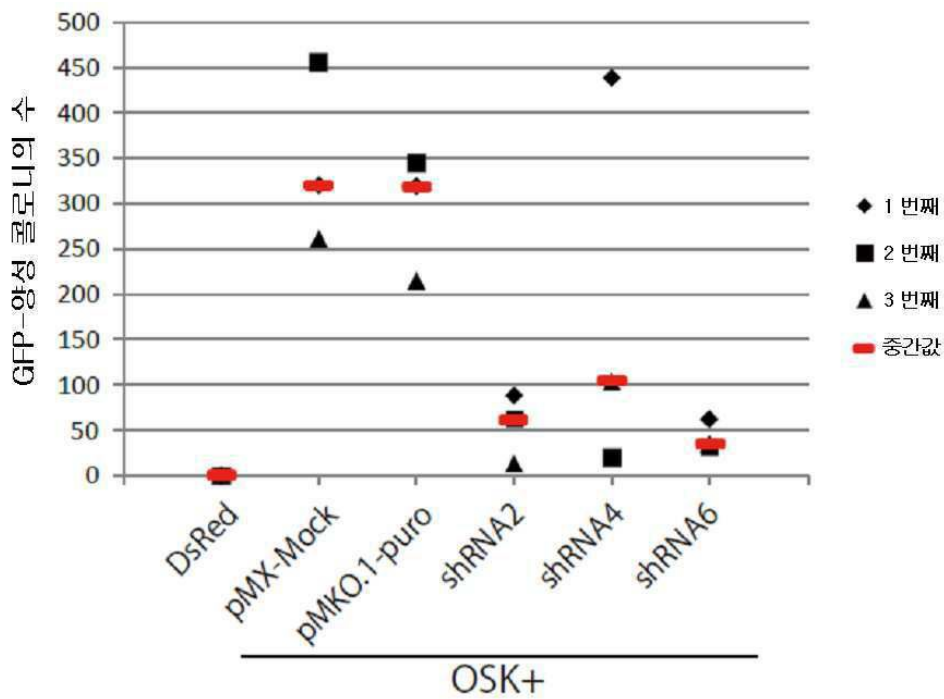




도면27



도면28



서열 목록

SEQUENCE LISTING

<110> Kyoto University

NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND TECHNOLOGY

JAPAN BIOLOGICAL INFORMATICS CONSORTIUM

<120> METHOD OF EFFICIENTLY ESTABLISHING INDUCED PLURIPOTENT STEM CELLS

<130> 091667

<150> US 61/305,107

<151> 2010-02-16

<150> US 61/379,949

<151> 2010-09-03

<160> 15

<170> PatentIn version 3.5

<210> 1

<211> 2816

<212> DNA

<213> Homo sapiens

<220><221> CDS

<222> (568)..(2430)

<400> 1

```

cactgtgtac tgagactgga tgcacacctg caataaaaa gaggttgatc acgacaaatg      60
tgaaccccgc cgttataaaa acagccatca tggctgtaaa tgccaaaaag cagtcagtct      120
tgtaacttga aaaaaaaaaa aaaggaattg tagattgtgc gcatggactc ggagtggggg      180
cgggtggacag taagtcatga tgttttggtg gtaccacctg gttgaatttc ttcacatgaa      240
taagaagctc ctgtgatggt ctggggaggc cttggaaggc tagcgcaccc ctcacatgaaa      300
gtgaatggga gctacggaca ccgtaccccg ggctcagaga agagcctgct ggacctggac      360
cttgctgagg gccttggecc cacctgctgc cagggcctgt ttctcctgc aggaagccca      420

cgccccggg ctcaccccca agcttgtgag aggctgctgc attccccca cctgacagg      480
tcacctagac ccaggccac gtatgtgaac ggcagcctcc caaccacaca acacatcaaa      540
caggagtctt tgcccgacta ccaagcc atg gca gag gcc cgc aca tcc ctg tet      594
                Met Ala Glu Ala Arg Thr Ser Leu Ser
                1                5
gcc cac tgt cgg ggc ccg ctg gcc act ggc ctg cac cca gac ctg gac      642
Ala His Cys Arg Gly Pro Leu Ala Thr Gly Leu His Pro Asp Leu Asp

10                15                20                25
ctc ccg ggc cga agc ctc gcc acc cct ggc cct tcc tgc tac ctt ctg      690
Leu Pro Gly Arg Ser Leu Ala Thr Pro Ala Pro Ser Cys Tyr Leu Leu
                30                35                40
ggc agc gaa ccc agc tct ggc ctg ggc ctc cag ccc gag acc cac ctc      738
Gly Ser Glu Pro Ser Ser Gly Leu Gly Leu Gln Pro Glu Thr His Leu

```

45	50	55	
ccc gag ggc agc ctg aag cgg tgc tgc gtc ttg ggc cta ccc ccc acc			786
Pro Glu Gly Ser Leu Lys Arg Cys Cys Val Leu Gly Leu Pro Pro Thr			
60	65	70	
tcc cca gcc tcc tcc tca ccc tgt gcc tcc tcc gac gtc acc tcc atc			834
Ser Pro Ala Ser Ser Ser Pro Cys Ala Ser Ser Asp Val Thr Ser Ile			
75	80	85	
atc cgc tcc tcc cag acg tct ctg gtc acc tgt gta aat gga ctc cgg			882
Ile Arg Ser Ser Gln Thr Ser Leu Val Thr Cys Val Asn Gly Leu Arg			
90	95	100	105
agc ccc cct ctg acg gga gat ctg ggg ggc cct tcc aag cgg gcc cgg			930
Ser Pro Pro Leu Thr Gly Asp Leu Gly Gly Pro Ser Lys Arg Ala Arg			
110	115	120	
cct gcc cct gca tcg acg gac agc cat gag ggc agc ttg caa ctt gaa			978
Pro Gly Pro Ala Ser Thr Asp Ser His Glu Gly Ser Leu Gln Leu Glu			
125	130	135	
gcc tgc cgg aag gcg agc ttc ctg aag cag gaa ccc gcg gat gag ttt			1026
Ala Cys Arg Lys Ala Ser Phe Leu Lys Gln Glu Pro Ala Asp Glu Phe			
140	145	150	
tca gag ctc ttt ggg cct cac cag cag ggc ctg ccg ccc ccc tat ccc			1074
Ser Glu Leu Phe Gly Pro His Gln Gln Gly Leu Pro Pro Pro Tyr Pro			
155	160	165	
ctg tct cag ttg ccg cct ggc cca agc ctt gga ggc ctg ggg ctg ggc			1122
Leu Ser Gln Leu Pro Pro Gly Pro Ser Leu Gly Gly Leu Gly Leu Gly			
170	175	180	185
ctg gca ggc agg gtg gtg gcc ggg cgg cag gcg tgc cgc tgg gtg gac			1170
Leu Ala Gly Arg Val Val Ala Gly Arg Gln Ala Cys Arg Trp Val Asp			
190	195	200	
tgc tgt gca gcc tat gag cag cag gag gag ctg gtg cgg cac atc gag			1218

Cys Cys Ala Ala Tyr Glu Gln Gln Glu Glu Leu Val Arg His Ile Glu  
 205 210 215  
 aag agc cac atc gac cag cgc aag ggc gag gac ttc acc tgc ttc tgg 1266  
 Lys Ser His Ile Asp Gln Arg Lys Gly Glu Asp Phe Thr Cys Phe Trp  
 220 225 230  
 gct ggc tgc gtg cgc cgc tac aag ccc ttc aac gcc cgc tac aag ctg 1314  
 Ala Gly Cys Val Arg Arg Tyr Lys Pro Phe Asn Ala Arg Tyr Lys Leu  
 235 240 245  
 ctc atc cac atg cga gtg cac tgc ggc gag aag ccc aac aag tgc atg 1362  
 Leu Ile His Met Arg Val His Ser Gly Glu Lys Pro Asn Lys Cys Met  
 250 255 260 265  
 ttt gaa ggc tgc agc aag gcc ttc tca cgg ctg gag aac ctc aag atc 1410  
 Phe Glu Gly Cys Ser Lys Ala Phe Ser Arg Leu Glu Asn Leu Lys Ile  
 270 275 280  
 cac ctg agg agc cac acg ggc gag aag ccg tac ctg tgc cag cac ccg 1458  
 His Leu Arg Ser His Thr Gly Glu Lys Pro Tyr Leu Cys Gln His Pro  
 285 290 295  
 ggt tgc cag aag gcc ttc agc aac tcc agc gac cgc gcc aag cac cag 1506  
 Gly Cys Gln Lys Ala Phe Ser Asn Ser Ser Asp Arg Ala Lys His Gln  
 300 305 310  
 cgc acc cac cta gac acg aag ccg tac gcc tgt cag atc cct ggc tgc 1554  
 Arg Thr His Leu Asp Thr Lys Pro Tyr Ala Cys Gln Ile Pro Gly Cys  
 315 320 325  
 tcc aag cgc tac aca gac ccc agc tcc ctc cgc aag cac gtc aag gcc 1602  
 Ser Lys Arg Tyr Thr Asp Pro Ser Ser Leu Arg Lys His Val Lys Ala  
 330 335 340 345  
 cat tca gcc aaa gag cag cag gtg cgt aag aag ctg cat gcg ggc cct 1650  
 His Ser Ala Lys Glu Gln Gln Val Arg Lys Lys Leu His Ala Gly Pro  
 350 355 360  
 gac acc gag gcc gac gtc ctg acc gag tgt ctg gtc ctg cag cag etc 1698  
 Asp Thr Glu Ala Asp Val Leu Thr Glu Cys Leu Val Leu Gln Gln Leu

365	370	375	
cac acg tcc aca cag ctg gct gcc agc gac ggc aag ggt ggc tgt ggc			1746
His Thr Ser Thr Gln Leu Ala Ala Ser Asp Gly Lys Gly Gly Cys Gly			
380	385	390	
ctg ggc cag gag ctg ctc cca ggt gtg tat cct ggc tcc atc acc ccc			1794
Leu Gly Gln Glu Leu Leu Pro Gly Val Tyr Pro Gly Ser Ile Thr Pro			
395	400	405	
cat aac gga ctt gca tcg ggc ctc ctg ccc cca gcg cac gac gta cct			1842
His Asn Gly Leu Ala Ser Gly Leu Leu Pro Pro Ala His Asp Val Pro			
410	415	420	425
tcc agg cac cac ccg ctg gat gcc acc acc agt tcc cac cac cat ctg			1890
Ser Arg His His Pro Leu Asp Ala Thr Thr Ser Ser His His His Leu			
430	435	440	
tcc cct ctg ccc atg gct gag agc acc cgg gat ggg ttg ggg ccc ggc			1938
Ser Pro Leu Pro Met Ala Glu Ser Thr Arg Asp Gly Leu Gly Pro Gly			
445	450	455	
ctc ctc tca cca ata gtc agc ccc ctg aag ggg ctg ggg cca ccg ccg			1986
Leu Leu Ser Pro Ile Val Ser Pro Leu Lys Gly Leu Gly Pro Pro Pro			
460	465	470	
ctg ccc cca tcc tct cag agc cat tct ccg ggg ggc cag ccc ttc ccc			2034
Leu Pro Pro Ser Ser Gln Ser His Ser Pro Gly Gly Gln Pro Phe Pro			
475	480	485	
aca ctc ccc agc aag ccg tcc tac cca ccc ttc cag agc cct cca ccc			2082
Thr Leu Pro Ser Lys Pro Ser Tyr Pro Pro Phe Gln Ser Pro Pro Pro			
490	495	500	505
ccg cct ctg ccc agc cca caa ggt tac cag ggc agt ttc cac tcc atc			2130
Pro Pro Leu Pro Ser Pro Gln Gly Tyr Gln Gly Ser Phe His Ser Ile			
510	515	520	
cag agt tgc ttc ccc tat ggc gac tgc tac cgg atg gct gaa cca gca			2178
Gln Ser Cys Phe Pro Tyr Gly Asp Cys Tyr Arg Met Ala Glu Pro Ala			
525	530	535	

gcc ggt ggg gac gga ctg gtc ggg gag acc cac ggt ttc aac ccc ctg 2226

Ala Gly Gly Asp Gly Leu Val Gly Glu Thr His Gly Phe Asn Pro Leu  
 540 545 550

cgg ccc aat ggc tac cac agc ctc agc acg ccc ttg cct gcc aca ggc 2274  
 Arg Pro Asn Gly Tyr His Ser Leu Ser Thr Pro Leu Pro Ala Thr Gly  
 555 560 565

tat gag gcc ctg gct gag gcc tca tgc ccc aca gcg ctg cca cag cag 2322  
 Tyr Glu Ala Leu Ala Glu Ala Ser Cys Pro Thr Ala Leu Pro Gln Gln

570 575 580 585

cca tct gaa gat gtg gtg tcc agc ggc ccc gag gac tgt ggc ttc ttc 2370  
 Pro Ser Glu Asp Val Val Ser Ser Gly Pro Glu Asp Cys Gly Phe Phe  
 590 595 600

ccc aat gga gcc ttt gac cac tgc ctg ggc cac atc ccc tcc atc tac 2418  
 Pro Asn Gly Ala Phe Asp His Cys Leu Gly His Ile Pro Ser Ile Tyr  
 605 610 615

aca gac acc tga aggagccccc acatgcgcct gcccatccag cactgcagat 2470  
 Thr Asp Thr  
 620

gccacctcgc ccacctgctg tcgctccccc cctccgtgca cctagcagga gtgccaggcc 2530  
 acagccggaa cagccaggcc atgaccagg ggagccagcg ctgccacccc acccagcgct 2590  
 gccaggggagc cgccatccga gcttgagctg ggcgcacaga ggtgcccc aggatctgtg 2650  
 gccctgtaac attccctcga tcttgtcttc cgttctctcc ccgcagtggg tttgaaatca 2710

cagacctcgt gtatataaaa tatgcagaac ttgttttccg ttcccctgcc agttttatat 2770  
 ttttggtttt acaagaaaaa acattaaaaa ctggaaagga gatgtg 2816

<210> 2  
 <211> 620  
 <212> PRT  
 <213> Homo sapiens  
 <400> 2

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

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

260 265 270  
 Phe Ser Arg Leu Glu Asn Leu Lys Ile His Leu Arg Ser His Thr Gly  
 275 280 285  
 Glu Lys Pro Tyr Leu Cys Gln His Pro Gly Cys Gln Lys Ala Phe Ser  
  
 290 295 300  
 Asn Ser Ser Asp Arg Ala Lys His Gln Arg Thr His Leu Asp Thr Lys  
 305 310 315 320  
 Pro Tyr Ala Cys Gln Ile Pro Gly Cys Ser Lys Arg Tyr Thr Asp Pro  
 325 330 335  
 Ser Ser Leu Arg Lys His Val Lys Ala His Ser Ala Lys Glu Gln Gln  
 340 345 350  
 Val Arg Lys Lys Leu His Ala Gly Pro Asp Thr Glu Ala Asp Val Leu  
  
 355 360 365  
 Thr Glu Cys Leu Val Leu Gln Gln Leu His Thr Ser Thr Gln Leu Ala  
 370 375 380  
 Ala Ser Asp Gly Lys Gly Gly Cys Gly Leu Gly Gln Glu Leu Leu Pro  
 385 390 395 400  
 Gly Val Tyr Pro Gly Ser Ile Thr Pro His Asn Gly Leu Ala Ser Gly  
 405 410 415  
 Leu Leu Pro Pro Ala His Asp Val Pro Ser Arg His His Pro Leu Asp  
  
 420 425 430  
 Ala Thr Thr Ser Ser His His His Leu Ser Pro Leu Pro Met Ala Glu  
 435 440 445  
 Ser Thr Arg Asp Gly Leu Gly Pro Gly Leu Leu Ser Pro Ile Val Ser  
 450 455 460  
 Pro Leu Lys Gly Leu Gly Pro Pro Pro Leu Pro Pro Ser Ser Gln Ser  
 465 470 475 480  
 His Ser Pro Gly Gly Gln Pro Phe Pro Thr Leu Pro Ser Lys Pro Ser  
  
 485 490 495  
 Tyr Pro Pro Phe Gln Ser Pro Pro Pro Pro Pro Leu Pro Ser Pro Gln  
 500 505 510

Gly Tyr Gln Gly Ser Phe His Ser Ile Gln Ser Cys Phe Pro Tyr Gly  
 515 520 525

Asp Cys Tyr Arg Met Ala Glu Pro Ala Ala Gly Gly Asp Gly Leu Val  
 530 535 540

Gly Glu Thr His Gly Phe Asn Pro Leu Arg Pro Asn Gly Tyr His Ser

545 550 555 560  
 Leu Ser Thr Pro Leu Pro Ala Thr Gly Tyr Glu Ala Leu Ala Glu Ala  
 565 570 575

Ser Cys Pro Thr Ala Leu Pro Gln Gln Pro Ser Glu Asp Val Val Ser  
 580 585 590

Ser Gly Pro Glu Asp Cys Gly Phe Phe Pro Asn Gly Ala Phe Asp His  
 595 600 605

Cys Leu Gly His Ile Pro Ser Ile Tyr Thr Asp Thr

610 615 620

<210> 3

<211> 2904

<212> DNA

<213> Mus musculus

<220><221> CDS

<222> (222)..(2591)

<400> 3

ggggaccag tggcgtccga atccgggagc tctggggtgg cgcggggctc gccgaggggc 60

gaggcgaatt tgggggcctt gaggcctcgc tctcgcggga atgatgctgg aatgatgct 120

gaggctccgg cgtgagactt cggcgtgccg gcggagcggg gtgtgagccg gtgaatgggg 180

agcctggcgc gacccccagc cgtgcgcccc gccccggcgc c atg cat tgc gag gtg 236

Met His Cys Glu Val

1 5

gcc gag gca ctt tcg gac aag agg cca aag gag gcc cct ggt gct cct 284

Ala Glu Ala Leu Ser Asp Lys Arg Pro Lys Glu Ala Pro Gly Ala Pro

10 15 20

ggc cag ggc cgc ggg cct gtc agc ctg gga gcg cac atg gcc ttc agg 332

Gly Gln Gly Arg Gly Pro Val Ser Leu Gly Ala His Met Ala Phe Arg

25	30	35	
att gct gtg agt ggt ggc ggc tgc ggg gac ggg aac ccg cta gac ctg			380
Ile Ala Val Ser Gly Gly Gly Cys Gly Asp Gly Asn Pro Leu Asp Leu			
40	45	50	
ctg cct cgg cta ccg gtg cca cca cca cgt gcc cac gat ctc ctt cgg			428
Leu Pro Arg Leu Pro Val Pro Pro Pro Arg Ala His Asp Leu Leu Arg			
55	60	65	
ccc cgg agc cct cga gac tat ggt gtg tcc aag acc ggc agc ggg aag			476
Pro Arg Ser Pro Arg Asp Tyr Gly Val Ser Lys Thr Gly Ser Gly Lys			
70	75	80	85
gtg aac ggg agc tac ggg cac agc tca gag aag agc ctg ctg gac ctg			524
Val Asn Gly Ser Tyr Gly His Ser Ser Glu Lys Ser Leu Leu Asp Leu			
90	95	100	
gac ctg gcc gag ggt ccc agc ccc tcc tgc cac cag ggt ctg ttt ctt			572
Asp Leu Ala Glu Gly Pro Ser Pro Ser Cys His Gln Gly Leu Phe Leu			
105	110	115	
cct gca ggg acc cca cca ccc cgg ggt cac ccc cct gtc tgt gag aag			620
Pro Ala Gly Thr Pro Pro Pro Arg Gly His Pro Pro Val Cys Glu Lys			
120	125	130	
ctg ctg cac ttc ccc cac cca aac agg tca ccc aga cct cag gct acg			668
Leu Leu His Phe Pro His Pro Asn Arg Ser Pro Arg Pro Gln Ala Thr			
135	140	145	
ttt gtg aac ggc agc ctc cca gcc gct cag cac atc aag caa gaa gcc			716
Phe Val Asn Gly Ser Leu Pro Ala Ala Gln His Ile Lys Gln Glu Ala			
150	155	160	165
cta ccg gac tac cag gcc atg gtc agc gcc cac aca ccc ctg ccc acc			764
Leu Pro Asp Tyr Gln Ala Met Val Ser Ala His Thr Pro Leu Pro Thr			
170	175	180	
cac tgc cga gcc cca tgc tcc atg ggt ctg ccc tca gac ctg gac ttt			812

His Cys Arg Ala Pro Ser Ser Met Gly Leu Pro Ser Asp Leu Asp Phe  
 185 190 195  
 cca gac cga ggc ctc acc aac cct gca cct tcc tgc tac ctt ctg ggc 860  
 Pro Asp Arg Gly Leu Thr Asn Pro Ala Pro Ser Cys Tyr Leu Leu Gly  
 200 205 210  
 aat gaa ccc atc tca gac ctg ggt ccc caa ccc gag gcc cac ctc ccc 908  
 Asn Glu Pro Ile Ser Asp Leu Gly Pro Gln Pro Glu Ala His Leu Pro  
 215 220 225  
 gag ggc agc ctg aaa cgc tgc tgc ctc ctg ggc ctg ccc ccc acc tct 956  
 Glu Gly Ser Leu Lys Arg Cys Cys Leu Leu Gly Leu Pro Pro Thr Ser  
 230 235 240 245  
 tca gcc tcc tcc tca ccc tgt gcc tcc tca gat atc aat cct gtc atc 1004  
 Ser Ala Ser Ser Ser Pro Cys Ala Ser Ser Asp Ile Asn Pro Val Ile  
 250 255 260  
 cac tcc tcc cag aca gct cta gtt agc tgt gta aat gga ctc cga agc 1052  
 His Ser Ser Gln Thr Ala Leu Val Ser Cys Val Asn Gly Leu Arg Ser  
 265 270 275  
 cca cct ctg ccg gga gac ctg ggg ggc cct ccc aag cgg tca cgg ccc 1100  
 Pro Pro Leu Pro Gly Asp Leu Gly Gly Pro Pro Lys Arg Ser Arg Pro  
 280 285 290  
 ggg cct gca tcc agt gac ggc cag gag ggc agc ttg cag ctt gaa gca 1148  
 Gly Pro Ala Ser Ser Asp Gly Gln Glu Gly Ser Leu Gln Leu Glu Ala  
 295 300 305  
 tgc cgg aag tca ggc ttc ctg aag cag gag ccc atg gac gag ttt tca 1196  
 Cys Arg Lys Ser Gly Phe Leu Lys Gln Glu Pro Met Asp Glu Phe Ser  
 310 315 320 325  
 gag ctt ttt gct cca cac cac cag ggt ttg cca ccc cct tac ccc ttg 1244  
 Glu Leu Phe Ala Pro His His Gln Gly Leu Pro Pro Pro Tyr Pro Leu  
 330 335 340  
 cct cag ttg cca act ggc ccc ggc ctc gga ggc cta ggg ctg ggc ctg 1292  
 Pro Gln Leu Pro Thr Gly Pro Gly Leu Gly Gly Leu Gly Leu Gly Leu

345	350	355	
gca ggt agg atg gtt gcc ggt cgg cag gca tgc cgc tgg gtg gac tgc			1340
Ala Gly Arg Met Val Ala Gly Arg Gln Ala Cys Arg Trp Val Asp Cys			
360	365	370	
tgc gca gcc tac gag cag cag gag gag ctg gtg cgg cac atc gag aag			1388
Cys Ala Ala Tyr Glu Gln Gln Glu Glu Leu Val Arg His Ile Glu Lys			
375	380	385	
agc cac atc gac cag cgc aag ggc gaa gac ttc acc tgc ttc tgg gcc			1436
Ser His Ile Asp Gln Arg Lys Gly Glu Asp Phe Thr Cys Phe Trp Ala			
390	395	400	405
ggg tgt gtg cgg cgc tac aag ccc ttc aat gcc cgc tac aag ctg ctc			1484
Gly Cys Val Arg Arg Tyr Lys Pro Phe Asn Ala Arg Tyr Lys Leu Leu			
410	415	420	
atc cac atg agg gta cac tca ggc gag aag ccc aac aag tgc atg ttc			1532
Ile His Met Arg Val His Ser Gly Glu Lys Pro Asn Lys Cys Met Phe			
425	430	435	
gaa ggc tgc agt aaa gcc ttt tcc cgt ctg gag aac ctg aag atc cat			1580
Glu Gly Cys Ser Lys Ala Phe Ser Arg Leu Glu Asn Leu Lys Ile His			
440	445	450	
ctg cgg agc cac aca ggc gag aaa cca tac ctg tgc cag cac cca ggc			1628
Leu Arg Ser His Thr Gly Glu Lys Pro Tyr Leu Cys Gln His Pro Gly			
455	460	465	
tgc cag aag gcc ttc agc aac tcc agc gac cgt gcc aag cac caa cgc			1676
Cys Gln Lys Ala Phe Ser Asn Ser Ser Asp Arg Ala Lys His Gln Arg			
470	475	480	485
acc cac ctc gac acg aag cca tat gct tgt cag atc cct ggc tgc tcc			1724
Thr His Leu Asp Thr Lys Pro Tyr Ala Cys Gln Ile Pro Gly Cys Ser			
490	495	500	
aag cgc tac acg gac ccc agc tcc ctc cgc aag cac gtg aag gcc cac			1772
Lys Arg Tyr Thr Asp Pro Ser Ser Leu Arg Lys His Val Lys Ala His			
505	510	515	

tca gcc aaa gag cag cag gtg cgt aag aag ctg cac aca ggt gcc gac 1820

Ser Ala Lys Glu Gln Gln Val Arg Lys Lys Leu His Thr Gly Ala Asp  
 520 525 530

cca gag gct gat gtt ctg tcc gag tgt ctg tcc ctg cag cag ctc caa 1868

Pro Glu Ala Asp Val Leu Ser Glu Cys Leu Ser Leu Gln Gln Leu Gln  
 535 540 545

gca tcc aca ctg ttg ccg gcc agc aga ggg aag ggc agc caa acc ctg 1916

Ala Ser Thr Leu Leu Pro Ala Ser Arg Gly Lys Gly Ser Gln Thr Leu

550 555 560 565

agc cag gag ctc ctc cca ggt gtg tat cct ggc tcc gtc acc cca caa 1964

Ser Gln Glu Leu Leu Pro Gly Val Tyr Pro Gly Ser Val Thr Pro Gln  
 570 575 580

aac ggg ctt gct tca ggc atc ctg tcc ccc tcc cac gat gtc cct tcc 2012

Asn Gly Leu Ala Ser Gly Ile Leu Ser Pro Ser His Asp Val Pro Ser  
 585 590 595

agg cac cac cca ctg gag gtc ccc act ggt tcc cac cac cac ctg tcc 2060

Arg His His Pro Leu Glu Val Pro Thr Gly Ser His His His Leu Ser  
 600 605 610

cct ctg ccc aca gct gag agc acc agg gat ggc ctg ggg ccc agt ctc 2108

Pro Leu Pro Thr Ala Glu Ser Thr Arg Asp Gly Leu Gly Pro Ser Leu  
 615 620 625

ctt tca ccc atg gtc agc cca ctg aag ggg ctt ggt ccc cca ccg cta 2156

Leu Ser Pro Met Val Ser Pro Leu Lys Gly Leu Gly Pro Pro Pro Leu  
 630 635 640 645

cca cca gcc tcc cag agt cag tct cca ggg gga cag tca ttc tct aca 2204

Pro Pro Ala Ser Gln Ser Gln Ser Pro Gly Gly Gln Ser Phe Ser Thr  
 650 655 660

gtc ccc agc aag cct acc tac cca tcc ttc caa agc cca cca cct ctg 2252

Val Pro Ser Lys Pro Thr Tyr Pro Ser Phe Gln Ser Pro Pro Pro Leu

665	670	675	
ccc agc ccc caa ggc tac caa ggc agt ttc cat tcc atc cag aac tgc			2300
Pro Ser Pro Gln Gly Tyr Gln Gly Ser Phe His Ser Ile Gln Asn Cys			
680	685	690	
ttc ccc tac gct gac tgc tac cgg gcc act gag cca gca gcc tcc agg			2348
Phe Pro Tyr Ala Asp Cys Tyr Arg Ala Thr Glu Pro Ala Ala Ser Arg			
695	700	705	
gat gga ctg gtg ggt gat gcc cac ggt ttc aac ccc ttg cga ccc agc			2396
Asp Gly Leu Val Gly Asp Ala His Gly Phe Asn Pro Leu Arg Pro Ser			
710	715	720	725
aca tac tcc agc ctc agc aca cct tta tcc gca cca ggc tac gag acc			2444
Thr Tyr Ser Ser Leu Ser Thr Pro Leu Ser Ala Pro Gly Tyr Glu Thr			
730	735	740	
ctg gca gaa acg ccg tgt ccc cca gcg ctg cag cca cag cca gct gaa			2492
Leu Ala Glu Thr Pro Cys Pro Pro Ala Leu Gln Pro Gln Pro Ala Glu			
745	750	755	
gac ctg gta cct agt ggt cct gag gac tgt ggc ttc ttc ccc aat ggg			2540
Asp Leu Val Pro Ser Gly Pro Glu Asp Cys Gly Phe Phe Pro Asn Gly			
760	765	770	
gcc ttt gac cac tgt ctg agt cac atc ccg tcc atc tac act gac acc			2588
Ala Phe Asp His Cys Leu Ser His Ile Pro Ser Ile Tyr Thr Asp Thr			
775	780	785	
tga aggaaggggc gctgctctgc ctgcctgcct ggctcctgag ctacttcacc			2641
tacctgccat ctgctgggtc ttcccacag ggcagcaag gccacaccac aggttacttc			2701
cctacctgga ggctgtctg gtccagagct gctgccagg agctatggcc ctctgacagc			2761
cccatggctg tgtcttctc tctctccata aggttctcaa atcacagacc tcgtgtatat			2821
acaatgtaca ggacctttt tccgcgcc tgcaagtttt atatttttg ttttacaaga			2881
aaaacattaa aaactgaaa cta			2904

<210> 4

<211> 789

<212> PRT

<213> Mus musculus

<400> 4

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

Leu Pro Pro Thr Ser Ser Ala Ser Ser Ser Pro Cys Ala Ser Ser Asp  
 245 250 255  
 Ile Asn Pro Val Ile His Ser Ser Gln Thr Ala Leu Val Ser Cys Val  
 260 265 270  
 Asn Gly Leu Arg Ser Pro Pro Leu Pro Gly Asp Leu Gly Gly Pro Pro  
 275 280 285  
 Lys Arg Ser Arg Pro Gly Pro Ala Ser Ser Asp Gly Gln Glu Gly Ser  
 290 295 300  
 Leu Gln Leu Glu Ala Cys Arg Lys Ser Gly Phe Leu Lys Gln Glu Pro  
 305 310 315 320  
 Met Asp Glu Phe Ser Glu Leu Phe Ala Pro His His Gln Gly Leu Pro  
 325 330 335  
 Pro Pro Tyr Pro Leu Pro Gln Leu Pro Thr Gly Pro Gly Leu Gly Gly  
 340 345 350  
 Leu Gly Leu Gly Leu Ala Gly Arg Met Val Ala Gly Arg Gln Ala Cys  
 355 360 365  
 Arg Trp Val Asp Cys Cys Ala Ala Tyr Glu Gln Gln Glu Glu Leu Val  
 370 375 380  
 Arg His Ile Glu Lys Ser His Ile Asp Gln Arg Lys Gly Glu Asp Phe  
 385 390 395 400  
 Thr Cys Phe Trp Ala Gly Cys Val Arg Arg Tyr Lys Pro Phe Asn Ala  
 405 410 415  
 Arg Tyr Lys Leu Leu Ile His Met Arg Val His Ser Gly Glu Lys Pro  
 420 425 430  
 Asn Lys Cys Met Phe Glu Gly Cys Ser Lys Ala Phe Ser Arg Leu Glu  
 435 440 445  
 Asn Leu Lys Ile His Leu Arg Ser His Thr Gly Glu Lys Pro Tyr Leu  
 450 455 460  
 Cys Gln His Pro Gly Cys Gln Lys Ala Phe Ser Asn Ser Ser Asp Arg  
 465 470 475 480  
 Ala Lys His Gln Arg Thr His Leu Asp Thr Lys Pro Tyr Ala Cys Gln

485 490 495  
 Ile Pro Gly Cys Ser Lys Arg Tyr Thr Asp Pro Ser Ser Leu Arg Lys

500 505 510  
 His Val Lys Ala His Ser Ala Lys Glu Gln Gln Val Arg Lys Lys Leu

515 520 525  
 His Thr Gly Ala Asp Pro Glu Ala Asp Val Leu Ser Glu Cys Leu Ser

530 535 540  
 Leu Gln Gln Leu Gln Ala Ser Thr Leu Leu Pro Ala Ser Arg Gly Lys

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

565 570 575  
 Ser Val Thr Pro Gln Asn Gly Leu Ala Ser Gly Ile Leu Ser Pro Ser

580 585 590  
 His Asp Val Pro Ser Arg His His Pro Leu Glu Val Pro Thr Gly Ser

595 600 605  
 His His His Leu Ser Pro Leu Pro Thr Ala Glu Ser Thr Arg Asp Gly

610 615 620  
 Leu Gly Pro Ser Leu Leu Ser Pro Met Val Ser Pro Leu Lys Gly Leu

625 630 635 640  
 Gly Pro Pro Pro Leu Pro Pro Ala Ser Gln Ser Gln Ser Pro Gly Gly

645 650 655  
 Gln Ser Phe Ser Thr Val Pro Ser Lys Pro Thr Tyr Pro Ser Phe Gln

660 665 670  
 Ser Pro Pro Pro Leu Pro Ser Pro Gln Gly Tyr Gln Gly Ser Phe His

675 680 685  
 Ser Ile Gln Asn Cys Phe Pro Tyr Ala Asp Cys Tyr Arg Ala Thr Glu

690 695 700  
 Pro Ala Ala Ser Arg Asp Gly Leu Val Gly Asp Ala His Gly Phe Asn

705 710 715 720  
 Pro Leu Arg Pro Ser Thr Tyr Ser Ser Leu Ser Thr Pro Leu Ser Ala

725 730 735

Pro Gly Tyr Glu Thr Leu Ala Glu Thr Pro Cys Pro Pro Ala Leu Gln  
 740 745 750

Pro Gln Pro Ala Glu Asp Leu Val Pro Ser Gly Pro Glu Asp Cys Gly  
 755 760 765

Phe Phe Pro Asn Gly Ala Phe Asp His Cys Leu Ser His Ile Pro Ser  
 770 775 780

Ile Tyr Thr Asp Thr  
 785

<210> 5

<211> 2949

<212> DNA

<213> Homo sapiens

<220><221> CDS

<222> (595)..(2034)

<400> 5

agtttcccga ccagagagaa cgaacgtgtc tgcgggcgcg cggggagcag aggcggtggc 60  
 gggcgggcggc ggcaccggga gccgccgagt gaccctcccc cgcccctctg gcccccacc 120  
 ctcccaccgg cccgtggccc gcgcccatgg ccgcgcgcgc tccacacaac tcaccggagt 180

ccgcgccttg cgcgccgac cagttcgcag ctccgcgcca cggcagccag tctcacctgg 240  
 cggcaccgce cgcccaccgc cccggccaca gccctcgcgc ccacggcagc actcgaggcg 300  
 accgcgacag tgggtggggga cgctgctgag tggaagagag cgcagcccgg ccaccggacc 360  
 tacttactcg ccttgctgat tgtctatfff tgcgtttaca acttttctaa gaacttttgt 420  
 atacaaagga actttttaa aaagacgctt ccaagttata tttaatccaa agaagaagga 480  
 tctcgcccaa tttggggttt tgggttttgg ctccgtttct tctcttcgtt gactttgggg 540  
 ttcaggtgcc ccagctgctt cgggctgccg aggaccttct gggccccac atta atg 597

Met

1

agg cag cca cct ggc gag tct gac atg gct gtc agc gac gcg ctg etc 645

Arg Gln Pro Pro Gly Glu Ser Asp Met Ala Val Ser Asp Ala Leu Leu  
 5 10 15

cca tct ttc tcc acg ttc gcg tct ggc ccg gcg gga agg gag aag aca 693

Pro Ser Phe Ser Thr Phe Ala Ser Gly Pro Ala Gly Arg Glu Lys Thr  
 20 25 30  
 ctg cgt caa gca ggt gcc ccg aat aac cgc tgg cgg gag gag etc tcc 741  
 Leu Arg Gln Ala Gly Ala Pro Asn Asn Arg Trp Arg Glu Glu Leu Ser  
 35 40 45  
 cac atg aag cga ctt ccc cca gtg ctt ccc ggc cgc ccc tat gac ctg 789  
 His Met Lys Arg Leu Pro Pro Val Leu Pro Gly Arg Pro Tyr Asp Leu  
 50 55 60 65  
 gcg gcg gcg acc gtg gcc aca gac ctg gag agc ggc gga gcc ggt gcg 837  
 Ala Ala Ala Thr Val Ala Thr Asp Leu Glu Ser Gly Gly Ala Gly Ala  
 70 75 80  
 gct tgc ggc ggt agc aac ctg gcg ccc cta cct cgg aga gag acc gag 885  
 Ala Cys Gly Gly Ser Asn Leu Ala Pro Leu Pro Arg Arg Glu Thr Glu  
 85 90 95  
 gag ttc aac gat ctc ctg gac ctg gac ttt att ctc tcc aat tcg ctg 933  
 Glu Phe Asn Asp Leu Leu Asp Leu Asp Phe Ile Leu Ser Asn Ser Leu  
 100 105 110  
 acc cat cct ccg gag tca gtg gcc gcc acc gtg tcc tcg tca gcg tca 981  
 Thr His Pro Pro Glu Ser Val Ala Ala Thr Val Ser Ser Ser Ala Ser  
 115 120 125  
 gcc tcc tct tcg tcg tcg ccg tcg agc agc ggc cct gcc agc gcg ccc 1029  
 Ala Ser Ser Ser Ser Ser Pro Ser Ser Ser Gly Pro Ala Ser Ala Pro  
 130 135 140 145  
 tcc acc tgc agc ttc acc tat ccg atc cgg gcc ggg aac gac ccg ggc 1077  
 Ser Thr Cys Ser Phe Thr Tyr Pro Ile Arg Ala Gly Asn Asp Pro Gly  
 150 155 160  
 gtg gcg ccg ggc ggc acg ggc gga ggc ctc ctc tat ggc agg gag tcc 1125  
 Val Ala Pro Gly Gly Thr Gly Gly Gly Leu Leu Tyr Gly Arg Glu Ser  
 165 170 175

gct ccc cct ccg acg gct ccc ttc aac ctg gcg gac atc aac gac gtg 1173  
 Ala Pro Pro Pro Thr Ala Pro Phe Asn Leu Ala Asp Ile Asn Asp Val  
 180 185 190

agc ccc tcg ggc ggc ttc gtg gcc gag ctc ctg cgg cca gaa ttg gac 1221  
 Ser Pro Ser Gly Gly Phe Val Ala Glu Leu Leu Arg Pro Glu Leu Asp  
 195 200 205

ccg gtg tac att ccg ccg cag cag ccg cag ccg cca ggt ggc ggg ctg 1269

Pro Val Tyr Ile Pro Pro Gln Gln Pro Gln Pro Pro Gly Gly Gly Leu  
 210 215 220 225

atg ggc aag ttc gtg ctg aag gcg tcg ctg agc gcc cct ggc agc gag 1317  
 Met Gly Lys Phe Val Leu Lys Ala Ser Leu Ser Ala Pro Gly Ser Glu  
 230 235 240

tac ggc agc ccg tcg gtc atc agc gtc agc aaa ggc agc cct gac ggc 1365  
 Tyr Gly Ser Pro Ser Val Ile Ser Val Ser Lys Gly Ser Pro Asp Gly  
 245 250 255

agc cac ccg gtg gtg gtg gcg ccc tac aac ggc ggg ccg ccg cgc acg 1413  
 Ser His Pro Val Val Val Ala Pro Tyr Asn Gly Gly Pro Pro Arg Thr  
 260 265 270

tgc ccc aag atc aag cag gag gcg gtc tct tcg tgc acc cac ttg ggc 1461  
 Cys Pro Lys Ile Lys Gln Glu Ala Val Ser Ser Cys Thr His Leu Gly  
 275 280 285

gct gga ccc cct ctc agc aat ggc cac cgg ccg gct gca cac gac ttc 1509  
 Ala Gly Pro Pro Leu Ser Asn Gly His Arg Pro Ala Ala His Asp Phe  
 290 295 300 305

ccc ctg ggg cgg cag ctc ccc agc agg act acc ccg acc ctg ggt ctt 1557  
 Pro Leu Gly Arg Gln Leu Pro Ser Arg Thr Thr Pro Thr Leu Gly Leu  
 310 315 320

gag gaa gtg ctg agc agc agg gac tgt cac cct gcc ctg ccg ctt cct 1605

Glu Glu Val Leu Ser Ser Arg Asp Cys His Pro Ala Leu Pro Leu Pro  
 325 330 335

ccc ggc ttc cat ccc cac ccg ggg ccc aat tac cca tcc ttc ctg ccc 1653

Pro Gly Phe His Pro His Pro Gly Pro Asn Tyr Pro Ser Phe Leu Pro  
 340 345 350  
 gat cag atg cag ccg caa gtc ccg ccg ctc cat tac caa gag ctc atg 1701  
 Asp Gln Met Gln Pro Gln Val Pro Pro Leu His Tyr Gln Glu Leu Met  
 355 360 365  
 cca ccc ggt tcc tgc atg cca gag gag ccc aag cca aag agg gga aga 1749  
 Pro Pro Gly Ser Cys Met Pro Glu Glu Pro Lys Pro Lys Arg Gly Arg  
 370 375 380 385  
 cga tcg tgg ccc cgg aaa agg acc gcc acc cac act tgt gat tac gcg 1797  
 Arg Ser Trp Pro Arg Lys Arg Thr Ala Thr His Thr Cys Asp Tyr Ala  
 390 395 400  
 ggc tgc ggc aaa acc tac aca aag agt tcc cat ctc aag gca cac ctg 1845  
 Gly Cys Gly Lys Thr Tyr Thr Lys Ser Ser His Leu Lys Ala His Leu  
 405 410 415  
 cga acc cac aca ggt gag aaa cct tac cac tgt gac tgg gac ggc tgt 1893  
 Arg Thr His Thr Gly Glu Lys Pro Tyr His Cys Asp Trp Asp Gly Cys  
 420 425 430  
 gga tgg aaa ttc gcc cgc tca gat gaa ctg acc agg cac tac cgt aaa 1941  
 Gly Trp Lys Phe Ala Arg Ser Asp Glu Leu Thr Arg His Tyr Arg Lys  
 435 440 445  
 cac acg ggg cac cgc ccg ttc cag tgc caa aaa tgc gac cga gca ttt 1989  
 His Thr Gly His Arg Pro Phe Gln Cys Gln Lys Cys Asp Arg Ala Phe  
 450 455 460 465  
 tcc agg tcg gac cac ctc gcc tta cac atg aag agg cat ttt taa 2034  
 Ser Arg Ser Asp His Leu Ala Leu His Met Lys Arg His Phe  
 470 475  
 atcccagaca gtggatatga cccacactgc cagaagagaa ttcagtatgt tttacttttc 2094  
 aactgtcttt cccgatgagg gaaggagccc agccagaaag cactacaatc atggtaagt 2154  
 tcccaactga gtcactttgt gactggataa tcaggaaaaa tgaggaatcc aaaagacaaa 2214  
 aatcaaagaa catgagggt ctgtgactgg atctttctatc attccaatc taaatccgac 2274  
 ttgaatattc ctggacttac aaaatgcca ggggtgact ggaagtgtg gatatcaggg 2334

tataaattat atccgtgagt tgggggaggg aagaccagaa ttccttgaa ttgtgtattg 2394

atgcaatata agcataaaag atcaccttgt attctcttta ccttctaaaa gccattatta 2454

tgatgttaga agaagaggaa gaaattcagg tacagaaaac atgtttaaat agcctaaatg 2514

atggtgcttg gtgagtcttg gttctaaagg taccaaaca ggaagccaaa gttttcaaac 2574

tgctgcatac tttgacaagg aaaatctata tttgtcttcc gatcaacatt tatgacctaa 2634

gtcaggtaat atacctgggt tacttcttta gcatttttat gcagacagtc tgttatgcac 2694

tgtggtttca gatgtgcaat aatttgtaca atggtttatt cccaagtag ccttaagcag 2754

aacaaatgtg tttttctata tagttccttg ccttaataaa tatgtaatat aaatttaagc 2814

aaacgtctat tttgtatatt tgtaactac aaagtaaaat gaacattttg tggagtttgt 2874

attttgcata ctcaaggiga gaattaagt ttaaataaac ctataatatt ttatctgaaa 2934

aaaaaaaaaa aaaaa 2949

<210> 6

<211> 479

<212> PRT

<213> Homo sapiens

<400> 6

Met Arg Gln Pro Pro Gly Glu Ser Asp Met Ala Val Ser Asp Ala Leu

1                    5                    10                    15

Leu Pro Ser Phe Ser Thr Phe Ala Ser Gly Pro Ala Gly Arg Glu Lys

                  20                    25                    30

Thr Leu Arg Gln Ala Gly Ala Pro Asn Asn Arg Trp Arg Glu Glu Leu

                  35                    40                    45

Ser His Met Lys Arg Leu Pro Pro Val Leu Pro Gly Arg Pro Tyr Asp

                  50                    55                    60

Leu Ala Ala Ala Thr Val Ala Thr Asp Leu Glu Ser Gly Gly Ala Gly

65                    70                    75                    80

Ala Ala Cys Gly Gly Ser Asn Leu Ala Pro Leu Pro Arg Arg Glu Thr

                  85                    90                    95

Glu Glu Phe Asn Asp Leu Leu Asp Leu Asp Phe Ile Leu Ser Asn Ser

                  100                    105                    110

Leu Thr His Pro Pro Glu Ser Val Ala Ala Thr Val Ser Ser Ser Ala



Met Pro Pro Gly Ser Cys Met Pro Glu Glu Pro Lys Pro Lys Arg Gly  
 370 375 380  
 Arg Arg Ser Trp Pro Arg Lys Arg Thr Ala Thr His Thr Cys Asp Tyr  
 385 390 395 400  
 Ala Gly Cys Gly Lys Thr Tyr Thr Lys Ser Ser His Leu Lys Ala His  
 405 410 415  
  
 Leu Arg Thr His Thr Gly Glu Lys Pro Tyr His Cys Asp Trp Asp Gly  
 420 425 430  
 Cys Gly Trp Lys Phe Ala Arg Ser Asp Glu Leu Thr Arg His Tyr Arg  
 435 440 445  
 Lys His Thr Gly His Arg Pro Phe Gln Cys Gln Lys Cys Asp Arg Ala  
 450 455 460  
 Phe Ser Arg Ser Asp His Leu Ala Leu His Met Lys Arg His Phe  
 465 470 475

<210> 7

<211> 3057

<212> DNA

<213> Mus musculus

<220><221> CDS

<222> (605)..(2056)

<400> 7

agttccccgg ccaagagagc gagcgcggct cggggcgcgc ggggagcaga ggcggtggcg 60  
 ggcggcggcg gcacccggag ccgccgagtg ccctccccg cccctccagc cccccacca 120  
 gcaaccgcc cgtgaccgc gcccatggcc gcgcgcaccc ggcacagtcc ccaggactcc 180  
 gcaccccgcg ccaccgcca gctcgcagtt ccgcgccacc gcggccattc tcacctggcg 240  
 gcgccgccc cccaccgcc ggaccacagc ccccgcccg ccgacagcca cagtggccgc 300  
 gacaacggtg ggggacactg ctgagtccaa gagcgtgcag cctggccatc ggacctactt 360  
  
 atctgccttg ctgattgtct atttttataa gagtttacia cttttctaag aatttttgta 420  
 taaaaggaa cttttttaa gacatcgccg gtttatattg aatccaaaga agaaggatct 480  
 cgggcaatct gggggttttg gtttgaggtt ttgtttctaa agtttttaat cttcgttgac 540  
 tttggggctc aggtaccctt ctctcttctt cggactccgg aggaccttct gggccccac 600  
 atta atg agg cag cca cct ggc gag tct gac atg gct gtc agc gac gct 649



gct gcc agc aac aca ggt gga ggg ctc ctc tac agc cga gaa tct gcg 1129

Ala Ala Ser Asn Thr Gly Gly Gly Leu Leu Tyr Ser Arg Glu Ser Ala  
 160 165 170 175

cca cct ccc acg gcc ccc ttc aac ctg gcg gac atc aat gac gtg agc 1177  
 Pro Pro Pro Thr Ala Pro Phe Asn Leu Ala Asp Ile Asn Asp Val Ser  
 180 185 190

ccc tcg ggc ggc ttc gtg gct gag ctc ctg cgg ccg gag ttg gac cca 1225  
 Pro Ser Gly Gly Phe Val Ala Glu Leu Leu Arg Pro Glu Leu Asp Pro

195 200 205

gta tac att ccg cca cag cag cct cag ccg cca ggt ggc ggg ctg atg 1273  
 Val Tyr Ile Pro Pro Gln Gln Pro Gln Pro Pro Gly Gly Gly Leu Met  
 210 215 220

ggc aag ttt gtg ctg aag gcg tct ctg acc acc cct ggc agc gag tac 1321  
 Gly Lys Phe Val Leu Lys Ala Ser Leu Thr Thr Pro Gly Ser Glu Tyr  
 225 230 235

agc agc cct tcg gtc atc agt gtt agc aaa gga agc cca gac ggc agc 1369  
 Ser Ser Pro Ser Val Ile Ser Val Ser Lys Gly Ser Pro Asp Gly Ser  
 240 245 250 255

cac ccc gtg gta gtg gcg ccc tac agc ggt ggc ccg ccg cgc atg tgc 1417  
 His Pro Val Val Val Ala Pro Tyr Ser Gly Gly Pro Pro Arg Met Cys  
 260 265 270

ccc aag att aag caa gag gcg gtc ccg tcc tgc acg gtc agc cgg tcc 1465

Pro Lys Ile Lys Gln Glu Ala Val Pro Ser Cys Thr Val Ser Arg Ser  
 275 280 285

cta gag gcc cat ttg agc gct gga ccc cag ctc agc aac ggc cac cgg 1513  
 Leu Glu Ala His Leu Ser Ala Gly Pro Gln Leu Ser Asn Gly His Arg  
 290 295 300

ccc aac aca cac gac ttc ccc ctg ggg cgg cag ctc ccc acc agg act 1561  
 Pro Asn Thr His Asp Phe Pro Leu Gly Arg Gln Leu Pro Thr Arg Thr

305	310	315	
acc cct aca ctg agt ccc gag gaa ctg ctg aac agc agg gac tgt cac			1609
Thr Pro Thr Leu Ser Pro Glu Glu Leu Leu Asn Ser Arg Asp Cys His			
320	325	330	335
cct ggc ctg cct ctt ccc cca gga ttc cat ccc cat ccg ggg ccc aac			1657
Pro Gly Leu Pro Leu Pro Pro Gly Phe His Pro His Pro Gly Pro Asn			
	340	345	350
tac cct cct ttc ctg cca gac cag atg cag tca caa gtc ccc tct ctc			1705
Tyr Pro Pro Phe Leu Pro Asp Gln Met Gln Ser Gln Val Pro Ser Leu			
	355	360	365
cat tat caa gag ctc atg cca ccg ggt tcc tgc ctg cca gag gag ccc			1753
His Tyr Gln Glu Leu Met Pro Pro Gly Ser Cys Leu Pro Glu Glu Pro			
	370	375	380
aag cca aag agg gga aga agg tcg tgg ccc cgg aaa aga aca gcc acc			1801
Lys Pro Lys Arg Gly Arg Arg Ser Trp Pro Arg Lys Arg Thr Ala Thr			
	385	390	395
cac act tgt gac tat gca ggc tgt ggc aaa acc tat acc aag agt tct			1849
His Thr Cys Asp Tyr Ala Gly Cys Gly Lys Thr Tyr Thr Lys Ser Ser			
400	405	410	415
cat ctc aag gca cac ctg cga act cac aca ggc gag aaa cct tac cac			1897
His Leu Lys Ala His Leu Arg Thr His Thr Gly Glu Lys Pro Tyr His			
	420	425	430
tgt gac tgg gac ggc tgt ggg tgg aaa ttc gcc cgc tcc gat gaa ctg			1945
Cys Asp Trp Asp Gly Cys Gly Trp Lys Phe Ala Arg Ser Asp Glu Leu			
	435	440	445
acc agg cac tac cgc aaa cac aca ggg cac cgg ccc ttt cag tgc cag			1993
Thr Arg His Tyr Arg Lys His Thr Gly His Arg Pro Phe Gln Cys Gln			
	450	455	460
aag tgt gac agg gcc ttt tcc agg tcg gac cac ctt gcc tta cac atg			2041
Lys Cys Asp Arg Ala Phe Ser Arg Ser Asp His Leu Ala Leu His Met			
	465	470	475

aag agg cac ttt taa atcccacgta gtggatgtga cccacactgc caggagagag 2096

Lys Arg His Phe

480

agttcagtat ttttttttct aacctttcac actgtcttcc cagagggga ggagcccagc 2156

tggcaagcgc tacaatcatg gtcaagttcc cagcaagtca gcttgtgaat ggataatcag 2216

gagaaaggaa gagtccaaga gacaaaacag aaatactaaa aacaacaaaa caaaaaaaca 2276

aacaaaaaaa acaagaaaaa aaaatcacag aacagatggg gtctgatact ggatggatct 2336

tctatcattc caatacaaaa tccaacttga acatgcccg acttacaaaa tgccaagggg 2396

tgactggaag tttgtggata tcagggtata cactaaatca gtgagcttgg ggggagggaa 2456

gaccaggatt cccttgaatt gtgtttcgat gatgcaatac acacgtaaag atcaccttgt 2516

atgctctttg ccttcttaaa aaaaaaaaaa gccattattg tgtcggagga agaggaagcg 2576

atcaggtac agaacatggt ctaacagcct aaatgatggt gcttggtag tcgtggttct 2636

aaaggtacca aacgggggag ccaaagtct ccaactgctg catacttttg acaaggaaaa 2696

tctagttttg tcttccgac tacattgatg acctaagcca ggtaataaag cctggtttat 2756

ttctgtaaca tttttatgca gacagtctgt tatgactgt ggtttcagat gtgcaataat 2816

ttgtacaatg gtttattccc aagtatgcct ttaagcagaa caaatgtgtt tttctatata 2876

gttcttgc ttaataaata tgtaataata atttaagcaa acttctattt tgtatatttg 2936

taactacaa agtaaaaaa aatgaacatt ttgtggagtt tgtattttgc atactcaagg 2996

tgagaaataa gttttaaata aacctataat attttatctg aacgacaaaa aaaaaaaaaa 3056

a 3057

<210> 8

<211> 483

<212> PRT

<213> Mus musculus

<400> 8

Met Arg Gln Pro Pro Gly Glu Ser Asp Met Ala Val Ser Asp Ala Leu

1 5 10 15

Leu Pro Ser Phe Ser Thr Phe Ala Ser Gly Pro Ala Gly Arg Glu Lys

20 25 30

Thr Leu Arg Pro Ala Gly Ala Pro Thr Asn Arg Trp Arg Glu Glu Leu

35

40

45

Ser His Met Lys Arg Leu Pro Pro Leu Pro Gly Arg Pro Tyr Asp Leu  
 50 55 60  
 Ala Ala Thr Val Ala Thr Asp Leu Glu Ser Gly Gly Ala Gly Ala Ala  
 65 70 75 80  
 Cys Ser Ser Asn Asn Pro Ala Leu Leu Ala Arg Arg Glu Thr Glu Glu  
 85 90 95  
 Phe Asn Asp Leu Leu Asp Leu Asp Phe Ile Leu Ser Asn Ser Leu Thr  
 100 105 110  
 His Gln Glu Ser Val Ala Ala Thr Val Thr Thr Ser Ala Ser Ala Ser  
 115 120 125  
 Ser Ser Ser Ser Pro Ala Ser Ser Gly Pro Ala Ser Ala Pro Ser Thr  
 130 135 140  
 Cys Ser Phe Ser Tyr Pro Ile Arg Ala Gly Gly Asp Pro Gly Val Ala  
 145 150 155 160  
 Ala Ser Asn Thr Gly Gly Gly Leu Leu Tyr Ser Arg Glu Ser Ala Pro  
 165 170 175  
 Pro Pro Thr Ala Pro Phe Asn Leu Ala Asp Ile Asn Asp Val Ser Pro  
 180 185 190  
 Ser Gly Gly Phe Val Ala Glu Leu Leu Arg Pro Glu Leu Asp Pro Val  
 195 200 205  
 Tyr Ile Pro Pro Gln Gln Pro Gln Pro Pro Gly Gly Gly Leu Met Gly  
 210 215 220  
 Lys Phe Val Leu Lys Ala Ser Leu Thr Thr Pro Gly Ser Glu Tyr Ser  
 225 230 235 240  
 Ser Pro Ser Val Ile Ser Val Ser Lys Gly Ser Pro Asp Gly Ser His  
 245 250 255  
 Pro Val Val Val Ala Pro Tyr Ser Gly Gly Pro Pro Arg Met Cys Pro  
 260 265 270  
 Lys Ile Lys Gln Glu Ala Val Pro Ser Cys Thr Val Ser Arg Ser Leu  
 275 280 285  
 Glu Ala His Leu Ser Ala Gly Pro Gln Leu Ser Asn Gly His Arg Pro

290 295 300  
 Asn Thr His Asp Phe Pro Leu Gly Arg Gln Leu Pro Thr Arg Thr Thr  
 305 310 315 320  
 Pro Thr Leu Ser Pro Glu Glu Leu Leu Asn Ser Arg Asp Cys His Pro  
 325 330 335  
 Gly Leu Pro Leu Pro Pro Gly Phe His Pro His Pro Gly Pro Asn Tyr  
 340 345 350  
 Pro Pro Phe Leu Pro Asp Gln Met Gln Ser Gln Val Pro Ser Leu His

355 360 365  
 Tyr Gln Glu Leu Met Pro Pro Gly Ser Cys Leu Pro Glu Glu Pro Lys  
 370 375 380  
 Pro Lys Arg Gly Arg Arg Ser Trp Pro Arg Lys Arg Thr Ala Thr His  
 385 390 395 400  
 Thr Cys Asp Tyr Ala Gly Cys Gly Lys Thr Tyr Thr Lys Ser Ser His  
 405 410 415  
 Leu Lys Ala His Leu Arg Thr His Thr Gly Glu Lys Pro Tyr His Cys

420 425 430  
 Asp Trp Asp Gly Cys Gly Trp Lys Phe Ala Arg Ser Asp Glu Leu Thr  
 435 440 445  
 Arg His Tyr Arg Lys His Thr Gly His Arg Pro Phe Gln Cys Gln Lys  
 450 455 460  
 Cys Asp Arg Ala Phe Ser Arg Ser Asp His Leu Ala Leu His Met Lys  
 465 470 475 480  
 Arg His Phe

<210> 9

<211> 81

<212> DNA

<213> Foot-and-mouth disease virus

<400> 9

aaaattgtcg ctctgtcaa acaaactctt aactttgatt tactcaaact ggctggggat 60

gtagaaagca atccaggtcc a 81

<210> 10	
<211> 34	
<212> DNA	
<213> Bacteriophage P1	
<400> 10	
ataacttcgt atagcataca ttatacgaag ttat	34
<210> 11	
<211> 34	
<212> DNA	
<213> Artificial Sequence	
<220><223> mutant loxP (lox71) sequence	
<400> 11	
taccgttcgt atagcataca ttatacgaag ttat	34
<210> 12	
<211> 34	
<212> DNA	
<213> Artificial Sequence	
<220><223> mutant loxP (lox66) sequence	
<400> 12	
ataacttcgt atagcataca ttatacgaac ggta	34
<210> 13	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> shRNA	
<400> 13	
ggcctcacca accctgcacc t	21
<210> 14	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> shRNA	
<400> 14	

gcccttcaat gcccgctaca a

21

<210> 15

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> shRNA

<400> 15

gggcaatgaa cccatctcag a

21