

(11) 공개번호 10-2022-0035793
(43) 공개일자 2022년03월22일

(71) 출원인
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자
김다함
서울특별시 서초구 잠원로8길 35, 106동 603호
남기현
서울특별시 강남구 압구정로 201, 87동 304호
(뒷면에 계속)

(74) 대리인
파도특허법인유한회사, 이재영

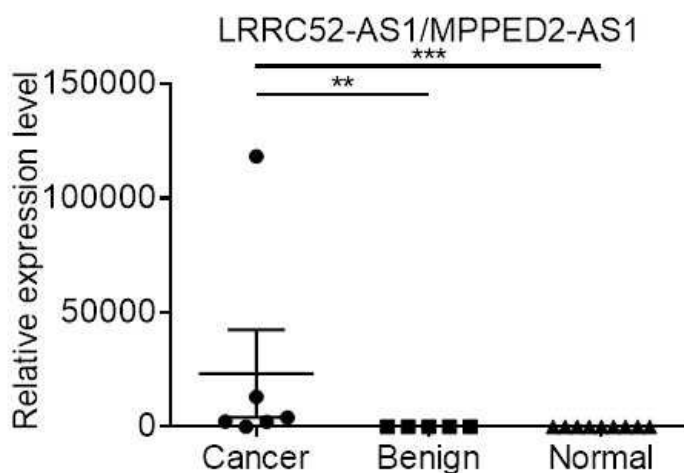
전체 청구항 수 : 총 19 항

(54) 발명의 명칭 암 진단의 정확도를 높이기 위한 유전자 발현 수준의 보정 방법

(57) 요약

본 발명은 암 진단의 정확도를 높이기 위한 방법에 관한 것이다. 특히 본 발명의 보정된 발현 수준 값을 대조군과 비교하여 암을 진단하는 경우, 단일 lncRNA의 발현 수준을 측정하는 것에 비하여 암 진단의 정확도를 높일 수 있다.

대표도 - 도11



(52) CPC특허분류

C12Q 2600/158 (2013.01)

C12Q 2600/178 (2013.01)

(72) 발명자

윤정현

서울특별시 서초구 서초대로65길 13-10, 105동
2403호

이은직

서울특별시 마포구 창전로 26, 103동 2103호

이 발명을 지원한 국가연구개발사업

과제고유번호 1345317629

과제번호 2018R1D1A1B07050637

부처명 교육부

과제관리(전문)기관명 한국연구재단

연구사업명 이공학 개인기초연구지원사업 기본연구지원사업 기본연구(1년~5년)

연구과제명 난치성 갑상선암 진행에 관여하는 새로운 long non-coding RNA 발굴 및 역할 규명

기 여 율 1/1

과제수행기관명 연세대학교

연구기간 2018.06.01 ~ 2021.05.31

명세서

청구범위

청구항 1

목적하는 개체로부터 분리된 생물학적 시료에서, 제1 lncRNA 및 제2 lncRNA의 발현 수준을 측정하는 단계; 및 측정된 상기 제1 lncRNA 및 제2 lncRNA의 발현 수준을 하기 식 1에 대입하는 단계를 포함하고,

상기 제1 lncRNA 및 제2 lncRNA는 각각 독립적으로 암에서 발현되는 lncRNA인, 암 진단의 정확도를 높이기 위한 보정 방법:

[식 1]

보정된 발현 수준 값 = 제1 lncRNA의 발현 수준/제2 lncRNA의 발현 수준.

청구항 2

제1항에 있어서,

상기 생물학적 시료는 전혈(whole blood), 백혈구(leukocytes), 말초혈액 단핵 세포(peripheral blood mononuclear cells), 백혈구 연층(buffy coat), 혈장(plasma), 혈청(serum), 객담(sputum), 눈물(tears), 점액(mucus), 세비액(nasal washes), 비강 흡인물(nasal aspirate), 호흡(breath), 소변(urine), 정액(semen), 침(saliva), 복강 세척액(peritoneal washings), 복수(ascites), 낭종액(cystic fluid), 뇌척수막 액(meningeal fluid), 양수(amniotic fluid), 선액(glandular fluid), 췌장액(pancreatic fluid), 림프액(lymph fluid), 흉수(pleural fluid), 유두 흡인물(nipple aspirate), 기관지 흡인물(bronchial aspirate), 활액(synovial fluid), 관절 흡인물(joint aspirate), 기관 분비물(organ secretions), 세포(cell), 세포 추출물(cell extract) 또는 뇌척수액(cerebrospinal fluid)인, 암 진단의 정확도를 높이기 위한 보정 방법.

청구항 3

제1항에 있어서,

상기 제1 lncRNA는 정상 대조군 대비 암에서 발현 수준이 증가하거나, 양성 종양 대비 암에서 발현 수준이 증가한 lncRNA인, 암 진단의 정확도를 높이기 위한 보정 방법.

청구항 4

제3항에 있어서,

상기 제1 lncRNA는 류신 풍부 반복 함유 52 단백질의 안티센스 RNA 1(LRRC52-AS1), 긴 유전자간 비-단백질 코딩 RNA 2082(LINC02082), Unc-5 네트린 수용체 B 안티센스 RNA 1(UNC5B-AS1), 긴 유전자간 비-단백질 코딩 RNA 2471(LINC02471) 및 긴 유전자간 비-단백질 코딩 RNA 2408(LINC02408)로 이루어진 군에서 선택된 1종 이상인, 암 진단의 정확도를 높이기 위한 보정 방법.

청구항 5

제1항에 있어서,

상기 제2 lncRNA는 정상 대조군 대비 암에서 발현 수준이 감소하거나, 양성 종양 대비 암에서 발현 수준이 감소한 lncRNA인, 암 진단의 정확도를 높이기 위한 보정 방법.

청구항 6

제5항에 있어서,

상기 제2 lncRNA는 메탈로포스포에스터라아제 도메인 함유 안티센스 RNA 1(MPPED2-AS1), lncRNA 인접 증강제(LINCNEF), LOC100129129 및 LOC642484으로 이루어진 군에서 선택된 1종 이상인, 암 진단의 정확도를 높이기 위

한 보정 방법.

청구항 7

제1항에 있어서,

상기 제1 lncRNA 및 제2 lncRNA의 발현 수준을 측정하는 단계는 상기 제1 lncRNA 및 제2 lncRNA 각각에 특이적으로 결합하는 프라이머, 프로브 및 안티센스 뉴클레오티드로 이루어진 군에서 선택된 1종 이상을 포함하는 재제를 이용하여 수행되는, 암 진단의 정확도를 높이기 위한 보정 방법.

청구항 8

제1항에 있어서,

상기 제1 lncRNA 및 제2 lncRNA의 발현 수준의 측정은 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) 또는 DNA 칩에 의해 수행되는, 암 진단의 정확도를 높이기 위한 보정 방법.

청구항 9

제1항에 있어서,

상기 암은 유방암, 난소암, 대장암, 위암, 간암, 췌장암, 자궁경부암, 갑상선암, 부갑상선암, 폐암, 비소세포성 폐암, 전립선암, 담낭암, 담도암, 비호지킨 림프종, 호지킨 림프종, 혈액암, 방광암, 신장암, 흑색종, 결장암, 골암, 피부암, 두부암, 자궁암, 직장암, 뇌종양, 향문부근암, 나팔관암종, 자궁내막암종, 질암, 음문암종, 식도암, 소장암, 내분비선암, 부신암, 연조직 육종, 요도암, 음경암, 수뇨관암, 신장세포 암종, 신장골반 암종, 중추신경계(central nervous system; CNS) 종양, 1차 CNS 림프종, 척수 종양, 뇌간 신경교종 또는 뇌하수체 선종인, 암 진단의 정확도를 높이기 위한 보정 방법.

청구항 10

목적하는 개체로부터 분리된 생물학적 시료에서, 제1 lncRNA 및 제2 lncRNA의 발현 수준을 측정하는 단계; 및 측정된 상기 제1 lncRNA 및 제2 lncRNA의 발현 수준을 하기 식 1에 대입하는 단계를 포함하고,

상기 제1 lncRNA 및 제2 lncRNA는 각각 독립적으로 암에서 발현되는 lncRNA인, 암 진단을 위한 정보 제공 방법:

[식 1]

보정된 발현 수준 값 = 제1 lncRNA의 발현 수준/제2 lncRNA의 발현 수준.

청구항 11

제10항에 있어서,

상기 생물학적 시료는 전혈(whole blood), 백혈구(leukocytes), 말초혈액 단핵 세포(peripheral blood mononuclear cells), 백혈구 연층(buffy coat), 혈장(plasma), 혈청(serum), 객담(sputum), 눈물(tears), 점액(mucus), 세비액(nasal washes), 비강 흡인물(nasal aspirate), 호흡(breath), 소변(urine), 정액(semen), 침(saliva), 복강 세척액(peritoneal washings), 복수(ascites), 낭종액(cystic fluid), 뇌척수막 액(meningeal fluid), 양수(amniotic fluid), 선액(glandular fluid), 췌장액(pancreatic fluid), 림프액(lymph fluid), 흉수(pleural fluid), 유두 흡인물(nipple aspirate), 기관지 흡인물(bronchial aspirate), 활액(synovial fluid), 관절 흡인물(joint aspirate), 기관 분비물(organ secretions), 세포(cell), 세포 추출물(cell extract) 또는 뇌척수액(cerebrospinal fluid)인, 암 진단을 위한 정보 제공 방법.

청구항 12

제10항에 있어서,

상기 제1 lncRNA는 정상 대조군 대비 암에서 발현 수준이 증가하거나, 양성 종양 대비 암에서 발현 수준이 증가한 lncRNA인, 암 진단을 위한 정보 제공 방법.

청구항 13

제12항에 있어서,

상기 제1 lncRNA는 류신 풍부 반복 함유 52 단백질의 안티센스 RNA 1(LRRC52-AS1), 긴 유전자간 비-단백질 코딩 RNA 2082(LINC02082), Unc-5 네트린 수용체 B 안티센스 RNA 1(UNC5B-AS1), 긴 유전자간 비-단백질 코딩 RNA 2471(LINC02471) 및 긴 유전자간 비-단백질 코딩 RNA 2408(LINC02408)으로 이루어진 군에서 선택된 1종 이상인, 암 진단을 위한 정보 제공 방법.

청구항 14

제10항에 있어서,

상기 제2 lncRNA는 정상 대조군 대비 암에서 발현 수준이 감소하거나, 양성 종양 대비 암에서 발현 수준이 감소한 lncRNA인, 암 진단을 위한 정보 제공 방법.

청구항 15

제14항에 있어서,

상기 제2 lncRNA는 메탈로포스포에스터라아제 도메인 함유 안티센스 RNA 1(MPPED2-AS1), lncRNA 인접 증강제(LINCNEF), LOC100129129 및 LOC642484으로 이루어진 군에서 선택된 1종 이상인, 암 진단을 위한 정보 제공 방법.

청구항 16

제10항에 있어서,

상기 제1 lncRNA 및 제2 lncRNA의 발현 수준의 측정은 상기 제1 lncRNA 및 제2 lncRNA 각각에 특이적으로 결합하는 프라이머, 프로브 및 안티센스 뉴클레오티드로 이루어진 군에서 선택된 1종 이상을 포함하는 체제를 이용하여 수행되는, 암 진단을 위한 정보 제공 방법.

청구항 17

제10항에 있어서,

상기 제1 lncRNA 및 제2 lncRNA의 발현 수준의 측정은 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) 또는 DNA 칩에 의해 수행되는, 암 진단을 위한 정보 제공 방법.

청구항 18

제10항에 있어서,

상기 암은 유방암, 난소암, 대장암, 위암, 간암, 췌장암, 자궁경부암, 갑상선암, 부갑상선암, 폐암, 비소세포성 폐암, 전립선암, 담낭암, 담도암, 비호지킨 림프종, 호지킨 림프종, 혈액암, 방광암, 신장암, 흑색종, 결장암, 골암, 피부암, 두부암, 자궁암, 직장암, 뇌종양, 항문부근암, 나팔관암종, 자궁내막암종, 질암, 음문암종, 식도암, 소장암, 내분비선암, 부신암, 연조직 육종, 요도암, 음경암, 수뇨관암, 신장세포 암종, 신장골반 암종, 중추신경계(central nervous system; CNS) 종양, 1차 CNS 림프종, 척수 종양, 뇌간 신경교종 또는 뇌하수체 선종인, 암 진단을 위한 정보 제공 방법.

청구항 19

제10항에 있어서,

상기 보정된 발현 수준 값이 대조군에 비하여 증가한 경우, 상기 암의 발병 가능성이 높은 것으로 예측하는, 암 진단을 위한 정보 제공 방법.

발명의 설명

기술 분야

[0001] 본 발명은 암 진단의 정확도를 높이기 위하여 유전자의 발현 수준을 보정하는 방법에 관한 것이다.

배경 기술

[0002] 암은 인류가 해결해야 할 난치병 중의 하나로, 전 세계적으로 이를 치유하기 위한 개발에 막대한 자본이 투자되고 있는 실정이며, 우리나라의 경우, 질병 사망원인 중 제1위의 질병으로서 연간 약 10만 명 이상이 진단되고, 약 6만 명 이상이 사망하고 있다.

[0003] 이러한 암의 유발 인자인 발암물질로는 흡연, 자외선, 화학물질, 음식물, 기타 환경인자들이 있으나, 그 유발 원인이 다양하여 더욱 발전된 진단 방법 및 치료제의 개발이 절실히 필요하다.

[0004] 그중에, 갑상선암은 남성보다는 특히 여성에게 많이 발병되는 암으로써 국가 암 정보센터의 암 발병률을 살펴보면 2002년 기준 여성 전체 암의 7.4%를 차지하고 있다. 최근 대한 갑상선 질병학회에 보고된 자료에 의하면 국내 암 발병률의 4위를 차지하고 있으며 여성의 경우는 1위를 차지하고 있다.

[0005] 이러한 갑상선암은 내분비계의 가장 흔한 악성종양으로 유두암(Papillary carcinoma), 여포암(Follicular carcinoma), 허들세포암(여포 선암종, Hurthle cell neoplasm.), 역형성암(저분화 갑상선암, Anaplastic carcinoma), 수질암(Medullary thyroid carcinoma) 등이 포함된다.

[0006] 갑상선암의 원인으로는 방사선 노출, 성호르몬 등으로 알려져 있으나 아직 정확한 원인에 대한 규명은 이루어지지 않고 있다. 갑상선암은 다른 암에 비해 발병률이 높기는 하지만 조기에 수술을 할 수 있다면 생존률이 높은 편에 속하므로, 조기 발견이 중요하여 갑상선암을 조기에 탐지하기 위한 시도가 다양하게 이루어지고 있다.

[0007] 반면, 갑상선의 양성 종양의 경우 간혹 아주 커졌을 때 주위 조직을 압박하여 음식물을 삼키기가 어렵거나 호흡 곤란을 느낄 수 있으나 이런 경우는 드물고, 미용상 문제가 될 뿐 몸에 해가 없다. 따라서, 갑상선 종양이 악성인 갑상선암인지 양성 종양인지 여부를 구별하는 것이 중요하다. 갑상선암과 갑상선의 양성 종양을 감별진단하기 위해서는 세침흡인세포검사(Fine needle aspiration; FNA)가 이용되는데, 이는 세포 형태가 병리학적으로 판별 불가능한 비진단적 또는 비정형인 경우가 존재하므로, 갑상선암을 양성 종양과 정확하게 구별하기 위해 분자표지가 검사가 이용될 수 있다.

[0008] 정확도를 높이기 위한 검사를 위하여 지금까지 개발된 분자표지자 검사는 유전자 돌연변이 검사, mRNA발현검사, miRNA발현검사 등이 있으나, 유전자 돌연변이만으로 갑상선암을 설명할 수 없으며 mRNA의 경우 mRNA가 단백질을 발현과 항상 일치하는 것은 아니며 전체 유전체의 2%만을 차지하여 진단 마커로서의 한계를 보인다.

[0009] 반면 긴 비-코딩 RNA(Long non-coding RNA; lncRNA)는 단백질로 번역되지 않는 비암호화 RNA 분자 중 상대적으로 길이가 긴 200 뉴클레오타이드 이상의 크기를 가지는 RNA로, 최근 세포의 암화 과정에서 중요한 것으로 주목받고 있다. 이에 따라 암 진단을 위한 RNA 발현 수준을 측정할 때 각 샘플에서 양을 보정하기 위해서는 GAPDH와 같은 양이 잘 변하지 않는 하우스 키핑 유전자 컨트롤이 필요한데 lncRNA는 대부분 발현양이 적고, GAPDH와 같은 대부분의 하우스 키핑 유전자들은 발현양이 상대적으로 높아 lncRNA 정량 결과를 분석하기에 부적합하며 lncRNA 중에서 하우스 키핑 유전자로 알려진 lncRNA는 아직까지 없다.

발명의 내용

해결하려는 과제

[0010] 본 발명의 일 목적은 암 중에서도 특히 갑상선암의 진단의 정확도를 높이기 위하여 바이오마커 유전자의 발현 수준을 보정하는 방법을 제공하고자 한다.

[0011] 본 발명의 또 다른 일 목적은 본 발명에서 제공하는 보정 방법을 이용하여 암 중에서도 특히 갑상선암을 진단하기 위한 정보를 제공하는 방법을 제공하고자 한다.

[0012] 그러나 본 발명이 이루고자 하는 기술적 과제는 이상에서 언급한 과제에 제한되지 않으며, 언급되지 않은 또 다른 과제들은 아래의 기재로부터 당 업계에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0013] 이하, 본원에 기재된 다양한 구현예가 도면을 참조로 기재된다. 하기 설명에서, 본 발명의 완전한 이해를 위해

서, 다양한 특이적 상세사항, 예컨대, 특이적 형태, 조성물 및 공정 등이 기재되어 있다. 그러나, 특정의 구현예에는 이들 특이적 상세 사항 중 하나 이상 없이, 또는 다른 공지된 방법 및 형태와 함께 실행될 수 있다. 다른 예에서, 공지된 공정 및 제조 기술은 본 발명을 불필요하게 모호하게 하지 않게 하기 위해서, 특정의 상세사항으로 기재되지 않는다. "한 가지 구현예" 또는 "구현예"에 대한 본 명세서 전체를 통한 참조는 구현예와 결부되어 기재된 특별한 특징, 형태, 조성 또는 특성이 본 발명의 하나 이상의 구현예에 포함됨을 의미한다. 따라서, 본 명세서 전체에 걸친 다양한 위치에서 표현된 "한 가지 구현예에서" 또는 "구현예"의 상황은 반드시 본 발명의 동일한 구현예를 나타내지는 않는다. 추가로, 특별한 특징, 형태, 조성, 또는 특성은 하나 이상의 구현예에서 어떠한 적합한 방법으로 조합될 수 있다.

[0014] 본 발명 내 특별한 정의가 없으면 본 명세서에 사용된 모든 과학적 및 기술적인 용어는 본 발명이 속하는 기술 분야에서 당업자에 의하여 통상적으로 이해되는 것과 동일한 의미를 가진다.

[0016] 본 발명의 일 구현예에 따르면, 목적하는 개체로부터 분리된 생물학적 시료에서, 제1 lncRNA 및 제2 lncRNA의 발현 수준을 측정하는 단계; 및 측정된 제1 lncRNA 및 제2 lncRNA의 발현 수준을 하기 식 1에 대입하는 단계를 포함하고, 상기 제1 lncRNA 및 제2 lncRNA는 각각 독립적으로 암에서 발현되는 lncRNA인, 암 진단의 정확도를 높이기 위한 보정 방법에 관한 것이다:

[0017] [식 1]

[0018] 보정된 발현 수준 값 = 제1 lncRNA의 발현 수준/제2 lncRNA의 발현 수준

[0019] 본 발명에서 상기 "목적하는 개체" 내지 "개체"란, 상기 암의 발병 여부가 불확실한 개체로, 발병 가능성이 높은 개체를 의미하거나, 종양 또는 결절이 발생한 개체로 상기 종양 또는 결절이 양성 종양인지 암인지의 여부가 불확실한 개체를 의미할 수 있으나, 이에 제한되지는 않는다.

[0020] 본 발명에서 상기 "생물학적 시료"는 개체로부터 얻어지거나 개체로부터 유래된 임의의 물질, 생물학적 체액, 조직 또는 세포를 의미하는 것으로, 예를 들면, 전혈(whole blood), 백혈구(leukocytes), 말초혈액 단핵 세포(peripheral blood mononuclear cells), 백혈구 연층(buffy coat), 혈장(plasma), 혈청(serum), 객담(sputum), 눈물(tears), 점액(mucus), 세비액(nasal washes), 비강 흡인물(nasal aspirate), 호흡(breath), 소변(urine), 정액(semen), 침(saliva), 복강 세척액(peritoneal washings), 복수(ascites), 낭종액(cystic fluid), 뇌척수막 액(meningeal fluid), 양수(amniotic fluid), 선액(glandular fluid), 췌장액(pancreatic fluid), 림프액(lymph fluid), 흉수(pleural fluid), 유두 흡인물(nipple aspirate), 기관지 흡인물(bronchial aspirate), 활액(synovial fluid), 관절 흡인물(joint aspirate), 기관 분비물(organ secretions), 세포(cell), 세포 추출물(cell extract) 또는 뇌척수액(cerebrospinal fluid)을 포함할 수 있고, 바람직하게는 발병 가능성이 높은 환자의 피부를 절개하지 않고 중공침 등을 생체 내 기관에 삽입하여 병리조직학적 검사용으로 채취한 샘플(예를 들면, 환자의 조직, 세포, 혈액, 혈청, 혈장, 타액, 객담 또는 복수(ascites) 등)일 수 있으나, 이에 제한되는 것은 아니다.

[0021] 본 발명에서 상기 제1 lncRNA는 정상 대조군 대비 암에서 발현 수준이 증가하거나, 양성 종양 대비 암에서 발현 수준이 증가한 lncRNA일 수 있다. 여기서 상기 양성 종양에서 상기 제1 lncRNA의 발현 수준은 정상 대조군 대비 증가된 것일 수 있다.

[0022] 본 발명에서, 상기 제1 lncRNA는 류신 풍부 반복 함유 52 단백질의 안티센스 RNA 1(LRRC52-AS1), 긴 유전자간 비-단백질 코딩 RNA 2082(LINC02082), Unc-5 네트린 수용체 B 안티센스 RNA 1(UNC5B-AS1), 긴 유전자간 비-단백질 코딩 RNA 2471(LINC02471) 및 긴 유전자간 비-단백질 코딩 RNA 2408(LINC02408)으로 이루어진 군에서 선택된 1종 이상일 수 있으나, 이에 제한되지는 않는다.

[0023] 본 발명에서, 류신 풍부 반복 함유 52 단백질 안티센스 RNA 1(Leucine-rich repeat-containing protein 52 Antisense RNA 1; LRRC52-AS1)은 1번 염색체에 존재하는 긴 비-코딩 RNA로, 서열번호 1로 표시되는 것일 수 있으나, 이에 제한되는 것은 아니다.

[0024] 본 발명에서, 긴 유전자간 비-단백질 코딩 RNA 2082(Long Intergenic Non-Protein Coding RNA 2082; LINC02082)은 3번 염색체에 존재하는 긴 비-코딩 RNA로, 서열번호 2로 표시되는 것일 수 있으나, 이에 제한되는 것은 아니다.

[0025] 본 발명에서, Unc-5 네트린 수용체 B 안티센스 RNA 1(Unc-5 Netrin Receptor B Antisense RNA 1; UNC5B-AS1)는

10번 염색체에 존재하는 긴 비-코딩 RNA로, 서열번호 3으로 표시되는 것일 수 있으나, 이에 제한되는 것은 아니다.

[0026] 본 발명에서, 긴 유전자간 비-단백질 코딩 RNA 2471(Long Intergenic Non-Protein Coding RNA 2471; LINC02471)은 12번 염색체에 존재하는 긴 비-코딩 RNA로, 서열번호 4로 표시되는 것일 수 있으나, 이에 제한되는 것은 아니다.

[0027] 본 발명에서, 긴 유전자간 비-단백질 코딩 RNA 2408(Long Intergenic Non-Protein Coding RNA 2408; LINC02408)은 12번 염색체에 존재하는 긴 비-코딩 RNA로, 서열번호 5로 표시되는 것일 수 있으나, 이에 제한되는 것은 아니다.

[0028] 본 발명에서 상기 제2 lncRNA는 정상 대조군 대비 암에서 발현 수준이 감소하거나, 양성 종양 대비 암에서 발현 수준이 감소한 lncRNA일 수 있다. 여기서 상기 양성 종양에서 상기 제2 lncRNA의 발현 수준은 정상 대조군 대비 감소된 것일 수 있다.

[0029] 본 발명에서, 상기 제2 lncRNA는 메탈로포스포에스테라아제 도메인 함유 안티센스 RNA 1(MPPED2-AS1), FOXA2의 lncRNA 인접 증강제(LNCNEF), LOC100129129 및 LOC642484으로 이루어진 군에서 선택된 1종 이상일 수 있으나, 이에 제한되지는 않는다.

[0030] 본 발명에서, 메탈로포스포에스테라아제 도메인 함유 안티센스 RNA 1(Metallophosphoesterase Domain Containing 2 Antisense RNA 1; MPPED2-AS1)은 11번 염색체에 존재하는 긴 비-코딩 RNA로, 서열번호 6으로 표시되는 것일 수 있으나, 이에 제한되는 것은 아니다.

[0031] 본 발명에서, FOXA2의 lncRNA 인접 증강제(LncRNA Neighboring Enhancer Of FOXA2; LNCNEF)는, 긴 유전자간 비-단백질 코딩 RNA 1384(Long Intergenic Non-Protein Coding RNA 1384; LINC01384)라고도 하며, 20번 염색체에 존재하는 긴 비-코딩 RNA로, 서열번호 7로 표시되는 것일 수 있으나, 이에 제한되는 것은 아니다.

[0032] 본 발명에서, LOC100129129은 8번 염색체에 존재하는 긴 비-코딩 RNA로, 서열번호 8로 표시되는 것일 수 있으나, 이에 제한되는 것은 아니다.

[0033] 본 발명에서, LOC642484는 18번 염색체에 존재하는 긴 비-코딩 RNA로, 서열번호 9로 표시되는 것일 수 있으나, 이에 제한되는 것은 아니다.

[0034] 본 발명에 사용되는 용어 "발현" 또는 "발현하는"은 유전자의 전사를 포함하는 생성 과정을 지칭한다. 발현 수준 또는 발현량의 증가는 예를 들어, 비-코딩 RNA 내지 이를 코딩하는 유전자의 수의 증가, 유전자의 전사의 증가(예를 들어, 구성성 프로모터의 제어 하에 유전자를 배치함으로써), 경쟁 유전자의 낙아웃(knockout) 또는 이들 및/또는 다른 방법의 조합을 포함하는 다수의 방법에 의할 수 있고, 발현 수준 또는 발현량의 감소는 유전자 수의 감소, 유전자의 전사의 감소, 경쟁 유전자의 발현 등에 의할 수 있다.

[0035] 본 발명에서 상기 제1 lncRNA 및 제2 lncRNA의 발현 수준을 측정하는 단계는 상기 제1 lncRNA 및 제2 lncRNA 각각에 특이적으로 결합하는 프라이머, 프로브 및 안티센스 뉴클레오티드로 이루어진 군에서 선택된 1종 이상을 포함하는 제제를 이용하여 수행될 수 있으나, 이에 제한되지는 않는다.

[0036] 본 발명에서 상기 "프라이머"는 표적 유전자 서열을 인지하는 단편으로서, 정방향 및 역방향의 프라이머 쌍을 포함하나, 바람직하게는, 특이성 및 민감성을 가지는 분석 결과를 제공하는 프라이머 쌍이다. 프라이머의 핵산 서열이 시료 내 존재하는 비-표적 서열과 불일치하는 서열이어서, 상보적인 프라이머 결합 부위를 함유하는 표적 유전자 서열만 증폭하고 비특이적 증폭을 유발하지 않는 프라이머일 때, 높은 특이성이 부여될 수 있다.

[0037] 본 발명에서 상기 "프로브"란 시료 내의 검출하고자 하는 표적 물질과 특이적으로 결합할 수 있는 물질을 의미하며, 상기 결합을 통하여 특이적으로 시료 내의 표적 물질의 존재를 확인할 수 있는 물질을 의미한다. 프로브의 종류는 당업계에서 통상적으로 사용되는 물질로서 제한은 없으나, 바람직하게는 PNA(peptide nucleic acid), LNA(locked nucleic acid), 펩티드, 폴리펩티드, 단백질, RNA 또는 DNA일 수 있으며, 가장 바람직하게는 PNA이다. 보다 구체적으로, 상기 프로브는 바이오 물질로서 생물에서 유래되거나 이와 유사한 것 또는 생체 외에서 제조된 것을 포함하는 것으로, 예를 들어, 효소, 단백질, 항체, 미생물, 동식물 세포 및 기관, 신경세포, DNA, 및 RNA일 수 있으며, DNA는 cDNA, 게놈 DNA, 올리고뉴클레오타이드를 포함하며, RNA는 게놈 RNA, mRNA, 올리고뉴클레오타이드를 포함하며, 단백질의 예로는 항체, 항원, 효소, 펩티드 등을 포함할 수 있다.

[0038] 본 발명에 상기 "PNA(Peptide Nucleic Acid)"는 인공적으로 합성된, DNA 또는 RNA와 비슷한 중합체를

가리키며, 1991년 덴마크 코펜하겐 대학교의 Nielsen, Egholm, Berg와 Buchardt 교수에 의해 처음으로 소개되었다. DNA는 인산-리보스당 골격을 갖는데 반해, PNA는 펩티드 결합에 의해 연결된 반복된 N-(2-아미노에틸)-글리신 골격을 가지며, 이로 인해 DNA 또는 RNA에 대한 결합력과 안정성이 크게 증가되어 분자 생물학, 진단 분석 및 안티센스 치료법에 사용되고 있다. PNA는 문헌[Nielsen PE, Egholm M, Berg RH, Buchardt O (December 1991). "Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide". Science 254(5037): 1497-1500]에 상세하게 개시되어 있다.

[0039] 본 발명에서 상기 "LNA(Locked nucleic acids)"란, 2'-O, 4'-C 메틸렌 브릿지를 포함하는 핵산 아날로그를 의미한다 [J Weiler, J Hunziker and J Hall Gene Therapy (2006) 13, 496.502]. LNA 뉴클레오사이드는 DNA와 RNA의 일반적 핵산 염기를 포함하며, Watson-Crick 염기 쌍 규칙에 따라 염기 쌍을 형성할 수 있다. 하지만, 메틸렌 브릿지로 인한 분자의 'locking'으로 인해, LNA는 Watson-Crick 결합에서 이상적 형상을 형성하지 못하게 된다. LNA가 DNA 또는 RNA 올리고뉴클레오타이드에 포함되면, LNA는 보다 빠르게 상보적 뉴클레오타이드 사슬과 쌍을 이루어 이중 나선의 안정성을 높일 수 있다. 본 발명에서 상기 "안티센스"는 안티센스 올리고머가 왓슨-크릭 염기쌍 형성에 의해 RNA 내의 표적 서열과 혼성화되어, 표적서열 내에서 전형적으로 mRNA와 RNA:올리고머 헤테로이중체의 형성을 허용하는, 뉴클레오타이드 염기의 서열 및 서브유닛간 백본을 갖는 올리고머를 의미한다. 올리고머는 표적 서열에 대한 정확한 서열 상보성 또는 근사 상보성을 가질 수 있다.

[0040] 본 발명의 일 예시에서, 상기 제1 lncRNA 또는 제2 lncRNA에 해당하는 상기 LRRC52-AS1, LINC02082, UNC5B-AS1, LINC02471, LINC02408, MPPED2-AS1, LINCNEF, LOC100129129 및 LOC642484 lncRNA 서열 정보는 이미 알려져 있으므로, 당업자라면 이를 바탕으로 상기 유전자에 특이적으로 결합하는 프라이머, 프로브 또는 안티센스 뉴클레오타이드를 용이하게 디자인할 수 있을 것이다.

[0041] 본 발명에서, 상기 제1 lncRNA 및 제2 lncRNA의 발현 수준의 측정은 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) 또는 DNA 칩에 의해 수행될 수 있으나, 이에 제한되는 것은 아니다.

[0042] 본 발명에서, 상기 "진단"은 특정 질병 또는 질환에 대한 대상(subject)의 감수성(susceptibility)을 판정하는 것, 대상이 특정 질병 또는 질환을 현재 가지고 있는지 여부를 판정하는 것, 특정 질병 또는 질환에 걸린 대상의 예후(prognosis)(예컨대, 전-전이성 또는 전이성 암 상태의 동정, 암의 단계 결정 또는 치료에 대한 암의 반응성 결정)를 판정하는 것, 테라메트릭스(therametrics)(예컨대, 치료 효능에 대한 정보를 제공하기 위하여 객체의 상태를 모니터링하는 것) 또는 암을 양성 종양 또는 결절과 구별하는 것일 수 있으나, 이에 제한되는 것은 아니다.

[0043] 본 발명에서, 용어 "종양(tumor)"은 체내의 세포가 개체의 전체성과는 관계없이 발육하는 것으로서, 상기 종양의 발육은 항상 종양 세포 자체의 분열 및 증식에 의한다. 상기 종양의 발육 형식은 주로 양성 종양에서 볼 수 있는 팽창성 또는 압박성 발육 또는 주로 악성 종양에서 볼 수 있는 침윤성 발육이 있다.

[0044] 본 발명에서, 상기 "악성 종양(malignant tumor)", "악성 결절(malignant nodule)" 또는 "암(cancer)"은 세포 주기가 조절되지 않아 세포 분열을 계속하는 질병으로서, 발생 부위에 따라 암종(Carcinoma)과 육종(Sarcoma)으로 나뉜다. 암종(Carcinoma)은 점막, 피부 같은 상피성 세포에서 발생한 악성 종양을 뜻하고, 육종(Sarcoma)은 근육, 결합 조직, 뼈, 연골, 혈관 등의 비상피성 세포에서 발생한 악성 종양을 뜻한다. 상기 악성 종양은 양성 종양에 비하여 증식 속도, 침윤, 재발 및 전이 등의 악성도가 높다.

[0045] 본 발명에서 진단의 대상이 되는 암은 유방암, 난소암, 대장암, 위암, 간암, 췌장암, 자궁경부암, 갑상선암, 부갑상선암, 폐암, 비소세포성폐암, 전립선암, 담낭암, 담도암, 비호지킨 림프종, 호지킨 림프종, 혈액암, 방광암, 신장암, 흑색종, 결장암, 골암, 피부암, 두부암, 자궁암, 직장암, 뇌종양, 항문부근암, 나팔관암종, 자궁내막암종, 질암, 음문암종, 식도암, 소장암, 내분비선암, 부신암, 연조직 육종, 요도암, 음경암, 수노관암, 신장세포 암종, 신장골반 암종, 중추신경계(central nervous system; CNS) 종양, 1차 CNS 림프종, 척수 종양, 뇌간 신경교종 또는 뇌하수체 선종일 수 있으나, 바람직하게는 갑상선암일 수 있다.

[0046] 본 발명에서, 상기 "양성 종양(benign tumor)" 내지 "양성 결절(benign nodule)"은 세포가 과도하게 증식해 덩어리가 된 것으로, 커지는 속도가 느릴 뿐 더러 일정한 크기 이상 자라지 않는다는 특징이 있다. 또한 암처럼 다른 조직에 전이되지 않으며, 쉽게 치유할 수 있다. 양성 종양은 인체에 해는 거의 없으나, 주요 기관에 압박을 가할 경우에는 문제가 될 수 있다.

- [0047] 본 발명에서 상기 양성 종양 또는 결절은 난소, 유방, 폐, 전립선, 결장, 신장, 피부, 갑상선, 뇌, 심장, 간, 위, 신경, 혈관, 뼈, 상부 소화관 또는 체장의 부위에서 발생한 것일 수 있고, 바람직하게는 갑상선 부위에서 발생한 것일 수 있다.
- [0048] 일반적으로 암 등의 질환을 진단하기 위하여 특정 바이오마커의 발현 수준을 측정할 때, 각 시료에서 발현양을 보정하기 위해서는 GAPDH 등과 같은 하우스 키팅 유전자의 컨트롤을 수행하고 있다. 하지만, lncRNA는 각 조직에서 매우 낮은 수준으로 발현되고, lncRNA 중에서 하우스 키팅 유전자로 알려진 lncRNA는 아직까지 없으며, GAPDH와 같이 종래에 일반적으로 사용되던 하우스 키팅 유전자들은 lncRNA 발현양에 비하여 매우 높은 수준으로 발현되어 lncRNA 정량 결과를 분석하기에 부적합하다. 하지만, 본 발명의 보정 방법을 이용하는 경우, 하우스 키팅 유전자를 이용한 컨트롤 없이도 정상 조직, 양성 종양 또는 결절 조직, 및 암 조직의 각 조직 별 본 발명에 따르는 lncRNA 바이오마커의 발현 수준 차이를 극대화하여, 암을 진단함에 있어서 정확도를 현저히 높일 수 있다.
- [0050] 본 발명의 다른 구현 예에 따르면, 목적하는 개체로부터 분리된 생물학적 시료에서, 제1 lncRNA 및 제2 lncRNA의 발현 수준을 측정하는 단계; 및 측정된 제1 lncRNA 및 제2 lncRNA의 발현 수준을 하기 식 1에 대입하는 단계를 포함하고, 상기 제1 lncRNA 및 제2 lncRNA는 각각 독립적으로 암에서 발현되는 lncRNA인, 암 진단을 위한 정보 제공 방법에 관한 것이다:
- [0051] [식 1]
- [0052]
$$\text{보정된 발현 수준 값} = \frac{\text{제1 lncRNA의 발현 수준}}{\text{제2 lncRNA의 발현 수준}}$$
- [0053] 본 발명에서 상기 "목적하는 개체" 내지 "개체"란, 상기 암의 발병 여부가 불확실한 개체로, 발병 가능성이 높은 개체를 의미하거나, 종양 또는 결절이 발생한 개체로 상기 종양 또는 결절이 양성 종양인지 암인지의 여부가 불확실한 개체를 의미할 수 있으나, 이에 제한되지는 않는다.
- [0054] 본 발명에서 상기 "생물학적 시료"는 개체로부터 얻어지거나 개체로부터 유래된 임의의 물질, 생물학적 체액, 조직 또는 세포를 의미하는 것으로, 예를 들면, 전혈(whole blood), 백혈구(leukocytes), 말초혈액 단핵 세포(peripheral blood mononuclear cells), 백혈구 연층(buffy coat), 혈장(plasma), 혈청(serum), 객담(sputum), 눈물(tears), 점액(mucus), 세비액(nasal washes), 비강 흡인물(nasal aspirate), 호흡(breath), 소변(urine), 정액(semen), 침(saliva), 복강 세척액(peritoneal washings), 복수(ascites), 낭종액(cystic fluid), 뇌척수막 액(meningeal fluid), 양수(amniotic fluid), 선액(glandular fluid), 췌장액(pancreatic fluid), 림프액(lymph fluid), 흉수(pleural fluid), 유두 흡인물(nipple aspirate), 기관지 흡인물(bronchial aspirate), 활액(synovial fluid), 관절 흡인물(joint aspirate), 기관 분비물(organ secretions), 세포(cell), 세포 추출물(cell extract) 또는 뇌척수액(cerebrospinal fluid)을 포함할 수 있고, 바람직하게는 발병 가능성이 높은 환자의 피부를 절개하지 않고 중공침 등을 생체 내 기관에 자입하여 병리조직학적 검사용으로 채취한 샘플(예를 들면, 환자의 조직, 세포, 혈액, 혈청, 혈장, 타액, 객담 또는 복수(ascites) 등)일 수 있으나, 이에 제한되는 것은 아니다.
- [0055] 본 발명에서 상기 제1 lncRNA는 정상 대조군 대비 암에서 발현 수준이 증가하거나, 양성 종양 대비 암에서 발현 수준이 증가한 lncRNA일 수 있다. 여기서 상기 양성 종양에서 상기 제1 lncRNA의 발현 수준은 정상 대조군 대비 증가된 것일 수 있다.
- [0056] 본 발명에서, 상기 제1 lncRNA는 류신 풍부 반복 함유 52 단백질의 안티센스 RNA 1(LRRC52-AS1), 긴 유전자간 비-단백질 코딩 RNA 2082(LINC02082), Unc-5 네트린 수용체 B 안티센스 RNA 1(UNC5B-AS1), 긴 유전자간 비-단백질 코딩 RNA 2471(LINC02471) 및 긴 유전자간 비-단백질 코딩 RNA 2408(LINC02408)으로 이루어진 군에서 선택된 1종 이상일 수 있으나, 이에 제한되지는 않는다.
- [0057] 본 발명에서 상기 제2 lncRNA는 정상 대조군 대비 암에서 발현 수준이 감소하거나, 양성 종양 대비 암에서 발현 수준이 감소한 lncRNA일 수 있다. 여기서 상기 양성 종양에서 상기 제2 lncRNA의 발현 수준은 정상 대조군 대비 감소된 것일 수 있다.
- [0058] 본 발명에서, 상기 제2 lncRNA는 메탈로포스포에스터라아제 도메인 함유 안티센스 RNA 1(MPPED2-AS1), FOXA2의 lncRNA 인접 증강제(LNCNEF), LOC100129129 및 LOC642484으로 이루어진 군에서 선택된 1종 이상일 수 있으나, 이에 제한되지는 않는다.

- [0059] 본 발명의 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LRRC52-AS1 이고, 상기 제2 lncRNA는 MPPED2-AS1 일 수 있으나, 이에 제한되는 것은 아니다.
- [0060] 본 발명의 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LRRC52-AS1 이고, 상기 제2 lncRNA는 LNCNEF 일 수 있으나, 이에 제한되는 것은 아니다.
- [0061] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LRRC52-AS1 이고, 상기 제2 lncRNA는 LOC100129129 일 수 있으나, 이에 제한되는 것은 아니다.
- [0062] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LRRC52-AS1 이고, 상기 제2 lncRNA는 LOC642484 일 수 있으나, 이에 제한되는 것은 아니다.
- [0063] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LINC02082 이고, 상기 제2 lncRNA는 MPPED2-AS1 일 수 있으나, 이에 제한되는 것은 아니다.
- [0064] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LINC02082 이고, 상기 제2 lncRNA는 LNCNEF 일 수 있으나, 이에 제한되는 것은 아니다.
- [0065] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LINC02082 이고, 상기 제2 lncRNA는 LOC100129129 일 수 있으나, 이에 제한되는 것은 아니다.
- [0066] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LINC02082 이고, 상기 제2 lncRNA는 LOC642484 일 수 있으나, 이에 제한되는 것은 아니다.
- [0067] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 UNC5B-AS1 이고, 상기 제2 lncRNA는 MPPED2-AS1 일 수 있으나, 이에 제한되는 것은 아니다.
- [0068] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 UNC5B-AS1 이고, 상기 제2 lncRNA는 LNCNEF 일 수 있으나, 이에 제한되는 것은 아니다.
- [0069] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 UNC5B-AS1 이고, 상기 제2 lncRNA는 LOC100129129 일 수 있으나, 이에 제한되는 것은 아니다.
- [0070] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 UNC5B-AS1 이고, 상기 제2 lncRNA는 LOC642484 일 수 있으나, 이에 제한되는 것은 아니다.
- [0071] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LINC02471 이고, 상기 제2 lncRNA는 MPPED2-AS1 일 수 있으나, 이에 제한되는 것은 아니다.
- [0072] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LINC02471 이고, 상기 제2 lncRNA는 LNCNEF 일 수 있으나, 이에 제한되는 것은 아니다.
- [0073] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LINC02471 이고, 상기 제2 lncRNA는 LOC100129129 일 수 있으나, 이에 제한되는 것은 아니다.
- [0074] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LINC02471 이고, 상기 제2 lncRNA는 LOC642484 일 수 있으나, 이에 제한되는 것은 아니다.
- [0075] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LINC02408 이고, 상기 제2 lncRNA는 MPPED2-AS1 일 수 있으나, 이에 제한되는 것은 아니다.
- [0076] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LINC02408 이고, 상기 제2 lncRNA는 LNCNEF 일 수 있으나, 이에 제한되는 것은 아니다.
- [0077] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LINC02408 이고, 상기 제2 lncRNA는 LOC100129129 일 수 있으나, 이에 제한되는 것은 아니다.
- [0078] 본 발명의 또 다른 일 예시에서, 상기 발현 수준을 측정하는 단계 시, 상기 제1 lncRNA는 LINC02408 이고, 상기 제2 lncRNA는 LOC642484 일 수 있으나, 이에 제한되는 것은 아니다.
- [0079] 본 발명에서, 상기 제1 lncRNA 및 제2 lncRNA의 발현 수준의 측정은 상기 제1 lncRNA 및 제2 lncRNA 각각에 특이적으로 결합하는 프라이머, 프로브 및 안티센스 뉴클레오티드로 이루어진 군에서 선택된 1종 이상을 포함하는

제제를 이용하여 수행될 수 있다.

- [0080] 본 발명의 일 예시에서, 상기 제1 lncRNA 또는 제2 lncRNA에 해당하는 상기 LRRC52-AS1, LINC02082, UNC5B-AS1, LINC02471, LINC02408, MPPED2-AS1, LINCNEF, LOC100129129 및 LOC642484 lncRNA 서열 정보는 이미 알려져 있으므로, 당업자라면 이를 바탕으로 상기 유전자에 특이적으로 결합하는 프라이머, 프로브 또는 안티센스 뉴클레오타이드를 용이하게 디자인할 수 있을 것이다.
- [0081] 본 발명에서, 상기 제1 lncRNA 및 제2 lncRNA의 발현 수준의 측정은 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) 또는 DNA 칩에 의해 수행될 수 있으나 이에 제한되는 것은 아니다.
- [0082] 본 발명에서, 상기 "진단"은 특정 질병 또는 질환에 대한 대상(subject)의 감수성(susceptibility)을 판정하는 것, 대상이 특정 질병 또는 질환을 현재 가지고 있는지 여부를 판정하는 것, 특정 질병 또는 질환에 걸린 대상의 예후(prognosis)(예컨대, 전-전이성 또는 전이성 암 상태의 동정, 암의 단계 결정 또는 치료에 대한 암의 반응성 결정)를 판정하는 것, 테라메트릭스(therametrics)(예컨대, 치료 효능에 대한 정보를 제공하기 위하여 객체의 상태를 모니터링하는 것), 또는 암을 양성 종양 또는 결절과 구별하는 것일 수 있으나, 이에 제한되는 것은 아니다.
- [0083] 본 발명에서, 상기 식 1에 대입하는 단계 시 상기 식 1에서 상기 제1 lncRNA 발현 수준의 값은 목적하는 개체로부터 분리된 어느 일 생물학적 시료로부터 1회 측정된 제1 lncRNA의 발현 수준의 값일 수 있고, 혹은 목적하는 개체로부터 분리된 복수 개의 생물학적 시료로부터 각각 측정된 복수 개의 제1 lncRNA의 발현 수준의 평균 값 또는 중앙 값일 수 있고, 혹은 목적하는 개체로부터 분리된 어느 일 생물학적 시료로부터 복수 회 측정된 복수 개의 제1 lncRNA의 발현 수준의 평균 값 또는 중앙 값일 수 있고, 혹은 목적하는 개체로부터 분리된 복수 개의 생물학적 시료로부터 복수 회 측정된 복수 개의 제1 lncRNA의 발현 수준의 평균 값 또는 중앙 값일 수 있으나, 이에 제한되는 것은 아니다.
- [0084] 본 발명에서, 상기 식 1에 대입하는 단계 시 상기 식 1에서 상기 제2 lncRNA 발현 수준의 값은 목적하는 개체로부터 분리된 어느 일 생물학적 시료로부터 1회 측정된 제2 lncRNA의 발현 수준의 값일 수 있고, 혹은 목적하는 개체로부터 분리된 복수 개의 생물학적 시료로부터 각각 측정된 복수 개의 제2 lncRNA의 발현 수준의 평균 값 또는 중앙 값일 수 있고, 혹은 목적하는 개체로부터 분리된 어느 일 생물학적 시료로부터 복수 회 측정된 복수 개의 제2 lncRNA의 발현 수준의 평균 값 또는 중앙 값일 수 있고, 혹은 목적하는 개체로부터 분리된 복수 개의 생물학적 시료로부터 복수 회 측정된 복수 개의 제2 lncRNA의 발현 수준의 평균 값 또는 중앙 값일 수 있으나, 이에 제한되는 것은 아니다.
- [0085] 본 발명에서, 상기 식 1에 대입하는 단계 시 얻어진 보정된 발현 수준 값이 대조군에 비하여 증가하거나 감소한 경우, 바람직하게는 증가한 경우, 암의 발병 가능성이 높은 것으로 예측할 수 있고, 바람직하게는 개체에 발생한 상기 종양 또는 결절이 암일 가능성이 높은 것으로 예측할 수 있다. 여기서 상기 대조군은 암이 발병하지 않은 개체에서의 보정된 발현 수준 값, 또는 이의 평균 값 또는 중앙 값; 혹은 양성 종양 또는 결절에서의 보정된 발현 수준 값, 또는 이의 평균 값 또는 중앙 값일 수 있으나, 이에 제한되는 것은 아니다.
- [0086] 본 발명에서, 상기 식 1에 대입하는 단계 시 상기 제1 lncRNA는 LRRC52-AS1 이고, 상기 제2 lncRNA는 MPPED2-AS1, LINCNEF, LOC100129129 및 LOC642484으로 이루어진 군에서 선택된 어느 하나이며, 얻어진 보정된 발현 수준의 값이 대조군에 비하여 증가한 경우, 암의 발병 가능성이 높은 것으로 예측할 수 있고, 바람직하게는 개체에 발생한 상기 종양 또는 결절이 암일 가능성이 높은 것으로 예측할 수 있다. 여기서 상기 대조군은 암이 발병하지 않은 개체에서의 보정된 발현 수준 값, 또는 이의 평균 값 또는 중앙 값; 혹은 양성 종양 또는 결절에서의 보정된 발현 수준 값, 또는 이의 평균 값 또는 중앙 값일 수 있으나, 이에 제한되는 것은 아니다.
- [0087] 본 발명에서, 상기 식 1에 대입하는 단계 시 상기 제1 lncRNA는 LINC02082이고, 상기 제2 lncRNA는 MPPED2-AS1, LINCNEF, LOC100129129 및 LOC642484으로 이루어진 군에서 선택된 어느 하나이며, 얻어진 보정된 발현 수준의 값이 대조군에 비하여 증가한 경우, 암의 발병 가능성이 높은 것으로 예측할 수 있고, 바람직하게는 개체에 발생한 상기 종양 또는 결절이 암일 가능성이 높은 것으로 예측할 수 있다. 여기서 상기 대조군은 암이 발병하지 않은 개체에서의 보정된 발현 수준 값, 또는 이의 평균 값 또는 중앙 값; 혹은 양성 종양 또는 결절에서의 보정된 발현 수준 값, 또는 이의 평균 값 또는 중앙 값일 수 있으나, 이에 제한되는 것은 아니다.
- [0088] 본 발명에서, 상기 식 1에 대입하는 단계 시 상기 제1 lncRNA는 UNC5B-AS1 이고, 상기 제2 lncRNA는 MPPED2-

AS1, LNCNEF, LOC100129129 및 LOC642484으로 이루어진 군에서 선택된 어느 하나이며, 얻어진 보정된 발현 수준의 값이 대조군에 비하여 증가한 경우, 암의 발병 가능성이 높은 것으로 예측할 수 있고, 바람직하게는 개체에 발생한 상기 종양 또는 결절이 암일 가능성이 높은 것으로 예측할 수 있다. 여기서 상기 대조군은 암이 발병하지 않은 개체에서의 보정된 발현 수준 값, 또는 이의 평균 값 또는 종양 값; 혹은 양성 종양 또는 결절에서의 보정된 발현 수준 값, 또는 이의 평균 값 또는 종양 값일 수 있으나, 이에 제한되는 것은 아니다.

[0089] 본 발명에서, 상기 식 1에 대입하는 단계 시 상기 제1 lncRNA는 LINC02471 이고, 상기 제2 lncRNA는 MPPED2-AS1, LNCNEF, LOC100129129 및 LOC642484으로 이루어진 군에서 선택된 어느 하나이며, 얻어진 보정된 발현 수준의 값이 대조군에 비하여 증가한 경우, 암의 발병 가능성이 높은 것으로 예측할 수 있고, 바람직하게는 개체에 발생한 상기 종양 또는 결절이 암일 가능성이 높은 것으로 예측할 수 있다. 여기서 상기 대조군은 암이 발병하지 않은 개체에서의 보정된 발현 수준 값, 또는 이의 평균 값 또는 종양 값; 혹은 양성 종양 또는 결절에서의 보정된 발현 수준 값, 또는 이의 평균 값 또는 종양 값일 수 있으나, 이에 제한되는 것은 아니다.

[0090] 본 발명에서, 상기 식 1에 대입하는 단계 시 상기 제1 lncRNA는 LINC02408이고, 상기 제2 lncRNA는 MPPED2-AS1, LNCNEF, LOC100129129 및 LOC642484으로 이루어진 군에서 선택된 어느 하나이며, 얻어진 보정된 발현 수준의 값이 대조군에 비하여 증가한 경우, 암의 발병 가능성이 높은 것으로 예측할 수 있고, 바람직하게는 개체에 발생한 상기 종양 또는 결절이 암일 가능성이 높은 것으로 예측할 수 있다. 여기서 상기 대조군은 암이 발병하지 않은 개체에서의 보정된 발현 수준 값, 또는 이의 평균 값 또는 종양 값; 혹은 양성 종양 또는 결절에서의 보정된 발현 수준 값, 또는 이의 평균 값 또는 종양 값일 수 있으나, 이에 제한되는 것은 아니다.

[0091] 본 발명에서 진단의 대상이 되는 암은 유방암, 난소암, 대장암, 위암, 간암, 췌장암, 자궁경부암, 갑상선암, 부갑상선암, 폐암, 비소세포성폐암, 전립선암, 담낭암, 담도암, 비호지킨 림프종, 호지킨 림프종, 혈액암, 방광암, 신장암, 흑색종, 결장암, 골암, 피부암, 두부암, 자궁암, 직장암, 뇌종양, 향문부근암, 나팔관암종, 자궁내막암종, 질암, 음문암종, 식도암, 소장암, 내분비선암, 부신암, 연조직 육종, 요도암, 음경암, 수노관암, 신장세포 암종, 신장골반 암종, 중추신경계(central nervous system; CNS) 종양, 1차 CNS 림프종, 척수 종양, 뇌간 신경교종 또는 뇌하수체 선종일 수 있으나, 바람직하게는 갑상선암일 수 있다.

[0092] 본 발명에서 상기 양성 종양 또는 결절은 난소, 유방, 폐, 전립선, 결장, 신장, 피부, 갑상선, 뇌, 심장, 간, 위, 신경, 혈관, 뼈, 상부 소화관 또는 췌장의 부위에서 발생한 것일 수 있고, 바람직하게는 갑상선 부위에서 발생한 것일 수 있다.

[0093] 본 발명에서 상기 진단, 종양, 발현 수준을 측정하는 제제 분석 방법에 관한 설명은 앞서 기재된 바와 중복되어 명세서의 과도한 복잡을 피하기 위하여 그 자세한 기재를 생략한다.

발명의 효과

[0094] 본 발명의 보정 방법을 이용하는 경우, 하우스 키핑 유전자를 이용한 컨트롤 없이도 정상 조직, 양성 종양 또는 결절 조직, 및 암 조직의 각 조직 별 본 발명에 따르는 lncRNA 바이오마커의 발현 수준 차이를 극대화하여, 암을 진단함에 있어서 정확도를 현저히 높일 수 있다.

도면의 간단한 설명

[0095] 도 1은 실시예 1에서 본 발명의 일 실시예에 따른 갑상선암 진단을 위한 lncRNA 후보 바이오마커를 선정하기 위한 분석 과정을 나타낸 것이다.

도 2는 실시예 2에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LRRC52-AS1의 발현 수준의 차이를 분석한 결과를 나타낸 것이다.

도 3은 실시예 2에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LINC02082의 발현 수준의 차이를 분석한 결과를 나타낸 것이다.

도 4는 실시예 2에서 갑상선암, 양성 결절 조직 및 주변 정상 갑상선 조직에서의 UNC5B-AS1의 발현 수준의 차이를 분석한 결과를 나타낸 것이다.

도 5는 실시예 2에서 갑상선암, 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LINC02471의 발현 수준의 차이를 분석한 결과를 나타낸 것이다.

도 6은 실시예 2에서 갑상선암, 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LINC02408의 발현 수준의 차이

를 분석한 결과를 나타낸 것이다.

도 7은 실시예 2에서 갑상선암, 양성 결절 조직 및 주변 정상 갑상선 조직에서의 MPPED2-AS1의 발현 수준의 차이를 분석한 결과를 나타낸 것이다.

도 8은 실시예 2에서 갑상선암, 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LNCNEF의 발현 수준의 차이를 분석한 결과를 나타낸 것이다.

도 9는 실시예 2에서 갑상선암, 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LOC100129129의 발현 수준의 차이를 분석한 결과를 나타낸 것이다.

도 10은 실시예 2에서 갑상선암, 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LOC642484의 발현 수준의 차이를 분석한 결과를 나타낸 것이다.

도 11은 실시예 3에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LRRC52-AS1 발현 수준을 MPPED2-AS1 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 12는 실시예 3에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LRRC52-AS1 발현 수준을 LNCNEF 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 13은 실시예 3에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LRRC52-AS1 발현 수준을 LOC100129129 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 14는 실시예 3에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LRRC52-AS1 발현 수준을 LOC642484 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 15는 실시예 4에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LINC02082 발현 수준을 MPPED2-AS1 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 16은 실시예 4에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LINC02082 발현 수준을 LNCNEF 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 17은 실시예 4에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LINC02082 발현 수준을 LOC100129129 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 18은 실시예 4에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LINC02082 발현 수준을 LOC642484 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 19는 실시예 5에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 UNC5B-AS1 발현 수준을 MPPED2-AS1 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 20은 실시예 5에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 UNC5B-AS1 발현 수준을 LNCNEF 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 21은 실시예 5에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 UNC5B-AS1 발현 수준을 LOC100129129 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 22는 실시예 5에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 UNC5B-AS1 발현 수준을 LOC642484 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 23은 실시예 6에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LINC02471 발현 수준을 MPPED2-AS1 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

다.

도 24는 실시예 6에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LINC02471 발현 수준을 LNCNEF 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 25는 실시예 6에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LINC02471 발현 수준을 LOC100129129 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 26은 실시예 6에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LINC02471 발현 수준을 LOC642484 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 27은 실시예 7에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LINC02408 발현 수준을 MPPED2-AS1 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 28은 실시예 7에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LINC02408 발현 수준을 LNCNEF 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 29는 실시예 7에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LINC02408 발현 수준을 LOC100129129 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

도 30은 실시예 7에서 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서의 LINC02408 발현 수준을 LOC642484 발현 수준으로 나눈 후 얻어진 보정된 발현 수준의 값의 차이를 분석한 결과를 나타낸 것이다.

발명을 실시하기 위한 구체적인 내용

[0096] 이하, 본 발명을 하기의 실시예에 의해 상세히 설명한다. 단, 하기 실시예는 본 발명을 예시하는 것일 뿐, 본 발명의 내용이 하기 실시예에 의해 한정되는 것은 아니다.

[0098] 실시예

[0100] [실시예 1] 갑상선암 진단을 위한 lncRNA 후보 선정

[0101] 갑상선암 진단을 위한 lncRNA 후보를 선정하기 위하여, 도 1과 같이 ncRNA 데이터와 갑상선암 환자의 ncRNA 발현을 비교분석 하였다. 구체적으로는 암에서의 ncRNA 아틀라스(The Atlas of ncRNA in Cancer; TANRIC) 갑상선암 데이터에서 갑상선암 환자 59명의 갑상선암 조직과 주변 정상 갑상선 조직에서 12,727개의 ncRNA를 비교하여 갑상선암 조직에서 2배 이상으로 의미 있게 증가한 901개 ncRNA와 갑상선암 조직에서 0.5배 이하로 의미 있게 감소한 1,710개 ncRNA를 선별하였다. 그 중 갑상선암 조직과 주변 정상 갑상선 조직에서 0.5 이상 차이가 나는 값을 가지고 미국 국가생물공학센터(National Center for Biotechnology Information; NCBI)에 주석이 달린 각각 상위 5개의 lncRNA를 후보로 선정하였다.

[0102] 그 중에서 도 1에 나타난 것처럼 갑상선암 조직에서 발현이 의미 있게 증가한 ncRNA 중 LRRC52-AS1, LINC02471, LINC02082, UNC5B-AS1 및 LINC02408, 그리고 갑상선암 조직에서 발현이 의미 있게 감소한 ncRNA 중 MPPED2-AS1, LNCNEF, LOC642484 및 LOC100129129를 갑상선암 진단을 위한 최종 후보로 선정하였다.

[0104] [실시예 2] lncRNA 후보에서 갑상선암 진단 능력 확인

[0105] [2-1] 갑상선암 조직에서 발현이 증가한 ncRNA 분석

[0106] 갑상선암 환자 59명으로부터 분리한 갑상선암 조직(암 조직)과 주변 정상 갑상선 조직(정상 조직)에 있어서, 상기 실시예 1에서 최종 선정된, 갑상선암 조직에서 발현이 의미 있게 증가한 5개의 lncRNA의 발현 수준을 측정하

여 그 결과를 하기 표 1에 나타내었다. 표 1에서, 'T59'는 갑상선암 조직에서 각 lncRNA의 평균 발현 수준 값이고, 'N59'는 주변 정상 갑상선 조직에서 각 lncRNA의 평균 발현 수준 값에 해당한다.

표 1

명칭	유전자 ID	P 값	N59	T59	T59-N59	T59/N59
LRRC52-AS1	ENSG00000237463.1	9.5782E-10	0.016932	2.76254	2.745608	163.1579
LINC02471	ENSG00000223914.1	8.23157E-15	0.250378	10.0865	9.836122	40.28509
LINC02082	ENSG00000242268.2	1.31763E-05	0.016668	0.659175	0.642507	39.54662
UNC5B-AS1	ENSG00000237512.2	6.17203E-06	0.035556	1.38515	1.349594	38.95697
LINC02408	ENSG00000203585.3	2.10601E-06	0.020193	0.666412	0.646219	33.00213

상기 표 1에 나타난 바와 같이, 실시예 1에서 주변 정상 갑상선 조직에 비하여 갑상선암 조직에서 발현 수준이 의미 있게 증가한 것으로 확인된 lncRNA인 LRRC52-AS1, LINC02471, LINC02082, UNC5B-AS1 및 LINC02408은 모두 주변 정상 갑상선 조직에 비하여 갑상선암 조직에서 발현 수준의 평균 값이 높게 발현되는 바, 갑상선암을 진단하는데 유효한 마커임을 확인할 수 있었다.

[2-2] 갑상선암 조직에서 발현이 감소한 ncRNA 분석

갑상선암 환자 59명으로부터 분리한 갑상선암 조직(암 조직)과 주변 정상 갑상선 조직(정상 조직)에 있어서, 상기 실시예 1에서 최종 선정된, 갑상선암 조직에서 발현이 의미 있게 감소한 4개의 lncRNA의 발현 수준을 측정하여 그 결과를 하기 표 2에 나타내었다. 표 2에서, 'T59'는 갑상선암 조직에서 각 lncRNA의 평균 발현 수준 값이고, 'N59'는 주변 정상 갑상선 조직에서 각 lncRNA의 평균 발현 수준 값에 해당한다.

표 2

명칭	유전자 ID	P 값	N59	T59	T59-N59	T59/N59
MPPED2-AS1	ENSG00000254489.1	1.24675E-21	1.42859	0.153562	-1.27503	0.107492
LNCNEF	ENSG00000237396.1	3.2035E-09	1.7388	0.201443	-1.53736	0.115852
LOC642484	ENSG00000206129.3	3.82645E-16	0.663489	0.11583	-0.54766	0.174577
LOC100129129	ENSG00000255020.1	6.71445E-13	4.23287	1.21616	-3.01671	0.287313

상기 표 2에 나타난 바와 같이, 실시예 1에서 갑상선암 조직에서 발현이 의미 있게 감소된 것으로 확인된 lncRNA인 MPPED2-AS1, LNCNEF, LOC642484 및 LOC100129129는 모두 정상 조직에 비하여 갑상선암에서 평균 값이 낮게 발현되는 바, 갑상선암을 진단하는데 유효한 마커임을 확인할 수 있었다.

[2-3] LRRC52-AS1 lncRNA의 갑상선암 진단 능력 확인

LRRC52-AS1의 갑상선암 여부 및 악성도에 대한 진단 능력을 확인하기 위하여, 병원에서 수술을 받은 9명의 갑상선 결절 환자에서 얻은 6개의 갑상선암(갑상선 악성 결절) 조직, 5개의 양성 결절 조직, 9개의 주변 정상 갑상선 조직에서 상기 lncRNA의 발현 수준을 확인하여 그 결과를 도 2에 나타내었다.

도 2에서 나타난 것처럼, LRRC52-AS1은 갑상선 양성 결절 조직이나 주변 정상 갑상선 조직에 비하여 갑상선암(갑상선 악성 결절) 조직에서 발현 수준이 의미 있게 증가하였다. 이를 통하여 상기 LRRC52-AS1은 갑상선 종양 또는 결절이 갑상선암(악성)인지 여부를 진단하는데 유효한 마커임을 확인할 수 있었다.

[2-4] LINC02082 lncRNA의 갑상선암 진단 능력 확인

LINC02082의 갑상선암 여부 및 악성도에 대한 진단 능력을 확인하기 위하여, 병원에서 수술을 받은 9명의 갑상선 결절 환자에서 얻은 6개의 갑상선암(갑상선 악성 결절) 조직, 5개의 양성 결절 조직, 9개의 주변 정상 갑상

선 조직에서 상기 lncRNA의 발현 수준을 확인하여 그 결과를 도 3에 나타내었다.

[0123] 도 3에서 나타난 것처럼, LINC02082는 갑상선 양성 결절이나 주변 정상 갑상선 조직에 비하여 갑상선암(갑상선 악성 결절) 조직에서 발현 수준이 의미 있게 증가하였다. 이를 통하여 상기 LINC02082는 갑상선 종양 또는 결절 조직이 갑상선암(악성)인지 여부를 진단하는데 유효한 마커임을 확인할 수 있었다.

[0125] [2-5] UNC5B-AS1 lncRNA의 갑상선암 진단 능력 확인

[0126] UNC5B-AS1의 갑상선암 여부 및 악성도에 대한 진단 능력을 확인하기 위하여, 병원에서 수술을 받은 9명의 갑상선 결절 환자에서 얻은 6개의 갑상선암(갑상선 악성 결절) 조직, 5개의 양성 결절 조직, 9개의 주변 정상 갑상선 조직에서 상기 lncRNA의 발현 수준을 확인하여 그 결과를 도 4에 나타내었다.

[0127] 도 4에서 나타난 것처럼, UNC5B-AS1은 갑상선 양성 결절이나 주변 정상 갑상선 조직에 비하여 갑상선암(갑상선 악성 결절) 조직에서 발현 수준이 의미 있게 증가하였다. 이를 통하여 상기 UNC5B-AS1은 갑상선 결절 또는 종양 조직이 갑상선암(악성)인지 여부를 진단하는데 유효한 마커임을 확인할 수 있었다.

[0129] [2-6] LINC02471 lncRNA의 갑상선암 진단 능력 확인

[0130] LINC02471의 갑상선암 여부 및 악성도에 대한 진단 능력을 확인하기 위하여, 병원에서 수술을 받은 9명의 갑상선 결절 환자에서 얻은 6개의 갑상선암(갑상선 악성 결절) 조직, 5개의 양성 결절 조직, 9개의 주변 정상 갑상선 조직에서 상기 lncRNA의 발현 수준을 확인하여 그 결과를 도 5에 나타내었다.

[0131] 도 5에서 나타난 것처럼, LINC02471은 주변 정상 갑상선 조직에 비하여 갑상선암(갑상선 악성 결절) 조직에서 발현 수준이 증가하였다. 이를 통하여 상기 LINC02471은 갑상선암인지 여부를 진단하는데 유효한 마커임을 확인할 수 있었다.

[0133] [2-7] LINC02408 lncRNA의 갑상선암 진단 능력 확인

[0134] LINC02408의 갑상선암 여부 및 악성도에 대한 진단 능력을 확인하기 위하여, 병원에서 수술을 받은 9명의 갑상선 결절 환자에서 얻은 6개의 갑상선암(갑상선 악성 결절) 조직, 5개의 양성 결절 조직, 9개의 주변 정상 갑상선 조직에서 상기 lncRNA의 발현 수준을 확인하여 그 결과를 도 6에 나타내었다.

[0135] 도 6에서 나타난 것처럼 LINC02408은 주변 정상 갑상선 조직에 비하여 갑상선암(갑상선 악성 결절) 조직에서 발현 수준이 증가하였다. 이를 통하여 상기 LINC02408은 갑상선암인지 여부를 진단하는데 유효한 마커임을 확인할 수 있었다.

[0137] [2-8] MPPED2-AS1 lncRNA의 갑상선암 진단 능력 확인

[0138] MPPED2-AS1의 갑상선암 여부 및 악성도에 대한 진단 능력을 확인하기 위하여, 병원에서 수술을 받은 9명의 갑상선 결절 환자에서 얻은 6개의 갑상선암(갑상선 악성 결절) 조직, 5개의 양성 결절 조직, 9개의 주변 정상 갑상선 조직에서 상기 lncRNA의 발현 수준을 확인하여 그 결과를 도 7에 나타내었다.

[0139] 도 7에서 나타난 것처럼, MPPED2-AS1은 갑상선 양성 결절 조직이나 주변 정상 갑상선 조직에 비하여 갑상선암(갑상선 악성 결절) 조직에서 발현 수준이 의미 있게 감소하였다. 이를 통하여 상기 MPPED2-AS1은 갑상선 종양 또는 결절 조직이 갑상선암(악성)인지 여부를 진단하는데 유효한 마커임을 확인할 수 있었다.

[0141] [2-9] LNCNEF lncRNA의 갑상선암 진단 능력 확인

[0142] LNCNEF의 갑상선암 여부 및 악성도에 대한 진단 능력을 확인하기 위하여, 병원에서 수술을 받은 9명의 갑상선 결절 환자에서 얻은 6개의 갑상선암(갑상선 악성 결절) 조직, 5개의 양성 결절 조직, 9개의 주변 정상 갑상선 조직에서 상기 lncRNA의 발현 수준을 확인하여 그 결과를 도 8에 나타내었다.

[0143] 도 8에서 나타난 것처럼, LNCNEF는 갑상선 양성 결절이나 주변 정상 갑상선 조직에 비하여 갑상선암(갑상선 악

성 결절) 조직에서 발현 수준이 의미 있게 감소하였다. 이를 통하여 상기 LNCNEF는 갑상선 종양 또는 결절 조직이 갑상선암(악성)인지 여부를 진단하는데 유효한 마커임을 확인할 수 있었다.

[0145] [2-10] LOC100129129 lncRNA의 갑상선암 진단 능력 확인

[0146] LOC100129129의 갑상선암 여부 및 악성도에 대한 진단 능력을 확인하기 위하여, 병원에서 수술을 받은 9명의 갑상선 결절 환자에서 얻은 6개의 갑상선암(갑상선 악성 결절) 조직, 5개의 양성 결절 조직, 9개의 주변 정상 갑상선 조직에서 상기 lncRNA의 발현 수준을 확인하여 그 결과를 도 9에 나타내었다.

[0147] 도 9에서 나타난 것처럼, LOC100129129는 갑상선 양성 결절 조직이나 주변 정상 갑상선 조직에 비하여 갑상선암(갑상선 악성 결절) 조직에서 발현 수준이 의미 있게 감소하였다. 이를 통하여 상기 LOC100129129는 갑상선 종양 또는 결절 조직이 갑상선암(악성)인지 여부를 진단하는데 유효한 마커임을 확인할 수 있었다.

[0149] [2-11] LOC642484 lncRNA의 갑상선암 진단 능력 확인

[0150] LOC642484의 갑상선암 여부 및 악성도에 대한 진단 능력을 확인하기 위하여, 병원에서 수술을 받은 9명의 갑상선 결절 환자에서 얻은 6개의 갑상선암(갑상선 악성 결절) 조직, 5개의 양성 결절 조직, 9개의 주변 정상 갑상선 조직에서 상기 lncRNA의 발현 수준을 확인하여 그 결과를 도 10에 나타내었다.

[0151] 도 10에서 나타난 것처럼, LOC642484는 주변 정상 갑상선 조직에 비하여 갑상선암(갑상선 악성 결절) 조직에서 발현 수준이 감소하였다. 이를 통하여 상기 LOC642484는 갑상선암인지 여부를 진단하는데 유효한 마커임을 확인할 수 있었다.

[0153] [실시예 3] lncRNA 발현 수준을 보정한 경우에서 갑상선암 진단 능력 확인(1)

[0154] [3-1] LRRC52-AS1/MPPED2-AS1의 갑상선암 스크리닝 검사 능력 확인

[0155] LRRC52-AS1 발현 수준을 MPPED2-AS1 발현 수준으로 나누는 보정 방법을 이용할 경우, 갑상선암 여부 및 악성도에 대한 스크리닝 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LRRC52-AS1 발현 수준과 MPPED2-AS1 발현 수준을 각각 측정한 후, LRRC52-AS1 발현 수준을 MPPED2-AS1 발현 수준으로 나누는 보정된 발현 수준 값을 도 11에 나타내었다. 또한, 갑상선암 조직과 주변 정상 조직에서 LRRC52-AS1 발현 수준과 MPPED2-AS1 발현 수준을 측정한 데이터에서, LRRC52-AS1 발현 수준을 MPPED2-AS1 발현 수준으로 나누어 얻어진 보정된 발현 수준의 값이 기준치 이상일 때 상기 결절을 갑상선암으로 진단할 경우, 상기 진단의 민감도(Sensitivity)와 음성 예측치(Negative predictive value; NPV)를 확인하여, 그 결과를 하기 표 3에 나타내었다.

표 3

[0157]

lncRNA	RPKM 컷오프 값	민감도	음성예측치
LRRC52-AS1	0.035766786	88%	89%
MPPED2-AS1	0.577267417	95%	95%
LRRC52-AS1/MPPED2-AS1	0.020621914	95%	94%

[0158] 도 11 및 상기 표 3에 나타난 바와 같이, 단일 LRRC52-AS1의 발현 수준의 값이나, 단일 MPPED2-AS1 발현 수준의 값만을 이용하여 갑상선암을 진단하는 경우에 비하여, LRRC52-AS1 발현 수준을 MPPED2-AS1 발현 수준으로 나누어 보정된 발현 수준 값을 이용하여 갑상선암을 진단하는 경우가 진단의 민감도 및 음성 예측치가 매우 증가하는 것을 볼 수 있었다. 이를 통해 상기 LRRC52-AS1 발현 수준을 MPPED2-AS1 발현 수준으로 나누는 보정에 의해 갑상선암과 갑상선 양성 결절 사이의 발현 수준의 차이를 극대화하여 갑상선 결절이 갑상선암(악성)인지 여부를 진단하는 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.

[0160] [3-2] LRRC52-AS1/LNCNEF의 갑상선암 스크리닝 검사 능력 확인

[0161] LRRC52-AS1 발현 수준을 LNCNEF 발현 수준으로 나누는 보정 방법을 이용할 경우 갑상선암 여부 및 악성도에 대한 스크리닝 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LRRC52-AS1 발현 수준과 LNCNEF 발현 수준을 각각 측정된 후, LRRC52-AS1 발현 수준을 LNCNEF 발현 수준으로 나누는 보정된 발현 수준 값을 도 12에 나타내었다. 또한, 갑상선암 조직과 주변 정상 조직에서 LRRC52-AS1 발현 수준과 LNCNEF 발현 수준을 측정된 데이터에서, LRRC52-AS1 발현 수준을 LNCNEF 발현 수준으로 나누는 보정된 발현 수준 값이 기준치 이상일 때 상기 결절을 갑상선암으로 진단할 경우, 상기 진단의 민감도(Sensitivity)와 음성 예측치(NPV)를 확인하여, 그 결과를 하기 표 4에 나타내었다.

표 4

lncRNA	RPKM 컷오프 값	민감도	음성예측치
LRRC52-AS1	0.035766786	88%	89%
LNCNEF	0.2183943845	85%	87%
LRRC52-AS1/LNCNEF	0.039972309	93%	93%

[0164] 도 12 및 상기 표 4에 나타난 바와 같이, 단일 LRRC52-AS1 발현 수준의 값이나 단일 LNCNEF 발현 수준의 값을 이용하여 갑상선암을 진단하는 경우에 비하여, LRRC52-AS1 발현 수준을 LNCNEF 발현 수준으로 나누어 보정된 발현 수준 값을 이용하여 갑상선암을 진단하는 경우가 진단의 민감도 및 음성 예측치가 매우 증가하는 것을 볼 수 있었다. 이를 통해 상기 LRRC52-AS1 발현 수준을 LNCNEF 발현 수준으로 나누는 보정에 의해 갑상선암과 갑상선 양성 결절 사이의 발현 수준의 차이를 극대화함으로써, 갑상선 결절이 갑상선암(악성)인지 여부를 진단함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.

[0166] [3-3] LRRC52-AS1/LOC100129129의 갑상선암 스크리닝 검사 능력 확인

[0167] LRRC52-AS1 발현 수준을 LOC100129129 발현 수준으로 나누는 보정 방법을 이용할 경우 갑상선암 여부 및 악성도에 대한 스크리닝 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LRRC52-AS1 발현 수준과 LOC100129129 발현 수준을 각각 측정된 후, LRRC52-AS1 발현 수준을 LOC100129129 발현 수준으로 나누는 보정된 발현 수준 값을 도 13에 나타내었다. 또한, 갑상선암 조직과 주변 정상 조직에서 LRRC52-AS1 발현 수준과 LOC100129129 발현 수준을 측정된 데이터에서, LRRC52-AS1 발현 수준을 LOC100129129 발현 수준으로 나누는 보정된 발현 수준 값이 기준치 이상일 때 상기 결절을 갑상선암으로 진단할 경우, 상기 진단의 민감도(Sensitivity)와 음성 예측치(NPV)를 확인하여, 그 결과를 하기 표 5에 나타내었다.

표 5

lncRNA	RPKM 컷오프 값	민감도	음성예측치
LRRC52-AS1	0.035766786	88%	89%
LOC100129129	1.9145184725	81%	84%
LRRC52-AS1/LOC100129129	0.0114036165	90%	90%

[0170] 도 13 및 상기 표 5에 나타난 바와 같이, 단일 LRRC52-AS1 발현 수준의 값이나 단일 LOC100129129 발현 수준의 값을 이용하여 갑상선암을 진단하는 경우에 비하여, LRRC52-AS1 발현 수준을 LOC100129129 발현 수준으로 나누어 보정된 발현 수준 값을 이용하여 갑상선암을 진단하는 경우가 진단의 민감도 및 음성 예측치가 매우 증가하는 것을 볼 수 있었다. 이를 통해 상기 LRRC52-AS1 발현 수준을 LOC100129129 발현 수준으로 나누는 보정에 의해 갑상선암과 갑상선 양성 결절 사이의 발현 수준의 차이를 극대화함으로써, 갑상선 결절이 갑상선암(악성)인지 여부를 진단함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.

[0172] [3-4] LRRC52-AS1/LOC642484의 갑상선암 스크리닝 검사 능력 확인

[0173] LRRC52-AS1 발현 수준을 LOC642484 발현 수준으로 나눔으로써 보정한 경우 갑상선암 여부 및 악성도에 대한 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LRRC52-AS1 및 LOC642484의 발현 수준을 측정 한 후, LRRC52-AS1의 발현 수준을 LOC642484 발현 수준으로 나눈 보정된 발현 수준 값을 도 14에 나타내었다.

[0174] 도 14에 나타난 것처럼, LRRC52-AS1 발현 수준을 LOC642484 발현 수준으로 나눈 보정 값은, 갑상선양성 결절 조직이나 주변 정상 갑상선 조직에 비하여 갑상선암 조직에서 의미 있게 증가하였는 바, LRRC52-AS1 발현 수준을 LOC642484 발현 수준으로 나눈 보정 값을 이용해 갑상선 결절이 갑상선암(악성)인지 여부를 유효하게 진단할 수 있음을 알 수 있었다.

[0176] [실시예 4] lncRNA 발현 수준을 보정한 경우에서 갑상선암 진단 능력 확인(2)

[0177] [4-1] LINC02082/MPPED2-AS1의 갑상선암 제외진단 능력 확인

[0178] LINC02082 발현 수준을 MPPED2-AS1 발현 수준으로 나누는 보정 방법을 이용하여, 갑상선암 여부 및 악성도에 대한 제외진단(rule out; R/O) 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LINC02082 발현 수준과 MPPED2-AS1 발현 수준을 각각 측정한 후, LINC02082 발현 수준을 MPPED2-AS1 발현 수준으로 나눈 보정된 발현 수준 값을 도 15에 나타내었다. 또한, 갑상선암 조직과 주변 정상 갑상선 조직에서 LINC02082 발현 수준과 MPPED2-AS1 발현 수준을 측정한 데이터에서, LINC02082 발현 수준을 MPPED2-AS1 발현 수준으로 나눈 보정된 발현 수준 값이 기준치 이상일 때 상기 결절을 암으로 진단할 경우, 상기 진단의 특이도(Specificity)와 양성 예측치(Positive predictive value; PPV)를 확인하여, 그 결과를 하기 표 6에 나타내었다.

표 6

[0180]

lncRNA	RPKM 컷오프 값	특이도	양성 예측치
LINC02082	0.0662616705	95%	94%
MPPED2-AS1	0.577267417	92%	92%
LINC02082/MPPED2-AS1	0.2809314735	100%	100%

[0181] 도 15 및 상기 표 6에 나타난 것처럼, 단일 LINC02082 발현 수준의 값이나 단일 MPPED2-AS1 발현 수준의 값을 이용하여 갑상선암을 진단한 경우에 비하여 LINC02082 발현 수준을 MPPED2-AS1 발현 수준으로 나누어 보정된 발현 수준 값을 이용하여 갑상선암을 진단하는 경우, 진단의 특이도 및 양성 예측치가 100%에 달하는 것을 볼 수 있었다. 이를 통해 상기 LINC02082 발현 수준을 MPPED2-AS1 발현 수준으로 나누는 보정에 의해 갑상선암 조직과 갑상선 양성 결절 조직 사이의 발현 수준의 차이를 극대화함으로써, 갑상선 결절이 갑상선암(악성)인지 여부를 제외진단(R/O)함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.

[0183] [4-2] LINC02082/LNCNEF의 갑상선암 제외진단 능력 확인

[0184] LINC02082 발현 수준을 LNCNEF 발현 수준으로 나누는 보정 방법을 이용할 경우 갑상선암 여부 및 악성도에 대한 제외진단(rule out; R/O) 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LINC02082 발현 수준과 LNCNEF 발현 수준을 각각 측정한 후, LINC02082 발현 수준을 LNCNEF 발현 수준으로 나눈 보정된 발현 수준 값을 도 16에 나타내었다. 또한, 갑상선암 조직과 주변 정상 갑상선 조직에서 LINC02082 발현 수준과 LNCNEF 발현 수준을 측정한 데이터에서, LINC02082 발현 수준을 LNCNEF 발현 수준으로 나눈 보정된 발현 수준 값이 기준치 이상일 때 상기 결절을 암으로 진단할 경우, 상기 진단의 특이도(Specificity)와 양성 예측치(PPV)를 확인하여, 그 결과를 하기 표 7에 나타내었다.

표 7

[0186]

lncRNA	RPKM 컷오프 값	특이도	양성 예측치
LINC02082	0.0662616705	95%	94%
LNCNEF	0.2183943845	100%	100%
LINC02082/LNCNEF	0.2491690695	100%	100%

[0187]

도 16 및 상기 표 7에 나타난 것처럼, 단일 LINC02082의 발현 수준의 값이나 단일 LNCNEF의 발현 수준의 값만을 이용하여 갑상선암을 진단하는 경우에 비하여 LINC02082 발현 수준을 MPED2-AS1 발현 수준으로 나누어 보정된 발현 수준 값을 이용하여 갑상선암을 진단하는 경우가 진단의 특이도 및 양성 예측치가 매우 증가하는 것을 볼 수 있었다. 이를 통해 상기 LINC02082 발현 수준을 LNCNEF 발현 수준으로 나누는 보정에 의해 갑상선암 조직과 갑상선 양성 결절 조직 사이의 발현 수준의 차이를 극대화함으로써, 갑상선 결절이 갑상선암(악성)인지 여부를 제외진단(R/O)함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.

[0189]

[4-3] LINC02082/LOC100129129의 갑상선암 진단 능력 확인

[0190]

LINC02082 발현 수준을 LOC100129129 발현 수준으로 나눈 보정 방법을 이용할 경우의 갑상선암 여부 및 악성도에 대한 진단 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LINC02082 발현 수준과 LOC100129129 발현 수준을 각각 측정한 후, LINC02082 발현 수준을 LOC100129129 발현 수준으로 나눈 보정된 발현 수준 값을 도 17에 나타내었다.

[0191]

도 17에 나타난 것처럼, LINC02082 발현 수준을 LOC100129129 발현 수준으로 나누는 보정을 수행함으로써 갑상선암 조직과, 갑상선 양성 결절 조직이나 주변 정상 갑상선 조직 사이에서의 발현 수준의 값의 차이를 현저히 증가시킬 수 있는 바, 하우스 키팅 유전자를 사용하지 않고도 본 발명에 따르는 보정 방법에 의해 갑상선암 조직과 갑상선 양성 종양 조직 또는 주변 정상 갑상선 조직의 각 군 사이의 발현 수준의 차이를 극대화하여 갑상선 종양 또는 결절이 갑상선암(악성)인지 여부를 진단하는 정확도를 향상시킬 수 있다.

[0193]

[4-4] LINC02082/LOC642484의 갑상선암 진단 능력 확인

[0194]

LINC02082 발현 수준을 LOC642484 발현 수준으로 나눔으로써 보정한 경우 갑상선암 여부 및 악성도에 대한 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LINC02082 및 LOC642484 발현 수준을 측정한 후, LINC02082 발현 수준을 LOC642484 발현 수준으로 나눈 보정된 발현 수준 값을 도 18에 나타내었다.

[0195]

도 18에 나타난 것처럼, LINC02082 발현 수준을 LOC642484 발현 수준으로 나누는 보정을 수행함으로써, 하우스 키팅 유전자를 사용하지 않고도 갑상선암 조직과 갑상선 양성 종양 조직 또는 주변 정상 갑상선 조직의 각 군 사이의 발현 수준의 차이를 극대화하여 갑상선 종양 또는 결절이 갑상선암(악성)인지 여부를 진단함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.

[0197]

[실시예 5] lncRNA 발현 수준을 보정한 경우에서 갑상선암 진단 능력 확인(3)

[0198]

[5-1] UNC5B-AS1/MPED2-AS1의 갑상선암 스크리닝 검사 능력 확인

[0199]

UNC5B-AS1 발현 수준을 MPED2-AS1 발현 수준으로 나누는 보정 방법을 이용할 경우 갑상선암 여부 및 악성도에 대한 스크리닝 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 UNC5B-AS1 발현 수준과 MPED2-AS1 발현 수준을 각각 측정한 후, UNC5B-AS1 발현 수준을 MPED2-AS1 발현 수준으로 나눈 보정된 발현 수준 값을 도 19에 나타내었다. 또한, 갑상선암 조직과 주변 정상 조직에서 UNC5B-AS1 발현 수준과 MPED2-AS1 발현 수준을 측정한 데이터에서, UNC5B-AS1 발현 수준을 MPED2-AS1 발현 수준으로 나눈 보정된 발현 수준 값이 기준치 이상일 때 상기 결절을 갑상선암으로 진단할 경우, 상기 진단의 민감도 (Sensitivity)와 음성 예측치(NPV)를 확인하여, 그 결과를 하기 표 8에 나타내었다.

표 8

lncRNA	RPKM 컷오프 값	민감도	음성 예측치
UNC5B-AS1	0.089951671	88%	89%
MPPED2-AS1	0.577267417	95%	95%
UNC5B-AS1/MPPED2-AS1	0.1133624365	95%	95%

도 19 및 상기 표 8에 나타난 것처럼, 단일 UNC5B-AS1의 발현 수준의 값이나 단일 MPPED2-AS1의 발현 수준의 값을 이용하여 갑상선암을 진단하는 경우에 비하여 UNC5B-AS1 발현 수준을 MPPED2-AS1 발현 수준으로 나누어 보정된 발현 수준 값을 이용하여 갑상선암을 진단하는 경우가 진단의 민감도 및 음성 예측치가 매우 증가하는 것을 볼 수 있었다. 이를 통해 상기 UNC5B-AS1 발현 수준을 MPPED2-AS1 발현 수준으로 나누는 보정에 의해 갑상선암 조직과 갑상선 양성 결절 조직 사이의 발현 수준의 차이를 극대화함으로써, 갑상선 종양 또는 결절이 갑상선암(악성)인지 여부를 진단함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.

[5-2] UNC5B-AS1/LNCNEF의 갑상선암 스크리닝 검사 능력 확인

UNC5B-AS1 발현 수준을 LNCNEF 발현 수준으로 나누는 보정 방법을 이용할 경우 갑상선암 여부 및 악성도에 대한 스크리닝 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 UNC5B-AS1 발현 수준과 LNCNEF 발현 수준을 각각 측정된 후, UNC5B-AS1 발현 수준을 LNCNEF 발현 수준으로 나눈 보정된 발현 수준 값을 도 20에 나타내었다. 또한, 갑상선암 조직과 주변 정상 조직에서 UNC5B-AS1 발현 수준과 LNCNEF 발현 수준을 측정된 데이터에서, UNC5B-AS1 발현 수준을 LNCNEF 발현 수준으로 나눈 보정된 발현 수준 값이 기준치 이상일 때 상기 결절을 갑상선암으로 진단할 경우, 상기 진단의 민감도(Sensitivity)와 음성 예측치(NPV)를 확인하여, 그 결과를 하기 표 9에 나타내었다.

표 9

lncRNA	RPKM 컷오프 값	민감도	음성 예측치
UNC5B-AS1	0.089951671	88%	89%
LNCNEF	0.2183943845	85%	87%
UNC5B-AS1/LNCNEF	0.242830794	92%	92%

도 20 및 상기 표 9에 나타난 것처럼, 단일 UNC5B-AS1 발현 수준의 값이나 단일 LNCNEF 발현 수준의 값을 이용하여 갑상선암을 진단한 경우에 비하여 UNC5B-AS1 발현 수준을 LNCNEF 발현 수준으로 나누어 보정된 발현 수준 값을 이용하여 갑상선암을 진단한 경우가 진단의 특이도 및 음성 예측치가 92%에 달하는 것을 볼 수 있었다. 이를 통해 상기 UNC5B-AS1 발현 수준을 LNCNEF 발현 수준으로 나누는 보정에 의해 갑상선암 조직과 갑상선 양성 결절 조직 사이의 발현 수준의 차이를 극대화함으로써, 갑상선 종양 또는 결절이 갑상선암(악성)인지 여부를 진단함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.

[5-3] UNC5B-AS1/LOC100129129의 갑상선암 스크리닝 검사 능력 확인

UNC5B-AS1 발현 수준을 LOC100129129 발현 수준으로 나누는 보정 방법을 이용할 경우 갑상선암 여부 및 악성도에 대한 스크리닝 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 UNC5B-AS1 발현 수준과 LOC100129129 발현 수준을 각각 측정된 후, UNC5B-AS1 발현 수준을 LOC100129129 발현 수준으로 나눈 보정된 발현 수준 값을 도 21에 나타내었다. 또한, 갑상선암 조직과 주변 정상 조직에서 UNC5B-AS1 발현 수준과 LOC100129129 발현 수준을 측정된 데이터에서, UNC5B-AS1 발현 수준을 LOC100129129 발현 수준으로 나눈 보정된 발현 수준 값이 기준치 이상일 때 상기 결절을 갑상선암으로 진단할 경우, 상기 진단의 민감도(Sensitivity)와 음성 예측치(NPV)를 확인하여, 그 결과를 하기 표 10에 나타내었다.

표 10

lncRNA	RPKM 컷오프 값	민감도	음성 예측치
UNC5B-AS1	0.089951671	88%	89%
LOC100129129	1.9145184725	81%	84%
UNC5B-AS1/LOC100129129	0.0316275305	90%	90%

[0213]

[0214]

도 21 및 상기 표 10에 나타난 것처럼, 단일 UNC5B-AS1 발현 수준의 값이나 단일 LOC100129129 발현 수준의 값을 이용하여 갑상선암을 진단하는 경우에 비하여 UNC5B-AS1 발현 수준을 LOC100129129 발현 수준으로 나누어 보정된 발현 수준 값을 이용하여 갑상선암을 진단하는 경우가 진단의 특이도 및 음성 예측치가 90%에 달하는 것을 볼 수 있었다. 이를 통해 상기 UNC5B-AS1 발현 수준을 LOC100129129 발현 수준으로 나누는 보정에 의해 갑상선암 조직과 갑상선 양성 결절 조직 사이의 발현 수준의 차이를 극대화함으로써, 갑상선 종양 또는 결절이 갑상선암(악성)인지 여부를 진단함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.

[0216]

[5-4] UNC5B-AS1/LOC642484의 갑상선암 진단 능력 확인

[0217]

UNC5B-AS1 발현 수준을 LOC642484 발현 수준으로 나눔으로써 보정한 경우 갑상선암 여부 및 악성도에 대한 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 UNC5B-AS1 발현 수준 및 LOC642484의 발현 수준을 측정 후, UNC5B-AS1 발현 수준을 LOC642484 발현 수준으로 나눈 보정된 발현 수준 값을 도 22에 나타내었다.

[0218]

도 22에 나타난 것처럼, 상기 UNC5B-AS1 발현 수준을 LOC642484 발현 수준으로 나누는 보정을 수행함으로써, 하우스 키핑 유전자를 사용하지 않고도 갑상선암 조직과 갑상선 양성 종양 조직 또는 주변 정상 갑상선 조직의 각 군 사이의 발현 수준의 차이를 극대화하여 갑상선 종양 또는 결절이 갑상선암(악성)인지 여부를 진단함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.

[0220]

[실시예 6] lncRNA 발현 수준을 보정한 경우에서 갑상선암 진단 능력 확인(4)

[0221]

[6-1] LINC02471/MPPED2-AS1의 갑상선암 진단 능력 확인

[0222]

LINC02471 발현 수준을 MPPED2-AS1 발현 수준으로 나눔으로써 보정한 경우 갑상선암 여부 및 악성도에 대한 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LINC02471 발현 수준 및 MPPED2-AS1 발현 수준을 측정 후, LINC02471 발현 수준을 MPPED2-AS1 발현 수준으로 나눈 보정된 발현 수준 값을 도 23에 나타내었다.

[0223]

도 23에 나타난 것처럼, LINC02471 발현 수준을 MPPED2-AS1 발현 수준으로 나누는 보정을 수행함으로써, 하우스 키핑 유전자를 사용하지 않고도 갑상선암 조직과 갑상선 양성 종양 조직 또는 주변 정상 갑상선 조직의 각 군 사이의 발현 수준의 차이를 극대화하여 갑상선 종양 또는 결절이 갑상선암(악성)인지 여부를 진단함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.

[0225]

[6-2] LINC02471/LNCNEF의 갑상선암 진단 능력 확인

[0226]

LINC02471 발현 수준을 LNCNEF 발현 수준으로 나눔으로써 보정한 경우 갑상선암 여부 및 악성도에 대한 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LINC02471 발현 수준 및 LNCNEF 발현 수준을 측정 후, LINC02471 발현 수준을 LNCNEF 발현 수준으로 나눈 보정된 발현 수준 값을 도 24에 나타내었다.

[0227]

도 24에 나타난 것처럼, LINC02471 발현 수준을 LNCNEF 발현 수준으로 나누는 보정을 수행함으로써, 하우스 키핑 유전자를 사용하지 않고도 갑상선암 조직과 갑상선 양성 결절 조직 또는 주변 정상 갑상선 조직의 각 군 사이의 발현 수준의 차이를 극대화하여 갑상선 종양 또는 결절이 갑상선암(악성)인지 여부를 진단함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.

- [0229] [6-3] LINC02471/LOC100129129의 갑상선암 진단 능력 확인
- [0230] LINC02471 발현 수준을 LOC100129129 발현 수준으로 나눔으로써 보정한 경우 갑상선암 여부 및 악성도에 대한 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LINC02471 발현 수준 및 LOC100129129 발현 수준을 측정 한 후, LINC02471 발현 수준을 LOC100129129 발현 수준으로 나눈 보정된 발현 수준 값을 도 25에 나타내었다.
- [0231] 도 25에 나타난 것처럼, LINC02471 발현 수준을 LOC100129129 발현 수준으로 나누는 보정을 수행함으로써, 하우스 키핑 유전자를 사용하지 않고도 갑상선암 조직과 갑상선 양성 결절 조직 또는 주변 정상 갑상선 조직의 각 군 사이의 발현 수준의 차이를 극대화하여 갑상선 종양 또는 결절이 갑상선암(악성)인지 여부를 진단함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.
- [0233] [6-4] LINC02471/LOC642484의 갑상선암 진단 능력 확인
- [0234] LINC02471 발현 수준을 LOC642484 발현 수준으로 나눔으로써 보정한 경우 갑상선암 여부 및 악성도에 대한 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LINC02471 및 LOC642484 발현 수준을 측정 한 후, LINC02471 발현 수준을 LOC642484 발현 수준으로 나눈 보정된 발현 수준 값을 도 26에 나타내었다.
- [0235] 도 26에 나타난 것처럼, LINC02471 발현 수준을 LOC642484 발현 수준으로 나누는 보정을 수행함으로써, 하우스 키핑 유전자를 사용하지 않고도 갑상선암 조직과 갑상선 양성 결절 조직 또는 주변 정상 갑상선 조직의 각 군 사이의 발현 수준의 차이를 극대화하여 갑상선 종양 또는 결절이 갑상선암(악성)인지 여부를 진단함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.
- [0237] [실시에 7] lncRNA 발현 수준을 보정한 경우에서 갑상선암 진단 능력 확인(5)
- [0238] [7-1] LINC02408/MPPED2-AS1의 갑상선암 진단 능력 확인
- [0239] LINC02408 발현 수준을 MPPED2-AS1 발현 수준으로 나눔으로써 보정한 경우 갑상선암 여부 및 악성도에 대한 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LINC02408 발현 수준 및 MPPED2-AS1 발현 수준을 측정 한 후, LINC02408 발현 수준을 MPPED2-AS1 발현 수준으로 나눈 보정된 발현 수준 값을 도 27에 나타내었다.
- [0240] 도 27에 나타난 것처럼, LINC02408 발현 수준을 MPPED2-AS1 발현 수준으로 나누는 보정을 수행함으로써, 하우스 키핑 유전자를 사용하지 않고도 갑상선암 조직과 갑상선 양성 결절 조직 또는 주변 정상 갑상선 조직의 각 군 사이의 발현 수준의 차이를 극대화하여 갑상선 종양 또는 결절이 갑상선암(악성)인지 여부를 진단함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.
- [0242] [7-2] LINC02408/LNCNEF의 갑상선암 진단 능력 확인
- [0243] LINC02408 발현 수준을 LNCNEF 발현 수준으로 나눔으로써 보정한 경우 갑상선암 여부 및 악성도에 대한 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LINC02408 발현 수준 및 LNCNEF 발현 수준을 측정 한 후, LINC02408 발현 수준을 LNCNEF 발현 수준으로 나눈 보정된 발현 수준 값을 도 28에 나타내었다.
- [0244] 도 28에 나타난 것처럼, LINC02408 발현 수준을 LNCNEF 발현 수준으로 나누는 보정을 수행함으로써, 하우스 키핑 유전자를 사용하지 않고도 갑상선암 조직과 갑상선 양성 결절 조직 또는 주변 정상 갑상선 조직의 각 군 사이의 발현 수준의 차이를 극대화하여 갑상선 종양 또는 결절이 갑상선암(악성)인지 여부를 진단함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.
- [0246] [7-3] LINC02408/LOC100129129의 갑상선암 진단 능력 확인
- [0247] LINC02408 발현 수준을 LOC100129129 발현 수준으로 나눔으로써 보정한 경우 갑상선암 여부 및 악성도에 대한

검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LINC02408 발현 수준 및 LOC100129129 발현 수준을 측정 한 후, LINC02408 발현 수준을 LOC100129129 발현 수준으로 나눈 보정된 발현 수준 값을 도 29에 나타내었다.

[0248] 도 29에 나타난 것처럼, LINC02408 발현 수준을 LOC100129129 발현 수준으로 나누는 보정을 수행함으로써, 하우스 키핑 유전자를 사용하지 않고도 갑상선암 조직과 갑상선 양성 종양 조직 및 주변 정상 갑상선 조직 각 군 사이의 발현 수준의 차이를 극대화하여 갑상선 종양 또는 결절이 갑상선암(악성)인지 여부를 진단함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.

[0250] [7-4] LINC02408/LOC642484의 갑상선암 진단 능력 확인

[0251] LINC02408 발현 수준을 LOC642484 발현 수준으로 나눔으로써 보정한 경우 갑상선암 여부 및 악성도에 대한 검사 능력을 확인하였다. 갑상선암 조직, 갑상선 양성 결절 조직 및 주변 정상 갑상선 조직에서 LINC02408 및 LOC642484 발현 수준을 측정 한 후, LINC02408 발현 수준을 LOC642484 발현 수준으로 나눈 보정된 발현 수준 값을 도 30에 나타내었다.

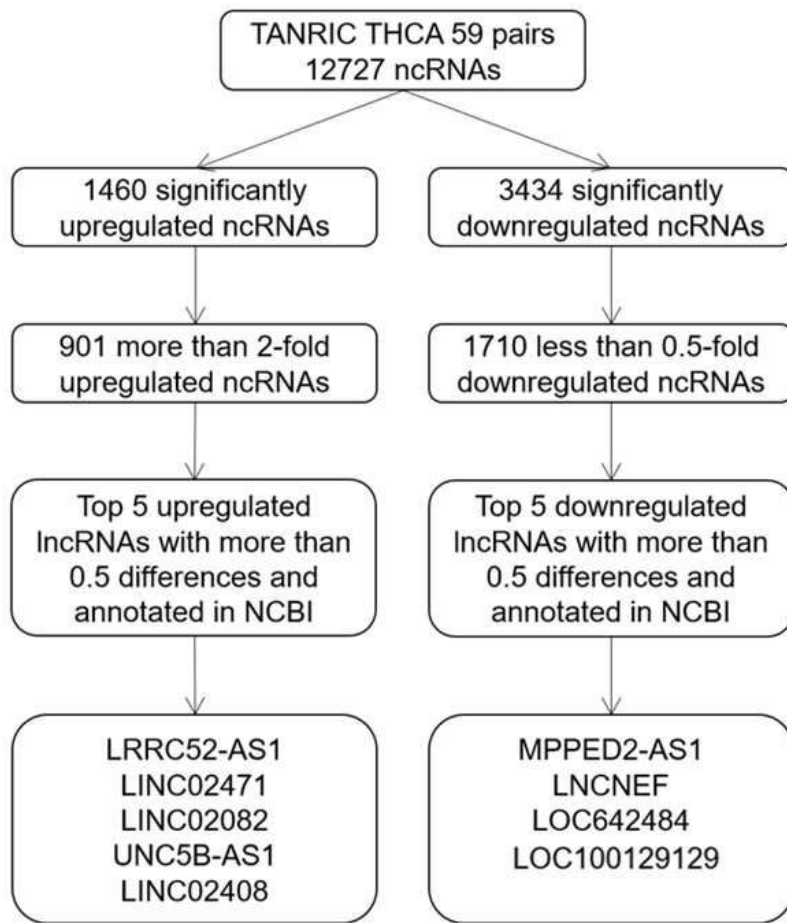
[0252] 도 30에 나타난 것처럼, LINC02408 발현 수준을 LOC642484 발현 수준으로 나누는 보정을 수행함으로써, 하우스 키핑 유전자를 사용하지 않고도 갑상선암 조직과 갑상선 양성 결절 조직 또는 주변 정상 갑상선 조직의 각 군 사이의 발현 수준의 차이를 극대화하여 갑상선 종양 또는 결절이 갑상선암(악성)인지 여부를 진단함에 있어 정확도를 현저히 향상시킬 수 있음을 알 수 있었다.

[0253]

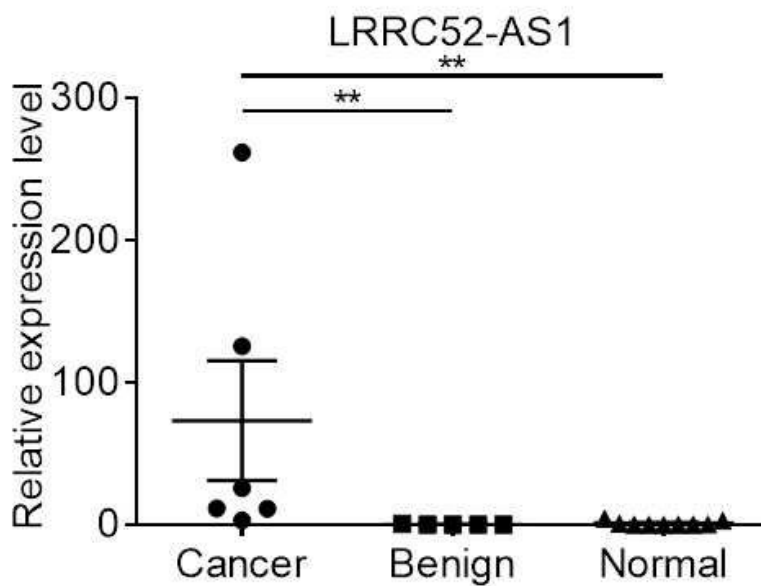
[0254] 이처럼 본 발명의 lncRNA 발현 수준을 보정한 값들을 이용하는 경우 높은 민감도와 특이도로 갑상선암 여부 내지 갑상선 종양 또는 결절의 악성 여부를 진단할 수 있고, 특히 양의 보정을 위한 하우스 키핑 유전자 컨트롤을 생략할 수 있으며, 갑상선암, 양성 종양 또는 결절, 및 정상 조직의 각 군에서의 발현 수준 값 간에 차이를 높여 진단의 정확도를 높일 수 있었다.

도면

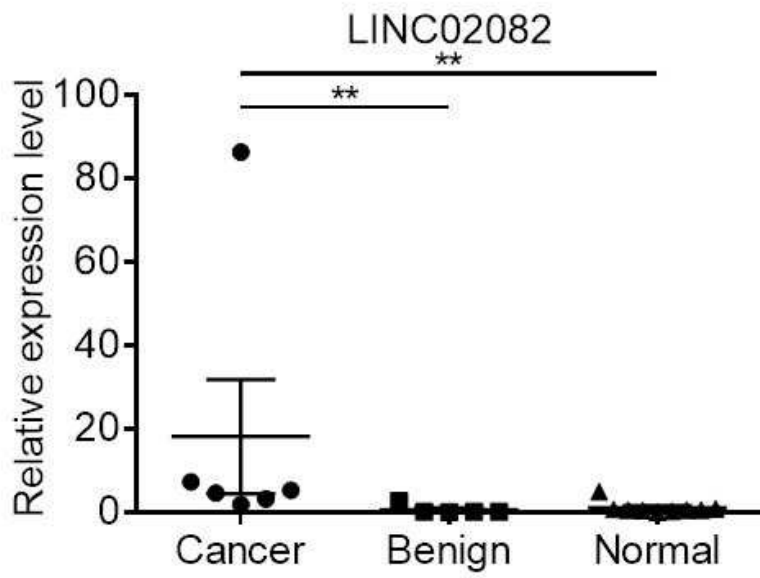
도면1



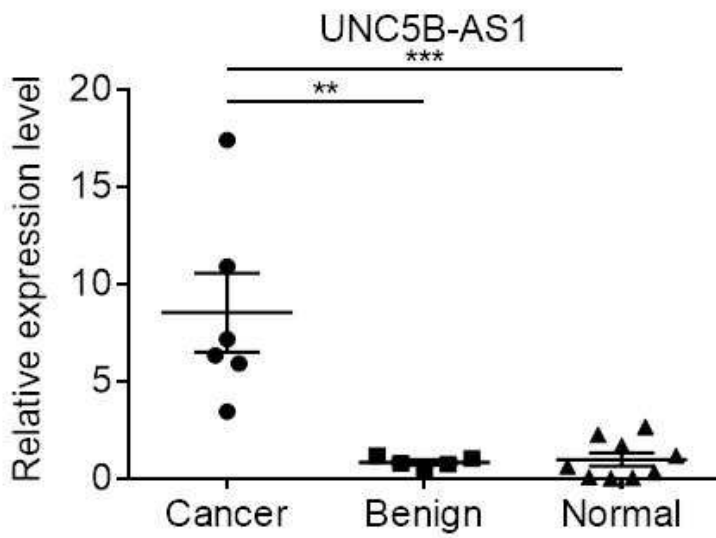
도면2



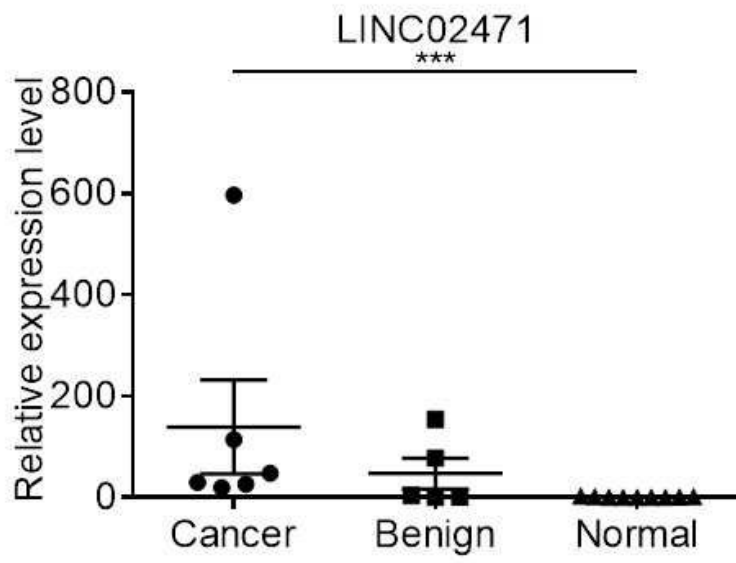
도면3



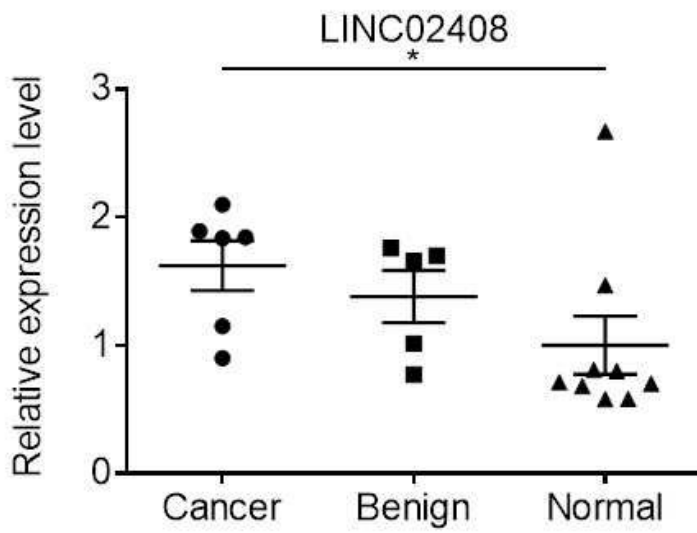
도면4



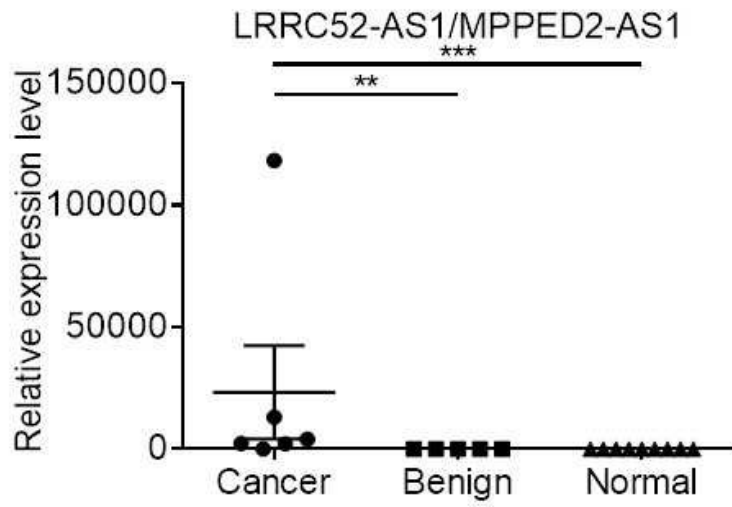
도면5



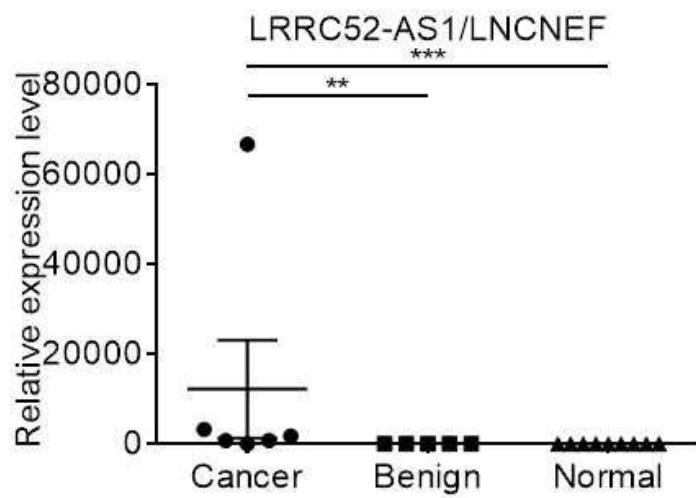
도면6



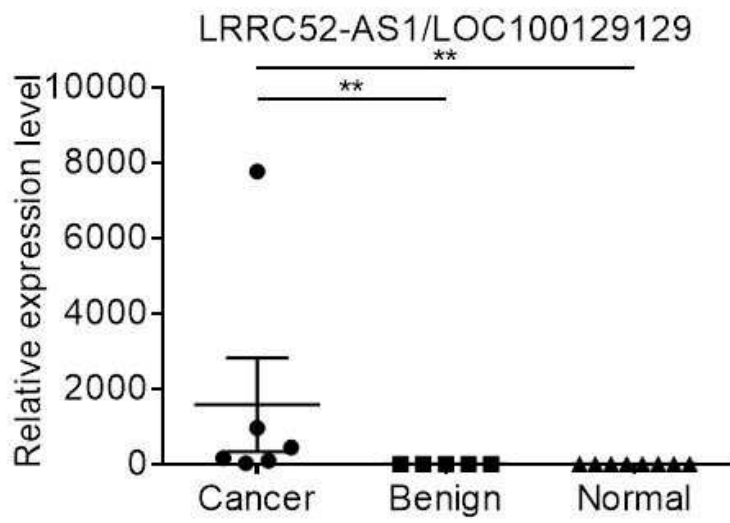
도면11



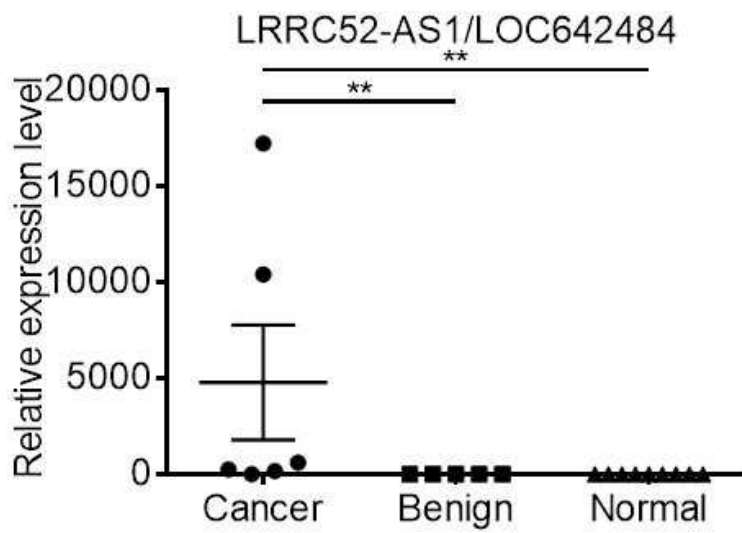
도면12



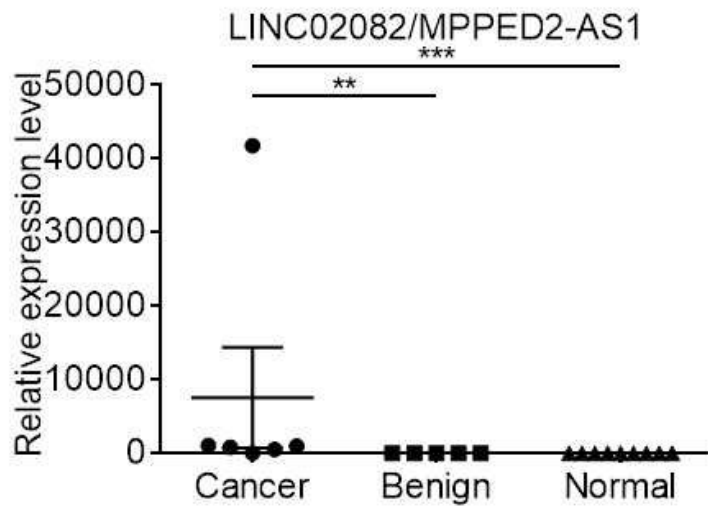
도면13



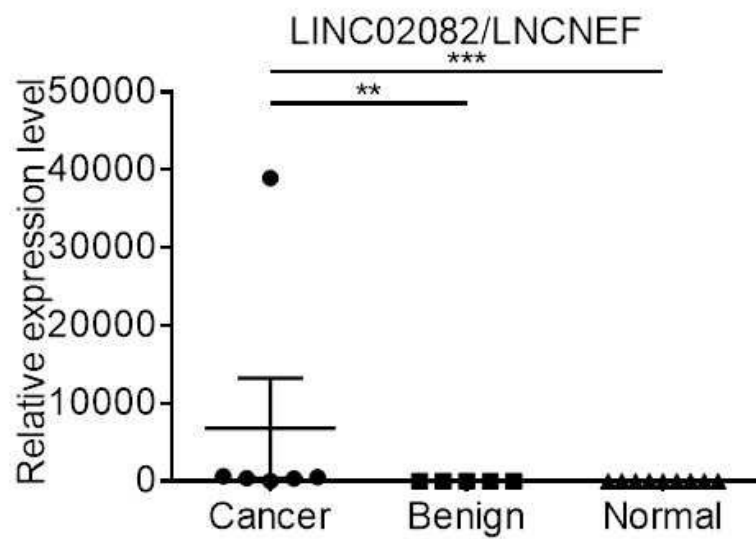
도면14



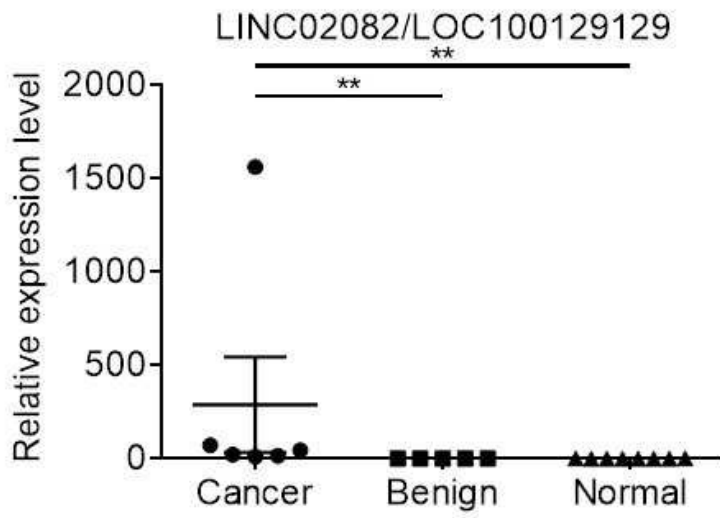
도면15



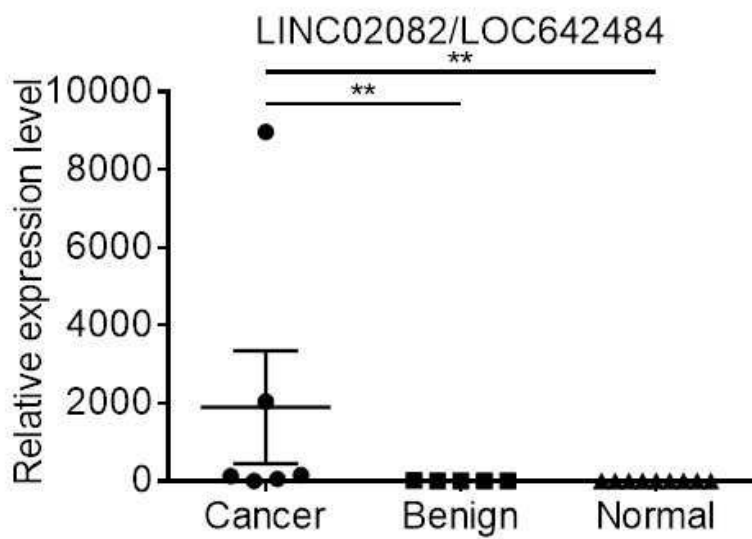
도면16



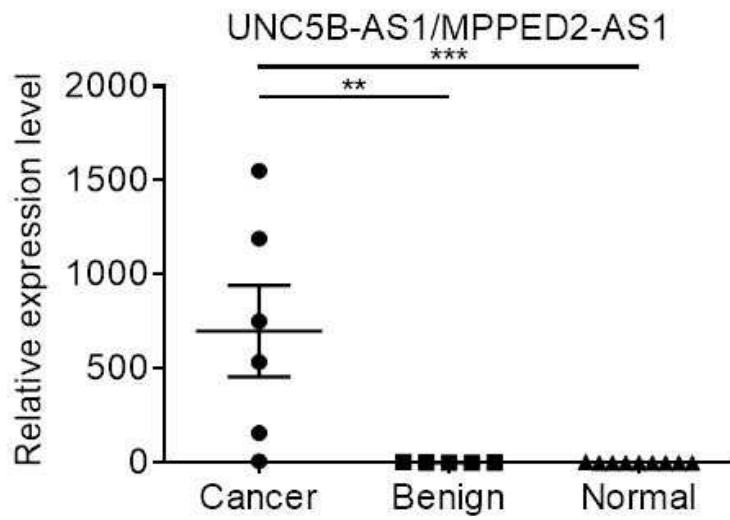
도면17



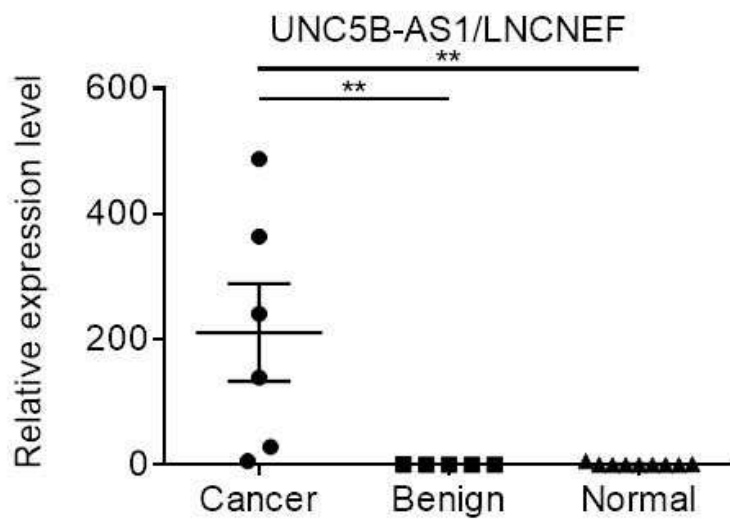
도면18



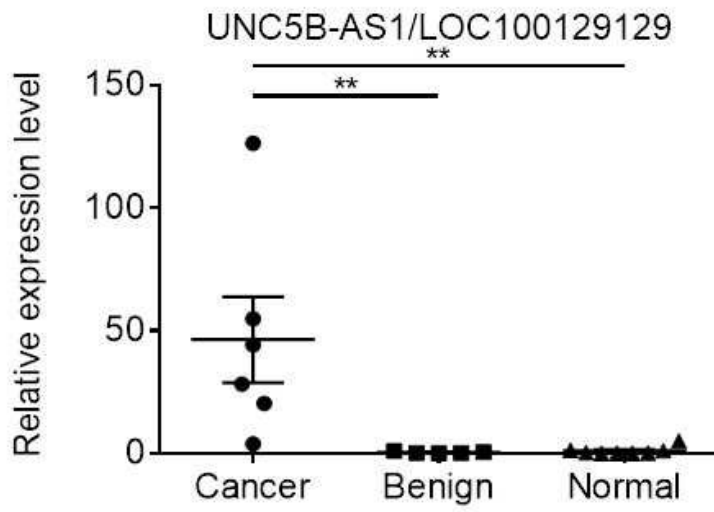
도면19



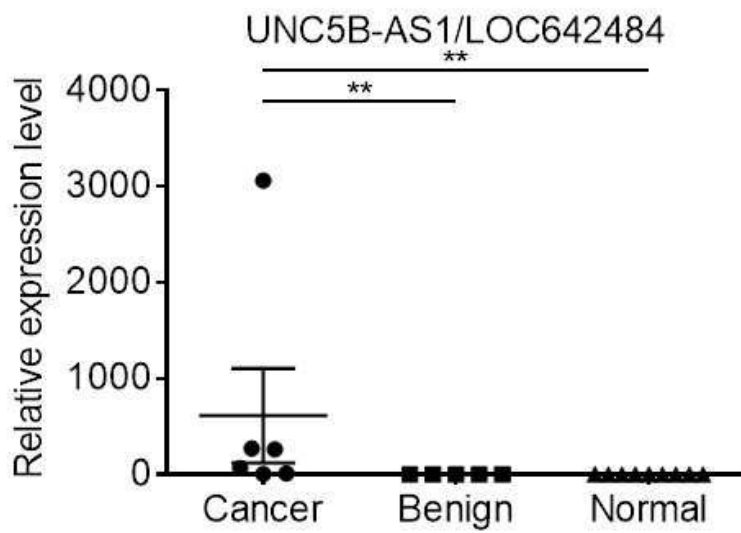
도면20



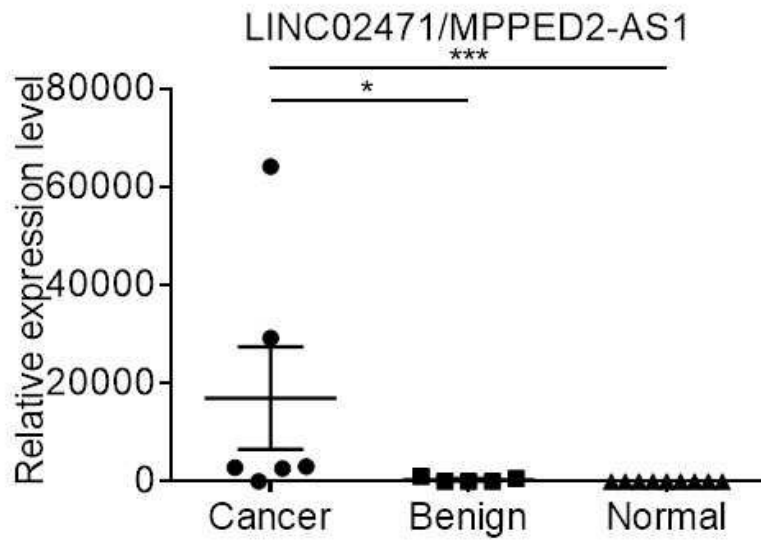
도면21



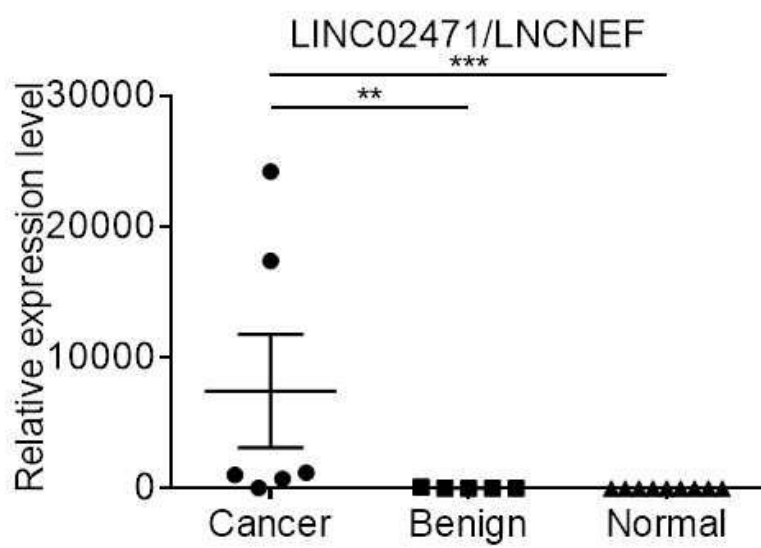
도면22



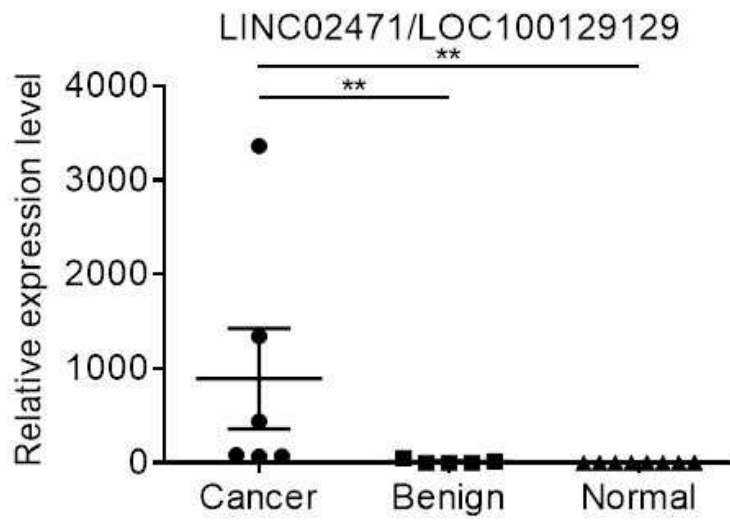
도면23



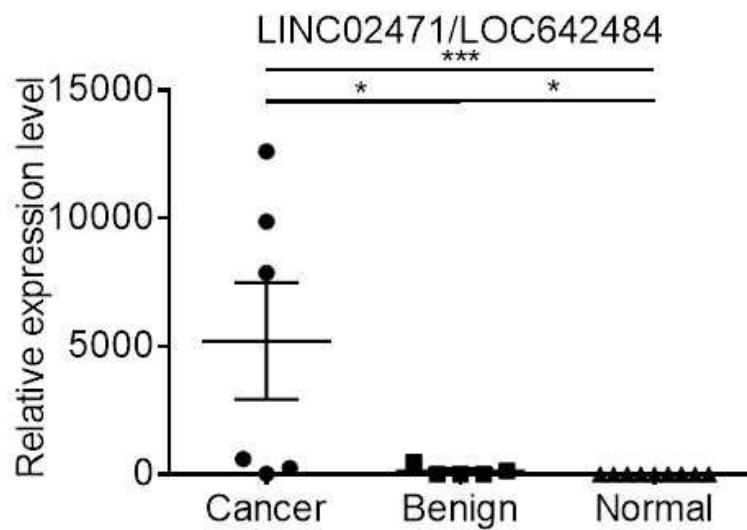
도면24



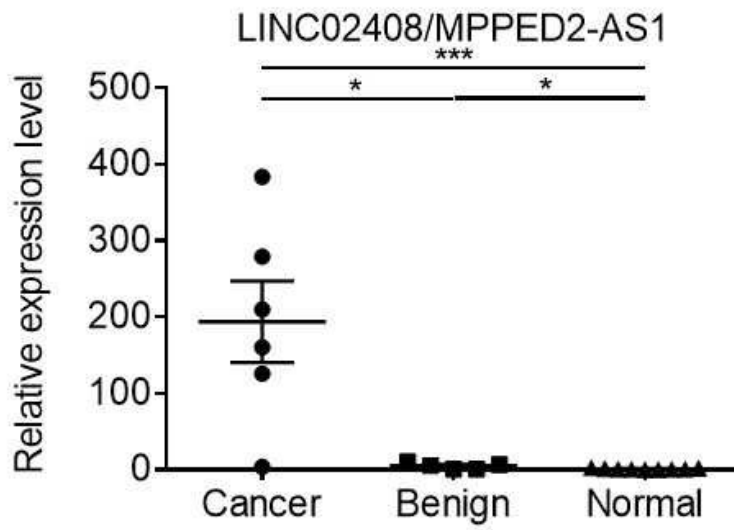
도면25



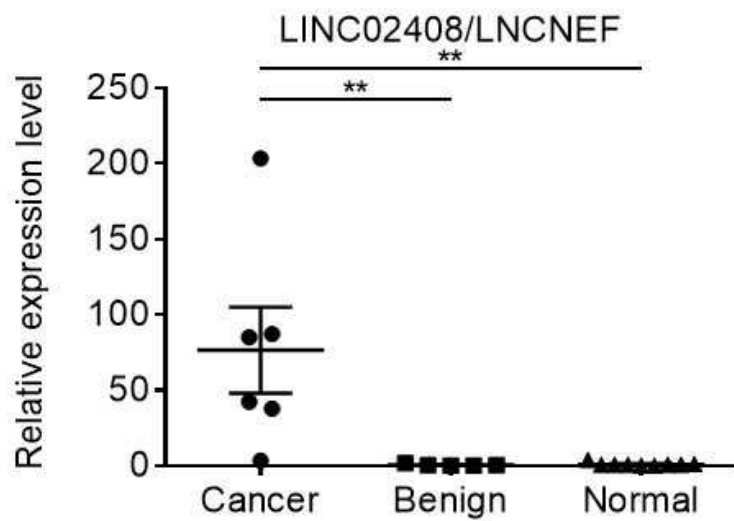
도면26



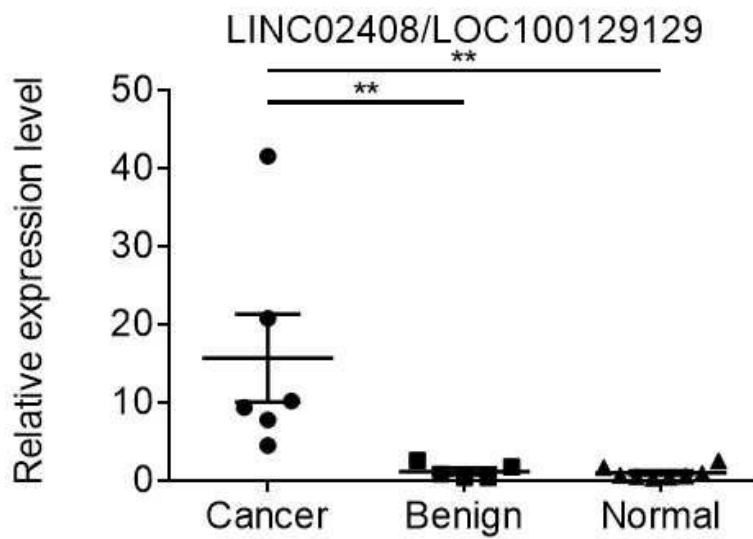
도면27



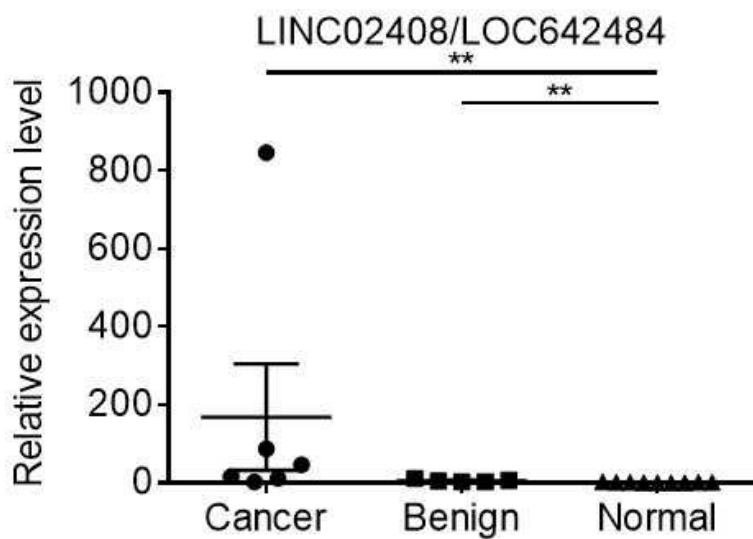
도면28



도면29



도면30



서열 목록

- <110> Industry-Academic Cooperation Foundation, Yonsei University
- <120> Method for correcting the expression level of gene to increase the accuracy of thyroid cancer diagnosis
- <130> PDPB204079
- <160> 9
- <170> KoPatent In 3.0
- <210> 1
- <211> 1803

<212> DNA

<213> Homo sapiens

<400> 1

agcccgctcac ctigcaacgg atgggttagg gcttgccgat cttgttcccc tagtaacctc	60
tgcgcatggg gacaatggag agctaggcca ggatgatggc cctgtggatg gcagtggcca	120
cctccttggg gcacttaaca cccagaccga cgtgaccgtt gtagcccgca atggccacaa	180
acgccttgaa cctggttcgc tggccggctc cggtctgctt ctgcaccggc ataactttca	240
aaacctaggc ctcgagagag gccccccagg aaaaaatcaa tgatctcaga ttccgtgatg	300
ggcagggaga agagacagat ctctccagg gaagtgatct tcatgtgctt gaccaggcag	360
cccagcttgc tgacgggcat ccactcctta tccttggcct tgcctccgag agctccgtgg	420
cttcggcccc agcctggctc acggcagtga ccccgccctc ggatgccact gccgaaacct	480
ccacggaagc caccatggtt ccccatccca gggccccggg gcctccggga ctccccccgc	540
aaccaatgc cggtcgcgc ggcgtcatcc gccatttggg gttttctcgg agaagaagct	600
acatgtggaa tatattaacg caacaaattg tgcgtacttg ctagacatta ggcatgatag	660
caagtgctag ggttacaag tattgctttt tggggaacct atggttgttg gggagaaacc	720
agaagatcag tcatcacagt gcggcaagat cagacttggt gtttgttcag gtcttgggag	780
ccgtcactgc agtgagaatg gacccacaa tgtctcataa ggggatctgc aaggcagccc	840
taaattccaa gcaggacatg taactcctca atggacaaat ggacactgga ccctggttga	900
aggaacctgt tgttcttatg ttggagggca gaaaccaag agggctggag ccaatccata	960
actcaggcac gtggccacat ctaacagcaa gggagtctgg aaaacacagt tcagctgtgt	1020
gaaggatatt tataaacatt tagcaagatc caaaggaatg gaccatagca cagccctgtt	1080
ttcttttagc gctgactcct gctgtctacc agatcctcgt ggtacaatga ctgggttaga	1140
ctggcttgtt gaagtcactc tctggatctg agcaacgtg gctgagccgt gatgaagtat	1200
tcctatccat ccactatgct tcacagtggg aatctctcac tgttcatcac cgctgcccta	1260
gcagccgcag ggattcttca cccatggttg cttaggttta taagcaagga agctgaatga	1320
agtccttttc ctctctggat gagcccttg gttcttgga gagtcttaa atcaccccta	1380
aatggcttcc aagcactgaa gcccagatg ttagctggcc tgtgtgaggc cttagacag	1440
aatgccacag ccacacccca aggatgctga aggggatttc cctcctctgt ggtttttctc	1500
ctttcttccc gcagacttcc tctacgtgc taacatcttc tagtcaaaca actctccttt	1560
tcaaaggac caggcacagt ttctgcttat cctgagtag cagtgttcag ttctctgcca	1620
gcctgtggag tttttcaaac aagccaatca cctcctctg tgggaactga gaagacccc	1680

accctctcga tactacaaag cccctgcttg ttccctagtg caactcccat gtggtcctgt 1740
ggagcatatg gtatcttctt cttctgggca gtgagcatat gtgactaata aactatctat 1800
ctc 1803

<210> 2
<211> 745
<212> DNA
<213> Homo sapiens
<400> 2

taccatcatg gtcactgca acctccgcct cccgggagaa ttgcttgaac cctggaggcg 60
gaggttgccg tgagctaaca ttgcgccctt acactccagc ctgccaacag agcaagactc 120
cgtctcaaaa ataaaaaag gatctacttt tgagaatact gtcattggtt gataaaatac 180
agcagaaact tggaagaag ctattttcca tattcacctg gcatgtttgg agatgattgc 240
catcaacttt gtgactgtga aggagaaacc ttctgccacc caaaaactga aaaatgcctc 300
tgccccgtg ggagaactgg agccagatgt gatgctggtt agtctctgt attgggcgtt 360

cagcatggat gtigccittt tctgacctca agacaactat ccaactttcc ctgtgtcctt 420
cacgatgcaa ccactaaagc tctactgaag tctctaattc tgagaaggca ccttaaaaat 480
ttgaatctaa taacttgctt ggatcatatg gctcctaagt ggagcagaac tcaggtttgc 540
ctcattccag aggccacgtt cttgactaca gcatgcatgc aattataaaa cactaccatc 600
gaggatgaaa atctataccc gtaaattttg gaaagtaata aaattttgat gaatccaagt 660
cagaatgagt tcatactcca gttgctaata aagacacaca caggtagtca cattttcaac 720
cacctttatt ttgatgcat aaaaa 745

<210> 3
<211> 652
<212> DNA
<213> Homo sapiens
<400> 3

ccctcctcaa acacacatcc atctccggc acacaccag tccatgcctc gccccacac 60
acacctgaca cccagtgca acaccacacc caccctcccc tccctgcaaa ctcccacctc 120
cgccacctg ctttaatacac attctcacc cccacacact ctttaataca tactctcaca 180
cccacaagcc tgccttcttg gagaagttag ccgagccgtg cagcgccgag aagggcaccc 240
ccgaagaccg ggaggaacgc cgcggggacc tgtggcttag cgcgtccgc cgggcttgt 300

ctgccgcggg gggcgcagcg gctgaggcgg ctccgggccg gagttccaat caagcgccac 360

ccaactccca gtcgggggcc gaggccagcg ccgggatgcc agcttcccc aaaaagatcc 420

tgcctcaggg aaatgcatgg agccggcgga aaagcccgcg gcgcccccg cgatcgag 480

accctaaggg ggcgggaggt ggcgccccag tccaacctc ttgagccaac ccagtgggtg 540

ggaagtgcc ttacctagg ccttcgcaa agtgttctc cttgtatta ttctaattac 600

ggtattttta atttccttaa aaaaataaga aacagaaaag cacagaaaa aa 652

<210> 4

<211> 705

<212> DNA

<213> Homo sapiens

<400> 4

gggaagctgt ttgtcagtg taacctggct aactttggac agctgcaaca acaacatcta 60

tcccttgcca tatggtgtgt ttagatgcta aacctgcct catctcttac ctaaagttaa 120

catctgatga catgatcgca actccatata ctgagttgat ttcaacctat gatgttgaa 180

ttttccacat ggagataaac actgggaagc agctaatac tctctctgga ttataagga 240

acaaatggat aagtgaact tgaaaacaat ggccaaagta ctaagaacta atgtttcttc 300

aacaaaaaa atacttttac tttaagaaca gaagataaaa ggccaaagtt gtaataaag 360

acaaattacg gaaatatata cttatacata taaaccact cctcttttta aaaaatcatg 420

ctttttaaga agaattctct catatcagtg agaaagaaag ttatcttaac aaacaaaaag 480

tattttgaaa aaaagcgttc acatgtataa gaaagaaaat gaacaaataa ataggaaaaa 540

agatctctgg ttcaaccata tcaaagcttc ttcattgttg ttattttact ctaaaggaca 600

cctccctgat cttgaactga taatacagtt aaaagaatta acagaaagag agagggatag 660

aatgggtcac gtacagaatg aaaaataaag gaaatttttt tctta 705

<210> 5

<211> 1771

<212> DNA

<213> Homo sapiens

<400> 5

agcaggaaag gcaggggagg cagcggcggg aagccttcct tagctgtgtg atcctagatg 60

gctgttaact tagcatttga ggactgtctc agatagacac aattggacca gattttgtac 120

catagaaaaa aaggctttac ccactcaaac atccactgtt ttcagcagac agagtcctaa 180

gtggatgcct gattggctga tcctgggaca cgtgcctgct cactggatgt actgaggcta	240
tgagaagaag gatctggcaa aacaaacctc agggacctca tcaagaaggc agaggacctg	300
gactaacat gacatccaaa ctgaacacaa ttttagagat aattcccaca gaagagacag	360
gatttctaga aaatggaaat agcccgtctt tggaaaagaa tgggcattta ctgagatgca	420
actataagct ggtgaaatat aattttgttt aaaatcattt aattttattc cacatcattt	480
tcctttgaag atatttgaat tgccaggacgac attaatattg tcaaggagat ttgctgtttc	540
acccaaatat gagatgtcca ctttgggatt agatccctgc ctcatcctaaa tttccttttc	600
ccttagtctt ttcttgagtt acggacattt gttcattttt ttcaacttat tcaacaagta	660
tttattgaac aactactatg tgccatatac tgtttttggt ttggtatgga aagcctgtag	720
gaatctctgt tgtaattttg ggaatttgcc caacctgctc tttagggtgt gctatgaact	780
gacttgtgtg ctcccaaaat ttatctgttg aagctctaag ctccaatgtg atgatatttg	840
gagatgggat ctttgggaat taattaaggt tagatgaggt catgaggatg aggcccttat	900
gatgggatta gtgcccttat aagaaaacac accagagtgc ctgcttgctc tctctctctg	960
ccatgtgaag acacagtatg aaggcagctt attgcaagcc taagagagaa cccttgccag	1020
aattggactt ccggtctcca gaaccatgaa aacataaact tctgttgttt aagccacca	1080
gtctacggaa ttttgttcta gaagcccaag ctgactaata tgggttgagt ccagtgtttt	1140
gcgtgaagaa ggtacatctc taaaactaaa catcttgctc attgaagctt gatctgaaat	1200
aattaagact ttggtcgtga aaatcttaga cagaatccat tcagtgacta cctttattct	1260
gaattggcag gcaggcagtt ctgatagatc tgagaagatt cttttgacaa caatttctcc	1320
aacttcagat atgttgaaaa acaaaataaa tatgttaata attaacatga tgtaatgaac	1380
ttcagtagta aattaagact agaagagatt aggggtacca ggaatcttgg ttgtaaagtg	1440
gcataataaa acatagtgc atcaaaactt gctaaagtgg ccaggtgcag tggctcatgc	1500
ctgtaatccc agcactttgg gaagccgaga caggtggatc acctgaggtc aggagtttga	1560
gattagcctg gccaacaggg tgaaacctg tctctactga taatacaaaa attagccaga	1620
catggtggca cagactgta atcccagcta cttaggagge tgaggcagga gaatcgcttg	1680
aacctgggag gcagaggttg cagtgagccg ggatcgacc attgcactcc agcctgggca	1740
acacgagcaa aactctgtct caaaacaaaa c	1771

<210> 6

<211> 4226

<212> DNA

<213> Homo sapiens

<400> 6

gtctccagca ctgcggacac cgggtgaca cattctttcg gagaggagg gcaagacttt	60
tctccggagc ctctggcaac tcgtggagtc tatataagat cttcaaagaa agagatatcc	120
tttcgaggca acacaagaag ccaactacaa atgagaatat tattctcttg gtgcagggat	180
tgttgcatgt ggaccccatc agccttgaaa agtccatgca atggagtga acttgacaaa	240
gcagttgcag tcgttcacca gtgaagcaaa gtcttcttcc tagtcacgca ggaatTTTT	300
tgggggagat aacttgatgc ctggagctgc tgctgcatct tatgtaaatg acaactaagg	360
ataaaaacca cacacattac ccaaggtgca gtaaaagaag atattttggg aggaagagag	420
aaagcaagag agagaaacca cattttcata acttttgta cagtatatg ttataattgg	480
tctattttta ttattattgt tgttcataat ttactttaat ttataaatta aacttcatta	540
taggtatgta caggaaaaaa tatggtatta tatatgattt ggtactatcc ttggtttcag	600
gcatccactg agggctcttg aacatatcac ctggatagg ggagactact gacttcctta	660
ctgctctatg tgacaaatga agtttggctt tatttgcagc caggaaaatt ttattatcaa	720
gactttaata ttaatacaga aataggaact tccagggata gacaaaaatg aagagagcat	780
ctgttcagc ttctttatTT taaagatttg aggcctcagag aggagatata ctcttcaaa	840
atcaacagcc tattagtcca caactgcgat tacaatttcg tctcctggct ctgggtttaa	900
taatattccc ccacatcata aaaccttgga actatcgaat tgatagcttt gaggggaaaa	960
ataggcaaag aagtaatatg ttaaaccttc cctgggatgc aggccttaat cagggttcaa	1020
attcttccaa cataatataa tttaatggtg agagctcttg acattaagag gtgaaattta	1080
tatttaaaag gctaagatcg tgacatatta aagcacttca aaaatgttgg caacaaatac	1140
caagagtaac tgatggttgt tattatacaa gaacaaagaa gttcatcctc tggaattaaa	1200
ggtgtgttag tgcattgtt gctcatgggg aggagtaggc acctgacttt ctttttttcc	1260
ttacgacagc tttatccaaa taaagcacgg tgtgtttcat caccataat ggaaagtgcc	1320
tgtcactcct tgaacaacat catatggttg gtggccggtg gataatggac cactgagtgt	1380
caagtgact ggctttcaat gcacacaaaa ctaaaaatta tagacctgtc attccacagg	1440
caggccacaa ttagctgtct atgcaacttg aaatttccat ttgccagct gagtaatcat	1500
ttgactagag tttttcattt ggaagcattt ctaagccttt tgtgtgcaag tcttttaaaa	1560
tgtggatcct gaagtaata aacagtgcac ttaatcagat gtattctttt tagaaacacc	1620
ccttcggaat cacacaggca aacccaaca ttgttgtaat agtaacaaca ataatgtgac	1680
aaaacaatat aagccttggg tattttatTT agggatctta aaagttcaca gcgaaaggta	1740
agaatcattt cttaatacct cagagtgcag aatttagaaa aagaaactga gatttttagca	1800

gtcagtaact agataaggct accaaaggag aagtaaacct agagctaggg gtcaaaactt	1860
gggaatttca gtgtcagaaa cattactga ttcaaaacta tccatttacc agttttctat	1920
ttacatttgg ttccaactc catatctgag ttctgtcatt tctctgattc tttaggggat	1980
agagaaaatg ataccaagtg gagtctgttg gaaaattgtg ttgtaatttc cattcagttg	2040
ctgtactgta agtaaacctc aggattctac aatcccagat cagattctag actgtctaata	2100
tgaacctgaa tcttctttta ccctcttctt gccctgtccc tggccacat tactctctct	2160
ccttctccct cggatgattt atatctacac ttaaattgcc gagccacaag gatacaagtg	2220
aagttgacaa tcaatatgaa ttaaggaaaa aatcagaagt ctgagatcag aagtgtttga	2280
gtgaaattct ccctacaagt ctgggcaacc tcatgatttt gtttaatttt ccaacatgtt	2340
ttggattatt ggaagagtt acatgggata aaaaggattt gcctggcaga gtaaaagaac	2400
tgaatttggg ttgtactgaa gtaggtccta actaggctgt tgatggtgaa atattgacga	2460
ttctattgtt ttgaatgaat gagtgtcata ttttactggg aattgatttc tgaacttcag	2520
tcaaggagtg actgaatttg caattagagt cactctgttt tgtaactcag attatgcaag	2580
gatcattctg aaacagcaac caagatataa gcccttgttg catggctttt ctttaatttc	2640
taattttgtg tgctctgttt cagtttttaa aggacaaaa agtttttttt ccccatctc	2700
tctagaaaaa aaaggaacat accctaggat tacaaggggc ttgattatct gttttataag	2760
aatcttaacc caggaattgg ttaatatctt ttctccatt tagaaaaata gtaggctttg	2820
tcgttgattc ctaatgatta gcacaaaaag ggaaggctt tttaaaacca agataaaatc	2880
actttaattc caaaagtgtc catttaagaa atataaactg tgaagataag caatggctat	2940
tttctctaaa ggaaagtatt ttcaaatgaa cagaacaggc aaatgtagac ttaattaggt	3000
ggagtcatg gggtattaat tattcacctt tccagggtga gaataatagc attgttgcta	3060
aagaaatgaa ataaggcatg acatttcttt cttcttcatt ctgtgaaagc tgctctgac	3120
tgaacagggt aaagatggga taacaggagg acagactgaa aatttgctct tcccttcaca	3180
attgagggtg gatagtcttt cgagagtgtg gtatcctcca ttgaataaaa atgttattga	3240
ctacctgctt ggagacataa aaaccttctc atgtttggaa tgtacatgta cttctctgtt	3300
ttgtctctt gaacattttc aaaatagcta aataaaattt gtcctcattc ttagagttct	3360
gtatattcta gacttgaata tagtggccaa cagtgagtag taattcacac tgatgataat	3420
cacaccagac taggtttgtt taattttgtt atagcctcta aaaagcattt gtttaagcatc	3480
tgctctgtta cacaggtaag atctcatgat ggatgtggag tctacaaatg aaaaatactt	3540
aaacgcactg cagtaaaaga tcaacactta ctctaataa atttgtagt accgagtgtc	3600
aaaaagtggg gtttgaataa acctaagggt aagatttagc aggtgtgggg tctagtcctg	3660

acactgatgg tgactacttg ttgcctacct caattctggt cataagagct atgagtgtga	3720
atgtgtatgt taagtagtgt gaggttgtgt gactggaagg ggctggagga aatagaggca	3780
gagaaactca gcaacctga gttgcagaga agctttttta gtgtccctga gaagtgaaat	3840
gactctgtta ttatgtctgc atgctaatta tacctctact ctacatactt ttaactttag	3900
gtaacctcat gttactgcc tttttaagct gaggaacca ataccatgg agataaaatg	3960
acttgaccca aatcggacag cttgtcaata gcaaagctgg aaccagactg taaagcttct	4020
gctttcttgt gtaagataca ttttccct agatgacttc ttaaaccaag gggctctgaa	4080
aataaacaat ctttcaaag tagcactgta tcctgtagg ttgttgcaa tagttcttca	4140
aaacttgaaa taattatata atcttaaaaa gtcccttata aattccaatt cagtataaaa	4200
atatgattta ggtatatgtt aattaa	4226
<210> 7	
<211> 675	
<212> DNA	
<213> Homo sapiens	
<400> 7	
atgaagaaac agagggatat actgccgtct taaaccaacc caaatgcctc tgccgcagtc	60
gcagttattt aggggtggaga agcgggctct ggcaactaga gaaggtggcc gcagcatggg	120
gacacccgaa ttgaggagct gtttgggcaa taagctgatt tagagaatta agacctccac	180
agatgggagt ggaatccgca actgctcaat ccatccactg agaagtecca taaaagatgg	240
acaatccata actgaaaccc aaccttcggc acaacgatca aattctacat gcataaagca	300
agagatctgt gtccagatc catcttaaag gcacccacc cagaggacag agtctgtctc	360
cttcccctaa acaatatgca gaaaccacc tggatattca tgggttctca cttaggaaaa	420
ttgagttctg ttcaatgatg ccccttctgc actggttggg ttcatgtgac aaagagagct	480
ctaggaggga cttttgctag gggaacgcag attgttccag caagatgaag gcctaggaga	540
cctggatggc ctgggtgttt tcttcccagg gagagcagag gaggcctgga ttgaaactcg	600
agtcttatga ctttgccatt gaggcaacat ccaaaataaa aatagcatgc taaccccaaa	660
aaaaaaaaaa aaaaa	675
<210> 8	
<211> 1113	
<212> DNA	
<213> Homo sapiens	

<400> 8

cagtcagggc cccgggagag tccgcgcgca tcgcccggct ccgcggcgcc agccatggcg	60
tccgtgcgtg gccccgccag ggatggggcg acgcggtcag agccgcacgc gaccgaaatc	120
cgcactctgg agccgcagag cgcgcggtct tgcgttttag cggctccctg gcaagtgcac	180
tggggaagaa acgcagggcg caggagagac agctggaaag cgggctagcc agggcccgtt	240
ttcttcagct tctcccttga cttgatccct gaagatggct ctgcctctgg cgcccttgaa	300
atatgtcgac taggtcatga tcaccgtgca ttctgctcag cgtagtctcc ccgcctgaag	360
tgtttattaat gcagtctaaa atgactacgg ttctgcggtc taccgaactg cccacggagg	420
taggggtttc agacttttgg catctcaaga gacttgagat taccgagacc acgatacccc	480
cagagacgtg acgagatggg atcaaggcca cgggggacca gaatgctgac ccaggcaacc	540
tgtccctgcc tcgctgccct ttgtaatact ccttatcata tcatgaggag ctattttccg	600
ttctccacaa acttggtggt ggatggaaga ggatatcaag ctacacctga gtgggattga	660
ggtaggaggc ggggctcaga caccagacca gattgaggac aaactaaagc aaggtcaggg	720
ccaaagcagc ttccaacag ataggaccac cagggtgcca cgtcagtta cgtttgccat	780
gtcaatacca gggagttacc gcccttttcc gtggcagtaa cccaataatt accaccctt	840
ccctagaaat ttctgcataa accgcccctt aatctgtatg caattaaaag tgtgtgtgtg	900
tgagtgtgtg tgtgtgtgtg tgtgtgtgta tgtgtgtgtg tgtgtgtata tatatatata	960
tatatatata tatatatata tatatatata tatatatata taaaactgca aaactgccct	1020
gagctgctcc tctctgccta tgggggagcc ctgctctgca ggagcagtca cggagctgta	1080
accctgcctc ttcaataaag ctgttttctt cta	1113

<210> 9

<211> 2558

<212> DNA

<213> Homo sapiens

<400> 9

aattagcgtg ggaggaagag acagtgaagc agccccitaa aaatggtctt tcgggtggaa	60
acactttatt tgggaagcaa agtgcaagat tcagcaatag aaaccactca gtgtccttgg	120
tgaatgattt gatcctggcc atcagaacct ttacttcagc cagcaataaa ctcttagaaa	180
cagaagcttc gaggtatgta ctctgttct gcacgaacaa gttttctggt tgaacatgct	240
gaacctctta aaaaattcag aacaagtttt ctcagagttt acaaagggat ttcaatcaca	300
tgaagagagg aaaacacact gagtcgtaac agtgacatca aaatactttt gaagttgact	360

gaaaatgtgg ttgcttttgt ttgagaatgg caaattaaag aattggatat tgggttctta	420
tctcagaaga aacaaaggag gagaccctga aaaactgaac atgaagagga agaagtggag	480
ttatgcaact cttcataaga aaatctccat ctgtttggag aagtactgcg tccttccgga	540
ttatcttcta cactatgcct tattgttatac tggctgggtc caccaccggt cagaagaatg	600
aaaattgcag atacctatc tgtgtgtttc tctccagaat ggtggatgtt agagacccta	660
ctaaggattt gtggactcac taacgtagtc tctaaaaggt tggacatccc tacccaaggc	720
tcaaatggta acagggtccat gccttacttg tgcagttct cttattgaaa gagcatctcc	780
taaccatcct accaataagg gtcttactct gttgccctgg ctggagtgcg gtggcccaat	840
catagctcac tgcagccttg acttcccagc ctcagttgat cctctacct taatctcccg	900
agtagctggg actgcaggtt tacttatact aagattttct tctgcctctg ccaaccctga	960
gacagcaaga ccaattcctc ctctcccctg tcctctcctc cgcctactcc acatgaagac	1020
aacgaggatg aagaccttca tgatgatcca ttccactta atgaacaaac ttcaggaaat	1080
tattcaagaa gagaggcatg gtggtatgct ggaaagtctg ctgagcttga aagtaggtaa	1140
cccaggctcc tgttctgtt ctcattgatg agttccctga cttcaggtga gtcacagctt	1200
ctctggactt tcatgtcctt ttgcaaagtg agtctgatca tttctaggat cttttccagc	1260
cctgagattg tatgagctgc aatagctaag agatatggga aagtgggtca ttggaaagtt	1320
gttttcttaa ctccacagcg aacaatgtgc ctcattctta gtcacaaag gcaaatacga	1380
agggatggga agcactctct tcaactgtcca ttgcaagca gaatcactct tcaactgatt	1440
gtccagtgtc tgtatagtct ctcacagagc ccagcacatg tgggcattct gtaaatggcc	1500
attgagaact gccaggctctg tggctgctac agagcacaaat gatgatgata atggtgttga	1560
tgatcatgat gatgatgggc tgggtatcgc aggcacagac ccaaggaagg ctaatgagaa	1620
tgtttctagg gaagcttctt gcatgtcact cttcttctt ctaacttctc taggttcac	1680
ttccctgact aacagggagc cagggtata ataacggga ggaagtcac atgtgattct	1740
acctagaaac gcaggcatca acttgaaatg ttttctctt ctcttgtaac tctggaatat	1800
ggctccttgg attggaggtg gcaaatacca cgggttaaac aatgtcattc tgagctcttt	1860
tagtactgaa taatgaata tataccatat tctacctct taaagaagga gtgtatcagc	1920
tatgtgctg atatggtttg gctctatgtc cccacccaaa tctcatcttg tagcacccat	1980
aattcccagg ttttgtggga gggacccggt gggagatgat tgaatcatgg gggctggtct	2040
ttcctgtgct ggtcttctga tagtaaatgg gtctcacaag atctgatggt ttttgaaaac	2100
aggagtttct ctgcatgagc tctctctgcc tctgcccac cacttaagat gtgacttgc	2160
cgtccatcgc ctctgccat gattgtgagg cctccccagc cacttggaac ttagctcagg	2220

acagcaacca gacctgagat tcttttcaac ttatacgctg ttgatgaatg tggtttctgg	2280
atgaaaaact gagaacagac tgacctctg gccacagcta ctgcgctcag ggacgcatag	2340
ttgcaggtgc atgctgtctg gtttggtttg ccaagtcaaa atagtgtgtt ttcccgtgt	2400
ggtgagagtt tccaacttca gacggggcaa gagaaattct aaacagtttg ccctttgcag	2460
atgtattgaa atagtactga aatttactga ttttcctca tgtttttttt ccgtgctttt	2520
atcactacc tgaaaataaa caaggagaaa aggcaagc	2558