(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2011-0128039
(43) 공개일자 2011년11월28일

(51) Int. Cl.
C12N 15/12 (2006.01) *C12Q 1/68* (2006.01)
(21) 출원번호 10-2010-0047631
(22) 출원일자 2010년05월20일
심사청구일자 없음

(71) 출원인
고려대학교 산학협력단
서울 성북구 안암동5가 1
(72) 발명자
송해룡
서울특별시 구로구 신도림동 동아1차아파트 105동 1102호
(74) 대리인
한양특허법인

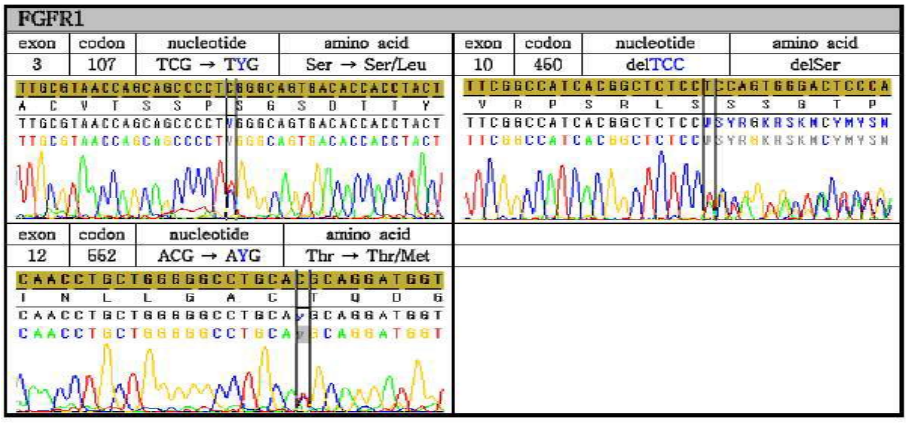
전체 청구항 수 : 총 8 항

(54) 연골저형성증 또는 연골무형성증의 예측 또는 진단에 유용한 FGFR 군의 유전자 돌연변이

(57) 요 약

본 발명은 연골저형성증 및 연골무형성증의 소인 및 감수성을 판단하기 위해 이용될 수 있는 한국인 특이적인 FGFR1, FGFR2, FGFR3 유전자 돌연변이 및 이의 이용에 관한 것이다.

대 표 도 - 도1



이 발명을 지원한 국가연구개발사업
과제고유번호 A080588091210501 00
부처명 한국보건산업진흥원
연구관리전문기관
연구사업명 보건의료기술개발사업
연구과제명 다발성 골단이형성증와 가연골 무형성증에서의 유전성 인자 연구
기여율
주관기관 한국보건산업진흥원
연구기간 2009-04-01 ~ 2010-03-31

특허청구의 범위

청구항 1

하기로 구성된 군으로부터 선택되는 연골저형성증 또는 연골무형성증의 진단을 위한 FGFR1 유전자 돌연변이 폴리뉴클레오티드:

(1) 서열번호 1의 서열로 구성된 폴리뉴클레오티드에서 2293번째 염기에서 2295번째 염기까지 TCC의 결실을 포함하는 20~100개의 연속 뉴클레오티드 서열로 구성되는 폴리뉴클레오티드;

(2) 서열번호 1의 서열로 구성된 폴리뉴클레오티드에서 2597번째 염기의 C에서 Y로의 치환을 포함하는 20~100개의 연속 뉴클레오티드 서열로 구성되는 폴리뉴클레오티드.

청구항 2

하기로 구성된 군으로부터 선택되는 연골저형성증 또는 연골무형성증의 진단을 위한 FGFR2 유전자 돌연변이 폴리뉴클레오티드:

(1) 서열번호 2의 서열로 구성된 폴리뉴클레오티드에서 1204번째 염기의 T에서 C로의 치환을 포함하는 20~100개의 연속 뉴클레오티드 서열로 구성되는 폴리뉴클레오티드;

(2) 서열번호 2의 서열로 구성된 폴리뉴클레오티드에서 1204번째 염기의 T에서 Y로의 치환을 포함하는 20~100개의 연속 뉴클레오티드 서열로 구성되는 폴리뉴클레오티드.

청구항 3

하기로 구성된 군으로부터 선택되는 연골저형성증 또는 연골무형성증의 진단을 위한 FGFR3 유전자 돌연변이 폴리뉴클레오티드:

(1) 서열번호 3의 서열로 구성된 폴리뉴클레오티드에서 449번째 염기의 G에서 R로의 치환을 포함하는 20~100개의 연속 뉴클레오티드 서열로 구성되는 폴리뉴클레오티드;

(2) 서열번호 3의 서열로 구성된 폴리뉴클레오티드에서 600번째 염기의 A에서 W로의 치환을 포함하는 20~100개의 연속 뉴클레오티드 서열로 구성되는 폴리뉴클레오티드.

청구항 4

제1항 내지 제3항 중 어느 한 항의 폴리뉴클레오티드 또는 이에 상보적인 폴리뉴클레오티드를 프로브로 포함하는 유전자 돌연변이 검출용 마이크로어레이.

청구항 5

제4항에 있어서, 상기 폴리뉴클레오티드는 방사선 표지, 형광 표지, 효소 표지, 서열 택, 또는 바이오틴에 의해 표지되는 것을 특징으로 하는 마이크로어레이.

청구항 6

제1항 내지 제3항 중 어느 한 항의 폴리뉴클레오티드 또는 이에 상보적인 폴리뉴클레오티드를 포함하는 연골저형성증 또는 연골무형성증 진단용 조성물.

청구항 7

제1항 내지 제3항 중 어느 한 항의 폴리뉴클레오티드 또는 이에 상보적인 폴리뉴클레오티드를 포함하는 연골저형성증 또는 연골무형성증 진단용 키트.

청구항 8

제7항에 있어서, 상기 진단용 키트는 마이크로어레이, 또는 ELISA 분석법을 이용하는 것을 특징으로 하는 진단용 키트.

명 세 서

기 술 분 야

[0001] 본 발명은 연골저형성증 및 연골무형성증의 소인 및 감수성을 판단하기 위해 이용될 수 있는 한국인 특이적인 FGFR1, FGFR2, FGFR3 유전자 돌연변이 및 이의 이용에 관한 것이다.

배 경 기 술

[0002] 인간 유전체 염기서열 중에 99.9%가 동일하지만, 일부분 0.1%의 차이에 의해 개인간의 모습이나, 행동, 그리고 질환감수성에 차이가 생긴다. 30억개 인간 유전체 염기서열 중에서 대략 1,000bp 마다 1회의 빈도로 서로 다른 염기가 올 수 있고, 이는 총 3백만 개가 되며, 이를 단일 염기 다형성(Single Nucleotide Polymorphism, SNP)이라고 부른다. 즉, 인간 유전체의 전체 염기서열 중에서 다형성을 보이는 3백만 개 정도의 염기를 분석한다면, 개인간이나 집단간의 차이, 또는 질환군과 정상인의 차이를 알 수 있을 것이다.

[0003] 한편, 선천성 골격계 이상질환(skeletal dysplasia)은 200여종 이상으로 알려져 있으며 연골 골성의 이상으로 왜소증, 사지 장관골 및 척추 기형, 조기 관절염 등을 동반하게 된다. 대부분의 골격계 질환은 유전자의 돌연변이, 상염색체 혹은 성염색체의 열성 및 우성 등 여러 가지 유전방식으로 발생하는데 조기에 진단과 치료가 늦어지면 골절 및 불유합 등의 합병증으로 골기형 및 퇴행성 관절염으로 보행이 불가능하거나 휠체어를 사용하는 등 영구적인 장애를 갖게 된다.

[0004] 상기 골격계 질환 중 하나인 연골저형성증은 상염색체 우성으로 유전되며, FGFR3 유전자의 이형접합체 돌연변이(heterozygous mutation)와 관련되는 비교적 흔한 연골형성 장애이다. 연골무형성증은 유전자의 돌연변이로 인하여 성장판의 연골세포가 성장호르몬에 정상적으로 반응하지 않아 뼈가 위로 자라지 못하게 되어 키가 120-135cm되는 왜소증을 유발하는 질환으로 알려져 있다.

[0005] 이러한 연골저형성증과 연골무형성증은 선천적 왜소증의 대표적 질환으로서, 외형상 매우 유사하며, X-ray를 통한 방사선학적 진단을 통한 진단이 어렵다는 문제가 있다. 상기 질환을 진단하기 위한 유전자 검사와 관련하여, 연골무형성증의 경우 FGFR-3 유전자의 코돈 380번의 돌연변이, 연골저형성증의 경우 FGFR-3 유전자의 코돈 540번의 돌연변이가 일반적으로 알려져 있으나, 연골저형성증 또는 연골무형성증에 관한 한국인 특유의 SNP 데이터베이스가 구축되어 있지 않고, 이를 진단할 수 있는 유전학적 특징을 찾는 연구는 미미한 상태이다.

[0006] 따라서, 현재 당업계에서는 연골저형성증 및 연골무형성증과 관련한 한국인 특이적인 유전자 돌연변이를 연구하고 파악하는 것이 요구된다. 이에 본 발명자는 다수의 한국인 연골저형성증/연골무형성증 환자들 및 정상인들로부터 수득한 시료를 분석하여, 특히 FGFR1, FGFR2, 또는 FGFR3 유전자 상에 존재하는 돌연변이 유형 및 빈도를 조사하여 연골저형성 및 연골무형성증의 예측 및 진단에 활용할 수 있는 유전자 돌연변이를 확인하고, 본 발명을 완성하였다.

발명의 내용

해결하려는 과제

[0007] 본 발명은 상기한 기술적 배경하에서 고안된 것으로,

[0008] 본 발명의 목적은 연골저형성증 또는 연골 무형성증의 진단에 활용될 수 있는 한국인 특이적인 FGFR1, FGFR2, 또는 FGFR3 유전자 돌연변이를 포함하는 폴리뉴클레오티드를 제공하는 것이다.

[0009] 본 발명의 다른 목적은 FGFR1, FGFR2, 또는 FGFR3 유전자 돌연변이를 포함하는 폴리뉴클레오티드 또는 이에 상보적인 서열을 프로브로 포함하는 연골저형성증 또는 연골무형성증 진단용 마이크로어레이를 제공하는 것이다.

[0010] 본 발명의 또 다른 목적은 연골저형성증 또는 연골무형성증 진단용 조성물 및 진단용 키트를 제공하는 것이다.

과제의 해결 수단

[0011] 상기 목적을 위하여, 본 발명자는 연골저형성증 또는 연골무형성증 환자들과 정상인들로부터 수득한 시료에서, 연골저형성증 또는 연골무형성증과 관련된 것으로 알려진 FGFR1, FGFR2, FGFR3 유전자의 엑손 전역과 엑손-인트론 접경부위(exon-intron junction)에 대한 염기서열을 분석하여 상기 유전자에 대한 돌연변이 여부를 분석하였

다.

[0012] 구체적으로, 상기 유전자의 분석을 위하여, NCBI (National Center for Biotechnology Information) 데이터베이스에서 유전자 서열(genomic sequence)을 얻어, 단일염기 다형성(SNP)을 찾기 위한 프라이머(primer)를 디자인하고 상기 프라이머들을 이용하여, 유전자의 엑손(exon) 부위를 PCR로 증폭하여 염기서열을 분석하였다.

[0013] 그 결과, 연골저형성증 또는 연골무형성증을 가진 한국인 환자에서는 FGFR1, FGFR2, 또는 FGFR3에서 특이적으로 유전자 돌연변이가 검출된 반면(표4 참조), 정상군에서는 이러한 돌연변이가 검출되지 않았다(표7 참조).

[0014] 이하, 상기 유전자 분석 결과 찾아낸 돌연변이 유전자 및 이를 이용한 연골저형성증 및 연골무형성증의 진단법 등에 대하여 구체적으로 설명하기로 한다.

[0015] 서열번호 1, 서열번호 2, 서열번호 3은 각각 FGFR1(Gen Bank Accession No. NG_007729), FGFR2(Gen Bank Accession No. NG_012449), FGFR3(Gen Bank Accession No. NG_012632)의 코딩 서열(CDS)을 포함하는 mRNA의 염기서열을 나타내며, FGFR1, FGFR2, FGFR3 유전자의 엑손 내의 돌연변이는 각각 서열번호 1, 서열번호 2, 서열번호 3을 기준으로 표시된다.

[0016] FGFR(Fibroblast Growth Factor Receptor) 1 유전자는 염색체 8번에 위치하고 18개의 엑손으로 구성되며, FGFR2 유전자는 염색체 10번에 위치하고 18개의 엑손으로 구성된다. FGFR3 유전자는 염색체 4번에 위치하고 17개의 엑손으로 구성된다.

[0017] 본 발명의 일 양태는 하기로 구성된 군으로부터 선택되는 연골저형성증 또는 연골무형성증의 진단을 위한 FGFR1 유전자 돌연변이 폴리뉴클레오티드를 제공한다:

[0018] (1) 서열번호 1의 서열로 구성된 폴리뉴클레오티드에서 2293번째 염기에서 2295번째 염기까지 TCC의 결실을 포함하는 20~100개의 연속 뉴클레오티드 서열로 구성되는 폴리뉴클레오티드;

[0019] (2) 서열번호 1의 서열로 구성된 폴리뉴클레오티드에서 2597번째 염기의 C에서 Y로의 치환을 포함하는 20~100개의 연속 뉴클레오티드 서열로 구성되는 폴리뉴클레오티드.

[0020] 본 발명은 또한 하기로 구성된 군으로부터 선택되는 FGFR2 유전자 돌연변이 폴리뉴클레오티드를 제공한다:

[0021] (1) 서열번호 2의 서열로 구성된 폴리뉴클레오티드에서 1204번째 염기의 T에서 C로의 치환을 포함하는 20~100개의 연속 뉴클레오티드 서열로 구성되는 폴리뉴클레오티드;

[0022] (2) 서열번호 2의 서열로 구성된 폴리뉴클레오티드에서 1204번째 염기의 T에서 Y로의 치환을 포함하는 20~100개의 연속 뉴클레오티드 서열로 구성되는 폴리뉴클레오티드.

[0023] 본 발명은 또한 하기로 구성된 군으로부터 선택되는 FGFR3 유전자 돌연변이 폴리뉴클레오티드를 제공한다:

[0024] (1) 서열번호 3의 서열로 구성된 폴리뉴클레오티드에서 449번째 염기의 G에서 R로의 치환을 포함하는 20~100개의 연속 뉴클레오티드 서열로 구성되는 폴리뉴클레오티드;

[0025] (2) 서열번호 3의 서열로 구성된 폴리뉴클레오티드에서 600번째 염기의 A에서 W로의 치환을 포함하는 20~100개의 연속 뉴클레오티드 서열로 구성되는 폴리뉴클레오티드.

[0026] 본 발명에서, 상기 20~100개의 연속 뉴클레오티드 서열은 예를 들면, 20~80개, 20~60개, 20~50개 또는 20~40개의 연속 뉴클레오티드 서열일 수 있다.

[0027] 본 발명에 따르면, FGFR1, FGFR2, 또는 FGFR3 유전자에 상기와 같은 돌연변이를 갖는 사람은 정상인에 비해 연골저형성증 및/또는 연골무형성증의 발생 가능성이 높다.

[0028] 본 발명의 폴리뉴클레오티드 또는 이에 상보적인 폴리뉴클레오티드는 프로브 또는 프라이머로 사용될 수 있다.

[0029] 상기 프라이머 또는 프로브는 전술된 FGFR1, FGFR2, 또는 FGFR3 유전자 돌연변이의 존재 유무를 확인하여 연골저형성증 또는 연골무형성증의 소인 또는 감수성을 확인하기 위해 이용될 수 있다. 이를 위해 이하의 실시예에 상세히 기재된 DNA 서열분석(sequencing)을 통해 확인하는 기술 외에 하기와 같이 여러 방법이 이용될 수 있다:

[0030] (가) 엑손 및 인트론 경계 부분을 포함하는 표적 절편들을 각각 돌연변이 유무를 확인할 수 있는 한 쌍의 프라이머로 증폭하여 F-CSGE를 수행하지 않고, 바로 디데옥시(dideoxy) 반응을 통한 서열분석으로 돌연변이의 존재 유무를 확인할 수 있다. 이 경우, 본 발명의 상기 돌연변이가 일어난 부위 및 이를 포함하는 염기서열로 구성된 올리고뉴클레오티드 프로브를 쓰지 않고, 바로 돌연변이의 존재 유무를 확인할 수 있다. 또는, FGFR1, FGFR2,

또는 FGFR3를 클로닝하여 벡터에 삽입한 후 서열분석을 수행할 수도 있다. 이 경우, 정상 FGFR1, FGFR2, FGFR3 유전자의 염기서열과 서열분석 결과를 서로 비교하여 돌연변이의 존재 유무를 확인할 수 있다.

[0031] (나) 유전자 돌연변이의 존재 유무는 또한 제한효소절단법을 이용하여 확인할 수 있다. 이는 FGFR1, FGFR2, FGFR3 유전자의 정상 유전자와 돌연변이 유전자의 염기서열이 서로 다르다는 점에 착안하여, 적당한 제한효소 또는 제한효소와 동일한 결과를 도출하는 화합물을 이용하여 정상 유전자는 절단되고 돌연변이 유전자는 절단되지 않거나 또는 역으로 정상 유전자는 절단되지 않고, 돌연변이 유전자는 절단되는 원리를 이용하는 방법이다. 이와 유사한 방법으로는 돌연변이 위치 부근에 결합한 프라이머를 절단하는 엔도뉴클레아제(endonuclease) 효소가 돌연변이에 의한 부정합 부위는 절단할 수 없는 원리를 이용한 인베이더 등이 있으며, 이들은 모두 본 발명의 돌연변이의 존재 유무를 확인하기 위해 이용될 수 있다.

[0032] (다) 돌연변이가 일어날 수 있는 유전자 상의 염기 위치 전후에 결합하는 프라이머가 돌연변이 위치의 염기에 따라서 결합하는 정도가 다른 것을 이용하여 돌연변이 양상을 확인하는 방법이 있다. 리버스 도트 블롯(reverse dot blot) 등이 있으며, 형광표지자, 비드 등을 사용한 다양한 방법들을 통해 상기 원리를 이용하여 본 발명의 돌연변이를 포함한 돌연변이의 존재 유무 및 그 양상을 확인할 수 있다.

[0033] (라) 대립형질-특이적 PCR(Allele-specific PCR)은 돌연변이가 일어날 수 있는 유전자 상의 염기 위치 부위에 결합하는 프라이머를 이용한 PCR을 수행함으로써, 돌연변이 위치의 염기에 따라 선택적으로 PCR 반응이 일어나는 원리를 이용하여 돌연변이 양상을 확인하는 방법이다. FRET와 AlphaScan 등이 있으며, 이는 모두 본 발명의 돌연변이의 존재 유무를 확인하기 위해 이용될 수 있다.

[0034] (마) 또 다른 방법으로서, SSCP(Single Strand Conformation Polymorphism)는 유전자의 절편 중에서 단쇄(single strand)를 분석하여 돌연변이 유무를 분석할 수 있다. 이는 돌연변이의 유무에 따라 단쇄의 이중구조가 다르다는 점에 착안하여 전기영동 또는 이와 유사한 결과를 도출할 수 있는 시스템을 이용하여 진단에 이용할 수 있다.

[0035] (바) 또 다른 방법으로서, 이형이중가닥(Heteroduplex) 분석법은 이형이중가닥을 생성시켜 이를 염기서열분석기(예를 들면, ABI377)와 같은 장비를 이용하여 분석하는 방법이다. 이와 유사한 방법으로는 생성된 이형이중가닥을 2-D로 분리하는 TDGS(Two-Dimensional Gene Scanning)와 D-HPLC, DGGE 등이 있으며, 이는 모두 본 발명의 돌연변이의 존재 유무를 확인하기 위해 이용될 수 있다.

[0036] (사) 또 다른 방법으로서, 프라이머 연장(primer extension) 분석법은 돌연변이의 유무를 확인하고자 할 때, 돌연변이가 일어날 수 있는 유전자 상의 염기 위치의 5' 방향에 선택적으로 결합하는 프라이머를 제작하여 부착시킨 후, 중합효소를 이용하여 프라이머를 연장시켜 돌연변이의 유무를 확인하는 방법이다.

[0037] (아) 또 다른 방법으로서, 실시간 PCR(real time PCR) 분석법은 돌연변이의 유무를 확인하고자 할 때, 돌연변이가 일어날 수 있는 유전자 상의 염기 위치 부근에 상보적으로 결합하는 프라이머를 제작하고, 프라이머의 양 말단에 형광표지자와 퀀처(quencher)를 표지하여 PCR의 진행과 함께 유리되는 형광표지자의 양을 측정하여 돌연변이 양상을 확인하는 방법이다. Taqman 등의 방법이 이에 해당되며 이와 유사한 방법으로는 Molecualr beacon 등이 있다. 이는 모두 본 발명의 돌연변이의 존재 유무를 확인하기 위해 이용될 수 있다.

[0038] (자) 또 다른 방법으로서, 올리고뉴클레오티드의 결합을 이용한 분석법(OLA: Oligonucleotide ligation assay)은 돌연변이의 유무를 확인하고자 할 때, 돌연변이가 일어날 수 있는 유전자 상의 염기 위치 전후에 결합하는 프라이머를 제작하여, 돌연변이 위치에 정확히 상보 결합한 프라이머끼리만 결합효소(ligase)에 의한 결합이 이루어지는 원리를 이용하여 돌연변이 양상을 확인하는 방법이다. OLA와 Colorimetric OLA, OLA에 기반한 DNA 칩 등이 있다. 이는 모두 본 발명의 돌연변이의 존재 유무를 확인하기 위해 이용될 수 있다.

[0039] 또한, 본 발명의 일 양태는 상기 폴리뉴클레오티드 또는 이에 상보적인 서열의 폴리뉴클레오티드를 프로브로 포함하는 유전자 돌연변이 검출용 마이크로어레이를 제공한다.

[0040] 마이크로어레이란 지지체인 기판 상에 복수의 프로브를 배열한 것으로, 목적 시료가 마이크로어레이 상의 프로브와 혼성화 되었는지 여부를 관찰함으로써 목적 시료가 소정의 염기 서열을 포함하는지 여부를 확인할 수 있는 수단이다.

[0041] 본 발명의 일 양태에 따른 마이크로어레이에서, 프로브인 폴리뉴클레오티드는 바람직하게는 20 내지 80개, 20 내지 60개, 20 내지 50개 또는 20 내지 40개의 연속된 뉴클레오티드 서열로 구성된 폴리뉴클레오티드이다.

[0042] 본 발명의 일 양태에 따른 마이크로어레이에서, 프로브인 폴리뉴클레오티드는 방사선 표지(radio label), 형광

표지(fluorescent label), 효소 표지(enzymatic label), 서열 택(sequence tag), 또는 바이오틴(biotin) 등에 의하여 표지될 수 있다.

[0043] 상기 마이크로어레이는 표면개질 유리, 실리콘, 또는 나일론 등의 재질로 된 작은 고형의 기판 표면에 FGFR1, FGFR2, 또는 FGFR3 유전자 돌연변이를 포함하는 폴리뉴클레오티드 또는 이에 상보적인 서열의 폴리뉴클레오티드인 프로브를 고밀도로 부착시킨 마이크로 칩일 수 있다. 마이크로어레이 상의 프로브와 시료 내의 표적 물질과의 혼성화(hybridization) 여부를 검출하여 시료의 FGFR1, FGFR2, 또는 FGFR3 유전자의 돌연변이 존재 여부를 확인하고 이를 통해 대상자의 연골저형성증 또는 연골무형성증의 소인(predisposition) 또는 감수성(susceptibility)을 예측할 수 있다.

[0044] 본 발명의 다른 양태는 상기 FGFR1, FGFR2, 또는 FGFR3 유전자 돌연변이에 관한 진단용 조성물에 관한 것이다. 구체적으로는, 상기 FGFR1, FGFR2, 또는 FGFR3 유전자 돌연변이 폴리뉴클레오티드 또는 이에 상보적인 폴리뉴클레오티드를 포함하는 연골저형성증 또는 연골무형성증 진단용 조성물에 관한 것이다.

[0045] 진단용 조성물에는 면역학적 분석에 사용되는 도구/시약이 포함될 수 있다. 이러한 도구/시약으로는 적합한 담체, 검출 가능한 신호를 생성할 수 있는 표지 물질, 용해제, 세정제 등이 포함된다. 또한 표지 물질이 효소인 경우에는 효소 활성을 측정할 수 있는 기질 및 반응 정지제를 포함할 수 있다.

[0046] 본 발명의 또 다른 양태는 상기 FGFR1, FGFR2, 또는 FGFR3 유전자 돌연변이 폴리뉴클레오티드 또는 이에 상보적인 폴리뉴클레오티드를 포함하는 연골저형성증 또는 연골무형성증 진단용 키트에 관한 것이다.

[0047] 상기 진단용 키트는 분석 방법에 적합한 한 종류 또는 그 이상의 다른 구성 성분 조성물, 용액 또는 장치로 구성된다.

[0048] 진단용 키트(kit)는 상기 FGFR1, FGFR2, 또는 FGFR3 유전자 돌연변이에 관한 마이크로어레이 또는 ELISA(Enzyme Liked ImmunoSorbent Assay)을 이용할 수 있다.

[0049] ELISA(효소 면역 분석법)은 측정하고자 하는 물질을 항원으로 하여 이에 대한 특정 항체를 만들어서 항원에 결합할 수 있도록 하고 항원 항체의 결합여부와 결합정도를 효소, 방사선 물질, 형광물질과 같은 표지를 사용하여 측정함으로써 생체물질을 정성, 정량하는 방법이다.

발명의 효과

[0050] 본 발명의 폴리뉴클레오티드는 연골저형성증 또는 연골무형성증 유발 유전자로 알려진 FGFR1, FGFR2, 또는 FGFR3의 뉴클레오티드 서열상에 존재하는 한국인 특이적인 돌연변이의 유형을 제공하여, 지금까지 진단이 어려웠던 연골저형성증 및 연골무형성증의 소인을 저비용, 고효율로 진단할 수 있게 한다.

도면의 간단한 설명

[0051] 도 1은 FGFR1 유전자의 변이부위의 염기서열분석 결과를 표시한다.

도 2는 FGFR2 유전자의 변이부위의 염기서열분석 결과를 표시한다.

도 3은 FGFR3 유전자의 변이부위의 염기서열분석 결과를 표시한다.

발명을 실시하기 위한 구체적인 내용

[0052] 이하 본 발명을 실시예를 통하여 보다 상세하게 설명한다. 그러나 이들 실시예는 본 발명을 예시적으로 설명하기 위한 것으로 본 발명의 범위가 이들 실시예에 한정되는 것은 아니다.

[0053] FGFR1, FGFR2, FGFR3 유전자의 엑손 전역과 엑손-인트론 접경분위(exon-intron junction)에 대한 염기서열을 분석함으로써 유전자에 대한 돌연변이 여부를 분석하였다. NCBI (National Center for Biotechnology Information) 데이터베이스에서 유전자 서열(genomic sequence)을 얻어서, 단일염기 다형성을 찾기 위한 프라이머를 디자인하고 상기 프라이머들을 이용하여, 유전자의 엑손 부위를 PCR로 증폭하고, 염기서열을 분석하여 유전적 다양성을 분석하였다.

[0054] **실시예 1: FGFR1, FGFR2, FGFR3 유전자의 돌연변이 검출 및 확인**

[0055] 1. 게놈 DNA 분리

[0056] 연골저형성증 환자 각 78명, 연골무형성증 환자 각 52명, 및 정상군 41명을 확보하고 이들의 말초혈액에서 다음의 순서로 게놈 DNA(genomic DNA)를 추출하였다.

[0057] 1) Whole blood 200~250ul 와 AP-1(Cell lysis buffer) 500ul를 넣고 10초 볼텍싱(vortexing)함.

[0058] 2) AP-2(Protein-depleting buffer) 100ul를 넣고 10초 볼텍싱(vortexing)함.

[0059] 3) 12000rpm에서 10분간 원심분리(centrifuge)

[0060] 4) 2ml 마이크로퓨지 튜브(microfuge tube)에 필터를 끼우고 원심분리한 상층액을 넣음.

[0061] 5) 6000rpm 에서 1분간 원심분리

[0062] 6) 세척(Washing) 단계

[0063] - w1(Wash buffer) 700ul 넣고 2분 상온보관 후 6000rpm에서 1분간 원심분리함.

[0064] - w2(Desalting buffer, Either 100% or 95% denatured ethanol) 800ul 넣고 12000rpm에서 1분간 원심분리함.

[0065] - w2(Desalting buffer, Either 100% or 95% denatured ethanol) 500ul 넣고 12000rpm에서 1분간 원심분리함.

[0066] - 버퍼를 버리고 빈 튜브를 1분간 12000rpm 원심분리함.

[0067] 7) 필터만 1.5ml 마이크로퓨지 튜브(microfuge tube)에 옮김.

[0068] 8) TE 버퍼(5mM Tris-HCl, 0.1mM EDTA, pH8.5) 80ul씩 넣고, 1분간 기다렸다가 12000rpm 원심분리함.

[0069] 9) 1~3ul정도 전기영동으로 DNA를 확인함.

[0070] 2. 중합효소 연쇄 반응(PCR)을 통한 FGFR1, FGFR2, FGFR3 유전자의 증폭

[0071] FGFR1, FGFR2, FGFR3 각각에 대하여, 하기 표 1 내지 3과 같이 프라이머를 설계하였다.

표 1

exon	FGFR1 primer sequence	
1	F	5'-CCA AAG GAG ACC AGG CTG TA-3'
	R	5'-AGG GAA GGG AGG GGA GAC-3'
2	F	5'-CAC TTT AAG CAG CCA CCA CA-3'
	R	5'-CCT GTT GAC CAC ATC ACC TG-3'
3	F	5'-ACT GAC CCA TGC TCC AAT TC-3'
	R	5'-TAG TGG GCA GCA GTT TCT GA-3'
4	F	5'-GAG GGT GTC CGT GTT CAT CT-3'
	R	5'-TTC CAT CTT TTC TGG GGA TG-3'
5	F	5'-CGG TGC CTG TTC TGC TAA GT-3'
	R	5'-CCT CCC CTG TTC CCA TTA CT-3'
6	F	5'-CCA GCT TGC AGA GTG TGA AC-3'
	R	5'-TCC TGA CTC TGC CCC TAA GA-3'
7	F	5'-CCC CAG CTC TGT TTT GAG AG-3'
	R	5'-CTC TGT TCC CAG CTG ACC TC-3'
8	F	5'-CCT GGT GGT TTT CGT TTG TT-3'
	R	5'-CCA TTC TTC CAG CTC TCC TG-3'
9	F	5'-CCT AGT GCC ACA ATC CCA GT-3'
	R	5'-GAT ATG GAG GGT GCA GGA AA-3'
10	F	5'-TGG GAA GCC CTG ACT AAG AA-3'
	R	5'-GCA GGC ATT TCA TGA ACC TT-3'
11	F	5'-TCG CAG TTG GGT GAA ATA CA-3'
	R	5'-ACG CCA AGG AAC AGA GTG AC-3'
12	F	5'-CTG CCC ACT CAC TAC CAC CT-3'
	R	5'-TAA CCC CCT TCC CTA GCT GT-3'
13	F	5'-GTG GGG TTT CTT TGA GGT GA-3'
	R	5'-GCT CCT GAC TCT GCA GTT CC-3'
14~15	F	5'-CTG GAG CTG GCA GAG AAT GT-3'
	R	5'-TCA GTC CAG GGA GAA AGC AG-3'
16~17	F	5'-ACT GCC CTG GGT AGA GGA TT-3'
	R	5'-AAA GCT GGG GGA GTA CTG GT-3'
18	F	5'-CCT TGA CCT CCA ACC AGG TA-3'
	R	5'-AAA CAG ACC AAA CCG ACA GG-3'

[0072]

표 2

exon	FGFR2 primer sequence	
1	F	5'-AAC GGC TCG GGT TTC AGT-3'
	R	5'-GAG GCA GAA CTT CCA GAA CG-3'
2	F	5'-CAC TTG GGC TGG AGT GAT TT-3'
	R	5'-AGC CCC ATT TAC AAT GCA AG-3'
3	F	5'-CCA TTT GCA GAG CAA GAA CA-3'
	R	5'-TTA ATG ATC GGC CTT TCT GG-3'
4	F	5'-CCT GGG TTG TTG ACT TTG CT-3'
	R	5'-CTG TGC AGC AAC ACT GGT TT-3'
5	F	5'-CAT GAA CCA TCC CTC TCT GG-3'
	R	5'-ATG GTG GCA CAC ACC TGT AG-3'
6	F	5'-TCC ATA CCT TTC TTG CCT CCT-3'
	R	5'-CGT TCA TGC TTT CAA ACG AG-3'
7	F	5'-CGG CAG TCT CCT TTG AAG TC-3'
	R	5'-CAT GAA GGA GAC CCC AGT TG-3'
8	F	5'-TCA GTG TTG CTC CGT GCT TC-3'
	R	5'-TGG GAT CTC ACT GTG TGT GAA-3'
9	F	5'-CAC CTT GCC TTG AGA ATG GT-3'
	R	5'-GCA GAT GGA AGC TCA CAG AA-3'
10	F	5'-TGG AGG GCA GAG GAG TTA GA-3'
	R	5'-CCT GCT GAC ATC ATC ACA CC-3'
11	F	5'-AGC CGT TCC ACC AGT AAT TG-3'
	R	5'-TTC ACA TGC CAC AAA AGG AA-3'
12	F	5'-TCC CTG TGT AGG CCT TTG TC-3'
	R	5'-TCT AGG CCC TCA AAT CCA AA-3'
13	F	5'-CCT CTG TAA GGG CTG GGA TT-3'
	R	5'-GGT TTA TGA GGC TGC TTT GG-3'
14	F	5'-ACT GCC CTG GGT AGA GGA TT-3'
	R	5'-AAA GCT GGG GGA GTA CTG GT-3'
15	F	5'-CCT TGA CCT CCA ACC AGG TA-3'
	R	5'-AAA CAG ACC AAA CCG ACA GG-3'
16	F	5'-ATC CCC GTG TCA TCA TCA TT-3'
	R	5'-CCA GAG AGC TTG AGC CAT TC-3'
17	F	5'-ATG GCT TCG ATG CCT TTC TA-3'
	R	5'-GGG GAT GTC CCT CTG TCT TT-3'
18	F	5'-CGT CCA ATA CCC ATC TCT CA-3'
	R	5'-TGC ATT GGG ACA TCC ATT TA-3'

[0073]

표 3

exon	FGFR3 primer sequence	
1	F	5'-TCT AAC GAG CTG CCT TCC TC-3'
	R	5'-CCT CGC ACT TGG AGG GTA G-3'
2	F	5'-AGA GAC CTT TGG CCC TTT TG-3'
	R	5'-TCC ACC AAT GAC CAG AGA GA-3'
3	F	5'-ACG CCA GCT ACA GAA AGG AG-3'
	R	5'-TTA GTC CCT CAG CTG CCT GT-3'
4~6	F	5'-TCC TAC ACA GGA CGG GAA AC-3'
	R	5'-CCT AGA CCC AAA TCC TCA CG-3'
7	F	5'-CTC ACC TGG GAG AGA GGA CT-3'
	R	5'-GGC CGT AAG TCA CAG GAT CT-3'
8	F	5'-TCA CTG GCG TTA CTG ACT GC-3'
	R	5'-TTC CAC CAG CAT TGA ATG AA-3'
9	F	5'-TTC ATT CAA TGCTGG TGG AA-3'
	R	5'-GAC GAA GAG TGT CAC CCA CA-3'
10~12	F	5'-AGG CTG TGG GTG ACA CTC TT-3'
	R	5'-CGT TCA TGC TTT CAA ACG AG-3'
13~14	F	5'-GTA ACC TGC GGG AGT TTC TG-3'
	R	5'-ACT CCC AGG ACA GAC ACC TG-3'
15~16	F	5'-GCT CCT GGG TGT GGT TTC TA-3'
	R	5'-AGG TCC AGG TAC TCC TGC TG-3'
17	F	5'-GAC CGT GTC CTT ACC GTG AC-3'
	R	5'-ACT AAC ACC CCA GGC ACA AG-3'

[0074]

[0075] DNA 200ng, 20pmol의 프라이머, 200uM dNTP, 3U의 Taq 폴리머라제(polymerase)와 버퍼(25mM TAPS(pH9.3), 50mM KCl, 2mM $MgCl_2$, 1mM 머캡토에탄올(mercaptoethanol)를 넣고 95℃에서 5분간 반응시키고 95℃서 1분, 68℃에서 1분, 72℃에서 1분간 30회 동안 PCR을 시행하였다. PCR 생성물 5ul와 Bigdye 8ul, 버퍼 4ul 증류수 3ul를 넣고 95℃에서 5분간 반응시키고 95℃에서 1분, 50℃에서 1분, 63℃에서 1분간 30회 동안 PCR을 하였다.

[0076] 3. DNA 염기서열분석

[0077] 생성된 PCR 생성물은 100% 에탄올 40ul를 넣고 15분간 실온에 방치 후 20분 동안 4℃에서 원심분리하고 상층액을 버린 후 포름아미드(formamide) 10ul로 재부유하고 95℃에서 3분간 노출시키고 얼음에서 5분간 방치 후 3100 염기서열분석기에서 분석하였다.

[0078] 4. 염기서열 분석 결과

[0079] 연골저형성증 환자의 유전자 검사는 총 78명(평균나이 28.3세)에게 시행되었고, 이 중 33명(남 14명(27.1세), 여 19명(29.2세))의 환자에게서 유전자 돌연변이가 검출되었다. 반면, FGFR 1, FGFR2, FGFR3 유전자의 돌연변이는 정상군 41명의 경우에서는 거의 대부분 검출되지 않았다. 이러한 유전자 돌연변이 발견과 관련하여, 환자간 가족관계가 발견되었다. 이러한 연골저형성증과 관련된 FGFR1, FGFR2, FGFR3 유전자 돌연변이의 염기서열 분석 결과를 하기 표4에 도시하였다. FGFR1, FGFR2, FGFR3 각각의 염기서열 분석 결과는 도1, 도2, 및 도3에 각각 도시하였다.

표 4

유전자	위치	엑손	코돈	뉴클레오티드	아미노산	케이스	정상군 41건중(%)
FGFR1	3	3	107	TCG → TYG	Ser → Ser/Leu	5	-
		10	450	delTCC	delSer	3	-
		12	552	ACG → AYG	Thr → Thr/Met	1	-
FGFR2	3	2	6	CGT → CST	Arg → Arg/Pro	1	1(2.4)
		6	186	ATG → ACG	Met → Thr	1	-
				ATG → AYG	Met → Met/Thr	9	-
		7	250	TCG → TYG	Ser → Ser/Leu	1	-
FGFR3	6	2	65	GGG → RGG	Gly → Gly/Arg	3	-
			115	CAG → CWG	Gln → Gln/Leu	3	-
		4	164	CTG → STG	Leu → Leu/Val	1	1(2.4)
		6	264	ACG → AYG	Thr → Thr/Met	2	-
			278	TAC → TRC	Tyr → Tyr/Cys	1	-
		11	540	AAC → AAM	Asn → Asn/Lys	3	-
total	12					34	2(4.9)

[0080]

[0081] 상기 실험을 통해, 이미 알려진 FGFR1, FGFR2, FGFR3 유전자의 돌연변이 이외에 신규한 FGFR1, FGFR2, FGFR3 돌연변이 유전자를 발견하였고, 이를 정리하면 하기 표5와 같다.

표 5

	Gene	알려진 돌연변이	신규 발견 유전자
1	FGFR1	Ser107Leu	450delSer, Thr552Met
2	FGFR2	Arg6Pro	Ser250Leu, Met186Thr
3	FGFR3	Gly65Arg, Trp278Cys, Gly380Arg, Asn540Lys	Gln115Leu, Leu164Val, Thr264Met

[0082]

[0083] 연골무형성증 환자의 유전자 검사는 총 52명(평균 나이 21세)의 환자의 샘플에서 시행되었고, 이 중 51명(평균 나이 20.4세)의 환자에게서 연골무형성증에게서 나타나는 FGFR3 유전자의 엑손 8에서 G380R 돌연변이가 검출되었으며, 그 분포는 각 남 21(19.5세)명, 여 30(21.4세)로 나타났다. 51명의 환자 중에 3명의 환자에게서 FGFR3 N540K 유전자와 FGFR2 유전자의 엑손 5 M186T 돌연변이가, 1명의 환자에게서 FGFR3 N540K 유전자와 FGFR1 엑손 3 S107L 돌연변이가 중복적으로 발견되었다. 이중 FGFR2 유전자의 엑손 5 M186T 돌연변이가 발견된 2명의 환자는 가족관계였다. 반면, FGFR 1, FGFR2, 및 FGFR3 유전자의 돌연변이는 정상군 41명의 경우에서는 검출되지 않았다. 하기 표6은 이러한 연골무형성증과 관련된 FGFR1, FGFR2, FGFR3 유전자 돌연변이의 염기서열 분석 결과를 도시한다.

표 6

유전자	위치	엑손	코돈	뉴클레오티드	아미노산	케이스	정상군 41건중(%)
FGFR1	1	3	107	TCG → TYG	Ser → Ser/Leu	1	-
FGFR2	1	6	186	ATG → AYG	Met → Met/Thr	3	-
FGFR3	1	8	380	GGG → RGG	Gly → Gly/Arg	48	-
total	3					52	

[0084]

[0085] **실시예 2: 정상군의 FGFR1, FGFR2, FGFR3 유전자의 돌연변이 존재 여부 확인**

[0086] 연골저형성증 및 연골무형성증과 상기 유전자의 돌연변이와의 상관관계를 재확인하기 위하여 추가로, 실시예 1과 동일한 방법으로 정상군 68명의 말초혈액에서 게놈 DNA를 분리하고, PCR을 통해 FGFR1, FGFR2, FGFR3 유전자를 증폭하여 DNA를 분석하였다.

[0087] 그러나, 연골저형성증이나 연골무형성증이 나타나지 않는 정상군을 대상으로 한 유전자 분석의 결과 상기와 같은 FGFR1, FGFR2, FGFR3 유전자의 돌연변이는 나타나지 않았다. 이러한 결과를 하기 표 7에 도시하였다.

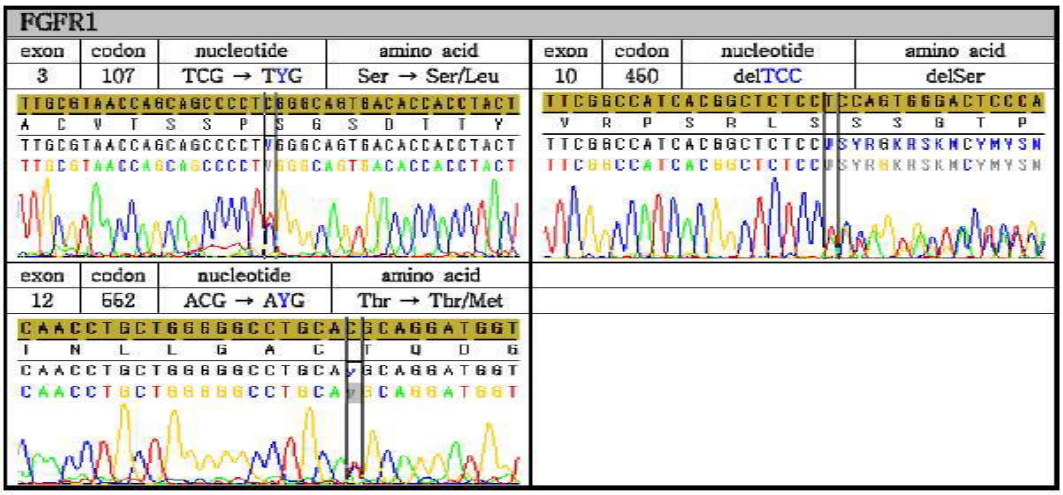
표 7

[0088]

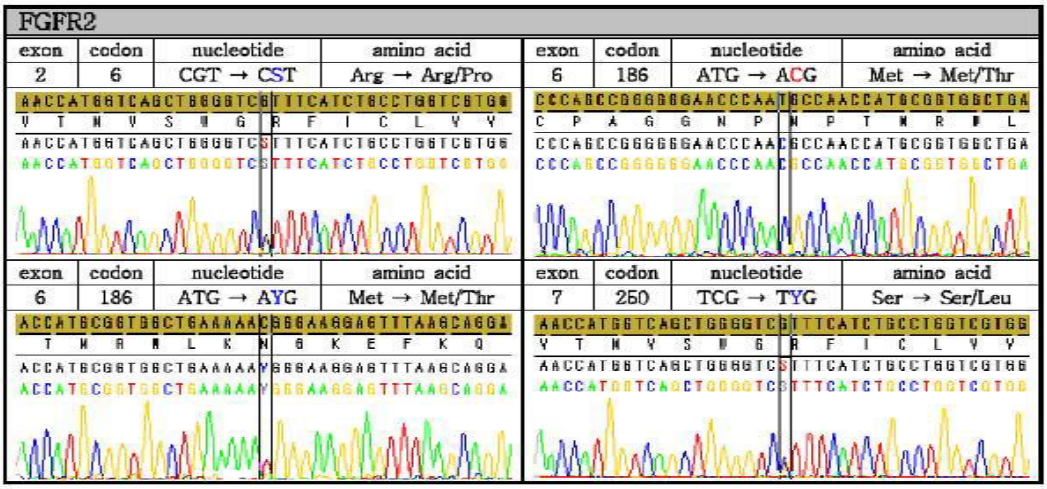
유전자	위치	엑손	코돈	아미노산	검출된 환자수	정상인(normal person)
FGFR1	3	10	450	delSer	3	총68명 중 돌연변이 0명
FGFR1	3	12	552	Thr → Thr/Met	1	총68명 중 돌연변이 0명
FGFR2	3	6	186	Met → Thr	1	총64명 중 돌연변이 0명
				Met → Met/Thr	13	
FGFR3	7	2	115	Gln → Gln/Leu	3	총68명 중 돌연변이 0명
FGFR3	7	2	65	Gly → Gly/Arg	3	총68명 중 돌연변이 1명

도면

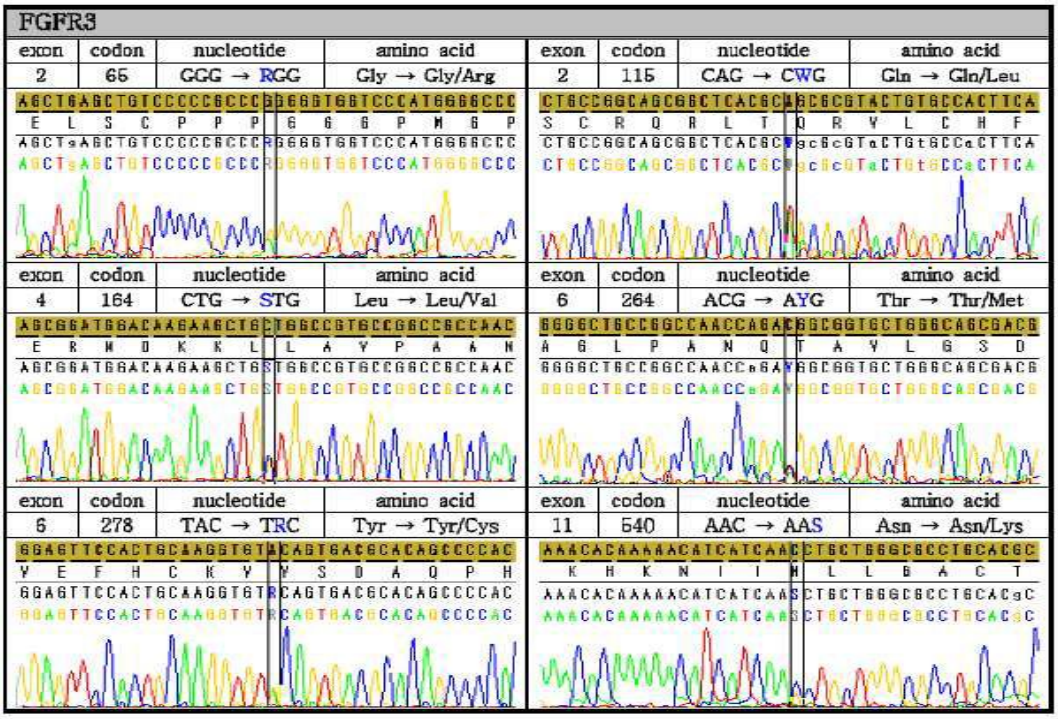
도면1



도면2



도면3



서 열 목 록

<110> Korea University Research and Business Foundation

<120> FGFR1, FGFR2, and FGFR3 Gene Mutations Useful For Predicting or Detecting Hypochondroplasia or Achondroplasia

<130> HY100477

<160> 3

<170> KopatentIn 1.71

<210> 1

<211> 5917

<212> DNA

<213> Homo sapiens

<400> 1

agatgcaggg gcgcaaacgc caaaggagac caggctgtag gaagagaagg gcagagcgcc 60

ggacagctcg gcccgctccc cgtcctttgg ggccgcggct ggggaactac aaggcccagc 120

aggcagctgc agggggcgga ggcggaggag ggaccagcgc gggtgggagt gagagagcga 180

gccctcgcgc cccgccggcg catagcgctc ggagcgctct tgcggccaca ggcgcggcgt 240

cctcggcggc gggcggcagc tagcgggagc cgggacgccg gtgcagccgc agcgcgcgga 300

ggaacccggg tgtgccggga gctgggcggc cacgtccgga cgggaccgag acccctcgta 360

gcgcattgcg gcgacctcgc cttccccggc cgcgagcgcg ccgctgcttg aaaagccgcg 420

gaacccaagg acttttctcc ggtccgagct cggggcgccc cgcagggcgc acggtacccg 480

tgctgcagtc gggcacgccg cggcgccggg gcctccgcag ggcgatggag cccggtctgc 540

aaggaaagtg aggcgccgcc gctgcgttct ggaggagggg ggcacaaggt ctggagaccc 600

cgggtggcgg acgggagccc tccccccgcc ccgcctccgg ggcaccagct ccggctccat 660

tgttcccgcc cgggctggag gcgccgagca ccgagcgccg ccgggagtcg agcgccggcc 720

gcggagctct tgcgaccccg ccaggacccg aacagagccc gggggcggcg ggccggagcc 780

ggggacgcgg gcacacgccc gctcgcacaa gccacggcgg actctcccga ggcggaacct 840

ccacgccgag cgagggtcag tttgaaaagg aggatcgagc tcactgtgga gtatccatgg 900

agatgtggag ccttgtcacc aacctctaac tgcagaactg ggatgtggag ctggaagtgc 960

ctcctcttct gggctgtgct ggtcacagcc acactctgca ccgctaggcc gtccccgacc 1020

ttgcctgaac aagcccagcc ctggggagcc cctgtggaag tggagtcctt cctggtccac 1080

cccggtgacc tgctgcagct tcgctgtcgg ctgcgggacg atgtgcagag catcaactgg 1140

ctgcgggacg gggtgcagct ggcggaaagc aaccgcaccc gcatcacagg ggaggaggtg 1200

gaggtgcagg actccgtgcc cgcagactcc ggcctctatg cttgcgtaac cagcagcccc 1260

tcgggcagtg acaccaccta cttctccgtc aatgtttcag atgctctccc ctcctcggag 1320

gatgatgatg atgatgatga ctcctcttca gaggagaaag aaacagataa caccaaacca 1380

aaccgtatgc ccgtagctcc atattggaca tccccagaaa agatggaaaa gaaattgcat 1440

gcagtgccgg ctgccaagac agtgaagttc aaatgccctt ccagtgggac cccaaacccc 1500

acactgcgct ggttgaaaaa tggcaaagaa ttcaaacctg accacagaat tggaggctac 1560

aaggtccgtt atgccacctg gagcatcata atggactctg tggtgccctc tgacaagggc 1620

aactacacct gcattgtgga gaatgagtac ggcagcatca accacacata ccagctggat 1680

gtcgtggagc ggtcccctca ccggcccatc ctgcaagcag ggttgcccgc caacaaaaca 1740

gtggccctgg gtagcaacgt ggagttcatg tgtaaggtgt acagtgaccc gcagccgcac 1800

atccagtggc taaagcacat cgaggtgaat gggagcaaga ttggcccaga caacctgcct 1860

tatgtccaga tcttgaagac tgctggagtt aataccaccg acaaagagat ggaggtgctt 1920

cacttaagaa atgtctcctt tgaggacgca ggggagtata cgtgcttggc gggtaactct 1980

atcggactct cccatcactc tgcatggttg accgttctgg aagccctgga agaggaggccg 2040

gcagtgatga cctcgcccct gtacctggag atcatcatct attgcacagg ggccttcctc 2100

atctcctgca tggtggggtc ggtcatcgtc tacaagatga agagtggtac caagaagagt 2160

gacttccaca gccagatggc tgtgcacaag ctggccaaga gcatccctct gcgcagacag 2220

gtaacagtgt ctgctgactc cagtgcatcc atgaactctg gggttcttct ggttcggcca 2280

tcacggctct cctccagtgg gactcccatg ctagcagggg tctctgagta tgagcttccc 2340

gaagaccctc gctgggagct gcctcgggac agactggtct taggcaaacc cctgggagag 2400

ggctgctttg ggcaggtggt gttggcagag gctatcgggc tggacaagga caaacccaac 2460

cgtgtgacca aagtggctgt gaagatgttg aagtcggacg caacagagaa agacttgtca 2520

gacctgatct cagaaatgga gatgatgaag atgatcggga agcataagaa tatcatcaac 2580

ctgctggggg cctgcacgca ggatggtccc ttgtatgtca tcgtggagta tgcctccaag 2640

ggcaacctgc gggagtacct gcaggcccgg aggccccccag ggctggaata ctgctacaac 2700

cccagccaca acccagagga gcagctctcc tccaaggacc tggtgtcctg cgcctaccag 2760

gtggcccgag gcatggagta tctggcctcc aagaagtgca tacaccgaga cctggcagcc 2820

aggaatgtcc tggtgacaga ggacaatgtg atgaagatag cagactttgg cctcgcacgg 2880

gacattcacc acatcgacta ctataaaaag acaaccaacg gccgactgcc tgtgaagtgg 2940

atggcacccg aggcattatt tgaccggatc tacacccacc agagtgatgt gtggtctttc 3000

ggggtgctcc tgtgggagat cttcactctg ggcggctccc cataccccgg tgtgcctgtg 3060

gaggaacttt tcaagctgct gaaggagggt caccgcatgg acaagcccag taactgcacc 3120

aacgagctgt acatgatgat gcgggactgc tggcatgcag tgccctcaca gagacccacc 3180

ttcaagcagc tggtggaaga cctggaccgc atcgtggcct tgacctccaa ccaggagtac 3240

ctggacctgt ccatgcccct ggaccagtac tcccccagct ttcccgacac ccggagctct 3300

acgtgctcct caggggagga ttccgtcttc tctcatgagc cgctgcccga ggagccctgc 3360

ctgccccgac acccagccca gcttgccaat ggcggactca aacgccgctg actgccaccc 3420

acacgccctc cccagactcc accgtcagct gtaaccctca cccacagccc ctgctgggcc 3480

caccacctgt ccgtccctgt ccccttttcct gctggcagga gccggctgcc taccaggggc 3540

cttcctgtgt ggcctgcctt caccccactc agctcacctc tccctccacc tcctctccac 3600

ctgctggtga gaggtgcaaa gaggcagatc tttgctgcca gccacttcat cccctcccag 3660

atgttggacc aacacccctc cctgccacca ggcactgcct ggagggcagg gagtgggagc 3720

caatgaacag gcatgcaagt gagagcttcc tgagctttct cctgtcggtt tggtctgttt 3780

tgccttcacc cataagcccc tcgcactctg gtggcaggtg ccttgtcctc agggctacag 3840

cagtagggag gtcagtgctt cgtgcctcga ttgaaggtga cctctgcccc agataggtgg 3900

tgccagtggc ttattaattc cgatactagt ttgctttgct gaccaaatgc ctggtaccag 3960

aggatggtga ggcgaaggcc aggttggggg cagtgttgtg gccctggggc ccagccccaa 4020

actgggggct ctgtatatag ctatgaagaa aacacaaagt gtataaatct gagtatatat 4080

ttacatgtct ttttaaaagg gtcgttacca gagatttacc catcgggtaa gatgctcctg 4140

gtggctggga ggcatcagtt gctatatatt aaaaacaaaa aagaaaaaaa aggaaaatgt 4200

ttttaaaaag gtcatatatt ttttgctact tttgctgttt tatttttta aattatgttc 4260

taaacctatt ttcagtttag gtccctcaat aaaaattgct gctgcttcat ttatctatgg 4320

gctgtatgaa aagggtggga atgtccactg gaaagaaggg acacccacgg gccctggggc 4380

taggtctgtc ccgagggcac cgcatgctcc cggcgcaggt tccttgtaac ctcttcttcc 4440

taggtcctgc acccagacct cacgacgcac ctcctgcctc tccgctgctt ttggaaagtc 4500

agaaaaagaa gatgtctgct tcgagggcag gaaccccatc catgcagtag aggcgctggg 4560

cagagagtca aggcccagca gccatcgacc atggatggtt tcctccaagg aaaccggtgg 4620

ggttgggctg gggaggggggc acctacctag gaatagccac ggggtagagc tacagtgatt 4680

aagaggaaag caagggcgcg gttgctcacg cctgtaatcc cagcactttg ggacaccgag 4740

gtgggcagat cacttcaggt caggagtttg agaccagcct ggccaactta gtgaaacccc 4800

atctctacta aaaatgcaaa aattatccag gcatggtggc acacgcctgt aatcccagct 4860

ccacaggagg ctgaggcaga atcccttgaa gctgggaggc ggaggttgca gtgagccgag 4920

attgcgccat tgcactccag cctgggcaac agagaaaaca aaaaggaaaa caaatgatga 4980

aggtctgcag aaactgaaac ccagacatgt gtctgccccc tctatgtggg catggttttg 5040

ccagtgcttc taagtgcagg agaacatgtc acctgaggct agttttgcat tcaggtccct 5100

ggcttcgttt cttgttggta tgcctcccca gatcgtcctt cctgtatcca tgtgaccaga 5160

ctgtatttgt tgggactgtc gcagatcttg gcttcttaca gttcttcctg tccaaactcc 5220

atcctgtccc tcaggaacgg ggggaaaatt ctccgaatgt ttttggtttt ttggctgctt 5280

ggaatttact tctgccacct gctggtcatc actgtcctca ctaagtggat tctggctccc 5340

ccgtacctca tggctcaaac taccactcct cagtcgctat attaaagctt atattttgct 5400

ggattactgc taaatacaaa agaaagttca atatgttttc atttctgtag ggaaaatggg 5460

attgctgctt taaatttctg agctagggat tttttggcag ctgcagtgtt ggcgactatt 5520

gtaaaattct ctttgtttct ctctgtaaat agcacctgct aacattacaa tttgtattta 5580

tgtttaaaga aggcatcatt tggtgaacag aactaggaaa tgaattttta gctcttaaaa 5640

gcatttgctt tgagaccgca caggagtgtc tttccttgta aaacagtgat gataatttct 5700

gccttggccc taccttgaag caatgttgtg tgaagggatg aagaatctaa aagtcttcat 5760

aagtccttgg gagaggtgct agaaaaatat aaggcactat cataattaca gtgatgtcct 5820

tgctgttact actcaaatca cccacaaatt tccccaaaga ctgcgctagc tgtcaaataa 5880

aagacagtga aattgacctg aaaaaaaaaa aaaaaaa 5917

<210> 2

<211> 4654

<212> DNA

<213> Homo sapiens

<400> 2

ggcggcggct ggaggagagc gcggtggaga gccgagcggg cgggcggcgg gtgcggagcg 60

ggcgagggag cgcgcgcggc cgccacaaag ctcgggcgcc gcggggctgc atgcggcgta 120

cctggcccgg cgcggcgact gctctccggg ctggcggggg ccggccgcga gccccggggg 180

ccccgaggcc gcagcttgcc tgcgcgctct gagccttcgc aactcgcgag caaagtttgg 240

tggaggcaac gccaagcctg agtcctttct tcctctcgtt ccccaaatcc gagggcagcc 300

cgcgggcgtc atgcccgcgc tcctccgcag cctggggtac gcgtgaagcc cgggaggctt 360

ggcgccggcg aagacccaag gaccactctt ctgcgtttgg agttgctccc cgcaaccccg 420

ggctcgtcgc tttctccatc ccgacccacg cggggcgcgg ggacaacaca ggtcgcggag 480

gagcgttgcc attcaagtga ctgcagcagc agcggcagcg cctcggttcc tgagcccacc 540

gcaggctgaa ggcattgcgc gtagtccatg cccgtagagg aagtgtgcag atgggattaa 600

cgtccacatg gagatatgga agaggaccgg ggattggtac cgtaaccatg gtcagctggg 660

gtcgtttcat ctgcctggtc gtggtcacca tggcaacctt gtccctggcc cggccctcct 720

tcagtttagt tgaggatacc acattagagc cagaagagcc accaaccaaa taccaaatct 780

ctcaaccaga agtgtacgtg gctgcgccag gggagtcgct agaggtgcgc tgcctgttga 840

aagatgccgc cgtgatcagt tggactaagg atggggtgca cttggggccc aacaatagga 900

cagtgcttat tggggagtac ttgcagataa agggcgccac gcctagagac tccggcctct 960

atgcttgtac tgccagtagg actgtagaca gtgaaacttg gtacttcatg gtgaatgtca 1020

cagatgccat ctcatccgga gatgatgagg atgacaccga tggtgcggaa gattttgtca 1080

gtgagaacag taacaacaag agagcaccat actggaccaa cacagaaaag atggaaaagc 1140

ggctccatgc tgtgcctgcg gccaacactg tcaagtttcg ctgcccagcc gggggggaacc 1200

caatgccaac catgcggtgg ctgaaaaacg ggaaggagtt taagcaggag catcgcattg 1260

gaggctacaa ggtacgaaac cagcactgga gcctcattat ggaaagtgtg gtcccatctg 1320

acaagggaaa ttatacctgt gtagtggaga atgaatacgg gtccatcaat cacacgtacc 1380

acctggatgt tgtggagcga tcgcctcacc ggcccatcct ccaagccgga ctgccggcaa 1440

atgcctccac agtggtcgga ggagacgtag agtttgtctg caaggtttac agtgatgccc 1500

agccccacat ccagtggatc aagcacgtgg aaaagaacgg cagtaaatac gggcccgacg 1560

ggctgcccta cctcaaggtt ctcaaggccg ccggtgttaa caccacggac aaagagattg 1620

aggttctcta tattcggaat gtaacttttg aggacgctgg ggaatatacg tgcttggcgg 1680

gtaattctat tgggatatcc tttcactctg catggttgac agttctgcca gcgcctggaa 1740

gagaaaagga gattacagct tccccagact acctggagat agccatttac tgcatagggg 1800

tcttcttaat cgcctgtatg gtggtaacag tcatcctgtg ccgaatgaag aacacgacca 1860

agaagccaga cttcagcagc cagccggctg tgcacaagct gaccaaacgt atccccctgc 1920

ggagacaggt aacagtttcg gctgagtcca gctcctccat gaactccaac accccgctgg 1980

tgaggataac aacacgcctc tcttcaacgg cagacacccc catgctggca ggggtctccg 2040

agtatgaact tccagaggac ccaaaatggg agtttccaag agataagctg acactgggca 2100

agcccctggg agaaggttgc tttgggcaag tggtcatggc ggaagcagtg ggaattgaca 2160

aagacaagcc caaggaggcg gtcaccgtgg ccgtgaagat gttgaaagat gatgccacag 2220

agaaagacct ttctgatctg gtgtcagaga tggagatgat gaagatgatt gggaaacaca 2280

agaatatcat aaatcttctt ggagcctgca cacaggatgg gcctctctat gtcatagttg 2340

agtatgcctc taaaggcaac ctccgagaat acctccgagc ccggaggcca cccgggatgg 2400

agtactccta tgacattaac cgtgttcctg aggagcagat gaccttcaag gacttggtgt 2460

catgcaccta ccagctggcc agaggcatgg agtacttggc ttcccaaaaa tgtattcatc 2520

gagatttagc agccagaaat gttttggtaa cagaaaacaa tgtgatgaaa atagcagact 2580

ttggactcgc cagagatatc aacaatatag actattacaa aaagaccacc aatgggcggc 2640

ttccagtcaa gtggatggct ccagaagccc tgtttgatag agtatacact catcagagtg 2700

atgtctggtc cttcggggtg ttaatgtggg agatcttcac tttaggggggc tcgccctacc 2760

cagggattcc cgtggaggaa ctttttaagc tgctgaagga aggacacaga atggataagc 2820

cagccaactg caccaacgaa ctgtacatga tgatgaggga ctgttggcat gcagtgccct 2880

cccagagacc aacgttcaag cagttggtag aagacttgga tcgaattctc actctcacaa 2940

ccaatgagga atacttggac ctcagccaac ctctcgaaca gtattcacct agttaccctg 3000

acacaagaag ttcttgttct tcaggagatg attctgtttt ttctccagac cccatgcctt 3060

acgaaccatg ccttcctcag tatccacaca taaacggcag tgttaaaaca tgaatgactg 3120

tgtctgcctg tccccaaaca ggacagcact gggaacctag ctacactgag cagggagacc 3180

atgcctccca gagcttgttg tctccacttg tatatatgga tcagaggagt aaataattgg 3240

aaaagtaatc agcatatgtg taaagattta tacagttgaa aacttgtaat cttccccagg 3300

aggagaagaa ggtttctgga gcagtggact gccacaagcc accatgtaac ccctctcacc 3360

tgccgtgcgt actggctgtg gaccagtagg actcaaggtg gacgtgcgtt ctgccttcct 3420

tgttaatttt gtaataattg gagaagattt atgtcagcac acacttacag agcacaaatg 3480

cagtatatag gtgctggatg tatgtaaata tattcaaatt atgtataaat atatattata 3540

tatttacaag gagttatttt ttgtattgat tttaaatgga tgtcccaatg cacctagaaa 3600

attggtctct ctttttttaa tagctatttg ctaaatgctg ttcttacaca taatttctta 3660

attttcaccg agcagaggtg gaaaaatact tttgctttca gggaaaatgg tataacgtta 3720

atttattaat aaattggtaa tatacaaaac aattaatcat ttatagtttt ttttgtaatt 3780

taagtggcat ttctatgcag gcagcacagc agactagtta atctattgct tggacttaac 3840

tagttatcag atcctttgaa aagagaatat ttacaatata tgactaattt ggggaaaatg 3900

aagttttgat ttatttgtgt ttaaatgctg ctgtcagacg attgttctta gacctcctaa 3960

atgccccata ttaaaagaac tcattcatag gaaggtgttt cattttggtg tgcaaccctg 4020

tcattacgtc aacgcaacgt ctaactggac ttcccaagat aaatggtacc agcgtcctct 4080

taaaagatgc cttaatccat tccttgagga cagaccttag ttgaaatgat agcagaatgt 4140

gcttctctct ggcagctggc cttctgcttc tgagttgcac attaatcaga ttagcctgta 4200

ttctcttcag tgaattttga taatggcttc cagactcttt ggcgttggag acgcctgtta 4260

ggatcttcaa gtcccatcat agaaaattga aacacagagt tgttctgctg atagttttgg 4320

ggatacgtcc atctttttaa gggattgctt tcatctaatt ctggcaggac ctcaccaaaa 4380

gatccagcct catacctaca tcagacaaaa tatcgccgtt gttccttctg tactaaagta 4440

ttgtgttttg ctttggaaac acccactcac tttgcaatag ccgtgcaaga tgaatgcaga 4500

ttacactgat cttatgtgtt acaaaattgg agaaagtatt taataaaacc tgttaatttt 4560

tatactgaca ataaaaatgt ttctacagat attaatgtta acaagacaaa ataaatgtca 4620

cgcaacttat ttttttaata aaaaaaaaaa aaaa 4654

<210> 3

<211> 4304

<212> DNA

<213> Homo sapiens

<400> 3

gtcgcgggca gctggcgccg cgcggtcctg ctctgccggt cgcacggacg caccggcggg 60

ccgccggccg gagggacggg gcgggagctg ggcccgcgga cagcgagccg gagcgggagc 120

cgcgcgtagc gagccgggct ccggcgctcg ccagtctccc gagcggcgcc cgcctcccgc 180

cggtgcccgc gccgggccgt ggggggcagc atgcccgcgc gcgctgcctg aggacgccgc 240

ggcccccgcc cccgccatgg gcgcccctgc ctgcgccctc gcgctctgcg tggccgtggc 300

catcgtggcc ggcgcctcct cggagtcctt ggggacggag cagcgcgtcg tggggcgagc 360

ggcagaagtc ccgggcccag agcccggcca gcaggagcag ttggtcttcg gcagcgggga 420

tgctgtggag ctgagctgtc ccccgcccgg gggtggtccc atggggccca ctgtctgggt 480

caaggatggc acagggctgg tgccctcgga gcgtgtcctg gtggggcccc agcggctgca 540

ggtgctgaat gcctcccacg aggactccgg ggcctacagc tgccggcagc ggctcacgca 600

gcgcgtactg tgccacttca gtgtgcgggt gacagacgct ccatcctcgg gagatgacga 660

agacggggag gacgaggctg aggacacagg tgtggacaca ggggcccctt actggacacg 720

gcccgagcgg atggacaaga agctgctggc cgtgccggcc gccaacaccg tccgcttccg 780

ctgcccagcc gctggcaacc ccactccctc catctcctgg ctgaagaacg gcagggagtt 840

ccgcggcgag caccgcattg gaggcatcaa gctgcggcat cagcagtgga gcctggtcat 900

ggaaagcgtg gtgccctcgg accgcggcaa ctacacctgc gtcgtggaga acaagtttgg 960

cagcatccgg cagacgtaca cgctggacgt gctggagcgc tccccgcacc ggcccatcct 1020

gcaggcgggg ctgccggcca accagacggc ggtgctgggc agcgacgtgg agttccactg 1080

caaggtgtac agtgacgcac agccccacat ccagtggctc aagcacgtgg aggtgaatgg 1140

cagcaaggtg ggcccggacg gcacacccta cgttaccgtg ctcaagacgg cgggcgctaa 1200

caccaccgac aaggagctag aggttctctc cttgcacaac gtcacctttg aggacgccgg 1260

ggagtacacc tgcctggcgg gcaattctat tgggttttct catcactctg cgtggctggt 1320

ggtgctgcca gccgaggagg agctggtgga ggctgacgag gcgggcagtg tgtatgcagg 1380

catcctcagc tacggggtgg gcttcttcct gttcatcctg gtggtggcgg ctgtgacgct 1440

ctgccgcctg cgcagccccc ccaagaaagg cctgggctcc cccaccgtgc acaagatctc 1500

ccgcttcccg ctcaagcgac aggtgtccct ggagtccaac gcgtccatga gctccaacac 1560

accactggtg cgcatcgcaa ggctgtcctc aggggagggc cccacgctgg ccaatgtctc 1620

cgagctcgag ctgcctgccg accccaaatg ggagctgtct cgggcccggc tgaccctggg 1680

caagcccctt ggggagggct gcttcggcca ggtggtcatg gcggaggcca tcggcattga 1740

caaggaccgg gccgccaagc ctgtcaccgt agccgtgaag atgctgaaag acgatgccac 1800

tgacaaggac ctgtcggacc tggtgtctga gatggagatg atgaagatga tcgggaaaca 1860

caaaaacatc atcaacctgc tgggcgcctg cacgcagggc gggcccctgt acgtgctggt 1920

ggagtacgcg gccaagggta acctgcggga gtttctgcgg gcgcggcggc ccccgggcct 1980

ggactactcc ttcgacacct gcaagccgcc cgaggagcag ctcaccttca aggacctggt 2040

gtcctgtgcc taccaggtgg cccggggcat ggagtacttg gcctcccaga agtgcatcca 2100
cagggacctg gctgcccgca atgtgctggt gaccgaggac aacgtgatga agatcgcaga 2160
cttcgggctg gcccgggacg tgcacaacct cgactactac aagaagacaa ccaacggccg 2220
gctgcccgtg aagtggatgg cgcctgaggc cttgtttgac cgagtctaca ctcaccagag 2280
tgacgtctgg tcctttgggg tcctgctctg ggagatcttc acgctggggg gctccccgta 2340
ccccggcatc cctgtggagg agctcttcaa gctgctgaag gagggccacc gcatggacaa 2400
gcccgccaac tgcacacacg acctgtacat gatcatgcgg gagtgctggc atgccgcgcc 2460

ctcccagagg cccaccttca agcagctggt ggaggacctg gaccgtgtcc ttaccgtgac 2520
gtccaccgac gagtacctgg acctgtcggc gcctttcgag cagtactccc cgggtggcca 2580
ggacaccccc agctccagct cctcagggga cgactccgtg tttgcccacg acctgctgcc 2640
cccggcccca cccagcagtg ggggctcgcg gacgtgaagg gccactggtc cccaacaatg 2700
tgaggggtcc ctagcagccc accctgctgc tggtgcacag ccactccccg gcatgagact 2760
cagtgcagat ggagagacag ctacacagag ctttggtctg tgtgtgtgtg tgtgcgtgtg 2820
tgtgtgtgtg tgtgcacatc cgcgtgtgcc tgtgtgcgtg cgcatcttgc ctccaggtgc 2880

agaggtaccc tgggtgtccc cgctgctgtg caacggtctc ctgactggtg ctgcagcacc 2940
gaggggcctt tgttctgggg ggacccagtg cagaatgtaa gtgggcccac ccggtgggac 3000
ccccgtgggg cagggagctg ggcccgacat ggctccggcc tctgcctttg caccacggga 3060
catcacaggg tgggcctcgg cccctcccac acccaaagct gagcctgcag ggaagcccca 3120
catgtccagc accttgtgcc tggggtgtta gtggcaccgc ctccccacct ccaggctttc 3180
ccacttccca ccctgcccct cagagactga aattacgggt acctgaagat gggagccttt 3240
acctttttatg caaaaggttt attccggaaa ctagtgtaca tttctataaa tagatgctgt 3300

gtatatggta tatatacata tatatatata acatatatgg aagaggaaaa ggctggtaca 3360
acggaggcct gcgaccctgg gggcacagga ggcaggcatg gccctgggcg gggcgtgggg 3420
gggcgtggag ggaggcccca gggggtctca cccatgcaag cagaggacca gggccttttc 3480
tggcaccgca gttttgtttt aaaactggac ctgtatattt gtaaagctat ttatgggccc 3540
ctggcactct tgttcccaca ccccaacact tccagcattt agctggccac atggcggaga 3600
gttttaattt ttaacttatt gacaaccgag aaggtttatc ccgccgatag agggacggcc 3660
aagaatgtac gtccagcctg ccccggagct ggaggatccc ctccaagcct aaaaggttgt 3720

taatagttgg aggtgattcc agtgaagata ttttatttcc tttgtccttt ttcaggagaa 3780

ttagatttct ataggatttt tctttaggag atttatttttt tggacttcaa agcaagctgg 3840

tattttcata caaattcttc taattgctgt gtgtcccagg cagggagacg gtttccaggg 3900

aggggccggc cctgtgtgca ggttccgatg ttattagatg ttacaagttt atatatatct 3960

atatatataa tttattgagt ttttacaaga tgtatttgtt gtagacttaa cacttcttac 4020

gcaatgcttc tagagtttta tagcctggac tgctaccttt caaagcttgg agggaagccg 4080

tgaattcagt tggttcgttc tgtactgtta ctgggccctg agtctgggca gctgtccctt 4140

gcttgcctgc agggccatgg ctcagggtgg tctcttcttg gggcccagtg catggtggcc 4200

agaggtgtca cccaaaccgg caggtgcgat tttgttaacc cagcgacgaa ctttccgaaa 4260

aataaagaca cctggttgct aacctggaaa aaaaaaaaaa aaaa 4304