



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0063333
(43) 공개일자 2019년06월07일

(51) 국제특허분류(Int. Cl.)
C12N 15/113 (2010.01) C12N 5/0793 (2010.01)
C12Q 1/68 (2018.01) G01N 33/50 (2017.01)
(52) CPC특허분류
C12N 15/113 (2013.01)
C12N 5/0619 (2013.01)
(21) 출원번호 10-2017-0162338
(22) 출원일자 2017년11월29일
심사청구일자 없음

(71) 출원인
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
김영준
서울특별시 강남구 논현로160길 31, 2층 2호(신사동, 청록빌라)
조승우
서울특별시 서초구 서초대로65길 13-10, 106동 2304호(서초동, 서초래미안아파트)
(74) 대리인
특허법인다나

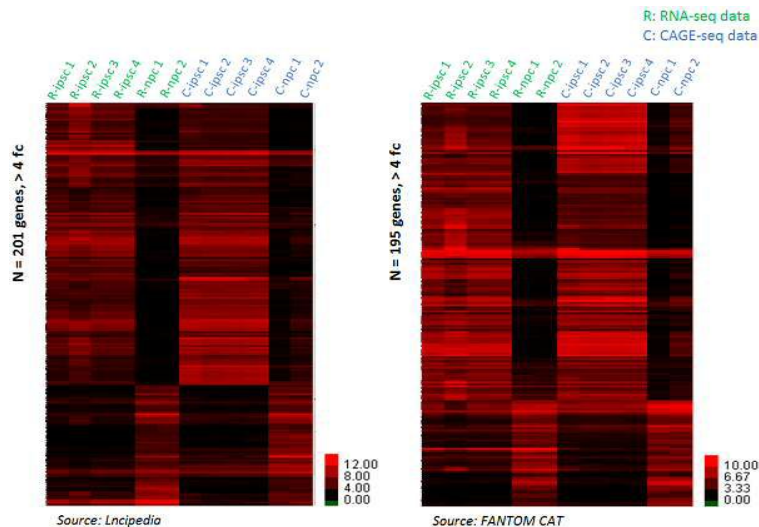
전체 청구항 수 : 총 6 항

(54) 발명의 명칭 유도만능 줄기세포 분화 특이적 lncRNA, 및 이의 용도

(57) 요약

본 발명은 유도만능 줄기세포 신경 분화 특이적 lncRNA에 관한 것으로, 특정 유전자로부터 전사된 lncRNA가 제공된다.

대표도 - 도2



(52) CPC특허분류

C12Q 1/6876 (2018.05)

G01N 33/5073 (2013.01)

C12N 2506/45 (2013.01)

(72) 발명자

황지현

서울특별시 서대문구 연세로 50, 706호(신촌동, 법
현학사)

김민정

서울특별시 은평구 연서로48길 12, 507동 801호(진
관동, 은평뉴타운 제각말)

박현지

경기도 고양시 덕양구 호국로 860, 216동 1901호(
성사동, 래미안휴레스트아파트)

조안나

서울특별시 서대문구 성산로17길 7-13(연희동)

이 발명을 지원한 국가연구개발사업

과제고유번호 2016M3C9A4921712

부처명 과학기술정보통신부

연구관리전문기관 한국연구재단

연구사업명 포스트게놈 다부처유전체 사업

연구과제명 세포 특이적 후성유전체마커 발굴 및 검증(총괄: 국제 협력 기반 암 특이적 후성유전체 마
커 개발)

기 여 율 1/1

주관기관 연세대학교 산학협력단

연구기간 2017.03.01 ~ 2017.12.31

명세서

청구범위

청구항 1

서열번호 1 내지 20으로 이루어진 군에서 선택된 하나 이상의 염기서열을 가지는 유전자로부터 전사된 유도만능 줄기세포 신경 분화 특이적 lncRNA(long non-coding RNA).

청구항 2

제1항의 lncRNA 염기서열, 상기 lncRNA에 대한 안티센스 염기서열, 또는 상기 lncRNA 염기서열의 단편인 유도만능 줄기세포의 분화 예측용 바이오 마커.

청구항 3

제1항의 염기서열, 상기 lncRNA 염기서열의 단편, 상기 lncRNA 염기서열에 특이적으로 결합하는 프라이머 또는 프로브를 포함하는 유도만능 줄기세포의 분화 예측용 키트.

청구항 4

제3항에 있어서,

마이크로어레이, 유전자 증폭 키트, RNA 시퀀싱(RNA sequencing) 또는 면역분석(immunoassay)용 키트인 유도만능 줄기세포의 분화 예측용 키트.

청구항 5

(a) 유도만능 줄기세포 및 시험 체제를 접촉시키는 단계;

(b) 서열번호 1 내지 20으로 이루어진 군에서 선택된 하나 이상의 염기서열을 가지는 유전자로부터 전사된 lncRNA의 발현에 대한 조절 효과를 결정하는 단계; 및

(c) 상기 시험 체제와 접촉된 세포 및 상기 시험 체제와 접촉되지 않은 대조군에서 상기 lncRNA의 발현 수준을 비교하는 단계;를 포함하는 유도만능 줄기세포의 분화 조절제 스크리닝 방법.

청구항 6

제1항의 신경 분화 특이적 lncRNA의 발현을 증가 또는 억제시키는 단계;를 포함하는 유도만능 줄기세포의 분화 제어 방법.

발명의 설명

기술 분야

[0001] 본 발명은 유도만능 줄기세포의 신경 분화 특이적인 lncRNA 및 이의 다양한 용도에 관한 것이다.

배경 기술

[0002] lncRNA는 지금까지 주 연구 대상이었던 단백질을 생성하는 coding DNA는 전체 유전체 중 2% 미만에 불과해 사람과 관련된 질병, 발달 및 생리현상을 모두 설명할 수 없는 한계점이 있었다.

[0003] 유전체의 암흑물질로 남아있는 많은 수의 long non-coding RNA(이하, lncRNA)가 세포 내에 존재할 가능성이 제시되고 그 기능이 유전자 발현 조절에 있다는 사실이 발견됨에 따라, 새로운 lncRNA를 찾아내고 그 조절 기능을 밝히는 것이 매우 중요한 연구로 부상하였다.

[0004] 현재 일부 lncRNA만이 연구되었는데 대부분 세포 분화 및 암 등의 질병 발생 과정에서 주요한 조절 인자로 작용하고 있음이 밝혀졌다.

[0005] lncRNA의 주요 특징으로는 세포 특이적인 발현 양상을 보이면서 세포 분화를 조절하며, 표적 유전자의 발현을

증진시키거나 억제할 수 있는 양방향으로의 조절을 가능하게 한다.

[0006] 또한, lncRNA는 유전자의 발현 조절뿐만 아니라 후성 유전 기전, 전사조절 인자들의 복합체 형성 및 염색체 구조 등에 대한 조절 가능성이 제시되고 있다.

발명의 내용

해결하려는 과제

[0007] 본 발명은 전술한 종래 기술의 문제점을 해결하기 위한 것으로, 본 발명의 목적은 유도만능 줄기세포의 신경 분화 특이적 lncRNA, 및 이의 산업적, 의학적 활용 가능성을 제공한다.

과제의 해결 수단

[0008] 본 발명의 일 측면에 따르면, 서열번호 1 내지 20으로 이루어진 군에서 선택된 하나 이상의 염기서열을 가지는 유전자로부터 전사된 유도만능 줄기세포 신경 분화 특이적 lncRNA(long non-coding RNA)가 제공된다.

[0009] 본 발명의 다른 측면에 따르면, 상기 lncRNA 염기서열, 상기 lncRNA에 대한 안티센스 염기서열, 또는 상기 lncRNA 염기서열의 단편인 유도만능 줄기세포의 분화 예측용 바이오 마커가 제공된다.

[0010] 본 발명의 다른 측면에 따르면, 상기 lncRNA 염기서열, 상기 lncRNA 염기서열의 단편, 상기 lncRNA 염기서열에 특이적으로 결합하는 프라이머 또는 프로브를 포함하는 유도만능 줄기세포의 분화 예측용 키트가 제공된다.

[0011] 일 실시예에 있어서, 상기 유도만능 줄기세포의 분화 예측용 키트는 마이크로어레이, 유전자 증폭 키트, RNA 시퀀싱(RNA sequencing) 또는 면역분석(immunoassay)용 키트일 수 있다.

[0012] 본 발명의 다른 측면에 따르면, (a) 유도만능 줄기세포 및 시험 체제를 접촉시키는 단계; (b) 서열번호 1 내지 20으로 이루어진 군에서 선택된 하나 이상의 염기서열을 가지는 유전자로부터 전사된 lncRNA의 발현에 대한 조절 효과를 결정하는 단계; 및 (c) 상기 시험 체제와 접촉된 세포 및 상기 시험 체제와 접촉되지 않은 대조군에서 상기 lncRNA의 발현 수준을 비교하는 단계;를 포함하는 유도만능 줄기세포의 분화 조절제 스크리닝 방법이 제공된다.

[0013] 본 발명의 다른 측면에 따르면, 상기 신경 분화 특이적 lncRNA의 발현을 증가 또는 억제시키는 단계;를 포함하는 유도만능 줄기세포의 분화 제어 방법이 제공된다.

발명의 효과

[0014] 본 발명에 따르면, 신경 분화 특이적 lncRNA의 발현량을 분석하거나, 발현량을 조절함으로써 퇴행성 뇌신경질환 세포치료를 위한 의학적 용도로 활용하거나, 관련된 연구에 있어서 기초 자료로서 활용 가능성이 높다.

[0015] 본 발명의 효과는 상기한 효과로 한정된 것은 아니며, 본 발명의 상세한 설명 또는 청구범위에 기재된 발명의 구성으로부터 추론 가능한 모든 효과를 포함하는 것으로 이해되어야 한다.

도면의 간단한 설명

[0016] 도 1은 본 발명의 일 실시예에 따른 신경 분화 특이적 lncRNA 탐색 방법을 도식화한 것이다.

도 2는 lncRNA 유전자 중 NPC 특이적 유전자를 규명한 것이다.

도 3은 발명의 일 실시예에 따른 신경 분화 특이적 lncRNA 후보군을 ZENBU browser에서 개별적으로 확인한 것이다.

도 4는 RT-qPCR을 이용하여 20개의 후보군에 대한 실험적 검증을 진행한 것이다.

발명을 실시하기 위한 구체적인 내용

[0017] 이하에서는 첨부한 도면을 참조하여 본 발명을 설명하기로 한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 따라서 여기에서 설명하는 실시예로 한정되는 것은 아니다.

[0018] 어떤 부분이 어떤 구성요소를 “포함” 한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 구비할 수 있다는 것을 의미한다.

- [0019] 달리 정의되지 않는 한, 분자 생물학, 미생물학, 단백질 정제, 단백질 공학, 및 DNA 서열 분석 및 당업자의 능력 범위 안에서 재조합 DNA 분야에서 흔히 사용되는 통상적인 기술에 의해 수행될 수 있다. 상기 기술들은 당업자에게 알려져 있고, 많은 표준화된 교재 및 참고서에 기술되어 있다.
- [0020] 본 명세서에 달리 정의되어 있지 않으면, 사용된 모든 기술 및 과학 용어는 당업계에 통상의 기술자가 통상적으로 이해하는 바와 같은 의미를 가진다.
- [0021] 본 명세서에 포함되는 용어를 포함하는 다양한 과학적 사건이 잘 알려져 있고, 당업계에서 이용가능하다. 본 명세서에 설명된 것과 유사 또는 등가인 임의의 방법 및 물질이 본원의 실행 또는 시험에 사용되는 것으로 발견되나, 몇몇 방법 및 물질이 설명되어 있다. 당업자가 사용하는 맥락에 따라, 다양하게 사용될 수 있기 때문에, 특정 방법학, 프로토콜 및 시약으로 본 발명이 제한되는 것은 아니다.
- [0022] 본 명세서에서 사용되는 바와 같이, 단수형은 문맥이 명확하게 달리 지시하지 않으면 복수의 대상을 포함한다. 또한, 달리 지시된 바가 없으면, 핵산은 각각 왼쪽에서 오른쪽, 5'에서 3' 방향으로 썬여지고, 아미노산 서열은 왼쪽에서 오른쪽, 아미노에서 카복실 방향으로 썬여진다. 이하 본 발명을 더욱 상세히 설명한다.
- [0023] 본 발명의 일 측면에 따르면, 서열번호 1 내지 20으로 이루어진 군에서 선택된 하나 이상의 염기서열을 가지는 유전자로부터 전사된 유도만능 줄기세포 신경 분화 특이적 lncRNA(long non-coding RNA)가 제공된다.
- [0024] 상기 “lncRNA(Long noncoding RNA, 긴 비-코딩 RNA)”는 게놈에 침투적으로 전사되는 최근 발견된 새로운 부류의 전사체로서 에피게놈의 중요한 조절인자이다. 상기 lncRNA는 몇몇 암에서 불량한 예후와 종양 전이의 유용한 바이오마커일 수 있다.
- [0025] 상기 lncRNA는 최근 Regulatory RNA로서 재조명되고 있으며, 특히 Epigenetic의 주요 역할을 수행하거나, Transcription의 조절, 암 형성이나 배아 발달, 신경계통 장애, 혹은 다른 필수적인 biological Process에 영향을 미치게 된다.
- [0026] 본 발명자들은 신경세포 분화에 관련된 lncRNA 규명 연구 및 lncRNA가 가지고 있는 조절 기능에 착안하여 non-coding 지역으로부터 lncRNA 유전자를 발굴할 수 있는 분석파이프라인을 개발하였다.
- [0027] RNA-seq과 CAGE-seq을 이용하여 인간 유도만능 줄기세포(hiPSC)로부터 신경줄기세포(NPC)로의 분화 단계별 전사체를 생산하고 상기 분석 파이프라인을 이용하여 신경세포 분화에 관여하는 50개의 lncRNA 후보군을 발굴하였다.
- [0028] 또한, RT-PCR 방법으로 실험적 검증을 진행하여 인간유도 만능줄기세포(hiPSC) 특이적 lncRNA를 3개, 신경줄기세포(NPC) 특이적 lncRNA 17개를 규명하였다.
- [0029] 상기 “줄기세포(stem cell)”란, 개체를 구성하는 세포나 조직의 근간이 되는 세포로서 반복 분열하여 자가 재생(self-renewal)할 수 있고, 환경에 따라 특정한 기능을 지닌 세포로 분화할 수 있는 다분화 능력을 갖는 세포를 의미한다. 태아의 발생과정 중 모든 조직에서 생겨나며, 성인이 되어서도 골수, 상피조직 등 세포가 활발히 교체되는 일부 조직에서 발견된다.
- [0030] 상기 줄기세포는 분화 가능한 세포의 종류에 따라 수정란이 첫 분열을 시작할 때 형성되는 전능성 줄기세포(totipotent stem cells)와, 이 세포들이 계속 분열해 만들어진 포배 내막에 있는 만능성 줄기세포(pluripotent stem cells), 그리고 성숙한 조직과 기관 속에 존재하는 다능성 줄기세포(multipotent stem cells)로 분류된다.
- [0031] 상기 다능성 줄기세포는, 이 세포가 포함되어 있는 조직 및 기관에 특이적인 세포로만 분화할 수 있는 세포로서 태아기, 신생아기 및 성체기의 각 조직 및 장기의 성장과 발달은 물론 성체조직의 항상성 유지와 조직 손상 시 재생을 유도하는 기능에 관여하고 있다. 상기 조직 특이적 다능성 세포들을 총칭하여 성체 줄기세포라고도 한다.
- [0032] 상기 “유도만능 줄기세포(induced pluripotent stem cell; iPSC)”는 역분화 줄기세포로서 분화가 끝난 체세포에 세포 분화 관련 유전자를 주입하여 분화 이전의 세포 단계로 되돌린, 배아 줄기세포처럼 만능성을 유도해낸 세포를 의미한다.
- [0033] 2006년 교토대 야마나카 신야 교수는 생쥐의 피부 섬유아세포에 몇 가지 유전자를 도입하여 배아줄기세포처럼 만능성을 가진 줄기세포를 만들었으며, 2007년 성인의 피부세포에 야마나카 인자라고도 불리는 SOX2, c-MYC,

OCT4, 및 KLF4유전자를 도입하여 유도만능 줄기세포를 만드는데 성공했다.

- [0034] 상기 유도만능 줄기세포는 체세포 또는 이미 분화된 세포를 처리하여 만능분화성을 가지며, 상기 처리 방법은 화합물, 유전적 변환 또는 특정 조건으로 배양하는 방법 등을 포함하나 이에 한정되는 것은 아니다.
- [0035] 상기 "인간 유도만능 줄기세포(hiPSC)" 는 인간의 체세포 또는 인간의 분화된 세포를 처리하여 만능분화성을 갖게 된 세포를 나타낸다.
- [0036] 상기 "신경전구세포(neural precursor cells, NPCs)"는 신경 세포의 형태 및 기능을 완전히 갖추기 전 단계의 세포를 의미한다.
- [0037] 상기 신경전구세포는 인간 유도만능 줄기세포로부터 분화된 것으로, 인간 유도만능 줄기세포로부터 신경전구세포로의 분화는 적절한 성장 인자 또는 화학물질을 처리하는 것에 의해 달성될 수 있다.
- [0038] 또한, 상기 인간 유도만능 줄기세포로부터 신경전구세포로의 분화는 적절한 세포와의 공배양에 의해 분화시킬 수 있다.
- [0039] 본 발명의 다른 측면에 따르면, 상기 lncRNA 염기서열, 상기 lncRNA에 대한 안티센스 염기서열, 또는 상기 lncRNA 염기서열의 단편인 유도만능 줄기세포의 분화 예측용 바이오 마커가 제공된다.
- [0040] 상기 "마커" 는 세포의 특징 및/또는 표현형을 나타내기 위해 사용된다. 마커는 관심 대상의 특징을 포함하는 세포를 선택하기 위해 사용될 수 있다. 상기 마커는 특정 세포형을 갖는 세포 또는 그 세포형에 의해 발현되는 분자의 형태학적, 기능적, 또는 생화학적(효소적) 특징을 가진다.
- [0041] 본 발명의 다른 측면에 따르면, 상기 lncRNA 염기서열, 상기 lncRNA 염기서열의 단편, 상기 lncRNA 염기서열에 특이적으로 결합하는 프라이머 또는 프로브를 포함하는 유도만능 줄기세포의 분화 예측용 키트가 제공된다.
- [0042] 상기 유도만능 줄기세포의 분화 예측용 키트는 마이크로어레이, 유전자 증폭 키트, RNA 시퀀싱(RNA sequencing) 또는 면역분석(immunoassay)용 키트일 수 있으며, 상기 유도만능 줄기세포의 분화 예측용 키트는 상기 신경 분화 특이적 lncRNA를 정량분석 또는 정성분석 함으로써 유도만능 줄기세포의 분화 가능성이나 정도를 예측할 수 있다.
- [0043] 상기 정량분석은 생물학적 시료에서 상기 lncRNA 존재 여부와 발현 정도를 확인하는 과정으로 mRNA의 양을 측정하는 과정을 의미한다. 상기 분석 방법은 RNA시퀀싱 (RNA-sequencing), 역전사 중합효소반응(RT-PCR), 경쟁적 역전사 중합효소반응(Competitive RT-PCR), 실시간 역전사 중합효소반응(Real-time RT-PCR), RNase 보호 분석법(RPA; RNase protection assay), 노던 블랏팅(Northern blotting), DNA 칩 등 다양한 방법이 사용될 수 있으나, 이에 제한되는 것은 아니다.
- [0044] 본 발명의 다른 측면에 따르면, (a) 유도만능 줄기세포 및 시험 제제를 접촉시키는 단계; (b) 서열번호 1 내지 20으로 이루어진 군에서 선택된 하나 이상의 염기서열을 가지는 유전자로부터 전사된 lncRNA의 발현에 대한 조절 효과를 결정하는 단계; 및 (c) 상기 시험 제제와 접촉된 세포 및 상기 시험 제제와 접촉되지 않은 대조군에서 상기 lncRNA의 발현 수준을 비교하는 단계;를 포함하는 유도만능 줄기세포의 분화 조절제 스크리닝 방법이 제공된다.
- [0045] 상기 "접촉" 은 통상적인 의미로서, 2개 이상의 제제(예: 2개의 폴리펩티드)를 결합(combine)시키거나, 제제와 세포(예; 단백질과 세포)를 결합시키는 것을 의미할 수 있다. 예컨대, 시험관(test tube) 또는 다른 컨테이너(container)에서 2개 이상의 제제를 결합시키거나 시험 제제와 세포 또는 세포 용해물과 시험 제제를 결합시킬 수 있으며, 2개의 폴리펩티드를 암호화하는 재조합 폴리뉴클레오티드를 세포 내에서 공동발현(coexpression)시킴으로써 세포 또는 세포 용해물에서 2개의 폴리펩티드를 접촉시킬 수 있다.
- [0046] 상기 "제제(agent)" 또는 "시험 제제(test agent)"는 임의의 물질(substance), 분자(molecule), 원소(element), 화합물(compound), 실체물(entity) 또는 이들의 조합을 포함할 수 있다. 예컨대, 단백질, 폴리펩티드, 소 유기 물질(small organic molecule), 다당류(polysaccharide), 폴리뉴클레오티드 등을 포함할 수 있으며, 자연 산물(natural product), 합성 화합물 또는 화학 화합물 또는 2개 이상의 물질의 조합일 수 있다. 달리 정의되지 않는 한, 상기 제제, 물질 및 화합물은 상호 교환적(interchangeably)으로 사용할 수 있다.
- [0047] 상기 스크리닝 방법은 당업계에 공지된 다양한 생화학적 및 분자생물학적 기술을 이용할 수 있으며[Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, N.Y., Second(1998) and Third(2000) Editions; Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons,

Inc., New York(1987-1999)], 예컨대, 시험관 내 단백질-단백질 결합 어세이(시험관 내 풀-다운 어세이), EMSA, 단백질 결합을 위한 면역 어세이, 기능적 어세이(인산화 어세이 등), 효모-2 하이브리드 어세이, 비면역 침전 어세이, 면역침전 웨스턴 블랏 어세이, 면역-공동-위치화 어세이 등 당업계에 공지된 다양한 방법으로 수행될 수 있다.

[0048] 상기 스크리닝에 사용되는 화합물은 저분자량의 화합물일 수 있다. 예컨대, 상기 저분자량의 화합물은 중량이 400Da, 600Da 또는 800Da과 같이 약 1000Da 내외일 수 있다. 목적에 따라 상기 화합물은 화합물 라이브러리의 일부를 구성할 수 있으며, 라이브러리를 구성하는 화합물의 개수도 수십개부터 수백만개까지 상이할 수 있다.

[0049] 상기 화합물 라이브러리는 펩타이드, 펩토이드 및 기타 환형 또는 선형의 올리고머성 화합물, 및 주형을 기본으로 하는 저분자 화합물, 예컨대 벤조디아제핀, 하이단토인, 바이아릴, 카보 사이클 및 폴리사이클 화합물(예컨대 나프탈렌, 페노티아진, 아크리딘, 스테로이드 등), 카보하이드레이트 및 아미노산 유도체, 디하이드로피리딘, 벤즈하이드릴 및 헤테로사이클(예컨대 트리아진, 인돌, 티아졸리딘 등)을 포함할 수 있으나, 이에 한정되는 것은 아니다.

[0050] 상기 스크리닝 방법에는 바이올로지스가 사용될 수 있다. 상기 바이올로지스는 세포 또는 바이오 분자를 일컫는 것으로, 단백질, 핵산, 탄수화물, 지질 또는 생체 내 및 생체 외에서 세포 시스템 등을 이용하여 생산된 물질을 의미한다. 상기 바이오분자는 단독으로 또는 다른 바이오분자 또는 세포와 조합되어 사용될 수 있다. 상기 바이오분자는 예컨대, 폴리뉴클레오타이드, 펩타이드, 항체, 또는 기타 혈장에서 발견되는 단백질 또는 생물학적 유기물질을 포함할 수 있다.

[0051] 본 발명의 다른 측면에 따르면, 상기 신경 분화 특이적 lncRNA의 발현을 증가 또는 억제시키는 단계;를 포함하는 유도만능 줄기세포의 분화 제어 방법이 제공된다.

[0052] 상기 lncRNA의 발현 증가 또는 억제는 lncRNA과 상호 작용하여 활성에 영향을 미칠 수 있으면 족하며, 항체 또는 이의 항원 결합 단편, 압타머, siRNA, shRNA, microRNA, 저해 화합물 및 이의 약학적으로 허용 가능한 염으로 이루어진 군에서 선택된 하나 이상이 사용될 수 있으나, 이에 한정되는 것은 아니다.

[0053] 이하, 실시예를 통해 본 발명을 더욱 상세히 기술한다.

[0054] 실험예 1 : 후보군 선별

[0055] RIKEN 으로 제공된 샘플 시퀀싱 데이터를 사용하여 자체 개발 분석 파이프라인으로 hiPSC 신경 분화-특이적인 lncRNA(50개)를 발굴하였다.

[0056] FANTOM에서 생산된 nAntTi-CAGE-seq(no-amplification non-tagging CAGE libraries for Illumina next-generation sequencers) 및 total RNA-seq 데이터를 분석하였다.

[0057] 신경 분화-특이적 lncRNA 발굴 분석 파이프라인을 구축하여 50개의 lncRNA 선별하였으며, 발현량을 기반으로 상위 20개의 lncRNA 후보군을 선별하였다.

[0058] 표 1을 참조하면, 2종의 annotation을 사용한 4배 이상 발현량에 차이가 있는 lncRNA 유전자 중 NPC 특이적 후보를 규명하였으며, lncRNA 유전자 대한 불충분한 annotation을 뒷받침하기 위해 각각의 파일이 포함하지 않는 정보를 보완하는 방식으로 진행하였다(도 2).

[0059] 1차 분석시 LNCipedia high-confidence set(version 4.0)를 사용하였고, 2차 분석시 FANTOM CAT(CAGE associated Transcriptome) gene annotation을 사용하였다.

표 1

[0060]	NPC 특이적 lncRNA 개수	1차 분석(LNCipedia)	2차 분석(FANTOM CAT)
	분석 결과	60	51
	공통으로 발견된 lncRNA	18	
	확인 후 선별된 lncRNA	36	14

[0061] 실험예 2 : 분석 파이프라인

[0062] FANTOM에서 인간유도 만능줄기세포(hiPSC) 및 신경줄기세포(NPC)에 대한 생산된 nAnt-iCAGE-seq 및 total RNA-seq ZENBU browser를 통해 데이터를 확보하였다.

[0063] LncRNA의 발현량을 확인하고자 lncRNA annotation 파일에 CAGE-seq 및 RNA-seq 데이터를 intersection 했으므로 NPC 및 hiPSC 간의 차등 발현 lncRNA를 조사하였다.

[0064] Fold-change 값이 4배 이상을 기준으로 후보군을 선별하였으며 그 중 NPC 특이적인 lncRNA를 규명하였다(표 2).

표 2

구분	단계	내용
1	데이터 다운로드	RNA-seq & CAGE-seq 데이터 ZENBU browser에서 내려 받음(binning size: 128bp)
2	파일 intersection	LncRNA발현 데이터와 annotation 파일 intersection(bedtools)
3	strand 유형 정리	LncRNA strand(Sense, Anti-sense) 정리
4	유전자 grouping	LncRNA 유전자별로의 grouping(bedtools)
5	데이터 표준화	각 샘플 값 normalization
6	Fold change 값 계산	RNA-seq 및 CAGE-seq 공통으로 iPSC와 NPC fold-change값이 4배 이상 차등발현 보이는 lncRNA 규명
7	ZENBU browser에서 선별	NPC 특이적인 lncRNA 후보를 개별적으로 ZENBU browser에서 확인함으로 high-confidence list 선별

[0066] 실험예 3 : lncRNA 발현량 검증

[0067] RIKEN에서 운영하는 genome browser인 ZENBU browser을 통해 NPC 특이적인 lncRNA 후보군을 선정하였으며, ZENBU browser에서 개별적으로 확인하였다.

[0068] RNA-seq 데이터를 이용해 lncRNA 발현량과 correlation 여부를 확인하고, 동시에 NPC 특이성을 확인함으로써 lncRNA 후보군에 대한 목록을 간소화 시켰으며, 최종적으로 high-confidence set을 선별하였다(도 3).

[0069] 실험예 4 : 최종 lncRNA 선별

[0070] 후보군 선별을 위해 hiPSC 및 NPC 특이적인 후보군 50개 중에서 발현량(transcripts per million_total RNA-seq)이 가장 높은 상위 20개의 lncRNA(hiPSC-특이적인 후보군 3개, NPC-특이적인 후보군 17개)를 선별하였다(표 3).

표 3

서열 번호	후보 이름	유전자명	발현량		위치		
			NPC	hiPSC	염색체	start	end
1	npc1	lnc-TRAFD1-4	1088.89	236.96	chr12	112267077	112267394
2	npc2	lnc-LRRTM4-6	895.776	189.466	chr2	76445079	76445410
3	npc3	lnc-TMEM135-8	683.198	6.39425	chr11	86800544	86937818
4	npc4	lnc-ENC1-5	293.456	15.3835	chr5	74627406	74641424
5	npc5	lnc-PYG01-2	238.96	0.709	chr15	55611552	55614422
6	npc6	lnc-SLC10A6-4	221.71	19.0565	chr4	86164446	86594119
7	npc7	lnc-CDH11-12	171.95	6.7045	chr16	65004642	65126112
8	npc8	lnc-SIX2-3	134.825	1.058	chr2	44920193	44939206
9	npc9	lnc-TDP1-3	113.702	24.6073	chr14	89712511	89712823
10	npc10	lnc-WDR25-6	104.141	0.13625	chr14	100726893	100728690
11	npc11	lnc-STBD1-3	91.0335	8.0185	chr4	76436146	76559576
12	npc12	lnc-COL2A1-1	78.06	0.331	chr12	47997615	48004486
13	npc13	SPON1	69.5095	4.6435	chr11	13962723	14280155
14	npc14	lnc-GLIS3-2	65.9525	4.04975	chr9	3824127	4348392
15	npc15	lnc-SS18-7	59.472	3.47475	chr18	25227587	25352190
16	npc16	lnc-CHD2-4	36.6095	1.82425	chr15	93312557	93569483
17	npc17	FEZF1-AS1	33.5945	0.95975	chr7	122303658	122310077
18	ipsc1	lnc-C8orf83-8	0.214	38.0028	chr8	92005166	92103228
19	ipsc2	lnc-POM121L2-2	0.9175	10.7953	chr6	27357823	27375374
20	ipsc3	lnc-HS3ST5-7	0.1655	45.5813	chr6	114408889	114456194

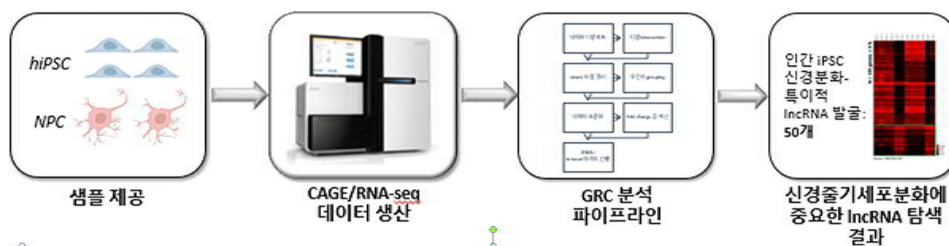
[0072] lncRNA 후보군을 검출하기 위한 프라이머 세트를 디자인한 후 RT- qPCR(reverse transcription quantitative PCR)을 이용하여 20개의 후보군에 대한 실험적 검증을 수행하였다(도 4).

[0073] 전술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.

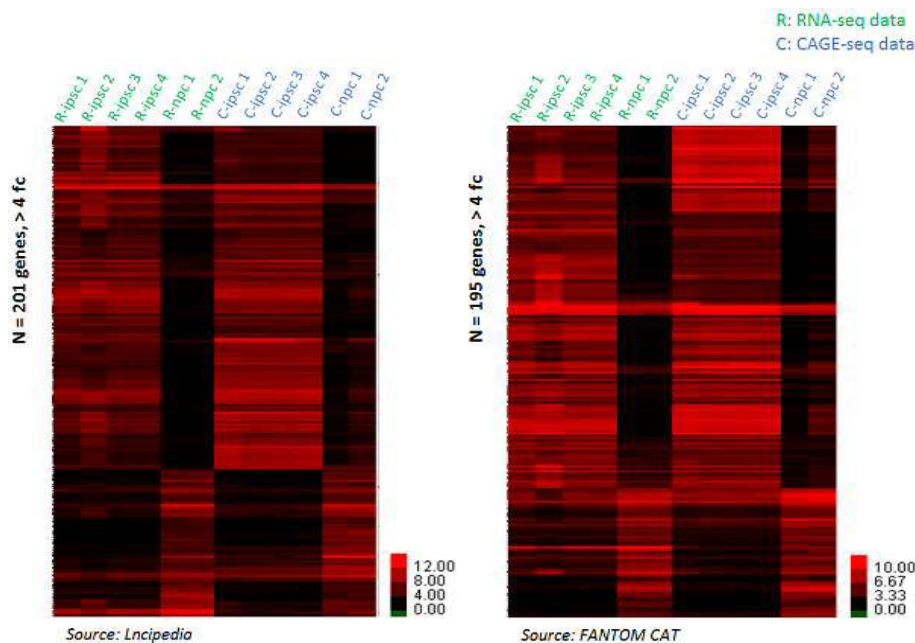
[0074] 본 발명의 범위는 후술하는 청구범위에 의하여 나타내어지며, 청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

도면

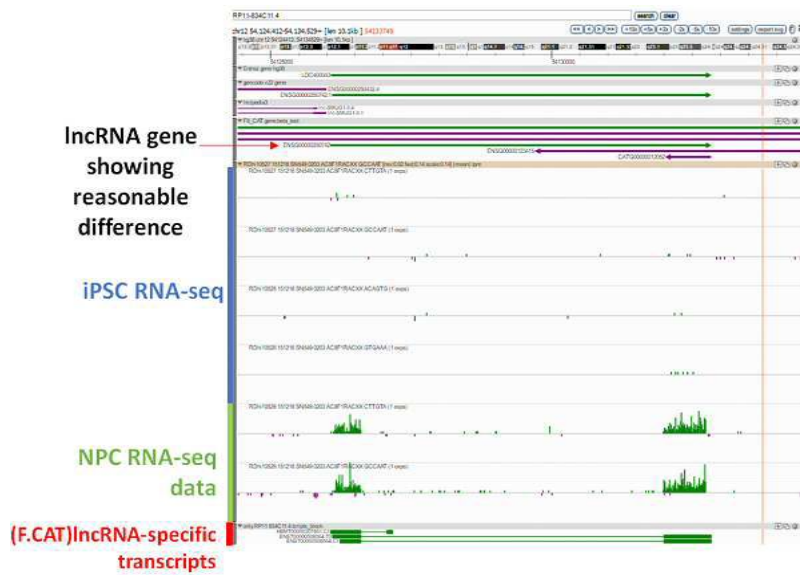
도면1



도면2



도면3



도면4

