



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2011-0036559
(43) 공개일자 2011년04월07일

- | | |
|--|---|
| <p>(51) Int. Cl.
C12Q 1/68 (2006.01) C12N 15/11 (2006.01)
G01N 33/574 (2006.01)</p> <p>(21) 출원번호 10-2011-0023976(분할)</p> <p>(22) 출원일자 2011년03월17일
심사청구일자 2011년03월17일</p> <p>(62) 원출원 특허 10-2007-0131453
원출원일자 2007년12월14일
심사청구일자 2008년11월11일</p> | <p>(71) 출원인
한국생명공학연구원
대전 유성구 어은동 52번지</p> <p>(72) 발명자
정경숙
대전 유성구 관평동 운암아파트 612-1902호
원미선
대전 서구 월평동 한아름아파트 110-1407
(뒷면에 계속)</p> <p>(74) 대리인
안소영</p> |
|--|---|

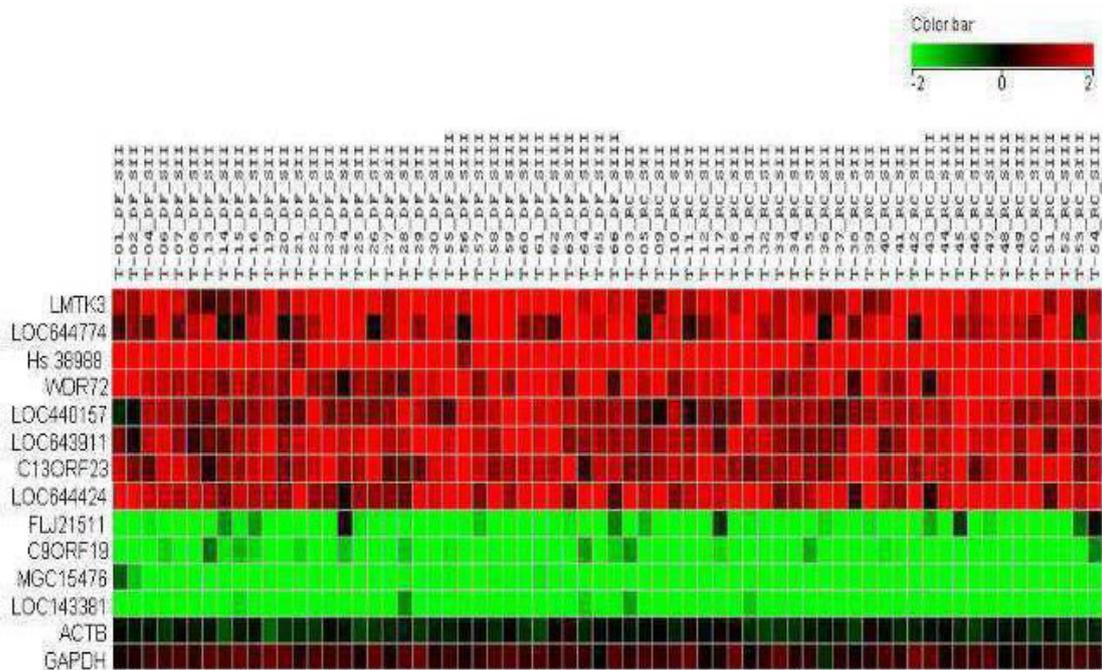
전체 청구항 수 : 총 6 항

(54) 대장암 진단용 조성물 및 그 용도

(57) 요약

본 발명은 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 및 LOC143381 유전자의 신규한 용도를 제공한다. 본 발명에 따른 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 또는 LOC143381 유전자는 대장암과 높은 관련성을 가지므로, 상기 유전자의 발현량을 확인함으로써 대장암의 진단, 약물스크리닝 등이 가능하고, 본 발명에 따른 유전자를 대장암 치료타겟으로 할 수 있다는 장점을 갖는다.

대표도



(72) 발명자

최정해

부산 사하구 당리동 325-5, 18통 6반

염영일

대전 동구 용전동 신동아아파트 8동 1302호

이희구

대전 동구 용전동 신동아아파트 14-405

송은영

서울 영등포구 여의도동 진주아파트 A-109

김선영

대전 유성구 도룡동 과기원교수아파트 2-205호

김영호

서울 강남구 대치3동 우성1차아파트 5-1303

전호경

서울 강남구 일원동 50번지

강창모

서울 노원구 공릉2동 삼익아파트 103동 1305호

특허청구의 범위

청구항 1

LOC644424(hypothetical protein LOC644424, GeneBank accession No.U79301 NT_113794) 유전자, 상기 유전자의 센스 및 안티센스 프라이머쌍, 상기 유전자로부터 발현된 단백질, 또는 상기 단백질에 대한 항체를 포함하는 대장암 진단용 조성물.

청구항 2

제1항에 있어서, 상기 센스 및 안티센스 프라이머쌍은 서열번호 29 및 30 인 대장암 진단용 조성물.

청구항 3

제1항에 있어서, 상기 조성물은 LOC644774(similar to phosphoglycerate kinase 1, GeneBank accession No.XM_927868.1), WDR72(WD repeat domain 72, GeneBank accession No.NM_182758.2), LOC440157(hypothetical gene supported by AK096951; BC066547, GeneBank accession No.NM_001013701.1) 및 LOC643911(hypothetical protein LOC643911, GeneBank accession No.XM_931911.2)로 이루어진 군에서 선택된 하나 이상의 유전자, 상기 유전자의 센스 및 안티센스 프라이머쌍, 상기 유전자로부터 발현된 단백질, 또는 상기 단백질에 대한 항체를 추가로 포함하는 대장암 진단용 조성물.

청구항 4

제3항에 있어서, 상기 유전자의 센스 및 안티센스 프라이머쌍중 LOC644774 유전자의 센스 및 안티센스 프라이머쌍은 서열번호 17 및 18인 센스 및 안티센스 프라이머쌍, WDR72 유전자의 센스 및 안티센스 프라이머쌍은 서열번호 21 및 22인 센스 및 안티센스 프라이머쌍, LOC440157 유전자의 센스 및 안티센스 프라이머쌍은 서열번호 23 및 24인 센스 및 안티센스 프라이머쌍, LOC643911 유전자의 센스 및 안티센스 프라이머쌍은 서열번호 25 및 26인 센스 및 안티센스 프라이머쌍인 대장암 진단용 조성물.

청구항 5

LOC644424(hypothetical protein LOC644424, GeneBank accession No.U79301 NT_113794) 유전자, 상기 유전자의 센스 및 안티센스 프라이머쌍, 상기 유전자로부터 발현된 단백질, 또는 상기 단백질에 대한 항체를 포함하는 대장암 진단용 키트.

청구항 6

제5항에 있어서, 상기 키트는 LOC644774(similar to phosphoglycerate kinase 1, GeneBank accession No.XM_927868.1), WDR72(WD repeat domain 72, GeneBank accession No.NM_182758.2), LOC440157(hypothetical gene supported by AK096951; BC066547, GeneBank accession No.NM_001013701.1) 및 LOC643911(hypothetical protein LOC643911, GeneBank accession No.XM_931911.2)로 이루어진 군에서 선택된 하나 이상의 유전자, 상기 유전자의 센스 및 안티센스 프라이머쌍, 상기 유전자로부터 발현된 단백질, 또는 상기 단백질에 대한 항체를 추가로 포함하는 대장암 진단용 키트.

명세서

기술분야

[0001] 본 발명은 대장암 진단용 조성물 및 그를 이용한 진단키트, 대장암 치료용 조성물, 및 대장암 치료제 스크리닝용 조성물 및 스크리닝 방법에 관한 것이다.

배경기술

[0002] 대장암은 일반적으로 소득수준이 높은 집단에서 발생률이 높아 일반적으로 '선진국형 암'으로 인식되어지고 있으며 서양의 경우 암 사망에 있어서 제2위를 차지하는 암으로 전체 암의 약 15% 정도를 차지하고 있다. 미국의 경우 연간 약 56,000명이 대장암으로 사망하고 130,000명의 새로운 대장암 환자가 발생한다고 보고되고 있으며

우리나라에서 대장암은 사망률에서 전체 암에서 차지하는 비율이 위암, 간암, 폐암에 이어 제4위를 차지하는 암이다. 1980년에 전체암의 5.8%를 차지하던 것이 1985년에는 6.1%, 1995년에는 8.2%, 가장 최근 자료인 2002년에는 11.2%로 지속적인 증가 추세를 보이고 있으며, 1991년도 자료와 비교해 볼 때 위암, 간암, 자궁경부암의 사망률이 감소한 데 비해 대장암에 의한 사망률은 인구 10만 명당 4.7명에서 11.4명으로 2배 이상 증가하였다. 이와 같은 추세가 계속된다면 2010년경에는 우리나라에서의 대장암 발생빈도가 서양의 수준에 도달할 것으로 예측되며 2006년 대장암 발병률은 2위에 해당한다.

[0003] 대장암 환자의 생존률을 높이기 위한 많은 노력이 진행되고 있으나 기존 치료법을 통하여 생존률을 향상시키기는 어려운 상태이다. 대장암 환자의 생존률을 높이기 위해서는 무엇보다도 조기진단과 새로운 대장암 맞춤형의 발굴이 시급하다. 대부분의 대장암은 선종에서 발생하며, 암으로의 변화에 약 10년이 소요되므로 조기진단을 통하여 전암성 병변인선종이나 조기암을 제거할 기회가 많다. 현재 사용되는 선별 검사로는 1) 대변 잠혈 검사, 2) 에스결장경 검사, 3) 대장 내시경 검사, 4) 대장 조영술이 사용되고 있는데 잠혈 검사는 조기진단용으로는 유용성이 떨어지며, 대장 조영술은 진단의 정확도가 떨어져 1cm 이하의 경우 20-50%, 1cm 이상일 때는 10-30%, 조기 대장암의 경우 15-45% 정도의 정확도를 보인다. 따라서 좀 더 정확하고 조기진단이 될 수 있는 진단시스템이 필요하며, 이는 대장암 관련 마커유전자의 발굴이 필요로 되는 부분이다. 또한 대장암이 발병했을 경우 예후를 예측하고 적절한 치료방침을 정하는데 대장암 마커가 적절히 활용될 수 있을 것으로 사료된다.

[0004] 대장암의 발생과정에는 여러종류의 암 유전자, 종양 억제유전자등 다양한 유전자 변화가 관여하는 것으로 알려져 있다. 실제 대장암은 발암과정에서 일어나는 유전적 변화가 가장 많이 밝혀진 암이다. 대장암은 한 개의 암 유전자 또는 종양억제 유전자의 변화가 단독으로 암을 유발시킬 수 있는 것이 아니고 정상 대장 점막세포가 선종의 단계를 거쳐 대장암으로 진행되기 위해서는 수년에 걸친 긴 세월을 통해 여러 개의 암 관련 유전자의 변화가 축적되어야 하는데 이를 대장암 발생에 있어서 유전자의 다단계적 변화라고 한다. 여기서 중요한 것은 각 단계에서의 유전자 변화 순서가 아니라 궁극적으로 누적되는 유전자들의 변화의 총합이다. 대장암 발생과정에 관여하는 유전적 변화로는 K-ras, APC, MCC 유전자, 18번 염색체의 DCC유전자, 17번 염색체의 p53 유전자, 그리고 DNA methylation의 이상 등이 있으며, hMSH2, hMSH1, hPMS1, hPMS2 등의 돌연변이도 관계된다.

[0005] 이와 같이, 암의 형성은 다양한 유전자들과 이들 유전자들의 발현 및 조절 기작이 복합적으로 연관되어 진행되므로 최근 들어 다량의 유전자를 사용하는 올리고 칩을 이용한 암 관련 유전자들의 발현률을 비교하여 암의 새로운 진단이나 치료의 마커를 발굴하기 위한 연구들이 이루어지고 있다. 암세포에서 발현이 증가하거나 감소하는 유전자들은 세포분열, 세포신호전달, 세포 골격, 세포 운동, 세포 방어, 유전자 및 단백질의 발현 그리고 세포 내 물질대사등 여러부분에 관여하는 것으로 환자 조직들에 따라 동일한 발현변화를 보이는 유전자가 있는 반면 다른 발현 변화를 보이는 유전자들이 많다. 이것은 각각 환자들 특이성 때문일 가능성이 크므로 연구하는 대상의 환자 조직들의 정확한 병리학적 소견과 분류에 따라야 하며 정확한 유전자를 이용한 진단에는 보다 많은 새로운 유전자들의 검색과 확인이 필요하다.

[0006] 특정 세포내에서 특정 유전자의 발현 빈도를 조사함으로써 대장암 관련된 유전자의 발굴이 가능하며, 이를 통하여 대장암 진행의 분자적 메커니즘을 이해하게 되고, 나아가 대장암 진단 및 치료 타겟으로의 사용이 가능하게 될 것이다.

[0007] 현재 LMTK3는 기능적으로는 Tyrosine kinase 도메인을 가지고 있다는 것이 알려져 있을 뿐, 다른 기능에 대해서 전혀 보고된 바가 없으며, LOC644774는 Phosphoglycerate kinase 1 의 유사 단백질을 코딩하고 있으나 기능 연구는 되어 있지 않다. HS.389988는 염기서열만 알려져 있고 WDR72는 WD repeat 도메인을 가진 미지의 단백질이다. LOC440157, LOC643911, C13ORF23, 및 LOC644424 또한 전혀 기능이 알려져 있지 않다. FLJ21511는 기능이 알려져 있지 않고 C9ORF19 는 Glioma pathogenesis와 관련되어 보고되었으나 암관련 보고는 없으며, MGC15476 는 DACT3 로도 알려져 있고 DACT3 단백질은 마우스의 뇌 분화시 발현되는 단백질로 알려져 있으며, HS.388347 는 유전자 서열만이 알려져 있다.

[0008] 본 발명자들은 상기 유전자들이 대장암관련성이 있음을 발견하여 본 발명을 완성하였다.

발명의 내용

해결하려는 과제

[0009] 본 발명은 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 및 LOC143381 유전자의 신규한 용도를 제공하는 것을 목적으로 한다.

[0010] 보다 구체적으로, 본 발명은 상기 유전자로부터 발현된 단백질을 포함하는 대장암 진단용 또는 대장암치료제 스크리닝용 조성물, 상기 유전자의 억제제 또는 상기 유전자로부터 발현된 단백질의 억제제 및 억제학적으로 허용되는 담체를 포함하는 대장암 치료용 조성물, 및 대장암 진단용 키트를 제공하고자 한다.

과제의 해결 수단

[0011] 본 발명은 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 및 LOC143381로 이루어진 군에서 선택된 하나 이상의 유전자를 포함하는 대장암 진단용 조성물을 제공한다.

[0012] 본 발명의 'LMTK3', 'LOC644774', 'HS.389988', 'WDR72', 'LOC440157', 'LOC643911', 'C13ORF23', 'LOC644424', 'FLJ21511', 'C9ORF19', 'MGC15476', 또는 'LOC143381'의 '유전자'라 함은 이들 유전자의 DNA 또는 mRNA를 말하며, DNA 또는 mRNA의 전부 또는 일부를 모두 포함하는 개념이다.

[0013] 본 발명의 대장암 진단용 조성물은 상기 유전자 외에도 핵산의 구조를 안정하게 유지시키는 증류수 또는 완충액을 포함할 수 있다.

[0014] 상기 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23 및 LOC644424 유전자는 대장암세포에서 특이적으로 발현이 증가하며, FLJ21511, C9ORF19, MGC15476 및 LOC143381 유전자는 대장암세포에서 특이적으로 발현이 감소한다. 따라서 상기 유전자의 발현정도를 조사하면 대장암을 진단할 수 있다.

[0015] 상기 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 및 LOC143381 유전자의 서열정보는 표 1에 나타낸 바와 같다.

표 1

서열번호	유전자이름	길이	CDS (단백질부위)	염색체위치	Genebank Number
고발현 유전자					
1	LMTK3	1,4972	1,4470	19q13.32	NM_001080434
2	LOC644774	1,282	1,282	X chromosome	XM_927888.1
3	HS.389988	1,488	undefine	Ch.10	CB133377 gi:28099007
4	WDR72	1,7326	43,3351	15q21.3	NM_182758.2
5	LOC440157	1,2779	362,994	14q11.1	NM_001013701.1
6	LOC643911	1,558	80,334	16q12.2	XM_931911.2
7	C13ORF23	1,5228	874,3708	13q13.3	NM_025138.3/ NM_170719.2
8	LOC644424	1,1610	336-743	1q12	U79301 NT_113794
저발현 유전자					
9	FLJ21511	1,2486	183,2282	4p12	NM_025087.2
10	C9ORF19	1,1900	35,499	9p13.3	NM_022343.2
11	MGC15476	1,2544	441,1655	19q13.32	NM_145056.1
12	LOC143381	1,4471	undefine	10q26.11	BX648964.1

[0016]

[0017] 본 발명은 또한 상기 유전자들의 대장암 진단 용도 및 상기 조성물과 대상체로부터 얻은 시료 간의 반응을 확인하여 대장암을 진단하는 방법을 제공한다.

[0018] 본 발명의 진단 방법에서 상기 유전자를 포함하는 조성물과 시료 간의 반응의 확인은 DNA-DNA, DNA-RNA, DNA-단백질 간의 반응 여부를 확인하는데 사용되는 통상적인 방법들, 예컨대 DNA 칩, 단백질 칩, 중합효소 연쇄반응(PCR), 노던 블롯팅, 서던 블롯팅, ELISA(Enzyme Linked Immunosorbent assay), 효모 이중 혼성법(yeast two-hybrid), 2-D 겔 분석 및 시험관 내 결합 에세이 (in vitro binding assay) 등을 이용할 수 있다. 예컨대 상기 유전자의 전부 또는 일부를 프로브로 사용하여 대상자의 체액으로부터 분리한 핵산과 하이브리드화한 후 당분야에 공지된 다양한 방법, 예컨대 역전사 중합효소 연쇄반응(reverse transcription polymerases chain reaction), 써던블로팅(southern blotting), 노던 블롯팅(Northern blotting) 등으로 이를 검출함으로써 대상자에서 상기 유전자가 고발현된 상태인지 또는 저발현된 상태인지 조사하면 대장암의 발생 여부를 판단할 수 있다. 상기 프로브를 방사선 동위원소 또는 효소 등으로 표지하면 용이하게 유전자의 존재를 확인할 수 있다. 상기 프로브의 염기서열은 상기 유전자의 염기서열과 70% 이상의 유사성이 있으면 족하다. 상기 프로브는 본 발명의 센스 및 안티센스 프라이머를 이용한 유전자 증폭법에 의해 제조할 수 있다.

[0019] 또한, 본 발명은 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 및 LOC143381로 이루어진 군에서 선택된 하나 이상의 유전자의 센스 및 안티센스 프라이머쌍을 포함하는 대장암 진단용 조성물을 제공한다.

[0020] 상기 센스 및 안티센스 프라이머쌍은 상기 유전자에 대한 상보적인 염기서열을 갖는 서열번호 15-38의 프라이머 쌍일 수 있다. 상기 프라이머쌍은 다음의 표 2에 정리하였다. 본 발명에서 '상보적'이란 프라이머의 염기서열에서 완전히 상보적인 것과 80% 이상의 상동성이 있는 것을 포함하는 개념이다.

표 2

서열번호	고발현 유전자	방향	프라이머 염기서열
15	LMTK3	sense	5'-ggaacagcagcagatcaaa-3'
16	LMTK3	antisense	5'-aggagaccatcttgcgcttc-3'
17	LOC644774	sense	5'-tcaatctgccacagaaggct-3'
18	LOC644774	antisense	5'-ccaggatggccaggaagggt-3'
19	Hs.389988	sense	5'-tggcaaatgaaaaattcagca-3'
20	Hs.389988	antisense	5'-gcctgggacctccagaataa-3'
21	WDR72	sense	5'-tccatgtacctgaggctga-3'
22	WDR72	antisense	5'-ttggaagttgcccaggttg-3'
23	LOC440157	sense	5'-gagaagaagcaggctccagg-3'
24	LOC440157	antisense	5'-agttccagggctctgtggtt-3'
25	LOC643911	sense	5'-aatcatccaaggctggtc-3'
26	LOC643911	antisense	5'-ccccaaacatattaaaccactcg-3'
27	C13ORF23	sense	5'-agcaccagctctcagcagcac-3'
28	C13ORF23	antisense	5'-gaactgaatgctgaggcga-3'
29	LOC644424	sense	5'-atcagtcctcactgcggtc-3'
30	LOC644424	antisense	5'-ttcgcctcttgatgagaag-3'
서열번호	저발현유전자	방향	프라이머 염기서열
31	FLJ21511	sense	5'-gtggctaggggaaaagttgg-3'
32	FLJ21511	antisense	5'-aatctctggagccagggtc-3'
33	C9ORF19	sense	5'-accgggaggctcaacagat-3'
34	C9ORF19	antisense	5'-tatctggccaccacaaggga-3'
35	MGC15476	sense	5'-cactgcggagatcgacgct-3'
36	MGC15476	antisense	5'-ttttcacgaagactttggcg-3'
37	LOC143381	sense	5'-ggcggctctgattttagg-3'
38	LOC143381	antisense	5'-tccaagggtgtggcaataa-3'

[0021]

- [0022] 상기 센스 프라이머 및 안티센스 프라이머를 이용한 공지의 방법으로 상기 12종의 유전자의 발현양을 측정할 수 있다. 예를 들면, RT-PCR방법 등으로 측정할 수 있다.
- [0023] 본 발명은 또한 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 및 LOC143381로 이루어진 군에서 선택된 하나 이상의 유전자로부터 발현된 단백질을 포함하는 대장암 진단용 조성물을 제공한다.
- [0024] 본 발명의 조성물은 상기 단백질 외에도 단백질의 구조를 안정하게 유지시키는 증류수 또는 완충액을 포함할 수 있다.
- [0025] 앞서 언급한 바와 같이, 상기 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23 또는 LOC644424 유전자는 대장암세포에서 특이적으로 발현이 증가하며, FLJ21511, C9ORF19, MGC15476 또는 LOC143381 유전자는 대장암세포에서 특이적으로 발현이 감소하므로 상기 유전자들로부터 발현된 단백질을 이용하여 상기 유전자 또는 단백질의 과발현 또는 저발현 여부를 조사하면 대장암을 진단할 수 있다.
- [0026] 본 발명은 또한 상기 단백질의 대장암 진단 용도 및 상기 조성물과 대상체로부터 얻은 시료 간의 반응을 확인하여 대장암을 진단하는 방법을 제공한다.
- [0027] 본 발명의 진단 방법에서 상기 단백질을 포함하는 조성물과 시료 간의 반응의 확인은 DNA-단백질, RNA-단백질, 단백질-단백질 간의 반응 여부를 확인하는데 사용되는 통상적인 방법들, 예컨대 DNA 칩, 단백질 칩, 중합효소 연쇄반응 (PCR), 노던 블롯팅, 서던 블롯팅, 웨스턴 블롯팅, ELISA(Enzyme Linked Immunosorbent assay), 특이적 면역염색(histoimmunostaining), 효모 이중 혼성법(yeast two-hybrid), 2-D 겔 분석 및 시험관 내 결합 에세이 (in vitro binding assay) 등을 이용할 수 있다. 예컨대 상기 유전자들로부터 발현된 단백질의 전부 또는 일부를 프로브로 사용하여 대상자의 체액으로부터 분리한 핵산 또는 단백질과 하이브리드화한 후 당분야에 공지된 다양한 방법, 예컨대 역전사 중합효소 연쇄반응(reverse transcription polymerases chain reaction), 웨스턴 블로팅(western blotting) 등으로 이를 검출함으로써 대상자에서 상기 유전자가 과발현된 상태인지 조사하면 대장암의 발생 여부를 판단할 수 있다. 상기 프로브를 방사선 동위원소 또는 효소 등으로 표지하면 용이하게 유전자의 존재를 확인할 수 있다.
- [0028] 또한 본 발명의 조성물은 상기 단백질 대신 상기 단백질에 대한 특이적 항체를 포함할 수 있다. LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23 또는 LOC644424 유전자는 대장암세포에서 특이적으로 발현이 증가하여 상기 유전자로부터 발현된 단백질의 양 또한 증가하게 되며, FLJ21511, C9ORF19, MGC15476, 또는 LOC143381 유전자는 대장암세포에서 특이적으로 발현이 감소하여 상기 유전자로부터 발현된 단백질의 양 또한 감소하게 된다. 따라서 상기 유전자로부터 발현된 단백질에 대한 항체를 이용하는 경우, 항원-항체 반응을 통해 상기 단백질을 검출해 내어 대장암을 진단할 수 있다.
- [0029] 상기 단백질에 대한 모노클로날 항체는 당업계에 통상적인 모노클로날 항체 제작 방법을 통해 제작되어 사용될 수도 있고, 시판되는 것을 사용할 수 있다. 상기 단백질에 대한 모노클로날 항체는 일반적으로 알칼라인 포스파타아제(alkaline phosphatase, AP) 또는 호올스래디쉬 퍼록시다제(horseradish peroxidase, HRP) 등의 효소가 컨쥬게이션된 2차 항체 및 이들의 기질을 사용하여 발색반응시킴으로써 정량분석할 수도 있고, 아니면 직접 상기 단백질에 대한 모노클로날 항체에 AP 또는 HRP 효소 등이 컨쥬게이션된 것을 사용하여 정량분석할 수도 있다. 또한, 모노클로날 항체 대신에 상기 단백질을 인식하는 폴리클로날 항체를 사용할 수도 있고 이는 당업계에 통상적인 항혈청 제작 방법을 통해 제작되어 사용될 수도 있으며, 항원결합성을 갖는 것이면 모노클로날 항체 또는 폴리클로날 항체의 일부도 본 발명의 항체에 포함되고, 모든 면역 글로불린 항체가 포함된다. 나아가, 본 발명의 항체에는 인간화 항체 등의 특수항체도 포함된다.
- [0030] 상기 항체는 본 발명의 12종의 각 유전자를 통상적인 방법에 따라 발현벡터에 클로닝하여 상기 유전자에 의해 코딩된 단백질을 얻고, 얻어진 단백질로부터 통상적인 방법에 의해 제조될 수 있다. 여기에는 12종의 단백질의 부분 펩타이드도 포함하며, 본 발명의 부분펩타이드로는 최소한 7개 이상, 바람직하게는 12개 이상의 아미노산을 포함한다.
- [0031] 또한 본 발명은 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 및 LOC143381로 이루어진 군에서 선택된 하나 이상의 유전자, 상기 유전자의 센스 및 안티센스 프라이머쌍, 상기 유전자로부터 발현된 단백질, 또는 상기 단백질에 대한 항체를 포함하는 대장암 진단용 키트를 제공한다.

- [0032] 상기 대장암 진단용 키트는 지지체, 적당한 완충용액, 발색 효소 또는 형광물질로 표지된 2차 항체, 발색 기질액, 또는 1,6-N-아세틸글루코사민 당쇄가지 변화를 측정하기 위한 L4-PHA, 폴리(A) RNA 분리시약 등을 더 포함할 수 있다.
- [0033] 상기 지지체는 니트로셀룰로오즈막, 폴리비닐수지로 함성된 96웰플레이트(96 well plate), 폴리스티렌수지로 함성된 96웰플레이트, 또는 유리로 된 슬라이드글라스 등일 수 있고, 상기 발색효소는 퍼옥시다아제(oxidase), 또는 알칼라인 포스파타아제(alkaline phosphatase) 등일 수 있으며, 상기 형광물질은 FITC, 또는 RITC 등일 수 있고, 상기 발색 기질액은 ABTS(2,2'-Azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)), OPD(o-Phenylenediamine), 또는 TMB(Tetramethyl Benzidine) 등일 수 있다.
- [0034] 본 발명은 또한 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 및 LOC143381로 이루어진 군에서 선택된 하나 이상의 유전자를 포함하는 대장암 치료제 스크리닝용 조성물을 제공한다. 또한 본 발명은 상기 조성물의 대장암치료제 스크리닝용 용도 및 상기 조성물을 표적물질로 이용하여 시험대상물질을 접촉시키고, 표적물질과 시험대상물질 간의 반응을 확인하여, 시험대상물질이 상기 유전자의 발현을 증진시키는 활성 또는 억제하는 활성을 나타내는지 결정하는 단계를 포함하는 대장암치료제 스크리닝 방법을 제공한다.
- [0035] 본 발명의 스크리닝 방법에서 상기 유전자를 포함하는 조성물과 시험대상물질 간의 반응 확인은, DNA-DNA, DNA-RNA, DNA-단백질, DNA-화합물 간의 반응 여부를 확인하는데 사용되는 통상적인 방법들을 사용할 수 있다. 예를 들면, 생체 외부에서(in vitro) 상기 유전자와 시험대상물질 사이의 결합 여부를 확인하기 위한 혼성화 시험, 포유류세포와 시험대상물질을 반응시킨 후 노던 분석, 정량적 PCR, 정량적 실시간 PCR 등을 통한 상기 유전자의 발현을 측정 방법, 또는 상기 유전자에 리포터 유전자를 연결시켜 세포 내로 도입한 후 시험대상물질과 반응시키고 리포터 단백질의 발현을 측정하는 방법 등을 사용할 수 있다. 이러한 경우 본 발명의 조성물은 상기 유전자 외에도, 핵산의 구조를 안정하게 유지시키는 증류수 또는 완충액을 포함할 수 있다.
- [0036] 또한, 본 발명은 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 및 LOC143381로 이루어진 군에서 선택된 하나 이상의 유전자로부터 발현된 단백질을 포함하는 대장암치료제 스크리닝용 조성물을 제공한다. 본 발명은 또한 상기 조성물의 대장암치료제 스크리닝용 용도 및 상기 조성물을 표적물질로 이용하여 시험대상물질을 접촉시키고, 표적물질과 시험대상물질 간의 반응을 확인하여, 시험대상물질이 상기 단백질의 기능을 증진시키는 활성 또는 억제하는 활성을 나타내는지 결정하는 단계를 포함하는 대장암치료제 스크리닝 방법을 제공한다.
- [0037] 본 발명의 스크리닝 방법에서 상기 단백질을 포함하는 조성물과 시험대상물질 간의 반응 확인은, 단백질-단백질, 단백질-화합물 간의 반응 여부를 확인하는데 사용되는 통상적인 방법들을 사용할 수 있다. 예를 들면, 상기 단백질과 시험대상물질을 반응시킨 후 활성을 측정하는 방법, 효모 이중 혼성법(yeast two-hybrid), 상기 단백질에 결합하는 파지 디스플레이 펩티드 클론(phage-displayed peptide clone)의 검색, 천연물 및 화학물질 라이브러리(chemical library) 등을 이용한 HTS(high throughput screening), 드럭 히트 HTS(drug hit HTS), 세포 기반 스크리닝(cell-based screening), 또는 DNA 어레이(DNA array)를 이용하는 스크리닝법 등을 사용할 수 있다. 이러한 경우 본 발명의 조성물은 상기 단백질 외에도, 단백질의 구조 또는 생리 활성을 안정하게 유지시키는 완충액 또는 반응액을 포함할 수 있다. 또한 본 발명의 조성물은 생체 내(in vivo) 실험을 위해, 상기 단백질을 발현하는 세포, 또는 전사율을 조절할 수 있는 프로모터 하에 상기 단백질을 발현하는 플라스미드를 함유하는 세포 등을 포함할 수 있다.
- [0038] 본 발명의 스크리닝 방법에서, 시험대상물질은 통상적인 선정방식에 따라 대장암전이 억제제로서의 가능성을 지닌 것으로 추정되거나 또는 무작위적으로 선정된 개별적인 핵산, 단백질, 기타 추출물 또는 천연물, 화합물 등이 될 수 있다.
- [0039] 본 발명의 스크리닝 방법을 통해 얻은, 대장암 고발현 유전자의 발현을 증진시키거나 단백질의 기능을 증진시키는 활성을 나타내는 시험대상물질 및 반대로 대장암 고발현 유전자의 발현을 억제시키거나 단백질의 기능을 억제시키는 활성을 나타내는 시험대상물질은, 전자의 경우, 시험대상물질에 대한 억제제를 개발함으로써 대장암치료제 후보물질이 될 수 있고, 후자의 경우는 대장암치료제 후보물질이 될 수 있다. 또한, 대장암 저발현 유전자의 발현을 증진시키거나 단백질의 기능을 증진시키는 활성을 나타내는 시험대상물질 및 반대로 대장암 저발현 유전자의 발현을 억제시키거나 단백질의 기능을 억제시키는 활성을 나타내는 시험대상 물질은, 전자의 경우, 대장암치료제 후보물질이 될 수 있고, 후자의 경우, 시험대상물질에 대한 억제제를 개발함으로써 대장암치료제 후보물질이 될 수 있다. 이와 같은 대장암치료제 후보물질은 이후의 대장암치료제 개발과정에서 선도물질(leading

compound)로서 작용하게 되며, 선도물질이 상기 유전자 또는 그로부터 발현되는 단백질의 기능 억제효과를 나타낼 수 있도록 그 구조를 변형시키고 최적화함으로써, 새로운 대장암치료제를 개발할 수 있다.

- [0040] 본 발명은 또한 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 및 LOC143381로 이루어진 군에서 선택된 하나 이상의 유전자의 siRNA를 제공한다.
- [0041] 상기 siRNA의 염기서열은 상기 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 및 LOC143381로 이루어진 군에서 선택된 하나 이상의 유전자(mRNA)의 염기서열 중 각각의 서열에서 연속된 19~23개의 염기서열일 수 있다.
- [0042] 상기 siRNA의 서열은 편의상 정방향 서열(sense strand)을 나타낸 것으로, 실제 유전자 억제효과를 나타내는 역방향 서열(antisense strand)과 함께 이중리보핵산쇄를 구성하게 된다.
- [0043] 본 발명의 siRNA는 짧은 19-23개의 이중 리보핵산쇄로 세포내에 도입하면 비특이적 저해(non-specific inhibition)없이 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 및 LOC143381로 이루어진 군에서 선택된 하나 이상의 유전자 발현만을 억제하는 효과를 나타내므로, 대장암 관련 유전자 기능연구에 이용할 수 있다. 또한, 상기 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, 및 LOC644424로 이루어진 군에서 선택된 하나 이상의 유전자 발현을 억제하여 대장암세포를 죽이는 효과도 나타낼 수 있다.
- [0044] 본 발명은 또한 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, 및 LOC644424로 이루어진 군에서 선택된 하나 이상의 유전자에 대한 억제제를 포함하는 대장암 치료용 조성물을 제공한다.
- [0045] 상기 조성물은 약제학적으로 허용되는 담체를 추가로 포함할 수 있다.
- [0046] 본 발명의 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, 및 LOC644424로 이루어진 군에서 선택된 하나 이상의 유전자는 대장암 세포에서 다량 발현되므로, 상기 유전자의 억제제를 투여하여 상기 유전자의 발현을 저해시키면 대장암을 억제할 수 있다.
- [0047] 따라서 본 발명은 또한 상기 유전자의 억제제의 대장암 치료 용도 및 유효량의 상기 유전자의 억제제를 환자에게 투여하는 단계를 포함하는 대장암 치료방법을 제공한다. 본 발명에 있어서 대장암 치료는 대장암의 예방 및 억제를 포함한다.
- [0048] 본 발명에 있어서, LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, 및 LOC644424로 이루어진 군에서 선택된 하나 이상의 유전자에 대한 억제제는 상기 유전자의 mRNA에 대한 안티센스 올리고뉴클레오타이드일 수 있다.
- [0049] 안티센스 올리고뉴클레오타이드는 생체 내 뿐만 아니라 생체 외에서도 유전자-특이적 억제를 달성하기 위해 성공적으로 사용되어 왔다. 안티센스 뉴클레오타이드는 특정 DNA 또는 RNA 타겟에 대해 안티센스(또는 상보)인 짧은 길이의 DNA 합성 가닥(또는 DNA 아날로그)이다. 안티센스 올리고뉴클레오타이드는 타겟에 결합하고 전사, 번역 또는 스플라이싱의 단계에서 발현을 멈추게 함으로써 DNA 또는 RNA 타겟에 의해 인코딩된 단백질의 발현을 막기 위해 제안되었다. 안티센스 올리고뉴클레오타이드는 세포 배양 및 질병의 동물 모델에서도 성공적으로 이용되어 왔다(Hogrefe, 1999). 올리고뉴클레오타이드가 분해되지 않도록 더욱 안정하고 저항적이 되게 하기 위한 안티센스 올리고뉴클레오타이드의 또 다른 변형이 당업자에게 알려져 있고 이해된다. 여기서 사용된 안티센스 올리고뉴클레오타이드는 이중나선 또는 단일나선 DNA, 이중나선 또는 단일나선 RNA, DNA/RNA 하이브리드, DNA 및 RNA 아날로그 및 염기, 당 또는 백본 변형을 지닌 올리고뉴클레오타이드를 포함한다. 올리고뉴클레오타이드는 안정성을 증가시키고, 뉴클레아제 분해에 대한 저항성을 증가시키기 위해 당분야에 알려진 방법에 의해 변형된다. 이들 변형은 당분야에 알려져 있는 올리고뉴클레오타이드 백본의 변형, 당 모이어티의 변형 또는 염기의 변형을 포함하나 이에 한정적인 것은 아니다.
- [0050] 또한, LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, 및 LOC644424로 이루어진 군에서 선택된 하나 이상의 유전자에 대한 억제제는 상기 유전자의 siRNA(Small Interfering RNA)일 수 있다.
- [0051] 즉, 상기 siRNA의 염기서열은 상기 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, 및 LOC644424로 이루어진 군에서 선택된 하나 이상의 유전자(mRNA)의 염기서열 중 각각의 서열에서 연속된 19-23개의 연속된 염기서열일 수 있다.

- [0052] 상기 siRNA의 서열은 편의상 정방향 서열(sense strand)을 나타낸 것으로, 실제 유전자 억제효과를 나타내는 역방향 서열(antisense strand)과 함께 이중리보핵산쇄를 구성하게 된다.
- [0053] siRNA는 세포 배양 및 생체 내에서 오래 지속되는 효과, 생체 내에서 세포를 트랜스펙션시키는 능력 및 혈청 내 분해에 대한 저항력의 측면에서 생체 내에서 특정한 유전자의 발현의 저해에 대해 매우 강한 약물이 되는 잠재력을 지닌다(Bertrand et al., 2002). siRNA의 전달 및 siRNA를 포함한 발현 컨스트럭트/백터는 당업자에게 알려져 있다. 예를 들어, 미국 출원 제2004/106567 및 2004/0086884는 바이러스성 백터, 비바이러스성 백터, 리포솜 전달 운반체, 플라스미드 주입 시스템, 인공 바이러스 엔벨로프 및 폴리라이신 컨쥬게이트를 포함한 전달 메커니즘뿐만 아니라 많은 발현 컨스트럭트/백터를 제공하고 있다.
- [0054] siRNA는 상대적으로 낮은 농도로도 안티센스 올리고뉴클레오타이드에 의해 얻을 수 있는 효과와 동등하거나 높은 효과를 얻을 수 있기 때문에 안티센스 올리고뉴클레오타이드의 대안으로 제시되고 있다(Thompson, 2002). siRNA의 이용은 질병의 동물 모델에 있어서 유전자 발현을 저해하는 데 대중성을 나타내고 있다. 당업자는 당해 기술 분야에 공지된 방법을 이용하여 원하는 방식대로 상기 안티센스 올리고뉴클레오타이드 및 siRNA를 합성하고 변형시킬 수 있다(예를 들어, Andreas Henschel, Frank Buchholz and Bianca Habermann (2004) DEQR: a web-based tool for the design and quality control of siRNAs. Nucleic Acids Research 32(Web Server Issue):W113-W120. 참조). 또한 당업자는 안티센스 올리고뉴클레오타이드 또는 siRNA를 지닌 발현 컨스트럭트/백터에 유용한 조절 서열(예컨대, 구성적 프로모터, 유도성 프로모터, 조직-특이적 프로모터 또는 그의 결합)을 잘 이해하고 있다.
- [0055] 대장암 치료를 위해 사용되는 본 발명의 안티센스 올리고뉴클레오타이드 또는 siRNA는 약제학적으로 허용되는 담체를 추가적으로 포함한 조성물의 형태로 투여될 수 있다. 적당한 약제학적으로 허용되는 담체는 예를 들어 하나 이상의 물, 식염수, 인산 완충 식염수, 텍스트린, 글리세롤, 에탄올뿐만 아니라 이들의 조합을 포함한다. 이러한 조성물은 투여 후 활성 성분의 빠른 방출, 또는 지속적이거나 지연된 방출을 제공하도록 제제화될 수 있다.
- [0056] 상기 유전자의 저해제는 안티센스 올리고뉴클레오타이드 또는 siRNA 외에도 상기 유전자의 발현을 억제하는 물질이면 어떤 것이든 가능하다. 따라서 종래 당해 기술 분야에서 상기 유전자의 저해제로 알려진 화합물 또한 이용가능하다.
- [0057] 본 발명은 또한 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, 및 LOC644424로 이루어진 군에서 선택된 하나 이상의 유전자로부터 발현된 단백질에 대한 억제제를 포함하는 대장암 치료용 조성물을 제공한다.
- [0058] 본 발명의 유전자를 억제하면 그에 따른 단백질의 발현이 억제되어 대장암이 억제되는 것이므로, 상기 유전자의 단백질을 저해하면 대장암을 억제할 수 있다.
- [0059] 따라서 본 발명은 상기 단백질에 대한 억제제의 대장암 치료 용도 및 유효량의 상기 단백질의 억제제를 환자에게 투여하는 단계를 포함하는 대장암 치료방법을 제공한다. 본 발명에 있어서 대장암 치료는 대장암의 예방 및 억제를 포함한다.
- [0060] 상기 단백질에 대한 억제제는 본 발명의 유전자로부터 발현된 단백질에 대한 항체일 수 있다. 상기 단백질에 대한 모노클로날 항체는 당업계에 통상적인 모노클로날 항체 제작 방법을 통해 제작되어 사용될 수도 있고, 시판되는 것을 사용할 수 있다. 또한, 모노클로날 항체 대신에 상기 단백질을 인식하는 폴리클로날 항체를 사용할 수도 있고 이는 당업계에 통상적인 항혈청 제작 방법을 통해 제작되어 사용될 수도 있다.
- [0061] 본 발명의 대장암 치료용 조성물은 투여를 위해서 상기 기재한 유효 성분 이외에 추가로 약제학적으로 허용되는 담체를 1종 이상 포함하여 약제학적 조성물로 바람직하게 제제화할 수 있다. 본 발명의 단백질에 대한 억제제가 항체일 경우 약제학적으로 허용되는 담체는 결합 단백질의 저장 수명 또는 유효성을 증가시키는 습윤제, 유화제, 방부제 또는 완충액과 같은 최소량의 보조 물질로 구성될 수 있다.
- [0062] 또한 본 발명의 대장암 치료용 조성물은 하나 또는 그 이상의 대장암치료제와 함께 사용될 수 있다. 본 발명의 대장암 치료용 조성물은 예를 들어 당업자에게 잘 알려진 화학요법약제(chemotherapeutic agent), 예컨대, 사이클로포스파마이드, 아지리딘, 알킬알코실포네이트, 나이트로소우레아, 다카르바진, 카르보플라틴, 시스플라틴 등과 같은 알킬화제(alkylating agent), 마이토마이신 C, 안트라사이클린, 독소루비신(아드리아마이신) 등과 같은 항생제, 메토트렉세이트, 5-플루오로우라실, 시타라빈 등과 같은 항대사제(antimetabolitic agent), 빈카 알

칼로이드와 같은 식물유래 약제 및 호르몬 등을 추가로 포함할 수 있다.

- [0063] 본 발명의 대장암 치료용 조성물은 상기 유효 성분 외에도 약제학적으로 적합하고 생리학적으로 허용되는 보조제를 포함할 수 있으며, 이러한 보조제로는 부형제, 붕해제, 감미제, 결합제, 피복제, 팽창제, 윤활제, 활택제, 또는 가용화제 등이 있다.
- [0064] 또한 본 발명의 조성물은 투여를 위해서 상기 기재한 유효 성분 이외에 추가로 약제학적으로 허용되는 담체를 1종 이상 포함하여 약제학적 조성물로 바람직하게 제제화할 수 있다.
- [0065] 액상 용액으로 제제화되는 조성물에 있어서 허용되는 약제학적 담체로는, 멸균 및 생체에 적합한 것으로서, 식염수, 멸균수, 링거액, 완충 식염수, 알부민 주사용액, 텍스트로즈 용액, 말토 텍스트린 용액, 글리세롤, 에탄올 및 이들 성분 중 1 성분 이상을 혼합하여 사용할 수 있으며, 필요에 따라 항산화제, 완충액, 정균제 등 다른 통상의 첨가제를 첨가할 수 있다. 또한 희석제, 분산제, 계면활성제, 결합제 및 윤활제를 부가적으로 첨가하여 수용액, 현탁액, 유탁액 등과 같은 주사용 제형, 환약, 캡슐, 과립 또는 정제로 제제화할 수 있다. 더 나아가 해당분야의 적절한 방법으로 Remington's Pharmaceutical Science, Mack Publishing Company, Easton PA에 개시되어 있는 방법을 이용하여 각 질환에 따라 또는 성분에 따라 바람직하게 제제화할 수 있다.
- [0066] 본 발명의 대장암 치료용 조성물의 약제 제제 형태는 과립제, 산제, 피복정, 정제, 캡슐제, 좌제, 시럽, 즙, 현탁제, 유제, 점적제 또는 주사 가능한 액제 및 활성 화합물의 서방출형 제제 등이 될 수 있다.
- [0067] 본 발명의 대장암 치료용 조성물은 정맥내, 동맥내, 복강내, 근육내, 동맥내, 복강내, 흉골내, 경피, 비측내, 흡입, 국소, 직장, 경구, 안구내 또는 피내 경로를 통해 통상적인 방식으로 투여할 수 있다.
- [0068] 본 발명의 치료 방법에 있어서, "유효량"은 대장암을 억제하는 효과를 이루는데 요구되는 양을 의미한다. 따라서, 본 발명의 유효 성분의 "유효량"은 질환의 종류, 질환의 중증도, 조성물에 함유된 유효 성분 및 다른 성분의 종류 및 함량, 제형의 종류 및 환자의 연령, 체중, 일반 건강 상태, 성별 및 식이, 투여 시간, 투여 경로 및 조성물의 분비율, 치료 기간, 동시 사용되는 약물을 비롯한 다양한 인자에 따라 조절될 수 있다. 성인의 경우, 상기 유전자 또는 단백질의 억제제를 1일 1회 내지 수회 투여시, siRNA일 경우 0.01ng/kg~10mg/kg, 상기 유전자의 mRNA에 대한 안티센스 올리고뉴클레오타이드인 경우 0.01ng/kg~10mg/kg, 화합물일 경우 0.1ng/kg~10mg/kg, 상기 단백질에 대한 모노클로날 항체일 경우 0.1ng/kg~10mg/kg의 용량으로 투여하는 것이 바람직하다.
- [0069] 본 발명에 있어서, 상기 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 및 LOC143381 유전자는 표 1의 염기서열 또는 상기 염기서열 중 하나 이상의 염기가 결실, 치환 또는 삽입된 염기서열일 수 있다.
- [0070] 상기 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 및 LOC143381 유전자, 상기 유전자의 센스 및 안티센스 프라이머쌍, 및 상기 유전자의 siRNA 서열은 예시일 뿐 이에 한정되는 것이 아님은 당업자에게 자명하다. 상기 서열들에 대해 실질적인 서열 동일성 또는 실질적인 서열 상동성을 지닌 서열 또한 본 발명의 범주에 포함된다. 여기서 사용된 "실질적인 서열 동일성" 또는 "실질적인 서열 상동성"이라는 용어는 서열이 또 다른 서열과의 실질적인 구조적 또는 기능적 동일성을 나타냄을 표현하기 위해 사용된다. 이러한 차이는 예를 들어 다른 종 간의 코돈 용법의 고유의 변이에 기인한다. 2 이상의 다른 서열 사이의 유의적인 양의 서열 중복 또는 유사성이 있는 경우 이들 서열의 길이 또는 구조가 다르더라도 유사한 물리적 특성을 지니는 경우 구조적 차이는 무시할만한 정도가 된다.
- [0071] 본 발명에서 유전공학적 기술과 관련된 사항은 샘브룩 등의 문헌(Sambrook, et al. Molecular Cloning, A Laboratory Manual, Cold Spring Harbor laboratory Press, Cold Spring Harbor, N. Y. (2001)) 및 프레드릭 등의 문헌(Frederick M. Ausubel et al., Current protocols in molecular biology volume 1, 2, 3, John Wiley & Sons, Inc. (1994))에 개시되어 있는 내용에 의해 보다 명확하게 된다.

발명의 효과

- [0072] LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476 또는 LOC143381 유전자는 대장암과 높은 관련성을 가지므로, 상기 유전자의 발현량을 확인함으로써 대장암의 진단, 약물스크리닝 등이 가능하고, 본 발명의 유전자를 대장암 치료타겟으로 할 수 있다.

도면의 간단한 설명

- [0073] 도 1은 대장암 유전자의 마이크로어레이 결과를 나타낸 사진이다.
- 도 2는 임상조직에서 대장암 고발현 유전자의 발현량을 나타낸 RT-PCR결과 사진(A)과 LOC440157 유전자와 LOC643911 유전자의 발현비를 나타낸 그래프(B)이다.
- 도 3은 임상조직에서 대장암 저발현 유전자의 발현량을 나타낸 RT-PCR결과 사진(A)과 MGC15476 유전자의 발현비를 나타낸 그래프(B)이다.
- 도 4는 대장암세포주에서 대장암 고발현 유전자의 발현비(대장암세포주/정상조직)를 나타낸 그래프이다.
- 도 5는 대장암세포주에서 대장암 저발현 유전자의 발현비(대장암세포주/정상조직)를 나타낸 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0074] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 것이며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다.

[0075] <실시예 1> 대장암진단용 유전자 확인

[0076] 1-1. 대장암 임상조직과 정상조직, 6개의 대장암 세포주에서 총 RNA 분리

[0077] (1) 대장암 임상조직의 준비

[0078] 66쌍의 환자의 대장암 조직과 정상 조직은 삼성의료원으로부터 공급 받았다. 임상조직은 환자로 부터 외과적으로 제거된 후, 분석때 까지 이를 액체질소에 보관하였다.

[0079] (2) 대장암 세포주의 준비

[0080] 10% 소혈청(Fetal Bovine Serum, GIBCO사), 페니실린 (10000U/ml)과 스트렙토마이신(10mg/ml) 을 첨가한 RPMI1640 (GIBCO) 배양액을 10mm 디쉬에 10ml씩 분주한 후 대장암 세포주인 HT29, SW480, DLD1, HCT116, SW620, Colo205(각각의 입수처 기재바랍니다.) 를 디쉬당 세포수가 1×10^6 이 되도록 접종하고 이를 37℃, 5% CO₂ 가 존재하는 배양기에서 배양하였다.

[0081] (3) 총 RNA 분리

[0082] 총 RNA는 QIAGEN 키트 (RNeasy Maxi kit: cat #75162)를 사용하여 분리하였고, Experion RNA StdSens(Bio-Rad사) 칩을 이용하여 정량하였다. 우선 상기 대장암 임상조직과 정상조직을 적절한 크기로 자른 후 150ul 의 베타 멀갑토 에탄올을 첨가한 15ml의 키트 내 분해 완충액에 용해시켰다. 여기에 15ml의 70% 에탄올을 넣어 잘 섞은 후, 3000g에서 5분간 원심분리하여 총 RNA를 막에 부착시켰다. 두 차례의 세척을 한 후 1.2ml의 RNase가 없는 물을 첨가하여 총 RNA를 분리하였다. 부착된 세포주의 경우는 트립신, EDTA를 이용하여 회수한 후 키트 내 분해 완충액 RLN (50mM TrisCl, pH 8.0, 140mM NaCl, 1.5mM MgCl2, 0.5% NP-40) 1ml에 베타 멀갑토 에탄올 10ul 를 첨가하여 용해시켰다. 여기에 1ml의 70% 에탄올을 넣어 잘 섞은 후, 3000g에서 5분간 원심분리하여 총 RNA를 막에 부착시켰다. 두 차례의 세척을 한 후 100ul의 RNase가 없는 물을 첨가하여 총 RNA를 분리하였다.

[0083] 1-2. 총 RNA를 이용한 마이크로어레이 실시 및 대장암 진단용 유전자 확인

[0084] (1) 마이크로어레이 실시

[0085] 하이브리드화를 위해 실시예 1-1에서 추출된 총 RNA를 Illumina TotalPrep RNA Amplification Kit (Ambion사) 을 이용하였다. T7 Oligo(dT) primer를 이용하여 cDNA를 합성하고, biotin-UTP를 이용하여 *in vitro* transcription을 실시하여 biotin-labeled cRNA를 제조하였다. 제조된 cRNA는 NanoDrop(Nanodrop사, ND-1000)을 이용하여 정량하였다. 정상 대장 상피세포 및 대장암 세포에서 제조된 cRNA를 Human-6 V2 (Illumina사) 칩에 하이브리드화 하였다. 하이브리드화 후 비특이적 하이브리드화를 제거하기 위하여 Illumina Gene

Expression System 세척액 (Illumina사)을 이용하여 Illumina Human-6 V2칩을 세척하였고, 세척된 Illumina Human-6 V2칩은 streptavidin-Cy3(Amersham사) 형광 염색약으로 표지하였다. 형광 표지된 DNA 칩은 공초점 (confocal) 레이저 스캐너 (Illumina사)를 이용하여 스캐닝하여 각 스팟에 존재하는 형광의 데이터를 얻어서 TIFF 형태의 이미지 파일로 저장하였다. TIFF 이미지 파일을 BeadStudio version 3(Illumina사)으로 정량하여 각 스팟의 형광값을 정량하였다. 정량된 결과는 Avadis Prophetic version 3.3(Strand Genomics사) 프로그램으로 'quantile' 기능을 이용하여 보정하였다.

[0086] (2) 대장암 진단용 유전자 확인

[0087] 상기의 과정으로 얻어진 1,601개의 유전자 발현 정도 분석을 통하여 정상 대장 상피세포와 대장암 세포의 유전자 발현 양상을 비교 분석하여, 대장암 진단용 유전자를 확인하였다. 군집화 분석(hierarchical clustering analysis)를 이용하여 유전자 발현양상을 분석하였으며, 그 결과, 정상 대장 상피세포와 대장암 세포는 크게 두 개의 군집으로 나누어지는 것을 알 수 있었다 (정상과 대장암조직). 또한, 정상 대장 상피세포와 대장암 세포를 비교하여 60% 이상의 환자에서 2배 이상의 발현차이를 나타내는 유전자들을 발견할 수 있었으며, 이중 고발현과 저발현이 각각 281개 및 605개였고, 현재 대장암에서 유전자 발현량의 변화가 보고되지 않은 대장암 마커 유전자 후보를 선별할 수 있었다.

[0088] 정상조직을 대조군으로 한 대장암 조직에서의 각 유전자의 발현량을 고발현되는 유전자의 경우 붉은색으로, 저발현되는 유전자의 경우 초록색으로 나타내었다 (도 1). 도 1에 나타난 바와 같이, 고발현 유전자로 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, 및 LOC644424을 확인하였고, 저발현 유전자로 FLJ21511, C9ORF19, MGC15476 및 LOC143381을 확인하였다.

[0089] 또한, 상기 각각의 유전자에 대해 마이크로어레이 상의 구체적인 변화율과 유의성(P값), 2배이상 변화한 시료의 수, 그리고 유전자의 정의 및 알려진 기능을 표 3에 나타내었다. 표 3에 나타난 바와 같이 고발현 유전자와 저발현 유전자는 정상 대장 상피세포와 비교하여 발현차이를 크게 나타내므로, 대장암 진단, 약물 스크리닝, 또는 치료타겟 등으로 사용할 수 있음을 알 수 있다.

표 3

프로브 ID	유전자	변화율 (log2)	p 값 (유의성)	2배이상 증가한 수	2배이상 감소한 수	유전자 정의 및 기능
290491	LMTK3	2.1444444	1.06E-07	55	0	PREDICTED: lemur tyrosine kinase 3 (LMTK3), mRNA.
6660017	LOC644774	1.2223737	1.38E-05	42	1	PREDICTED: similar to Phosphoglycerate kinase 1 (LOC644774), mRNA.
3520491	Hs.38988	3.28803030	8.46E-18	63	0	K-EST0184252 L9SNU354 cDNA clone L9SNU354-11-G07 5, mRNA sequence
4490215	WDR72	2.0005051	3.32E-06	36	0	WD repeat domain 72 (WDR72), mRNA.
4220129	LOC440157	1.6494949	1.24E-09	35	0	hypothetical gene supported by AK096951; BC066547 (LOC440157), mRNA.
1400402	LOC643911	1.3365657	4.47E-14	40	0	PREDICTED: hypothetical protein LOC643911 (LOC643911), mRNA.
3710477	C13orf23	1.6339899	4.76E-05	55	0	chromosome 13 open reading frame 23 (C13orf23), transcript variant 1, mRNA.
6510338	LOC644424	1.44040404	1.386E-05	48	0	PREDICTED: hypothetical protein LOC644424 (LOC644424), mRNA.
5080044	FLJ21511	-3.026212	3.005E-06	0	62	hypothetical protein FLJ21511 (FLJ21511), mRNA.
4280253	C9orf19	-2.798232	9.95E-06	0	66	chromosome 9 open reading frame 19 (C9orf19), mRNA.
4890148	MGC15476	-3.004293	0.0004726	0	66	thymus expressed gene 3-like (MGC15476), mRNA.
7550072	LOC143381	-2.957323	3.14E-06	0	66	mRNA: cDNA DKFZp686J0156 (from clone DKFZp686J0156)

[0090]

[0091]

<실시에 2> 임상시료에서 대장암진단용 유전자 발현 확인

[0092]

17쌍의 임상환자 샘플(실시에 1의 66쌍의 대장암 환자로 부터 채취한 대장암조직(T1~T66)과 정상조직(N1~N66) 중 임상2기에 해당하는 환자조직 17쌍(T1~T17, N1~N17))을 이용하였으며, 실시에 1에서 확인된 8종의 대장암 고발현 유전자와 5종의 대장암 저발현 유전자의 발현량을 RT-PCR 방법을 통하여 분석하였다.

[0093]

2-1. 역전사 효소 반응에 의한 cDNA 합성

[0094]

17쌍의 임상환자 샘플(실시에 1의 66쌍의 대장암 환자로 부터 채취한 대장암조직(T1~T66)과 정상조직(N1~N66) 중 임상2기에 해당하는 환자조직 17쌍(T1~T17, N1~N17))의 RNA 를 이용하여 cDNA를 합성하였다. 시료 각각의 총 RNA 5ug, 프라이머인 50uM Oligo(dT)20 1ul와 10mM dNTP 2.5ul, 를 넣고 RNase 저해제인 DEPC 가 들어 있는 멸균수로 전체가 25ul 가 되도록 하여 RNA/primer 혼합용액을 만들었다. 65℃에서 5분간 반응시킨 후 55℃로 옮겨 보관하였다. 다음 10X RT buffer 5ul, 25mM MgCl₂ 10ul, 0.1M DTT 5ul, RNase inhibitor 1ul, SuperScriptIII RT 효소를 1ul 넣고 전체가 25ul 가 되도록 한 후 55℃에서 보관 중인 RNA/primer 혼합용액과 섞어준 후, 55℃에서 50분간 반응시켰다. 그 후 85℃에서 5분간 반응시켜 RT 효소를 불활성화 한후 얼음에 넣어 반응을 종결시켰다. PCR 을 하기 전에 cDNA sample에 RNase 1ul 를 처리하여 37℃에서 20분간 반응 시켜 RNA 를 제거한 후 PCR 반응을 하였다.

[0095]

2-2. PCR 을 통한 cDNA 증폭과 발현량 확인

[0096]

(1) 주형의 농도 보정

[0097] 마커유전자를 정량하기 위한 표준 유전자로서 GAPDH를 사용하였다. 표준 유전자의 프라이머를 이용하여 PCR 반응을 수행하고 표준 유전자 GAPDH의 발현량이 동일해지도록 cDNA의 농도를 보정하였다.

[0098] 우선 각각의 cDNA를 20배 희석한 후 희석된 샘플 2ul를 이용하여 PCR 반응을 수행하였다. PCR은 2X PCR premix (바이오니아사) 15ul, 2ul의 GAPDH 5' 프라이머(20pmole), 2ul 의 3' 프라이머(20pmole), 11ul 의 증류수를 넣어 사용하였고 20 cycle, 23cycle, 25cycle을 수행하였다. 이 때 PCR 반응 조건은 94℃ 30초, 50℃ 30초, 72℃ 1분으로 수행하였으며, 산물의 크기는 457bp 이다.

[0099] PCR 산물을 2% 아가로스 젤에 로딩하여 전기영동한 후 젤 사진을 찍고, 이미지를 TotalLab v1.0 프로그램 [Nonlinear Dynamix사] 으로 정량 한 후, 다시 보정하여 PCR을 수행하여 정량하는 방식으로 각 시료의 농도를 동일하게 보정하였다.

[0100] (2) PCR 에 의한 대장암 마커 유전자 증폭

[0101] 상기 (1)에서 보정된 시료를 20배 희석한 cDNA 를 각 유전자의 센스 및 안티센스 프라이머를 이용하여 PCR 하였다. PCR반응은 94도 1분, 54도 30초, 72도 1분으로 하였으며 cycle 수는 각 유전자의 샘플내의 농도에 따라 보정하면서 실시하였다. 적게는 25cycle을 수행하였으며 많게는 38cycle을 수행하였다. PCR 반응용액 조성은 표 4 와 같고, 사용된 프라이머는 표 5와 같다.

표 4

cDNA	3ul
2X Premix (바이오니아)	10ul
정방향 프라이머(10pmole/ul)	2ul
역방향 프라이머(10pmole/ul)	2ul
증류수	2ul
총부피	20ul

[0102]

표 5

서열번호	표준 유전자	방향	프라이머 염기서열
13	GAPDH	sense	5'-tcatgaccacagtcctatgcc-3'
14	GAPDH	antisense	5'-tccaccaccctgttgctgta-3'
서열번호	고발현 유전자	방향	프라이머 염기서열
15	LMTK3	sense	5'-ggaacagcagcagatcaaa-3'
16	LMTK3	antisense	5'-aggagaccatcttgcgcttc-3'
17	LOC644774	sense	5'-tcaatctgccacagaaggct-3'
18	LOC644774	antisense	5'-ccaggatggccaggaagggt-3'
19	Hs.38988	sense	5'-gcctgggacctccagaataa-3'
20	Hs.38988	antisense	5'-tggcaaatgaaaattcagca-3'
21	WDR72	sense	5'-tccatgtacctgaggctga-3'
22	WDR72	antisense	5'-ttggaagtgtccgagtttg-3'
23	LOC440157	sense	5'-gagaagaagcagggtcccagg-3'
24	LOC440157	antisense	5'-agttccaggctctgtggtt-3'
25	LOC643911	sense	5'-aattcatccaaggctgtgc-3'
26	LOC643911	antisense	5'-cccaaacatattaaaccactcg-3'
27	C13ORF23	sense	5'-agcaccagctctcagcagcac-3'
28	C13ORF23	antisense	5'-gaactgaatgctgaggcga-3'
29	LOC644424	sense	5'-attcagctcactcggct-3'
30	LOC644424	antisense	5'-ttccgcctcttgatgagaag-3'
서열번호	저발현 유전자	방향	프라이머 염기서열
31	FLJ21511	sense	5'-gtggctaggggaaaagtgg-3'
32	FLJ21511	antisense	5'-aatctctggagccagggtct-3'
33	C9ORF19	sense	5'-accgggaggctcaacagtat-3'
34	C9ORF19	antisense	5'-tatctggccaccacaaggga-3'
35	MGC15476	sense	5'-cactcggagatcgacgct-3'
36	MGC15476	antisense	5'-ttttcacgaagactttggcg-3'
37	LOC143381	sense	5'-ggcggctctgattttagg-3'
38	LOC143381	antisense	5'-tccaaggggttggcaataa-3'

[0103]

[0104]

대장암 관련 8종의 고발현 유전자인 LMTK3, LOC644774, HS.38988, WDR72, LOC440157, LOC643911, C13ORF23, 및 LOC644424 와 저발현 유전자인 FLJ21511, C9ORF19, MGC15476 및 LOC143381의 프라이머는 단백질 코딩 내부에서 디자인 하였으며 각 프라이머의 길이는 19~24mer 이고 Tm 값은 55도 근처이다.

[0105]

(3) 발현량 확인

[0106]

PCR 산물을 확인하기 위하여 2% 아가로스 젤을 이용하여 전기영동하고 이미지 장비를 이용하여 분석하였다.

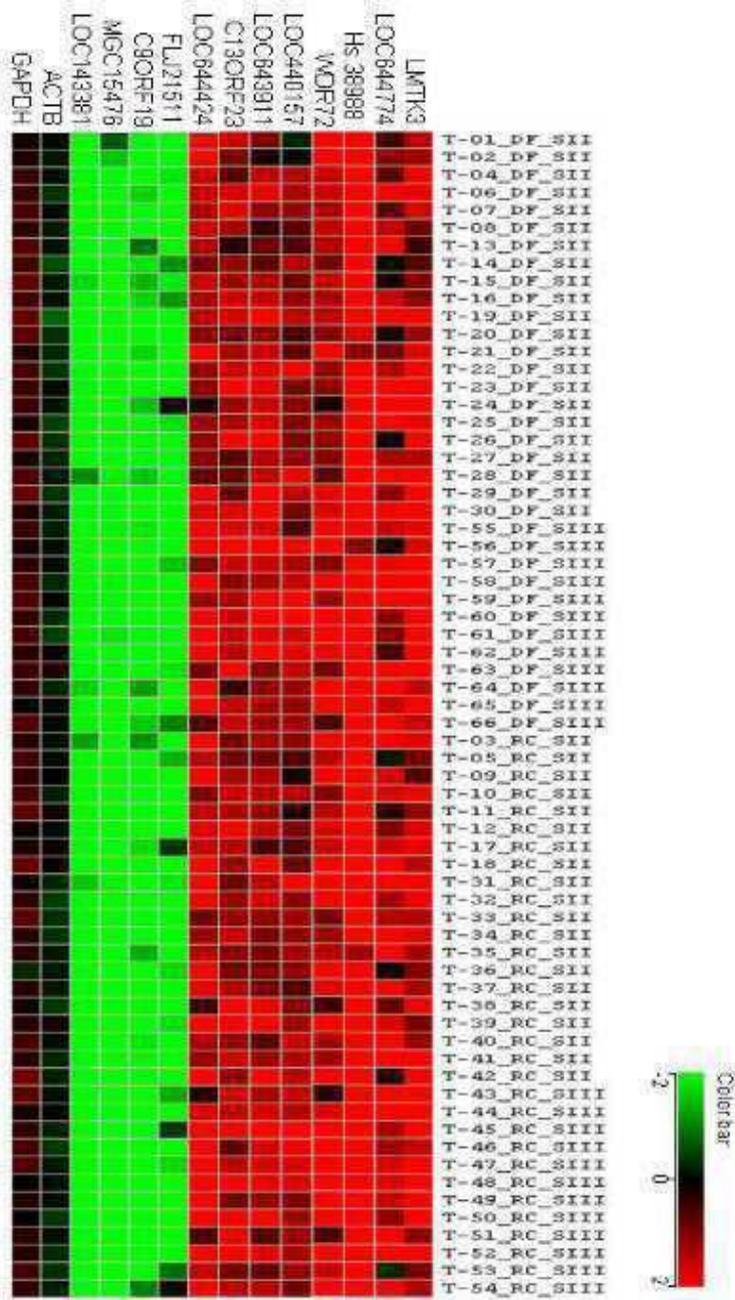
[0107]

그 결과를 도 2(A, B) 및 도 3(A, B)에 나타내었으며, 도 2는 임상조직에서 대장암 고발현 유전자의 발현량을 나타낸 RT-PCR 결과사진(A)과 LOC440157과 LOC643911의 발현비를 나타낸 그래프(B)이다. 도 3은 임상조직에서 대장암 저발현 유전자의 발현량을 나타낸 RT-PCR 결과사진(A)과 MGC15476의 발현비를 나타낸 그래프(B)이다. 사진에서 N은 정상조직(nontumer 조직)을 의미하며, T는 그에 해당하는 대장암 조직을 의미한다. 확인한 차를 보이는 유전자에 대해서는 아가로스젤 이미지 사진으로 나타내었고, 발현량의 차이가 적은 LOC440157과 LOC643911, 저발현유전자 MGC15476의 경우는 이미지 사진을 TotalLab v1.0 프로그램[Nonlinear Dynamix사]으로 정량 한 후 GAPDH로 보정하고 그래프로 다시 나타내었다. 상기 그래프의 X축은 임상시료를 나타낸 것이며 Y축은 발현비(대장암조직/정상조직)를 나타낸다. 상기 발현비는 대장암 조직에서 발현되는 양(각각의 마커유전자의 발현량을 표준유전자인 GAPDH의 발현량으로 보정하여 준 값)을 각각의 셋트의 정상 대장 조직에서 발현되는 발현량으로 나누어 준 값이다.

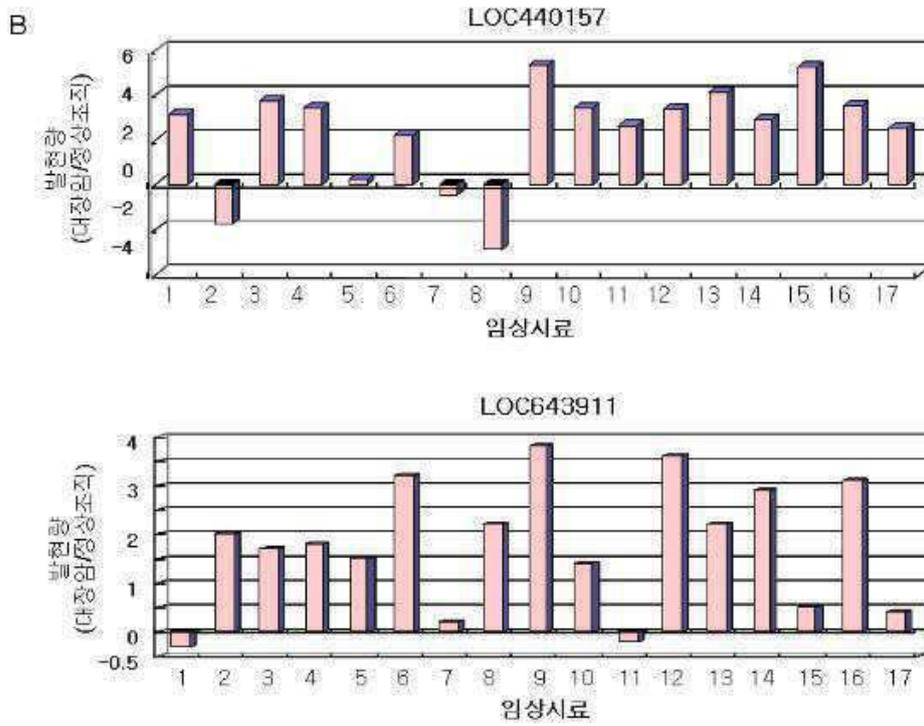
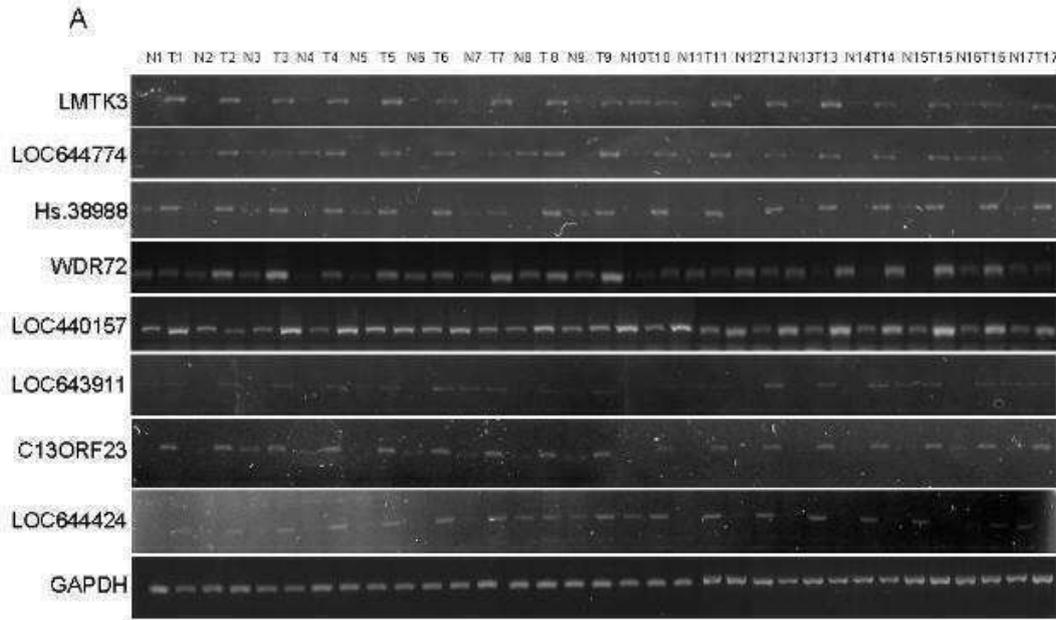
- [0108] 그 결과 13종의 유전자들은 확연히 대장암 시료에서 고발현되거나 저발현됨을 확인하였고, 고발현되는 대장암 마커 유전자 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, 또는 LOC644424는 대장암의 진단, 또는 약물스크리닝 등을 위한 대장암마커, 또는 치료타겟 등으로 사용할 수 있고, 저발현되는 유전자 FLJ21511, C9ORF19, MGC15476, 또는 LOC143381는 대장암의 진단, 또는 약물스크리닝 등을 위한 대장암억제 마커로 사용가능할 것으로 보인다.
- [0109] <실시예 3> 대장암세포주에서의 대장암 진단용 유전자의 발현량 확인
- [0110] 실시예 2에서 발현량의 변화가 확인된 유전자에 대해서 대장암 세포주를 이용하여 그 발현량을 조사하였다. 실시예2와 동일하게 추출한 RNA를 이용하여 cDNA를 만들고 RT-PCR로 발현량을 조사하였다.
- [0111] 3-1. cDNA 합성
- [0112] 실시예 1의 6개의 대장암 세포주(HT29, SW480, DLD1, HCT116, SW620, Colo205) 각각의 총 RNA를 사용한 것을 제외하고는 실시예 2와 동일하게 수행하였다.
- [0113] 3-2. PCR 을 통한 cDNA 증폭과 발현량 확인
- [0114] 실시예 2와 동일하게 주형의 농도 보정, PCR 에 의한 유전자 증폭, 및 발현량 확인을 하였다.
- [0115] 그 결과를 도 4와 도 5에 나타냈으며, 도 4는 대장암세포주에서 대장암 고발현 유전자의 발현량을 나타낸 그래프이고, 도 5는 대장암세포주에서 대장암 저발현 유전자의 발현량을 나타낸 그래프이다. 도 4 및 도 5의 y축에는 표준 유전자인 GAPDH로 보정한 PCR 산물의 량을 실시예 2의 정상 대장조직에서의 발현량의 평균값으로 나누어 준 값을 나타낸 것이며, x축은 대장암세포주를 나타낸 것이다.
- [0116] 그 결과 8종의 대장암 고발현 유전자 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, 또는 LOC644424의 발현량은 대장암 세포주에서 대체적으로 높게 나타났으며, 저발현유전자 FLJ21511, C9ORF19, MGC15476 또는 LOC143381의 발현량은 대장암 세포주에서 대체적으로 낮게 나타났다.
- [0117] 이와 같은 결과는 LMTK3, LOC644774, HS.389988, WDR72, LOC440157, LOC643911, C13ORF23, LOC644424, FLJ21511, C9ORF19, MGC15476, 또는 LOC143381 유전자와 대장암의 관련성을 재확인해 주는 것으로써, 상기 유전자의 발현량을 확인함으로써 대장암의 진단, 약물스크리닝 등이 가능하고, 상기 유전자를 대장암 치료타겟으로 할 수 있음을 나타내는 것이다.

도면

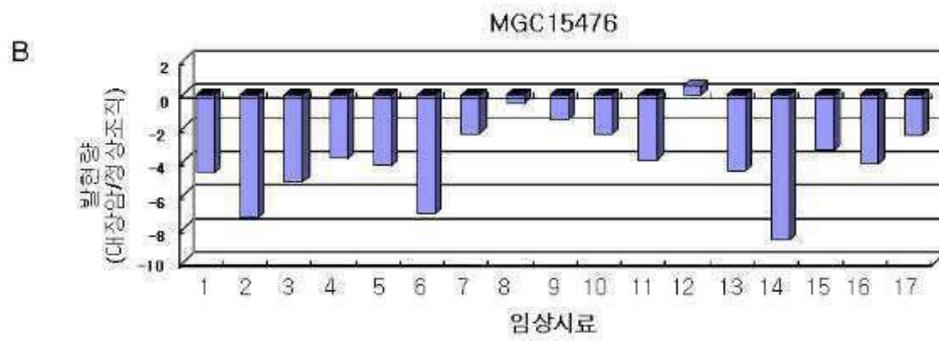
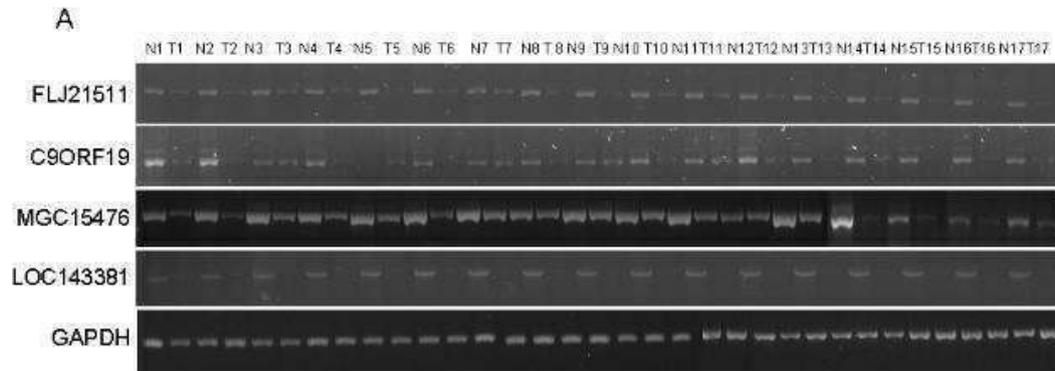
도면1



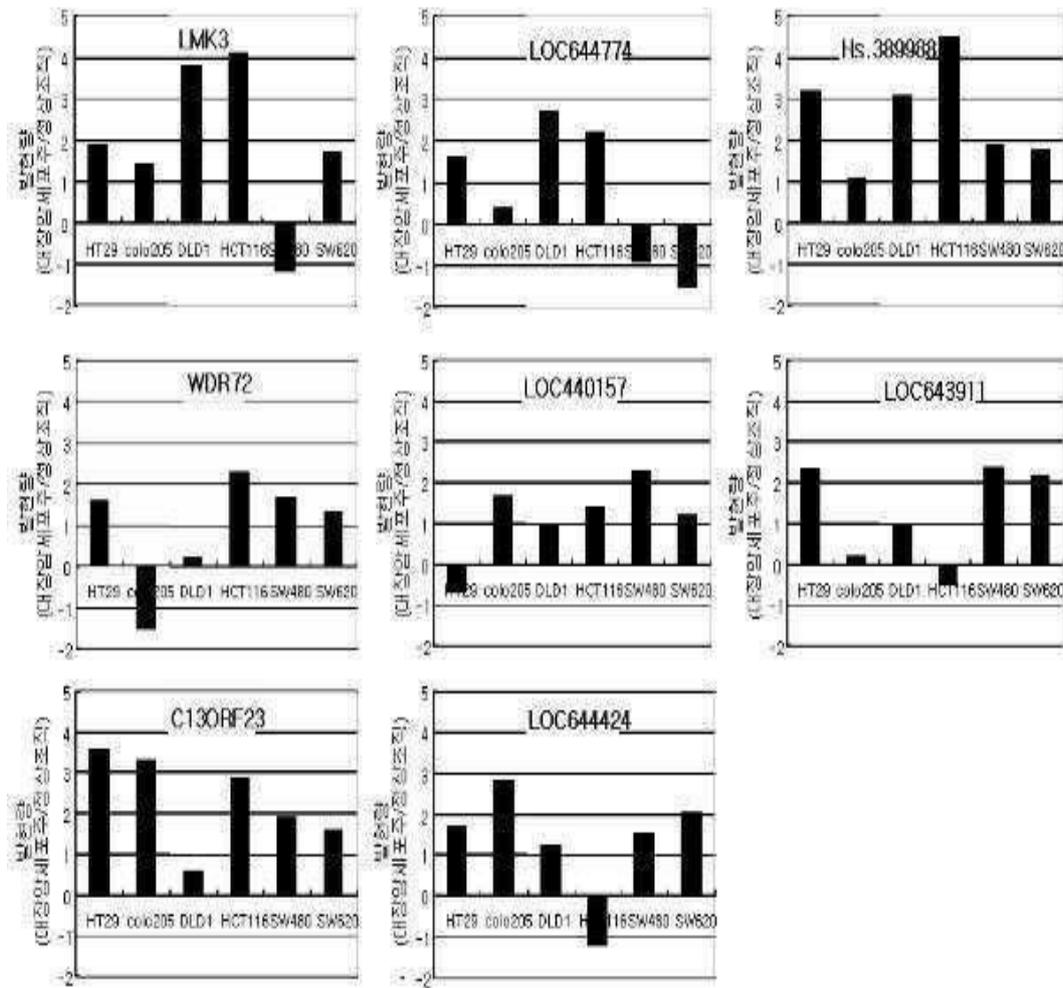
도면2



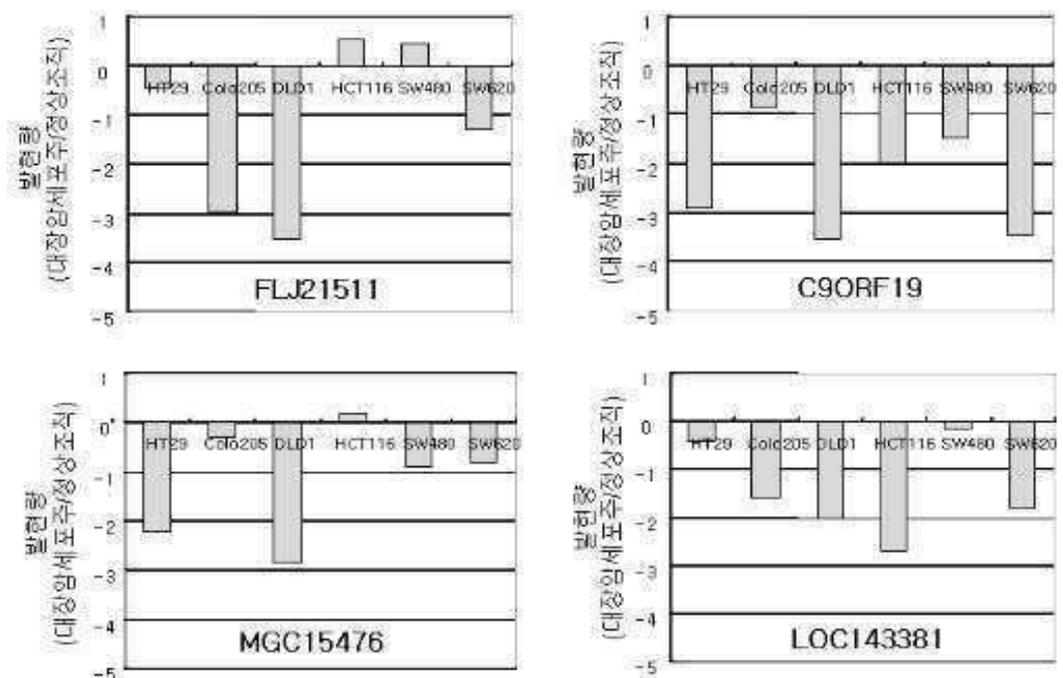
도면3



도면4



도면5



서열 목록

<110> Korea Research Institute of Bioscience and Biotechnology
 <120> Composition for diagnosing colorectal cancer and use thereof
 <130> P07-075-KRI-DA1
 <160> 38
 <170> Kopatent In 1.71
 <210> 1
 <211> 4972
 <212> DNA
 <213> Homo sapiens
 <220><221> mRNA
 <222> (1)..(4470)
 <223> PREDICTED: lemur tyrosine kinase 3(LMTK3), mRNA.

<400> 1
 atgaggcaag tgctgtggtt gtgtaatgtc tgcgtaaccg cacgggaaac cgccaccac 60
 ctccacctcc ctgccatcct cgacaagatg cctgcccccg gcgcctcat ctccttgcg 120

 gccgtctccg cctccggttg cctggcgtcc ccggcccacc ccgatggatt cgccctgggc 180
 cgggctcttc tggctcctcc ctacgtgtg gtctcattt cctgctccgg cctgctggcc 240
 ttcatcttcc tctcctcac ctgtctgtgc tgcaaacggg gcgatgtcgg cttcaaggaa 300
 tttgagaacc ctgaagggga ggactgctcc ggggagtaca ctcccctgc ggaggagacc 360
 tctcctcac agtcgtgcc tgatgtctac attctccgc tggctgaggt ctccctgcca 420
 atgcctgccc cgcagccttc aactcagac atgaccaccc ccctgggcct tagccggcag 480
 cacctgagct acctgcagga gattgggagt ggctggtttg ggaaggtgat cctgggagag 540

 attttctccg actacacccc cgcccagggtg gtggtgaagg agctccgagc cagcgcgggg 600
 cccctgggagc aacgcaagtt catctcgga gacagaccgt acaggagcct gcagcacccc 660
 aatgtcctcc agtgcctggg tctgtgcgtg gagacgtgc cgtttctgct gattatggag 720
 ttctgtcaac tgggggacct gaagcgttac ctccgagccc agcggccccc cgagggcctg 780
 tccccgagc taccctcctg agacctgcgg acgctgcaga ggatgggcct ggagatcgcc 840
 cgcgggctgg cgcacctgca ttcccacaac tacgtgcaca gcgacctggc cctgcgcaac 900
 tgctgtgta cctctgacct gacctgcgc atcggagact acgggctggc ccacagcaac 960

 tacaaggagg actactacct gaccccagag cgctgtgga tcccactgcg ctgggcggcg 1020
 cccgagctcc tcggggagct ccacgggacc ttcatggtgg tggaccagag ccgcgagagc 1080

aacatctggt ccctgggggt gacctgtgg gagctgtttg agtttggggc ccagccctac 1140
cgccactgt cacagcagga ggtcctcgcc ttcgtggtcc gccagcagca tgtgaagctg 1200
gcccggccga ggtcaagct gccttacgcg gactactggt atgacattct tcagtctctgc 1260
tggcggccac ctgcccagcg ccttcagcc tctgatctcc aattgcagct cacctacttg 1320
ctctccgagc ggctccccg gccccaccg ccgccacccc caccgagaga cggtccttc 1380

ccctggccct ggccccctgc acacagtgcg cccgcccgg ggacctctc ctaccgttc 1440
ccctactgg atggcttccc tggagccgac cccgacgatg tgctcacggt caccgagagt 1500
agccgcgcc tcaacctga gtgcctgtgg gagaaggccc ggctggggc cggccggggt 1560
gggggggac ctgcctggca gccggctcg gccccccgg cccccacgc caaccctcc 1620
aacctttct acgaggcgt gtccacgcc agcgtgctgc ctgtcatcag cggccgagc 1680
ccctccgta gcagcagta ctacatccg ttggaggagc acggctcccc tctgagccc 1740
ctttcccca acgactggga cccctggac ccaggagtgc ccgccctca ggccccag 1800

gccccctcg aggtcccca gctggtgtcc gagacctggg cctccccct ctccctgcg 1860
ccccggcct tccagccca gtccctcagc tcaggcagct tctgctgag cggctgggac 1920
cccgaggcc gggcgccgg ggagacctg gcgggagacc ctgccgaggt ctgggggag 1980
cgggggacc ccccgtgggt ggaagaagaa gaggaggagg aggaggcag ctccccagg 2040
gaagacagca gcagccttg aggtggccca agccgcccgg gtccctacc ctgtccctg 2100
tcagccgcg agggggcctg ctctgctg cactggagc ggggggacgc cgtagcaggc 2160
tggggaggcc acctgctct tggctgccc cccccccg aggacgact ctgctgagg 2220

gcagagcggg gctccctggc cgacttgccc atggccccc ccgcctcggc ccccccgag 2280
tttctggacc cctcatggg ggcggcggcg cccagtagc ccgggcgggg gccacctcc 2340
gtccccccc cccgcccgc acctctcgg gcccccggg acccgcccgc gtccccgac 2400
ccccctcg ccgtggccag tcccgttca ggctctcgt cggcggccc caagccgggg 2460
gacagcggt acgagaccga gacctttt tccccagagg gaccttcc aggtggggg 2520
gcggccgagg aggaaggggt ccctcgccc cggctcccc ccgagccacc cgaccagga 2580
gcgccccgc cacctcaga cccgggtccg ctccactcc cggggcccc ggagaagccg 2640

accttctgg ttcaagtgag cacggaacag ctgctgatgt cctcgggga ggatgtgaca 2700
aggaacctcc tgggggagaa gggggcgaca gcccgggaga caggaccag gaaggcggg 2760
agaggcccc ggaacagaga gaaagtccc ggctgaaca gggaccgac agtctgggc 2820
aacgggaaac aagccccag cctgagcctc ccagtgaac gggtagact gctggagaac 2880
ggggaccaga gagccccag catcgaggag aaggcggcg agaatgggc cctggggtcc 2940

cccgagagag aagagaaagt gctggagaat ggggagctga cacccccaag gagggaggag 3000
 aaagcgtgg agaatgggga gctgaggtcc ccagagcccg gggagaaggt gctggtgaat 3060

 gggggcctga ccccccaaa gagcgaggac aaggtgtcag agaatggggg cctgagattc 3120
 cccaggaaca cggagaggcc accagagact gggccttggga gagccccagg gccttgggag 3180
 aagacgcccc agagtggggg tccagcccc acgatcgggg agccagcccc agagacctct 3240
 ctggagagag ccctgcacc cagcgcagtg gtctcctccc ggaacggcgg ggagacagcc 3300
 cctggcccc ttggcccagc cccaagaac gggacgctgg aaccggggac cgagaggaga 3360
 gccccgaga ctgggggggc gccgagagcc ccaggggtg ggaggctgga cctcgggagt 3420
 gggggccgag cccagtggg cacggggacg gccccggcg gcggccccgg aagcggcgtg 3480

 gacgcaaagg ccgatgggt agacaacacg agcccgacg caccgccgc accgctgcca 3540
 ccgccaccgg aggcacagcc gaggaggctg gagccagcgc ccccagagc caggccggag 3600
 gtggccccg agggagagcc cggggcccca gacagcaggg ccggcggaga cacggcactc 3660
 agcggagacg gggaccccc caagcccag aggaaggcc ccgagatgcc acgactatc 3720
 ttggaactgg gacccccca ggggaacagc gagcagatca aagccagct ctcccggctc 3780
 tcgctggcgc tgcgccgt cagctcacg ccattcccgg ggccgggccc gcggcggccc 3840
 ccgtgggagg gcgcggacgc cggggcggct ggcggggagg ccggcggggc gggagcggc 3900

 gggccggcgg aggaggacgg ggaggacgag gacgaggacg aggaggagga cgaggaggcg 3960
 gcggcggcgg gcgcggcggc ggggcccggc ggccccggga gggcgcgagc agccccggtg 4020
 cccgtcgtgg tgagcagcgc cgacgggac gcggcccccc cgctcgggg gctgctcaag 4080
 tctccgcgc gggccgacga gccagaggac agcagctgg agaggaagcg caagatggtc 4140
 tccttccag gggacgtgac cgtctacctc ttgaccagg agacccaac caacgagctg 4200
 agcgtccagg cccccccga gggggacacg gaccctcaa cgctccagc gcccccgaca 4260
 cctccccacc ccgcccccc cggagatggg tttcccagca acgacagcgg ctttgaggc 4320

 agtttcgagt gggcggagga tttcccctc ctcccctc caggcccccc gctgtgcttc 4380
 tcccgttct ccgtctgcc tgcgtggag acccggggc caccgcccg gcccccgac 4440
 gcccgcccc caggccccgt ggagaattga ttccccgaag accgacccc gctgcacct 4500
 cagaagagg gttgagaatg gaatcctctg tggatgacgg cgccactgcc accaccgag 4560
 acgccctc tggggaggcc cccgaggctg ggccctccc ctcccactcc cctaccatgt 4620
 gccaaacggg aggccccggg ccccccccc ccagcccccc agatggctcc cctgaccccc 4680

ctgaccccct cggagccaaa tgaggcagga atccccccgc ccctccatag agagccgcct 4740
 ttctcggaac tgaactgaac tcttttgggc ctggagcccc tcgacacagc ggaggtccct 4800
 cctcaccacac tectggccca agacaggggc cgcaggettc ggggacccgg accccccatt 4860
 tcgctctcc cctttccctc cccagcccgg ccctggagg ggcctctggt tcaaaccttc 4920
 gcgtggcatt ttcacattat ttaaaaaaga caaaaacaac tttttggagg aa 4972

<210> 2
 <211> 282
 <212> DNA
 <213> Homo sapiens
 <220><221> mRNA
 <222> (1)..(282)
 <223> PREDICTED: similar to Phosphoglycerate kinase 1(LOC644774), mRNA.
 <400> 2

atgatgcttt gggcactggt tcacagagcc cacagctcca tggtaggagt caatctgcca 60
 cagaaggctg gtgggttttt gatgaagaag gagctgaact actttgcaa ggccttggag 120
 agcccagagc gacccttctt ggccatcctg ggcggagcta aagttgcaga caagatccag 180
 ctcatcaata atatgctgga caaagtcaat gagatgatta ttggtggtgg aatggctttt 240
 aagttcctta aggtgctcaa caacatggag gtaggaactt aa 282

<210> 3
 <211> 488
 <212> DNA
 <213> Homo sapiens
 <220><221> mRNA
 <222> (1)..(488)
 <223> K-EST0184252 L9SNU354 cDNA clone L9SNU354-11G075, mRNA sequence

<400> 3
 cttattttct ctgaacgac ctgattccag tcactttggt gaatacccta gttctaataa 60
 ttgactcttg cttttctaga gaaatatttc caaatgatgc tagttttgtc tcttcctttc 120
 aaagttgtat accacttctt tttctgtca ttttgcatg cctgggacct ccagaataat 180
 gtttcatgaa gtagcatgta tccatatctg gttcttgact ttttcatcat tataattggt 240
 ttctatgggt tacttatcag ttaagaatg ctaattcct agatgaacta agagtgttta 300
 ttacatggtg agatttatgg tatgcttttt cttctcaag ataatgcatt ttttgtatta 360

tctgttaatg tgataggta tccatttgig tattttcaat cattgaacaa cccttgattt 420

ttttggataa acctatttg gtcattatgc atcattctat aaacctgct gaattttca 480

tttgccaa 488

<210> 4

<211> 7326

<212> DNA

<213> Homo sapiens

<220><221> mRNA

<222> (1)..(7326)

<223> WD repeat domain 72(WDR72), mRNA.

<400> 4

cctctgccgc tctccgccc cacggctcag gattcgccca aaatgaggac ttccctgcag 60

gcagtggcac tctggggaca gaagcccct cccacagca tcactgccat catgatcact 120

gatgaccagc gaacgattgt gactggaagt caagagggtc agctctgtct ctggaatctc 180

tcacatgaac taaagatttc agcgaagaa ctctatttg gtcattcagc ttcggtaaca 240

tgtttggcaa gagcaagga cttctctaaa cagccctaca ttgttagtgc tgctgaaat 300

ggggagatgt gtgtttgaa tgtcaccaat ggacagtgcg tggagaaggc tacacttctc 360

tacaggcaca ctgcaatctg ttattaccac tgctcattcc ggatgacagg agaaggctgg 420

cttctttgtt gtggagaata tcaagatgic cttataattg atgcaaaac tttggctgtt 480

gttcacagtt ttagatcacc tcagtttctc gactggatca actgcatgtg cattgttcac 540

tccatgagaa ttcaagaaga ttctctcttg gtggtatcag tagctggtga gctcaaagta 600

tgggatcttt cctcatctat caacagcatt caggaaaagc aagatgtcta tgaaaaagaa 660

tccaagtttc ttgagtcctt gaactgccag acaattcgat tttgcacata tactgagaga 720

cttctattgg tggatatttc taaatgttg aaggtttatg attattgtga tttttcctt 780

ctgctgactg aagttagtag aaatgggcag ttccttgctg gtggagaagt gattgctgct 840

cacagaatcc tcatctggac agaagatggt cacagttaca tctatcagct gctgaacagt 900

gggctttcaa aaagcatata cctctctgat ggaagagtgc ttaaagagac cttttatcct 960

catttactgt gcctacttc tgtgcaggaa aataaggaac agagccgtcc ctttgttatg 1020

ggctacatga atgaaaggaa agagcctttt tacaaggtac ttttctctgg agaagtctca 1080

ggaagaatta ctttgtggca catcctgat gttcctgtat ccaagtttga tggttctcct 1140

agagagatac cagtaactgc cacctggact cttcaagata attttgataa gcatgatact 1200

atgtcacaaa gtattattga ctatttctct gggcttaaag atggggcagg aactgctgta 1260
 gtcacttcat cagagtatat tccaagtctt gataaactaa tatgtggctg tgaagatggg 1320
 acaattatca ttaccaggc tttgaatgct gccaaagcaa gacttctgga aggtggttct 1380
 ttagtaaaag attctcccc tcataaagt ttaaaggcc accaccaaag tgtcacttca 1440

 ttactctatc cacatggtct ctcttcgaaa ttagaccaa gttggatggt gctctggggac 1500
 ctggactcat gtgtgatctt gtgggatac tttactgaag aaattttgca taaattcttt 1560
 ttggaagctg gtccagtaac aagtcttttg atgtcaccag agaagttta actaaggggt 1620
 gagcagataa tttgctgtgt gtgcggtgac cattccgtgg ctctccttca ccttgaggga 1680
 aagagtggc tctgcatgc ccggaagcac ctttttctg tgaggatgat aaaatggcac 1740
 ccggttgaga attttttaat tgttggatgt gcagatgact cagtttatat ctgggaatt 1800
 gaaacaggca ctttgaaag acatgagaca ggagaaagag cacgaattat tcttaattgt 1860

 tgtgatgatt cacagctgt gaagtctgta ctccattg cctcagagac acttaagcac 1920
 aaaagtatag aacagagatc ctccagcccc taccagcttg ggccattacc ttgccctggt 1980
 ctgcaggtag agtcttcatg taaggttact gatgccaaat ttgcccag accttttaat 2040
 gctttgcctg tgaagacaaa atggagtaac gttggctttc atattcttct atttgatctg 2100
 gaaaacctg ttgaactttt gctaccaact ccactcagt atgttgactc ttccagttea 2160
 ttctatggtg gtgaggctct gagaagagcc aagagcacag tggagaagaa gacactgaca 2220
 ctgagaaaaa gtaaaactgc ctgtggtcct ctttcagcag aggcactagc caagcctatt 2280

 actgaaagcc tggcccaagg agataatacc atcaaattct cagaagaaaa tgatggcatt 2340
 aaaaggcaga agaaaatgaa gatctccaaa aaaatgcagc ctaagccatc aagaaaagta 2400
 gatgccagtc tcacaataga cacagcaaaa ttgtttctgt cttgcctttt gccatgggga 2460
 gtggataaag atttagatta tctttgcatt aagcacetca atattttaa gettcagggt 2520
 cctatttctt tgggaatttc tttgaatgaa gataatttct cactgatggt gccaggttgg 2580
 gatttatgca atagtggaat gataaaagac tattcaggag taaatttatt ttccaggaaa 2640
 gttttggact tgcagataa atacacagcc actcttccaa atcaggttgg aattccaaga 2700

 ggattggaat ataattgiga ttctttgca gagtcagata ctatagtta tttgttgagc 2760
 agactatttt tagttaataa attagttaac atgcctttag aattggcatg tagagttggc 2820
 agttctttca gaatggaag tatacataat aagatgagag gtgctgggaa tgacatttta 2880
 aatatgtcaa gcttctacag ttgcttacga aatggtaaga atgaatcca tgtacctgag 2940
 gctgaccttt cacttttgaa gctaatttcc tgttgagag accagtctgt gcaggttaact 3000
 gaagcaatac aagctgttct cttggcgaa gttcaacaac acatgaagag tttgggaaag 3060

ataccctca atagtcaacc agtgtccatg gcagagaatg gtaactgtga gatgaagcag 3120

 atgctgccaa agctggaatg gacagaagaa ctagagttac agtgtgttag aaacactttg 3180
 cctctgcaaa ctccagtcag cccggtcaag catgacagca actcaaactc ggcaaaactc 3240
 caagacgtgg aggacatgcc tgacagatgt gccttgggaag agtctgagag tccaggtgag 3300
 ccaaggcatc attcatggat agcaaaggtc tgcccctgca aggtgtctta aatggaatct 3360
 catcagtagg agctgaatit ggacaaatta agaaatccaa aagatgcat ttgtttatta 3420
 ctgtataaaa gcattgttgt tatttggtcaa gttattagc tgtagtggat ttgctaatac 3480
 tttagccaac atgtattaaa gtgattttaa tacatgctga ttacaatgca atacatactg 3540

 attgaaaata ttcatattca tctaatttta gaaaaatatt gcctagatca ttctctattc 3600
 ctgtttctta cttttctgt taatatttcc aacagggat gccagtccac agacacaaat 3660
 ttaaactgc tttaaatit ctcctatcct tttagtcct gaattatata ataaacaatg 3720
 ttaaaccaa tglagtacac aatacttact tacaattta atactgctc aaggtattta 3780
 atctaaaatt ttaccaactt tgatttgtct ggttaggata tttgtttta gtggatatgc 3840
 tttaatcgg atcaactact gcagtaaatc tcaccctaa gcatgaaatg ttgtcaacaa 3900
 ataccagtt ccatttagtt atcaattagc ccaataaga gatacaaagt ataacagtga 3960

 ccaacctgt actgttagt taatttgaac ttctaagac attgaggcta atgtctttag 4020
 ctcaaggtg atctgttgg ccatatagat gtgaactagg gaaggggaat caacttacag 4080
 catatcaaa ttgatcctta ttaagtataa actctttag gtcctttccc agaaagaagc 4140
 ttgactagca ggaattctaa aactgaaata tatcaaacag cataaatagg aatagacata 4200
 aagtgtctt ctattaaagc ctttggat ctatttacta tgatttatat tgtacagttc 4260
 ctcgatttac agaaaatcat caaaattatt aatctacata tcttatgtat ataaatattg 4320
 cctaaccat agaaaaagg atataaagta ttaaatatgt gatatatagc tatatctatc 4380

 tatctatgta tctaataagg aagtccaagt cacttcaatt gaagaaacat atctctgagc 4440
 ataggagcag cctcaggtcc tatggtggga tgcagtggac aggagagggg gaaattagaa 4500
 aagagaacta tataattgaa aaaggatata aaagcattaa atatatgata tatagctata 4560
 tctatgatg tatctaacag agaagttaa gtcacttcaa ttaaagaaac atttttgagc 4620
 atgggaccag cctcaggtct tatgctggga tgcagttagc aggagatggg gaaattagaa 4680
 aagagaactg tgaattgaa atgactggg ctgcaccctt aaggaaacta taattaaatga 4740
 tgatctgaat aaacatacca ggataaagat gtcaaatgag tgtgactccc ttaaagtaga 4800

ttaaagtgtg cattctttgt ttcctaaaat atgattttac tgcttgaaat tacatttgag 4860
 ttgaagttta gaaactaaca tagcattaat atgaataatc atggaaaatt attatccttt 4920
 gaaaactgat tgataaatat attccccctc ctttagaaac agtcaaaagc cacttcaaac 4980
 aagtttcaaa ataaaggaag gtagcaagtt aggcgatgga ttatattttc ttgcttgttg 5040
 tataccagtt gtcaaggaca ttataaggac tcccaaagc attttgaagg atggcaatat 5100
 caaataagtg tatgtcctct caaatgaggc atttttaatt gttaaaatct atttggacgc 5160
 tcaggttatg atatgtttat gaaaaataag cttcattatt tttatagcta catcctatta 5220

 ttcctttta gaaacaagaa taacaatagg ttttaatagt tgccatactt agcatttacc 5280
 agttctaag aaaccaatat tgaatctctg ataaatattt tctgatgtta ctagctatgg 5340
 gaaattagaa ctggcacaac cctgacatta ctaagtggaa atgttaggat ttttcggcat 5400
 cgcagttag aatctctaaa atttaacat tccgtttaa tgactaaggt ttgcttttat 5460
 caatatgaat tcigaaggcc aatatcatic cattaactat gaaagctttt aattcctaaa 5520
 aatagttta gagatattca agcaatgctc tcctaatac caticgcaag tgtgtttatg 5580
 acacaaatc actagtctgt ttaaaaatga atctttata ttgactgggt ttccacatat 5640

 ttcagtaatt tctgttatga gaggacttga aatagcaaat tgccacacag ttaactggat 5700
 agaccacgta cgtgggtgac ataaccactt ggtactacac ccagaaactc aaaattgtct 5760
 ttctctgat gagatatggg tgtccttttg tacgtctagg cctaggtaac cagtggagtg 5820
 attatattag caaatgtgtt tgtatccgga gtcttctgt cattgtaata aaaaatttat 5880
 ttaaaaattt tgtcttgtgt ttcttttttg gtaatcatct tggatgatatt tatgaatgta 5940
 tttgtgtgtg tgtgtgtgtg tgtatacata catatatatg tgtgtatata tatacacaca 6000
 tatatatata catatatatg tgtgtatata tatatatata tggcatatcc cccaaaggaa 6060

 tatcatattt tcatttgcga gccttgattc tttgacttgc tgttgagttt cacattgget 6120
 ccattgtata ttgcaggaaa aagtacctgg aaagcctgga gggtaaaaag cagaggctcc 6180
 tcggcagcgt gtccaggccg tgtgtgcttg ctgtgcttgt cagagcctca ttcctcccc 6240
 tgtgttggag agctttctca cagggtctcat ccagaggtgc tagaaagcaa tgtgagcact 6300
 attaaacact tcaaatataa ggaaaaatgt ataatgttag aaacacacag ttaaaaagta 6360
 gccccaggta tcaaacatc tgttactatt ttagaaacta aattaaggct gctttgaaca 6420
 tgatgttatt agaaaaatgg tagaatacag tttgttaaat attaataatc tatctttaga 6480

 tgatagcatt tggtagcttt tatttctttt ctacacatt tttccagtg ttctgtattt 6540
 ttaaaatcct ttcataatta ttataatgaa cacttttagaa aagatactat ttctcttcca 6600
 tccttagagg gcacattata ataattgagt attttaaaat taaattttat gccatgtatt 6660

taaaatgaaa gttcaggtta cattaattcc aactgatttt aaagccttat tttcacagag 6720
 tggactcatt caatTTTTTA aattttatat tcatacattt tgattgctca gttatataaa 6780
 agcgggtgta tagactcctt tattctccac aaaatgttgt cttatatgta aagtgcctaaa 6840
 agtttggttt aatttcctaa gaaagtcatt gctttgtaaa catttgcata atacttagtg 6900

ttttttgtaa ttttaaaaaa ttagcttaac caaaagagtt gcaaaattaa aacttacttg 6960
 agtaggtatc tttcagacaa agctgcatgg cttcatgtac aatagtcctc agatcacctc 7020
 gcactgtagt ttgctcaaga aggagcctac agatggcagg taggagaagg taaactggaa 7080
 attaacaggt ctgtctgtaa gagcatgcca ctggaaacac ccatccaatt tgggaccctc 7140
 ggaagacagc atgtgtctaa tgggaggcct gtccctagtag aactccttta ctcttttcac 7200
 cttcatcttg accttctcag ataccatctt tatgtaaaat gttgatgttg tggagttttt 7260
 gcagtaattt aatttttata aatattgtat taaatataat ttatcttgaa aaaaaaaaaa 7320

aaaaaa 7326

<210> 5
 <211> 2779
 <212> DNA
 <213> Homo sapiens
 <220><221> mRNA
 <222> (1)..(2779)
 <223> hypothetical gene supported by AK096951; BC066547(LOC440157),
 mRNA.
 <400> 5

cagcgatcag cttggctgcc cttgcttctg agtggccacag tagaggctag gggagcaact 60
 ggctttcctc cccaaaaggc gggcaggggt atccacactt tgcccaggtc cctgaagcct 120
 gcggctgagc tcggggataa caggggctaa gtcaccggtc ccagacacct aggaactatt 180

agagacagga accagcatat gacacagggg ctgttaagta ggaggttgga gaacacacgt 240
 ttttggtcta aaccgggggc ccctctcttt gccactgag cccgcggcct gcgtggtgct 300
 gagactgcct ctggccgctt ccgcttggga caaggcctga gcggtggctg atccacactg 360
 gatgtcccgg gccggtctcc acccgaagcc cgccatcccg ggacgcggtg gggagaagct 420
 ggcactgctc cttgccatgc ttggcggccg ctgctgcccg gctgggggtc ccgagtcgca 480
 cagccctgc aagccctggc caccgatccg gagggaacgc cctgggctgc ggtccccgaa 540
 gccaaagaaa gaagcaggtc ccagggccga ctccaaagcc gcatctccag ctttgttcat 600

ggggtccggga agcagaggcc gccgccggcc accgtcgtgg gcgagaagaa gggcacgagg 660
 cggccggggc tctgccccg aaccacatgt gcgcgccggg ccccgttct tcatcgact 720
 tgcggccccg gctgccccgg gcctgcgagt ttccagccag ggccccggac tctggcgagg 780
 tccggccgcg aggaaggaag gcgtggcccc ggtgggggta gcggcaggcc tgcggctcag 840
 gccatggggc aggggcagaa aaacgacccc ggcgtgtcc gggcatccag ctcggttccc 900
 gctgcagcca ggagactccc gggagcgctc taggaaccac agagccctgg aactcacctg 960
 gcagcctcgc ggcgctaaag ccggcgggagc ctgagacagc gcgcggcgag gcggtcacgc 1020

 tccacccccg cgtggcgcca ggactcggat ttcgccctg gttttaaata tgtgccggtg 1080
 gagccccgga cgctgggaag agcgttctgc gccctccag tgcgggtctc cgccctaac 1140
 cgacttccag agccgcctct gctccctgga ggggcgcagt ggcggacacc ggcgtccac 1200
 gaagtgcag gtctcagtc tgagggtgc cccgcacgct cggaatgcag gagggtctcc 1260
 gcctcgtgc gctgccctg ggggcggagg cgtgcgctgc aggcgagaga ggcggccccg 1320
 tatcgatgga gaagcacaga gggctttgag gtcgcaact cccggttgct gagcggagtc 1380
 aggagtccag ttccaaagg acagcgctca gggttgtaat caccacccgg cccaccgctt 1440

 ccgcagctgc gactctagg cggagctgtt gggtagaccg agcaggcgag gcgcaggcag 1500
 gcagcggctc cgcctcgaa tccgcctga cggggcaca ggtgcccc ccacctgtcc 1560
 ctcggtcacc ccaacctgt ttctcgacc cccagcactc ctccaggcct agttcgcttc 1620
 agaggcga gaccggaaa acaaggaaga agcagctca gcctcaatcc ccgtccccac 1680
 cccactttcg ggaccgctaa gctggagaat tgaagggggc ggacccccga ttaaagccgc 1740
 tcccttccca gcctgcccc gctttcctaa tgtccgtgat gatttcgtta ttggcaggg 1800
 agagccagac tcctgcct cccaagacgg ggcgattggg agggggttct ggagctcatg 1860

 cctggggtcg gcccgcgagg ggtgacccc cgcctcgcc ggtgcaagga gaacagctgg 1920
 ttcccgccgg ggcaggaag cgtggacggt gtgggctcag gcgcctggca ggcacacggg 1980
 gcctctaaag cttggtcact gtcacagatc gttggttgt ttcttcgct cccgccacgc 2040
 ctctctctg ggatggggat tcattcccta gcagggtcgc gagaactggc gccttgcag 2100
 ggtaggcgc ccggagcctg aggcgggaac tttaaaatca gacgcttggg gcccgggctg 2160
 ggaaaaactg gcggaaaata ttataactga actctcaatg ccagctgttg tagaagctcc 2220
 tgggacaagc ggtggaagtc ccctcaggag gcttccgca tgtcctaggt ggctgctccg 2280

 cccgccaggg tcatttccat tgactcacac gcgccctg gaggaggagg ctgcctgga 2340
 cacgccggtg gcgcttctc ctgggggagc gcagcctgga gctctggcgg tagcctggg 2400
 agcggggcct cggaggctgg gcctggggac ccaaggttgg gcggggcgca ggaggtgggc 2460

tcagggttct ccagagaatc cccatgagct gacccgcagg gcggccgtgc cagtaggcac 2520
 cgggcccccg cggtgacctg cggccccgaa gctggagcag ccactgcaaa tgctgcgctg 2580
 accccaaatg ctgtgtcctt taaatgtttt aattaagaat aattaatagg tccgggtgtg 2640
 gaggtcaag ccttaatccc cagcacctgg cgaggccgag gagggaggat cccttgagcc 2700

cagagttcg agactagcct gggcaacaca gtcagactcc atccttccaa aacaaacaaa 2760
 caaaaaaaaa aaaaaaaaaa 2779

<210> 6
 <211> 558
 <212> DNA
 <213> Homo sapiens
 <220><221> mRNA
 <222> (1)..(558)
 <223> PREDICTED: hypothetical protein LOC643911(LOC643911), mRNA.
 <400> 6

aggtgttaag tggatgctt ccataataca tttggatgct gtcagctaag ttcacttctg 60
 aactaagggg ttctccaaa tgttggtga aattcatccc aaggtggtc tgcaaagtct 120
 gcaattcata atggagctac tgtactggct attggaagga ggagattctg aagataagga 180

ggatgccact gaaatgttg aatgaaaaa tattcagccg ttggtctttg aaatttctg 240
 tgatgtgttt caatctagat gcaaagaaca tggaaaaatc aaagtgctcg agtggtttaa 300
 atatgttttg ggtattcctg tttatagact ataatacttt tccaattaaa atcctcagtt 360
 gtcacgcaga agaaggttaa gctgtattg attgccagtt ttactgaaaa tgcttagtat 420
 tttacagtat caccaaatat atttgttta gccaaaggtat aggaaaaata aaataaattg 480
 tataggttga ctttttcta aaatgtcttt attggattga atgaatgttt atacctgaaa 540
 aaaaaaggtt caaaaaaa 558

<210> 7
 <211> 5228
 <212> DNA
 <213> Homo sapiens
 <220><221> mRNA
 <222> (1)..(5228)
 <223> chromosome 13 open reading frame 23(C13orf23), transcript variant 1, mRNA.

<400> 7

acaccaggac cccacctctg gctttcagga aagaggggcc caagctggcg gcgcagtgtt 60
gagtttactt gggcgcgact ggctttattg taggtcccca tgagctggtg ggggctgtag 120
gggtaaccag gcagggcgctc tggcgtgtt ggggctgcgt ttttccggag tccccctatg 180
cgaggccgtg ccctcttcg cccggtcct ggtgctcccc agccgcccgc ccgtcccgg 240

aagctgcage agctggtaac aaagagcctg cgggcccgc tgctgccggg gccgaggtat 300
ccaggagtgc ggcggagcgg ctggggatga gcgcgagccg gaccatgcgt ccgggttcgc 360
ggcggtgagg gtcgcccgc gcggccgagg cgcaggtgcg ggtcgagaga ccggcgcctc 420
cctccccgc gccctggac tctgagcga ggtgttcta ctctgccgc cgctccacct 480
gggaaagcgc ctcaagagc tacctctggc ggcacctggc cctcagctgc agggaaaccg 540
cactcaaacc tgcgccgag cccgggtttt tctgcaaac tcagaggacc tggctcttgc 600
ccgaggtgcc ctgaaccatt gtgtgaataa gctggggaag cctgactgga gtgtttttc 660

ttttaataa aaagagttt tctgaatctc aatttttaga aattttttt gcgaagttt 720
caatactctc tacctcttcg cactctctct tctctgtgga gtctttttcc tectcagcta 780
taaataatit ttactatcgg cggagcagac caggatgaa cgtcttttaa ttgcaagtat 840
aactgttaaa accacgtcgg gatagtagtc aagatggata aaaagtcctt tgaatggtg 900
ctggatgaaa ttagaaagcg tgttttgaca gaatacaaat taaaagcaat tgaatatgtg 960
catggatact ttctagtag acaggtggtt gatttgctga gatatttttc ctgggccgag 1020
ccccagttga aggcaatgaa agcattacag cataaaatgg tggctgtcca gccaacagaa 1080

gtggtcaata tactcaactg tttcactttc agtaaagaca aactagtgc tcttgaactg 1140
ttagcctcga acattattga tgcacagaat tctcgtccta ttgaagattt attcagggta 1200
aatatgtctg agaagaaacg gtgcaagaga aacttgaac aggctttcaa gggggctgc 1260
aaagctctc atgctatgat atcttctgt ggaacaatcc caggaaatcc atatccaaa 1320
ggaagaccta gccgcataaa tggaaatctc ccaggaactc ctttgaaaaa agatggtgaa 1380
gaatgtaacta acgaaggcaa aggaatagct gcacgaatc ttggccatc caaacacct 1440
cctcaacat ataatccaca taaacctgtt ccttatccga tacctccatg ccgaccacat 1500

gcaactattg caccaagtgc ttataacaat gcaggtctgg taccattagc gaatgtcata 1560
gctccacctc cactccata tactccta cctgtaggaa cagagaatga agacctttcg 1620
aatccgtcaa aacctafaca gaatcaaaca ttttccacc cagcaagtca actctttct 1680
cctcatggtt ctaatectc aacctctgt gcaactctg ttectactgc atccccagtc 1740
aaggcaatta atcatccatc agcatcagca gctgccaccg tttctggaat gaacctgctg 1800

aatactgtcc ttctgtggt cccagggcag gtctctcag ccgttcacac acctcagcca 1860
tcaataccaa acccaacagt tatcagaacc ctttcattgc ccactgcacc tgttacatcc 1920

atccacagta caaccaccac tctgttctt tccatttttt ctggcctagt gtcactgcca 1980
ggtccttctg ccactcctac cgcagccact cctaccccag gacctacacc acggtcact 2040
cttgggtcca gtgaagcatt tgccttctact tctgcacctt tctactagcct ccccttttcc 2100
accagctctt ctgctgcttc taccagcaac ccaaattctg cttcattgtc atcagttttt 2160
gcagggtcc ctttgcctt accaccaaca tccaaggcc tatccaacc gactcctgta 2220
attgtgggtg gctctactcc cagcgttgcc ggtccacttg gtgtgaacag tctcttttg 2280
tctgcgttaa aaggttttct gacatccaat gacaccaatt taatcaactc ctctgcttta 2340

tctctgctg tcacaagtgg gctggttca ctatcttctc ttactcttca gaactctgac 2400
tcttctgctt cagccctaa caagtctat gcccacag ccatcctac cccacagagg 2460
acttccactc cagggttggc cctgttcca ggctgcccgt ctcccgtggc taactcaact 2520
tccactccc tgacattgcc tgtacagtct cctttagcca ctgctgcatc agcttccacg 2580
tcagtccag ttagctgtgg ctctcagcc tcccttttgc gtggcccaca cccaggtacc 2640
tcagatctgc atatttcac taccctgct gcaacaactc ttcctgttat gatcaaaact 2700
gagcccaca gtctactcc ctggccttc aaaggtccat ctattcttg gaatcctct 2760

catggcactt taggtttgct agggacattg ggccgtgcat atacttcaac atccgtgccc 2820
atcagtttat ctgcttgctt taatctgca ttgtcaggtc tctccagctt gactactct 2880
ttaaagtgtt caaatctct tctctatt tccctccac cacatggttc ctccactccc 2940
attgcaccag tattcactgc tcttcttct tttacttctt tgaccaacaa ttttcttta 3000
actggcaacc catctcttaa tccgtcagta tctctcccag ggtcattaat agccactca 3060
tctaccgtg ccactccac atctctccct catcttagct caacggcagc tgttctctca 3120
gggttttctg cttcagcacc agtctcagca gcaccttcc cctcaacct gtccactgct 3180

gttccctcac ttttctctgt tactcaagga cctctgcat cttcaaatcc ctctatcca 3240
ggcttttctg tcttaatac cccaagcgtt acccctgctc ttcctcatt cccggggctg 3300
caggcgcct ctacagtcgc agctgtcaca ccaactctg tggctgccac agcccctcc 3360
ccagctccag tctcccagg attgcctca gattcagtt ccaatttcaa ctccgtctt 3420
gttgacaag ccggtttatc atctggactt caagctgcag gcagttctgt tttccaggc 3480
cttttgtccc tcccgggtat cctgggttt cctcagaatc cttcacaatc atccttgc 3540

gaattacagc ataatgcggc tgcgcagtca gcattgttac agcaggtcca ttcagcttcg 3600

gctctggaaa gctatccagc tcagcctgat gggtttccta gttatccttc agcgcaggga 3660

acaccatfff ctttgaacc aagcctgtcc cagagtgggt ggcaagtgaat acttttaact 3720

tttattctcc ttcagagcaa catcagaatt gcctgagaac tgcaatgaac aatctgacaa 3780

atgtgaagct ggccaaaagt cggaaaatga gaatgagggt aatcctggag aaattgtgac 3840

aacaatttga aaattgtggt tgcattttaa agtgtgaaca ctcccctatg taaatatgct 3900

gacaataaat tgtatggaga atggatttta aaaagtgttt ggagactttt cacctgtcct 3960

ataaaatfff gaattgtgta tgtgatctac atagaaagaa tattaagag taggttgaac 4020

tctttatagc cgaatacagc cttaaatatg cttgtatagc atccactggc agaagiataa 4080

gttgtgcctc agacttgggg gttgcatgtg gcctggggg agttactacc cttggtatgc 4140

atgagcgggt cctattagca tcagtgggaa ctcagtactc tgtatgtatc cacaaaaggg 4200

aacttgagac ccacagtat tcttaatttc tgatattaac aaccgtacat actgctgaat 4260

ttaactcaaa atatttcagg taagtгааag tggtgcttaa tgtagactat agaatgactt 4320

tcagggtfff tcaactgaaa gtatatatcc agaactgcat ccttatagaa atacaagtaa 4380

gacttaggat aatttgcctt caaaacagt ttcctaactc cagcagtatc cagtgagtga 4440

agaacacttg actgactctt gggccacctc tgttacttac tgtactatgg aagctcctgg 4500

tgaatgttta caattatggg atgtagtatt tctatttga ctttaagtca aatgcttata 4560

tgaaatatgt gacaacaaat agagaagact ggctctgtta gtaattatgc agtatgtact 4620

ctatttaagg atctgtggtg gtataacatg agtgaatgtc attaatfff aagtaataac 4680

tgccacatgt gggaagtagg ggagtaagga gaatgaattc caatctgtga ttaaaagtgt 4740

aaactataga ctctactgta gtacatttca ggatctagaa gttttacttt tataaagatg 4800

gtgtccggaa gatgttgcta atgtatfff cttcaacata gggaacaaac tttttaagta 4860

tattaataaa cctgtatggg tagtttttaa cagtttttta aaataaacta tggatatgac 4920

aaatatctg tgttttacta agtgccttga taggctttct aattttgtat acgtgctaga 4980

gttaattatt gaacatfff atccaaattt agttgtaact ctgtttatac tactgattgc 5040

tcattcgttt aaatgatatt ttaatgtaa agtcataacc aacatatgaa cagacagatt 5100

tatgtcttta aacacagaat gtaagctata gtttaactg ataccagttg ctggaagttg 5160

ccatttgttt ttcttaaatc tataccata aaacttcttt taagattaaa aaaaaaaaaa 5220

aaaaaaaaa 5228

<210> 8
 <211> 1610
 <212> DNA
 <213> Homo sapiens
 <220><221> mRNA
 <222> (1)..(1610)
 <223> PREDICTED: hypothetical protein LOC644424(LOC64424), mRNA.
 <400> 8

```

caaaaggcca aagtaaaaaa aaaaaaaaaa ggaaaaaatg aaaaagaagc cattgaatac      60
cctattgacc ataagaaaaa ggcagggaaa gttgttcaac ggcaaccag gcaactgaagg      120
aaaaagaaca tggagaatga caaaagccca gagggaggag taaaaggaca caaacaccat      180
tctcagggac caggactggg cgctgttctc agggaccagg actgggcact aatcacagaa      240
ctgtaacagg cgccccatgg gaatgaccaa ctgttagatg gggcctgcag ggcagcacct      300

ccctcctgcc tcccaccaac agcttctaaa gggaaatgcc gactcttttc acaccagtcc      360
cctcactgcg gctgagtggt tgggcgcaga tgataggtca caacaacctg attcagtcct      420
cactcgggct gagtgtgtcg ggtgcagatg acaggccacc acaacctgat tcagtcccca      480
ctgtggctgt gtgtgggggt gcagatgaca ggccaccaca acctgattca ggattcagtt      540
ggctaccagc cagtgccata aggaaaacca ttctggggct cttgagaggg gcaaagcata      600
at ttgcatgt gggagaaatg ttaatagttt gtgaccagag gacaagctgt ggtttattaa      660
agactgctgc aggttcttac tatgttctc atcaagaggc ggaatctaata caccttcccc      720

ccttgaatca tggctggctc cagtgatgag tacgactgga cagtgtggca ggagagatgc      780
tctgggactt ctgaggggcg atcatgagag gccttacagc ttctgcctgg gcctcttggga      840
cacacacctt gggagaagcc agacaaacct gactacctga cactgccaga ctgggaggaa      900
gtccatgctg gccacaaaga gagggctggg tgctgtctcc atgtccccag ccaactagat      960
ccttctgggt gctgcttca cgteccccag cactagagtc ettccagatg agaccaggga      1020
catcatgaag caaccaacct acacagccct gtctgtgtc ttgaccaga aaattgtgac      1080
atgtaaaaag aataaattcc tggtttaagc cagtaaggtt actggtacat tgttacatct      1140

cagataatta aaaccttga aactcatga gagatcacia gtagaacctt gatctgaaac      1200
atggcatgtg gcgatttata ttgagtatta ggttaaaaat gcaagaaggg agcatagtta      1260
atattttacg ttaaagctaa aactataatt gcctacttaa aattttcagt taattaggtt      1320
gtcacttttt gticttaacg aagaaatcaa ctagttttat tccataaaca gttagaactg      1380
atgcacacat ccgtttctcc ttactcattt taacagcta tctgaaatag gaagtgtaat      1440
    
```

ataattttta aagaatctga aaacatgaca gaaatgttta aactataaac atatattgta 1500
 tatgttagca tattgtatac attgcatatt aacataagct agaatcattg acataaattt 1560

atataaaca aaggtataaa ataaaaaaaa aaaaaaaaaa aaaaaaaaaa 1610

<210> 9
 <211> 2486
 <212> DNA
 <213> Homo sapiens
 <220><221> mRNA
 <222> (1)..(2486)
 <223> hypothetical protein FLJ21511(FLJ21511), mRNA.
 <400> 9

acactgcctc ggttcggcaa gtgggtcagt tggctggggc tcaacttgca acgggacgcg 60
 ggaacgaggg gcgcggacgc aggcccggga ggacgcggcg gcgggaacct gggggcgcag 120
 ggctagggca gcgggcccga cccgcacggc tttcctggaa agcgtgccc ctgcgcgagg 180
 cgatgcctc gcigtggaga gaaatcctct tggagtcgct gctgggatgt gtttcttggg 240

ctctctacca tgacctggga ccgatgatct attactttcc tttgcaaaca ctagaactca 300
 ctgggcttga aggttttagt atagcatttc tttctccaat attcctaaca attactcctt 360
 tctggaaatt ggtaacaag aagtggatgc taaccctgct gaggataatc actattggca 420
 gcatagcctc cttccaggct ccaaatgcca aacttcgact gatggttctt gcgcttgggg 480
 tgtcttctc actgatagtga caagctgtga cttggtggtc aggaagtcat ttgcaaaggt 540
 acctcagaat ttggggattc attttaggac agattgttct tgttgttcta cgcatatggt 600
 atacttcaact aaaccaatc tggagttatc agatgtccaa caaagtgata ctgacattaa 660

gtgcatagc cacacttgat cgtattggca cagatggatga ctgcagtaaa cctgaagaaa 720
 agaagactgg tgaggtagcc acggggatgg cctctagacc caactggctg ctggcagggg 780
 ctgcttttgg tagccttgtg ttctcacc accgggtttt tggagaagtc tetcttgttt 840
 ccagatgggc agtgagtggg catccacatc cagggccaga tcctaacca tttggaggtg 900
 cagtactgct gtgcttggca agtggattga tgcttccatc ttgtttgtgg tttcgtggta 960
 ctggtttgat ctggtgggtt acaggaacag cttcagctgc ggggctcctt tacctgcaca 1020
 catgggcagc tgctgtgtct ggctgtgtct tcgccatctt tactgcatcc atgtggcccc 1080

aaacacttgg acaccttatt aactcaggga caaacctgg gaaaacctg accattgcca 1140
 tgatatttta tcttctagaa atatttttct gtgcctgggtg cacagctttt aagtttgtcc 1200

caggaggtgt ctacgctaga gaaagatcag atgtgctttt ggggacaatg atgttaatta 1260
 tcgggctgaa tatgctattt ggtcctaaga aaaaccttga cttgcttctt caaacaaaaa 1320
 acagtcttaa agtgcttttc agaaagagtg aaaaatacat gaaacttttt ctgtggctgc 1380
 ttgttgggtg gggattgttg ggattaggac tacggcataa agcctatgag agaaaactgg 1440
 gcaaagtggc accaaccaaa gaggtctctg ctgccatctg gcctttcagg tttggatatg 1500

acaatgaagg gtggtctagt ctagaagat cagctcacct gctcaatgaa acagggtcag 1560
 atttcataac aattttggag agtgatgctt ctaagcccta tatggggaac aatgacttaa 1620
 ccatgtggct aggggaaaag ttgggtttct atacagactt tggccaagc acaaggtatc 1680
 acacttgggg gattatggct ttgtcaagat acccaattgt gaaatctgag catcaccttc 1740
 ttccgtcacc agagggcgag atcgcaccag ccatcacatt gaccgttaac atttcgggca 1800
 agctggtgga ttttgcctg acacactttg ggaaccacga agatgacctc gacaggaaac 1860
 tgcaggctat tgctgtttca aaactactga aaagtagctc taatcaagtg atatttctgg 1920

gatataatcac ttcagcacct ggctccagag attatctaca gctcactgaa catggcaatg 1980
 tgaaggatat cgacagcact gatcatgaca gatggtgtga atacattatg tatcgagggc 2040
 tgatcagggtt gggttatgca agaatctccc atgctgaact gagtattca gaaattcaga 2100
 tggcaaaatt taggatccct gatgaccca ctaattatag agacaaccag aaagtgtca 2160
 tagaccacag agaagtttct gagaaaatic attttaatcc cagatttga tectacaaag 2220
 aaggacacaa ttatgaaaac aaccatcatt tccatgtaa tactcccaa tactttttat 2280
 gaaacattta aaacaagaag ttattggctg ggaaaatcta agaaaaaag tatgtaagat 2340

aaaaagaaga gattaatgaa agtgggaaaa tacacatgaa gaacctcaac ttaaaaaaca 2400
 catggtatct atgcagtggg aaattacctc cttttgtaa ctatgttgct taataaaaac 2460
 atttctctaa aaaaaaaaaa aaaaaa 2486

<210> 10
 <211> 1900
 <212> DNA
 <213> Homo sapiens
 <220><221> mRNA
 <222> (1)..(1900)
 <223> chromosome 9 open reading frame 19(C9orf19), mRNA.
 <400> 10

agcccgggg agcgaggagc gcgcgagcc ggccatgggc aagtcagctt ccaaacagtt 60

tcataatgag gtcttgaagg cccacaatga gtaccggcag aagcacggcg tccccact 120

 gaagctctgc aagaacctca accgggaggc tcaacagtat tctgaggccc tggccagcac 180
 gaggatcctc aagcacagcc cggagtccag ccgtggccag tgtggggaga accttgcattg 240
 ggcatcctat gatcagacag gaaaggaggt ggctgataga tggtagactg aatcaagaa 300
 ctataacttc cagcagcctg gcttcacctc ggggactgga cacttcacgg ccatggtatg 360
 gaagaacacc aagaagatgg gcgtggggaa ggcgtccgca agtgacgggt cctcctttgt 420
 ggtggccaga tacttcccag cggggaatgt tgtcaatgag ggcttcttcg aagaaaactg 480
 cctgccgccg aagaagtaac ttgttaaatg taatgggaag gtggcagact taagaactg 540

 gatataaagt gcctagaacc accacaacct ggctgtgcgt ctgtccctgt ggggtgatgt 600
 gcttgtgtgt gtgatgcatg tgagcgtctc tggcacacac acttggacat acagtctctg 660
 gtgcgctcat tcttattaca ggagtgagca aaggaagcat ttaccccgat ggttacctag 720
 accacgatta ttggattgg ggggaggggg gatccgtttt ttttttttaa tttttgtta 780
 tttctaagca aacctctttt gtacttttct tacttctaata atccatccct ggactttttg 840
 tattccaat gtttgtgatg ctgagaagtg aagttcattt tatgtgatct tcatgcgtcg 900
 taatctactt ttggtagata attaagatta ttaaacctc atttaaatgt gacataaaat 960

 acagctttaa gcacataaat ataaagcagc ttccatcagg aacatggagc aggcagggac 1020
 tccattttac agaattactg agatttctca gttgtaaac atgatgtcat cctgcatgcc 1080
 tcctggaatt ctccaatggg gtcgcaaac acaaatgga gaaaaaagt tttacttct 1140
 tgcatcttc tacctttaaa tagcaaaagta cactaccac caccacctct tgccccctc 1200
 cctcttttct taaacttctg gcatttcaga gctcagcagg ctaccctgg tttctggaga 1260
 gttggcctag gcctgaagct cccctcccc cacctctgct aggcagccca ggctgtgtct 1320
 gggagacagc cctcacct gcctgggctc ttggccaagc ggcttggat ggatgaagtc 1380

 agagaggtgg ggtgagggtg agcttactca gggccccag aggaagccct cagcctctgc 1440
 cctccccca cacagggcgg gagcccaggc ctgttctctg cagctgtggc tgcagctgtg 1500
 ctctgctcc ctctggaat gtgcgacaag ccaaatgtt ccaggaggc ggccggggca 1560
 gggggcttag aagtgctaat atggttctgt gttttgctg aaacgatacc aggttccct 1620
 gaatagcaac tttaacaagt ccatgtggga gggaccaacc cagatgccct gctgagtgtc 1680
 cctgaaacca tggcagctcc atctgtcaag atggcagggg ccggagttag gggctgctg 1740
 gcttaacagc aggcactcgg gcagccagc cctcaaagca gctcctgaag gctctgtttg 1800

cactgtcacc agtctcaagc tatgcctcta atttcaccag ggatattgac taagaagaca 1860
ataaaatctt tttctttgtg taaaaaaaa aaaaaaaaaa 1900

<210> 11
<211> 2544
<212> DNA
<213> Homo sapiens
<220><221> mRNA
<222> (1)..(2544)
<223> thymus expressed gene 3-like(MGC15476), mRNA.
<400> 11

ctcactcaga cccatgaggc cctgcctggt ctcgtctggg acctgggaca gcagctggga 60
gacctgagcc tggagtctgg gggcctggaa caggagagcg ggcgtagctc gggcttctat 120
gaagatccca gctctacagg aggtccagat tcaccacct caaccttctg tggggacagt 180

ggcttctctg gatccagctc ctatggtcgc ctgggtccct ctgagccccg gggcatctat 240
gccagtgaga ggccaagtc cctaggagac gccagtccca gcgctccgga ggtggigggc 300
gcgcgggcag cggtgccgcg gtccttctca gcgccctacc cgacggcagg tgggtcgccg 360
gcccggagge ctgctctctg gcggagcggc gggccccgcg cgggcccctt ctgacgcca 420
gccccctgca cgccgtggcg atgcgcagcc cgcggccctg cggcgcacct cccaccgact 480
cgcccgacgc gggggcgca gggcggcccc tggacggcta catctcgcg ctctctcgca 540
ggcgccgccc ccggggggcg ggccagcccc ggaccagtcc tgggggcgcg gacggcggcc 600

cgcggcgcca gaacagcgtg cgccagcggc cgcccgacgc gtctccgtcc cccggcagcg 660
cgcgacccgc gcgggagccc tcgttggagc gcgtcggggg ccaccccacc agccctgccg 720
ccttgagccg cgcttggcg tcgtcgtggg agtcggaggc ggcacccgag cccgctgcgc 780
cgcccgcgc cccctcacc cccgacagcc cggctgaggg ccgcttgggt aaggcgcagt 840
acatcccggg cgcgcaggcg gccacccgag gcctccctgg ccgcgcgcc cgccgcaaac 900
cgcccact gacccgccc cgcagcgtgg agcagtcacc accccgggag cgtcccggg 960
ccgccggccc ccgtggacgc atggccgagg ctccgggccc ccgcggctcg cccagggccc 1020

gcaaggcctc gcgtcccag tctgagacca gctgtctggg ccgcgcctcc gcggtccctt 1080
cggggcccc taagtacccc acggcggagc gggaagagcc tcggcctcca cggccagcc 1140
gcggcccagc gcccacgtg gcggcccagg ccgcagggtc ctgccgtcgc tggcgtcca 1200
ctgcggagat cgacgtgcc gatggcgccc gcgtgcggcc ccgagcccct gcggcgcgtg 1260
ttcccggccc cggcccgtcc ccgtcagctc cccagcgtcg tctgctttac ggctgcgccc 1320

gcagcgactc cgagtgctcg gctgggcgcc tggggccctt gggacgccgg gggcctgcgg 1380
gaggcgtcgg cgggggttac ggggagagcg aatcgagcgc cagcgaggga gaatgcctg 1440

ccttcagctc tgcctccagc gactcagacg gcagcgggtg cctcgtgtgg cgcgacgacg 1500
tgggtggcggc caccgcggcc tctgggggtg gacgaggtgc aggggcgccc gcaggccccg 1560
ccaaagtctt cgtgaaaate aaagcttccc acgcgctcaa gaaaaagata ctgcgtttcc 1620
gttcgggttc tctcaaggtc atgactacag tgtgagtttg gggatttctt tgggtcccc 1680
cttcattggc tctgcacctc cactctcca accactgacc cttccacatc taccttccaa 1740
agaccatcgt tttctctgct tccaaagacc cccctcactc tcccactcc taacagtctt 1800
ggttgaaaag gctccccac caccaccgag aggaatgggg aggagccctg tttgaccag 1860

ttcagcttct agcttgaag cccttgggca agacagtcc ccttctctgg gcgtcacttt 1920
cctcatctgt acagtaagtg tccatgtatg caaaaggggt aattcggttt gaatttcccc 1980
gttttagttt agaagcctag tctgtttgtt cccttcacc gctctccctc tcattctga 2040
tgagccctct cattctctct ttcttgccc agctatggcc ccctctcatt cacaaagtgc 2100
ccctccatg tccttgacc cttaagatat cccttggca ccttggtcag agactctgtg 2160
tctgactcag gtgtccctg cagagtgcc tgggaaggga aggagcactg atttgggggt 2220
tttgagggtc aagtaggggt tggtaacacc tggaaagaag gactctttca cttcgatccc 2280

tggacaatta tggaggattc ggaggtagaa gaggggaagg aagatggttt ctatctcatg 2340
acccccactc cctgtgagag ggaatggggg aagcctgatg accctcagct gtccaatct 2400
agtatTTTTT ttctTTTTT aaattactgt atttattatg acgatggtga ctccccagtg 2460
caaagggggg ccagattctg tgtgtttctc taacctcttt gtaataaat gcacagtgta 2520
acataaaaaa aaaaaaaaaa aaaa 2544

<210> 12
<211> 4471
<212> DNA
<213> Homo sapiens
<220><221> mRNA
<222> (1)..(4471)
<223>
mRNA; cDNA DKFZp686J0156(from clone DKFZp686J0156)
<400> 12
caattacata tttatttttc catacctgat ttttttcaag tctgtaataa aaaaagtata 60
agttgagatt aacataggtt atttttcatg aagtatagca aacgatctag aatgtgatag 120

gagtgtggtt tccatttctt tttttttttt tttaacagac ttttgtgctg tagtaagaat 180
 gtcttagaag ttgtgtcttc gagttcctca aacaccggac agggctcccc cagagctgcc 240
 tcagcctttg atttggtcca tattcagaca cttcctgtct tgcatacact ctgaggatgg 300
 gccgcttcta gctgccgccc gtctccctct tgccttcttg tctctaaatc agacactccc 360

aagcactcca tctgctatac taggagagtg ggagagccaa agcatttttg agtccatgta 420
 aagtgcctgg gaaaagagat gtaagagcag tgggagttta tataattagc gaattctttt 480
 ctcagaattc agcggaaatg tatttctgtt gttattattht acaaaattgt gccctcagag 540
 gagctgtctt ttggaaaata agttttcatt catgcagtta gttagacatg cacgatgtac 600
 ttgaatcaat aaactgtgga ccagaaaagg gctgcttaaa tgaattcagc tgttctaatt 660
 ctctttcagag gctgttgatg acctcctgga ttcatacccc catgatctgt cttgctcttg 720
 cctctgttct ttgataaagc acatcagact atgtgacacc catcttttca agaacgacct 780

aaaagcaatg acaacagcca ggagagaatt accactggcc tgggcacagt gacagcagaa 840
 ggcccacctg agcttggaga gttggcgaga taaggcaaaa tcaaggtgca acagtgggcc 900
 atgttccaag ggtgaatgga tctgagagtc cgtgatgtt ccttactgat cagaacaggg 960
 accttctaac gaaacagcca gagttccttg aatcaggcag aagtcttita gagacaggaa 1020
 gccagagcta ttgataatat ccgaacgaca gaggaaggaa tttcttatta gaaattctaa 1080
 aataaaacaa gtggggattt ggttggatac ccttagaagc tacatttttc tctcatccct 1140
 aagatagata tctcaagtg atatctgaca cctcttgct ttgacacggc tctgcttctc 1200

aagtgacat ttgttttaag tctgttctc agaattcttt gtggggctg gacctctgaa 1260
 aaagattctc actgcagagg ctttaaatct gtgttttttt ccaactcctc cgtcctggga 1320
 tgtattcggg aggacatttc atttatatca ggatattggc tgattcagct ctctttcacc 1380
 gttattatte aggccctcct ccaccacctt gttctttatt gcctaagat tgcctcccc 1440
 aattcttcat cglaatagct ctaaatttag tttgacccaa attaattttt gttccaacag 1500
 atatatthta agacagagaa gcctctctca ggagccagat gatctgaaga tgatgctaat 1560
 ccccaaatc gaagttctc cagctaactc tctctgagc tccctcttct tagctttggg 1620

attggggatg gtctcatcct tgccaggaat cctttgttg ctcccagtgt ctctgatgag 1680
 cttctectec ctcccagaga caatgcaggg gtggttttca gcagaatgtg ccacctgac 1740
 ttctgattgt tctgagacc agcccagagg ggtggagaag ggcacaaac cctattccaa 1800
 taaactgcct gcagagaacc acagcctgct aatagtcgtt gaggtctcag ccccctaagg 1860
 tctccaagct catttgacca gcgcagctgc gacttcggct gtctgcctct ggcaggtgcc 1920
 aatatcagca tcttgtccca ctgccatggg gagttaattt ttgctccagg gcagctgcca 1980

agcccagcac cggcatcaat agaaggtttt ctctgatttt cttctcttcc tctcaggcta 2040

ggctttctta ccaatcccca atttcttctt gtagggcgcc aggagactgg ccagagatct 2100

ttagactctt acttagtttg aacatctctt gctgctcaca cccaggcttc ttcagcagaa 2160

agcatgacct atagtaggta ctcaaaaact ggttgaatga aaagccttga catttaaggg 2220

ctttattctg tcatacggtg gcaagaactg gcctattggc atcgtattcc tggcattgag 2280

ctgtattgac cctgaatgtg tcctctgag gtttttgagg gtgacctagg agaaagggtc 2340

tgaaaaagca ctgggatcca agtgacttta aatctgcccg tggcatttcc tgctgagagg 2400

cgagaaacta gcatctctcc caacaggctc acatccgcca ggaaggacag agtccaaaac 2460

gggggtagac ttigtgggaa attcggcagt tagtccgctg agattttaat ctgcctcaag 2520

cacaggatgc tgacaagaaa ggtagaccac tccccgatc ccagacaaaac cccaggggtg 2580

acaacttctc catcagtatc ctgagaagca acatcaaatg cttaccaagt cagcacgttt 2640

actgagtgag tactctgacg gaataaacca gcatttggag ctgaacttga gaacacaggg 2700

ccctctgaca gcagcttagg gaaacacca tccttccag attggcagaa tccttcatt 2760

tcatgtcctt cggagacgct gtattctctt tacaatcat tcatgaaact gagactcagt 2820

ctcccagggg taagtcctgg gccacgtgca aacaatggca aatccagctc ttccatgata 2880

gagaaggcaa acacaccac cttcatctg agcctaaaag gccacctctg agcaactggg 2940

cagccactcc tctgggctc agagggcat gagcttggcc aggtaggcac agcggcgggg 3000

aagtcacagc tgtcaggtac cggccatggt gcaggtggga ataggagatg ccagagctgc 3060

tttagctgag gaaagcaaac agtcagcagt gctcaaagga gcaaaacttc gaatgtgac 3120

attgaccctc gacacctgca agcataaac agatcctaag actagagtga agtaggaaga 3180

agaattagaa aatccagttg atgtctgag tataggaac cagggccgtt gaaaatcagt 3240

aaaggttgat tacctggggc gagaccgggt gactgtggca gtgcaggtga aggtaccctg 3300

gaccttctca gttcgtggc acataaggct ccgcaataa agcgtggttc tctctgtcac 3360

acacacacac acacacacac acacacacac acaatgattg gagggctata tgatccagca 3420

ttagcttctt ggigtgcca gcatgcttga tcgggaattt tttttatta ttattatctt 3480

ttagctgtag ctgaaggcat ttctcggatg tggagaggag aatggaaatc gcagaaccaa 3540

atcagtttgc cctgccatat ttggctgtgg tctgtcattg ggcatctctg atgtgctttt 3600

ctggattcag gaagagctga ttgtcctccg agggttgaaa aaaaaaaaaac agtttcagaa 3660

acctgaatcc agggccttat agttctctc attatcgatc ttcttctccc ttccctcgcc 3720

caagggagtg ggggaaacac ttttactgc agagtttgc ttaaagttt cccatcttgc 3780
 gtgcattatc ccttgatatt aaaattatt tctcagtta atccaactcc tgctgagaag 3840
 ctgtgtgaga tttaggctgt ggggttttt tctgttactc ttttggatgg tgttgcat 3900
 ttcactctta acccgagggt gtgtttcagc ttatgttcgt tctgtttcat gcaggttat 3960
 agcacggtag agtagaaggc ggcttctgat ttttaggta tttttagaat tcattcctga 4020
 gtgtggggtt cagacacca gtctcctcgg aacaggggtg aggggtcgac tgagcttgt 4080
 tgagaagcct ccagttaggc ttcgggcggg tctccatgtt atattgtgtg tttactgagc 4140

ttccactgg tagaagatga cacattgtc catcgtctg tgtatctgct ttccagagga 4200
 caccggagca ttctcctggg gtcactcca catggctgcc tcacatagct gttttgcaac 4260
 agcctttatt gccaacacc ttggaagaac atctcctgta gaagtacaag ttttgcaata 4320
 tggactttgg aagtggtttt gtgatgttcg aattttctgt tagcctattg ctcaagcact 4380
 acagaatgta catagtcct ttggccatcc tctcaccca gaatataccta ataaacattg 4440
 ttttctttt taaaaaaaa aaaaaaaaa a 4471

- <210> 13
- <211> 20
- <212> DNA

<213> Artificial Sequence
 <220><223> Nucleic acid used as forward primer for amplifying GAPDH

<400> 13
 tcatgaccac agtccatgcc 20

- <210> 14
- <211> 20
- <212> DNA

<213> Artificial Sequence
 <220><223> Nucleic acid used as reverse primer for amplifying GAPDH

<400> 14
 tccaccacce tgttctgta 20

- <210> 15
- <211> 20
- <212> DNA

<213> Artificial Sequence

<220

><223> Nucleic acid used as forward primer for amplifying LMTK3

<400> 15
ggaacagcga gcagatcaaa 20

<210> 16
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> Nucleic acid used as reverse primer for amplifying LMTK3

<400> 16
aggagaccat cttgcgcttc 20

<210> 17
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> Nucleic acid used as forward primer for amplifying LOC644774

<400> 17
tcaatctgcc acagaaggct 20

<210> 18
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> Nucleic acid used as reverse primer for amplifying LOC644774

<400> 18
ccaggatggc caggaagggt 20

<210> 19
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> Nucleic acid used as forward primer for amplifying Hs.389988

<400> 19
gcctgggacc tccagaataa 20

<210> 20
<211> 21
<212> DNA

<213> Artificial Sequence

<220><223> Nucleic acid used as reverse primer for amplifying Hs.389988

<400> 20

tggcaaatga aaaattcagc a 21

<210> 21

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Nucleic acid used as forward primer for amplifying WDR72

<400> 21

tcccatgtac ctgaggctga 20

<210> 22

<211> 20

<

212> DNA

<213> Artificial Sequence

<220><223> Nucleic acid used as reverse primer for amplifying WDR72

<400> 22

ttggaagttt gccgagtttg 20

<210> 23

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Nucleic acid used as forward primer for amplifying LOC440157

<400> 23

gagaagaagc aggtcccagg 20

<210> 24

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Nucleic acid used as reverse primer for amplifying LOC440157

<400> 24

agttccaggg ctctgtggtt 20

<210> 25

<211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Nucleic acid used as forward primer for amplifying LOC643911
 <400> 25
 aattcatccc aaggctggtc 20
 <210> 26
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> Nucleic acid used as reverse primer for amplifying LOC643911

 <400> 26
 cccaaaacat atttaaacca ctcg 24
 <210> 27
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Nucleic acid used as forward primer for amplifying C13ORF23
 <400> 27
 agcaccagtc tcagcagcac 20
 <210> 28
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Nucleic acid used as reverse primer for amplifying C13ORF23
 <400> 28
 gaactgaatg ctgaggcgaa 20

 <210> 29
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Nucleic acid used as forward primer for amplifying LOC644424
 <400> 29

attcagtcct cactgcggt 20

<210> 30

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Nucleic acid used as reverse primer for amplifying LOC644424

<400> 30

ttccgctct tgatgagaag 20

<210> 31

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Nucleic acid used as forward primer for amplifying FLJ21511

<400> 31

gtggctaggg gaaaagttgg 20

<210> 32

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Nucleic acid used as reverse primer for amplifying FLJ21511

<400> 32

aatctctgga gccagtgct 20

<210> 33

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Nucleic acid used as forward primer for amplifying C9ORF19

<400> 33

accgggagc tcaacagtat 20

<210> 34

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Nucleic acid used as reverse primer for amplifying C9ORF19
 <400> 34
 tatctggcca ccacaaagga 20
 <210> 35
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220><223> Nucleic acid used as forward primer for amplifying MGC15476
 <400> 35
 cactgcggag atcgacgct 19
 <210> 36
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Nucleic acid used as reverse primer for amplifying MGC15476
 <400> 36
 ttttcacgaa gactttggcg 20
 <210> 37
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Nucleic acid used as forward primer for amplifying LOC143381
 <400> 37
 ggcggcttct gatttttagg 20
 <210> 38
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Nucleic acid used as reverse primer for amplifying LOC143381
 <400> 38
 tccaagggtg ttggcaataa 20