(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0141470
(43) 공개일자 2020년12월18일

(51) 국제특허분류(Int. Cl.)
C12N 5/16 (2006.01) *A01K 67/027* (2006.01)
C12N 5/075 (2010.01) *C12N 9/02* (2006.01)
(52) CPC특허분류
C12N 5/16 (2013.01)
A01K 67/027 (2013.01)
(21) 출원번호 10-2020-7031724
(22) 출원일자(국제) 2019년04월05일
심사청구일자 없음
(85) 번역문제출일자 2020년11월03일
(86) 국제출원번호 PCT/US2019/026074
(87) 국제공개번호 WO 2019/195738
국제공개일자 2019년10월10일
(30) 우선권주장
62/654,199 2018년04월06일 미국(US)

(71) 출원인
칠드런'즈 메디컬 센터 코포레이션
미국 매사추세츠 보스턴 새턱 스트리트 55 (우편번호:02115)
(72) 발명자
장, 이
미국 02115 매사추세츠주 보스턴 새턱 스트리트 55 칠드런'즈 메디컬 센터 코포레이션 내
마토바, 쇼고
미국 02115 매사추세츠주 보스턴 새턱 스트리트 55 칠드런'즈 메디컬 센터 코포레이션 내
(74) 대리인
양영준, 이상남

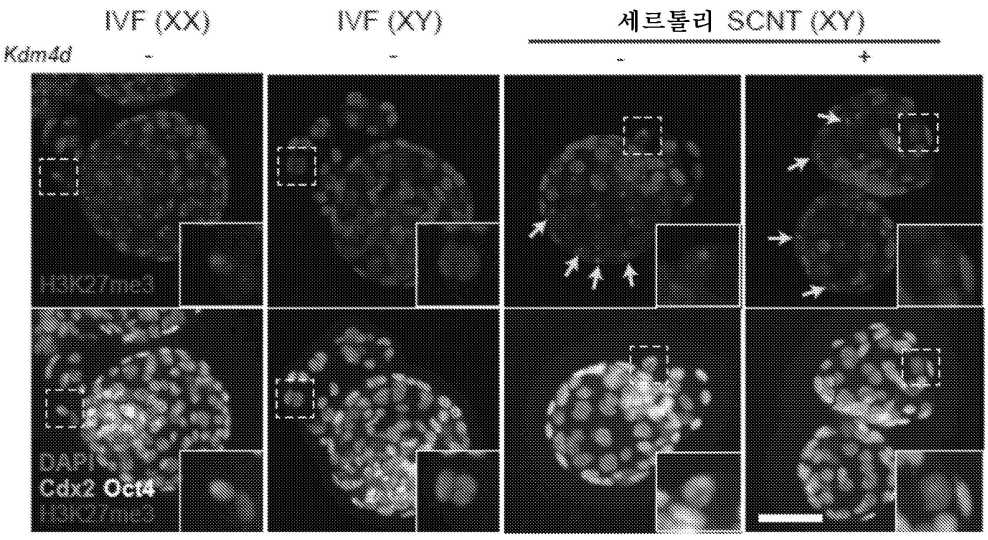
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 **체세포 재프로그래밍 및 각인의 조정을 위한 조성물 및 방법**

(57) 요 약

본 발명은 클로닝 효율을 개선시키고, 각인 제어 영역을 조정하는 방법을 제공한다. 특정 실시양태에서, 본 발명은 각인 제어 영역 내에서 억제된 대립유전자를 활성화시킴으로써 각인 연관된 장애를 치료하는 방법을 제공한다. 다른 실시양태에서, 본 발명은 *Kdm4d* 과발현이 *Xist* 넉아웃 공여자 세포인 것을 포함하는 체세포 핵 전달 효율을 개선시키는 방법을 제공한다.

대 표 도



(52) CPC특허분류
C12N 5/0609 (2013.01)
C12N 9/0071 (2013.01)

명 세 서

청구범위

청구항 1

Xist 활성이 결여된 체세포로부터 얻어진 공여자 핵을 제핵된 난모세포 내로 전달하고, 난모세포에서 *Kdm4d*를 발현시킴으로써 클로닝된 배반포를 얻는 것을 포함하는, 클로닝된 배반포를 얻는 방법.

청구항 2

제1항에 있어서, 난모세포가 *Kdm4d* mRNA로 주사되는 것인 방법.

청구항 3

제1항에 있어서, 공여자 세포 핵이 Xist에서의 결실을 포함하거나 Xist의 불활성 형태를 포함하는 배아 섬유모세포로부터 얻어지는 것인 방법.

청구항 4

제1항에 있어서, 공여자 핵이 인간, 고양이, 소, 개, 돼지, 또는 말로부터 얻어지는 것인 방법.

청구항 5

제1항에 있어서, 배반포를 임신을 위해 숙주 자궁 내로 전달하는 것을 더 포함하는 방법.

청구항 6

제5항에 있어서, 방법이 출생률을 통상적인 체세포 핵 전달에 비해 적어도 약 10 내지 20% 증가시키는 것인 방법.

청구항 7

(a) 대상체로부터 얻어진 배양된 세포에서 *Xist*를 불활성화시키거나 Xist 활성 또는 발현을 감소시키는 단계;

(b) 배양된 세포로부터의 핵을 제핵된 난모세포 내로 전달함으로써 난모세포를 활성화시키는 단계; 및

(c) 단계 (b)에서 얻어진 활성화된 난모세포를 *Kdm4d* mRNA로 주사하고, 생성된 세포를 배양함으로써 대상체 내로의 이식에 적합한 세포 또는 조직을 얻는 단계

를 포함하는, 대상체 내로의 이식을 위한 세포 또는 조직을 얻는 방법.

청구항 8

제7항에 있어서, Xist가 게놈 편집에 의해 불활성화되는 것인 방법.

청구항 9

제7항에 있어서, CRISPR 시스템이 결실 또는 불활성화 돌연변이를 게놈 *Xist* 폴리뉴클레오티드에 도입하기 위해 사용되는 것인 방법.

청구항 10

제7항에 있어서, Xist 폴리뉴클레오티드 발현 또는 활성이 siRNA 또는 shRNA를 사용하여 감소되는 것인 방법.

청구항 11

제1항의 방법에 따라 제조된 배반포.

청구항 12

Xist에서의 결실을 포함하거나 감소된 수준의 Xist 발현을 갖고, *Kdm4d*를 코딩하는 이종 폴리뉴클레오티드를 포

함하는 세포.

청구항 13

제7항의 방법에 따라 제조된 세포 또는 조직.

청구항 14

제1항의 배반포를 숙주 자궁 내로 착상시킴으로써 제조된 클로닝된 유기체.

청구항 15

Xist 활성이 결여된 체세포로부터 얻어진 공여자 핵을 포함하고, 통상적인 난모세포에 비해 증가된 수준의 Kdm4d를 발현하는 난모세포.

발명의 설명

기 술 분 야

[0001] **관련 출원에 대한 상호 참조**

[0002] 이 출원은 그의 전체 내용이 본원에 참조로 포함되는 2018년 4월 6일에 출원된 미국 가출원 제62/654,199호의 우선권을 주장한다.

[0003] **연방정부 지원된 연구 하에 이루어진 발명에 대한 권리의 진술**

[0004] 본 발명은 국립 보건원 (National Institutes of Health)에 의해 수여된 허가 번호 HD092465 하의 정부 지원으로 이루어졌다. 정부는 본 발명에서 특정 권리를 갖는다.

배 경 기 술

[0005] 포유동물 난모세포는 체세포 핵 전달 (SCNT)을 통해 체세포를 전능한 상태로 재프로그래밍할 수 있다. SCNT는 의도되는 숙주와 유전적으로 동일하거나 유사한 공여자 유기체로부터의 조직의 생성을 포함하는 치료적 클로닝에 사용된다. SCNT는 또한 동물의 클로닝을 가능하게 한다. 이 기법은 농생명공학에서 뿐만 아니라 멸종위기종의 보존에서 큰 잠재성을 갖는다. 그러나, 극도로 낮은 클로닝의 성공률은 이 기법의 실제 사용을 곤란하게 만든다. 예를 들어, 마우스의 경우, SCNT 배아의 단지 약 30%만이 배반포로 발달하고, 대리모에 전달된 배아의 단지 1 내지 2%만이 만기에 도달할 수 있다. 더욱이, 생존한 배아에서, 거의 모든 클로닝된 포유동물 종에서, 배외 조직, 예컨대 태반 및 탯줄에서 이상이 빈번히 관찰된다. 이들 관찰은 SCNT 재프로그래밍이 배아 발생 프로세스를 방해하는 일부 결함을 가짐을 시사한다. DNA 메틸화, 히스톤 변형, 및 게놈 각인에 있어서의 다양한 후생적 이상은 SCNT의 낮은 성공률에 연루되었다. 클로닝의 효율을 개선시키기 위한 상당한 필요가 있다.

발명의 내용

[0006] **발명의 요약**

[0007] 본 발명은 클로닝 효율을 개선시키는 방법을 제공한다. 본 발명은 클로닝 효율을 개선시키는 방법을 제공한다. 특정 실시양태에서, 본 발명은 *Kdm4d* 과발현이 *Xist* 넉아웃 공여자 세포인 것을 포함하는 체세포 핵 전달 효율을 개선시키는 방법을 제공한다.

[0008] 한 측면에서, 본 발명은 Xist 활성이 결여된 체세포로부터 얻어진 공여자 핵을 제핵된 난모세포 내로 전달하고, 난모세포에서 *Kdm4d*를 발현시킴으로써 클로닝된 배반포를 얻는 것을 포함하는, 클로닝된 배반포를 얻는 방법을 제공한다. 방법의 일부 실시양태에서, 난모세포는 *Kdm4d* mRNA로 주사된다. 일부 실시양태에서, 공여자 세포 핵은 Xist에서의 결실을 포함하거나 Xist의 불활성 형태를 포함하는 배아 섬유모세포로부터 얻어진다. 일부 실시양태에서, 공여자 핵은 인간, 고양이, 소, 개, 돼지, 또는 말로부터 얻어진다. 일부 실시양태에서, 방법은 또한 배반포를 임신을 위해 숙주 자궁 내로 전달하는 것을 포함한다. 일부 실시양태에서, 방법은 출생률을 통상적인 체세포 핵 전달에 비해 적어도 약 10 내지 20% 증가시킨다. 본 발명의 일부 측면은 상기 기재된 방법에 의해 제조된 배반포를 포함한다. 본 발명의 일부 측면은 상기 기재된 방법에 의해 제조된 배반포를 착상시킴으로써 제조된 클로닝된 유기체를 포함한다.

[0009] 또다른 측면에서, 본 발명은 대상체로부터 얻어진 배양된 세포에서 *Xist*를 불활성화시키거나 Xist 활성 또는 발현을 감소시키는 단계; 배양된 세포로부터의 핵을 제핵된 난모세포 내로 전달함으로써 난모세포를 활성화시키는 단계; 및 활성화된 난모세포를 *Kdm4d* mRNA로 주사하고, 생성된 세포를 배양함으로써 대상체 내로의 이식에 적합한 세포 또는 조직을 얻는 단계를 포함하는, 대상체 내로의 이식을 위한 세포 또는 조직을 얻는 방법을 제공한다. 본 발명의 일부 측면에서, 이 방법에 의해 제조된 세포 또는 조직이 제공된다. 일부 실시양태에서, Xist는 게놈 편집에 의해 불활성화된다. 예를 들어, 일부 실시양태에서, CRISPR 시스템은 결실 또는 불활성화 돌연변이를 게놈 *Xist* 폴리뉴클레오티드에 도입하기 위해 사용된다. 방법의 다른 실시양태에서, Xist 폴리뉴클레오티드 발현 또는 활성은 siRNA 또는 shRNA를 사용하여 감소된다.

[0010] 본 발명의 다른 측면에서, Xist에서의 결실 또는 감소된 수준의 Xist 발현을 갖고, *Kdm4d*를 코딩하는 이종 폴리뉴클레오티드를 갖는 세포가 제공된다.

[0011] 추가의 측면은 Xist 활성이 결여된 체세포로부터 얻어진 공여자 핵을 포함하고, 통상적인 난모세포에 비해 증가된 수준의 Kdm4d를 발현하는 난모세포를 포함한다.

[0012] **정의**

[0013] 달리 정의되지 않는다면, 본원에 사용된 모든 기술 및 과학 용어는 본 발명이 속하는 관련 기술분야의 통상의 기술자에 의해 통상적으로 이해되는 의미를 갖는다. 하기 참고문헌은 통상의 기술자에게 본 발명에 사용된 많은 용어의 일반적 정의를 제공한다: 문헌 [Singleton et al., Dictionary of Microbiology and Molecular Biology (2nd ed. 1994)]; [The Cambridge Dictionary of Science and Technology (Walker ed., 1988)]; [The Glossary of Genetics, 5th Ed., R. Rieger et al. (eds.), Springer Verlag (1991)]; 및 [Hale & Marham, The Harper Collins Dictionary of Biology (1991)]. 본원에 사용된 바와 같은 하기 용어는 달리 명시되지 않는다면 하기에 그들에게 할당된 의미를 갖는다.

[0014] "KDM4D 폴리펩티드"란 NCBI 참조 번호 Q6B0I6과 적어도 약 85% 아미노산 서열 동일성을 갖고, 데메틸라제 활성을 갖는 폴리펩티드 또는 그의 단편을 의미한다. 예시적인 KDM4D 아미노산 서열은 하기에 제공된다:

[0015] >sp|Q6B0I6|KDM4D_인간 리신-특이적 데메틸라제 4D OS=호모 사피엔스 (Homo sapiens) OX=9606 GN=KDM4D PE=1 SV=3

METMKSKANCAQNPNCNIMIFHPTKEEFNDFDKYIAYMESQGAHRAGLAKIIPPKEWKAR
ETYDNISEILIATPLQQVASGRAGVFTQYHKKKKAMTVGEYRHLANSKKYQTPPHQNFED
LERKYWKNRIYNSPIYGADISGSLFDENTKQWNLGHLGTIQDLLEKECGVVIEGVNTPYL
YFGMWKTTFAWHTEDMDLYSINYLHLGEPKTWYVVPPEHGQRLERLARELFPGSSRGCGA
FLRHKVALISPTVLKENGIPFNRITQEAGEFMVTFPYGYHAGFNHGFNCAEAINFATPRW
IDYGKMASQCSCGEARVTFSMDAFVRILQPERYDLWKRGQDRAVVDHMEPRVPASQELST
QKEVQLPRRAALGLRQLPSHWARHSPWPMAARSGTRCHTLVCSSLPRRSAVSGTATQPRA
AAVHSSKKPSSTPSSTPGPSAQIIHPSNGRRGRGRPPQKLRAQELTLQTPAKRPLLAGTT
CTASGPEPEPLPEDGALMDKPVPLSPGLQHPVKASGCSWAPVP

[0016]

[0017] "KDM4D 폴리뉴클레오티드"란 KDM4D 폴리펩티드를 코딩하는 핵산 분자를 의미한다. 예시적인 KDM4D 핵산은 하기에 제공된다:

```
   1 aaggggcggg gccgaagcgg cccagggggc gggcgtttga aatcagtgcc ttagagtaga
  61 ccctaaacct cattttatac cttcaagaac caattactta atgtctcttc cgtcttttcc
 121 gtccccgacc ccctcccaga ctccttcatt ccggtactgc gtggacggaa agccccgggt
 181 agccgacacc acgtccccgg ctagcgggag agagcgtgga aaaggattac accaaactgt
 241 ttaaatccaa cgactcctgc ttccatcctt tctcctgagc tagaaccaac aaacctagag
 301 agttgggctt cggaaaaact agtgttttca tttaattgga tatgaagaaa gaacaaatat
 361 gtacggggca accacgatct ttacaaagaa cataagttcc aggaaagcag gaaccttgtc
 421 tctcttgttc actgggtgta tcctctgcat atagaacagt gcctggcaca taataggtgc
 481 tgaatttttgt tctaaacact gaggacattc tctgctacat ttgggtcgta cccccaggtc
 541 tgagtaattc aatagactta agaagacaga gcccagcagc aaccgaaaca taacagagtt
 601 gcaggatcag ctaacgtcaa tgcctgggca aagctgctgc ccagagtgga atctcactag
 661 tgaataaaca agcccaagaa agattatcat ctcatttgca aaaaaaaaag tacgctggta
```

[0018]

```
 721 gatcctgcta cctcatagat aacaccagtc aaatttttttt ttaaagtagc attttcctac
 781 attgtcaact atctagaaca tacctaaaaa ctaagagttt actgcttatt aaatggaaac
 841 tatgaagtct aaggccaact gtgcccagaa tccaaattgt aacataatga tatttcatcc
 901 aaccaaagaa gagtttaatg attttgataa atatattgct tacatggaat cccaaggtgc
 961 acacagagct ggcttggcta agataattcc acccaaagaa tggaaagcca gagagaccta
1021 tgataatatc agtgaaatct taatagccac tcccctccag caggtggcct ctgggcgggc
1081 aggggtgttt actcaatacc ataaaaaaaa gaaagccatg actgtggggg agtatcgcca
1141 tttggcaaac agtaaaaaat atcagactcc accacaccag aatttcgaag atttggagcg
1201 aaaatactgg aagaaccgca tctataattc accgatttat ggtgctgaca tcagtggctc
1261 cttgtttgat gaaaacacta aacaatggaa tcttgggcac ctgggaacaa ttcaggacct
1321 gctggaaaag gaatgtgggg ttgtcataga aggcgtcaat acaccctact tgtactttgg
1381 catgtggaaa accacgtttg cttggcatac agaggacatg gacctttaca gcatcaacta
1441 cctgcacctt ggggagccca aaacttggta tgtggtgccc ccagaacatg gccagcgcct
1501 ggaacgcctg gccagggagc tcttcccagg cagttcccgg ggttgtgggg ccttcctgcg
1561 gcacaaggtg gccctcatct cgcctacagt tctcaaggaa aatgggattc ccttcaatcg
1621 cataactcag gaggctggag agttcatggt gacctttccc tatggctacc atgctggctt
1681 caaccatggt ttcaactgcg cagaggccat caattttgcc actccgcgat ggattgatta
1741 tggcaaaatg gcctcccagt gtagctgtgg ggaggcaagg gtgacctttt ccatggatgc
1801 cttcgtgcgc atcctgcaac ctgaacgcta tgacctgtgg aaacgtgggc aagaccgggc
1861 agttgtggac cacatggagc ccagggtacc agccagccaa gagctgagca cccagaagga
1921 agtccagtta cccaggagag cagcgctggg cctgagacaa ctcccttccc actgggcccg
1981 gcattcccct tggcctatgg ctgcccgcag tgggacacgg tgccacaccc ttgtgtgctc
2041 ttcactccca cgccgatctg cagttagtgg cactgctacg cagccccggg ctgctgctgt
2101 ccacagctct aagaagccca gctcaactcc atcatccacc cctggtccat ctgcacagat
2161 tatccacccg tcaaatggca gacgtggtcg tggtcgccct cctcagaaac tgagagctca
2221 ggagctgacc ctccagactc cagccaagag gcccctcttg gcgggcacaa catgcacagc
2281 ttcgggccca gaacctgagc ccctacctga ggatggggct ttgatggaca agcctgtacc
2341 actgagccca gggctccagc atcctgtcaa ggcttctggg tgcagctggg cccctgtgcc
2401 ctaagtccac gggctgtctt tatatcccac tgccctgctg tgtgacagtt tgatgaaact
2461 ggttacattt acatcccaaa actttggttg agtttgcagg actctaggca tgcatgaaag
2521 agccccctg gtgatgccct tggatgctgc caagtccatg gtagttttca attttgccat
2581 acttttgttc ttcctaccgg accctggaat gtctttggat attgctaaaa tctatttctg
2641 cagctgaggt tttatccact ggacacattt gtgtgtgaga actaggtctt gttgaggtta
2701 gcgtaacctg gtatatgcaa ctaccatcct ctgggccaac tgtggaagct gctgcacttg
2761 tgaagaatcc tgagctttga ttcctcttca gtctacgcat ttctctcttc ccctccctca
2821 ccccctttt cttataaaac taggttcttt atacagataa ggtcagtaga gttccagaat
2881 aaaagatatg acttttctga gttatttatg tacttaaaat atgttgtcac agtatttgtt
2941 cccaaatata ttaaaggtaa ccaaaatgtt aaaaaaaaaa aaaaaaaa
```

[0019]

[0020] "EZH1 폴리펩티드" (히스톤-리신 N-메틸트랜스퍼라제 EZH1)란 NCBI 참조 서열: NP_001982에 제공된 서열, 또는 그의 단편과 적어도 약 85% 아미노산 동일성을 갖고, 메틸트랜스퍼라제 활성을 갖는 단백질을 의미한다. 예시적인 H3K27 메틸트랜스퍼라제 아미노산 서열은 하기에 제공된다:

```
   1 meipnpptsk citywkrkvk seymrlrqlk rlqanmgaka lyvanfakvq ektqilneew
  61 kklrvqpvqs mkpvsghpfl kkctiesifp gfasqhmlmr slntvalvpi myswsplqqn
 121 fmvedetvlc nipymgdevk eedetfieel innydgkvhg eeemipgsvl isdavflelv
 181 dalnqysdee eeghndtsdg kqddskedlp vtrkrkrhai egnkksskkq fpndmifsai
 241 asmfpengvp ddmkeryrel temsdpnalp pqctpnidgp naksvqreqs lhsfhtlfcr
 301 rcfkydcflh pfhatpnvyk rknkeikiep epcgtdcfll legakeyaml hnprskcsgr
 361 rrrrhhivsa scsnasasav aetkegdsdr dtgndwasss seansrcqtp tkqkaspapp
 421 qlcvveapse pvewtgaees lfrvfhgtyf nnfcsiarll gtktckqvfq favkeslilk
 481 lptdelmnps qkkkrkhrlw aahcrkiqlk kdnsstqvyn yqpcdhpdrp cdstcpcimt
 541 qnfcekfcqc npdcqnrfpg crcktqcntk qcpcylavre cdpdlcltcg asehwdckvv
 601 scknсsiqrg lkkhlllaps dvagwgtfik esvqknefis eycgelisqd eadrrgkvyd
 661 kymssflfnl nndfvvdatr kgnkirfanh svnpncyakv vmvngdhrig ifakraiqag
 721 eelffdyrys qadalkyvgi eretdvl
```

[0021]

[0022] "EZH1 폴리뉴클레오티드"란 EZH1 폴리펩티드를 코딩하는 핵산 분자를 의미한다. 예시적인 EZH1 폴리뉴클레오티드 서열은 NM_001991.4에 제공되고, 하기에 재현된다:

```
   1 aggaggcgcg gggcggggca cggcgcaggg gtggggccgc ggcgcgcatg cgtcctagca
  61 gcgggacccg cggctcggga tggaggctgg acacctgttc tgctgttgtg tcctgccatt
 121 ctcctgaaga acagaggcac actgtaaaac ccaacacttc cccttgcatt ctataagatt
 181 acagcaagat ggaaatacca aatcccccta cctccaaatg tatcacttac tggaaaagaa
 241 aagtgaaatc tgaatacatg cgacttcgac aacttaaacg gcttcaggca aatatgggtg
 301 caaaggcttt gtatgtggca aattttgcaa aggttcaaga aaaaacccag atcctcaatg
 361 aagaatggaa gaagcttcgt gtccaacctg ttcagtcaat gaagcctgtg agtggacacc
 421 cttttctcaa aaagtgtacc atagagagca ttttcccggg atttgcaagc caacatatgt
 481 taatgaggtc actgaacaca gttgcattgg ttcccatcat gtattcctgg tcccctctcc
 541 aacagaactt tatggtagaa gatgagacgg ttttgtgcaa tattccctac atgggagatg
 601 aagtgaaaga agaagatgag acttttattg aggagctgat caataactat gatgggaaag
 661 tccatggtga agaagagatg atccctggat ccgttctgat tagtgatgct gtttttctgg
 721 agttggtcga tgccctgaat cagtactcag atgaggagga ggaagggcac aatgacacct
 781 cagatggaaa gcaggatgac agcaaagaag atctgccagt aacaagaaag agaaagcgac
 841 atgctattga aggcaacaaa aagagttcca agaaacagtt cccaaatgac atgatcttca
 901 gtgcaattgc ctcaatgttc cctgagaatg gtgtcccaga tgacatgaag gagaggtatc
 961 gagaactaac agagatgtca gaccccaatg cacttccccc tcagtgcaca cccaacatcg
1021 atggccccaa tgccaagtct gtgcagcggg agcaatctct gcactccttc cacacacttt
1081 tttgccggcg ctgctttaaa tacgactgct tccttcaccc ttttcatgcc acccctaatg
1141 tatataaacg caagaataaa gaaatcaaga ttgaaccaga accatgtggc acagactgct
1201 tccttttgct ggaaggagca aaggagtatg ccatgctcca caacccccgc tccaagtgct
1261 ctggtcgtcg ccggagaagg caccacatag tcagtgcttc ctgctccaat gcctcagcct
1321 ctgctgtggc tgagactaaa gaaggagaca gtgacaggga cacaggcaat gactgggcct
1381 ccagttcttc agaggctaac tctcgctgtc agactcccac aaaacagaag gctagtccag
1441 ccccacctca actctgcgta gtggaagcac cctcggagcc tgtggaatgg actggggctg
1501 aagaatctct ttttcgagtc ttccatggca cctacttcaa caacttctgt tcaatagcca
1561 ggcttctggg gaccaagacg tgcaagcagg tctttcagtt tgcagtcaaa gaatcactta
```

[0023]

```
1621 tcctgaagct gccaacagat gagctcatga acccctcaca gaagaagaaa agaaagcaca
1681 gattgtgggc tgcacactgc aggaagattc agctgaagaa agataactct tccacacaag
1741 tgtacaacta ccaaccctgc gaccacccag accgcccctg tgacagcacc tgcccctgca
1801 tcatgactca gaatttctgt gagaagttct gccagtgcaa cccagactgt cagaatcgtt
1861 tccctggctg tcgctgtaag acccagtgca ataccaagca atgtccttgc tatctggcag
1921 tgcgagaatg tgaccctgac ctgtgtctca cctgtggggc ctcagagcac tgggactgca
1981 aggtggtttc ctgtaaaaac tgcagcatcc agcgtggact taagaagcac ctgctgctgg
2041 ccccctctga tgtggccgga tggggcacct tcataaagga gtctgtgcag aagaacgaat
2101 tcatttctga atactgtggt gagctcatct ctcaggatga ggctgatcga cgcggaaagg
2161 tctatgacaa atacatgtcc agcttcctct tcaacctcaa taatgatttt gtagtggatg
2221 ctactcggaa aggaaacaaa attcgatttg caaatcattc agtgaatccc aactgttatg
2281 ccaaagtggt catggtgaat ggagaccatc ggattgggat ctttgccaag agggcaattc
2341 aagctggcga agagctcttc tttgattaca ggtacagcca agctgatgct ctcaagtacg
2401 tggggatcga gagggagacc gacgtccttt agccctccca ggccccacgg cagcacttat
2461 ggtagcggca ctgtcttggc tttcgtgctc acaccactgc tgctcgagtc tcctgcactg
2521 tgtctcccac actgagaaac cccccaaccc actccctctg tagtgaggcc tctgccatgt
2581 ccagagggca caaaactgtc tcaatgagag gggagacaga ggcagctagg gcttggtctc
2641 ccaggacaga gagttacaga aatgggagac tgtttctctg gcctcagaag aagcgagcac
2701 aggctggggt ggatgactta tgcgtgattt cgtgtcggct ccccaggctg tggcctcagg
2761 aatcaactta ggcagttccc aacaagcgct agcctgtaat tgtagctttc cacatcaaga
2821 gtccttatgt tattgggatg caggcaaacc tctgtggtcc taagacctgg agaggacagg
2881 ctaagtgaag tgtggtccct ggagcctaca agtggtctgg gttagaggcg agcctggcag
2941 gcagcacaga ctgaactcag aggtagacag gtcaccttac tacctcctcc ctcgtggcag
3001 ggctcaaact gaaagagtgt gggttctaag tacaggcatt caaggctggg ggaaggaaag
3061 ctacgccatc cttccttagc cagagaggga gaaccagcca gatgatagta gttaaactgc
3121 taagcttggg cccaggaggc tttgagaaag ccttctctgt gtactctgga gatagatgga
3181 gaagtgtttt cagattcctg ggaacagaca ccagtgctcc agctcctcca aagttctggc
3241 ttagcagctg caggcaagca ttatgctgct attgaagaag cattaggggt atgcctggca
3301 ggtgtgagca tcctggctcg ctggatttgt gggtgttttc aggccttcca ttccccatag
3361 aggcaaggcc caatggccag tgttgcttat cgcttcaggg taggtgggca caggcttgga
3421 ctagagagga gaaagattgg tgtaatctgc tttcctgtct gtagtgcctg ctgtttggaa
3481 agggtgagtt agaatatgtt ccaaggttgg tgaggggcta aattgcacgc gtttaggctg
3541 gcaccccgtg tgcagggcac actggcagag ggtatctgaa gtgggagaag aagcaggtag
3601 accacctgtc ccaggctgtg gtgccaccct ctctggcatt catgcagagc aaagcacttt
3661 aaccatttct tttaaaaggt ctatagattg gggtagagtt tggcctaagg tctctagggt
3721 ccctgcctaa atcccactcc tgagggaggg ggaagaagag agggtgggag attctcctcc
3781 agtcctgtct catctcctgg gagaggcaga cgagtgagtt tcacacagaa gaatttcatg
3841 tgaatggggc cagcaagagc tgccctgtgt ccatggtggg tgtgccgggc tggctgggaa
3901 caaggagcag tatgttgagt agaaagggtg tgggcgggta tagattggcc tgggagtgtt
3961 acagtaggga gcaggcttct cccttctttc tgggactcag agccccgctt cttcccactc
4021 cacttgttgt cccatgaagg aagaagtggg gttcctcctg acccagctgc ctcttacggt
4081 ttggtatggg acatgcacac acactcacat gctctcactc accacactgg agggcacaca
4141 cgtaccccgc acccagcaac tcctgacaga aagctcctcc cacccaaatg ggccaggccc
```

[0024]

```
4201 cagcatgatc ctgaaatctg catccgccgt ggtttgtatt cattgtgcat atcagggata
4261 ccctcaagct ggactgtggg ttccaaatta ctcatagagg agaaaaccag agaaagatga
4321 agaggaggag ttaggtctat ttgaaatgcc aggggctcgc tgtgaggaat aggtgaaaaa
4381 aaacttttca ccagcctttg agagactaga ctgaccccac ccttccttca gtgagcagaa
4441 tcactgtggt cagtctcctg tcccagcttc agttcatgaa tactcctgtt cctccagttt
4501 cccatccttt gtccctgctg tcccccactt ttaaagatgg gtctcaaccc ctccccacca
4561 cgtcatgatg gatggggcaa ggtggtgggg actaggggag cctggtatac atgcggcttc
4621 attgccaata aatttcatgc actttaaagt cctgtggctt gtgacctctt aataaagtgt
4681 tagaatccaa aaaaaaa
```

[0025]

[0026] "EZH2 폴리펩티드" (히스톤-리신 N-메틸트랜스퍼라제 EZH2)란 유니프롯 (UniProt)KB/스위스-프롯 (Swiss-Prot): Q15910.2에 제공된 서열, 또는 그의 단편과 적어도 약 85% 아미노산 동일성을 갖고, 메틸트랜스퍼라제 활성을 갖는 단백질을 의미한다. 예시적인 H3K27 메틸트랜스퍼라제 아미노산 서열은 하기에 제공된다:

```
  1 mgqtgkksek gpvcwrkrvk seymrlrqlk rfrradevks mfssnrqkil erteilnqew
 61 kqrriqpvhi ltsvsslrgt recsvtsdld fptqviplkt lnavasvpim yswsplqqnf
121 mvedetvlhn ipymgdevld qdgtfieeli knydgkvhgd recgfindei fvelvnalgq
181 yndddddddg ddpeereekq kdledhrddk esrpprkfps dkifeaissm fpdkgtaeel
241 kekykelteq qlpgalppec tpnidgpnak svqreqslhs fhtlfcrrcf kydcflhpfh
301 atpntykrkn tetaldnkpc gpqcyqhleg akefaaalta eriktppkrp ggrrrgrlpn
361 nssrpstpti nvleskdtds dreagtetgg enndkeeeek kdetssssea nsrcqtpikm
421 kpnieppenv ewsgaeasmf rvligtyydn fcaiarligt ktcrqvyefr vkessiiapa
481 paedvdtppr kkkrkhrlwa ahcrkiqlkk dgssnhvyny qpcdhprqpc dsscpcviaq
541 nfcekfcqcs secqnrfpgc rckaqcntkq cpcylavrec dpdlcltcga adhwdsknvs
601 ckncsiqrgs kkhlllapsd vagwgifikd pvqknefise ycgeiisqde adrrgkvydk
661 ymcsflfnln ndfvvdatrk gnkirfanhs vnpncyakvm mvngdhrigi fakraiqtge
721 elffdyrysq adalkyvgie remeip
```

[0027]

[0028] "EZH2 폴리뉴클레오티드"란 EZH2 폴리펩티드를 코딩하는 핵산 분자를 의미한다. 예시적인 EZH2 폴리뉴클레오티드 서열은 NM_001203248.1에 제공되고, 하기에 제공된다:

```
  1 ggcggcgctt gattgggctg ggggggccaa ataaaagcga tggcgattgg gctgccgcgt
 61 ttggcgctcg gtccggtcgc gtccgacacc cggtgggact cagaaggcag tggagccccg
121 gcggcggcgg cggcggcgcg cgggggcgac gcgcgggaac aacgcgagtc ggcgcgcggg
181 acgaagaata atcatgggcc agactgggaa gaaatctgag aagggaccag tttgttggcg
241 gaagcgtgta aaatcagagt acatgcgact gagacagctc aagaggttca gacgagctga
301 tgaagtaaag agtatgttta gttccaatcg tcagaaaatt ttggaaagaa cggaaatctt
361 aaaccaagaa tggaaacagc gaaggataca gcctgtgcac atcctgactt cttgttcggt
421 gaccagtgac ttggattttc caacacaagt catcccatta aagactctga atgcagttgc
481 ttcagtaccc ataatgtatt cttggtctcc cctacagcag aattttatgg tggaagatga
541 aactgtttta cataacattc cttatatggg agatgaagtt ttagatcagg atggtacttt
601 cattgaagaa ctaataaaaa attatgatgg gaaagtacac ggggatagag aatgtgggtt
661 tataaatgat gaaatttttg tggagttggt gaatgccctt ggtcaatata atgatgatga
721 cgatgatgat gatggagacg atcctgaaga aagagaagaa aagcagaaag atctggagga
781 tcaccgagat gataaagaaa gccgcccacc tcggaaattt ccttctgata aaatttttga

841 agccatttcc tcaatgtttc cagataaggg cacagcagaa gaactaaagg aaaaatataa
901 agaactcacc gaacagcagc tcccaggcgc acttcctcct gaatgtaccc ccaacataga
961 tggaccaaat gctaaatctg ttcagagaga gcaaagctta cactcctttc atacgctttt
1021 ctgtaggcga tgttttaaat atgactgctt cctacatcct tttcatgcaa cacccaacac
1081 ttataagcgg aagaacacag aaacagctct agacaacaaa ccttgtggac cacagtgtta
1141 ccagcatttg gagggagcaa aggagtttgc tgctgctctc accgctgagc ggataaagac
1201 cccaccaaaa cgtccaggag gccgcagaag aggacggctt cccaataaca gtagcaggcc
1261 cagcaccccc accattaatg tgctggaatc aaaggataca gacagtgata gggaagcagg
1321 gactgaaacg gggggagaga acaatgataa agaagaagaa gagaagaaag atgaaacttc
1381 gagctcctct gaagcaaatt ctcggtgtca aacaccaata aagatgaagc caaatattga
1441 acctcctgag aatgtggagt ggagtggtgc tgaagcctca atgtttagag tcctcattgg
1501 cacttactat gacaatttct gtgccattgc taggttaatt gggaccaaaa catgtagaca
1561 ggtgtatgag tttagagtca aagaatctag catcatagct ccagctcccg ctgaggatgt
1621 ggatactcct ccaaggaaaa agaagaggaa acaccggttg tgggctgcac actgcagaaa
1681 gatacagctg aaaaaggacg gctcctctaa ccatgtttac aactatcaac cctgtgatca
1741 tccacggcag ccttgtgaca gttcgtgccc ttgtgtgata gcacaaaatt tttgtgaaaa
1801 gttttgtcaa tgtagttcag agtgtcaaaa ccgctttccg ggatgccgct gcaaagcaca
1861 gtgcaacacc aagcagtgcc cgtgctacct ggctgtccga gagtgtgacc ctgacctctg
1921 tcttacttgt ggagccgctg accattggga cagtaaaaat gtgtcctgca agaactgcag
1981 tattcagcgg ggctccaaaa agcatctatt gctggcacca tctgacgtgg caggctgggg
2041 gatttttatc aaagatcctg tgcagaaaaa tgaattcatc tcagaatact gtggagagat
2101 tatttctcaa gatgaagctg acagaagagg gaaagtgtat gataaataca tgtgcagctt
2161 tctgttcaac ttgaacaatg attttgtggt ggatgcaacc cgcaagggta acaaaattcg
2221 ttttgcaaat cattcggtaa atccaaactg ctatgcaaaa gttatgatgg ttaacggtga
2281 tcacaggata ggtatttttg ccaagagagc catccagact ggcgaagagc tgtttttga
2341 ttacagatac agccaggctg atgccctgaa gtatgtcggc atcgaaagag aaatggaaat
2401 cccttgacat ctgctacctc ctccccccctc ctctgaaaca gctgccttag cttcaggaac
2461 ctcgagtact gtgggcaatt tagaaaaaga acatgcagtt tgaaattctg aatttgcaaa
2521 gtactgtaag aataatttat agtaatgagt ttaaaaatca acttttttatt gccttctcac
2581 cagctgcaaa gtgttttgta ccagtgaatt tttgcaataa tgcagtatgg tacatttttc
2641 aactttgaat aaagaatact tgaacttgtc cttgttgaat c
```

[0029]

[0030]

[0031] "KDM6A 폴리펩티드" (리신-특이적 데메틸라제 6A, 또한 히스톤 데메틸라제 UTX로 지칭됨)란 NCBI 참조 서열: 015550.2에 제공된 서열, 또는 그의 단편과 적어도 약 85% 아미노산 동일성을 갖고, 데메틸라제 활성을 갖는 단백질을 의미한다. 예시적인 KDM6A 아미노산 서열은 하기에 제공된다:

[0032]
```
   1 mkscgvslat aaaaaaafgd eekkmaagka sgeseeasps ltaeerealg gldsrlfgfv
  61 rfhedgartk allgkavrcy eslilkaegk vesdffcqlg hfnllledyp kalsayqryy
 121 slqsdywkna aflyglglvy fhynafqwai kafqevlyvd psfcrakeih lrlglmfkvn
 181 tdyesslkhf qlalvdcnpc tlsnaeiqfh iahlyetqrk yhsakeayeq llqtenlsaq
 241 vkatvlqqlg wmhhtvdllg dkatkesyai qylqkslead pnsgqswyfl grcyssigkv

 301 qdafisyrqs idkseasadt wcsigvlyqq qnqpmdalqa yicavqldhg haaawmdlgt
 361 lyescnqpqd aikcylnatr skscsntsal aarikylqaq lcnlpqgslq nktkllpsie
 421 eawslpipae ltsrqgamnt aqqntsdnws gghavshppv qqqahswclt pqklqhleql
 481 ranrnnlnpa qklmleqles qfvlmqqhqm rptgvaqvrs tgipngptad sslptnsvsg
 541 qqpqlaltrv psvsqpgvrp acpgqplang pfsaghvpcs tsrtlgstdt iligpnnhitg
 601 sgsngnvpyl qrnaltlphn rtnltssaee pwknqlsnst qglhkgqssh sagpngerpl
 661 sstgpsqhlq aagsgiqnqn ghptlpsnsv tqgaalnhls shtatsggqq gitltkeskp
 721 sgniltvpet srhtgetpns tasveglpnh vhqmtadavc spshgdsksp gllssdnpql
 781 sallmgkann nvgtgtcdkv nnihpavhtk tdnsvassps saistatpsp ksteqtttns
 841 vtslnsphsg lhtingegme esqspmktdl llvnhkpspq iipsmsvsiy pssaevlkac
 901 rnlgknglsn ssilldkcpp prppsspypp lpkdklnppt psiylenkrd affpplhqfc
 961 tnpnnpvtvi rglagalkld lglfstktlv eannehmvev rtqllqpade nwdptgtkki
1021 whcesnrsht tiakyaqyqa ssfqeslree nekrshhkdh sdsestssdn sgrrrkgpfk
1081 tikfgtnidl sddkkwklql heltklpafv rvvsagnlls hvghtilgmn tvqlymkvpg
1141 srtpghqenn nfcsvninig pgdcewfvvp egywgvlndf ceknnlnflm gswwpnledl
1201 yeanvpvyrf iqrpgdlvwi nagtvhwvqa igwcnniawn vgpltacqyk laveryewnk
1261 lqsvksivpm vhlswnmarn ikvsdpklfe mikycllrtl kqcqtlreal iaagkeiiwh
1321 grtkeepahy csicevevfd llfvtnesns rktyivhcqd carktsgnle nfvvleqykm
1381 edlmqvydqf tlapplpsas s
```

[0033]

[0034] "KDM6A 폴리뉴클레오티드"란 KDM6A 폴리펩티드를 코딩하는 핵산 분자를 의미한다. 예시적인 KDM6A 폴리뉴클레오티드 서열은 NM_001291415.1에 제공된다.

[0035] "KDM6B 폴리펩티드" (리신-특이적 데메틸라제 6, 또한 JmjC 도메인-함유 단백질 3으로 지칭됨)란 NCBI 참조 서열: 015054.4에 제공된 서열, 또는 그의 단편과 적어도 약 85% 아미노산 동일성을 갖고, 데메틸라제 활성을 갖는 단백질을 의미한다. 예시적인 KDM6B 아미노산 서열은 하기에 제공된다:

[0036]
```
   1 mhravdppga raareafalg glscagawss cpphppprsa wlpggrcsas igqpplpapl
  61 ppshgsssgh pskpyyapga ptprplhgkl eslhgcvqal lrepaqpglw eqlgqlyese
 121 hdseeatrcy hsalryggsf aelgprigrl qqaqlwnfht gscqhrakvl ppleqvwnll
 181 hlehkrnyga krggppvkra aeppvvqpvp paalsgpsge eglspggkrr rgcnseqtgl
 241 ppglplpppp lppppppppp pppplpglat sppfqltkpg lwstlhgdaw gperkgsapp
 301 erqeqrhslp hpypypapay tahppghrlv paappgpgpr ppgaeshgcl patrppgsdl
 361 resrvqrsrm dssvspaatt acvpyapsrp pglpgtttss sssssntgl rgvepnpgip
 421 gadhyqtpal evshhgrlgp sahssrkpfl gapaatphls lppgpssppp ppcprllrpp
 481 pppawlkgpa craaredgei leelffgteg pprpappplp hregflgppa srfsvgtqds
 541 htpptpptpt tsssnsnsgs hssspagpvs fppppylars idplprppsp aqnpqdpplv
```

```
 601 pltlalppap psschqntsg sfrrpesprp rvsfpktpev gpgpppgpls kapqpvppgv
 661 gelpargprl fdfpptpled qfeepaefki lpdglanimk mldesirkee eqqqheagva
 721 pqpplkepfa slqspfptdt aptttapava vttttttttt ttatqeeekk pppalppppp
 781 lakfpppsqp qpppppppsp asllkslasv legqkycyrg tgaavstrpg plpttqyspg
 841 ppsgatalpp tsaapsaqgs pqpsassssq fstsggpwar errageepvp gpmtptqppp
 901 plslpparse sevleeisra cetlvervgr satdpadpvd taepadsgte rllppaqake
 961 eaggvaavsg sckrrqkehq kehrrhrrac kdsvgrrpre grakakakvp keksrrvlgn
1021 ldlqseeiqg reksrpdlgg askakpptap appsapapsa qptppsasvp gkkareeapg
1081 ppgvsradml klrslsegpp kelkirlikv esgdketfia seveerrlrm adltishcaa
1141 dvvrasrnak vkgkfresyl spaqsvkpki nteeklprek lnpptpsiyl eskrdafspv
1201 llqfctdprn pitvirglag slrlnlglfs tktlveasge htvevrtqvq qpsdenwdlt
1261 gtrqiwpces srshttiaky aqyqassfqe slqeekesed eeseepdstt gtppssapdp
1321 knhhiikfgt nidlsdakrw kpqlqellkl pafmrvtstg nmlshvghti lgmntvqlym
1381 kvpgsrtpgh qennnfcsvn inigpgdcew favhehywet isafcdrhgv dyltgswwpi
1441 lddlyasnip vyrfvqrpgd lvwinagtvh wvqatgwcnn iawnvgplta yqyqlalery
1501 ewnevknvks ivpmihvswn vartvkisdp dlfkmikfcl lqsmkhcqvq reslvragkk
1561 iayqgrvkde payycnecdv evfnilfvts engsrntylv hcegcarrrs aglqgvvvle
1621 qyrteelaqa ydaftlapas tsr
```

[0037]

[0038] "KDM6B 폴리뉴클레오티드"란 KDM6B 폴리펩티드를 코딩하는 핵산 분자를 의미한다. 예시적인 KDM6B 폴리뉴클레오티드 서열은 NM_001080424.2에 제공되고, 하기에 재현된다:

```
   1 ggcaacatgc cagccccgta gcactgccca ccccacccac tgtggtctgt tgtaccccac
  61 tgctggggtg gtggttccaa tgagacaggg cacaccaaac tccatctggc tgttactgag
 121 gcggagacac gggtgatgat tggctttctg gggagagagg aagtcctgtg attggccaga
 181 tctctggagc ttgccgacgc ggtgtgagga cgctcccacg gaggccggaa ttggctgtga
 241 aaggactgag gcagccatct gggggtagcg ggcactctta tcagagcggc tggagccgga
 301 ccatcgtccc agagagctgg ggcagggggc cgtgcccaat ctccagggct cctggggcca
 361 ctgctgacct ggctggatgc atcgggcagt ggaccctcca ggggcccgcg ctgcacggga
 421 agcctttgcc cttgggggcc tgagctgtgc tggggcctgg agctcctgcc cgcctcatcc
 481 ccctcctcgt agcgcatggc tgcctggagg cagatgctca gccagcattg ggcagccccc
 541 gcttcctgct cccctacccc cttcacatgg cagtagttct gggcacccca gcaaaccata
 601 ttatgctcca ggggcgccca ctccaagacc cctccatggg aagctggaat ccctgcatgg
 661 ctgtgtgcag gcattgctcc gggagccagc ccagccaggg ctttgggaac agcttgggca
 721 actgtacgag tcagagcacg atagtgagga ggccacacgc tgctaccaca gcgcccttcg
 781 atacggagga agcttcgctg agctggggcc ccgcattggc cgactgcagc aggcccagct
 841 ctggaacttt catactggct cctgccagca ccgagccaag gtcctgcccc cactggagca
 901 agtgtggaac ttgctacacc ttgagcacaa acggaactat ggagccaagc ggggaggtcc
 961 cccggtgaag cgagctgctg aacccccagt ggtgcagcct gtgcctcctg cagcactctc
1021 aggcccctca ggggaggagg gcctcagccc tggaggcaag cgaaggagag gctgcaactc
1081 tgaacagact ggccttcccc cagggctgcc actgcctcca ccaccattac caccaccacc
```

[0039]

```
1141 accaccacca ccaccaccac caccacccct gcctggcctg gctaccagcc ccccatttca
1201 gctaaccaag ccagggctgt ggagtaccct gcatggagat gcctggggcc cagagcgcaa
1261 gggttcagca cccccagagc gccaggagca gcggcactcg ctgcctcacc catatccata
1321 cccagctcca gcgtacaccg cgcacccccc tggccaccgg ctggtcccgg ctgctccccc
1381 aggcccaggc ccccgccccc caggagcaga gagccatggc tgcctgcctg ccacccgtcc
1441 ccccggaagt gaccttagag agagcagagt tcagaggtcg cggatggact ccagcgtttc
1501 accagcagca accaccgcct gcgtgcctta cgccccttcc cggccccctg gcctccccgg
1561 caccaccacc agcagcagca gtagcagcag cagcaacact ggtctccggg gcgtggagcc
1621 gaacccaggc attcccggcg ctgaccatta ccaaactccc gcgctggagg tctctcacca
1681 tggccgcctg gggccctcgg cacacagcag tcggaaaccg ttcttggggg ctcccgctgc
1741 cactccccac ctatccctgc cacctggacc ttcctcaccc cctccacccc cctgtccccg
1801 cctcttacgc cccccaccac cccctgcctg gttgaagggt ccggcctgcc gggcagcccg
1861 agaggatgga gagatcttag aagagctctt ctttgggact gagggacccc cccgccctgc
1921 cccaccaccc ctcccccatc gcgagggctt cttggggcct ccggcctccc gcttttctgt
1981 gggcactcag gattctcaca cccctcccac tcccccaacc ccaaccacca gcagtagcaa
2041 cagcaacagt ggcagccaca gcagcagccc tgctgggcct gtgtcctttc ccccaccacc
2101 ctatctggcc agaagtatag acccccttcc ccggcctccc agcccagcac agaaccccca
2161 ggacccacct cttgtacccc tgactcttgc cctgcctcca gcccctcctt cctcctgcca
2221 ccaaaatacc tcaggaagct tcaggcgccc ggagagcccc cggcccaggg tctccttccc
2281 aaagaccccc gaggtggggc cggggccacc cccaggcccc ctgagtaaag ccccccagcc
2341 tgtgccgccc ggggttgggg agctgcctgc ccgaggccct cgactctttg attttccccc
2401 cactccgctg gaggaccagt ttgaggagcc agccgaattc aagatcctac ctgatgggct
2461 ggccaacatc atgaagatgc tggacgaatc cattcgcaag gaagaggaac agcaacaaca
2521 cgaagcaggc gtggcccccc aacccccgct gaaggagccc tttgcatctc tgcagtctcc
2581 tttccccacc gacacagccc ccaccactac tgctcctgct gtcgccgtca ccaccaccac
2641 caccaccacc accaccacca cggccaccca ggaagaggag aagaagccac caccagccct
2701 accaccacca ccgcctctag ccaagttccc tccaccctct cagccacagc caccaccacc
2761 cccacccccc agcccggcca gcctgctcaa atccttggcc tccgtgctgg agggacaaaa
2821 gtactgttat cgggggactg gagcagctgt ttccacccgg cctgggccct tgcccaccac
2881 tcagtattcc cctggccccc catcaggtgc taccgccctg ccgcccacct cagcggcccc
2941 tagcgcccag ggctccccac agccctctgc ttcctcgtca tctcagttct ctacctcagg
3001 cgggccctgg gcccgggagc gcagggcggg cgaagagcca gtcccgggcc ccatgacccc
3061 cacccaaccg cccccacccc tatctctgcc ccctgctcgc tctgagtctg aggtgctaga
3121 agagatcagc cgggcttgcg agacccttgt ggagcgggtg ggccggagtg ccactgaccc
3181 agccgaccca gtggacacag cagagccagc ggacagtggg actgagcgac tgctgccccc
3241 cgcacaggcc aaggaggagg ctggcggggt ggcggcagtg tcaggcagct gtaagcggcg
3301 acagaaggag catcagaagg agcatcggcg gcacaggcgg gcctgtaagg acagtgtggg
3361 tcgtcggccc cgtgagggca gggcaaaggc caaggccaag gtccccaaag aaaagagccg
3421 ccgggtgctg gggaacctgg acctgcagag cgaggagatc cagggtcgtg agaagtcccg
3481 gcccgatctt ggcggggcct ccaaggccaa gccacccaca gctccagccc ctccatcagc
3541 tcctgcacct tctgcccagc ccacaccccc gtcagcctct gtccctggaa agaaggctcg
3601 ggaggaagcc ccagggccac cgggtgtcag ccgggccgac atgctgaagc tgcgctcact
3661 tagtgagggg ccccccaagg agctgaagat ccggctcatc aaggtagaga gtggtgacaa
```

[0040]

3721 ggagaccttt atcgcctctg aggtggaaga gcggcggctg cgcatggcag acctcaccat
3781 cagccactgt gctgctgacg tcgtgcgcgc cagcaggaat gccaaggtga aagggaagtt
3841 tcgagagtcc tacctttccc ctgcccagtc tgtgaaaccg aagatcaaca ctgaggagaa
3901 gctgccccgg gaaaaactca acccccctac acccagcatc tatctggaga gcaaacggga
3961 tgccttctca cctgtcctgc tgcagttctg tacagaccct cgaaatccca tcacagtgat
4021 ccggggcctg gcgggctccc tgcggctcaa cttgggcctc ttctccacca agaccctggt
4081 ggaagcgagt ggcgaacaca ccgtggaagt tcgcacccag gtgcagcagc cctcagatga
4141 gaactgggat ctgacaggca ctcggcagat ctggccttgt gagagctccc gttcccacac
4201 caccattgcc aagtacgcac agtaccaggc ctcatccttc caggagtctc tgcaggagga
4261 gaaggagagt gaggatgagg agtcagagga gccagacagc accactggaa cccctcctag
4321 cagcgcacca gacccgaaga accatcacat catcaagttt ggcaccaaca tcgacttgtc
4381 tgatgctaag cggtggaagc cccagctgca ggagctgctg aagctgcccg ccttcatgcg
4441 ggtaacatcc acgggcaaca tgctgagcca cgtgggccac accatcctgg gcatgaacac
4501 ggtgcagctg tacatgaagg tgcccggcag ccgaacgcca ggccaccagg agaataacaa
4561 cttctgctcc gtcaacatca acattggccc aggcgactgc gagtggttcg cggtgcacga
4621 gcactactgg gagaccatca gcgctttctg tgatcggcac ggcgtggact acttgacggg
4681 ttcctggtgg ccaatcctgg atgatctcta tgcatccaat attcctgtgt accgcttcgt
4741 gcagcgaccc ggagacctcg tgtggattaa tgcggggact gtgcactggg tgcaggccac
4801 cggctggtgc aacaacattg cctggaacgt ggggcccctc accgcctatc agtaccagct
4861 ggccctggaa cgatacgagt ggaatgaggt gaagaacgtc aaatccatcg tgcccatgat
4921 tcacgtgtca tggaacgtgg ctcgcacggt caaaatcagc gaccccgact tgttcaagat
4981 gatcaagttc tgcctgctgc agtccatgaa gcactgccag gtgcaacgcg agagcctggt
5041 gcgggcaggg aagaaaatcg cttaccaggg ccgtgtcaag gacgagccag cctactactg
5101 caacgagtgc gatgtggagg tgtttaacat cctgttcgtg acaagtgaga atggcagccg
5161 caacacgtac ctggtacact gcgagggctg tgcccggcgc cgcagcgcag gcctgcaggg
5221 cgtggtggtg ctggagcagt accgcactga ggagctggct caggcctacg acgccttcac
5281 gctggtgagg gcccggcggg cgcgcgggca gcggaggagg gcactggggc aggctgcagg
5341 gacgggcttc gggagcccgg ccgcgccttt ccctgagccc ccgccggctt tctcccccca
5401 ggccccagcc agcacgtcgc gatgaggccg gacgccccgc ccgcctgcct gcccgcgcaa
5461 ggcgccgcgg ggccaccagc acatgcctgg gctggaccta ggtcccgcct gtggccgaga
5521 agggggtcgg gcccagccct tccaccccat tggcagctcc cctcacttaa tttattaaga
5581 aaaacttttt ttttttttt agcaaatatg aggaaaaaag gaaaaaaaat gggagacggg
5641 ggaggggggct ggcagccccct cgcccaccag cgcctcccct caccgacttt ggccttttta
5701 gcaacagaca caaggaccag gctccggcgg cggcgggggt cacatacggg ttccctcacc
5761 ctgccagccg cccgcccgcc cggcgcagat gcacgcggct cgtgtatgta catagacgtt
5821 acggcagccg aggtttttaa tgagattctt tctatgggct ttacccctcc cccggaacct
5881 cctttttac ttccaatgct agctgtgacc cctgtacatg tctctttatt cacttggtta
5941 tgatttgtat tttttgttct tttcttgttt ttttgttttt aatttataac agtcccactc
6001 acctctattt attcattttt gggaaaaccc gacctcccac acccccaagc catcctgccc
6061 gcccctccag ggaccgcccg tcgccgggct ctccccgcgc cccagtgtgt gtccgggccc
6121 ggcccgaccg tctccacccg tccgcccgcg gctccagccg ggttctcatg gtgctcaaac
6181 ccgctccccc cccctacgtc ctgcactttc tcggaccagt ccccccactc ccgacccgac
6241 cccagcccca cctgagggtg agcaactcct gtactgtagg ggaagaagtg ggaactgaaa

[0041]

6301 tggtattttg taaaaaaaat aaataaaata aaaaaattaa aggttttaaa gaaagaacta
6361 tgaggaaaag gaaccccgtc cttcccagcc ccggccaact ttaaaaaaca cagaccttca
6421 cccccacccc cttttctttt taagtgtgaa acaacccagg gccagggcct cactggggca
6481 gggacacccc ggggtgagtt tctctggggc tttatttttcg ttttgttggt tgttttttct
6541 ccacgctggg gctgcggagg ggtgggggggt ttacagtccc gcaccctcgc actgcactgt
6601 ctctctgccc caggggcaga ggggtcttcc caaccctacc cctattttcg gtgatttttg
6661 tgtgagaata ttaatattaa aaataaacgg agaaaaaaaa aaaaaaaaaa aaaaaaaaaa
6721 aaaaaaaaaa a

[0042]

[0043] "KDM6C 폴리펩티드" (히스톤 데메틸라제 UTY, 또한 Y 염색체 상의 편재적으로-전사된 TPR 단백질로 지칭됨)란 NCBI 참조 서열: O14607.2에 제공된 서열, 또는 그의 단편과 적어도 약 85% 아미노산 동일성을 갖고, 데메틸라제 활성을 갖는 단백질을 의미한다. 예시적인 KDM6C 아미노산 서열은 하기에 제공된다:

```
   1 mkscavsltt aavafgdeak kmaegkasre seeesvsltv eerealggmd srlfgfvrlh
  61 edgartktll gkavrcyesl ilkaegkves dffcqlghfn llledyskal sayqryyslq
 121 adywknaafl yglglvyfyy nafhwaikaf qdvlyvdpsf crakeihlrl glmfkvntdy
 181 ksslkhfqla lidcnpctls naeiqfhiah lyetqrkyhs akeayeqllq tenlpaqvka
 241 tvlqqlgwmh hnmdlvgdka tkesyaiqyl qksleadpns gqswyflgrc yssigkvqda
 301 fisyrqsidk seasadtwcs igvlyqqqnq pmdalqayic avqldhghaa awmdlgtlye
 361 scnqpqdaik cylnaarskr csntstlaar ikflqngsdn wnggqslshh pvqqvyslcl
 421 tpqklqhleq lranrdnlnp aqkhqleqle sqfvlmqqmr hkevaqvrtt gihngaitds
 481 slptnsvsnr qphgaltrvs svsqpgvrpa cvekllssga fsagcipcgt skilgstdti
 541 llgsnciags esngnvpylq qnthtlphnh tdlnssteep wrkqlsnsaq glhksqsscl
 601 sgpneeqplf stgsaqyhqa tstgikkane hltlpsnsvp qgdadshlsc htatsggqqg
 661 imftkeskps knrslvpets rhtgdtsngc advkglsnhv hqliadavss pnhgdspnll
 721 iadnpqlsal ligkangnvg tgtcdkvnni hpavhtktdh svasspssai statpspkst
 781 eqrsinsvts lnsphsglht vngeglgksq sstkvdlpla shrstsqilp smsvsicpss
 841 tevlkacrnp gknglsnsci lldkcppprp ptspypplpk dklnpptpsi ylenkrdaff
 901 pplhqfctnp knpvtvirgl agalkldlgl fstktlvean nehmvevrtq llqpadenwd
 961 ptgtkkiwrc esnrshttia kyaqyqassf qeslreenek rtqhkdhsdn estssensgr
1021 rrkgpfktik fgtnidlsdn kkwklqlhel tklpafarvv sagnllthvg htilgmntvq
1081 lymkvpgsrt pghqennnfc svninigpgd cewfvvpedy wgvlndfcek nnlnflmssw
1141 wpnledlyea nvpvyrfiqr pgdlvwinag tvhwvqavgw cnniawnvgp ltacqyklav
1201 eryewnklks vkspvpmvhl swnmarnikv sdpklfemik ycllkilkqy qtlrealvaa
1261 gkeviwhgrt ndepahycsi cevevfnllf vtnesntqkt yivhchdcar ktskslenfv
1321 vleqykmedl iqvydqftla lslssss
```

[0044]

[0045] "KDM6C 폴리뉴클레오티드란 KDM6C 폴리펩티드를 코딩하는 핵산 분자를 의미한다. 예시적인 KDM6A 폴리뉴클레오티드 서열은 NM_001258249.1에 제공되고, 이 서열은 하기에 재현된다:

[0046]

```
   1 gctcatcgtt tgttgtttag ataatatcat gaactgataa atgcagttgc cacgttgatt
  61 ccctagggcc tggcttaccg actgaggtca taagatatta tgccttctct ttagacttgg
 121 tcagtggaga ggaaatgggc aaagaaccag cctatggagg tgacaaggcc ttagggccaa
 181 aagtcttgag ggtgaaggtt tagggcctgc gcagcttccc tgccatgccc cgcaaggtct
 241 cgcattcgca aggcttgtga cagtgggagc ctcattacgg actctcctaa agtccatggt
 301 gtcctctttt cgcatttgcg ccccgtgggt gatgcccgat gccgcccttc ccatcgctct
 361 cttccccttc aagcgtatcg caactgcaaa aacacccagc acagacactc cattttctat
 421 cttaatgcat ttaactagca caacctacag gttgttccat cccagagact acccttttct
 481 ccatagacgt gaccatcaac caaccagcgg tcagaatcag tcagcctctg tcatgttcct
 541 aggtccttgg cgaactggct gggcggggtc ccagcagcct aggagtacag tggagcaatg
 601 cctgacgtaa gtcaacaaag atcacgtgag acgaatcagt cgcctagatt ggctacaact
 661 aagtggttgg gagcggggag gtcgcggcgg ctgcgtgggg ttcgcccgtg acacaattac
 721 aactttgtgc tggtgctggc aaagtttgtg attttaagaa attctgctgt gctctccagc
 781 actgcgagct tctgccttcc ctgtagtttc ccagatgtga tccaggtagc cgagttccgc
 841 tgcccgtgct tcggtagctt aagtctttgc ctcagctttt ttccttgcag ccgctgagga
 901 ggcgataaaa ttggcgtcac agtctcaagc agcgattgaa ggcgtctttt caactactcg
 961 attaaggttg ggtatcgtcg tgggacttgg aaatttgttg tttccatgaa atcctgcgca
1021 gtgtcgctca ctaccgccgc tgttgccttc ggtgatgagg caaagaaaat ggcggaagga
1081 aaagcgagcc gcgagagtga agaggagtct gttagcctga cagtcgagga aagggaggcg
1141 cttggtggca tggacagccg tctcttcggg ttcgtgaggc ttcatgaaga tggcgccaga
1201 acgaagaccc tactaggcaa ggctgttcgc tgctacgaat ctttaatctt aaaagctgaa
1261 ggaaaagtgg agtctgactt cttttgccaa ttaggtcact tcaacctctt gttggaagat
1321 tattcaaaag cattatctgc atatcagaga tattacagtt tacaggctga ctactggaag
1381 aatgctgcgt ttttatatgg ccttggtttg gtctacttct actacaatgc atttcattgg
1441 gcaattaaag catttcaaga tgtcctttat gttgacccca gcttttgtcg agccaaggaa
1501 attcatttac gacttgggct catgttcaaa gtgaacacag actacaagtc tagtttaaag
1561 cattttcagt tagccttgat tgactgtaat ccatgtactt tgtccaatgc tgaaattcaa
1621 tttcatattg cccatttgta tgaaacccag aggaagtatc attctgcaaa ggaggcatat
1681 gaacaacttt tgcagacaga aaaccttcct gcacaagtaa aagcaactgt attgcaacag
1741 ttaggttgga tgcatcataa tatggatcta gtaggagaca aagccacaaa ggaaagctat
1801 gctattcagt atctccaaaa gtctttggag gcagatccta attctggcca atcgtggtat
1861 tttcttggaa ggtgttattc aagtattggg aaagttcagg atgcctttat atcttacagg
1921 caatctattg ataaatcaga agcaagtgca gatacatggt gttcaatagg tgtgttgtat
1981 cagcagcaaa atcagcctat ggatgcttta caggcatata tttgtgctgt acaattggac
2041 catgggcatg ccgcagcctg gatggaccta ggtactctct atgaatcctg caatcaacct
2101 caagatgcca ttaaatgcta cctaaatgca gctagaagca aacgttgtag taatacctct
2161 acgcttgctg caagaattaa atttctacag gctcagttgt gtaaccttcc acaaagtagt
2221 ctacagaata aaactaaatt acttcctagt attgaggagg catggagcct accaatcccc
2281 gcagagctta cctccaggca gggtgccatg aacacagcac agcaggctta tagagctcat
```

2341 gatccaaata ctgaacatgt attaaaccac agtcaaacac caattttaca gcaatccttg
2401 tcactacaca tgattacttc tagccaagta gaaggcctgt ccagtcctgc caagaagaaa
2461 agaacatcta gtccaacaaa gaatggttct gataactgga atggtggcca gagtctttca
2521 catcatccag tacagcaagt ttattcgttg tgtttgacac cacagaaatt acagcacttg
2581 gaacaactgc gagcaaatag agataattta aatccagcac agaagcatca gctggaacag
2641 ttagaaagtc agtttgtctt aatgcagcaa atgagacaca aagaagttgc tcaggtacga
2701 actactggaa ttcataacgg ggccataact gattcatcac tgcctacaaa ctctgtctct
2761 aatcgacaac cacatggtgc tctgaccaga gtatctagcg tctctcagcc tggagttcgc
2821 cctgcttgtg ttgaaaaact tttgtccagt ggagcttttt ctgcaggctg tattccttgt
2881 ggcacatcaa aaattctagg aagtacagac actatcttgc taggcagtaa ttgtatagca
2941 ggaagtgaaa gtaatggaaa tgtgccttac ctgcagcaaa atacacacac tctacctcat
3001 aatcatacag acctgaacag cagcacagaa gagccatgga gaaaacagct atctaactcc
3061 gctcaggggc ttcataaaag tcagagttca tgtttgtcag gacctaatga agaacaacct
3121 ctgttttcca ctgggtcagc ccagtatcac caggcaacta gcactggtat taagaaggcg
3181 aatgaacatc tcactctgcc tagtaattca gtaccacagg gggatgctga cagtcacctc
3241 tcctgtcata ctgctacctc aggtggacaa caaggcatta tgtttaccaa agagagcaag
3301 ccttcaaaaa atagatcctt ggtgcctgaa acaagcaggc atactggaga cacatctaat
3361 ggctgtgctg atgtcaaggg actttctaat catgttcatc agttgatagc agatgctgtt
3421 tccagtccta accatggaga ttcaccaaat ttattaattg cagacaatcc tcagctctct
3481 gctttgttga ttggaaaagc caatggcaat gtgggtactg gaacctgtga caaagtgaat
3541 aatattcacc cagctgttca tacaaagact gatcattctg ttgcctcttc accctcttca
3601 gccatttcca cagcaacacc ttctcctaaa tccactgagc agagaagcat aaacagtgtt
3661 accagcctta acagtcctca cagtggatta cacacagtca atggagaggg gctggggaag
3721 tcacagagct ctacaaaagt agacctgcct ttagctagcc acagatctac ttctcagatc
3781 ttaccatcaa tgtcagtgtc tatatgcccc agttcaacag aagttctgaa agcatgcagg
3841 aatccaggta aaaatggctt gtctaatagc tgcattttgt tagataaatg tccacctcca
3901 agaccaccaa cttcaccata cccacccttg ccaaaggaca agttgaatcc acccacacct
3961 agtatttact tggaaaataa acgtgatgct ttctttcctc cattacatca attttgtaca
4021 aatccaaaaa accctgttac agtaatacgt ggccttgctg gagctcttaa attagatctt
4081 ggacttttct ctaccaaaac tttggtagaa gctaacaatg aacatatggt agaagtgagg
4141 acacagttgc tgcaaccagc agatgaaaac tgggatccca ctggaacaaa gaaaatctgg
4201 cgttgtgaaa gcaatagatc tcatactaca attgccaaat acgcacaata ccaggcttcc
4261 tccttccagg aatcattgag agaagaaaat gagaaaagaa cacaacacaa agatcattca
4321 gataacgaat ccacatcttc agagaattct ggaaggagaa ggaaaggacc ttttaaaacc
4381 ataaaatttg ggaccaacat tgacctctct gataacaaaa agtggaagtt gcagttacat
4441 gaactgacta aacttcctgc ttttgcgcgt gtggtgtcag caggaaatct tctaacccat
4501 gttgggcata ccattctggg catgaataca gtacaactgt atatgaaagt tccagggagt
4561 cggacaccag gtcaccaaga aaataacaac ttctgctctg ttaacataaa tattggtcca
4621 ggagattgtg aatggtttgt tgtacctgaa gattattggg gtgttctgaa tgacttctgt
4681 gaaaaaaata atttgaattt tttaatgagt tcttggtggc ccaaccttga agatctttat
4741 gaagcaaatg tccctgtgta tagatttatt cagcgacctg gagatttggt ctggataaat
4801 gcaggcactg tgcattgggt tcaagctgtt ggctggtgca ataacattgc ctggaatgtt
4861 ggtccactta cagcctgcca gtataaattg gcagtggaac ggtatgaatg gaacaaattg

[0047]

```
4921 aaaagtgtga agtcaccagt acccatggtg catctttcct ggaatatggc acgaaatatc
4981 aaagtctcag atccaaagct ttttgaaatg attaagtatt gtcttttgaa aattctgaag
5041 caatatcaga cattgagaga agctcttgtt gcagcaggaa aagaggttat atggcatggg
5101 cggacaaatg atgaaccagc tcattactgt agcatttgtg aggtggaggt ttttaatctg
5161 ctttttgtca ctaatgaaag caatactcaa aaaacctaca tagtacattg ccatgattgt
5221 gcacgaaaaa caagcaaaag tttggaaaat tttgtggtgc tcgaacagta caaaatggag
5281 gacctaatcc aagtttatga tcaatttaca ctagctcttt cattatcatc ctcatcttga
5341 tatagttcca tgaatattaa atgagattat ttctgctctt caggaaattt ctgcaccact
5401 ggttttgtag ctgtttcata aaactgttga ctaaaagcta tgtctatgca accttccaag
5461 aatagtatgt caagcaactg gacacagtgc tgcctctgct tcaggactta acatgctgat
5521 ccagctgtac ttcagaaaaa taatattaat catatgtttt gtgtacgtat gacaaactgt
5581 caaagtgaca cagaatactg atttgaagat agccttttt atgtttctct atttctgggc
5641 tgatgaatta atattcattt gtattttaac cctgcagaat tttccttagt taaaaacact
5701 ttcctagctg gtcatttctt cataagatag caaatttaaa tctctcctcg atcagctttt
5761 aaaaaatgtg tactattatc tgaggaagtt ttttactgct ttatgttttt gtgtgttttg
5821 aggccatgat gattacattt gtggttccaa aataatttttt ttaaatatta atagcccata
5881 tacaaagata atggattgca catagacaaa gaaataaact tcagatttgt gatttttgtt
5941 tctaaacttg atacagattt acactattta taaatacgta tttattgcct gaaaatattt
6001 gtgaatggaa tgttgttttt ttccagacgt aactgccatt aaatactaag gagttctgta
6061 gttttaaaca ctactcctat tacattttat atgtgtagat aaaactgctt agtattatac
6121 agaaatttttt attaaaattg ttaaatgttt aaagggtttc ccaatgtttg agtttaaaaa
6181 agactttctg aaaaaatcca ctttttgttc attttcaaac ctaatgatta tatgtatttt
6241 atatgtgtgt gtatgtgtac acacatgtat aatatataca gaaacctcga tatataattg
6301 tatagatttt aaaagtttta ttttttacat ctatggtagt ttttgaggtg cctattataa
6361 agtattacgg aagtttgctg tttttaaagt aaatgtcttt tagtgtgatt tattaagttg
6421 tagtcaccat agtgatagcc cataaataat tgctggaaaa ttgtatttta taacagtaga
6481 aaacatatag tcagtgaagt aaatatttta aaggaaacat tatatagatt tgataaatgt
6541 tgtttataat taagagtttc ttatggaaaa gagattcaga atgataacct cttttagaga
6601 acaaataagt gacttatttt tttaaagcta gatgactttg aaatgctata ctgtcctgct
6661 tgtacaacat ggtttggggt gaaggggagg aaagtattaa aaaatctata tcgctagtaa
6721 attgtaataa gttctattaa aacttgtatt tcatatgaaa aatttgctaa tttaatatta
6781 actcatttga taataatact tgtcttttct acctctc
```

[0048]

[0049] "Gab1 폴리펩티드" (GRB2-연관된-결합 단백질 1)란 NCBI 참조 서열: NP_997006.1에 제공된 서열과 적어도 약 85% 아미노산 동일성을 갖는 단백질, 또는 그의 단편을 의미한다. 예시적인 Gab1 아미노산 서열은 하기에 제공된다:

[0050]

```
  1 msggevvcsg wlrksppekk lkryawkrrw fvlrsgrltg dpdvleyykn dhakkpirii
 61 dlnlcqqvda gltfnkkefe nsyifdinti drifylvads eeemnkwvrc icdicgfnpt
121 eedpvkppgs slqapadlpl aintappstq adsssatlpp pyqlinvpph letlgiqedp
181 qdylllincq skkpeptrth adsakstsse tdcndnvpsh knpassqskh gmngffqqqm
241 iydsppsrap sasvdsslyn lprsyshdvl pkvspsstea dgelyvfntp sgtssvetqm

301 rhvsisydip ptpgntyqip rtfpegtlgq tskldtipdi pppprppkphp ahdrspvetc
361 siprtasdtd ssyciptagm spsrsntist vdlnklrkda ssqdcydipr afpsdrsssl
421 egfhnhfkvk nvltvgsvss eeldenyvpm npnspprqhs ssftepiqea nyvpmtpgtf
481 dfssfgmqvp ppahmgfrss pktpprrpvp vadcepppvd rnlkpdrkgq spkilrlkph
541 glertdsqti gdfatrrkvk papleikplp eweelqapvr spitrsfard ssrfpmsprp
601 dsvhsttsss dshdseenyv pmnpnlssed pnlfgsnsld ggsspmikpk gdkqveyldl
661 dldsgkstpp rkqkssgsgs svadervdyv vvdqqktlal kstreawtdg rqstesetpa
721 ksvk
```

[0051]

[0052] "Gab1 폴리뉴클레오티드"란 Gab1 폴리펩티드를 코딩하는 핵산 분자를 의미한다. 예시적인 Gab1 폴리뉴클레오티드 서열은 NM_002039.3에 제공되고, 이는 하기에 재현된다:

[0053]

```
   1 aggggcgga gcgcaaagga cagaagctcc ggcaccgagt cggggcagag tcccgctgag
  61 tccgagcgct gctgaggcag ctggcgagac ggcacgtctg gaggcgaggc gggcgcactg
 121 aaaggaggcc ggcgcgcccg cggccccggc tcgcgttctg ttcaggttcg tgggcctgca
 181 gaggagagac tcgaactcgt ggaacccgcg caccgtggag tctgtccgcc cagtccgtcc
 241 ggggtgcgcg accaggagag ctaggttctc gccactgcgc gctcggcagg cgtcggctgt
 301 gtcgggagcg cgcccgccgc ccctcagctg cccggcccgg agcccgagac gcgcgcacca
 361 tgagcggtgg tgaagtggtc tgctccggat ggctccgcaa gtccccccg gagaaaaagt
 421 tgaagcgtta tgcatggaag aggagatggt tcgtgttacg cagtggccgt ttaactggag
 481 atccagatgt tttggaatat tacaaaaatg atcatgccaa gaagcctatt cgtattattg
 541 atttaaattt atgtcaacaa gtagatgctg gattgacatt taacaaaaaa gagtttgaaa
 601 acagctacat ttttgatatc aacactattg accggatttt ctacttggta gcagacagcg
 661 aggaggagat gaataagtgg gttcgttgta tttgtgacat ctgtgggttt aatccaacag
 721 aagaagatcc tgtgaagcca cctggcagct ctttacaagc accagctgat ttacctttag
 781 ctataaatac agcaccacca tccacccagg cagattcatc ctctgctact ctacctcctc
 841 catatcagct aatcaatgtt ccaccacacc tggaaactct tggcattcag gaggatcctc
 901 aagactacct gttgctcatc aactgtcaaa gcaagaagcc cgaacccacc agaacgcatg
 961 ctgattctgc aaaatccacc tcttctgaaa cagactgcaa tgataacgtc ccttctcata
1021 aaaatcctgc ttcctcccag agcaaacatg gaatgaatgg ctttttcag cagcaaatga
1081 tatacgactc tccaccttca cgtgccccat ctgcttcagt tgactccagc ctttataacc
1141 tgcccaggag ttattcccat gatgttttac caaaggtgtc tccatcaagt actgaagcag
1201 atggagaact ctatgttttt aataccccat ctgggacatc gagtgtagag actcaaatga
1261 ggcatgtatc tattagttat gacattcctc caacacctgg taatacttat cagattccac
1321 gaacatttcc agaaggaacc ttgggacaga catcaaagct agacactatt ccagatattc
1381 ctccacctcg gccaccgaaa ccacatccag ctcatgaccg atctcctgtg gaaacgtgta
1441 gtatcccacg caccgcctca gacactgaca gtagttactg tatccctaca gcagggatgt
1501 cgccttcacg tagtaatacc atttccactg tggatttaaa caaattgcga aaagatgcta
1561 gttctcaaga ctgctatgat attccacgag catttccaag tgatagatct agttcacttg
1621 aaggcttcca taaccacttt aaagtcaaaa atgtgttgac agtgggaagt gtttcaagtg
1681 aagaactgga tgaaaattac gtcccaatga atcccaattc accaccacga caacattcca
```

1741 gcagttttac agaaccaatt caggaagcaa attatgtgcc aatgactcca ggaacatttg
1801 atttttcctc atttggaatg caagttcctc ctcctgctca tatgggcttc aggtccagcc
1861 caaaaacccc tcccagaagg ccagttcctg ttgcagactg tgaaccaccc cccgtggata
1921 ggaacctcaa gccagacaga aaagtcaagc cagcgccttt agaaataaaa cctttgccag
1981 aatgggaaga attacaagcc ccagttagat ctcccatcac taggagtttt gctcgagact
2041 cttccaggtt tcccatgtcc ccccgaccag attcagtgca tagcacaact tcaagcagtg
2101 actcacacga cagtgaagag aattatgttc ccatgaaccc aaacctgtcc agtgaagacc
2161 caaatctctt tggcagtaac agtcttgatg gaggaagcag ccctatgatc aagcccaaag
2221 gagacaaaca ggtggaatac ttagatctcg acttagattc tgggaaatcc acaccaccac
2281 gtaagcaaaa gagcagtggc tcaggcagca gtgtagcaga tgagagagtg gattatgttg
2341 ttgttgacca acagaagacc ttggctctaa agagtacccg ggaagcctgg acagatggga
2401 gacagtccac agaatcagaa acgccagcga agagtgtgaa atgaaaatat tgccttgcca
2461 tttctgaaca aaagaaaact gaattgtaaa gataaatccc ttttgaagaa tgacttgaca
2521 cttccactct aggtagatcc tcaaatgagt agagttgaag tcaaaggacc tttctgacat
2581 aatcaagcaa tttagactta agtggtgctt tgtggtatct gaacaattca taacatgtaa
2641 ataatgtggg aaaatagtat tgtttagctc ccagagaaac atttgttcca cagttaacac
2701 actcgtagta ttactgtatt tatgcacttt ttcatctaaa acattgttct gggttttccc
2761 aatgtacctt accataattc ctttgggagt tcttgttttt tgtcacacta ctttatataa
2821 caatactaag tcaactaagc tacttttaga tttggaaatt gctgtttaca gtctaacaac
2881 attaaaatga gaggtagatt cacaagttag ctttctacct gaagcttcag gtgataacca
2941 ttagcttata cttggactca tcatttgttg ccttccaaaa tgctgaggat aatgtatgta
3001 ctggtgtcag gacctagttc tctggttaat gtacatttag tttttaatgg tggaactttg
3061 ttatattttg ttaattacag tgtttttggt tcattgagtg aagattctgc cgggtgggat
3121 cttgcacctt tgaaagactg aataattaca ctaccaagta agcctgcaaa tcattgatgg
3181 catgcagtga tgatgtgctc ttacacttgt taacatgtat taagtgttat ttgcaaaagg
3241 tagattatgt aaccaatcag gtacgtacca ggcagtgatg tgctaataca ctgatcaggt
3301 ttagacaatg agctttggtt gtgttcttgt tagtcctaat attggttttc agtttggaat
3361 taataaagca gttgacattc actgttagtt acagcaacat actgtgattt ttaattagat
3421 agtaattcag atttattact ctatgaaatt ctgtcttttg acaccatagt gccctttcta
3481 tgattttttt tacttaatat tcttcttggc cttatattta attccctatg caattaatat
3541 tttatatctg catttttta aaaaaaatag atgttatata agtgattctc gtatgtagca
3601 cctgttgctt ttccactgaa agaattacgg attttgtact gtgatttata ttcactgccc
3661 caattcaaga aatattggag ccttgctaca atgtgaaatg ttatagtcat ggactccttc
3721 caaccagatt tctgaaaaca ccagagggat ggtataattc tgtctcacct ataacatggt
3781 cctgtgacat agatattaag accacaagtt gtagtgaggc tacaattata ttcgtctgtc
3841 ttggctttgc aacataattt agaaagcacg tatagttgtt ttttaaccaa gttacataca
3901 atctcatgta ctgatttgag acttataaca atttttggag ggggcataga gaaaggagtg
3961 cccacagttg aggcatgacc ccctccattc agacctctaa ctgttgcctg agtacacaga
4021 tgtgccctga tttctggccc attggccata gtactgtgcc taatcaatgt aataggttta
4081 ttttcccaat cctcaaacta aaaatgttca taacaagatg aattgtagac tagtaacatt
4141 tgatgctttt aaatatttgc ttctttttaa acaaaaacta aaacccagaa gtgaattttt
4201 aggtggattt ttaaataaaa aagattgatt gagtttggtg tgcaagctgt tttataatga
4261 aacaacaaaa tgaaatctaa aatcctgaaa tgtgcctaaa ctatcaaaac acacgataca

[0054]

4321 gctaatgtgt aaagatgcta aattctgtta cttggaggat gaatatattt aagatttaaa
4381 acacaataat aaatacatga ttaattcaaa aataaaaatc tttacagctg cctatcaagg
4441 gtctaaagca cttaatgaat gtttttagtc taacttatca ttaacttttt acaagtcacc
4501 atatttgaag atctgtagca ctctgatttt cagaaaattt ttcattctga ataatttaaa
4561 aatggtgatg tattagaaag gcagtttgct ttagaaaact aaatcacatt gaacattgta
4621 ttagagaatt aaattaaaag tttcttacag agcagtattt tccaaacatt tttagcacta
4681 gaatcttttt agatgaaatt ttatgtataa ccccaataca taaagcctga aaactcaatt
4741 ttatcaatat aaatgtattt tgggttcaca tttatgctta ttcattttgg ctcattacta
4801 agcataataa gattctgagt tatttctgaa taacacaaat gtggagttat acatagttga
4861 tgaaaccagc agccaattta tagctatgcc ctgttttatt tgtatactat caagaaaatt
4921 ttgattcaca caaatgtaag caaaaataat aggttttaaa catacatctc aggaaattct
4981 ttaattagag atagctaaag ttattcaagg tctatacaaa aataagttat cctggtagtg
5041 gaagttaata cataagcagt ctccagtgtg gtaaagtagg gtatgtaaca catcagaatg
5101 tgcgttttta ttaggtttta aaatatgcac gtataaaaac taaatttgaa tcaaaccctt
5161 ttaactcacc tccaagaagc tagactttgg ccaggaatgg gctaaaaacc actggttaac
5221 gatgtgacag ttatgatctt ggagattgga aatctttctt ccacattaga gttctttacc
5281 ttaattcctt attctgaaaa attgtaagat tttatgaagg tttgaatact gaagcacagt
5341 tctgctttca aaaattaaaa ttcaaacttg aaaaagctgt ttaacccatg gaagatatca
5401 tttagtaaga tgtaaaagat tttttaaatc tacacttcag tttatacatc tttatcatta
5461 tcaatactat ataagttact gtgagcattt tagagaattc cataaaggta ctatgagtgt
5521 gtctgtatgt gtgtgtatat atagcattgt atttaatcat agactaaatt taatttgata
5581 tagaaatact actttacttg tacattaagg tcataatttc tgctggactc ttttatattt
5641 aattaatggg gattatagtc ttccttcata aatgcattta aacctgaaat tgaacaccag
5701 tgtttttctt tttctactta tgggaagttg tctgcttccc cctttagaga aaacagtatt
5761 tttatatttt gttaaaatat taactacttt atgcctacac actatgctgt agatactgat
5821 cataattctt gggtgttcac aaacactcct agtgcctctt ttttggcccg ttgaaagtgt
5881 tggtattact actttcacta cagagccttt ggccctctaa taatgctgag gtgggctgat
5941 ccttcccatt tctgtcttcg ggtcattctg gtaggtcttc tcctccactg tcaagtaagc
6001 aatcaggtcc gtgacaggga ttggacatat gaacaaatta agtggataca cacagtgaga
6061 aagatacatg cattctatgg taacaactac tgtcaataac atctgatgtt acatgcacat
6121 ttatatatat ataattttaa aaactgaact atgagaagcc atggtataaa tgaatattgt
6181 ggacatcatg gacttgatat gatagaaatc aattgtcagc ttgagaaagt tgtttttaat
6241 ctgtctaaat agttcatgca ttactacagt taaaaatagt ttcatttgtc ttctatagac
6301 ttaattttat tccggttcag tataatctct gttaacagag tttcagcaaa ctgattggtc
6361 aaggtattaa catagcttct acttccttta cttaaaaaga tgtggtttta tgtaagttct
6421 tgattactga tgatcatccc aaattttgac aacaaaatca tatgtataaa tttatttctc
6481 ccctcttgtt catcatcttt tgtaaaggtc ccattgtaga tcttttctgc taccaaataa
6541 aacttttcaa acaatttggt ttcaagacct taaatagaca agttggatac taagattgtg
6601 aactgataag gacatataaa tttatatttc cagcccttcc ttagagtctt tatctgcatc
6661 aaaaacccaa ttctgccatt aactgtgctt cccagtccca cctctatatg tcactcattt
6721 tctgcaacaa agatctcact aaatcatgtt gaaacacaag tcatgatcct ctctaagtaa
6781 atagaaaaag ctccctggaa aaactctgtt gccacatgca cgtgccctgt tactcctcca
6841 gccagccagt gctgccagca ttttattgtg taaaagtcca aataaataag ggcctgcatg

[0055]

6901 caacctttat cttcagaaac taggttttat atgtaaaatg tgacttggga aatgattctg
6961 tttattaact ggctgggatt tttcatttct atgaaagttt caaacatctc cagtacttta
7021 taaaatccca acaattgctg taagtcagca ctttggtcca ctcagcccac ccagcccact
7081 tgcaactctg actcttcact gaatcatatt tgggaagttt gggtagggtg aggctatctt
7141 cttcaagatt attttctcat atgtctgtct gtcaccttgt aaaccatgag actcctgggt
7201 atttgcatgt aacttctttg aggaagttac caccatctct gatatagaca cacttttga
7261 gttgcagttt ctgttagaat tttttggaga ctaacttgcc aattctgtga atgttattga
7321 atatttaaaa agctgggtct gtaatgggag gcattttatt agctgttgtg attgggtaac
7381 atgtcccctt agatttcctg atttaaaatt atacaaaatt actatttttg ataaaataaa
7441 ggaacaccta cagaaaatta agtttctaag atgtttctat acttcattag aaaagatttt
7501 attactatta cttatggtta ttggtgatta acacttaatg cgtctcctct gattttgtgt
7561 tccatgaggt gcttggaaca tttggagtgc tctgtgcgag ggacatacag tgatatagga
7621 aatttaaaaa ttaaaataat acccaaaacc cactttatca gatatggtat tgtgatggtt
7681 aatattatgt gtcaacttgg tgaggctatg gcgcccatgt gtttggtcaa acactagcct
7741 agatgttgct gtgaatatat tttgtagatg tgattaacat ttacaatcag ttgattttaa
7801 gtaaagcaga ttctcatcca aaaaaaaaaa aaaaaa

[0056]

[0057] "Sfmbt2 폴리펩티드" (4개의 MBT 도메인을 갖는 scm-유사 단백질 2)란 NCBI 참조 서열: NP_001018049.1에 제공된 서열과 적어도 약 85% 아미노산 동일성을 갖는 단백질, 또는 그의 단편을 의미한다. 예시적인 Sfmbt2 아미

노산 서열은 하기에 제공된다:

```
  1 mestlsasnm qdpsssplek clgsangngd ldseegssle etgfnwgeyl eetgasaaph
 61 tsfkhveisi qsnfqpgmkl evanknnpdt ywvatiittc gqllllrycg ygedrradfw
121 cdvviadlhp vgwctqnnkv lmppdaikek ytdwteflir dltgsrtapa nllegplrgk
181 gpidlitvgs lielqdsqnp fqywivsvie nvggrlrlry vgledtesyd qwlfyldyrl
241 rpvgwcqenk yrmdppseiy plkmasewkc tlekslidaa kfplpmevfk dhadlrshff
301 tvgmkletvn mcepfyispa svtkvfnnhf fqvtiddlrp epsklsmlch adslgilpvq
361 wclkngvslt ppkgysgqdf dwadyhkqhg aqeappfcfr ntsfsrgftk nmkleavnpr
421 npgelcvasv vsvkgrlmwl hleglqtpvp evivdvesmd ifpvgwcean sypltaphkt
481 vsqkkrkiav vqpekqlppt vpvkkiphdl clfphldttg tvngkyccpq lfinhrcfsg
541 pylnkgriae lpqsvgpgkc vlvlkevlsm iinaaykpgr vlrelqlved phwnfqeetl
601 kakyrgktyr avvkivrtsd qvanfcrrvc akleccpnlf spvlisencp encsihtktk
661 ytyyygkrkk iskppigesn pdsghpkpar rrkrrksifv qkkrrssavd ftagsgeese
721 eedadamddd taseetgsel rddqtdtssa evpsarprra vtlrsgsepv rrpppertrr
781 grgapaassa eegekcpptk pegtedtkqe eeerlvlesn plewtvtdvv rfikltdcap
841 lakifqeqdi dgqalllltl ptvqecmelk lgpaiklchq iervkvafya qyan
```

[0058]

[0059] "Sfmbt2 폴리뉴클레오티드"란 Sfmbt2 폴리펩티드를 코딩하는 폴리펩티드를 의미한다. 예시적인 Sfmbt2 폴리뉴클레오티드 서열은 NM_001018039.1에 제공되고, 이는 하기에 재현된다:

```
1 cgccttgtgt gtgctggatc ctgcgcgggt agatccccga gtaatttttt ctgcaggatg
       61 aattaagaga agagacactt gctcatcagg catggagagc actttgtcag cttccaatat
      121 gcaagaccct tcatcttcac ccttggaaaa gtgtctcggc tcagctaatg gaaatggaga
      181 ccttgattct gaagaaggct caagcttgga ggaaactggc tttaactggg gagaatattt
      241 ggaagagaca ggagcaagtg ctgctcccca cacatcattc aaacacgttg aaatcagcat
      301 tcagagcaac ttccagccag gaatgaaatt ggaagtggct aataagaaca acccggacac
      361 gtactgggtg gccacgatca ttaccacgtg cgggcagctg ctgcttctgc gctactgcgg
      421 ttacggggag gaccgcaggg ccgacttctg gtgtgacgta gtcatcgcgg atttgcaccc
      481 cgtggggtgg tgcacacaga acaacaaggt gttgatgccg ccggacgcaa tcaaagagaa
      541 gtacacagac tggacagaat ttctcatacg tgacttgact ggttcgagga cagcacccgc
      601 caacctcctg gaaggtcctc tgcgagggaa aggccctata gacctcatta cagttggttc
      661 cttaatagaa cttcaggatt cccagaaccc ttttcagtac tggatagtta gtgtgattga
      721 aaatgttgga ggaagattac gccttcgcta tgtgggattg gaggacactg aatcctatga
      781 ccagtggttg ttttacttgg attacagact tcgaccagtt ggttggtgtc aagagaataa
      841 atacagaatg gacccacctt cagaaatcta tcctttgaag atggcctctg aatggaaatg
      901 tactctggaa aaatccctta ttgatgctgc caaatttcct cttccaatgg aagtgtttaa
      961 ggatcacgca gatttgcgaa gccatttctt cacagttggg atgaagcttg agacagtgaa
     1021 tatgtgcgag ccctttaca tctctcctgc gtcggtgact aaggttttta acaatcactt
     1081 ttttcaagtg actattgatg acctaagacc tgaaccaagt aaactgtcaa tgctgtgcca
     1141 tgcagattct ttggggattt tgccagtaca gtggtgcctt aaaaatggag tcagcctcac
     1201 tcctcccaaa ggttactctg gccaggactt cgactgggca gattatcaca agcagcatgg
     1261 ggcgcaggaa gcccctccct tctgcttccg aaatacatca ttcagtcgag gtttcacaaa
     1321 gaacatgaaa cttgaagctg tgaaccccag gaatccagga gaactgtgtg tggcctccgt
     1381 tgtgagtgtg aaggggcggc taatgtggct tcacctggaa gggctgcaga ctcctgttcc
     1441 agaggtcatt gttgatgtgg aatccatgga catcttccca gtgggctggt gtgaagccaa
     1501 ttcttatcct ttgactgcac cacacaaaac agtctcacaa aagaagagaa agattgcagt
     1561 cgtgcaacca gagaaacaat tgccgcccac agtgcctgtt aagaaaatac ctcatgacct
     1621 ttgtttattc cctcacctgg acaccacagg aaccgtcaac gggaaatact gctgtcctca
     1681 gctcttcatc aaccacaggt gtttctcagg cccttacctg aacaaaggaa ggattgcaga
     1741 gctacctcag tcggtgggac cgggcaaatg cgtgctggtt cttaaagagg ttcttagcat
     1801 gataatcaac gcagcctaca agcctggaag ggtattaaga gaattacagc tggtagaaga
     1861 tccccactgg aatttccagg aagagacgct gaaggccaaa tacagaggca aaacatacag
     1921 ggctgtggtc aaaatcgtac ggacatctga ccaagtcgca aatttctgcc gccgagtctg
     1981 tgccaagcta gagtgctgtc caaatttgtt tagtcctgtg ctgatatctg aaaactgccc
     2041 agagaactgc tccattcata ccaaaaccaa atacacctat tactatggaa agagaaagaa
     2101 gatctccaag ccccccatcg gggaaagcaa ccccgacagc ggacacccca aacccgccag
     2161 gcggaggaag cgacggaaat ccattttcgt gcagaagaaa cggaggtctt ctgccgtgga
     2221 cttcaccgcg ggctcggggg aggaaagtga agaggaggac gctgacgcca tggacgatga
     2281 caccgccagt gaggagaccg gctccgagct ccgggatgac cagacggaca cctcgtcggc
     2341 ggaggtgccc tcggcccggc cccggagggc cgtcaccctg cggagcggct cagagcccgt
     2401 gcgccggcca cccccagaga ggacacgaag gggccgcggg gcgccggctg cctcctcagc
     2461 agaggaaggg gagaagtgcc cgccgaccaa gcccgagggg acagaggaca cgaaacagga
     2521 ggaggaggag agactggttc tggagagcaa cccgttggag tggacggtca ccgacgtggt
```

[0060]

2581 gaggttcatt aagctgacag actgtgcccc cttggccaag atatttcagg agcaggatat
2641 tgacggccaa gcactcctgc ttctgaccct tccgacggtg caggagtgca tggagctgaa
2701 gctggggcct gccatcaagt tatgccacca gatcgagaga gtcaaagtgg ctttctacgc
2761 ccagtacgcc aactgagtct gccctcggga ggtggcccat tattgctggg atgcggtgtt
2821 ggtaaaggtt tccaggactg aaactttgat tttccgggat atgttaaatg gtacagccac
2881 taagtatcac cagaaaacca gaagcccagg atcttctgcc tccgccagcc tgtgagctgt
2941 ttccatgttt tcaaagcaca gcagcagtcg cttctgggga gtgccagtta aagtcatgca
3001 tcagaccctg ccagacgtgg gcctgcttct tggctcaccc acgttttgcc tttctcctgc
3061 cccaaatcag gcagctccct tggagcaggg tttcctcaga tgaggactgc attctttgaa
3121 aacaaagaat gtcgccaagg aagaaacctc acgccatgct gtagtgtttc ctgtaatcac
3181 acgagcacat ttatatatgc agtttcccat ggataggcgt gtgaccctgg ttgagtggca
3241 cttgcggttt catcttggtg gcaactcctt tgcaatgcag ctggcagcga catccttata
3301 aaaacatgtg ctaaagctct gtcctctgtt agaggtgcct tttaggaata cggggagtga
3361 aggaaggccg gcaggcatct ccatgcaact agatggtttg tttgtttgtt tgtttgtttg
3421 ttgttcattt tgttgtgttt tttgagacag ggtcttgctc tgtcgcccag gttgtaatgc
3481 agtggcgcaa tctcagctca ctgcaacctc tctctcccgg gttcaagtga ttctcctgcc
3541 tcagcctccc aagtagctgg gattacaggc acccaccacc atgcctggct aatttttgta
3601 tttttggtag agacagggtt tcaccatgtt ggtcaggcta gtcttgaact cccaacctca
3661 agtgatctgc ccgcctcggc ctcccaacgt gctgggatta caggtgtgag ccactacgcc
3721 ccggcccaac tggatggttt ttgattgaag cctagaacat ctgtagagac aaactctacc
3781 cagtcttttc tagaccctca actatctcca gtgttgttgt ttaatcgtag ccggatcagg
3841 gagtgagtct tttaggcaaa tgttggatta tatatcaaag gaaaagctta gtttcagaga
3901 ggaggaaggg aaagagatgt gagggaagca tttcatcaac cagctacgtc ccccttagaa
3961 ggatcactgc agcaggtcac cgagcaggag tccctctgag cgtcccttct gtctcgttct
4021 gccctagctg gcagcatatg aaccaggcat gatgcagcag gagcagtgaa tctggagtca
4081 gccacttggc accctggttt cgctgagaac aaactctgag atcttgggtg acttctcatc
4141 actctggacc tccattcctg tgaagtgaca ggtgtggacc ctgagggtgc ggtggtgagc
4201 acactgtctc ctgctggcat tcacccccact catgctggaa aggaagatcc agatcgtaca
4261 aaaattagaa aaagaaagaa taagaagggt ctggtcccag ttctgactcg gccattctta
4321 cagctctttc tggctttgag tttgcttgtg gaatttcctg ggcagttgtg ttaaatccgc
4381 caggtcacgt gcagacaaag ctgtggctgc gagagttggc tggcctcttg gaccagaagc
4441 catctccata tcctcatgag cgattccata tctccactca gaccctgtgg actacagtgt
4501 tccgctgtgg tggctgccaa gatgccttct taaacttatg caaggaaacc aaaccctccc
4561 acagttccca agcagacact ggaagcagag gcttctcacc cttcctgctt tttcaccaca
4621 atcaccttga gctcgtccct tggactagag tctccacagt tccagtaaaa ttctgcggtg
4681 ggctgatgag ctgcttgcat ttctgtgaca tttccagata tgattctcag tgggattttg
4741 gaaactttga ttgctcaagc tcacccttct taacattctg taatggttac agatgagaat
4801 ggaaaacaca tattttatgg atgaggcgtt ttggtctccc ctgcagtcga tttctagaat
4861 caagttttag agttcggctg atgcatctgc ctggggacct cagatgggag gagtgtgtca
4921 gttgtacccc gacagaaatg tctctgggat ctgtggctgg cttgccccgg gcatctctcc
4981 tttaagctca agttttgaac tctctgcggt tttccacccc tgccttctca gccacatgct
5041 tttggcctta aacgctcagt cttgtggagt tcaactctgt caaacgattg gaaagggcat
5101 ccatttccag atctttggca ttttccccgc gctgactctt tgatgatcct tcactgtggc

[0061]

```
5161 cttttcaagc tcagctgttc ctgttgtatt tgagacgagg gtgagggaat gtggtggcca
5221 caaaagaaca gggacttgca gcacaaatgt cacttctgtc tccctttca gtggtagcac
5281 ggaggaggag gtgctgcgtt ggagggaggg gatcctccag gagctctctg gagcccatct
5341 aggaagctag agtgtgtggc ccgccaggag ctcaggaagg atacagccac tgtcgcaggg
5401 gaaagtgttt gcttcccgtg gagccaagcg cccaagactc tccgtatcct tcaccctgac
5461 agtttaactt cagcgtttct ctgtgcagtt gcggtcacca tgggtgagca ctgtctgtgc
5521 acgtgccagg gaggagatgg ctgggaccac tgcacaggag ggcgcagcct ggcgtcgcca
5581 tgaaagttgt ctctgtgcca tctctccggt ccttgaggag agcccagaaa gattttagga
5641 cccaggaggt gcttttcctc cagctgttgc cagtgtcctt ctgagcctgg attctccggg
5701 gatttccgtc gtggtggatg gacttcacat cagcagcagt tctggtacag aattgtaatg
5761 tgttttcatt tctctgtagg attcacctct caccagcgtc tgtcttaaag gtagggccaa
5821 tttcatggag cattttctg tgtgtgtcct tgttgctttt gccagaaaaa gtggatttga
5881 catgcgtgcc ccgatgccac catagcccct aggccaacaa tgtcatggtc taaacaccaa
5941 aaagtgatgc cccgcattcc ttccctggat ggtaccgttt cttctccgtc tctctttgat
6001 gattctttgg gaccaaagtc ctctccttag tgcgcctact tcctgtgggc atcatgccac
6061 ttggaactta ttggaactgg cccgggagac tctgcagtct gcgccgtttg aaaaccctga
6121 gaaagagatg ccacctcaac ttgaatcatg acagcccatc gctcagtctc accctaaact
6181 catggagctt gtttcagctc ctcacttctt gactgtattt gtactatgtt gaaaaaatat
6241 cctgtccaca aagacataag cctaacaacc tagaaaaaca acagggtact actggcatta
6301 cagaacttct ttgcctttca aaacaaaagc aaaacacagt gaacttcacc acggagctgc
6361 acagcgtggg gaactcatcc atcactttca aaattagagt catttgatcc aagttggagt
6421 cagacacagt atttgagctg cacggcttct gggttctccc accttatttg atcatattcg
6481 aaagattatt tcctgtgttt gctttgattt gttcctcagt acattaaaat gatccacacc
6541 ttgaacactg ccctctctag aaggttgatt ttgatcagcc ttttgaagat gggtgtcgtt
6601 tccctaactt atctcacaga attttgagtg ttgtatttgg caagttctga gatttgcctt
6661 ctgtcttatg ccaaacaccc ctttctaaga gctgtccccg cttagtttta gaagtactag
6721 gggttttcat acttatttta tagaacaccc atttatattt atttctgtat atagaactaa
6781 aaaaaacagt agtgttaaaa atctttgttg tggtttgagc atctttgctg cttttggatt
6841 gagatggcga atcaaggctt cacttcctct ctcttctgtc tttagaaagc tgtgatcgtg
6901 cgtgcaatta tttgaaaggc aacatagtca attaagaaac ctgtagttgt taaggaagaa
6961 attgttggca agatatccat actgcccata tctcgttggt gcaataatta aatagcaaag
7021 gaaatctgta ttggcaacta ttataattca ataattcttt tgtttactgc ccttttctgt
7081 tcaagaattt tctggaaatt actccctttc acatggttga actcttaagt tgaccagttc
7141 tcatagctct atcactagaa tggtttgcag ataccccaaa catactatga taaaatcaaa
7201 ttgtgctact tttgacccat gtaatttacc taaaagttgt aattgctgac agagtactgc
7261 cttgaatttt ggtttaaaac ctctctagtt tcaatgacaa gtaacaactc aaataattcc
7321 atattgtttg aggaagaggc cataatcctt ctgaattgtt ggcactaagt aatgggattt
7381 ggcccagtaa gtatgacggt cgtgtcgcct aaccaacgca gagcagtgct ttttgtgtgg
7441 ctgaagcgat gtgctgacga aaaaaggaaa attctaggac aatcgttggc taaaaatcac
7501 cttaggatga aaaatttgag gcaaatttttt ttaaatgaca gaaaaagata atcatctcac
7561 ttgcttgaaa caggagccag catgatctct ggaagcatca actatccctc gtcgtgattg
7621 ttgaaagctc tttcactgtt ttgcattcta gtttgaatag tttgtattga aattggattc
7681 ctatcttgtg tatgtttttg gtgcgtaaaa gggaaaaatt ggtgtcatta cttttgaaat
```

[0062]

```
7741 ttgcaggacg aagggcatgc ttttggtttg ctgtaagatt gtattctgta tatatgtttt
7801 catgtaaata aatgaaaatc tatatcagag ttatatttta atttttattc taaatgaaaa
7861 aaacccttttt tacttcaaaa aaattgtaag ccacattgtt aataaagtaa aaataaattc
7921 ta
```

[0063]

[0064] "Smoc1 폴리펩티드" (SPARC 관련된 모듈 칼슘 결합 1)란 NCBI 참조 서열: NP_001030024에 제공된 서열과 적어도 약 85% 아미노산 동일성을 갖는 단백질, 또는 그의 단편을 의미한다. 예시적인 Smoc1 아미노산 서열은 하기에 제공된다:

```
  1 mlparcarll tphlllvlvq lsparghrtt gprflisdrd pqcnlhcsrt qpkpicasdg
 61 rsyesmceyq rakcrdptlg vvhrgrckda gqskcrlera qaleqakkpq eavfvpecge
121 dgsftqvqch tytgycwcvt pdgkpisgss vqnktpvcsg svtdkplsqg nsgrkddgsk
181 ptptmetqpv fdgdeitapt lwikhlvikd sklnntnirn sekvyscdqe rqsaleeaqq
241 npregivipe capgglykpv qchqstgycw cvlvdtgrpl pgtstryvmp scesdarakt
301 teaddpfkdr elpgcpegkk mefitsllda lttdmvqain saaptgggrf sepdpshtle
361 ervvhwyfsq ldsnssndin kremkpfkry vkkkakpkkc arrftdycdl nkdkvislpe
421 lkgclgvske vgrlv
```

[0065]

[0066] "Smoc1 폴리뉴클레오티드"란 Smoc1 폴리펩티드를 코딩하는 핵산 분자를 의미한다. 예시적인 Smoc1 폴리뉴클레오티드 서열은 XM_005267995.1에 제공되고, 이는 하기에 재현된다:

```
   1 ataacgggaa ttcccatggc ccgggctcag gcgtccaacc tgctgccgcc tgggccccgc
  61 cgagcggagc tagcgccgcg cgcagagcac acgctcgcgc tccagctccc ctcctgcgcg
 121 gttcatgact gtgtcccctg accgcagcct ctgcgagccc ccgccgcagg accacggccc
 181 gctccccgcc gccgcgaggg ccccgagcga aggaaggaag ggaggcgcgc tgtgcgcccc
 241 gcggagcccg cgaaccccgc tcgctgccgg ctgcccagcc tggctggcac catgctgccc
 301 gcgcgctgcg cccgcctgct cacgccccac ttgctgctgg tgttggtgca gctgtcccct
 361 gctcgcggcc accgcaccac aggccccagg tttctaataa gtgaccgtga cccacagtgc
 421 aacctccact gctccaggac tcaacccaaa cccatctgtg cctctgatgg caggtcctac
 481 gagtccatgt gtgagtacca gcgagccaag tgccgagacc cgaccctggg cgtggtgcat
 541 cgaggtagat gcaaagatgc tggccagagc aagtgtcgcc tggagcgggc tcaagccctg
 601 gagcaagcca agaagcctca ggaagctgtg tttgtcccag agtgtggcga ggatggctcc
 661 tttacccagg tgcagtgcca tacttacact gggtactgct ggtgtgtcac cccggatggg
 721 aagcccatca gtggctcttc tgtgcagaat aaaactcctg tatgttcagg ttcagtcacc
 781 gacaagccct tgagccaggg taactcagga aggaaagtct cctttcgatt ctttttaacc
 841 ctcaattcag atgacgggtc taagccgaca cccacgatgg agacccagcc ggtgttcgat
 901 ggagatgaaa tcacagcccc aactctatgg attaaacact tggtgatcaa ggactccaaa
 961 ctgaacaaca ccaacataag aaattcagag aaagtctatt cgtgtgacca ggagaggcag
1021 agtgccctgg aagaggccca gcagaatccc cgtgagggta ttgtcatccc tgaatgtgcc
1081 cctgggggac tctataagcc agtgcaatgc caccagtcca ctggctactg ctggtgtgtg
```

[0067]

```
1141 ctggtggaca cagggcgccc gctgcctggg acctccacac gctacgtgat gcccagttgt
1201 gagagcgacg ccagggccaa gactacagag gcggatgacc ccttcaagga cagggagcta
1261 ccaggctgtc cagaagggaa gaaaatggag tttatcacca gcctactgga tgctctcacc
1321 actgacatgg ttcaggccat taactcagca gcgcccactg gaggtgggag gttctcagag
1381 ccagacccca gccacaccct ggaggagcgg gtagtgcact ggtatttcag ccagctggac
1441 agcaatagca gcaacgacat taacaagcgg gagatgaagc ccttcaagcg ctacgtgaag
1501 aagaaagcca agcccaagaa atgtgcccgg cgtttcaccg actactgtga cctgaacaaa
1561 gacaaggtca tttcactgcc tgagctgaag ggctgcctgg gtgttagcaa agaagtagga
1621 cgcctcgtct aaggagcaga aaacccaagg gcaggtggag agtccaggga ggcaggatgg
1681 atcaccagac acctaacctt cagcgttgcc catggccctg ccacatcccg tgtaacataa
1741 gtggtgccca ccatgtttgc acttttaata actcttactt gcgtgttttg tttttggttt
1801 cattttaaaa caccaatatc taataccaca gtgggaaaag gaaagggaag aaagacttta
1861 ttctctctct tattgtaagt ttttggatct gctactgaca acttttagag ggttttgggg
1921 gggtgggggа gggtgttgtt ggggctgaga agaaagagat ttatatgctg tatataaata
1981 tatatgtaaa ttgtatagtt cttttgtaca ggcattggca ttgctgtttg tttatttctc
2041 tccctctgcc tgctgtgggt ggtgggcact ctggacacat agtccagctt tctaaaatcc
2101 aggactctat cctgggccta ctaaacttct gtttggagac tgacccttgt gtataaagac
2161 gggagtcctg caattgtact gcggactcca cgagttcttt tctggtggga ggactatatt
2221 gccccatgcc attagttgtc aaaattgata agtcacttgg ctctcggcct tgtccaggga
2281 ggttgggcta aggagagatg gaaactgccc tgggagagga agggagtcca gatcccatga
2341 atagcccaca caggtaccgg ctctcagagg gtccgtgcat tcctgctctc cggacccccа
2401 aagggcccag cattggtggg tgcaccagta tcttagtgac cctcggagca aattatccac
2461 aaaggatttg cattacgtca ctcgaaacgt tttcatccat gcttagcatc tactctgtat
2521 aacgcatgag aggggaggca aagaagaaaa agacacacag aagggccttt aaaaaagtag
2581 atatttaata tctaagcagg ggaggggaca ggacagaaag cctgcactga ggggtgcggt
2641 gccaacaggg aaactcttca cctccctgca aacctaccag tgaggctccc agagacgcag
2701 ctgtctcagt gccaggggca gattgggtgt gacctctcca ctcctccatc tcctgctgtt
2761 gtcctagtgg ctatcacagg cctgggtggg tgggttgggg gaggtgtcag tcaccttgtt
2821 ggtaacacta aagttgtttt gttggttttt taaaaaccca atactgaggt tcttcctgtt
2881 ccctcaagtt ttcttatggg cttccaggct ttaagctaat tccagaagta aaactgatct
2941 tgggtttcct attctgcctc ccctagaagg gcaggggtga taacccagct acagggaaat
3001 cccggcccag ctttccacag gcatcacagg catcttccgc ggattctagg gtgggctgcc
3061 cagccttctg gtctgaggcg cagctccctc tgcccaggtg ctgtgcctat tcaagtggcc
3121 ttcaggcaga gcagcaagtg gcccttagcg ccccttccca taagcagctg tggtggcagt
3181 gagggaggtt gggtagccct ggactggtcc cctcctcaga tcacccttgc aaatctggcc
3241 tcatcttgta ttccaacccg acatccctaa aagtacctcc acccgttccg ggtctggaag
3301 gcgttggcac cacaagcact gtccctgtgg gaggagcaca accttctcgg gacaggatct
3361 gatggggtct tgggctaaag gaggtccctg ctgtcctgga gaaagtccta gaggttatct
3421 caggaatgac tggtggccct gccccaacgt ggaaaggtgg gaaggaagcc ttctcccatt
3481 agccccaatg agagaactca acgtgccgga gctgagtggg ccttgcacga gacactggcc
3541 ccactttcag gcctggagga agcatgcaca catggagacg gcgcctgcct gtagatgttt
3601 ggatcttcga gatctcccca ggcatcttgt ctcccacagg atcgtgtgtg taggtggtgt
3661 tgtgtggttt tcctttgtga aggagagagg gaaactattt gtagcttgtt ttataaaaaa
3721 taaaaaatgg gtaaatcttg
```

[0068]

[0069] "리신 27에서의 트리-메틸화 히스톤 H3 (H3K27me3)"이란 히스톤 H3 단백질 서브유닛 상의 리신 27의 트리메틸화

를 의미한다. H3K27me3 변형은 일반적으로 유전자 억제와 연관된다.

[0070] "작용제"란 펩티드, 핵산 분자, 또는 작은 화합물을 의미한다.

[0071] "대립유전자"란 염색체 상의 동일한 장소에서 발견되는 유전자의 2개 이상의 대안적 형태 중 하나를 의미한다.

[0072] "변경"이란 본원에 기재된 것들과 같은 표준 기술 공지된 방법에 의해 검출된 바로, 유전자 또는 폴리펩티드의 발현 수준 또는 활성의 변화 (증가 또는 감소)를 의미한다. 본원에 사용된 바와 같은 변경은 발현 수준의 10% 변화, 바람직하게는 발현 수준의 25% 변화, 보다 바람직하게는 40% 변화, 및 가장 바람직하게는 50% 이상의 변화를 포함한다.

[0073] "개선시키다"란 질환의 발달 또는 진행을 감소시키거나, 억제하거나, 약화시키거나, 줄이거나, 정지시키거나, 안정화하는 것을 의미한다.

[0074] 이 개시내용에서, "포함하다 (comprise)", "포함하는 (comprising)", "함유하는" 및 "갖는" 등은 미국 특허법에서 그들에게 할당된 의미를 가질 수 있고, "포함하다 (include)", "포함하는 (including)" 등을 의미할 수 있으며; "로 본질적으로 이루어진" 또는 "본질적으로 이루어지다"는 마찬가지로 미국 특허법에서 할당된 의미를 갖고, 상기 용어는 인용된 것의 기본적 또는 신규한 특징이 인용된 것 초과의 존재에 의해 변화되지 않는 한, 인용된 것 초과의 존재를 허용하는 개방-말단이지만, 선행 기술 실시양태는 배제한다.

[0075] "검출하다"는 검출되는 분석물의 존재, 부재 또는 양을 확인하는 것을 지칭한다.

[0076] "질환"이란 세포, 조직, 또는 기관의 정상적인 기능을 손상시키거나, 방해하는 임의의 상태 또는 장애를 의미한다. 장애의 예로는 H3K27me3-의존성 각인에 의한 대립유전자의 비바람직한 억제와 연관된 것들을 들 수 있다. 소안구증은 각인 장애와 관련된 H3K27me3-의존성 각인과 연관된 예시적인 장애이다.

[0077] "DNA"란 데옥시리보핵산을 의미한다. 다양한 실시양태에서, 용어 DNA는 게놈 DNA, 재조합 DNA, 또는 cDNA를 지칭한다. 특정 실시양태에서, DNA는 "표적 영역"을 포함한다. 본원에서 고려되는 DNA 라이브러리는 게놈 DNA 라이브러리, 및 RNA로부터 구축된 cDNA 라이브러리, 예를 들어, RNA 발현 라이브러리를 포함한다. 다양한 실시양태에서, DNA 라이브러리는 1개 이상의 추가의 DNA 서열 및/또는 태그를 포함한다.

[0078] "유효량"이란 비치료된 환자에 비해 질환의 증상을 개선시키는 데 요구되는 양을 의미한다. 질환의 치료적 치료를 위한 본 발명을 실시하는 데 사용되는 활성 화합물(들)의 유효량은 투여의 방식, 대상체의 연령, 체중, 및 일반적 건강에 따라 다양하다. 궁극적으로, 담당 의사 또는 수의사는 적절한 양 및 투여량 처방을 결정할 것이다. 이러한 양은 "유효"량으로 지칭된다.

[0079] "단편"이란 폴리펩티드 또는 핵산 분자의 부분을 의미한다. 이 부분은 참조 핵산 분자 또는 폴리펩티드의 전체 길이의 바람직하게는, 적어도 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 또는 90%를 함유한다. 단편은 10, 20, 30, 40, 50, 60, 70, 80, 90, 또는 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000개의 뉴클레오티드 또는 아미노산을 함유할 수 있다.

[0080] 용어 "단리된", "정제된", 또는 "생물학적으로 순수한"은 그의 천연 상태에서 발견되는 바와 같은 통상적으로 이를 수반하는 성분이 다양한 정도로 없는 물질을 지칭한다. "단리하다"는 원래 공급원 또는 환경으로부터의 분리의 정도를 나타낸다. "정제하다"는 단리보다 더 높은 분리의 정도를 나타낸다. "정제된" 또는 "생물학적으로 순수한" 단백질은 임의의 불순물이 단백질의 생물학적 특성에 물질적으로 영향을 미치지 않거나, 다른 유해 결과를 유발하지 않도록 다른 물질이 충분히 없는 것이다. 즉, 본 발명의 핵산 또는 펩티드는 그것이 재조합 DNA 기법에 의해 생성되는 경우 세포성 물질, 바이러스성 물질 또는 배양 배지, 또는 화학적으로 합성되는 경우 화학적 전구체 또는 다른 화학물질이 실질적으로 없는 경우 정제된다. 순도 및 균질성은 전형적으로 분석적 화학 기법, 예를 들어, 폴리아크릴아미드 겔 전기영동 또는 고성능 액체 크로마토그래피를 사용하여 결정된다. 용어 "정제된"은 핵산 또는 단백질이 전기영동 겔에서 본질적으로 한 밴드를 발생시키는 것을 나타낼 수 있다. 변형, 예를 들어, 인산화 또는 글리코실화될 수 있는 단백질에 대해, 상이한 변형은 별개로 정제될 수 있는 상이한 단리된 단백질을 발생시킬 수 있다.

[0081] "단리된 폴리뉴클레오티드"란 본 발명의 핵산 분자가 유래되는 유기체의 천연-발생 게놈에서 유전자를 플랭킹하는 유전자가 없는 핵산 (예를 들어, DNA)을 의미한다. 따라서, 상기 용어는 예를 들어, 벡터 내로; 자율적으로 복제하는 플라스미드 또는 바이러스 내로; 또는 원핵생물 또는 진핵생물의 게놈 DNA 내로 혼입된; 또는 다른 서열과 독립적인 별개의 분자 (예를 들어, PCR 또는 제한 엔도뉴클레아제 소화에 의해 생성된 cDNA 또는 게놈 또는 cDNA 단편)로서 존재하는 재조합 DNA를 포함한다. 또한, 상기 용어는 DNA 분자로부터 전사된 RNA 분자, 뿐

만 아니라 추가의 폴리펩티드 서열을 코딩하는 하이브리드 유전자의 일부인 재조합 DNA를 포함한다.

[0082] "단리된 폴리펩티드"란 자연적으로 그것을 수반하는 성분으로부터 분리된 본 발명의 폴리펩티드를 의미한다. 전형적으로, 폴리펩티드는 그것이, 그것이 자연적으로 회합되는 단백질 및 천연-발생 유기 분자로부터 적어도 60 중량% 유리되는 경우 단리된다. 바람직하게는, 제제는 적어도 75 중량%, 보다 바람직하게는 적어도 90 중량%, 및 가장 바람직하게는 적어도 99 중량% 본 발명의 폴리펩티드이다. 본 발명의 단리된 폴리펩티드는 예를 들어, 천연 공급원으로부터의 추출에 의해, 이러한 폴리펩티드를 코딩하는 재조합 핵산의 발현에 의해; 또는 단백질을 화학적으로 합성함으로써 얻어질 수 있다. 순도는 임의의 적절한 방법, 예를 들어, 컬럼 크로마토그래피, 폴리아크릴아미드 겔 전기영동에 의해, 또는 HPLC 분석에 의해 측정될 수 있다.

[0083] 본원에서 제공된 범위는 범위 내의 모든 값에 대한 약칭인 것으로 이해된다. 예를 들어, 1 내지 50의 범위는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 또는 50으로 이루어진 군으로부터의 임의의 수, 수의 조합, 또는 하위-범위를 포함하는 것으로 이해된다.

[0084] "감소시키다"란 적어도 10%, 25%, 50%, 75%, 또는 100%의 음성 변경을 의미한다.

[0085] "참조물"이란 표준 또는 대조군 상태를 의미한다.

[0086] "참조 서열"은 서열 비교를 위한 기초로서 사용되는 한정된 서열이다. 참조 서열은 명시된 서열의 하위세트 또는 그의 전체; 예를 들어, 전장 cDNA 또는 유전자 서열의 절편, 또는 완전한 cDNA 또는 유전자 서열일 수 있다. 폴리펩티드에 대해, 참조 폴리펩티드 서열의 길이는 일반적으로 적어도 약 16개 아미노산, 바람직하게는 적어도 약 20개 아미노산, 보다 바람직하게는 적어도 약 25개 아미노산, 및 보다 더 바람직하게는 약 35개 아미노산, 약 50개 아미노산, 또는 약 100개 아미노산일 것이다. 핵산에 대해, 참조 핵산 서열의 길이는 일반적으로 적어도 약 50개 뉴클레오티드, 바람직하게는 적어도 약 60개 뉴클레오티드, 보다 바람직하게는 적어도 약 75개 뉴클레오티드, 및 보다 더 바람직하게는 약 100개 뉴클레오티드 또는 약 300개 뉴클레오티드 또는 그 부근의 또는 그 사이의 임의의 정수일 것이다.

[0087] 본 발명의 방법에 유용한 핵산 분자는 본 발명의 폴리펩티드 또는 그의 단편을 코딩하는 임의의 핵산 분자를 포함한다. 이러한 핵산 분자는 내인성 핵산 서열과 100% 동일할 필요는 없지만, 전형적으로 실질적 동일성을 나타낼 것이다. 내인성 서열과 "실질적 동일성"을 갖는 폴리뉴클레오티드는 전형적으로 이중-가닥 핵산 분자의 적어도 한 가닥과 혼성화할 수 있다. 본 발명의 방법에 유용한 핵산 분자는 본 발명의 폴리펩티드 또는 그의 단편을 코딩하는 임의의 핵산 분자를 포함한다. 이러한 핵산 분자는 내인성 핵산 서열과 100% 동일할 필요는 없지만, 전형적으로 실질적 동일성을 나타낼 것이다. 내인성 서열과 "실질적 동일성"을 갖는 폴리뉴클레오티드는 전형적으로 이중-가닥 핵산 분자의 적어도 한 가닥과 혼성화할 수 있다. "혼성화하다"란 다양한 엄격성의 조건 하에서 상보성 폴리뉴클레오티드 서열 (예를 들어, 본원에 기재된 유전자), 또는 그의 부분 사이에 이중-가닥 분자를 형성하는 쌍을 의미한다. (예를 들어, 문헌 [Wahl, G. M. and S. L. Berger (1987) Methods Enzymol. 152:399]; [Kimmel, A. R. (1987) Methods Enzymol. 152:507] 참조).

[0088] 예를 들어, 엄격한 염 농도는 통상적으로 약 750 mM NaCl 및 75 mM 시트르산삼나트륨 미만, 바람직하게는 약 500 mM NaCl 및 50 mM 시트르산삼나트륨 미만, 및 보다 바람직하게는 약 250 mM NaCl 및 25 mM 시트르산삼나트륨 미만일 것이다. 낮은 엄격성 혼성화는 유기 용매, 예를 들어, 포름아미드의 부재 하에서 얻어질 수 있는 반면, 높은 엄격성 혼성화는 적어도 약 35% 포름아미드, 및 보다 바람직하게는 적어도 약 50% 포름아미드의 존재 하에서 얻어질 수 있다. 엄격한 온도 조건은 통상적으로 적어도 약 30℃의, 보다 바람직하게는 적어도 약 37℃의, 및 가장 바람직하게는 적어도 약 42℃의 온도를 포함할 것이다. 추가의 파라미터, 예컨대 혼성화 시간, 세정제, 예를 들어, 나트륨 도데실 술페이트 (SDS)의 농도, 및 운반체 DNA의 포함 또는 배제를 다양화하는 것은 관련 기술분야의 통상의 기술자에게 널리 공지되어 있다. 다양한 수준의 엄격성은 필요에 따라 이들 다양한 조건을 조합함으로써 달성된다. 바람직한: 실시양태에서, 혼성화는 30℃에서 750 mM NaCl, 75 mM 시트르산삼나트륨, 및 1% SDS에서 일어날 것이다. 보다 바람직한 실시양태에서, 혼성화는 37℃에서 500 mM NaCl, 50 mM 시트르산삼나트륨, 1% SDS, 35% 포름아미드, 및 100 μg/ml 변성된 연어 정자 DNA (ssDNA)에서 일어날 것이다. 가장 바람직한 실시양태에서, 혼성화는 42℃에서 250 mM NaCl, 25 mM 시트르산삼나트륨, 1% SDS, 50% 포름아미드, 및 200 μg/ml ssDNA에서 일어날 것이다. 이들 조건에 대한 유용한 변화는 관련 기술분야의 통상의 기술자에게 용이하게 명백할 것이다.

[0089] 대부분의 적용을 위해, 혼성화에 이어지는 세척 단계는 또한 엄격성에 있어서 다양할 것이다. 세척 엄격성 조

건은 염 농도에 의해 및 온도에 의해 정의될 수 있다. 상기와 같이, 세척 엄격성은 염 농도를 감소시킴으로써 또는 온도를 증가시킴으로써 증가될 수 있다. 예를 들어, 세척 단계를 위한 엄격한 염 농도는 바람직하게는 약 30 mM NaCl 및 3 mM 시트르산삼나트륨 미만, 및 가장 바람직하게는 약 15 mM NaCl 및 1.5 mM 시트르산삼나트륨 미만일 것이다. 세척 단계를 위한 엄격한 온도 조건은 통상적으로 적어도 약 25℃의, 보다 바람직하게는 적어도 약 42℃의, 및 보다 더 바람직하게는 적어도 약 68℃의 온도를 포함할 것이다. 바람직한 실시양태에서, 세척 단계는 25℃에서 30 mM NaCl, 3 mM 시트르산삼나트륨, 및 0.1% SDS에서 일어날 것이다. 보다 바람직한 실시양태에서, 세척 단계는 42℃에서 15 mM NaCl, 1.5 mM 시트르산삼나트륨, 및 0.1% SDS에서 일어날 것이다. 보다 바람직한 실시양태에서, 세척 단계는 68℃에서 15 mM NaCl, 1.5 mM 시트르산삼나트륨, 및 0.1% SDS에서 일어날 것이다. 이들 조건에 대한 추가의 변화는 관련 기술분야의 통상의 기술자에게 용이하게 명백할 것이다. 혼성화 기법은 관련 기술분야의 통상의 기술자에게 널리 공지되어 있으며, 예를 들어, 문헌 [Benton and Davis (Science 196:180, 1977)]; [Grunstein and Hogness (Proc. Natl. Acad. Sci., USA 72:3961, 1975)]; [Ausubel et al. (Current Protocols in Molecular Biology, Wiley Interscience, New York, 2001)]; [Berger and Kimmel (Guide to Molecular Cloning Techniques, 1987, Academic Press, New York)]; 및 [Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York]에 기재되어 있다.

[0090] "실질적으로 동일한"이란 참조 아미노산 서열 (예를 들어, 본원에 기재된 아미노산 서열 중 어느 하나) 또는 핵산 서열 (예를 들어, 본원에 기재된 핵산 서열 중 어느 하나)과 적어도 50% 동일성을 나타내는 폴리펩티드 또는 핵산 분자를 의미한다. 바람직하게는, 이러한 서열은 비교에 사용되는 서열과 아미노산 수준 또는 핵산에서 적어도 60%, 보다 바람직하게는 80% 또는 85%, 및 보다 바람직하게는 90%, 95% 또는 심지어 99% 동일하다.

[0091] 서열 동일성은 전형적으로 서열 분석 소프트웨어 (예를 들어, 미국 53705 위스콘신주 매디슨 유니버시티 애브뉴 1710에 소재하는 유니버시티 오브 위스콘슨 바이오테크놀로지 센터 (University of Wisconsin Biotechnology Center), 제네틱스 컴퓨터 그룹 (Genetics Computer Group)의 서열 분석 소프트웨어 패키지 (Sequence 분석 Software Package), BLAST, BESTFIT, GAP, 또는 PILEUP/PRETTYBOX 프로그램)를 사용하여 측정된다. 이러한 소프트웨어는 상동성의 정도를 다양한 치환, 결실, 및/또는 다른 변형에 할당함으로써 동일하거나 유사한 서열을 매칭시킨다. 보존적 치환은 전형적으로 하기 군 내의 치환을 포함한다: 글리신, 알라닌; 발린, 이소류신, 류신; 아스파르트산, 글루탐산, 아스파라긴, 글루타민; 세린, 트레오닌; 리신, 아르기닌; 및 페닐알라닌, 티로신. 동일성의 정도를 측정하는 예시적인 접근법에서, BLAST 프로그램이 사용될 수 있으며, e^{-3} 내지 e^{-100}의 확률 점수는 가깝게 관련된 서열을 지시한다.

[0092] "체세포 핵 전달" 또는 "SCNT"란 제핵된 난모세포 내로의 체세포로부터의 공여자 핵의 전달을 의미한다. 상기 프로세스는 생식적 또는 치료적 클로닝 중 어느 하나에 사용될 수 있으며, 체세포와 제핵된 난모세포의 융합, 제핵된 난모세포 내로의 핵의 주사에 의해, 또는 임의의 다른 방법에 의해 달성될 수 있다.

[0093] 체세포의 핵은 유전 정보를 제공하는 반면, 난모세포는 배아의 발생에 필요한 영양분 및 다른 에너지-생성 물질을 제공한다. 융합이 일어나면, 세포는 전능성이고, 결국 배반포로 발달하며, 이 시점에서 내부 세포 덩이가 단리된다.

[0094] 본원에 사용된 바와 같은 용어 "핵 전달"은 제핵된 난모세포를 체세포의 세포 핵 유전 물질 또는 핵과 인공적으로 조합함으로써 획득된 동일한 특징 및 품질을 허용하는 유전자 조작 기법을 지칭한다. 일부 실시양태에서, 핵 전달 절차는 공여자 체세포로부터의 핵 또는 핵 유전 물질이 제핵된 난 또는 난모세포 (핵/전핵이 제거된 난 또는 난모세포) 내로 전달되는 경우이다. 공여자 핵은 체세포로부터 올 수 있다.

[0095] 용어 "핵 유전 물질"은 개체에 관한 정보를 코딩하는 폴리뉴클레오티드 (예를 들어, DNA)를 포함하는 핵에서 발견되는 구조 및/또는 분자를 지칭한다. 핵 유전 물질은 염색체 및 염색질을 포함한다. 상기 용어는 또한 세포 분열, 예컨대 모 세포의 딸 세포로의 분열에 의해 생산된 핵 유전 물질 (예를 들어, 염색체)을 지칭한다. 핵 유전 물질은 미토콘드리아 DNA를 포함하지 않는다.

[0096] 용어 "SCNT 배아"는 체세포의 핵 또는 핵 유전 물질과 융합된 제핵된 난모세포의 세포, 또는 그의 전능성 자손을 지칭한다. SCNT 배아는 배반포로 발달하고, 착상후에 살아 있는 후손으로 발달할 수 있다. SCNT 배아는 1-세포 배아, 2-세포 배아, 4-세포 배아, 또는 배반포가 되기 전의 임의의 단계의 배아일 수 있다.

[0097] 용어 "공여자 인간 세포" 또는 "공여자 인간 체세포"는 핵 수용체 또는 수용자로서 수용자 난모세포 내로 전달

된 인간 세포의 체세포 또는 핵을 지칭한다.

[0098] 용어 "체세포"는 생식 세포 또는 생식 세포 전구체가 아닌 식물 또는 동물 세포를 지칭한다. 일부 실시양태에서, 분화된 세포는 배 세포가 아니다. 체세포는 다능성 또는 전능성 세포와 관련되지 않는다. 일부 실시양태에서, 체세포는 "비-배아 체세포"이며, 이는 배아에 존재하거나 그로부터 얻어지지 않으며, 시험관내에서 이러한 세포의 증식으로부터 발생되지 않는 체세포를 의미한다. 일부 실시양태에서, 체세포는 "성체 체세포"이며, 이는 배아 또는 태아 이외의 유기체에 존재하거나 그로부터 얻어지거나, 시험관내에서 이러한 세포의 증식으로부터 발생되는 세포를 의미한다.

[0099] 본원에 사용된 바와 같은 용어 "난모세포"는 감수분열의 중기 II에 도달한 성숙한 난모세포를 지칭한다. 난모세포는 또한 생식에 관여하는 자성 배우자 또는 배 세포를 기재하는 데 사용되며, 통상적으로 또한 난으로 지칭된다. 성숙한 난은 모계 염색체의 단일 세트 (23개, 인간 영장류에서 X)를 가지며, 중기 II에서 중단된다.

[0100] "하이브리드 난모세포"는 제1 인간 난모세포 ("수용자"로 용어화됨)로부터의 세포질을 갖지만, 수용자 난모세포의 핵 유전 물질을 갖지 않는 제핵된 난모세포를 지칭하며; 이는 "공여자"로 용어화되는 또다른 인간 세포로부터의 핵 유전 물질을 갖는다. 일부 실시양태에서, 하이브리드 난모세포는 또한 수용자 난모세포로부터의 것이 아니지만, 공여자 세포로부터의 것인 (이는 핵 유전 물질과 동일한 공여자 세포, 또는 상이한 공여자로부터의, 예를 들어 공여자 난모세포로부터의 것일 수 있음) 미토콘드리아 DNA (mtDNA)를 포함할 수 있다.

[0101] 본원에 사용된 바와 같은 용어 "제핵된 난모세포"는 그의 핵이 제거된 인간 난모세포를 지칭한다.

[0102] 본원에 사용된 바와 같은 용어 "제핵"은 세포의 핵 물질이 제거되어, 세포질만 남기는 프로세스를 지칭한다. 난에 적용되는 경우, 제핵은 핵막에 의해 둘러싸이지 않은 모계 염색체의 제거를 지칭한다. 용어 "제핵된 난모세포"는 핵 물질 또는 핵이 제거된 난모세포를 지칭한다.

[0103] 본원에 사용된 바와 같은 "수용자 인간 난모세포"는 그의 원래 핵을 제거한 후에 인간 핵 공여자 세포로부터 핵을 받는 인간 난모세포를 지칭한다.

[0104] 본원에 사용된 바와 같은 용어 "융합"은 핵 공여자 세포 및 수용자 난모세포의 지질 막의 조합을 지칭한다. 예를 들어, 지질 막은 세포의 형질막 또는 핵막일 수 있다. 융합은 핵 공여자 세포 및 수용자 난모세포 사이의 전기적 자극의 인가 시 일어날 수 있는데, 이 때 이들은 서로 인접하게 위치하거나, 핵 공여자 세포는 수용자 난모세포의 난황주위 공간에 위치한다.

[0105] 본원에 사용된 바와 같은 용어 "살아 있는 후손"은 자궁 외에서 생존할 수 있는 동물을 의미한다. 바람직하게는, 이는 1초, 1분, 1일, 1주, 1개월, 6개월 또는 1년 초과 동안 생존할 수 있는 동물이다. 동물은 생존을 위해 자궁내 환경을 요구하지 않을 수 있다.

[0106] 용어 "출생전"은 출생 전에 존재하거나 일어나는 것을 지칭한다. 유사하게, 용어 "출생후"는 출생 후에 존재하거나 일어나는 것이다.

[0107] 본원에 사용된 바와 같은 용어 "배반포"는 약 30 내지 150개의 세포의 태반 포유동물에서의 착상전 배아 (마우스에서 수정 후 약 3일, 인간에서 수정 후 약 5일)를 지칭한다. 배반포 단계는 상실배 단계에 이어지며, 그의 고유한 형태에 의해 구별될 수 있다. 배반포는 세포의 층 (영양외배엽), 유체-충전된 강 (포배공간 또는 배반포강), 및 내부 상의 세포의 클러스터 (내부 세포 덩이, 또는 ICM)로 구성된 구체로 이루어진다. 비분화된 세포로 이루어진 ICM은 배반포가 자궁에 착상되는 경우 태아가 될 것을 발생시킨다. 이들 동일한 ICM 세포는, 배양에서 성장되는 경우, 배아 줄기 세포주를 발생시킬 수 있다. 착상 시, 마우스 배반포는 약 70개의 영양막 세포 및 30개의 ICM 세포로 구성된다.

[0108] 본원에 사용된 바와 같은 용어 "포배"는 포배공간으로 지칭되는 유체-충전된 강을 둘러싸는 세포의 중공 구체로 이루어진 배아의 발생에서의 초기 단계를 지칭한다. 용어 포배는 때때로 배반포와 상호교환가능하게 사용된다.

[0109] 용어 "할구"는 전반에 걸쳐 착상전 배아로부터 얻어진 적어도 1개의 할구 (예를 들어, 1개, 2개, 3개, 4개 등)를 지칭하기 위해 사용된다. 용어 "2개 이상의 할구의 클러스터"는 할구의 시험관내 배양 동안 생성되는 세포를 지칭하기 위해 "할구-유래된 증식물"과 상호교환가능하게 사용된다. 예를 들어, 할구가 SCNT 배아로부터 얻어지고, 초기에 배양된 후, 이는 일반적으로 적어도 1회 분열하여 2개 이상의 할구의 클러스터 (또한 할구-유래된 증식물로 공지됨)를 생성한다. 클러스터는 배아 또는 태아 세포와 함께 더 배양될 수 있다. 궁극적으로, 할구-유래된 증식물은 분열하기를 계속할 것이다. 이들 구조로부터, ES 세포, 전능성 줄기 (TS) 세포, 및 부분

적으로 분화된 세포 유형이 배양 방법의 과정에 걸쳐 발달할 것이다.

[0110] 본원에 사용된 바와 같은 용어 "클로닝된 (또는 클로닝)"은 또다른 개체 단위와 동일한 유전자 세트를 갖는 새로운 개체 단위를 제조하기 위한 유전자 조작 기법을 지칭한다. 본 발명에서, 본원에 사용된 바와 같은 용어 "클로닝된"은 또다른 세포, 배아 세포, 태아 세포, 분화된 세포, 및/또는 동물 세포의 핵 DNA 서열과 실질적으로 유사하거나 동일한 핵 DNA 서열을 갖는 세포, 배아 세포, 태아 세포, 및/또는 동물 세포를 지칭한다. 용어 "실질적으로 유사한" 및 "동일한"은 본원에서 기재된다. 클로닝된 SCNT 배아는 하나의 핵 전달로부터 발생할 수 있거나, 대안적으로, 클로닝된 SCNT 배아는 적어도 1회의 재-클로닝 단계를 포함하는 클로닝 프로세스로부터 발생할 수 있다.

[0111] 본원에 사용된 바와 같은 용어 "트랜스제닉 유기체"는 또다른 유기체로부터의 유전 물질이 실험적으로 전달되어, 숙주가 그의 염색체 조성에서 전달된 유전자의 유전적 소질을 획득하는 유기체를 지칭한다.

[0112] 본원에 개시된 바와 같은 SCNT 배아와 관련하여 본원에 사용된 바와 같은 용어 "착상"은 본원에 기재된 SCNT 배아를 갖는 대리 암컷 동물을 수태시키는 것을 지칭한다. 이 기법은 관련 기술분야의 통상의 기술자에게 널리 공지되어 있다. 예를 들어, 문헌 [Seidel and Elsden, 1997, Embryo Transfer in Dairy Cattle, W. D. Hoard & Sons, Co., Hoards Dairyman]을 참조한다. 배아는 자궁내에서 발달하도록 허용될 수 있거나, 대안적으로, 태아는 분만 전에 자궁 환경으로부터 제거될 수 있다.

[0113] "대상체"란 인간 또는 비-인간 포유동물을 포함하나 이에 제한되지는 않는 포유동물, 예컨대 농업적으로 중요한 포유동물 (예를 들어, 소, 말, 양, 돼지), 애완동물 (예를 들어, 개, 고양이), 또는 희귀하거나 멸종위기 포유동물 (예를 들어, 팬더)을 의미한다.

[0114] 본원에 사용된 바와 같은 용어 "치료하다", "치료하는", "치료" 등은 그와 연관된 장애 및/또는 증상을 감소시키거나 개선시키는 것을 지칭한다. 불가능하지는 않지만, 장애 또는 상태를 치료하는 것은 장애, 상태 또는 그와 연관된 증상이 완전히 제거되는 것을 요구하지는 않음이 인정될 것이다.

[0115] 구체적으로 언급되거나 맥락으로부터 명백하지 않다면, 본원에 사용된 바와 같은 용어 "또는"은 포함적인 것으로 이해된다. 구체적으로 언급되거나 맥락으로부터 명백하지 않다면, 본원에 사용된 바와 같은 단수 형태 용어는 단수 또는 복수 (즉, 적어도 하나)인 것으로 이해된다. 예로서, "요소"는 1개의 요소 또는 1개 초과의 요소를 의미한다.

[0116] 구체적으로 언급되거나 맥락으로부터 명백하지 않다면, 본원에 사용된 바와 같은 용어 "약"은 관련 기술분야에서 정상적인 내성의 범위 내, 예를 들어 평균의 2 표준 편차 내로서 이해된다. 약은 언급된 값의 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.05%, 또는 0.01% 내로서 이해될 수 있다. 맥락으로부터 달리 명백하지 않다면, 본원에서 제공된 모든 수치 값은 용어 약에 의해 변형된다.

[0117] 본원에서 변수의 임의의 정의에서 화학기의 목록의 나열은 열거된 기의 임의의 단일 기 또는 조합으로서의 그 변수의 정의를 포함한다. 본원에서 변수 또는 측면에 대한 실시양태의 나열은 임의의 단일 실시양태로서 또는 임의의 다른 실시양태 또는 그의 부분과의 조합으로의 그 실시양태를 포함한다.

[0118] 본원에서 제공된 임의의 조성물 및 방법은 본원에서 제공된 다른 조성물 및 방법 중 임의의 것 중 1종 이상과 조합될 수 있다.

도면의 간단한 설명

[0119] 도 1a 내지 1f는 *Xist* KO 공여자 세포 및 *Kdm4d* mRNA 주사의 조합된 사용이 SCNT 배아의 발생 잠재성을 완전히 복원하지 않음을 나타낸다.

도 1a는 항-H3K27me3, 항-Cdx2, 항-Oct4 항체 및 DAPI로 염색된 IVF 및 SCNT 배반포의 대표적인 화상을 포함한다. 화살표는 이소적으로 불활성화된 X 염색체를 나타내는 점상 H3K27me3 신호를 지시한다. 이소적 XCI는 *Kdm4d* mRNA 주사와 무관하게 관찰될 수 있음을 주목한다. 스케일 바, 50 μm.

도 1b는 IVF 및 SCNT 배반포에서 점상 H3K27me3 신호 (불활성화된 X 염색체를 나타냄)를 갖는 또는 갖지 않는 세포의 비를 나타내는 막대 그래프를 제공한다. 각각의 열은 단일 배반포를 나타낸다.

도 1c는 E19.5에서의 제왕 절개에 의한 조사된 SCNT 배아의 새끼 비율을 나타내는 막대 그래프를 제공한다. *Xist* KO 공여자 세포와 *Kdm4d* mRNA 주사를 사용하는 것의 조합은 공여자로서 난구 세포, 세르톨리 세포 및 MEF

세포로의 SCNT 배아의 만기 비율을 상가적으로 개선시킴을 주목한다.

도 1d는 *Kdm4d* mRNA 주사와 조합된 *Xist* KO 세르톨리 세포를 사용하여 SCNT에 의해 유래된 성체 수컷 마우스, 및 야생형 암컷과의 자연 교배를 통해 생성된 그의 새끼의 화상을 나타낸다.

도 1e는 E19.5에서의 제왕 절개에 의한 조사된 태반의 중량을 나타내는 박스 플롯을 제공한다. 수염모양은 최대값 및 최소값을 나타낸다. ***p < 0.001. ns, 유의하지 않음.

도 1f는 과요오드산-쉬프 (Periodic acid-Schiff) (PAS: 우측)로 염색된 만기 태반의 조직학적 섹션의 대표적인 화상을 제공한다. PAS-양성 해면영양막 층은 공여자 세포에서의 *Xist* 대립유전자의 유전자형과 무관하게 SCNT 태반에서 미로 층 내로 침습되었음을 주목한다. 스케일 바, 1 mm.

도 2a 내지 2c는 SCNT 배아의 착상후 발생 정지를 나타낸다.

도 2a는 지시된 시점에서 Kdm4d mRNA 주사와 조합된 Xist KO MEF 세포를 사용하여 생성된 SCNT 배아의 발생률을 나타내는 막대 그래프를 제공한다.

도 2b는 E4.5에서 수집된 SCNT 배아의 화상이다.

도 2c는 E10.5에서 수집된 SCNT 배아의 화상이다. SCNT 배아는 각각의 단계에서 배아/신체 크기의 큰 변이를 나타냄을 주목한다. 스케일 바, (도 2b)에서 100 μm 및 (도 2c)에서 1 mm.

도 3a 내지 d는 SCNT 배반포에서의 DNA 메틸화의 광범위한 재프로그래밍을 나타낸다.

도 3a는 실험적 접근법의 개략적 예시이다. IVF 또는 SCNT (*Xist* KO 공여자 및 *Kdm4d* 주사의 조합)에 의해 생성된 배반포를 전체-게놈 비술파이트 시퀀싱 (WGBS) 및 RNA-seq에 사용하였다.

도 3b는 SCNT 및 IVF 배반포, 뿐만 아니라 MEF, 접합자, 정자 및 난모세포의 게놈에 걸쳐 모두 커버된 CpG의 DNA 메틸화 수준을 비교하는 박스 플롯을 포함한다. 박스에서 두꺼운 선은 중위값을 지시하고, 십자표는 평균을 나타낸다. 수염모양은 제2.5 및 제97.5 백분위수를 나타낸다. Sp+Oo는 정자 및 난모세포의 평균 값을 나타낸다. MEF, 정자 및 난모세포의 WGBS 데이터세트는 GSE56151 및 GSE56697로부터 얻었다.

도 3c는 각각의 샘플 사이의 DNA 메틸화 수준을 비교하는 플롯이다. 심하게 메틸화된 공여자 MEF 세포 게놈은 SCNT에 의해 전반적으로 재프로그래밍되어 IVF 배반포의 그것과 유사한 DNA 메틸화 프로파일을 발생시킴을 주목한다.

도 3d는 IVF 및 SCNT 배반포의 유전자 발현 프로파일을 비교하는 산포도이다. SCNT 배아에서의 상향-조절된 유전자 (n=37 (어두운 색 클러스터); 배수 변화 (FC) > 3.0) 및 하향-조절된 유전자 (n=55, 더 밝은 색 클러스터; FC > 3.0).

도 4a 및 4b는 SCNT 및 IVF 배반포가 유사한 DNA 메틸롬 및 전사체를 가짐을 나타낸다.

도 4a는 IVF 및 SCNT 배반포에서의 반복부를 비롯한 다양한 게놈 특색에서의 평균 메틸화 수준을 비교하는 막대 그래프를 제공한다.

도 4b는 IVF 및 SCNT 배반포의 생물학적 반복실험의 전사체를 비교하는 산포도를 포함한다.

도 5a 내지 5h는 SCNT 배반포에서의 차등적으로 메틸화된 영역 (DMR)의 확인 및 특징규명을 나타낸다.

도 5a는 과다- 및 과소-DMR에서의 SCNT 및 IVF 배반포의 DNA 메틸화 수준을 나타내는 박스 플롯을 나타낸다. 박스에서 두꺼운 선은 중위값을 지시하고, 십자표는 평균을 나타낸다. DMR의 수가 또한 지시된다.

도 5b는 과다- 및 과소-DMR의 길이를 비교하는 박스 플롯을 포함한다.

도 5c는 게놈에서의 과다- 및 과소-DMR의 파이 차트 분포이다.

도 5d는 그들의 플랭킹 영역에 비해 과소DMR에서의 지시된 샘플의 평균 DNA 메틸화 수준을 나타내는 그래프이다.

도 5e는 그들의 플랭킹 영역에 비해 과소DMR에서의 IVF 및 SCNT 배반포의 부계 (Pat) 및 모계 (Mat) 대립유전자-특이적 DNA 메틸화 수준을 나타내는 그래프이다.

도 5f는 그들의 플랭킹 영역에 비해 과소DMR에서의 지시된 발생 단계에서의 IVF 및 SCNT 배아의 부계 및 모계

대립유전자-특이적 DNA 메틸화 수준을 나타내는 그래프이다.

도 5g는 그들의 플랭킹 영역에 비해 과다DMR에서의 지시된 샘플의 평균 DNA 메틸화 수준을 나타내는 그래프이다.

도 5h는 그들의 플랭킹 영역에 비해 과다DMR에서의 지시된 샘플의 평균 DNA 메틸화 수준을 나타내는 그래프이다. 사용된 데이터세트는 GSE11034로부터의 것이었다.

도 6a 내지 6d는 SCNT 배반포에서의 과소- 및 과다-DMR의 특색을 제공한다.

도 6a는 과다- 및 과소-DMR의 대표적인 게놈 브라우저 보기이다.

도 6b는 난모세포에서의 메틸화 피크가 IVF 배반포에서의 그것들과 중첩함을 나타내는 대표적인 게놈 브라우저 보기이다.

도 6c는 과다DMR-연관된 유전자의 유전자 존재론 분석이다.

도 6d는 PGC 발생 동안 과다DMR에서의 평균 메틸화 (5mC) 및 히드록시메틸화 (5hmC) 수준을 나타내는 피크 플롯을 포함한다.

도 7a 내지 d는 SCNT 배반포에서의 H3K27me3-의존성 각인의 소실을 나타낸다.

도 7a는 SCNT 배반포에서의 H3K27me3-각인된 유전자의 상대 유전자 발현 수준을 나타내는 막대 그래프를 제공한다. 신뢰성 있게 검출가능한 수준 (백만개 맵핑된 단편 당 엑손의 킬로베이스 당 단편 (FPKM) > 1)으로 IVF 배반포에서 발현된 26개의 유전자가 나타내어진다. IVF 배반포의 발현 수준은 1로 설정되었다. 유전자를 IVF 배반포에서의 그것에 비해 SCNT에서의 발현 변화에 의해 상향, 하향 및 비변화됨으로 분류하였다 (FC >1.5).

도 7b는 IVF 및 SCNT 배반포에서의 H3K27me3-각인된 유전자의 대립유전자 발현의 비 (Pat/Mat)를 나타내는 막대 그래프를 제공한다. 26개의 발현된 유전자 중에서 (FPKM >1), 어느 샘플에서든 10개 초과의 SNP 판독물을 갖는 17개의 유전자가 나타내어진다. 별표는 부계 대립유전자로 100% 편향됨을 나타낸다. 모든 17개의 유전자는 SCNT 배반포에서 그들의 부계 대립유전자 편향을 소실하였음을 주목한다.

도 7c는 2개의 대표적인 H3K27me3-각인된 유전자에서의 H3K27me3 ChIP-seq 신호의 게놈 브라우저 보기를 나타낸다.

도 7d는 3 Mb 플랭킹 영역에 비해 76개의 H3K27me3-각인된 유전자에서의 다양한 세포 유형 (난모세포, 정자, MEF, ESC) 및 조직의 평균 H3K27me3 ChIP-seq 강도를 나타낸다.

도 8a 내지 8f는 공지된 126개의 각인된 유전자 및 그들의 공지된 ICR의 각인 상태를 예시한다.

도 8a는 SCNT 배반포에서의 23개의 공지된 각인 제어 영역 (ICR)의 상대 DNA 메틸화 수준을 나타내는 막대 그래프를 제공한다. IVF 배반포의 메틸화 수준은 1로 설정되었다. 파선은 IVF 배반포 메틸화 수준의 50%를 지시한다. 23개의 ICR 중 21개는 SCNT 배반포에서 IVF 메틸화 수준의 그것의 적어도 50%를 유지하였지만, *S1c38a4* 및 *Snrpn* ICR (적색으로 표시됨)은 IVF 수준의 그것의 50% 미만을 나타내었음을 주목한다.

도 8b는 충분한 대립유전자-특이적 메틸화 정보 (IVF 및 SCNT 배반포 둘 다의 둘 다의 대립유전자에서 5 초과의 검출된 CpG)를 갖는 20개의 ICR에서의 DNA 메틸화의 대립유전자 편향을 나타내는 막대 그래프를 제공한다. 모든 20개의 ICR은 SCNT 배반포에서 대립유전자 편향된 DNA 메틸화를 유지하였음을 주목한다.

도 8c는 SCNT 배반포에서의 공지된 각인된 유전자의 상대 유전자 발현 수준을 나타내는 막대 그래프를 제공한다. IVF 배반포에서 신뢰성 있게 검출가능한 45개의 각인된 유전자가 나타내어진다 (FPKM > 1). IVF 배반포의 발현 수준은 1로 설정되었다. 유전자를 IVF 배아에 비해 SCNT 배아에서의 그들의 발현 수준에 기초하여 상향, 하향, 및 비변화됨으로 분류하였다 (FC >1.5).

도 8d는 IVF 및 SCNT 배반포에서의 공지된 각인된 유전자의 대립유전자 발현의 비 (Mat/Pat)를 나타내는 막대 그래프를 제공한다. IVF 배반포에서 충분한 SNP 추적된 판독물 (어느 샘플에서든 FPKM>1, 평균 SNP 판독물 > 10)을 갖는 신뢰성 있게 검출가능한 수준으로 발현된 6개의 모계적으로 발현된 유전자 (MEG; Mat/Pat > 2.0)가 나타내어진다. 별표는 모계 대립유전자에 대한 100% 편향을 나타낸다. 모든 6개의 MEG는 SCNT 배반포에서 모계 대립유전자 편향을 유지하였음을 주목한다.

도 8e는 IVF 및 SCNT 배반포에서의 공지된 각인된 유전자의 대립유전자 발현의 비 (Pat/Mat)를 나타내는 막대

그래프를 제공한다. IVF 배반포에서 충분한 SNP 추적된 판독물 (어느 샘플에서든 FPKM>1, 평균 SNP 판독물 > 10)을 갖는 신뢰성 있게 검출가능한 수준으로 발현된 13개의 부계적으로 발현된 유전자 (PEG; Pat/Mat > 2.0)가 나타내어진다. 별표는 부계 대립유전자에 대한 100% 편향을 나타낸다. 화살표는 SCNT 배반포에서 부계 편향된 발현을 소실한 유전자를 지시한다. Slc38a4, Sfmbt2, Phf17, 및 Gab1은 H3K27me3-의존성 각인된 유전자이다.

도 8f는 비-정규 각인된 유전자에서의 H3K27me3 ChIP-seq 신호의 대표적인 게놈 브라우저 보기를 제시한다.

발명을 실시하기 위한 구체적인 내용

[0120] **상세한 설명**

[0121] 본 발명은 클로닝 효율을 개선시키는 방법을 제공한다. 특정 실시양태에서, 본 발명은 *Kdm4d* 과발현이 *Xist* 넉아웃 공여자 세포인 것을 포함하는 체세포 핵 전달 효율을 개선시키는 방법을 제공한다.

[0122] 본 발명은 적어도 부분적으로, *Kdm4d* mRNA 주사와 결합된 *Xist* 넉아웃 공여자 세포가 체세포 핵 전달 효율을 개선시킬 수 있다는 발견에 기초한다. 이 조합된 접근법은 분화된 체세포 공여자 세포를 사용한 마우스 클로닝에서 지금껏 보고된 가장 높은 효율을 발생시켰다. 그러나, SCNT 배아 중 많은 것은 여전히 착상후 발생 정지를 나타내고, 생존한 배아는 비정상적으로 큰 태반을 가지며, 이는 일부 재프로그래밍 결함이 여전히 존속함을 시사한다. 비교적 메틸롬 및 전사체 분석은 SCNT 배반포 배아에서의 비정상적인 DNA 메틸화 및 H3K27me3-의존성 각인의 소실을 밝혀내었으며, 이는 관찰된 발생적 결함의 원인일 가능성이 있다.

[0123] **H3K27me3은 DNA 메틸화-독립성 각인 메커니즘이다**

[0124] 포유동물 정자 및 난모세포는 상이한 후생적 경관을 가지며, 상이한 방식으로 조직화된다. 수정 후, 초기에 별개의 모 에피게놈은 각인 제어 영역 (ICR)을 비롯한 특정 좌위를 제외하고는 매우 동등화된다. 어떻게 모 염색질이 동등화되고 어떻게 ICR이 이 재프로그래밍으로부터 탈출하는 지는 거의 잘 알려져 있지 않다. 여기서 모 대립유전자-특이적 DN아제 I 과민성 부위 (DHS)를 마우스 접합자 및 상실배 배아에서 특징규명하였으며, 대립유전자 DHS 기저에 있는 후생적 메커니즘을 조사하였다. DNA 메틸롬 및 H3K27me3 ChIP-seq 데이터 세트의 통합된 분석은 DNA 메틸화가 없지만, 모계 대립유전자-특이적 H3K27me3을 가진 부계 대립유전자-특이적 DHS를 갖는 76개의 유전자 (표 1)를 밝혀내었다.

[0125] <표 1> H3K27me3-의존성 각인된 유전자

유전자_명칭	유전자_chr	유전자_시작	유전자_종료
Rbp2	chr9	98390956	98410190
Runx1	chr16	92601711	92826311
Sfmbt2	chr2	10292078	10516880
Slc38a2	chr15	96517823	96530129
Slc38a4	chr15	96825254	96886387
Gramd1b	chr9	40105492	40263349
Bbx	chr16	50191957	50432502
Sox21	chr14	118632456	118636252
Mbnl2	chr14	120674891	120830920
Prdm11	chr2	92815063	92886301
1700067G17Rik	chr1	90912688	90918785
1700095B10Rik	chr5	113222312	113230721
Mir692-2b	chr4	125181992	125182101
Sh3gl3	chr7	89319728	89455927
Etv6	chr6	133985725	134220165
Tle3	chr9	61220173	61266304
Hunk	chr16	90386642	90499798
Gab1	chr8	83288333	83404378
Matn1	chr4	130500300	130511391
Chst1	chr2	92439864	92455409
Clic6	chr16	92498392	92541486
1700110K17Rik	chr9	40141057	40150922
Foxl1	chr8	123651585	123654544
Mir6241	chr14	118657855	118657958
Otog	chr7	53496357	53566804
1700017J07Rik	chr2	168803769	168804406
4930404H11Rik	chr12	72641594	72657120
Gm5086	chr13	98329955	98353949
Tshz2	chr2	169459146	169714004
Bmp7	chr2	172695189	172765794
G730013B05Rik	chr16	50526358	50559572
Rftn1	chr17	50132632	50329822
C430002E04Rik	chr3	41291603	41297121
Myoz2	chr3	122709124	122737905
Six3os1	chr17	86001272	86017736
Slc38a1	chr15	96401849	96473344
Rbms1	chr2	60590010	60801261
Flt1	chr5	148373772	148537564
Sall3	chr18	81163113	81183317

[0126]

Otx2os1	chr14	49288963	49413023
1700006F04Rik	chr14	120148449	120150786
2300005B03Rik	chr15	74573269	74577117
4931430N09Rik	chr6	118830176	118835561
Gas7	chr11	67346500	67502494
Phf17	chr3	41359656	41420786
Igsf21	chr4	139582767	139802726
Otx2	chr14	49277859	49282547
Klhdc7a	chr4	139518088	139523941
1700125H03Rik	chr8	70892358	70899609
Lpar3	chr3	145883925	145949178
Mir6239	chr14	118352964	118353069
Epas1	chr17	87153204	87232750
Slc6a1	chr6	114232629	114267519
Cdh26	chr2	178165312	178222071
1700025C18Rik	chr2	164904193	164916250
Prox1	chr1	191945658	191994559
1700121N20Rik	chr12	107680862	107685876
Adamts2	chr11	50415587	50617551
Gadl1	chr9	115818573	115985294
Dnase2b	chr3	146244337	146278562
Inhbb	chr1	121312042	121318825
E2f3	chr13	29998444	30077932
Ajap1	chr4	152747330	152856939
BC049762	chr11	51067153	51076453
Edn3	chr2	174586274	174609543
Enc1	chr13	98011060	98022995
4930465M20Rik	chr12	108961953	108973698
9630028H03Rik	chr2	135406266	135408956
Cd44	chr2	102651300	102741822
Epgn	chr5	91456543	91464238
Syt13	chr2	92755258	92796208
Myb	chr10	20844736	20880790
Lrig3	chr10	125403275	125452415
Fam198b	chr3	79689852	79750200
Smoc1	chr12	82127795	82287401
1700084F23Rik	chr13	70142928	70167226

[0127]

[0128] 흥미롭게도, 이들 유전자는 착상전 배아에서 부계적으로 발현되며, H3K27me3의 이소적 제거는 모계 대립유전자 발현을 유도하였다. H3K27me3-의존성 각인은 배아 세포 계통에서 크게 소실되었지만, 적어도 5개의 유전자는 배외 세포 계통에서 그들의 각인을 유지하였다. 5개의 유전자는 모두 이전에 확인된 DNA 메틸화-독립성 각인된 상염색체 유전자를 포함한다. 모계 H3K27me3은 DNA 메틸화-독립성 각인 메커니즘이다. 한 실시양태에서, 본 발명의 방법은 H3K27me3 선택적 메틸라제의 사용을 포함한다.

[0129] **H3K27me3은 X 염색체 불활성화에 중요하다**

[0130] 설치류를 비롯한 특정 진수류 포유동물의 암컷에서, 2개의 X 염색체 중 1개는 유전자 용량 보상을 달성하기 위해 불활성화된다. X 염색체 불활성화 (XCI)로 지칭되는 이 현상은 후생적 침묵화의 메커니즘을 이해하기 위한 우수한 모델을 제공한다. 발생 동안, XCI는 각인된 또는 무작위 방식 중 어느 하나로 일어날 수 있다. 각인된 XCI에 대해, 부계 X 염색체 (Xp)는 착상전 발생 동안 선택적으로 불활성화된다. 각인된 XCI는 배외 세포 계통에서 유지되지만, 이는 후기 배반포의 내부 세포 덩이 (ICM)에서 소실된다. 착상주위 단계에서, 위판 세포는 무작위 XCI를 겪어 Xp 또는 모계 X 염색체 (Xm) 중 어느 하나의 침묵화를 발생시킨다. 이전의 연구는 각인된 및 무작위 XCI 둘 다에서 X-연관 긴 비-코딩 RNA인 *Xist*의 중요한 역할을 입증하였다. *Xist* RNA는 시스로 X 염색체를 코팅하고 불활성화시킴으로써 XCI에 참여한다.

[0131] 게놈 각인은 기원의 모체 특이적 유전자 조절을 허용한다. 각인된 XCI 동안 Xp를 선택적으로 침묵시키기 위해, *Xist* 유전자는 긴 탐구 후에, 그러나 아직 확인되지 않은 메커니즘을 갖는 Xm에서의 침묵화를 위해 각인된다. 핵 전달 접근법을 사용한 이전의 연구는 *Xist*의 게놈 각인이 상염색체 각인된 유전자의 그것과 같이, 난자형성

동안 확립됨을 시사하였다. 마우스 착상전 배아 및 배외 세포에서, 단지 부계 X 염색체 (Xp)만이 불활성화된다. 각인된 부계 X 염색체 불활성화 (XCI)에 대한 중심은 긴 비-코딩 RNA, *Xist*이며, 이는 Xp로부터 발현되고, 시스로 작용하여 전체 Xp를 코팅하고 침묵시킨다. Xp-특이적 불활성화를 달성하기 위해, 모계 *Xist* 유전자는 침묵화되어야 하지만, 침묵화 메커니즘은 아직 명확하지 않다. 본원에서 보고된 바와 같이, *Xist* 좌위는 마우스 난모세포에서 폭넓은 H3K27me3 도메인으로 코팅되며, 이는 착상전 발생을 통해 지속된다. H3K27me3의 이소적 제거는 모계 *Xist* 발현 및 모계 XCI를 유도한다. 따라서, 모계 H3K27me3은 *Xist*의 각인 마크로서 기능한다.

[0132] 일부 실시양태에서, 본 발명의 방법은 선택적 H3K27me3 데메틸라제 억제제를 포함하는 제약 조성물을 투여하는 것을 포함한다.

[0133] **H3K9me3 및 SCNT**

[0134] 공여자 체세포에서의 히스톤 H3 리신 9 트리메틸화 (H3K9me3)는 SCNT 재프로그래밍에 대한 후생적 장벽이다. 공여자 세포에서의 H3K9me3은 접합자 게놈 활성화에서의 연관된 영역의 전사 활성화를 방지하며, 마우스 및 인간 둘 다에서 착상전 단계에서 SCNT 배아의 발생 정지를 발생시킨다. 중요하게는, H3K9me3-특이적 데메틸라제, *Kdm4d*의 과발현에 의한 H3K9me3 장벽의 제거는 SCNT 배아가 IVF의 그것과 유사한 속도로 배반포 단계로 발달하는 것을 허용한다. 결과적으로, 만기 비율을 위한 전체 클로닝 효율은 8 내지 9배 증가된다. SCNT에서 *Kdm4d*의 사용은 IVF의 그것과 필적하는 착상률을 발생시키지만, 착상된 SCNT 배아의 15% 미만이 만기까지 발생한다. 더욱이, 비정상적으로 큰 태반은 *Kdm4d*-주사된 SCNT 배아에서 여전히 관찰된다. 이들 결과는 H3K9me3 재프로그래밍 장벽이 주로 착상전 발생을 방해하며, 다른 장벽은 착상후 발생에 영향을 미침을 시사한다.

[0135] *Xist*는 마우스 SCNT 배아의 착상후 발생에 중요하다. 모계 X 염색체로부터의 *Xist*의 비정상적 발현은 이소적 X 염색체 불활성화 (XCI) 및 착상전 배아의 전반적 전사 변경을 발생시키며, 이는 SCNT 배아의 착상후 발생 실패를 발생시킨다. 중요하게는, 이소적 *Xist* 발현에 의해 유발된 이 발생 실패는 공여자 세포로서 *Xist* 넉아웃 (KO) 체세포를 사용함으로써 또는 *Xist*에 대한 작은 간섭 RNA를 1-세포 수컷 SCNT 배아 내로 주사하여 만기 비율의 8 내지 10배 증가를 발생시킴으로써 극복될 수 있다.

[0136] **억제성 핵산**

[0137] 억제성 핵산 분자는 폴리펩티드 또는 폴리뉴클레오티드 (예를 들어, Xist 폴리뉴클레오티드)의 발현 또는 활성을 억제하는 올리고뉴클레오티드이다. 이러한 올리고뉴클레오티드는 Xist 폴리뉴클레오티드를 코딩하는 핵산 분자 (예를 들어, 안티센스 분자, siRNA, shRNA)에 결합하는 단일 및 이중 가닥 핵산 분자 (예를 들어, DNA, RNA, 및 그의 유사체)를 포함한다.

[0138] *siRNA*

[0139] 짧은 21 내지 25개의 뉴클레오티드 이중-가닥 RNA는 유전자 발현을 하향-조절하는 데 유효하다 (Zamore et al., Cell 101: 25-33; Elbashir et al., Nature 411: 494-498, 2001, 본원에 참조로 포함됨). 포유동물에서 siRNA 접근법의 유효성은 맥카페리 (McCaffrey) 등 (Nature 418: 38-39.2002)에 의해 생체내에서 입증되었다.

[0140] 표적 유전자의 서열을 고려하여, siRNA는 그 유전자를 불활성화시키도록 디자인될 수 있다. 이러한 siRNA는 예를 들어, 질환에 걸린 조직에 직접적으로 투여되거나, 전신적으로 투여될 수 있다. 유전자의 핵산 서열은 작은 간섭 RNA (siRNA)를 디자인하는 데 사용될 수 있다. 21 내지 25개의 뉴클레오티드 siRNA는 예를 들어, Xist 발현을 감소시키는 데 사용될 수 있다.

[0141] 본 발명의 억제성 핵산 분자는 발현의 RNA 간섭 (RNAi)-매개된 넉-다운을 위한 이중-가닥 RNA로서 채용될 수 있다. 한 실시양태에서, Xist 유전자의 발현은 체세포에서 감소된다. RNAi는 관심의 특이적 단백질의 세포 발현을 감소시키는 방법이다 (문헌 [Tuschl, Chembiochem 2:239-245, 2001]; [Sharp, Genes & Devel. 15:485-490, 2000]; [Hutvagner and Zamore, Curr. Opin. Genet. Devel. 12:225-232, 2002]; 및 [Hannon, Nature 418:244-251, 2002]에서 검토됨). dsRNA의 형질감염에 의한 또는 플라스미드-기재 발현 시스템을 사용한 siRNA의 발현을 통한 세포 내로의 siRNA의 도입은 포유동물 세포에서 기능의 소실 표현형을 생성하기 위해 점점 더 사용되고 있다.

[0142] 본 발명의 한 실시양태에서, 본 발명의 핵염기 올리고머의 8 내지 19개의 연속적 핵염기를 포함하는 이중-가닥 RNA (dsRNA) 분자가 제조된다. dsRNA는 듀플렉스화된 RNA의 2개의 별개의 가닥, 또는 자기-듀플렉스화된 단일 RNA 가닥 (작은 헤어핀 (sh)RNA)일 수 있다. 전형적으로, dsRNA는 약 21개 또는 22개 염기 쌍이지만, 바람직할

경우 더 짧거나 더 길 수 있다 (약 29개 이하의 핵염기). dsRNA는 표준 기법 (예를 들어, 화학적 합성 또는 시험관내 전사)을 사용하여 제조될 수 있다. 예를 들어, 앰비온 (Ambion) (미국 텍사스주 오스틴) 및 에피센터 (Epicentre) (미국 위스콘신주 매디슨)로부터의 키트가 이용가능하다. 포유동물 세포에서 dsRNA를 발현시키는 방법은 문헌 [Brummelkamp et al. Science 296:550-553, 2002]; [Paddison et al. Genes & Devel. 16:948-958, 2002]. [Paul et al. Nature Biotechnol. 20:505-508, 2002]; [Sui et al. Proc. Natl. Acad. Sci. USA 99:5515-5520, 2002]; [Yu et al. Proc. Natl. Acad. Sci. USA 99:6047-6052, 2002]; [Miyagishi et al. Nature Biotechnol. 20:497-500, 2002]; 및 [Lee et al. Nature Biotechnol. 20:500-505 2002]에 기재되어 있으며, 이들의 각각은 본원에 참조로 포함된다.

[0143] 작은 헤어핀 RNA (shRNA)는 줄기-루프 구조를 갖는 RNA 서열을 포함한다. "줄기-루프 구조"는 우세하게 단일-가닥 뉴클레오티드의 영역 (루프 부분)에 의해 한 측에 연결된 이중 가닥 또는 듀플렉스를 형성하는 것으로 공지되거나 예측된 뉴클레오티드의 영역 (줄기 부분)을 포함하는 2차 구조를 갖는 핵산을 지칭한다. 용어 "헤어핀"은 또한 본원에서 줄기-루프 구조를 지칭하기 위해 사용된다. 이러한 구조는 관련 기술분야에 널리 공지되어 있으며, 상기 용어는 관련 기술분야에서의 그의 공지된 의미와 일치하게 사용된다. 관련 기술분야에 공지된 바와 같이, 2차 구조는 정확한 염기-쌍형성을 요구하지 않는다. 따라서, 줄기는 1개 이상의 염기 미스매치 또는 벌지 (bulge)를 포함할 수 있다. 대안적으로, 염기-쌍형성은 정확할 수 있으며, 즉, 임의의 미스매치를 포함하지 않는다. 다수의 줄기-루프 구조는 링커, 예컨대, 예를 들어, 핵산 링커, miRNA 플랭킹 서열, 다른 분자, 또는 이들의 일부 조합을 통해 서로에 연결될 수 있다.

[0144] 본원에 사용된 바와 같은 용어 "작은 헤어핀 RNA"는 전구체 miRNA (프리-miRNA)를 형성하는 통상적인 줄기-루프 shRNA를 포함한다. 범위에 있어서 일부 변이가 있을 수 있지만, 통상적인 줄기-루프 shRNA는 19 내지 29 bp 범위의 줄기, 및 4 내지 30 bp 범위의 루프를 포함할 수 있다. "shRNA"는 또한 마이크로-RNA 포매된 shRNA (miRNA-기재 shRNA)를 포함하며, 여기서, miRNA 듀플렉스의 가이드 가닥 및 패신저 가닥은 기존의 (또는 천연) miRNA 내로 또는 변형된 또는 합성 (디자인된) miRNA 내로 혼입된다. 일부 예에서, 전구체 miRNA 분자는 1개 초과의 줄기-루프 구조를 포함할 수 있다. 마이크로RNA는 약 22-뉴클레오티드 길이이며, 일반적으로 고도로 조직- 또는 발생-단계-특이적 방식으로 발현되고, 표적 유전자를 전사후로 조절하는 내인적으로 코딩된 RNA 분자이다. 200개 초과의 별개의 miRNA가 식물 및 동물에서 확인되었다. 이들 작은 조절성 RNA는 2가지 우세한 작용의 방식에 의해 중요한 생물학적 기능을 하는 것으로 믿어진다: (1) 표적 mRNA의 번역을 억제함으로써, 및 (2) RNA 간섭 (RNAi), 즉, mRNA의 절단 및 분해를 통해. 후자의 경우, miRNA는 작은 간섭 RNA (siRNA)와 유사하게 기능한다. 따라서, 기존의 miRNA 유전자의 특색에 기초하여 인공 miRNA를 디자인하고 발현시킬 수 있다.

[0145] shRNA는 지속된 침묵화 및 거의 임의의 세포 유형 내로의 고수율 전달을 제공하도록 DNA 벡터로부터 발현될 수 있다. 일부 실시양태에서, 벡터는 바이러스 벡터이다. 예시적인 바이러스 벡터로는 렌티바이러스, 아데노바이러스, 바쿨로바이러스 및 조류 바이러스 벡터를 비롯한, 및 안정한, 단일-카피 게놈 통합을 허용하는 이러한 벡터를 비롯한 레트로바이러스를 들 수 있다. 레트로바이러스 플라스미드 벡터가 유래될 수 있는 레트로바이러스는 몰로니 뮤린 백혈병 바이러스, 비장 괴사 바이러스, 라우스 육종 바이러스, 하비 육종 바이러스, 조류 백혈증 바이러스, 긴팔 원숭이 백혈병 바이러스, 인간 면역결핍 바이러스, 골수증식성 육종 바이러스, 및 유방 종양 바이러스를 포함하나, 이에 제한되지는 않는다. 레트로바이러스 플라스미드 벡터는 패키징 세포주를 형질도입하여 생산자 세포주를 형성하는 데 채용될 수 있다. 형질감염될 수 있는 패키징 세포의 예로는 그 전문이 본원에 참조로 포함되는 문헌 [Miller, Human Gene Therapy 1:5-14 (1990)]에 기재된 바와 같은 PE501, PA317, R-2, R-AM, PA12, T19-14x, VT-19-17-H2, RCRE, RCRIP, GP+E-86, GP+envAm12, 및 DAN 세포주를 들 수 있으나, 이에 제한되지는 않는다. 벡터는 패키징 세포를 관련 기술분야에 공지된 임의의 수단을 통해 형질도입할 수 있다. 생산자 세포주는 DNA 복제 단백질을 코딩하는 폴리뉴클레오티드를 포함하는 감염성 레트로바이러스 벡터 입자를 생성한다. 그 후, 이러한 레트로바이러스 벡터 입자는 시험관내에서 또는 생체내에서 진핵생물 세포를 형질도입하는 데 채용될 수 있다. 형질도입된 진핵생물 세포는 DNA 복제 단백질을 발현할 것이다.

[0146] 본 발명의 안티센스 서열을 포함하는 촉매 RNA 분자 또는 리보자임은 생체내에서 핵산 분자 (예를 들어, Xist)의 발현을 억제하는 데 사용될 수 있다. 안티센스 RNA 내의 리보자임 서열의 포함은 그들에게 RNA-절단 활성을 부여함으로써, 구축물의 활성을 증가시킨다. 표적 RNA-특이적 리보자임의 디자인 및 용도는 문헌 [Haseloff et al., Nature 334:585-591. 1988], 및 미국 특허 출원 공개 제2003/0003469 A1호에 기재되어 있으며, 이들의 각각은 참조로 포함된다.

[0147] 따라서, 본 발명은 또한 결합 아암에 8 내지 19개의 연속적 핵염기를 갖는 안티센스 RNA를 포함하는 촉매 RNA 분자를 특색으로 한다. 본 발명의 바람직한 실시양태에서, 촉매 핵산 분자는 해머헤드 또는 헤어핀 모티프에서

형성된다. 이러한 해머헤드 모티프의 예는 문헌 [Rossi et al., Aids Research and Human Retroviruses, 8:183, 1992]에 의해 기재되어 있다. 헤어핀 모티프의 예는 1988년 9월 20일에 출원된 미국 일련 번호 07/247,100의 일부계속출원인 1989년 9월 20일에 출원된 햄펠 (Hampel) 등, "특이적 RNA 서열을 절단하기 위한 RNA 촉매"; 문헌 [Hampel and Tritz, Biochemistry, 28:4929, 1989]; 및 [Hampel et al., Nucleic Acids Research, 18: 299, 1990]에 의해 기재되어 있다. 이들 특이적 모티프는 본 발명에서 제한적이지 않으며, 관련 기술분야의 통상의 기술자는 본 발명의 효소적 핵산 분자에서 중요한 모든 것이, 그것이 표적 유전자 RNA 영역 중 1개 이상에 상보적인 특이적 기질 결합 부위를 갖고, 그것이 분자에 RNA 절단 활성을 부여하는 그 기질 결합 부위 내의 또는 주위의 뉴클레오티드 서열을 갖는 것임을 인식할 것이다.

[0148] 세포 내로 핵산 구축물을 도입하는 본질적으로 임의의 방법이 채용될 수 있다. 핵산을 도입하는 물리적 방법은 구축물을 함유하는 용액의 주사, 구축물에 의해 덮인 입자에 의한 포격, 핵산의 용액에서의 세포, 조직 샘플 또는 유기체의 흡인, 또는 구축물의 존재 하에서의 세포 막의 전기천공을 포함한다. 바이러스 입자 내로 패키징된 바이러스 구축물은 세포 내로의 발현 구축물의 효율적인 도입 및 코딩된 shRNA의 전사 둘 다를 달성하는 데 사용될 수 있다. 핵산을 세포에 도입하는 관련 기술분야에 공지된 다른 방법, 예컨대 지질-매개된 운반체 수송, 화학물질 매개된 수송, 예컨대 인산칼슘 등이 사용될 수 있다. 따라서, shRNA-코딩 핵산 구축물은 하기 활성 중 하나 이상을 수행하는 성분과 함께 도입될 수 있다: 세포에 의한 RNA 흡수를 향상시키거나, 듀플렉스 가닥의 어닐링을 촉진시키거나, 어닐링된 가닥을 안정화시키거나, 다르게는 표적 유전자의 억제를 증가시킴.

[0149] 세포 내에서의 발현을 위해, DNA 벡터, 예를 들어 RNA 폴리머라제 II 또는 RNA 폴리머라제 III 프로모터 중 어느 하나를 포함하는 플라스미드 벡터가 채용될 수 있다. 내인성 miRNA의 발현은 RNA 폴리머라제 II (Pol II) 프로모터에 의해 제어되며, 일부의 경우, shRNA는 RNA 폴리머라제 III 프로모터에 비해 Pol II 프로모터에 의해 가장 효율적으로 유도된다 (Dickins et al., 2005, Nat. Genet. 39: 914-921). 일부 실시양태에서, shRNA의 발현은 제한 없이, RNA 폴리머라제 유형 II 프로모터를 비롯한 유도성 프로모터 또는 조건적 발현 시스템에 의해 제어될 수 있다. 본 발명의 맥락에서 유용한 프로모터의 예는 테트라시클린-유도성 프로모터 (TRE-타이트를 포함함), IPTG-유도성 프로모터, 테트라시클린 전사활성화제 시스템, 및 리버스 테트라시클린 전사활성화제 (rtTA) 시스템이다. 구성적 프로모터는 또한 사용될 수 있으며, 세포- 또는 조직-특이적 프로모터도 사용될 수 있다. 많은 프로모터는 이들이 모든 세포 및 조직 유형에서 발현되도록 편재적일 것이다. 특정 실시양태는 시험관내 및 생체내 연구에서 가장 효과적인 조건적 유전자 발현 시스템 중 하나인 테트라시클린-반응성 프로모터를 사용한다. 유도성 shRNA의 설명에 대해서는, 국제 특허 출원 PCT/US2003/030901 (공개 제WO 2004-029219 A2호) 및 문헌 [Fewell et al., 2006, Drug Discovery Today 11: 975-982]을 참조한다.

[0150] **폴리뉴클레오티드의 전달**

[0151] 네이키드 폴리뉴클레오티드, 또는 그의 유사체는 포유동물 세포에 진입하고, 관심의 유전자의 발현을 억제할 수 있다. 그럼에도 불구하고, 세포에의 올리고뉴클레오티드 또는 다른 핵염기 올리고머의 전달을 보조하는 제형을 이용하는 것이 바람직할 수 있다 (예를 들어, 미국 특허 제5,656,611호; 제5,753,613호; 제5,785,992호; 제6,120,798호; 제6,221,959호; 제6,346,613호; 및 제6,353,055호 참조; 이들의 각각은 본원에 참조로 포함됨).

[0152] **올리고뉴클레오티드 및 다른 핵염기 올리고머**

[0153] 적어도 2가지 유형의 올리고뉴클레오티드는 RN아제 H에 의한 RNA의 절단을 유도한다: 포스포디에스테르 (PO) 또는 포스포로티오에이트 (PS) 연결을 갖는 폴리데옥시뉴클레오티드. 2'-OMe-RNA 서열은 RNA 표적에 대한 높은 친화도를 나타내지만, 이들 서열은 RN아제 H에 대한 기질이 아니다. 바람직한 올리고뉴클레오티드는 뉴클레아제 저항성을 위해 포스포로티오에이트에 대해 변형된 일부 또는 전부의 뉴클레오티드간 연결을 갖는 올리고데옥시뉴클레오티드 갭을 함유하는 2'-변형된 올리고뉴클레오티드에 기초한 것이다. 메틸포스포네이트 변형의 존재는 그의 표적 RNA에 대한 올리고뉴클레오티드의 친화도를 증가시키고, 따라서 IC_{50}을 감소시킨다. 이 변형은 또한 변형된 올리고뉴클레오티드의 뉴클레아제 저항성을 증가시킨다. 본 발명의 방법 및 시약은 공유-폐쇄된 다중 안티센스 (CMAS) 올리고뉴클레오티드 (Moon et al., Biochem J. 346:295-303, 2000; PCT 공개 제WO 00/61595호), 리본-유형 안티센스 (RiAS) 올리고뉴클레오티드 (Moon et al., J. Biol. Chem. 275:4647-4653, 2000; PCT 공개 제WO 00/61595호), 및 큰 원형 안티센스 올리고뉴클레오티드 (미국 특허 출원 공개 제US 2002/0168631 A1호)를 비롯한 개발될 수 있는 임의의 기술과 함께 사용될 수 있음이 이해된다.

[0154] 관련 기술분야에 공지된 바와 같이, 뉴클레오시드는 핵염기-당 조합이다. 뉴클레오시드의 염기 부분은 통상적으로 헤테로시클릭 염기이다. 이러한 헤테로시클릭 염기의 2가지 가장 통상적인 부류는 퓨린 및 피리미딘이다. 뉴클레오티드는 뉴클레오시드의 당 부분에 공유 연결된 포스페이트 기를 더 포함하는 뉴클레오시드이다. 펜토

푸라노실 당을 포함하는 그러한 뉴클레오시드에 대해, 포스페이트 기는 당의 2', 3' 또는 5' 히드록실 모이어티 중 어느 하나에 연결될 수 있다. 올리고뉴클레오티드를 형성하는 데 있어서, 포스페이트 기는 인접한 뉴클레오시드를 서로에게 공유 연결하여 선형 중합체성 화합물을 형성한다. 다시, 이 선형 중합체성 구조의 각각의 말단은 더 결합되어 원형 구조를 형성할 수 있으며; 개방된 선형 구조는 일반적으로 바람직하다. 올리고뉴클레오티드 구조 내에서, 포스페이트 기는 통상적으로 올리고뉴클레오티드의 백본을 형성하는 것으로 지칭된다. RNA 및 DNA의 통상적인 연결 또는 백본은 3' 내지 5' 포스포디에스테르 연결이다.

[0155] 본 발명에서 유용한 바람직한 핵염기 올리고머의 구체적인 예로는 변형된 백본 또는 비-천연 뉴클레오시드간 연결을 함유하는 올리고뉴클레오티드를 들 수 있다. 이 명세서에서 정의된 바와 같이, 변형된 백본을 갖는 핵염기 올리고머는 백본에 인 원자를 보유하는 것들 및 백본에 인 원자를 갖지 않는 것들을 포함한다. 이 명세서의 목적상, 그들의 뉴클레오시드간 백본에 인 원자를 갖지 않는 변형된 올리고뉴클레오티드는 또한 핵염기 올리고머인 것으로 간주된다.

[0156] 변형된 올리고뉴클레오티드 백본을 갖는 핵염기 올리고머는 예를 들어, 통상적인 3'-5' 연결을 갖는 포스포로티오에이트, 키랄 포스포로티오에이트, 포스포로디티오에이트, 포스포트리에스테르, 아미노알킬-포스포트리에스테르, 메틸 및 다른 알킬 포스포네이트, 예컨대 3'-알킬렌 포스포네이트 및 키랄 포스포네이트, 포스피네이트, 포스포르아미데이트, 예컨대 3'-아미노 포스포르아미데이트 및 아미노알킬포스포르아미데이트, 티오노포스포르아미데이트, 티오노알킬포스포네이트, 티오노알킬포스포트리에스테르, 및 보라노포스페이트, 이들의 2'-5' 연결된 유사체, 및 반전된 극성을 갖는 것들을 포함하며, 여기서, 뉴클레오시드 단위의 인접한 쌍은 3'-5' 내지 5'-3' 또는 2'-5' 내지 5'-2' 연결된다. 다양한 염, 혼합된 염 및 유리산 형태는 또한 포함된다. 상기 인-함유 연결의 제조를 교시하는 대표적인 미국 특허는 미국 특허 제3,687,808호; 제4,469,863호; 제4,476,301호; 제5,023,243호; 제5,177,196호; 제5,188,897호; 제5,264,423호; 제5,276,019호; 제5,278,302호; 제5,286,717호; 제5,321,131호; 제5,399,676호; 제5,405,939호; 제5,453,496호; 제5,455,233호; 제5,466,677호; 제5,476,925호; 제5,519,126호; 제5,536,821호; 제5,541,306호; 제5,550,111호; 제5,563,253호; 제5,571,799호; 제5,587,361호; 및 제5,625,050호를 포함하나, 이에 제한되지는 않으며, 이들의 각각은 본원에 참조로 포함된다.

[0157] 그 안에 인 원자를 포함하지 않는 변형된 올리고뉴클레오티드 백본을 갖는 핵염기 올리고머는 짧은 쇄 알킬 또는 시클로알킬 뉴클레오시드간 연결, 혼합된 헤테로원자 및 알킬 또는 시클로알킬 뉴클레오시드간 연결, 또는 1개 이상의 짧은 쇄 헤테로원자 또는 헤테로시클릭 뉴클레오시드간 연결에 의해 형성되는 백본을 갖는다. 이들은 모르폴리노 연결 (부분적으로 뉴클레오시드의 당 부분으로부터 형성됨); 실록산 백본; 술피드, 술폭시드 및 술폰 백본; 포름아세틸 및 티오포름아세틸 백본; 메틸렌 포름아세틸 및 티오포름아세틸 백본; 알켄 함유 백본; 술파메이트 백본; 메틸렌이미노 및 메틸렌히드라지노 백본; 술포네이트 및 술폰아미드 백본; 아미드 백본을 갖는 것들; 및 혼합된 N, O, S 및 CH_2 성분 부분을 갖는 기타의 것들을 포함한다. 상기 올리고뉴클레오티드의 제조를 교시하는 대표적인 미국 특허는 미국 특허 제5,034,506호; 제5,166,315호; 제5,185,444호; 제5,214,134호; 제5,216,141호; 제5,235,033호; 제5,264,562호; 제5,264,564호; 제5,405,938호; 제5,434,257호; 제5,466,677호; 제5,470,967호; 제5,489,677호; 제5,541,307호; 제5,561,225호; 제5,596,086호; 제5,602,240호; 제5,610,289호; 제5,602,240호; 제5,608,046호; 제5,610,289호; 제5,618,704호; 제5,623,070호; 제5,663,312호; 제5,633,360호; 제5,677,437호; 및 제5,677,439호를 포함하나, 이에 제한되지는 않으며, 이들의 각각은 본원에 참조로 포함된다.

[0158] 다른 핵염기 올리고머에서, 당 및 뉴클레오시드간 연결, 즉, 백본 둘 다는 신규한 기로 대체된다. 핵염기 단위는 열거된 *Xist* 유전자와의 혼성화를 위해 유지된다. 한 이러한 핵염기 올리고머는 펩티드 핵산 (Peptide Nucleic Acid) (PNA)으로 지칭된다. PNA 화합물에서, 올리고뉴클레오티드의 당-백본은 아미드 함유 백본, 특히 아미노에틸글리신 백본으로 대체된다. 핵염기는 보유되고, 백본의 아미드 부분의 아자 질소 원자에 직접적으로 또는 간접적으로 결합된다. 이들 핵염기 올리고머를 제조하고 사용하는 방법은 예를 들어, 문헌 ["Peptide Nucleic Acids: Protocols and Applications" Ed. P. E. Nielsen, Horizon Press, Norfolk, United Kingdom, 1999]에 기재되어 있다. PNA의 제조를 교시하는 대표적인 미국 특허는 미국 특허 제5,539,082호; 제5,714,331호; 및 제5,719,262호를 포함하나, 이에 제한되지는 않으며, 이들의 각각은 본원에 참조로 포함된다. PNA 화합물의 추가의 교시는 문헌 [Nielsen et al., Science, 1991, 254, 1497-1500]에서 발견될 수 있다.

[0159] 본 발명의 특정 실시양태에서, 핵염기 올리고머는 포스포로티오에이트 백본 및 헤테로원자를 갖는 뉴클레오시드 백본, 및 특히 $-CH_2-NH-O-CH_2-$, $-CH_2-N(CH_3)-O-CH_2-$ (메틸렌 (메틸이미노) 또는 MMI 백본으로 공지됨), $-CH_2-O-$

$N(CH_3)-CH_2-$, $-CH_2-N(CH_3)-N(CH_3)-CH_2-$, 및 $-O-N(CH_3)-CH_2-CH_2-$를 갖는다. 다른 실시양태에서, 올리고뉴클레오티드는 미국 특허 제5,034,506호에 기재된 모르폴리노 백본 구조를 갖는다.

[0160] 핵염기 올리고머는 또한 1개 이상의 치환된 당 모이어티를 함유할 수 있다. 핵염기 올리고머는 2' 위치에 하기 중 하나를 포함한다: OH; F; O-, S-, 또는 N-알킬; O-, S-, 또는 N-알케닐; O-, S- 또는 N-알키닐; 또는 O-알킬-O-알킬 (여기서, 알킬, 알케닐, 및 알키닐은 치환되거나 비치환된 C_1 내지 C_{10} 알킬 또는 C_2 내지 C_{10} 알케닐 및 알키닐일 수 있음). 특히 바람직한 것은 $O[(CH_2)_nO]_mCH_3$, $O(CH_2)_nOCH_3$, $O(CH_2)_nNH_2$, $O(CH_2)_nCH_3$, $O(CH_2)_nONH_2$, 및 $O(CH_2)_nON[(CH_2)_nCH_3)]_2$이며, 여기서, n 및 m은 1 내지 약 10이다. 다른 바람직한 핵염기 올리고머는 2' 위치에 하기 중 하나를 포함한다: C_1 내지 C_{10} 저급 알킬, 치환된 저급 알킬, 알카릴, 아르알킬, O-알카릴, 또는 O-아르알킬, SH, SCH_3, OCN, Cl, Br, CN, CF_3, OCF_3, $SOCH_3$, SO_2CH_3, ONO_2, NO_2, NH_2, 헤테로시클로알킬, 헤테로시클로알카릴, 아미노알킬아미노, 폴리알킬아미노, 치환된 실릴, RNA 절단기, 리포터 기, 삽입제, 핵염기 올리고머의 약동학적 특성을 개선시키는 기, 또는 핵염기 올리고머의 약역학적 특성을 개선시키는 기, 및 유사한 특성을 갖는 다른 치환기. 바람직한 변형은 2'-O-메틸 및 2'-메톡시에톡시 (2'-O-$CH_2CH_2OCH_3$, 또한 2'-O-(2-메톡시에틸) 또는 2'-MOE로 공지됨)이다. 또다른 바람직한 변형은 또한 2'-DMAOE로 공지된 2'-디메틸아미노옥시에톡시 (즉, $O(CH_2)_2ON(CH_3)_2$)이다. 다른 변형은 2'-아미노프로폭시 (2'-$OCH_2CH_2CH_2NH_2$) 및 2'-플루오로 (2'-F)를 포함한다. 유사한 변형은 또한 올리고뉴클레오티드 또는 다른 핵염기 올리고머의 다른 위치, 특히 3' 말단 뉴클레오티드 상의 또는 또는 2'-5' 연결된 올리고뉴클레오티드에서의 당의 3' 위치 및 5' 말단 뉴클레오티드의 5' 위치에서 이루어질 수 있다. 핵염기 올리고머는 또한 펜토푸라노실 당 대신 당 모방체, 예컨대 시클로부틸 모이어티를 가질 수 있다. 이러한 변형된 당 구조의 제조를 교시하는 대표적인 미국 특허는 미국 특허 제4,981,957호; 제5,118,800호; 제5,319,080호; 제5,359,044호; 제5,393,878호; 제5,446,137호; 제5,466,786호; 제5,514,785호; 제5,519,134호; 제5,567,811호; 제5,576,427호; 제5,591,722호; 제5,597,909호; 제5,610,300호; 제5,627,053호; 제5,639,873호; 제5,646,265호; 제5,658,873호; 제5,670,633호; 및 제5,700,920호를 포함하나, 이에 제한되지는 않으며, 이들의 각각은 그 전문이 본원에 참조로 포함된다.

[0161] 핵염기 올리고머는 또한 핵염기 변형 또는 치환을 포함할 수 있다. 본원에 사용된 바와 같은 "비변형된" 또는 "천연" 핵염기는 퓨린 염기 아데닌 (A) 및 구아닌 (G), 및 피리미딘 염기 티민 (T), 시토신 (C) 및 우라실 (U)을 포함한다. 변형된 핵염기는 다른 합성 및 천연 핵염기, 예컨대 5-메틸시토신 (5-me-C), 5-히드록시메틸 시토신, 크산틴, 히포크산틴, 2-아미노아데닌, 아데닌 및 구아닌의 6-메틸 및 다른 알킬 유도체; 아데닌 및 구아닌의 2-프로필 및 다른 알킬 유도체; 2-티오우라실, 2-티오티민 및 2-티오시토신; 5-할로우라실 및 시토신; 5-프로피닐 우라실 및 시토신; 6-아조 우라실, 시토신 및 티민; 5-우라실 (슈도우라실); 4-티오우라실; 8-할로, 8-아미노, 8-티올, 8-티오알킬, 8-히드록실 및 다른 8-치환된 아데닌 및 구아닌; 5-할로 (예를 들어, 5-브로모), 5-트리플루오로메틸 및 다른 5-치환된 우라실 및 시토신; 7-메틸구아닌 및 7-메틸아데닌; 8-아자구아닌 및 8-아자아데닌; 7-데아자구아닌 및 7-데아자아데닌; 및 3-데아자구아닌 및 3-데아자아데닌을 포함한다. 추가의 핵염기는 미국 특허 제3,687,808호에 개시된 것들, 문헌 [The Concise Encyclopedia Of Polymer Science And Engineering, pages 858-859, Kroschwitz, J. I., ed. John Wiley & Sons, 1990]에 개시된 것들, 문헌 [Englisch et al., Angewandte Chemie, International Edition, 1991, 30, 613]에 의해 개시된 것들, 및 문헌 [Sanghvi, Y. S., Chapter 15, Antisense Research and Applications, pages 289-302, Crooke, S. T. and Lebleu, B., ed., CRC Press, 1993]에 의해 개시된 것들을 포함한다. 특정의 이들 핵염기는 본 발명의 안티센스 올리고뉴클레오티드의 결합 친화도를 증가시키는 데 특히 유용하다. 이들은 5-치환된 피리미딘, 6-아자피리미딘, 및 N-2, N-6 및 O-6 치환된 퓨린, 예컨대 2-아미노프로필아데닌, 5-프로피닐우라실 및 5-프로피닐시토신을 포함한다. 5-메틸시토신 치환은 핵산 듀플렉스 안정성을 0.6 내지 1.2℃ 증가시키는 것으로 나타났으며 (Sanghvi, Y. S., Crooke, S. T. and Lebleu, B., eds., Antisense Research and Applications, CRC Press, Boca Raton, 1993, pp. 276-278), 심지어 보다 특히 2'-O-메톡시에틸 또는 2'-O-메틸 당 변형과 조합되는 경우에도 바람직한 염기 치환이다. 특정의 상기 언급된 변형된 핵염기 뿐만 아니라 다른 변형된 핵염기의 제조를 교시하는 대표적인 미국 특허는 미국 특허 제4,845,205호; 제5,130,302호; 제5,134,066호; 제5,175,273호; 제5,367,066호; 제5,432,272호; 제5,457,187호; 제5,459,255호; 제5,484,908호; 제5,502,177호; 제5,525,711호; 제5,552,540호; 제5,587,469호; 제5,594,121호, 제5,596,091호; 제5,614,617호; 제5,681,941호; 및 제5,750,692호를 포함하며, 이들의 각각은 본원에 참조로 포함된다.

[0162] 본 발명의 핵염기 올리고머의 또다른 변형은 올리고뉴클레오티드의 활성, 세포 분포, 또는 세포 흡수를 향상시키는 1개 이상의 모이어티 또는 접합체를 핵염기 올리고머에 화학적으로 연결시키는 것을 포함한다. 이러한 모이어티는 지질 모이어티, 예컨대 콜레스테롤 모이어티 (Letsinger et al., Proc. Natl. Acad. Sci. USA, 86:6553-6556, 1989), 콜산 (Manoharan et al., Bioorg. Med. Chem. Let, 4:1053-1060, 1994), 티오에테르, 예를 들어, 헥실-S-트리틸티올 (Manoharan et al., Ann. N.Y. Acad. Sci., 660:306-309, 1992; Manoharan et al., Bioorg. Med. Chem. Let., 3:2765-2770, 1993), 티오콜레스테롤 (Oberhauser et al., Nucl. Acids Res., 20:533-538: 1992), 지방족 쇄, 예를 들어, 도데칸디올 또는 운데실 잔기 (Saison-Behmoaras et al., EMBO J., 10:1111-1118, 1991; Kabanov et al., FEBS Lett., 259:327-330, 1990; Svinarchuk et al., Biochimie, 75:49-54, 1993), 인지질, 예를 들어, 디-헥사데실-rac-글리세롤 또는 트리에틸암모늄 1,2-디-O-헥사데실-rac-글리세로-3-H-포스포네이트 (Manoharan et al., Tetrahedron Lett., 36:3651-3654, 1995; Shea et al., Nucl. Acids Res., 18:3777-3783, 1990), 폴리아민 또는 폴리에틸렌 글리콜 쇄 (Manoharan et al., Nucleosides & Nucleotides, 14:969-973, 1995), 아다만탄 아세트산 (Manoharan et al., Tetrahedron Lett., 36:3651-3654, 1995), 팔미틸 모이어티 (Mishra et al., Biochim. Biophys. Acta, 1264:229-237, 1995), 또는 옥타데실아민 또는 헥실아미노-카르보닐-옥시콜레스테롤 모이어티 (Crooke et al., J. Pharmacol. Exp. Ther., 277:923-937, 1996)를 포함하나, 이에 제한되지는 않는다. 이러한 핵염기 올리고머 접합체의 제조를 교시하는 대표적인 미국 특허는 미국 특허 제4,587,044호; 제4,605,735호; 제4,667,025호; 제4,762,779호; 제4,789,737호; 제4,824,941호; 제4,828,979호; 제4,835,263호; 제4,876,335호; 제4,904,582호; 제4,948,882호; 제4,958,013호; 제5,082,830호; 제5,109,124호; 제5,112,963호; 제5,118,802호; 제5,138,045호; 제5,214,136호; 제5,218,105호; 제5,245,022호; 제5,254,469호; 제5,258,506호; 제5,262,536호; 제5,272,250호; 제5,292,873호; 제5,317,098호; 제5,371,241호, 제5,391,723호; 제5,414,077호; 제5,416,203호, 제5,451,463호; 제5,486,603호; 제5,510,475호; 제5,512,439호; 제5,512,667호; 제5,514,785호; 제5,525,465호; 제5,541,313호; 제5,545,730호; 제5,552,538호; 제5,565,552호; 제5,567,810호; 제5,574,142호; 제5,578,717호; 제5,578,718호; 제5,580,731호; 제5,585,481호; 제5,587,371호; 제5,591,584호; 제5,595,726호; 제5,597,696호; 제5,599,923호; 제5,599,928호; 제5,608,046호; 및 제5,688,941호를 포함하며, 이들의 각각은 본원에 참조로 포함된다.

[0163] 본 발명은 또한 키메라 화합물인 핵염기 올리고머를 포함한다. "키메라" 핵염기 올리고머는 각각 적어도 1개의 단량체 단위, 즉, 올리고뉴클레오티드의 경우 뉴클레오티드로 구성된 2개 이상의 화학적으로 별개의 영역을 함유하는 핵염기 올리고머, 특히 올리고뉴클레오티드이다. 이들 핵염기 올리고머는 전형적으로 핵염기 올리고머가 핵염기 올리고머 상에 뉴클레아제 분해에 대한 증가된 저항성, 증가된 세포 흡수, 및/또는 표적 핵산에 대한 증가된 결합 친화도를 부여하도록 변형된 적어도 1개의 영역을 함유한다. 핵염기 올리고머의 추가의 영역은 RNA:DNA 또는 RNA:RNA 하이브리드를 절단할 수 있는 효소에 대한 기질로서 기능할 수 있다. 예로서, RN아제 H는 RNA:DNA 듀플렉스의 RNA 가닥을 절단하는 세포성 엔도뉴클레아제이다. 따라서, RN아제 H의 활성화는 RNA 표적의 절단을 발생시킴으로써, 유전자 발현의 핵염기 올리고머 억제의 효율을 크게 향상시킨다. 결과적으로, 키메라 핵염기 올리고머가 사용되는 경우, 동일한 표적 영역에 혼성화하는 포스포로티오에이트 데옥시올리고뉴클레오티드에 비해, 필적하는 결과는 종종 보다 짧은 핵염기 올리고머로 얻어질 수 있다.

[0164] 본 발명의 키메라 핵염기 올리고머는 상기 기재된 바와 같은 2개 이상의 핵염기 올리고머의 복합 구조로서 형성될 수 있다. 이러한 핵염기 올리고머는, 올리고뉴클레오티드의 경우, 또한 관련 기술분야에서 하이브리드 또는 갭머로 지칭되었다. 이러한 하이브리드 구조의 제조를 교시하는 대표적인 미국 특허는 미국 특허 제5,013,830호; 제5,149,797호; 제5,220,007호; 제5,256,775호; 제5,366,878호; 제5,403,711호; 제5,491,133호; 제5,565,350호; 제5,623,065호; 제5,652,355호; 제5,652,356호; 및 제5,700,922호를 포함하며, 이들의 각각은 그 전문이 본원에 참조로 포함된다.

[0165] 본 발명에 따라 사용되는 핵염기 올리고머는 편리하게는 및 통상적으로 고체 상 합성의 널리 공지된 기법을 통해 제조될 수 있다. 이러한 합성을 위한 장비는 예를 들어, 어플라이드 바이오시스템스 (Applied Biosystems) (미국 캘리포니아주 포스터 시티)를 비롯한 몇몇 판매자에 의해 판매된다. 관련 기술분야에 공지된 이러한 합성을 위한 임의의 다른 수단은 추가적으로 또는 대안적으로 채용될 수 있다. 올리고뉴클레오티드, 예컨대 포스포로티오에이트 및 알킬화된 유도체를 제조하는 유사한 기법을 사용하는 것은 널리 공지되어 있다.

[0166] 본 발명의 핵염기 올리고머는 또한 흡수 (uptake), 분포 및/또는 흡수 (absorption)를 보조하기 위해, 예를 들어, 리포솜, 수용체 표적화된 분자, 경구, 직장, 국소 또는 다른 제형으로서 다른 분자, 분자 구조 또는 화합물의 혼합물과 혼합되거나, 캡슐화되거나, 접합되거나, 다르게는 회합될 수 있다. 이러한 흡수, 분포 및/또는 흡

수 보조 제형의 제조를 교시하는 대표적인 미국 특허는 미국 특허 제5,108,921호; 제5,354,844호; 제5,416,016호; 제5,459,127호; 제5,521,291호; 제5,543,158호; 제5,547,932호; 제5,583,020호; 제5,591,721호; 제4,426,330호; 제4,534,899호; 제5,013,556호; 제5,108,921호; 제5,213,804호; 제5,227,170호; 제5,264,221호; 제5,356,633호; 제5,395,619호; 제5,416,016호; 제5,417,978호; 제5,462,854호; 제5,469,854호; 제5,512,295호; 제5,527,528호; 제5,534,259호; 제5,543,152호; 제5,556,948호; 제5,580,575호; 및 제5,595,756호를 포함하며, 이들의 각각은 본원에 참조로 포함된다.

[0167] **Xist를 넉아웃시키기 위한 게놈 편집**

[0168] 유전자 편집은 기본 및 임상 과학 사이의 계면을 포괄하는 생물의학적 연구의 주요한 초점이다. 신규한 "유전자 편집" 도구의 발달은 게놈의 다른 부위에 돌연변이를 도입하지 않고, 특이적 염색체 좌위에서 세포의 DNA 서열을 조작하는 능력을 제공한다. 이 기술은 연구자들이 대상체의 세포의 게놈을 시험관내에서 또는 생체내에서 조작하는 것을 효과적으로 가능하게 한다. SCNT의 맥락에서, Xist에 KO를 포함하는 세포는 CRISPR을 사용하여 생성된다.

[0169] 한 실시양태에서, 유전자 편집은 엔도뉴클레아제 (DNA 분자 내에서 내부적으로 DNA 파단을 유발하는 효소)를 게놈의 특이적 부위에 표적화하고, 그에 의해 선택된 부위에서 염색체 이중 가닥 파단 (DSB)의 형성을 촉발시키는 것을 포함한다. 염색체 파단의 도입과 수반하여, 공여자 DNA 분자가 (예를 들어, 플라스미드 또는 올리고뉴클레오티드 도입에 의해) 도입되는 경우, 파단된 염색체 및 도입된 DNA 사이의 상호작용은 특히 2개의 서열이 상동성을 공유하는 경우 일어날 수 있다. 이 예에서, "유전자 표적화"로 용어화된 프로세스가 일어날 수 있으며, 여기서 염색체의 DNA 말단은 상동성 재조합 (HR)에 의해 공여자 DNA의 상동성 서열을 침입한다. 공여자 플라스미드 서열을 HR을 위한 주형으로서 사용함으로써, 관심의 유전자의 이음매 없는 넉아웃이 달성될 수 있다. 중요하게는, 공여자 DNA 분자가 표적 유전자 (예를 들어, *Xist*) 내에 결실을 포함하는 경우, HR-매개된 DSB 복구는 공여자 서열을 염색체 내로 도입하여, 결실이 염색체 좌위 내에 도입되는 것을 발생시킬 것이다. 뉴클레아제를 표적 유전자를 함유하는 게놈 부위에 표적화함으로써, 개념은 HR을 자극하고, 그에 의해 기능적 표적 유전자를 유전자의 결실된 형태로 대체하는 DSB 형성을 사용하는 것이다. HR 경로의 이점은 그것이 이전의 야생형 대립유전자 대신 유전자의 넉아웃을 이음매 없이 생성하는 잠재성을 갖는다는 것이다.

[0170] 현재의 게놈 편집 도구는 세포의 유전자 조작을 향상시키는 이중 가닥 파단 (DSB)의 도입을 사용한다. 이러한 방법으로는 아연 핑거 뉴클레아제 (ZFN; 예를 들어 본원에 참조로 포함되는 미국 특허 제6,534,261호; 제6,607,882호; 제6,746,838호; 제6,794,136호; 제6,824,978호; 제6,866,997호; 제6,933,113호; 제6,979,539호; 제7,013,219호; 제7,030,215호; 제7,220,719호; 제7,241,573호; 제7,241,574호; 제7,585,849호; 제7,595,376호; 제6,903,185호; 및 제6,479,626호; 및 미국 특허 공개 제20030232410호 및 제US2009020314호에 기재됨), 전사 활성화제-유사 이펙터 뉴클레아제 (Transcription Activator-Like Effector Nuclease) (TALEN; 예를 들어 본원에 참조로 포함되는 미국 특허 제8,440,431호, 제8,440,432호, 제8,450,471호, 제8,586,363호, 및 제8,697,853호, 및 미국 특허 공개 제20110145940호, 제20120178131호, 제20120178169호, 제20120214228호, 제20130122581호, 제20140335592호, 및 제20140335618호에 기재됨), 및 CRISPR (군집화된 규칙적으로 간격화된 짧은 회문구조 반복부)/Cas9 시스템 (예를 들어 본원에 참조로 포함되는 미국 특허 제8,697,359호, 제8,771,945호, 제8,795,965호, 제8,871,445호, 제8,889,356호, 제8,906,616호, 제8,932,814호, 제8,945,839호, 제8,993,233호, 및 제8,999,641호, 및 미국 특허 공개 제20140170753호, 제20140227787호, 제20140179006호, 제20140189896호, 제20140273231호, 제20140242664호, 제20140273232호, 제20150184139호, 제20150203872호, 제20150031134호, 제20150079681호, 제20150232882호, 및 제20150247150호에 기재됨)을 들 수 있다. 예를 들어, ZFN DNA 서열 인식 능력 및 특이성은 예측불가능할 수 있다. 유사하게, TALEN 및 CRISPR/Cas9는 바람직한 부위에서 뿐만 아니라 다른 "오프-타겟" 부위에서도 또한 절단한다. 이들 방법은 오프-타겟 이중-가닥 파단 유도, 및 이들 오프-타겟 효과와 연관된 indel, 게놈 재배열, 및 염색체 재배열을 비롯한 유해한 돌연변이에 대한 잠재성과 관련된 상당한 문제를 갖는다. ZFN 및 TALEN은 게놈에서 약 18 bp 서열에 대한 특이성을 생성하는 모듈식 서열-특이적 DNA 결합 단백질의 사용을 수반한다.

[0171] 유형 2 CRISPR (군집화된 규칙적으로 간격화된 짧은 회문구조 반복부)/Cas (CRISPR 연관된) 시스템에 기초한 RNA-가이드된 뉴클레아제-매개된 게놈 편집은 게놈을 변경시키기 위한 가치있는 접근법을 제공한다. 간략하게, 단일-가이드 RNA (sgRNA)에 의해 가이드된 뉴클레아제인 Cas9는 프로토스페이서 인접 모티프 (PAM) 다음의 표적화된 게놈 좌위에 결합하며, 이중-가닥 파단 (DSB)을 생성한다. 그 후, DSB는 삽입/결실 (indel) 돌연변이를 초래하는 비-상동성 말단 연결 (NHEJ)에 의해, 또는 외인성 주형을 요구하고, 표적 좌위에서 정확한 변형을 생성할 수 있는 상동성-지정 복구 (HDR)에 의해 복구된다 (Mali et al., Science. 2013 Feb 15;339(6121):823-

6). 기능적, 또는 부분적으로 기능적 유전자의 카피를 환자의 세포에 첨가하지만, 유전자의 원래 기능이상적 카피를 보유하는 다른 유전자 요법 방법과는 달리, 이 시스템은 결함을 제거할 수 있다. 조작된 뉴클레아제를 사용한 유전적 교정은 조직 배양 세포 및 희귀 질환의 설치류 모델에서 입증되었다.

[0172] CRISPR은 빵 효모 (에스. 세레비지아에 (S. cerevisiae)), 제브라 피쉬, 선충 (씨. 엘레간스 (C. elegans)), 식물, 마우스, 및 몇몇 다른 유기체를 비롯한 폭넓은 범위의 유기체에서 사용되었다. 추가적으로, CRISPR은 과학자들이 특이적 유전자를 표적화하고 활성화시키거나 침묵시키는 것을 허용하는 프로그래밍가능한 전사 인자를 제조하기 위해 변형되었다. 수 만 개의 가이드 RNA의 라이브러리가 현재 이용가능하다.

[0173] 2012년 이래로, CRISPR/Cas 시스템은 심지어 마우스 및 영장류와 같은 진핵생물에서도 작동하는 유전자 편집 (특이적 유전자를 침묵시키거나, 증진시키거나, 변화시킴)에 사용되었다. cas 유전자 및 특이적으로 설계된 CRISPR을 함유하는 플라스미드를 삽입함으로써, 유기체의 게놈은 임의의 바람직한 위치에서 커팅될 수 있다.

[0174] CRISPR 반복부는 크기로 24 내지 48 염기 쌍의 범위이다. 이들은 통상적으로 일부 이원 대칭을 나타내며, 이는 2차 구조, 예컨대 헤어핀의 형성을 암시하지만, 진정하게 회문구조는 아니다. 반복부는 유사한 길이의 스페이서에 의해 분리된다. 일부 CRISPR 스페이서 서열은 플라스미드 및 파지로부터의 서열과 정확하게 매칭하지만, 일부 스페이서는 원핵생물의 게놈과 매칭한다 (자기-표적화 스페이서). 새로운 스페이서는 파지 감염에 반응하여 급속하게 첨가될 수 있다.

[0175] CRISPR-연관된 (cas) 유전자는 종종 CRISPR 반복부-스페이서 어레이와 연관된다. 2013년 현재, 40개 초과의 상이한 Cas 단백질 족이 기재되었다. 이들 단백질 족 중에서, Cas1은 상이한 CRISPR/Cas 시스템 중에서 편재하는 것으로 보인다. cas 유전자 및 반복부 구조의 특정 조합은 8가지 CRISPR 하위유형 (Ecoli, Ypest, Nmeni, Dvulg, Tneap, Hmari, Apern, 및 M튜브)을 정의하는 데 사용되었으며, 이들 중 일부는 반복부-연관된 신비한 단백질 (RAMP)을 코딩하는 추가의 유전자 모듈과 연관된다. 1개 초과의 CRISPR 하위유형은 단일 게놈에서 발생할 수 있다. CRISPR/Cas 하위유형의 산발적 분포는 시스템이 미생물 진화 동안 수평적 유전자 전달에 처함을 시사한다.

[0176] 외인성 DNA는 일견하여 Cas 유전자에 의해 코딩되는 단백질에 의해 작은 요소 (길이로 약 30 염기 쌍)로 프로세싱되며, 이는 그 후 왜 그런지 리더 서열 부근의 CRISPR 좌위 내로 삽입된다. CRISPR 좌위로부터의 RNA는 구성적으로 발현되며, Cas 단백질에 의해 플랭킹 반복 서열을 갖는 개별적, 외인적으로-유래된 서열 요소로 구성된 작은 RNA로 프로세싱된다. RNA는 다른 Cas 단백질을 가이드하여 외인성 유전적 요소를 RNA 또는 DNA 수준에서 침묵시킨다. 증거는 CRISPR 하위유형 중에서의 기능적 다양성을 시사한다. Cse (Cas 하위유형 이콜라이 (Ecoli)) 단백질 (이. 콜라이 (E. coli)에서 CasA-E로 지칭됨)은 CRISPR RNA 전사체를 캐스케이드 (Cascade)가 보유하는 스페이서-반복 단위로 프로세싱하는 기능적 복합체, 캐스케이드를 형성한다. 다른 원핵생물에서, Cas6은 CRISPR 전사체를 프로세싱한다. 흥미롭게도, 이. 콜라이에서의 CRISPR-기재 파지 불활성화는 캐스케이드 및 Cas3을 요구하지만, Cas1 및 Cas2는 요구하지 않는다. 피로코쿠스 푸리오수스 (Pyrococcus furiosus) 및 다른 원핵생물에서 발견되는 Cmr (Cas RAMP 모듈) 단백질은 상보성 표적 RNA를 인식하고 절단하는 작은 CRISPR RNA와 기능적 복합체를 형성한다. RNA-가이드된 CRISPR 효소는 유형 V 제한 효소로서 분류된다.

[0177] 또한, 그 전문이 참조로 포함되는 미국 특허 공개 2014/0068797을 참조한다.

[0178] *Cas9*

[0179] Cas9는 DNA를 커팅하기 위해 특수화된 효소인 뉴클레아제이며, 하나가 이중 나선의 각각의 가닥에 대한 것인 2개의 활성 커팅 부위를 갖는다. 팀은 이들이 위치된 그의 표적 DNA로 귀소하는 Cas9의 능력을 보존하면서, 하나 또는 둘 다의 부위를 망가뜨릴 수 있음을 입증하였다. 지넥 (Jinek) 등 (2012)은 tracrRNA 및 스페이서 RNA를 Cas9와 혼합되어 정확한 DNA 표적을 발견하고 커팅할 수 있는 "단일-가이드 RNA" 분자로 조합하였다. 이러한 합성 가이드 RNA는 유전자 편집에 사용될 수 있음이 제안되었다 (Jinek et al., Science. 2012 Aug 17;337(6096):816-21).

[0180] Cas9 단백질은 병원성 및 공생 박테리아에 매우 풍부하다. CRISPR/Cas-매개된 유전자 조절은 특히 진핵생물 숙주와의 박테리아 상호작용 동안 내인성 박테리아 유전자의 조절에 기여할 수 있다. 예를 들어, 프란시셀라 노비시다 (Francisella novicida)의 Cas 단백질 Cas9는 에프. 논비시다가 숙주 반응을 약화시키고 병독성을 촉진시키는 데 중요한 박테리아 지질단백질을 코딩하는 내인성 전사체를 억제하는 독특한 작은 CRISPR/Cas-연관된 RNA (scaRNA)를 사용한다. 배선 (접합자) 내로의 Cas9 mRNA 및 sgRNA의 공동주사는 돌연변이를 갖는 마우스를 생성하였다. Cas9 DNA 서열의 전달은 또한 고려된다.

[0181] *gRNA*

[0182] RNA 가이드된 단백질로서, Cas9는 DNA 표적의 인식을 지정하는 짧은 RNA를 요구한다. Cas9는 PAM 서열 NGG를 함유하는 DNA 서열을 우선적으로 추궁하지만, 이는 여기서 프로토스페이서 표적 없이 결합할 수 있다. 그러나, Cas9-gRNA 복합체는 이중 가닥 파단을 생성하기 위해 gRNA에 가까운 매칭을 요구한다. 박테리아에서 CRISPR 서열은 다중 RNA에서 발현되고, 그 후 프로세싱되어 RNA에 대한 가이드 가닥을 생성한다. 진핵생물 시스템은 CRISPR RNA를 프로세싱하는 데 요구되는 단백질의 일부가 결여되기 때문에, 발현된 단일 RNA 내로의 Cas9 표적화를 위한 RNA의 필수 조각을 RNA 폴리머라제 유형 2I 프로모터 U6)과 조합하는 합성 구축물 gRNA를 생성하였다. 합성 gRNA는 최소 길이에서 100 bp 약간 초과이며, PAM 서열 NGG 바로 앞의 20 프로토스페이서 뉴클레오티드를 표적화하는 부분을 함유한다; gRNA는 PAM 서열을 함유하지 않는다.

[0183] 한 접근법에서, 대상체의 1개 이상의 세포는 CRISPR-Cas 시스템을 사용하여 *Xist*를 결실시키거나 불활성화시키도록 변경된다. Cas9는 *Xist* 유전자를 표적화하는 데 사용될 수 있다. 표적 인식 시, Cas9는 *Xist* 표적 유전자에서 이중 가닥 파단을 유도한다. 이중-가닥 파단 부위에서 상동성-지정 복구 (HDR)는 *Xist* 서열의 불활성 또는 결실된 형태의 삽입을 허용할 수 있다.

[0184] 하기 US 특허 및 특허 공개는 본원에 참조로 포함된다: 특허 번호 8,697,359, 20140170753, 20140179006, 20140179770, 20140186843, 20140186958, 20140189896, 20140227787, 20140242664, 20140248702, 20140256046, 20140273230, 20140273233, 20140273234, 20140295556, 20140295557, 20140310830, 20140356956, 20140356959, 20140357530, 20150020223, 20150031132, 20150031133, 20150031134, 20150044191, 20150044192, 20150045546, 20150050699, 20150056705, 20150071898, 20150071899, 20150071903, 20150079681, 20150159172, 20150165054, 20150166980, 및 20150184139.

[0185] **치료 방법**

[0186] 각인 제어 영역에 존재하는 H3K27me3 각인을 조정하는 작용제는 SCNT를 사용하여 클로닝된 완전한 만기 유기체를 생성하는 데 유용하다. H3K27me3 각인을 첨가하는 작용제는 *Kdm4d* 폴리뉴클레오티드로 주사된 Xist KO 세포와 조합으로 사용될 수 있다.

[0187] 한 접근법에서, H3K27me3 데메틸라제를 억제하는 작용제는 SCNT와 조합으로 사용된다. 투여되는 작용제의 투여량은 개별적 환자의 크기 및 건강을 비롯한 다수의 인자에 의존한다. 임의의 특정 대상체에 대해, 구체적인 투여량 처방은 개별적 필요 및 조성물의 투여를 투여하거나 감독하는 사람의 전문적인 판단에 따라 시간 경과에 따라 조정되어야 한다.

[0188] 하기 실시예는 관련 기술분야의 통상의 기술자에게 본 발명의 검정, 스크리닝, 및 치료 방법을 수행하고 사용하는 방법의 완전한 개시 및 설명을 제공하기 위해 제시되며, 본 발명자들이 그들의 발명으로 간주하는 것의 범위를 제한하는 것으로 의도되지 않는다.

[0189] **체세포 핵 전달**

[0190] 체세포 핵 전달 (SCNT)은 예를 들어, 가축 (예를 들어, 소, 말, 양, 염소, 돼지)의 생식적 클로닝에, 또는 바람직한 조직이 세포 대체 요법을 위해 생성되는 치료적 클로닝에 사용될 수 있는 기법이다. 불행하게도, 클로닝된 동물은 부적절한 각인으로부터 발생하는 특정 결함, 예컨대 히스톤 H3 단백질 서브유닛 상의 리신 27의 트리메틸화의 결핍을 겪는다. 이 결핍은 SCNT 절차 동안 트리메틸화 사건을 수행하는 효소를 코딩하는 mRNA를 제공함으로써 개선될 수 있다. 한 실시양태에서, 트리메틸화 사건을 수행할 수 있는 효소 (예를 들어, EZH1, EZH2, PRC2)를 코딩하는 mRNA는 SCNT 절차 전에 또는 동안 수용자 세포 또는 핵 공여자 세포 내로 주사된다.

[0191] 체세포 핵 전달은 핵 공여자 세포를 얻고, 그 후 이 핵 공여자 세포를 제핵된 수용자 세포, 가장 바람직하게는 제핵된 난모세포 내로 융합시켜 핵 전달 배아를 형성하고, 이 배아를 활성화시키고, 마지막으로 배아를 배양하거나, 이 배아를 모계 숙주 내로 전달하는 것을 포함한다. 핵 전달 동안, 한 세포로부터의 핵 DNA의 완전한 상보체는 제핵된 세포에 도입된다. 핵 전달 방법은 관련 기술분야의 통상의 기술자에게 널리 공지되어 있다. 1991년 2월 19일에 허여된 발명의 명칭이 "소 배아의 증식"인 프라더 (Prather) 등의 미국 특허 제4,994,384호; 1991년 10월 15일에 허여된 발명의 명칭이 "소 핵 이식"인 매시 (Massey)의 미국 특허 제5,057,420호; 1999년 11월 30일에 스티스 (Stice) 등에게 허여된 발명의 명칭이 "배양된 내부 세포 덩이 세포를 사용한 키메라 소 또는 돼지 동물의 제조"인 미국 특허 제5,994,619호; 2000년 1월 19일에 캠벨 (Campbell) 등 및 윌무트 (Wilmut) 등에게 각각 허여된 발명의 명칭이 "핵 전달을 위한 정지 세포 집단"인 영국 특허 제GB 2,318,578호, 제GB

2,331,751호; 2000년 1월 4일에 허여된 발명의 명칭이 "재프로그래밍된 비-배아 소 세포를 사용하여 소를 클로닝하는 방법"인 스트렐켄코 (Strelchenko) 등의 미국 특허 제6,011,197호; 및 발명의 명칭이 "돼지 동물을 클로닝하는 방법"인 미국 특허 출원 일련 번호 09/753,323 (대리인 정리 번호 030653.0026.CIP1, 2000년 12월 28일에 출원됨)을 참조하며, 이들의 각각은 모든 도면, 표 및 도면을 비롯하여 그 전문이 본원에 참조로 포함된다. 핵 전달은 투명대에 의해 둘러싸이지 않은 난모세포를 사용함으로써 달성될 수 있다.

[0192] 핵 전달 절차에서, 핵 공여자 세포, 또는 그의 핵은 수용자 세포 내로 도입된다. 수용자 세포는 바람직하게는 난모세포이고, 바람직하게는 제핵된다. 그러나, 본 발명은 부분적으로 난모세포의 핵이 난모세포로부터 물리적으로 추출되지 않는 핵 전달에 관한 것이다. 공여자 세포로부터의 핵 DNA가 세포 분열 동안 복제되는 핵 전달 배아를 확립하는 것이 가능하다. 예를 들어, 문헌 [Wagoner et al., 1996, "Functional enucleation of bovine oocytes: effects of centrifugation and ultraviolet light," Theriogenology 46: 279-284]을 참조한다. 또한, 핵 전달은 1개의 핵 공여자 및 1개 초과의 제핵된 난모세포를 조합함으로써 달성될 수 있다. 또한, 핵 전달은 1개의 핵 공여자, 1개 이상의 제핵된 난모세포, 및 1개 이상의 제핵된 난모세포의 세포질을 조합함으로써 달성될 수 있다. 핵 공여자 세포 및 수용자 세포의 생성된 조합은 "하이브리드 세포"로 지칭될 수 있다.

[0193] 본원에 사용된 바와 같은 용어 "핵 공여자"는 난모세포 내로 전위될 수 있는 핵 DNA를 갖는 임의의 세포, 또는 그의 핵을 지칭한다. 핵 공여자는 세포로부터 단리된 핵일 수 있다. 세포로부터 핵을 단리한 후, 핵을 핵 공여자로서 이용하기 위한 다수의 기법이 관련 기술분야의 통상의 기술자에게 이용가능하다. 예를 들어, 미국 특허 제4,664,097호, 제6,011,197호, 및 제6,107,543호를 참조하며, 이들의 각각은 그 전문이 본원에 참조로 포함된다. 임의의 유형의 세포는 핵 공여자로서 기능할 수 있다. 핵 공여자 세포의 예로는 시험관내에서 또는 생체내에서 2개의 배우자의 결합으로부터 발생한 배아로부터 단리된 배양된 및 비-배양된 세포; 배양된 배아 세포 (예를 들어, 사전-배반포 세포 및 내부 세포 덩이 세포)로부터 발생한 배아 줄기 세포 (ES 세포); 배아로부터 단리된 내부 세포 덩이 세포로부터 발생한 배양된 및 비-배양된 세포; 배양된 및 비-배양된 사전-배반포 세포; 배양된 및 비-배양된 태아 세포; 배양된 및 비-배양된 성체 세포; 배양된 및 비-배양된 원시 배 세포; 배양된 및 비-배양된 배 세포 (예를 들어, 배아 배 세포); 동물로부터 단리된 배양된 및 비-배양된 체세포; 배양된 및 비-배양된 난구 세포; 배양된 및 비-배양된 양막 세포; 배양된 및 비-배양된 태아 섬유모세포 세포; 배양된 및 비-배양된 생식 융기 세포; 배양된 및 비-배양된 분화된 세포; 동기 집단에서의 배양된 및 비-배양된 세포; 비동기 집단에서의 배양된 및 비-배양된 세포; 배양된 및 비-배양된 혈청-기아 세포; 배양된 및 비-배양된 영구 세포; 및 배양된 및 비-배양된 전능성 세포를 들 수 있으나, 이에 제한되지는 않는다. 예를 들어, 문헌 [Piedrahita et al., 1998, Biol. Reprod. 58: 1321-1329]; [Shim et al., 1997, Biol. Reprod. 57: 1089-1095]; [Tsung et al., 1995, Shih Yen Sheng Wu Hsueh Pao 28: 173-189]; 및 [Wheeler, 1994, Reprod. Fertil. Dev. 6: 563-568]을 참조하며, 이들의 각각은 모든 도면, 도면, 및 표를 비롯한 그 전문이 본원에 참조로 포함된다. 또한, 핵 공여자는 이전에 동결되거나 동결보존된 세포일 수 있다.

[0194] 핵 전달의 프로세스에 의해 제조된 하이브리드 세포는 예를 들어, 생식적 클로닝에 또는 재생적 클로닝에 사용될 수 있다.

[0195] SCNT 실험은 성체 분화된 체세포로부터의 핵이 전능한 상태로 재프로그래밍될 수 있음을 나타내었다. 따라서, 본원에 개시된 바와 같은 방법을 사용하여 생성된 SCNT 배아는 전능성 또는 배아 줄기 세포 또는 줄기-유사 세포 및 세포 콜로니의 생성을 위해 적합한 시험관내 배양 배지에서 배양될 수 있다. 배아의 배양 및 성숙에 적합한 배양 배지는 관련 기술분야에 널리 공지되어 있다. 소 배아 배양 및 유지에 사용될 수 있는 공지된 배지의 예로는 햄 (Ham) F-10+10% 소 태아 혈청 (FCS), 조직 배양 배지 (Tissue Culture Medium)-199 (TCM-199)+10% 소 태아 혈청, 티로드-알부민-락테이트-피루베이트 (Tyrodes-Albumin-Lactate-Pyruvate) (TALP), 둘베코 포스페이트 완충 염수 (Dulbecco's Phosphate Buffered Saline) (PBS), 이글 (Eagle) 및 위튼 (Whitten) 배지를 들 수 있다. 난모세포의 수집 및 성숙에 사용되는 가장 통상적인 배지 중 하나는 TCM-199, 및 소 태아 혈청, 신생아 혈청, 발정기 소 혈청, 양 혈청 또는 거세수소 혈청을 비롯한 1 내지 20% 혈청 보충물이다. 바람직한 유지 배지는 얼 (Earl) 염, 10% 소 태아 혈청, 0.2 Ma 피루베이트 및 50 ug/ml 겐타미신 술페이트를 갖는 TCM-199를 포함한다. 상기 중 임의의 것은 또한 다양한 세포 유형, 예컨대 과립막 세포, 난관 세포, BRL 세포 및 자궁 세포 및 STO 세포와의 공동-배양을 포함할 수 있다.

[0196] 특히, 자궁내막의 상피 세포는 착상전 및 착상 기간 동안 백혈병 억제성 인자 (LIF)를 분비한다. 따라서, 일부 실시양태에서, 배양 배지에의 LIF의 첨가는 SCNT-유래된 배아의 시험관내 발생을 향상시키기 위해 포함된다. 배아 또는 줄기-유사 세포 배양을 위한 LIF의 사용은 미국 특허 제5,712,156호에 기재되었으며, 이는 본원에 참

조로 포함된다.

[0197] 또다른 유지 배지는 본원에 참조로 포함되는 로센크란스, 주니어 (Rosenkrans, Jr.) 등의 미국 특허 제 5,096,822호에 기재되어 있다. CR1로 명명된 이 배아 배지는 배아를 지지하는 데 필요한 영양 물질을 함유한다. CR1은 1.0 mM 내지 10 mM, 바람직하게는 1.0 mM 내지 5.0 mM 범위의 양으로 헤미칼슘 L-락테이트를 함유한다. 헤미칼슘 L-락테이트는 그 상에 혼입된 헤미칼슘 염을 갖는 L-락테이트이다. 또한, 문헌 [Thomson et al., Science, 282:1145-1147 (1998)] 및 [Proc. Natl. Acad. Sci., USA, 92:7844-7848 (1995)]에 논의된 바와 같은 인간 배아 줄기 세포를 유지하기 위한 적합한 배양 배지.

[0198] 일부 실시양태에서, 공급자 세포는 마우스 배아 섬유모세포를 포함할 것이다. 적합한 섬유모세포 공급자 층의 제조를 위한 수단은 이어지는 실시예에 기재되어 있으며, 통상의 기술자의 기술 내에 있다.

[0199] 배반포-단계 SCNT 배아 (또는 그의 등가물)로부터 ES 세포 (예를 들어, NT-ESC 또는 hNT-ESC)를 유도하는 방법은 관련 기술분야에 널리 공지되어 있다. 이러한 기법은 SCNT 배아로부터 ES 세포 (예를 들어, hNT-ESC)를 유도하는 데 사용될 수 있으며, 여기서, hNT-ESC를 생성하는 데 사용되는 SCNT 배아는 KDM4 데메틸라제 족의 구성원 및/또는 히스톤 메틸트랜스퍼라제 SUV39h1/SUV39h2의 억제제로 처리되지 않은 SCNT에 비해, 체세포 공여자 세포로부터 공여된 핵 유전 물질에서 감소된 수준의 H3K9me3을 갖는다. 추가적으로 또는 대안적으로, hNT-ESC는 발생의 보다 초기 단계 동안 클로닝된 SCNT 배아로부터 유도될 수 있다.

[0200] 특정 실시양태에서, 본원에 개시된 바와 같은 방법, 조성물 및 키트를 사용하여 생성된 SCNT 배아로부터 생성된 할구는 유리 피펫을 사용하여 분리되어 전능성 세포를 얻을 수 있다. 일부 실시양태에서, 분리는 0.25% 트립신의 존재 하에서 일어날 수 있다 (Collas and Robl, 43 BIOL. REPROD. 877-84, 1992; Stice and Robl, 39 BIOL. REPROD. 657-664, 1988; Kanka et al., 43 MOL. REPROD. DEV. 135-44, 1996).

[0201] 특정 실시양태에서, SCNT 배아로부터의 생성된 배반포, 또는 배반포-유사 클러스터는 배아 줄기 세포주, 예를 들어, 핵 전달 ESC (ntESC) 세포주를 얻는 데 사용될 수 있다. 이러한 주는 예를 들어, 그 전문이 본원에 참조로 포함되는 문헌 [Thomson et al., Science, 282:1145-1147 (1998)] 및 [Thomson et al., Proc. Natl. Acad. Sci., USA, 92:7544-7848 (1995)]에 의해 보고된 배양 방법에 따라 얻어질 수 있다.

[0202] 다능성 배아 줄기 세포는 또한 배아의 정상적인 발생에서 출생까지를 방해하지 않고 SCNT 배아로부터 제거된 단일 할구로부터 생성될 수 있다. 2004년 11월 4일에 출원된 미국 출원 제60/624,827호; 2005년 3월 14일에 출원된 제60/662,489호; 2005년 6월 3일에 출원된 제60/687,158호; 2005년 10월 3일에 출원된 제60/723,066호; 2005년 10월 14일에 출원된 제60/726,775호; 2005년 11월 4일에 출원된 제11/267,555호; 2005년 11월 4일에 출원된 PCT 출원 제PCT/US05/39776호를 참조하며, 이들의 개시내용은 그 전문이 본원에 참조로 포함된다; 또한 문헌 [Chung et al., Nature, Oct. 16, 2005 (인쇄물에 앞서 전자적으로 공개됨) 및 [Chung et al., Nature V. 439, pp. 216-219 (2006)]을 참조하며, 이들의 각각의 전체 개시내용은 그 전문이 참조로 포함된다). 이러한 경우, SCNT 배아는 다능성 줄기 세포의 생성을 위해 파괴되지 않는다.

[0203] 본 발명의 한 측면에서, 방법은 연구에서의 및 요법에서의 SCNT 배아로부터 유래된 세포의 이용을 포함한다. 이러한 다능성 줄기 세포 (PSC) 또는 전능성 줄기 세포 (TSC)는 제한 없이, 피부, 연골, 골, 골격근, 심장근, 신장, 간, 혈액 및 혈액 형성, 혈관 전구체 및 혈관 내피, 췌장 베타, 뉴런, 교질, 망막, 내이 모낭, 장, 또는 폐 세포를 비롯한 신체 내의 임의의 세포로 분화될 수 있다.

[0204] 본 발명의 또다른 실시양태에서, SCNT 배아, 또는 배반포, 또는 SCNT 배아로부터 얻어진 다능성 또는 전능성 세포 (예를 들어, NT-ESC)는 1종 이상의 분화의 유도제에 노출되어 다른 치료적으로 유용한 세포, 예컨대 망막 색소 상피, 조혈 전구체 및 혈관모세포 전구체 뿐만 아니라 외배엽, 중배엽, 및 내배엽의 많은 다른 유용한 세포 유형을 수득할 수 있다. 이러한 유도제는 시토카인, 예컨대 인터류킨-알파 A, 인터페론-알파 A/D, 인터페론-베타, 인터페론-감마, 인터페론-감마-유도성 단백질-10, 인터류킨-1 내지 17, 각질세포 성장 인자, 렙틴, 백혈병 억제성 인자, 대식세포 콜로니-자극 인자, 및 대식세포 염증성 단백질-1 알파, 1-베타, 2, 3 알파, 3 베타, 및 단핵구 화학주성 단백질 1 내지 3, 6카인, 액티빈 A, 암피레귤린, 안지오게닌, B-내피 세포 성장 인자, 베타 셀룰린, 뇌-유래된 신경영양 인자, C10, 카디오트로핀-1, 섬모 신경영양 인자, 시토카인-유도된 호중구 화학유인 물질-1, 에오탁신, 표피 성장 인자, 상피 호중구 활성화 펩티드-78, 에리트로포이에틴, 에스트로겐 수용체-알파, 에스트로겐 수용체-베타, 섬유모세포 성장 인자 (산성 및 염기성), 헤파린, FLT-3/FLK-2 리간드, 아교 세포주-유래된 신경영양 인자, Gly-His-Lys, 과립구 콜로니 자극 인자, 과립구대식세포 콜로니 자극 인자, GRO-알파/MGSA, GRO-베타, GRO-감마, HCC-1, 헤파린-결합 표피 성장 인자, 간세포 성장 인자, 헤레귤린-알파,

인슐린, 인슐린 성장 인자 결합 단백질-1, 인슐린-유사 성장 인자 결합 단백질-1, 인슐린-유사 성장 인자, 인슐린-유사 성장 인자 II, 신경 성장 인자, 뉴로토핀-3,4, 온코스타틴 M, 태반 성장 인자, 플레이오트로핀, 란테스, 줄기 세포 인자, 기질 세포-유래된 인자 1B, 트로모포이에틴, 전환 성장 인자--(알파, 베타 1,2,3,4,5), 종양 괴사 인자 (알파 및 베타), 혈관 내피 성장 인자, 및 골 형태형성 단백질, 호르몬 및 호르몬 길항제, 예컨대 17B-에스트라디올, 부신피질자극 호르몬, 아드레노메둘린, 알파-멜라닌세포 자극 호르몬, 융모막성 고나도트로핀, 코르티코스테로이드-결합 글로불린, 코르티코스테론, 덱사메타손, 에스트리올, 난포 자극 호르몬, 가스트린 1, 글루카곤, 고나도트로핀, L-3,3',5'-트리아이오도티로닌, 황체형성 호르몬, L-티록신, 멜라토닌, MZ-4, 옥시토신, 부갑상선 호르몬, PEC-60, 뇌하수체 성장 호르몬, 프로게스테론, 프로락틴, 세크레틴, 성 호르몬 결합 글로불린, 갑상선 자극 호르몬, 티로트로핀 방출 인자, 티록신-결합 글로불린, 및 바소프레신의 발현을 변경시키는 효소, 세포외 매트릭스 성분, 예컨대 피브로넥틴, 피브로넥틴의 단백질분해적 단편, 라미닌, 테나신, 트롬보스폰딘, 및 프로테오글리칸, 예컨대 아그레칸, 헤파란 술페이트 프로테오글리칸, 콘트로이틴 술페이트 프로테오글리칸, 및 신데칸을 포함하나, 이에 제한되지는 않는다. 다른 유도제는 본 발명의 재프로그래밍된 세포로부터 유래된 분화 세포에 대한 유도성 신호를 제공하는 데 사용되는 한정된 조직으로부터의 세포로부터 유래된 세포 또는 성분을 포함한다. 이러한 유도제 세포는 인간, 비-인간 포유동물, 또는 조류, 예컨대 특이적 병원체-무함유 (SPF) 배아 또는 성체 세포로부터 유래될 수 있다.

[0205] 할구 배양. 한 실시양태에서, SCNT 배아는 할구를 생성하고, SCNT 배아를 파괴하거나, 다르게는 그들의 생존성을 유의하게 변경시키지 않고, 본원에 개시된 바와 같은 방법에 의해 생성된 SCNT 배아로부터 단일 할구를 단리하기 위해 착상전 유전적 진단 (PGD)에 현재 사용되는 것들과 관련된 시험관내 기법을 이용하는 데 사용될 수 있다. 본원에서 입증된 바와 같이, 다능성 배아 줄기 (hES) 세포 및 세포주는 배아의 정상적인 발생에서 출생까지를 방해하지 않고 본원에 개시된 바와 같은 SCNT 배아로부터 제거된 단일 할구로부터 생성될 수 있다.

[0206] hESC를 유도하는 데 있어서 톰슨 (Thomson) 등 (Thomson et al., Science 282, 1145 (1998))과 함께 "돌리 (Dolly)"의 양 클로닝에서 윌무트 (Wilmut) 등 (Wilmut, et al., Nature 385, 810 (1997)의 발견은 SCNT-배아 또는 환자 자신의 핵으로부터 생성된 SCNT-조작된 세포 덩이로부터 유래된 환자-특이적 hESC의 확립에 기초한 재생적 세포 이식을 위한 상당한 열광을 발생시켰다. 자가 이식을 통해 면역 거부를 회피하는 데 목표를 둔 이 전략은 아마 SCNT에 대한 가장 강한 임상적 근거이다. 마찬가지로, 복잡한 질환-특이적 SCNT-hESC의 유도는 질환 메커니즘의 발견을 가속화할 수 있다. 세포 이식을 위해, 개별적 마우스 자신의 SCNT-유래된 mESC로의 뮤린 SCID 및 PD 모델의 혁신적인 치료는 고무적이다 (Rideout et al., Cell 109, 17 (2002); Barberi, Nat. Biotechnol. 21, 1200 (2003)). 궁극적으로, 폭넓은 조직 적합성을 갖는 SCNT-유래된 줄기 세포의 뱅크를 생성하는 능력은 새로운 난모세포의 계속되는 공급에 대한 필요를 감소시킬 것이다.

[0207] 본 발명의 특정 실시양태에서, SCNT 배아로부터 얻어진 다능성 또는 전능성 세포 (예를 들어, hNT-ESC)는 치료적 유용성을 나타내기 위해 임의로 분화되고, 이들이 정상적으로 거주하는 조직 내로 도입될 수 있다. 예를 들어, SCNT 배아로부터 얻어진 다능성 또는 전능성 세포는 조직 내로 도입될 수 있다. 특정 다른 실시양태에서, SCNT 배아로부터 얻어진 다능성 또는 전능성 세포는 전신적으로 또는 치료적 유용성이 바람직한 부위로부터 거리를 두고 도입될 수 있다. 이러한 실시양태에서, SCNT 배아로부터 얻어진 다능성 또는 전능성 세포는 거리를 두고 작용할 수 있거나, 바람직한 부위로 향할 수 있다.

[0208] 본 발명의 특정 실시양태에서, SCNT 배아로부터 얻어진 클로닝된 세포, 다능성 또는 전능성 세포는 다른 다능성 줄기 세포의 분화를 유도하는 데 이용될 수 있다. 유전자 발현의 배아 패턴을 유지하면서 시험관내에서 번식할 수 있는 세포의 단일 세포-유래된 집단의 생성은 다른 다능성 줄기 세포의 분화를 유도하는 데 유용하다. 세포-세포 유도는 초기 배아에서의 분화를 지정하는 통상적인 수단이다. 많은 잠재적으로 의학적으로 유용한 세포 유형은 척수 뉴런, 심장 세포, 췌장 베타 세포, 및 최종적인 조혈 세포를 비롯한 정상적인 배아 발생 동안 유도적 신호에 의해 영향을 받는다. 유전자 발현의 배아 패턴을 유지하면서 시험관내에서 번식될 수 있는 세포의 단일 세포-유래된 집단은 다른 다능성 줄기 세포의 분화를 유도하는 다양한 시험관내, 난자내, 또는 생체내 배양 조건에서 배양되어 바람직한 세포 또는 조직 유형이 될 수 있다.

[0209] SCNT 배아로부터 얻어진 다능성 또는 전능성 세포 (예를 들어, ntESC)는 임의의 바람직한 분화된 세포 유형을 얻는 데 사용될 수 있다. 이러한 분화된 세포의 치료적 용법은 독보적이다. 본원에서 논의된 바와 같이, 공여자 세포, 또는 수용자 난모세포, 하이브리드 난모세포 또는 SCNT 배아는 본원에서 논의된 바와 같은 방법에 따라 KDM4 히스톤 디메틸라제 활성화제 및/또는 H3K9 메틸트랜스퍼라제 억제제로 처리될 수 있다.

[0210] 대안적으로, 공여자 세포는 장애를 갖는 대상체로부터의 성체 체세포일 수 있으며, 생성된 SCNT 배아는 질환의

동물 모델 또는 질환의 세포 모델을 생성하는 분화 조건 하에서 배양될 수 있는 질환-특이적 다능성 또는 전능성 세포를 생성하는 데 사용될 수 있다. 본 발명의 큰 이점은 SCNT의 효율을 증가시킴으로써, 그것이 특히 유도된 다능성 줄기 세포가 아닌 (예를 들어, iPSC가 아닌) 다능성인 동종 또는 동계 ES 세포의 본질적으로 제한없는 공급을 제공한다는 것이다. 이러한 NT-ESC는, 이들이 부분적으로 다능성이 아니며, 그들의 재프로그래밍을 지정하기 위해 바이러스 트랜스진 또는 재프로그래밍 인자의 강제된 발현을 갖지 않기 때문에, iPSC에 비해 이점을 가지며, 이식에 적합하다.

[0211] 일부 실시양태에서, SCNT로부터 생성된 NT-ESC는 SCNT 배아로부터 얻어진 환자-특이적 전능성이며, 여기서, 공여자 세포는 다능성 줄기 세포 또는 그의 분화된 자손으로 처리되는 대상체로부터 얻어졌다. 따라서, 이는 현재의 이식 방법과 연관된 상당한 문제, 즉, 숙주-대-이식편 또는 이식편-대-숙주 거부 때문에 일어날 수 있는 이식된 조직의 거부를 제거할 것이다. 통상적으로, 거부는 항-거부 약물, 예컨대 시클로스포린의 투여에 의해 방지되거나 감소된다. 그러나, 이러한 약물은 상당한 유해한 부작용, 예를 들어, 면역억제, 발암 특성을 가질 뿐만 아니라 매우 고비용이다. 본 발명은 항-거부 약물, 예컨대 시클로스포린, 이뮬란, FK-506, 글루코코르티코이드, 및 라파마이신, 및 그의 유도체에 대한 필요를 제거하거나, 적어도 크게 감소시킬 것이다.

[0212] 본 발명의 실시는 달리 지시되지 않는다면 분자 생물학 (재조합 기법을 포함함), 미생물학, 세포 생물학, 생화학, 및 면역학의 통상적인 기법을 채용하며, 이는 널리 통상의 기술자의 범위 내에 있다. 이러한 기법은 문헌, 예컨대, 문헌 ["Molecular Cloning: A Laboratory Manual", second edition (Sambrook, 1989)]; ["Oligonucleotide Synthesis" (Gait, 1984)]; ["Animal Cell Culture" (Freshney, 1987)]; ["Methods in Enzymology" "Handbook of Experimental Immunology" (Weir, 1996)]; ["Gene Transfer Vectors for Mammalian Cells" (Miller and Calos, 1987)]; ["Current Protocols in Molecular Biology" (Ausubel, 1987)]; ["PCR: The Polymerase Chain Reaction", (Mullis, 1994)]; ["Current Protocols in Immunology" (Coligan, 1991)]에 충분히 설명되어 있다. 이들 기법은 본 발명의 폴리뉴클레오티드 및 폴리펩티드의 제조에 적용하능하며, 따라서, 본 발명을 수행하고 실시하는 데 있어서 고려될 수 있다. 특정 실시양태에 특히 유용한 기법은 이어지는 섹션에서 논의될 것이다.

[0213] 하기 실시예는 관련 기술분야의 통상의 기술자에게 본 발명의 검정, 스크리닝, 및 치료 방법을 수행하고 사용하는 방법의 완전한 개시 및 설명을 제공하기 위해 제시되며, 본 발명자들이 그들의 발명으로 간주하는 것의 범위를 제한하는 것으로 의도되지 않는다.

[0214] **실시예**

[0215] **실시예 1: *Kdm4d* 주사는 SCNT-연관된 비정상적 *Xist* 활성화를 완화시키지 않는다**

[0216] 불활성 X 염색체는 항-H3K27me3 항체로의 그의 점상 염색에 의해 표시된다. 일관되게, 이러한 점상 염색은 시험관내에서 수정된 (IVF) 배아에서 단지 자성 (XX) 세포에서만 검출가능하지만, 웅성 (XY) 세포에서는 그렇지 않다. 대조적으로, 이러한 점상 염색은 웅성 세르톨리 세포가 공여자 세포로서 사용된 경우 SCNT 배아에서 관찰되었으며 (도 1a), 이는 SCNT 배아에서의 비정상적 *Xist* 활성화를 시사한다. 중요하게는, *Kdm4d* 주사는 주사 없는 대조군 SCNT 배아에 비해 점상 염색 패턴 또는 그의 빈도를 변경시키지 않는다 (도 1a, 1b). 이들 결과는 공여자 세포에서의 H3K9me3 및 SCNT 배아에서의 *Xist*의 이소적 활성화가 SCNT 재프로그래밍에 있어서 2가지 독립적인 장벽임을 입증하였다.

[0217] **실시예 2: *Xist* 돌연변이체 공여자 세포 및 *Kdm4d* mRNA 주사의 조합적 사용은 클로닝 효율을 크게 개선시킨다**

[0218] 2가지 재프로그래밍 장벽이 서로에 독립적이라는 사실은 *Xist* KO 공여자 세포를 사용하고 *Kdm4d*를 주사하는 것의 조합된 접근법이 증가된 클로닝 효율을 달성하는 상승적 또는 상가적 효과를 가질 것인지 여부의 질문을 촉발시켰다. 공여자로서 난구 세포를 사용한 SCNT가 시도되었다. B6D2F1 배경을 갖는 야생형 대조군에서, 대리모에 전달된 배아 중 단지 1.2%만이 만기에 도달하였다 (도 1c; 표 2).

[0219] <표 2> SCNT 배아의 착상후 발생, 도 1과 관련됨

수용자의 수	전달된 2-세포 배아의 수	착상된 수 (ET 당 %)	새끼의 수 (ET 당 %)	출생시 체중 (g ± SD)	출생시 태반 중량 (g ± SD)
6	171	52 (30.4)	2 (1.2)	1.56 ± 0.04	0.36 ± 0.06
8	179	110 (61.5)	15 (8.4)	1.59 ± 0.14	0.34 ± 0.06
4	75	46 (61.3)	14 (18.7)	1.50 ± 0.12	0.30 ± 0.04
3	55	25 (45.5)	1 (1.8)	1.50	0.28
4	77	47 (61.0)	7 (9.1)	1.48 ± 0.11	0.26 ± 0.10
4	85	57 (67.1)	20 (23.5)	1.46 ± 0.16	0.27 ± 0.08
5	40	15 (37.5)	0 (0.0)	N/A	N/A
5	53	36 (67.9)	2 (3.8)	1.70 ± 0.18	0.38 ± 0.04
5	29	23 (79.3)	2 (6.9)	1.89 ± 0.25	0.40 ± 0.13

[0220]

[0221] 주사된 Kdm4d mRNA의 농도는 1500 ng/ul이었다. N/A, 적용가능하지 않음. ET, 배아 전달

[0222] *Kdm4d* mRNA 주사는 이전의 관찰 (Matoba et al., Proc. Natl. Acad. Sci. U. S. A. *108*, 20621-20626, 2014)과 일치하게, 새끼 비율을 8.4%로 증가시켰다. *Xist* 이형접합성 마우스로부터 유래된 난구 세포가 공여자 세포로서 사용되고, *Kdm4d* mRNA 주사와 조합되는 경우, 새끼 비율은 18.7%로 증가하였다 (도 1c; 표 S1). 유사하게, 세르톨리 세포-유래된 SCNT 배아의 새끼 비율은 *Kdm4d* mRNA 주사에 의해 1.8%에서 9.1%로 개선되었고, *Xist* KO를 *Kdm4d* mRNA 주사와 조합함으로써 23.5%로 더 증가하였으며 (도 1c; 표 2), 이는 지금껏 보고된 가장 높은 마우스 클로닝률이다 (Ogura et al., Phil. Trans. R. Soc. B *368*, 20110329, 2013). 중요하게는, 이 상가적 효과는 또한 하이브리드 (129S1/Svj x CAST/EiJ) 유전적 배경에서 상이한 *Xist* 돌연변이체 주의 MEF 세포를 사용하여 관찰되었다 (도 1c; 표 2). 조합된 접근법을 사용하여 생성된 새끼는 성체기까지 성장하였으며, 정상적인 생식력을 나타내었다 (도 1d).

[0223] 상기 결과는 본 발명이 조합된 접근법을 사용함으로써 가장 높은 마우스 클로닝 효율을 제공함을 지시한다. 그러나, 심지어 23.5%의 가장 높은 클로닝 효율은 여전히 IVF 새끼 비율의 절반 미만이며, 이는 전달된 배아의 50% 초과이다. 사실, 착상된 배아 중 65% (57개 중 37개)는 조합된 세르톨리 세포 SCNT 군에서 착상후 발생 동안 정지하였다 (표 2). 상이한 배아 단계에서의 주의깊은 형태학적 조사는 배아 정지가 착상 직후 시작하며, 발생이 진행됨에 따라 점차적으로 증가함을 밝혀내었다 (도 2a). 더욱이, 형태학적 및 조직학적 분석은 미로층 내로의 PAS-양성 해면영양막 세포의 침습과 연관된 큰 태반 표현형 (도 1e)이 *Kdm4d* mRNA 주사 또는 조합된 접근법에 의해 구제되지 않았음을 밝혀내었다 (도 1e 및 1f; 표 2). 따라서, SCNT 배아 발생을 개선시키는 데 있어서 *Xist* KO 및 *Kdm4d* mRNA 주사의 조합된 양성 효과에도 불구하고, 다른 재프로그래밍 장벽은 이들 SCNT 배아의 발생 실패에 기여할 수 있다.

[0224] **실시예 3: 광범위한 DNA 메틸화 재프로그래밍은 조합적 재프로그래밍된 SCNT 배반포에서 달성된다**

[0225] 조합적 재프로그래밍된 SCNT 배아는 착상 직후 발생적 결함을 나타내기 때문에 (도 2a, 2b, 2c), 재프로그래밍 관련된 후생적 결함은 SCNT 배반포에서 이미 존재하였을 것이지만, 이들은 형태적으로 정상인 것으로 보인다고 상정되었다. 이러한 후생적 결함을 확인하기 위해, *Kdm4d* mRNA 주사와 조합된 *Xist* KO MEF 세포 (129S1/Svj x CAST/EiJ 배경)로부터 유래된 둘 다의 SCNT 배반포의 전체 게놈 비술파이트 시퀀싱 (WGBS) 데이터를 생성하고, 유전적으로 매칭된 IVF 배반포의 그것과 비교하였다 (도 3a).

[0226] 각각 IVF 및 SCNT 배반포로부터의 2060만개 및 2090만개 CpG 부위의 DNA 메틸화 정보 (표 3).

[0227] <표 3> WGBS 라이브러리의 요약, 도 2와 관련됨

샘플	총 시퀀싱 판독물	맵핑된 판독물	맵핑된 판독물의 백분율(%)	1x 커버된 CpG	5x 커버된 CpG	비술파이트 전환율 (%)
IVF 배반포	416,465,544	297,678,300	71.5	20,611,901	12,717,260	99.2
SCNT 배반포	412,706,010	298,966,969	72.4	20,676,648	12,506,862	99.2

[0228]

[0229] 비교를 위해, WGBS 데이터세트를 또한 공개적 데이터베이스로부터 MEF (Yu et al., Proc. Natl. Acad. Sci. U. S. A. *111*, 5890-5895, 2014), 정자 및 난모세포 (Wang et al., Cell *157*, 979-991, 2014)에 대해 얻었다. 먼저, 모든 커버된 CpG의 평균 DNA 메틸화 수준을 계산하였다. 정자 및 난모세포에서의 평균 메틸화 수준은 각각 82.2% 및 58.8%인 것으로 밝혀졌다. IVF 접합자의 출발 메틸화 수준으로서 기능하는 정자 및 난모세포의 평균 메틸화 수준은 70.5%이다 (도 3b). 그러나, 고도로 메틸화된 배우자는 가능성 있게는 착상전 발생 동안 일어난 능동 및 수동 탈메틸화 프로세스로 인해 배반포 단계에 의해 낮은 메틸화 수준으로 전반적으로 재프로그래밍되었다 (19.1%).

[0230] 다음으로, 통상적으로 커버된 CpG의 MEF 세포 및 SCNT 배반포의 DNA 메틸화 수준을 계산하였으며, 공여자 MEF 세포는 고도로 메틸화되지만 (78.0%), SCNT 배반포는 IVF 배반포의 그것 (19.1%)과 유사한 메틸화 수준 (15.6%)으로 낮게 메틸화된 것으로 밝혀졌으며, 이는 DNA 메틸화 상태의 성공적인 전반적 재프로그래밍을 지시한다 (도 3b). 사실, DNA 메틸화의 쌍별 비교는 IVF 및 SCNT 배반포 둘 다가 배우자 또는 MEF 세포에서의 그것에 비해 극도로 낮은 DNA 메틸화를 가짐을 밝혀내었다 (도 3c). 전반적 메틸화 수준 뿐만 아니라 메틸화된 CpG의 분포는 유사하다 (도 5a). 유사한 전반적 DNA 메틸화 패턴과 일치하게, RNA-seq는 IVF 및 SCNT 배반포 사이에 매우 유사한 전사체를 밝혀내었다 (R = 0.988) (도 3d, 4a, 4b, 및 5b). 사실, 검출된 8,921개의 유전자 중에서 (적어도 1개의 샘플에서 FPKM>3), 단지 92개의 유전자만이 SCNT 배반포에서 차등적으로 발현되었다 (FC>3) (도 5d, 표 4).

[0231] <표 4> SCNT 배반포에서 차등적으로 조절된 유전자

기호	상태	Chr
Msc	하향	Chr1
Crygd	하향	Chr1
Ren1	하향	Chr1
Mael	하향	Chr1
Car4	하향	Chr11
Cbr2	하향	Chr11
Cd7	하향	Chr11
Scarna3b	하향	Chr12
Acot1	하향	Chr12
Cox8c	하향	Chr12
Pi1	하향	Chr13
Bhmt2	하향	Chr13
Sult4a1	하향	Chr15
Tnp2	하향	Chr16
Fetub	하향	Chr16
Kng1	하향	Chr16
Impact	하향	Chr18
Pde6a	하향	Chr18

[0232]

1700123I01Rik	하향	Chr19
Fabp9	하향	Chr3
Fabp4	하향	Chr3
Slc25a31	하향	Chr3
S100a3	하향	Chr3
Tdpoz4	하향	Chr3
Gstm6	하향	Chr3
Scarna2	하향	Chr3
Fam154a	하향	Chr4
Cox7b2	하향	Chr5
Asz1	하향	Chr6
Npy	하향	Chr6
Tuba3b	하향	Chr6
Tex101	하향	Chr7
Klk7	하향	Chr7
Rbmxl2	하향	Chr7
Adrb3	하향	Chr8
Tat	하향	Chr8
Tex12	하향	Chr9
2410076I21Rik	하향	Chr9
Rbp2	하향	Chr9
AU022751	하향	ChrX
1700013H16Rik	하향	ChrX
Xlr	하향	ChrX
Gm773	하향	ChrX
4930550L24Rik	하향	ChrX
Xlr3a	하향	ChrX
Xlr5a	하향	ChrX
Xlr5b	하향	ChrX
Xlr3b	하향	ChrX
Xlr4b	하향	ChrX
Xlr4c	하향	ChrX
Xlr3c	하향	ChrX
Magea8	하향	ChrX
Magea5	하향	ChrX
Gm6432	하향	Chr9
LOC100861571	하향	Chr12
Scarna3a	상향	Chr1
Derl3	상향	Chr10
Ddit3	상향	Chr10

[0233]

Krt19	상향	Chr11
Jdp2	상향	Chr12
Serpina3m	상향	Chr12
BC051665	상향	Chr13
Slc39a2	상향	Chr14
Slc7a8	상향	Chr14
Entpd1	상향	Chr19
Aifl1	상향	Chr2
Elf5	상향	Chr2
Chac1	상향	Chr2
Trib3	상향	Chr2
Defb25	상향	Chr2
Myl9	상향	Chr2
Ptk6	상향	Chr2
Glipr2	상향	Chr4
Tspan1	상향	Chr4
Plac8	상향	Chr5
Apoc2	상향	Chr7
LOC100302567	상향	Chr7
Parva	상향	Chr7
Nupr1	상향	Chr7
H19	상향	Chr7
Gm514	상향	Chr9
Ostb	상향	Chr9
Praf2	상향	ChrX
Rhox6	상향	ChrX
Plac1	상향	ChrX
0610009L18Rik	상향	Chr11
1700019E08Rik	상향	Chr2
4930461G14Rik	상향	Chr9
4930591A17Rik	상향	Chr2
AF357426	상향	Chr12
Gm5480	상향	Chr16
LOC100861570	상향	Chr12

[0234]

[0235] 이들 결과는 DNA 메틸화 및 전사체가 배반포 단계에서 SCNT에 의해 크게 재프로그래밍됨을 지시한다.

[0236] **실시예 4: SCNT 배반포에서의 차등적으로 메틸화된 영역 (DMR)의 확인 및 특징규명**

[0237] DNA 메틸롬의 성공적인 전반적 재프로그래밍에도 불구하고, SCNT 배반포의 평균 메틸화 수준 (15.6%)은 IVF 배반포의 그것 (19.1%)보다 약간 그러나 유의하게 더 낮았다 (p 값 < 2.2e-16) (도 5b). 게놈-범위 스캐닝 분석 (최소 창 크기로서 10 CpG)을 수행하여 SCNT 및 IVF 배반포 사이의 차등적으로 메틸화된 영역 (DMR)을 확인하였으며, 이는 10% 초과의 절대 메틸화 차이를 갖는 56,240개의 DMR을 밝혀내었다 (도 5a). 이들 DMR의 대다수 (48,315개: 85.9%)는 IVF의 그것에 비해 SCNT에서 더 낮은 DNA 메틸화를 나타내었으며, 과소DMR로 용어화되었다. IVF의 그것에 비해 SCNT에서 더 높은 DNA 메틸화를 갖는 7,925개의 영역을 확인하고, 과다DMR로 용어화하였다. 흥미롭게도, 과다DMR의 평균 길이 (741 bp)는 과소DMR의 그것 (5,743 bp)보다 훨씬 더 짧다 (도 5b). 사실, 대표적인 과다DMR은 프로모터/인핸서 영역에 샤프한 피크로서 존재하는 반면, 대표적인 과소DMR은 전체 유전자-코딩 영역을 커버한다 (도 6a, 6b). 일관되게, 과다DMR 및 과소DMR은 유전자간 영역에 풍부화된 과다DMR, 반면 유전자 체에 풍부화된 과소DMR을 갖는 별개의 게놈 분포를 나타낸다 (도 5c).

[0238] 어떻게 이들 DMR이 형성되는지 및 이들이 SCNT 배아의 착상후 발생 실패에 기여할 수 있는지 여부를 이해하기

위해, 분석은 과소DMR에 초점을 맞추었다. 정자에서, 과소DMR에서의 메틸화 수준은 플랭킹 영역보다 유의하게 더 높았다 (약 90% 대 약 80%) (도 5d). 난모세포에서, 과소DMR 및 플랭킹 영역 사이의 메틸화 차이는 훨씬 더 크다 (약 75% 대 약 60%) (도 3d). 대조적으로, 과소DMR 및 플랭킹 영역 사이의 메틸화 차이는 MEF에서 훨씬 더 작다 (도 5d). 따라서, SCNT 배반포의 과소DMR은, SCNT 및 IVF 배아 둘 다가 동일한 수의 복제-의존성 희석을 거치는 경우, IVF 배반포에서의 보다 높은 수준에서인 것으로 잔류하는 배우자에서의 그들의 상대적으로 더 높은 DNA 메틸화 수준과 잘 상관된다. 이 개념과 일치하게, 게놈 브라우저 보기에서의 대표적인 과소DMR의 가시적 검사는 난모세포에서의 메틸화 피크가 IVF 배반포에서의 그것들과 명백하게 중첩됨을 밝혀내었다 (도 6b). 대립유전자 DNA 메틸화 분석은 또한 IVF 배반포에서의 메틸화 패턴이 모계 대립유전자로 강하게 편향되기 때문에 이 개념을 뒷받침한다 (도 5e). 사실, 착상전 배아의 공개된 WGBS 데이터세트의 분석 (Wang et al., Cell *157*, 979-991, 2014)은 모계 대립유전자가 4 세포-단계까지 과소DMR에서 그의 높은 DNA 메틸화 수준을 유지하는 반면, 부계 대립유전자는 이들 영역에서 그의 메틸화를 빠르게 소실함을 밝혀내었다 (도 5f).

[0239] 다음으로, 과다DMR을 분석하였다. 과다DMR 및 플랭킹 영역에서의 메틸화 수준은 난모세포에서 유사한 반면 (약 50%), 이는 정자에서 플랭킹 영역보다 유의하게 더 낮다 (약 55% 대 약 80%) (도 5g). 그들의 메틸화 수준의 차이에도 불구하고, 과다DMR 및 플랭킹 영역 둘 다는 IVF 배반포에서 매우 낮은 수준 (약 20%)으로 탈메틸화되었다 (도 5g). 대조적으로, 과다DMR은 MEF에서 플랭킹 영역의 그것보다 훨씬 더 높은 메틸화 수준으로 심하게 메틸화되었다 (약 80%) (도 5g). 과다DMR이 MEF에서 심하게 메틸화되지만, 배우자에서는 그렇지 않다는 사실은 이들 영역에서의 낮은 메틸화가 배선의 고유한 특색일 수 있음을 시사한다. 사실, 상이한 세포 유형의 공개적인 DNA 메틸롬 데이터세트의 분석은 과다DMR이 사실 분석된 모든 체세포 유형에서 심하게 메틸화되지만, 단지 정모세포, 정세포 및 난모세포에서만 덜 메틸화됨을 밝혀내었다 (도 5h). 일관되게, 과다DMR과 연관된 유전자의 GO 분석은 배선 관련된 기능, 예컨대 정자형성 및 배우자형성의 유의한 풍부화를 밝혀내었다 (도 6c). 과다DMR은 Tet1에 의한 원시 배 세포 (PGC) 발생 동안 탈메틸화되는 것으로 보이는데 (Yamaguchi et al., Nature *504*, 460-464, 2013), 이는 히드록시메틸시토신 (5hmC)이 PGC에서 과다DMR에서 유의하게 풍부화되기 때문이다 (도 6d). 이는 과다DMR이 대부분 배선 발생과 관련되지만, 배아 발생과는 그렇지 않음을 시사한다.

[0240] **실시예 5: SCNT 배반포에서의 H3K27me3-의존성 각인의 소실**

[0241] 결함성 태반 발달은 SCNT 배아에서 중심적 특색이다. 이전의 연구는 게놈 각인이 태반 발달에서 중요한 역할을 함을 확립하였다. 따라서, SCNT 배아의 착상후 결함에 잠재적으로 기여할 수 있는 확인된 DRM 외에도, 게놈 각인에 대한 조합적 접근법의 영향을 평가하는 것은 중요하다. 이 목적으로, 23개의 공지된 각인 제어 영역 (ICR)의 DNA 메틸화 수준을 분석하였는데, 이는 각인된 유전자의 대립유전자 발현이 대립유전자 ICR 메틸화에 의해 크게 제어되기 때문이다. 비교적 DNA 메틸화 분석은 DNA 메틸화 수준이 대부분의 ICR에서 약간 감소되지만, 23개의 ICR 중 21개는 IVF 배반포 수준의 그것의 적어도 반을 유지하였음을 밝혀내었으며 (도 8a), 이는 DNA 메틸화-매개된 게놈 각인이 크게 유지되었음을 지시한다. 사실, 충분한 대립유전자 특이적 메틸화 정보 (IVF 및 SCNT 배반포 둘 다의 둘 다의 대립유전자에서 5 초과의 검출된 CpG)를 갖는 모든 20개의 ICR은 IVF 및 SCNT 배반포 사이의 일관된 대립유전자 특이적 DNA 메틸화를 나타내었다 (도 8b).

[0242] SCNT 결함에서의 DNA 메틸화-매개된 게놈 각인의 잠재적 역할을 추가로 평가하기 위해, RNAseq 데이터세트를 126개의 공지된 각인된 유전자에 초점을 맞추어 분석하였다. IVF 배반포에서 신뢰성 있게 검출가능한 (FPKM > 1) 45개의 각인된 유전자 중에서, 단지 6개만이 IVF 배반포에서의 그것에 비해 SCNT 배반포에서 유의하게 상향조절되었다 (FC > 1.5) (도 8c). 대립유전자 발현 분석 (FPKM>1, 어느 샘플에서든 평균 SNP 판독물 > 10)은 충분한 수의 SNP 판독물을 갖는 36개의 각인된 유전자 중에서, IVF 배반포에서 6개는 모계-대립유전자 편향된 (Mat/Pat > 2.0) 발현을 나타내었고, 13개는 부계-대립유전자 편향된 (Pat/Mat >2.0) 발현을 나타내었음을 밝혀내었다 (도 8d 및 8e, 보다 밝은 막대). 모든 6개의 모계적으로 발현된 유전자 (MEG)는 SCNT 배반포에서 그들의 모계 편향된 발현을 유지하였다 (도 8d). 13개의 부계적으로 발현된 유전자 (PEG) 중에서, 7개는 대립유전자 편향을 소실하였고, SCNT 배반포에서 이중대립유전자적으로 발현되게 되었다 (도 8e에서 화살표). 흥미롭게도, SCNT 배반포에서 각인된 발현을 소실한 7개의 PEG는 *Slc38a4*, *Sfmbt2*, *Phf17* 및 *Gab1*을 포함하였으며 (도 8e에서 보다 어두운 막대), 그의 각인된 발현은 DNA 메틸화에 독립적이지만, 모계적으로 침착된 H3K27me3에 의존하는 것으로 공지되어 있다 (Inoue et al., Nature *547*, 419-424, 2017).

[0243] 분석은 최근 확인되었지만 상기 분석에 포함되지 않은 H3K27me3-의존성 각인된 유전자에 초점을 맞추었다. 상실배 배아에서 H3K27me3-의존성 각인된 발현을 나타내는 76개의 유전자 중에서 (Inoue et al., Nature *547*, 419-424, 2017), 26개는 IVF 배반포에서 신뢰성 있게 검출가능한 수준 (FPKM >1)으로 발현된다. 흥미롭게도, 이들 중 대다수 (15개/26개)는 SCNT 배반포에서 유의하게 상향조절된다 (FC>1.5) (도 7a). 대립유전자 발현 분

석은 충분한 SNP 판독물을 갖는 23개의 유전자 중에서 (FPKM>1, 어느 샘플에서든 평균 SNP 판독물 > 10), 17개가 분석된 유전적 배경 (129S1/Svj x CAST/EiJ) 하에서 IVF 배반포에서 부계적으로 편향된 발현 (Pat/Mat > 2.0)을 나타내었음을 밝혀내었다. 현저하게, 모든 17개의 PEG는 부계 대립유전자-편향된 발현을 소실하였고, 이중대립유전자 발현을 나타내었다 (도 7b). 이들 결과는 H3K27me3-의존성 각인된 유전자가 SCNT 배반포에서 그들의 각인을 완전히 소실함을 명백하게 입증하였다.

[0244] 왜 이들 H3K27me3-의존성 각인된 유전자는 SCNT 배반포에서 각인을 소실하는가? 이들 유전자의 각인 상태는 난자형성 동안 침착된 모계 대립유전자-특이적 H3K27me3 도메인에 의해 조절되기 때문에, 공여자 MEF에서의 H3K27me3 패턴은 난모세포에서의 그것과 상이할 수 있음이 가능하다. 충분히 성장된 난모세포 및 MEF 세포의 이용가능한 H3K27me3 ChIP-seq 데이터세트의 분석은 난모세포에서 이들 각인된 유전자에서의 H3K27me3 도메인이 MEF 세포에 완전히 부재하였음을 밝혀내었다 (도 7c 및 8f). 분석은 다른 체세포 유형을 포함하도록 확장되었다. 이 분석은 이들 각인된 유전자에서의 H3K27me3 도메인이 일반적으로 분석된 체세포 유형으로부터 부재하였으며, 따라서 난모세포 게놈에 고유하였음을 밝혀내었다 (도 8d). 이들 결과는 공여자 체세포에서의 이들 각인된 유전자의 모계 대립유전자에서의 H3K27me3 메틸화의 결여가 SCNT 후의 각인의 소실 (LOI)의 원인일 가능성이 있음을 지시한다.

[0245] SCNT에 의한 20개 초과의 포유동물 종의 성공적인 클로닝에도 불구하고, 클로닝 효율은 한결같이 낮으며, 태반 과다성장을 비롯한 발생적 이상이 본질적으로 모든 클로닝된 포유동물 종에서 관찰된다. 후생적 이상은 클로닝된 동물의 발생 실패의 원인이 되는 것으로 추측되었다. 이 연구에서, 2가지 접근법을 조합하여 마우스 SCNT 배아의 발생을 방해하는 2가지 이전에 확인된 재프로그래밍 장벽을 극복하였다. *Xist* KO 공여자 체세포 및 *Kdm4d* mRNA 주사의 조합적 사용은 사실 전체 클로닝 효율 (만기 비율)을 20배 증가시켜 체세포 공여자 세포를 사용한 마우스 생식적 클로닝에서 지금껏 보고된 가장 높은 새끼 비율 (예를 들어, 세르톨리 세포를 사용하여 23%)을 생성하였다. 이 효율은 그것이 유사한 핵 주사가 포함된 세포질내 정자- 또는 정세포 주위-주사 (ICSI/ROSI)의 그것과 가깝기 때문에 주목할 만하다 (Ogonuki et al., PLoS One *5,* 2010; Ogura et al., International Review of Cytology, pp. 189-2292005). 이 달성은 공여자 세포에서의 H3K9me3 및 *Xist*의 비정상적 활성화가 성공적인 클로닝을 위한 2가지 주요한 장벽을 나타내며, 따라서 SCNT-매개된 재프로그래밍의 분자적 메커니즘을 이해하기 위한 토대를 확립함을 명백하게 입증한다.

[0246] 주목할 만한 개선에도 불구하고, 조합적 접근법을 사용하여 생성된 SCNT 배아 중 많은 것은 착상 후에 발생하는데 실패하였다. 더욱이, 태반 과다성장은 공여자 세포 유형과 무관하게 여전히 관찰되었으며, 이는 고효율 동물 클로닝을 위해 추가의 장벽이 존재함을 지시한다. 이들 추가의 재프로그래밍 장벽을 확인하기 위해, 발생 실패가 시작되는 시간을 확인하고, 제1 WGBS 데이터세트를 발생적 표현형이 나타나기 직전에 SCNT 배반포로부터 생성하였다. 비교적 DNA 메틸롬 분석은 SCNT에 의한 성공적인 전반적 DNA 메틸롬 재프로그래밍을 밝혀내었으며, 이는 난모세포질 및 절단 배아에서의 DNA 탈메틸화 기구가 IVF 배아에서의 그것에 비해 SCNT 배아에서 유사하게 기능적임을 지시한다. 그럼에도 불구하고, DNA 메틸롬의 상세한 비교적 분석은 IVF 및 SCNT 배반포 사이의 게놈에 걸쳐 많은 DMR을 밝혀내었다. 흥미롭게도, 과다DMR은 배선에서 탈메틸화된 게놈 영역에 풍부하며, 이는 배 세포-특이적 유전자가 배 세포 발생 동안, 특히 원시 배 세포 (PGC) 단계에서 Tet1에 의해 탈메틸화된다는 사실과 일치한다. 그러나, SCNT는 이 탈메틸화 프로세스를 우회한다. 과다DMR-연관된 유전자의 목록은 1세포 SCNT 배아에서 빠르게 탈메틸화되는 것으로 보고된 소수의 배선 유전자를 포함하지 않으며, 이는 대다수는 아니지만 일부 배선 유전자가 SCNT 후에 탈메틸화됨을 지시한다. 한편, 과소DMR은 주로 난모세포에서 메틸화된 영역과 중첩한다. 이들 영역에서의 모계 DNA 메틸화는 IVF 배아에서 특히 8-세포 단계 전에 탈메틸화 프로세스를 벗어나는 것으로 보인다. 8-세포 단계 전의 DNA 메틸화의 모계 대립유전자-특이적 유지에 대한 기저의 메커니즘은 추가의 연구를 위해 흥미롭다. 집합적으로, SCNT 특이적 DMR은, 높거나 낮거나, 정상적인 수정을 통해 배반포에 유전되는 배우자형성의 고유한 특색으로 인해 형성되는 것으로 보인다. 수정을 통해 난모세포로부터 배아로 전해진 모계 DNA 메틸화는 초기 단계 영양막 발달에서 중요한 역할을 하는 것으로 나타났다 (Branco et al., Dev. Cell *36*, 152-163, 2016). 따라서, SCNT 배반포에서의 난모세포-유사 DNA 메틸화 패턴의 소실은 SCNT 배아의 발생적 표현형에 기여할 수 있다.

[0247] 본원에서 보고된 바와 같이, DNA 메틸화-각인된 유전자의 DNA 메틸화 및 전사체 분석은 대부분의 ICR이 그들의 정상적인 각인 상태를 크게 유지하며, 대부분의 정규 각인된 유전자가 사실 SCNT 배반포에서 대립유전자 발현 패턴을 유지함을 밝혀내었다. 대조적으로, 최근 발견된 H3K27me3-매개된 비-정규 각인된 유전자 (Inoue et al., 2017)는 전체적으로 조절이상되며, SCNT 배반포에서 이중대립유전자 발현을 나타낸다. SCNT 배반포에서의 조절이상된 비-정규 각인된 유전자의 목록은 *Slc38a4*, *Sfmbt2* 및 *Gab1*을 포함하였으며, 이는 E13.5 SCNT 배아의

태반에서의 이들 3가지 유전자의 각인의 소실 (LOI)의 이전의 보고와 일치한다 (Okae et al., Hum. Mol. Genet. *23*, 992-1001, 2014). 3가지 유전자 모두가 태반 성장에 중요한 역할을 하는 것으로 나타났음을 고려하여, 이들 유전자의 LOI는 SCNT 배아의 태반 과다성장 표현형에 기여할 가능성이 있다. 더욱이, *Runx1*, *Otx2* 및 *Etv6*은 마우스 초기 배아 발생에서 중요한 역할을 하는 것으로 나타났으며, 따라서 배반포 단계에서 이들 유전자의 LOI는 착상후 SCNT 배아의 배아 치사 표현형에 기여할 수 있다. SCNT에서의 비-정규 각인된 유전자의 LOI의 원인은 공여자 체세포에서 이들 좌위에서의 H3K27me3의 부재에 기인할 가능성이 크다. H3K27me3-각인된 유전자의 조절 메커니즘에 대한 추가의 상세한 분석은 SCNT 배아 발생을 개선시키기 위한 단서를 제공할 것이다.

[0248] 요약하면, *Kdm4d* mRNA 주사를 *Xist* KO 공여자 세포의 사용과 조합함으로써 가장 효율적인 마우스 클로닝 방법을 확립하는 것 외에도, H3K27me3 각인은 효율적인 동물 클로닝을 방지하는 잠재적인 장벽으로서 밝혀졌다. 이론에 의해 구애되기를 의도하지는 않지만, 마우스 SCNT 배반포에서의 H3K27me3-의존성 각인된 유전자에서의 LOI의 명백한 연관 및 배아 발생에서의 그들의 중요한 기능에 기초하여, H3K27me3-각인된 유전자에서의 LOI는 SCNT 배아의 착상후 표현형을 해명할 가능성이 크지만, 이 연구에서 확인된 비정상적 DNA 메틸화의 잠재적 기여의 가능성은 배제될 수 없다. SCNT 배아에서 결함성 착상후 발생 및 비정상적 태반 표현형이 포유동물 종에서 통상적으로 관찰됨을 고려하여, 다른 종의 클로닝된 배아에서의 H3K27me3-의존성 각인 상태의 조사는 추가의 조사를 보증할 수 있다.

[0249] 기재된 결과는 하기 방법 및 물질을 사용하여 얻어졌다.

[0250] **PN5 단계 접합자로부터의 모계 및 부계 전핵의 단리**

[0251] 모든 동물 연구는 하바드 메디칼 스쿨 (Harvard Medical School)에서의 기관 동물 관리 및 사용 위원회 (Institutional Animal Care and Use Committee)의 지침에 따라 수행되었다. MII-단계 난모세포를 7.5 I.U.의 PMSG (밀리포어 (Millipore)) 및 hCG (밀리포어)를 주사함으로써 과배란된 8주령 B6D2F1/J (BDF1) 암컷으로부터 수집하였다. 시험관내 수정 (IVF)을 위해, MII 난모세포를 10 mg/ml 소 혈청 알부민 (BSA; 시그마-알드리치 (Sigma-Aldrich))으로 보충된 HTF 배지에서 성체 BDF1 수컷 마우스의 꼬리 부고환으로부터 얻어진 활성화된 정자로 수정시켰다. 정자 수정능은 HTF 배지에서 1시간 인큐베이션에 의해 획득되었다. 접합자를 37.8℃에서 5 % CO_2/95% 공기를 갖는 습윤 분위기에서 배양하였다. 수정후 10시간 (hpf)에서, 접합자를 10 μg/ml 시토칼라신 B (시그마-알드리치)를 함유하는 M2 배지 내로 옮겼다. 투명대를 피에조 (Piezo) 충격-유도된 미세조작장치 (프라임 테크 리미티드 (Prime Tech Ltd.), 일본 이바라키켄)에 의해 커팅하고, 전핵을 접합자로부터 단리하였다. 12 hpf (PN5-단계)에서, 단리된 전핵을 0.2% BSA/PBS로 세척하고, 에펜도르프 로바인드 (Eppendorf LoBind) 1.5 ml 튜브 내로 옮기고, DN아제 I 처리까지 얼음 상에 정치하였다. 각각의 실험에 대해, 150 내지 200개의 전핵을 수집하고, liDN아제-seq를 위해 준비하였다. 모 전핵을 (1) 제2 극체로부터의 거리 및 (2) 전핵의 크기에 의해 구별하였다.

[0252] **웅성발생성 (AG) 및 자성발생성 (GG) 배아의 제조**

[0253] MII 난모세포를 8주령 과배란된 BDF1 암컷으로부터 수집하고, BDF1 정자와 수정시켰다. 7 hpf에서, 접합자를 5 μg/ml 시토칼라신 B를 함유하는 M2 배지 내로 옮기고, 모 전핵을 피에조 충격-유도된 미세조작장치를 사용함으로써 교환하였다. 센다이 바이러스 (HVJ, 코스모-바이오 (Cosmo-bio))를 사용하여 이전에 기재된 바와 같이 핵질체를 세포질과 융합시켰다. 재구축 후, 배아를 KSOM에서 배양하였다.

[0254] RNA-seq 또는/및 liDN아제-seq를 위한 배아를 수집할 경우, 투명대 (ZP)를 산 티로이드 용액 (시그마-알드리치)에의 간단한 노출에 의해 제거한 후, 배아를 M2 배지, 및 그 후 0.2% BSA/PBS로 세척하였다. liDN아제-seq를 위해, 10개의 상실배 배아를 에펜도르프 로바인드 1.5 ml 튜브 내로 옮기고, DN아제 I 처리까지 얼음 상에 정치하였다. RNA-seq를 위해, 7 내지 10개의 배아를 얇은 벽 RN아제-무함유 PCR 튜브 (앰비온) 내로 옮겼다. 2-세포 및 상실배 배아를 각각 30 및 78 hpf에서 수집하였다. α-아마니틴 처리된 2-세포 배아를 제조하기 위해, 5 hpf 접합자를 25 μg/ml α-아마니틴 (시그마-알드리치)을 함유하는 KSOM 내로 옮기고, 수집까지 (30 hpf) α-아마니틴의 존재 하에서 배양하였다. ICM 및 TE를 단리하였다. 간략하게, 120 hpi에서의 AG 및 GG 배아를 산 티로이드 용액으로 처리하여 ZP를 제거하였다. M2 배지에서 세척한 후, 배아를 토끼 항-마우스 림프구 혈청 (씨더레인 (Cedarlane), 1:8 희석)을 함유하는 KSOM에서 37℃에서 45분 동안 인큐베이션하였다. M2 배지에서 세척한 후, 이들을 기니아 피그 보체 (엠피 바이오메디칼스 (MP Biomedicals), 1:3.3 희석)를 함유하는 KSOM 내로 옮겼다. 37℃에서 30분 동안 인큐베이션 후, 용해된 TE 세포를 유리 모세관으로 피펫팅함으로써 제거하였다. 잔류의 ICM 클럼프를 0.25% 트립신/EDTA (써모 피셔 (Thermo Fisher), 25200)에서 37℃에서 10분 동안 인큐베이션한 후, 용해된 TE 세포의 오염을 회피하기 위해 단일 세포로 분리하였다. 100 내지 200개의 세

포를 RNA-seq를 위해 수집하였다.

[0255] **완전히 성장된 난모세포로부터의 GV 핵의 단리**

[0256] 완전히 성장된 GV-단계 난모세포를 5 I.U. PMSG로의 주사 후 44 내지 48시간에 3주령 BDF1 마우스로부터 얻었다. 난소를 M2 배지 내로 옮겼다. 난소 여포를 30-게이지 바늘로 천공하고, 난구 세포를 좁은 구멍 유리 피펫을 사용하여 난구-난모세포 복합체로부터 부드럽게 제거하였다. 그 후, 난모세포를 5% 소 태아 혈청 (FBS) (시그마-알드리치, F0926), 10 ng/ml 표피 성장 인자 (시그마-알드리치, E4127), 및 0.2 mM 3-이소부틸-1-메틸크산틴 (IBMX; 시그마-알드리치)으로 보충된 α-MEM (라이프 테크놀로지스 (Life technologies), 12571-063) 내로 옮겼다. 수집 후 1시간에, 핵소체 주위 (SN)-유형 염색질을 갖는 가시적 난황주위 공간을 나타내는 GV 난모세포를 도태시켰다. 그 후, 이들을 10 μg/ml 시토칼라신 B, 0.1 μg/ml 콜세미드 (시그마-알드리치), 및 0.2 mM IBMX를 함유하는 M2 배지에서 15분 동안 인큐베이션하였다. 그 후, GV 핵을 피에조-유도된 미세조작 장치를 사용함으로써 단리하였다. 0.2% BSA/PBS로 세척한 후, GV 핵을 에펜도르프 로바인드 1.5 ml 튜브 내로 옮겼다. 각각의 실험에 대해, 115 내지 150개의 GV 핵을 liDN아제-seq를 위해 수집하였다.

[0257] **E6.5 배아의 절개 및 GFP-양성 E9.5 태반 세포의 FACS 분류**

[0258] C57BL6(B6)/PWK 하이브리드 배아를 얻기 위해, 자연 교배 방식을 사용하였다. PWK/B6 하이브리드 배아를 얻기 위해, PWK 난모세포와 B6 정자의 시험관내 수정을 사용하고, 2-세포 배아를 ICR 주 대리모 내로 전달하였다. E6.5 배아의 EPI, EXE, 및 VE로의 절개를 수행하였다. E9.5 태반 세포를 수집하기 위해, 잭슨 래버러토리 (Jackson laboratory)로부터 $B6^{GFP}$ 마우스를 구입하였다 [C57BL/6-Tg(CAG-EGFP)1310sb/LeySopJ, 스톡 번호 006567]. MII 난모세포 및 정자를 과배란된 8주령 $B6^{GFP}$ 또는 PWK 마우스로부터 수집하였다. 시험관내 수정 후, 2-세포 배아를 ICR 주 대리모 내로 전달하였다. E9.5에서, 태반을 수확하고, 약 0.5 mm 조각으로 커팅하고, 50 ml 튜브 내로 옮기고, 2 ml의 0.25% 트립신-EDTA (써모 피셔 사이언티픽 (Thermo Fisher Scientific), 25200)로 30℃에서 15분 동안 200 rpm의 진탕기에서 처리하여 태반 세포를 분리하였다. 트립신 처리를 10% FBS를 함유하는 2 ml DMEM의 첨가에 의해 중지시켰다. 피펫팅 후, 튜브를 원심분리하고, 펠릿화된 세포를 0.2%BSA/PBS로 3회 세척하였다. DAPI를 최종 세포 현탁액 중 1 μM의 최종 농도로 첨가하였다. GFP-양성 세포를 BD FACS아리아 기계 (비디 바이오사이언시스 (BD Biosciences))를 사용하여 분류하고, DAPI 양성 세포는 죽은 세포로서 제외하였다. 대략 10,000 내지 20,000개의 GFP-양성 세포를 각각의 태반으로부터 수집하였으며, 이는 총 태반 세포의 40 내지 60%에 상응하였다.

[0259] **플라스미드 구축 및 mRNA 제조**

[0260] *Kdm6b*^{WT} 구축물을 생성하기 위해, 촉매 도메인을 함유하는 카르복실-말단 부분 (아미노산 1025-말단)을 코딩하는 cDNA를 증폭시켰다. PCR 앰플리콘을 pcDNA3.1-플래그 (Flag)-폴리(A)83 플라스미드의 플래그 태그 및 폴리(A) 사이에서 클로닝하였다. H1390A *Kdm6b*^{MUT} 구축물을 프라임스타 (PrimeSTAR) 돌연변이유발 (타카라 (TAKARA))을 사용함으로써 생성하였다. 돌연변이유발에 사용된 프라이머는 5'-CCAGGCgctCAAGAGAATAACAATTTCTGCTCAGTCAACATCAAC-3' 및 5'- CTCTTGagcGCCTGGCGTTCGGCTGCCAGGGACCTTCATG-3'이다. 모든 구축물을 DNA 시퀀싱에 의해 확인하였다. 야생형 및 H189A 돌연변이체 *Kdm4d*에 대한 플라스미드는 이전에 기재되었다.

[0261] 제한 효소에 의한 선형화 후, 구축물을 페놀-클로로포름 추출로 정제하였다. mRNA를 mMESSAGE mMACHINE T7 울트라 키트 (Ultra Kit) (라이프 테크놀로지스)를 제조업자의 지시서에 따라 사용하여 시험관내 전사에 의해 합성하였다. 합성된 mRNA를 염화리튬 침전에 의해 정제하고, 뉴클레아제-무함유수로 희석하였다. mRNA 분취물을 사용시까지 -80℃에서 저장하였다.

[0262] **mRNA 주사**

[0263] MII 난모세포를 과배란된 8주령 BDF1 암컷으로부터 수집하고, BDF1 정자와 수정시켰다. 2.5 hpf에서, 수정된 난모세포를 M2 배지 내로 옮기고, mRNA를 피에조 충격-유도된 미세조작장치를 사용하여 주사하였다. mRNA 주사를 4 hpf까지 완결하였다. *Kdm6b*^{WT} 및 *Kdm6b*^{MUT}의 mRNA 농도는 1.8 μg/μl이고, *Kdm4d*^{WT} 및 *Kdm4d*^{MUT}의 그것은 1.5 μg/μl이었다. *Kdm6b*-주사된 PG 배아를 제조하는 경우, MII 난모세포를 5 μg/ml 시토칼라신 B를 함유하는 Ca^{2+}-무함유 KSOM에서 3 mM $SrCl_2$로 처리함으로써 화학적으로 활성화시켰다. 활성화후 4시간 (hpa)에서, 배

아를 KSOM으로 세척하였다. 5 hpa에서, 이들을 mRNA로 주사하였다.

[0264] **전조직 표본 면역염색**

[0265] 접합자를 0.2% 트리톤 (Triton)을 함유하는 PBS 중 3.7% 파라포름알데히드 (PFA)에서 20분 동안 고정시켰다. 10 mg/ml BSA를 함유하는 PBS (PBS/BSA)로 4회 세척 후, 접합자를 1차 항체로 4℃에서 밤새 처리하였다. 이 연구에 사용된 1차 항체는 마우스-항-H3K27me3 (1/500, 액티브 모티프 (Active Motif), 61017), 토끼 항-H3K9me3 (1/500, 밀리포어, 07-442), 및 토끼 항-FLAG (1/2000, 시그마-알드리치, F7524)이었다. PBS/BSA로 3회 세척 후, 샘플을 플루오레세인 이소티오시아네이트-접합된 항-마우스 IgG (잭슨 이뮤노-리서치 (Jackson Immuno-Research)) 또는 알렉사 플루오르 568 당나귀 항-토끼 IgG (라이프 테크놀로지스)의 1:250 희석물과 함께 1시간 동안 인큐베이션하였다. 그 후, 접합자를 4',6-디아미디노-2-페닐인돌 (DAPI)을 갖는 벡타쉴드 (Vectashield) 항-표백 용액 (벡터 래버러토리스 (Vector Laboratories), 미국 캘리포니아주 벌링게임)에서 유리 슬라이드 상에 마운팅하였다. 형광을 회전 디스크 (CSU-10, 요코가와 (Yokogawa)) 및 EM-CCD 카메라 (ImagEM, 하마마츠 (Hamamatsu)) 또는 자이스 (Zeiss) LSM800을 갖는 레이저-스캐닝 공초점 현미경 하에서 검출하였다.

[0266] 모든 화상을 액시오비젼 (Axiovision) 소프트웨어 (칼 자이스 (Carl Zeiss))를 사용하여 획득하고, 분석하였다. 형광 신호 강도를 액시오비젼 소프트웨어로 정량화하였다. 간략하게, 모계 전핵 내의 신호 강도를 측정하고, 세포질성 신호를 배경으로서 차감하였다. 그 후, 주사 없는 대조군 접합자의 평균된 신호 강도를 1.0으로 설정하였다.

[0267] **저-투입 DN아제-seq**

[0268] 저-투입 DN아제-seq 라이브러리를 약간의 변형으로 이전에 기재된 바와 같이 제조하였다. 1.5 ml 튜브에서 수집된 배아 또는 핵을 36 μl 용해 완충제 (10 mM 트리스 (Tris)-HCl, pH 7.5, 10 mM NaCl, 3 mM MgCl2, 0.1% 트리톤 X-100)에 재현탁시키고, 얼음 상에서 5분 동안 인큐베이션하였다. DN아제 I (10 U/μl, 로슈 (Roche))을 80 U/ml (GV 핵 샘플에 대해) 또는 40 U/ml (모든 다른 샘플에 대해)의 최종 농도로 첨가하고, 37 ℃에서 정확히 5분 동안 인큐베이션하였다. 반응을 2 μl 프로테이나제 (Proteinase) K (20 mg/ml, 라이프 테크놀로지스)를 함유하는 80 μl 정지 완충제 (Stop Buffer) (10 mM 트리스-HCl, pH 7.5, 10 mM NaCl, 0.15% SDS, 10 mM EDTA)를 첨가함으로써 정지시켰다. 그 후, 20 ng의 원형 운반체 DNA [작은 DNA 단편을 제거하기 위해 0.5x 베크만 SPRI셀렉트 (Beckman SPRIselect) 비드 (베크만 쿨터 (Beckman Coulter))로 정제된 임의의 포유동물 유전자가 없는 순수한 플라스미드 DNA]를 첨가하였다. 혼합물을 50 ℃에서 1시간 동안 인큐베이션한 후, DNA를 페놀-클로로포름으로의 추출에 의해 정제하고, 선형 아크릴아미드 (라이프 테크놀로지스)의 존재 하에서 에탄올에 의해 -20 ℃에서 밤새 침전시켰다. 침전된 DNA를 50 μl TE (2.5 mM 트리스, pH 7.6, 0.05 mM EDTA)에 재현탁시키고, 전체 부피를 시퀀싱 라이브러리 구축에 사용하였다.

[0269] 어댑터 라이게이션을 20 ℃에서 30분 동안 라이게이션 반응에서 0.03 μM 어댑터로 수행하고, PCR 증폭을 카파 하이파이 (Kapa Hifi) 핫스타트 레디믹스 (카파 바이오시스템스 (Kapa Biosystems))를 사용하여 8-사이클 동안 수행한 것을 제외하고는, 일루미나를 위한 NEB넥스트 울트라 II DNA 라이브러리 프렙 키트 (NEBNext Ultra II DNA Library Prep Kit for Illumina) (뉴 잉글랜드 바이오랩스 (New England Biolabs))를 제조업자의 지시서에 따라 사용하여 시퀀싱 라이브러리를 제조하였다. PCR 생성물을 x1.3 부피의 SPRI셀렉트 비드 (베크만 쿨터)로 정제한 후, x0.65 부피, 이어서 x0.7 부피의 SPRI셀렉트 비드로 크기 선택하였다. 샘플을 24 μl TE에 용리하였다. 제2 PCR 증폭에 필요한 사이클의 수를 1 μl의 1:1,000 희석된 샘플을 사용하여 qPCR에 의해 결정하였다. 그 후, 잔류의 23 μl의 샘플을 카파 하이파이 핫스타트 레디믹스로 증폭시켰다 (본 발명자들은 이 연구에서 모든 샘플에 대해 7 사이클을 사용하였다). PCR 생성물을 x1.3 부피의 SPRI셀렉트 비드로 정제한 후, x0.65 부피, 이어서 x0.7 부피의 SPRI셀렉트 비드로 크기 선택하였다. DNA를 30 μl의 TE에 용리하고, 큐비트 (Qubit) dsDNA HS 검정 키트 (써모 피셔 사이언티픽, Q32854) 및 애질런트 (Agilent) 고민감도 검정 키트 (애질런트 테크놀로지스 (Agilent Technologies))에 의해 정량화하였다. 라이브러리를 단일-말단 100 bp 판독물을 갖는 하이섹 (Hiseq)2500 (일루미나 (Illumina)) 상에서 시퀀싱하였다.

[0270] **RNA-시퀀싱**

[0271] RNA-seq 라이브러리를 이전에 기재된 바와 같이 제조하였다. 간략하게, 역전사 및 cDNA 증폭을 스마터 울트라 저 투입 (SMARTer Ultra Low Input) RNA cDNA 제조 키트 (클론테크 (Clontech), 634890)로 전체 배아 용해물을 사용하여 수행하였다. 2-세포 AG, GG 및 α-아마니틴-처리된 IVF 배아 샘플을 프로세싱하는 경우, 1 μl의 1:40,000 희석된 ERCC (외부 RNA 제어 컨소시엄 (External RNA Controls Consortium)) 표준 RNA (라이프 테크

놀로지스)를 세포 용해의 단계에서 각각의 튜브에 첨가하였다. 그 후, cDNA를 마이크로튜브 (microTUBE)-50 (코바리스 (Covaris))을 갖는 코바리스 M220 초음파발생기 (코바리스)를 사용하여 평균 150 내지 160 bp 단편으로 단편화하였다. 단편화된 cDNA를 말단-복구시키고, 어댑터 라이게이션하고, 일루미나를 위한 NEB넥스트 울트라 DNA 라이브러리 프렙 키트를 제조업자의 지시서 (뉴 잉글랜드 바이오랩스)에 따라 사용하여 증폭시켰다. 단일 말단 100 bp 시퀀싱을 하이섹2500 시퀀서 (일루미나) 상에서 수행하였다.

[0272] **liDN아제-seq 데이터 분석**

[0273] liDN아제-seq 데이터의 판독물을 먼저 trim_galore를 갖는 저 품질 및 어댑터로 트리밍한 후, 고유한 맵핑 히트를 유지하기 위해 보타이 (Bowtie) v0.12.9. '-m 1' 파라미터를 사용하여 마우스 게놈 (mm9)에 대해 맵핑하였다. 10 이하의 맵핑 품질 (MAPQ)을 갖는 판독물 또는 동일한 배향을 갖는 동일한 위치에 대해 맵핑된 중복 판독물을 SAM툴스 (SAMtools)로 제거하였다. liDN아제-seq 데이터에서의 DHS 피크를 FDR <= 0.01을 갖는 핫스팟 (Hotspot) 프로그램에 의해 확인하였다. 모든 33개의 라이브러리로부터의 DHS 피크를 베드툴스로부터의 '베드툴스 머지 (bedtools merge)'를 사용하여 병합하였다. 각각의 라이브러리에 대한 각각의 DHS에서의 판독물의 수를 딥툴스 (deepTools)로부터의 '멀티밤서머리 (multiBamSummary)'를 사용하여 계산하고, 맵핑된 판독물의 총 수에 대해 및 DHS의 길이에 대해 정규화하였다 (백만개 맵핑된 판독물 당 1 kb 당 위치 상에 위치한 태그의 가능성). 성 염색체의 수는 모 전핵 사이에 및 웅성발생성 및 자성발생성 배아 사이에 상이하기 때문에, 성 염색체의 판독물을 제거하였다. 게놈-범위 DHS에서의 태그 밀도의 피어슨 상관 계수 (r)를 계산하여 반복실험 사이의 상관을 측정하였다. 접합자 및 상실배 배아에서의 모 대립유전자-특이적 DHS의 확인을 위해, 엄격한 컷오프를 사용하였다 (편향된 대립유전자에서 모든 반복실험에서 RPKM 평균>2, RPKM>1, 및 2개의 대립유전자 사이에 4 초과의 평균 값 배수 변화). 431개의 가장 신뢰성 있는 Ps-DHS를 미세주사된 접합자의 부계 PN의 모든 반복실험에서 추가의 기준 'RPKM>1을 Ps-DHS에 적용함으로써 확인하였다. UCSC 게놈 브라우저 데이터베이스로부터의 RefSeq 유전자 어셈블리 (mm9) 및 이전에 정의된 CGI를 도 2d 및 2e에서 게놈 특색 분포 분석으로서 사용하였다.

[0274] **RNA-seq 데이터 분석**

[0275] 마우스 게놈 (mm9)을 ERCC 대조군과 조합한 주문제작 참조 서열을 구축하였다. RNA-seq의 판독물을 톱햇 (TopHat) v2.0.6 또는 STAR (github.com/alexdobin/STAR)로 참조 게놈에 대해 맵핑하였다. 모든 프로그램을 달리 명시되지 않는다면 디폴트 파라미터로 실행하였다. 고유하게 맵핑된 판독물을 이어서 하위판독물-v1.5.1로부터의 피처카운트 (featureCounts)로 참조 주석 (UCSC 유전자 모델)에 의해 가이드된 전사체 내로 어셈블리하였다. 모든 2-세포 RNA-seq 라이브러리에 대해, 라이브러리 크기 인자를 단지 ERCC 판독물 카운트를 사용하여 R 패키지 DESeq로부터의 'estimateSizeFactors' 기능으로 추정하였다. 라이브러리 크기를 정규화한 후, 각각의 유전자의 발현 수준을 정규화된 FPKM (백만개 맵핑된 단편 당 엑손의 킬로베이스 당 단편)으로 정량화하였다. 이중실험 사이의 상관을 지시하는 유전자 발현 수준의 피어슨 상관 계수 (r)를 계산하였다. 2-세포 단계에서 새롭게 합성된 전사체의 확인을 위해, 통계적으로 비-유의한 유전자를 AG 또는 GG 및 α-아마니틴 처리된 2-세포 배아 사이에서 필터링하였다. 이 목적으로, 조정된 P 값을 음성 이항식 모델을 사용하여 R 패키지 DESeq로부터의 'nbinomTest' 기능으로 계산하고, 단지 FDR<0.05를 갖는 유전자를 선택하였다. 그 후, 추가의 컷오프 [평균 FPKM (AG 또는 GG)>2 및 배수-변화 (FC) (AG/Ama 또는 GG/Ama)>2]를 적용하였다. 그 결과, 4,381개 및 3,916개의 유전자가 AG 및 GG 2-세포 배아에서 각각 새롭게 합성된 유전자로서 확인되었다. 2-세포 배아에서 AG- 및 GG-특이적 DEG를 확인하기 위해, α-아마니틴 2-세포 배아에서 각각의 유전자의 유전자 발현 수준 (FPKM)을 AG 및 GG 배아의 그것으로부터 차감하였다. FC (AG/GG 또는 GG/AG)>10을 나타내는 유전자를 DEG로서 확인하였다.

[0276] **WGBS 및 H3K27me3 ChIP-seq 데이터 분석**

[0277] DHS에서의 DNA 메틸화 수준을 메트파이프 v3.4.2를 사용하여 계산하였다. 각각의 DHS에서의 DNA 메틸화 수준을 계산하는 경우, WGBS 판독물의 충분한 커버리지를 얻기 위해, 각각의 DHS를 2 kb 위 및 아래 둘 다로 연장하여 보다 가까운 CpG 부위를 포함시켰다. 난모세포-메틸화된 gDMR은 난모세포에서 >80% 메틸화 및 정자에서 <20%에 의해 정의되었다. 도 5a에 대해, "베드툴스 메이크윈도우스 (bedtools makewindows)"를 사용하여 Ps-DHS의 ±100 kb 플랭킹 영역에 대해 비-중첩된 1 kb 빈의 세트를 생성하였다. H3K27me3 ChIP-seq 분석을 위해, 베드 파일을 문헌 [Zheng et al., 2016]으로부터 다운로드하고, UCSC 게놈 브라우저 데이터베이스로부터의 '베드클립 (bedClip)' 및 '베드그래프투빅윅 (bedGraphToBigWig)'을 사용하여 빅윅 (bigWig) 포맷으로 전환시켰다. 딥툴스로부터의 '멀티빅윅서머리 (multiBigwigSummary)'를 사용하여 DHS 및 주위의 영역에 비한 H3K27me3 신호를 산

출하였다.

[0278] **통계적 분석 및 데이터 가시화**

[0279] 통계적 분석을 R (www.r-project.org/)로 실행하였다. 피어슨 r 계수를 디폴트 파라미터를 갖는 'cor' 기능을 사용하여 계산하였다. 도 6b 및 10d는 R 기능 'heatmap.2'로 생성되었다. 도 7d, 10c, 및 12a 내지 12d는 R 기능 'pheatmap'로 생성되었다. 도 1b 및 7b는 딥툴스에서 '컴퓨트매트릭스 (computeMatrix)' 및 '플롯히트맵 (plotHeatmap)' 기능을 사용하여 생성되었다. 시퀀싱 판독물에 의한 게놈의 위치-범위 커버리지를, macs2 v2.1.0을 사용하여 라이브러리에서의 총 고유한 맵핑된 판독물에 대해 정규화함으로써 결정하고, IGV 게놈 브라우저에서 커스텀 트랙으로서 가시화하였다.

[0280] **공지된 각인 유전자 정보**

[0281] 공지된 각인 정보를 www.geneimprint.com/site/genes-by-species.Mus+musculus로부터 다운로드하였다.

[0282] **코드 가용성**

[0283] 주문제작된 파이프라인을 사용하여 하이브리드 RNA-seq 데이터를 SNP 정보에 기초하여 그들의 모 기원으로 분할하였다. 코드는 github.com/lanjiangboston/UniversalSNPsplit에서 발견될 수 있다.

[0284] **데이터 가용성 진술**

[0285] 이 연구에서 생성된 모든 liDN아제-seq 및 RNA-seq 데이터세트를 GEO 데이터베이스에 수탁 번호 GSE92605 하에 기탁하였다. 정자 liDN아제-seq 데이터세트는 이전의 공개 (GSE76642)로부터의 것이었다. 정자 및 GV 난모세포에 대한 WGBS 데이터세트를 www.nodai-genome.org/mouse.html?lang=en으로부터 다운로드하였다. 정자, MII 난모세포, 및 1-세포 배아의 SNP-추적된 모계 및 부계 대립유전자의 H3K27me3 ChIP-seq 데이터세트를 이전의 공개 (GSE76687)로부터 다운로드하였다.

[0286] **마우스 착상전 배아의 수집**

[0287] 모든 동물 연구는 하바드 메디칼 스쿨에서의 기관 동물 관리 및 사용 위원회의 지침에 따라 수행되었다. MII-단계 난모세포를 7.5 I.U.의 PMSG (밀리포어) 및 hCG (밀리포어)를 주사함으로써 과배란된 8주령 B6D2F1/J (BDF1) 암컷으로부터 수집하였다. 시험관내 수정 (IVF)을 위해, MII 난모세포를 10 mg/ml 소 혈청 알부민 (BSA; 시그마-알드리치)으로 보충된 HTF 배지에서 성체 BDF1 또는 PWK (잭슨 래버러토리, 003715) 수컷의 꼬리 부고환으로부터 얻어진 활성화된 정자로 수정시켰다. 정자 수정능은 HTF 배지에서 1시간 인큐베이션에 의해 획득되었다. 접합자를 KSOM에 옮기고, 37.8℃에서 5% CO_2/95% 공기를 갖는 습윤 분위기에서 배양하였다.

[0288] **mRNA 주사**

[0289] 수정후 4시간 (hpf)에서, 접합자를 M2 배지 내로 옮기고, mRNA를 피에조 충격-유도된 미세조작장치 (프라임 테크 리미티드, 일본 이바라키켄)를 사용하여 주사하였다. mRNA의 구축 및 제조는 상기 기재되었다. *Kdm6b*^{WT} 및 *Kdm6b*^{MUT}의 주사된 mRNA의 농도는 1.8 μg/μl이고, *Kdm4d*^{WT} 및 *Kdm4d*^{MUT}의 그것은 1.5 μg/μl이었다.

[0290] **형광 계내 혼성화를 위한 프로브**

[0291] *Xist* RNA를 위한 프로브를 Cy3-dCTP (지이 헬스케어 (GE healthcare), PA53021)를 갖는 닉 (Nick) 번역 시약 키트 (애보트 몰리큘라 (Abbott Molecular), 07J00-001)를 제조업자의 지시서에 따라 사용함으로써 제조하였다. 프로브 제조에 사용된 주형 DNA는 루돌프 재니쉬 (Rudolf Jaenisch)로부터의 기증품인 전장 마우스 *Xist* 유전자를 코딩하는 플라스미드였다 (pCMV-Xist-PA, 26760) (Wutz and Jaenisch, 2000). DNA FISH를 위한 프로브를 그린 (Green)-dUTP를 갖는 동일한 키트 (애보트 몰리큘라, 02N32-050)를 사용하여 제조하였다. 주형 DNA는 *Rnf12* 좌위를 함유하는 BAC 클론 (RP23-36C20)이었다. 형광 프로브를 5 μg Cot-1 DNA (라이프 테크놀로지스), 5 μg 청어 정자 DNA (써모 피셔 사이언티픽), 및 2.5 μg 효모 tRNA (써모 피셔 사이언티픽, AM7119)로 에탄올 침전시킨 후, 20 μl 포름아미드 (써모 피셔 사이언티픽, 17899)로 용해시켰다. 프로브를 4℃에서 저장하였다. 사용하기 전에, 프로브 (각각 0.75 μl)를 0.75 μl Cot-1 DNA와 혼합하고, 이를 에탄올 침전시키고, 포름아미드, 및 2.25 μl의 4xSSC/20% 덱스트란 (Dextran) (밀리포어 S4030)에 용해시켰다. 프로브 혼합물을 80℃에서 30분 동안 가열한 후, 37℃ 인큐베이터에 옮겼다 ('사전-어닐링된 프로브').

[0292] **전조직 표본 RNA/DNA 형광 계내 혼성화**

[0293] 상실배 배아를 78 hpf에서 0.5% 트리톤 X-100을 함유하는 PBS 중 2% 파라포름알데히드 (PFA)에서 실온에서 20분 동안 고정시켰다. 1 mg/ml BSA를 함유하는 PBS (PBS/BSA)로 3회 세척 후, 배아를 0.02% 트리톤 X-100을 함유하는 0.1 N HCl로 4℃에서 15분 동안 처리하였다. 0.1% BSA를 함유하는 2xSSC로 3회 세척 후, 배아를 유리 디쉬 (일렉트론 마이크로스코피 사이언스 (Electron Microscopy Science), 705430-30)에서 15 μl의 10% 포름아미드/2xSSC에서 인큐베이션하였다. 모든 배아를 가라앉히고, 부드러운 피펫팅에 의해 유리 디쉬의 바닥에 부착시켰다. 5분 후, 15 μl의 30% 포름아미드/2xSSC를 첨가하였다. 5분 후, 90 μl의 60% 포름아미드/2xSSC를 첨가하여 최종 포름아미드 농도 50%를 만들고, 배아를 실온에서 추가의 30분 동안 인큐베이션하였다. 배아를 함유하는 포름아미드 용액을 광물유로 덮었다. 샘플을 80℃에서 30분 동안 가열한 후, 37℃ 인큐베이터에 적어도 30분 동안 옮겼다. 배아를 유리 피펫에서 집어내고, 또다른 유리 디쉬 상의 광물유로 덮힌 4.5 μl의 '사전-어닐링된 프로브' 내로 옮기고, 37℃에서 적어도 24시간 동안 인큐베이션하였다. 배아를 0.1%BSA를 함유하는 사전-가온된 (42℃) 2xSSC로 세척하고, 마지막 액적에서 30분 동안 방치하였다. 1%BSA/PBS로 3회 세척 후, 이들을 4',6-디아미디노-2-페닐인돌 (DAPI)을 갖는 벡타쉴드 항-표백 용액 (벡터 래버러토리스, 미국 캘리포니아주 벌링게임) 중 유리 슬라이드 상에 마운팅하였다. 형광을 레이저-스캐닝 공초점 현미경 자이스 LSM800 하에서 검출하였다.

[0294] **전조직 표본 면역염색**

[0295] 면역염색 및 정량화의 절차는 상기 기재되었다.

[0296] **모계 대립유전자-편향된 H3K27me3의 컴퓨터 확인**

[0297] 내부 세포 덩이 (ICM)에서 H3K27me3 ChIP-seq를 위한 100 bp 빈으로의 RPKM 값을 포함하는 베드 파일을 GEO로부터 번호 GSE76687 하에 다운로드하였다. 2개의 모 대립유전자 및 대립유전자 판독물에 대한 RPKM 값을 함유하는 모계 또는 부계로 표지된 베드 파일을 총 판독물 수에 대해 정규화하였다. '베드툴스 메이크윈도우스'를 사용하여 mm9 게놈에 대한 1000 bp 빈을 생성한 후, 각각의 빈에 대한 RPKM 값을 '베드툴스 맵 (bedtools map)'에 의해 계산하였다. 모든 빈을 1의 신호 컷오프 및 4의 배수 변화 컷오프를 사용하여 신호 없음, 이중대립유전자, 모계 편향의 3개의 범주로 분류하였다. 슬라이딩 창 접근법을 사용하여 20 kb의 창 크기, 3의 최소 빈 수 및 50% 초과의 모계 편향된 H3K27me3 빈의 백분율의 기준을 갖는 모계 편향된 H3K27me3 빈을 함유하는 창을 확인하였다. 중첩된 창을 "베드툴스 머지"로 병합하였다. 총 5986개의 창을 게놈에서 확인하였다.

[0298] **RNA-시퀀싱**

[0299] RNA-seq 라이브러리를 약간의 변형으로 상기 기재된 바와 같이 제조하였다. 간략하게, 역전사 및 cDNA 증폭을 스마터 울트라 저 투입 RNA cDNA 제조 키트 (클론테크, 634890)로 전체 배아 용해물을 사용하여 수행하였다. 그 후, cDNA를 넥스테라 (Nextera) XT DNA 라이브러리 프렙 키트 (일루미나)로의 태그단편화를 사용하여 단편화하였다. 단편화된 cDNA를 넥스테라 PCR 마스터 믹스를 제조업자의 지시서에 따라 사용하여 증폭시켰다. 단일 말단 100 bp 시퀀싱을 하이섹2500 시퀀서 (일루미나) 상에서 수행하였다.

[0300] **RNA-seq 데이터 분석**

[0301] RNA-seq의 판독물을 STAR (github.com/alexdobin/STAR)로 참조 게놈에 대해 맵핑하였다. 모든 프로그램은 달리 명시되지 않는다면 디폴트 파라미터로 실행되었다. 고유하게 맵핑된 판독물을 이어서 하위판독물-v1.5.1로부터의 피처카운트로 참조 주석 (UCSC 유전자 모델)에 의해 가이드된 전사체 내로 어셈블리하였다. 라이브러리 크기를 정규화한 후, 각각의 유전자의 발현 수준을 정규화된 FPKM (백만개 맵핑된 단편 당 엑손의 킬로베이스 당 단편)으로 정량화하였다. 이중실험 사이의 상관을 지시하는 유전자 발현 수준의 피어슨 상관 계수 (r)를 계산하였다.

[0302] 통계적 분석을 R (www.r-project.org/)로 실행하였다. 피어슨 r 계수를 디폴트 파라미터를 갖는 'cor' 기능을 사용하여 계산하였다.

[0303] **코드 가용성**

[0304] 주문제작된 파이프라인을 사용하여 하이브리드 RNA-seq 데이터를 SNP 정보에 기초하여 그들의 모 기원으로 분할하였다. 코드는 github.com/lanjiangboston/UniversalSNPsplit에서 발견될 수 있다.

[0305] **데이터 가용성**

[0306] 이 연구에서 생성된 RNA-seq 데이터세트를 GEO 데이터베이스에 수탁 번호 GSEXXXXX 하에 기탁하였다. GV 난모

세포에 대한 WGBS 데이터세트를 www.nodai-genome.org/mouse.html?lang=en으로부터 다운로드하였다. 동일한 100 bp 빈으로부터의 WGBS 판독물을 함께 풀링하여 평균 메틸화 수준을 계산하였으며, 10 판독물의 최소 커버리지가 요구되었다. 정자, MII 난모세포, 및 1-세포, 2-세포, 및 배반포 배아의 내부 세포 덩이의 SNP-추적된 모계 및 부계 대립유전자의 H3K27me3 ChIP-seq 데이터세트를 이전의 연구 (GSE76687)로부터 다운로드하였다. 난모세포 DN아제I-seq 데이터세트는 상기 (GSE92605)로부터의 것이었다.

[0307] **마우스**

[0308] B6D2F1/J (BDF1) 마우스를 SCNT를 위한 수용자 난모세포의 수집에 사용하였다. 마우스 배아 섬유모세포 (MEF) 세포 제조를 위해, 129S1/SvImj 배경 (Marahrens et al., 1997)에서 유지된 *Xist* KO 암컷 마우스를 CAST/EiJ 수컷과 교배시켜 129/CAST F1 배경에서 *Xist* 이형접합성 KO 배아를 생성하였다. 난구 및 세르톨리 세포 제조를 위해, C57BL/6N 배경 (Sado et al., 2005)에서의 *Xist* KO 암컷 마우스를 DBA/2N 수컷과 교배시켜 BDF1 배경에서 *Xist 이형접합성* KO 배아를 생성하였다. 모든 동물 실험은 하바드 메디칼 스쿨 및 리켄 츠쿠바 인스티튜트 (RIKEN Tsukuba Institute)의 기관 동물 관리 및 사용 위원회에 의해 승인되었다.

[0309] **공여자 세포 제조**

[0310] 1차 MEF 세포는 13.5 dpc에서 *Xist* KO 수컷 마우스 배아로부터 유래되었다. 머리 및 모든 기관의 제거 후, 나머지 몸통으로부터의 갈려진 조직을 1 mM EDTA (써모 피셔 사이언티픽 # 25200056)를 갖는 500 μl의 0.25% 트립신에서 37℃에서 10분 동안 분리하였다. 세포 현탁액을 10% FBS 및 페니실린/스트렙토마이신 (써모 피셔 사이언티픽 # 15140-022)을 함유하는 동등한 양의 DMEM (써모 피셔 사이언티픽 # 11995-073)으로 희석하고, 상하로 20회 피펫팅하였다. 세포 현탁액을 새로운 배지로 희석하고, 100 mm 디쉬 상으로 플레이팅하고, 37℃에서 배양하였다. 2일 후, MEF 세포를 수확하고, 동결시켰다. MEF 세포의 동결된 스톡을 해동시키고, 한 계대 후에 실험에 사용하였다.

[0311] 난구 세포를 7.5 IU의 임신 말 혈청 고나도트로핀 (PMSG; 밀리포어 # 367222) 및 7.5 IU의 인간 융모막성 고나도트로핀 (hCG; 밀리포어 # 230734)을 주사함으로써 과배란을 통해 야생형 (WT) 및 *Xist* 이형접합성 KO 성체 암컷 (리켄 바이오리소스 센터 (RIKEN BioResource Center), RBRC01260)으로부터 수집하였다. hCG 주사 후 15 내지 17시간에, 난구-난모세포 복합체 (COC)를 난관으로부터 수집하고, 300 U/ml 소 고환 히알루로니다제 (칼바이오켐 (Calbiochem) # 385931)를 함유하는 헤페스 (Hepes)-완충 칼륨 단순-최적화 배지 (KSOM)로 간단히 처리하여 분리된 난구 세포를 얻었다.

[0312] 세르톨리 세포를 기재된 바와 같이 (Matoba et al., 2011) 3 내지 5일령 WT 또는 *Xist* KO 수컷 마우스의 고환으로부터 수집하였다. 고환 덩이를 0.1 mg/ml 콜라게나제 (써모 피셔 사이언티픽 # 17104-019)를 함유하는 PBS에서 37℃에서 30분 동안 인큐베이션하고, 이어서 1 mM EDTA를 갖는 0.25% 트립신으로 실온에서 5분 처리하였다. 3 mg/ml 소 혈청 알부민을 함유하는 PBS로 4회 세척한 후, 분리된 세포를 헤페스-KSOM 배지에 현탁시켰다.

[0313] ***Kdm4d* mRNA 합성**

[0314] *Kdm4d* mRNA를 이전에 기재된 바와 같이 (Matoba et al., 2014) 시험관내 전사 (IVT)에 의해 합성하였다. 간략하게, 전장 마우스 *Kdm4d*, 이어서 폴리(A)83을 함유하는 pcDNA3.1 플라스미드 (애드진 (Addgene) # 61553)를 *Xba*I에 의해 선형화하였다. 정제 후, 선형화된 플라스미드 DNA를 mMESSAGE mMACHINE T7 울트라 키트 (써모 피셔 사이언티픽 # AM1345)를 사용하여 IVT를 위한 주형으로서 사용하였다. 합성된 mRNA를 뉴클레아제-무함유수에 용해시키고, 나노드롭 (NanoDrop) ND-1000 분광광도계 (나노드롭 테크놀로지스 (NanoDrop Technologies))에 의해 정량화하였다. mRNA를 1500 ng/μl로 희석한 후, 분취물을 -80℃에서 저장하였다.

[0315] **SCNT**

[0316] 마우스 체세포 핵 전달을 이전에 기재된 바와 같이 수행하였다 (Matoba et al., 2014; Ogura et al., 2000). 간략하게, 수용자 MII 난모세포를 7.5 IU의 PMSG 및 7.5 IU의 hCG를 주사함으로써 과배란을 통해 성체 BDF1 암컷 마우스로부터 수집하였다. hCG 주사 후 15 내지 17시간에, 난구-난모세포 복합체 (COC)를 난관으로부터 수집하고, 300 U/ml 소 고환 히알루로니다제를 함유하는 헤페스-KSOM으로 간단히 처리하여 MII 난모세포를 얻었다. 단리된 MII 난모세포를 피에조-유도된 미세조작장치 (프라임테크 # PMM-150FU)를 사용함으로써 7.5 μg/ml의 시토칼라신 B (칼바이오켐 # 250233)를 함유하는 헤페스-완충 KSOM 배지에서 제핵시켰다. 난구 또는 세르톨리 세포의 핵을 제핵된 난모세포 내로 주사하였다. MEF 세포를 불활성화된 센다이 바이러스 외피 (게놈원 (GenomOne) CF); 이시하라 산교 (Ishihara Sangyo) #CF001)에 의해 제핵된 난모세포와 융합시켰다. KSOM에서 1시간 인큐베이션 후, 재구축물된 SCNT 난모세포를 3 mM 염화스트론튬 (SrCl2) 및 5 μg/ml 시토칼라신 B를 함

유하는 Ca-무함유 KSOM에서 1시간 동안 인큐베이션함으로써 활성화시키고, 5 μg/ml 시토칼라신 B를 갖는 KSOM에서 4시간 동안 추가로 배양하였다. 활성화된 SCNT 배아를 $SrCl_2$ 처리의 개시 후 5시간 (활성화후 시간, hpa)에 세척하고, 37.8℃에서 5% CO_2를 갖는 습윤 분위기에서 KSOM에서 배양하였다. 일부 SCNT 배아를 피에조-유도된 미세조작장치를 사용함으로써 5 내지 6 hpa에서 약 10 pl의 1500 ng/μl 마우스 Kdm4d mRNA로 주사하였다.

[0317] **배아 전달**

[0318] 2-세포 단계 SCNT 배아를 가임신된 (E0.5) ICR 암컷의 난관으로 전달하였다. 새끼를 분만일 (E19.5)에 제왕 절개에 의해 회수하고, 수유 ICR 암컷에 의해 길렀다. 일부 암컷을 E4.5 및 E10.5에서 배아 발생을 조사하기 위해 희생시켰다.

[0319] **태반의 조직학적 분석**

[0320] E19.5에서의 태반을 4% 파라포름알데히드 (PFA)에서 4℃에서 밤새 고정시키고, 파라핀에 통상적으로 포매시켰다. 일련의 섹션 (두께로 4 μm)을 과요오드산 쉬프 (PAS) 염색으로 처리하였다.

[0321] **배반포에서의 H3K27me3에 대한 면역염색**

[0322] 배반포를 4% 파라포름알데히드 (PFA)로 실온에서 20분 동안 고정시켰다. 10 mg/ml BSA를 함유하는 PBS (PBS/BSA)로 세척한 후, 고정된 배아를 0.5% 트리톤-X 100으로 15분 인큐베이션에 의해 투과화하였다. 실온에서 1시간 동안 PBS/BSA에서 차단시킨 후, 이들을 토끼 항-H3K27me3 항체 (1/500, 밀리포어, 07-449), 염소 항-Oct4 항체 (1/500, 산타크루즈 (SantaCruz), sc-8628) 및 마우스 항-Cdx2 항체 (1/100, 바이오제넥스 (BioGenex), AM392-5M)를 포함하는 1차 항체의 혼합물에서 4℃에서 밤새 인큐베이션하였다. PBS/BSA로 3회 세척 후, 배아를 플루오레세인 이소티오시아네이트-접합된 항-마우스 IgG (1/400, 잭슨 이뮤노-리서치), 알렉사 플루오르 546 당나귀 항-토끼 IgG (1/400, 써모 피셔 사이언티픽) 및 알렉사 플루오르 647 당나귀 항-염소 IgG (1/400, 써모 피셔 사이언티픽)를 포함하는 2차 항체의 혼합물과 함께 실온에서 1시간 동안 인큐베이션하였다. 마지막으로, 이들을 4',6-디아미디노-2-페닐인돌 (DAPI)을 갖는 벡타쉴드 (벡터 래버러토리스 # H-1200)로 마운팅하였다. 형광 신호를 레이저-스캐닝 공초점 현미경 (자이스 LSM510) 및 EM-CCD 카메라 (하마마츠 ImagEM)를 사용하여 관찰하였다.

[0323] **WGBS**

[0324] 초기 배반포 단계 (수정 또는 활성화 후 96시간)의 IVF 및 SCNT 배아를 EZ DNA 메틸화-다이렉트 키트 (EZ DNA Methylation-Direct Kit) (자이모 리서치 (Zymo Research), D5020)를 사용하여 비술파이트 전환으로 직접적으로 처리하였다. 39개 및 36개의 배아를 각각 IVF 및 SCNT 샘플을 제조하는 데 사용하였다. 소량 (0.01 ng)의 비메틸화된 람다 (Lambda) DNA (프로메가 (Promega), D152A)를 비술파이트 전환 전에 각각의 샘플에 첨가하여 비술파이트 전환 효율을 평가하기 위한 스파이크-인 컨트롤로 기능하게 하였다. 시퀀싱 라이브러리를 에피게놈 메틸-Seq (EpiGnome Methyl-Seq) 키트 (에피센터 (Epicenter), EGMK81312)를 제조업자의 지시서에 따라 사용하여 제조하였다. 라이브러리를 단지 12 사이클 동안 증폭시킨 후, 아젠코트 AM퓨어 (Agencourt AMPure) XP 비드 (베크만 쿨터, A63880)를 사용하여 정제하였다. 최종 라이브러리를 PhiX 스파이크-인 컨트롤로 하이섹 2500 시퀀서 (일루미나) 상에서 단일-판독물 (100 bp) 시퀀싱으로 처리하였다.

[0325] **RNA-seq**

[0326] 초기 배반포 단계 (수정 또는 활성화 후 96시간)의 6개의 IVF 또는 SCNT 배아를 직접적으로 용해시키고, 스마터 울트라 저 투입 RNA cDNA 제조 키트 (클론테크, 634936)를 사용한 cDNA 합성에 사용하였다. 증폭 후, cDNA 샘플을 코바리스 M220 초음파발생기 (코바리스)를 사용하여 단편화하였다. 시퀀싱 라이브러리를 일루미나를 위한 NEB넥스트 울트라 DNA 라이브러리 프렙 키트를 제조업자의 지시서 (뉴 잉글랜드 바이오랩스, E7370)에 따라 사용하여 단편화된 cDNA로 제조하였다. 단일-판독물 50 bp 시퀀싱을 하이섹 2500 시퀀서 (일루미나) 상에서 수행하였다.

[0327] **정량화 및 통계적 분석**

[0328] **WGBS 및 RRBS 데이터 분석**

[0329] WGBS 및 감소된 제시 비술파이트 시퀀싱 (RRBS) 판독물을 먼저 trim_galore를 사용하여 트리밍하여 저-품질 서열 및 어댑터 서열을 제거하였다. 비스마크 (버전 0.15.0)를 사용하여 판독물을 비술파이트 전환된 참조 게놈

(mm9)으로 정렬하였다. 각각의 시토신의 커버리지 깊이 및 메틸화 수준을 비스마크_메틸화_추출기(bismark_methylation_extractor)로 정렬된 판독물로부터 추출하였다. CpG 부위에 대한 메틸화 수준을 계산하는 경우, 둘 다의 가닥으로부터의 정보가 조합되고, 적어도 5개의 판독물의 커버리지가 요구되었다. DMR을 메트파이프 (버전 3.4.3)를 사용하여 확인하고, 적어도 10개의 CpG 부위 및 적어도 10% 메틸화 차이를 요구하여 추가로 필터링하였다. DMR 연관된 유전자 (즉, TSS±3kb 영역에 위치한 DMR을 갖는 유전자)의 기능적 주석을 R에서 클러스터프로파일러 (버전 2.4.3)로 수행하였다. 공지된 ICR의 대립유전자 특이적 메틸화 분석을 위해, 1개의 ICR 내의 모든 검출된 CpG를 함께 풀링하였으며, 둘 다의 대립유전자에서 적어도 5개의 검출된 CpG의 커버리지는 추가의 메틸화 비교에 요구되었다.

[0330] **RNA-seq 데이터 분석**

[0331] RNA-seq 데이터를 파라미터 "--no-coverage-search --no-novel-juncs --library-type=fr-unstranded"를 갖는 톱햇 (버전 2.0.14)으로 마우스 게놈 (mm9)에 대해 맵핑하였다. 고유하게 맵핑된 판독물을 이어서 커프링크스 (Cufflinks) (버전 2.2.1)로 참조 주석 (UCSC 유전자 모델)에 의해 가이드된 전사체 내로 어셈블리하였다. 각각의 유전자의 발현 수준을 정규화된 FPKM (백만개 맵핑된 단편 당 엑손의 킬로베이스 당 단편)으로 정량화하였다. 이중실험 사이의 상관을 지시하는 유전자 발현 수준의 피어슨 상관 계수를 계산하였다.

[0332] **대립유전자 특이적 분석**

[0333] 하이브리드 WGBS 및 RNA-seq 데이터 (129S1/Svj x CAST/EiJ)를 마우스 게놈 (mm9)에 대해 맵핑한 후, 커스텀 Perl 스크립트를 사용하여 맵핑된 판독물을 마우스 게놈 프로젝트 (Mouse Genomes Project) (ftp://ftp-mouse.sanger.ac.uk/REL-1211-SNPs_Indels/)로부터 다운로드된 SNP 정보에 기초하여 그들의 모 기원으로 분할하였다. 그 후, 대립유전자 판독물을 개별적으로 프로세싱하였다.

[0334] **ChIP-seq 및 DIP-seq 데이터 분석**

[0335] 다운로드된 ChIP-seq 및 DIP-seq 판독물을 파라미터 "-D 20 -R 3 -N 1 -L 20 -i S,1,0.50"을 갖는 보타이 (버전 2.1.0)를 사용하여 마우스 게놈 (mm9)에 대해 맵핑하여 단지 많아야 3개의 미스매치를 갖는 고유하게 맵핑된 판독물을 얻었다. 게놈 브라우저에서 신호를 가시화하기 위해, 본 발명자들은 고유하게 맵핑된 판독물을 (동일한 게놈 위치에서 많아야 2개의 판독물을 유지하면서) 3' 말단을 향해 200 bp로 연장하고, 판독물 카운트를 50 bp 간격으로 비닝함으로써 MACS2 (버전 2.1.1)로 각각의 데이터 세트에 대한 위그 트랙 파일을 생성하였다. 태그 카운트를 각각의 빈에서 고유하게 맵핑된 판독물의 총 수에 대해 정규화하였다 (백만개 판독물 당 판독물, RPM). 딥툴스로부터의 '컴퓨트매트릭스 (computeMatrix)' 프로그램을 사용하여 DMR 또는 ICR 및 그들의 플랭킹 영역에 비한 ChIP-seq 및 DIP-seq 신호를 산출하였다.

[0336] **통계적 분석 및 데이터 가시화**

[0337] 통계적 분석 및 플롯을 R (버전 3.4.1, http://www.r-project.org)로 실행하였다. 피어슨 상관 계수를 디폴트 파라미터를 갖는 'cor' 기능을 사용하여 계산하였다. 스튜던트 (Student) *t*-검정 (양측, 동등 분산)을 디폴트 파라미터를 갖는 't.검정' 기능을 사용하여 수행하였다. ChIP-seq 신호 및 DNA 메틸화 수준을 통합 게노믹스 뷰어 (Integrative Genomics Viewer) 게놈 브라우저에서 커스텀 트랙으로서 가시화하였다.

[0338] **데이터 및 소프트웨어 가용성**

[0339] 이 연구에서 생성된 WGBS 및 RNA-seq 데이터세트는 유전자 발현 옴니버스 (Gene Expression Omnibus) (GEO)에 수탁 번호 GSE109214 하에 기탁되었다.

[0340] **이 연구에 사용된 공개된 데이터세트**

[0341] 착상전 배아 (2-세포, 4-세포, 및 ICM)의 모계 및 부계 DNA 메틸화는 GSE56697 (Wang et al., 2014)로부터 얻었다. 상이한 세포 및 체세포 조직의 RRBS 데이터는 GSE11034 및 GSE43719 (Soumillon et al., 2013)로부터 얻었다. H3K27me3 ChIP-seq 데이터는 GSE49847 (Yue et al., 2014) 및 GSE76687 (Zheng et al., 2016)로부터 얻었다. PGC 발생 동안 5mC 및 5hmC의 DIP-seq 데이터는 SRP016940 (Hackett et al., 2013)으로부터 다운로드하였다.

[0342] **다른 실시양태**

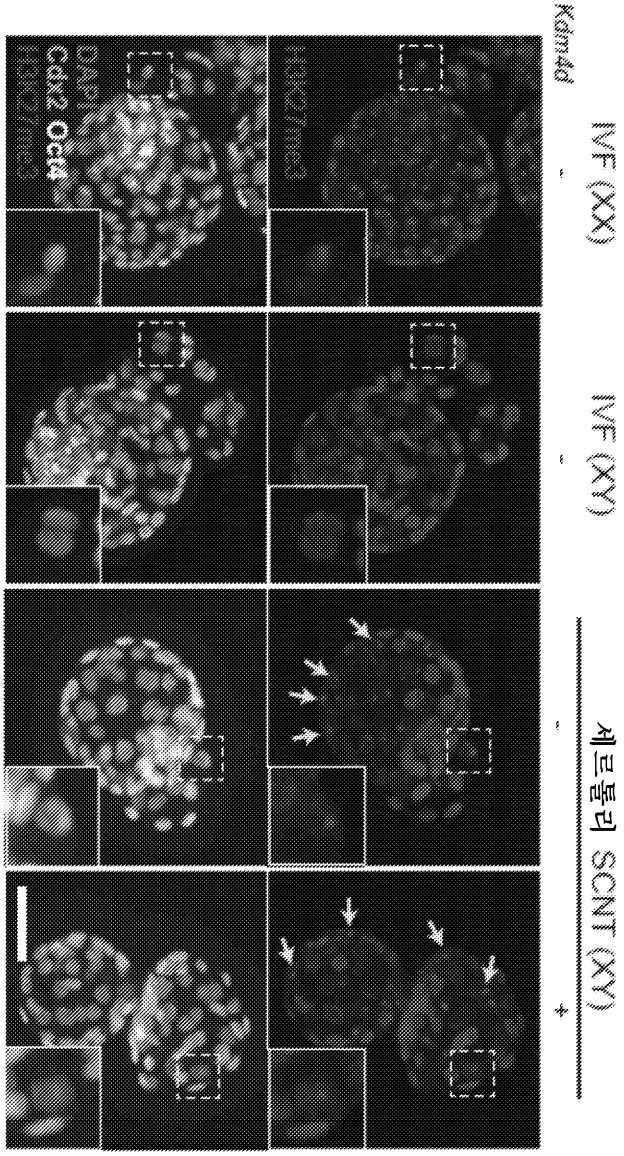
[0343] 상기 설명으로부터, 변화 및 변형이 이를 다양한 용법 및 조건에 채택하기 위해 본원에 기재된 본 발명에 이루어질 수 있음이 명백할 것이다. 이러한 실시양태는 또한 하기 청구범위의 범위 내에 있다.

[0344] 본원에서 변수의 임의의 정의에서 요소의 목록의 나열은 열거된 요소의 임의의 단일 요소 또는 조합 (또는 하위 조합)으로서 그 변수의 정의를 포함한다. 본원에서 실시양태의 나열은 임의의 단일 실시양태로서 또는 임의의 다른 실시양태 또는 그의 부분과 조합으로 그 실시양태를 포함한다.

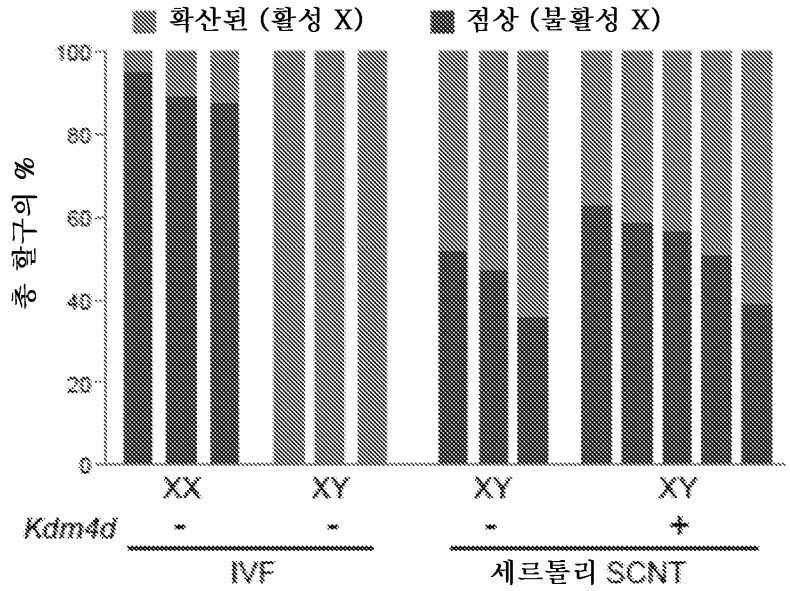
[0345] 이 명세서에 언급된 모든 특허 및 간행물은 각각의 독립적인 특허 및 간행물이 구체적으로 및 개별적으로 참조로 포함되는 것으로 지시되는 것과 동일한 정도로 본원에 참조로 포함된다.

도면

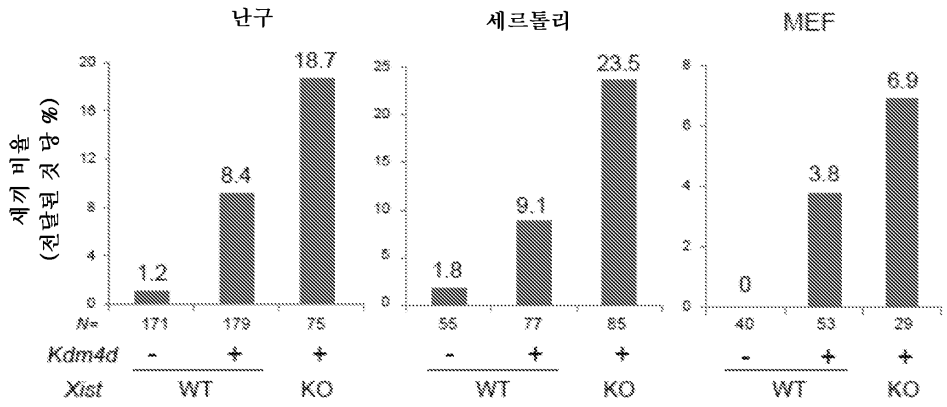
도면1a



도면1b



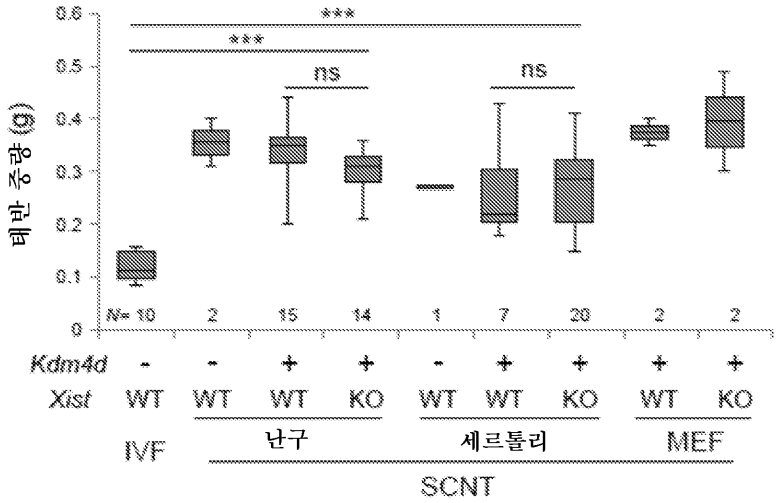
도면1c



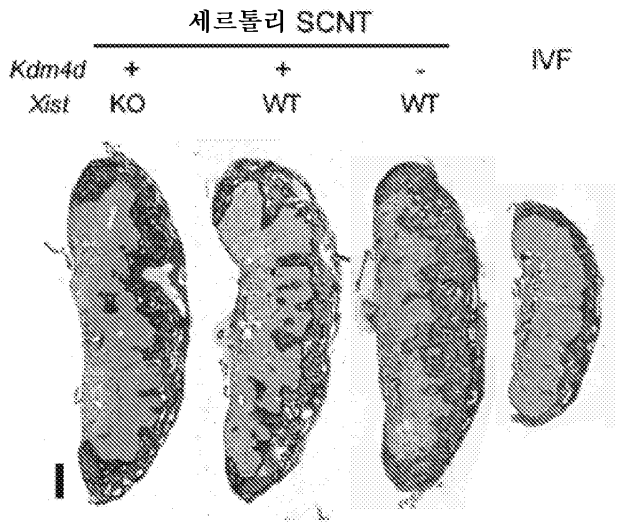
도면1d



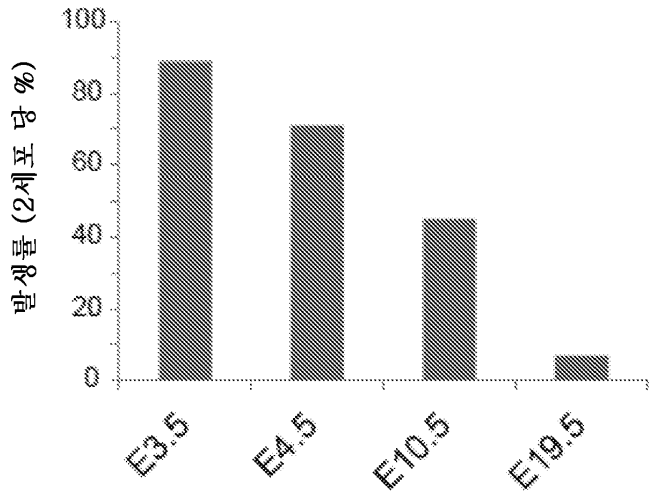
도면1e


태반 중량 (g)
0.6
0.5
0.4
0.3
0.2
0.1
0
***
***
ns
ns
N= 10 2 15 14 1 7 20 2 2
Kdm4d - - + + - + + + +
Xist WT WT WT KO WT WT KO WT KO
IVF
난구
세르톨리
MEF
SCNT

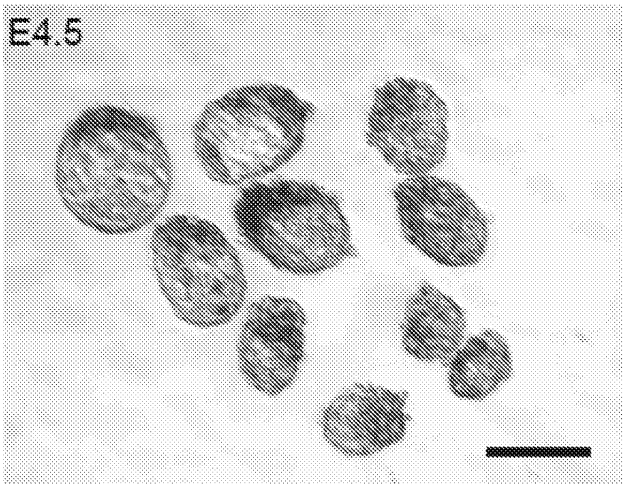
도면1f


세르톨리 SCNT
IVF
Kdm4d + + -
Xist KO WT WT

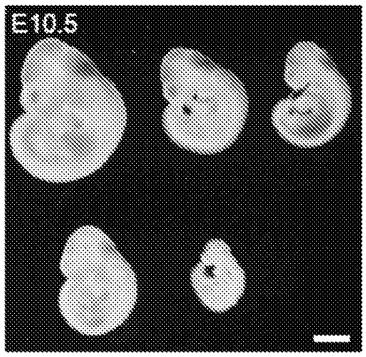
도면2a



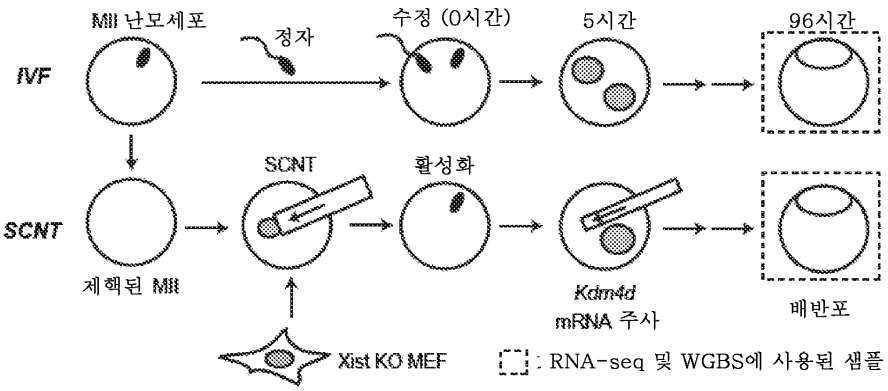
도면2b



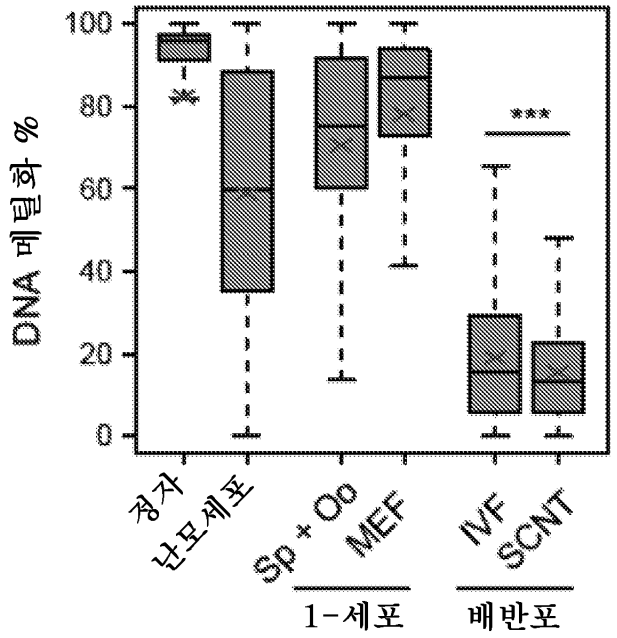
도면2c



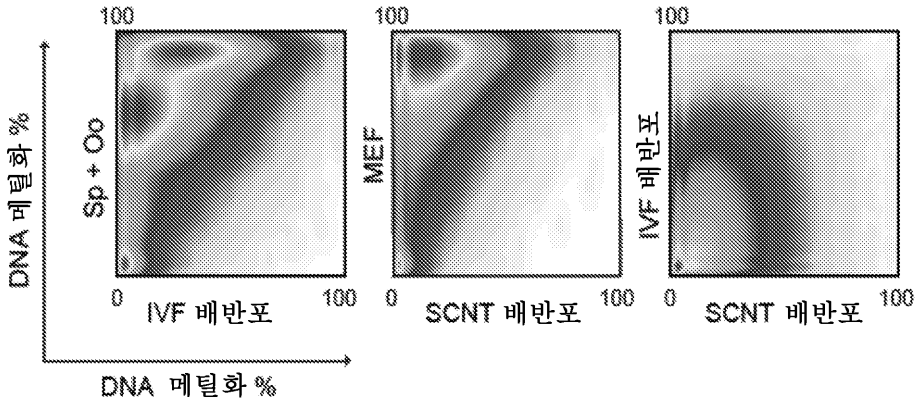
도면3a



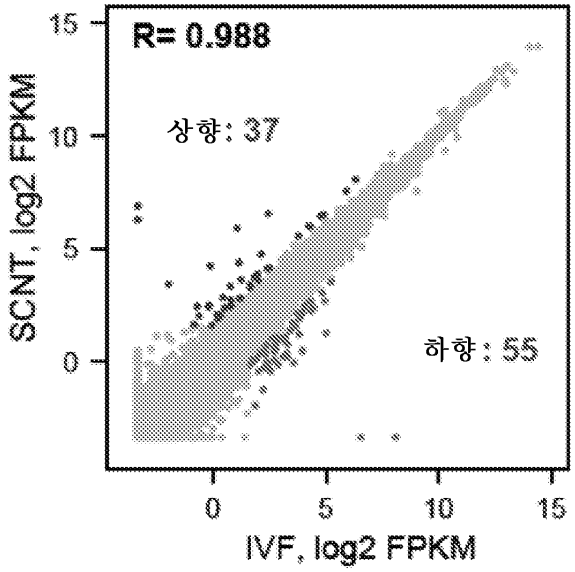
도면3b



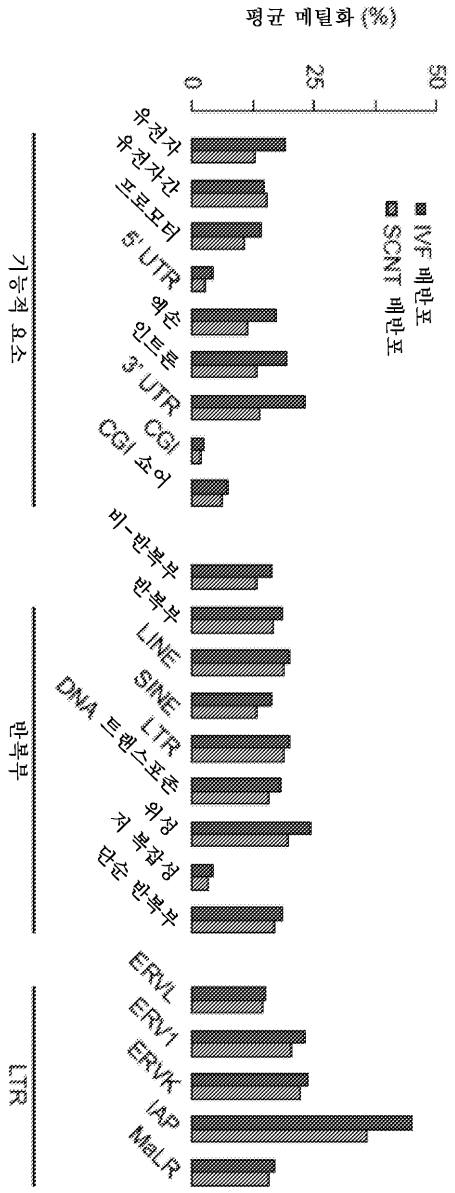
도면3c



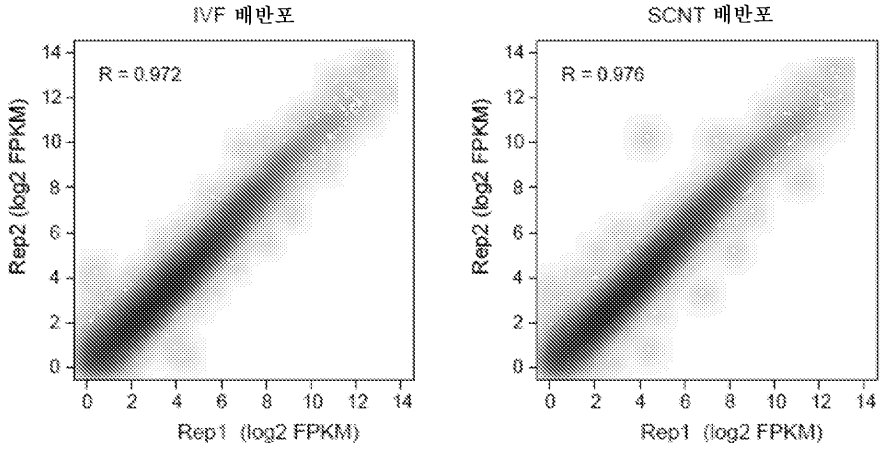
도면3d



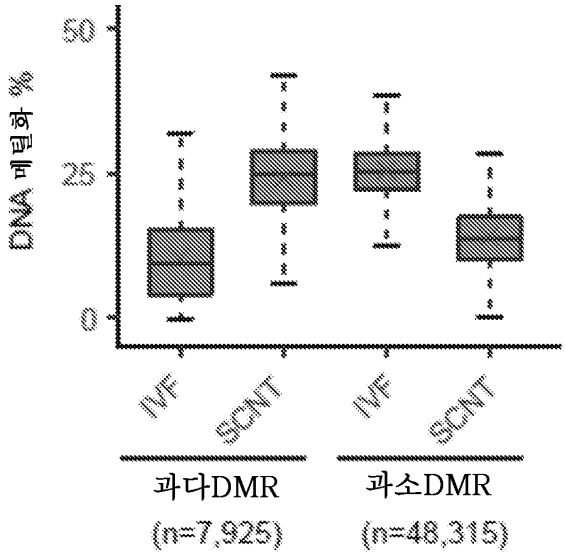
도면4a


평균 메틸화 (%)
0
25
50
IVF 배반포
SCNT 배반포
기능적 요소
반복부
LTR
유전자
유전자간
프로모터
5' UTR
엑손
인트론
3' UTR
CGI
CGI 쇼어
비-반복부
반복부
LINE
SINE
LTR
DNA 트랜스포존
위성
저 복잡성
단순 반복부
ERVL
ERV1
ERVK
IAP
MaLR

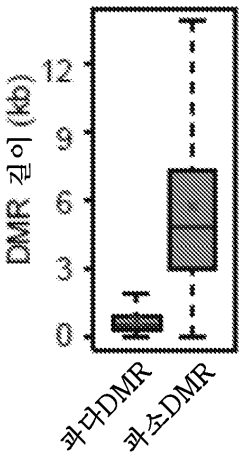
도면4b


IVF 배반포
R = 0.972
Rep2 (log2 FPKM)
Rep1 (log2 FPKM)
SCNT 배반포
R = 0.976
Rep2 (log2 FPKM)
Rep1 (log2 FPKM)

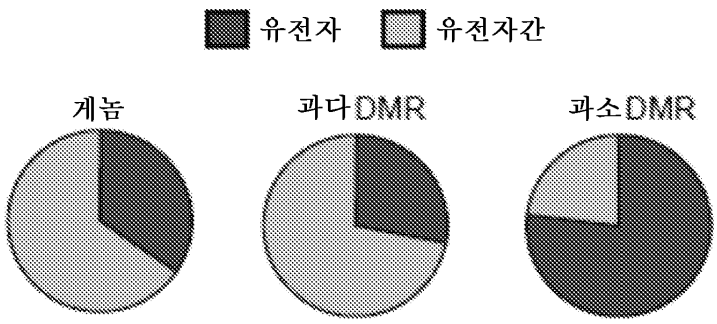
도면5a



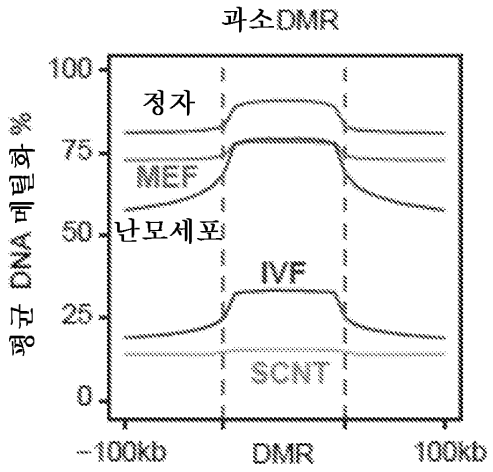
도면5b



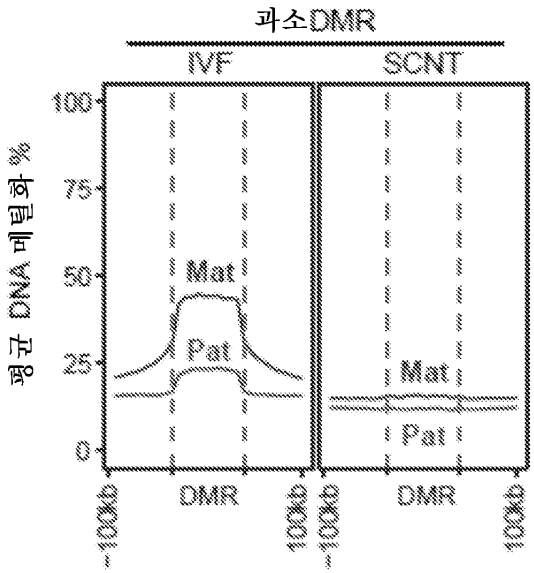
도면5c



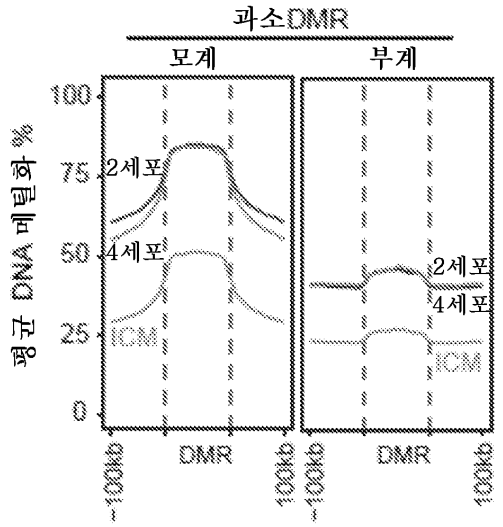
도면5d


과소DMR
100
75
50
25
0
평균 DNA 메틸화 %
정자
MEF
난모세포
IVF
SCNT
-100kb
DMR
100kb

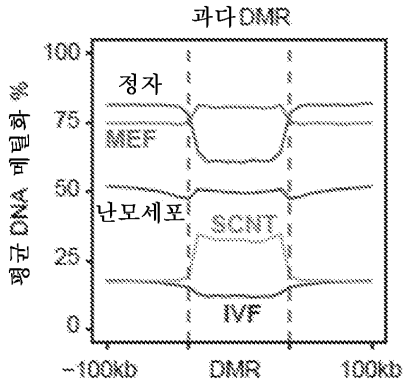
도면5e


과소DMR
IVF
SCNT
100
75
50
25
0
평균 DNA 메틸화 %
Mat
Pat
Mat
Pat
-100kb
DMR
100kb
-100kb
DMR
100kb

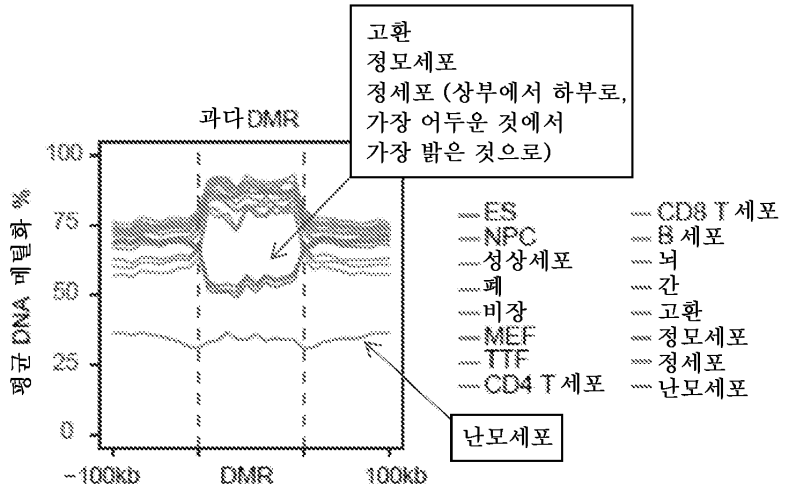
도면5f



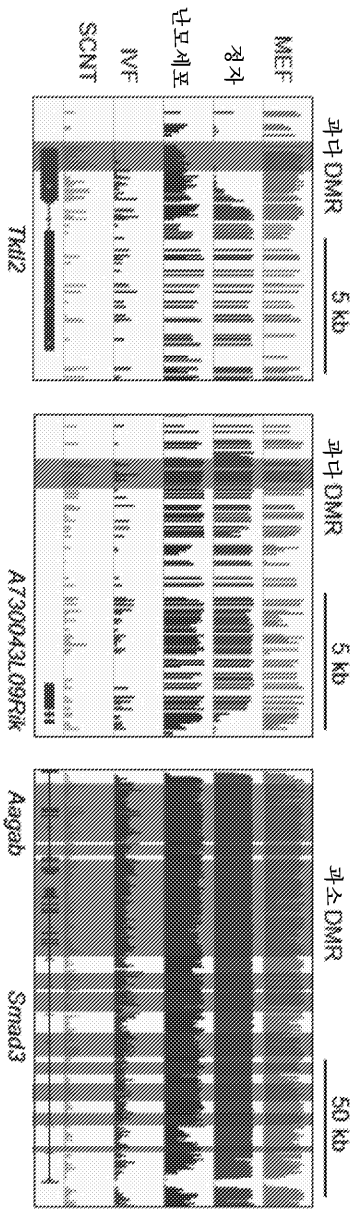
도면5g



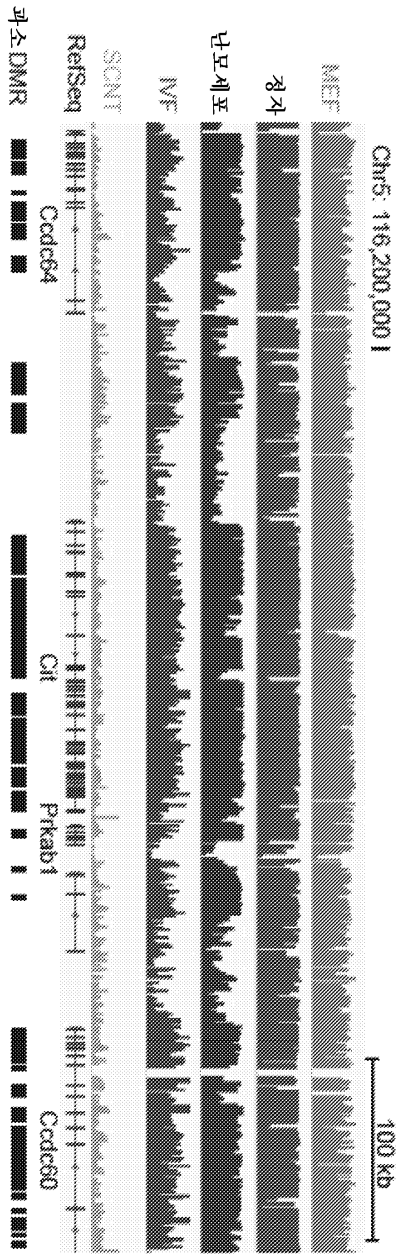
도면5h



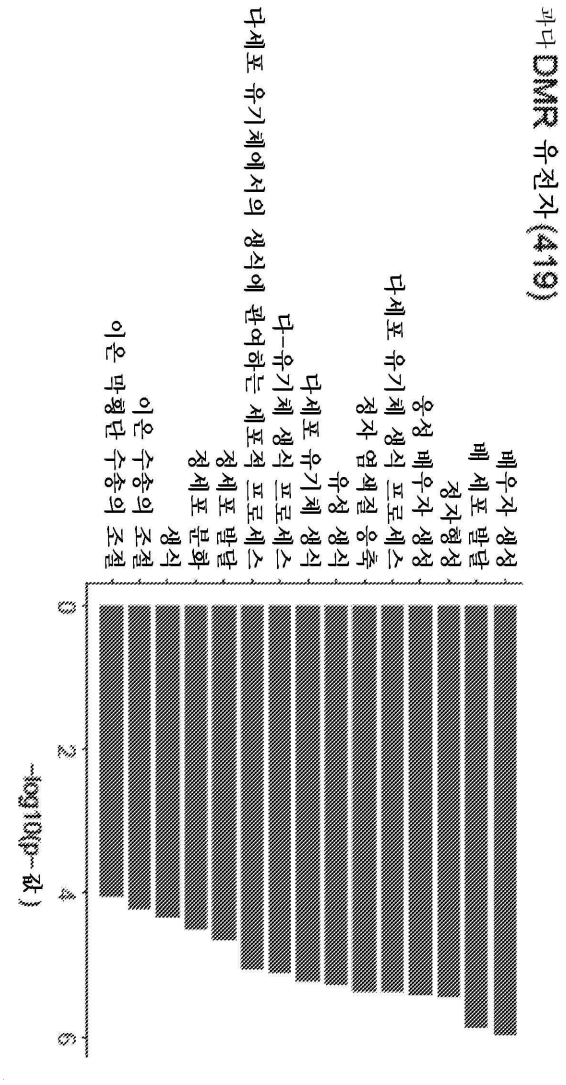
도면6a


Tkti2
과다 DMR
5 kb
MEF
정자
난모세포
IVF
SCNT
A730043L09Rik
과다 DMR
5 kb
Aagab
Smad3
과소 DMR
50 kb

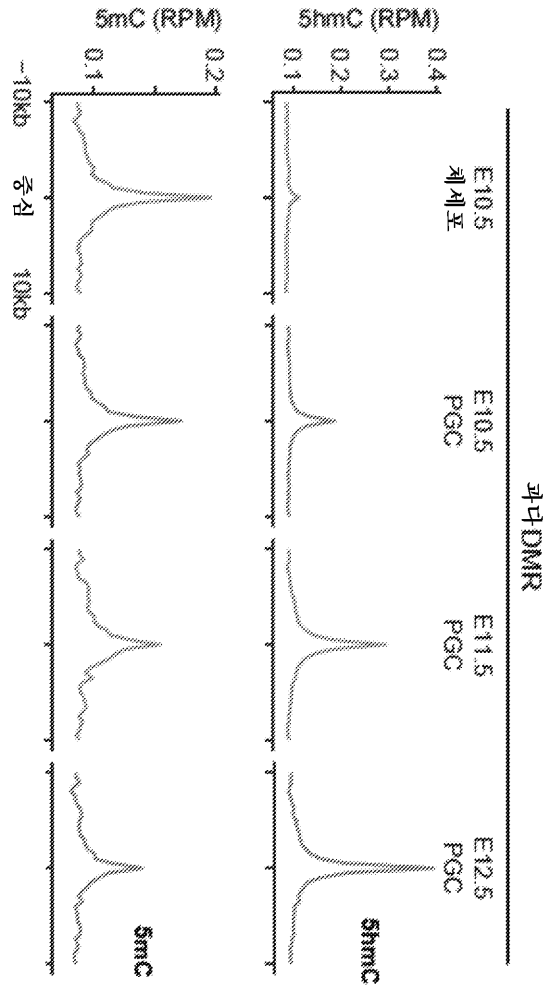
도면6b


Chr5: 116,200,000
100 kb
MEF
정자
난모세포
IVF
SCNT
RefSeq
파소DMR
Ccdc64
Cit
Prkab1
Ccdc60

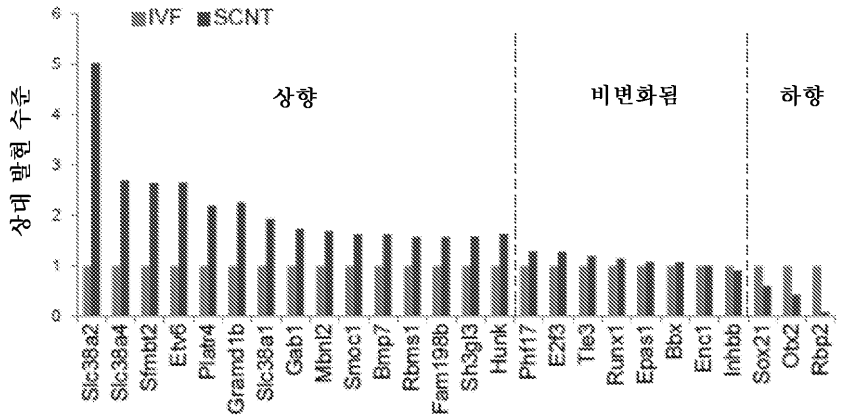
도면6c


과다 DMR 유전자(419)
배우자 생성
배 세포 발달
정자형성
다세포 유기체 생식 프로세스
웅성 배우자 생성
정자 염색질 응축
유성 생식
다세포 유기체 생식
다-유기체 생식 프로세스
다세포 유기체에서의 생식에 관여하는 세포적 프로세스
정세포 발달
정세포 분화
생식
이은 수송의 조절
이은 막횡단 수송의 조절
0
2
4
6
-log10p-값)

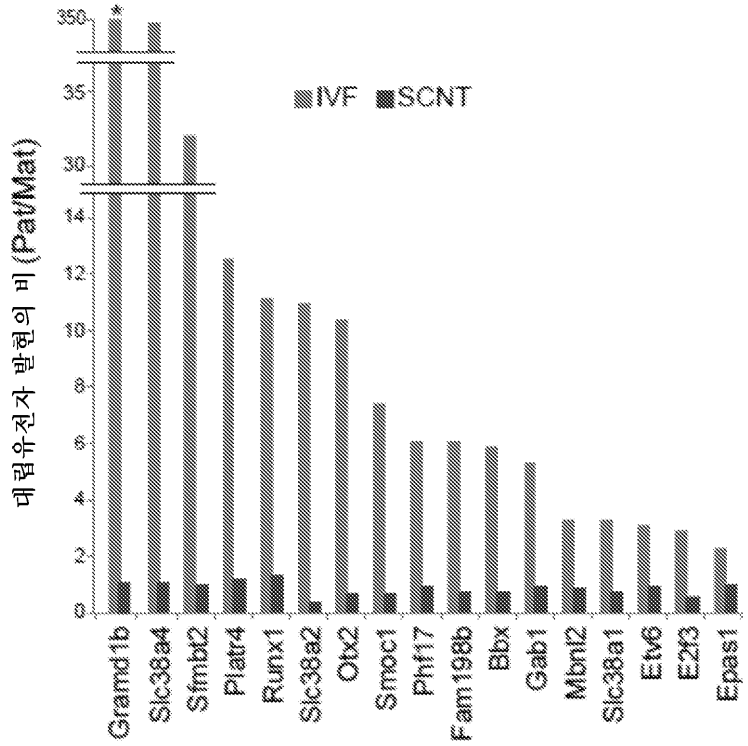
도면6d


5mC (RPM)
5hmC (RPM)
0.1
0.2
0.3
0.4
-10Kb
중심
10Kb
E10.5 체세포
E10.5 PGC
E11.5 PGC
E12.5 PGC
파닥DMR
5mC
5hmC

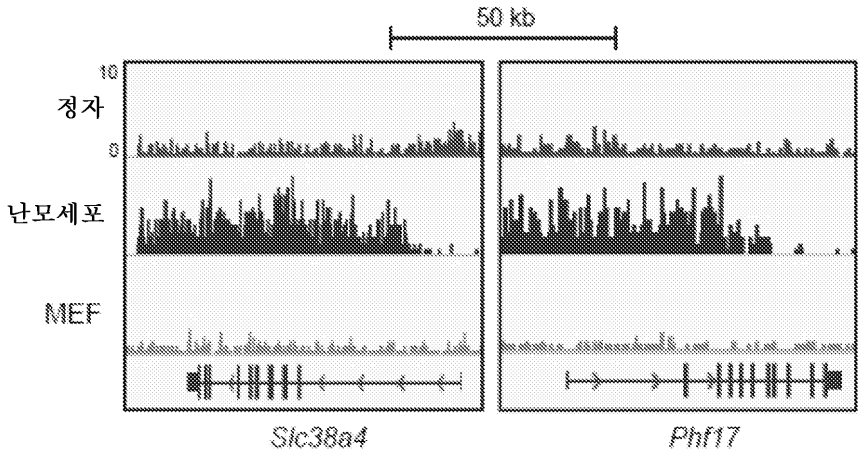
도면7a


IVF
SCNT
상대 발현 수준
0
1
2
3
4
5
6
상향
비변화됨
하향
Slc38a2
Slc38a4
Sfmbt2
Etv6
Ptar4
Gramd1b
Slc38a1
Gab1
Mbnl2
Smoc1
Bmp7
Rbms1
Fam198b
Sh3gl3
Hunk
Phf17
E2f3
Tle3
Runx1
Epas1
Bbx
Enc1
Inhbb
Sox21
Otx2
Rbp2

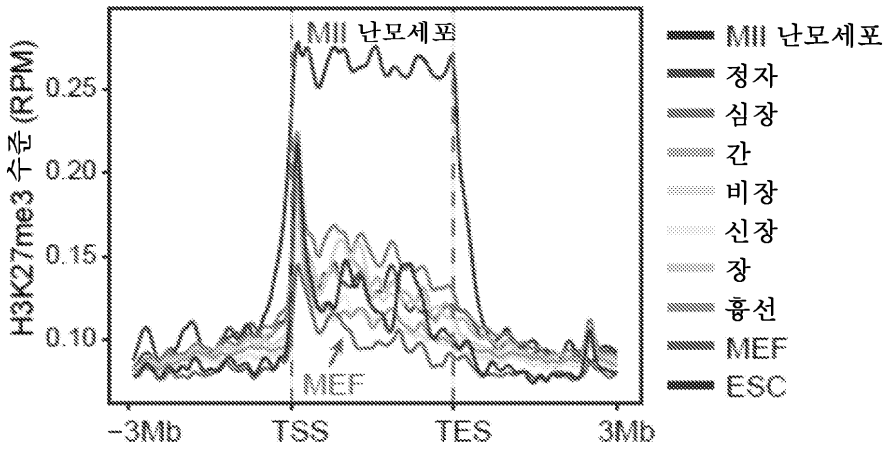
도면7b


IVF
SCNT
대립유전자 발현의 비 (Pat/Mat)
Gramd1b
Slc38a4
Sfmbt2
Platr4
Runx1
Slc38a2
Otx2
Smoc1
Phf17
Fam198b
Bbx
Gab1
Mbnl2
Slc38a1
Etv6
E2f3
Epas1

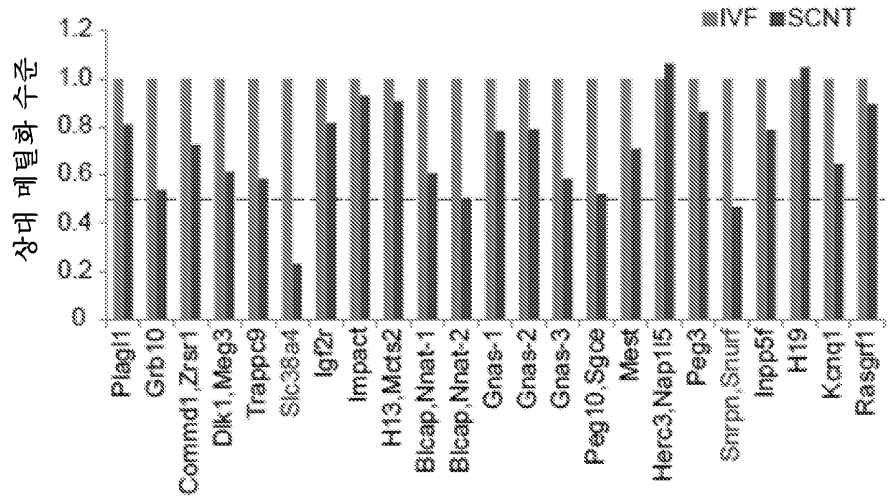
도면7c


50 kb
정자
난모세포
MEF
Slc38a4
Phf17

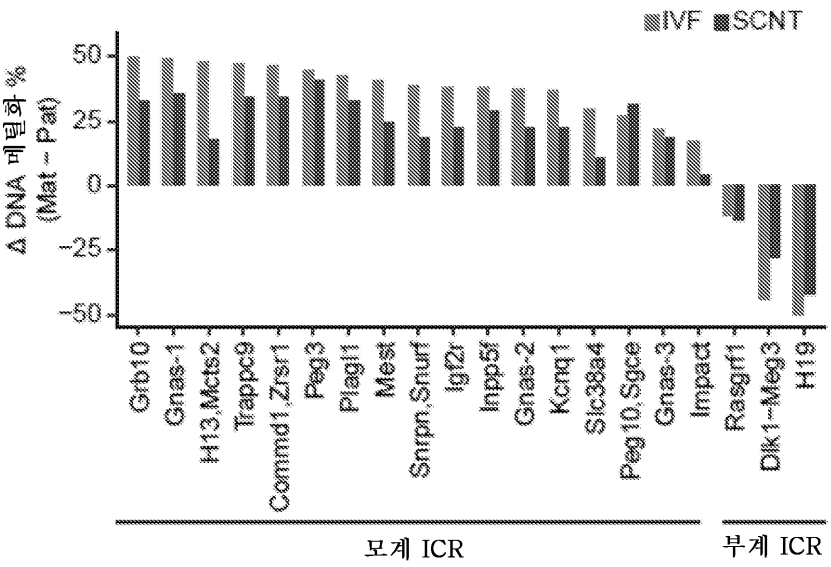
도면7d


MII 난모세포
H3K27me3 수준 (RPM)
0.25
0.20
0.15
0.10
MEF
-3Mb
TSS
TES
3Mb
MII 난모세포
정자
심장
간
비장
신장
장
흉선
MEF
ESC

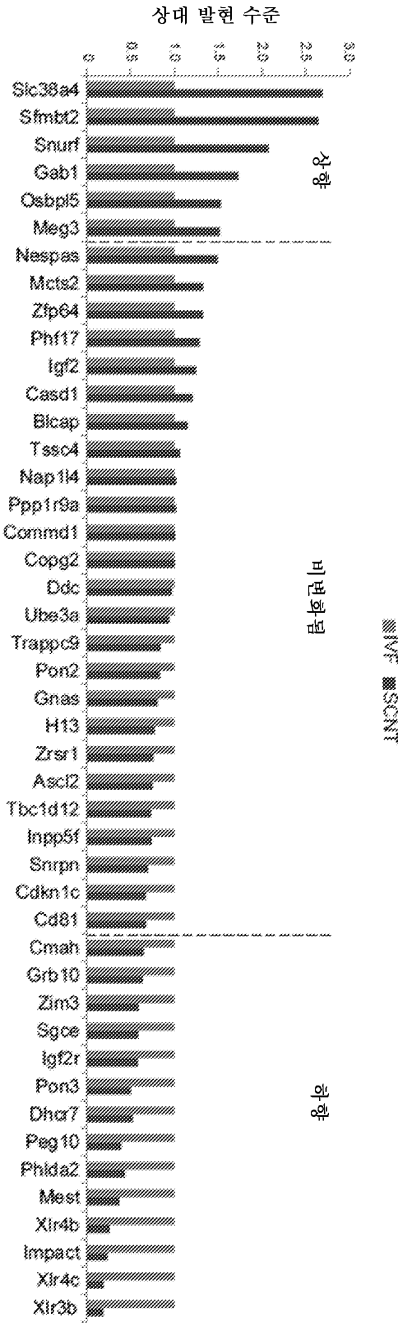
도면8a


IVF SCNT
상대 메틸화 수준
1.2
1.0
0.8
0.6
0.4
0.2
0
Plagl1
Grb10
Commd1,Zrsr1
Dlk1,Meg3
Trappc9
Slc38a4
Igf2r
Impact
H13,Mcts2
Blcap,Nnat-1
Blcap,Nnat-2
Gnas-1
Gnas-2
Gnas-3
Peg10,Sgce
Mest
Herc3,Nap1l5
Peg3
Snrpn,Snurf
Inpp5f
H19
Kcnq1
Rasgrf1

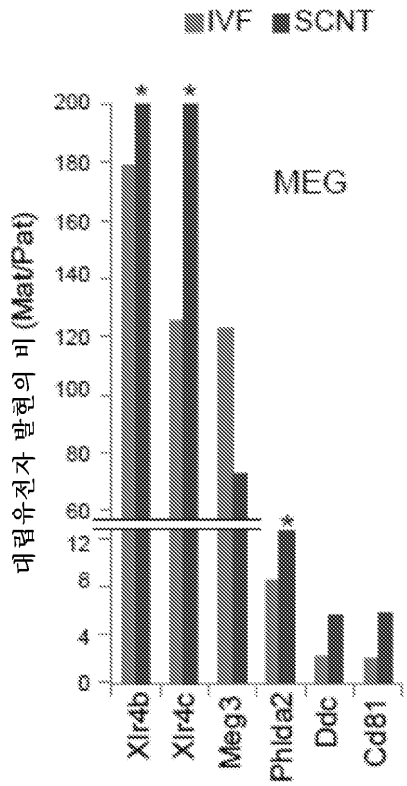
도면8b


IVF
SCNT
Δ DNA 메틸화 %
(Mat – Pat)
50
25
0
-25
-50
Grb10
Gnas-1
H13,Mcts2
Trappc9
Commd1,Zrsr1
Peg3
Plagl1
Mest
Snrpn,Snurf
Igf2r
Inpp5f
Gnas-2
Kcnq1
Slc38a4
Peg10,Sgce
Gnas-3
Impact
Rasgrf1
Dlk1-Meg3
H19
모계 ICR
부계 ICR

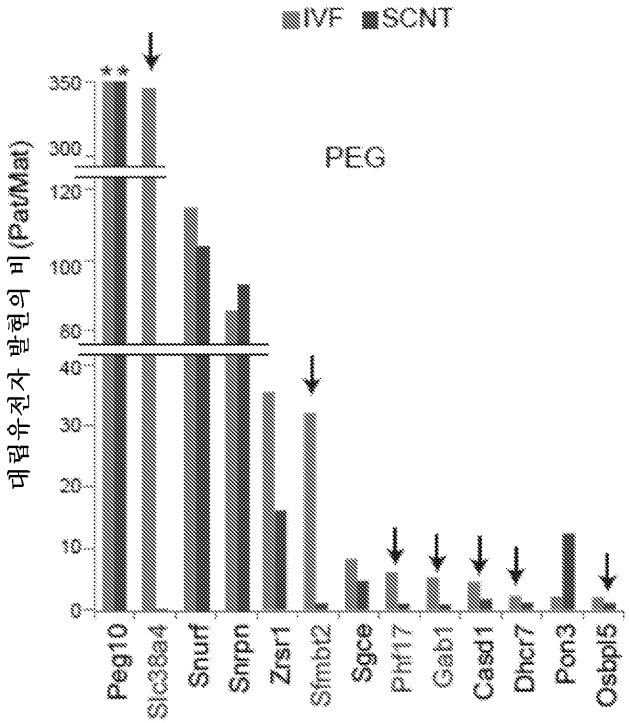
도면8c


상대 발현 수준
0
5
10
15
20
25
30
Slc38a4
Sfmbt2
Snurf
Gab1
Osbpl5
Meg3
Nespas
Mcts2
Zfp64
Phf17
Igf2
Casd1
Blcap
Tssc4
Nap1l4
Ppp1r9a
Commd1
Copg2
Ddc
Ube3a
Trappc9
Pon2
Gnas
H13
Zrsr1
Ascl2
Tbc1d12
Inpp5f
Snrpn
Cdkn1c
Cd81
Cmah
Grb10
Zim3
Sgce
Igf2r
Pon3
Dhcr7
Peg10
Phlda2
Mest
Xlr4b
Impact
Xlr4c
Xlr3b
상향
비변화
하향
iVF
SCNT

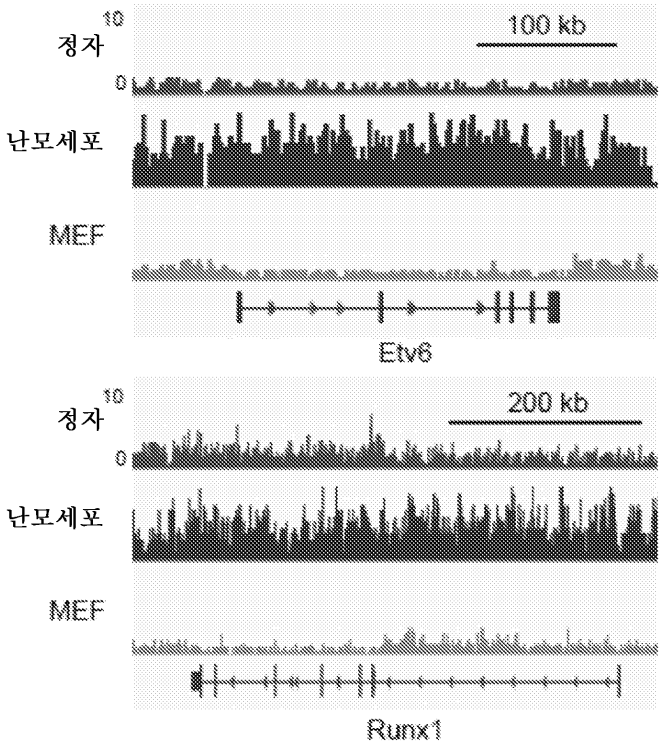
도면8d


IVF
SCNT
MEG
대립유전자 발현의 비 (Mat/Pat)
200
160
140
120
100
80
60
12
8
4
0
*
*
*
Xlr4b
Xlr4c
Meg3
Phlda2
Ddc
Cd81

도면8e


IVF
SCNT
PEG
대립유전자 발현의 비 (Pat/Mat)
350
300
120
100
80
40
30
20
10
0
**
Peg10
Slc38a4
Snurf
Snrpn
Zrsr1
Sfmbt2
Sgce
Peg13
Gab1
Casd1
Dhcr7
Pon3
Osbpl5

도면8f



서 열 목 록

SEQUENCE LISTING

<110> CHILDREN'S MEDICAL CENTER CORPORATION

<120> COMPOSITIONS AND METHODS FOR SOMATIC CELL REPROGRAMMING AND MODULATING IMPRINTING

<130> 167705.015801/PCT

<140> PCT/US2019/026074

<141> 2019-04-05

<150> 62/654,199

<151> 2018-04-06

<160> 19

<170> PatentIn version 3.5

<210> 1

<211> 523

<212> PRT

<213> Homo sapiens

<400> 1

Met Glu Thr Met Lys Ser Lys Ala Asn Cys Ala Gln Asn Pro Asn Cys
1 5 10 15

Asn Ile Met Ile Phe His Pro Thr Lys Glu Glu Phe Asn Asp Phe Asp
20 25 30

Lys Tyr Ile Ala Tyr Met Glu Ser Gln Gly Ala His Arg Ala Gly Leu
35 40 45

Ala Lys Ile Ile Pro Pro Lys Glu Trp Lys Ala Arg Glu Thr Tyr Asp
50 55 60

Asn Ile Ser Glu Ile Leu Ile Ala Thr Pro Leu Gln Gln Val Ala Ser
65 70 75 80

Gly Arg Ala Gly Val Phe Thr Gln Tyr His Lys Lys Lys Lys Ala Met
85 90 95

Thr Val Gly Glu Tyr Arg His Leu Ala Asn Ser Lys Lys Tyr Gln Thr
100 105 110

Pro Pro His Gln Asn Phe Glu Asp Leu Glu Arg Lys Tyr Trp Lys Asn
115 120 125

Arg Ile Tyr Asn Ser Pro Ile Tyr Gly Ala Asp Ile Ser Gly Ser Leu
130 135 140

Phe Asp Glu Asn Thr Lys Gln Trp Asn Leu Gly His Leu Gly Thr Ile
145 150 155 160

Gln Asp Leu Leu Glu Lys Glu Cys Gly Val Val Ile Glu Gly Val Asn
165 170 175

Thr Pro Tyr Leu Tyr Phe Gly Met Trp Lys Thr Thr Phe Ala Trp His
180 185 190

Thr Glu Asp Met Asp Leu Tyr Ser Ile Asn Tyr Leu His Leu Gly Glu
195 200 205

Pro Lys Thr Trp Tyr Val Val Pro Pro Glu His Gly Gln Arg Leu Glu
210 215 220

Arg Leu Ala Arg Glu Leu Phe Pro Gly Ser Ser Arg Gly Cys Gly Ala
225 230 235 240

Phe Leu Arg His Lys Val Ala Leu Ile Ser Pro Thr Val Leu Lys Glu

245 250 255

Asn Gly Ile Pro Phe Asn Arg Ile Thr Gln Glu Ala Gly Glu Phe Met

260 265 270

Val Thr Phe Pro Tyr Gly Tyr His Ala Gly Phe Asn His Gly Phe Asn

275 280 285

Cys Ala Glu Ala Ile Asn Phe Ala Thr Pro Arg Trp Ile Asp Tyr Gly

290 295 300

Lys Met Ala Ser Gln Cys Ser Cys Gly Glu Ala Arg Val Thr Phe Ser

305 310 315 320

Met Asp Ala Phe Val Arg Ile Leu Gln Pro Glu Arg Tyr Asp Leu Trp

325 330 335

Lys Arg Gly Gln Asp Arg Ala Val Val Asp His Met Glu Pro Arg Val

340 345 350

Pro Ala Ser Gln Glu Leu Ser Thr Gln Lys Glu Val Gln Leu Pro Arg

355 360 365

Arg Ala Ala Leu Gly Leu Arg Gln Leu Pro Ser His Trp Ala Arg His

370 375 380

Ser Pro Trp Pro Met Ala Ala Arg Ser Gly Thr Arg Cys His Thr Leu

385 390 395 400

Val Cys Ser Ser Leu Pro Arg Arg Ser Ala Val Ser Gly Thr Ala Thr

405 410 415

Gln Pro Arg Ala Ala Ala Val His Ser Ser Lys Lys Pro Ser Ser Thr

420 425 430

Pro Ser Ser Thr Pro Gly Pro Ser Ala Gln Ile Ile His Pro Ser Asn

435 440 445

Gly Arg Arg Gly Arg Gly Arg Pro Pro Gln Lys Leu Arg Ala Gln Glu

450 455 460

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

465 470 475 480

Cys Thr Ala Ser Gly Pro Glu Pro Glu Pro Leu Pro Glu Asp Gly Ala

485 490 495

Leu Met Asp Lys Pro Val Pro Leu Ser Pro Gly Leu Gln His Pro Val

500 505 510

Lys Ala Ser Gly Cys Ser Trp Ala Pro Val Pro

515 520

<210> 2

<211> 2988

<212> DNA

<213> Homo sapiens

<400> 2

aaggggcggg gccgaagcgg cccaggggggc gggcgtttga aatcagtgcc ttagagtaga 60

ccctaaacct cattttatac cttcaagaac caattactta atgtctcttc cgtcttttcc 120

gtccccgacc ccctcccaga ctccttcatt ccggtactgc gtggacggaa agccccgggt 180

agccgacacc acgtccccgg ctagcgggag agagcgtgga aaaggattac accaaactgt 240

ttaaatccaa cgactcctgc ttccatcctt tctcctgagc tagaaccaac aaacctagag 300

agttgggctt cggaaaaact agtgttttca tttaattgga tatgaagaaa gaacaaatat 360

gtacggggca accacgatct ttacaaagaa cataagttcc aggaaagcag gaaccttgtc 420

tctcttgttc actgggtgta tcctctgcat atagaacagt gcctggcaca taataggtgc 480

tgaattttgt tctaaacact gaggacattc tctgctacat ttgggtcgta cccccaggtc 540

tgagtaattc aatagactta agaagacaga gcccagcagc aaccgaaaca taacagagtt 600

gcaggatcag ctaacgtcaa tgcctgggca aagctgctgc ccagagtgga atctcactag 660

tgaataaaca agcccaagaa agattatcat ctcatttgca aaaaaaaaag tacgctggta 720

gatcctgcta cctcatagat aacaccagtc aaatttttt ttaaagtagc attttcctac 780

attgtcaact atctagaaca tacctaaaaa ctaagagttt actgcttatt aaatggaaac 840

tatgaagtct aaggccaact gtgcccagaa tccaaattgt aacataatga tatttcatcc 900

aaccaaagaa gagtttaatg attttgataa atatattgct tacatggaat cccaaggtgc 960

acacagagct ggcttggcta agataattcc acccaaagaa tggaaagcca gagagaccta 1020

tgataatatc agtgaaatct taatagccac tcccctccag caggtggcct ctgggcgggc 1080

aggggtgttt actcaatacc ataaaaaaaa gaaagccatg actgtggggg agtatcgcca 1140

tttggcaaac agtaaaaaat atcagactcc accacaccag aatttcgaag atttggagcg 1200

aaaatactgg aagaaccgca tctataattc accgatttat ggtgctgaca tcagtggctc 1260

cttgtttgat gaaaacacta aacaatggaa tcttgggcac ctgggaacaa ttcaggacct 1320

gctggaaaag gaatgtgggg ttgtcataga aggcgtcaat acaccctact tgtactttgg 1380

catgtggaaa accacgtttg cttggcatac agaggacatg gacctttaca gcatcaacta 1440

cctgcacctt ggggagccca aaacttggta tgtggtgccc ccagaacatg gccagcgcct 1500

ggaacgcctg gccagggagc tcttcccagg cagttcccgg ggttgtgggg ccttcctgcg 1560

gcacaaggtg gccctcatct cgcctacagt tctcaaggaa aatgggattc ccttcaatcg 1620

cataactcag gaggctggag agttcatggt gacctttccc tatggctacc atgctggctt 1680

caaccatggt ttcaactgcg cagaggccat caattttgcc actccgcgat ggattgatta 1740

tggcaaaatg gcctcccagt gtagctgtgg ggaggcaagg gtgacctttt ccatggatgc 1800

cttcgtgcgc atcctgcaac ctgaacgcta tgacctgtgg aaacgtgggc aagaccgggc 1860

agttgtggac cacatggagc ccagggtacc agccagccaa gagctgagca cccagaagga 1920

agtccagtta cccaggagag cagcgctggg cctgagacaa ctcccttccc actgggcccg 1980

gcattcccct tggcctatgg ctgcccgcag tgggacacgg tgccacaccc ttgtgtgctc 2040

ttcactccca cgccgatctg cagttagtgg cactgctacg cagccccggg ctgctgctgt 2100

ccacagctct aagaagccca gctcaactcc atcatccacc cctggtccat ctgcacagat 2160

tatccacccg tcaaatggca gacgtggtcg tggtcgccct cctcagaaac tgagagctca 2220

ggagctgacc ctccagactc cagccaagag gcccctcttg gcgggcacaa catgcacagc 2280

ttcgggccca gaacctgagc ccctacctga ggatggggct ttgatggaca agcctgtacc 2340

actgagccca gggctccagc atcctgtcaa ggcttctggg tgcagctggg cccctgtgcc 2400

ctaagtccac gggctgtctt tatatcccac tgccctgctg tgtgacagtt tgatgaaact 2460

ggttacattt acatcccaaa actttggttg agtttgcagg actctaggca tgcatgaaag 2520

agccccccctg gtgatgccct tggatgctgc caagtccatg gtagttttca attttgccat 2580

acttttgttc ttcctaccgg accctggaat gtctttggat attgctaaaa tctatttctg 2640

cagctgaggt tttatccact ggacacattt gtgtgtgaga actaggtctt gttgaggtta 2700

gcgtaacctg gtatatgcaa ctaccatcct ctgggccaac tgtggaagct gctgcacttg 2760

tgaagaatcc tgagctttga ttcctcttca gtctacgcat ttctctcttc ccctccctca 2820

ccccctttt cttataaaac taggttcttt atacagataa ggtcagtaga gttccagaat 2880

aaaagatatg acttttctga gttatttatg tacttaaaat atgttgtcac agtatttgtt 2940

cccaaatata ttaaaggtaa ccaaaatgtt aaaaaaaaaa aaaaaaaa 2988

<210> 3

<211> 747

<212> PRT

<213> Homo sapiens

<400> 3

```
Met Glu Ile Pro Asn Pro Pro Thr Ser Lys Cys Ile Thr Tyr Trp Lys
1               5                  10                  15
Arg Lys Val Lys Ser Glu Tyr Met Arg Leu Arg Gln Leu Lys Arg Leu
            20                  25                  30
Gln Ala Asn Met Gly Ala Lys Ala Leu Tyr Val Ala Asn Phe Ala Lys
        35                  40                  45

Val Gln Glu Lys Thr Gln Ile Leu Asn Glu Glu Trp Lys Lys Leu Arg
    50                  55                  60
Val Gln Pro Val Gln Ser Met Lys Pro Val Ser Gly His Pro Phe Leu
65                  70                  75                  80
Lys Lys Cys Thr Ile Glu Ser Ile Phe Pro Gly Phe Ala Ser Gln His
                85                  90                  95
Met Leu Met Arg Ser Leu Asn Thr Val Ala Leu Val Pro Ile Met Tyr
            100                 105                 110

Ser Trp Ser Pro Leu Gln Gln Asn Phe Met Val Glu Asp Glu Thr Val
        115                 120                 125
Leu Cys Asn Ile Pro Tyr Met Gly Asp Glu Val Lys Glu Glu Asp Glu
    130                 135                 140
Thr Phe Ile Glu Glu Leu Ile Asn Asn Tyr Asp Gly Lys Val His Gly
145                 150                 155                 160
Glu Glu Glu Met Ile Pro Gly Ser Val Leu Ile Ser Asp Ala Val Phe
                165                 170                 175

Leu Glu Leu Val Asp Ala Leu Asn Gln Tyr Ser Asp Glu Glu Glu Glu
            180                 185                 190
Gly His Asn Asp Thr Ser Asp Gly Lys Gln Asp Asp Ser Lys Glu Asp
        195                 200                 205
Leu Pro Val Thr Arg Lys Arg Lys Arg His Ala Ile Glu Gly Asn Lys
    210                 215                 220
Lys Ser Ser Lys Lys Gln Phe Pro Asn Asp Met Ile Phe Ser Ala Ile
```

225 230 235 240

Ala Ser Met Phe Pro Glu Asn Gly Val Pro Asp Asp Met Lys Glu Arg
245 250 255
Tyr Arg Glu Leu Thr Glu Met Ser Asp Pro Asn Ala Leu Pro Pro Gln
260 265 270
Cys Thr Pro Asn Ile Asp Gly Pro Asn Ala Lys Ser Val Gln Arg Glu
275 280 285
Gln Ser Leu His Ser Phe His Thr Leu Phe Cys Arg Arg Cys Phe Lys
290 295 300

Tyr Asp Cys Phe Leu His Pro Phe His Ala Thr Pro Asn Val Tyr Lys
305 310 315 320
Arg Lys Asn Lys Glu Ile Lys Ile Glu Pro Glu Pro Cys Gly Thr Asp
325 330 335
Cys Phe Leu Leu Leu Glu Gly Ala Lys Glu Tyr Ala Met Leu His Asn
340 345 350
Pro Arg Ser Lys Cys Ser Gly Arg Arg Arg Arg Arg His His Ile Val
355 360 365

Ser Ala Ser Cys Ser Asn Ala Ser Ala Ser Ala Val Ala Glu Thr Lys
370 375 380
Glu Gly Asp Ser Asp Arg Asp Thr Gly Asn Asp Trp Ala Ser Ser Ser
385 390 395 400
Ser Glu Ala Asn Ser Arg Cys Gln Thr Pro Thr Lys Gln Lys Ala Ser
405 410 415
Pro Ala Pro Pro Gln Leu Cys Val Val Glu Ala Pro Ser Glu Pro Val
420 425 430

Glu Trp Thr Gly Ala Glu Glu Ser Leu Phe Arg Val Phe His Gly Thr
435 440 445
Tyr Phe Asn Asn Phe Cys Ser Ile Ala Arg Leu Leu Gly Thr Lys Thr
450 455 460
Cys Lys Gln Val Phe Gln Phe Ala Val Lys Glu Ser Leu Ile Leu Lys
465 470 475 480

Leu Pro Thr Asp Glu Leu Met Asn Pro Ser Gln Lys Lys Lys Arg Lys
485 490 495

His Arg Leu Trp Ala Ala His Cys Arg Lys Ile Gln Leu Lys Lys Asp
500 505 510

Asn Ser Ser Thr Gln Val Tyr Asn Tyr Gln Pro Cys Asp His Pro Asp
515 520 525

Arg Pro Cys Asp Ser Thr Cys Pro Cys Ile Met Thr Gln Asn Phe Cys
530 535 540

Glu Lys Phe Cys Gln Cys Asn Pro Asp Cys Gln Asn Arg Phe Pro Gly
545 550 555 560

Cys Arg Cys Lys Thr Gln Cys Asn Thr Lys Gln Cys Pro Cys Tyr Leu
565 570 575

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

Glu His Trp Asp Cys Lys Val Val Ser Cys Lys Asn Cys Ser Ile Gln
595 600 605

Arg Gly Leu Lys Lys His Leu Leu Leu Ala Pro Ser Asp Val Ala Gly
610 615 620

Trp Gly Thr Phe Ile Lys Glu Ser Val Gln Lys Asn Glu Phe Ile Ser
625 630 635 640

Glu Tyr Cys Gly Glu Leu Ile Ser Gln Asp Glu Ala Asp Arg Arg Gly
645 650 655

Lys Val Tyr Asp Lys Tyr Met Ser Ser Phe Leu Phe Asn Leu Asn Asn
660 665 670

Asp Phe Val Val Asp Ala Thr Arg Lys Gly Asn Lys Ile Arg Phe Ala
675 680 685

Asn His Ser Val Asn Pro Asn Cys Tyr Ala Lys Val Val Met Val Asn
690 695 700

Gly Asp His Arg Ile Gly Ile Phe Ala Lys Arg Ala Ile Gln Ala Gly
705 710 715 720

Glu Glu Leu Phe Phe Asp Tyr Arg Tyr Ser Gln Ala Asp Ala Leu Lys

725 730 735

Tyr Val Gly Ile Glu Arg Glu Thr Asp Val Leu

740 745

<210> 4

<211> 4697

<212> DNA

<213> Homo sapiens

<400> 4

aggaggcgcg gggcggggca cggcgcaggg gtggggccgc ggcgcgcatg cgtcctagca 60

gcgggacccg cggctcggga tggaggctgg acacctgttc tgctgttgtg tcctgccatt 120

ctcctgaaga acagaggcac actgtaaaac ccaacacttc cccttgcatt ctataagatt 180

acagcaagat ggaaatacca aatcccccta cctccaaatg tatcacttac tggaaaagaa 240

aagtgaaatc tgaatacatg cgacttcgac aacttaaacg gcttcaggca aatatgggtg 300

caaaggcttt gtatgtggca aattttgcaa aggttcaaga aaaaacccag atcctcaatg 360

aagaatggaa gaagcttcgt gtccaacctg ttcagtcaat gaagcctgtg agtggacacc 420

cttttctcaa aaagtgtacc atagagagca ttttcccggg atttgcaagc caacatatgt 480

taatgaggtc actgaacaca gttgcattgg ttcccatcat gtattcctgg tcccctctcc 540

aacagaactt tatggtagaa gatgagacgg ttttgtgcaa tattccctac atgggagatg 600

aagtgaaaga agaagatgag acttttattg aggagctgat caataactat gatgggaaag 660

tccatggtga agaagagatg atccctggat ccgttctgat tagtgatgct gtttttctgg 720

agttggtcga tgccctgaat cagtactcag atgaggagga ggaagggcac aatgacacct 780

cagatggaaa gcaggatgac agcaaagaag atctgccagt aacaagaaag agaaagcgac 840

atgctattga aggcaacaaa aagagttcca agaaacagtt cccaaatgac atgatcttca 900

gtgcaattgc ctcaatgttc cctgagaatg gtgtcccaga tgacatgaag gagaggtatc 960

gagaactaac agagatgtca gaccccaatg cacttccccc tcagtgcaca cccaacatcg 1020

atggccccaa tgccaagtct gtgcagcggg agcaatctct gcactccttc cacacacttt 1080

tttgccggcg ctgctttaaa tacgactgct tccttcaccc ttttcatgcc acccctaatg 1140

tatataaacg caagaataaa gaaatcaaga ttgaaccaga accatgtggc acagactgct 1200

tccttttgct ggaaggagca aaggagtatg ccatgctcca caacccccgc tccaagtgct 1260

ctggtcgtcg ccggagaagg caccacatag tcagtgcttc ctgctccaat gcctcagcct 1320

ctgctgtggc tgagactaaa gaaggagaca gtgacaggga cacaggcaat gactgggcct 1380

ccagttcttc agaggctaac tctcgctgtc agactcccac aaaacagaag gctagtccag 1440

ccccacctca actctgcgta gtggaagcac cctcggagcc tgtggaatgg actggggctg 1500

aagaatctct ttttcgagtc ttccatggca cctacttcaa caacttctgt tcaatagcca 1560

ggcttctggg gaccaagacg tgcaagcagg tctttcagtt tgcagtcaaa gaatcactta 1620

tcctgaagct gccaacagat gagctcatga acccctcaca gaagaagaaa agaaagcaca 1680

gattgtgggc tgcacactgc aggaagattc agctgaagaa agataactct tccacacaag 1740

tgtacaacta ccaaccctgc gaccacccag accgcccctg tgacagcacc tgcccctgca 1800

tcatgactca gaatttctgt gagaagttct gccagtgcaa cccagactgt cagaatcgtt 1860

tccctggctg tcgctgtaag acccagtgca ataccaagca atgtccttgc tatctggcag 1920

tgcgagaatg tgaccctgac ctgtgtctca cctgtggggc ctcagagcac tgggactgca 1980

aggtggtttc ctgtaaaaac tgcagcatcc agcgtggact taagaagcac ctgctgctgg 2040

ccccctctga tgtggccgga tggggcacct tcataaagga gtctgtgcag aagaacgaat 2100

tcatttctga atactgtggt gagctcatct ctcaggatga ggctgatcga cgcggaaagg 2160

tctatgacaa atacatgtcc agcttcctct tcaacctcaa taatgatttt gtagtggatg 2220

ctactcggaa aggaaacaaa attcgatttg caaatcattc agtgaatccc aactgttatg 2280

ccaaagtggt catggtgaat ggagaccatc ggattgggat ctttgccaag agggcaattc 2340

aagctggcga agagctcttc tttgattaca ggtacagcca agctgatgct ctcaagtacg 2400

tggggatcga gagggagacc gacgtccttt agccctccca ggccccacgg cagcacttat 2460

ggtagcggca ctgtcttggc tttcgtgctc acaccactgc tgctcgagtc tcctgcactg 2520

tgtctcccac actgagaaac cccccaaccc actccctctg tagtgaggcc tctgccatgt 2580

ccagagggca caaaactgtc tcaatgagag gggagacaga ggcagctagg gcttggtctc 2640

ccaggacaga gagttacaga aatgggagac tgtttctctg gcctcagaag aagcgagcac 2700

aggctggggt ggatgactta tgcgtgattt cgtgtcggct ccccaggctg tggcctcagg 2760

aatcaactta ggcagttccc aacaagcgct agcctgtaat tgtagctttc cacatcaaga 2820

gtccttatgt tattgggatg caggcaaacc tctgtggtcc taagacctgg agaggacagg 2880

ctaagtgaag tgtggtccct ggagcctaca agtggtctgg gttagaggcg agcctggcag 2940

gcagcacaga ctgaactcag aggtagacag gtcaccttac tacctcctcc ctcgtggcag 3000

ggctcaaact gaaagagtgt gggttctaag tacaggcatt caaggctggg ggaaggaaag 3060

ctacgccatc cttccttagc cagagaggga gaaccagcca gatgatagta gttaaactgc 3120

taagcttggg cccaggaggc tttgagaaag ccttctctgt gtactctgga gatagatgga 3180

gaagtgtttt cagattcctg ggaacagaca ccagtgctcc agctcctcca aagttctggc 3240

ttagcagctg caggcaagca ttatgctgct attgaagaag cattaggggt atgcctggca 3300

ggtgtgagca tcctggctcg ctggatttgt gggtgttttc aggccttcca ttccccatag 3360

aggcaaggcc caatggccag tgttgcttat cgcttcaggg taggtgggca caggcttgga 3420

ctagagagga gaaagattgg tgtaatctgc tttcctgtct gtagtgcctg ctgtttggaa 3480

agggtgagtt agaatatgtt ccaaggttgg tgaggggcta aattgcacgc gtttaggctg 3540

gcaccccgtg tgcagggcac actggcagag ggtatctgaa gtgggagaag aagcaggtag 3600

accacctgtc ccaggctgtg gtgccaccct ctctggcatt catgcagagc aaagcacttt 3660

aaccatttct tttaaaaggt ctatagattg gggtagagtt tggcctaagg tctctagggt 3720

ccctgcctaa atcccactcc tgagggaggg ggaagaagag agggtgggag attctcctcc 3780

agtcctgtct catctcctgg gagaggcaga cgagtgagtt tcacacagaa gaatttcatg 3840

tgaatggggc cagcaagagc tgccctgtgt ccatggtggg tgtgccgggc tggctgggaa 3900

caaggagcag tatgttgagt agaaagggtg tgggcgggta tagattggcc tgggagtgtt 3960

acagtaggga gcaggcttct cccttctttc tgggactcag agccccgctt cttcccactc 4020

cacttgttgt cccatgaagg aagaagtggg gttcctcctg acccagctgc ctcttacggt 4080

ttggtatggg acatgcacac acactcacat gctctcactc accacactgg agggcacaca 4140

cgtaccccgc acccagcaac tcctgacaga aagctcctcc cacccaaatg ggccaggccc 4200

cagcatgatc ctgaaatctg catccgccgt ggtttgtatt cattgtgcat atcagggata 4260

ccctcaagct ggactgtggg ttccaaatta ctcatagagg agaaaaccag agaaagatga 4320

agaggaggag ttaggtctat ttgaaatgcc aggggctcgc tgtgaggaat aggtgaaaaa 4380

aaacttttca ccagcctttg agagactaga ctgaccccac ccttccttca gtgagcagaa 4440

tcactgtggt cagtctcctg tcccagcttc agttcatgaa tactcctgtt cctccagttt 4500

cccatccttt gtccctgctg tcccccactt ttaaagatgg gtctcaaccc ctccccacca 4560

cgtcatgatg gatggggcaa ggtggtgggg actaggggag cctggtatac atgcggcttc 4620

attgccaata aatttcatgc actttaaagt cctgtggctt gtgacctctt aataaagtgt 4680

tagaatccaa aaaaaaa 4697

<210> 5

<211> 746

<212> PRT

<213> Homo sapiens

<400> 5

Met Gly Gln Thr Gly Lys Lys Ser Glu Lys Gly Pro Val Cys Trp Arg

1 5 10 15

Lys Arg Val Lys Ser Glu Tyr Met Arg Leu Arg Gln Leu Lys Arg Phe

20 25 30

Arg Arg Ala Asp Glu Val Lys Ser Met Phe Ser Ser Asn Arg Gln Lys

35 40 45

Ile Leu Glu Arg Thr Glu Ile Leu Asn Gln Glu Trp Lys Gln Arg Arg

50 55 60

Ile Gln Pro Val His Ile Leu Thr Ser Val Ser Ser Leu Arg Gly Thr

65 70 75 80

Arg Glu Cys Ser Val Thr Ser Asp Leu Asp Phe Pro Thr Gln Val Ile

85 90 95

Pro Leu Lys Thr Leu Asn Ala Val Ala Ser Val Pro Ile Met Tyr Ser

100 105 110

Trp Ser Pro Leu Gln Gln Asn Phe Met Val Glu Asp Glu Thr Val Leu

115 120 125

His Asn Ile Pro Tyr Met Gly Asp Glu Val Leu Asp Gln Asp Gly Thr

130 135 140

Phe Ile Glu Glu Leu Ile Lys Asn Tyr Asp Gly Lys Val His Gly Asp

145 150 155 160

Arg Glu Cys Gly Phe Ile Asn Asp Glu Ile Phe Val Glu Leu Val Asn

165 170 175

Ala Leu Gly Gln Tyr Asn Asp Asp Asp Asp Asp Asp Asp Gly Asp Asp

180 185 190

Pro Glu Glu Arg Glu Glu Lys Gln Lys Asp Leu Glu Asp His Arg Asp

195 200 205

Asp Lys Glu Ser Arg Pro Pro Arg Lys Phe Pro Ser Asp Lys Ile Phe

210 215 220

Glu Ala Ile Ser Ser Met Phe Pro Asp Lys Gly Thr Ala Glu Glu Leu

225 230 235 240

Lys Glu Lys Tyr Lys Glu Leu Thr Glu Gln Gln Leu Pro Gly Ala Leu

245 250 255

Pro Pro Glu Cys Thr Pro Asn Ile Asp Gly Pro Asn Ala Lys Ser Val
260 265 270
Gln Arg Glu Gln Ser Leu His Ser Phe His Thr Leu Phe Cys Arg Arg
275 280 285
Cys Phe Lys Tyr Asp Cys Phe Leu His Pro Phe His Ala Thr Pro Asn

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

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

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

485 490 495
Arg Leu Trp Ala Ala His Cys Arg Lys Ile Gln Leu Lys Lys Asp Gly

500 505 510

Ser Ser Asn His Val Tyr Asn Tyr Gln Pro Cys Asp His Pro Arg Gln

515 520 525

Pro Cys Asp Ser Ser Cys Pro Cys Val Ile Ala Gln Asn Phe Cys Glu

530 535 540

Lys Phe Cys Gln Cys Ser Ser Glu Cys Gln Asn Arg Phe Pro Gly Cys

545 550 555 560

Arg Cys Lys Ala Gln Cys Asn Thr Lys Gln Cys Pro Cys Tyr Leu Ala

565 570 575

Val Arg Glu Cys Asp Pro Asp Leu Cys Leu Thr Cys Gly Ala Ala Asp

580 585 590

His Trp Asp Ser Lys Asn Val Ser Cys Lys Asn Cys Ser Ile Gln Arg

595 600 605

Gly Ser Lys Lys His Leu Leu Leu Ala Pro Ser Asp Val Ala Gly Trp

610 615 620

Gly Ile Phe Ile Lys Asp Pro Val Gln Lys Asn Glu Phe Ile Ser Glu

625 630 635 640

Tyr Cys Gly Glu Ile Ile Ser Gln Asp Glu Ala Asp Arg Arg Gly Lys

645 650 655

Val Tyr Asp Lys Tyr Met Cys Ser Phe Leu Phe Asn Leu Asn Asn Asp

660 665 670

Phe Val Val Asp Ala Thr Arg Lys Gly Asn Lys Ile Arg Phe Ala Asn

675 680 685

His Ser Val Asn Pro Asn Cys Tyr Ala Lys Val Met Met Val Asn Gly

690 695 700

Asp His Arg Ile Gly Ile Phe Ala Lys Arg Ala Ile Gln Thr Gly Glu

705 710 715 720

Glu Leu Phe Phe Asp Tyr Arg Tyr Ser Gln Ala Asp Ala Leu Lys Tyr

725 730 735

Val Gly Ile Glu Arg Glu Met Glu Ile Pro

740 745

<210> 6

<211> 2681

<212> DNA

<213> Homo sapiens

<400> 6

ggcggcgctt gattgggctg gggggggccaa ataaaagcga tggcgattgg gctgccgcgt 60

ttggcgctcg gtccggtcgc gtccgacacc cggtgggact cagaaggcag tggagccccg 120

gcggcggcgg cggcggcgcg cgggggcgac gcgcgggaac aacgcgagtc ggcgcgcggg 180

acgaagaata atcatgggcc agactgggaa gaaatctgag aagggaccag tttgttggcg 240

gaagcgtgta aaatcagagt acatgcgact gagacagctc aagaggttca gacgagctga 300

tgaagtaaag agtatgttta gttccaatcg tcagaaaatt ttggaaagaa cggaaatctt 360

aaaccaagaa tggaaacagc gaaggataca gcctgtgcac atcctgactt cttgttcggt 420

gaccagtgac ttggattttc caacacaagt catcccatta aagactctga atgcagttgc 480

ttcagtaccc ataatgtatt cttggtctcc cctacagcag aattttatgg tggaagatga 540

aactgtttta cataacattc cttatatggg agatgaagtt ttagatcagg atggtacttt 600

cattgaagaa ctaataaaaa attatgatgg gaaagtacac ggggatagag aatgtgggtt 660

tataaatgat gaaatttttg tggagttggt gaatgccctt ggtcaatata atgatgatga 720

cgatgatgat gatggagacg atcctgaaga aagagaagaa aagcagaaag atctggagga 780

tcaccgagat gataaagaaa gccgcccacc tcggaaattt ccttctgata aaatttttga 840

agccatttcc tcaatgtttc cagataaggg cacagcagaa gaactaaagg aaaaatataa 900

agaactcacc gaacagcagc tcccaggcgc acttcctcct gaatgtaccc ccaacataga 960

tggaccaaat gctaaatctg ttcagagaga gcaaagctta cactcctttc atacgctttt 1020

ctgtaggcga tgttttaaat atgactgctt cctacatcct tttcatgcaa cacccaacac 1080

ttataagcgg aagaacacag aaacagctct agacaacaaa ccttgtggac cacagtgtta 1140

ccagcatttg gagggagcaa aggagtttgc tgctgctctc accgctgagc ggataaagac 1200

cccaccaaaa cgtccaggag gccgcagaag aggacggctt cccaataaca gtagcaggcc 1260

cagcaccccc accattaatg tgctggaatc aaaggataca gacagtgata gggaagcagg 1320

gactgaaacg gggggagaga acaatgataa agaagaagaa gagaagaaag atgaaacttc 1380

gagctcctct gaagcaaatt ctcggtgtca aacaccaata aagatgaagc caaatattga 1440

acctcctgag aatgtggagt ggagtggtgc tgaagcctca atgtttagag tcctcattgg 1500

cacttactat gacaatttct gtgccattgc taggttaatt gggaccaaaa catgtagaca 1560

ggtgtatgag tttagagtca aagaatctag catcatagct ccagctcccg ctgaggatgt 1620

ggatactcct ccaaggaaaa agaagaggaa acaccggttg tgggctgcac actgcagaaa 1680
gatacagctg aaaaaggacg gctcctctaa ccatgtttac aactatcaac cctgtgatca 1740
tccacggcag ccttgtgaca gttcgtgccc ttgtgtgata gcacaaaatt tttgtgaaaa 1800
gttttgtcaa tgtagttcag agtgtcaaaa ccgctttccg ggatgccgct gcaaagcaca 1860
gtgcaacacc aagcagtgcc cgtgctacct ggctgtccga gagtgtgacc ctgacctctg 1920
tcttacttgt ggagccgctg accattggga cagtaaaaat gtgtcctgca agaactgcag 1980
tattcagcgg ggctccaaaa agcatctatt gctggcacca tctgacgtgg caggctgggg 2040

gatttttatc aaagatcctg tgcagaaaaa tgaattcatc tcagaatact gtggagagat 2100
tatttctcaa gatgaagctg acagaagagg gaaagtgtat gataaataca tgtgcagctt 2160
tctgttcaac ttgaacaatg attttgtggt ggatgcaacc cgcaagggta acaaaattcg 2220
ttttgcaaat cattcggtaa atccaaactg ctatgcaaaa gttatgatgg ttaacggtga 2280
tcacaggata ggtatttttg ccaagagagc catccagact ggcgaagagc tgtttttttga 2340
ttacagatac agccaggctg atgccctgaa gtatgtcggc atcgaaagag aaatggaaat 2400
cccttgacat ctgctacctc ctccccccctc ctctgaaaca gctgccttag cttcaggaac 2460

ctcgagtact gtgggcaatt tagaaaaaga acatgcagtt tgaaattctg aatttgcaaa 2520
gtactgtaag aataatttat agtaatgagt ttaaaaatca acttttttatt gccttctcac 2580
cagctgcaaa gtgttttgta ccagtgaatt tttgcaataa tgcagtatgg tacatttttc 2640
aactttgaat aaagaatact tgaacttgtc cttgttgaat c 2681

<210> 7
<211> 1401
<212> PRT
<213> Homo sapiens
<400> 7

Met Lys Ser Cys Gly Val Ser Leu Ala Thr Ala Ala Ala Ala Ala Ala
1 5 10 15
Ala Phe Gly Asp Glu Glu Lys Lys Met Ala Ala Gly Lys Ala Ser Gly

20 25 30
Glu Ser Glu Glu Ala Ser Pro Ser Leu Thr Ala Glu Glu Arg Glu Ala
35 40 45
Leu Gly Gly Leu Asp Ser Arg Leu Phe Gly Phe Val Arg Phe His Glu

50 55 60

Asp Gly Ala Arg Thr Lys Ala Leu Leu Gly Lys Ala Val Arg Cys Tyr

65 70 75 80

Glu Ser Leu Ile Leu Lys Ala Glu Gly Lys Val Glu Ser Asp Phe Phe

85 90 95

Cys Gln Leu Gly His Phe Asn Leu Leu Leu Glu Asp Tyr Pro Lys Ala

100 105 110

Leu Ser Ala Tyr Gln Arg Tyr Tyr Ser Leu Gln Ser Asp Tyr Trp Lys

115 120 125

Asn Ala Ala Phe Leu Tyr Gly Leu Gly Leu Val Tyr Phe His Tyr Asn

130 135 140

Ala Phe Gln Trp Ala Ile Lys Ala Phe Gln Glu Val Leu Tyr Val Asp

145 150 155 160

Pro Ser Phe Cys Arg Ala Lys Glu Ile His Leu Arg Leu Gly Leu Met

165 170 175

Phe Lys Val Asn Thr Asp Tyr Glu Ser Ser Leu Lys His Phe Gln Leu

180 185 190

Ala Leu Val Asp Cys Asn Pro Cys Thr Leu Ser Asn Ala Glu Ile Gln

195 200 205

Phe His Ile Ala His Leu Tyr Glu Thr Gln Arg Lys Tyr His Ser Ala

210 215 220

Lys Glu Ala Tyr Glu Gln Leu Leu Gln Thr Glu Asn Leu Ser Ala Gln

225 230 235 240

Val Lys Ala Thr Val Leu Gln Gln Leu Gly Trp Met His His Thr Val

245 250 255

Asp Leu Leu Gly Asp Lys Ala Thr Lys Glu Ser Tyr Ala Ile Gln Tyr

260 265 270

Leu Gln Lys Ser Leu Glu Ala Asp Pro Asn Ser Gly Gln Ser Trp Tyr

275 280 285

Phe Leu Gly Arg Cys Tyr Ser Ser Ile Gly Lys Val Gln Asp Ala Phe

290 295 300

Ile Ser Tyr Arg Gln Ser Ile Asp Lys Ser Glu Ala Ser Ala Asp Thr
305 310 315 320
Trp Cys Ser Ile Gly Val Leu Tyr Gln Gln Gln Asn Gln Pro Met Asp
325 330 335
Ala Leu Gln Ala Tyr Ile Cys Ala Val Gln Leu Asp His Gly His Ala

340 345 350
Ala Ala Trp Met Asp Leu Gly Thr Leu Tyr Glu Ser Cys Asn Gln Pro
355 360 365
Gln Asp Ala Ile Lys Cys Tyr Leu Asn Ala Thr Arg Ser Lys Ser Cys
370 375 380
Ser Asn Thr Ser Ala Leu Ala Ala Arg Ile Lys Tyr Leu Gln Ala Gln
385 390 395 400
Leu Cys Asn Leu Pro Gln Gly Ser Leu Gln Asn Lys Thr Lys Leu Leu

405 410 415
Pro Ser Ile Glu Glu Ala Trp Ser Leu Pro Ile Pro Ala Glu Leu Thr
420 425 430
Ser Arg Gln Gly Ala Met Asn Thr Ala Gln Gln Asn Thr Ser Asp Asn
435 440 445
Trp Ser Gly Gly His Ala Val Ser His Pro Pro Val Gln Gln Gln Ala
450 455 460
His Ser Trp Cys Leu Thr Pro Gln Lys Leu Gln His Leu Glu Gln Leu

465 470 475 480
Arg Ala Asn Arg Asn Asn Leu Asn Pro Ala Gln Lys Leu Met Leu Glu
485 490 495
Gln Leu Glu Ser Gln Phe Val Leu Met Gln Gln His Gln Met Arg Pro
500 505 510
Thr Gly Val Ala Gln Val Arg Ser Thr Gly Ile Pro Asn Gly Pro Thr
515 520 525
Ala Asp Ser Ser Leu Pro Thr Asn Ser Val Ser Gly Gln Gln Pro Gln

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

545 550 555 560

Ala Cys Pro Gly Gln Pro Leu Ala Asn Gly Pro Phe Ser Ala Gly His

565 570 575

Val Pro Cys Ser Thr Ser Arg Thr Leu Gly Ser Thr Asp Thr Ile Leu

580 585 590

Ile Gly Asn Asn His Ile Thr Gly Ser Gly Ser Asn Gly Asn Val Pro

595 600 605

Tyr Leu Gln Arg Asn Ala Leu Thr Leu Pro His Asn Arg Thr Asn Leu

610 615 620

Thr Ser Ser Ala Glu Glu Pro Trp Lys Asn Gln Leu Ser Asn Ser Thr

625 630 635 640

Gln Gly Leu His Lys Gly Gln Ser Ser His Ser Ala Gly Pro Asn Gly

645 650 655

Glu Arg Pro Leu Ser Ser Thr Gly Pro Ser Gln His Leu Gln Ala Ala

660 665 670

Gly Ser Gly Ile Gln Asn Gln Asn Gly His Pro Thr Leu Pro Ser Asn

675 680 685

Ser Val Thr Gln Gly Ala Ala Leu Asn His Leu Ser Ser His Thr Ala

690 695 700

Thr Ser Gly Gly Gln Gln Gly Ile Thr Leu Thr Lys Glu Ser Lys Pro

705 710 715 720

Ser Gly Asn Ile Leu Thr Val Pro Glu Thr Ser Arg His Thr Gly Glu

725 730 735

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

740 745 750

Gln Met Thr Ala Asp Ala Val Cys Ser Pro Ser His Gly Asp Ser Lys

755 760 765

Ser Pro Gly Leu Leu Ser Ser Asp Asn Pro Gln Leu Ser Ala Leu Leu

770 775 780

Met Gly Lys Ala Asn Asn Asn Val Gly Thr Gly Thr Cys Asp Lys Val

785 790 795 800

Asn Asn Ile His Pro Ala Val His Thr Lys Thr Asp Asn Ser Val Ala
805 810 815
Ser Ser Pro Ser Ser Ala Ile Ser Thr Ala Thr Pro Ser Pro Lys Ser
820 825 830
Thr Glu Gln Thr Thr Thr Asn Ser Val Thr Ser Leu Asn Ser Pro His
835 840 845
Ser Gly Leu His Thr Ile Asn Gly Glu Gly Met Glu Glu Ser Gln Ser

850 855 860
Pro Met Lys Thr Asp Leu Leu Leu Val Asn His Lys Pro Ser Pro Gln
865 870 875 880
Ile Ile Pro Ser Met Ser Val Ser Ile Tyr Pro Ser Ser Ala Glu Val
885 890 895
Leu Lys Ala Cys Arg Asn Leu Gly Lys Asn Gly Leu Ser Asn Ser Ser
900 905 910
Ile Leu Leu Asp Lys Cys Pro Pro Pro Arg Pro Pro Ser Ser Pro Tyr

915 920 925
Pro Pro Leu Pro Lys Asp Lys Leu Asn Pro Pro Thr Pro Ser Ile Tyr
930 935 940
Leu Glu Asn Lys Arg Asp Ala Phe Phe Pro Pro Leu His Gln Phe Cys
945 950 955 960
Thr Asn Pro Asn Asn Pro Val Thr Val Ile Arg Gly Leu Ala Gly Ala
965 970 975
Leu Lys Leu Asp Leu Gly Leu Phe Ser Thr Lys Thr Leu Val Glu Ala

980 985 990
Asn Asn Glu His Met Val Glu Val Arg Thr Gln Leu Leu Gln Pro Ala
995 1000 1005
Asp Glu Asn Trp Asp Pro Thr Gly Thr Lys Lys Ile Trp His Cys
1010 1015 1020
Glu Ser Asn Arg Ser His Thr Thr Ile Ala Lys Tyr Ala Gln Tyr
1025 1030 1035
Gln Ala Ser Ser Phe Gln Glu Ser Leu Arg Glu Glu Asn Glu Lys

1040 1045 1050

Arg Ser His His Lys Asp His Ser Asp Ser Glu Ser Thr Ser Ser

1055 1060 1065

Asp Asn Ser Gly Arg Arg Arg Lys Gly Pro Phe Lys Thr Ile Lys

1070 1075 1080

Phe Gly Thr Asn Ile Asp Leu Ser Asp Asp Lys Lys Trp Lys Leu

1085 1090 1095

Gln Leu His Glu Leu Thr Lys Leu Pro Ala Phe Val Arg Val Val

1100 1105 1110

Ser Ala Gly Asn Leu Leu Ser His Val Gly His Thr Ile Leu Gly

1115 1120 1125

Met Asn Thr Val Gln Leu Tyr Met Lys Val Pro Gly Ser Arg Thr

1130 1135 1140

Pro Gly His Gln Glu Asn Asn Asn Phe Cys Ser Val Asn Ile Asn

1145 1150 1155

Ile Gly Pro Gly Asp Cys Glu Trp Phe Val Val Pro Glu Gly Tyr

1160 1165 1170

Trp Gly Val Leu Asn Asp Phe Cys Glu Lys Asn Asn Leu Asn Phe

1175 1180 1185

Leu Met Gly Ser Trp Trp Pro Asn Leu Glu Asp Leu Tyr Glu Ala

1190 1195 1200

Asn Val Pro Val Tyr Arg Phe Ile Gln Arg Pro Gly Asp Leu Val

1205 1210 1215

Trp Ile Asn Ala Gly Thr Val His Trp Val Gln Ala Ile Gly Trp

1220 1225 1230

Cys Asn Asn Ile Ala Trp Asn Val Gly Pro Leu Thr Ala Cys Gln

1235 1240 1245

Tyr Lys Leu Ala Val Glu Arg Tyr Glu Trp Asn Lys Leu Gln Ser

1250 1255 1260

Val Lys Ser Ile Val Pro Met Val His Leu Ser Trp Asn Met Ala

1265 1270 1275

Arg Asn Ile Lys Val Ser Asp Pro Lys Leu Phe Glu Met Ile Lys

1280 1285 1290

Tyr Cys Leu Leu Arg Thr Leu Lys Gln Cys Gln Thr Leu Arg Glu

1295 1300 1305

Ala Leu Ile Ala Ala Gly Lys Glu Ile Ile Trp His Gly Arg Thr

1310 1315 1320

Lys Glu Glu Pro Ala His Tyr Cys Ser Ile Cys Glu Val Glu Val

1325 1330 1335

Phe Asp Leu Leu Phe Val Thr Asn Glu Ser Asn Ser Arg Lys Thr

1340 1345 1350

Tyr Ile Val His Cys Gln Asp Cys Ala Arg Lys Thr Ser Gly Asn

1355 1360 1365

Leu Glu Asn Phe Val Val Leu Glu Gln Tyr Lys Met Glu Asp Leu

1370 1375 1380

Met Gln Val Tyr Asp Gln Phe Thr Leu Ala Pro Pro Leu Pro Ser

1385 1390 1395

Ala Ser Ser

1400

<210> 8

<211> 1643

<212> PRT

<213> Homo sapiens

<400> 8

Met His Arg Ala Val Asp Pro Pro Gly Ala Arg Ala Ala Arg Glu Ala

1 5 10 15

Phe Ala Leu Gly Gly Leu Ser Cys Ala Gly Ala Trp Ser Ser Cys Pro

20 25 30

Pro His Pro Pro Pro Arg Ser Ala Trp Leu Pro Gly Gly Arg Cys Ser

35 40 45

Ala Ser Ile Gly Gln Pro Pro Leu Pro Ala Pro Leu Pro Pro Ser His

50 55 60

Gly Ser Ser Ser Gly His Pro Ser Lys Pro Tyr Tyr Ala Pro Gly Ala
65 70 75 80
Pro Thr Pro Arg Pro Leu His Gly Lys Leu Glu Ser Leu His Gly Cys
85 90 95
Val Gln Ala Leu Leu Arg Glu Pro Ala Gln Pro Gly Leu Trp Glu Gln
100 105 110
Leu Gly Gln Leu Tyr Glu Ser Glu His Asp Ser Glu Glu Ala Thr Arg
115 120 125

Cys Tyr His Ser Ala Leu Arg Tyr Gly Gly Ser Phe Ala Glu Leu Gly
130 135 140
Pro Arg Ile Gly Arg Leu Gln Gln Ala Gln Leu Trp Asn Phe His Thr
145 150 155 160
Gly Ser Cys Gln His Arg Ala Lys Val Leu Pro Pro Leu Glu Gln Val
165 170 175
Trp Asn Leu Leu His Leu Glu His Lys Arg Asn Tyr Gly Ala Lys Arg
180 185 190

Gly Gly Pro Pro Val Lys Arg Ala Ala Glu Pro Pro Val Val Gln Pro
195 200 205
Val Pro Pro Ala Ala Leu Ser Gly Pro Ser Gly Glu Glu Gly Leu Ser
210 215 220
Pro Gly Gly Lys Arg Arg Arg Gly Cys Asn Ser Glu Gln Thr Gly Leu
225 230 235 240
Pro Pro Gly Leu Pro Leu Pro Pro Pro Pro Leu Pro Pro Pro Pro Pro
245 250 255

Pro Pro Pro Pro Pro Pro Pro Pro Leu Pro Gly Leu Ala Thr Ser Pro
260 265 270
Pro Phe Gln Leu Thr Lys Pro Gly Leu Trp Ser Thr Leu His Gly Asp
275 280 285
Ala Trp Gly Pro Glu Arg Lys Gly Ser Ala Pro Pro Glu Arg Gln Glu
290 295 300
Gln Arg His Ser Leu Pro His Pro Tyr Pro Tyr Pro Ala Pro Ala Tyr

305 310 315 320

Thr Ala His Pro Pro Gly His Arg Leu Val Pro Ala Ala Pro Pro Gly
325 330 335
Pro Gly Pro Arg Pro Pro Gly Ala Glu Ser His Gly Cys Leu Pro Ala
340 345 350
Thr Arg Pro Pro Gly Ser Asp Leu Arg Glu Ser Arg Val Gln Arg Ser
355 360 365
Arg Met Asp Ser Ser Val Ser Pro Ala Ala Thr Thr Ala Cys Val Pro
370 375 380

Tyr Ala Pro Ser Arg Pro Pro Gly Leu Pro Gly Thr Thr Thr Ser Ser
385 390 395 400
Ser Ser Ser Ser Ser Ser Asn Thr Gly Leu Arg Gly Val Glu Pro Asn
405 410 415
Pro Gly Ile Pro Gly Ala Asp His Tyr Gln Thr Pro Ala Leu Glu Val
420 425 430
Ser His His Gly Arg Leu Gly Pro Ser Ala His Ser Ser Arg Lys Pro
435 440 445

Phe Leu Gly Ala Pro Ala Ala Thr Pro His Leu Ser Leu Pro Pro Gly
450 455 460
Pro Ser Ser Pro Pro Pro Pro Pro Cys Pro Arg Leu Leu Arg Pro Pro
465 470 475 480
Pro Pro Pro Ala Trp Leu Lys Gly Pro Ala Cys Arg Ala Ala Arg Glu
485 490 495
Asp Gly Glu Ile Leu Glu Glu Leu Phe Phe Gly Thr Glu Gly Pro Pro
500 505 510

Arg Pro Ala Pro Pro Pro Leu Pro His Arg Glu Gly Phe Leu Gly Pro
515 520 525
Pro Ala Ser Arg Phe Ser Val Gly Thr Gln Asp Ser His Thr Pro Pro
530 535 540
Thr Pro Pro Thr Pro Thr Thr Ser Ser Ser Asn Ser Asn Ser Gly Ser
545 550 555 560

His Ser Ser Ser Pro Ala Gly Pro Val Ser Phe Pro Pro Pro Pro Tyr
565 570 575

Leu Ala Arg Ser Ile Asp Pro Leu Pro Arg Pro Pro Ser Pro Ala Gln
580 585 590
Asn Pro Gln Asp Pro Pro Leu Val Pro Leu Thr Leu Ala Leu Pro Pro
595 600 605
Ala Pro Pro Ser Ser Cys His Gln Asn Thr Ser Gly Ser Phe Arg Arg
610 615 620
Pro Glu Ser Pro Arg Pro Arg Val Ser Phe Pro Lys Thr Pro Glu Val
625 630 635 640

Gly Pro Gly Pro Pro Pro Gly Pro Leu Ser Lys Ala Pro Gln Pro Val
645 650 655
Pro Pro Gly Val Gly Glu Leu Pro Ala Arg Gly Pro Arg Leu Phe Asp
660 665 670
Phe Pro Pro Thr Pro Leu Glu Asp Gln Phe Glu Glu Pro Ala Glu Phe
675 680 685
Lys Ile Leu Pro Asp Gly Leu Ala Asn Ile Met Lys Met Leu Asp Glu
690 695 700

Ser Ile Arg Lys Glu Glu Glu Gln Gln Gln His Glu Ala Gly Val Ala
705 710 715 720
Pro Gln Pro Pro Leu Lys Glu Pro Phe Ala Ser Leu Gln Ser Pro Phe
725 730 735
Pro Thr Asp Thr Ala Pro Thr Thr Thr Ala Pro Ala Val Ala Val Thr
740 745 750
Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Ala Thr Gln Glu Glu Glu
755 760 765

Lys Lys Pro Pro Pro Ala Leu Pro Pro Pro Pro Pro Leu Ala Lys Phe
770 775 780
Pro Pro Pro Ser Gln Pro Gln Pro Pro Pro Pro Pro Pro Pro Ser Pro
785 790 795 800
Ala Ser Leu Leu Lys Ser Leu Ala Ser Val Leu Glu Gly Gln Lys Tyr

805 810 815

Cys Tyr Arg Gly Thr Gly Ala Ala Val Ser Thr Arg Pro Gly Pro Leu

820 825 830

Pro Thr Thr Gln Tyr Ser Pro Gly Pro Pro Ser Gly Ala Thr Ala Leu

835 840 845

Pro Pro Thr Ser Ala Ala Pro Ser Ala Gln Gly Ser Pro Gln Pro Ser

850 855 860

Ala Ser Ser Ser Ser Gln Phe Ser Thr Ser Gly Gly Pro Trp Ala Arg

865 870 875 880

Glu Arg Arg Ala Gly Glu Glu Pro Val Pro Gly Pro Met Thr Pro Thr

885 890 895

Gln Pro Pro Pro Pro Leu Ser Leu Pro Pro Ala Arg Ser Glu Ser Glu

900 905 910

Val Leu Glu Glu Ile Ser Arg Ala Cys Glu Thr Leu Val Glu Arg Val

915 920 925

Gly Arg Ser Ala Thr Asp Pro Ala Asp Pro Val Asp Thr Ala Glu Pro

930 935 940

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

945 950 955 960

Glu Ala Gly Gly Val Ala Ala Val Ser Gly Ser Cys Lys Arg Arg Gln

965 970 975

Lys Glu His Gln Lys Glu His Arg Arg His Arg Arg Ala Cys Lys Asp

980 985 990

Ser Val Gly Arg Arg Pro Arg Glu Gly Arg Ala Lys Ala Lys Ala Lys

995 1000 1005

Val Pro Lys Glu Lys Ser Arg Arg Val Leu Gly Asn Leu Asp Leu

1010 1015 1020

Gln Ser Glu Glu Ile Gln Gly Arg Glu Lys Ser Arg Pro Asp Leu

1025 1030 1035

Gly Gly Ala Ser Lys Ala Lys Pro Pro Thr Ala Pro Ala Pro Pro

1040 1045 1050

Ser Ala Pro Ala Pro Ser Ala Gln Pro Thr Pro Pro Ser Ala Ser
1055 1060 1065
Val Pro Gly Lys Lys Ala Arg Glu Glu Ala Pro Gly Pro Pro Gly
1070 1075 1080

Val Ser Arg Ala Asp Met Leu Lys Leu Arg Ser Leu Ser Glu Gly
1085 1090 1095
Pro Pro Lys Glu Leu Lys Ile Arg Leu Ile Lys Val Glu Ser Gly
1100 1105 1110
Asp Lys Glu Thr Phe Ile Ala Ser Glu Val Glu Glu Arg Arg Leu
1115 1120 1125
Arg Met Ala Asp Leu Thr Ile Ser His Cys Ala Ala Asp Val Val
1130 1135 1140

Arg Ala Ser Arg Asn Ala Lys Val Lys Gly Lys Phe Arg Glu Ser
1145 1150 1155
Tyr Leu Ser Pro Ala Gln Ser Val Lys Pro Lys Ile Asn Thr Glu
1160 1165 1170
Glu Lys Leu Pro Arg Glu Lys Leu Asn Pro Pro Thr Pro Ser Ile
1175 1180 1185
Tyr Leu Glu Ser Lys Arg Asp Ala Phe Ser Pro Val Leu Leu Gln
1190 1195 1200

Phe Cys Thr Asp Pro Arg Asn Pro Ile Thr Val Ile Arg Gly Leu
1205 1210 1215
Ala Gly Ser Leu Arg Leu Asn Leu Gly Leu Phe Ser Thr Lys Thr
1220 1225 1230
Leu Val Glu Ala Ser Gly Glu His Thr Val Glu Val Arg Thr Gln
1235 1240 1245
Val Gln Gln Pro Ser Asp Glu Asn Trp Asp Leu Thr Gly Thr Arg
1250 1255 1260

Gln Ile Trp Pro Cys Glu Ser Ser Arg Ser His Thr Thr Ile Ala
1265 1270 1275
Lys Tyr Ala Gln Tyr Gln Ala Ser Ser Phe Gln Glu Ser Leu Gln

1280 1285 1290

Glu Glu Lys Glu Ser Glu Asp Glu Glu Ser Glu Glu Pro Asp Ser

1295 1300 1305

Thr Thr Gly Thr Pro Pro Ser Ser Ala Pro Asp Pro Lys Asn His

1310 1315 1320

His Ile Ile Lys Phe Gly Thr Asn Ile Asp Leu Ser Asp Ala Lys

1325 1330 1335

Arg Trp Lys Pro Gln Leu Gln Glu Leu Leu Lys Leu Pro Ala Phe

1340 1345 1350

Met Arg Val Thr Ser Thr Gly Asn Met Leu Ser His Val Gly His

1355 1360 1365

Thr Ile Leu Gly Met Asn Thr Val Gln Leu Tyr Met Lys Val Pro

1370 1375 1380

Gly Ser Arg Thr Pro Gly His Gln Glu Asn Asn Asn Phe Cys Ser

1385 1390 1395

Val Asn Ile Asn Ile Gly Pro Gly Asp Cys Glu Trp Phe Ala Val

1400 1405 1410

His Glu His Tyr Trp Glu Thr Ile Ser Ala Phe Cys Asp Arg His

1415 1420 1425

Gly Val Asp Tyr Leu Thr Gly Ser Trp Trp Pro Ile Leu Asp Asp

1430 1435 1440

Leu Tyr Ala Ser Asn Ile Pro Val Tyr Arg Phe Val Gln Arg Pro

1445 1450 1455

Gly Asp Leu Val Trp Ile Asn Ala Gly Thr Val His Trp Val Gln

1460 1465 1470

Ala Thr Gly Trp Cys Asn Asn Ile Ala Trp Asn Val Gly Pro Leu

1475 1480 1485

Thr Ala Tyr Gln Tyr Gln Leu Ala Leu Glu Arg Tyr Glu Trp Asn

1490 1495 1500

Glu Val Lys Asn Val Lys Ser Ile Val Pro Met Ile His Val Ser

1505 1510 1515

Trp Asn Val Ala Arg Thr Val Lys Ile Ser Asp Pro Asp Leu Phe
1520 1525 1530
Lys Met Ile Lys Phe Cys Leu Leu Gln Ser Met Lys His Cys Gln
1535 1540 1545
Val Gln Arg Glu Ser Leu Val Arg Ala Gly Lys Lys Ile Ala Tyr
1550 1555 1560

Gln Gly Arg Val Lys Asp Glu Pro Ala Tyr Tyr Cys Asn Glu Cys
1565 1570 1575
Asp Val Glu Val Phe Asn Ile Leu Phe Val Thr Ser Glu Asn Gly
1580 1585 1590
Ser Arg Asn Thr Tyr Leu Val His Cys Glu Gly Cys Ala Arg Arg
1595 1600 1605
Arg Ser Ala Gly Leu Gln Gly Val Val Val Leu Glu Gln Tyr Arg
1610 1615 1620

Thr Glu Glu Leu Ala Gln Ala Tyr Asp Ala Phe Thr Leu Ala Pro
1625 1630 1635
Ala Ser Thr Ser Arg
1640

<210> 9
<211> 6731
<212> DNA
<213> Homo sapiens
<400> 9

ggcaacatgc cagccccgta gcactgccca ccccacccac tgtggtctgt tgtaccccac 60
tgctggggtg gtggttccaa tgagacaggg cacaccaaac tccatctggc tgttactgag 120
gcggagacac gggtgatgat tggctttctg gggagagagg aagtcctgtg attggccaga 180
tctctggagc ttgccgacgc ggtgtgagga cgctcccacg gaggccggaa ttggctgtga 240

aaggactgag gcagccatct gggggtagcg ggcactctta tcagagcggc tggagccgga 300
ccatcgtccc agagagctgg ggcaggggggc cgtgcccaat ctccagggct cctggggcca 360
ctgctgacct ggctggatgc atcgggcagt ggaccctcca ggggcccgcg ctgcacggga 420
agcctttgcc cttgggggcc tgagctgtgc tggggcctgg agctcctgcc cgcctcatcc 480
ccctcctcgt agcgcatggc tgcctggagg cagatgctca gccagcattg ggcagccccc 540

gcttcctgct cccctacccc cttcacatgg cagtagttct gggcacccca gcaaaccata 600

ttatgctcca ggggcgccca ctccaagacc cctccatggg aagctggaat ccctgcatgg 660

ctgtgtgcag gcattgctcc gggagccagc ccagccaggg ctttgggaac agcttgggca 720

actgtacgag tcagagcacg atagtgagga ggccacacgc tgctaccaca gcgcccttcg 780

atacggagga agcttcgctg agctggggcc ccgcattggc cgactgcagc aggcccagct 840

ctggaacttt catactggct cctgccagca ccgagccaag gtcctgcccc cactggagca 900

agtgtggaac ttgctacacc ttgagcacaa acggaactat ggagccaagc ggggaggtcc 960

cccggtgaag cgagctgctg aacccccagt ggtgcagcct gtgcctcctg cagcactctc 1020

aggcccctca ggggaggagg gcctcagccc tggaggcaag cgaaggagag gctgcaactc 1080

tgaacagact ggccttcccc cagggctgcc actgcctcca ccaccattac caccaccacc 1140

accaccacca ccaccaccac caccaccccct gcctggcctg gctaccagcc ccccatttca 1200

gctaaccaag ccagggctgt ggagtaccct gcatggagat gcctggggcc cagagcgcaa 1260

gggttcagca cccccagagc gccaggagca gcggcactcg ctgcctcacc catatccata 1320

cccagctcca gcgtacaccg cgcacccccc tggccaccgg ctggtcccgg ctgctccccc 1380

aggcccaggc ccccgccccc caggagcaga gagccatggc tgcctgcctg ccacccgtcc 1440

ccccggaagt gaccttagag agagcagagt tcagaggtcg cggatggact ccagcgtttc 1500

accagcagca accaccgcct gcgtgcctta cgccccttcc cggcccccctg gcctccccgg 1560

caccaccacc agcagcagca gtagcagcag cagcaacact ggtctccggg gcgtggagcc 1620

gaacccaggc attcccggcg ctgaccatta ccaaactccc gcgctggagg tctctcacca 1680

tggccgcctg gggccctcgg cacacagcag tcggaaaccg ttcttggggg ctcccgctgc 1740

cactccccac ctatccctgc cacctggacc ttcctcaccc cctccacccc cctgtccccg 1800

cctcttacgc cccccaccac cccctgcctg gttgaagggt ccggcctgcc gggcagcccg 1860

agaggatgga gagatcttag aagagctctt ctttgggact gagggacccc cccgccctgc 1920

cccaccaccc ctcccccatc gcgagggctt cttggggcct ccggcctccc gcttttctgt 1980

gggcactcag gattctcaca cccctcccac tcccccaacc ccaaccacca gcagtagcaa 2040

cagcaacagt ggcagccaca gcagcagccc tgctgggcct gtgtcctttc ccccaccacc 2100

ctatctggcc agaagtatag acccccttcc ccggcctccc agcccagcac agaaccccca 2160

ggacccacct cttgtacccc tgactcttgc cctgcctcca gcccctcctt cctcctgcca 2220

ccaaaatacc tcaggaagct tcaggcgccc ggagagcccc cggcccaggg tctccttccc 2280

aaagaccccc gaggtggggc cggggccacc cccaggcccc ctgagtaaag ccccccagcc 2340

tgtgccgccc ggggttgggg agctgcctgc ccgaggccct cgactctttg attttccccc 2400
cactccgctg gaggaccagt ttgaggagcc agccgaattc aagatcctac ctgatgggct 2460
ggccaacatc atgaagatgc tggacgaatc cattcgcaag gaagaggaac agcaacaaca 2520
cgaagcaggc gtggcccccc aacccccgct gaaggagccc tttgcatctc tgcagtctcc 2580
tttccccacc gacacagccc ccaccactac tgctcctgct gtcgccgtca ccaccaccac 2640
caccaccacc accaccacca cggccaccca ggaagaggag aagaagccac caccagccct 2700
accaccacca ccgcctctag ccaagttccc tccaccctct cagccacagc caccaccacc 2760

cccacccccc agcccggcca gcctgctcaa atccttggcc tccgtgctgg agggacaaaa 2820
gtactgttat cgggggactg gagcagctgt ttccacccgg cctgggccct tgcccaccac 2880
tcagtattcc cctggccccc catcaggtgc taccgccctg ccgcccacct cagcggcccc 2940
tagcgcccag ggctccccac agccctctgc ttcctcgtca tctcagttct ctacctcagg 3000
cgggccctgg gcccgggagc gcagggcggg cgaagagcca gtcccgggcc ccatgacccc 3060
cacccaaccg cccccacccc tatctctgcc ccctgctcgc tctgagtctg aggtgctaga 3120
agagatcagc cgggcttgcg agacccttgt ggagcgggtg ggccggagtg ccactgaccc 3180

agccgaccca gtggacacag cagagccagc ggacagtggg actgagcgac tgctgccccc 3240
cgcacaggcc aaggaggagg ctggcggggt ggcggcagtg tcaggcagct gtaagcggcg 3300
acagaaggag catcagaagg agcatcggcg gcacaggcgg gcctgtaagg acagtgtggg 3360
tcgtcggccc cgtgagggca gggcaaaggc caaggccaag gtccccaaag aaaagagccg 3420
ccgggtgctg gggaacctgg acctgcagag cgaggagatc cagggtcgtg agaagtcccg 3480
gcccgatctt ggcggggcct ccaaggccaa gccacccaca gctccagccc ctccatcagc 3540
tcctgcacct tctgcccagc ccacaccccc gtcagcctct gtccctggaa agaaggctcg 3600

ggaggaagcc ccagggccac cgggtgtcag ccgggccgac atgctgaagc tgcgctcact 3660
tagtgagggg ccccccaagg agctgaagat ccggctcatc aaggtagaga gtggtgacaa 3720
ggagaccttt atcgcctctg aggtggaaga gcggcggctg cgcatggcag acctcaccat 3780
cagccactgt gctgctgacg tcgtgcgcgc cagcaggaat gccaaggtga aagggaagtt 3840
tcgagagtcc tacctttccc ctgcccagtc tgtgaaaccg aagatcaaca ctgaggagaa 3900
gctgccccgg gaaaaactca acccccctac acccagcatc tatctggaga gcaaacggga 3960
tgccttctca cctgtcctgc tgcagttctg tacagaccct cgaaatccca tcacagtgat 4020

ccggggcctg gcgggctccc tgcggctcaa cttgggcctc ttctccacca agaccctggt 4080

ggaagcgagt ggcgaacaca ccgtggaagt tcgcacccag gtgcagcagc cctcagatga 4140

gaactgggat ctgacaggca ctcggcagat ctggccttgt gagagctccc gttcccacac 4200

caccattgcc aagtacgcac agtaccaggc ctcatccttc caggagtctc tgcaggagga 4260

gaaggagagt gaggatgagg agtcagagga gccagacagc accactggaa cccctcctag 4320

cagcgcacca gacccgaaga accatcacat catcaagttt ggcaccaaca tcgacttgtc 4380

tgatgctaag cggtggaagc cccagctgca ggagctgctg aagctgcccg ccttcatgcg 4440

ggtaacatcc acgggcaaca tgctgagcca cgtgggccac accatcctgg gcatgaacac 4500

ggtgcagctg tacatgaagg tgcccggcag ccgaacgcca ggccaccagg agaataacaa 4560

cttctgctcc gtcaacatca acattggccc aggcgactgc gagtggttcg cggtgcacga 4620

gcactactgg gagaccatca gcgctttctg tgatcggcac ggcgtggact acttgacggg 4680

ttcctggtgg ccaatcctgg atgatctcta tgcatccaat attcctgtgt accgcttcgt 4740

gcagcgaccc ggagacctcg tgtggattaa tgcggggact gtgcactggg tgcaggccac 4800

cggctggtgc aacaacattg cctggaacgt ggggcccctc accgcctatc agtaccagct 4860

ggccctggaa cgatacgagt ggaatgaggt gaagaacgtc aaatccatcg tgcccatgat 4920

tcacgtgtca tggaacgtgg ctcgcacggt caaaatcagc gaccccgact tgttcaagat 4980

gatcaagttc tgcctgctgc agtccatgaa gcactgccag gtgcaacgcg agagcctggt 5040

gcgggcaggg aagaaaatcg cttaccaggg ccgtgtcaag gacgagccag cctactactg 5100

caacgagtgc gatgtggagg tgtttaacat cctgttcgtg acaagtgaga atggcagccg 5160

caacacgtac ctggtacact gcgagggctg tgcccggcgc cgcagcgcag gcctgcaggg 5220

cgtggtggtg ctggagcagt accgcactga ggagctggct caggcctacg acgccttcac 5280

gctggtgagg gcccggcggg cgcgcgggca gcggaggagg gcactggggc aggctgcagg 5340

gacgggcttc gggagcccgg ccgcgccttt ccctgagccc ccgccggctt tctcccccca 5400

ggccccagcc agcacgtcgc gatgaggccg gacgccccgc ccgcctgcct gcccgcgcaa 5460

ggcgccgcgg ggccaccagc acatgcctgg gctggaccta ggtcccgcct gtggccgaga 5520

agggggtcgg gcccagccct tccaccccat tggcagctcc cctcacttaa tttattaaga 5580

aaaacttttt ttttttttt agcaaatatg aggaaaaaag gaaaaaaaat gggagacggg 5640

ggaggggggct ggcagcccct cgcccaccag cgcctcccct caccgacttt ggccttttta 5700

gcaacagaca caaggaccag gctccggcgg cggcgggggt cacatacggg ttccctcacc 5760

ctgccagccg cccgcccgcc cggcgcagat gcacgcggct cgtgtatgta catagacgtt 5820

acggcagccg aggtttttaa tgagattctt tctatgggct ttacccctcc cccggaacct 5880

ccttttttac ttccaatgct agctgtgacc cctgtacatg tctctttatt cacttggtta 5940

tgatttgtat tttttgttct tttcttgttt ttttgttttt aatttataac agtcccactc 6000

acctctattt attcattttt gggaaaaccc gacctcccac acccccaagc catcctgccc 6060

gcccctccag ggaccgcccg tcgccgggct ctccccgcgc cccagtgtgt gtccgggccc 6120

ggcccgaccg tctccacccg tccgcccgcg gctccagccg ggttctcatg gtgctcaaac 6180

ccgctcccct cccctacgtc ctgcactttc tcggaccagt ccccccactc ccgacccgac 6240

cccagcccca cctgagggtg agcaactcct gtactgtagg ggaagaagtg ggaactgaaa 6300

tggtattttg taaaaaaaat aaataaaata aaaaaattaa aggttttaaa gaaagaacta 6360

tgaggaaaag gaaccccgtc cttcccagcc ccggccaact ttaaaaaaca cagaccttca 6420

cccccacccc cttttctttt taagtgtgaa acaacccagg gccagggcct cactggggca 6480

gggacacccc ggggtgagtt tctctggggc tttattttcg ttttgttggt tgtttttttct 6540

ccacgctggg gctgcggagg ggtgggggggt ttacagtccc gcaccctcgc actgcactgt 6600

ctctctgccc caggggcaga ggggtcttcc caaccctacc cctattttcg gtgatttttg 6660

tgtgagaata ttaatattaa aaataaacgg agaaaaaaaa aaaaaaaaaa aaaaaaaaaa 6720

aaaaaaaaaa a 6731

<210> 10

<211> 1347

<212> PRT

<213> Homo sapiens

<400> 10

Met Lys Ser Cys Ala Val Ser Leu Thr Thr Ala Ala Val Ala Phe Gly

1 5 10 15

Asp Glu Ala Lys Lys Met Ala Glu Gly Lys Ala Ser Arg Glu Ser Glu

20 25 30

Glu Glu Ser Val Ser Leu Thr Val Glu Glu Arg Glu Ala Leu Gly Gly

35 40 45

Met Asp Ser Arg Leu Phe Gly Phe Val Arg Leu His Glu Asp Gly Ala

50 55 60

Arg Thr Lys Thr Leu Leu Gly Lys Ala Val Arg Cys Tyr Glu Ser Leu

65 70 75 80

Ile Leu Lys Ala Glu Gly Lys Val Glu Ser Asp Phe Phe Cys Gln Leu

85 90 95

Gly His Phe Asn Leu Leu Leu Glu Asp Tyr Ser Lys Ala Leu Ser Ala

100 105 110

Tyr Gln Arg Tyr Tyr Ser Leu Gln Ala Asp Tyr Trp Lys Asn Ala Ala

115 120 125

Phe Leu Tyr Gly Leu Gly Leu Val Tyr Phe Tyr Tyr Asn Ala Phe His

130 135 140

Trp Ala Ile Lys Ala Phe Gln Asp Val Leu Tyr Val Asp Pro Ser Phe

145 150 155 160

Cys Arg Ala Lys Glu Ile His Leu Arg Leu Gly Leu Met Phe Lys Val

165 170 175

Asn Thr Asp Tyr Lys Ser Ser Leu Lys His Phe Gln Leu Ala Leu Ile

180 185 190

Asp Cys Asn Pro Cys Thr Leu Ser Asn Ala Glu Ile Gln Phe His Ile

195 200 205

Ala His Leu Tyr Glu Thr Gln Arg Lys Tyr His Ser Ala Lys Glu Ala

210 215 220

Tyr Glu Gln Leu Leu Gln Thr Glu Asn Leu Pro Ala Gln Val Lys Ala

225 230 235 240

Thr Val Leu Gln Gln Leu Gly Trp Met His His Asn Met Asp Leu Val

245 250 255

Gly Asp Lys Ala Thr Lys Glu Ser Tyr Ala Ile Gln Tyr Leu Gln Lys

260 265 270

Ser Leu Glu Ala Asp Pro Asn Ser Gly Gln Ser Trp Tyr Phe Leu Gly

275 280 285

Arg Cys Tyr Ser Ser Ile Gly Lys Val Gln Asp Ala Phe Ile Ser Tyr

290 295 300

Arg Gln Ser Ile Asp Lys Ser Glu Ala Ser Ala Asp Thr Trp Cys Ser

305 310 315 320

Ile Gly Val Leu Tyr Gln Gln Gln Asn Gln Pro Met Asp Ala Leu Gln

325 330 335

Ala Tyr Ile Cys Ala Val Gln Leu Asp His Gly His Ala Ala Ala Trp

340 345 350

Met Asp Leu Gly Thr Leu Tyr Glu Ser Cys Asn Gln Pro Gln Asp Ala

355 360 365

Ile Lys Cys Tyr Leu Asn Ala Ala Arg Ser Lys Arg Cys Ser Asn Thr

370 375 380

Ser Thr Leu Ala Ala Arg Ile Lys Phe Leu Gln Asn Gly Ser Asp Asn

385 390 395 400

Trp Asn Gly Gly Gln Ser Leu Ser His His Pro Val Gln Gln Val Tyr

405 410 415

Ser Leu Cys Leu Thr Pro Gln Lys Leu Gln His Leu Glu Gln Leu Arg

420 425 430

Ala Asn Arg Asp Asn Leu Asn Pro Ala Gln Lys His Gln Leu Glu Gln

435 440 445

Leu Glu Ser Gln Phe Val Leu Met Gln Gln Met Arg His Lys Glu Val

450 455 460

Ala Gln Val Arg Thr Thr Gly Ile His Asn Gly Ala Ile Thr Asp Ser

465 470 475 480

Ser Leu Pro Thr Asn Ser Val Ser Asn Arg Gln Pro His Gly Ala Leu

485 490 495

Thr Arg Val Ser Ser Val Ser Gln Pro Gly Val Arg Pro Ala Cys Val

500 505 510

Glu Lys Leu Leu Ser Ser Gly Ala Phe Ser Ala Gly Cys Ile Pro Cys

515 520 525

Gly Thr Ser Lys Ile Leu Gly Ser Thr Asp Thr Ile Leu Leu Gly Ser

530 535 540

Asn Cys Ile Ala Gly Ser Glu Ser Asn Gly Asn Val Pro Tyr Leu Gln

545 550 555 560

Gln Asn Thr His Thr Leu Pro His Asn His Thr Asp Leu Asn Ser Ser

565 570 575

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

580 585 590

His Lys Ser Gln Ser Ser Cys Leu Ser Gly Pro Asn Glu Glu Gln Pro

595 600 605

Leu Phe Ser Thr Gly Ser Ala Gln Tyr His Gln Ala Thr Ser Thr Gly

610 615 620

Ile Lys Lys Ala Asn Glu His Leu Thr Leu Pro Ser Asn Ser Val Pro

625 630 635 640

Gln Gly Asp Ala Asp Ser His Leu Ser Cys His Thr Ala Thr Ser Gly

645 650 655

Gly Gln Gln Gly Ile Met Phe Thr Lys Glu Ser Lys Pro Ser Lys Asn

660 665 670

Arg Ser Leu Val Pro Glu Thr Ser Arg His Thr Gly Asp Thr Ser Asn

675 680 685

Gly Cys Ala Asp Val Lys Gly Leu Ser Asn His Val His Gln Leu Ile

690 695 700

Ala Asp Ala Val Ser Ser Pro Asn His Gly Asp Ser Pro Asn Leu Leu

705 710 715 720

Ile Ala Asp Asn Pro Gln Leu Ser Ala Leu Leu Ile Gly Lys Ala Asn

725 730 735

Gly Asn Val Gly Thr Gly Thr Cys Asp Lys Val Asn Asn Ile His Pro

740 745 750

Ala Val His Thr Lys Thr Asp His Ser Val Ala Ser Ser Pro Ser Ser

755 760 765

Ala Ile Ser Thr Ala Thr Pro Ser Pro Lys Ser Thr Glu Gln Arg Ser

770 775 780

Ile Asn Ser Val Thr Ser Leu Asn Ser Pro His Ser Gly Leu His Thr

785 790 795 800

Val Asn Gly Glu Gly Leu Gly Lys Ser Gln Ser Ser Thr Lys Val Asp

805 810 815

Leu Pro Leu Ala Ser His Arg Ser Thr Ser Gln Ile Leu Pro Ser Met

820 825 830

Ser Val Ser Ile Cys Pro Ser Ser Thr Glu Val Leu Lys Ala Cys Arg

835 840 845

Asn Pro Gly Lys Asn Gly Leu Ser Asn Ser Cys Ile Leu Leu Asp Lys

850 855 860

Cys Pro Pro Pro Arg Pro Pro Thr Ser Pro Tyr Pro Pro Leu Pro Lys

865 870 875 880

Asp Lys Leu Asn Pro Pro Thr Pro Ser Ile Tyr Leu Glu Asn Lys Arg

885 890 895

Asp Ala Phe Phe Pro Pro Leu His Gln Phe Cys Thr Asn Pro Lys Asn

900 905 910

Pro Val Thr Val Ile Arg Gly Leu Ala Gly Ala Leu Lys Leu Asp Leu

915 920 925

Gly Leu Phe Ser Thr Lys Thr Leu Val Glu Ala Asn Asn Glu His Met

930 935 940

Val Glu Val Arg Thr Gln Leu Leu Gln Pro Ala Asp Glu Asn Trp Asp

945 950 955 960

Pro Thr Gly Thr Lys Lys Ile Trp Arg Cys Glu Ser Asn Arg Ser His

965 970 975

Thr Thr Ile Ala Lys Tyr Ala Gln Tyr Gln Ala Ser Ser Phe Gln Glu

980 985 990

Ser Leu Arg Glu Glu Asn Glu Lys Arg Thr Gln His Lys Asp His Ser

995 1000 1005

Asp Asn Glu Ser Thr Ser Ser Glu Asn Ser Gly Arg Arg Arg Lys

1010 1015 1020

Gly Pro Phe Lys Thr Ile Lys Phe Gly Thr Asn Ile Asp Leu Ser

1025 1030 1035

Asp Asn Lys Lys Trp Lys Leu Gln Leu His Glu Leu Thr Lys Leu

1040 1045 1050

Pro Ala Phe Ala Arg Val Val Ser Ala Gly Asn Leu Leu Thr His

1055 1060 1065

Val Gly His Thr Ile Leu Gly Met Asn Thr Val Gln Leu Tyr Met

1070 1075 1080

Lys Val Pro Gly Ser Arg Thr Pro Gly His Gln Glu Asn Asn Asn

1085 1090 1095

Phe Cys Ser Val Asn Ile Asn Ile Gly Pro Gly Asp Cys Glu Trp

1100 1105 1110

Phe Val Val Pro Glu Asp Tyr Trp Gly Val Leu Asn Asp Phe Cys

1115 1120 1125

Glu Lys Asn Asn Leu Asn Phe Leu Met Ser Ser Trp Trp Pro Asn

1130 1135 1140

Leu Glu Asp Leu Tyr Glu Ala Asn Val Pro Val Tyr Arg Phe Ile

1145 1150 1155

Gln Arg Pro Gly Asp Leu Val Trp Ile Asn Ala Gly Thr Val His

1160 1165 1170

Trp Val Gln Ala Val Gly Trp Cys Asn Asn Ile Ala Trp Asn Val

1175 1180 1185

Gly Pro Leu Thr Ala Cys Gln Tyr Lys Leu Ala Val Glu Arg Tyr

1190 1195 1200

Glu Trp Asn Lys Leu Lys Ser Val Lys Ser Pro Val Pro Met Val

1205 1210 1215

His Leu Ser Trp Asn Met Ala Arg Asn Ile Lys Val Ser Asp Pro

1220 1225 1230

Lys Leu Phe Glu Met Ile Lys Tyr Cys Leu Leu Lys Ile Leu Lys

1235 1240 1245

Gln Tyr Gln Thr Leu Arg Glu Ala Leu Val Ala Ala Gly Lys Glu

1250 1255 1260

Val Ile Trp His Gly Arg Thr Asn Asp Glu Pro Ala His Tyr Cys

1265 1270 1275

Ser Ile Cys Glu Val Glu Val Phe Asn Leu Leu Phe Val Thr Asn

1280 1285 1290

Glu Ser Asn Thr Gln Lys Thr Tyr Ile Val His Cys His Asp Cys

1295 1300 1305

Ala Arg Lys Thr Ser Lys Ser Leu Glu Asn Phe Val Val Leu Glu
1310 1315 1320

Gln Tyr Lys Met Glu Asp Leu Ile Gln Val Tyr Asp Gln Phe Thr
1325 1330 1335

Leu Ala Leu Ser Leu Ser Ser Ser Ser
1340 1345

<210> 11

<211> 6817

<212> DNA

<213> Homo sapiens

<400> 11

gctcatcgtt tgttgtttag ataatatcat gaactgataa atgcagttgc cacgttgatt 60

ccctagggcc tggcttaccg actgaggtca taagatatta tgccttctct ttagacttgg 120

tcagtggaga ggaaatgggc aaagaaccag cctatggagg tgacaaggcc ttagggccaa 180

aagtcttgag ggtgaaggtt tagggcctgc gcagcttccc tgccatgccc cgcaaggtct 240

cgcattcgca aggcttgtga cagtgggagc ctcattacgg actctcctaa agtccatggt 300

gtcctctttt cgcatttgcg ccccgtgggt gatgcccgat gccgcccttc ccatcgctct 360

cttccccttc aagcgtatcg caactgcaaa aacacccagc acagacactc cattttctat 420

cttaatgcat ttaactagca caacctacag gttgttccat cccagagact acccttttct 480

ccatagacgt gaccatcaac caaccagcgg tcagaatcag tcagcctctg tcatgttcct 540

aggtccttgg cgaactggct gggcggggtc ccagcagcct aggagtacag tggagcaatg 600

cctgacgtaa gtcaacaaag atcacgtgag acgaatcagt cgcctagatt ggctacaact 660

aagtggttgg gagcggggag gtcgcggcgg ctgcgtgggg ttcgcccgtg acacaattac 720

aactttgtgc tggtgctggc aaagtttgtg attttaagaa attctgctgt gctctccagc 780

actgcgagct tctgccttcc ctgtagtttc ccagatgtga tccaggtagc cgagttccgc 840

tgcccgtgct tcggtagctt aagtctttgc ctcagctttt ttccttgcag ccgctgagga 900

ggcgataaaa ttggcgtcac agtctcaagc agcgattgaa ggcgtctttt caactactcg 960

attaaggttg ggtatcgtcg tgggacttgg aaatttgttg tttccatgaa atcctgcgca 1020

gtgtcgctca ctaccgccgc tgttgccttc ggtgatgagg caaagaaaat ggcggaagga 1080

aaagcgagcc gcgagagtga agaggagtct gttagcctga cagtcgagga aagggaggcg 1140

cttggtggca tggacagccg tctcttcggg ttcgtgaggc ttcatgaaga tggcgccaga 1200

acgaagaccc tactaggcaa ggctgttcgc tgctacgaat ctttaatctt aaaagctgaa 1260

ggaaaagtgg agtctgactt cttttgccaa ttaggtcact tcaacctctt gttggaagat 1320

tattcaaaag cattatctgc atatcagaga tattacagtt tacaggctga ctactggaag 1380

aatgctgcgt ttttatatgg ccttggtttg gtctacttct actacaatgc atttcattgg 1440

gcaattaaag catttcaaga tgtcctttat gttgacccca gcttttgtcg agccaaggaa 1500

attcatttac gacttgggct catgttcaaa gtgaacacag actacaagtc tagtttaaag 1560

cattttcagt tagccttgat tgactgtaat ccatgtactt tgtccaatgc tgaaattcaa 1620

tttcatattg cccatttgta tgaaacccag aggaagtatc attctgcaaa ggaggcatat 1680

gaacaacttt tgcagacaga aaaccttcct gcacaagtaa aagcaactgt attgcaacag 1740

ttaggttgga tgcatcataa tatggatcta gtaggagaca aagccacaaa ggaaagctat 1800

gctattcagt atctccaaaa gtctttggag gcagatccta attctggcca atcgtggtat 1860

tttcttggaa ggtgttattc aagtattggg aaagttcagg atgcctttat atcttacagg 1920

caatctattg ataaatcaga agcaagtgca gatacatggt gttcaatagg tgtgttgtat 1980

cagcagcaaa atcagcctat ggatgcttta caggcatata tttgtgctgt acaattggac 2040

catgggcatg ccgcagcctg gatggaccta ggtactctct atgaatcctg caatcaacct 2100

caagatgcca ttaaatgcta cctaaatgca gctagaagca aacgttgtag taatacctct 2160

acgcttgctg caagaattaa atttctacag gctcagttgt gtaaccttcc acaaagtagt 2220

ctacagaata aaactaaatt acttcctagt attgaggagg catggagcct accaatcccc 2280

gcagagctta cctccaggca gggtgccatg aacacagcac agcaggctta tagagctcat 2340

gatccaaata ctgaacatgt attaaaccac agtcaaacac caattttaca gcaatccttg 2400

tcactacaca tgattacttc tagccaagta gaaggcctgt ccagtcctgc caagaagaaa 2460

agaacatcta gtccaacaaa gaatggttct gataactgga atggtggcca gagtctttca 2520

catcatccag tacagcaagt ttattcgttg tgtttgacac cacagaaatt acagcacttg 2580

gaacaactgc gagcaaatag agataattta aatccagcac agaagcatca gctggaacag 2640

ttagaaagtc agtttgtctt aatgcagcaa atgagacaca aagaagttgc tcaggtacga 2700

actactggaa ttcataacgg ggccataact gattcatcac tgcctacaaa ctctgtctct 2760

aatcgacaac cacatggtgc tctgaccaga gtatctagcg tctctcagcc tggagttcgc 2820

cctgcttgtg ttgaaaaact tttgtccagt ggagcttttt ctgcaggctg tattccttgt 2880

ggcacatcaa aaattctagg aagtacagac actatcttgc taggcagtaa ttgtatagca 2940

ggaagtgaaa gtaatggaaa tgtgccttac ctgcagcaaa atacacacac tctacctcat 3000

aatcatacag acctgaacag cagcacagaa gagccatgga gaaaacagct atctaactcc 3060

gctcaggggc ttcataaaag tcagagttca tgtttgtcag gacctaatga agaacaacct 3120

ctgttttcca ctgggtcagc ccagtatcac caggcaacta gcactggtat taagaaggcg 3180

aatgaacatc tcactctgcc tagtaattca gtaccacagg gggatgctga cagtcacctc 3240

tcctgtcata ctgctacctc aggtggacaa caaggcatta tgtttaccaa agagagcaag 3300

ccttcaaaaa atagatcctt ggtgcctgaa acaagcaggc atactggaga cacatctaat 3360

ggctgtgctg atgtcaaggg actttctaat catgttcatc agttgatagc agatgctgtt 3420

tccagtccta accatggaga ttcaccaaat ttattaattg cagacaatcc tcagctctct 3480

gctttgttga ttggaaaagc caatggcaat gtgggtactg gaacctgtga caaagtgaat 3540

aatattcacc cagctgttca tacaaagact gatcattctg ttgcctcttc accctcttca 3600

gccatttcca cagcaacacc ttctcctaaa tccactgagc agagaagcat aaacagtgtt 3660

accagcctta acagtcctca cagtggatta cacacagtca atggagaggg gctggggaag 3720

tcacagagct ctacaaaagt agacctgcct ttagctagcc acagatctac ttctcagatc 3780

ttaccatcaa tgtcagtgtc tatatgcccc agttcaacag aagttctgaa agcatgcagg 3840

aatccaggta aaaatggctt gtctaatagc tgcattttgt tagataaatg tccacctcca 3900

agaccaccaa cttcaccata cccacccttg ccaaaggaca agttgaatcc acccacacct 3960

agtatttact tggaaaataa acgtgatgct ttctttcctc cattacatca attttgtaca 4020

aatccaaaaa accctgttac agtaatacgt ggccttgctg gagctcttaa attagatctt 4080

ggacttttct ctaccaaaac tttggtagaa gctaacaatg aacatatggt agaagtgagg 4140

acacagttgc tgcaaccagc agatgaaaac tgggatccca ctggaacaaa gaaaatctgg 4200

cgttgtgaaa gcaatagatc tcatactaca attgccaaat acgcacaata ccaggcttcc 4260

tccttccagg aatcattgag agaagaaaat gagaaaagaa cacaacacaa agatcattca 4320

gataacgaat ccacatcttc agagaattct ggaaggagaa ggaaaggacc ttttaaaacc 4380

ataaaatttg ggaccaacat tgacctctct gataacaaaa agtggaagtt gcagttacat 4440

gaactgacta aacttcctgc ttttgcgcgt gtggtgtcag caggaaatct tctaacccat 4500

gttgggcata ccattctggg catgaataca gtacaactgt atatgaaagt tccagggagt 4560

cggacaccag gtcaccaaga aaataacaac ttctgctctg ttaacataaa tattggtcca 4620

ggagattgtg aatggtttgt tgtacctgaa gattattggg gtgttctgaa tgacttctgt 4680

gaaaaaaata atttgaattt tttaatgagt tcttggtggc ccaaccttga agatctttat 4740

gaagcaaatg tccctgtgta tagatttatt cagcgacctg gagatttggt ctggataaat 4800

gcaggcactg tgcattgggt tcaagctgtt ggctggtgca ataacattgc ctggaatgtt 4860

ggtccactta cagcctgcca gtataaattg gcagtggaac ggtatgaatg gaacaaattg 4920

aaaagtgtga agtcaccagt acccatggtg catctttcct ggaatatggc acgaaatatc 4980

aaagtctcag atccaaagct ttttgaaatg attaagtatt gtcttttgaa aattctgaag 5040
caatatcaga cattgagaga agctcttgtt gcagcaggaa aagaggttat atggcatggg 5100
cggacaaatg atgaaccagc tcattactgt agcatttgtg aggtggaggt ttttaatctg 5160
ctttttgtca ctaatgaaag caatactcaa aaaacctaca tagtacattg ccatgattgt 5220
gcacgaaaaa caagcaaaag tttggaaaat tttgtggtgc tcgaacagta caaaatggag 5280
gacctaatcc aagtttatga tcaatttaca ctagctcttt cattatcatc ctcatcttga 5340
tatagttcca tgaatattaa atgagattat ttctgctctt caggaaattt ctgcaccact 5400

ggttttgtag ctgtttcata aaactgttga ctaaaagcta tgtctatgca accttccaag 5460
aatagtatgt caagcaactg gacacagtgc tgcctctgct tcaggactta acatgctgat 5520
ccagctgtac ttcagaaaaa taatattaat catatgtttt gtgtacgtat gacaaactgt 5580
caaagtgaca cagaatactg atttgaagat agccttttt atgtttctct atttctgggc 5640
tgatgaatta atattcattt gtattttaac cctgcagaat tttccttagt taaaaacact 5700
ttcctagctg gtcatttctt cataagatag caaatttaaa tctctcctcg atcagctttt 5760
aaaaaatgtg tactattatc tgaggaagtt ttttactgct ttatgttttt gtgtgttttg 5820

aggccatgat gattacattt gtggttccaa aataattttt ttaaatatta atagcccata 5880
tacaaagata atggattgca catagacaaa gaaataaact tcagatttgt gatttttgtt 5940
tctaaacttg atacagattt acactattta taaatacgta tttattgcct gaaaatattt 6000
gtgaatggaa tgttgttttt ttccagacgt aactgccatt aaatactaag gagttctgta 6060
gttttaaaca ctactcctat tacattttat atgtgtagat aaaactgctt agtattatac 6120
agaaattttt attaaaattg ttaaatgttt aaagggtttc ccaatgtttg agtttaaaaa 6180
agactttctg aaaaaatcca ctttttgttc attttcaaac ctaatgatta tatgtatttt 6240

atatgtgtgt gtatgtgtac acacatgtat aatatataca gaaacctcga tatataattg 6300
tatagatttt aaaagtttta ttttttacat ctatggtagt ttttgaggtg cctattataa 6360
agtattacgg aagtttgctg tttttaaagt aaatgtcttt tagtgtgatt tattaagttg 6420
tagtcaccat agtgatagcc cataaataat tgctggaaaa ttgtatttta taacagtaga 6480
aaacatatag tcagtgaagt aaatatttta aaggaaacat tatatagatt tgataaatgt 6540
tgtttataat taagagtttc ttatggaaaa gagattcaga atgataacct cttttagaga 6600
acaaataagt gacttatttt tttaaagcta gatgactttg aaatgctata ctgtcctgct 6660

tgtacaacat ggtttggggt gaaggggagg aaagtattaa aaaatctata tcgctagtaa 6720

attgtaataa gttctattaa aacttgtatt tcatatgaaa aatttgctaa tttaatatta 6780

actcatttga taataatact tgtcttttct acctctc 6817

<210> 12

<211> 724

<212> PRT

<213> Homo sapiens

<400> 12

Met Ser Gly Gly Glu Val Val Cys Ser Gly Trp Leu Arg Lys Ser Pro
1 5 10 15

Pro Glu Lys Lys Leu Lys Arg Tyr Ala Trp Lys Arg Arg Trp Phe Val
20 25 30

Leu Arg Ser Gly Arg Leu Thr Gly Asp Pro Asp Val Leu Glu Tyr Tyr
35 40 45

Lys Asn Asp His Ala Lys Lys Pro Ile Arg Ile Ile Asp Leu Asn Leu
50 55 60

Cys Gln Gln Val Asp Ala Gly Leu Thr Phe Asn Lys Lys Glu Phe Glu
65 70 75 80

Asn Ser Tyr Ile Phe Asp Ile Asn Thr Ile Asp Arg Ile Phe Tyr Leu
85 90 95

Val Ala Asp Ser Glu Glu Glu Met Asn Lys Trp Val Arg Cys Ile Cys
100 105 110

Asp Ile Cys Gly Phe Asn Pro Thr Glu Glu Asp Pro Val Lys Pro Pro
115 120 125

Gly Ser Ser Leu Gln Ala Pro Ala Asp Leu Pro Leu Ala Ile Asn Thr
130 135 140

Ala Pro Pro Ser Thr Gln Ala Asp Ser Ser Ser Ala Thr Leu Pro Pro
145 150 155 160

Pro Tyr Gln Leu Ile Asn Val Pro Pro His Leu Glu Thr Leu Gly Ile
165 170 175

Gln Glu Asp Pro Gln Asp Tyr Leu Leu Leu Ile Asn Cys Gln Ser Lys
180 185 190

Lys Pro Glu Pro Thr Arg Thr His Ala Asp Ser Ala Lys Ser Thr Ser
195 200 205
Ser Glu Thr Asp Cys Asn Asp Asn Val Pro Ser His Lys Asn Pro Ala
210 215 220

Ser Ser Gln Ser Lys His Gly Met Asn Gly Phe Phe Gln Gln Gln Met
225 230 235 240
Ile Tyr Asp Ser Pro Pro Ser Arg Ala Pro Ser Ala Ser Val Asp Ser
245 250 255
Ser Leu Tyr Asn Leu Pro Arg Ser Tyr Ser His Asp Val Leu Pro Lys
260 265 270
Val Ser Pro Ser Ser Thr Glu Ala Asp Gly Glu Leu Tyr Val Phe Asn
275 280 285

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

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

Ser Ser Ser Leu Glu Gly Phe His Asn His Phe Lys Val Lys Asn Val
420 425 430
Leu Thr Val Gly Ser Val Ser Ser Glu Glu Leu Asp Glu Asn Tyr Val

435 440 445

Pro Met Asn Pro Asn Ser Pro Pro Arg Gln His Ser Ser Ser Phe Thr

450 455 460

Glu Pro Ile Gln Glu Ala Asn Tyr Val Pro Met Thr Pro Gly Thr Phe

465 470 475 480

Asp Phe Ser Ser Phe Gly Met Gln Val Pro Pro Pro Ala His Met Gly

485 490 495

Phe Arg Ser Ser Pro Lys Thr Pro Pro Arg Arg Pro Val Pro Val Ala

500 505 510

Asp Cys Glu Pro Pro Pro Val Asp Arg Asn Leu Lys Pro Asp Arg Lys

515 520 525

Gly Gln Ser Pro Lys Ile Leu Arg Leu Lys Pro His Gly Leu Glu Arg

530 535 540

Thr Asp Ser Gln Thr Ile Gly Asp Phe Ala Thr Arg Arg Lys Val Lys

545 550 555 560

Pro Ala Pro Leu Glu Ile Lys Pro Leu Pro Glu Trp Glu Glu Leu Gln

565 570 575

Ala Pro Val Arg Ser Pro Ile Thr Arg Ser Phe Ala Arg Asp Ser Ser

580 585 590

Arg Phe Pro Met Ser Pro Arg Pro Asp Ser Val His Ser Thr Thr Ser

595 600 605

Ser Ser Asp Ser His Asp Ser Glu Glu Asn Tyr Val Pro Met Asn Pro

610 615 620

Asn Leu Ser Ser Glu Asp Pro Asn Leu Phe Gly Ser Asn Ser Leu Asp

625 630 635 640

Gly Gly Ser Ser Pro Met Ile Lys Pro Lys Gly Asp Lys Gln Val Glu

645 650 655

Tyr Leu Asp Leu Asp Leu Asp Ser Gly Lys Ser Thr Pro Pro Arg Lys

660 665 670

Gln Lys Ser Ser Gly Ser Gly Ser Ser Val Ala Asp Glu Arg Val Asp

675 680 685

Tyr Val Val Val Asp Gln Gln Lys Thr Leu Ala Leu Lys Ser Thr Arg
690 695 700

Glu Ala Trp Thr Asp Gly Arg Gln Ser Thr Glu Ser Glu Thr Pro Ala
705 710 715 720

Lys Ser Val Lys

<210> 13

<211> 7836

<212> DNA

<213> Homo sapiens

<400> 13

aggggcgga gcgcaaagga cagaagctcc ggcaccgagt cggggcagag tcccgctgag 60

tccgagcgct gctgaggcag ctggcgagac ggcacgtctg gaggcgaggc gggcgcactg 120

aaaggaggcc ggcgcgcccg cggccccggc tcgcgttctg ttcaggttcg tgggcctgca 180

gaggagagac tcgaactcgt ggaacccgcg caccgtggag tctgtccgcc cagtccgtcc 240

ggggtgcgcg accaggagag ctaggttctc gccactgcgc gctcggcagg cgtcggctgt 300

gtcgggagcg cgcccgccgc ccctcagctg cccggcccgg agcccgagac gcgcgcacca 360

tgagcggtgg tgaagtggtc tgctccggat ggctccgcaa gtccccccg gagaaaaagt 420

tgaagcgtta tgcatggaag aggagatggt tcgtgttacg cagtggccgt ttaactggag 480

atccagatgt tttggaatat tacaaaaatg atcatgccaa gaagcctatt cgtattattg 540

atttaaattt atgtcaacaa gtagatgctg gattgacatt taacaaaaaa gagtttgaaa 600

acagctacat ttttgatatc aacactattg accggatttt ctacttggta gcagacagcg 660

aggaggagat gaataagtgg gttcgttgta tttgtgacat ctgtgggttt aatccaacag 720

aagaagatcc tgtgaagcca cctggcagct ctttacaagc accagctgat ttacctttag 780

ctataaatac agcaccacca tccacccagg cagattcatc ctctgctact ctacctcctc 840

catatcagct aatcaatgtt ccaccacacc tggaaactct tggcattcag gaggatcctc 900

aagactacct gttgctcatc aactgtcaaa gcaagaagcc cgaacccacc agaacgcatg 960

ctgattctgc aaaatccacc tcttctgaaa cagactgcaa tgataacgtc ccttctcata 1020

aaaatcctgc ttcctcccag agcaaacatg gaatgaatgg ctttttttcag cagcaaatga 1080

tatacgactc tccaccttca cgtgccccat ctgcttcagt tgactccagc ctttataacc 1140

tgcccaggag ttattcccat gatgttttac caaaggtgtc tccatcaagt actgaagcag 1200

atggagaact ctatgttttt aataccccat ctgggacatc gagtgtagag actcaaatga 1260

ggcatgtatc tattagttat gacattcctc caacacctgg taatacttat cagattccac 1320

gaacatttcc agaaggaacc ttgggacaga catcaaagct agacactatt ccagatattc 1380

ctccacctcg gccaccgaaa ccacatccag ctcatgaccg atctcctgtg gaaacgtgta 1440

gtatcccacg caccgcctca gacactgaca gtagttactg tatccctaca gcagggatgt 1500

cgccttcacg tagtaatacc atttccactg tggatttaaa caaattgcga aaagatgcta 1560

gttctcaaga ctgctatgat attccacgag catttccaag tgatagatct agttcacttg 1620

aaggcttcca taaccacttt aaagtcaaaa atgtgttgac agtgggaagt gtttcaagtg 1680

aagaactgga tgaaaattac gtcccaatga atcccaattc accaccacga caacattcca 1740

gcagttttac agaaccaatt caggaagcaa attatgtgcc aatgactcca ggaacatttg 1800

atttttcctc atttggaatg caagttcctc ctcctgctca tatgggcttc aggtccagcc 1860

caaaaacccc tcccagaagg ccagttcctg ttgcagactg tgaaccaccc cccgtggata 1920

ggaacctcaa gccagacaga aaagtcaagc cagcgccttt agaaataaaa cctttgccag 1980

aatgggaaga attacaagcc ccagttagat ctcccatcac taggagtttt gctcgagact 2040

cttccaggtt tcccatgtcc ccccgaccag attcagtgca tagcacaact tcaagcagtg 2100

actcacacga cagtgaagag aattatgttc ccatgaaccc aaacctgtcc agtgaagacc 2160

caaatctctt tggcagtaac agtcttgatg gaggaagcag ccctatgatc aagcccaaag 2220

gagacaaaca ggtggaatac ttagatctcg acttagattc tgggaaatcc acaccaccac 2280

gtaagcaaaa gagcagtggc tcaggcagca gtgtagcaga tgagagagtg gattatgttg 2340

ttgttgacca acagaagacc ttggctctaa agagtacccg ggaagcctgg acagatggga 2400

gacagtccac agaatcagaa acgccagcga agagtgtgaa atgaaaatat tgccttgcca 2460

tttctgaaca aaagaaaact gaattgtaaa gataaatccc ttttgaagaa tgacttgaca 2520

cttccactct aggtagatcc tcaaatgagt agagttgaag tcaaaggacc tttctgacat 2580

aatcaagcaa tttagactta agtggtgctt tgtggtatct gaacaattca taacatgtaa 2640

ataatgtggg aaaatagtat tgtttagctc ccagagaaac atttgttcca cagttaacac 2700

actcgtagta ttactgtatt tatgcacttt ttcatctaaa acattgttct gggttttccc 2760

aatgtacctt accataattc ctttgggagt tcttgttttt tgtcacacta ctttatataa 2820

caatactaag tcaactaagc tacttttaga tttggaaatt gctgtttaca gtctaacaac 2880

attaaaatga gaggtagatt cacaagttag ctttctacct gaagcttcag gtgataacca 2940

ttagcttata cttggactca tcatttgttg ccttccaaaa tgctgaggat aatgtatgta 3000

ctggtgtcag gacctagttc tctggttaat gtacatttag tttttaatgg tggaactttg 3060

ttatattttg ttaattacag tgtttttggt tcattgagtg aagattctgc cgggtgggat 3120

cttgcacctt tgaaagactg aataattaca ctaccaagta agcctgcaaa tcattgatgg 3180

catgcagtga tgatgtgctc ttacacttgt taacatgtat taagtgttat ttgcaaaagg 3240

tagattatgt aaccaatcag gtacgtacca ggcagtgatg tgctaataca ctgatcaggt 3300

ttagacaatg agctttggtt gtgttcttgt tagtcctaat attggttttc agtttggaat 3360

taataaagca gttgacattc actgttagtt acagcaacat actgtgattt ttaattagat 3420

agtaattcag atttattact ctatgaaatt ctgtcttttg acaccatagt gccctttcta 3480

tgatttttt tacttaatat tcttcttggc cttatattta attccctatg caattaatat 3540

tttatatctg catttttta aaaaaaatag atgttatata agtgattctc gtatgtagca 3600

cctgttgctt ttccactgaa agaattacgg attttgtact gtgatttata ttcactgccc 3660

caattcaaga aatattggag ccttgctaca atgtgaaatg ttatagtcat ggactccttc 3720

caaccagatt tctgaaaaca ccagagggat ggtataattc tgtctcacct ataacatggt 3780

cctgtgacat agatattaag accacaagtt gtagtgaggc tacaattata ttcgtctgtc 3840

ttggctttgc aacataattt agaaagcacg tatagttgtt ttttaaccaa gttacataca 3900

atctcatgta ctgatttgag acttataaca atttttggag ggggcataga gaaaggagtg 3960

cccacagttg aggcatgacc ccctccattc agacctctaa ctgttgcctg agtacacaga 4020

tgtgccctga tttctggccc attggccata gtactgtgcc taatcaatgt aataggttta 4080

ttttcccaat cctcaaacta aaaatgttca taacaagatg aattgtagac tagtaacatt 4140

tgatgctttt aaatatttgc ttctttttaa acaaaaacta aaacccagaa gtgaattttt 4200

aggtggattt ttaaataaaa aagattgatt gagtttggtg tgcaagctgt tttataatga 4260

aacaacaaaa tgaaatctaa aatcctgaaa tgtgcctaaa ctatcaaaac acacgataca 4320

gctaatgtgt aaagatgcta aattctgtta cttggaggat gaatatattt aagatttaaa 4380

acacaataat aaatacatga ttaattcaaa aataaaaatc tttacagctg cctatcaagg 4440

gtctaaagca cttaatgaat gtttttagtc taacttatca ttaacttttt acaagtcacc 4500

atatttgaag atctgtagca ctctgatttt cagaaaattt ttcattctga ataatttaaa 4560

aatggtgatg tattagaaag gcagtttgct ttagaaaact aaatcacatt gaacattgta 4620

ttagagaatt aaattaaaag tttcttacag agcagtattt tccaaacatt tttagcacta 4680

gaatcttttt agatgaaatt ttatgtataa ccccaataca taaagcctga aaactcaatt 4740

ttatcaatat aaatgtattt tgggttcaca tttatgctta ttcattttgg ctcattacta 4800

agcataataa gattctgagt tatttctgaa taacacaaat gtggagttat acatagttga 4860

tgaaaccagc agccaattta tagctatgcc ctgttttatt tgtatactat caagaaaatt 4920

ttgattcaca caaatgtaag caaaaataat aggttttaaa catacatctc aggaaattct 4980

ttaattagag atagctaaag ttattcaagg tctatacaaa aataagttat cctggtagtg 5040

gaagttaata cataagcagt ctccagtgtg gtaaagtagg gtatgtaaca catcagaatg 5100

tgcgttttta ttaggtttta aaatatgcac gtataaaaac taaatttgaa tcaaaccctt 5160

ttaactcacc tccaagaagc tagactttgg ccaggaatgg gctaaaaacc actggttaac 5220

gatgtgacag ttatgatctt ggagattgga aatctttctt ccacattaga gttctttacc 5280

ttaattcctt attctgaaaa attgtaagat tttatgaagg tttgaatact gaagcacagt 5340

tctgctttca aaaattaaaa ttcaaacttg aaaaagctgt ttaacccatg gaagatatca 5400

tttagtaaga tgtaaaagat tttttaaatc tacacttcag tttatacatc tttatcatta 5460

tcaatactat ataagttact gtgagcattt tagagaattc cataaaggta ctatgagtgt 5520

gtctgtatgt gtgtgtatat atagcattgt atttaatcat agactaaatt taatttgata 5580

tagaaatact actttacttg tacattaagg tcataatttc tgctggactc ttttatattt 5640

aattaatggg gattatagtc ttccttcata aatgcattta aacctgaaat tgaacaccag 5700

tgtttttctt tttctactta tgggaagttg tctgcttccc cctttagaga aaacagtatt 5760

tttatatttt gttaaaatat taactacttt atgcctacac actatgctgt agatactgat 5820

cataattctt gggtgttcac aaacactcct agtgcctctt ttttggcccg ttgaaagtgt 5880

tggtattact actttcacta cagagccttt ggccctctaa taatgctgag gtgggctgat 5940

ccttcccatt tctgtcttcg ggtcattctg gtaggtcttc tcctccactg tcaagtaagc 6000

aatcaggtcc gtgacaggga ttggacatat gaacaaatta agtggataca cacagtgaga 6060

aagatacatg cattctatgg taacaactac tgtcaataac atctgatgtt acatgcacat 6120

ttatatatat ataatttaa aaactgaact atgagaagcc atggtataaa tgaatattgt 6180

ggacatcatg gacttgatat gatagaaatc aattgtcagc ttgagaaagt tgtttttaat 6240

ctgtctaaat agttcatgca ttactacagt taaaaatagt ttcatttgtc ttctatagac 6300

ttaattttat tccggttcag tataatctct gttaacagag tttcagcaaa ctgattggtc 6360

aaggtattaa catagcttct acttccttta cttaaaaaga tgtggtttta tgtaagttct 6420

tgattactga tgatcatccc aaattttgac aacaaaatca tatgtataaa tttatttctc 6480

ccctcttgtt catcatcttt tgtaaaggtc ccattgtaga tcttttctgc taccaaataa 6540

aacttttcaa acaatttggt ttcaagacct taaatagaca agttggatac taagattgtg 6600

aactgataag gacatataaa tttatatttc cagcccttcc ttagagtctt tatctgcatc 6660

aaaaacccaa ttctgccatt aactgtgctt cccagtccca cctctatatg tcactcattt 6720

tctgcaacaa agatctcact aaatcatgtt gaaacacaag tcatgatcct ctctaagtaa 6780

atagaaaaag ctccctggaa aaactctgtt gccacatgca cgtgccctgt tactcctcca 6840
gccagccagt gctgccagca ttttattgtg taaaagtcca aataaataag ggcctgcatg 6900
caacctttat cttcagaaac taggttttat atgtaaaatg tgacttggga aatgattctg 6960
tttattaact ggctgggatt tttcatttct atgaaagttt caaacatctc cagtacttta 7020
taaaatccca acaattgctg taagtcagca ctttggtcca ctcagcccac ccagcccact 7080
tgcaactctg actcttcact gaatcatatt tgggaagttt gggtagggtg aggctatctt 7140
cttcaagatt attttctcat atgtctgtct gtcaccttgt aaaccatgag actcctgggt 7200

atttgcatgt aacttctttg aggaagttac caccatctct gatatagaca cacttttga 7260
gttgcagttt ctgttagaat tttttggaga ctaacttgcc aattctgtga atgttattga 7320
atatttaaaa agctgggtct gtaatgggag gcattttatt agctgttgtg attgggtaac 7380
atgtcccctt agatttcctg atttaaaatt atacaaaatt actatttttg ataaaataaa 7440
ggaacaccta cagaaaatta agtttctaag atgtttctat acttcattag aaaagatttt 7500
attactatta cttatggtta ttggtgatta acacttaatg cgtctcctct gattttgtgt 7560
tccatgaggt gcttggaaca tttggagtgc tctgtgcgag ggacatacag tgatatagga 7620

aatttaaaaa ttaaaataat acccaaaacc cactttatca gatatggtat tgtgatggtt 7680
aatattatgt gtcaacttgg tgaggctatg gcgcccatgt gtttggtcaa acactagcct 7740
agatgttgct gtgaatatat tttgtagatg tgattaacat ttacaatcag ttgattttaa 7800
gtaaagcaga ttctcatcca aaaaaaaaaa aaaaaa 7836

<210> 14
<211> 894
<212> PRT
<213> Homo sapiens
<400> 14

Met Glu Ser Thr Leu Ser Ala Ser Asn Met Gln Asp Pro Ser Ser Ser
1 5 10 15
Pro Leu Glu Lys Cys Leu Gly Ser Ala Asn Gly Asn Gly Asp Leu Asp

20 25 30
Ser Glu Glu Gly Ser Ser Leu Glu Glu Thr Gly Phe Asn Trp Gly Glu
35 40 45
Tyr Leu Glu Glu Thr Gly Ala Ser Ala Ala Pro His Thr Ser Phe Lys

50 55 60

His Val Glu Ile Ser Ile Gln Ser Asn Phe Gln Pro Gly Met Lys Leu

65 70 75 80

Glu Val Ala Asn Lys Asn Asn Pro Asp Thr Tyr Trp Val Ala Thr Ile

85 90 95

Ile Thr Thr Cys Gly Gln Leu Leu Leu Leu Arg Tyr Cys Gly Tyr Gly

100 105 110

Glu Asp Arg Arg Ala Asp Phe Trp Cys Asp Val Val Ile Ala Asp Leu

115 120 125

His Pro Val Gly Trp Cys Thr Gln Asn Asn Lys Val Leu Met Pro Pro

130 135 140

Asp Ala Ile Lys Glu Lys Tyr Thr Asp Trp Thr Glu Phe Leu Ile Arg

145 150 155 160

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

165 170 175

Leu Arg Gly Lys Gly Pro Ile Asp Leu Ile Thr Val Gly Ser Leu Ile

180 185 190

Glu Leu Gln Asp Ser Gln Asn Pro Phe Gln Tyr Trp Ile Val Ser Val

195 200 205

Ile Glu Asn Val Gly Gly Arg Leu Arg Leu Arg Tyr Val Gly Leu Glu

210 215 220

Asp Thr Glu Ser Tyr Asp Gln Trp Leu Phe Tyr Leu Asp Tyr Arg Leu

225 230 235 240

Arg Pro Val Gly Trp Cys Gln Glu Asn Lys Tyr Arg Met Asp Pro Pro

245 250 255

Ser Glu Ile Tyr Pro Leu Lys Met Ala Ser Glu Trp Lys Cys Thr Leu

260 265 270

Glu Lys Ser Leu Ile Asp Ala Ala Lys Phe Pro Leu Pro Met Glu Val

275 280 285

Phe Lys Asp His Ala Asp Leu Arg Ser His Phe Phe Thr Val Gly Met

290 295 300

Lys Leu Glu Thr Val Asn Met Cys Glu Pro Phe Tyr Ile Ser Pro Ala
305 310 315 320
Ser Val Thr Lys Val Phe Asn Asn His Phe Phe Gln Val Thr Ile Asp
325 330 335
Asp Leu Arg Pro Glu Pro Ser Lys Leu Ser Met Leu Cys His Ala Asp

340 345 350
Ser Leu Gly Ile Leu Pro Val Gln Trp Cys Leu Lys Asn Gly Val Ser
355 360 365
Leu Thr Pro Pro Lys Gly Tyr Ser Gly Gln Asp Phe Asp Trp Ala Asp
370 375 380
Tyr His Lys Gln His Gly Ala Gln Glu Ala Pro Pro Phe Cys Phe Arg
385 390 395 400
Asn Thr Ser Phe Ser Arg Gly Phe Thr Lys Asn Met Lys Leu Glu Ala

405 410 415
Val Asn Pro Arg Asn Pro Gly Glu Leu Cys Val Ala Ser Val Val Ser
420 425 430
Val Lys Gly Arg Leu Met Trp Leu His Leu Glu Gly Leu Gln Thr Pro
435 440 445
Val Pro Glu Val Ile Val Asp Val Glu Ser Met Asp Ile Phe Pro Val
450 455 460
Gly Trp Cys Glu Ala Asn Ser Tyr Pro Leu Thr Ala Pro His Lys Thr

465 470 475 480
Val Ser Gln Lys Lys Arg Lys Ile Ala Val Val Gln Pro Glu Lys Gln
485 490 495
Leu Pro Pro Thr Val Pro Val Lys Lys Ile Pro His Asp Leu Cys Leu
500 505 510
Phe Pro His Leu Asp Thr Thr Gly Thr Val Asn Gly Lys Tyr Cys Cys
515 520 525
Pro Gln Leu Phe Ile Asn His Arg Cys Phe Ser Gly Pro Tyr Leu Asn

530 535 540
Lys Gly Arg Ile Ala Glu Leu Pro Gln Ser Val Gly Pro Gly Lys Cys

545 550 555 560

Val Leu Val Leu Lys Glu Val Leu Ser Met Ile Ile Asn Ala Ala Tyr

565 570 575

Lys Pro Gly Arg Val Leu Arg Glu Leu Gln Leu Val Glu Asp Pro His

580 585 590

Trp Asn Phe Gln Glu Glu Thr Leu Lys Ala Lys Tyr Arg Gly Lys Thr

595 600 605

Tyr Arg Ala Val Val Lys Ile Val Arg Thr Ser Asp Gln Val Ala Asn

610 615 620

Phe Cys Arg Arg Val Cys Ala Lys Leu Glu Cys Cys Pro Asn Leu Phe

625 630 635 640

Ser Pro Val Leu Ile Ser Glu Asn Cys Pro Glu Asn Cys Ser Ile His

645 650 655

Thr Lys Thr Lys Tyr Thr Tyr Tyr Tyr Gly Lys Arg Lys Lys Ile Ser

660 665 670

Lys Pro Pro Ile Gly Glu Ser Asn Pro Asp Ser Gly His Pro Lys Pro

675 680 685

Ala Arg Arg Arg Lys Arg Arg Lys Ser Ile Phe Val Gln Lys Lys Arg

690 695 700

Arg Ser Ser Ala Val Asp Phe Thr Ala Gly Ser Gly Glu Glu Ser Glu

705 710 715 720

Glu Glu Asp Ala Asp Ala Met Asp Asp Asp Thr Ala Ser Glu Glu Thr

725 730 735

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

740 745 750

Pro Ser Ala Arg Pro Arg Arg Ala Val Thr Leu Arg Ser Gly Ser Glu

755 760 765

Pro Val Arg Arg Pro Pro Pro Glu Arg Thr Arg Arg Gly Arg Gly Ala

770 775 780

Pro Ala Ala Ser Ser Ala Glu Glu Gly Glu Lys Cys Pro Pro Thr Lys

785 790 795 800

Pro Glu Gly Thr Glu Asp Thr Lys Gln Glu Glu Glu Glu Arg Leu Val
805 810 815

Leu Glu Ser Asn Pro Leu Glu Trp Thr Val Thr Asp Val Val Arg Phe
820 825 830

Ile Lys Leu Thr Asp Cys Ala Pro Leu Ala Lys Ile Phe Gln Glu Gln
835 840 845

Asp Ile Asp Gly Gln Ala Leu Leu Leu Leu Thr Leu Pro Thr Val Gln

850 855 860

Glu Cys Met Glu Leu Lys Leu Gly Pro Ala Ile Lys Leu Cys His Gln
865 870 875 880

Ile Glu Arg Val Lys Val Ala Phe Tyr Ala Gln Tyr Ala Asn
885 890

<210> 15

<211> 7922

<212> DNA

<213> Homo sapiens

<400> 15

cgccttgtgt gtgctggatc ctgcgcgggt agatccccga gtaatttttt ctgcaggatg 60

aattaagaga agagacactt gctcatcagg catggagagc actttgtcag cttccaatat 120

gcaagaccct tcatcttcac ccttggaaaa gtgtctcggc tcagctaatg gaaatggaga 180

ccttgattct gaagaaggct caagcttgga ggaaactggc tttaactggg gagaatattt 240

ggaagagaca ggagcaagtg ctgctcccca cacatcattc aaacacgttg aaatcagcat 300

tcagagcaac ttccagccag gaatgaaatt ggaagtggct aataagaaca acccggacac 360

gtactgggtg gccacgatca ttaccacgtg cgggcagctg ctgcttctgc gctactgcgg 420

ttacggggag gaccgcaggg ccgacttctg gtgtgacgta gtcatcgcgg atttgcaccc 480

cgtggggtgg tgcacacaga acaacaaggt gttgatgccg ccggacgcaa tcaaagagaa 540

gtacacagac tggacagaat ttctcatacg tgacttgact ggttcgagga cagcacccgc 600

caacctcctg gaaggtcctc tgcgagggaa aggccctata gacctcatta cagttggttc 660

cttaatagaa cttcaggatt cccagaaccc ttttcagtac tggatagtta gtgtgattga 720

aaatgttgga ggaagattac gccttcgcta tgtgggattg gaggacactg aatcctatga 780

ccagtggttg ttttacttgg attacagact tcgaccagtt ggttggtgtc aagagaataa 840

atacagaatg gacccacctt cagaaatcta tcctttgaag atggcctctg aatggaaatg 900

tactctggaa aaatccctta ttgatgctgc caaatttcct cttccaatgg aagtgtttaa 960

ggatcacgca gatttgcgaa gccatttctt cacagttggg atgaagcttg agacagtgaa 1020

tatgtgcgag ccctttttaca tctctcctgc gtcggtgact aaggttttta acaatcactt 1080

ttttcaagtg actattgatg acctaagacc tgaaccaagt aaactgtcaa tgctgtgcca 1140

tgcagattct ttggggattt tgccagtaca gtggtgcctt aaaaatggag tcagcctcac 1200

tcctcccaaa ggttactctg gccaggactt cgactgggca gattatcaca agcagcatgg 1260

ggcgcaggaa gcccctccct tctgcttccg aaatacatca ttcagtcgag gtttcacaaa 1320

gaacatgaaa cttgaagctg tgaaccccag gaatccagga gaactgtgtg tggcctccgt 1380

tgtgagtgtg aaggggcggc taatgtggct tcacctggaa gggctgcaga ctcctgttcc 1440

agaggtcatt gttgatgtgg aatccatgga catcttccca gtgggctggt gtgaagccaa 1500

ttcttatcct ttgactgcac cacacaaaac agtctcacaa aagaagagaa agattgcagt 1560

cgtgcaacca gagaaacaat tgccgcccac agtgcctgtt aagaaaatac ctcatgacct 1620

ttgtttattc cctcacctgg acaccacagg aaccgtcaac gggaaatact gctgtcctca 1680

gctcttcatc aaccacaggt gtttctcagg cccttacctg aacaaaggaa ggattgcaga 1740

gctacctcag tcggtgggac cgggcaaatg cgtgctggtt cttaaagagg ttcttagcat 1800

gataatcaac gcagcctaca agcctggaag ggtattaaga gaattacagc tggtagaaga 1860

tccccactgg aatttccagg aagagacgct gaaggccaaa tacagaggca aaacatacag 1920

ggctgtggtc aaaatcgtac ggacatctga ccaagtcgca aatttctgcc gccgagtctg 1980

tgccaagcta gagtgctgtc caaatttgtt tagtcctgtg ctgatatctg aaaactgccc 2040

agagaactgc tccattcata ccaaaaccaa atacacctat tactatggaa agagaaagaa 2100

gatctccaag ccccccatcg gggaaagcaa ccccgacagc ggacacccca aacccgccag 2160

gcggaggaag cgacggaaat ccattttcgt gcagaagaaa cggaggtctt ctgccgtgga 2220

cttcaccgcg ggctcggggg aggaaagtga agaggaggac gctgacgcca tggacgatga 2280

caccgccagt gaggagaccg gctccgagct ccgggatgac cagacggaca cctcgtcggc 2340

ggaggtgccc tcggcccggc cccggagggc cgtcaccctg cggagcggct cagagcccgt 2400

gcgccggcca cccccagaga ggacacgaag gggccgcggg gcgccggctg cctcctcagc 2460

agaggaaggg gagaagtgcc cgccgaccaa gcccgagggg acagaggaca cgaaacagga 2520

ggaggaggag agactggttc tggagagcaa cccgttggag tggacggtca ccgacgtggt 2580

gaggttcatt aagctgacag actgtgcccc cttggccaag atatttcagg agcaggatat 2640

tgacggccaa gcactcctgc ttctgaccct tccgacggtg caggagtgca tggagctgaa 2700

gctggggcct gccatcaagt tatgccacca gatcgagaga gtcaaagtgg ctttctacgc 2760

ccagtacgcc aactgagtct gccctcggga ggtggcccat tattgctggg atgcggtgtt 2820

ggtaaaggtt tccaggactg aaactttgat tttccgggat atgttaaatg gtacagccac 2880

taagtatcac cagaaaacca gaagcccagg atcttctgcc tccgccagcc tgtgagctgt 2940

ttccatgttt tcaaagcaca gcagcagtcg cttctgggga gtgccagtta aagtcatgca 3000

tcagaccctg ccagacgtgg gcctgcttct tggctcaccc acgttttgcc tttctcctgc 3060

cccaaatcag gcagctccct tggagcaggg tttcctcaga tgaggactgc attctttgaa 3120

aacaaagaat gtcgccaagg aagaaacctc acgccatgct gtagtgtttc ctgtaatcac 3180

acgagcacat ttatatatgc agtttcccat ggataggcgt gtgaccctgg ttgagtggca 3240

cttgcggttt catcttggtg gcaactcctt tgcaatgcag ctggcagcga catccttata 3300

aaaacatgtg ctaaagctct gtcctctgtt agaggtgcct tttaggaata cggggagtga 3360

aggaaggccg gcaggcatct ccatgcaact agatggtttg tttgtttgtt tgtttgtttg 3420

ttgttcattt tgttgtgttt tttgagacag ggtcttgctc tgtcgcccag gttgtaatgc 3480

agtggcgcaa tctcagctca ctgcaacctc tctctcccgg gttcaagtga ttctcctgcc 3540

tcagcctccc aagtagctgg gattacaggc acccaccacc atgcctggct aatttttgta 3600

tttttggtag agacagggtt tcaccatgtt ggtcaggcta gtcttgaact cccaacctca 3660

agtgatctgc ccgcctcggc ctcccaacgt gctgggatta caggtgtgag ccactacgcc 3720

ccggcccaac tggatggttt ttgattgaag cctagaacat ctgtagagac aaactctacc 3780

cagtcttttc tagaccctca actatctcca gtgttgttgt ttaatcgtag ccggatcagg 3840

gagtgagtct tttaggcaaa tgttggatta tatatcaaag gaaaagctta gtttcagaga 3900

ggaggaaggg aaagagatgt gagggaagca tttcatcaac cagctacgtc ccccttagaa 3960

ggatcactgc agcaggtcac cgagcaggag tccctctgag cgtcccttct gtctcgttct 4020

gccctagctg gcagcatatg aaccaggcat gatgcagcag gagcagtgaa tctggagtca 4080

gccacttggc accctggttt cgctgagaac aaactctgag atcttgggtg acttctcatc 4140

actctggacc tccattcctg tgaagtgaca ggtgtggacc ctgagggtgc ggtggtgagc 4200

acactgtctc ctgctggcat tcaccccact catgctggaa aggaagatcc agatcgtaca 4260

aaaattagaa aaagaaagaa taagaagggt ctggtcccag ttctgactcg gccattctta 4320

cagctctttc tggctttgag tttgcttgtg gaatttcctg ggcagttgtg ttaaatccgc 4380

caggtcacgt gcagacaaag ctgtggctgc gagagttggc tggcctcttg gaccagaagc 4440

catctccata tcctcatgag cgattccata tctccactca gaccctgtgg actacagtgt 4500

tccgctgtgg tggctgccaa gatgccttct taaacttatg caaggaaacc aaaccctccc 4560

acagttccca agcagacact ggaagcagag gcttctcacc cttcctgctt tttcaccaca 4620

atcaccttga gctcgtccct tggactagag tctccacagt tccagtaaaa ttctgcggtg 4680

ggctgatgag ctgcttgcat ttctgtgaca tttccagata tgattctcag tgggattttg 4740

gaaactttga ttgctcaagc tcacccttct taacattctg taatggttac agatgagaat 4800

ggaaaacaca tattttatgg atgaggcgtt ttggtctccc ctgcagtcga tttctagaat 4860

caagttttag agttcggctg atgcatctgc ctggggacct cagatgggag gagtgtgtca 4920

gttgtacccc gacagaaatg tctctgggat ctgtggctgg cttgccccgg gcatctctcc 4980

tttaagctca agttttgaac tctctgcggt tttccacccc tgccttctca gccacatgct 5040

tttggcctta aacgctcagt cttgtggagt tcaactctgt caaacgattg gaaagggcat 5100

ccatttccag atctttggca ttttccccgc gctgactctt tgatgatcct tcactgtggc 5160

cttttcaagc tcagctgttc ctgttgtatt tgagacgagg gtgagggaat gtggtggcca 5220

caaaagaaca gggacttgca gcacaaatgt cacttctgtc tccctttca gtggtagcac 5280

ggaggaggag gtgctgcgtt ggagggaggg gatcctccag gagctctctg gagcccatct 5340

aggaagctag agtgtgtggc ccgccaggag ctcaggaagg atacagccac tgtcgcaggg 5400

gaaagtgttt gcttcccgtg gagccaagcg cccaagactc tccgtatcct tcaccctgac 5460

agtttaactt cagcgtttct ctgtgcagtt gcggtcacca tgggtgagca ctgtctgtgc 5520

acgtgccagg gaggagatgg ctgggaccac tgcacaggag ggcgcagcct ggcgtcgcca 5580

tgaaagttgt ctctgtgcca tctctccggt ccttgaggag agcccagaaa gattttagga 5640

cccaggaggt gcttttcctc cagctgttgc cagtgtcctt ctgagcctgg attctccggg 5700

gatttccgtc gtggtggatg gacttcacat cagcagcagt tctggtacag aattgtaatg 5760

tgttttcatt tctctgtagg attcacctct caccagcgtc tgtcttaaag gtagggccaa 5820

tttcatggag cattttctg tgtgtgtcct tgttgctttt gccagaaaaa gtggatttga 5880

catgcgtgcc ccgatgccac catagcccct aggccaacaa tgtcatggtc taaacaccaa 5940

aaagtgatgc cccgcattcc ttccctggat ggtaccgttt cttctccgtc tctctttgat 6000

gattctttgg gaccaaagtc ctctccttag tgcgcctact tcctgtgggc atcatgccac 6060

ttggaactta ttggaactgg cccgggagac tctgcagtct gcgccgtttg aaaaccctga 6120

gaaagagatg ccacctcaac ttgaatcatg acagcccatc gctcagtctc accctaaact 6180

catggagctt gtttcagctc ctcacttctt gactgtattt gtactatgtt gaaaaaatat 6240

cctgtccaca aagacataag cctaacaacc tagaaaaaca acagggtact actggcatta 6300

cagaacttct ttgcctttca aaacaaaagc aaaacacagt gaacttcacc acggagctgc 6360

acagcgtggg gaactcatcc atcactttca aaattagagt catttgatcc aagttggagt 6420

cagacacagt atttgagctg cacggcttct gggttctccc accttatttg atcatattcg 6480
aaagattatt tcctgtgttt gctttgattt gttcctcagt acattaaaat gatccacacc 6540
ttgaacactg ccctctctag aaggttgatt ttgatcagcc ttttgaagat gggtgtcgtt 6600
tccctaactt atctcacaga attttgagtg ttgtatttgg caagttctga gatttgcctt 6660
ctgtcttatg ccaaacaccc ctttctaaga gctgtccccg cttagtttta gaagtactag 6720
gggttttcat acttatttta tagaacaccc atttatattt atttctgtat atagaactaa 6780
aaaaaacagt agtgttaaaa atctttgttg tggtttgagc atctttgctg cttttggatt 6840

gagatggcga atcaaggctt cacttcctct ctcttctgtc tttagaaagc tgtgatcgtg 6900
cgtgcaatta tttgaaaggc aacatagtca attaagaaac ctgtagttgt taaggaagaa 6960
attgttggca agatatccat actgcccata tctcgttggt gcaataatta aatagcaaag 7020
gaaatctgta ttggcaacta ttataattca ataattcttt tgtttactgc ccttttctgt 7080
tcaagaattt tctggaaatt actccctttc acatggttga actcttaagt tgaccagttc 7140
tcatagctct atcactagaa tggtttgcag ataccccaaa catactatga taaaatcaaa 7200
ttgtgctact tttgacccat gtaatttacc taaaagttgt aattgctgac agagtactgc 7260

cttgaatttt ggtttaaaac ctctctagtt tcaatgacaa gtaacaactc aaataattcc 7320
atattgtttg aggaagaggc cataatcctt ctgaattgtt ggcactaagt aatgggattt 7380
ggcccagtaa gtatgacggt cgtgtcgcct aaccaacgca gagcagtgct ttttgtgtgg 7440
ctgaagcgat gtgctgacga aaaaaggaaa attctaggac aatcgttggc taaaaatcac 7500
cttaggatga aaaatttgag gcaaatttttt ttaaatgaca gaaaaagata atcatctcac 7560
ttgcttgaaa caggagccag catgatctct ggaagcatca actatccctc gtcgtgattg 7620
ttgaaagctc tttcactgtt ttgcattcta gtttgaatag tttgtattga aattggattc 7680

ctatcttgtg tatgtttttg gtgcgtaaaa gggaaaaatt ggtgtcatta cttttgaaat 7740
ttgcaggacg aagggcatgc ttttggtttg ctgtaagatt gtattctgta tatatgtttt 7800
catgtaaata aatgaaaatc tatatcagag ttatatttta atttttattc taaatgaaaa 7860
aaaccctttt tacttcaaaa aaattgtaag ccacattgtt aataaagtaa aaataaattc 7920
ta 7922

<210> 16

<211> 435

<212> PRT

<213> Homo sapiens

<400> 16

Met Leu Pro Ala Arg Cys Ala Arg Leu Leu Thr Pro His Leu Leu Leu

1 5 10 15

Val Leu Val Gln Leu Ser Pro Ala Arg Gly His Arg Thr Thr Gly Pro

20 25 30

Arg Phe Leu Ile Ser Asp Arg Asp Pro Gln Cys Asn Leu His Cys Ser

35 40 45

Arg Thr Gln Pro Lys Pro Ile Cys Ala Ser Asp Gly Arg Ser Tyr Glu

50 55 60

Ser Met Cys Glu Tyr Gln Arg Ala Lys Cys Arg Asp Pro Thr Leu Gly

65 70 75 80

Val Val His Arg Gly Arg Cys Lys Asp Ala Gly Gln Ser Lys Cys Arg

85 90 95

Leu Glu Arg Ala Gln Ala Leu Glu Gln Ala Lys Lys Pro Gln Glu Ala

100 105 110

Val Phe Val Pro Glu Cys Gly Glu Asp Gly Ser Phe Thr Gln Val Gln

115 120 125

Cys His Thr Tyr Thr Gly Tyr Cys Trp Cys Val Thr Pro Asp Gly Lys

130 135 140

Pro Ile Ser Gly Ser Ser Val Gln Asn Lys Thr Pro Val Cys Ser Gly

145 150 155 160

Ser Val Thr Asp Lys Pro Leu Ser Gln Gly Asn Ser Gly Arg Lys Asp

165 170 175

Asp Gly Ser Lys Pro Thr Pro Thr Met Glu Thr Gln Pro Val Phe Asp

180 185 190

Gly Asp Glu Ile Thr Ala Pro Thr Leu Trp Ile Lys His Leu Val Ile

195 200 205

Lys Asp Ser Lys Leu Asn Asn Thr Asn Ile Arg Asn Ser Glu Lys Val

210 215 220

Tyr Ser Cys Asp Gln Glu Arg Gln Ser Ala Leu Glu Glu Ala Gln Gln

225 230 235 240

Asn Pro Arg Glu Gly Ile Val Ile Pro Glu Cys Ala Pro Gly Gly Leu
245 250 255
Tyr Lys Pro Val Gln Cys His Gln Ser Thr Gly Tyr Cys Trp Cys Val

260 265 270
Leu Val Asp Thr Gly Arg Pro Leu Pro Gly Thr Ser Thr Arg Tyr Val
275 280 285
Met Pro Ser Cys Glu Ser Asp Ala Arg Ala Lys Thr Thr Glu Ala Asp
290 295 300
Asp Pro Phe Lys Asp Arg Glu Leu Pro Gly Cys Pro Glu Gly Lys Lys
305 310 315 320
Met Glu Phe Ile Thr Ser Leu Leu Asp Ala Leu Thr Thr Asp Met Val

325 330 335
Gln Ala Ile Asn Ser Ala Ala Pro Thr Gly Gly Gly Arg Phe Ser Glu
340 345 350
Pro Asp Pro Ser His Thr Leu Glu Glu Arg Val Val His Trp Tyr Phe
355 360 365
Ser Gln Leu Asp Ser Asn Ser Ser Asn Asp Ile Asn Lys Arg Glu Met
370 375 380
Lys Pro Phe Lys Arg Tyr Val Lys Lys Lys Ala Lys Pro Lys Lys Cys

385 390 395 400
Ala Arg Arg Phe Thr Asp Tyr Cys Asp Leu Asn Lys Asp Lys Val Ile
405 410 415
Ser Leu Pro Glu Leu Lys Gly Cys Leu Gly Val Ser Lys Glu Val Gly
420 425 430
Arg Leu Val
435

<210> 17
<211> 3740
<212> DNA
<213> Homo sapiens
<400> 17
ataacgggaa ttcccatggc ccgggctcag gcgtccaacc tgctgccgcc tgggccccgc 60

cgagcggagc tagcgccgcg cgcagagcac acgctcgcgc tccagctccc ctcctgcgcg 120

gttcatgact gtgtcccctg accgcagcct ctgcgagccc ccgccgcagg accacggccc 180
gctccccgcc gccgcgaggg ccccgagcga aggaaggaag ggaggcgcgc tgtgcgcccc 240
gcggagcccg cgaaccccgc tcgctgccgg ctgcccagcc tggctggcac catgctgccc 300
gcgcgctgcg cccgcctgct cacgccccac ttgctgctgg tgttggtgca gctgtcccct 360
gctcgcggcc accgcaccac aggccccagg tttctaataa gtgaccgtga cccacagtgc 420
aacctccact gctccaggac tcaacccaaa cccatctgtg cctctgatgg caggtcctac 480
gagtccatgt gtgagtacca gcgagccaag tgccgagacc cgaccctggg cgtggtgcat 540

cgaggtagat gcaaagatgc tggccagagc aagtgtcgcc tggagcgggc tcaagccctg 600
gagcaagcca agaagcctca ggaagctgtg tttgtcccag agtgtggcga ggatggctcc 660
tttacccagg tgcagtgcca tacttacact gggtactgct ggtgtgtcac cccggatggg 720
aagcccatca gtggctcttc tgtgcagaat aaaactcctg tatgttcagg ttcagtcacc 780
gacaagccct tgagccaggg taactcagga aggaaagtct cctttcgatt ctttttaacc 840
ctcaattcag atgacgggtc taagccgaca cccacgatgg agacccagcc ggtgttcgat 900
ggagatgaaa tcacagcccc aactctatgg attaaacact tggtgatcaa ggactccaaa 960

ctgaacaaca ccaacataag aaattcagag aaagtctatt cgtgtgacca ggagaggcag 1020
agtgccctgg aagaggccca gcagaatccc cgtgagggta ttgtcatccc tgaatgtgcc 1080
cctgggggac tctataagcc agtgcaatgc caccagtcca ctggctactg ctggtgtgtg 1140
ctggtggaca cagggcgccc gctgcctggg acctccacac gctacgtgat gcccagttgt 1200
gagagcgacg ccagggccaa gactacagag gcggatgacc ccttcaagga cagggagcta 1260
ccaggctgtc cagaagggaa gaaaatggag tttatcacca gcctactgga tgctctcacc 1320
actgacatgg ttcaggccat taactcagca gcgcccactg gaggtgggag gttctcagag 1380

ccagacccca gccacaccct ggaggagcgg gtagtgcact ggtatttcag ccagctggac 1440
agcaatagca gcaacgacat taacaagcgg gagatgaagc ccttcaagcg ctacgtgaag 1500
aagaaagcca agcccaagaa atgtgcccgg cgtttcaccg actactgtga cctgaacaaa 1560
gacaaggtca tttcactgcc tgagctgaag ggctgcctgg gtgttagcaa agaagtagga 1620
cgcctcgtct aaggagcaga aaacccaagg gcaggtggag agtccaggga ggcaggatgg 1680
atcaccagac acctaacctt cagcgttgcc catggccctg ccacatcccg tgtaacataa 1740
gtggtgccca ccatgtttgc acttttaata actcttactt gcgtgttttg tttttggttt 1800

cattttaaaa caccaatatc taataccaca gtgggaaaag gaaagggaag aaagacttta 1860

ttctctctct tattgtaagt ttttggatct gctactgaca acttttagag ggttttgggg 1920

gggtggggga gggtgttgtt ggggctgaga agaaagagat ttatatgctg tatataaata 1980

tatatgtaaa ttgtatagtt cttttgtaca ggcattggca ttgctgtttg tttatttctc 2040

tccctctgcc tgctgtgggt ggtgggcact ctggacacat agtccagctt tctaaaatcc 2100

aggactctat cctgggccta ctaaacttct gtttggagac tgacccttgt gtataaagac 2160

gggagtcctg caattgtact gcggactcca cgagttcttt tctggtggga ggactatatt 2220

gccccatgcc attagttgtc aaaattgata agtcacttgg ctctcggcct tgtccaggga 2280

ggttgggcta aggagagatg gaaactgccc tgggagagga agggagtcca gatcccatga 2340

atagcccaca caggtaccgg ctctcagagg gtccgtgcat tcctgctctc cggacccca 2400

aagggcccag cattggtggg tgcaccagta tcttagtgac cctcggagca aattatccac 2460

aaaggatttg cattacgtca ctcgaaacgt tttcatccat gcttagcatc tactctgtat 2520

aacgcatgag aggggaggca aagaagaaaa agacacacag aagggccttt aaaaaagtag 2580

atatttaata tctaagcagg ggaggggaca ggacagaaag cctgcactga ggggtgcggt 2640

gccaacaggg aaactcttca cctccctgca aacctaccag tgaggctccc agagacgcag 2700

ctgtctcagt gccaggggca gattgggtgt gacctctcca ctcctccatc tcctgctgtt 2760

gtcctagtgg ctatcacagg cctgggtggg tgggttgggg gaggtgtcag tcaccttgtt 2820

ggtaacacta aagttgtttt gttggttttt taaaaaccca atactgaggt tcttcctgtt 2880

ccctcaagtt ttcttatggg cttccaggct ttaagctaat tccagaagta aaactgatct 2940

tgggtttcct attctgcctc ccctagaagg gcaggggtga taacccagct acagggaaat 3000

cccggcccag ctttccacag gcatcacagg catcttccgc ggattctagg gtgggctgcc 3060

cagccttctg gtctgaggcg cagctccctc tgcccaggtg ctgtgcctat tcaagtggcc 3120

ttcaggcaga gcagcaagtg gcccttagcg ccccttccca taagcagctg tggtggcagt 3180

gagggaggtt gggtagccct ggactggtcc cctcctcaga tcacccttgc aaatctggcc 3240

tcatcttgta ttccaacccg acatccctaa aagtacctcc acccgttccg ggtctggaag 3300

gcgttggcac cacaagcact gtccctgtgg gaggagcaca accttctcgg gacaggatct 3360

gatggggtct tgggctaaag gaggtccctg ctgtcctgga gaaagtccta gaggttatct 3420

caggaatgac tggtggccct gccccaacgt ggaaaggtgg gaaggaagcc ttctcccatt 3480

agccccaatg agagaactca acgtgccgga gctgagtggg ccttgcacga gacactggcc 3540

ccactttcag gcctggagga agcatgcaca catggagacg gcgcctgcct gtagatgttt 3600

ggatcttcga gatctcccca ggcatcttgt ctcccacagg atcgtgtgtg taggtggtgt 3660

tgtgtggttt tcctttgtga aggagagagg gaaactattt gtagcttgtt ttataaaaaa 3720

taaaaaatgg gtaaatcttg 3740

<210> 18

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 18

ccaggcgctc aagagaataa caatttctgc tcagtcaaca tcaac 45

<210> 19

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 19

ctcttgagcg cctggcgttc ggctgccagg gaccttcatg 40