



**(19) 대한민국특허청(KR)**  
**(12) 등록특허공보(B1)**

(45) 공고일자 2014년07월04일  
 (11) 등록번호 10-1415736  
 (24) 등록일자 2014년06월30일

(51) 국제특허분류(Int. Cl.)  
*C12Q 1/68* (2006.01) *G01N 33/574* (2006.01)  
 (21) 출원번호 10-2012-0055348  
 (22) 출원일자 2012년05월24일  
 심사청구일자 2012년05월24일  
 (65) 공개번호 10-2012-0132665  
 (43) 공개일자 2012년12월07일  
 (30) 우선권주장  
 1020110049541 2011년05월25일 대한민국(KR)  
 (56) 선행기술조사문헌  
 WO2010081001 A2

(73) 특허권자  
**한국생명공학연구원**  
 대전광역시 유성구 과학로 125 (어은동)  
 (72) 발명자  
**박홍석**  
 대전광역시 유성구 과학로 125  
**김룡남**  
 대전광역시 유성구 과학로 125  
 (뒷면에 계속)  
 (74) 대리인  
**이원희**

전체 청구항 수 : 총 10 항

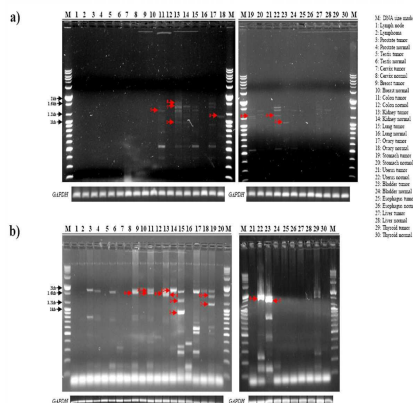
심사관 : 이효진

**(54) 발명의 명칭 신규한 인간 융합 유전자 또는 이의 융합 유전자 전사 변이체, 및 이의 용도**

**(57) 요약**

본 발명은 신규한 인간 융합 유전자(Conjoined gene, CGs) 또는 이의 융합 유전자 전사 변이체(Conjoined gene transcript variants, CGTVs), 및 이의 용도에 관한 것으로, 구체적으로, 신규한 5개의 인간 융합유전자 (*ZNF343-SNRPB-Hu*, *COX17-POPDC2-Hu*, *RBM7-REXO2-Hu*, *PRKRIP1-ORAI2-Hu*, *ATP5F1-Clorf162-Hu*)와 이의 57개의 인간 융합 유전자 전사 변이체에 대한 시퀀싱(염기서열분석) 및 생물정보학적 해석들을 수행하여 융합유전자의 형성 기전을 확인하였으며, 인간 종양 조직에서 특이적으로 발현되는 융합 유전자 전사 변이체를 동정함으로써, 융합 유전자 전사 변이체와 암 발생과의 연관성을 확인한 바, 상기 융합 유전자 전사 변이체는 암의 발병 기전 및 암 진단에 대한 표적으로서 유용하게 사용될 수 있다.

**대표도 - 도6**



(72) 발명자

**강아람**

대전광역시 유성구 과학로 125

**김대수**

대전광역시 유성구 과학로 125

**김동욱**

대전광역시 유성구 과학로 125

**김애리**

대전광역시 유성구 과학로 125

**남성혁**

대전광역시 유성구 과학로 125

**최상행**

대전광역시 유성구 과학로 125

이 발명을 지원한 국가연구개발사업

과제고유번호 20100020659

부처명 교육과학기술부

연구사업명 미래기반기술개발사업

연구과제명 유전체 정보생산 및 활용화 시스템개발

기여율 1/1

주관기관 한국생명공학연구원

연구기간 2010.08.01 ~ 2011.07.31

---

**특허청구의 범위**

**청구항 1**

하기 모든 유전자 각각의 핵산 서열 올리고뉴클레오티드 전부 또는 그의 상보 가닥 분자가 집적된, 암 진단용 DNA 마이크로어레이:

서열번호 1로 기재되는 ZNF343-SNRPB-Hu-P1 유전자, 서열번호 2로 기재되는 ZNF343-SNRPB-Hu-P2 유전자, 서열번호 3으로 기재되는 ZNF343-SNRPB-Hu-P3 유전자, 서열번호 4로 기재되는 ZNF343-SNRPB-Hu-P4 유전자, 서열번호 5로 기재되는 ZNF343-SNRPB-Hu-P5 유전자, 서열번호 6으로 기재되는 ZNF343-SNRPB-Hu-P6 유전자, 서열번호 7로 기재되는 ZNF343-SNRPB-Hu-P7 유전자, 서열번호 8로 기재되는 ZNF343-SNRPB-Hu-P8 유전자, 서열번호 9로 기재되는 ZNF343-SNRPB-Hu-P9 유전자, 서열번호 10으로 기재되는 ZNF343-SNRPB-Hu-P10 유전자, 서열번호 11로 기재되는 ZNF343-SNRPB-Hu-P11 유전자, 서열번호 12로 기재되는 ZNF343-SNRPB-Hu-P12 유전자, 서열번호 13으로 기재되는 ZNF343-SNRPB-Hu-F1 유전자, 서열번호 14로 기재되는 ZNF343-SNRPB-Hu-F2 유전자, 서열번호 15로 기재되는 ZNF343-SNRPB-Hu-F3 유전자, 서열번호 16으로 기재되는 ZNF343-SNRPB-Hu-F4 유전자, 서열번호 17로 기재되는 ZNF343-SNRPB-Hu-F5 유전자, 서열번호 18로 기재되는 ZNF343-SNRPB-Hu-F6 유전자, 서열번호 19로 기재되는 ZNF343-SNRPB-Hu-F7 유전자, 서열번호 20으로 기재되는 ZNF343-SNRPB-Hu-F8 유전자, 서열번호 21로 기재되는 COX17-POPDC2-Hu-P1 유전자, 서열번호 22로 기재되는 COX17-POPDC2-Hu-P2 유전자, 서열번호 23으로 기재되는 COX17-POPDC2-Hu-P3 유전자, 서열번호 24로 기재되는 COX17-POPDC2-Hu-P4 유전자, 서열번호 25로 기재되는 COX17-POPDC2-Hu-P5 유전자, 서열번호 26으로 기재되는 COX17-POPDC2-Hu-P6 유전자, 서열번호 27로 기재되는 COX17-POPDC2-Hu-P7 유전자, 서열번호 28로 기재되는 COX17-POPDC2-Hu-P8 유전자, 서열번호 29로 기재되는 COX17-POPDC2-Hu-P9 유전자, 서열번호 30으로 기재되는 COX17-POPDC2-Hu-P10 유전자, 서열번호 31로 기재되는 COX17-POPDC2-Hu-P11 유전자, 서열번호 32로 기재되는 COX17-POPDC2-Hu-F1 유전자, 서열번호 33으로 기재되는 COX17-POPDC2-Hu-F2 유전자, 서열번호 34로 기재되는 COX17-POPDC2-Hu-F3 유전자, 서열번호 35로 기재되는 COX17-POPDC2-Hu-F4 유전자, 서열번호 36으로 기재되는 COX17-POPDC2-Hu-F5 유전자, 서열번호 37로 기재되는 COX17-POPDC2-Hu-F6 유전자, 서열번호 38로 기재되는 COX17-POPDC2-Hu-F7 유전자, 서열번호 39로 기재되는 COX17-POPDC2-Hu-F8 유전자, 서열번호 40으로 기재되는 COX17-POPDC2-Hu-F9 유전자, 서열번호 41로 기재되는 COX17-POPDC2-Hu-F10 유전자, 서열번호 42로 기재되는 COX17-POPDC2-Hu-F11 유전자, 서열번호 43으로 기재되는 COX17-POPDC2-Hu-F12 유전자, 서열번호 44로 기재되는 RBM7-REX02-Hu-P1 유전자, 서열번호 45로 기재되는 RBM7-REX02-Hu-P2 유전자, 서열번호 46으로 기재되는 RBM7-REX02-Hu-P3 유전자, 서열번호 47로 기재되는 RBM7-REX02-Hu-P4 유전자, 서열번호 48로 기재되는 PRKRIP1-ORAI2-Hu-P1 유전자, 서열번호 49로 기재되는 PRKRIP1-ORAI2-Hu-P2 유전자, 서열번호 50으로 기재되는 PRKRIP1-ORAI2-Hu-P3 유전자, 서열번호 51로 기재되는 PRKRIP1-ORAI2-Hu-P4 유전자, 서열번호 52로 기재되는 PRKRIP1-ORAI2-Hu-P5 유전자, 서열번호 53으로 기재되는 PRKRIP1-ORAI2-Hu-P6 유전자, 서열번호 54로 기재되는 ATP5F1-C1orf162-Hu-P1 유전자, 서열번호 55로 기재되는 ATP5F1-C1orf162-Hu-P2 유전자, 서열번호 56으로 기재되는 ATP5F1-C1orf162-Hu-P3 유전자 및 서열번호 57로 기재되는 ATP5F1-C1orf162-Hu-P4 유전자.

**청구항 2**

제 1항에 있어서, 상기 암은 림프종, 전립선암, 자궁경부암, 유방암, 대장암, 신장암, 폐암, 난소암, 위암, 자궁암, 방광암, 식도암, 간암 및 갑상선암으로 이루어진 군으로부터 선택되는 어느 하나인 것을 특징으로 하는 DNA 마이크로어레이.

**청구항 3**

제 1항에 있어서, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-P11 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-P5, COX17-POPDC2-Hu-P8 유전자, COX17-POPDC2-Hu-P9 유전자, COX17-POPDC2-Hu-P10 유전자, RBM7-REX02-Hu-P1 유전자, 및 PRKRIP1-ORAI2-Hu-P1 유전자는 림프종에서 발현하고, COX17-POPDC2-Hu-P9 유전자, PRKRIP1-ORAI2-Hu-P1 유전자, KRIP1-ORAI2-Hu-P3 유전자, 및 ATP5F1-C1orf162-

Hu-P3 유전자는 전립선암에서 발현하고,

ZNF343-SNRPB-Hu-P1 유전자, ZNF343-SNRPB-Hu-P3 유전자, ZNF343-SNRPB-Hu-P4 유전자, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-P11 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-P10 유전자, PRKRIP1-ORAI2-Hu-P3 유전자, 및 ATP5F1-C1orf162-Hu-P4 유전자는 자궁경부암에서 발현하고,

ZNF343-SNRPB-Hu-P8 유전자, COX17-POPDC2-Hu-P7 유전자, COX17-POPDC2-Hu-P11 유전자, COX17-POPDC2-Hu-F9 유전자, COX17-POPDC2-Hu-F11 유전자, COX17-POPDC2-Hu-F12 유전자, RBM7-REX02-Hu-P4 유전자, PRKRIP1-ORAI2-Hu-P6 유전자 및 ATP5F1-C1orf162-Hu-P4 유전자는 유방암에서 발현하고,

ZNF343-SNRPB-Hu-P5 유전자, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-F5 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-F7 유전자, COX17-POPDC2-Hu-F11 유전자 및 COX17-POPDC2-Hu-F12 유전자는 대장암에서 발현하고,

ZNF343-SNRPB-Hu-P1 유전자, ZNF343-SNRPB-Hu-P2 유전자, ZNF343-SNRPB-Hu-P3 유전자, ZNF343-SNRPB-Hu-P4 유전자, ZNF343-SNRPB-Hu-P11 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-P4 유전자, COX17-POPDC2-Hu-P6 유전자, COX17-POPDC2-Hu-P8 유전자, COX17-POPDC2-Hu-F3 유전자, COX17-POPDC2-Hu-F4 유전자, COX17-POPDC2-Hu-F6 유전자, PRKRIP1-ORAI2-Hu-P5 유전자 및 ATP5F1-C1orf162-Hu-P3 유전자는 신장암에서 발현하고,

ZNF343-SNRPB-Hu-P1 유전자, ZNF343-SNRPB-Hu-F5 유전자, ZNF343-SNRPB-Hu-F7 유전자, COX17-POPDC2-Hu-F1 유전자, COX17-POPDC2-Hu-F3 유전자 및 PRKRIP1-ORAI2-Hu-P6 유전자, 6개는 폐암에서 발현하고,

ZNF343-SNRPB-Hu-P4 유전자, ZNF343-SNRPB-Hu-P11 유전자, ZNF343-SNRPB-Hu-P12 유전자, ZNF343-SNRPB-Hu-F7 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-F10 유전자, COX17-POPDC2-Hu-F12 유전자, PRKRIP1-ORAI2-Hu-P4 유전자, PRKRIP1-ORAI2-Hu-P6 유전자 및 ATP5F1-C1orf162-Hu-P1 유전자는 난소암에서 발현하고,

ZNF343-SNRPB-Hu-P11 유전자, ZNF343-SNRPB-Hu-F6 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-F2 유전자 및 COX17-POPDC2-Hu-F10 유전자는 위암에서 발현하고,

PRKRIP1-ORAI2-Hu-P2 유전자 및 PRKRIP1-ORAI2-Hu-P5 유전자는 자궁암에서 발현하고,

ZNF343-SNRPB-Hu-P7 유전자, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-F2 유전자, COX17-POPDC2-Hu-P1 유전자, COX17-POPDC2-Hu-F1 유전자, COX17-POPDC2-Hu-F5 유전자 및 RBM7-REX02-Hu-P3 유전자는 방광암에서 발현하고,

ZNF343-SNRPB-Hu-P8 유전자는 식도암에서 발현하고,

ZNF343-SNRPB-Hu-F7 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-P10 유전자, ATP5F1-C1orf162-Hu-P2 유전자, ATP5F1-C1orf162-Hu-P3 유전자 및 ATP5F1-C1orf162-Hu-P4 유전자는 간암에서 발현하고,

ZNF343-SNRPB-Hu-P8 유전자는 갑상선암에서 발현하는 것을 특징으로 하는 DNA 마이크로어레이.

#### 청구항 4

제 1항의 DNA 마이크로어레이를 포함하는 암 진단용 키트.

#### 청구항 5

하기의 단계를 포함하는 개체 내 암 진단의 정보를 제공하기 위한 유전자 발현 여부 확인 방법:

1) 분리된 시료로부터 하기 모든 유전자의 발현 여부를 확인하는 단계:

서열번호 1로 기재되는 ZNF343-SNRPB-Hu-P1 유전자, 서열번호 2로 기재되는 ZNF343-SNRPB-Hu-P2 유전자, 서열번호 3으로 기재되는 ZNF343-SNRPB-Hu-P3 유전자, 서열번호 4로 기재되는 ZNF343-SNRPB-Hu-P4 유전자, 서열번호 5로 기재되는 ZNF343-SNRPB-Hu-P5 유전자, 서열번호 6으로 기재되는 ZNF343-SNRPB-Hu-P6 유전자, 서열번호 7로 기재되는 ZNF343-SNRPB-Hu-P7 유전자, 서열번호 8로 기재되는 ZNF343-SNRPB-Hu-P8 유전자, 서열번호 9로 기재되는 ZNF343-SNRPB-Hu-P9 유전자, 서열번호 10으로 기재되는 ZNF343-SNRPB-Hu-P10 유전자, 서열번호 11로

기재되는 ZNF343-SNRPB-Hu-P11 유전자, 서열번호 12로 기재되는 ZNF343-SNRPB-Hu-P12 유전자, 서열번호 13으로 기재되는 ZNF343-SNRPB-Hu-F1 유전자, 서열번호 14로 기재되는 ZNF343-SNRPB-Hu-F2 유전자, 서열번호 15로 기재되는 ZNF343-SNRPB-Hu-F3 유전자, 서열번호 16으로 기재되는 ZNF343-SNRPB-Hu-F4 유전자, 서열번호 17로 기재되는 ZNF343-SNRPB-Hu-F5 유전자, 서열번호 18로 기재되는 ZNF343-SNRPB-Hu-F6 유전자, 서열번호 19로 기재되는 ZNF343-SNRPB-Hu-F7 유전자, 서열번호 20으로 기재되는 ZNF343-SNRPB-Hu-F8 유전자, 서열번호 21로 기재되는 COX17-POPDC2-Hu-P1 유전자, 서열번호 22로 기재되는 COX17-POPDC2-Hu-P2 유전자, 서열번호 23으로 기재되는 COX17-POPDC2-Hu-P3 유전자, 서열번호 24로 기재되는 COX17-POPDC2-Hu-P4 유전자, 서열번호 25로 기재되는 COX17-POPDC2-Hu-P5 유전자, 서열번호 26으로 기재되는 COX17-POPDC2-Hu-P6 유전자, 서열번호 27로 기재되는 COX17-POPDC2-Hu-P7 유전자, 서열번호 28로 기재되는 COX17-POPDC2-Hu-P8 유전자, 서열번호 29로 기재되는 COX17-POPDC2-Hu-P9 유전자, 서열번호 30으로 기재되는 COX17-POPDC2-Hu-P10 유전자, 서열번호 31로 기재되는 COX17-POPDC2-Hu-P11 유전자, 서열번호 32로 기재되는 COX17-POPDC2-Hu-F1 유전자, 서열번호 33으로 기재되는 COX17-POPDC2-Hu-F2 유전자, 서열번호 34로 기재되는 COX17-POPDC2-Hu-F3 유전자, 서열번호 35로 기재되는 COX17-POPDC2-Hu-F4 유전자, 서열번호 36으로 기재되는 COX17-POPDC2-Hu-F5 유전자, 서열번호 37로 기재되는 COX17-POPDC2-Hu-F6 유전자, 서열번호 38로 기재되는 COX17-POPDC2-Hu-F7 유전자, 서열번호 39로 기재되는 COX17-POPDC2-Hu-F8 유전자, 서열번호 40으로 기재되는 COX17-POPDC2-Hu-F9 유전자, 서열번호 41로 기재되는 COX17-POPDC2-Hu-F10 유전자, 서열번호 42로 기재되는 COX17-POPDC2-Hu-F11 유전자, 서열번호 43으로 기재되는 COX17-POPDC2-Hu-F12 유전자, 서열번호 44로 기재되는 RBM7-REX02-Hu-P1 유전자, 서열번호 45로 기재되는 RBM7-REX02-Hu-P2 유전자, 서열번호 46으로 기재되는 RBM7-REX02-Hu-P3 유전자, 서열번호 47로 기재되는 RBM7-REX02-Hu-P4 유전자, 서열번호 48로 기재되는 PRKRIP1-ORAI2-Hu-P1 유전자, 서열번호 49로 기재되는 PRKRIP1-ORAI2-Hu-P2 유전자, 서열번호 50으로 기재되는 PRKRIP1-ORAI2-Hu-P3 유전자, 서열번호 51로 기재되는 PRKRIP1-ORAI2-Hu-P4 유전자, 서열번호 52로 기재되는 PRKRIP1-ORAI2-Hu-P5 유전자, 서열번호 53으로 기재되는 PRKRIP1-ORAI2-Hu-P6 유전자, 서열번호 54로 기재되는 ATP5F1-C1orf162-Hu-P1 유전자, 서열번호 55로 기재되는 ATP5F1-C1orf162-Hu-P2 유전자, 서열번호 56으로 기재되는 ATP5F1-C1orf162-Hu-P3 유전자 및 서열번호 57로 기재되는 ATP5F1-C1orf162-Hu-P4 유전자; 및

2) 상기 유전자의 발현이 확인되는 개체를 암 위험이 있는 개체로 판정하되,

이때,

ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-P11 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-P5 유전자, COX17-POPDC2-Hu-P8 유전자, COX17-POPDC2-Hu-P9 유전자, COX17-POPDC2-Hu-P10 유전자, RBM7-REX02-Hu-P1 유전자, 및 PRKRIP1-ORAI2-Hu-P1 유전자가 발현하면 상기 암을 림프종으로 판정하고,

COX17-POPDC2-Hu-P9 유전자, PRKRIP1-ORAI2-Hu-P1 유전자, KRIP1-ORAI2-Hu-P3 유전자, 및 ATP5F1-C1orf162-Hu-P3 유전자가 발현하면 상기 암을 전립선암으로 판정하고,

ZNF343-SNRPB-Hu-P1 유전자, ZNF343-SNRPB-Hu-P3 유전자, ZNF343-SNRPB-Hu-P4 유전자, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-P11 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-P10 유전자, PRKRIP1-ORAI2-Hu-P3 유전자, 및 ATP5F1-C1orf162-Hu-P4 유전자가 발현하면 상기 암을 자궁경부암으로 판정하고,

ZNF343-SNRPB-Hu-P8 유전자, COX17-POPDC2-Hu-P7 유전자, COX17-POPDC2-Hu-P11 유전자, COX17-POPDC2-Hu-F9 유전자, COX17-POPDC2-Hu-F11 유전자, COX17-POPDC2-Hu-F12 유전자, RBM7-REX02-Hu-P4 유전자, PRKRIP1-ORAI2-Hu-P6 유전자 및 ATP5F1-C1orf162-Hu-P4 유전자가 발현하면 상기 암을 유방암으로 판정하고,

ZNF343-SNRPB-Hu-P5 유전자, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-F5 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-F7 유전자, COX17-POPDC2-Hu-F11 유전자 및 COX17-POPDC2-Hu-F12 유전자가 발현하면 상기 암은 대장암으로 판정하고,

ZNF343-SNRPB-Hu-P1 유전자, ZNF343-SNRPB-Hu-P2 유전자, ZNF343-SNRPB-Hu-P3 유전자, ZNF343-SNRPB-Hu-P4 유전자, ZNF343-SNRPB-Hu-P11 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-P4 유전자, COX17-POPDC2-Hu-P6 유전자, COX17-POPDC2-Hu-P8 유전자, COX17-POPDC2-Hu-F3 유전자, COX17-POPDC2-Hu-F4 유전자, COX17-POPDC2-Hu-F6 유전자, PRKRIP1-ORAI2-Hu-P5 유전자 및 ATP5F1-C1orf162-Hu-P3 유전자가 발현하면 상기 암은 신장암으로 판정하고,

ZNF343-SNRPB-Hu-P1 유전자, ZNF343-SNRPB-Hu-F5 유전자, ZNF343-SNRPB-Hu-F7 유전자, COX17-POPDC2-Hu-F1 유전자, COX17-POPDC2-Hu-F3 유전자 및 PRKRIP1-ORAI2-Hu-P6 유전자가 발현하면 상기 암은 폐암으로 판정하고,

ZNF343-SNRPB-Hu-P4 유전자, ZNF343-SNRPB-Hu-P11 유전자, ZNF343-SNRPB-Hu-P12 유전자, ZNF343-SNRPB-Hu-F7 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-F10 유전자, COX17-POPDC2-Hu-F12 유전자, PRKRIP1-ORAI2-Hu-P4 유전자, PRKRIP1-ORAI2-Hu-P6 유전자 및 ATP5F1-C1orf162-Hu-P1 유전자가 발현하면 상기 암은 난소암으로 판정하고,

ZNF343-SNRPB-Hu-P11 유전자, ZNF343-SNRPB-Hu-F6 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-F2 유전자 및 COX17-POPDC2-Hu-F10 유전자가 발현하면 상기 암은 위암으로 판정하고,

PRKRIP1-ORAI2-Hu-P2 유전자 및 PRKRIP1-ORAI2-Hu-P5 유전자가 발현하면 상기 암은 자궁암으로 판정하고,

ZNF343-SNRPB-Hu-P7 유전자, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-F2 유전자, COX17-POPDC2-Hu-P1 유전자, COX17-POPDC2-Hu-F1 유전자, COX17-POPDC2-Hu-F5 유전자 및 RBM7-REXO2-Hu-P3 유전자가 발현하면 상기 암은 방광암으로 판정하고,

ZNF343-SNRPB-Hu-P8 유전자가 발현하면 상기 암은 식도암으로 판정하고,

ZNF343-SNRPB-Hu-F7 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-P10 유전자, ATP5F1-C1orf162-Hu-P2 유전자, ATP5F1-C1orf162-Hu-P3 유전자 및 ATP5F1-C1orf162-Hu-P4 유전자가 발현하면 상기 암은 간암으로 판정하고,

ZNF343-SNRPB-Hu-P8 유전자가 발현하면 상기 암은 갑상선암으로 판정하는 단계.

## 청구항 6

제 5항에 있어서, 상기 단계 1)의 시료는 조직, 세포, 전혈, 혈장, 혈청, 타액 및 뇨로 이루어진 군으로부터 선택되는 어느 하나인 것을 특징으로 하는 유전자 발현 여부 확인 방법.

## 청구항 7

제 5항에 있어서, 유전자 발현은 RT-PCR, 경쟁적 RT-PCR, 실시간 RT-PCR, RNase 보호 분석법(RNase protection assay), 노던블랏팅(Northern blotting) 및 DNA 칩 중에서 선택되는 어느 하나의 방법을 이용하여 확인하는 것을 특징으로 하는 유전자 발현 여부 확인 방법.

## 청구항 8

하기의 단계를 포함하는 개체 내 암 진단의 정보를 제공하기 위한 단백질 발현 여부 확인 방법:

1) 분리된 시료로부터 하기 모든 유전자를 코딩하는 단백질의 발현 여부를 확인하는 단계:

서열번호 1로 기재되는 ZNF343-SNRPB-Hu-P1 유전자, 서열번호 2로 기재되는 ZNF343-SNRPB-Hu-P2 유전자, 서열번호 3으로 기재되는 ZNF343-SNRPB-Hu-P3 유전자, 서열번호 4로 기재되는 ZNF343-SNRPB-Hu-P4 유전자, 서열번호 5로 기재되는 ZNF343-SNRPB-Hu-P5 유전자, 서열번호 6으로 기재되는 ZNF343-SNRPB-Hu-P6 유전자, 서열번호 7로 기재되는 ZNF343-SNRPB-Hu-P7 유전자, 서열번호 8로 기재되는 ZNF343-SNRPB-Hu-P8 유전자, 서열번호 9로 기재되는 ZNF343-SNRPB-Hu-P9 유전자, 서열번호 10으로 기재되는 ZNF343-SNRPB-Hu-P10 유전자, 서열번호 11로 기재되는 ZNF343-SNRPB-Hu-P11 유전자, 서열번호 12로 기재되는 ZNF343-SNRPB-Hu-P12 유전자, 서열번호 13으로 기재되는 ZNF343-SNRPB-Hu-F1 유전자, 서열번호 14로 기재되는 ZNF343-SNRPB-Hu-F2 유전자, 서열번호 15로 기재되는 ZNF343-SNRPB-Hu-F3 유전자, 서열번호 16으로 기재되는 ZNF343-SNRPB-Hu-F4 유전자, 서열번호 17로 기재되는 ZNF343-SNRPB-Hu-F5 유전자, 서열번호 18로 기재되는 ZNF343-SNRPB-Hu-F6 유전자, 서열번호 19로 기재되는 ZNF343-SNRPB-Hu-F7 유전자, 서열번호 20으로 기재되는 ZNF343-SNRPB-Hu-F8 유전자, 서열번호 21로 기재되는 COX17-POPDC2-Hu-P1 유전자, 서열번호 22로 기재되는 COX17-POPDC2-Hu-P2 유전자, 서열번호 23으로 기재되는 COX17-POPDC2-Hu-P3 유전자, 서열번호 24로 기재되는 COX17-POPDC2-Hu-P4 유전자, 서열번호 25로 기재되는 COX17-POPDC2-Hu-P5 유전자, 서열번호 26으로 기재되는 COX17-POPDC2-Hu-P6 유전자, 서열번호 27로 기재되는 COX17-POPDC2-Hu-P7 유전자, 서열번호 28로 기재되는 COX17-POPDC2-Hu-P8 유전자, 서열번호 29로 기재되는 COX17-POPDC2-Hu-P9 유전자, 서열번호 30으로 기재되는 COX17-POPDC2-Hu-P10 유전자, 서열번호 31로 기재되는

COX17-POPDC2-Hu-P11 유전자, 서열번호 32로 기재되는 COX17-POPDC2-Hu-F1 유전자, 서열번호 33으로 기재되는 COX17-POPDC2-Hu-F2 유전자, 서열번호 34로 기재되는 COX17-POPDC2-Hu-F3 유전자, 서열번호 35로 기재되는 COX17-POPDC2-Hu-F4 유전자, 서열번호 36으로 기재되는 COX17-POPDC2-Hu-F5 유전자, 서열번호 37로 기재되는 COX17-POPDC2-Hu-F6 유전자, 서열번호 38로 기재되는 COX17-POPDC2-Hu-F7 유전자, 서열번호 39로 기재되는 COX17-POPDC2-Hu-F8 유전자, 서열번호 40으로 기재되는 COX17-POPDC2-Hu-F9 유전자, 서열번호 41로 기재되는 COX17-POPDC2-Hu-F10 유전자, 서열번호 42로 기재되는 COX17-POPDC2-Hu-F11 유전자, 서열번호 43으로 기재되는 COX17-POPDC2-Hu-F12 유전자, 서열번호 44로 기재되는 RBM7-REX02-Hu-P1 유전자, 서열번호 45로 기재되는 RBM7-REX02-Hu-P2 유전자, 서열번호 46으로 기재되는 RBM7-REX02-Hu-P3 유전자, 서열번호 47로 기재되는 RBM7-REX02-Hu-P4 유전자, 서열번호 48로 기재되는 PRKRIP1-ORAI2-Hu-P1 유전자, 서열번호 49로 기재되는 PRKRIP1-ORAI2-Hu-P2 유전자, 서열번호 50으로 기재되는 PRKRIP1-ORAI2-Hu-P3 유전자, 서열번호 51로 기재되는 PRKRIP1-ORAI2-Hu-P4 유전자, 서열번호 52로 기재되는 PRKRIP1-ORAI2-Hu-P5 유전자, 서열번호 53으로 기재되는 PRKRIP1-ORAI2-Hu-P6 유전자, 서열번호 54로 기재되는 ATP5F1-C1orf162-Hu-P1 유전자, 서열번호 55로 기재되는 ATP5F1-C1orf162-Hu-P2 유전자, 서열번호 56으로 기재되는 ATP5F1-C1orf162-Hu-P3 유전자 및 서열번호 57로 기재되는 ATP5F1-C1orf162-Hu-P4 유전자; 및

2) 상기 단백질의 발현이 확인되는 개체를 암 위험이 있는 개체로 판정하되,  
이때,

ZNF343-SNRPB-Hu-P8 단백질, ZNF343-SNRPB-Hu-P11 단백질, COX17-POPDC2-Hu-P3 단백질, COX17-POPDC2-Hu-P5 단백질, COX17-POPDC2-Hu-P8 단백질, COX17-POPDC2-Hu-P9 단백질, COX17-POPDC2-Hu-P10 단백질, RBM7-REX02-Hu-P1 단백질, 및 PRKRIP1-ORAI2-Hu-P1 단백질이 발현하면 상기 암을 림프종으로 판정하고,

COX17-POPDC2-Hu-P9 단백질, PRKRIP1-ORAI2-Hu-P1 단백질, KRIP1-ORAI2-Hu-P3 단백질, 및 ATP5F1-C1orf162-Hu-P3 단백질이 발현하면 상기 암을 전립선암으로 판정하고,

ZNF343-SNRPB-Hu-P1 단백질, ZNF343-SNRPB-Hu-P3 단백질, ZNF343-SNRPB-Hu-P4 단백질, ZNF343-SNRPB-Hu-P8 단백질, ZNF343-SNRPB-Hu-P11 단백질, COX17-POPDC2-Hu-P3 단백질, COX17-POPDC2-Hu-P10 단백질, PRKRIP1-ORAI2-Hu-P3 단백질, 및 ATP5F1-C1orf162-Hu-P4 단백질이 발현하면 상기 암을 자궁경부암으로 판정하고,

ZNF343-SNRPB-Hu-P8 단백질, COX17-POPDC2-Hu-P7 단백질, COX17-POPDC2-Hu-P11 단백질, COX17-POPDC2-Hu-F9 단백질, COX17-POPDC2-Hu-F11 단백질, COX17-POPDC2-Hu-F12 단백질, RBM7-REX02-Hu-P4 단백질, PRKRIP1-ORAI2-Hu-P6 단백질 및 ATP5F1-C1orf162-Hu-P4 단백질이 발현하면 상기 암을 유방암으로 판정하고,

ZNF343-SNRPB-Hu-P5 단백질, ZNF343-SNRPB-Hu-P8 단백질, ZNF343-SNRPB-Hu-F5 단백질, COX17-POPDC2-Hu-P3 단백질, COX17-POPDC2-Hu-F7 단백질, COX17-POPDC2-Hu-F11 단백질 및 COX17-POPDC2-Hu-F12 단백질이 발현하면 상기 암은 대장암으로 판정하고,

ZNF343-SNRPB-Hu-P1 단백질, ZNF343-SNRPB-Hu-P2 단백질, ZNF343-SNRPB-Hu-P3 단백질, ZNF343-SNRPB-Hu-P4 단백질, ZNF343-SNRPB-Hu-P11 단백질, ZNF343-SNRPB-Hu-F8 단백질, COX17-POPDC2-Hu-P4 단백질, COX17-POPDC2-Hu-P6 단백질, COX17-POPDC2-Hu-P8 단백질, COX17-POPDC2-Hu-F3 단백질, COX17-POPDC2-Hu-F4 단백질, COX17-POPDC2-Hu-F6 단백질, PRKRIP1-ORAI2-Hu-P5 단백질 및 ATP5F1-C1orf162-Hu-P3 단백질이 발현하면 상기 암은 신장암으로 판정하고,

ZNF343-SNRPB-Hu-P1 단백질, ZNF343-SNRPB-Hu-F5 단백질, ZNF343-SNRPB-Hu-F7 단백질, COX17-POPDC2-Hu-F1 단백질, COX17-POPDC2-Hu-F3 단백질 및 PRKRIP1-ORAI2-Hu-P6 단백질이 발현하면 상기 암은 폐암으로 판정하고,

ZNF343-SNRPB-Hu-P4 단백질, ZNF343-SNRPB-Hu-P11 단백질, ZNF343-SNRPB-Hu-P12 단백질, ZNF343-SNRPB-Hu-F7 단백질, ZNF343-SNRPB-Hu-F8 단백질, COX17-POPDC2-Hu-F10 단백질, COX17-POPDC2-Hu-F12 단백질, PRKRIP1-ORAI2-Hu-P4 단백질, PRKRIP1-ORAI2-Hu-P6 단백질 및 ATP5F1-C1orf162-Hu-P1 단백질이 발현하면 상기 암은 난소암으로 판정하고,

ZNF343-SNRPB-Hu-P11 단백질, ZNF343-SNRPB-Hu-F6 단백질, ZNF343-SNRPB-Hu-F8 단백질, COX17-POPDC2-Hu-F2 단백질 및 COX17-POPDC2-Hu-F10 단백질이 발현하면 상기 암은 위암으로 판정하고,

PRKRIP1-ORAI2-Hu-P2 단백질 및 PRKRIP1-ORAI2-Hu-P5 단백질이 발현하면 상기 암은 자궁암으로 판정하고,

ZNF343-SNRPB-Hu-P7 단백질, ZNF343-SNRPB-Hu-P8 단백질, ZNF343-SNRPB-Hu-F2 단백질, COX17-POPDC2-Hu-P1 단

백질, COX17-POPDC2-Hu-F1 단백질, COX17-POPDC2-Hu-F5 단백질 및 RBM7-REX02-Hu-P3 단백질이 발현하면 상기 암은 방광암으로 판정하고,

ZNF343-SNRPB-Hu-P8 단백질이 발현하면 상기 암은 식도암으로 판정하고,

ZNF343-SNRPB-Hu-F7 단백질, ZNF343-SNRPB-Hu-F8 단백질, COX17-POPDC2-Hu-P3 단백질, COX17-POPDC2-Hu-P10 단백질, ATP5F1-C1orf162-Hu-P2 단백질, ATP5F1-C1orf162-Hu-P3 단백질 및 ATP5F1-C1orf162-Hu-P4 단백질이 발현하면 상기 암은 간암으로 판정하고,

ZNF343-SNRPB-Hu-P8 단백질이 발현하면 상기 암은 갑상선암으로 판정하는 단계.

### 청구항 9

제 8항에 있어서, 상기 단계 1)의 시료는 조직, 세포, 전혈, 혈장, 혈청, 타액 및 뇨로 이루어진 군으로부터 선택되는 어느 하나인 것을 특징으로 하는 단백질 발현 여부 확인 방법.

### 청구항 10

제 8항에 있어서, 단백질의 발현은 면역형광법, 효소면역분석법(ELISA), 웨스턴 블롯(Western Blot) 및 단백질 마이크로어레이 칩으로 이루어진 군으로부터 선택되는 어느 하나의 방법을 이용하여 확인하는 것을 특징으로 하는 단백질 발현 여부 확인 방법.

## 명세서

### 기술분야

[0001] 본 발명은 신규한 인간 융합 유전자(Conjoined gene, CG) 또는 이의 융합 유전자 전사 변이체(Conjoined gene transcript variants, CGTV), 및 이의 용도에 관한 것이다.

### 배경기술

[0002] 융합 유전자(conjoined gene, CG)는 동일한 방향성을 가지고 같은 염색체상에 존재하는 2가지 이상의 분리된 각각의 단일 유전자(모체 유전자)로부터 유래되는 적어도 하나의 엑손 부분이 결합된 전사체로서, read-through transcript 또는 공동-전사 유전자(co-transcribed gene)로 알려져 있다. 인간 게놈 데이터베이스로부터의 컴퓨터 데이터 마이닝 및 큐레이팅을 통하여, 융합 유전자 형성 동안에 생성된 신규한 엑손은 두 모체 유전자 사이에 유전자간 부위(intergenic regions)에서 주로 발생한다.

[0003] 최근에, 융합 유전자는 인간 게놈 이해를 위해 필수적으로 중요한 유전적 요인으로 알려졌지만, 그 정확한 구조 및 기능은 잘 알려져 있지 않다. 또한, 차세대 염기서열 분석(next generation sequence, NGS)을 사용한 최근의 전사체 연구에서는, 조직-특이적인 발현이 아닌, 양성 및 몇몇 암 세포주에 걸쳐 폭넓게 발현되는 부분적 융합 유전자를 동정하였다. 대부분의 염색체 사이의 융합 현상이 혈액종양, 육종, 전립선암 및 폐암과 같은 다양한 암과 관련되어 있음이 알려졌으나, 암과 융합 유전자와의 기능적 관련성은 정확하게 밝혀지지 않았다. 또한, 융합 유전자의 형성 메커니즘뿐만 아니라 인간 조직에서의 융합 유전자 발현 역시 알려진 바 없고, 이러한 융합 유전자 연구의 상당한 어려움은 인간 조직에서 그것들의 발현이 극히 낮기 때문이므로, 극미량으로 발현하는 유전자의 검출을 통한 새로운 진단 방법의 도입이 필요한 실정이다.

[0004] 미국 공개특허 제2010-0092984호, 미국 공개특허 제2009-0061454호 및 미국 공개특허 제2009-0280491호에는 *ZNF*, *SNRPB*, *COX*, *RBM*, *REX02*, *PRKRIP*, *ORAI2*, *ATP5F1* 및 *C11orf162* 단일 유전자의 발현을 검출하여 암을 진단하는 방법이 기재되어 있으나, 서로 다른 2 종류의 유전자가 융합된 유전자 및 이러한 융합 유전자가 암 진단에 이용될 수 있음에 대해서는 아직까지 밝혀진 바 없다.



[0005] 이에, 본 발명자들은 신규한 5개의 인간 융합 유전자(*ZNF343-SNRPB-Hu*, *COX17-POPDC2-Hu*, *RBM7-REXO2-Hu*, *PRKRIP1-ORAI2-Hu*, 및 *ATP5F1-Clorf162-Hu*)와 57개의 인간 융합 유전자 전사 변이체(Conjoined gene transcript variants, CGTV)를 발굴하고, 이에 대한 시퀀싱(염기서열분석) 및 생물정보학적 해석들을 수행하여 융합 유전자의 형성 기전을 확인하였다. 또한, 상기 5개의 인간 융합 유전자 전사 변이체들 중에서, 인간 종양 조직에서 특이적으로 발현하는 융합 유전자 전사 변이체를 동정하고, 융합 유전자 전사 변이체와 암 발생과의 연관성을 확인하였으며, 이를 통해 각 암종에서 특이적으로 발현하는 융합 유전자 전사 변이체를 암 진단을 위한 마커 또는 DNA 마이크로어레이, 암 진단용 키트, 및 암 진단 방법에 이용할 수 있음을 밝힘으로써, 본 발명을 완성하였다.

**발명의 내용**

**해결하려는 과제**

[0006] 본 발명의 목적은 신규한 인간 융합 유전자(Conjoined gene, CG) 또는 이의 융합 유전자 전사 변이체(Conjoined gene transcript variants, CGTVs), 상기 유전자 또는 변이체에 의해 코딩되는 단백질, 상기 융합 유전자 또는 변이체를 포함하는 발현 벡터, 및 상기 발현 벡터로 형질전환된 형질전환체를 제공하는 것이다.

[0007] 본 발명의 또 다른 목적은 신규한 인간 융합 유전자 또는 이의 융합 유전자 전사 변이체를 이용한 암 진단용 DNA 마이크로어레이, 암 진단용 키트 및 암 진단 방법을 제공하는 것이다.

**과제의 해결 수단**

[0008] 상기 목적을 달성하기 위하여, 본 발명은 하기로 구성된 군으로부터 선택되는 어느 하나의 융합 유전자 전사 변이체(Conjoined Gene Transcript Variants, CGTV)를 제공한다:

[0009] 서열번호 1로 기재되는 *ZNF343-SNRPB-Hu-P1*, 서열번호 2로 기재되는 *ZNF343-SNRPB-Hu-P2*, 서열번호 3으로 기재되는 *ZNF343-SNRPB-Hu-P3*, 서열번호 4로 기재되는 *ZNF343-SNRPB-Hu-P4*, 서열번호 5로 기재되는 *ZNF343-SNRPB-Hu-P5*, 서열번호 6으로 기재되는 *ZNF343-SNRPB-Hu-P6*, 서열번호 7로 기재되는 *ZNF343-SNRPB-Hu-P7*, 서열번호 8로 기재되는 *ZNF343-SNRPB-Hu-P8*, 서열번호 9로 기재되는 *ZNF343-SNRPB-Hu-P9*, 서열번호 10으로 기재되는 *ZNF343-SNRPB-Hu-P10*, 서열번호 11로 기재되는 *ZNF343-SNRPB-Hu-P11*, 서열번호 12로 기재되는 *ZNF343-SNRPB-Hu-P12*, 서열번호 13으로 기재되는 *ZNF343-SNRPB-Hu-F1*, 서열번호 14로 기재되는 *ZNF343-SNRPB-Hu-F2*, 서열번호 15로 기재되는 *ZNF343-SNRPB-Hu-F3*, 서열번호 16으로 기재되는 *ZNF343-SNRPB-Hu-F4*, 서열번호 17로 기재되는 *ZNF343-SNRPB-Hu-F5*, 서열번호 18로 기재되는 *ZNF343-SNRPB-Hu-F6*, 서열번호 19로 기재되는 *ZNF343-SNRPB-Hu-F7*, 서열번호 20으로 기재되는 *ZNF343-SNRPB-Hu-F8*, 서열번호 21로 기재되는 *COX17-POPDC2-Hu-P1*, 서열번호 22로 기재되는 *COX17-POPDC2-Hu-P2*, 서열번호 23으로 기재되는 *COX17-POPDC2-Hu-P3*, 서열번호 24로 기재되는 *COX17-POPDC2-Hu-P4*, 서열번호 25로 기재되는 *COX17-POPDC2-Hu-P5*, 서열번호 26으로 기재되는 *COX17-POPDC2-Hu-P6*, 서열번호 27로 기재되는 *COX17-POPDC2-Hu-P7*, 서열번호 28로 기재되는 *COX17-POPDC2-Hu-P8*, 서열번호 29로 기재되는 *COX17-POPDC2-Hu-P9*, 서열번호 30으로 기재되는 *COX17-POPDC2-Hu-P10*, 서열번호 31로 기재되는 *COX17-POPDC2-Hu-P11*, 서열번호 32로 기재되는 *COX17-POPDC2-Hu-F1*, 서열번호 33으로 기재되는 *COX17-POPDC2-Hu-F2*, 서열번호 34로 기재되는 *COX17-POPDC2-Hu-F3*, 서열번호 35로 기재되는 *COX17-POPDC2-Hu-F4*, 서열번호 36으로 기재되는 *COX17-POPDC2-Hu-F5*, 서열번호 37로 기재되는 *COX17-POPDC2-Hu-F6*, 서열번호 38로 기재되는 *COX17-POPDC2-Hu-F7*, 서열번호 39로 기재되는 *COX17-POPDC2-Hu-F8*, 서열번호 40으로 기재되는 *COX17-POPDC2-Hu-F9*, 서열번호 41로 기재되는 *COX17-POPDC2-Hu-F10*, 서열번호 42로 기재되는 *COX17-POPDC2-Hu-F11*, 서열번호 43으로 기재되는 *COX17-POPDC2-Hu-F12*, 서열번호 44로 기재되는 *RBM7-REXO2-Hu-P1*, 서열번호 45로 기재되는 *RBM7-REXO2-Hu-P2*, 서열번호 46으로 기재되는 *RBM7-REXO2-Hu-P3*, 서열번호 47로 기재되는 *RBM7-REXO2-Hu-P4*, 서열번호 48로 기재되는 *PRKRIP1-ORAI2-Hu-P1*, 서열번호 49로 기재되는 *PRKRIP1-ORAI2-Hu-P2*, 서열번호 50으로 기재되는 *PRKRIP1-ORAI2-Hu-P3*, 서열번호 51로 기재되는 *PRKRIP1-ORAI2-Hu-P4*, 서열번호 52로 기재되는 *PRKRIP1-ORAI2-Hu-P5*, 서열번호 53으로 기재되는 *PRKRIP1-ORAI2-Hu-P6*, 서열번호 54로 기재되는 *ATP5F1-Clorf162-Hu-P1*, 서열번호 55로 기재되는 *ATP5F1-Clorf162-Hu-P2*, 서열번호 56으로 기재되는 *ATP5F1-Clorf162-*

Hu-P3 및 서열번호 57로 기재되는 ATP5F1-C1orf162-Hu-P4.

- [0010] 또한, 본 발명은 본 발명에 따른 융합 유전자 전사 변이체에 의해 코딩되는 단백질, 상기 융합 유전자 전사 변이체를 포함하는 발현 벡터, 및 상기 발현 벡터로 형질전환된 형질전환체를 제공한다.
- [0011] 또한, 본 발명은 하기로 구성된 군으로부터 선택되는 어느 하나의 유전자의 올리고뉴클레오티드 또는 그의 상보 가닥 분자가 집적된, 암 진단용 DNA 마이크로어레이를 제공한다:
- [0012] 서열번호 1로 기재되는 ZNF343-SNRPB-Hu-P1 유전자, 서열번호 2로 기재되는 ZNF343-SNRPB-Hu-P2 유전자, 서열번호 3으로 기재되는 ZNF343-SNRPB-Hu-P3 유전자, 서열번호 4로 기재되는 ZNF343-SNRPB-Hu-P4 유전자, 서열번호 5로 기재되는 ZNF343-SNRPB-Hu-P5 유전자, 서열번호 6으로 기재되는 ZNF343-SNRPB-Hu-P6 유전자, 서열번호 7로 기재되는 ZNF343-SNRPB-Hu-P7 유전자, 서열번호 8로 기재되는 ZNF343-SNRPB-Hu-P8 유전자, 서열번호 9로 기재되는 ZNF343-SNRPB-Hu-P9 유전자, 서열번호 10으로 기재되는 ZNF343-SNRPB-Hu-P10 유전자, 서열번호 11로 기재되는 ZNF343-SNRPB-Hu-P11 유전자, 서열번호 12로 기재되는 ZNF343-SNRPB-Hu-P12 유전자, 서열번호 13으로 기재되는 ZNF343-SNRPB-Hu-F1 유전자, 서열번호 14로 기재되는 ZNF343-SNRPB-Hu-F2 유전자, 서열번호 15로 기재되는 ZNF343-SNRPB-Hu-F3 유전자, 서열번호 16으로 기재되는 ZNF343-SNRPB-Hu-F4 유전자, 서열번호 17로 기재되는 ZNF343-SNRPB-Hu-F5 유전자, 서열번호 18로 기재되는 ZNF343-SNRPB-Hu-F6 유전자, 서열번호 19로 기재되는 ZNF343-SNRPB-Hu-F7 유전자, 서열번호 20으로 기재되는 ZNF343-SNRPB-Hu-F8 유전자, 서열번호 21로 기재되는 COX17-POPDC2-Hu-P1 유전자, 서열번호 22로 기재되는 COX17-POPDC2-Hu-P2 유전자, 서열번호 23으로 기재되는 COX17-POPDC2-Hu-P3 유전자, 서열번호 24로 기재되는 COX17-POPDC2-Hu-P4 유전자, 서열번호 25로 기재되는 COX17-POPDC2-Hu-P5 유전자, 서열번호 26으로 기재되는 COX17-POPDC2-Hu-P6 유전자, 서열번호 27로 기재되는 COX17-POPDC2-Hu-P7 유전자, 서열번호 28로 기재되는 COX17-POPDC2-Hu-P8 유전자, 서열번호 29로 기재되는 COX17-POPDC2-Hu-P9 유전자, 서열번호 30으로 기재되는 COX17-POPDC2-Hu-P10 유전자, 서열번호 31로 기재되는 COX17-POPDC2-Hu-P11 유전자, 서열번호 32로 기재되는 COX17-POPDC2-Hu-F1 유전자, 서열번호 33으로 기재되는 COX17-POPDC2-Hu-F2 유전자, 서열번호 34로 기재되는 COX17-POPDC2-Hu-F3 유전자, 서열번호 35로 기재되는 COX17-POPDC2-Hu-F4 유전자, 서열번호 36으로 기재되는 COX17-POPDC2-Hu-F5 유전자, 서열번호 37로 기재되는 COX17-POPDC2-Hu-F6 유전자, 서열번호 38로 기재되는 COX17-POPDC2-Hu-F7 유전자, 서열번호 39로 기재되는 COX17-POPDC2-Hu-F8 유전자, 서열번호 40으로 기재되는 COX17-POPDC2-Hu-F9 유전자, 서열번호 41로 기재되는 COX17-POPDC2-Hu-F10 유전자, 서열번호 42로 기재되는 COX17-POPDC2-Hu-F11 유전자, 서열번호 43으로 기재되는 COX17-POPDC2-Hu-F12 유전자, 서열번호 44로 기재되는 RBM7-REX02-Hu-P1 유전자, 서열번호 45로 기재되는 RBM7-REX02-Hu-P2 유전자, 서열번호 46으로 기재되는 RBM7-REX02-Hu-P3 유전자, 서열번호 47로 기재되는 RBM7-REX02-Hu-P4 유전자, 서열번호 48로 기재되는 PRKRIP1-ORAI2-Hu-P1 유전자, 서열번호 49로 기재되는 PRKRIP1-ORAI2-Hu-P2 유전자, 서열번호 50으로 기재되는 PRKRIP1-ORAI2-Hu-P3 유전자, 서열번호 51로 기재되는 PRKRIP1-ORAI2-Hu-P4 유전자, 서열번호 52로 기재되는 PRKRIP1-ORAI2-Hu-P5 유전자, 서열번호 53으로 기재되는 PRKRIP1-ORAI2-Hu-P6 유전자, 서열번호 54로 기재되는 ATP5F1-C1orf162-Hu-P1 유전자, 서열번호 55로 기재되는 ATP5F1-C1orf162-Hu-P2 유전자, 서열번호 56으로 기재되는 ATP5F1-C1orf162-Hu-P3 유전자 및 서열번호 57로 기재되는 ATP5F1-C1orf162-Hu-P4 유전자.
- [0013] 또한, 본 발명은 본 발명에 따른 DNA 마이크로어레이를 포함하는 암 진단용 키트를 제공한다.
- [0014] 또한, 본 발명은 본 발명에 따른 인간 융합 유전자 또는 융합 유전자 전사 변이체를 특이적으로 증폭시키는 프라이머 쌍을 포함하는 암 진단용 키트를 제공한다.
- [0015] 또한, 본 발명은 하기의 단계를 포함하는 개체 내 암 진단의 정보를 제공하기 위한 유전자 발현 여부를 확인하는 방법을 제공한다:
- [0016] 1) 개체로부터 시료를 분리하는 단계;
- [0017] 2) 상기 분리된 시료로부터 하기로 구성된 군으로부터 선택되는 어느 하나의 유전자의 발현 여부를 확인하는 단계;
- [0018] 서열번호 1로 기재되는 ZNF343-SNRPB-Hu-P1 유전자, 서열번호 2로 기재되는 ZNF343-SNRPB-Hu-P2 유전자, 서열번호 3으로 기재되는 ZNF343-SNRPB-Hu-P3 유전자, 서열번호 4로 기재되는 ZNF343-SNRPB-Hu-P4 유전자, 서열번호 5로 기재되는 ZNF343-SNRPB-Hu-P5 유전자, 서열번호 6으로 기재되는 ZNF343-SNRPB-Hu-P6 유전자, 서열번호 7로 기재되는 ZNF343-SNRPB-Hu-P7 유전자, 서열번호 8로 기재되는 ZNF343-SNRPB-Hu-P8 유전자, 서열번호 9로 기재되는 ZNF343-SNRPB-Hu-P9 유전자, 서열번호 10으로 기재되는 ZNF343-SNRPB-Hu-P10 유전자, 서열번호 11로

기재되는 ZNF343-SNRPB-Hu-P11 유전자, 서열번호 12로 기재되는 ZNF343-SNRPB-Hu-P12 유전자, 서열번호 13으로 기재되는 ZNF343-SNRPB-Hu-F1 유전자, 서열번호 14로 기재되는 ZNF343-SNRPB-Hu-F2 유전자, 서열번호 15로 기재되는 ZNF343-SNRPB-Hu-F3 유전자, 서열번호 16으로 기재되는 ZNF343-SNRPB-Hu-F4 유전자, 서열번호 17로 기재되는 ZNF343-SNRPB-Hu-F5 유전자, 서열번호 18로 기재되는 ZNF343-SNRPB-Hu-F6 유전자, 서열번호 19로 기재되는 ZNF343-SNRPB-Hu-F7 유전자, 서열번호 20으로 기재되는 ZNF343-SNRPB-Hu-F8 유전자, 서열번호 21로 기재되는 COX17-POPDC2-Hu-P1 유전자, 서열번호 22로 기재되는 COX17-POPDC2-Hu-P2 유전자, 서열번호 23으로 기재되는 COX17-POPDC2-Hu-P3 유전자, 서열번호 24로 기재되는 COX17-POPDC2-Hu-P4 유전자, 서열번호 25로 기재되는 COX17-POPDC2-Hu-P5 유전자, 서열번호 26으로 기재되는 COX17-POPDC2-Hu-P6 유전자, 서열번호 27로 기재되는 COX17-POPDC2-Hu-P7 유전자, 서열번호 28로 기재되는 COX17-POPDC2-Hu-P8 유전자, 서열번호 29로 기재되는 COX17-POPDC2-Hu-P9 유전자, 서열번호 30으로 기재되는 COX17-POPDC2-Hu-P10 유전자, 서열번호 31로 기재되는 COX17-POPDC2-Hu-P11 유전자, 서열번호 32로 기재되는 COX17-POPDC2-Hu-F1 유전자, 서열번호 33으로 기재되는 COX17-POPDC2-Hu-F2 유전자, 서열번호 34로 기재되는 COX17-POPDC2-Hu-F3 유전자, 서열번호 35로 기재되는 COX17-POPDC2-Hu-F4 유전자, 서열번호 36으로 기재되는 COX17-POPDC2-Hu-F5 유전자, 서열번호 37로 기재되는 COX17-POPDC2-Hu-F6 유전자, 서열번호 38로 기재되는 COX17-POPDC2-Hu-F7 유전자, 서열번호 39로 기재되는 COX17-POPDC2-Hu-F8 유전자, 서열번호 40으로 기재되는 COX17-POPDC2-Hu-F9 유전자, 서열번호 41로 기재되는 COX17-POPDC2-Hu-F10 유전자, 서열번호 42로 기재되는 COX17-POPDC2-Hu-F11 유전자, 서열번호 43으로 기재되는 COX17-POPDC2-Hu-F12 유전자, 서열번호 44로 기재되는 RBM7-REX02-Hu-P1 유전자, 서열번호 45로 기재되는 RBM7-REX02-Hu-P2 유전자, 서열번호 46으로 기재되는 RBM7-REX02-Hu-P3 유전자, 서열번호 47로 기재되는 RBM7-REX02-Hu-P4 유전자, 서열번호 48로 기재되는 PRKRIP1-ORAI2-Hu-P1 유전자, 서열번호 49로 기재되는 PRKRIP1-ORAI2-Hu-P2 유전자, 서열번호 50으로 기재되는 PRKRIP1-ORAI2-Hu-P3 유전자, 서열번호 51로 기재되는 PRKRIP1-ORAI2-Hu-P4 유전자, 서열번호 52로 기재되는 PRKRIP1-ORAI2-Hu-P5 유전자, 서열번호 53으로 기재되는 PRKRIP1-ORAI2-Hu-P6 유전자, 서열번호 54로 기재되는 ATP5F1-C1orf162-Hu-P1 유전자, 서열번호 55로 기재되는 ATP5F1-C1orf162-Hu-P2 유전자, 서열번호 56으로 기재되는 ATP5F1-C1orf162-Hu-P3 유전자 및 서열번호 57로 기재되는 ATP5F1-C1orf162-Hu-P4 유전자; 및

- [0019] 3) 상기 유전자의 발현이 확인되는 개체를 암 위험이 있는 개체로 판정하되,
- [0020] 이때,
- [0021] ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-P11 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-P5 유전자, COX17-POPDC2-Hu-P8 유전자, COX17-POPDC2-Hu-P9 유전자, COX17-POPDC2-Hu-P10 유전자, RBM7-REX02-Hu-P1 유전자, 및 PRKRIP1-ORAI2-Hu-P1 유전자가 발현하면 상기 암을 림프종으로 판정하고,
- [0022] COX17-POPDC2-Hu-P9 유전자, PRKRIP1-ORAI2-Hu-P1 유전자, KRIP1-ORAI2-Hu-P3 유전자, 및 ATP5F1-C1orf162-Hu-P3 유전자가 발현하면 상기 암을 전립선암으로 판정하고,
- [0023] ZNF343-SNRPB-Hu-P1 유전자, ZNF343-SNRPB-Hu-P3 유전자, ZNF343-SNRPB-Hu-P4 유전자, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-P11 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-P10 유전자, PRKRIP1-ORAI2-Hu-P3 유전자, 및 ATP5F1-C1orf162-Hu-P4 유전자가 발현하면 상기 암을 자궁경부암으로 판정하고,
- [0024] ZNF343-SNRPB-Hu-P8 유전자, COX17-POPDC2-Hu-P7 유전자, COX17-POPDC2-Hu-P11 유전자, COX17-POPDC2-Hu-F9 유전자, COX17-POPDC2-Hu-F11 유전자, COX17-POPDC2-Hu-F12 유전자, RBM7-REX02-Hu-P4 유전자, PRKRIP1-ORAI2-Hu-P6 유전자 및 ATP5F1-C1orf162-Hu-P4 유전자가 발현하면 상기 암을 유방암으로 판정하고,
- [0025] ZNF343-SNRPB-Hu-P5 유전자, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-F5 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-F7 유전자, COX17-POPDC2-Hu-F11 유전자 및 COX17-POPDC2-Hu-F12 유전자가 발현하면 상기 암은 대장암으로 판정하고,
- [0026] ZNF343-SNRPB-Hu-P1 유전자, ZNF343-SNRPB-Hu-P2 유전자, ZNF343-SNRPB-Hu-P3 유전자, ZNF343-SNRPB-Hu-P4 유전자, ZNF343-SNRPB-Hu-P11 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-P4 유전자, COX17-POPDC2-Hu-P6 유전자, COX17-POPDC2-Hu-P8 유전자, COX17-POPDC2-Hu-F3 유전자, COX17-POPDC2-Hu-F4 유전자, COX17-POPDC2-Hu-F6 유전자, PRKRIP1-ORAI2-Hu-P5 유전자 및 ATP5F1-C1orf162-Hu-P3 유전자가 발현하면 상기 암은 신장암으로 판정하고,
- [0027] ZNF343-SNRPB-Hu-P1 유전자, ZNF343-SNRPB-Hu-F5 유전자, ZNF343-SNRPB-Hu-F7 유전자, COX17-POPDC2-Hu-F1 유전자, COX17-POPDC2-Hu-F3 유전자 및 PRKRIP1-ORAI2-Hu-P6 유전자가 발현하면 상기 암은 폐암으로 판정하고,

- [0028] ZNF343-SNRPB-Hu-P4 유전자, ZNF343-SNRPB-Hu-P11 유전자, ZNF343-SNRPB-Hu-P12 유전자, ZNF343-SNRPB-Hu-F7 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-F10 유전자, COX17-POPDC2-Hu-F12 유전자, PRKRIP1-ORAI2-Hu-P4 유전자, PRKRIP1-ORAI2-Hu-P6 유전자 및 ATP5F1-C1orf162-Hu-P1 유전자가 발현하면 상기 암은 난소암으로 판정하고,
- [0029] ZNF343-SNRPB-Hu-P11 유전자, ZNF343-SNRPB-Hu-F6 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-F2 유전자 및 COX17-POPDC2-Hu-F10 유전자가 발현하면 상기 암은 위암으로 판정하고,
- [0030] PRKRIP1-ORAI2-Hu-P2 유전자 및 PRKRIP1-ORAI2-Hu-P5 유전자가 발현하면 상기 암은 자궁암으로 판정하고,
- [0031] ZNF343-SNRPB-Hu-P7 유전자, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-F2 유전자, COX17-POPDC2-Hu-P1 유전자, COX17-POPDC2-Hu-F1 유전자, COX17-POPDC2-Hu-F5 유전자 및 RBM7-REXO2-Hu-P3 유전자가 발현하면 상기 암은 방광암으로 판정하고,
- [0032] ZNF343-SNRPB-Hu-P8 유전자가 발현하면 상기 암은 식도암으로 판정하고,
- [0033] ZNF343-SNRPB-Hu-F7 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-P10 유전자, ATP5F1-C1orf162-Hu-P2 유전자, ATP5F1-C1orf162-Hu-P3 유전자 및 ATP5F1-C1orf162-Hu-P4 유전자가 발현하면 상기 암은 간암으로 판정하고,
- [0034] ZNF343-SNRPB-Hu-P8 유전자가 발현하면 상기 암은 갑상선암으로 판정하는 단계.
- [0035] 아울러, 본 발명은 하기의 단계를 포함하는 개체 내 암 진단의 정보를 제공하기 위한 단백질의 발현 여부를 확인하는 방법을 제공한다:
- [0036] 1) 개체로부터 시료를 분리하는 단계;
- [0037] 2) 상기 분리된 시료로부터 하기로 구성된 군으로부터 선택되는 어느 하나의 유전자에 의해 코딩되는 단백질의 발현 여부를 확인하는 단계;
- [0038] 서열번호 1로 기재되는 ZNF343-SNRPB-Hu-P1 유전자, 서열번호 2로 기재되는 ZNF343-SNRPB-Hu-P2 유전자, 서열번호 3으로 기재되는 ZNF343-SNRPB-Hu-P3 유전자, 서열번호 4로 기재되는 ZNF343-SNRPB-Hu-P4 유전자, 서열번호 5로 기재되는 ZNF343-SNRPB-Hu-P5 유전자, 서열번호 6으로 기재되는 ZNF343-SNRPB-Hu-P6 유전자, 서열번호 7로 기재되는 ZNF343-SNRPB-Hu-P7 유전자, 서열번호 8로 기재되는 ZNF343-SNRPB-Hu-P8 유전자, 서열번호 9로 기재되는 ZNF343-SNRPB-Hu-P9 유전자, 서열번호 10으로 기재되는 ZNF343-SNRPB-Hu-P10 유전자, 서열번호 11로 기재되는 ZNF343-SNRPB-Hu-P11 유전자, 서열번호 12로 기재되는 ZNF343-SNRPB-Hu-P12 유전자, 서열번호 13으로 기재되는 ZNF343-SNRPB-Hu-F1 유전자, 서열번호 14로 기재되는 ZNF343-SNRPB-Hu-F2 유전자, 서열번호 15로 기재되는 ZNF343-SNRPB-Hu-F3 유전자, 서열번호 16으로 기재되는 ZNF343-SNRPB-Hu-F4 유전자, 서열번호 17로 기재되는 ZNF343-SNRPB-Hu-F5 유전자, 서열번호 18로 기재되는 ZNF343-SNRPB-Hu-F6 유전자, 서열번호 19로 기재되는 ZNF343-SNRPB-Hu-F7 유전자, 서열번호 20으로 기재되는 ZNF343-SNRPB-Hu-F8 유전자, 서열번호 21로 기재되는 COX17-POPDC2-Hu-P1 유전자, 서열번호 22로 기재되는 COX17-POPDC2-Hu-P2 유전자, 서열번호 23으로 기재되는 COX17-POPDC2-Hu-P3 유전자, 서열번호 24로 기재되는 COX17-POPDC2-Hu-P4 유전자, 서열번호 25로 기재되는 COX17-POPDC2-Hu-P5 유전자, 서열번호 26으로 기재되는 COX17-POPDC2-Hu-P6 유전자, 서열번호 27로 기재되는 COX17-POPDC2-Hu-P7 유전자, 서열번호 28로 기재되는 COX17-POPDC2-Hu-P8 유전자, 서열번호 29로 기재되는 COX17-POPDC2-Hu-P9 유전자, 서열번호 30으로 기재되는 COX17-POPDC2-Hu-P10 유전자, 서열번호 31로 기재되는 COX17-POPDC2-Hu-P11 유전자, 서열번호 32로 기재되는 COX17-POPDC2-Hu-F1 유전자, 서열번호 33으로 기재되는 COX17-POPDC2-Hu-F2 유전자, 서열번호 34로 기재되는 COX17-POPDC2-Hu-F3 유전자, 서열번호 35로 기재되는 COX17-POPDC2-Hu-F4 유전자, 서열번호 36으로 기재되는 COX17-POPDC2-Hu-F5 유전자, 서열번호 37로 기재되는 COX17-POPDC2-Hu-F6 유전자, 서열번호 38로 기재되는 COX17-POPDC2-Hu-F7 유전자, 서열번호 39로 기재되는 COX17-POPDC2-Hu-F8 유전자, 서열번호 40으로 기재되는 COX17-POPDC2-Hu-F9 유전자, 서열번호 41로 기재되는 COX17-POPDC2-Hu-F10 유전자, 서열번호 42로 기재되는 COX17-POPDC2-Hu-F11 유전자, 서열번호 43으로 기재되는 COX17-POPDC2-Hu-F12 유전자, 서열번호 44로 기재되는 RBM7-REXO2-Hu-P1 유전자, 서열번호 45로 기재되는 RBM7-REXO2-Hu-P2 유전자, 서열번호 46으로 기재되는 RBM7-REXO2-Hu-P3 유전자, 서열번호 47로 기재되는 RBM7-REXO2-Hu-P4 유전자, 서열번호 48로 기재되는 PRKRIP1-ORAI2-Hu-P1 유전자, 서열번호 49로 기재되는 PRKRIP1-ORAI2-Hu-P2 유전자, 서열번호 50으로 기재되는 PRKRIP1-ORAI2-Hu-P3 유전자, 서열번호 51로 기재되는 PRKRIP1-ORAI2-Hu-P4 유전자, 서열번호 52로 기재되는 PRKRIP1-ORAI2-Hu-P5 유전자, 서열번호 53으로 기재되는

PRKRIP1-ORAI2-Hu-P6 유전자, 서열번호 54로 기재되는 ATP5F1-C1orf162-Hu-P1 유전자, 서열번호 55로 기재되는 ATP5F1-C1orf162-Hu-P2 유전자, 서열번호 56으로 기재되는 ATP5F1-C1orf162-Hu-P3 유전자 및 서열번호 57로 기재되는 ATP5F1-C1orf162-Hu-P4 유전자; 및

- [0039] 3) 상기 단백질의 발현이 확인되는 개체를 암 위험이 있는 개체로 판정하되,
- [0040] 이때,
- [0041] ZNF343-SNRPB-Hu-P8 단백질, ZNF343-SNRPB-Hu-P11 단백질, COX17-POPDC2-Hu-P3 단백질, COX17-POPDC2-Hu-P5 단백질, COX17-POPDC2-Hu-P8 단백질, COX17-POPDC2-Hu-P9 단백질, COX17-POPDC2-Hu-P10 단백질, RBM7-REXO2-Hu-P1 단백질, 및 PRKRIP1-ORAI2-Hu-P1 단백질이 발현하면 상기 암을 림프종으로 판정하고,
- [0042] COX17-POPDC2-Hu-P9 단백질, PRKRIP1-ORAI2-Hu-P1 단백질, KRIP1-ORAI2-Hu-P3 단백질, 및 ATP5F1-C1orf162-Hu-P3 단백질이 발현하면 상기 암을 전립선암으로 판정하고,
- [0043] ZNF343-SNRPB-Hu-P1 단백질, ZNF343-SNRPB-Hu-P3 단백질, ZNF343-SNRPB-Hu-P4 단백질, ZNF343-SNRPB-Hu-P8 단백질, ZNF343-SNRPB-Hu-P11 단백질, COX17-POPDC2-Hu-P3 단백질, COX17-POPDC2-Hu-P10 단백질, PRKRIP1-ORAI2-Hu-P3 단백질, 및 ATP5F1-C1orf162-Hu-P4 단백질이 발현하면 상기 암을 자궁경부암으로 판정하고,
- [0044] ZNF343-SNRPB-Hu-P8 단백질, COX17-POPDC2-Hu-P7 단백질, COX17-POPDC2-Hu-P11 단백질, COX17-POPDC2-Hu-F9 단백질, COX17-POPDC2-Hu-F11 단백질, COX17-POPDC2-Hu-F12 단백질, RBM7-REXO2-Hu-P4 단백질, PRKRIP1-ORAI2-Hu-P6 단백질 및 ATP5F1-C1orf162-Hu-P4 단백질이 발현하면 상기 암을 유방암으로 판정하고,
- [0045] ZNF343-SNRPB-Hu-P5 단백질, ZNF343-SNRPB-Hu-P8 단백질, ZNF343-SNRPB-Hu-F5 단백질, COX17-POPDC2-Hu-P3 단백질, COX17-POPDC2-Hu-F7 단백질, COX17-POPDC2-Hu-F11 단백질 및 COX17-POPDC2-Hu-F12 단백질이 발현하면 상기 암은 대장암으로 판정하고,
- [0046] ZNF343-SNRPB-Hu-P1 단백질, ZNF343-SNRPB-Hu-P2 단백질, ZNF343-SNRPB-Hu-P3 단백질, ZNF343-SNRPB-Hu-P4 단백질, ZNF343-SNRPB-Hu-P11 단백질, ZNF343-SNRPB-Hu-F8 단백질, COX17-POPDC2-Hu-P4 단백질, COX17-POPDC2-Hu-P6 단백질, COX17-POPDC2-Hu-P8 단백질, COX17-POPDC2-Hu-F3 단백질, COX17-POPDC2-Hu-F4 단백질, COX17-POPDC2-Hu-F6 단백질, PRKRIP1-ORAI2-Hu-P5 단백질 및 ATP5F1-C1orf162-Hu-P3 단백질이 발현하면 상기 암은 신장암으로 판정하고,
- [0047] ZNF343-SNRPB-Hu-P1 단백질, ZNF343-SNRPB-Hu-F5 단백질, ZNF343-SNRPB-Hu-F7 단백질, COX17-POPDC2-Hu-F1 단백질, COX17-POPDC2-Hu-F3 단백질 및 PRKRIP1-ORAI2-Hu-P6 단백질이 발현하면 상기 암은 폐암으로 판정하고,
- [0048] ZNF343-SNRPB-Hu-P4 단백질, ZNF343-SNRPB-Hu-P11 단백질, ZNF343-SNRPB-Hu-P12 단백질, ZNF343-SNRPB-Hu-F7 단백질, ZNF343-SNRPB-Hu-F8 단백질, COX17-POPDC2-Hu-F10 단백질, COX17-POPDC2-Hu-F12 단백질, PRKRIP1-ORAI2-Hu-P4 단백질, PRKRIP1-ORAI2-Hu-P6 단백질 및 ATP5F1-C1orf162-Hu-P1 단백질이 발현하면 상기 암은 난소암으로 판정하고,
- [0049] ZNF343-SNRPB-Hu-P11 단백질, ZNF343-SNRPB-Hu-F6 단백질, ZNF343-SNRPB-Hu-F8 단백질, COX17-POPDC2-Hu-F2 단백질 및 COX17-POPDC2-Hu-F10 단백질이 발현하면 상기 암은 위암으로 판정하고,
- [0050] PRKRIP1-ORAI2-Hu-P2 단백질 및 PRKRIP1-ORAI2-Hu-P5 단백질이 발현하면 상기 암은 자궁암으로 판정하고,
- [0051] ZNF343-SNRPB-Hu-P7 단백질, ZNF343-SNRPB-Hu-P8 단백질, ZNF343-SNRPB-Hu-F2 단백질, COX17-POPDC2-Hu-P1 단백질, COX17-POPDC2-Hu-F1 단백질, COX17-POPDC2-Hu-F5 단백질 및 RBM7-REXO2-Hu-P3 단백질이 발현하면 상기 암은 방광암으로 판정하고,
- [0052] ZNF343-SNRPB-Hu-P8 단백질이 발현하면 상기 암은 식도암으로 판정하고,
- [0053] ZNF343-SNRPB-Hu-F7 단백질, ZNF343-SNRPB-Hu-F8 단백질, COX17-POPDC2-Hu-P3 단백질, COX17-POPDC2-Hu-P10 단백질, ATP5F1-C1orf162-Hu-P2 단백질, ATP5F1-C1orf162-Hu-P3 단백질 및 ATP5F1-C1orf162-Hu-P4 단백질이 발현하면 상기 암은 간암으로 판정하고,
- [0054] ZNF343-SNRPB-Hu-P8 단백질이 발현하면 상기 암은 갑상선암으로 판정하는 단계.

**발명의 효과**

[0055] 본 발명은 신규한 인간 융합 유전자(Conjoined gene, CG) 또는 이의 융합 유전자 전사 변이체(Conjoined gene transcript variants, CGTV)를 제공하고, 상기 융합 유전자 전사 변이체들 중 각 암 종에서 특이적으로 발현하는 융합 유전자 전사 변이체를 다양한 암 종의 진단을 위한 표적으로 유용하게 사용할 수 있다.

**도면의 간단한 설명**

[0056] 도 1은 업스트림 및 다운스트림의 모체 유전자들(parent genes) 사이 연결된 인간 및 침팬지 융합 유전자들(conjoined genes, CGs)의 엑손 구조를 나타내는 그림이다.

여기서, 상응하는 참조 업스트림 및 다운스트림의 모체 유전자들은 점이 찍힌 녹색선으로 나타낸다. **a)**는 인간 융합 유전자 *ZNF343-SNRPB-Hu* 및 침팬지 융합 유전자 *ZNF343-SNRPB-Ch* (Accession number: AK305048)의 엑손 구조를 나타낸다. 20명의 다른 인간 CGTVs (*ZNF343-SNRPB-Hu-P1-12* 및 *ZNF343-SNRPB-Hu-F1-8*)를 나타낸다. 파란색 및 오렌지색 화살 막대는 각각 업스트림 유전자 *ZNF343* 및 다운스트림 유전자 *SNRPB*를 나타낸다. 마지막 엑손을 제외하고 동일한 엑손 구조를 가지는 두 개의 *SNRPB* 변이체 (Accession numbers: NM\_198216 및 NM\_003091)는, 이런 CG 형성에 참가한다. 얇은 선은 유전자내 및 유전자 간 부위 내에서 인트론 서열을 나타낸다. 노란색 및 밝은 파란색 박스는 각각 업스트림 및 다운스트림 유전자에서의 엑손, 인간 CGTVs 및 침팬지 CG mRNA를 나타낸다. 얇은 검정색 화살표는 업스트림 및 다운스트림의 모체 유전자들에서 CDS 부위들을 나타낸다. 인간 CGTVs (*ZNF343-SNRPB-Hu-F1-8*)에 대한 cDNAs는 업스트림 및 다운스트림의 모체 유전자들의 두 CDS 부위를 덮는 것으로 고안된 PCR 프라이머들을 이용하여 증폭되었다. 다른 인간 CGTVs (*ZNF343-SNRPB-Hu-P1-12*)는 엑손 "2" (업스트림 유전자 *ZNF343*) 및 엑손 "3" (오렌지색 박스)(다운스트림 유전자 *SNRPB*) 사이 증폭되었다. 녹색 박스들 ("a", "b", "c", "d", "e", "f", "a<sup>v</sup>" 및 "c<sup>v</sup>")는 이런 CG 형성 동안 생성되는 신규한 엑손 및 연장을 나타낸다. 엑손 "a<sup>v</sup>", "c<sup>v</sup>", "4<sup>v</sup>", "5<sup>v1</sup>", "5<sup>v2</sup>", "6<sup>v1</sup>" 및 "6<sup>v2</sup>"는 각각 엑손 "a", "c", "4" (*ZNF343*), "5" (*ZNF343*) 및 "6" (*ZNF343*)로부터 유래된 변이 형태를 나타낸다. **a-1)**은 엑손 "6" (*ZNF343*)으로부터 변이 형태 "6<sup>v1</sup>" 및 "6<sup>v2</sup>"의 도출 동안 절단 부위의 서열 수준에 대한 상세한 설명을 나타낸다. 다중 서열 정렬은 절단 부위를 강조하기 위해 *ZNF343-SNRPB-Hu-F8*, *-P5*, CCDS (*ZNF343*) 및 *ZNF343*로부터 유래된 엑손 서열들 (밝은 파란색 박스, 엑손 "6<sup>v1</sup>", "6<sup>v2</sup>" 및 "6") 간을 보여준다.

**b)**는 인간 융합 유전자 *COX17-POPDC2-Hu* 및 침팬지 융합 유전자 *COX17-POPDC2-Ch* (Accession number: AK306023)의 엑손 구조를 나타낸다. 23명의 다른 인간 CGTVs (*COX17-POPDC2-Hu-F1-12* 및 *COX17-POPDC2-Hu-P1-12*)를 나타낸다. 인간 CGTVs (*COX17-POPDC2-Hu-F1-12*)에 대한 cDNAs는 업스트림 유전자 *COX17* 및 다운스트림 유전자 *POPDC2*의 두 개의 CDS 부위를 덮는 것으로 고안된 PCR 프라이머들을 이용하여 증폭되었다. 다른 인간 CGTVs (*COX17-POPDC2-Hu-P1-12*)는 엑손 "1" (*COX17*) 및 엑손 "3" (오렌지 박스)(*POPDC2*) 사이 증폭되었다. 녹색 박스들 ("a", "b", "c", "d" 및 "e")은 새롭게 생성된 신규한 엑손 및 연장을 나타낸다. 엑손 "1<sup>v1</sup>" 및 "1<sup>v2</sup>"는 엑손 "1" (*POPDC2*)로부터 유래된 변이 형태이다. **b-1)**은 엑손 "1" (*POPDC2*)로부터 변이 형태 엑손 "1<sup>v1</sup>" 및 "1<sup>v2</sup>"의 도출 동안 이용되는 절단 부위의 서열 수준에 대한 상세한 설명을 나타낸다. 다중 서열 정렬은 절단 부위를 강조하기 위해 *COX17-POPDC2-Hu-P1*, *-P3*, *-P6-11*, *-F2*, *-F6-8* 및 *-F10-12*, *COX17-POPDC2-Ch*, CCDS (*POPDC2*) 및 *POPDC2*로부터 유래된 엑손 서열들 (노란색 박스들, 엑손 "1<sup>v1</sup>", "1<sup>v2</sup>" 및 "1") 간을 보여준다.

**c)**는 인간 융합 유전자 *RBM7-REXO2-Hu* 및 침팬지 융합 유전자 *RBM7-REXO2-Ch* (Accession number: AK305532)의 엑손 구조를 나타낸다. 4명의 다른 인간 CGTVs (*RBM7-REXO2-Hu-P1-4*)를 나타낸다. 얇은 자주색 화살표는 모체 유전자 *RBM7* 및 *REXO2*의 업스트림에 위치한 세 번째 유전자 *C11orf71*을 나타낸다. 녹색 박스들 ("a", "b", "c", "d", "e", "f" 및 "g")은 새롭게 생성된 신규한 엑손을 나타낸다. 별표 표시는 두 개의 신규한 엑손 생성 ("f" 및 "g")이 발생하는 곳의 위치를 나타내고, 이는 침팬지 융합 유전자 *RBM7-REXO2-Ch*가 인간 게놈에서 세 개의 연속적인 인접한 유전자 좌위보다 폭이 멀리 걸쳐져 있는 것을 나타낸다. 엑손 "4<sup>v</sup>" 및 "5<sup>v</sup>"는 엑손 "4" 및 "5" (*RBM7*)로부터 유래된 변이 형태이다. **c-1)**은 엑손 "5" (*RBM7*)로부터 변이 형태 엑손 "5<sup>v</sup>"의 도출 동안 이용되는 절단 부위의 서열 수준에 대한 상세한 설명을 나타낸다. 다중 서열 정렬은 절단 부위를 강조하기 위해 *RBM7-REXO2-Hu-P4* 및 *RBM7-REXO2-Ch*, CCDS (*RBM7*) 및 *RBM7*로부터 유래된 엑손 서열들 (노란색 박스들, 엑손

"5<sup>v</sup>" 및 "5") 간을 보여준다. **c-2)**는 침팬지 융합 유전자 *RBM7-REXO2-Ch*의 신규한 엑손 "g" 및 "f"의 상세한 엑손 구조를 나타낸다. 이런 짧은 설명은 UCSC Genome Browser로부터 획득하였다.

**d)**는 인간 융합 유전자 *PRKRIP1-ORAI2-Hu* 및 침팬지 융합 유전자 *PRKRIP1-ORAI2-Ch*의 엑손 구조를 나타낸다. 6명의 인간 CGTVs (*PRKRIP1-ORAI2-Hu-P1-6*)를 나타낸다. 녹색 박스는 침팬지의 신규한 엑손 "a"를 나타내고, 엑손 "1<sup>v</sup>"는 엑손 "1"(*PRKRIP1*)로부터 유래된 변이 형태이다. 엑손 "3<sup>v</sup>", "6<sup>v1</sup>" 및 "6<sup>v2</sup>"는 엑손 "3" (*ORAI2*) 및 엑손 "6" (*PRKRIP1*)으로부터 유래된 변이 형태이다. **d-1)**은 엑손 "6" (*PRKRIP1*)으로부터 변이 형태 엑손 "6<sup>v1</sup>" 및 "6<sup>v2</sup>"의 도출 동안 이용되는 절단 부위의 서열 수준에 대한 상세한 설명을 나타낸다. 다중 서열 정렬은 절단 부위를 강조하기 위해 *PRKRIP1-ORAI2-Hu-P4*, *-P6*, CCDS (*PRKRIP1*) 및 *PRKRIP1*으로부터 유래된 엑손 서열들 (노란색 박스들, 엑손 "6<sup>v1</sup>", "6<sup>v2</sup>" 및 "6") 간을 보여준다.

**e)**는 인간 융합 유전자 *ATP5F1-C1orf162-Hu* 및 침팬지 융합 유전자 *ATP5F1-C1orf162-Ch*의 엑손 구조를 나타낸다. 4명의 다른 인간 CGTVs (*ATP5F1-C1orf162-Hu-P1-4*)를 나타낸다. 얇은 붉은 보라색 화살 막대는 또다른 유전자 *WDR77*의 축소된 엑손 구조를 나타내며, 상기 *WDR77*의 전사 방향은 업스트림 유전자 *ATP5F1* 및 다운스트림 유전자 *C1orf162*의 전사 방향에 상반된다. 침팬지 CG mRNA (*ATP5F1-C1orf162-Ch*)는 인간 게놈에서 세 개의 연속적인 인접한 유전자 좌 보다 폭이 멀리 걸쳐져 있는 것을 나타낸다. 녹색 박스들 ("a", "b", "b<sup>v</sup>" 및 "c")은 새롭게 생성된 신규한 엑손 및 연장을 나타낸다. 엑손 "1", "7<sup>v1</sup>", "7<sup>v2</sup>", "3<sup>v</sup>" (빨간색 경계선의 박스) 및 "b<sup>v</sup>"는 각각 엑손 1 (*ATP5F1*), "7" (*ATP5F1*), "3" (*C1orf162*) 및 "b"로부터 유래된 변이 형태이다. **e-1)**은 엑손 "7" (*ATP5F1*)로부터 변이 형태 엑손 "7<sup>v1</sup>" 및 "7<sup>v2</sup>"의 도출 동안 이용되는 절단 부위의 서열 수준에 대한 상세한 설명을 나타낸다. 다중 서열 정렬은 절단 부위를 강조하기 위해 *ATP5F1-C1orf162-Hu-P2*, *ATP5F1-C1orf162-Ch*, CCDS (*ATP5F1*) 및 *ATP5F1*으로부터 유래된 엑손 서열들 (엑손 "7<sup>v1</sup>", "7<sup>v2</sup>" 및 "7") 간을 보여준다. **e-2)**는 침팬지 융합 유전자 *ATP5F1-C1orf162-Ch*의 엑손 "e+1"의 상세한 엑손 구조를 나타낸다. 이런 짧은 설명은 UCSC Genome Browser로부터 획득하였다. 엑손 "e+1<sup>v</sup>" (*ATP5F1-C1orf162-Ch*)은 엑손 "1" (*WDR77*: 빨간색 화살촉으로 나타냄) 및 엑손 "1"의 일부 (*ATP5F1*: 빨간색 화살촉으로 나타냄)를 포함한다. **e-3)**는 엑손 "3<sup>v</sup>" (*ATP5F1-C1orf162-Ch*)의 말단에서 poly-A 신호 서열의 상세한 설명을 나타낸다. 흥미롭게도, CG 형성에서 다중-엑손 스킵핑(multi-exon skipping) ("4", "5" 및 "6") 때문에, 침팬지 CG는 엑손 "3"의 마지막 A 및 엑손화된 인트론 서열의 첫 번째 5-bp (GTAAA)를 결합함에 의해 생성된 새로운 poly-A 신호 서열을 갖는다.

**도 2**는 전형적인 융합 유전자의 가장 일반적이고 필수적인 구조적 특징을 나타낸다. CG 형성에서, 유전자 간 부위에 인접한, 업스트림 모체 유전자의 최종 엑손 (왼쪽 부위) 및 다운스트림 모체 유전자의 첫 번째 엑손 (오른쪽 부위)는 스킵핑되거나 절단된다. 특히, 업스트림 모체 유전자의 poly-A 신호는 스킵핑 또는 절단을 통해 완전히 제거된다. 3' UTR의 말단에서 파란색 부호 (4개의 뾰족한 모서리를 가진)는 업스트림 모체 유전자의 poly-A 신호 서열을 나타낸다. 업스트림 및 다운스트림 모체 유전자 사이 유전자 간 부위는 신규한 엑손 생성을 위한 핫스팟(hot spot)이다. 작은 막대 (구배의 검은색-파란색 강도를 갖는)는 엑손 생성 경향을 보여준다; 신규한 엑손 생성 빈도는 유전자 간 부위 내에서 높으나, 상기 빈도는 유전자 간 부위로부터 떨어진 위치는 낮다. 긴 막대 (구배의 검은색-파란색 강도를 갖는)는 엑손 보존 경향을 보여준다; 유전자 간 부위로부터 더 멀리 있는 엑손은 전형적인 CG 형성 동안 더 높게 보존된다.

**도 3**은 전이 인자(transposable element, TE)로부터 유래된 새롭게 생성된 신규한 엑손을 나타내는 그림이다.

**a)**는 TE로부터 유래된 신규한 엑손 "c" (ZNF343-SNRPB-Hu)를 나타내는 그림이다. 이런 스냅샷은 UCSC Genome Browser로부터 획득하였다. 첫 번째 트랙(track)은 GC 퍼센트를 나타낸다. 두 번째 트랙은 20명의 인간 CGTVs 및 침팬지 mRNA (ZNF343-SNRPB-Ch)의 엑손 구조를 나타낸다. 세 번째 및 네 번째 트랙은 각각 업스트림 및 다운스트림 참조 CCDS 및 유전자들의 엑손 구조를 나타낸다. 다섯 번째 내지 열네 번째 트랙은 전이 인자, 게놈 DNA 인자 및 그 외 특징을 나타낸다. 자주색 직사각형은 전이 인자 AluSp로부터 유래된 신규한 엑손 "c"를 나타낸다. 다운스트림의 녹색 박스는 신규한 엑손 "c"의 위치를 나타낸다.

**b)**는 TE로부터 유래된 신규한 엑손 "a", "b" 및 "c" (COX17-POPDC2-Hu)를 나타내는 그림이다. 이런 스냅샷은 UCSC Genome Browser로부터 획득하였다. 첫 번째 트랙(track)은 GC 퍼센트를 나타낸다. 두 번째 트랙은 23명의 인간 CGTVs 및 침팬지 mRNA (COX17-POPDC2-Ch)의 엑손 구조를 나타낸다. 세 번째 및 네 번째 트랙은 각각

업스트림 및 다운스트림 참조 CCDS 및 유전자들의 엑손 구조를 나타낸다. 다섯 번째 내지 열네 번째 트랙은 전이 인자, 계놈 DNA 인자 및 그 외 특징을 나타낸다. 자주색 직사각형은 전이 인자 L2a, MIRc 및 MLT1L로부터 각각 유래된 신규한 엑손 "a", "b" 및 "c"를 나타낸다. 다운스트림의 녹색 박스는 신규한 엑손 "a", "b" 및 "c"의 위치를 나타낸다.

c)는 TE로부터 유래된 신규한 엑손 "a", "c", "d", "e" 및 "g" (RBM7-REX02-Hu and RBM7-REX02-Ch)를 나타내는 그림이다. 이런 스냅샷은 UCSC Genome Browser로부터 획득하였다. 첫 번째 트랙(track)은 GC 퍼센트를 나타낸다. 두 번째 트랙은 4명의 인간 CGTVs 및 침팬지 mRNA의 엑손 구조를 나타낸다. 세 번째 및 네 번째 트랙은 각각 업스트림 및 다운스트림 참조 CCDS 및 유전자들의 엑손 구조를 나타낸다. 다섯 번째 내지 열네 번째 트랙은 전이 인자, 계놈 DNA 인자 및 그 외 특징을 나타낸다. 자주색 직사각형은 전이 인자 L1MB3, Charlie13b, Charlie13b, MER41B 및 MIRb로부터 각각 유래된 신규한 엑손 "a", "c", "d", "e" 및 "g"를 나타낸다. 다운스트림의 녹색 박스는 신규한 엑손 "a", "c", "d", "e" 및 "g"의 위치를 나타낸다.

d)는 TE로부터 유래된 신규한 엑손 "a" (PRKRIP1-ORAI2-Ch)를 나타내는 그림이다. 이런 스냅샷은 UCSC Genome Browser로부터 획득하였다. 첫 번째 트랙(track)은 GC 퍼센트를 나타낸다. 두 번째 트랙은 6명의 인간 CGTVs (PRKRIP1-ORAI2-Hu) 및 침팬지 mRNA의 엑손 구조를 나타낸다. 세 번째 및 네 번째 트랙은 각각 업스트림 및 다운스트림 참조 CCDS 및 유전자들의 엑손 구조를 나타낸다. 다섯 번째 내지 열네 번째 트랙은 전이 인자, 계놈 DNA 인자 및 그 외 특징을 나타낸다. 자주색 직사각형은 전이 인자 MER로부터 각각 유래된 신규한 엑손 "a"를 나타낸다. 다운스트림의 녹색 박스는 신규한 엑손 "a"의 위치를 나타낸다.

e)는 TE로부터 유래된 신규한 엑손 "a" 및 "b" (ATP5F1-C1orf162-Hu 및 ATP5F1-C1orf162-Ch)를 나타내는 그림이다. 이런 스냅샷은 UCSC Genome Browser로부터 획득하였다. 첫 번째 트랙(track)은 GC 퍼센트를 나타낸다. 두 번째 트랙은 4명의 인간 CGTVs 및 침팬지 mRNA의 엑손 구조를 나타낸다. 세 번째 및 네 번째 트랙은 각각 업스트림 및 다운스트림 참조 CCDS 및 유전자들의 엑손 구조를 나타낸다. 다섯 번째 내지 열네 번째 트랙은 전이 인자, 계놈 DNA 인자 및 그 외 특징을 나타낸다. 자주색 직사각형은 전이 인자 L2a 및 MLT1E1A로부터 각각 유래된 신규한 엑손 "a" 및 "b"를 나타낸다. 다운스트림의 녹색 박스는 신규한 엑손 "a" 및 "b"의 위치를 나타낸다.

도 4는 인간 계놈 전반에 걸친 융합 유전자의 신규 엑손 생성 분포를 나타낸 그림이다. 신규 엑손 중에서 유전자간 부위(intergenic:215, intergenic:482)는 SINE, LINE, LTR 및 DAN에 상응하는 비례 퍼센트로 융합 유전자를 만들어낸다. a는 인트라제닉(intergenic) 엑손 생성 분포이고, b는 인터제닉(intergenic) 엑손 생성 분포를 나타낸다.

도 5는 인간 계놈에서 융합 유전자 형성의 신규 메커니즘을 나타낸 그림이다.

a)는 정상 인간 계놈에서 단일 부분 유전자의 전사를 나타낸다. 갈색-주황색 구는 RNA 폴리머라제(polymerase)가 붉은색 화살표 방향으로 두 가닥 계놈 DNA를 이동하며 세 개의 단일 센스 유전자를 전사함을 가리킨다(주황, 녹색 및 붉은-주황 부분은 인트론 부위 또는 유전자간 부위를 나타내는 검은색 줄에 의해 분리되었다). RNA 폴리머라제 II 및 두 계놈 DNA 가닥 사이에 접촉 부위에서 단일 가닥에서 불룩 튀어나온 곳은, 전사 장치에 의해 접촉 부위에서 두 가닥 계놈 DNA가 멜팅(melting)될 수 있는 부위에서의 전사된 가닥의 노출을 나타낸다. 분홍색 별은 각 유전자 내 폴리(A) 신호 서열을 나타낸다.

b)는 3' UTRs의 결실 또는 절단 돌연변이를 갖는 비정상적 인간 계놈에서 융합 유전자의 전사가 폴리(A) 신호 부위의 제거를 유발함을 나타낸다. 주황색 및 초록색 유전자 내 폴리(A) 신호 부위의 계놈 결실에 의해, 융합 유전자의 전사 장치는 첫 번째(주황색) 유전자로부터 유전자가 부위를 통해 마지막(세 번째, 붉은-주황색) 유전자까지 그 전사를 끊임없이 지속하고, 두 번째(초록색) 유전자 및 또 다른 유전자간 부위는 융합 유전자 전-mRNA 전사체를 만들어내게 된다.

도 6은 인간 정상 조직 및 이에 상응하는 종양 조직으로부터 유래된 총 RNAs를 이용하여 증폭된 RT-PCR 산물을 나타내는 그림이다.

a)는 인간 융합 유전자 *ZNF343-SNRPB-Hu*의 8개의 전장 CDS 스페닝(폭) CGTVs의 RT-PCR 산물을 나타낸다. 이들의 발현은 30종류의 인간 정상 조직 및 이에 상응하는 종양 조직에서 조사하였다. 화살표 및 숫자는 각각 *ZNF343-SNRPB-Hu-F1-8*의 RT-PCR 산물인 밴드를 나타낸다. *GAPDH* RT-PCR 산물은 RNA 투입에 대한 대조군으로 사용되었다. b)는 인간 융합 유전자 *COX17-POPDC2-Hu*의 12개의 전장 CDS 스페닝(폭) CGTVs의 RT-PCR 산물을 나타낸다. 이들의 발현은 상기 언급된 조직들에서 조사하였다. 화살표 및 숫자는 각각 *COX17-POPDC2-Hu-F1-12*



의 RT-PCR 산물인 밴드를 나타낸다. GAPDH RT-PCR 산물은 RNA 투입에 대한 대조군으로 사용되었다.

도 7은 30종의 인간 암 조직 및 그에 상응하는 정상 조직으로부터 분리된 전체 RNA를 이용한 RT-PCR에 의해 의해 증폭된 부분 융합 유전자 cDNA 서열을 나타내는 겔 사진이다.

도 8은 HEK293세포주에서 인간 융합 유전자 ZNF343-SNRPB-Hu-F5의 발현 결과를 나타내는 겔 사진이다.

**발명을 실시하기 위한 구체적인 내용**

이하, 본 발명을 상세히 설명한다.

[0057]

본 발명은 하기로 구성된 군으로부터 선택되는 어느 하나의 융합 유전자 전사 변이체(Conjoined Gene Transcript Variants, CGTV)를 제공한다: 서열번호 1로 기재되는 ZNF343-SNRPB-Hu-P1, 서열번호 2로 기재되는 ZNF343-SNRPB-Hu-P2, 서열번호 3으로 기재되는 ZNF343-SNRPB-Hu-P3, 서열번호 4로 기재되는 ZNF343-SNRPB-Hu-P4, 서열번호 5로 기재되는 ZNF343-SNRPB-Hu-P5, 서열번호 6으로 기재되는 ZNF343-SNRPB-Hu-P6, 서열번호 7로 기재되는 ZNF343-SNRPB-Hu-P7, 서열번호 8로 기재되는 ZNF343-SNRPB-Hu-P8, 서열번호 9로 기재되는 ZNF343-SNRPB-Hu-P9, 서열번호 10으로 기재되는 ZNF343-SNRPB-Hu-P10, 서열번호 11로 기재되는 ZNF343-SNRPB-Hu-P11, 서열번호 12로 기재되는 ZNF343-SNRPB-Hu-P12, 서열번호 13으로 기재되는 ZNF343-SNRPB-Hu-F1, 서열번호 14로 기재되는 ZNF343-SNRPB-Hu-F2, 서열번호 15로 기재되는 ZNF343-SNRPB-Hu-F3, 서열번호 16으로 기재되는 ZNF343-SNRPB-Hu-F4, 서열번호 17로 기재되는 ZNF343-SNRPB-Hu-F5, 서열번호 18로 기재되는 ZNF343-SNRPB-Hu-F6, 서열번호 19로 기재되는 ZNF343-SNRPB-Hu-F7, 서열번호 20으로 기재되는 ZNF343-SNRPB-Hu-F8, 서열번호 21로 기재되는 COX17-POPDC2-Hu-P1, 서열번호 22로 기재되는 COX17-POPDC2-Hu-P2, 서열번호 23으로 기재되는 COX17-POPDC2-Hu-P3, 서열번호 24로 기재되는 COX17-POPDC2-Hu-P4, 서열번호 25로 기재되는 COX17-POPDC2-Hu-P5, 서열번호 26으로 기재되는 COX17-POPDC2-Hu-P6, 서열번호 27로 기재되는 COX17-POPDC2-Hu-P7, 서열번호 28로 기재되는 COX17-POPDC2-Hu-P8, 서열번호 29로 기재되는 COX17-POPDC2-Hu-P9, 서열번호 30으로 기재되는 COX17-POPDC2-Hu-P10, 서열번호 31로 기재되는 COX17-POPDC2-Hu-P11, 서열번호 32로 기재되는 COX17-POPDC2-Hu-F1, 서열번호 33으로 기재되는 COX17-POPDC2-Hu-F2, 서열번호 34로 기재되는 COX17-POPDC2-Hu-F3, 서열번호 35로 기재되는 COX17-POPDC2-Hu-F4, 서열번호 36으로 기재되는 COX17-POPDC2-Hu-F5, 서열번호 37로 기재되는 COX17-POPDC2-Hu-F6, 서열번호 38로 기재되는 COX17-POPDC2-Hu-F7, 서열번호 39로 기재되는 COX17-POPDC2-Hu-F8, 서열번호 40으로 기재되는 COX17-POPDC2-Hu-F9, 서열번호 41로 기재되는 COX17-POPDC2-Hu-F10, 서열번호 42로 기재되는 COX17-POPDC2-Hu-F11, 서열번호 43으로 기재되는 COX17-POPDC2-Hu-F12, 서열번호 44로 기재되는 RBM7-REX02-Hu-P1, 서열번호 45로 기재되는 RBM7-REX02-Hu-P2, 서열번호 46으로 기재되는 RBM7-REX02-Hu-P3, 서열번호 47로 기재되는 RBM7-REX02-Hu-P4, 서열번호 48로 기재되는 PRKRIP1-ORAI2-Hu-P1, 서열번호 49로 기재되는 PRKRIP1-ORAI2-Hu-P2, 서열번호 50으로 기재되는 PRKRIP1-ORAI2-Hu-P3, 서열번호 51로 기재되는 PRKRIP1-ORAI2-Hu-P4, 서열번호 52로 기재되는 PRKRIP1-ORAI2-Hu-P5, 서열번호 53으로 기재되는 PRKRIP1-ORAI2-Hu-P6, 서열번호 54로 기재되는 ATP5F1-C1orf162-Hu-P1, 서열번호 55로 기재되는 ATP5F1-C1orf162-Hu-P2, 서열번호 56으로 기재되는 ATP5F1-C1orf162-Hu-P3 및 서열번호 57로 기재되는 ATP5F1-C1orf162-Hu-P4.

[0059]

또한, 본 발명은 상기 융합 유전자 전사 변이체에 의해 코딩되는 단백질, 상기 융합 유전자 전사 변이체를 포함하는 발현 벡터, 및 상기 발현 벡터로 형질전환된 형질전환체를 제공한다.

[0060]

본 발명은 이전 연구에서 얻어진 1,933개의 고품질의 전장 칩랜지 정소 cDNA 서열을 통해 14개의 신규 칩랜지 융합 유전자를 발굴하였다. 상기 칩랜지의 융합 유전자와 인간 융합 유전자의 상동성을 확인하기 위하여, 인간 정상 조직 및 이에 상응하는 중앙 조직 30종의 전체 RNA를 이용하여 역전사 증합효소 연쇄 반응(RT-PCR)을 수행하였다. 상기 PCR 프라이머(표 1 참조)는 칩랜지 융합 서열에 상응하는 두 개의 이웃한 인간 모체 유전자 사이의 유전자간 부위(intergenic region)에 놓여진 코딩 서열(CDS) 또는 UTR 서열에 대해 주로 고안되었다. 일반적인 30회의 PCR 증폭 사이클(cycle)에서는 어떠한 시각화된 밴드도 검출할 수 없었으나, PCR 사이클 수를 점차적으로 증가시키면서 실험을 수행한 결과, 43회의 사이클 횟수에서 이후 실험에 이용하기 위해 충분한 양인 2,450 개의 PCR 산물을 얻었고, 이를 클로닝하고 서열을 분석하였다. 서열 분석을 통해 160개의 PCR 산물이 융합 유전자 서열(융합 유전자 전사 변이체)로 예상되었으며, 나머지 산물들은 예상되는 서열을 가지고 있지 않았다. 실험 조건에서, 표적화된 융합 유전자를 검출할 수 있는 성공률은 0.06%로 측정되었으며, 이는 융합 유전자가 인간 세포에서 극히 낮은 수준으로 발현된다는 것을 의미한다. 얻어진 160개의 융합 유전자 전사 변이체

(CGTV) 서열을 분석한 결과, 이들이 5개의 인간 융합 유전자(*ZNF343-SNRPB-Hu*, *COX17-POPDC2-Hu*, *RBM7-REXO2-Hu*, *PRKRIP1-ORAI2-Hu* 및 *ATP5F1-C1orf162-Hu*)에 속한 57개의 단일 융합 유전자 형태 내로 분류됨을 확인하였다.

- [0061] 첫 번째로, *ZNF343-SNRPB-Hu*는 인간 20번 염색체에서 10.968 kb의 간격으로 분리되어 있는 모체 유전자 *ZNF343*(업스트림, upstream) 및 *SNRPB*(다운스트림, downstream)을 갖는 융합유전자이다. 총 20개의 융합 유전자 전사 변이체(CGTV) 중 12개는 부분적 CGTV를 갖는, *ZNF343-SNRPB-Hu-P1~12*이고, 8개는 전체 CDS 부위를 갖는 CGTV인 *ZNF343-SNRPB-Hu-F1~8*이다(도 1a 및 도 6a 참조). 20개의 CGTV의 서열의 구조적 분석을 통하여, 신규 엑손 생성, 엑손 스킵(skip) 및/또는 절단, 엑손 신장 및 서열 결실과 같은, 다양한 변형적 스플라이싱(splicing)이 일어남을 관찰하였다.
- [0062] 새롭게 만들어진 신규 엑손(엑손 "a", "b", "c", "e" 및 "f") 5종이 9개의 CGTV 구조(*ZNF343-SNRPB-Hu-P1, -P2, -P3, -P4, -P6, -P11, -P12, -F4* 및 *-F6*)의 유전자간 부위(intragenic region)내에서 확인되었다. 신규 엑손 "c"는 전이 인자(transposable element) 서열인 AluSp로 동정되었다. 특히, 상기 신규 엑손인 "b" 및 "f"는 지금까지 동정된 것 중 가장 짧은 엑손인 3개의 염기(TCC)로 구성되어 있었다. 이전 연구에서 새로 만들어진 3개의 염기의 엑손을 갖는 변형된 스플라이싱 전사 변이체가 상기 유전자의 다양한 변형적 스플라이싱 형태와 대조하여, 조직 특이적으로 발현됨이 제시되었다. 놀랍게도, 두 개의 융합 유전자 전사 변이체(CGTVs)(*ZNF343-SNRPB-Hu-P3* 및 *P4*)는 자궁경부, 신장 및 난소 종양 조직에서 조직 특이적으로 발현함을 확인하였다(표 2 참조).
- [0063] 또 다른 새로운 현상은 세 개의 이웃한 엑손간 융합에 의해 새로운 엑손이 형성된다는 것이다(*ZNF343-SNRPB-Hu-P12*의 "4+a+3"). 구조적인 관점에서, 새로운 엑손 "a"는 세 개의 엑손("4", "a" 및 "3")의 특이적 융합에 의해 형성된, 업스트림 모체 유전자(*ZNF343*)의 "3" 및 "4" 엑손 사이의 인트론 서열로부터 기원된 변형된 엑손이다.
- [0064] 또한 엑손 연장에 의해 생성된 두 개의 신규 엑손 ("1+d") 및 "a<sup>v</sup>+3" 를 발견하였다. 이 중 하나는 *ZNF343-SNRPB-Hu-P10*의 엑손 "1"의 수용체 부위에 47bp의 인트론 서열("d")이 삽입되어 만들어졌다. 또 다른 하나는 *ZNF343-SNRPB-Hu-P9*의 공여체 부위에 85 bp의 인트론 서열("a<sup>v</sup>")이 삽입되어 만들어졌고, 새롭게 만들어진 신규 엑손 "a"의 절단된 형태를 가진다. 추가로, 각각 엑손 "6"(*ZNF343*)으로부터 2661 bp 및 2664 bp가 결실된 "6<sup>v1</sup>" 및 "6<sup>v2</sup>"(*ZNF343-SNRPB-Hu-P5* 및 *-F8*), 및 신규 엑손 "c"로부터 93 bp가 결실된 "6<sup>v2</sup>"(*ZNF343-SNRPB-Hu-P2*)의 세 개의 CGTV 구조 내 세 가지 형태의 엑손 절단을 관찰하였다(도 1a-1 참조).
- [0065] 특히, 21개 CGTV 구조를 갖는 업스트림 단일 유전자 *ZNF343*의 마지막 엑손("6")은 절단된 두 개(*ZNF343-SNRPB-Hu-P5* 및 *-F8*)를 제외하고, 거의 모두가 융합 유전자 형성에서 모두 스킵되었다. 엑손 "6" 내에서, 이러한 절단은 그 CDS("6<sup>v1</sup>" 및 "6<sup>v2</sup>")내에 일어났다. 이러한 현상은 상기 업스트림 모체 유전자의 5'UTR의 완전한 소실을 의미하고, 추가로 업스트림 유전자의 전사 종결을 유발하는 폴리-A 신호 서열의 제거를 의미한다. 융합 유전자가 업스트림의 폴리-A 신호가 제거되었음에도 불구하고 완벽한 유전자 구조를 갖기 위해서는 또 다른 신호 서열이 다운스트림에 존재해야만 한다. 이러한 관점에서, 임의의 폴리-A 서열이 없는 불완전한 mRNA 구조를 갖는, 본 발명의 인간 융합 유전자와 비교하여, 폴리-A 서열(AK305048)의 완벽한 mRNA 구조를 갖는, 칩펜지 융합 유전자(CG) mRNA를 이용하여 신호 서열을 분석하였다. 생물정보학 프로그램 "HCpolya" ([http://zeus2.itb.cnr.it/~webgene/wwwHC\\_polya.html](http://zeus2.itb.cnr.it/~webgene/wwwHC_polya.html))에 의한 폴리-A 신호 서열 예측을 통해, 1747 내지 1752 bp 부위에 위치한 적당한 폴리-A 신호 서열의 후보를 확인하였다(AATAAA). 이러한 결과는 융합 유전자 형성 기전 내 업스트림 모체 유전자 내 폴리-A 신호 서열이 완벽히 제거되고, 다운스트림 유전자의 신호 서열이 융합 유전자 mRNA의 전사 종결에 중요한 역할을 한다는 관점을 제공한다.
- [0066] 다음으로, 융합 유전자 형성에서 엑손 보존에 대해 연구하였다. 12개의 부분적 융합 유전자 전사 변이체(*ZNF343-SNRPB-Hu-P1~12*)의 구조적 분석을 통해, 21개의 CGTV 구조 내 다운스트림 유전자의 시작 엑손(*SNRPB*의 엑손 "1")이 연장된 한 개(*ZNF343-SNRPB-Hu-P10*)를 제외하고, 업스트림 유전자 *ZNF343*의 마지막 엑손("6")과 유사하게, 완벽히 스킵되었음을 확인하였다. 게다가, 인트론 부위에서 멀리 떨어진 엑손(*SNRPB*의 엑손 "2" 및 "3", 및 *ZNF343*의 엑손 "2")은 두 가지 다른 엑손(*ZNF343*의 엑손 "3" 및 "4")의 낮은 보존성과 비교할 때 매우 높게 보존되어 나타났다. 이러한 보존 경향은 4개의 엑손 "4", "5", "6" 및 "7"이 완전히 보존되어 있는, 8 개의 CGTVs(*ZNF343-SNRPB-Hu-F1~8*)의 전장 CDS의 구조적 분석을 통해 명확해졌다.

- [0067] *SNRPB*는 UCSC 유전체 브라우저(Genome Browser)에 따른 4개의 다양한 변형된 스플라이싱(splicing) 변이체를 갖는데, 이들 중 두 개는 그 마지막 엑손의 길이가 길거나(uc002wfz.1) 짧은(uc002wga.1) 차이가 있는 것을 제외하고 같은 유전자 구조를 갖는다. 추가적 분석을 통하여, 본 발명의 8개의 전장 CDS CGTV 중 5개는 변이체(uc002wfz.1)의 엑손 "7"과 같은 서열을 갖고, 나머지 3개는 다른 변이체(uc002wga.1)와 같은 서열을 가짐을 확인하였다(도 1 및 도 3a 참조). 이것은 *SNRPB*로부터 유래한 두 개의 다양한 변이체가 다양한 스플라이싱 배열 간 선택을 통해 융합 유전자를 형성에 경쟁적으로 관여할 수 있음을 의미한다.
- [0068] 상기 결과를 통해 두 가지 사실을 얻어낼 수 있다. 첫째, 인트론 부위로부터 멀리 떨어진 엑손은 인트론 부위에 가까운 엑손보다 더 높은 보존성을 나타내었다. 두 번째, 이러한 엑손 보존성의 강도는 융합 유전자 형성과 연관된 두 개의 모체 유전자 사이에 차이가 날 수 있다.
- [0069] 두 번째는 인간 3번 염색체에서 8.968 kb의 간격으로 분리되어 있는 모체 유전자 *COX17* 및 *POPDC2*를 갖는 융합 유전자 *COX17-POPDC2-Hu*이다(도 1b 참조). 총 23개의 다양한 융합 유전자 전사 변이체(CGTV) 중 11개는 부분적 CGTV를 갖는, *COX17-POPDC2-Hu-P1-11*이고, 12개는 전체 CDS 부위를 갖는 CGTV인 *COX17-POPDC2-Hu-F1-12*이다(도 6b 참조). 23개의 CGTV 중 21개(91%)의 구조 내 인트론에서 만들어진 두 가지 형태의 신규 엑손인 "b" 및 "c"를 확인하였다. 흥미롭게도, 상기 두 개의 신규 엑손은 각각 전이 인자 서열인 MIRc(SINE) 및 MLT1L(LTR)로부터 유래되었다(도 3b 참조). 추가적으로, 두 개의 CGTV(*COX17-POPDC2-Hu-P9* 및 *-11*)의 인트론 부위 내에 새롭게 만들어진 두 개의 엑손("a" 및 "d")을 확인하였고, 상기 "a"는 전이 인자 L2a로부터 유래하였다(도 3b 참조). 이러한 결과는 전이 인자(TE)가 융합 유전자의 형성에 중요함을 의미한다.
- [0070] 엑손 생성 빈도를 비교하여, 엑손의 생성이 하나의 융합 유전자로부터 유래한 융합 유전자 전사 변이체(CGTV) 사이에 존재하는 인트론 부위보다 유전자간 부위 내에서 더 자주 일어난다는 것을 확인하였다. 이러한 결과는 유전자간 부위가 새로운 엑손의 생체 내 합성의 가장 중요한 부분이 될 수 있음을 의미하고, 융합 유전자가 단일 유전자의 변형된 스플라이싱(splicing)과 명확히 구별될 수 있는 유일한 현상일 수 있다. 추가로, 7 개의 인간 전체 CDS CGTV 구조 및 칩젠지 융합 유전자 mRNA(AK306023) 내 공여체 부위("3")로부터 147 bp 서열 연장("e")에 의해 새롭게 생성된 엑손("e"+"3")을 확인하였다.
- [0071] 엑손 보존 분석을 통해 23개의 *COX17-POPDC2-Hu* CGTVs의 형성에서 유전자간 부위로부터 멀리 떨어진 엑손이 프록시말(proximal) 엑손보다 더 높은 수준으로 보존되어 있음을 확인하였다. 23개의 CGTV 구조 내에서, 엑손 보존성은 엑손 "1" 및 "2"(*COX17*) 뿐 아니라, 다운스트림 유전자의 엑손 "4"에서도 완벽하였다. 다른 엑손들에 있어서, 보존성은 유전자간 부위로부터 멀리 떨어진 엑손에서 높게 나타났다; *POPDC2* 에서 엑손 "3"(91%) 및 엑손 "2"(61%)는 높은 보존성을 나타내었다. 하지만, 이러한 경향은 유전자간 부위에 인접한 두 개의 엑손에서 더욱 명확히 나타났다(업스트림 유전자의 마지막 엑손 및 다운스트림 유전자의 첫 번째 엑손). 놀랍게도, 융합 유전자의 형성에 있어서 두 개의 엑손이 완벽히 스킵(skip)되거나 절단되었음을 확인하였다. 절단 위치("1<sup>v1</sup>")는 하나의 CGTV(*COX17-POPDC2-Hu-F11*)의 엑손 "1<sup>v1</sup>"("1<sup>v2</sup>")에서 12bp가 결실된 것을 제외하고, 14개의 CGTV 및 칩젠지 융합 유전자 mRNA에서 동일하였다. 업스트림 유전자의 마지막 엑손의 완벽한 스킵(skip)은 폴리-A 신호의 제거를 의미한다. 상기 융합 유전자 분석에서 언급한 바와 같이, 완벽한 mRNA 구조(AK306023)를 갖는 칩젠지 CG mRNA를 이용하여 다운스트림 부위의 폴리-A 신호 서열을 분석하였다. 이러한 신호 서열은 이러한 융합 유전자에서 1574 내지 1579 bp(AATAAA)에 위치하고 있다. 이러한 결과는 첫 번째 융합 유전자로부터 관찰된 바를 뒷받침한다: 업스트림 유전자의 폴리-A 신호 서열을 갖는 3'UTR이 완벽히 제거되고, 다운스트림의 폴리-A 신호가 융합 유전자의 전사를 종결하는데 이용되어야만 한다.
- [0072] *POPDC2*는 PDP 패밀리의 한 종류로, 인간 유전체 내 *POP2*(uc003ecy.1)와 거의 같은 좌위(locus)에 위치하고 있다. 흥미롭게도, 융합 유전자의 형성에 자주 참여하는 엑손 "1<sup>v1</sup>"는 다운스트림 유전자의 엑손 "1"의 절단된 형태이고, *POP2* 유전자 구조 내 두 번째 엑손의 서열과 동일하다. 이러한 결과는 구조적인 관점에서, *POP2*의 엑손 "2"가 *COX17-POPDC2-Hu*의 형성에 참여하기 위해 *POPDC2*의 엑손 "1" 보다 더 강력한 경향성을 가짐을 의미한다. 게다가, 상기 CGTV에서 신규 엑손 "c"가 *POP2*의 첫 번째 엑손으로 사용되었다. 상기 두 개의 엑손 이용 패턴으로부터, 융합 유전자 형성에서 엑손의 선택은 두 개의 단일 유전자 내 엑손 이용에 비의존적이고, 스플라이싱(splicing) 부위의 선택에서 강한 경쟁의 결과로 나타난다.
- [0073] 세 번째, 융합 유전자 *RBM7-REXO2-Hu*는 인간 11번 염색체에서 30.542 kb의 간격으로 분리되어 있는 모체 유전자 *RBM7* 및 *REXO2*를 갖는 융합유전자이다(도 1c 참조). 이러한 융합 유전자는 총 4개의 CGTVs(*RBM7-REXO-Hu-P1-4*)로부터 전사되었다. 대부분의 융합 유전자는 두 개의 이웃한 모체 유전자 사이에 융합된 형태로 나타나지

만, 세 개의 유전자가 융합된 융합 유전자는 인간에서 매우 드물다. 본 발명의 참조인 침팬지 융합 전사체 (*RBM7-REXO2-Ch*)는 세 개의 유전자 좌위(loci)에 걸쳐져 있다(*RBM7*, *REXO2* 및 *C11orf71*). 이러한 융합 유전자는 *C11orf71* 및 UCSC 데이터베이스(database)에서 확인되지 않는 또 다른 EST(BG718298) 사이에 유전자간 부위 내 새롭게 만들어진 두 개의 신규 엑손("f" 및 "g")을 나타낸다.

[0074] 특히, *RBM7-REXO2-Hu* 내 유전자간 부위와 인접한 두 개의 엑손이 이러한 융합 유전자의 형성에서 완벽히 스킵핑되거나 절단되었음을 관찰하였다. 구체적으로, 업스트림 유전자의 마지막 엑손의 절단은 CDS 부위내에서 일어났다. 흥미롭게도, 상기 절단은 상기 언급된 두 개의 융합 유전자와 마찬가지로, 업스트림 유전자의 폴리-A 신호 서열의 제거의 결과였다(도 1c-1 참조). 추가적 분석을 통해 또 다른 폴리-A 신호가 상기 두 개의 융합 유전자와 일치하게, 다운스트림 모체 유전자의 1887 내지 1891 bp(AATAAG)에 위치함을 확인하였다.

[0075] 또한, 엑손 "4"의 변이체("4<sup>v</sup>")가 수용체 부위에 3 염기 서열(CAG)의 작은 인트론의 삽입에 의해 야기된 서열 연장을 갖음에도, 프록시말(proximal) 엑손(*REXO2*의 "2", "3" 및 "4")은 융합 유전자의 형성에서 낮은 보존성을 보이는 반면, 유전자간 부위로부터 멀리 떨어진 엑손 "3"(*RBM7*) 및 "5"(*REXO2*)는 CGTV 구조 내에서 매우 높은 보존성을 보였다. 침팬지 융합 유전자 mRNA를 포함한 모든 5개의 CGTV의 유전자간 부위 내에서 새롭게 만들어진 6개의 신규 엑손("a", "c", "d", "e", "f" 및 "g")을 구조적 분석을 통해 확인하였다. 또한, TE로부터 유래한 5개의 새롭게 만들어진 신규 엑손 서열을 확인하였다: L1MB3(LINE), Charlie13b(DNA), LIMeC(LINE), MER41B(LTR) 및 MIRb(SINE)(도 3c 참조). 추가로, 신규 엑손 "2"가 융합 유전자 형성에 자주 참여하였다.

[0076] 네 번째 *PRKRIP1-ORAI2-Hu* 융합 유전자는 인간 7번 염색체에서 6.867 kb의 간격으로 분리되어 있는 모체 유전자 *PRKRIP1* 및 *ORAI2*를 갖는 융합 유전자로, 6 개의 CGTV(*PRKRIP1-ORAI2-Hu-P1-6*)로 전사되었다(도 1d 참조). 구조적 분석을 통해, 스트림에 인접한 엑손이 절단된 두 개의 CGTV(*PRKRIP1-ORAI2-Hu-P4* 및 *-P6*)를 제외하고, 모든 CGTV의 유전자간 부위에 인접한 엑손이 완벽히 스킵핑되었음을 확인하였다. 이러한 두 개의 CGTV 구조 내 절단된 마지막 업스트림 엑손의 서열 수준에 대한 구조적 분석에 의해, 절단이 UTR 내에서 일어나고, 폴리-A 신호 서열 부위는 완벽히 결실되었음을 확인하였다(도 1d-1 참조). 상기 세 개의 융합 유전자와 마찬가지로, 폴리-A 신호가 다운스트림 부위에 유지되는 지의 여부를 측정하였으며, 침팬지 mRNA 서열(AK306465) 내 3005 내지 3010 bp(AATACA)에 있는 신호 부위를 확인하였다. 이러한 현상은 상기 세 개의 융합 유전자 결과와 일치한다.

[0077] 또한 침팬지 융합 유전자 mRNA의 유전자간 부위(모체 유전자 *ORAI2*) 내 새롭게 만들어진 신규 엑손("a")이 본 발명의 6개 인간 CGTV 및 UCSC 유전체 데이터베이스(database)에는 존재하지 않음을 확인하였다. 이러한 결과는 신규 엑손 "a"가 전이 인자 서열 MER2(LINE)에서 유래한 침팬지 특이적 엑손일 수 있음을 나타낸다(도 3d 참조).

[0078] 다섯 번째 *ATP5F1-C1orf162-Hu* 융합 유전자는 인간 1번 염색체에서 12.079 kb의 간격으로 분리되어 있는 모체 유전자 *ATP5F1* 및 *C1orf162*를 갖는 융합 유전자이고, 4개의 다른 인간 CGTV(*ATP5F1-C1orf162-Hu-P1-4*)로 전사되었다(도 1e 참조). 이러한 분석에서, 침팬지의 새로운 형태의 융합 유전자(*ATP5F1-C1orf162-Ch*)가 세 개의 이웃한 모체 유전자 *ATP5F1*, *C1orf162* 및 *WDR77* 사이에서 융합되었음을 확인하였다. 첫 번째 침팬지 엑손, "c+1<sup>v</sup>"은 *WDR77*의 첫 번째 엑손 및 *ATP5F1*의 첫 번째 엑손의 일부분을 포함하고 있다(도 1e-2 및 도 3e 참조). UCSC 유전체 브라우저(Browser)로부터는 엑손 구조를 동정할 수 없었다.

[0079] 흥미로운 관점에서, 상기 4개의 융합 유전자에서 나타낸 바와 같이, 유전자간 부위 내 새롭게 만들어진 신규 엑손은 전이 인자로부터 유래한 서열이었다. 이러한 융합 유전자에서, 유전자간 부위 내 새롭게 만들어진 엑손 "b"는 인간과 침팬지 모두에서 높게 보존되어 있다. 서열 분석을 통해 각각, 전이 인자 서열 MLT1E1A(LTR) 및 L2a(비-LTR 이동 유전자)에서 유래된 엑손 "b" 뿐 아니라, 새로운 유전체내 엑손 "a" (*ATP5F1-C1orf162-Ch*)를 밝혔다(도 3e 참조). 침팬지 특이적 엑손 생성을 의미하는, UCSC 데이터베이스 조사에 의해 엑손 "a"와 상동성을 갖는 엑손을 찾지 못하였다. 신규 엑손 "a"를 갖는 침팬지 CG mRNA가 침팬지 특이적 유전자 변이체일 수 있음을 의미하는 결과로 해석하였다. 이전의 연구는 단일 유전자의 개별 변이체가 조직 분화의 측정에 중요한 역할을 수행할 수 있음을 제시하였다. 이러한 점에서, 침팬지 특이적 CG mRNA가 진화하는 동안에 침팬지 고유 특성에 이바지할 수 있다는 가설이 합리적임을 알 수 있다.

[0080] 상기 4개의 인간 융합 유전자의 구조적 분석을 통하여, 각각의 모체 유전자 내 마지막 엑손 및 첫 번째 엑손이 융합 유전자의 형성에서 완벽히 절단되거나 스킵핑되었음을 확인하였다. 이러한 관점에서, 이러한 구조가 융합 유전자 형성의 전형적인 현상으로 정의하였다. 흥미롭게도, 이러한 융합 유전자는 상기 4개의 융합 유전자와 비교하여 다운스트림 모체 유전자 내 첫 번째 엑손과 인접한 유전자내 부위가 완벽히 보존되는 이례적인 구조를

갖는다(도 1e 참조). 하지만, 상기 전형적인 현상과 일치하게, 업스트림 모체 유전자(*ATP5F1*)의 마지막 인접 엑손은 CDS(*ATP5F1-C1orf162-Hu-P3*) 또는 UTR(*ATP5F1-C1orf162-Hu-P2*, *-P4* 및 *ATP5F1-C1orf162-Ch*) 내로 완벽히 스킵핑되거나 절단되었다. 추가적 분석을 통해 상기 절단은 업스트림의 폴리-A 신호 서열에서 신호 부위가 절단되어 일어남을 확인하였다(도 1e-1 참조). 상기 4개의 융합 유전자 구조와 반대로, 침팬지 CG mRNA(AK306934)는 엑손 "3" 이후에 긴 폴리-A 서열을 갖는 완벽한 mRNA 구조를 갖지만, 상기 다운스트림 모체 유전자(*C1orf162*) 내 세 개 엑손("4", "5" 및 "6")의 완벽한 결실을 통해 구성되었다(도 1e-3 참조). 일반적으로, 이러한 구조적 패턴은 단일 유전자 경우에서 다중-엑손 스킵핑에 의해 생성된 변형된 변이체와 유사하다. 중요하게, 상기 4개의 융합 유전자와는 달리, 다운스트림 유전자의 폴리-A 신호 서열을 갖는 부위는 제거되었다. 따라서, 이러한 침팬지 융합 유전자 mRNA가 완벽한 mRNA가 전사되었지만 인공구조가 아니라면, 이후 폴리-A 신호 서열은 폴리-A 꼬리 서열 근처에서 발견될 수 있다. 본 서열 분석은 엑손 "3" 직후에 마지막 엑손 "3"이 18 bp 인트론 서열 연장(GTAAATGAGCATTCTCCT)을 가짐을 나타내었다. "HCpolya"에 의한 폴리-A 신호 서열의 예측을 통해, 1585 내지 1591 bp 위치에 새롭게 만들어진 6 bp의 폴리-A 신호 서열(AGTAAA)을 확인하였다(도 1e-3 참조). 놀랍게도, 새롭게 만들어진 폴리-A 신호는 엑손 "3"의 마지막 염기 및 추가로 엑손화된 인트론 서열의 최초 5개 염기(GTAAA)의 조합으로부터 구성되었다. 이러한 결과는 인트론 서열의 엑손화와 같은, 또 다른 이벤트에 의해 새롭게 만들어진 폴리-A 신호 서열이 다운스트림 유전자의 폴리-A 신호 서열이 절단된 곳에서 융합 유전자 전사 종결에 사용될 수 있다.

[0081] 침팬지 융합 유전자와 상동성을 보이는 새로운 5개의 인간 융합유전자들과 57개의 그 융합 유전자 전사 변이체(Conjoined Gene Alternative Splicing Transcript Variants, CGTVs)를 발견하였으며, 57개의 CGTV 구조내에서 각 업스트림(upstream) 유전자의 마지막 엑손(엑손, 폴리-A 신호 서열을 갖는)이 완벽히 스킵핑(skipping)되거나 절단되어 있음을 확인하였다. 이 연구에서 우리가 해명한 융합 유전자 형성 기전은 다음과 같다; 융합 유전자 형성에 참가하는 두 이웃 단일 유전자들 중에서 업스트림(5'-Upstream Region)에 존재하는 단일 유전자의 3'-말단 폴리-A 신호 서열의 제거가 융합유전자형성을 위한 필수적 전제조건으로 된다.

[0082] 그러한 폴리-A 신호 서열의 제거는 그 업스트림 단일 유전자(Upstream gene)에서 시작된 전사(Transcription)가 업스트림 단일 유전자의 3'-말단 배열에서 멈춤이 없이 두 단일 유전자들 사이에 존재하는 유전자간 부위(Intergenic region)와 다운스트림 단일 유전자(Downstream gene)에서 연속적으로 진행될 수 있는 전제 조건을 보장한다. 결과적으로 융합 유전자의 전사는 다운스트림 모체 단일 유전자(Downstream gene)의 3'-말단 폴리-A 신호 서열에 의해 정지되며, 이러한 융합유전자의 mRNA는 그 두 이웃 단일 유전자들의 엑손(exon)들과 이 두 단일 유전자들 사이에 유전자간 부위들에 존재하는 새로운 엑손들을 포함하게 된다. 만일 다운스트림 단일 유전자(Downstream gene)의 3'-말단 폴리-A 신호 서열마저도 제거된 경우 업스트림 단일 유전자(Upstream gene)에서 시작된 융합 유전자의 전사는 그 다운스트림 단일 유전자의 내부 인트론(intron) 배열에 존재하는 Poly-A 신호 서열을 엑손화함으로써 형성된 새로운 전사 정지 신호를 이용하여 종결된다. 또한 두 단일 유전자들 사이에 존재한 유전자간 부위(Intergenic region)로부터 멀리에 위치한 엑손들이 융합 유전자의 형성에 더 잘 참가한다는 것이 이 연구에서 발견되었다. 두 단일 유전자들 사이에 유전자간 부위(Intergenic region)에 존재하는 새로운 엑손들은 전이 인자(전이 인자(transposable element)s)로부터 기원되었다. 이 연구에서 우리는 세계 최초로 인간 암 조직들에서만 매우 특이적으로 발견되는 융합 유전자 전사 변이체(CGTVs)를 동정하였고, 이러한 융합유전자들이 인간의 암 발생과 깊이 연관되어 있음을 증명하였다. 따라서, 이렇게 동정된 본 발명의 융합 유전자 전사 변이체(CGTVs)들은 암 유발을 측정하는 데 중요한 표적으로 이용될 수 있다. 결과적으로, 본 발명은 인간 유전체의 유전자 구조 전경도에 새로운 관점을 제공한다.

[0083] 또한, 본 발명은 하기로 구성된 군으로부터 선택되는 어느 하나의 유전자의 올리고뉴클레오티드 또는 그의 상보 가닥 분자가 집적된, 암 진단용 DNA 마이크로어레이를 제공한다:

[0084] 서열번호 1로 기재되는 ZNF343-SNRPB-Hu-P1 유전자, 서열번호 2로 기재되는 ZNF343-SNRPB-Hu-P2 유전자, 서열번호 3으로 기재되는 ZNF343-SNRPB-Hu-P3 유전자, 서열번호 4로 기재되는 ZNF343-SNRPB-Hu-P4 유전자, 서열번호 5로 기재되는 ZNF343-SNRPB-Hu-P5 유전자, 서열번호 6으로 기재되는 ZNF343-SNRPB-Hu-P6 유전자, 서열번호 7로 기재되는 ZNF343-SNRPB-Hu-P7 유전자, 서열번호 8로 기재되는 ZNF343-SNRPB-Hu-P8 유전자, 서열번호 9로 기재되는 ZNF343-SNRPB-Hu-P9 유전자, 서열번호 10으로 기재되는 ZNF343-SNRPB-Hu-P10 유전자, 서열번호 11로 기재되는 ZNF343-SNRPB-Hu-P11 유전자, 서열번호 12로 기재되는 ZNF343-SNRPB-Hu-P12 유전자, 서열번호 13으로

기재되는 ZNF343-SNRPB-Hu-F1 유전자, 서열번호 14로 기재되는 ZNF343-SNRPB-Hu-F2 유전자, 서열번호 15로 기재되는 ZNF343-SNRPB-Hu-F3 유전자, 서열번호 16으로 기재되는 ZNF343-SNRPB-Hu-F4 유전자, 서열번호 17로 기재되는 ZNF343-SNRPB-Hu-F5 유전자, 서열번호 18로 기재되는 ZNF343-SNRPB-Hu-F6 유전자, 서열번호 19로 기재되는 ZNF343-SNRPB-Hu-F7 유전자, 서열번호 20으로 기재되는 ZNF343-SNRPB-Hu-F8 유전자, 서열번호 21로 기재되는 COX17-POPDC2-Hu-P1 유전자, 서열번호 22로 기재되는 COX17-POPDC2-Hu-P2 유전자, 서열번호 23으로 기재되는 COX17-POPDC2-Hu-P3 유전자, 서열번호 24로 기재되는 COX17-POPDC2-Hu-P4 유전자, 서열번호 25로 기재되는 COX17-POPDC2-Hu-P5 유전자, 서열번호 26으로 기재되는 COX17-POPDC2-Hu-P6 유전자, 서열번호 27로 기재되는 COX17-POPDC2-Hu-P7 유전자, 서열번호 28로 기재되는 COX17-POPDC2-Hu-P8 유전자, 서열번호 29로 기재되는 COX17-POPDC2-Hu-P9 유전자, 서열번호 30으로 기재되는 COX17-POPDC2-Hu-P10 유전자, 서열번호 31로 기재되는 COX17-POPDC2-Hu-P11 유전자, 서열번호 32로 기재되는 COX17-POPDC2-Hu-F1 유전자, 서열번호 33으로 기재되는 COX17-POPDC2-Hu-F2 유전자, 서열번호 34로 기재되는 COX17-POPDC2-Hu-F3 유전자, 서열번호 35로 기재되는 COX17-POPDC2-Hu-F4 유전자, 서열번호 36으로 기재되는 COX17-POPDC2-Hu-F5 유전자, 서열번호 37로 기재되는 COX17-POPDC2-Hu-F6 유전자, 서열번호 38로 기재되는 COX17-POPDC2-Hu-F7 유전자, 서열번호 39로 기재되는 COX17-POPDC2-Hu-F8 유전자, 서열번호 40으로 기재되는 COX17-POPDC2-Hu-F9 유전자, 서열번호 41로 기재되는 COX17-POPDC2-Hu-F10 유전자, 서열번호 42로 기재되는 COX17-POPDC2-Hu-F11 유전자, 서열번호 43으로 기재되는 COX17-POPDC2-Hu-F12 유전자, 서열번호 44로 기재되는 RBM7-REX02-Hu-P1 유전자, 서열번호 45로 기재되는 RBM7-REX02-Hu-P2 유전자, 서열번호 46으로 기재되는 RBM7-REX02-Hu-P3 유전자, 서열번호 47로 기재되는 RBM7-REX02-Hu-P4 유전자, 서열번호 48로 기재되는 PRKRIP1-ORAI2-Hu-P1 유전자, 서열번호 49로 기재되는 PRKRIP1-ORAI2-Hu-P2 유전자, 서열번호 50으로 기재되는 PRKRIP1-ORAI2-Hu-P3 유전자, 서열번호 51로 기재되는 PRKRIP1-ORAI2-Hu-P4 유전자, 서열번호 52로 기재되는 PRKRIP1-ORAI2-Hu-P5 유전자, 서열번호 53으로 기재되는 PRKRIP1-ORAI2-Hu-P6 유전자, 서열번호 54로 기재되는 ATP5F1-C1orf162-Hu-P1 유전자, 서열번호 55로 기재되는 ATP5F1-C1orf162-Hu-P2 유전자, 서열번호 56으로 기재되는 ATP5F1-C1orf162-Hu-P3 유전자 및 서열번호 57로 기재되는 ATP5F1-C1orf162-Hu-P4 유전자.

[0085] 또한, 본 발명은 상기 DNA 마이크로어레이를 포함하는 암 진단용 키트를 제공한다.

[0086] 또한, 본 발명은 하기로 구성된 군으로부터 선택되는 어느 하나의 유전자를 특이적으로 증폭시키는 프라이머 쌍을 포함하는 암 진단용 키트를 제공한다: 서열번호 1로 기재되는 ZNF343-SNRPB-Hu-P1 유전자, 서열번호 2로 기재되는 ZNF343-SNRPB-Hu-P2 유전자, 서열번호 3으로 기재되는 ZNF343-SNRPB-Hu-P3 유전자, 서열번호 4로 기재되는 ZNF343-SNRPB-Hu-P4 유전자, 서열번호 5로 기재되는 ZNF343-SNRPB-Hu-P5 유전자, 서열번호 6으로 기재되는 ZNF343-SNRPB-Hu-P6 유전자, 서열번호 7로 기재되는 ZNF343-SNRPB-Hu-P7 유전자, 서열번호 8로 기재되는 ZNF343-SNRPB-Hu-P8 유전자, 서열번호 9로 기재되는 ZNF343-SNRPB-Hu-P9 유전자, 서열번호 10으로 기재되는 ZNF343-SNRPB-Hu-P10 유전자, 서열번호 11로 기재되는 ZNF343-SNRPB-Hu-P11 유전자, 서열번호 12로 기재되는 ZNF343-SNRPB-Hu-P12 유전자, 서열번호 13으로 기재되는 ZNF343-SNRPB-Hu-F1 유전자, 서열번호 14로 기재되는 ZNF343-SNRPB-Hu-F2 유전자, 서열번호 15로 기재되는 ZNF343-SNRPB-Hu-F3 유전자, 서열번호 16으로 기재되는 ZNF343-SNRPB-Hu-F4 유전자, 서열번호 17로 기재되는 ZNF343-SNRPB-Hu-F5 유전자, 서열번호 18로 기재되는 ZNF343-SNRPB-Hu-F6 유전자, 서열번호 19로 기재되는 ZNF343-SNRPB-Hu-F7 유전자, 서열번호 20으로 기재되는 ZNF343-SNRPB-Hu-F8 유전자, 서열번호 21로 기재되는 COX17-POPDC2-Hu-P1 유전자, 서열번호 22로 기재되는 COX17-POPDC2-Hu-P2 유전자, 서열번호 23으로 기재되는 COX17-POPDC2-Hu-P3 유전자, 서열번호 24로 기재되는 COX17-POPDC2-Hu-P4 유전자, 서열번호 25로 기재되는 COX17-POPDC2-Hu-P5 유전자, 서열번호 26으로 기재되는 COX17-POPDC2-Hu-P6 유전자, 서열번호 27로 기재되는 COX17-POPDC2-Hu-P7 유전자, 서열번호 28로 기재되는 COX17-POPDC2-Hu-P8 유전자, 서열번호 29로 기재되는 COX17-POPDC2-Hu-P9 유전자, 서열번호 30으로 기재되는 COX17-POPDC2-Hu-P10 유전자, 서열번호 31로 기재되는 COX17-POPDC2-Hu-P11 유전자, 서열번호 32로 기재되는 COX17-POPDC2-Hu-F1 유전자, 서열번호 33으로 기재되는 COX17-POPDC2-Hu-F2 유전자, 서열번호 34로 기재되는 COX17-POPDC2-Hu-F3 유전자, 서열번호 35로 기재되는 COX17-POPDC2-Hu-F4 유전자, 서열번호 36으로 기재되는 COX17-POPDC2-Hu-F5 유전자, 서열번호 37로 기재되는 COX17-POPDC2-Hu-F6 유전자, 서열번호 38로 기재되는 COX17-POPDC2-Hu-F7 유전자, 서열번호 39로 기재되는 COX17-POPDC2-Hu-F8 유전자, 서열번호 40으로 기재되는 COX17-POPDC2-Hu-F9 유전자, 서열번호 41로 기재되는 COX17-POPDC2-Hu-F10 유전자, 서열번호 42로 기재되는 COX17-POPDC2-Hu-F11 유전자, 서열번호 43으로 기재되는 COX17-POPDC2-Hu-F12 유전자, 서열번호 44로 기재되는 RBM7-REX02-Hu-P1 유전자, 서열번호 45로 기재되는 RBM7-REX02-Hu-P2 유전자, 서열번호 46으로 기재되는 RBM7-REX02-Hu-P3 유전자, 서열번호 47로 기재되는 RBM7-REX02-Hu-P4 유전자, 서열번호 48로 기재되는 PRKRIP1-

ORAI2-Hu-P1 유전자, 서열번호 49로 기재되는 PRKRIP1-ORAI2-Hu-P2 유전자, 서열번호 50으로 기재되는 PRKRIP1-ORAI2-Hu-P3 유전자, 서열번호 51로 기재되는 PRKRIP1-ORAI2-Hu-P4 유전자, 서열번호 52로 기재되는 PRKRIP1-ORAI2-Hu-P5 유전자, 서열번호 53으로 기재되는 PRKRIP1-ORAI2-Hu-P6 유전자, 서열번호 54로 기재되는 ATP5F1-C1orf162-Hu-P1 유전자, 서열번호 55로 기재되는 ATP5F1-C1orf162-Hu-P2 유전자, 서열번호 56으로 기재되는 ATP5F1-C1orf162-Hu-P3 유전자, 서열번호 57로 기재되는 ATP5F1-C1orf162-Hu-P4 유전자.

- [0087] 상기 암은 림프종, 전립선암, 자궁경부암, 유방암, 대장암, 신장암, 폐암, 난소암, 위암, 자궁암, 방광암, 식도암, 간암 및 갑상선암으로 이루어진 군으로부터 선택되는 어느 하나인 것이 바람직하나, 이에 한정하지 않는다.
- [0088] 상기 ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-P11 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-P5, COX17-POPDC2-Hu-P8 유전자, COX17-POPDC2-Hu-P9 유전자, COX17-POPDC2-Hu-P10 유전자, RBM7-REXO2-Hu-P1 유전자, 및 PRKRIP1-ORAI2-Hu-P1 유전자는 림프종에서 발현하고,
- [0089] COX17-POPDC2-Hu-P9 유전자, PRKRIP1-ORAI2-Hu-P1 유전자, KRIP1-ORAI2-Hu-P3 유전자, 및 ATP5F1-C1orf162-Hu-P3 유전자는 전립선암에서 발현하고,
- [0090] ZNF343-SNRPB-Hu-P1 유전자, ZNF343-SNRPB-Hu-P3 유전자, ZNF343-SNRPB-Hu-P4 유전자, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-P11 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-P10 유전자, PRKRIP1-ORAI2-Hu-P3 유전자, 및 ATP5F1-C1orf162-Hu-P4 유전자는 자궁경부암에서 발현하고,
- [0091] ZNF343-SNRPB-Hu-P8 유전자, COX17-POPDC2-Hu-P7 유전자, COX17-POPDC2-Hu-P11 유전자, COX17-POPDC2-Hu-F9 유전자, COX17-POPDC2-Hu-F11 유전자, COX17-POPDC2-Hu-F12 유전자, RBM7-REXO2-Hu-P4 유전자, PRKRIP1-ORAI2-Hu-P6 유전자 및 ATP5F1-C1orf162-Hu-P4 유전자는 유방암에서 발현하고,
- [0092] ZNF343-SNRPB-Hu-P5 유전자, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-F5 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-F7 유전자, COX17-POPDC2-Hu-F11 유전자 및 COX17-POPDC2-Hu-F12 유전자는 대장암에서 발현하고,
- [0093] ZNF343-SNRPB-Hu-P1 유전자, ZNF343-SNRPB-Hu-P2 유전자, ZNF343-SNRPB-Hu-P3 유전자, ZNF343-SNRPB-Hu-P4 유전자, ZNF343-SNRPB-Hu-P11 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-P4 유전자, COX17-POPDC2-Hu-P6 유전자, COX17-POPDC2-Hu-P8 유전자, COX17-POPDC2-Hu-F3 유전자, COX17-POPDC2-Hu-F4 유전자, COX17-POPDC2-Hu-F6 유전자, PRKRIP1-ORAI2-Hu-P5 유전자 및 ATP5F1-C1orf162-Hu-P3 유전자는 신장암에서 발현하고,
- [0094] ZNF343-SNRPB-Hu-P1 유전자, ZNF343-SNRPB-Hu-F5 유전자, ZNF343-SNRPB-Hu-F7 유전자, COX17-POPDC2-Hu-F1 유전자, COX17-POPDC2-Hu-F3 유전자 및 PRKRIP1-ORAI2-Hu-P6 유전자는 폐암에서 발현하고,
- [0095] ZNF343-SNRPB-Hu-P4 유전자, ZNF343-SNRPB-Hu-P11 유전자, ZNF343-SNRPB-Hu-P12 유전자, ZNF343-SNRPB-Hu-F7 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-F10 유전자, COX17-POPDC2-Hu-F12 유전자, PRKRIP1-ORAI2-Hu-P4 유전자, PRKRIP1-ORAI2-Hu-P6 유전자 및 ATP5F1-C1orf162-Hu-P1 유전자는 난소암에서 발현하고,
- [0096] ZNF343-SNRPB-Hu-P11 유전자, ZNF343-SNRPB-Hu-F6 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-F2 유전자 및 COX17-POPDC2-Hu-F10 유전자는 위암에서 발현하고,
- [0097] PRKRIP1-ORAI2-Hu-P2 유전자 및 PRKRIP1-ORAI2-Hu-P5 유전자는 자궁암에서 발현하고,
- [0098] ZNF343-SNRPB-Hu-P7 유전자, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-F2 유전자, COX17-POPDC2-Hu-P1 유전자, COX17-POPDC2-Hu-F1 유전자, COX17-POPDC2-Hu-F5 유전자 및 RBM7-REXO2-Hu-P3 유전자는 방광암에서 발현하고,
- [0099] ZNF343-SNRPB-Hu-P8 유전자는 식도암에서 발현하고,
- [0100] ZNF343-SNRPB-Hu-F7 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-P10 유전자, ATP5F1-C1orf162-Hu-P2 유전자, ATP5F1-C1orf162-Hu-P3 유전자 및 ATP5F1-C1orf162-Hu-P4 유전자는 간암에서 발현하고,
- [0101] ZNF343-SNRPB-Hu-P8 유전자는 갑상선암에서 발현하는 것이 바람직하나, 이에 한정하지 않는다.

- [0102] 최근의 연구들은 단일 유전자의 개별 변형된 스플라이싱 변이체가 인간에서 조직 분화, 질병 및 종양 발생을 측정하는데 중요한 역할을 할 수 있음을 밝혔고, 이것은 유전자로부터 유래된 다양한 변이체가 다양한 기능을 가질 수 있음을 의미한다. 이러한 관점에서, 다섯 개의 상기 융합 유전자 각각에 속하는 다양한 융합 유전자 전사 변이체(CGTV)가 다양한 발현 패턴을 나타내는지 여부를 조사하였다. 흥미롭게도, 다양한 종양 특이적 발현 패턴을 갖는 융합 유전자 전사 변이체에 상응하는 세 가지 융합 유전자를 확인하였다: *RBM-REXO2-Hu* (-P1 및 -P4), *PRKRIP1-ORAI2-Hu* (-P3 및 -P5) 및 *ATP5F1-C1orf162-Hu* (-P1 및 -P2). 특히, 종양 특이적 발현을 나타내는 두 개의 동족 변이체(*RBM-REXO2-Hu*)는 -P1 (림프종) 및 -P4 (유방암), 및 두 개의 또 다른 동족 변이체(*PRKRIP1-ORAI2-Hu*)는 -P3 (전립선 및 자궁경부) 및 -P5 (신장 및 자궁)이다. 이러한 결과는 단일 유전자 변이체에서 관찰된 것과 일치하게, 상기 동족 CGTV가 발암의 메카니즘에서 기능적 차이를 나타낼 수 있음을 제시하고 있다.
- [0103] 30종의 인간 정상 및 이에 상응하는 종양 조직에서 증폭된 전체 2,450개 PCR 산물 중에서 160개의 표적화된 융합 유전자 전사 변이체를 얻었고, 이러한 조직에서 융합 유전자 전사 변이체의 발현 분포를 확인하였다. 실험한 조직 중에서, 신장 정상 및 종양 조직에서 가장 높은 수의 융합 유전자 전사 변이체가 검출되었다(160개 중 19%). 평균적으로, 정상 조직에서보다 종양 조직에서 융합 유전자 전사 변이체의 발현이 더욱 특이적이다; 종양 조직에서 115개(72%)가 발현된 반면, 정상 조직에서는 160개 중 45개(28%)의 융합 유전자 전사 변이체가 발현되었다(표 2 참조). 특히, 상기 종양-편향된 경향성은 방광암(92%), 유방암(85%), 대장암(82%) 및 난소암(80%) 조직에서 더욱 뚜렷하게 나타났다. 이러한 결과는 융합 유전자 전사 변이체가 종양-편향된 발현에 강한 경향성을 가짐을 의미한다. 이러한 경향성 더 많은 증거는 세 가지 정상 및 이에 상응하는 종양 조직(림프, 자궁경부 및 폐) 간 분포 빈도의 비교를 통해 확인할 수 있다. 비록 그들 중 몇몇이 몇 가지 다른 정상 조직에서도 발현됨에도 불구하고, 융합 유전자 전사 변이체(CGTV)는 이러한 세 가지 종양 조직(림프: 정상에서 0개 융합 유전자 전사 변이체 대 종양에서 9개 융합 유전자 전사 변이체; 자궁경부: 0개 대 10개; 폐: 0개 대 6개)에서만 관찰되었다. 이러한 발견은 본 발명자들이 깊이 있게 연구한 다섯 개의 융합 유전자가 구조적으로 발암에 있어 중요한 역할을 할 것이라는 것을 제시한다.
- [0104] 추가로, 본 발명자들이 깊이있게 연구한 57개의 융합 유전자 전사 변이체(CGTV)가 종양 특이적으로 발현되는지의 여부를 조사하였다. 이들 중에서, 12개의 융합 유전자 전사 변이체(CGTV)가 한 가지 조직 형태의 정상 및 이에 상응한 종양 샘플에서 동시에 발현되었다. 이러한 12개의 융합 유전자 전사 변이체(CGTV) 중에서, 세 개(*ZNF343-SNRPB-Hu-P1* (5/5), -P8 (5/13) 및 *ATP5F1-C1orf162-Hu-P3* (6/8))는 다양한 형태의 정상 및 종양 조직에서 가장 일반적으로 발현하는 것으로 나타났다.
- [0105] 본 발명의 결과가 융합 유전자 전사 변이체(VGTV)가 인간 조직에서 강한 경향성으로 종양-편향적으로 발현되어 나타남에도 불구하고, 이러한 융합 유전자 전사 변이체(CGTV)가 암과 연관되어 있는지의 여부는 알려져 있지 않다. 따라서, 본 발명에서 정의된 57개의 다양한 CGTV가 종양 특이적으로 발현되는지의 여부를 조사하였다. 놀랍게도, 35개의 다양한 융합 유전자 전사 변이체(CGTV)가 종양 조직에서만 독점적으로 발현되고, 이러한 분획은 57개 융합 유전자 전사 변이체에서 61.4%로 나타났다. 이러한 35개의 융합 유전자 전사 변이체 중에서, 13개는 다양한 종양 조직에서 발현된 것과 비교하여, 22개는 하나의 조직에서만 종양 특이적으로 발현되어 나타났다. 특히, 이는 융합 유전자 전사 변이체(CGTV) 중 37%가 상응하는 정상 조직에서는 발현되지 않지만, 다양한 종양 조직에서 특이적으로 발현됨을 의미한다. 이러한 융합 유전자 전사 변이체(CGTV) 중에서, 9개(*ZNF343-SNRPB-Hu-P3*, *COX17-POPCD2-Hu-P8-9*, *COX17-POPCD2-Hu-F1*, -F3, -F10, -F11, *PRKRIP1-ORAI2-Hu-P3* 및 -P5), 세 개(*ZNF343-SNRPB-Hu-P4*, *COX17-POPCD2-Hu-P10* 및 *PRKRIP1-ORAI2-Hu-P6*) 및 한 개(*ZNF343-SNRPB-Hu-F8*)는 각각 두 가지 종양-, 세 가지 종양-, 네 가지 종양 특이적 발현을 나타내었다(표 2 참조). *ZNF343-SNRPB-Hu-F8*의 특이적 발현을 보이는 네 가지 종양 조직은 신장, 난소, 위 및 간이다. 이러한 결과는 융합 유전자가 발암과 구조적으로 연관되고, 각 융합 유전자가 암종에 따라 특이적으로 발현됨으로, 본 발명의 신규한 인간 융합 유전자의 전사 변이체들은 암 진단을 위한 DNA 마이크로어레이 또는 키트에 유용하게 이용될 수 있다.
- [0106] 또한, 본 발명은 하기의 단계를 포함하는 개체 내 암 진단의 정보를 제공하기 위한 유전자 발현 여부를 확인하는 방법을 제공한다:
- [0107] 1) 개체로부터 시료를 분리하는 단계;
- [0108] 2) 상기 분리된 시료로부터 하기로 구성된 군으로부터 선택되는 어느 하나의 유전자의 발현 여부를 확인하는 단



계:

- [0109] 서열번호 1로 기재되는 ZNF343-SNRPB-Hu-P1 유전자, 서열번호 2로 기재되는 ZNF343-SNRPB-Hu-P2 유전자, 서열번호 3으로 기재되는 ZNF343-SNRPB-Hu-P3 유전자, 서열번호 4로 기재되는 ZNF343-SNRPB-Hu-P4 유전자, 서열번호 5로 기재되는 ZNF343-SNRPB-Hu-P5 유전자, 서열번호 6으로 기재되는 ZNF343-SNRPB-Hu-P6 유전자, 서열번호 7로 기재되는 ZNF343-SNRPB-Hu-P7 유전자, 서열번호 8로 기재되는 ZNF343-SNRPB-Hu-P8 유전자, 서열번호 9로 기재되는 ZNF343-SNRPB-Hu-P9 유전자, 서열번호 10으로 기재되는 ZNF343-SNRPB-Hu-P10 유전자, 서열번호 11로 기재되는 ZNF343-SNRPB-Hu-P11 유전자, 서열번호 12로 기재되는 ZNF343-SNRPB-Hu-P12 유전자, 서열번호 13으로 기재되는 ZNF343-SNRPB-Hu-F1 유전자, 서열번호 14로 기재되는 ZNF343-SNRPB-Hu-F2 유전자, 서열번호 15로 기재되는 ZNF343-SNRPB-Hu-F3 유전자, 서열번호 16으로 기재되는 ZNF343-SNRPB-Hu-F4 유전자, 서열번호 17로 기재되는 ZNF343-SNRPB-Hu-F5 유전자, 서열번호 18로 기재되는 ZNF343-SNRPB-Hu-F6 유전자, 서열번호 19로 기재되는 ZNF343-SNRPB-Hu-F7 유전자, 서열번호 20으로 기재되는 ZNF343-SNRPB-Hu-F8 유전자, 서열번호 21로 기재되는 COX17-POPDC2-Hu-P1 유전자, 서열번호 22로 기재되는 COX17-POPDC2-Hu-P2 유전자, 서열번호 23으로 기재되는 COX17-POPDC2-Hu-P3 유전자, 서열번호 24로 기재되는 COX17-POPDC2-Hu-P4 유전자, 서열번호 25로 기재되는 COX17-POPDC2-Hu-P5 유전자, 서열번호 26으로 기재되는 COX17-POPDC2-Hu-P6 유전자, 서열번호 27로 기재되는 COX17-POPDC2-Hu-P7 유전자, 서열번호 28로 기재되는 COX17-POPDC2-Hu-P8 유전자, 서열번호 29로 기재되는 COX17-POPDC2-Hu-P9 유전자, 서열번호 30으로 기재되는 COX17-POPDC2-Hu-P10 유전자, 서열번호 31로 기재되는 COX17-POPDC2-Hu-P11 유전자, 서열번호 32로 기재되는 COX17-POPDC2-Hu-F1 유전자, 서열번호 33으로 기재되는 COX17-POPDC2-Hu-F2 유전자, 서열번호 34로 기재되는 COX17-POPDC2-Hu-F3 유전자, 서열번호 35로 기재되는 COX17-POPDC2-Hu-F4 유전자, 서열번호 36으로 기재되는 COX17-POPDC2-Hu-F5 유전자, 서열번호 37로 기재되는 COX17-POPDC2-Hu-F6 유전자, 서열번호 38로 기재되는 COX17-POPDC2-Hu-F7 유전자, 서열번호 39로 기재되는 COX17-POPDC2-Hu-F8 유전자, 서열번호 40으로 기재되는 COX17-POPDC2-Hu-F9 유전자, 서열번호 41로 기재되는 COX17-POPDC2-Hu-F10 유전자, 서열번호 42로 기재되는 COX17-POPDC2-Hu-F11 유전자, 서열번호 43으로 기재되는 COX17-POPDC2-Hu-F12 유전자, 서열번호 44로 기재되는 RBM7-REX02-Hu-P1 유전자, 서열번호 45로 기재되는 RBM7-REX02-Hu-P2 유전자, 서열번호 46으로 기재되는 RBM7-REX02-Hu-P3 유전자, 서열번호 47로 기재되는 RBM7-REX02-Hu-P4 유전자, 서열번호 48로 기재되는 PRKRIP1-ORAI2-Hu-P1 유전자, 서열번호 49로 기재되는 PRKRIP1-ORAI2-Hu-P2 유전자, 서열번호 50으로 기재되는 PRKRIP1-ORAI2-Hu-P3 유전자, 서열번호 51로 기재되는 PRKRIP1-ORAI2-Hu-P4 유전자, 서열번호 52로 기재되는 PRKRIP1-ORAI2-Hu-P5 유전자, 서열번호 53으로 기재되는 PRKRIP1-ORAI2-Hu-P6 유전자, 서열번호 54로 기재되는 ATP5F1-C1orf162-Hu-P1 유전자, 서열번호 55로 기재되는 ATP5F1-C1orf162-Hu-P2 유전자, 서열번호 56으로 기재되는 ATP5F1-C1orf162-Hu-P3 유전자 및 서열번호 57로 기재되는 ATP5F1-C1orf162-Hu-P4 유전자; 및
- [0110] 3) 상기 유전자의 발현이 확인되는 개체를 암 위험이 있는 개체로 판정하되,
- [0111] 이때,
- [0112] ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-P11 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-P5 유전자, COX17-POPDC2-Hu-P8 유전자, COX17-POPDC2-Hu-P9 유전자, COX17-POPDC2-Hu-P10 유전자, RBM7-REX02-Hu-P1 유전자, 및 PRKRIP1-ORAI2-Hu-P1 유전자가 발현하면 상기 암을 림프종으로 판정하고,
- [0113] COX17-POPDC2-Hu-P9 유전자, PRKRIP1-ORAI2-Hu-P1 유전자, KRIP1-ORAI2-Hu-P3 유전자, 및 ATP5F1-C1orf162-Hu-P3 유전자가 발현하면 상기 암을 전립선암으로 판정하고,
- [0114] ZNF343-SNRPB-Hu-P1 유전자, ZNF343-SNRPB-Hu-P3 유전자, ZNF343-SNRPB-Hu-P4 유전자, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-P11 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-P10 유전자, PRKRIP1-ORAI2-Hu-P3 유전자, 및 ATP5F1-C1orf162-Hu-P4 유전자가 발현하면 상기 암을 자궁경부암으로 판정하고,
- [0115] ZNF343-SNRPB-Hu-P8 유전자, COX17-POPDC2-Hu-P7 유전자, COX17-POPDC2-Hu-P11 유전자, COX17-POPDC2-Hu-F9 유전자, COX17-POPDC2-Hu-F11 유전자, COX17-POPDC2-Hu-F12 유전자, RBM7-REX02-Hu-P4 유전자, PRKRIP1-ORAI2-Hu-P6 유전자 및 ATP5F1-C1orf162-Hu-P4 유전자가 발현하면 상기 암을 유방암으로 판정하고,
- [0116] ZNF343-SNRPB-Hu-P5 유전자, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-F5 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-F7 유전자, COX17-POPDC2-Hu-F11 유전자 및 COX17-POPDC2-Hu-F12 유전자가 발현하면 상기 암은 대장암으로 판정하고,
- [0117] ZNF343-SNRPB-Hu-P1 유전자, ZNF343-SNRPB-Hu-P2 유전자, ZNF343-SNRPB-Hu-P3 유전자, ZNF343-SNRPB-Hu-P4 유

전자, ZNF343-SNRPB-Hu-P11 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-P4 유전자, COX17-POPDC2-Hu-P6 유전자, COX17-POPDC2-Hu-P8 유전자, COX17-POPDC2-Hu-F3 유전자, COX17-POPDC2-Hu-F4 유전자, COX17-POPDC2-Hu-F6 유전자, PRKRIP1-ORAI2-Hu-P5 유전자 및 ATP5F1-C1orf162-Hu-P3 유전자가 발현하면 상기 암은 신장암으로 판정하고,

- [0118] ZNF343-SNRPB-Hu-P1 유전자, ZNF343-SNRPB-Hu-F5 유전자, ZNF343-SNRPB-Hu-F7 유전자, COX17-POPDC2-Hu-F1 유전자, COX17-POPDC2-Hu-F3 유전자 및 PRKRIP1-ORAI2-Hu-P6 유전자가 발현하면 상기 암은 폐암으로 판정하고,
- [0119] ZNF343-SNRPB-Hu-P4 유전자, ZNF343-SNRPB-Hu-P11 유전자, ZNF343-SNRPB-Hu-P12 유전자, ZNF343-SNRPB-Hu-F7 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-F10 유전자, COX17-POPDC2-Hu-F12 유전자, PRKRIP1-ORAI2-Hu-P4 유전자, PRKRIP1-ORAI2-Hu-P6 유전자 및 ATP5F1-C1orf162-Hu-P1 유전자가 발현하면 상기 암은 난소암으로 판정하고,
- [0120] ZNF343-SNRPB-Hu-P11 유전자, ZNF343-SNRPB-Hu-F6 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-F2 유전자 및 COX17-POPDC2-Hu-F10 유전자가 발현하면 상기 암은 위암으로 판정하고,
- [0121] PRKRIP1-ORAI2-Hu-P2 유전자 및 PRKRIP1-ORAI2-Hu-P5 유전자가 발현하면 상기 암은 자궁암으로 판정하고,
- [0122] ZNF343-SNRPB-Hu-P7 유전자, ZNF343-SNRPB-Hu-P8 유전자, ZNF343-SNRPB-Hu-F2 유전자, COX17-POPDC2-Hu-P1 유전자, COX17-POPDC2-Hu-F1 유전자, COX17-POPDC2-Hu-F5 유전자 및 RBM7-REX02-Hu-P3 유전자가 발현하면 상기 암은 방광암으로 판정하고,
- [0123] ZNF343-SNRPB-Hu-P8 유전자가 발현하면 상기 암은 식도암으로 판정하고,
- [0124] ZNF343-SNRPB-Hu-F7 유전자, ZNF343-SNRPB-Hu-F8 유전자, COX17-POPDC2-Hu-P3 유전자, COX17-POPDC2-Hu-P10 유전자, ATP5F1-C1orf162-Hu-P2 유전자, ATP5F1-C1orf162-Hu-P3 유전자 및 ATP5F1-C1orf162-Hu-P4 유전자가 발현하면 상기 암은 간암으로 판정하고,
- [0125] ZNF343-SNRPB-Hu-P8 유전자가 발현하면 상기 암은 갑상선암으로 판정하는 단계.
- [0126] 상기 단계 1)의 시료는 조직, 세포, 전혈, 혈장, 혈청, 타액 및 뇨로 이루어진 군으로부터 선택되는 어느 하나인 것이 바람직하나, 이에 한정하지 않는다.
- [0127] 상기 단계 2)의 유전자 발현은 RT-PCR, 경쟁적 RT-PCR, 실시간 RT-PCR, RNase 보호 분석법(RNase protection assay), 노던블랏팅(Northern blotting) 및 DNA 칩 중에서 선택되는 어느 하나의 방법을 이용하여 확인하는 것이 바람직하나, 이에 한정하지 않는다.
- [0128] 아울러, 본 발명은 하기의 단계를 포함하는 개체 내 암 진단의 정보를 제공하기 위한 단백질의 발현 여부를 확인하는 방법을 제공한다:
- [0129] 1) 개체로부터 시료를 분리하는 단계;
- [0130] 2) 상기 분리된 시료로부터 하기로 구성된 군으로부터 선택되는 어느 하나의 유전자에 의해 코딩되는 단백질의 발현 여부를 확인하는 단계;
- [0131] 서열번호 1로 기재되는 ZNF343-SNRPB-Hu-P1 유전자, 서열번호 2로 기재되는 ZNF343-SNRPB-Hu-P2 유전자, 서열번호 3으로 기재되는 ZNF343-SNRPB-Hu-P3 유전자, 서열번호 4로 기재되는 ZNF343-SNRPB-Hu-P4 유전자, 서열번호 5로 기재되는 ZNF343-SNRPB-Hu-P5 유전자, 서열번호 6으로 기재되는 ZNF343-SNRPB-Hu-P6 유전자, 서열번호 7로 기재되는 ZNF343-SNRPB-Hu-P7 유전자, 서열번호 8로 기재되는 ZNF343-SNRPB-Hu-P8 유전자, 서열번호 9로 기재되는 ZNF343-SNRPB-Hu-P9 유전자, 서열번호 10으로 기재되는 ZNF343-SNRPB-Hu-P10 유전자, 서열번호 11로 기재되는 ZNF343-SNRPB-Hu-P11 유전자, 서열번호 12로 기재되는 ZNF343-SNRPB-Hu-P12 유전자, 서열번호 13으로 기재되는 ZNF343-SNRPB-Hu-F1 유전자, 서열번호 14로 기재되는 ZNF343-SNRPB-Hu-F2 유전자, 서열번호 15로 기재되는 ZNF343-SNRPB-Hu-F3 유전자, 서열번호 16으로 기재되는 ZNF343-SNRPB-Hu-F4 유전자, 서열번호 17로 기재되는 ZNF343-SNRPB-Hu-F5 유전자, 서열번호 18로 기재되는 ZNF343-SNRPB-Hu-F6 유전자, 서열번호 19로 기재되는 ZNF343-SNRPB-Hu-F7 유전자, 서열번호 20으로 기재되는 ZNF343-SNRPB-Hu-F8 유전자, 서열번호 21로 기재되는 COX17-POPDC2-Hu-P1 유전자, 서열번호 22로 기재되는 COX17-POPDC2-Hu-P2 유전자, 서열번호 23으로 기재되는 COX17-POPDC2-Hu-P3 유전자, 서열번호 24로 기재되는 COX17-POPDC2-Hu-P4 유전자, 서열번호 25로 기재되는

는 COX17-POPDC2-Hu-P5 유전자, 서열번호 26으로 기재되는 COX17-POPDC2-Hu-P6 유전자, 서열번호 27로 기재되는 COX17-POPDC2-Hu-P7 유전자, 서열번호 28로 기재되는 COX17-POPDC2-Hu-P8 유전자, 서열번호 29로 기재되는 COX17-POPDC2-Hu-P9 유전자, 서열번호 30으로 기재되는 COX17-POPDC2-Hu-P10 유전자, 서열번호 31로 기재되는 COX17-POPDC2-Hu-P11 유전자, 서열번호 32로 기재되는 COX17-POPDC2-Hu-F1 유전자, 서열번호 33으로 기재되는 COX17-POPDC2-Hu-F2 유전자, 서열번호 34로 기재되는 COX17-POPDC2-Hu-F3 유전자, 서열번호 35로 기재되는 COX17-POPDC2-Hu-F4 유전자, 서열번호 36으로 기재되는 COX17-POPDC2-Hu-F5 유전자, 서열번호 37로 기재되는 COX17-POPDC2-Hu-F6 유전자, 서열번호 38로 기재되는 COX17-POPDC2-Hu-F7 유전자, 서열번호 39로 기재되는 COX17-POPDC2-Hu-F8 유전자, 서열번호 40으로 기재되는 COX17-POPDC2-Hu-F9 유전자, 서열번호 41로 기재되는 COX17-POPDC2-Hu-F10 유전자, 서열번호 42로 기재되는 COX17-POPDC2-Hu-F11 유전자, 서열번호 43으로 기재되는 COX17-POPDC2-Hu-F12 유전자, 서열번호 44로 기재되는 RBM7-REX02-Hu-P1 유전자, 서열번호 45로 기재되는 RBM7-REX02-Hu-P2 유전자, 서열번호 46으로 기재되는 RBM7-REX02-Hu-P3 유전자, 서열번호 47로 기재되는 RBM7-REX02-Hu-P4 유전자, 서열번호 48로 기재되는 PRKRIP1-ORAI2-Hu-P1 유전자, 서열번호 49로 기재되는 PRKRIP1-ORAI2-Hu-P2 유전자, 서열번호 50으로 기재되는 PRKRIP1-ORAI2-Hu-P3 유전자, 서열번호 51로 기재되는 PRKRIP1-ORAI2-Hu-P4 유전자, 서열번호 52로 기재되는 PRKRIP1-ORAI2-Hu-P5 유전자, 서열번호 53으로 기재되는 PRKRIP1-ORAI2-Hu-P6 유전자, 서열번호 54로 기재되는 ATP5F1-C1orf162-Hu-P1 유전자, 서열번호 55로 기재되는 ATP5F1-C1orf162-Hu-P2 유전자, 서열번호 56으로 기재되는 ATP5F1-C1orf162-Hu-P3 유전자 및 서열번호 57로 기재되는 ATP5F1-C1orf162-Hu-P4 유전자; 및

- [0132] 3) 상기 단백질의 발현이 확인되는 개체를 암 위험이 있는 개체로 판정하되,
- [0133] 이때,
- [0134] ZNF343-SNRPB-Hu-P8 단백질, ZNF343-SNRPB-Hu-P11 단백질, COX17-POPDC2-Hu-P3 단백질, COX17-POPDC2-Hu-P5 단백질, COX17-POPDC2-Hu-P8 단백질, COX17-POPDC2-Hu-P9 단백질, COX17-POPDC2-Hu-P10 단백질, RBM7-REX02-Hu-P1 단백질, 및 PRKRIP1-ORAI2-Hu-P1 단백질이 발현하면 상기 암을 림프종으로 판정하고,
- [0135] COX17-POPDC2-Hu-P9 단백질, PRKRIP1-ORAI2-Hu-P1 단백질, KRIP1-ORAI2-Hu-P3 단백질, 및 ATP5F1-C1orf162-Hu-P3 단백질이 발현하면 상기 암을 전립선암으로 판정하고,
- [0136] ZNF343-SNRPB-Hu-P1 단백질, ZNF343-SNRPB-Hu-P3 단백질, ZNF343-SNRPB-Hu-P4 단백질, ZNF343-SNRPB-Hu-P8 단백질, ZNF343-SNRPB-Hu-P11 단백질, COX17-POPDC2-Hu-P3 단백질, COX17-POPDC2-Hu-P10 단백질, PRKRIP1-ORAI2-Hu-P3 단백질, 및 ATP5F1-C1orf162-Hu-P4 단백질이 발현하면 상기 암을 자궁경부암으로 판정하고,
- [0137] ZNF343-SNRPB-Hu-P8 단백질, COX17-POPDC2-Hu-P7 단백질, COX17-POPDC2-Hu-P11 단백질, COX17-POPDC2-Hu-F9 단백질, COX17-POPDC2-Hu-F11 단백질, COX17-POPDC2-Hu-F12 단백질, RBM7-REX02-Hu-P4 단백질, PRKRIP1-ORAI2-Hu-P6 단백질 및 ATP5F1-C1orf162-Hu-P4 단백질이 발현하면 상기 암을 유방암으로 판정하고,
- [0138] ZNF343-SNRPB-Hu-P5 단백질, ZNF343-SNRPB-Hu-P8 단백질, ZNF343-SNRPB-Hu-F5 단백질, COX17-POPDC2-Hu-P3 단백질, COX17-POPDC2-Hu-F7 단백질, COX17-POPDC2-Hu-F11 단백질 및 COX17-POPDC2-Hu-F12 단백질이 발현하면 상기 암은 대장암으로 판정하고,
- [0139] ZNF343-SNRPB-Hu-P1 단백질, ZNF343-SNRPB-Hu-P2 단백질, ZNF343-SNRPB-Hu-P3 단백질, ZNF343-SNRPB-Hu-P4 단백질, ZNF343-SNRPB-Hu-P11 단백질, ZNF343-SNRPB-Hu-F8 단백질, COX17-POPDC2-Hu-P4 단백질, COX17-POPDC2-Hu-P6 단백질, COX17-POPDC2-Hu-P8 단백질, COX17-POPDC2-Hu-F3 단백질, COX17-POPDC2-Hu-F4 단백질, COX17-POPDC2-Hu-F6 단백질, PRKRIP1-ORAI2-Hu-P5 단백질 및 ATP5F1-C1orf162-Hu-P3 단백질이 발현하면 상기 암은 신장암으로 판정하고,
- [0140] ZNF343-SNRPB-Hu-P1 단백질, ZNF343-SNRPB-Hu-F5 단백질, ZNF343-SNRPB-Hu-F7 단백질, COX17-POPDC2-Hu-F1 단백질, COX17-POPDC2-Hu-F3 단백질 및 PRKRIP1-ORAI2-Hu-P6 단백질이 발현하면 상기 암은 폐암으로 판정하고,
- [0141] ZNF343-SNRPB-Hu-P4 단백질, ZNF343-SNRPB-Hu-P11 단백질, ZNF343-SNRPB-Hu-P12 단백질, ZNF343-SNRPB-Hu-F7 단백질, ZNF343-SNRPB-Hu-F8 단백질, COX17-POPDC2-Hu-F10 단백질, COX17-POPDC2-Hu-F12 단백질, PRKRIP1-ORAI2-Hu-P4 단백질, PRKRIP1-ORAI2-Hu-P6 단백질 및 ATP5F1-C1orf162-Hu-P1 단백질이 발현하면 상기 암은 난소암으로 판정하고,
- [0142] ZNF343-SNRPB-Hu-P11 단백질, ZNF343-SNRPB-Hu-F6 단백질, ZNF343-SNRPB-Hu-F8 단백질, COX17-POPDC2-Hu-F2 단백질 및 COX17-POPDC2-Hu-F10 단백질이 발현하면 상기 암은 위암으로 판정하고,

- [0143] PRKRIP1-ORAI2-Hu-P2 단백질 및 PRKRIP1-ORAI2-Hu-P5 단백질이 발현하면 상기 암은 자궁암으로 판정하고,
- [0144] ZNF343-SNRPB-Hu-P7 단백질, ZNF343-SNRPB-Hu-P8 단백질, ZNF343-SNRPB-Hu-F2 단백질, COX17-POPDC2-Hu-P1 단백질, COX17-POPDC2-Hu-F1 단백질, COX17-POPDC2-Hu-F5 단백질 및 RBM7-REX02-Hu-P3 단백질이 발현하면 상기 암은 방광암으로 판정하고,
- [0145] ZNF343-SNRPB-Hu-P8 단백질이 발현하면 상기 암은 식도암으로 판정하고,
- [0146] ZNF343-SNRPB-Hu-F7 단백질, ZNF343-SNRPB-Hu-F8 단백질, COX17-POPDC2-Hu-P3 단백질, COX17-POPDC2-Hu-P10 단백질, ATP5F1-C1orf162-Hu-P2 단백질, ATP5F1-C1orf162-Hu-P3 단백질 및 ATP5F1-C1orf162-Hu-P4 단백질이 발현하면 상기 암은 간암으로 판정하고,
- [0147] ZNF343-SNRPB-Hu-P8 단백질이 발현하면 상기 암은 갑상선암으로 판정하는 단계.

[0148] 본 발명의 신규한 5개의 인간 융합유전자(*ZNF343-SNRPB-Hu*, *COX17-POPDC2-Hu*, *RBM7-REX02-Hu*, *PRKRIP1-ORAI2-Hu*, *ATP5F1-C1orf162-Hu*)와 이의 57개의 인간 융합유전자 전사변이체(CGTV)에 대한 시퀀싱(염기서열분석) 및 생물정보학적 해석들을 수행하여 융합 유전자의 형성 기전을 확인하였으며, 인간 종양 조직에서 특이적으로 발현되는 CGTV를 동정함으로써, CGTV와 암 발생과의 연관성을 확인한 바, 본 발명에 따른 유전자 또는 상기 유전자에 의해 코딩되는 단백질의 발현 여부 확인을 통해 다양한 암 종을 진단할 수 있다.

[0149] 이하, 본 발명을 실시예에 의해 상세히 설명한다.

[0150] 단, 하기 실시예는 본 발명을 예시하는 것일 뿐, 본 발명의 내용이 하기 실시예에 의해 한정되는 것은 아니다.

[0151] <실시예 1> 본 발명의 신규 융합 유전자의 발굴

[0152] <1-1> 전체 RNA 준비

[0153] 인간 정상 조직 및 이에 상응하는 종양 조직 30 종류(립프질, 전립선, 정소, 자궁경부, 유방, 대장, 신장, 폐, 난소, 위, 자궁, 방광, 식도, 간, 및 갑상선)로부터 분리한 전체 RNA는 각각 Ambion 및 Clontech에서 구입하여 이용하였다.

[0154] <1-2> 역전사 중합효소 연쇄반응(RT-PCR)

[0155] 이전 연구에서 얻어진 1,933개의 고품질의 전장 칩앤지 정소 cDNA 서열을 통해 14개의 신규 칩앤지 융합 유전자를 발굴하였고, 상기 칩앤지의 융합 유전자와 인간 융합 유전자의 상동성을 확인하기 위하여, 인간 정상 조직 및 이에 상응하는 종양 조직 30종의 전체 RNA를 이용하여 역전사 중합효소 연쇄 반응(RT-PCR)을 수행하였다.

[0156] 구체적으로, RT-PCR을 위한 제 1가닥 cRNA는 올리고-dT 프라이머를 이용한 Sensiscript® RT 키트로 제조하였다. 30종류의 인간 정상 조직 및 이에 상응하는 종양 조직에서 각 융합 유전자 *ZNF343-SNRPB-Hu*, *COX17-POPDC2-Hu*, *RBM7-REX02-Hu*, *PRKRIP1-ORAI2-Hu* 및 *ATP5F1-C1orf162-Hu*의 부분적 또는 전장 CDS를 갖는 전사 변이체의 동정을 위한 RT-PCR은 하기 [표 1]에 기재된 프라이머를 이용하여 수행하였다. 상기 PCR은 극도로 낮게 발현되는 융합 유전자를 검출하기 위하여 43회의 증폭 사이클로 일반적 PCR 조건(94° C에서 5분간 초기방치; 94° C에서 30초간 DNA 변성단계, 60° C에서 30초간 어닐링단계, 72° C에서 1분간 신장 단계: DNA 변성 단계로부터 신장 단계까지 43회 사이클 반복; 72° C에서 7분간 최종 방치단계)하에서 수행되었다.

표 1

프라이머 리스트

| No. | CG name                   | CG Type  | Primer name | Primer sequence       |
|-----|---------------------------|----------|-------------|-----------------------|
| 1   | <i>ZNF343-SNRPB-Hu</i>    | Full-CDS | ZS-FF       | TTCTGTGCTCCCCGAGAACTT |
|     |                           |          | ZS-FR       | CAGTTTCATAGGCCACAAGGA |
|     |                           | Partial  | ZS-PF       | GAAAGACTTCTGTGCTCCCG  |
|     |                           |          | ZS-PR       | TGGGAGGAGGTCCCTCTACT  |
| 2   | <i>COX17-POPDC2-Hu</i>    | Full-CDS | CP-FF       | GAAGTGAAGTGCAGCAATC   |
|     |                           |          | CP-FR       | GGCATGTGCTAGGTGCTTAC  |
|     |                           | Partial  | CP-PF       | GACATGGTAGAGGCTGGGAA  |
|     |                           |          | CP-PR       | TCTAATTGAGGCCCAAGG    |
| 3   | <i>RBM7-REXO2-Hu</i>      | Partial  | RR-PF       | ATCAGGAAGTAGTCATGCCC  |
|     |                           |          | RR-PR       | TACCAGCGTCTGCACAGTTC  |
| 4   | <i>PRKRIP1-ORAI2-Hu</i>   | Partial  | PO-PF       | AATGAGTGAATGGGCACCTC  |
|     |                           |          | PO-PR       | GGAAGTTGATCCAGCAGAGC  |
| 5   | <i>ATP5F1-C1orf162-Hu</i> | Partial  | AC-PF       | AGGCTTCCATCCAACACATC  |
|     |                           |          | AC-PR       | ACATGTGGAGCCATTTCCTC  |

[0157]

[0158]

그 결과, 43회의 PCR 싸이클 횟수에서 이후 실험에 이용하기 위해 충분한 양인 2,450 개의 PCR 산물을 얻었고, 이를 클로닝하고 서열을 분석하였다.

[0159]

<실시예 2> 본 발명의 신규 융합 유전자의 서열 분석

[0160]

<2-1> 클로닝 및 시퀀싱(sequencing)

[0161]

겔 추출에 의해 회수된 상기 RT-PCR 산물은 4℃에서 하룻밤 동안 10 ul의 부피로 라이게이션(ligation)하여 RBC T&A 클로닝 벡터 내로 라이게이션(ligation)하였다. *E. coli* 균주 DH5α에 라이게이션(ligation) 산물 5 ul를 형질전환하였다. 인서트(insrt)가 삽입된 플라스미드(plasmid)의 서열 분석은 RBC T&A 클로닝 벡터내에 인식 서열이 존재하는, 두 개의 프라이머(M13 정방향 및 역방향 프라이머)를 이용하였다. 준비된 플라스미드 DNA를 주형으로써 이용한, 시퀀싱 반응은 BigDye Terminator chemistry(Ver. 3.1; Cycle Sequencing Kit; Applied Biosystems, Foster City, CA)를 이용하여 제조사의 지시에 따라 수행하였다. PCR 증폭은 250ng의 플라스미드 DNA 3 ul, 다범위 적용 프라이머 0.5 ul(3 pmol), 5X 시퀀싱(sequencing) 완충용액 0.87 ul, 멸균 증류수 1.38 ul 및 GeneAmp PCR System 2720(Applied Biosystems, Foster City, CA)를 이용한 BDT v3.1 0.25 ul로 수행하였다. 상기 반응액을 96℃에서 10초 변성, 50℃ 5초 어닐링(annealing), 및 60℃ 4분 신장 조건으로 35 싸이클(cycle)동안 증폭하였다. 증폭된 PCR 산물은 에탄올 침전법을 통해 정제하고, ABI 3730XL DNA Analyzer(Applied Biosystems)로 분석하였다.

[0162]

<2-2> 서열 열라인먼트(alignment)를 통한 결과 분석

[0163]

서열 분석을 통해 160개의 PCR 산물이 융합 유전자 서열(융합 유전자 전사 변이체)로 예상되었으며, 나머지 산물들은 예상되는 서열을 가지고 있지 않았다. 실험 조건에서, 표적화된 융합 유전자를 검출할 수 있는 성공률은 0.06%로 측정되었으며, 이는 융합 유전자가 인간 세포에서 극히 낮은 수준으로 발현된다는 것을 의미한다. 얻어진 160개의 융합 유전자 전사 변이체(CGTV) 서열을 분석한 결과, 이들이 5개의 인간 융합 유전자(*ZNF343-SNRPB-Hu*, *COX17-POPDC2-Hu*, *RBM7-REXO2-Hu*, *PRKRIP1-ORAI2-Hu* 및 *ATP5F1-C1orf162-Hu*)에 속한 57개의 단일

융합 유전자 형태 내로 분류됨을 확인하였다.

- [0164] 첫 번째, *ZNF343-SNRPB-Hu*는 모체 유전자 *ZNF343*(업스트림, upstream) 및 *SNRPB*(다운스트림, downstream)을 갖는 융합유전자로, 총 20개의 융합 유전자 전사 변이체(CGTV) 중 12개는 부분적 CGTV를 갖는, *ZNF343-SNRPB-Hu-P1~12*이고, 8개는 전체 CDS 부위를 갖는 CGTV인 *ZNF343-SNRPB-Hu-F1~8*이다(도 1a 및 도 6a).
- [0165] 임의의 폴리-A 서열이 없는 불완전한 mRNA 구조를 갖는, 본 발명의 인간 융합 유전자와 비교하여, 폴리-A 서열(AK305048)의 완벽한 mRNA 구조를 갖는, 칩팬지 융합 유전자(CG) mRNA를 이용하여 신호 서열을 분석하였다. 생물정보학 프로그램 "HCpolya" ([http://zeus2.itb.cnr.it/~webgene/wwwHC\\_polya.html](http://zeus2.itb.cnr.it/~webgene/wwwHC_polya.html))에 의한 폴리-A 신호 서열 예측을 통해, 1747 내지 1752 bp 부위에 위치한 적당한 폴리-A 신호 서열의 후보를 확인하였다(AATAAA).
- [0166] *SNRPB*는 UCSC 유전체 브라우저(genome browser)에 따른 4개의 다양한 변형된 스플라이싱(splicing) 변이체를 갖고, 이들 중 두 개는 그 마지막 엑손의 길이가 길거나(uc002wfz.1) 짧은(uc002wga.1) 차이가 있는 것을 제외하고 같은 유전자 구조를 갖는다. 추가적 분석을 통하여, 본 발명의 8개의 전장 CDS CGTV 중 5개는 변이체(uc002wfz.1)의 엑손 "7"과 같은 서열을 갖고, 나머지 3개는 다른 변이체(uc002wga.1)와 같은 서열을 가짐을 확인하였다(도 1 및 도 3a).
- [0167] 두 번째, 모체 유전자 *COX17* 및 *POPDC2*를 갖는 융합 유전자 *COX17-POPDC2-Hu*로(도 1b), 총 23개의 다양한 융합 유전자 전사 변이체(CGTV) 중 11개는 부분적 CGTV를 갖는, *COX17-POPDC2-Hu-P1~11*이고, 12개는 전체 CDS 부위를 갖는 CGTV인 *COX17-POPDC2-Hu-F1~12*이다(도 6b). 23개의 CGTV 중 21개(91%)의 구조 내 인트론에서 만들어진 두 가지 형태의 신규 엑손인 "b" 및 "c"를 확인하였다. 상기 두 개의 신규 엑손은 각각 전이 인자 서열인 MIRc(SINE) 및 MLT1L(LTR)로부터 유래되었다(도 3b). 추가적으로, 두 개의 CGTV(*COX17-POPDC2-Hu-P9* 및 *-11*)의 인트론 부위 내에 새롭게 만들어진 두 개의 엑손("a" 및 "d")을 확인하였고, 상기 "a"는 전이 인자 L2a로부터 유래함을 확인하였다(도 3b). 완벽한 mRNA 구조(AK306023)를 갖는 칩팬지 CG mRNA를 이용하여 다운스트림 부위의 폴리-A 신호 서열을 분석하였다. 이러한 신호 서열은 이러한 융합 유전자에서 1574 내지 1579 bp(AATAAA)에 위치하고 있다. 이러한 결과는 첫 번째 융합 유전자로부터 관찰된 바를 뒷받침한다: 업스트림 유전자의 폴리-A 신호 서열을 갖는 3'UTR이 완벽히 제거되고, 다운스트림의 폴리-A 신호가 융합 유전자의 전사를 종결하는데 이용되어야만 한다.
- [0168] *POPDC2*는 PDP 패밀리의 한 종류로, 인간 유전체 내 *POP2*(uc003ecy.1)와 거의 같은 좌위(locus)에 위치하고 있다. 흥미롭게도, 융합 유전자의 형성에 자주 참여하는 엑손 "1"<sup>v1</sup>는 다운스트림 유전자의 엑손 "1"의 절단된 형태이고, *POP2* 유전자 구조 내 두 번째 엑손의 서열과 동일하다.
- [0169] 세 번째, 융합 유전자 *RBM7-REXO2-Hu*는 모체 유전자 *RBM7* 및 *REXO2*를 갖는 융합유전자로(도 1c), 이러한 융합 유전자는 총 4개의 CGTVs(*RBM7-REXO-Hu-P1~4*)로부터 전사됨을 확인하였다. 특히, *RBM7-REXO-Hu* 내 유전자간 부위와 인접한 두 개의 엑손이 이러한 융합 유전자의 형성에서 완벽히 스킵되거나 절단되었음을 관찰하였다. 구체적으로, 업스트림 유전자의 마지막 엑손의 절단은 CDS 부위내에서 일어났으며, 상기 절단은 상기 언급된 두 개의 융합 유전자와 마찬가지로, 업스트림 유전자의 폴리-A 신호 서열의 제거의 결과임을 확인하였다(도 1c-1). 추가적 분석을 통해 또 다른 폴리-A 신호가 상기 두 개의 융합 유전자와 일치하게, 다운스트림 모체 유전자의 1887 내지 1891 bp(AATAAG)에 위치함을 확인하였다. 또한, TE로부터 유래한 5개의 새롭게 만들어진 신규 엑손 서열을 확인하였다: L1MB3(LINE), Charlie13b(DNA), L1MEc(LINE), MER41B(LTR) 및 MIRb(SINE)(도 3c).
- [0170] 네 번째 *PRKRIP1-ORAI2-Hu* 융합 유전자는 모체 유전자 *PRKRIP1* 및 *ORAI2*를 갖는 융합 유전자로, 6 개의 CGTV(*PRKRIP1-ORAI2-Hu-P1~6*)로 전사되었다(도 1d). 상기 세 개의 융합 유전자와 마찬가지로, 폴리-A 신호가 다운스트림 부위에 유지되는 지의 여부를 측정하였으며, 칩팬지 mRNA 서열(AK306465) 내 3005 내지 3010 bp(AATACA)에 있는 신호 부위를 확인하였으며, 이러한 현상은 상기 세 개의 융합 유전자 결과와 일치하였다. 또한, 칩팬지 융합 유전자 mRNA의 유전자간 부위(모체 유전자 *ORAI2*) 내 새롭게 만들어진 신규 엑손("a")이 본 발명의 6개 인간 CGTV 및 UCSC 유전체 데이터베이스(database)에는 존재하지 않음을 확인하였다. 이러한 결과는 신규 엑손 "a"가 전이 인자 서열 MER2(LINE)에서 유래한 칩팬지 특이적 엑손일 수 있음을 나타낸다(도 3d).

[0171] 다섯 번째 *ATP5F1-Clorf162-Hu* 융합 유전자는 모체 유전자 *ATP5F1* 및 *Clorf162*를 갖는 융합유전자이고, 4개의 다른 인간 CGTV(*ATP5F1-Clorf162-Hu-P1-4*)로 전사되었다(도 1e). 이러한 분석에서, 침팬지의 새로운 형태의 융합 유전자(*ATP5F1-Clorf162-Ch*)가 세 개의 이웃한 모체 유전자 *ATP5F1*, *Clorf162* 및 *WDR77* 사이에서 융합되었음을 확인하였다. 서열 분석을 통해 각각, 전이 인자 서열 MLT1E1A(LTR) 및 L2a(비-LTR 이동 유전자)에서 유래된 엑손 "b" 뿐 아니라, 새로운 유전체내 엑손 "a" (*ATP5F1-Clorf162-Ch*)를 밝혔다(도 3e). 추가적 분석을 통해 상기 절단은 엽스트림의 폴리-A 신호 서열에서 신호 부위가 절단되어 일어남을 확인하였다(도 1e-1). 본 서열 분석은 엑손 "3" 직후에 마지막 엑손 "3"이 18 bp 인트론 서열 연장(GTAAATGAGCATTCTCCT)을 가짐을 나타내었다. "HCpolya"에 의한 폴리-A 신호 서열의 예측을 통해, 1585 내지 1591 bp 위치에 새롭게 만들어진 6 bp의 폴리-A 신호 서열(AGTAAA)을 확인하였다(도 1e-3).

[0172] <실시예 3> 본 발명의 신규 융합 유전자의 구조 분석

[0173] 상기 <실시예 2>에서 본 발명의 융합 유전자의 서열 분석과 함께, 그 구조분석을 수행하였다.

[0174] 구체적으로, 융합유전자의 염기서열 분석 결과는 시퀀처 4.1.5 프로그램(Sequencher 4.1.5 program(미국 엔아버진코드회사로부터 구입된 소프트웨어)를 이용하였으며, UCSC 지놈브라우저(UCSC (University of California Santa Cruz) genome browser)를 이용하여 인간 융합 유전자들의 인간유전체에서 구조적 조직화 및 엑손 구조들을 고찰하고 비교분석하였다.

[0175] 첫 번째로, 20개의 *ZNF343-SNRPB-Hu* CGTV 서열의 구조적 분석을 통하여, 신규 엑손 생성, 엑손 스킵(skip) 및/또는 절단, 엑손 신장 및 서열 결실과 같은, 다양한 변형적 스플라이싱(splicing)이 일어남을 관찰하였다.

[0176] 새롭게 만들어진 신규 엑손(엑손 "a", "b", "c", "e" 및 "f") 5종이 9개의 CGTV 구조(*ZNF343-SNRPB-Hu-P1*, *-P2*, *-P3*, *-P4*, *-P6*, *-P11*, *-P12*, *-F4* 및 *-F6*)의 유전자간 부위(intragenic region)내에서 확인되었다. 신규 엑손 "c"는 전이 인자(transposable element) 서열인 AluSp로 동정되었다. 특히, 상기 신규 엑손인 "b" 및 "f"는 지금까지 동정된 것 중 가장 짧은 엑손인 3개의 염기(TCC)로 구성되어 있었다. 또한, 엑손 연장에 의해 생성된 두 개의 신규 엑손 ("1+d") 및 "a<sup>v</sup>+3" 를 발견하였다. 이 중 하나는 *ZNF343-SNRPB-Hu-P10*의 엑손 "1"의 수용체 부위에 47bp의 인트론 서열("d")이 삽입되어 만들어졌다. 또 다른 하나는 *ZNF343-SNRPB-Hu-P9*의 공여체 부위에 85 bp의 인트론 서열("a<sup>v</sup>")이 삽입되어 만들어졌고, 새롭게 만들어진 신규 엑손 "a"의 절단된 형태를 가진다. 추가로, 각각 엑손 "6"(*ZNF343*)으로부터 2661 bp 및 2664 bp가 결실된 "6<sup>v1</sup>" 및 "6<sup>v2</sup>"(*ZNF343-SNRPB-Hu-P5* 및 *-F8*), 및 신규 엑손 "c"로부터 93 bp가 결실된 "6<sup>v2</sup>"(*ZNF343-SNRPB-Hu-P2*)의 세 개의 CGTV 구조 내 세 가지 형태의 엑손 절단을 관찰하였다(도 1a-1).

[0177] 특히, 21개 CGTV 구조를 갖는 엽스트림 단일 유전자 *ZNF343*의 마지막 엑손("6")은 절단된 두 개(*ZNF343-SNRPB-Hu-P5* 및 *-F8*)를 제외하고, 거의 모두가 융합 유전자 형성에서 모두 스킵되었다. 엑손 "6" 내에서, 이러한 절단은 그 CDS("6<sup>v1</sup>" 및 "6<sup>v2</sup>")내에 일어났다.

[0178] 다음으로, 융합 유전자 형성에서 엑손 보존에 대해 연구하였다. 12개의 부분적 융합 유전자 전사 변이체(*ZNF343-SNRPB-Hu-P1-12*)의 구조적 분석을 통해, 21개의 CGTV 구조 내 다운스트림 유전자의 시작 엑손(*SNRPB*의 엑손 "1")이 연장된 한 개(*ZNF343-SNRPB-Hu-P10*)를 제외하고, 엽스트림 유전자 *ZNF343*의 마지막 엑손("6")과 유사하게, 완벽히 스킵되었음을 확인하였다.

[0179] 두 번째, *COX17-POPDC2-Hu-F1-12*의 23개 CGTV 중 21개(91%)의 구조 내 인트론에서 만들어진 두 가지 형태의 신규 엑손인 "b" 및 "c"를 확인하였다. 상기 두 개의 신규 엑손은 각각 전이 인자 서열인 MIRc(SINE) 및 MLT1L(LTR)로부터 유래되었다(도 3b). 추가적으로, 두 개의 CGTV(*COX17-POPDC2-Hu-P9* 및 *-11*)의 인트론 부위 내에 새롭게 만들어진 두 개의 엑손("a" 및 "d")을 확인하였고, 상기 "a"는 전이 인자 L2a로부터 유래함을 확인하였다(도 3b).

- [0180] 엑손 생성 빈도를 비교하여, 엑손의 생성이 하나의 융합 유전자로부터 유래한 융합 유전자 전사 변이체(CGTV) 사이에 존재하는 인트론 부위보다 유전자간 부위 내에서 더 자주 일어난다는 것을 확인하였다. 추가로, 7 개의 인간 전체 CDS CGTV 구조 및 침팬지 융합 유전자 mRNA(AK306023) 내 공여체 부위("3")로부터 147 bp 서열 연장("e")에 의해 새롭게 생성된 엑손("e"+"3")을 확인하였다.
- [0181] 엑손 보존 분석을 통해 23개의 *COX17-POPDC2-Hu* CGTVs의 형성에서 유전자간 부위로부터 멀리 떨어진 엑손이 프록시멀(proximal) 엑손보다 더 높은 수준으로 보존되어 있음을 확인하였다. 23개의 CGTV 구조 내에서, 엑손 보존성은 엑손 "1" 및 "2"(*COX17*) 뿐 아니라, 다운스트림 유전자의 엑손 "4"에서도 완벽하였다. 다른 엑손들에 있어서, 보존성은 유전자간 부위로부터 멀리 떨어진 엑손에서 높게 나타났다; *POPDC2*에서 엑손 "3"(91%) 및 엑손 "2"(61%)는 높은 보존성을 나타내었다. 하지만, 이러한 경향은 유전자간 부위에 인접한 두 개의 엑손에서 더욱 명확히 나타났다(업스트림 유전자의 마지막 엑손 및 다운스트림 유전자의 첫 번째 엑손). 놀랍게도, 융합 유전자의 형성에 있어서 두 개의 엑손이 완벽히 스킵(skip)되거나 절단되었음을 확인하였다. 절단 위치("1<sup>v1</sup>")는 하나의 CGTV(*COX17-POPDC2-Hu-F11*)의 엑손 "1<sup>v1</sup>"("1<sup>v2</sup>")에서 12bp가 결실된 것을 제외하고, 14개의 CGTV 및 침팬지 융합 유전자 mRNA에서 동일하였다.
- [0182] *POPDC2*는 PDP 패밀리의 한 종류로, 인간 유전체 내 *POP2*(uc003ecy.1)와 거의 같은 좌위(locus)에 위치하고 있다. 흥미롭게도, 융합 유전자의 형성에 자주 참여하는 엑손 "1<sup>v1</sup>"는 다운스트림 유전자의 엑손 "1"의 절단된 형태이고, *POP2* 유전자 구조 내 두 번째 엑손의 서열과 동일하다. 게다가, 상기 CGTV에서 신규 엑손 "c"가 *POP2*의 첫 번째 엑손으로 사용되었다.
- [0183] 세 번째, 융합 유전자 *RBM7-REXO2-Hu*는 *C11orf71* 및 UCSC 데이터베이스(database)에서 확인되지 않는 또 다른 EST(BG718298) 사이에 유전자간 부위 내 새롭게 만들어진 두 개의 신규 엑손("f" 및 "g")을 나타낸다.
- [0184] 특히, *RBM7-REXO-Hu* 내 유전자간 부위와 인접한 두 개의 엑손이 이러한 융합 유전자의 형성에서 완벽히 스킵되거나 절단되었음을 관찰하였다. 또한, 엑손 "4"의 변이체("4<sup>v</sup>")가 수용체 부위에 3 염기 서열(CAG)의 작은 인트론의 삽입에 의해 야기된 서열 연장을 갖음에도, 프록시멀(proximal) 엑손(*REXO2*의 "2", "3" 및 "4")은 융합 유전자의 형성에서 낮은 보존성을 보이는 반면, 유전자간 부위로부터 멀리 떨어진 엑손 "3"(*RBM7*) 및 "5"(*REXO2*)는 CGTV 구조 내에서 매우 높은 보존성을 보였다. 침팬지 융합 유전자 mRNA를 포함한 모든 5개의 CGTV의 유전자간 부위 내에서 새롭게 만들어진 6개의 신규 엑손("a", "c", "d", "e", "f" 및 "g")을 구조적 분석을 통해 확인하였다.
- [0185] 네 번째 *PRKRIP1-ORAI2-Hu* 융합 유전자의 구조적 분석을 통해, 스트림에 인접한 엑손이 절단된 두 개의 CGTV(*PRKRIP1-ORAI2-Hu-P4* 및 *-P6*)를 제외하고, 모든 CGTV의 유전자간 부위에 인접한 엑손이 완벽히 스킵되었음을 확인하였다. 이러한 두 개의 CGTV 구조 내 절단된 마지막 업스트림 엑손의 서열 수준에 대한 구조적 분석에 의해, 절단이 UTR 내에서 일어나고, 폴리-A 신호 서열 부위는 완벽히 결실되었음을 확인하였다(도 1d-1). 또한 침팬지 융합 유전자 mRNA의 유전자간 부위(모체 유전자 *ORAI2*) 내 새롭게 만들어진 신규 엑손("a")이 본 발명의 6개 인간 CGTV 및 UCSC 유전체 데이터베이스(database)에는 존재하지 않음을 확인하였다. 이러한 결과는 신규 엑손 "a"가 전이 인자 서열 MER2(LINE)에서 유래한 침팬지 특이적 엑손일 수 있음을 나타낸다(도 3d).
- [0186] 다섯 번째 *ATP5F1-C1orf162-Hu* 융합 유전자는 구조적 분석을 통해, 침팬지의 새로운 형태의 융합 유전자(*ATP5F1-C1orf162-Ch*)가 세 개의 이웃한 모체 유전자 *ATP5F1*, *C1orf162* 및 *WDR77* 사이에서 융합되었음을 확인하였다. 첫 번째 침팬지 엑손, "c+1<sup>v</sup>"은 *WDR77*의 첫 번째 엑손 및 *ATP5F1*의 첫 번째 엑손의 일부분을 포함하고 있다(도 1e-2 및 도 3e). *ATP5F1-C1orf162-Hu* 융합 유전자는 상기 4개의 융합 유전자와 비교하여 다운스트림 모체 유전자 내 첫 번째 엑손과 인접한 유전자내 부위가 완벽히 보존되는 이례적인 구조를 갖는다(도 1e). 하지만, 상기 전형적인 현상과 일치하게, 업스트림 모체 유전자(*ATP5F1*)의 마지막 인접 엑손은 CDS(*ATP5F1-C1orf162-Hu-P3*) 또는 UTR(*ATP5F1-C1orf162-Hu-P2*, *-P4* 및 *ATP5F1-C1orf162-Ch*) 내로 완벽히 스킵되거나 절단되었다. 추가적 분석을 통해 상기 절단은 업스트림의 폴리-A 신호 서열에서 신호 부위가 절단되어 일어남을 확인하였다(도 1e-1). 상기 4개의 융합 유전자 구조와 반대로, 침팬지 CG mRNA(AK306934)는 엑손 "3" 이후에



긴 폴리-A 서열을 갖는 완벽한 mRNA 구조를 갖지만, 상기 다운스트림 모체 유전자(*C1orf162*) 내 세 개 엑손 ("4", "5" 및 "6")의 완벽한 결실을 통해 구성되었음을 확인하였다(도 1e-3). 또한, 상기 4개의 융합 유전자와는 달리, 다운스트림 유전자의 폴리-A 신호 서열을 갖는 부위는 제거되었다.

[0187] <실시예 4> 본 발명의 신규 융합 유전자의 조직 내 발현 분포 확인

[0188] 30종의 인간 정상 및 이에 상응하는 종양 조직에서 증폭된 전체 2,450개 PCR 산물 중에서 160개의 표적화된 융합 유전자 전사 변이체를 얻었고, 이러한 조직에서 융합 유전자 전사 변이체의 발현 분포를 확인하였다.

[0189] 구체적으로, 30 종류의 인간 정상 및 암세포조직들로부터 분리된 총 RNA(Clontech과 Ambion으로부터 구입)를 이용하였으며, 융합유전자의 발현을 확인하기 위한 PCR의 조건은 하기와 같다: 94° C에서 5분간 초기 방치; 94° C에서 30초간 DNA 변성 단계, 60° C에서 30초간 어닐링 단계, 72° C에서 1분간 신장 단계: DNA 변성 단계로부터 신장 단계까지 43회 사이클 반복; 72° C에서 7분간 최종 방치단계. 상기 PCR 반응을 통해 얻어진 산물을 전기영동하여 그 밴드를 확인하였으며, 상기 밴드를 젤 추출법을 이용하여 획득한 후 TA 클로닝(TA cloning)하여, DNA염기 배열 분석 장치인 ABI 3730XL DNA Analyzer (Applied Biosystems, Foster City, CA)를 이용하여 융합 유전자들의 염기서열을 분석하였다.

[0190] 그 결과, 실험한 조직 중에서, 신장 정상 및 종양 조직에서 가장 높은 수의 융합 유전자 전사 변이체가 검출되었다(160개 중 19%). 평균적으로, 정상 조직에서보다 종양 조직에서 융합 유전자 전사 변이체의 발현이 더욱 특이적이었다; 종양 조직에서 115개(72%)가 발현된 반면, 정상 조직에서는 160개 중 45개(28%)의 융합 유전자 전사 변이체가 발현되었다(표 2). 특히, 상기 종양-편향된 경향성은 방광암(92%), 유방암(85%), 대장암(82%) 및 난소암(80%) 조직에서 더욱 뚜렷하게 나타났다. 이러한 결과는 융합 유전자 전사 변이체가 종양-편향된 발현에 강한 경향성을 가짐을 의미한다. 이러한 경향성 더 많은 증거는 세 가지 정상 및 이에 상응하는 종양 조직(립프, 자궁경부 및 폐) 간 분포 빈도의 비교를 통해 확인할 수 있었다. 비록 그들 중 몇몇이 몇 가지 다른 정상 조직에서도 발현되에도 불구하고, 융합 유전자 전사 변이체(CGTV)는 이러한 세 가지 종양 조직(립프: 정상에서 0개 융합 유전자 전사 변이체 대 종양에서 9개 융합 유전자 전사 변이체; 자궁경부: 0개 대 10개; 폐: 0개 대 6개)에서만 관찰되었다.

[0191] 추가로, 본 발명의 57개의 융합 유전자 전사 변이체(CGTV)가 종양 특이적으로 발현되는지의 여부를 조사하였다.

[0192] 그 결과, 이들 중에서, 12개의 융합 유전자 전사 변이체(CGTV)가 한 가지 조직 형태의 정상 및 이에 상응한 종양 샘플에서 동시에 발현되는 것을 확인하였다. 또한, 이러한 12개의 융합 유전자 전사 변이체(CGTV) 중에서, 세 개(*ZNF343-SNRPB-Hu-P1* (5/5), *-P8* (5/13) 및 *ATP5F1-C1orf162-Hu-P3* (6/8))는 다양한 형태의 정상 및 종양 조직에서 가장 일반적으로 발현하는 것으로 나타났다.

[0193] 따라서, 본 발명의 다섯 개의 융합 유전자 또는 이의 융합 유전자 전사 변이체는 구조적으로 발암에 있어 중요한 역할을 할 것이라는 것을 알 수 있다.

표 2

신규 인간 융합 유전자(CG)의 변형된 전사 변이체(CGTVs)의 종양 특이적 발현 양상

| 융합 유전자<br>(CG) 이름 | CGTV | 프림 |   | 전립선 |   | 정소 |    | 지방경부 |    | 유방 |   | 대장 |    | 신장 |   | 폐 |    | 난소 |   | 위 |   | 자궁 |    | 방광 |   | 식도 |   | 간 |   | 갑상선    |         | 총합(T) |
|-------------------|------|----|---|-----|---|----|----|------|----|----|---|----|----|----|---|---|----|----|---|---|---|----|----|----|---|----|---|---|---|--------|---------|-------|
|                   |      | T  | N | T   | N | T  | N  | T    | N  | T  | N | T  | N  | T  | N | T | N  | T  | N | T | N | T  | N  | T  | N | T  | N | T | N | T      | N       |       |
| ZNF343-SNRPB-Hu   | P1   |    |   |     |   |    |    | o    | o  | o  | o | o  | o  | o  | o | o |    |    |   |   | o |    |    |    |   |    |   |   |   | o      | 10 (5)  |       |
|                   | P2   |    |   |     |   |    |    |      |    |    |   |    |    |    | o |   |    |    |   |   |   |    |    |    |   |    |   |   |   |        | 1 (1)   |       |
|                   | P3   |    |   |     |   |    |    | o    |    |    |   |    |    |    | o |   |    |    |   |   |   |    |    |    |   |    |   |   |   |        | 2 (2)   |       |
|                   | P4   |    |   |     |   |    |    | o    |    |    |   |    |    |    | o |   |    |    |   |   |   |    |    |    |   |    |   |   |   |        | 3 (3)   |       |
|                   | P5   |    |   |     |   |    |    |      |    |    |   |    | o  |    |   |   |    |    |   |   |   |    |    |    |   |    |   |   |   |        | 1 (1)   |       |
|                   | P6   |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    |    | o  |   |    |   |   |   |        | 1 (0)   |       |
|                   | P7   |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    |    | o  |   |    |   |   |   |        | 1 (1)   |       |
|                   | P8   | o  |   | o   |   |    | o  | o    | o  | o  | o | o  | o  | o  | o | o | o  | o  | o | o | o | o  | o  | o  | o | o  | o | o | o | o      | 18 (12) |       |
|                   | P9   |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    |    |    |   |    |   | o |   |        | 1 (0)   |       |
|                   | P10  |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    |    | o  |   |    |   |   |   |        | 1 (0)   |       |
|                   | P11  | o  |   |     |   |    |    | o    |    |    |   |    |    |    |   | o | o  |    |   |   |   |    | o  | o  |   |    |   |   |   |        | 7 (6)   |       |
|                   | P12  |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   | o  |    |   |   |   |    |    |    |   |    |   |   |   |        | 1 (1)   |       |
|                   | F1   |    |   |     |   |    |    |      |    |    |   |    |    | o  | o |   |    |    |   |   |   |    |    |    |   |    |   |   |   |        | 2 (1)   |       |
|                   | F2   |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    |    | o  |   |    |   |   |   |        | 1 (1)   |       |
|                   | F3   |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    | o |   |   |    |    |    |   |    |   |   |   |        | 3 (1)   |       |
| F4                |      |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   | o  |    |    |   |    |   |   |   | 3 (1)  |         |       |
| F5                |      |    |   |     |   |    |    |      |    | o  | o | o  | o  |    |   |   |    |    |   |   |   |    |    |    |   |    |   |   |   | 4 (3)  |         |       |
| F6                |      |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    | o  |   |   |   |    |    |    |   |    |   |   |   | 1 (1)  |         |       |
| F7                |      |    |   |     |   |    |    |      |    |    |   |    |    | o  | o | o | o  |    |   |   |   |    |    |    |   |    | o |   |   | 5 (4)  |         |       |
| F8                |      |    |   |     |   |    |    |      |    |    |   |    |    | o  |   | o | o  |    |   |   |   |    |    |    |   |    | o |   |   | 4 (4)  |         |       |
| COX17-POPCD2-Hu   | P1   |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    | o  |    |   |    |   |   |   | 1 (1)  |         |       |
|                   | P2   |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    |    |    |   |    |   |   |   | o      | 1 (0)   |       |
|                   | P3   | o  |   |     |   |    | o  |      |    | o  | o | o  | o  | o  | o | o | o  | o  | o | o | o | o  | o  | o  | o | o  | o | o | o | 9 (6)  |         |       |
|                   | P4   |    |   |     |   |    |    |      |    |    |   |    |    |    | o |   |    |    |   |   |   |    |    | o  | o |    |   |   |   | o      | 2 (1)   |       |
|                   | P5   | o  |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    |    |    |   |    |   |   |   |        | 1 (1)   |       |
|                   | P6   |    |   |     |   |    | o  |      |    |    |   |    |    |    | o |   |    |    |   |   |   |    |    |    |   |    |   |   |   | o      | 3 (1)   |       |
|                   | P7   |    |   |     |   |    |    |      | o  |    |   |    |    |    |   |   |    |    |   |   |   |    |    |    |   |    |   |   |   |        | 1 (1)   |       |
|                   | P8   | o  |   |     |   |    |    |      |    |    |   |    |    | o  |   |   |    |    |   |   |   |    |    |    |   |    |   |   |   |        | 2 (2)   |       |
|                   | P9   | o  | o |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    |    |    |   |    |   |   |   |        | 2 (2)   |       |
|                   | P10  | o  |   |     |   |    | o  |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    |    |    |   |    | o |   |   |        | 3 (3)   |       |
|                   | P11  |    |   |     |   |    |    |      | o  |    |   |    |    |    |   |   |    |    |   |   |   |    |    |    |   |    |   |   |   |        | 1 (1)   |       |
|                   | F1   |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   | o  |    |   |   |   |    | o  |    |   |    |   |   |   |        | 2 (2)   |       |
|                   | F2   |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    | o  |    |   |    |   |   |   |        | 1 (1)   |       |
|                   | F3   |    |   |     |   |    |    |      |    |    |   |    |    |    | o | o |    |    |   |   |   |    |    |    |   |    |   |   |   |        | 2 (2)   |       |
|                   | F4   |    |   |     |   |    |    |      |    |    |   |    |    |    | o |   |    |    |   |   |   |    |    |    |   |    |   |   |   |        | 1 (1)   |       |
| F5                |      |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   | o  |    |    |   |    |   |   |   | 1 (1)  |         |       |
| F6                |      |    |   |     |   |    |    |      |    |    |   |    |    |    | o |   |    |    |   |   |   |    |    |    |   |    |   |   |   | 1 (1)  |         |       |
| F7                |      |    |   |     |   |    |    |      |    |    |   |    | o  |    |   |   |    |    |   |   |   |    |    |    |   |    |   |   |   | 1 (1)  |         |       |
| F8                |      |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   | o  |    |    |   |    |   |   |   | 1 (0)  |         |       |
| F9                |      |    |   |     |   |    |    |      | o  |    |   |    |    |    |   |   |    |    |   |   |   |    |    |    |   |    |   |   |   | 1 (1)  |         |       |
| F10               |      |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   | o  | o  |   |   |   |    |    |    |   |    |   |   |   | 2 (2)  |         |       |
| F11               |      |    |   |     |   |    |    |      | o  | o  |   |    |    |    |   |   |    |    |   |   |   |    |    |    |   |    |   |   |   | 2 (2)  |         |       |
| F12               |      |    |   |     |   |    |    | o    | o  | o  | o | o  | o  | o  | o | o | o  | o  | o | o | o | o  | o  | o  | o | o  | o | o | o | 7 (5)  |         |       |
| PRKRIP1-ORAI2-Hu  | P1   | o  |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    |    |    |   |    |   |   |   | 1 (1)  |         |       |
|                   | P2   |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    |    |    |   |    |   | o |   | 1 (0)  |         |       |
|                   | P3   |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    | o  |    |   |    |   | o |   | 3 (1)  |         |       |
|                   | P4   |    |   |     |   |    |    |      |    | o  |   |    |    |    |   |   |    |    |   |   |   |    |    |    |   |    |   |   |   | 1 (1)  |         |       |
| PRKRIP1-ORAI2-Hu  | P1   | o  | o |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   | o | o  | o  |    |   |    |   |   |   | 5 (3)  |         |       |
|                   | P2   |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   | o  |    |    |   |    |   |   |   | 1 (1)  |         |       |
|                   | P3   |    |   | o   |   |    | o  |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    |    |    |   |    |   |   |   | 2 (2)  |         |       |
|                   | P4   |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    | o  |   |   |   |    |    |    |   |    |   |   |   | 1 (1)  |         |       |
|                   | P5   |    |   |     |   |    |    |      |    |    |   |    |    |    |   | o |    |    |   |   |   | o  |    |    |   |    |   |   |   | 2 (2)  |         |       |
|                   | P6   |    |   |     |   |    |    |      | o  |    |   |    |    |    |   | o | o  | o  | o | o | o | o  | o  | o  | o | o  | o | o | o | 3 (3)  |         |       |
| ATP5F1-Clof162-Hu | P1   |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    |    |    |   |    |   |   |   | 1 (1)  |         |       |
|                   | P2   |    |   |     |   |    |    |      |    |    |   |    |    |    |   |   |    |    |   |   |   |    |    |    |   |    |   | o |   | 1 (1)  |         |       |
|                   | P3   |    |   | o   |   | o  | o  | o    | o  | o  | o | o  | o  | o  | o | o | o  | o  | o | o | o | o  | o  | o  | o | o  | o | o | o | 14 (8) |         |       |
|                   | P4   |    |   |     |   |    |    | o    | o  | o  | o | o  | o  | o  | o | o | o  | o  | o | o | o | o  | o  | o  | o | o  | o | o | o | 6 (4)  |         |       |
| 분포 빈도             | 9    | 0  | 4 | 1   | 0 | 2  | 10 | 0    | 11 | 2  | 9 | 2  | 22 | 9  | 6 | 0 | 12 | 3  | 6 | 3 | 3 | 10 | 11 | 1  | 2 | 2  | 8 | 4 | 2 | 6      | 160     |       |

[0194]

[0195] <실시에 5> 본 발명의 신규 융합 유전자의 종양 특이적 발현 확인

[0196] 본 발명의 융합 유전자 전사 변이체(VGTV)가 인간 조직에서 강한 경향성으로 종양-편향적으로 발현되어 나타남에도 불구하고, 이러한 융합 유전자 전사 변이체(CGTV)가 암과 연관되어 있는지의 여부는 알려져 있지 않다. 따라서, 본 연구에서 정의된 57개의 다양한 CGTV가 종양 특이적으로 발현되는지의 여부를 조사하였다.

[0197] 그 결과, 35개의 다양한 융합 유전자 전사 변이체(CGTV)가 종양 조직에서만 독점적으로 발현되고, 이러한 분석은 57개 융합 유전자 전사 변이체에서 61.4%로 나타남을 확인하였다. 이러한 35개의 융합 유전자 전사 변이체 중에서, 13개는 다양한 종양 조직에서 발현된 것과 비교하여, 22개는 하나의 조직에서만 종양 특이적으로 발현되어 나타났다. 특히, 이는 융합 유전자 전사 변이체(CGTV) 중 37%가 상응하는 정상 조직에서는 발현되지 않지만, 다양한 종양 조직에서 특이적으로 발현됨을 확인하였다. 이러한 융합 유전자 전사 변이체(CGTV) 중에서, 9개(ZNF343-SNRPB-Hu-P3, COX17-POPCD2-Hu-P8-9, COX17-POPCD2-Hu-F1, -F3, -F10, -F11, PRKRIP1-ORAI2-Hu-P3 및 -P5), 세 개(ZNF343-SNRPB-Hu-P4, COX17-POPCD2-Hu-P10 및 PRKRIP1-ORAI2-Hu-P6) 및 한 개(ZNF343-SNRPB-

*Hu-F8*)는 각각 두 가지 종양-, 세 가지 종양-, 네 가지 종양 특이적 발현을 나타내었다(표 2). *ZNF343-SNRPB-Hu-F8*의 특이적 발현을 보이는 네 가지 종양 조직은 신장, 난소, 위 및 간이었다(도 6 및 도 7).

[0198] 따라서, 종양 특이적으로 발현되는 본 발명의 융합 유전자 전사 변이체는 암을 진단하는데 유용하게 이용될 수 있다.

[0199] <실시예 6> 본 발명의 신규 융합 유전자의 세포주 수준에서 발현 양상 확인

[0200] 상기 살펴본 바와 같이, 본 발명자들은 인간 정상 조직 또는 암 세포 조직에서 본 발명의 신규 융합 유전자가 많이 발현됨을 확인하였다. 이러한 신규 융합 유전자들의 세포주 수준에서의 발현 양상을 확인하기 위하여 HEK293 세포주를 이용하였다.

[0201] <6-1> HEK293 세포 배양

[0202] HEK293(human embryonic kidney cell line; 인간신장세포주) 세포는 10%(v/v) 열-불활성화된 우태아 혈청, 100 U/ml 페니실린(penicillin) 및 100 U/ml 스트렙토마이신(streptomycin)이 포함된 DMEM에서 배양하였다. 세포 들은 37°C에서 가습된 5% CO<sub>2</sub> 상태에서 세포 밀도가 대수 증식 상태가 될 때까지 배양하였다.

[0203] <6-2> HEK293 세포 배양으로부터 전체 RNA의 분리 및 확인

[0204] 전체 RNA는 RNeasy Protect Cell Reagent Kit(QIAGEN) 및 RNeasy Protect Cell Mini Kit(QIAGEN)을 이용하여 HEK293 세포에서 분리하였고, 이를 통해 인간 융합 유전자의 발현 양상을 확인한 결과, 본 발명의 신규 인간 융합 유전자 중 하나인 *ZNF343-SNRPB-Hu-F5*이 발현됨을 확인하였다(도 8).

[0205] <6-3> 신규 융합 유전자 전사 변이체의 DDBJ 등록번호

[0206] 본원발명의 신규 인간 융합 유전자 전사 변이체는 DDBJ 데이터베이스에 등재되었고, 그 등록번호는 하기 [표 3]에 기재된 바와 같이 CJ999946~ CJ999991 및 AK377173-AK377184이다.

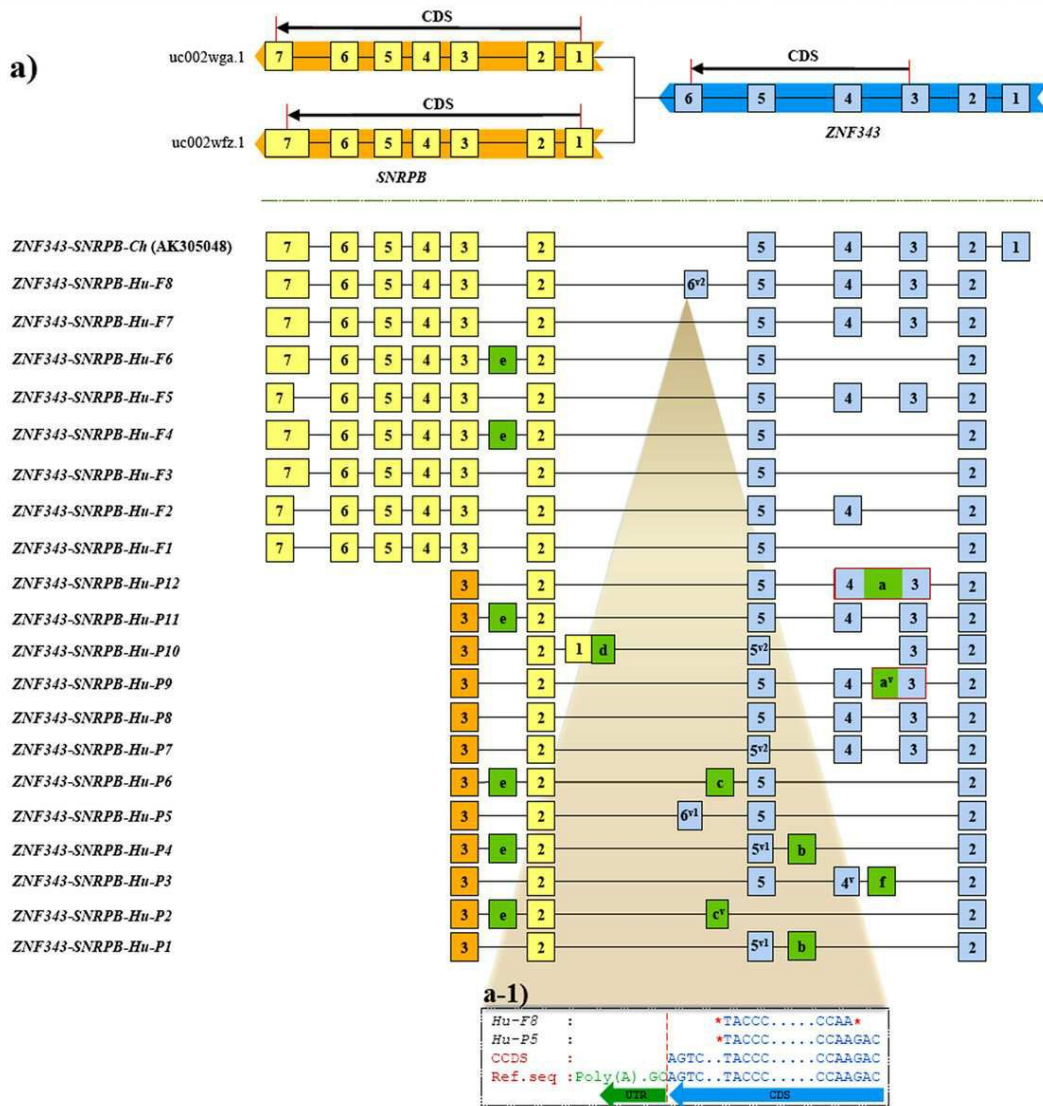
표 3

| CG name             | CGTV name             | Acc.No.  |
|---------------------|-----------------------|----------|
| ZNF343-SNRPB-Hu     | ZNF343-SNRPB-Hu-P1    | CJ999946 |
|                     | ZNF343-SNRPB-Hu-P2    | CJ999947 |
|                     | ZNF343-SNRPB-Hu-P3    | CJ999948 |
|                     | ZNF343-SNRPB-Hu-P4    | CJ999949 |
|                     | ZNF343-SNRPB-Hu-P5    | CJ999950 |
|                     | ZNF343-SNRPB-Hu-P6    | CJ999951 |
|                     | ZNF343-SNRPB-Hu-P7    | CJ999952 |
|                     | ZNF343-SNRPB-Hu-P8    | CJ999953 |
|                     | ZNF343-SNRPB-Hu-P9    | CJ999954 |
|                     | ZNF343-SNRPB-Hu-P10   | CJ999955 |
|                     | ZNF343-SNRPB-Hu-P11   | CJ999956 |
|                     | ZNF343-SNRPB-Hu-P12   | CJ999957 |
|                     | ZNF343-SNRPB-Hu-F1    | CJ999958 |
|                     | ZNF343-SNRPB-Hu-F2    | CJ999959 |
|                     | ZNF343-SNRPB-Hu-F3    | CJ999960 |
|                     | ZNF343-SNRPB-Hu-F4    | CJ999961 |
|                     | ZNF343-SNRPB-Hu-F5    | CJ999962 |
|                     | ZNF343-SNRPB-Hu-F6    | CJ999963 |
|                     | ZNF343-SNRPB-Hu-F7    | CJ999964 |
|                     | ZNF343-SNRPB-Hu-F8    | CJ999965 |
| COX17-POPDC2-Hu     | COX17-POPDC2-Hu-P1    | CJ999966 |
|                     | COX17-POPDC2-Hu-P2    | CJ999967 |
|                     | COX17-POPDC2-Hu-P3    | CJ999968 |
|                     | COX17-POPDC2-Hu-P4    | CJ999969 |
|                     | COX17-POPDC2-Hu-P5    | CJ999970 |
|                     | COX17-POPDC2-Hu-P6    | CJ999971 |
|                     | COX17-POPDC2-Hu-P7    | CJ999972 |
|                     | COX17-POPDC2-Hu-P8    | CJ999973 |
|                     | COX17-POPDC2-Hu-P9    | CJ999974 |
|                     | COX17-POPDC2-Hu-P10   | CJ999975 |
|                     | COX17-POPDC2-Hu-P11   | CJ999977 |
|                     | COX17-POPDC2-Hu-F1    | AK377173 |
|                     | COX17-POPDC2-Hu-F2    | AK377174 |
|                     | COX17-POPDC2-Hu-F3    | AK377175 |
|                     | COX17-POPDC2-Hu-F4    | AK377176 |
|                     | COX17-POPDC2-Hu-F5    | AK377177 |
|                     | COX17-POPDC2-Hu-F6    | AK377178 |
|                     | COX17-POPDC2-Hu-F7    | AK377179 |
|                     | COX17-POPDC2-Hu-F8    | AK377180 |
|                     | COX17-POPDC2-Hu-F9    | AK377181 |
| COX17-POPDC2-Hu-F10 | AK377182              |          |
| COX17-POPDC2-Hu-F11 | AK377183              |          |
| COX17-POPDC2-Hu-F12 | AK377184              |          |
| RBM7-REXO2-Hu       | RBM7-REXO2-Hu-P1      | CJ999978 |
|                     | RBM7-REXO2-Hu-P2      | CJ999979 |
|                     | RBM7-REXO2-Hu-P3      | CJ999980 |
|                     | RBM7-REXO2-Hu-P4      | CJ999981 |
| PRKRIP1-ORAI2-Hu    | PRKRIP1-ORAI2-Hu-P1   | CJ999982 |
|                     | PRKRIP1-ORAI2-Hu-P2   | CJ999983 |
|                     | PRKRIP1-ORAI2-Hu-P3   | CJ999984 |
|                     | PRKRIP1-ORAI2-Hu-P4   | CJ999985 |
|                     | PRKRIP1-ORAI2-Hu-P5   | CJ999986 |
|                     | PRKRIP1-ORAI2-Hu-P6   | CJ999987 |
| ATP5F1-C1orf162-Hu  | ATP5F1-C1orf162-Hu-P1 | CJ999988 |
|                     | ATP5F1-C1orf162-Hu-P2 | CJ999989 |
|                     | ATP5F1-C1orf162-Hu-P3 | CJ999990 |
|                     | ATP5F1-C1orf162-Hu-P4 | CJ999991 |

[0207]

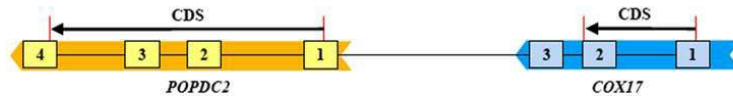
도면

도면1a

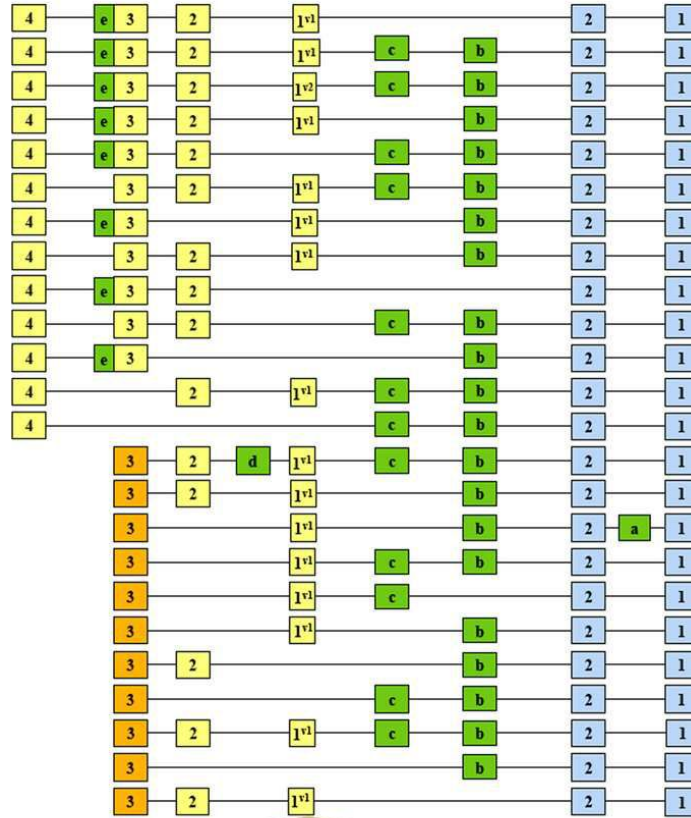


도면1b

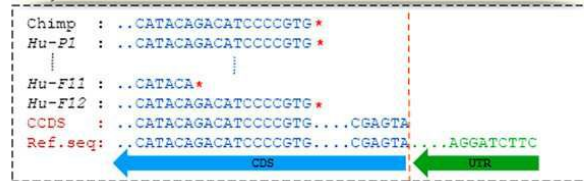
b)



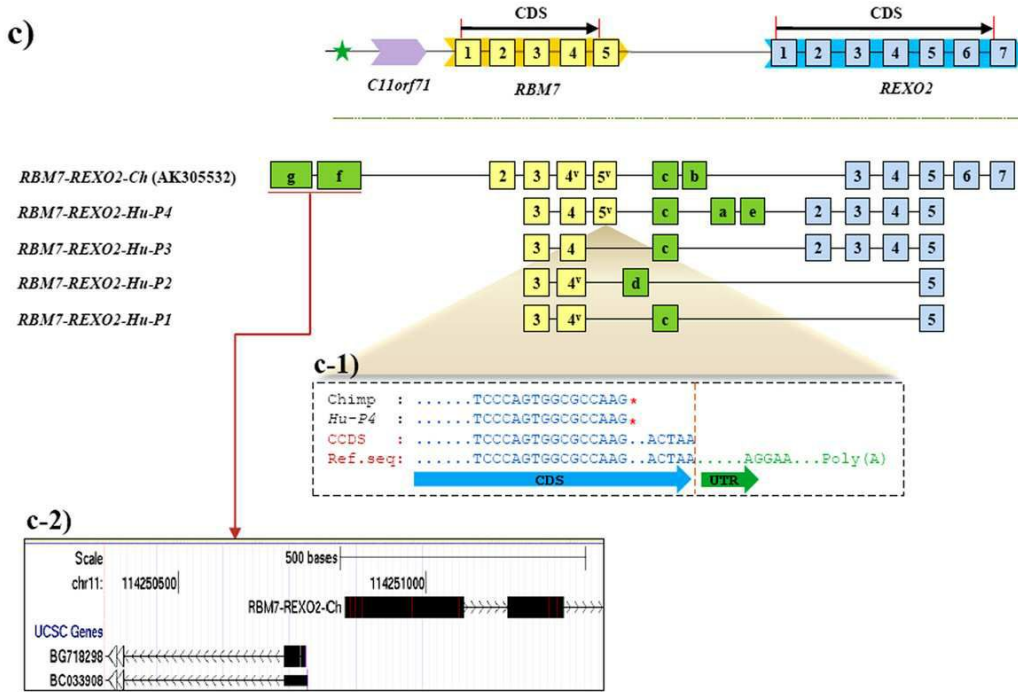
- COX17-POPDC2-Ch (AK306023)
- COX17-POPDC2-Hu-F12
- COX17-POPDC2-Hu-F11
- COX17-POPDC2-Hu-F10
- COX17-POPDC2-Hu-F9
- COX17-POPDC2-Hu-F8
- COX17-POPDC2-Hu-F7
- COX17-POPDC2-Hu-F6
- COX17-POPDC2-Hu-F5
- COX17-POPDC2-Hu-F4
- COX17-POPDC2-Hu-F3
- COX17-POPDC2-Hu-F2
- COX17-POPDC2-Hu-F1
- COX17-POPDC2-Hu-P11
- COX17-POPDC2-Hu-P10
- COX17-POPDC2-Hu-P9
- COX17-POPDC2-Hu-P8
- COX17-POPDC2-Hu-P7
- COX17-POPDC2-Hu-P6
- COX17-POPDC2-Hu-P5
- COX17-POPDC2-Hu-P4
- COX17-POPDC2-Hu-P3
- COX17-POPDC2-Hu-P2
- COX17-POPDC2-Hu-P1



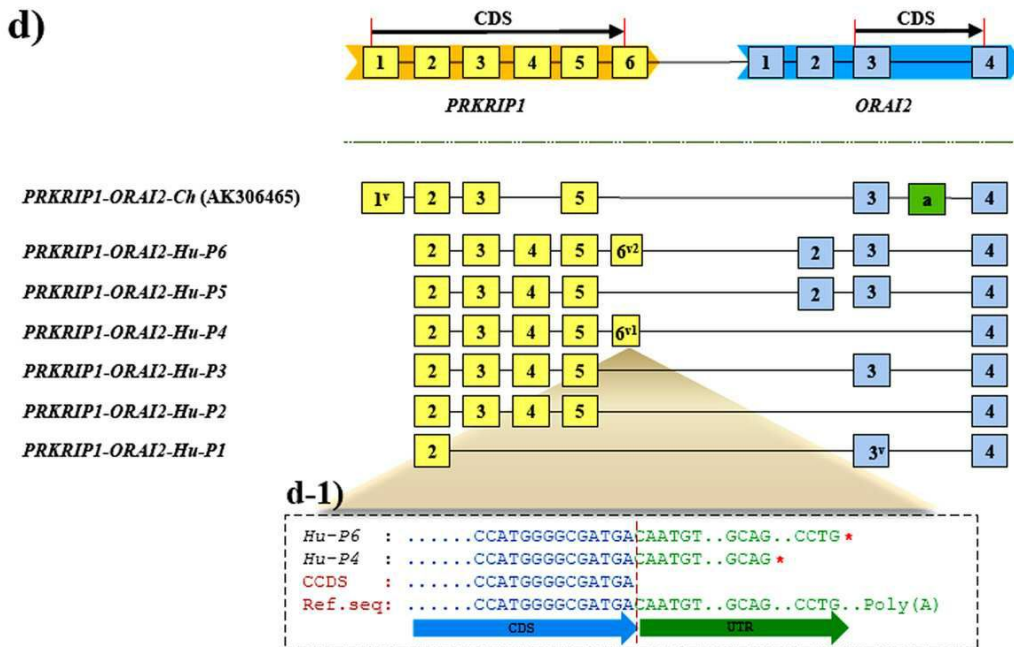
b-1)



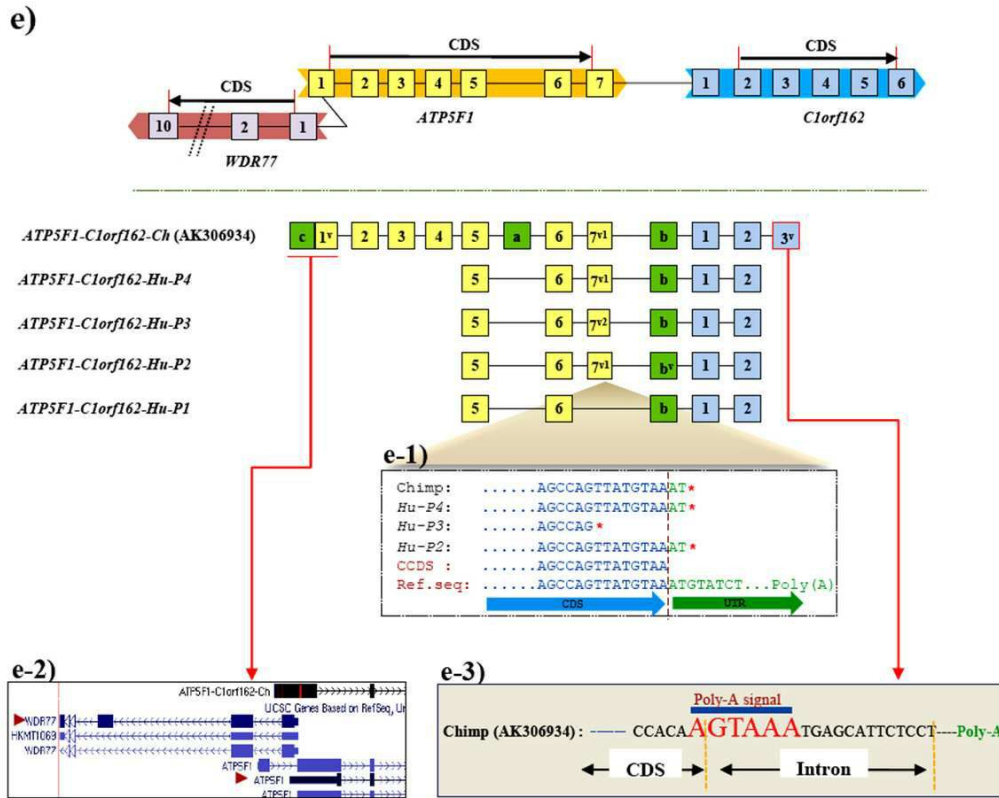
도면1c



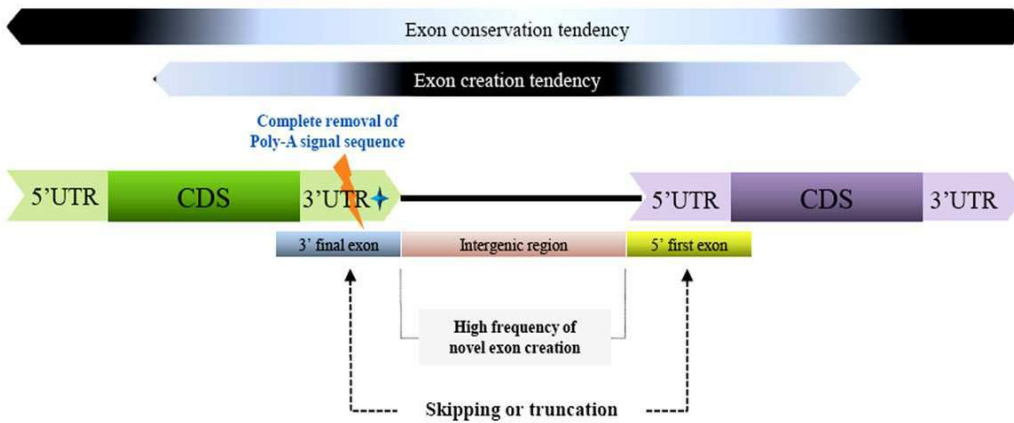
도면1d



도면1e

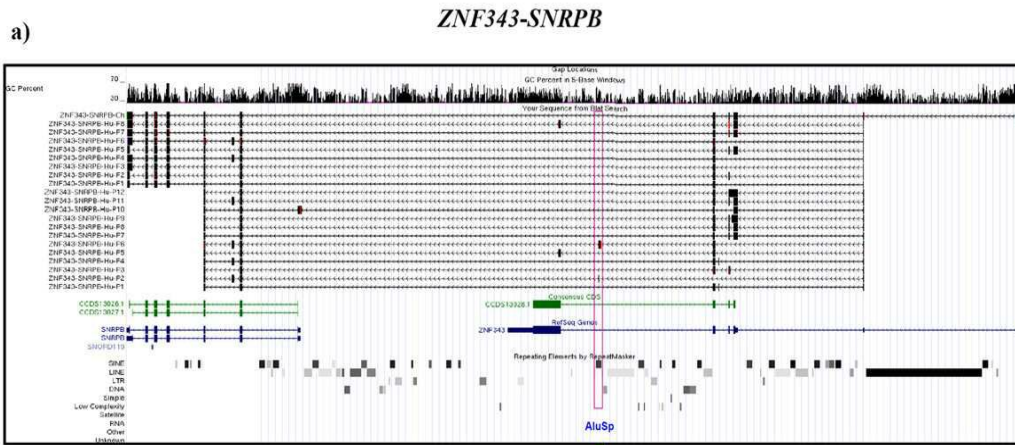


도면2

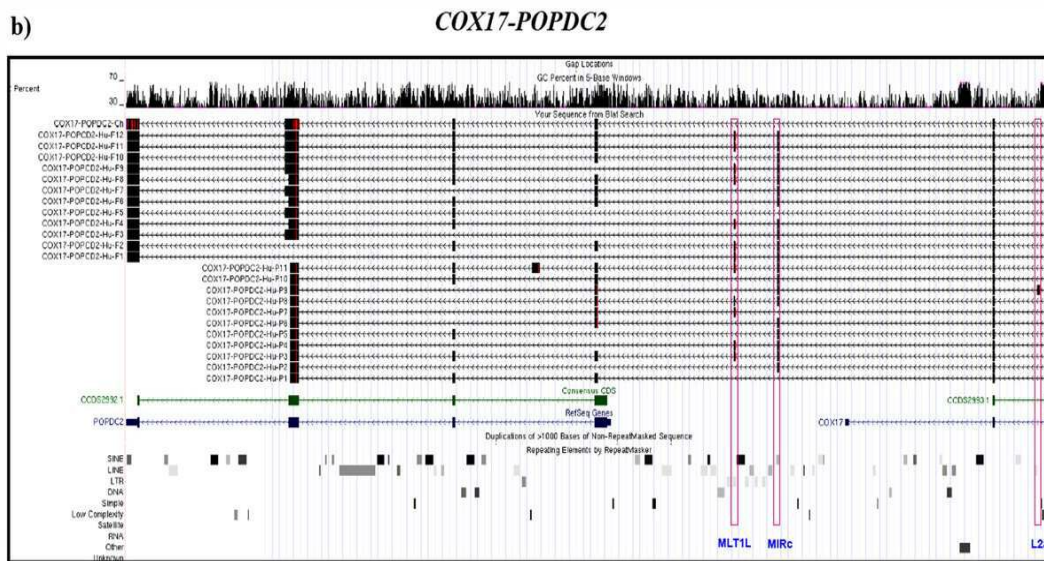




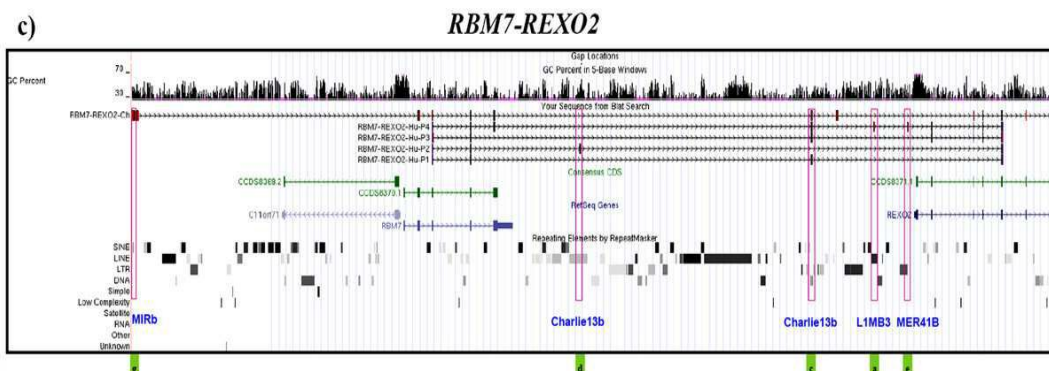
도면3a



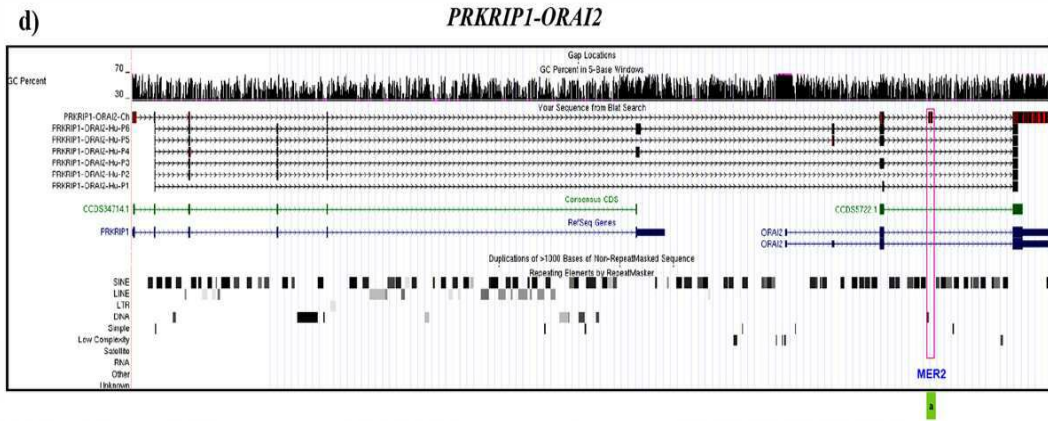
도면3b



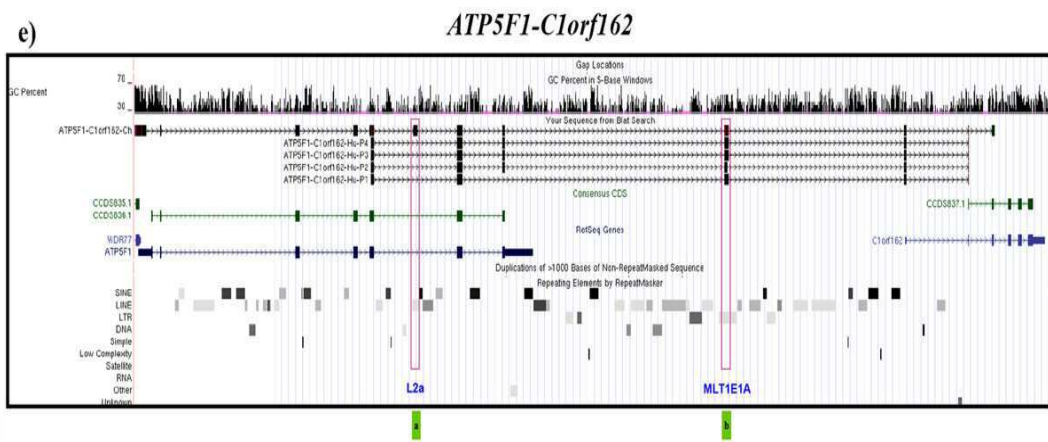
도면3c



도면3d

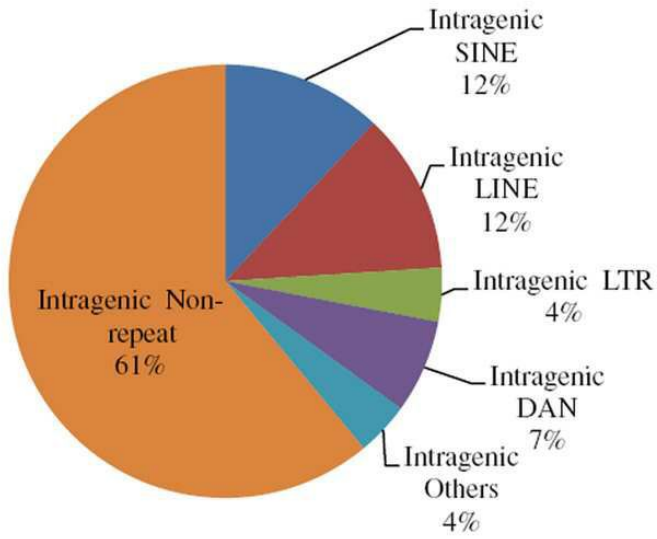


도면3e

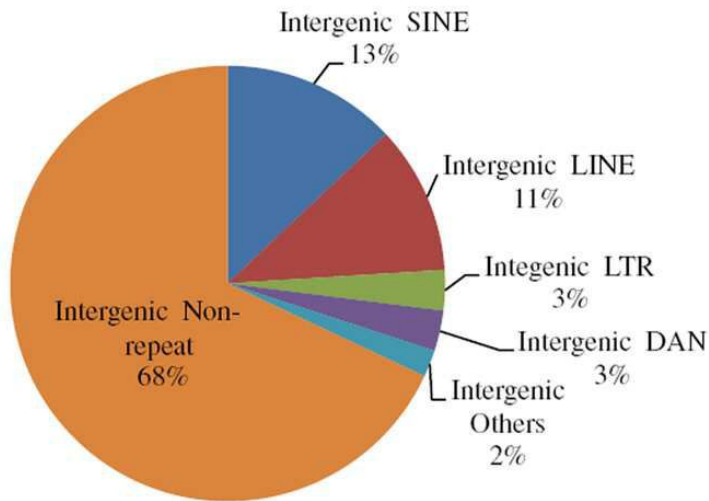


도면4

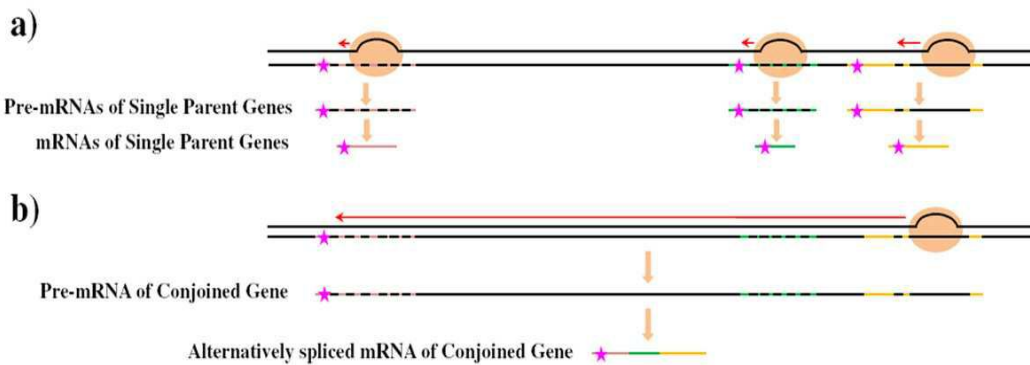
a) **Intragenic exoncreation (215)**



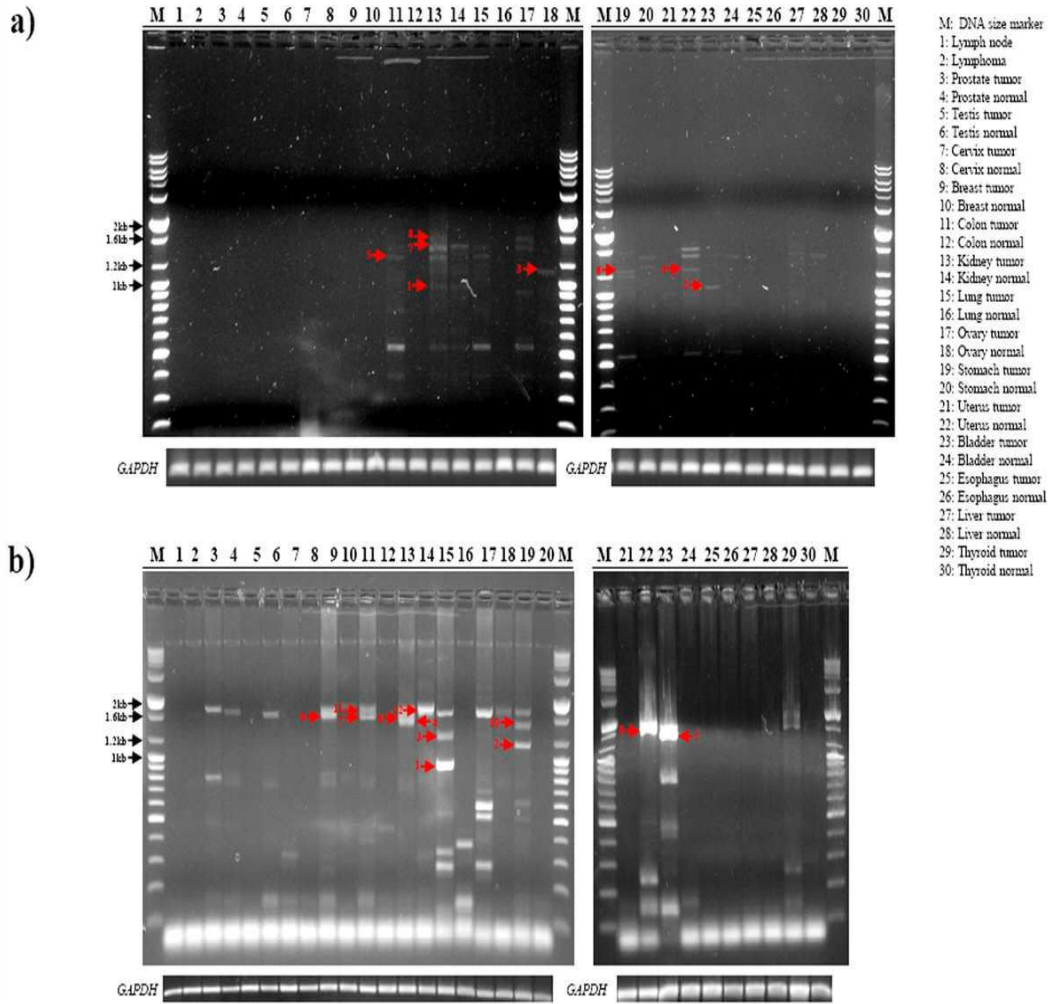
b) **Intergenic exoncreation (482)**



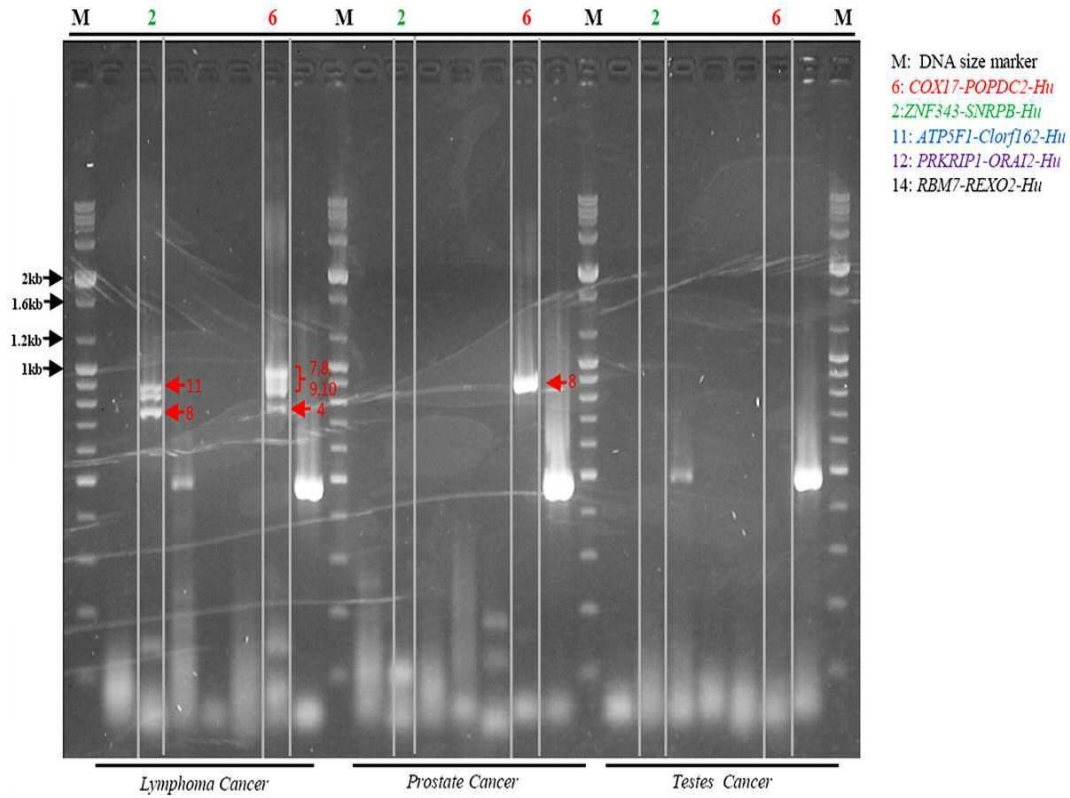
도면5



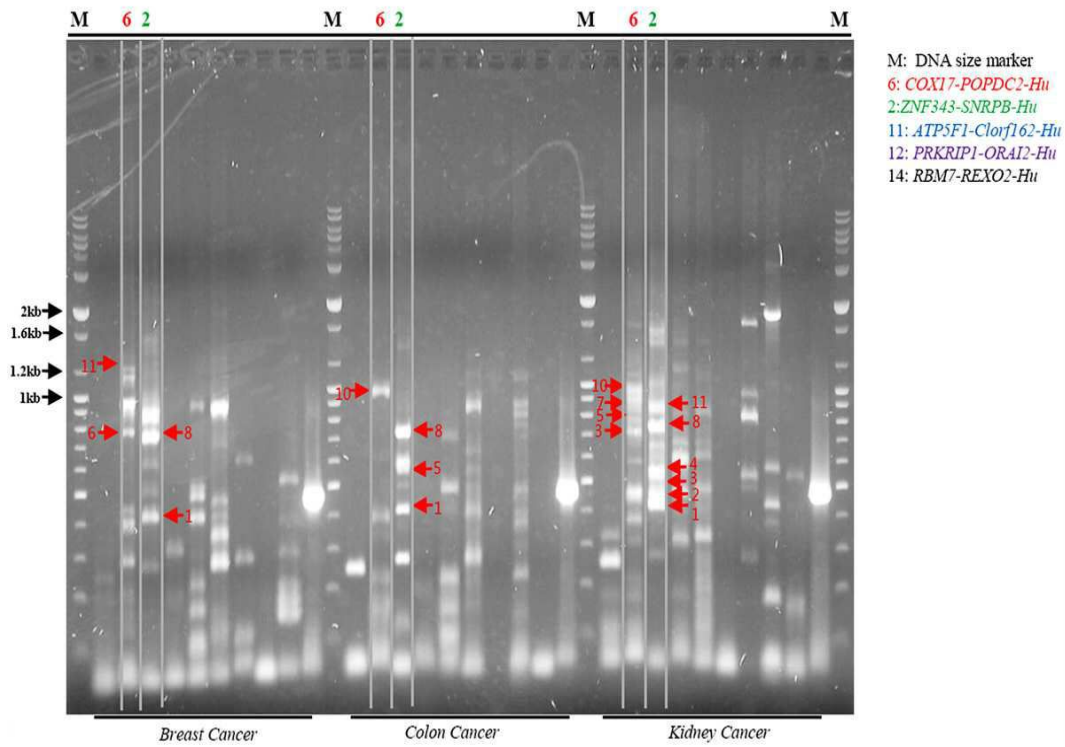
도면6



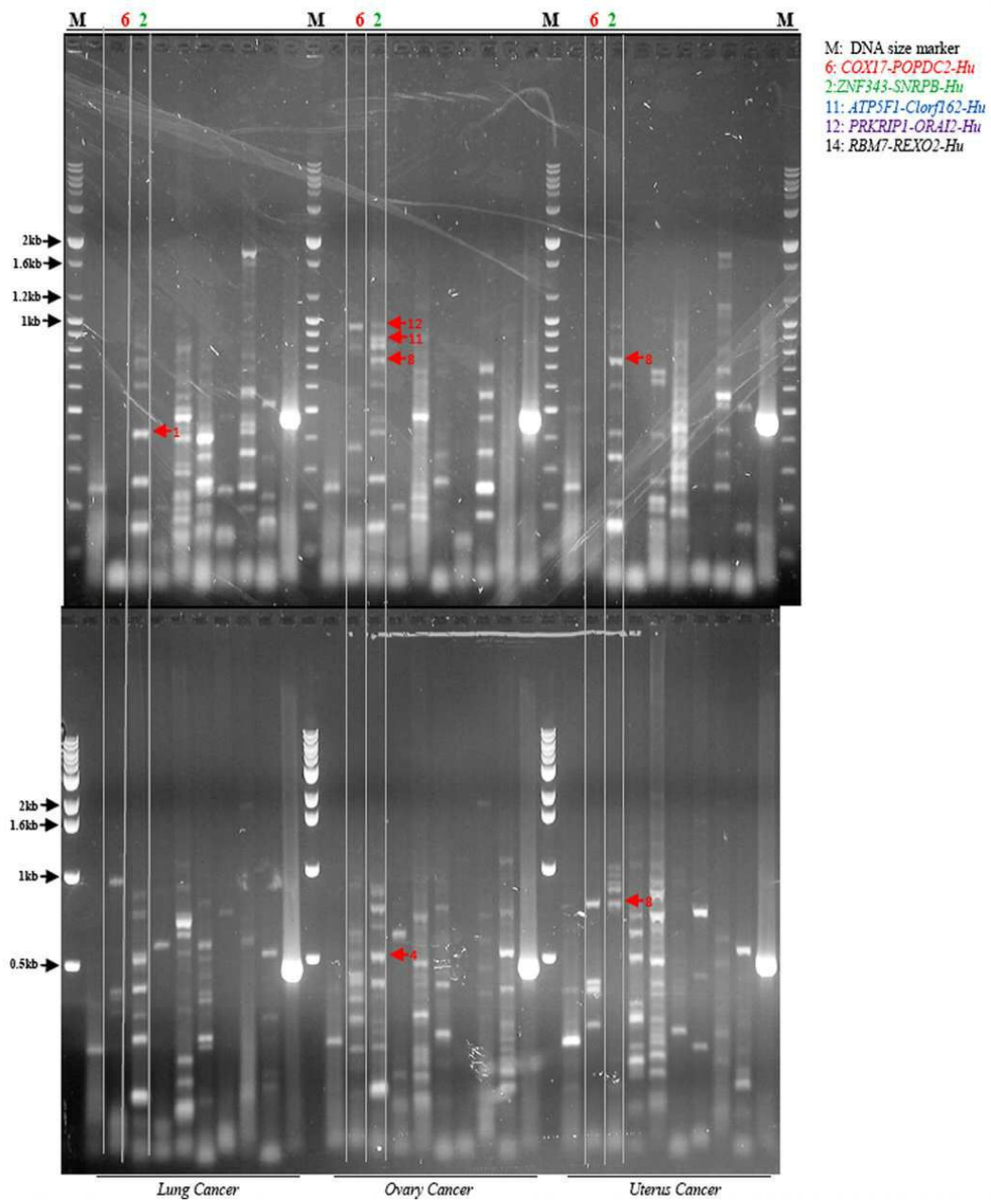
도면7a



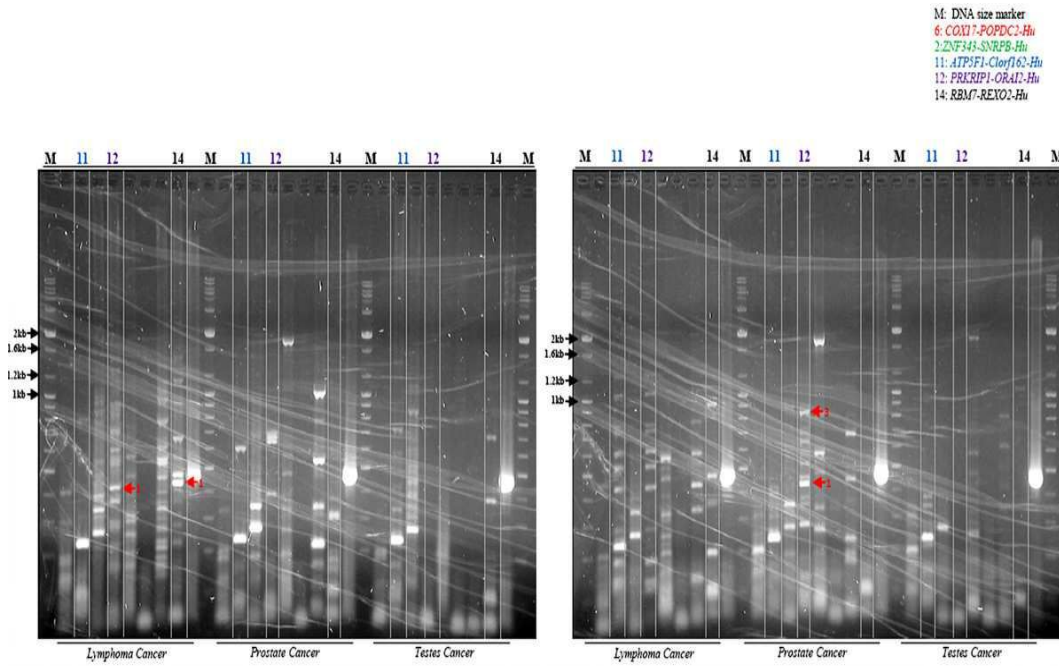
도면7b



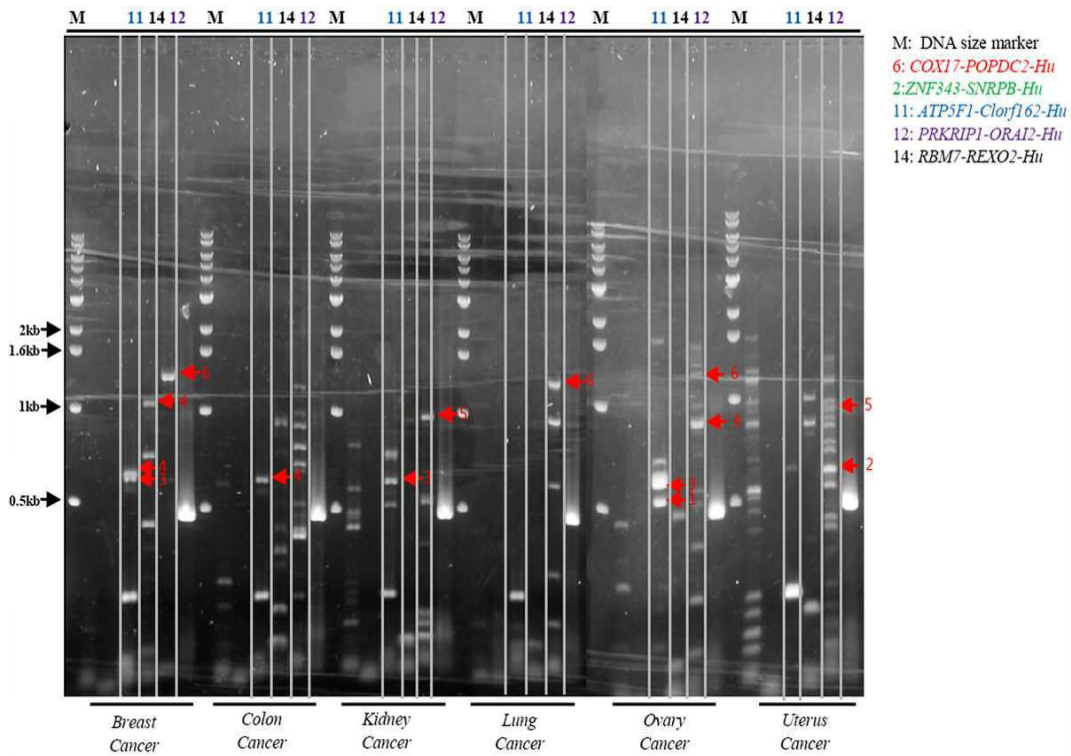
도면7c



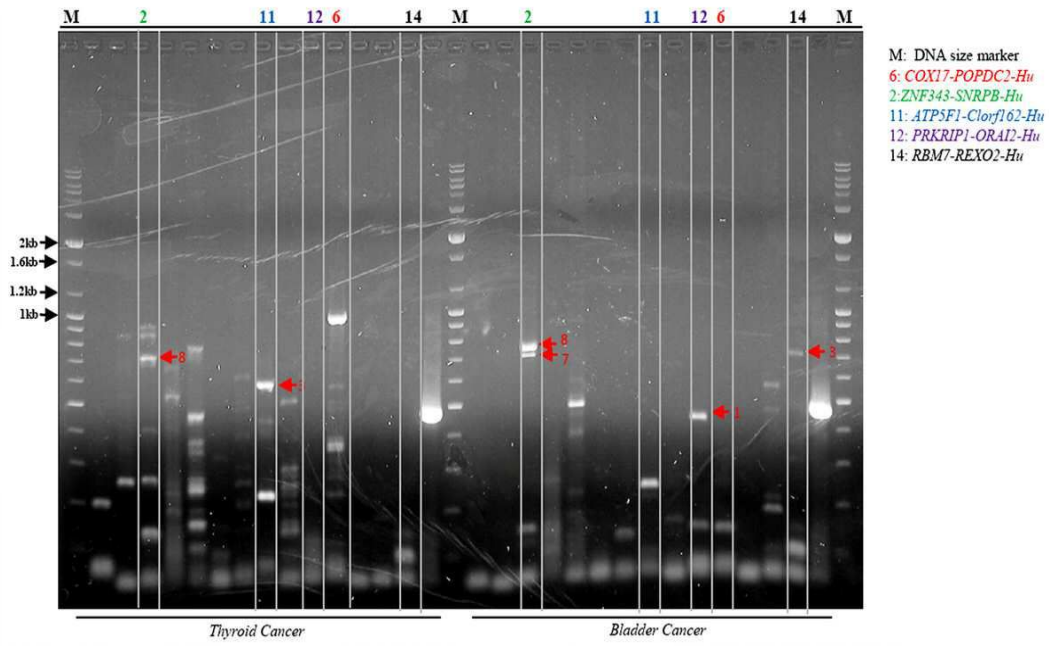
도면7d



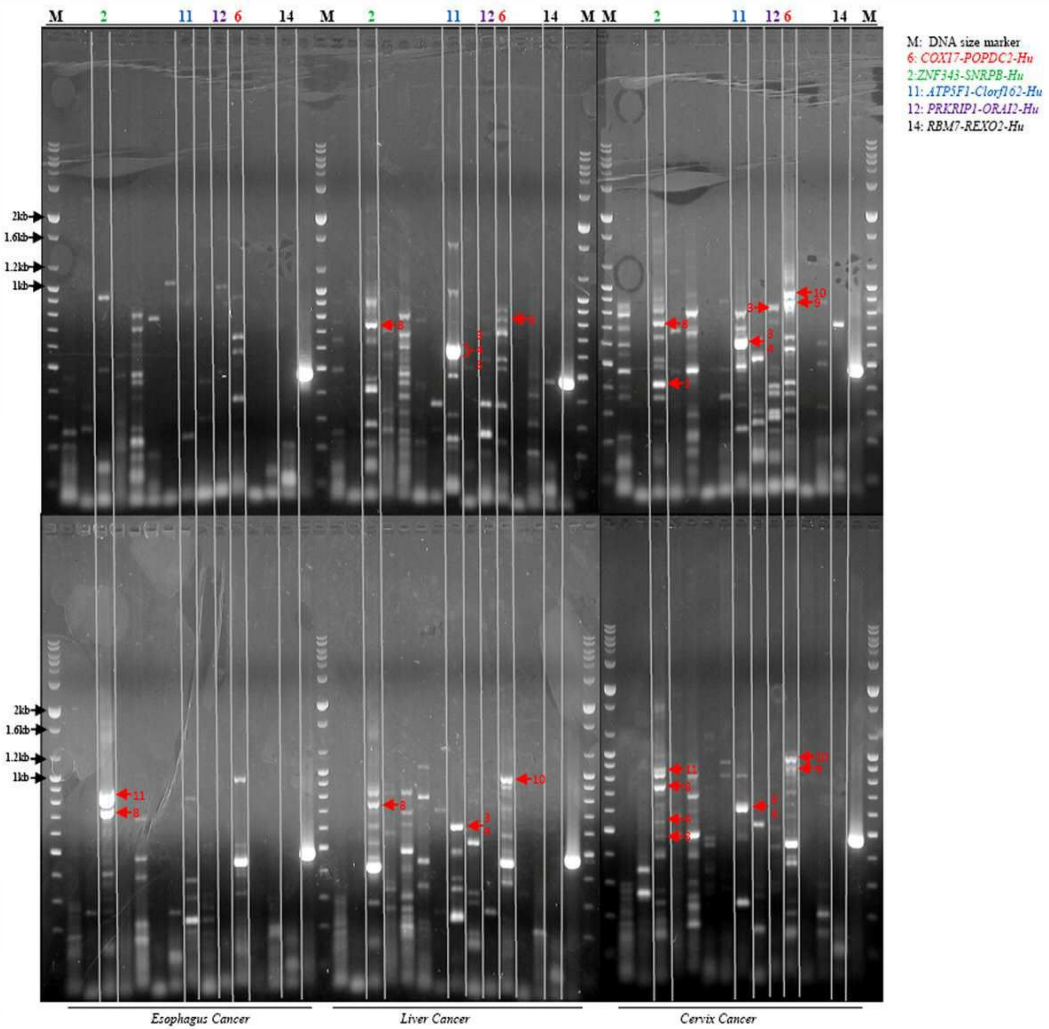
도면7e



도면7f

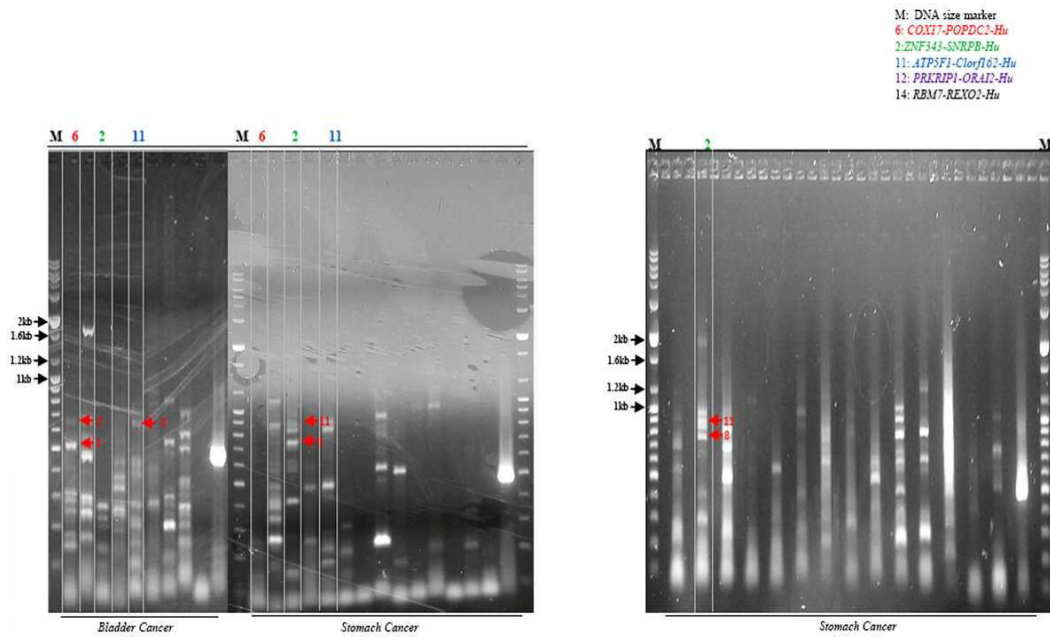


도면7g

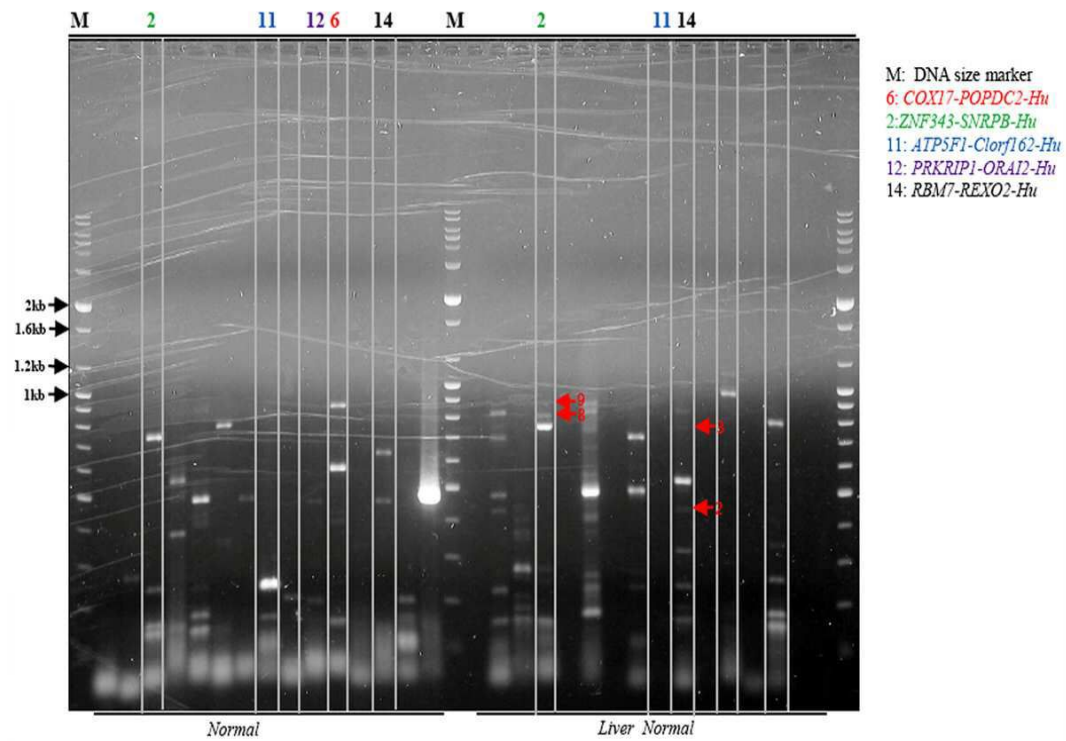




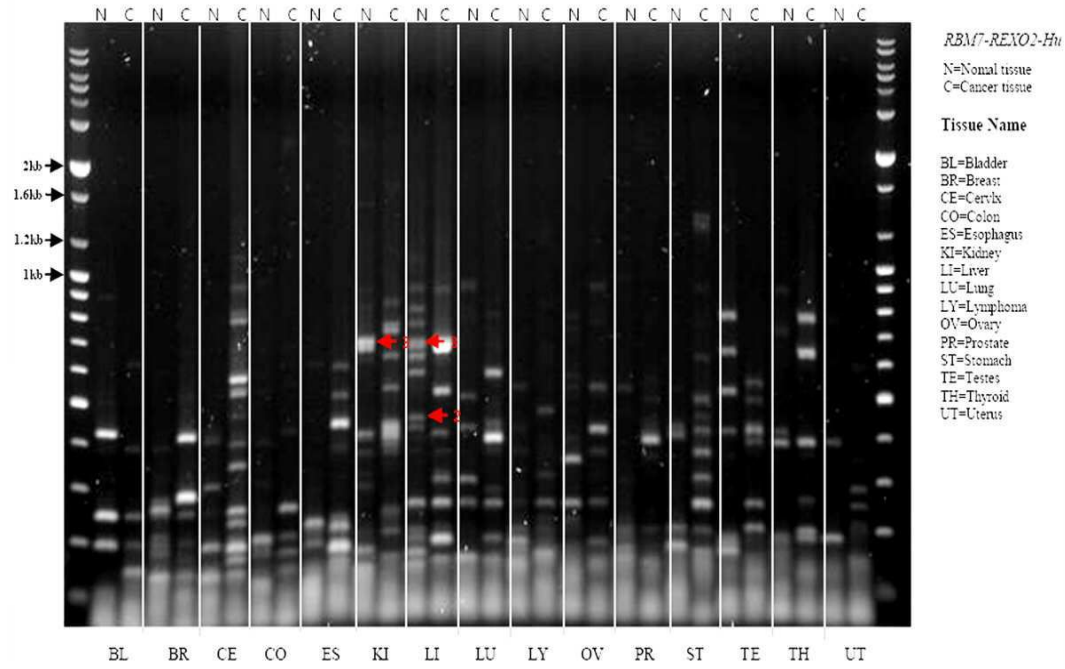
도면7h



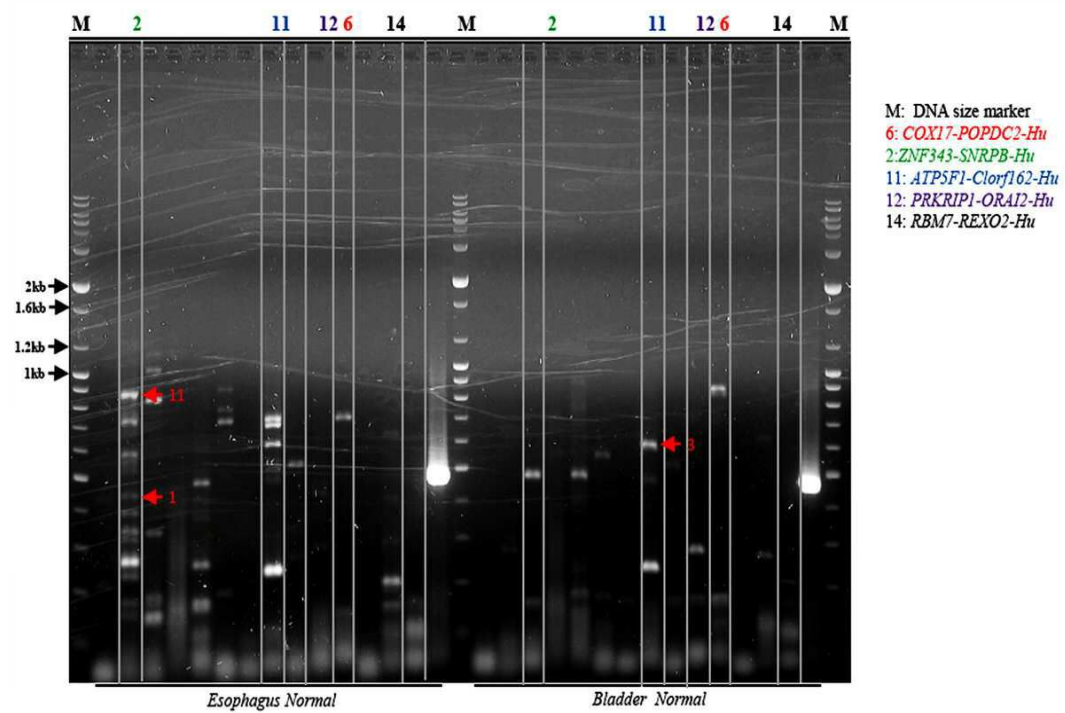
도면7i



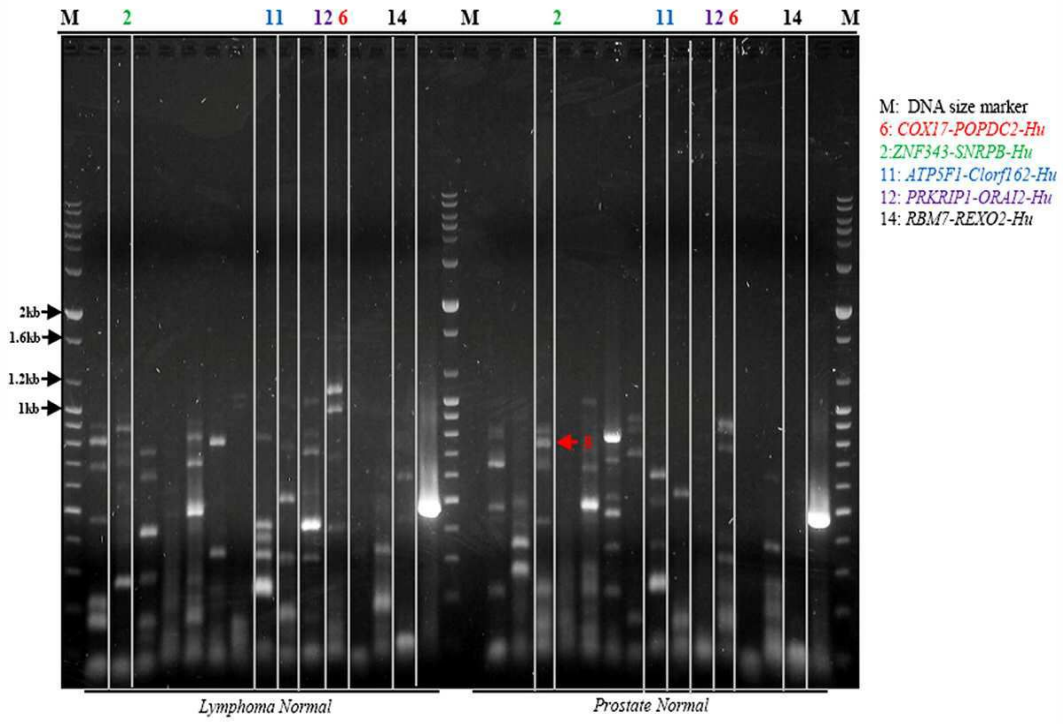
도면7j



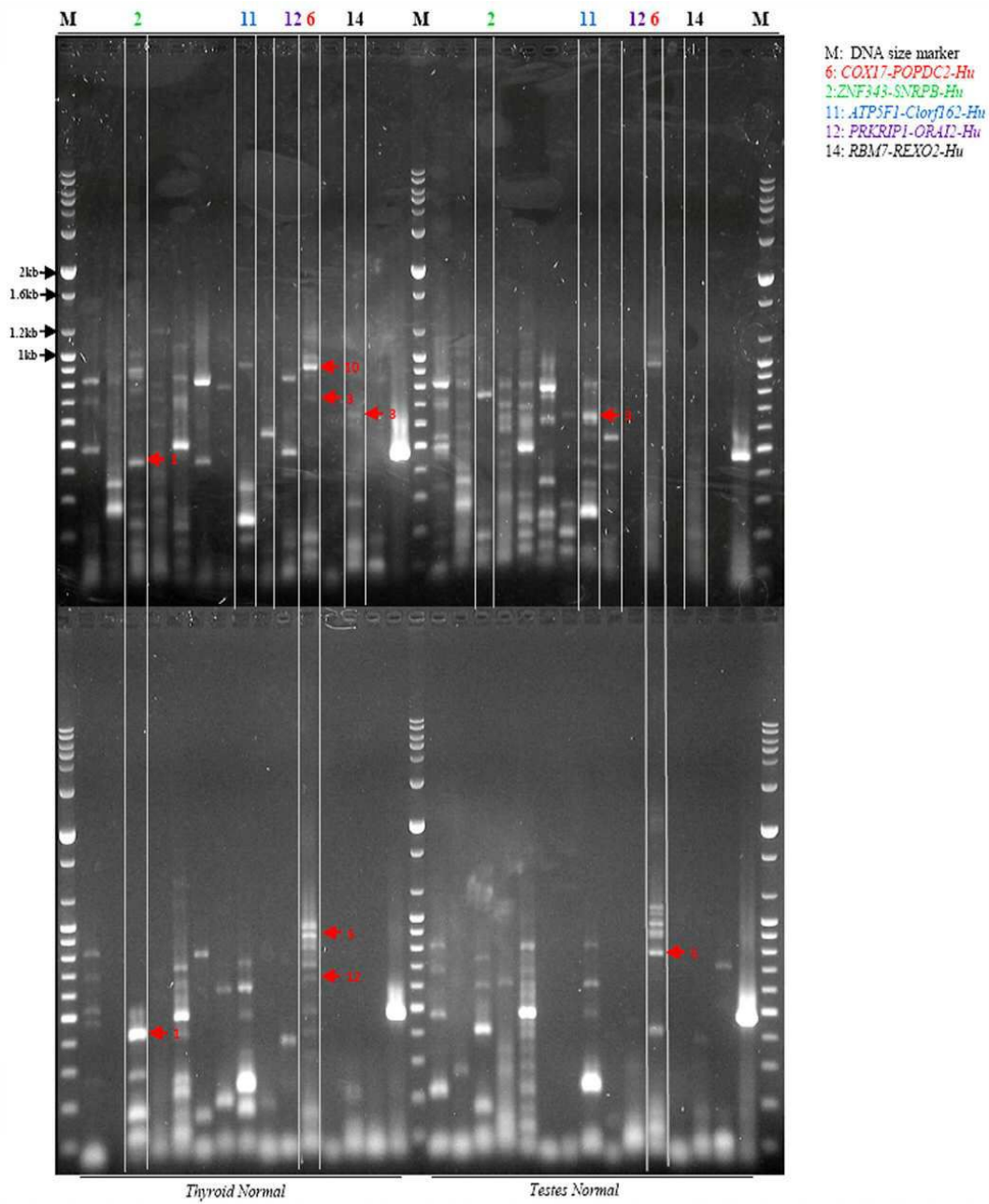
도면7k



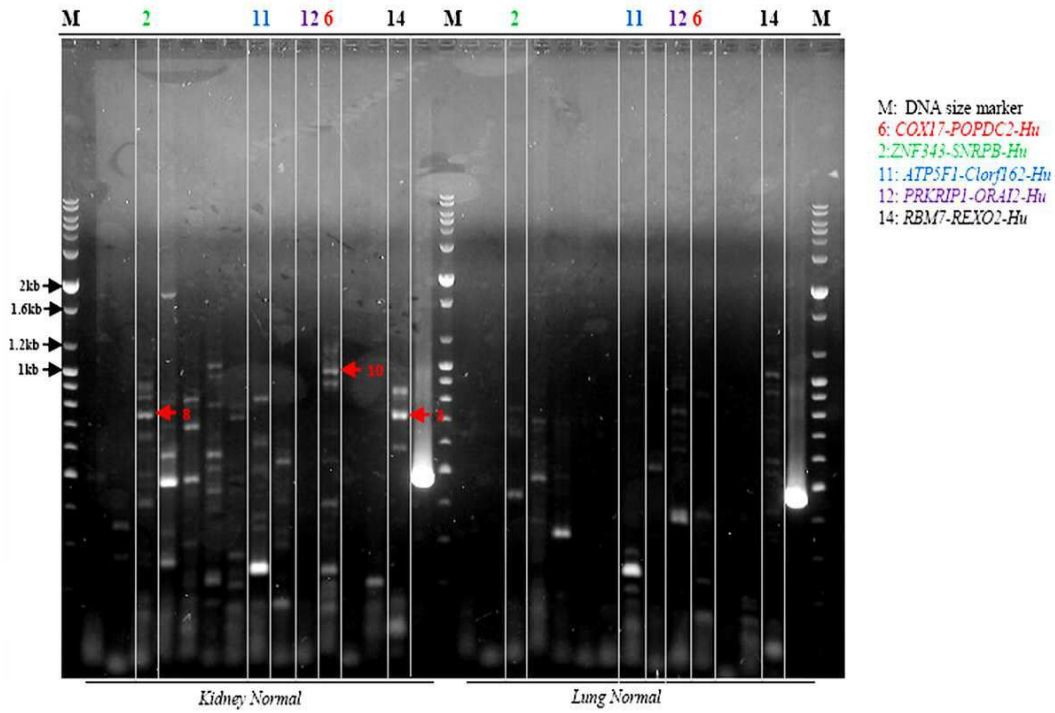
도면71



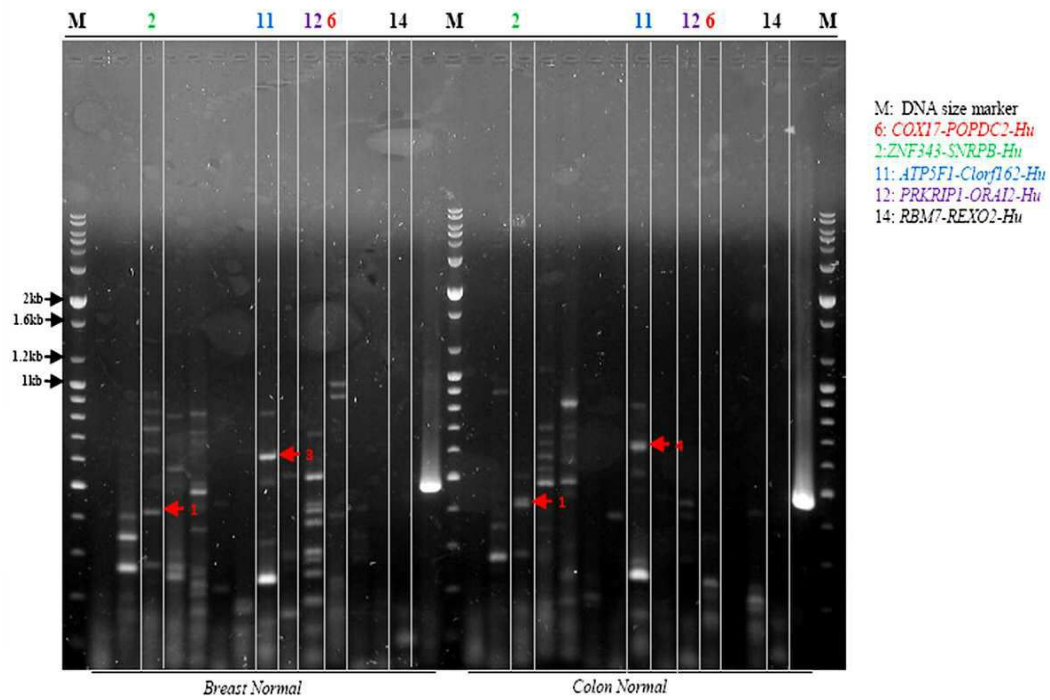
도면7m



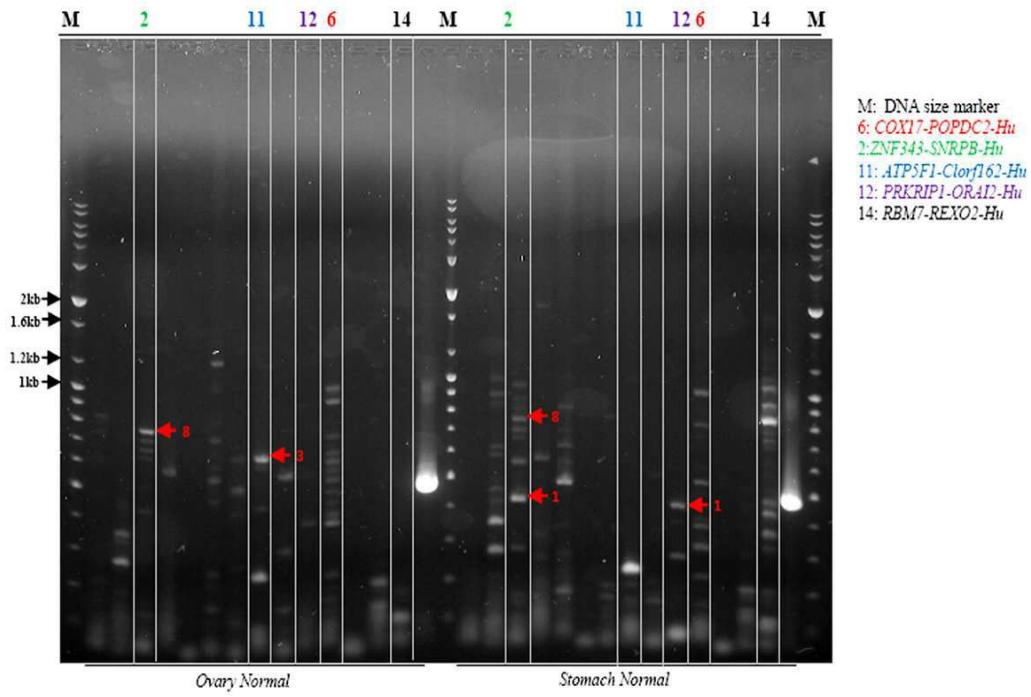
도면7n



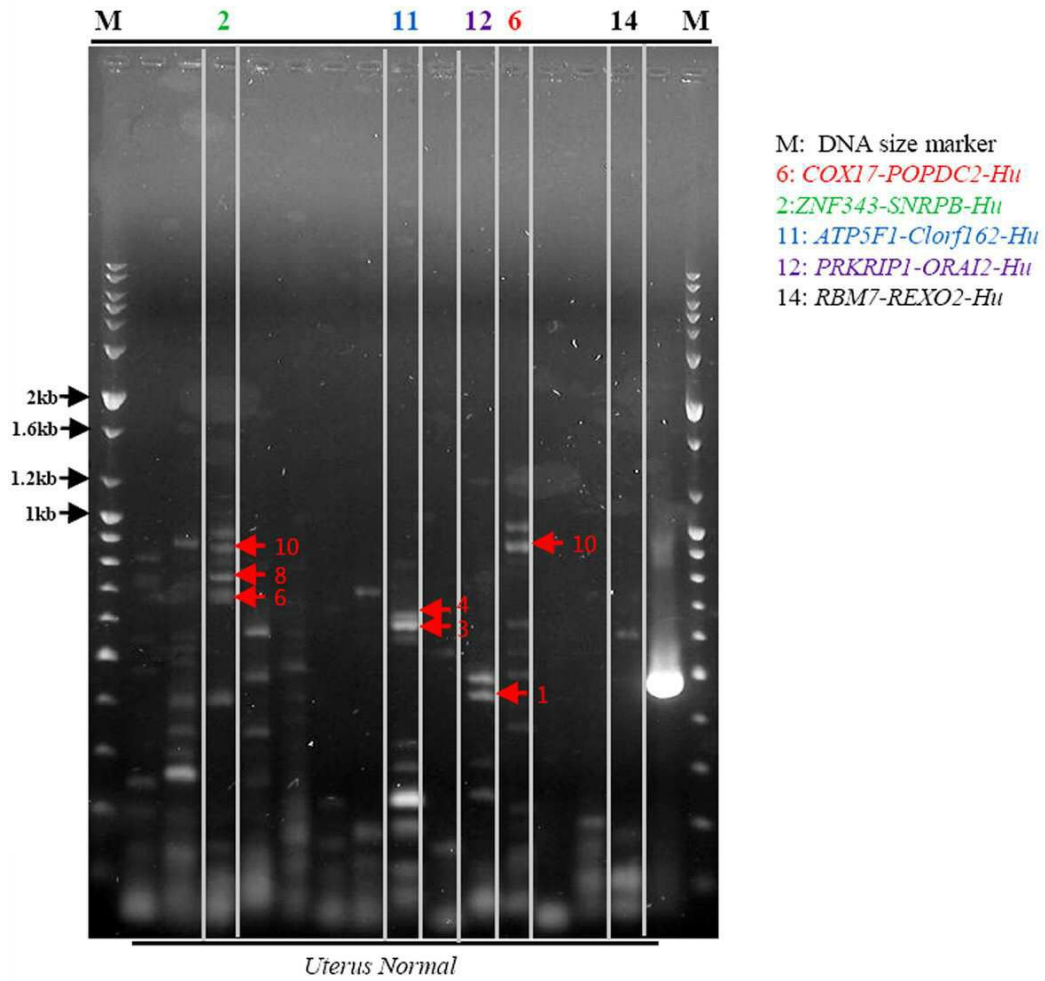
도면7o



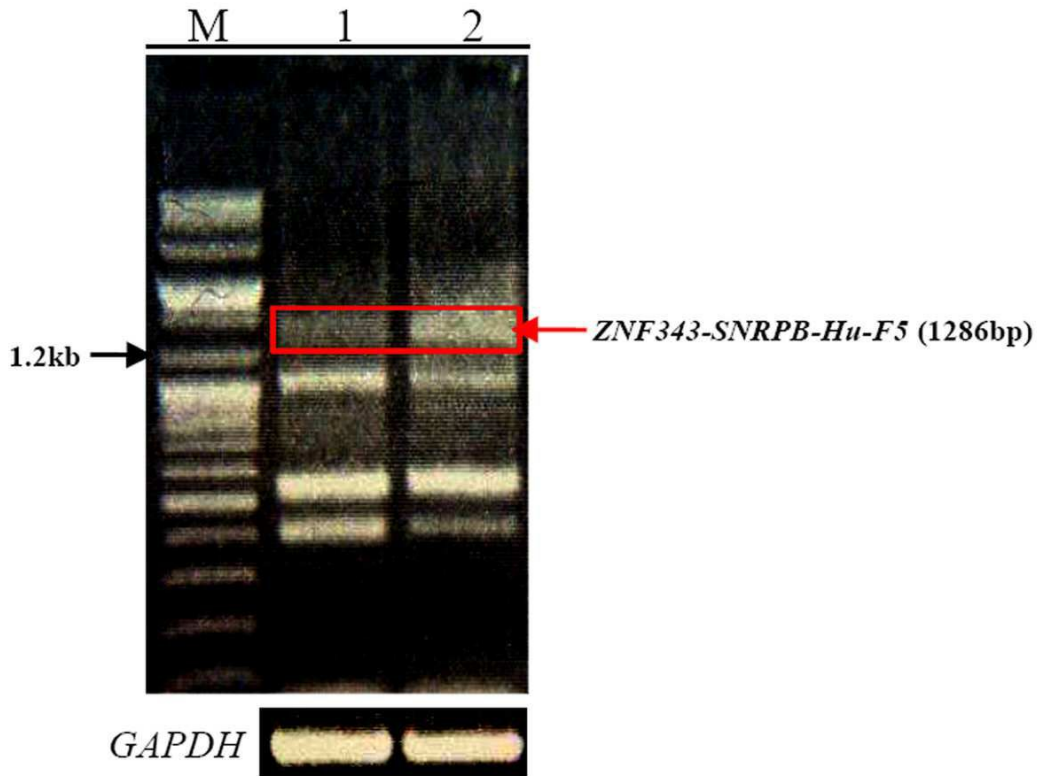
도면7p



도면7q



도면8



M: DNA size marker

1: HEK293 cell total RNA (1.9µg)

2: HEK293 cell total RNA (50ng)

서열목록

- <110> Korea Research Institute of Bioscience and Biotechnology
- <120> Novel human conjoined genes or conjoined gene transcript variants, and a use thereof
- <130> 12p-05-30
- <150> KR2011/0049541
- <151> 2011-05-25
- <160> 57
- <170> KopatentIn 1.71
- <210> 1
- <211> 421
- <212> DNA
- <213> Homo sapiens
- <400> 1



gaaagacttc tggctcccg agaacttct atcaggctct agttacattc agggatgtga 60  
 ctgtgatctt cacagaagca gaatggaaga gactgagtcc agagcagagg aatctataca 120

aagaagtgat gctggagaat tacaggaatc ttctctcatt ggacggtggg caagagcagc 180  
 aagatgctgc agcatattga ttacaggatg aggtgcatcc tgcaggacgg ccggatcttc 240  
 attggcacct tcaaggettt tgacaagcac atgaatttga tctctgtga ctgtgatgag 300  
 ttcagaaaga tcaagccaaa gaactccaaa caagcagaaa gggaagagaa gcgagtcctc 360  
 ggtctggtgc tgcctgcagg ggagaatctg gtctcaatga cagtagaggg acctctccc 420  
 a 421

<210> 2  
 <211> 454  
 <212> DNA

<213> Homo sapiens

<400> 2

gaaagacttc tggctcccg agaacttct atcaggctct gccacagcct ctcgagtagt 60  
 tgggattaca gacggtgggc aagagcagca agatgctgca gcatattgat tacaggatga 120  
 ggtgcatcct gcaggacggc cggatcttca ttggcacctt caaggctttt gacaagcaca 180  
 tgaatttgat cctctgtgac tgtgatgagt tcagaaagat caagtgtgtg aggggggaga 240  
 gcttgagttt ctggcctcat tgcctatttg taaagctgta gcctaaaggg ctctccaat 300  
 gcagcctggt ctgaaactgc ttttaagaagc ctctgaccct cttcagggcc aaagaactcc 360  
 aaacaagcag aaagggaaga gaagcgagtc ctcggtctgg tgctgctgcg aggggagaat 420

ctggtctcaa tgacagtaga gggacctcct ccca 454

<210> 3  
 <211> 479  
 <212> DNA

<213> Homo sapiens

<400> 3

gaaagacttc tggctcccg agaacttct atcaggctct ettctaatga tactgactgc 60  
 cccataaaa aggagggaaa ggcccaata gtggtaccag ttacattcag ggatgtgact 120  
 gtgatcttca cagaagcaga atggaagaga ctgagtccag agcagaggaa tctatacaaa 180  
 gaagtgatgc tggggaatta caggaatctt ctctcattgg acggtgggca agagcagcaa 240  
 gatgctgcag catattgatt acaggatgag gtgcatcctg caggacggcc ggatcttcat 300

tggcaccttc aaggcttttg acaagcacat gaatttgatc ctctgtgact gtgatgagtt 360  
 cagaaagatc aagccaaaga actccaaaca agcagaaagg gaagagaagc gagtctcgg 420  
 tctggtgctg ctgaggggg agaactctgt ctcaatgaca gtagaggac ctctccca 479  
 <210> 4  
 <211> 545  
 <212> DNA  
 <213> Homo sapiens  
 <400> 4  
 gaaagacttc tgtgctcccg agaacttct atcaggtcct agttacattc agggatgtga 60  
 ctgtgatctt cacagaagca gaatggaaga gactgagtcc agagcagagg aatctataca 120  
 aagaagtgat gctggagaat tacaggaatc ttctctcatt ggacggtggg caagagcagc 180  
 aagatgctgc agcatattga ttacaggatg aggtgcatcc tgcaggacgg ccgatcttc 240  
 attggcacct tcaaggcttt tgacaagcac atgaatttga tcctctgtga ctgtgatgag 300  
 ttcagaaaga tcaagtgtgt gaggggggag agcttgagtt tctggcctca ttgcctattt 360  
 gtaaagctgt agcctaaagg gctctcccaa tgcagcctgg tctgaaactg ctttaagaag 420  
 cctctgacce tcttcagggc caaagaactc caaacaagca gaaaggaag agaagcagat 480  
 cctcggtctg gtgctgctgc gaggggagaa tctggtctca atgacagtag agggacctcc 540  
 tcca 545  
 <210> 5  
 <211> 597  
 <212> DNA  
 <213> Homo sapiens  
 <400> 5  
 gaaagacttc tgtgctcccg agcacttct atcaggtacc agttacattc agggatgtga 60  
 ctgtgatctt cacagaagca gaatggaaga gactgagtcc agagcagagg aatctataca 120  
 aagaagtgat gctggagaat tacaggaatc ttctctcatt ggacagaacca aagccagaaa 180  
 tctacacttg ttctctctgc cttctggcct tctctgtca gcagttctc agtcaacatg 240  
 tacttcagat cttctgggc ttatgtgcag aaaateactt ccatccaggg aattctagcc 300  
 cagggcattg gaaacagcag gggcagcagc attcccatc ggtgggcaag agcagcaaga 360  
 tgctgcagca tattgattac aggatgaggt gcatcctgca ggacggccgg atcttcattg 420  
 gcaccttcaa ggcttttgac aagcacatga atttgatcct ctgtgactgt gatgagtcca 480  
 gaaagatcaa gccaaagaac tccaacaag cagaaagga agagaagcga gtctcggtc 540

tgggtgctgct gcgagggggag aatctggtct caatgacagt agagggacct cctccca 597  
 <210> 6  
 <211> 668  
 <212> DNA  
 <213> Homo sapiens  
 <400> 6  
 gaaagacttc tggctcccg agaacttct atcaggtacc agttacattc agggatgtga 60  
 ctgtgatctt cacagaagca gaatggaaga gactgagtcc agagcagagg aatctataca 120  
  
 aagaagtgat gctggagaat tacaggaatc ttctctcatt ggatggagtt tcgctcctgt 180  
 tgcccaggct ggagtgcaat cgtgcaatct cggtcacca caacctttgc ctcccgggtt 240  
 caagcaattc tctgccaca gcctctcgag tagttgggat tacagacggt gggcaagagc 300  
 agcaaatgc tgcagcatat tgattacagg atgaggtgca tcctgcagga cggccggatc 360  
 ttcatggca cctcaaggc tttgacaag cacatgaatt tgatcctctg tgactgtgat 420  
 gatttcagaa agatcaagt tgtgaggggg gagagcttga gtttctggcc tcattgccta 480  
 tttgtaaagc tgiagcctaa agggctctcc caatgcagcc tggctgaaa ctgctttaag 540  
  
 aagcctctga cctctttcag ggccaaagaa ctccaaacaa gcagaaagg aagagaagcg 600  
 agtcctcggt ctggtgctgc tgcgagggga gaatctagtc tcaatgacag tagagggacc 660  
 tcctccca 668  
 <210> 7  
 <211> 707  
 <212> DNA  
 <213> Homo sapiens  
 <400> 7  
 gaaagacttc tggctcccg agaacttct atcaggtcct ggctgcaggg aaacaagctg 60  
 ggctttttat aattaaggtt ggaagaagtc accacaggca gcagaactcc atcttgagat 120  
 gaaataacat ctacctggac ctctggcaga attcaaggc acacactggg ctgactctgg 180  
  
 cgccatgatg ttgccttacc cttcagcact gggagatcaa tactgggaag agatcttctgct 240  
 tccaaagaat ggggaaaatg tagagactat gaagaaattg acccaaaatc ataaagcgaa 300  
 aggcttgctt tctaatgata ctgactgcc ccagaaaag gagggaaagg cccaatagt 360  
 gaagcagaat ggaagagact gactccagag cagaggaatc tatacaaaga agtgatgctg 420  
 gagaattaca ggaatcttct ctattggac ggtgggcaag agcagcaaga tgctgcagca 480  
 tattgattac aggatgaggt gcatcctgca ggacggccgg atcttcattg gcacctcaa 540

ggcttttgac aagcacatga atttgatcct ctgtgactgt gatgagttca gaaagatcaa 600

gccaagaac tccaacaag cagaaagga agagaagcga gtcctcggtc tggctgctgct 660

gcgaggggag aatctggtct caatgacagt agaggacct cctccca 707

<210> 8

<211> 747

<212> DNA

<213> Homo sapiens

<400> 8

gaaagacttc tgtgctcccg agaacttct atcaggctct ggctgcaggg aaacaagctg 60

ggctttttat aattaaggtt ggaagaagtc accacaggca gcagaactcc atcttgagat 120

gaaataacat ctacctggac ctctggcaga atttcaaggc acacactggg ctgactctgg 180

cgccatgatg ttgccttacc cttcagcact gggagatcaa tactgggaag agattttgct 240

tccaagaat ggggaaaatg tagagactat gaagaaattg acccaaatc ataaagcgaa 300

aggcttgcct tctaatgata ctgactgcc ccagaaaaag gagggaaagg cccaatagt 360

ggtaccagtt acattcaggg atgtgactgt gatcttcaca gaagcagaat ggaagagact 420

gagtccagag cagaggaatc tatacaaaga agtgatgctg gagaattaca ggaatcttct 480

ctcattggac ggtgggcaag agcagcaaga tgctgcagca tattgattac aggatgaggt 540

gcatcctgca ggacggccgg atcttcattg gcacctcaa ggcttttgac aagcacatga 600

atttgatcct ctgtgactgt gatgagttca gaaagatcaa gccaagaac tccaacaag 660

cagaaagga agagaagcga gtcctcggtc tggctgctgct gcgaggggag aatctggtct 720

caatgacagt agaggacct cctccca 747

<210> 9

<211> 832

<212> DNA

<213> Homo sapiens

<400> 9

gaaagacttc tgtgctcccg agaacttct atcaggctct ggctgcaggg aaacaagctg 60

ggctttttat aattaaggtt ggaagaagtc accacaggca gcagaactcc atcttgagat 120

gaaataacat ctacctggac ctctggcaga atttcaaggc acacactggg ctgactctgg 180

cgccatgatg ttgccttacc cttcagcact gggagatcaa tactgggaag agattttgct 240

tccaagaat ggggaaaatg tagagactat gaagaaattg acccaaatc ataaagcgaa 300

aggtgagaat tgttgaggga gccggaggga tgaagaggaa gccaaggctc caggttcacc 360  
 tgaaaatgct tcagaggfca cctctgagct tgccttctaa tgatactgac tgccccaga 420  
 aaaaggaggg aaaggcccaa atagtggtag cagttacatt cagggatgtg actgtgatct 480  
 tcacagaagc agaatggaag agactgagtc cagagcagag gaatctatac aaagaagtga 540  
 tgctggagaa ttacaggaat cttctctcat tggacggtag gcaagagcag caagatgctg 600  
 cagcatattg attacaggat gaggtagcat ctgcaggacg gccggatctt cattggcacc 660  
  
 ttcaaggctt ttgacaagca catgaatttg atcctctgtg actgtgatga gttcagaaag 720  
 atcaagccaa agaactcaa acaagcagaa agggaagaga agcagatcct cggctctggtg 780  
 ctgctgcgag gggagaatct ggtctcaatg acagtagagg gacctctcc ca 832  
  
 <210> 10  
 <211> 861  
 <212> DNA  
 <213> Homo sapiens  
 <400> 10  
  
 gaaagacttc tigtctcccg agaacttctc atcaggctct ggctgcaggg aaacaagctg 60  
 ggctttttat aattaaggtt ggaagaagtc accacaggca gcagaactcc atcttgagat 120  
 gaaataacat ctacctggac ctctggcaga atttcaagc acacactggg ctgactctgg 180  
  
 cgccatgatg ttgccttacc cttcagcact gggagatcaa tactgggaag agattttgct 240  
 tccaaagaat ggggaaaatg tagagactat gaagaaattg acccaaaatc ataaagcgaa 300  
 agaagcagaa tggaagagac tgagtccaga gcagaggaat ctatacaaag aagtgatgct 360  
 ggagaattac aggaatcttc tctcattggt cccaggccc cgcctctgct cggcagaagc 420  
 tccgctccgg gaccggaact ccagggctag tgagctggac cggaagtagg tttctaccg 480  
 accgaatttt acgtggtgct gcatttccgg tagcggcggc gggaaatcgg ctgtgggaga 540  
 gaggctagcc ctctgaggag gcgaatccgg cgggtatcag agccatcaga accgccacca 600  
  
 tgacggtggg caagagcagc aagatgctgc agcatattga ttacaggatg aggtgcatcc 660  
 tgcaggacgg ccggatcttc attggcacct tcaaggcttt tgacaagcac atgaattga 720  
 tcctctgtga ctgtgatgag ttcagaaaga tcaagccaaa gaactccaaa caagcagaaa 780  
 gggaagagaa gcgagtcttc ggtctggtgc tgctgcgagg ggagaatctg gttctcaatga 840  
 cagtagaggg acctctctcc a 861  
  
 <210> 11  
 <211> 871  
 <212> DNA

<213> Homo sapiens  
 <400> 11  
 gaaagacttc tgtgctcccg agaacttctc atcaggtcct ggctgcaggg aaacaagctg 60  
  
 ggctttttat aattaaggtt ggaagaagtc accacaggca gcagaactcc atcttgagat 120  
 gaaataacat ctacctggac ctctggcaga atttcaaggc acacactggg ctgactctgg 180  
 cgccatgatg ttgccttacc cttcagcact gggagatcaa tactgggaag agattttgct 240  
 tccaaagaat ggggaaaatg tagagactat gaagaaattg acccaaatc ataaagcgaa 300  
 aggcttgctt tctaatagata ctgactgccc ccagaaaaag gagggaaaagg cccaatagt 360  
 ggtaccagtt acattcaggg atgtgactgt gatcttcaca gaagcagaat ggaagagact 420  
 gagtccagag cagaggaatc tatacaaaga agtgatgctg gagaattaca ggaatcttct 480  
  
 ctcatggac ggtgggcaag agcagcaaga tctgcagca tattgattac aggatgaggt 540  
 gcatcctgca ggacggccgg atcttcattg gcaccttcaa ggcttttgac aagccatga 600  
 atttgatcct ctgtgactgt gatgagtcca gaaagatcaa gtgtgtgagg ggggagagct 660  
 tgagtttctg gectcattgc ctatttgtaa agctgtagcc taaagggtc tcccaatgca 720  
 gcctggctctg aaactgcttt aagaagcctc tgaccctctt cagggccaaa gaactccaaa 780  
 caagcagaaa ggaagagaaa gcgagtctc ggtctggtgc tgctgagagg ggagaatctg 840  
 gtctcaatga cagtagaggg acctctctcc a 871  
  
 <210> 12  
 <211> 947  
 <212> DNA  
 <213> Homo sapiens  
 <400> 12  
 gaaagacttc tgtgctcccg agaacttctc atcaggtcct ggctgcaggg aaacaagctg 60  
 ggctttttat aattaaggtt ggaagaagtc accacaggca gcagaactcc atcttgagat 120  
 gaaataacat ctacctggac ctctggcaga atttcaaggc acacactggg ctgactctgg 180  
 cgccatgatg ttgccttacc cttcagcact gggagatcaa tactgggaag agattttgct 240  
 tccaaagaat ggggaaaatg tagagactat gaagaaattg acccaaatc ataaagcgaa 300  
 aggtgagaat tgttgagggg gccggagggg tgaagaggaa gccaaaggctc caggttcacc 360  
  
 tgaaaatgct tcagaggcca cctctgagtt aagtgactac ttccttacc ctaaggetca 420  
 gaagagctgc tcaggagcgc tcatggtggg ggtggggggg tgggggggtc tggtttctca 480  
 tagcttcttt ttcttctcct aggcttgctt tctaatagata ctgactgccc ccagaaaaag 540

gagggaaagg cccaaatagt ggtaccagt acattcaggg atgtgactgt gatcttcaca 600  
 gaagcagaat ggaagagact gagtccagag cagaggaatc tatacaaaga agtgatectg 660  
 gagaattaca ggaatcttct ctcatctggac ggtgggcaag agcagcaaga tgctgcagca 720  
 tattgattac aggatgaggt gcatcctgca ggacggccgg atcttcattg gcaccttcaa 780

ggcttttgac aagcacatga atttgatcct ctgtgactgt gatgagttca gaaagatcaa 840  
 gccaaagaac tccaaacaag cagaaaggga agagaagcga gtcctcggtc tgggtgctgct 900  
 gcgaggggag aatctggctct caatgacagt agagggacct cctccca 947

<210> 13

<211> 960

<212> DNA

<213> Homo sapiens

<400> 13

ttctgtgctc ccgagaactt cctatcaggt accagttaca ttcagggatg tgactgtgat 60  
 cttcacagaa gcagaatgga agagactgag tccagagcag aggaatctat acaaagaagt 120  
 gatgctggag aattacagga atcttctctc attggacggt gggcaagagc agcaagatgc 180

tgcagcatat tgattacagg atgaggtgca tctctcagga cggccggatc ttcattggca 240  
 ctttcaagge ttttgacaag cacatgaatt tgatcctctg tgactgtgat gatttcagaa 300  
 agatcaagcc aaagaactcc aaacaagcag aaaggggaaga gaagcgagtc ctcggtctgg 360  
 tgctgctgcg aggggagaat ctgggtctcaa tgacagtaga gggacctcct cccaaagata 420  
 ctggtattgc tcgagtcca cttgctggag ctgccggggg cccagggatc ggcagggctg 480  
 ctggcagagg aatcccagct ggggttcca tgccccaggc tcctgcagga cttgctgggc 540  
 cagtccgtgg ggttggcggg ccatccaac aggtgatgac cccacaagga agaggtactg 600

ttgcagccgc tgacgtgct gccacagcca gtattgccgg ggctccaacc cagtaccac 660  
 ctggccgtgg gggctctccc ccacctatgg gccgaggagc acccctcca ggcattgatg 720  
 gcccactec tggatgaga cctctatgg gtcccccaat ggggatcccc cctggaagag 780  
 ggactccaat gggcatgccc cctccgggaa tgcggcctcc tcccctggg atgcgagggc 840  
 cccctcccc gggatgccc ccaccaagge cctagactca tcttggcct cctcagctcc 900  
 ctgcctgttt cccgtaagge tgtacatagt ccttttatct ccttgggcc tatgaaactg 960

<210> 14

<211> 1019

<212> DNA  
 <213> Homo sapiens  
 <400> 14

ttctgtgctc ccgagaactt cctatcaggc ttgccttcta atgatactga ctgccccag 60  
 aaaaaggagg gaaaggccca aatagtggta ccagttacat tcagggatgt gactgtgatc 120  
 ttcacagaag cagaatggaa gagactgagt ccagagcaga ggaatctata caaagaagtg 180  
 atgctggaga attacaggaa tcttctctca ttggacggtg ggcaagagca gcaagatgct 240  
 gcagcatatt gattacagga tgaggatgcat cctgcaggac ggccggatct tcattggcac 300  
 cttaaggct ttgacaagc acatgaattt gatcctctgt gactgtgatg agttcagaaa 360

gatcaagcca aagaactcca aacaagcaga aagggaagag aagcgagtcc tcggtctggt 420  
 gctgtcgcga ggggagaatc tggcttcaat gacagtagag ggacctctc ccaaagatac 480  
 tggatttgct cgagttccac ttgctggagc tgccgggggc ccagggatcg gcagggtgctg 540  
 tggcagagga atcccagctg gggttcccat gccccaggct cctgcaggac ttgctgggcc 600  
 agtccgtggg gttggcgggc catccaaca ggtgatgacc ccacaaggaa gaggtactgt 660  
 tgcagccgct gcagctgctg ccacagccag tattgcccgg gctccaacce agtaccacc 720  
 tggccgtggg ggtcctcccc cacatatggg ccgaggagca ccccctccag gcatgatggg 780

cccacctctt ggtatgagac ctctatggg tcccccaatg gggatcccc ctggaagagg 840  
 gactccaatg ggcattgccc ctccgggaat gcggcctct ccccctggga tgcgagggcc 900  
 ccctccccg ggaatgcgcc caccaaggcc ctgactcat cttggcctc ctgagctccc 960  
 tgctgtttc ccgtaaggct gtacatagtc cttttatctc cttgtggcct atgaaactg 1019

<210> 15  
 <211> 1106  
 <212> DNA  
 <213> Homo sapiens  
 <400> 15

ttctgtgctc ccgagaactt cctatcaggc accagttaca ttcagggatg tgactgtgat 60  
 cttcacagaa gcagaatgga agagactgag tccagagcag aggaatctat acaaagaagt 120

gatgctggag aattacagga atcttctctc attggacggg gggcaagagc agcaagatgc 180  
 tgcagcatat tgattacagg atgaggtgca tctgcagga cggccggatc tcattggca 240  
 cttcaaggc ttttgacaag cacatgaatt tgatcctctg tgactgtgat gaggtcagaa 300  
 agatcaagcc aaagaactcc aaacaagcag aaagggaaga gaagcgagtc ctcggtctgg 360  
 tgctgtgctg agggggagaat ctggtctcaa tgacagtaga gggacctct ccaaagata 420



ctggtattgc tcgagttcca cttgctggag ctgccggggg cccagggatc ggcagggctg 480  
ctggcagagg aatcccagct ggggttcca tgccccaggc tctgcagga cttgctgggc 540

cagtcctggg ggttggcggg ccateccaac aggtgatgac cccacaagga agaggtactg 600  
ttgcagccgc tgcagctgct gccacagcca gtattgccgg ggctccaacc cagtaccac 660  
ctggccgtgg gggctctccc ccacctatgg gccgaggagc acccctcca ggcattgatg 720  
gcccacctcc tggatgaga cctctatgg gtccccaat ggggatcccc cctggaagag 780  
ggactccaat gggcatgccc cctccgggaa tgcggcctcc tcccctggg atgcgaggcc 840  
ttctttgacc cttggccaca gagtatgaa gtagctccgc agaggcgtgg gctcgattcc 900  
tcaggccac gttaccacag acctgtttgt ttcttatgct gttgttcgtg gagtctcatg 960

ggattgtctg gtttccctta cagggecccc tccccggga atgcgccac caaggccta 1020  
gactcatctt ggcctctc agctecctgc ctgttcccg taaggctgta catagtcctt 1080  
ttatctcctt gtggcctatg aaactg 1106

<210> 16  
<211> 1230  
<212> DNA  
<213> Homo sapiens  
<400> 16

ttctgtgctc ccgagaactt cctatcaggf accagttaca ttcagggatg tgactgtgat 60  
cttcacagaa gcagaatgga agagactgag tccagagcag aggaatctat acaaagaagt 120  
gatgctggag aattacagga atcttctctc attggacggt gggcaagagc agcaagatgc 180

tgcagcatat tgattacagg atgaggtgca tctgcagga cggccggatc ttcattggca 240  
ccttcaagge ttttgacaag cacatgaatt tgatcctctg tgactgtgat gatttcagaa 300  
agatcaagtg tglgaggggg gagagcttga gtttctggcc tcattgccta tttgtaaagc 360  
tgtagcctaa agggctctcc caatgcagcc tggctgaaa ctgctttaag aagcctctga 420  
ccctcttcag ggccaaagaa ctccaacaa gcagaaaggg aagagaagcg agtcctcggt 480  
ctggtgctgc tgcgagggga gaatctggtc tcaatgacag tagagggacc tctcccaaa 540  
gatactgga ttgctcgagt tccacttctt ggagctgccg ggggcccagg gatcggcagg 600

gctgctggca gaggaatccc agctgggggt cccatgcccc aggctcctgc aggacttctt 660  
gggccagtcc gtggggttgg cgggcatcc caacaggtga tgacccaca aggaagaggt 720  
actgttgcag ccctgcagc tgctgccaca gccagtattg cggggctcc aacceagtac 780  
ccacctggcc gtgggggtcc tccccacct atgggcccag gagcaccccc tccaggcatg 840

atgggcccac ctctgggat gagacctcct atgggtcccc caatggggat cccccctgga 900

agagggactc caatgggcat gccccctccg ggaatgcggc ctctccccc tgggatgcga 960

ggccttcttt gacccttggc cacagagtat ggaagtagct cgcagaggc gtgggctcga 1020

ttcctcaggg ccacgttacc acagacctgt ttgtttctta tgctgttgtt cgtggagtct 1080

catgggattg tctggtttcc cttacagggc cccctccccc gggaatgcgc ccaccaaggc 1140

cctagactca tcttggccct cctcagctcc ctgcctgttt cccgtaaggc tgtacatagt 1200

ccttttatct ccttgtggcc tatgaaactg 1230

<210> 17

<211> 1286

<212> DNA

<213> Homo sapiens

<400> 17

ttctgtgctc ccgagaactt cctatcaggt cctggctgca gggaacaag ctgggctttt 60

tataattaag gttggaagaa gtcaccacag gcagcagaac tccatcttga gatgaaataa 120

catctacctg gacctctggc agaatttcaa ggcacacact gggctgactc tggcgccatg 180

atgttgctt atccttcagc actgggagat caatactggg aagagatttt gcttccaaag 240

aatggggaaa atgtagagac tatgaagaaa ttgaccctaaa atcataaagc gaaaggcttg 300

ccttctaagt atactgactg cccccagaaa aaggagggaa aggcccaat agtggtagca 360

gttacattca gggatgtgac tgtgatcttc acagaagcag aatggaagag actgagtcca 420

gagcagagga atctatacaa agaagtgatg ctggagaatt acaggaatct tctctcattg 480

gacggtgggc aagagcagca agatgctgca gcatattgat tacaggatga ggtgcatcct 540

gcaggacggc cggatcttca ttggcacctt caaggctttt gacaagcaca tgaatttgat 600

cctctgtgac tigtatgagt tcagaaagat caagccaaag aactccaaac aagcagaaag 660

ggaagagaag cgagtctcgc gtctggtgct gctgcgaggg gagaatctgg tctcaatgac 720

agtagaggga cctcctccca aagatactgg tattgctcga gttccacttg ctggagctgc 780

cgggggcca gggatcgga gggctgctgg cagaggaatc ccagctgggg ttcccatgcc 840

ccaggctcct gcaggacttg ctgggccagt ccgtgggggt ggcgggcat cccaacaggt 900

gatgaccca caaggaagag gtactgttgc agccgctgca gctgctgcca cagccagtat 960

tgccggggct ccaaccagat accacactgg ccgtgggggt cctcccccac ctatgggccg 1020

aggagcacce cctccaggca tgatgggecc acctcctggg atgagacctc ctatgggtcc 1080

cccaatgggg atccccctg gaagagggac tccaatgggc atgccccctc cgggaatgcg 1140

gcctcctccc cctgggatgc gagggccccc tccccggga atgcgccac caaggccta 1200  
gactcatctt ggcctcctc agctcctgc ctgtttcccg taaggctgta catagtcctt 1260  
ttatctcctt gtggcctatg aaactg 1286

<210> 18  
<211> 1304  
<212> DNA

<213> Homo sapiens  
<400> 18

ttctgtgctc ccgagaactt cctatcaggt accagttaca ttcagggatg tgtctgtgat 60  
cttcacagaa gcagaatgga agagactgag tccagagcag aggaatctat acaaagaagt 120  
gatgctggag aattacagga atcttctctc attggacggt gggcaagagc agcaagatgc 180  
tgcagcatat tgattacagg atgaggtgca tcctgcagga cggccggatc ttcattggca 240  
ccttcaaggc tttgacaag cacatgaatt cgatcctctg tgaccgtgat gatttcagaa 300  
agatcaagtg tgtgaggggg gagagcttga gtttctggcc tcattgccta tttgtaaagc 360  
tgtagcctaa agggctctcc caatgcagcc tggctctgaaa ctgctttaag aagcctctga 420

ccctcttcag ggccaaagag ctccaacaa gcagaaaggg aagagaagcg agtcctcgg 480  
ctggtgctgc tgcgagggga gaatctggac tcaatgacag tagagggacc tcctcccaaa 540  
gatactggta ttgctcagat tccacttgct ggagctgccg ggggccagg gatcggcagg 600  
gctgctggca gaggaatccc agctgggggt cccatgcccc aggctcctgc aggacttct 660  
gggccagtcc gtggggttgg cgggccatcc caacaggtga tgactccaca aggaagaggt 720  
actgttcag ccgctgcagc tgctccaca gccagtattg ccggggctcc aaccagtac 780  
ccacctggcc gtgggggtcc tccccacct atgggccgag gagcaccccc tccagcatg 840

atgggcccac ctcttggtat gagacctct atgggtcccc caatgggat cccccctgga 900  
agagggactc caatgggcat gccccctccg ggaatgcggc ctctccccc tgggatgca 960  
ggccttcttt gacccttggc cacagagtat ggaagtagct ccgagaggc gtgggctcga 1020  
ttcctcaggg ccacgttacc acagacctgt ttgtttctta tgctgttgtt cgtggagtct 1080  
catgggattg tetggtttcc cttacaggc cccctcccc gggaatgcgc ccaccaaggc 1140  
cctagactca tcttggcct cctcagctcc ctgcctgttt cccgtaaggc tgtacatagt 1200  
ccttttatct ccttgtggcc tatgaaactg aagcttggaa ttcgcccggg taccgagctc 1260

gccttatagt gactgtata caatcactgc cctcgcacac cctc 1304

<210> 19

<211> 1432

<212> DNA

<213> Homo sapiens

<400> 19

|   |      |
|---|------|
| ttctgtgctc ccgagaactt cctatcaggt cctggctgca gggaacaag ctggcctttt  | 60   |
| tataattaag gtaggaagaa gtcaccacag gcagcagaac tccatcttga gatgaaataa | 120  |
| catctacctg gacctctggc agaatttcaa ggcacacact gggctgactc tggcgccatg | 180  |
| atgttgctt atccttcagc actgggagat caatactggg aagagatfff gcttccaaag  | 240  |
| aatggggaaa atgtagagac tatgaagaaa ttgacccaaa atcataaagc gaaaggcttg | 300  |
|   |      |
| ccttctaata atactgactg cccccagaaa aaggagggaa aggcccgaat agtggtagca | 360  |
| gttacattca gggatgtgac tgtgatcttc acagaagcag aatggaagag actgagtcca | 420  |
| gagcagagga atctatacaa agaagtgatg ctggagaatt acaggaatct tctctcattg | 480  |
| gacggtgggc aagagcagca agatgctgca gcatattgat tacaggatga ggtgcatcct | 540  |
| gcaggacggc cggatcttca ttggcacctt caaggctfff gacaagcaca tgaatttgat | 600  |
| cctctgtgac tigtatgagt tcagaaagat caagccaaag aactccaaac aagcagaaag | 660  |
| ggaagagaag cgagtcctcg gtctgggtct gctgcgaggg gagaatctgg tctcaatgac | 720  |
|   |      |
| agtagagggg cctcctccca aagatactgg tattgctcga gttccacttg ctggagctgc | 780  |
| agggggccca gggatcgcca gggctgctgg cagaggaatc ccagctgggg ttccatgcc  | 840  |
| ccagctcctc gcaggacttg ctgggccagt ccgtgggggt ggccggccat cccaacaggt | 900  |
| gatgacccca caaggaagag gtaactgttc agccgctgca gctgctgcca cagccagtat | 960  |
| tgccgggact ccaaccagat acccactgg ccgtgggggt cctccccac ctatgggccc   | 1020 |
| aggagcacc cctccaggca tgatgggccc acctcctggg atgagacctc ctatgggtcc  | 1080 |
| cccaatgggg atccccctg gaagagggac tccaatgggc atgccccctc cgggaatgcg  | 1140 |
|   |      |
| gcctcctccc cctgggatgc gaggccttct ttgaccttg gccacagagt atggaagtag  | 1200 |
| ctccgagag gcgtgggctc gattcctcag ggccacgtta ccacagacct gtttgtttct  | 1260 |
| tatgctgttg ttcgtggagt ctcatgggat tgtctggttt cccttacagg gccccctccc | 1320 |
| ccgggaatgc gcccaccaag gccctagact catcttggcc ctctcagct cctgctctgt  | 1380 |
| ttcccgtgag gctgtacata gtccttttat ctcttctggg cctatgaaac tg         | 1432 |

<210> 20

<211> 1604

<212> DNA

<213> Homo sapiens

<400> 20

ttctgtgctc ccgagaactt cctatcaggt cctggctgca gggaacaag ctgggctttt 60

tataattaag gttggaagaa gtcaccacag gcagcagaac tccatcttga gatgaaataa 120

catctacctg gacctctggc agaatttcaa ggcacacact gggctgactc tggcgccatg 180

atgttgctt atccttcagc actgggagat caatactggg aagagatfff gcttccaaag 240

aatggggaaa atgtagagac tatgaagaaa ttgacccaaa atcataaagc gaaaggcttg 300

ccttctaatt atactgactg cccccagaat aaggagggaa aggcccaagt agtggtagca 360

gttacattca gggatgtgac tgtgatcttc acagaagcag aatggaagag actgagtcca 420

gagcagagga atctatacaa agaagtgatg ctggagaatt acaggaatct tctctcattg 480

gaaccaaagc cagaaatcta cacttggtcc tectgccttc tggccttctc ctgtcagcag 540

ttctcagtc aacatgtact tcagatcttc ctgggcttat gtgcagaaaa tcaattccat 600

ccaggaatt ctagcccagg gcgttggaaa cagcaggggc agcagtattc ccatacggtg 660

ggcaagagca gcaagatgct gcagcatatt gattacagga tgagtgcac cctgcaggacg 720

gccgatctt cattggcacc ttcaaggctt ttgacaagca catgaatttg atcctccgtg 780

actgtgatga gttcagaaa atcaagccaa agaactcaa acaagcagaa agggaagaga 840

agcagtcct cggctctggtg ctgctgcgag gggagaatct ggtctcaatg acagtagagg 900

gacctctcc caagatact ggtattgctc gagttccact tgctggagct gccgggggcc 960

cagggatcgg cagggtctgt ggcagaggaa tcccagctgg ggttcccatg ccccaggctc 1020

ctgcaggact tgetgggcca gtccgtgggg ttggcgggcc atccaacag gtgatgacct 1080

cacaaggaag aggtactgtt gcagccgtg cagctgtctc cacagccagt attgccgggg 1140

ctccaacca gtaccacct ggccgtgggg gtctctcccc acctatgggc cgaggagcac 1200

ccccccagg catgatgggc ccacctctg gtatgagacc tctatgggt ccccaatgg 1260

ggatcccccc tggaaagagg actccaatgg gcatgcccc tccgggaatg cggcctctc 1320

ccctgggat gcgaggcctt cattgacct tggccacaga gtatggaagt agctccgag 1380

aggcgtgggc tcgattctc agggccacgt taccacagac ctgtttgttt cttatgtgt 1440

tgttcgtgga gtctcatggg attgtctggt ttcccttaca gggccccctc ccccgggaat 1500

gcgccacca aggcctaga ctcatcttgg cctctctcag ctcctgctt gtttcccgta 1560

aggctgtaca tagtctttt atctcttgt ggctatgaa actg 1604

<210> 21

<211> 506

<212> DNA

<213> Homo sapiens  
 <400> 21  
 tctaattgag gccacaagg aatgcatgag agccctagga tttaaaatat gaaatggtgc 60  
  
 ccctacagac atacaaggag attgttcaact gctgagagga gcaggtctta actctggcca 120  
 ctgaacagac ctatgctgtg gaggtgaga cacccatcaa ccgctgtcc ctgctgctct 180  
 ctggccgggt tegtgtgagc caggatgggc agtttctgca ctacatcttt ccataccagt 240  
 tcatggactc tctgagtgg gaatcactac agccttctga ggagggggtg ttccaggtca 300  
 ctctgactgc tgagacctca tgtagctaca tttctggcc ccgaaaaagt ctccatcttc 360  
 ttctgaccaa agagcgatac atctctgccc tcttctcggc tctgctggga tatgacatct 420  
 cggagaagct ctacactctc aatgacaagc tctttgctaa gtttgggctg cgctttgaca 480  
  
 tccgccttcc cagcctctac catgtc 506  
 <210> 22  
 <211> 637  
 <212> DNA  
 <213> Homo sapiens  
 <400> 22  
 tagatttggc tgctccgct catagctgct tttggcgga aagatgccgg gctctggtga 60  
 ctcaaacct gccccgctg agtctcagga gaagaagccg ctgaagccct gctgcgcttg 120  
 cccggagacc aagaaggcgc gcgatgcgtg tatcatcgag aaaggagaag aacctgtgg 180  
 acatctaatt gaggccaca aggaatgcat gagagcccta ggatttaaaa tatgaaatgg 240  
 tgagtatagg tgcttgaga ggataaggaa atcacctga tccaatagct ggtaacctgc 300  
  
 gaggtcactc tgactgetga gacctcatgt agetacattt cctggccccg gaaaagtctc 360  
 catcttcttc tgaccaaaga gcgatacatc tctgcctct tctcggctct gctgggatat 420  
 gacatctcgg agaagctcta cactctcaat gacaagctct ttgctaagtt tgggctgcgc 480  
 ttgacatcc gccttccag cctctacat gtctgggctc ccaactgctgc agatgetgga 540  
 ccagagtccg agaagggtga tgaggaagtc tgtgagccag ctgtgtcccc tctcaggcc 600  
 acaccacct ctctccagca aacacccct tgttcta 637  
 <210> 23  
 <211> 643  
 <212> DNA  
  
 <213> Homo sapiens

<400> 23  
 tctaattgag gccacaagg aatgcatgag agccctagga tttaaaatat gaaatggtga 60  
 gtataggtgc ttggagagga taaggaaatc accctgatcc aatagctggt aacctgccgag 120  
 gcatacata tcicatgact atggtaactt caagcattac atcgtcacgc aactattcaa 180  
 gcctggatga aagggtcccc tacagacata caaggagatt gttcactgct gcgaggagca 240  
 ggtcttaact ctggcactg aacagaccta tgctgtggag ggtgagacac ccatcaaccg 300  
 cctgtccctg ctgctctctg gccgggttcg tgtgagccag gatgggcagt ttctgcacta 360  
 catctttcca taccagtcca tggactctcc tgagtgggaa tcactacagc ctcttgagga 420

gggggtgttc caggteactc tgactgctga gacctcatgt agctacattt cctggccccg 480  
 gaaaagtctc catcttcttc tgaccaaaga gcgatacatc tcctgcctct tctcggctct 540  
 gctgggatat gacatctcgg agaagctcta cactctcaat gacaagctct ttgctaagtt 600  
 tgggctgcgc ttigacatcc gccttcccag cctctacat gtc 643

<210> 24  
 <211> 710  
 <212> DNA  
 <213> Homo sapiens  
 <400> 24

tagatttggc tgictccgct catagctgct tttggcgcga aagatgccgg gtctggttga 60  
 ctcaaaccct gccccgcctg agtctcagga gaagaagccg ctgaagccct gctgcgcttg 120  
  
 cccggagacc aagaaggcgc gcgatgcgtg tatcatcgag aaaggagaag aacactgtgg 180  
 acatctaatt gaggcccaca aggaatgcat gagagcccta ggatttaaaa tatgaaatgg 240  
 tgagtatagg tgcttggaga ggataaggaa atcacctga tccaatagct ggtaacctgc 300  
 gaggcatatc atatctcatg actatggtaa cttcaagcat tacatcgtca cgcaactatt 360  
 caagcctgga taaagggtca ctctgactgc tgagacctca thtagctaca tttcctggcc 420  
 ccggaagagt ctccatcttc ttctgaccaa agagcgatac atctcctgcc tcttctcggc 480  
 tctgctggga taigacatct cggagaagct ctacactctc aatgacaagc tctttgctaa 540  
  
 gtttgggctg cgctttgaca tccgccttcc cagcctctac catgtcctgg gtcccactgc 600  
 tgcagatgct ggaccagagt ccgagaaggg ttagtaggaa gtctgtgagc cagctgtgtc 660  
 ccctcctcag gccacacca cctctctcca gaaacaccc ccttgttcta 710

<210> 25  
 <211> 746  
 <212> DNA

<213> Homo sapiens  
 <400> 25  
 tagatttggc tgtctccgct catagctgct tttggcgcga aagatgccgg gctctggtga 60  
 ctcaaacct gccccgctg agtctcagga gaagaagccg ctgaagccct gctgcgcttg 120  
 cccggagacc aagaaggcgc gcgatgcgtg tatcatcgag aaaggagaag aacctgtgg 180  
  
 acatctaatt gaggcccaca aggaatgcat gagagcccta ggatttaaaa tatgaaatgg 240  
 tgagtatagg tgcttgaga ggataaggaa atcacctga tccaatagct ggtaacctgc 300  
 gaggttcgt gtgagccagg atgggcagt tctgcaactac atctttccat accagttcat 360  
 ggactctcct gagtgggaat cactacagcc tctgaggag ggggtgttcc aggtcactct 420  
 gactgctgag acctcatgta gctacattc ctggccccg aaaagtctcc atcttcttct 480  
 gaccaaaag cgatacatct cctgctctt ctggctctg ctgggatatg acatctcgga 540  
 gaagctctac acitcfaatg acaagctctt tgctaagttt gggctgcgct ttgacatccg 600  
  
 cttcccage ctctaccatg tcttgggtcc cactgctgca gatgctggac cagagtccga 660  
 gaagggtgat gaggaagtct gtgagccagc tgtgtccct cctcaggcca cacceacctc 720  
 tctccagcaa acaccccctt gttcta 746  
 <210> 26  
 <211> 768  
 <212> DNA  
 <213> Homo sapiens  
 <400> 26  
 tagatttggc tgtctccgct catagctgct tttggcgcga aagatgccgg gctctggtga 60  
 ctcaaacct gccccgctg agtctcagga gaagaagccg ctgaagccct gctgcgcttg 120  
 cccggagacc aagaaggcgc gcgatgcgtg tatcatcgag aaaggagaag aacctgtgg 180  
  
 acatctaatt gaggcccaca aggaatgcat gagagcccta ggatttaaaa tatgaaatgg 240  
 tgagtatagg tgcttgaga ggataaggaa atcacctga tccaatagct ggtaacctgc 300  
 gaggtgcccc tacagacata caaggagatt gttcactgct gtgaggagca ggtcttaact 360  
 ctggccactg aacagaccta tgctgtggag ggtgagacac ccatcaaccg cctgtccctg 420  
 ctgctctctg gccggtcact ctgactgctg agacctcatg tagctacatt tctggeccc 480  
 ggaaaagtct ccatcttctt ctgaccaaag agcgatacat ctctgcctc ttctcggctc 540  
 tgctgggata tgacatctcg gagaagctct acactctcaa tgacaagctc ttgtctaagt 600  
  
 ttgggctgcg ctttgacatc cgccttcca gcctctacca tgtctgggt cccactgctg 660



cagatgctgg accagagtcg gagaagggtg atgaggaagt ctgtgagcca gctgtgtccc 720  
ctcctcagcg cacaccacc tctctccagc aaacaccccc ttgttcta 768  
<210> 27  
<211> 777  
<212> DNA  
<213> Homo sapiens  
<400> 27  
tagatttggc tgtctccgct catagctgct ttggcgcgga aagatgccgg gctctggtga 60  
ctcaaaccct gccccgctg agtctcagga gaagaagccg ctgaagccct gctgcgcttg 120  
cccggagacc aagaaggcgc gcgatgcgtg tatcatcgag aaaggagaag aacactgtgg 180  
  
acatctaatt gaggcccaca aggaatgcat gagagcccta ggatttaaaa tatgaaatgg 240  
catatcatat ctcatgacta tggtaacttc aagcattaca tcgtcacgca actattcaag 300  
cctggatgaa aggtgccctt acagacatac aaggagattg ttcactgctg tgaggagcag 360  
gtcttaactc tggccactga acagacctat gctgtggagg gtgagacacc catcaaccgc 420  
ctgtccctgc tgctctctgg ccggtcactc tgactgctga gacctcatgt agctacattt 480  
cctggccccg gaaaagtctc catctctctc tgaccaaaga gcgatacatc tctgcctct 540  
tctcggctct gctgggatat gacatctcgg agaagctcta cactctcaat gacaagctct 600  
  
ttgctaagtt tgggctgcgc ttgacatcc gccttcccag cctctacat gtcttgggtc 660  
ccactgctgc agatgctgga ccagagtccg agaagggtga tgaggaagtc tgtgagccag 720  
ctgtgtcccc tctcaggcc acaccacct ctctccagca aacaccccct ttgttcta 777  
<210> 28  
<211> 841  
<212> DNA  
<213> Homo sapiens  
<400> 28  
tagatttggc tgtctccgct catagctgct ttggcgcgga aagatgccgg gctctggtga 60  
ctcaaaccct gccccgctg agtctcagga gaagaagccg ctgaagccct gctgcgcttg 120  
cccggagacc aagaaggcgc gcgatgcgtg tatcatcgag aaaggagaag aacactgtgg 180  
  
acatctaatt gaggcccaca aggaatgcat gagagcccta ggatttaaaa tatgaaatgg 240  
tgagtatagg tgcttggaga ggataaggaa atcacctga tccaatagct ggtaacctgc 300  
gaggcatatc ctatctcatg actatggtaa cttcaagcat tacatcgtca cgcaactatt 360  
caagcctgga tgaagggtgc ccctacagac atacaaggag attgttact gctgtgagga 420

gcaggcttta accttgccca ctgaacagac ctatgctgtg gagggtgaga cacccatcaa 480  
 ccgcctgtcc ctgctgctct ctggccggtc actctgactg ctgagacctc atgtagctac 540  
 atttcctggc cccggaaaag tctccatctt cttctgacca aagagcgata catctcctgc 600

ctcttctcgg ctctgctggg atatgacatc tccgagaagc tctacactct caatgacaag 660  
 ctctttgcta agtttggct gcgctttgac atccgccttc ccagcctcta ccatgtcctg 720  
 ggtcccactg ctgcagatgc tggaccagag tccgagaagg gtgatgagga agtctgtgag 780  
 ccagctgtgt cccctcctca ggccacacc acctctctcc agcaaacacc cccttgttct 840  
 a 841

<210> 29  
 <211> 873  
 <212> DNA  
 <213> Homo sapiens  
 <400> 29

tagatttggc tgtctccgct catagctgct tttggcgcga aagatgccgg gctctggtga 60  
  
 ctcaaaccct gccccgctg agtctcagga gaagaagccg ctgaagccct gctgcgcttg 120  
 cccggagacc aagaaggcgc gcgatgcgtg gtttcgtggt cggggctctg ttgacttacc 180  
 ccgtagctta gtaccgggaa acgatggaga gagaggagaa cggcaagctt caaccactca 240  
 ttcttctcta caaattatca tcgagaaagg agaagaacac tgtggacatc taattgaggc 300  
 ccacaaggaa tgcattgagag ccctaggatt taaaatatga aatggtgagt ataggtgctt 360  
 ggagaggata aggaaatcac cctgatccaa tagctggtaa cctgagaggt gccctacag 420  
 acatacaagg agattgttca ctgctgtgag gacgaggtct taactctggc cactgaacag 480

acctatgctg tggagggtga gacacccatc aaccgctgt cctgctgct ctctggccgg 540  
 tcaactctgac tgctgagacc tcatgtagct acatttctctg gccccgaaa agtctccatc 600  
 ttcttctgac caaagagcga tacatctctt gcctcttctc ggctctgctg ggatatgaca 660  
 tctcggagaa gcctacact ctcaatgaca agctctttgc taagtttggg ctgcgctttg 720  
 acatccgct tcccagctc taccatgtcc tgggtccac tgctgcagat gctggaccag 780  
 agtccgagaa gggatgatgag gaagtctgtg agccagctgt gtcctctct caggccacac 840  
 ccacctctct ccagcaaca ccccttgtt cta 873

<210> 30  
 <211> 877  
 <212> DNA

<213> Homo sapiens  
 <400> 30  
 tagatttggc tgtctccgct catagctgct tttggcgcga aagatgccgg gctctggttga 60  
 ctcaaaccct gccccgcctg agtctcagga gaagaagccg ctgaagccct gctgcgcttg 120  
 cccggagacc aagaaggcgc gcgatgcgtg tatcatcgag aaaggagaag aacctgtgg 180  
 acatctaatt gagggccaca aggaatgcat gagagcccta ggatttaaaa tatgaaatgg 240  
 tgagtatagg tgcttggaga ggataaggaa atcacctga tccaatagct ggtaacctgc 300  
 gaggtgcccc tacagacata caaggagatt gttcactgct gcgaggagca ggtcttaact 360  
  
 ctggccactg aacagaccta tgctgtggag ggtgagacac ccatcaaccg cctgtccctg 420  
 ctgctctctg gccgggttcg tgtgagccag gatgggcagt ttctgacta catctttcca 480  
 taccagtcca tggactctcc tgagtgggaa tctactacagc cttctgagga gggggtgttc 540  
 caggtcactc tgactgctga gacctcatgt agctacattt cctggccccg gaaaagtctc 600  
 catctttctc tgaccaaaaga gcgatacatc tctgcctct tctcggtct gctgggatat 660  
 gacatctcgg agaagctcta cactctcaat gacaagctct ttgctaagtt tgggctgcgc 720  
 tttgacatcc gccttcccag cctctacat gtctctgggtc ccactgctgc agatgctgga 780  
  
 ccagagtccg agaagggatga tgaggaagtc tgtgagccag ctgtgtcccc tctcaggcc 840  
 acaccacct ctctccagca aacacccct tgttcta 877  
  
 <210> 31  
 <211> 1223  
 <212> DNA  
 <213> Homo sapiens  
 <400> 31  
 tagatttggc tgtctccgct catagctgct tttggcgcga aagatgccgg gctctggttga 60  
 ctcaaaccct gccccgcctg agtctcagga gaagaagccg ctgaagccct gctgcgcttg 120  
 cccggagacc aagaaggcgc gcgatgcgtg tatcatcgag aaaggagaag aacctgtgg 180  
 acatctaatt gagggccaca aggaatgcat gagagcccta ggatttaaaa tatgaaatgg 240  
  
 tgagtatagg tgcttggaga ggataaggaa atcacctga tccaatagct ggtaacctgc 300  
 gaggcataatc atatctcatg actatggtaa cttcaagcat tacatcgtca cgcagctatt 360  
 caagcctgga tgaaggtgc ccctacagac atacaaggag attgttact gctgcgagga 420  
 gcaggcttta actctggcca ctgaacagac ctatgctgtg gagggtgaga cacccatcaa 480  
 ccgcctgtcc ctgctgctct ctggccggag aagatcatcc atcttccagt agatcaagaa 540  
 gggagaagaa acagatgata aaggccacag cattccgatt taaatggctt cagcgaggca 600

cgaagatcat tagacattat ctacactctc cggttgcac tctttggtgg cttgcctggg 660

tggccgaaat cctcgcctaa ggagctgaca tatgctgcag cagcgcatt ctcttctgct 720

gtcttccttt gctatggaag cttttcagtg aggttttgat cccctttat ggtgacacaa 780

ggttcgtgtg agccaggatg ggcagtttct gactacatc tttcatacc agttcatgga 840

ctctcctgag tgggaatcac tacagccttc tgaggagggg gtgttccagg tcaactctgac 900

tgctgagacc teatgtagct acatttcttg gccccgaaa agtctccatc ttcttctgac 960

caaagagcga tacatctctt gcctcttctc ggctctgctg ggatatggca tctcggagaa 1020

gctctacact ctcaatgaca agctctttgc taagtttggg ctgcgctttg acatccgct 1080

tcccagcctc taccatgtcc tgggtccac tgctgcagat gctggaccag agtccgagaa 1140

gggtgatgag gaagtctgtg agccagctgt gtcccctct cagccacac ccacctctct 1200

ccagcaaaca ccccctgtt cta 1223

<210> 32

<211> 880

<212> DNA

<213> Homo sapiens

<400> 32

gaagtgactg cggacgaatc ggcgtttgcc gagctggca tagatttggc tgtctccgct 60

catagctgct tttggcgcga aagatgccgg gtctggttga ctcaaacct gccccgctg 120

agtctcagga gaagaagccg ctgaagccct gctgcgcttg cccggagacc aagaaggcgc 180

gcatgctgtg tatcatcgag aaaggagaag aacctgtgg acatctaatt gaggcccaca 240

aggaatgcat gagagcccta ggatttaaaa tatgaaatgg tgagtatagg tgcttgaga 300

ggataaggaa atcacctga tccaatagct ggtaacctgc gaggcatac atatctcatg 360

actatggtaa cttaagcat tacatcgtca cgcaactatt caagcctgga tgaaagcttc 420

tagaattcct ctccagagct actctcaagt tatatccagg ggacaggccc ctttggtctc 480

aaccacacg cctgaacttt aaggatcatt ggactatctt ctctgtggcc agcgcagctc 540

tcttctgtgt tcacagaatg gccactgata ggcatgcctc tttcccacc cactggaagg 600

ctcacaggca aggtgagaga ggacacagaa ggtgccaaca ctgtcgctac agtaaggacc 660

tgaagtgact ttgagaaatt caccctcaca aaccttctt caggagcagg catttgtagt 720

gcagaggcac agattccgtc ctttaccagc tgcagaatct tgggcaagtt acatagcctc 780

tgtgagcctc atcggtaaac agtgggggtt atgaaacca cctcacaggg ttgttgtgag 840

gatccaatga gttgatttag gtaagcacct agcacatgcc 880

<210> 33  
 <211> 1120  
 <212> DNA  
 <213> Homo sapiens  
 <400> 33  
 gaagtgactg cggacgaatc ggcgtttgcc gaggctggca tagatttggc tgtctccgct 60  
 catagctgct ttiggcgcga aagatgccgg gtctggttga ctcaaaccct gccccgctg 120  
 agtctcagga gaagaagccg ctgaagccct gctgcgcttg cccggagacc aagaaggcgc 180  
 gcgatgcgtg taiccatcgag aaaggagaag aacctgtgg acatctaatt gaggccaca 240  
 aggaatgcac gagagcccta ggatttaaaa tatgaaatgg tgagtatagg tgcttgaga 300  
 ggataaggaa atcacctga tccaatagct ggtaacctgc gaggcataatc atatctcatg 360  
 actatggtaa ctcaagcat tacaccgtca cgcaactatt caagcctgga taaaagggtgc 420  
 cactacagac atacaaggag attgttcaact gctgcgagga gcaggtctta actctggcca 480  
 ctgaacagac ctatgctgtg gaggttgaga cacccatcaa cgcctgtcc ctgctgctct 540  
 ctggccgggt tcgtgtgagc caggatgggc agtttctgca ctacatcttt ccataaccagt 600  
 tcatggactc tctgagtgg gaatcactac agccttctga ggagggggtg ttccagcttc 660  
 tagaattcct ctccagagct actctcaagt tatatccagg ggacaggccc ctttggctcc 720  
 aaccacacg cctgaacttt aaggatcatt ggactatctt ctctgtggcc agcgcagcac 780  
 tcttctgtgt tcacagaatg gccactgata ggcatgcctc ttttcccacc cactggaagg 840  
 ctcacaggca aggtgagaga ggacacagaa ggtgccaaca ctgtcgtac agtaaggacc 900  
 tgaagtgact ttgagaaatt cacctcaca aaccttctt caggagcagg cattggtagt 960  
 gcagaggcac agattccgtc ctttaccagc tgcagaatct tgggcaagtt acatagcctc 1020  
 tgtgacctc atcggtaaac agtgggggtt atgaaacca cctcacaggg ttgttgtag 1080  
 gatccaatga gttgatttag gtaagcact agcacatgcc 1120  
 <210> 34  
 <211> 1357  
 <212> DNA  
 <213> Homo sapiens  
 <400> 34  
 gaagtgactg cggacgaatc ggcgtttgcc gaggctggca tagatttggc tgtctccgct 60  
 catagctgct ttiggcgcga aagatgccgg gtctggttga ctcaaaccct gccccgctg 120

agtctcagga gaagaagccg ctgaagccct gctgcgcttg cccggagacc aagaaggcgc 180  
 gcgatgcgtg tatcatcgag aaaggagaag aacctgtgg acatctaatt gaggccaca 240  
 aggaatgcat gagagcccta ggatttaaaa tatgaaatgg tgagtatagg tgcttggaga 300  
 ggataaggaa atcacctga tccaatagct ggtaacctgc gagtctactc tgactgctga 360  
 gacctcatgt agctacattt cctggccccg gaaaagtctc catcttcttc tgaccaaaga 420  
 gcgatacacc tectgectct tctcggtctc gctgggatat gacatctcgg agaagctcta 480  
 cactctcaat gacaagctct ttgctaagtt tgggctgcgc ttgacatcc gccttcccag 540

cctctacat gtctgggtc ccaactgctgc agatgctgga ccagagtccg agaagggtga 600  
 tgaggaagtc tggagccag ctgtgtcccc tctcaggcc acaccacct ctctccagca 660  
 aacacccct tgttctacc ctccagctac caccaacttt cctgcacctc ctaccgggc 720  
 caggttgtcc aggccagaca gtggcactact gggtagggac tccaccagt cgggtgctgga 780  
 ggattttgag gaggtgtcag gatcagaatc gtttatggat tataggagt atggggagta 840  
 catgaggtga agggagaact aacatgggca cagccaccgg ctcaggatcc tatcttctag 900  
 aattctctc cagagctact ctcaagttat atccagggga caggccccct tggctccaac 960

ccacacgct gaactttaag gatcattgga ctatcttctc tgtggccagc gcagctctct 1020  
 tctgtgttca cagaatggcc actgatagc atgctcttt tcccaccac tggaggctc 1080  
 acaggcaagg tgagagagga cacagaaggt gccaacactg tgcctacagt aaggacctga 1140  
 agtgactttg agaattcac cctcacaac ctctctcag gagcaggcat tggtagtgca 1200  
 gaggcacaga ttccgtcctt taccagctgc agaactcttg gcaagttaca tagcctctgt 1260  
 gagcctcacc ggtaaacagt gggggttatg aaaccacct cacagggttg ttgtgaggat 1320  
 ccaatgagtt gatttagta agcacctagc acatgcc 1357

- <210> 35
- <211> 1398
- <212> DNA
- <213> Homo sapiens
- <400> 35

gaagtgactg cggacgaatc ggcgtttgcc gaggttgca tagatttggc tgtctccgct 60  
 catagctgct ttggcgcgga aagatgccgg gtctggttga ctcaaacct gccccgctg 120  
 agtctcagga gaagaagccg ctgaagccct gctgcgcttg cccggagacc aagaaggcgc 180  
 gcgatgcgtg tatcatcgag aaaggagaag aacctgtgg acatctaatt gaggccaca 240  
 aggaatgcat gagagcccta ggatttaaaa tatgaaatgg tgagtatagg tgcttggaga 300

ggataaggaa atcacctga tccaatagct ggtaacctgc gaggcatatc atatctcatg 360  
  
 actatggtaa ctcaagcat tacatcgtca cgcaactatt caagcctgga tgaaagggtt 420  
 cgtgtgagcc aggatgggca gtttctgcac tacatcttcc cataccagtt catggactct 480  
 cctgagtggg aatcactaca gccttctgag gagggggtgt tccaggtcac tctgactgct 540  
 gagacctcat gtagctacat ttctggccc cggaaaagtc tccatcttct tctgacaaa 600  
 gagcgataca tctctgctt cttctcggt ctgctgggat atgacatctc ggagaagctc 660  
 tacactctca atgacaagct ctttgctaag ttgggctgc gcttgacat cgccttccc 720  
 agcctctacc atgtctggg tcccactgct gcagatctg gaccagagtc cgagaagggt 780  
  
 gatgaggaag tctgtgagcc agctgtgtcc cctctcagg ccacaccac ctctctccag 840  
 caaacacccc ctgtttctac cctccagct accaccaact ttctgcacc tctaccgg 900  
 gccagttgt ccagccaga cagtggcata ctgcttcta gaattctct ccagagctac 960  
 tctcaagtta tatccagggg acaggcccct ttggctcaa cccacacgcc tgaacttta 1020  
 ggatcattgg actatctct ctgtggccag cgcagctctc ttctgtgtc acagaatggc 1080  
 cactgatagg catgctctt tcccacca ctggaaggct cacaggcaag gtgagagagg 1140  
 acacagaagg tgccaacact gtcgtacag taaggacctg aagtgacttt gagaaattca 1200  
  
 cctcacaaa ccttcttca ggagcaggca ttggtagtgc agaggcacag attcctct 1260  
 ttaccagctg cagaatcttg ggcaagttac atagcctctg tgagctcat cggtaaacag 1320  
 tgggggttat gaaaccacc tcacagggtt gttgtgagga tccaatgagt tgatttaggt 1380  
 aagcacctag cacatgcc 1398  
  
 <210> 36  
 <211> 1402  
 <212> DNA  
 <213> Homo sapiens  
 <400> 36  
  
 gaagtgactg cggacgaatc ggcgtttgcc gaggtggca tagatttggc tgtctcct 60  
 catagctgct ttggcgca aagatgccgg gtctggtga ctcaaacct gccccctg 120  
  
 agtctcagga gaagaagccg ctgaagcct gctgccttg cccggagacc aagaaggcgc 180  
 gcgatcgtg tatcatcag aaaggagaag aacctgtgg acatctaatt gaggccaca 240  
 agaatgcat gagagccta ggatttaaaa tatgaaatgg gttcgtgtga gccagatgg 300  
 gcagtttctg cactacatct ttccatacca gttcatggac tctctgagt gggaactact 360  
 acagccttct gaggagggg tgttcagggt cactctgact gctgagacct catgtagcta 420

catttctg ccccgaaaa gtctccatct tcttctgacc aaagagcgat acatctcctg 480  
 cctcttctcg gctctgctgg gatatgacat ctccgagaag ctctacactc tcaatgacaa 540

gctcttttgc aagtttgggc tgcgctttga catccgectt cccagcctct accatgtcct 600  
 ggggccact gctgcagatg ctggaccaga gtccgagaag ggtgatgagg aagtctgtga 660  
 gccagctgtg tcccctctc aggccacacc cacctctctc cagcaaacac ccccttgttc 720  
 taccctcca gctaccacca actttcctgc acctctacc cgggccaggt tgtccaggcc 780  
 agacagtggc atactgggtg aggactccac cagtctgggt ctggaggatt ttgaggaggt 840  
 gtcaggatca gaatcgttta tggattatag gagtatggg gagtacatga ggtgaaggga 900  
 gaactaacat gggcacagcc accggctcag gatcctatct tctagaattc ctctccagag 960

ctactctcaa gttatatcca ggggacaggc ccttttggct ccaaccaca cgctgaact 1020  
 ttaaggatca ttggactatc ttctctgtgg ccagcgagc tctcttctgt gttcacagaa 1080  
 tggccactga taggcatgcc tcttttccca cccactggaa ggctcacagg caaggtgaga 1140  
 gaggacacag aaggtgcca cactgtcgt acagtaagga cctgaagtga ctttgagaaa 1200  
 ttcacctca caaaccttc ttcaggagca ggcatgtgta gtgcagagcc acagattccg 1260  
 tcctttacca gctgcagaat cttgggcaag ttacatagcc tctgtgagcc tcatcgtaa 1320  
 acagtggggg ttatgaaacc cacctcacag ggttgtgtg aggatccaat gaggttgatt 1380

aggtaagcac ctagcacatg cc 1402

<210> 37  
 <211> 1456  
 <212> DNA  
 <213> Homo sapiens  
 <400> 37

gaagtgactg cggacgaatc ggcgtttgcc gaggctggca tagatttggc tgtctccgct 60  
 catagctgct tttggcgcca aagatgccgg gctctggtga ctcaaacct gccccgctg 120  
 agtctcagga gaagaagccg ctgaagcct gctgcgcttg cccggagacc aagaaggcgc 180  
 gcgatgcgtg tatcatcgag aaaggagaag aacctgtgg acatctaatt gaggcccaca 240  
 aggaatgcat gagagcccta ggatttaaaa tatgaaatgg tgagtatagg tgcttgagaa 300

ggataaggaa atcacctga tccaatagct ggtaacctgc gagggtcccc tacagacata 360  
 caaggagatt gttcactgct gcgaggagca ggtcttaact ctggccactg aacagaccta 420  
 tgctgtggag ggtgagacac ccatcaaccg cctgtccctg ctgctctctg gccgggttcg 480  
 tgtgagccag gatgggcagt ttctgcaacta catctttcca taccagttca tggactctcc 540



tgagtgggaa tcactacagc cttctgagga gggggtgttc caggtcactc tgactgctga 600  
 gacctcatgt agctacattt cctggccccg gaaaagtctc catcttcttc tgaccaaaga 660  
 gcgatacatc tcttgcctct tctcggctct gctgggatat gacatctcgg agaagctcta 720  
  
 cactctcaat gacaagctct ttgctaagtt tgggctgcgc tttgacatcc gccttcccag 780  
 cctctacat gtcttgggtc ccaactgctgc agatgctgga ccagagtccg agaagggtga 840  
 tgaggaagtc tggagccag ctgtgtcccc tctcaggcc acaccacct ctctecagca 900  
 aacacccct tgttctaccc ctccagctac caccaacttt cctgcacctc ctaccgggc 960  
 caggttgtcc aggccagaca gtggcatact ggcttctaga attcctctcc agagctactc 1020  
 tcaagttata tccaggggac aggccctttt ggctccaacc cacacgcctg aactttaagg 1080  
 atcattggac tatcttctct gtggccagcg cagctctctt ctgtgttcac agaatggcca 1140  
  
 ctgataggca tgctctttt cccacccact ggaaggetca caggcaaggt gagagaggac 1200  
 acagaaggtg ccaacactgt cgctacagta aggacctgaa gtgactttga gaaattcacc 1260  
 ctcaaaaacc ttcttcagg agcaggcatt ggtagtgcag aggcacagat tccgtccttt 1320  
 accagctgca gaactttggg caagttacat agcctctgtg agcctcatcg gtaaacagtg 1380  
 ggggttatga aaccacctc acagggttgt tgtgaggatc caatgagttg atttagttaa 1440  
 gcacctagca catgcc 1456  
  
 <210> 38  
 <211> 1488  
 <212> DNA  
  
 <213> Homo sapiens  
 <400> 38  
  
 gaagtgactg cggacgaatc ggcgtttgcc gaggtggca tagatttggc tgtctccgct 60  
 catagctgct ttiggcgca aagatgccgg gtctggttga ctcaaaccct gccccgctg 120  
 agtctcagga gaagaagccg ctgaagccct gctgcgcttg cccggagacc aagaaggcgc 180  
 gcgatgcgtg taicactcag aaaggagaag aacctgtgg acatctaatt gaggeccaca 240  
 aggaatgcat gagagcccta ggatttaaaa tatgaaatgg tgagtatagg tgcttggaga 300  
 ggataaggaa atcacctga tccaatagct ggtaacctgc gaggtgcccc tacagacata 360  
 caaggagatt gttcactgct gcgaggagca ggtcttaact ctggccactg aacagaccta 420  
  
 tgctgtggag ggtgagacac ccatcaaccg cctgtccctg ctgctctctg gccggtcact 480  
 ctgactgctg agacctcatg tagctacatt tcttggcccc ggaaaagtct ccatcttctt 540  
 ctgaccaaag agcgatacat ctctgcctc ttctcggctc tgctgggata tgacatctcg 600

gagaagctct acactctcaa tgacaagctc ttgctaagt ttgggctgcg ctttgacatc 660  
 cgccctccca gcctctacca tgtcctgggt cccactgctg cagatgctgg accagagtcc 720  
 gagaagggtg atgaggaagt ctgtgagcca gctgtgtccc ctctcaggc cacaccacc 780  
 tctctccagc aaacaccccc ttgttctacc cctccagcta ccaccaactt tctgcacct 840  
  
 cctaccggg ccaggttgtc caggccagac agtggcatac tgggtgagga ctccaccagt 900  
 ctggtgctgg aggatttga ggaggtgtca ggatcagaat cgtttatgga ttataggagt 960  
 gatggggagt acatgaggtg aagggagaac taacatgggc acagccaccg gctcaggatc 1020  
 ctatcttcta gaattcctct ccagagctac tctcaagtta tatccagggg acaggeccct 1080  
 ttggctccaa cccacacgcc tgaactttaa ggatcattgg actatcttct ctgtggccag 1140  
 cgcagctctc ttctgtgttc acagaatggc cactgatagg catgcctctt tccccacca 1200  
 ctggaaggct cacaggcaag gtgagagagg acacagaagg tgccaacact gtcgctacag 1260  
  
 taaggacctg aagtgacttt gagaaattca ccctcacaaa ccttcttca ggagcaggca 1320  
 ttggtagtgc agaggcacag attccgtcct ttaccagctg cagaatcttg ggcaagttac 1380  
 atagcctctg tgagcctcat cggtaaacag tgggggttat gaaaccacc tcacagggtt 1440  
 gttgtgagga tccaatgagt tgatttaggt aagcacctag cacatgcc 1488  
  
 <210> 39  
 <211> 1529  
 <212> DNA  
 <213> Homo sapiens  
 <400> 39  
  
 gaagtgactg cggacgaatc ggcgtttgcc gaggctggca tagatttggc tgtctccgt 60  
 catagctgct tttggcgcga aagatgccgg gtctggttga ctcaaacct gcccccctg 120  
  
 agtctcagga gaagaagccg ctgaagccct gctgcgcttg cccggagacc aagaaggcgc 180  
 gcgatgcgtg tatcatcgag aaaggagaag aacctgtgg acatctaatt gaggcccaca 240  
 aggaatgcat gagagcccta ggatttaaaa tatgaaatgg tgagtatagg tgcttgagga 300  
 ggataaggaa atcacctga tccaatagct ggtaacctgc gaggcatatc atatctcatg 360  
 actatggtaa cttaagcat tacatgtca cgcaactatt caagcctgga tgaaaggtgc 420  
 ccctacagac atacaaggag attgttcact gctgcgagga gcaggtctta actctggcca 480  
 ctgaacagac ctatgctgtg gagggtgaga cacccatcaa ccgctgtcc ctgctgctct 540  
  
 ctggccgggt tcgtgtgagc caggatgggc agtttctgca ctacatcttt ccataccagt 600  
 tcatggactc tctgagtgg gaatcactac agccttctga ggaggggtg tccaggtca 660

ctctgactgc tgagacctca tgtagctaca tttcctggcc ccggaagagt ctccatcttc 720  
 ttctgaccaa agagcgatac atctctgcc tcttctcggc tctgctggga tatgacatct 780  
 cggagaagct ctacactctc aatgacaagc tctttgctaa gtttgggctg cgctttgaca 840  
 tccgccttcc cagcctctac catgtcctgg gtcccactgc tgcagatgct ggaccagagt 900  
 ccgagaaggg tgatgaggaa gtctgtgagc cagctgtgtc cctcctcag gccacacca 960

ccctctccca gcaaacacce ccttgttcta ccctccage taccaccaac tttctgcac 1020  
 ctctaccgg gccaggttg tccaggccag acagtggcat actggcttct agaattcctc 1080  
 tccagagcta ctctcaagtt atatccaggg gacaggcccc tttggctcca acccacagc 1140  
 ctgaacttta aggatcattg gactatcttc tctgtggcca gcgagctct ctctgtgtt 1200  
 cacagaatgg cactgatag gcatgcctct tttcccacc actggaagge tcacaggcaa 1260  
 ggtgagagag gacacagaag gtgccaacac tgtcgctaca gtaaggacct gaagtgactt 1320  
 tgagaaattc accctcaaa accttccttc aggagcaggc attggtagtg cagaggcaca 1380

gattccgtcc tttaccagct gcagaatctt gggcaagtta catagcctct gtgagcctca 1440  
 tcggtaaaca gtggggggtta tgaaacceac ctcacagggt tgttgtgagg atccaatgag 1500  
 ttgatttagg taagcaccta gcacatgcc 1529

<210> 40  
 <211> 1537  
 <212> DNA  
 <213> Homo sapiens  
 <400> 40

gaagtgactg cggacgaatc ggcgtttgcc gaggctggca tagatttggc tgtctccgt 60  
 catagctgct tttggcgcga aagatgccgg gtctgggtga ctcaaacct gccccgctg 120  
 agtctcagga gaagaagccg ctgaagccct gctgcgcttg cccggagacc aagaaggcgc 180

gcatgctgtg tatcatcgag aaaggagaag aacctgtgg acatctaatt gaggcccaca 240  
 aggaatgcat gagagcccta ggatttaaaa tatgaaatgg tgaglatagg tgcttgagaga 300  
 ggataaggaa atcacctga tccaatagct ggtaacctgc gaggcatatc atatctcatg 360  
 actatggtaa cttaagcat tacatgtca cgcaactatt caagcctgga tgaaagggtt 420  
 cgtgtgagcc aggatgggca gtttctgcac tacatcttc cataccagtt catggactct 480  
 cctgagtggg aatcactaca gccttctgag gagggggtgt tccaggtcac tctgactgct 540  
 gagacctcat gtagctacat ttcttgcccc cggaaaagtc tccatcttct tctgaccaa 600

gagcgataca tctctgctt cttctcgct ctgctgggat atgacatctc ggagaagctc 660

tacactctca atgacaagct ctttgctaag tttgggctgc gctttgacat cgccttccc 720  
 agcctctacc atgtcctggg tcccactgct gcagatctg gaccagagtc cgagaagggt 780  
 gatgaggaag tctgtgagcc agctgtgtcc cctcctcagg ccaccacact ctctccagca 840  
 aacaccccct tgttctacce ctccagctac caccaacttt cctgcacctc ctaccgggc 900  
 caggttgtcc aggccagaca gtggcatact gggtgaggac tccaccagtc tggctctgga 960  
 ggattttgag gaggtgtcag gatcagaatc gtttatggat tataggagtg atggggagta 1020

catgaggtga agggagaact aacatgggca cagccaccgg ctcaggatcc tatcttctag 1080  
 aattcctctc cagagctact ctcaagttat atccagggga caggcccctt tggtccaac 1140  
 ccacacgcct gaactttaag gatcattgga ctatcttctc tgtggccagc gcagctctct 1200  
 tctgtgttca cagaatggcc actgataggc atgcctcttt tcccaccac tggaggctc 1260  
 acaggcaagg tgagagagga cacagaaggt gccaacactg tcgctacagt aaggacctga 1320  
 agtgactttg agaaattcac cctcacaac cttccttcag gagcaggcat tggtagtgca 1380  
 gaggcacaga ttccgtcctt taccagctgc agaacttgg gcaagttaca tagcctctgt 1440

gagcctcacc ggtaaacagt gggggttatg aaaccacact cacagggttg ttgtgaggat 1500  
 ccaatgagtt gatttaggta agcacctagc acatgcc 1537

<210> 41  
 <211> 1597  
 <212> DNA  
 <213> Homo sapiens  
 <400> 41

gaagtgactg cggacgaatc ggcgtttgcc gaggtggca tagatttggc tgtctccgct 60  
 catagctgct tttggcgcga aagatgccgg gtctggttga ctcaaaccct gccccgctg 120  
 agtctcagga gaagaagccg ctgaagccct gctgcgcttg cccggagacc aagaaggcgc 180  
 gcgatgcgtg tatcatcgag aaaggagaag aacctgtgg acatctaatt gagggccaca 240

aggaatgcat gagagcccta ggatttaaaa tatgaaatgg tgagtatagg tgcttgagaga 300  
 ggataaggaa atcacctga tccaatagct ggtaacctgc gaggtgcccc tacagacata 360  
 caaggagatt gttcactgct gcgaggagca ggtcttaact ctggccactg aacagaccta 420  
 tgcttgggag ggtgagacac ccatcaaccg cctgtccctg ctgctctctg gccgggttcg 480  
 tgtgagccag gatgggcagt ttctgacta catctttcca taccagttca tggactctcc 540  
 tgagtgggaa tcaactacagc cttctgagga gggggtgttc caggtcactc tgactgctga 600  
 gacctcatgt agctacattt cctggccccg gaaaagtctc catcttcttc tgaccaaaaga 660

gcgatacatc tctgcctct tctcggctct gctgggatat gacatctcgg agaagctcta 720  
 cactctcaat gacaagctct ttgctaagtt tgggctgcgc ttgacatcc gccttcccag 780  
 cctctacat gtctcgggtc ccaactgtgc agatgctgga ccagagtccg agaagggtga 840  
 tgaggaagtc tggagccag ctgtgtcccc tctcaggcc acacccacct ctctccagca 900  
 aacacccct tgttctacce ctccagctac caccaacttt cctgcacctc ctaccgggc 960  
 caggttgtec agccagaca gtggcatact gggtagggac tccaccagtc tggtagctgga 1020  
 ggattttgag gaggtgtcag gatcagaatc gtttatggat tataggagtg atggggagta 1080

catgagtgga agggagaact aacatgggca cagccaccgg ctcaggatcc tatcttctag 1140  
 aattctctc cagagctact ctcaagttat atccagggga caggccctt tggctccaac 1200  
 ccacacgct gaactttaag gatcattgga ctatctctc tgtggccagc gcagctctct 1260  
 tctgtgttca cagaatggcc actgatagc atgctcttt tcccaccac tggaggctc 1320  
 acaggcaagg tgagagagga cacagaaggt gccaacactg tgcctacagt aaggacctga 1380  
 agtgactttg agaaattcac ctcacaaac ctctctcag gagcaggcat tggtagtgca 1440  
 gaggcacaga ttccgtcctt taccagctgc agaacttgg gcaagttaca tagcctctgt 1500

gagcctcacc ggtaaaccgt gggggttatg aaaccacct cacagggtg ttgtaggat 1560  
 ccaatgagtt gatttaggta agcacctagc acatgcc 1597

- <210> 42
- <211> 1658
- <212> DNA
- <213> Homo sapiens
- <400> 42

gaagtgactg cggacgaatc ggcgtttgcc gaggctggca tagatttggc tgtctccgct 60  
 catagctgct ttggcgcgca aagatgccgg gtctggttga ctcaaaccct gccccgctg 120  
 agtctcagga gaagaagccg ctgaagccct gctgcgcttg cccggagacc aagaaggcgc 180  
 gcgatgcgtg tatcatcgag aaaggagaag aacctgtgg acatctaatt gaggccaca 240

aggaatgcat gagagcccta ggatttaaaa tatgaaatgg tgagtatagg tgcttgaga 300  
 ggataaggaa atcacctga tccaatagct ggtaacctgc gaggcatatc atatctcatg 360  
 actatggtaa ctcaagcat tacatcgtca cgcaactatt caagcctgga taaaagacat 420  
 acaaggagat tgttactgc tgtgaggagc aggtcttaac tctggccact gaacagacct 480  
 atgctgtgga gggtaggaca cccatcaacc gctgtccct gctgctctct ggccgggttc 540  
 gtgtgagcca ggatgggcag tttctgcaact acatctttcc ataccagttc atggactctc 600

ctgagtggga atcactacag ccttctgagg agggggtgtt ccaggtcact ctgactgctg 660

agacctcatg tagctacatt tcttggcccc ggaaaagtct ccatcttctt ctgaccaaag 720

agcgatacat ctctgcctc ttctcggtc tcttgggata tgacatctcg gagaagctct 780

acactctcaa tgacaagctc ttgtctaagt ttgggtctcg ctttgacatc cgccttccca 840

gcctctacca tgtcctgggt cccactgctg cagatgtctg accagagtcc gagaagggtg 900

atgaggaagt ctgtgagcca gctgtgtccc ctctcagge cacaccacc tetctccagc 960

aaacaccccc ttgttctacc cctccagcta ccaccaactt tctgcacct cctaccgggg 1020

ccaggttgtc caggccagac agtgggtatac tgggtgagga ctccaccagt ctggtgctgg 1080

aggattttga ggaggtgtca ggatcagaat cgtttatgga ttataggagt gatggggagt 1140

acatgaggtg aagggagaac taacatgggc acagccaccg gctcaggatc ctatcttcta 1200

gaattctctt ccagagetac tctcaagtta tatccagggg acaggccctt ttggtccaa 1260

cccacacgcc tgaactttaa ggatcattgg actatcttct ctgtggccag cgcagctctc 1320

ttctgtgttc acagaatggc cactgatagg cacgcctctt ttcccacca ctggaaggct 1380

cacaggcaag gtgagagagg acacagaagg tgccaacact gtcgctacag taaggacctg 1440

aagtgacttt gagaaattca ccctcaciaa ccttccttca ggagcaggca ttggtagtgc 1500

agaggcacag attccgtcct ttaccagctg cagaatcttg ggcaagtac atagcctctg 1560

tgagcctcat cggtaaacag tgggggttat gaaaccacc tcacagggtt gttgtgagga 1620

tccaatgagt tgatttaggt aagcacctag cacatgcc 1658

<210> 43

<211> 1670

<212> DNA

<213> Homo sapiens

<400> 43

gaagtgactg cggacgaatc ggcgtttgcc gaggctggca tagatttggc tgtctccgt 60

catagctgct ttggcgcgga aagatgccgg gtctggttga ctcaaacct gccccgctg 120

agtctcagga gaagaagccg ctgaagccct gctgctcttg cccggagacc aagaaggcgc 180

gcgatgcgtg tatcatcgag aaaggagaag aacctgtgg acatctaatt gaggcccaca 240

aggaatgcat gagagcccta ggatttaaaa tatgaaatgg tgaglatagg tgcttgaga 300

ggataaggaa atcacctga tccaatagct ggtaacctgc gaggcatatc atatctcatg 360

actatggtaa cttcaagcat tacatctca cgcaactatt caagcctgga tgaaaggtgc 420

ccctacagac atacaaggag attgttact gctgcgagga gcaggtctta actctggcca 480

ctgaacagac ctatgctgtg gagggtgaga cacccatcaa cgcctgtcc ctgctgctct 540  
ctggccgggt tcgtgtgagc caggatgggc agtttctgca ctacatcttt ccataaccagt 600

tcattggactc tectgagtgg gaatcactac agccttctga ggagggggtg ttccaggtea 660  
ctctgactgc tgagacctca tgtagctaca tttctggcc cggaaaagt ctccatcttc 720  
ttctgaccaa agagcgatac atctctgcc tcttctggc tctgctggga tatgacatct 780  
cggagaagct ctacactctc aatgacaagc tctttgctaa gtttgggctg cgctttgaca 840  
tccgccttc cagcctctac catgtcctgg gtcccactgc tgcagatgct ggaccagagt 900  
ccgagaaggg tgatgaggaa gtctgtgagc cagctgtgtc cctcctcag gccacacca 960  
cctctctcca gcaaacacce ccttgttctc ccctccagc taccaccaac tttctgac 1020

ctcctaccg gccaggttg tccaggccag acagtggcat actgggtgag gactccacca 1080  
gtctggtgct ggaggatttt gaggaggtgt caggatcaga atcgtttat gattatagga 1140  
gtgatggga gtacatgagg tgaagggaga actaacatgg gcacagccac cggctcagga 1200  
tcctatcttc tagaattct ctccagagct actctcaagt tatatccagg ggacaggccc 1260  
ctttggctcc aaccacacag cctgaacttt aaggatcatt ggactatctt ctctgtggcc 1320  
agcgcagctc tcttctgtgt tcacagaatg gccactgata ggcatgcctc tttcccacc 1380  
cactggaagg ctacagcca aggtgagaga ggacacagaa ggtgccaaca ctgtcgtac 1440

agtaaggacc tgaagtgact ttgagaaatt cacctcaca aaccttctt caggagcagg 1500  
cattggtagt gcagaggcac agattccgtc cttaccagc tgcagaatct tgggcaagtt 1560  
acatagcctc tgtgagcctc atcggtaaac agtgggggtt atgaaacca cctcacaggg 1620  
ttgttgtgag gatccaatga gttgatttag gtaagcacct agcacatgcc 1670

<210> 44  
<211> 442  
<212> DNA  
<213> Homo sapiens  
<400> 44

atcaggaagt agtcatgccc cacaagatgt cagtttgtca tatcccaac atcatgttgg 60  
aaattcaage cctacctca catctcctag cagcaggtac gaaaggacta tggataacat 120

gacttcatca gcacagataa ttcagagatc tttctcttct ccagaaaatt ttcagagaca 180  
agcagtgatt actctagttt gatgcatata tttctcatgg agttaaaact ggactcttca 240  
aatgataga gttgtctgaa tttggctgca atgcgaacta ctaatcgatt caagcttgtt 300  
atcaagacac ctagatatca ccaggaaatt cagttcatga agataagaag tttcttgaca 360

aatacatgcc ccagttcatg aaacatcttc attatagaat aattgatgtg agcactgtta 420  
aagaactgtg cagacgctgg ta 442  
<210> 45  
<211> 474  
<212> DNA  
<213> Homo sapiens  
<400> 45  
atcaggaagt agtcatgccc cacaagatgt cagtttgtca tatccccaac atcatgttgg 60  
aaattcaagc cctacctcca catctcctag cagcaggtac gaaaggacta tggataacat 120  
gacttcatca gcacagataa ttcagagatc tttctcttct ccagaaaatt ttcagagaca 180  
agcagtggta gtttcttcaa acctgtatac tgatcagctg tctcatgaat acgggagaaa 240  
acctgcaggt ctccgagagt tctctccctg tgcaacttat cgcttttcag tgctgtgtac 300  
cggccattcc agctgccttt gtttcgctgg actctcagct gtgtcttctg aactcagaaa 360  
ttcagttcat gaagataaga agtttcttga caatacatg cccagttca tgaacatct 420  
tcattataga ataattgatg tgagcactgt taaagaactg tgcagacgct ggta 474  
<210> 46  
<211> 713  
<212> DNA  
<213> Homo sapiens  
<400> 46  
atcaggaagt agtcatgccc cacaagatgt cagtttgtca tatccccaac atcatgttgg 60  
aaattcaagc cctactgcca catctcctag caggtacgaa aggactatgg ataacatgac 120  
ttcatcagca cagataattc agagatcttt ctcttctcca gaaaatttc agagacaagc 180  
agtgattact ctagtttgat gtcataattt ctcatggagt taaaactgga ctcttcaaaa 240  
tgatagagtt gtctgaattt ggctgcaatg cgaactacta atcgattcaa gcttgttate 300  
aagacaccta gagatcacca gatgacagga ttggacattg agaaggacca gattattgag 360  
atggcctgtc tgataactga ctctgatctc aacattttgg ctgaaggfcc taacctgatt 420  
ataaaacaac cagatgagtt gctggacagc atgtcagatt ggtgtaagga gcatcacggg 480  
aagtctggcc ttaccaaggc agtgaaggag agtacaatta cattgcagca ggcagagtat 540  
gaatttctgt cctttgtacg acagcagact cctccagggc tctgtccact tgcaggaat 600  
tcagttcatg aagataagaa gtttcttgac aaatacatgc cccagttcat gaaacatctt 660



cattatagaa taattgatgt gggcactggt aaagaactgt gcagacgctg gta 713

<210> 47

<211> 1036

<212> DNA

<213> Homo sapiens

<400> 47

atcaggaagt agtcatgccc cacaagatgt cagtttgtca tatccccaac atcatgttgg 60

aaattcaage cctacctcca catctcctag caggtacgaa aggactatgg ataacatgac 120

ttcatcagca cagataattc agagatcttt ctcttctcca gaaaattttc agagacaagc 180

agtgatgaac agtgctttga gacaaatgtc atatggtgga aaatttggtt cttcacctct 240

ggatcaatca ggattttcac catcagttca atcacacagt catagtttca atcagtcctc 300

aagctcccag tggcgccaag attactctag ttgatgtca taatttctca tggagttaa 360

actggactct tcaaatgat agagtgtct gaatttggct gcaatgcaa ctactaatcg 420

attcaagctt gttatcaaga cacctagaga tcaccagctt gaaaaaggaa gaacattctg 480

acacatgcta caacatggat gaactttgag gagatgacag ttaagagaa ataagccaga 540

tatgaaaaga caaatactgc ataagtctac ttgtaagagg aggtccaaga acccactctt 600

tgggtctgga tctggaccgc ttteccataa gaaaacgaag gaagatgaca ggattggaca 660

ttgagaagga ccagattatt gagatggcct gtctgataac tgactctgat ctcaacattt 720

tggctgaagg tccctaacctg attataaac aaccagatga gttgctggac agcatgtcag 780

attggtgtaa ggagcatcac gggaagtctg gccttaccaa ggcagtgaag gagagtaca 840

ttacattgca gcaggcagag tatgaatttc tgcctttgt acgacagcag actcctccag 900

ggctctgtcc acttgcagga aattcagttc atgaagataa gaagtttctt gacaaataca 960

tgccccagtt catgaaacat cttcattata gaataattga tgtgagcact gttaaagaac 1020

tgtgcagacg ctggtgta 1036

<210> 48

<211> 421

<212> DNA

<213> Homo sapiens

<400> 48

aatgagtga tgggcacctc taaccaccac tcggtacagg cctgtcgtg gcggaagctc 60

tacctgagca gggccaagct gaaggcctcc agcaggacct ccgcctcct ctccggcttt 120

gccatggtgg ccatggtgga ggtgcagctg gagacgcagt accagtaccc gcggccgctg 180  
 ctgattgcct tcagcgctg caccacggtg ctggtggccg tgcacctgtt gcacctcctc 240  
 atcagcacct gcacctgcc caatgtggag gccgtgagca acatccaca cctgaactcc 300  
 atcagcgagt ccccgcatga gcgcatgcac cctacatcg agctggcctg gggttctcc 360  
 accgtgcttg gcacctact cttctggcc gaggtggtgc tgctctgctg gatcaagttc 420  
 c 421

<210> 49

<211> 600

<212> DNA

<213> Homo sapiens

<400> 49

aatgagttaa tgggcacctc gacctcccc agaatttgc cgagatgtca tgggttcaag 60  
 tgctggggcc ggcagtggag agttccacgt gtacagacat ctgcgccga gagaatatca 120  
 gcgacaggac tacatggatg ccatggctga gaagcaaaaa ttggatgcag agtttcagaa 180  
 aagactggaa aagaataaaa ttgctgcaga ggagcagacc gcaaagcgc ggaagaagcg 240  
 ccagaagtta aaagagaaga aattactggc aaagaagatg aaacttgaac agaagaaca 300  
 agaaggtggc catggtggag gtgcagctgg agacgcagta ccagtaccg cggccgctgc 360

tgattgcctt cagcgctgc accacggtgc tggggccgt gcacctgttc gcacctcctca 420  
 tcagcacctg catcctgcc aatgtggagg ccgtgagcaa catccacaac ctgaactcca 480  
 tcagcgagtc cccgcatgag cgcatgcacc cctacatcga gctggcctgg ggcttctcca 540  
 ccgtgcttgg catcctactc ttctggccg aggtggtgct gctctgctgg atcaagttcc 600  
 600

<210> 50

<211> 838

<212> DNA

<213> Homo sapiens

<400> 50

aatgagttaa tgggcacctc gacctcccc agaatttgc cgagatgtca tgggttcaag 60  
 tgctggggcc ggcagtggag agttccacgt gtacagacat ctgcgccga gagaatatca 120  
 gcgacaggac tacatggatg ccatggctga gaagcaaaaa ttggatgcag agtttcagaa 180  
 aagactggaa aagaataaaa ttgctgcaga ggagcagacc gcaaagcgc ggaagaagcg 240  
 ccagaagtta aaagagaaga aattactggc aaagaagatg aaacttgaac agaagaaca 300

agaagcctgg ctccccatc gagtctgag cttaacgtgc ctatcgacc ctctgctcct 360  
 gcctgcctg agccccgcca taagggcatg gattaccggg actgggtccg ccgcagctac 420  
 ctggaactgg tcacctctaa ccaccactcg gtacaggccc tgtcgtggcg gaagctctac 480

ctgagcaggg ccaagctgaa ggcctccagc aggacctccg cctcctctc cggtttgcc 540  
 atggtggcca tgggtggagt gcagctggag acgcagtacc agtaccgcg gccgctgctg 600  
 attgccttca ggcctgcac cacggtgctg gtggccgtgc acctgttcgc cctcctcacc 660  
 agcacctgca tctgcccac tgtggaggcc gtgagcaaca tccacaacct gaactccatc 720  
 agcagatccc cgcctgagc catgcacccc tacatcgagc tggcctgggg ctctccacc 780  
 gtgcttgcca tctactctt cctggccgag gtggtgctgc tctgctggat caagtcc 838

<210> 51

<211> 857

<212> DNA

<213> Homo sapiens

<400> 51

aatgagtgaa tgggcacctc gacctcccc agaatttgc cgagatgtca tgggttcaag 60  
 tctggtggcc ggcagtggag agttccactg gtacagacat ctgcccggga gagaatatca 120  
 gcgacaggac tacatggatg ccatggttga gaagcaaaaa ttggatgcag agtttcagaa 180  
 aagactggaa aagaataaaa ttgctgcaga ggagcagacc gcaaagcgc ggaagaagcg 240  
 ccagaagtta aaagagaaga aattactggc aaagaagatg aaacttgaac agaagaaca 300  
 agaaggacc ggtcagccca aggagcaggg gtccagcagc tctgcccagg catctggaac 360  
 agaggaggag gaggaagtgc ccagttcac catggggcga tgacaatgtt tgccacagcc 420

tctgcttggc acctggctcg tctgtgacc agaagggaaa ggcggctgtt tggctctttc 480  
 tccccgcaa ggaccgctg acccgctgga tggagagcaa aggagacccc tcccagccc 540  
 ctcacagtcc tgtatttggc aggtggccat ggtggagggt cagctggaga cgcagtacca 600  
 gtaccgcccg ccgtgctga ttgcttccag cgctgcacc accgtgctgg tggcctgca 660  
 cctgttcgcc ctctcatca gcacctgcat cctgcccac gtggaggccg tgagcaacat 720  
 ccacaacctg aactccatca gcgagtcccc gcatgagcgc atgcaccct acatcgagct 780  
 ggcttggggc ttctccaccg tgcttggcat cctactcttc ctggccgagg tgggtgctgt 840

ctgctggatc aagtcc 857

<210> 52

<211> 947

<212> DNA  
 <213> Homo sapiens  
 <400> 52

aatgagtgaatgggcacctc gacctcccc agaatttgtc cgagatgtca tgggttcaag 60  
 tgctggggcc ggcagtggag agttccacgt gtacagacat ctgcgccga gagaatatca 120  
 gcgacaggac tacatggatg ccatggctga gaagcaaaaa ttggatgcag agtttcagaa 180  
 aagactggaa aagaataaaa ttgctgcaga ggagcagacc gcaaagcgcc ggaagaagcg 240  
 ccagaagtta aaagagaaga aattactggc aaagaagatg aaactgaac agaagaaca 300

agaaggaag taactccca gggatggaag tgcttggatg cggctgctgt ggctgcggat 360  
 gtgcgaagg agatgggatg gagagcctga gttggcattc gtataaatga cctgcctggc 420  
 tcccaccatg agtgtgagc ttaactgcc tatcgacccc tctgctctg cctgcctga 480  
 gcccgccat aaggcatgg attaccggga ctgggtccgc cgcagctacc tggactggt 540  
 cacctctaac caccactcg tacaggcct gtcgtggcgg aagctctacc tgagcagggc 600  
 caagctgaag gctccagca ggacctcgc cctctctcc ggctttgcca tgggtggccat 660  
 ggtggaggtg cagctggaga cgcagtacca gtaccccgcg ccgctgctga ttgccttcag 720

cgctgcacc acggtgctgg tggcctgca cctgttcgcc ctctcatca gcacctgcat 780  
 cctgccaat gtggaggccg tgagcaacat ccacaactg aactccatca gcgagtcacc 840  
 gcatgagcgc atgcaccct acatcgagct ggctggggc ttctccaccg tgcttggcat 900  
 cctactcttc ctggccgagg tggctgctgt ctgctggatc aagtcc 947

<210> 53  
 <211> 1275  
 <212> DNA  
 <213> Homo sapiens  
 <400> 53

aatgagtgaatgggcacctc gacctcccc agaatttgtc cgagatgtca tgggttcaag 60  
 tgctggggcc ggcagtggag agttccacgt gtacagacat ctgcgccga gagaatatca 120

gcgacaggac tacatggatg ccatggctga gaagcaaaaa ttggatgcag agtttcagaa 180  
 aagactggaa aagaataaaa ttgctgcaga ggagcagacc gcaaagcgcc ggaagaagcg 240  
 ccagaagtta aaagagaaga aattactggc aaagaagatg aaactgaac agaagaaca 300  
 agaaggacc ggtcagccca aggagcagg gtcagcagc tctgcggagg catctggaac 360  
 agaggaggag gaggaagtgc ccagttcac catggggcga tgacaatgtt tgccacagcc 420  
 tctgcctgga acctggctcg tgctgtgacc agaagggaaa ggcggctgtt tgctctttc 480

tccccgcaa ggaccgctg acccgctgga tggagagcaa aggagacccc tcccagccg 540

ctcacagtcc tgtatttggc aggtttggga gcctgagggg ccatctccct gacactcaga 600

ggcactgcct tgcagacacc atccgtgctc ctgggaagga acgtcccagg gatggaagtg 660

cttgatgcg gtgctgctgg ctgcgatgt gcgcaaggag atgggatgga gacctgagt 720

tggcattcgt ataatgacc tgcctggctc ccaccatgag tgctgagctt aacgtgccta 780

tcgaccctc tgetcctgcc tgccctgagc cggccataa gggcatggat taccgggact 840

gggtccgccg cagctacctg gaactggctc cctctaacca cactcggtc caggccctgt 900

cgtggcgcaa gctctacctg agcaggcca agctgaaggc ctccagcagg acctccgcc 960

tctctccgg ctttgccatg gtggccatgg tggagtgca gctggagacg cagtaccagt 1020

accgcggcc gctgctgatt gccttcagcg cctgcaccac ggtgctggtg gccgtgcacc 1080

tgttgccct cctcatcagc acctgcatcc tgcccaatgt ggaggcctg agcaacatcc 1140

acaacctgaa ctccatcagc gagtccccgc atgagcgcac gcaccctac atcgagctgg 1200

cctggggctt ctccaccgtg cttggcatcc tactcttct ggccgaggtg gtgctgctct 1260

gctggatcaa gttcc 1275

<210> 54

<211> 533

<212> DNA

<213> Homo sapiens

<400> 54

aggcttccat ccaacacatc cagaatgcaa ttgatacgga gaagtcacaa caggcactgg 60

ttcagaagcg ccattacctt tttgatgtgc aaaggaataa cattgctatg gctttggaag 120

ttacttaccg ggaacgactg tatagagtat ataaggaagt aaagaatcgc ctggactatc 180

atatactgt gcagaacatg atgcgtcga aggaacaaga acacatgata aattgggtgg 240

agaagcactg ggtgcaaagc atctccacac agcaggctgg agagactctt ctgtggcatt 300

caaaactgcca tgttgtgaac tgctgtgga gagaaccaca aagcagggaa ctgtagacag 360

cctctaggac ctgcaaacag gctccagctg acaaccagta agaagccagg gcctcagtc 420

ataagccgta acggaagcaa gtgaaatgct cagtcttaga cgaactgcgtc gtgctatgac 480

cggacttttt ctigaaaggg gatgacagca tgggagggaaa tggctccaca tgt 533

<210> 55

<211> 588

<212> DNA

<213> Homo sapiens  
 <400> 55  
 aggcttccat ccaacacatc cagaatgcaa ttgatacggga gaagtcacaa caggcactgg 60  
 ttcagaagcg ccattacctt tttgatgtgc aaaggaataa cattgctatg gctttggaag 120  
 ttacttaccg ggaacgactg tatagagtat ataaggaagt aaagaatcgc ctggactatc 180  
 atatatctgt gcagaacatg atgcgtcgaa aggaacaaga acacatgata aattgggtgg 240  
  
 agaagcacgt ggtgcaaagc atctccacac agcaggaaaa ggagacaatt gccaagtgca 300  
 ttgcggacct aaagctgctg gcaaagaagg ctcaagcaca gccagttatg taaatgctgg 360  
 agagactctt ctgtggcatt caaactgcca tgttgtgaac tgcctgtgga gagaaccaca 420  
 aagcagggaa ctgtagacag cctctaggac ctgcaaacag gctccagctg acaaccaaag 480  
 ccgtaacgga agcaagtgaa atgctcagtc ttagacgact gcgtcgtgct atgaccggac 540  
 tttttcttga aaggggatga cagcatggga ggaatggct ccacatgt 588  
 <210> 56  
 <211> 603  
 <212> DNA  
  
 <213> Homo sapiens  
 <400> 56  
 aggcttccat ccaacacatc cagaatgcaa ttgatacggga gaagtcacaa caggcactgg 60  
 ttcagaagcg ccattacctt tttgatgtgc aaaggaataa cattgctatg gctttggaag 120  
 ttacttaccg ggaacgactg tatagagtat ataaggaagt aaagaatcgc ctggactatc 180  
 atatatctgt gcagaacatg atgcgtcgaa aggaacaaga acacatgata aattgggtgg 240  
 agaagcacgt ggtgcaaagc atctccacac agcaggaaaa ggagacaatt gccaagtgca 300  
 ttgcggacct aaagctgctg gcaaagaagg ctcaagcaca gccaggctgg agagactctt 360  
 ctgtggcatt caaactgcca tgttgtgaac tgcctgtgga gagaaccaca aagcagggaa 420  
  
 ctgtagacag cctctaggac ctgcaaacag gctccagctg acaaccagta agaagccagg 480  
 gccctcagtc ataagccgta acggaagcaa gtgaaatgct cagtcttaga cgactgcgtc 540  
 gtgctatgac cggacttttt cttgaaaggg gatgacagca tgggaggaaa tggtccaca 600  
 tgt 603  
 <210> 57  
 <211> 613  
 <212> DNA  
 <213> Homo sapiens

<400> 57

|   |     |
|---|-----|
| aggcttccat ccaacacatc cagaatgcaa ttgatacggg gaagtcacaa caggcactgg | 60  |
| ttcagaagcg ccattacctt tttgatgtgc aaaggaataa cattgctatg gctttggaag | 120 |
| ttacttaccg ggaacgactg tatagagtat ataaggaagt aaagaatcgc ctggactatc | 180 |
| atatactgtg gcagaacatg atgcgtcgaa aggaacaaga acacatgata aattgggtgg | 240 |
| agaagcacgt ggtgcaaagc atctccacac agcaggaaaa ggagacaatt gccaagtgca | 300 |
| ttgcggacct aaagctgctg gcaaagaagg ctcaagcaca gccagttagt taaatgctgg | 360 |
| agagactctt ctgtggcatt caaactgcca tgttgtgaac tgcctgtgga gagaaccaca | 420 |
| aagcagggaa ctgtagacag cctctaggac ctgcaaacag gctccagctg acaaccagta | 480 |
| agaagccagg gccctcagtc ataagccgta acggaagcaa gtgaaatgct cagtcttaga | 540 |
| cgactgcgtc gtgctatgac cggacttttt cttgaaaggg gatgacagca tgggaggaaa | 600 |
| tggtccaca tgt   | 613 |