



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0032439
(43) 공개일자 2020년03월26일

(51) 국제특허분류(Int. Cl.)
C12Q 1/6883 (2018.01)
(52) CPC특허분류
C12Q 1/6883 (2018.05)
C12Q 2600/156 (2013.01)
(21) 출원번호 10-2018-0111472
(22) 출원일자 2018년09월18일
심사청구일자 2018년09월18일

(71) 출원인
재단법인 아산사회복지재단
서울특별시 송파구 올림픽로43길 88 (풍납동)
울산대학교 산학협력단
울산광역시 남구 대학로 93(무거동)
(72) 발명자
홍수중
서울특별시 강남구 압구정로29길 71, 24동 902호
(압구정동, 현대아파트)
(74) 대리인
특허법인 이룸리온

전체 청구항 수 : 총 8 항

(54) 발명의 명칭 아토피 피부염 진단을 위한 마커 조성물 및 이를 이용한 아토피 피부염 예측 또는 진단 방법

(57) 요약

본 발명은 PTGR2 유전자 내의 rs1968105#G 또는 rs9646165#A의 단일염기다형성(SNP)를 포함하는 연속된 DNA 서열로 구성된 아토피 피부염 진단을 위한 마커 조성물을 제공한다. 본 발명은 다양한 원인에 의해 발생하는 아토피 피부염 관련 유전자의 단일염기다형성 및 해플로타입들의 연관성 분석을 통해 유전적 특성을 확인하였으며, 이를 통해 발굴한 마커 조성물은 아토피 피부염을 구분할 수 있는 기능을 보이며, 이 기능을 이용하여, 아토피 피부염에 대한 예측 또는 진단 조성물로서 유용하게 활용될 수 있다.

대표도

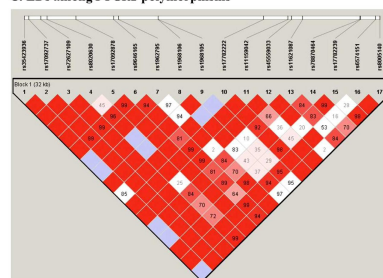
A. Map of *PTGR2* (prostaglandin reductase 2) on chromosome 14q24.3 (35 kb) NM_001146154.1



B. Haplotypes in *PTGR2*

Haplotype block																		
ht	rs35423936	rs17092737	rs72627109	rs8020630	rs17092878	rs9646165	rs1962795	rs1968106	rs1968105	rs17782222	rs11259042	rs4559003	rs11621087	rs78870464	rs17782239	rs6574151	rs80051440	Freq.
ht1	C	T	C	G	A	G	T	T	T	T	A	C	A	C	G	G	A	0.426
ht2	C	C	T	G	A	G	T	T	T	A	G	C	A	C	A	G	A	0.269
ht3	C	C	T	A	G	A	T	T	T	T	A	C	A	C	T	G	A	0.064
ht4	G	C	T	G	G	G	T	T	T	T	A	C	A	C	G	A	C	0.062
Others																		0.179

C. LDs among *PTGR2* polymorphisms



(52) CPC특허분류

C12Q 2600/158 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 1711072976

부처명 과학기술정보통신부

연구관리전문기관 한국연구재단

연구사업명 바이오.의료기술개발

연구과제명 아토피피부염 질환군의 면역기전연구를 통한 장내 마이크로바이옴 기반 파마바이오틱스 발
굴

기 여 율 1/1

주관기관 울산대학교 산학협력단

연구기간 2018.03.01 ~ 2018.12.31

명세서

청구범위

청구항 1

서열번호 1의 PTGR2 유전자의 염기서열에 있어서, rs1968105#G 또는 rs9646165#A의 단일염기다형성(SNP)를 포함하는 20 내지 100 개의 연속된 DNA 서열로 구성되는 아토피 피부염 진단을 위한 마커 조성물.

청구항 2

서열번호 1의 PTGR2 유전자의 염기서열에 있어서, rs1968105#G 또는 rs9646165#A의 단일염기다형성(SNP)를 검출 또는 증폭할 수 있는 제제를 포함하는 아토피 피부염 예측 또는 진단용 조성물.

청구항 3

제1항의 조성물을 포함하는 아토피 피부염 예측 또는 진단용 키트.

청구항 4

아토피 피부염 질환의 진단에 필요한 정보를 제공하기 위하여,

서열번호 1의 PTGR2 유전자의 염기서열에 있어서, rs1968105#G 또는 rs9646165#A의 단일염기다형성(SNP)을 증폭하는 단계;를 포함하는 아토피 피부염에 대한 유전자 PTGR2의 유전자형 분석방법.

청구항 5

서열번호 1의 PTGR2 유전자의 염기서열에 있어서, 73849885 번째 염기가 C이고, 73850297 번째 염기가 C이며, 73851433 번째 염기가 T이고, 73851970 번째 염기가 G이며, 73854619 번째 염기가 G이고, 73858104 번째 염기가 G이고, 73858548 번째 염기가 G이고, 73862199 번째 염기가 G이고, 73862202 번째 염기가 G이고, 73871846 번째 염기가 T이고, 73874094 번째 염기가 A이고, 73874421 번째 염기가 G이고, 73875453 번째 염기가 C이고, 73880710 번째 염기가 A이고, 73880736 번째 염기가 C이고, 73882116 번째 염기가 A이고, 73882524 번째 염기가 G인 해플로타입을 구성하는 SNP 증폭을 위한 20 내지 100 개의 연속된 DNA 서열로 구성되는 아토피 피부염 진단을 위한 마커 조성물.

청구항 6

서열번호 1의 PTGR2 유전자의 염기서열에 있어서, 73849885 번째 염기가 C이고, 73850297 번째 염기가 C이며, 73851433 번째 염기가 T이고, 73851970 번째 염기가 G이며, 73854619 번째 염기가 G이고, 73858104 번째 염기가 G이고, 73858548 번째 염기가 G이고, 73862199 번째 염기가 G이고, 73862202 번째 염기가 G이고, 73871846 번째 염기가 T이고, 73874094 번째 염기가 A이고, 73874421 번째 염기가 G이고, 73875453 번째 염기가 C이고, 73880710 번째 염기가 A이고, 73880736 번째 염기가 C이고, 73882116 번째 염기가 A이고, 73882524 번째 염기가 G인 해플로타입을 구성하는 SNP 증폭을 위한 20 내지 100 개의 연속된 DNA 서열을 검출 또는 증폭할 수 있는 제제를 포함하는 아토피 피부염 예측 또는 진단용 조성물.

청구항 7

제6항에 따른 조성물을 포함하는 아토피 피부염 예측 또는 진단용 키트.

청구항 8

아토피 피부염 질환의 진단에 필요한 정보를 제공하기 위하여,

서열번호 1의 PTGR2 유전자의 염기서열에 있어서, 73849885 번째 염기가 C이고, 73850297 번째 염기가 C이며, 73851433 번째 염기가 T이고, 73851970 번째 염기가 G이며, 73854619 번째 염기가 G이고, 73858104 번째 염기가 G이고, 73858548 번째 염기가 G이고, 73862199 번째 염기가 G이고, 73862202 번째 염기가 G이고, 73871846 번째 염기가 T이고, 73874094 번째 염기가 A이고, 73874421 번째 염기가 G이고, 73875453 번째 염기가 C이고, 73880710 번째 염기가 A이고, 73880736 번째 염기가 C이고, 73882116 번째 염기가 A이고, 73882524 번째 염기가

가 G인 해플로타입을 구성하는 SNP 증폭하는 단계;를 포함하는 아토피 피부염에 대한 유전자 PTGR2의 유전자형 분석방법.

발명의 설명

기술 분야

[0001] 본 발명은 PTGR2 유전자 내의 아토피 피부염 진단을 위한 마커 조성물 및 이를 이용한 아토피 피부염 예측 또는 진단 방법에 관한 것으로서, 보다 상세하게는 본 발명의 마커 조성물은 PTGR2 다형성이 아토피 피부염과 연관되어 있음을 확인함으로써, 이의 진단용도로 사용될 수 있다.

배경 기술

[0002] 아토피 피부염(AD, Atopic Dermatitis)은 정상인에게는 해가 되지 않은 유발인자들에 피부가 과민반응상태를 보임으로써 만성적으로 피부에 염증이 생기는 만성 염증성 피부질환이다. 아토피 피부염은 세균, 바이러스, 곰팡이 등에 의한 감염, 꽃가루 등과 같은 특정 물질, 음식물, 인공적 화합물질 등에 의하여 면역체계가 과민반응하여 발생하는 것으로 알려져있으나, 정확한 원인은 확실하게 밝혀진 바 없다.

[0003] 현재 아토피 피부염은 전세계적으로 가장 흔한 피부 질환으로, 전세계 20% 어린이와 3% 성인이 고통받고 있는 것으로 추정되고 있으며, 점차 그 발병율이 상승하고 있다.

[0004] 아토피 피부염 환자들은 일반적으로 외인성 아토피 피부염 및 내인성 아토피 피부염의 두 그룹으로 분류된다. 외인성 아토피 피부염(extrinsic atopic dermatitis, EAD) 환자들은 혈청(serum)의 전체 이뮤노글로불린(immunoglobulin) E (IgE) 및 알레르겐(allergen)-특이적 IgE 레벨이 증가한다. 알레르겐-특이적 IgE가 없는 정상 혈청의 전체 IgE 레벨을 가진 환자들은 내인성 아토피 피부염(intrinsic atopic dermatitis, IAD) 또는 아토피양 습진(atopiform dermatitis)이라 한다.

[0005] 현재까지 아토피피부염에서 가장 흔하게 측정되는 바이오마커는 IgE로, 그 결과를 바탕으로 아토피피부염의 표현형을 내인성과 외인성으로 분류해 왔다. 하지만 중증 아토피피부염 환자의 총 IgE 수치가 상승하지 않은 경우가 많고 총 IgE가 아토피피부염의 중증도와 상관성이 높지 않아, 객관적인 판단이 어렵고, 병변의 경증을 판단하는데 많은 시간이 요구되며 개인별로 다르게 판단할 위험성이 있었다.

[0006] 상술한 문제점을 해결하기 위해, 아토피 진단과 관련된 다양한 방법들이 개발되었다. 특히, 환자와 건강한 대조군에서의 T 림프구 역할 연구를 통해, 혈액 내 갈렉틴(galectin)-10(특허문헌 1), 호산구 양이온 단백질(eosinophil cationic protein, ECP), 인터류킨(interleukin, IL)-2R, IL-18, IL-31, 흥선 및 활성화-조절 케모카인(thymus and activation-regulated chemokine)(TARC/CCL17), 젖산 탈수소효소(lactate dehydrogenase, LDH)와 같은 새로운 바이오마커들이 아토피 피부염의 중증도를 추적 관찰할 수 있을 것으로 보고되었다. 그러나 이들은 병태 생리의 다양성을 제대로 반영하지 못하고, 아토피 피부염을 효과적으로 진단하는데 많은 시간이 요구되거나, 정확도 혹은 신뢰도가 떨어지는 문제점이 있었다.

[0007] 따라서, 아토피 피부염을 진단 및 검진하는 것은 본 기술 분야에서 매우 중요한 기술적 과제이나, 아토피 피부염에 관여하는 병인학적 소인은 매우 광범위하기 때문에 기존에 사용해오던 바이오마커 발굴전략으로는 한계가 있다.

[0008] 상기한 배경기술로서 설명된 사항들은 본 발명의 배경에 대한 이해 증진을 위한 것일 뿐, 이 기술분야에서 통상의 지식을 가진 자에게 이미 알려진 종래기술에 해당함을 인정하는 것으로 받아들여져서는 안 될 것이다.

선행기술문헌

특허문헌

[0009] (특허문헌 0001) 특허문헌 1. 대한민국 등록특허공보 제10-1512121호

발명의 내용

해결하려는 과제

- [0010] 본 발명자들은 아토피 피부염을 진단하는 새로운 방법의 개발에 유용한 PTGR2 다형성을 발굴하고자 예의 노력을 하였다. 그 결과, PTGR2 유전자의 단일염기다형성 또는 해플로타입이 아토피 피부염과 연관성을 가지고 있음을 분석하는 전략을 고안하였고, 이를 이용하여 아토피 피부염을 예측 또는 진단함으로써 본 발명을 완성하였다.
- [0011] 따라서, 본 발명의 목적은 PTGR2 유전자의 단일염기다형성 또는 해플로타입으로 구성되는 아토피 피부염 진단을 위한 마커 조성물 및 이를 이용하여 아토피 피부염 예측 또는 진단방법을 제공하는데 있다.
- [0012] 본 발명의 다른 목적은 PTGR2 유전자의 단일염기다형성 또는 해플로타입을 검출 또는 증폭할 수 있는 체제를 포함하는 진단용 조성물 또는 이를 포함하는 키트를 제공하는데 있다.
- [0013] 본 발명의 다른 목적 및 이점은 하기의 발명의 상세한 설명, 청구범위 및 도면에 의해 보다 명확하게 된다.

과제의 해결 수단

- [0014] 본 발명의 일 양태에 따르면, 본 발명은 아토피 피부염을 예측 또는 진단할 수 있는 단일염기다형성(single nucleotide polymorphism, SNP)과 해플로타입(haplotype)으로 구성된 마커 조성물을 제공한다.
- [0015] 본 발명자들은 장내 미생물이 면역발달과 면역 반응과 관련이 있으며, 알레르기 질환의 발병에 중요한 역할을 하고 있음을 통해, 이로부터 아토피 피부염을 예측 또는 진단할 수 있는 신규 마커를 발굴하고자 노력한 결과, 아토피 피부염 환자와 정상 대조군과의 회귀 분석을 이용하여 PTGR2 유전자의 다형성과 해플로타입이 아토피 피부염과 유의한 연관성을 가지고 있음을 확인하였다.
- [0016] 구체적으로 대변 샘플에서 대장 세포의 유전체수준의 유전자 발현을 프로파일링한 결과, PTGR2가 아토피 피부염의 잠재적 위험인자임을 확인하였고, PTGR2 유전자 내의 서열을 분석하고, 한국인 아토피 피부염 환자와 정상 대조군에서 아토피 피부염 표현형과 PTGR2의 연관성을 분석하였다. 한국인 군집(아토피 피부염 환자 420명, 정상 대조군 444명)에서 17개의 단일염기다형성을 직접적인 분석을 통해 유전형을 확인하였다. PTGR2 다형성들과의 연관성을 연령, 성별을 조정하면서 로지스틱과 선형 회귀를 사용하여 분석하였고, 그 결과 PTGR2 유전자 내 유전적 다형성 사이에 현저한 연관성이 있음을 검출하였다. 추가적인 해플로타입 분석을 살펴보면 ht2 해플로타입이 관련성을 가지고 있음을 확인하였다(haplotype ht2에서 $P=0.026$). 특히, 2 개의 SNP(rs1968105#G 또는 rs9646165#A)는 아토피 피부염 발병과 상당한 유의성을 가지고 있음을 확인하였다. 또한 PTGR2의 해플로타입 중에서 ht2(해플로타입2#)는 아토피 피부염의 발병과 현저한 연관성을 가지는 것으로 확인되었다. 또한 선형 회귀 분석에서 상기 rs9646165 SNP와 해플로타입 ht2는 SCORAD 값에 유의한 영향을 미치고 있는 것으로 확인되었다. 즉 PTGR2 유전자의 다형성과 아토피 피부염의 강력한 연관성을 통해, 개체로부터 유전적 역학 분석으로, 아토피 피부염 질환을 정확히 예측 또는 진단하는데 유용하게 활용할 수 있다.
- [0017] 상기 마커 조성물은 바람직하게는 PTGR2 유전자의 SNP를 포함하는 전체 또는 일부 폴리뉴클레오타이드일 수 있고, 보다 바람직하게는 서열번호 1로 구성된 폴리뉴클레오타이드의 전체 또는 일부 폴리뉴클레오타이드일 수 있으며, 가장 바람직하게는 서열번호 1의 PTGR2 유전자의 염기서열에 있어서, rs1968105#G 또는 rs9646165#A의 단일염기다형성(SNP)를 포함하는 20 내지 100 개의 연속된 DNA 서열로 구성되는, 아토피 피부염 진단용 마커 조성물일 수 있다.
- [0018] 본 명세서에서 용어, "뉴클레오타이드"는 단일가닥 또는 이중가닥 형태로 존재하는 디옥시리보뉴클레오타이드 또는 리보뉴클레오타이드이며, 다르게 특별하게 언급되어 있지 않은 한 자연의 뉴클레오타이드의 유사체를 포함한다(Scheit, Nucleotide Analogs, John Wiley, New York(1980); Uhlman 및 Peyman, Chemical Reviews, 90:543-584(1990)).
- [0019] 본 발명은 PTGR2 유전자의 다형성을 이용해, 아토피 피부염 예측 또는 진단 마커로서의 용도에 관한 것으로, PTGR2 유전자의 특정 SNP가 아토피 피부염과 연관성이 있음을 처음으로 규명하였다.
- [0020] 보다 구체적으로 서열번호 1의 핵산서열로 구성된 DNA에 있어서, 73862202 번째 G 염기를 포함하거나(rs1968105), 73858104 번째 G 염기(본 발명에서 표시된 minor A의 major 염기)를 포함하는(rs9646165) 20~100 개의 연속서열로 구성되는 폴리뉴클레오타이드, 아토피 피부염을 진단할 수 있는 마커 조성물일 수 있다. 본 발명의 서열번호 1의 핵산서열은 Homo sapiens prostaglandin reductase 2 (PTGR2) sequence 중에서 hg38_ncbiRefSeqCurated_NM_001146154.1의 chr14에 대한 73849831 번째부터 73885465 번째까지의 서열이고, 본 원발명에서 서열번호 1의 핵산서열에서의 위치는 이를 기준으로 나타내었다.
- [0021] 본 발명에서 PTGR2 유전자란 prostaglandin reductase 2로, 프로스타글란딘(prostaglandins) 촉매 작용에 관여

하는 효소이다. 상기 PTGR2 유전자의 염기서열은 NCBI의 GenBank 및 UCSC의 Genome Browser 등 공지의 데이터베이스에서 얻을 수 있는데, 그 예로서 GenBank Accession No. NM_001146154.1로 표시되는 유전자가 될 수 있으며, 바람직하게는 서열번호 1의 폴리뉴클레오타이드를 포함할 수 있다.

- [0022] 상기 마커 조성물은 상기 SNP 각각으로 아토피 피부염 예측 또는 진단을 위한 정보를 제공할 수 있을 뿐만 아니라, 상기 SNP의 haplotype 조합으로 더욱 우수한 마커로서 작용할 수도 있다.
- [0023] "유전적 다형성(genetic polymorphism)"은 하나의 유전자 좌위(locus)에 두가지 이상의 대립 유전자(allele)가 존재하는 경우를 말하며, 인구집단에서 적어도 1% 이상의 빈도로 유전자 변이가 나타나는 경우를 말한다. DNA에서 한 개의 뉴클레오타이드의 삽입, 소실, 또는 치환이 일어나는 것을 단일염기다형성(single nucleotide polymorphism, SNP) 이라고 한다.
- [0024] "다형성"이라는 용어는 군집 내에서 변하는 유전자의 서열에서의 배치를 지칭한다. 다형성은 상이한 "대립유전자"로 구성된다. 이러한 다형성의 배치는 유전자에서의 그의 위치 및 그에서 발견되는 상이한 아미노산 또는 염기에 의해 확인될 수 있다. 이러한 아미노산 변이는 2개의 상이한 대립유전자인, 2개의 가능한 변이체 염기, C 및 T의 결과이다. 유전자형은 2개의 다른 별개의 대립유전자로 구성되기 때문에, 여러 가능한 변이체 중 임의의 변이체가 어느 한 개체에서 관찰될 수 있다 (예를 들어, 이 예에서, CC, CT 또는 TT). 개개의 다형성은 또한 당업자에게 공지되어 있고, 예를 들면, NCBI 웹사이트 상에서 이용가능한 뉴클레오타이드 염기 변이의 단일 뉴클레오타이드 다형성 데이터베이스(Single Nucleotide Polymorphism Database (dbSNP) of Nucleotide Sequence Variation)에서 사용되는 것인, 지정된 독특한 식별자 ("기준 SNP", "refSNP" 또는 "rs#")이다.
- [0025] "단일염기다형성(single nucleotide polymorphism, SNP)"은 계통에서 단일염기(A, T, C 또는 G)가 종의 멤버들 간 또는 한 개체(individual)의 쌍 염색체 간에 다른 경우에 발생하는 DNA 서열의 다양성을 의미한다. 예를 들어, 서로 다른 개체의 세 개의 DNA 단편들(예: AAGT[A/A]AG, AAGT[A/G]AG, AAGT[G/G]AG)처럼 단일염기에서 차이를 포함하는 경우, 두 개의 대립 유전자(C 또는 T)라고 부르며, 일반적으로 거의 모든 SNPs는 두 개의 대립 유전자를 가진다. 한 집단(population)내에서, SNP는 소수 대립인자 빈도(minor allele frequency, MAF; 특정 집단에서 발견되는 유전자위치(locus)에서 가장 낮은 대립인자 빈도)로 할당될 수 있다. 단일염기는 폴리뉴클레오타이드 서열에 변화(대체), 제거(결실) 또는 첨가(삽입)될 수 있다. SNP는 번역 프레임의 변화를 유발할 수 있다.
- [0026] "유전자형(genotype)"이라는 용어는 세포 또는 조직 샘플에서 특정 유전자의 특이적 대립유전자를 지칭한다.
- [0027] "대립 인자" 또는 '대립 유전자(allele)'란 같은 염색체 위치(same chromosomal locus)를 점유하는 한 유전자의 둘 또는 그 이상의 선택적인 형태(alternative forms) 중 하나를 뜻한다. 대립유전자는 다형성을 나타내는데 사용되기도 하며, 예를 들어 SNP는 두 종류의 대립인자(biallele)를 갖는다.
- [0028] "대립유전자 빈도"는 대립유전자가 개체 내, 계통 내, 또는 계통의 집단 내에 존재하는 빈도(비율 또는 백분율)을 지칭한다. 계통 또는 집단 내에서의 대립유전자 빈도를, 계통 또는 집단으로부터의 개체들의 샘플의 대립유전자 빈도를 평균화함으로써 추정할 수 있다.
- [0029] "진단"은 병리 상태의 존재 또는 특징을 확인하는 것을 의미한다. 그 중에서도 본 발명은 특히 아토피 피부염(AD) 진단에 유용하다. 본 발명에서는 아토피 피부염 진행, 발병을 예측하는 것을 포함한다. 이에, 아토피 피부염의 진행과 발병을 정확히 예측할 수 있는 지표가 매우 중요하며 임상적 지표를 보완하면서 치료의 반응을 예측할 수 있는 인자가 필요한데, 본 발명의 PTGR2의 SNP가 이러한 지표 기능을 하므로 아토피 피부염 진단 인자로 이용할 수 있다. 즉, 이러한 유전자들의 다형성 특성 측정은 아토피 피부염을 예측 또는 진단하는데 유용한 지표(진단 마커)로 사용될 수 있다.
- [0030] "진단용 마커 또는 진단 마커(diagnosis marker)란 아토피 피부염을 가진 세포를 정상 세포와 구분하여 진단할 수 있는 물질로, 정상 세포에 비하여 아토피 피부염이 발병한 세포에서 증가양상을 보이는 폴리펩타이드 또는 핵산(예: mRNA 등), 지질, 당지질, 당단백질, 당(단당류, 이당류, 다당류 등) 등과 같은 유기 생체 분자 등을 포함한다.
- [0031] "아토피 피부염 예측용 마커"란 아토피 피부염 발병 위험성, 발병한 아토피 피부염의 완치 여부, 혹은 경과를 예측할 수 있는 다형성을 가진 마커를 의미하며, 바람직하게는 상기에서 서술한 뉴클레오타이드를 의미한다. 또한, 상기 환자는 아토피 피부염 발병 위험성을 판별하기 위한 환자 또는 가족병력상에 아토피 피부염이 존재하는 환자를 의미한다. "위험"은 특정 다형성 대립유전자를 보유하지 않은 개체의 구성원에서의 질환 또는 상태의 발생 빈도에 비교하여, 특정 다형성 대립유전자를 보유한 개체에서의 질환 또는 상태의 발생 빈도가 통계적으로

높음을 나타낸다.

- [0032] "대상" 또는 "환자"는 인간, 소, 개, 기니아 피그, 토끼, 닭, 파충류, 어류, 양서류, 곤충 등을 포함하여 치료가 요구되는 임의의 단일 개체를 의미한다. 또한, 임의의 질병 임상 소견을 보이지 않는 임상 연구 시험에 참여한 임의의 대상 또는 역학연구에 참여한 대상 또는 대조군으로 사용된 대상이 대상에 포함된다.
- [0033] "조직 또는 세포 샘플"은 대상 또는 환자의 조직으로부터 얻은 유사한 세포의 집합체를 의미한다. 조직 또는 세포 샘플의 공급원은 신선한, 동결된 및/또는 보존된 장기 또는 조직 샘플 또는 생검 또는 흡인물로부터의 고정 조직; 혈액 또는 임의의 혈액 구성분; 대상의 임신 또는 발생의 임의의 시점의 세포일 수 있다. 조직 샘플은 또한 1차 또는 배양 세포 또는 세포주일 수 있다.
- [0034] "핵산"은 임의의 DNA 또는 RNA, 예를 들어, 조직 샘플에 존재하는 염색체, 미토콘드리아, 바이러스 및/또는 세균 핵산을 포함하는 의미이다. 이중가닥 핵산 분자의 하나 또는 두개 모두의 가닥을 포함하고, 무손상 핵산 분자의 임의의 단편 또는 일부를 포함한다.
- [0035] "유전자"는 단백질 코딩 또는 전사시에 또는 다른 유전자 발현의 조절시에 기능적 역할을 갖는 임의의 핵산 서열 또는 그의 일부를 의미한다. 유전자는 기능적 단백질을 코딩하는 모든 핵산 또는 단백질을 코딩 또는 발현하는 핵산의 일부만으로 이루어질 수 있다. 핵산 서열은 엑손, 인트론, 개시 또는 종료 영역, 프로모터 서열, 다른 조절 서열 또는 유전자에 인접한 특유한 서열 내에 유전자 이상을 포함할 수 있다.
- [0036] 본 발명은 PTGR2 유전자 관련하여 아토피 피부염 진단용 마커를 개발하였고, 일 실시예에 따르면, 본 발명에 따른 마커 조성물은 아토피 피부염과 밀접한 연관성이 있음을 확인하였다.
- [0037] 따라서, 본 발명의 마커 조성물의 유전자형을 검출하거나 또는 상기 유전자형을 검출할 수 있는 프라이머를 이용하면, 아토피 피부염을 갖거나 가지고 있을 것으로 예상되는 개체로부터 아토피 피부염을 정확하게 예측 또는 진단할 수 있을 것으로 기대되며, 상기 아토피 피부염 및 임상지표와 연관성이 확인된 SNP는 본 발명자들에 의해 최초로 규명된 것이다.
- [0038] 본 발명의 다른 측면은 서열번호 1의 PTGR2 유전자의 염기서열에 있어서, rs1968105#G 또는 rs9646165#A의 단일 염기다형성(SNP)를 검출 또는 증폭할 수 있는 체제를 포함하는 아토피 피부염 예측 또는 진단용 조성물에 관한 것이다.
- [0039] 본 발명의 용어 "SNP 마커를 검출 또는 증폭할 수 있는 체제"란, 상기와 같은 유전자의 단일염기다형성(SNP) 부위를 증폭을 통해 확인하여 아토피 피부염을 판단할 수 있는 조성물을 의미하며, 바람직하게는 상기 SNP 마커의 폴리뉴클레오타이드를 특이적으로 증폭할 수 있는 프라이머를 의미한다. 상기 SNP 마커 증폭에 사용되는 프라이머는, 적절한 버퍼 중의 적절한 조건(예를 들면, 4개의 다른 뉴클레오타이드 트리포스페이트 및 DNA, RNA 폴리머라제 또는 역전사 효소와 같은 중합제) 및 적당한 온도 하에서 주형-지시 DNA 합성(template-directed DNA synthesis)의 시작점으로서 작용할 수 있는 단일가닥 올리고뉴클레오타이드가 될 수 있는데, 상기 프라이머의 적절한 길이는 사용 목적에 따라 달라질 수 있으나, 통상 15 내지 30 뉴클레오타이드이다. 짧은 프라이머 분자는 일반적으로 주형과 안정한 혼성체를 형성하기 위해서는 더 낮은 온도를 필요로 한다. 상기 프라이머 서열은 상기 SNP 마커와 완전하게 상보적일 필요는 없으나, 상기 SNP 마커와 혼성화 할 정도로 충분히 상보적이어야 한다.
- [0040] 본 발명의 용어 "프라이머"란, 짧은 자유 3' 말단 수산화기(free 3' hydroxyl group)를 가지는 염기 서열로 상보적인 주형(template)과 염기쌍(base pair)을 형성할 수 있고 주형 가닥 복사를 위한 시작 지점으로 기능을 하는 짧은 서열을 의미한다. 프라이머는 적절한 완충용액 및 온도에서 중합반응(즉, DNA 폴리머레이즈 또는 역전사효소)을 위한 시약 및 상이한 4가지 뉴클레오타이드 트리포스페이트의 존재하에서 DNA 합성을 개시할 수 있다. PCR 증폭을 실시하여 원하는 생성물의 생성 여부를 통해 아토피 피부염을 예측할 수 있다. PCR 조건, 센스 및 안티센스 프라이머의 길이는 당업계에 공지된 것을 기초로 변형할 수 있다.
- [0041] 본 발명의 프라이머는 포스포로아미다이트(phosphoramidite) 고체 지지체 방법, 또는 기타 널리 공지된 방법을 사용하여 화학적으로 합성할 수 있다. 이러한 핵산 서열은 또한 당해 분야에 공지된 많은 수단을 이용하여 변형시킬 수 있다. 이러한 변형의 비-제한적인 예로는 메틸화, "캡화", 천연 뉴클레오타이드 하나 이상의 동족체로의 치환, 및 뉴클레오타이드 간의 변형, 예를 들면, 하전되지 않은 연결체(예: 메틸 포스포네이트, 포스포트리에스테르, 포스포로아미데이트, 카바메이트 등) 또는 하전된 연결체(예: 포스포로티오에이트, 포스포로디티오에이트 등)로의 변형이 있다.

- [0042] 또한, 본 발명의 또 다른 측면은 서열번호 1의 PTGR2 유전자의 염기서열에 있어서, rs1968105#G 또는 rs9646165#A의 단일염기다형성(SNP)를 검출 또는 증폭할 수 있는 체제를 포함하는 아토피 피부염 예측 또는 진단용 키트에 관한 것이다.
- [0043] 본 발명의 키트는 아토피 피부염을 진단할 수 있는 PTGR2 유전자의 SNP를 증폭을 통해 확인하거나, SNP 마커의 발현 수준을 확인함으로써, 아토피 피부염을 예측 또는 진단할 수 있게하는 것이다.
- [0044] 본 발명에 있어서, 상기 키트를 사용하고자 하는 대상은 한국인, 일본인 또는 중국인을 대상으로 하는 것이 더욱 바람직하나 이에 한정되지 않는다.
- [0045] 상기 진단용 키트에는 본 발명의 PTGR2 유전자의 SNP(rs1968105#G 또는 rs9646165#A)를 증폭시킬 수 있는 프라이머 뿐만 아니라, 중합 반응에 필요한 시약, 예를 들면 dNTP, 각종의 중합효소 및 발색제 등을 포함할 수 있다.
- [0046] 상기 "증폭시킬 수 있는 프라이머"란 적절한 버퍼 중의 적절한 조건(예를 들면, 4개의 다른 뉴클레오타이드 트리포스페이트 및 DNA, RNA 폴리머라제 또는 역전사 효소와 같은 중합제) 및 적당한 온도 하에서 주형-지시 DNA 합성의 시작점으로서 작용할 수 있는 단일가닥 올리고뉴클레오타이드를 말한다. 상기 프라이머의 적절한 길이는 사용 목적에 따라 달라질 수 있으나, 통상 15 내지 30 뉴클레오타이드이다. 짧은 프라이머 분자는 일반적으로 주형과 안정한 혼성체를 형성하기 위해서는 더 낮은 온도를 필요로 한다. 프라이머 서열은 주형과 완전하게 상보적일 필요는 없으나, 주형과 혼성화 할 정도로 충분히 상보적이어야 한다.
- [0047] 본 발명에서 아토피 피부염 예측 또는 진단용 조성물의 mRNA 발현 수준을 측정하기 위한 키트는 RT-PCR을 수행하기 위해 필요한 필수 요소를 포함하는 키트일 수 있다. RT-PCR 키트는, 상기 SNP 마커의 유전자에 대한 특이적인 각각의 프라이머 쌍 외에도 RT-PCR 키트는 테스트 튜브 또는 다른 적절한 컨테이너, 반응 완충액(pH 및 마그네슘 농도는 다양), 데옥시뉴클레오타이드(dNTPs), Taq-폴리머라아제 및 역전사효소와 같은 효소, DNase, RNase 억제제, DEPC-수(DEPCwater), 멸균수 등을 포함할 수 있다. 또한 정량 대조군으로 사용되는 유전자에 특이적인 프라이머 쌍을 포함할 수 있다. 또한 바람직하게는, 본 발명의 키트는 DNA 칩을 수행하기 위해 필요한 필수 요소를 포함하는 아토피 피부염 예측 또는 진단용 키트일 수 있다. DNA 칩 키트는, 일반적으로 편평한 고체 지지판, 전형적으로는 현미경용 슬라이드보다 크지않은 유리 표면에 핵산 중을 격자형 배열(grided array)로 부착한 것으로, 칩 표면에 핵산이 일정하게 배열되어, DNA 칩 상의 핵산과 칩 표면에 처리된 용액 내에 포함된 상보적인 핵산 간에 다중 혼성화(hybridization) 반응이 일어나 대량 병렬 분석이 가능하도록 하는 도구이다.
- [0048] 본 발명의 또 다른 측면은 아토피 피부염 질환의 진단에 필요한 정보를 제공하기 위하여, 서열번호 1의 PTGR2 유전자의 염기서열에 있어서, rs1968105#G 또는 rs9646165#A의 단일염기다형성(SNP)을 증폭하는 단계를 포함하는 아토피 피부염에 대한 유전자 PTGR2의 유전자형 분석방법에 관한 것이다.
- [0049] 구체적으로 피검자로부터 체외로 분리된 샘플을 취득하고, 상기 샘플에서 서열번호 1의 PTGR2 유전자의 염기서열에 있어서, rs1968105#G 또는 rs9646165#A의 단일염기다형성(SNP)을 증폭한 후, 아토피 피부염에 대한 유전자 PTGR2의 유전자형을 분석함으로써, 아토피 피부염 예측 또는 진단에 필요한 정보를 제공할 수 있다.
- [0050] 우선, 피검자로부터 체외로 분리된 샘플은 특별히 한정된 것은 아니며, 일례로, 털, 뇨, 타액, 각종 체액, 근육, 표피, 혈액, 뼈, 장기로부터 얻은 것일 수 있고, 바람직하게는 근육 또는 혈액일 수 있다.
- [0051] 상기 SNP 유전자형을 분석하기 위해, 상기 분리된 샘플로부터 DNA를 추출할 수 있다. gDNA의 분리는 당업계에 공지된 통상의 방법에 따라 실시될 수 있으며(참조: Rogers & Bendich (1994)), 출발물질이 mRNA인 경우에는, 역전사효소를 이용하여 cDNA로 합성될 수 있다.
- [0052] 다음으로, 상기 샘플로부터, 서열번호 1의 PTGR2 유전자의 염기서열에서의 rs1968105#G 또는 rs9646165#A의 단일염기다형성(SNP) 부위를 증폭하거나 프로브와 혼성화할 수 있는데, 이는 당업자에게 알려진 어떠한 방법이든 사용가능하다. 예를 들면, 표적 핵산을 PCR을 통하여 증폭하고 이를 정제하여 얻을 수 있다. 그 외 리가제 연쇄 반응(LCR)(Wu 및 Wallace, Genomics 4, 560(1989), Landegren 등, Science 241, 1077(1988)), 전사증폭(transcription amplification)(Kwoh 등, Proc. Natl.Acad. Sci. USA 86, 1173(1989)) 및 자가유지 서열 복제(Guatelli 등, Proc. Natl. Acad. Sci. USA 87, 1874(1990)) 및 핵산서열 기반 증폭(nucleic acid sequence-based amplification, NASBA)이 사용될 수 있다.
- [0053] 이후, 상기 산물로부터 유전자형을 확인함으로써, 아토피 피부염에 대한 유전자 PTGR2의 유전자형 분석방법을

제공하는 것을 특징으로 할 수 있다. 상기 산물로부터 유전자형을 분석하기 위해서는 당업자에게 알려진 어떠한 방법이든 사용가능하나, 예를 들어 서열 분석, 마이크로어레이(microarray)에 의한 혼성화, 대립유전자 특이적인 PCR(allele specific PCR), 동적 대립유전자-특이적 혼성화 기법(dynamic allele-specific hybridization, DASH), PCR 연장 분석, PCR-SSCP(Polymerase Chain Reaction-Single Strand Conformation Polymorphism, 중합효소 연쇄 반응-단일가닥 구조 다형성), PCR-RFLP(Polymerase Chain Reaction-Restriction Fragment Length Polymorphism, 중합효소 연쇄 반응-제한효소 단편 길이 다형성) 분석 또는 TaqMan 기법, SNPlex 플랫폼(Applied Biosystems), 질량 분석법(예를 들면, Sequenom의 MassARRAY 시스템), 미니-시퀀싱(mini-sequencing) 방법, Bio-Plex 시스템(BioRad), CEQ and SNPstream 시스템(Beckman), Molecular Inversion Probe 어레이 기술(예를 들면, Affymetrix GeneChip), Fluidigm 시스템 및 Bead Array Technologies(예를 들면, Illumina GoldenGate 및 Infinium 분석법)를 포함하나, 이에 한정되지 않는다. 상기 방법들 또는 본 발명이 속하는 기술분야의 당업자에게 이용 가능한 다른 방법에 의해, 미소부수체(microsatellite), SNP 또는 다른 종류의 다형성 마커를 포함한, 다형성 마커에서의 하나 이상의 대립유전자가 확인될 수 있다. 이와 같은 다형성 부위의 염기를 결정하는 것은 바람직하게는 SNP 칩을 통해 수행할 수 있다.

[0054] 상기에서, 시퀀싱 분석은 염기서열 결정을 위한 통상적인 방법을 사용할 수 있으며, 자동화된 유전자분석기를 이용하여 수행될 수 있다. 또한, 대립유전자 특이적 PCR은 SNP가 위치하는 염기를 3' 말단으로 하여 고안한 프라이머를 포함한 프라이머 세트로 상기 SNP가 위치하는 DNA 단편을 증폭하는 PCR 방법을 의미한다. 상기 방법의 원리는, 예를 들어, 특정 염기가 C 또는 A로 치환된 경우, 상기 치환되지 않은 원래 서열을 포함하는 프라이머 및 적당한 크기의 DNA 단편을 증폭할 수 있는 반대 방향 프라이머를 고안하여 PCR 반응을 수행할 경우, 상기 SNP 위치의 염기가 치환되지 않은 경우에는 증폭반응이 정상적으로 수행되어 원하는 위치의 밴드가 관찰되고, 상기 염기가 C 또는 A로 치환된 경우에는 프라이머는 주형 DNA에 상보결합할 수 있으나, 3' 말단 쪽이 상보결합을 하지 못함으로써 증폭반응이 제대로 수행되지 않는 점을 이용한 것이다. DASH는 통상적인 방법으로 수행될 수 있고, 바람직하게는 프린스 등에 의한 방법에 의하여 수행될 수 있다.

[0055] 한편, PCR 연장 분석은 먼저 단일염기 다형성이 위치하는 염기를 포함하는 DNA 단편을 프라이머 쌍으로 증폭한 다음, 반응에 첨가된 모든 뉴클레오타이드를 탈인산화시킴으로써 불활성화시키고, 여기에 SNP 특이적 연장 프라이머, dNTP 혼합물, 디디옥시뉴클레오타이드, 반응 완충액 및 DNA 중합효소를 첨가하여 프라이머 연장반응을 수행함으로써 이루어진다. 이때, 연장 프라이머는 SNP가 위치하는 염기의 5' 방향의 바로 인접한 염기를 3' 말단으로 삼으며, dNTP 혼합물에는 디디옥시뉴클레오타이드와 동일한 염기를 갖는 핵산이 제외되고, 상기 디디옥시뉴클레오타이드는 SNP를 나타내는 염기 종류 중 하나에서 선택된다. 예를 들어, A에서 G로의 치환이 있는 경우, dGTP, dCTP 및 TTP 혼합물과 ddATP를 반응에 첨가할 경우, 상기 치환이 일어난 염기에서 프라이머는 DNA 중합효소에 의하여 연장되고, 몇 염기가 지난 후 A 염기가 최초로 나타나는 위치에서 ddATP에 의하여 프라이머 연장반응이 종결된다. 만일 상기 치환이 일어나지 않았다면, 그 위치에서 연장반응이 종결되므로, 상기 연장된 프라이머의 길이를 비교함으로써 SNP를 나타내는 염기 종류를 판별할 수 있게 된다.

[0056] 이때, 검출방법으로는 연장 프라이머 또는 디디옥시뉴클레오타이드를 형광 표지한 경우에는 일반적인 염기서열 결정에 사용되는 유전자 분석기(예를 들어, ABI사의 Model 3700 등)를 사용하여 형광을 검출함으로써 상기 SNP를 검출할 수 있으며, 무-표지된 연장 프라이머 및 디디옥시뉴클레오타이드를 사용할 경우에는 MALDI-TOF(matrix assisted laser desorption ionization-time of flight, 매트릭스 보조 레이저 탈착 이온화-비행 시간) 기법을 이용하여 분자량을 측정함으로써 상기 SNP를 검출할 수 있다.

[0057] 또한 상기 산물의 유전자형에서, 서열번호 1의 PTGR2 유전자의 유전자형이 rs1968105#G 또는 rs9646165#A 또는 이들 모두를 포함하고 있는 경우, 피검자는 아토피 피부염이 발병한 것으로 예측 또는 진단할 수 있다.

[0058] 본 발명의 또 다른 측면은 서열번호 1의 PTGR2 유전자의 염기서열에 있어서, 73849885 번째 염기가 C이고, 73850297 번째 염기가 C이며, 73851433 번째 염기가 T이고, 73851970 번째 염기가 G이며, 73854619 번째 염기가 G이고, 73858104 번째 염기가 G이고, 73858548 번째 염기가 G이고, 73862199 번째 염기가 G이고, 73862202 번째 염기가 G이고, 73871846 번째 염기가 T이고, 73874094 번째 염기가 A이고, 73874421 번째 염기가 G이고, 73875453 번째 염기가 C이고, 73880710 번째 염기가 A이고, 73880736 번째 염기가 C이고, 73882116 번째 염기가 A이고, 73882524 번째 염기가 G인 해플로타입을 구성하는 SNP 증폭을 위한 20 내지 100 개의 연속된 DNA 서열로 구성되는 아토피 피부염 진단을 위한 마커 조성물에 관한 것이다.

[0059] 본 발명에서 PTGR2 유전자란 prostaglandin reductase 2로, 프로스타글란딘(prostaglandins) 촉매 작용에 관여하는 효소이다. 상기 PTGR2 유전자의 염기서열은 NCBI의 GenBank 및 UCSC의 Genome Browser 등 공지의 데이터

베이스에서 얻을 수 있는데, 그 예로서 GenBank Accession No. NM_001146154.1로 표시되는 유전자가 될 수 있으며, 바람직하게는 서열번호 1의 폴리뉴클레오타이드를 포함할 수 있다.

- [0060] 상기 해플로타입은 PTGR2에서의 SNP들의 조합으로, 더욱 정확도가 높은 아토피 피부염의 예측 또는 진단을 위한 정보를 제공할 수 있는데, 이는 SNP간에 연관불평형을 구성하고 있기 때문이다. 상기 연관불평형(Linkage Disequilibrium)이란 집단 유전학에서 반드시 동일 염색체상에 존재하지는 않는 둘 이상의 유전자좌(loci)에서의 대립 유전자의 비-무작위적 연관(non-random association)을 의미한다.
- [0061] 본 발명에서 PTGR2 유전자에 있어서, 각각의 SNP 간의 완전한 또는 절대적인 결합으로 인해, 0.05 초과와 빈도를 갖는 4개 이상의 해플로타입이 구성될 수 있었다. 이들 4개의 해플로타입중에서, ht2는 아토피 피부염 발병과 상당히 유의한 연관성을 가지고 있음을 확인하였다(표 3). 게다가 아토피 피부염의 SCORAD와의 연관성 분석에서도 확인하였다. 이 분석은 피검자의 수에 따라 통계적 유의성이 0.02~0.06으로 보전되었다. 상술한 결과들을 통해 PTGR2 유전자의 SNP 조합인 해플로타입 2(ht2)이 아토피 피부염과 강하게 연관성을 가지고 있음을 시사한다 할 것이다.
- [0062] 본 발명자들은 PTGR2 유전자 다형성을 지노타이핑(genotyping)하고, 아토피 피부염 환자들에서 아토피 피부염과 관련있는 이들 유전자의 해플로타입을 확인하였다. 이에 본 발명의 ht2 해플로타입일 때 더욱 현재해짐을 확인한 바, 이는 아토피 피부염의 진단을 위한 새로운 방법을 제시하는데 매우 유용하다 할 것이다.
- [0063] 따라서, 본 발명의 마커 조성물의 유전자형을 검출하거나 또는 상기 유전자형을 검출할 수 있는 프라이머를 이용하면, 아토피 피부염을 갖거나 가지고 있을 것으로 예상되는 개체로부터 아토피 피부염을 정확하게 예측 또는 진단할 수 있을 것으로 기대되며, 상기 SNP는 본 발명자들에 의해 최초로 규명된 것이다.
- [0064] 본 발명의 다른 측면은 서열번호 1의 PTGR2 유전자의 염기서열에 있어서, 73849885 번째 염기가 C이고, 73850297 번째 염기가 C이며, 73851433 번째 염기가 T이고, 73851970 번째 염기가 G이며, 73854619 번째 염기가 G이고, 73858104 번째 염기가 G이고, 73858548 번째 염기가 G이고, 73862199 번째 염기가 G이고, 73862202 번째 염기가 G이고, 73871846 번째 염기가 T이고, 73874094 번째 염기가 A이고, 73874421 번째 염기가 G이고, 73875453 번째 염기가 C이고, 73880710 번째 염기가 A이고, 73880736 번째 염기가 C이고, 73882116 번째 염기가 A이고, 73882524 번째 염기가 G인 해플로타입을 포함하는 SNP 증폭을 위한 20 내지 100 개의 연속된 DNA 서열을 검출 또는 증폭할 수 있는 체제를 포함하는 아토피 피부염 예측 또는 진단용 조성물에 관한 것이다.
- [0065] 공개 데이터베이스 NCBI에서는 일부 SNP의 경우 아시안 자료가 없거나 인종간의 차이로 인하여 minor allele이 뒤바뀌어 표시될 수 있으나, 본 발명에서 표시된 해플로타입 2에서 minor allele로 치환된 한국인 유전자형 기준의 SNP 부위가, 73858104 번째 염기가 G이고(rs9646165), 73874094 번째 염기가 A이고(rs11159042), 73874421 번째 염기가 G(rs45559033)이고, 73882116 번째 염기가 A(rs6574151)인 것이 바람직하다. 상기 73874094, 73874421, 73882116의 경우, minor가 서열번호 1의 reference 염기와 동일하다 여길 수 있으나, 상기 3개의 경우는 NCBI DB에 아시안 빈도가 없거나 인종간의 차이로 인하여 minor allele이 상이한 경우가 있기 때문에, 상술한 해플로타입을 구성하는 SNP 증폭을 위한 20 내지 100 개의 연속된 DNA 서열을 검출 또는 증폭할 수 있는 체제를 포함하는 아토피 피부염 예측 또는 진단용 조성물인 것이 바람직하다.
- [0066] 본 발명의 용어 "SNP 마커를 검출 또는 증폭할 수 있는 체제"란, 상기와 같은 유전자의 단일염기다형성(SNP) 부위를 증폭을 통해 확인하여 아토피 피부염을 판단할 수 있는 조성물을 의미하며, 바람직하게는 상기 SNP 마커의 폴리뉴클레오타이드를 특이적으로 증폭할 수 있는 프라이머를 의미한다. 상기 SNP 마커 증폭에 사용되는 프라이머는, 적절한 버퍼 중의 적절한 조건(예를 들면, 4개의 다른 뉴클레오타이드 트리포스페이트 및 DNA, RNA 폴리머라제 또는 역전사 효소와 같은 중합제) 및 적당한 온도 하에서 주형-지시 DNA 합성의 시작점으로서 작용할 수 있는 단일가닥 올리고뉴클레오타이드가 될 수 있는데, 상기 프라이머의 적절한 길이는 사용 목적에 따라 달라질 수 있으나, 통상 15 내지 30 뉴클레오타이드이다. 짧은 프라이머 분자는 일반적으로 주형과 안정한 혼성체를 형성하기 위해서는 더 낮은 온도를 필요로 한다. 상기 프라이머 서열은 상기 SNP 마커와 완전하게 상보적일 필요는 없으나, 상기 SNP 마커와 혼성화할 정도로 충분히 상보적이어야 한다.
- [0067] 본 발명의 용어 "프라이머"란, 짧은 자유 3' 말단 수산화기(free 3' hydroxyl group)를 가지는 염기 서열로 상보적인 주형(template)과 염기쌍(base pair)을 형성할 수 있고 주형 가닥 복사를 위한 시작 지점으로 기능을 하는 짧은 서열을 의미한다. 프라이머는 적절한 완충용액 및 온도에서 중합반응(즉, DNA 폴리머레이즈 또는 역전사효소)을 위한 시약 및 상이한 4가지 뉴클레오타이드 트리포스페이트의 존재하에서 DNA 합성을 개시할 수 있다. PCR 증폭을 실시하여 원하는 생성물의 생성 여부를 통해 아토피 피부염을 예측할 수 있다. PCR 조건, 센

스 및 안티센스 프라이머의 길이는 당업계에 공지된 것을 기초로 변형할 수 있다.

- [0068] 본 발명의 프라이머는 포스포르아미다이트 고체 지지체 방법, 또는 기타 널리 공지된 방법을 사용하여 화학적으로 합성할 수 있다. 이러한 핵산 서열은 또한 당해 분야에 공지된 많은 수단을 이용하여 변형시킬 수 있다. 이러한 변형의 비-제한적인 예로는 메틸화, "캡화", 천연 뉴클레오타이드 하나 이상의 동족체로의 치환, 및 뉴클레오타이드 간의 변형, 예를 들면, 하전되지 않은 연결체(예: 메틸 포스포네이트, 포스포트리에스테르, 포스포로아미데이트, 카바메이트 등) 또는 하전된 연결체(예: 포스포로티오에이트, 포스포로디티오에이트 등)로의 변형이 있다.
- [0069] 또한, 본 발명의 또 다른 측면은 서열번호 1의 PTGR2 유전자의 염기서열에 있어서, 73849885 번째 염기가 C이고, 73850297 번째 염기가 C이며, 73851433 번째 염기가 T이고, 73851970 번째 염기가 G이며, 73854619 번째 염기가 G이고, 73858104 번째 염기가 G이고, 73858548 번째 염기가 G이고, 73862199 번째 염기가 G이고, 73862202 번째 염기가 G이고, 73871846 번째 염기가 T이고, 73874094 번째 염기가 A이고, 73874421 번째 염기가 G이고, 73875453 번째 염기가 C이고, 73880710 번째 염기가 A이고, 73880736 번째 염기가 C이고, 73882116 번째 염기가 A이고, 73882524 번째 염기가 G인 해플로타입을 구성하는 SNP 증폭을 위한 20 내지 100 개의 연속된 DNA 서열을 검출 또는 증폭할 수 있는 제제를 포함하는 아토피 피부염 예측 또는 진단용 키트에 관한 것이다.
- [0070] 본 발명의 키트는 아토피 피부염을 진단할 수 있는 PTGR2 유전자의 SNP를 증폭을 통해 확인하거나, SNP 마커의 발현 수준을 확인함으로써, 아토피 피부염을 예측 또는 진단할 수 있게 하는 것이다.
- [0071] 본 발명에 있어서, 상기 키트를 사용하고자 하는 대상은 동아시아인을 대상으로 하는 것이 바람직하고, 한국인, 일본인 또는 중국인을 대상으로 하는 것이 더욱 바람직하나 이에 한정되지 않는다.
- [0072] 상기 진단용 키트에는 본 발명의 PTGR2 유전자의 SNP(rs1968105#G 또는 rs9646165#A)를 증폭시킬 수 있는 프라이머 뿐만 아니라, 중합 반응에 필요한 시약, 예를 들면 dNTP, 각종의 중합효소 및 발색제 등을 포함할 수 있다.
- [0073] 상기 "증폭시킬 수 있는 프라이머"란 적절한 버퍼 중의 적절한 조건(예를 들면, 4개의 다른 뉴클레오타이드 트리포스페이트 및 DNA, RNA 폴리머라제 또는 역전사 효소와 같은 중합제) 및 적당한 온도 하에서 주형-지시 DNA 합성의 시작점으로서 작용할 수 있는 단일가닥 올리고뉴클레오타이드를 말한다. 상기 프라이머의 적절한 길이는 사용 목적에 따라 달라질 수 있으나, 통상 15 내지 30 뉴클레오타이드이다. 짧은 프라이머 분자는 일반적으로 주형과 안정한 혼성체를 형성하기 위해서는 더 낮은 온도를 필요로 한다. 프라이머 서열은 주형과 완전하게 상보적일 필요는 없으나, 주형과 혼성화 할 정도로 충분히 상보적이어야 한다.
- [0074] 본 발명에서 아토피 피부염 예측 또는 진단용 조성물의 mRNA 발현 수준을 측정하기 위한 키트는 RT-PCR을 수행하기 위해 필요한 필수 요소를 포함하는 키트일 수 있다. RT-PCR 키트는, 상기 SNP 마커의 유전자에 대한 특이적인 각각의 프라이머 쌍 외에도 RT-PCR 키트는 테스트 튜브 또는 다른 적절한 컨테이너, 반응 완충액(pH 및 마그네슘 농도는 다양), 데옥시뉴클레오타이드(dNTPs), Taq-폴리머라아제 및 역전사효소와 같은 효소, DNase, RNase 억제제, DEPC-수(DEPCwater), 멸균수 등을 포함할 수 있다. 또한 정량 대조군으로 사용되는 유전자에 특이적인 프라이머 쌍을 포함할 수 있다. 또한 바람직하게는, 본 발명의 키트는 DNA 칩을 수행하기 위해 필요한 필수 요소를 포함하는 아토피 피부염 예측 또는 진단용 키트일 수 있다. DNA 칩 키트는, 일반적으로 편평한 고체 지지판, 전형적으로는 현미경용 슬라이드보다 크지않은 유리 표면에 핵산 종을 격자형 배열(grided array)로 부착한 것으로, 칩 표면에 핵산이 일정하게 배열되어, DNA 칩 상의 핵산과 칩 표면에 처리된 용액 내에 포함된 상보적인 핵산 간에 다중 혼성화(hybridization) 반응이 일어나 대량 병렬 분석이 가능하도록 하는 도구이다.
- [0075] 본 발명의 또 다른 측면은 아토피 피부염 질환의 진단에 필요한 정보를 제공하기 위하여, 서열번호 1의 PTGR2 유전자의 염기서열에 있어서, 73849885 번째 염기가 C이고, 73850297 번째 염기가 C이며, 73851433 번째 염기가 T이고, 73851970 번째 염기가 G이며, 73854619 번째 염기가 G이고, 73858104 번째 염기가 G이고, 73858548 번째 염기가 G이고, 73862199 번째 염기가 G이고, 73862202 번째 염기가 G이고, 73871846 번째 염기가 T이고, 73874094 번째 염기가 A이고, 73874421 번째 염기가 G이고, 73875453 번째 염기가 C이고, 73880710 번째 염기가 A이고, 73880736 번째 염기가 C이고, 73882116 번째 염기가 A이고, 73882524 번째 염기가 G인 해플로타입을 증폭하는 단계;를 포함하는 아토피 피부염에 대한 유전자 PTGR2의 유전자형 분석방법에 관한 것이다.
- [0076] 구체적으로 피검자로부터 체외로 분리된 샘플을 취득하고, 상기 샘플에서 서열번호 1의 PTGR2 유전자의 염기서

열에 있어서, 73849885 번째 염기가 C이고, 73850297 번째 염기가 C이며, 73851433 번째 염기가 T이고, 73851970 번째 염기가 G이며, 73854619 번째 염기가 G이고, 73858104 번째 염기가 G이고, 73858548 번째 염기가 G이고, 73862199 번째 염기가 G이고, 73862202 번째 염기가 G이고, 73871846 번째 염기가 T이고, 73874094 번째 염기가 A이고, 73874421 번째 염기가 G이고, 73875453 번째 염기가 C이고, 73880710 번째 염기가 A이고, 73880736 번째 염기가 C이고, 73882116 번째 염기가 A이고, 73882524 번째 염기가 G인 해플로타입을 증폭한 후, 아토피 피부염에 대한 유전자 PTGR2의 유전자형을 분석함으로써, 아토피 피부염 예측 또는 진단에 필요한 정보를 제공할 수 있다.

[0077] 우선, 피검자로부터 채취로 분리된 샘플은 특별히 한정된 것은 아니며, 일례로, 털, 뇨, 타액, 각종 체액, 근육, 표피, 혈액, 뼈, 장기로부터 얻은 것일 수 있고, 바람직하게는 근육 또는 혈액일 수 있다.

[0078] 상기 SNP 유전자형을 분석하기 위해, 상기 분리된 샘플로부터 DNA를 추출할 수 있다. gDNA의 분리는 당업계에 공지된 통상의 방법에 따라 실시될 수 있으며(참조: Rogers & Bendich (1994)), 출발물질이 mRNA인 경우에는, 역전사효소를 이용하여 cDNA로 합성될 수 있다.

[0079] 다음으로, 상기 샘플로부터, 서열번호 1의 PTGR2 유전자의 염기서열에 있어서, 73849885 번째 염기가 C이고, 73850297 번째 염기가 C이며, 73851433 번째 염기가 T이고, 73851970 번째 염기가 G이며, 73854619 번째 염기가 G이고, 73858104 번째 염기가 G이고, 73858548 번째 염기가 G이고, 73862199 번째 염기가 G이고, 73862202 번째 염기가 G이고, 73871846 번째 염기가 T이고, 73874094 번째 염기가 A이고, 73874421 번째 염기가 G이고, 73875453 번째 염기가 C이고, 73880710 번째 염기가 A이고, 73880736 번째 염기가 C이고, 73882116 번째 염기가 A이고, 73882524 번째 염기가 G인 해플로타입 부위를 증폭하거나 프로브와 혼성화할 수 있는데, 이는 당업자에게 알려진 어떠한 방법이든 사용가능하다. 예를 들면, 표적 핵산을 PCR을 통하여 증폭하고 이를 정제하여 얻을 수 있다. 그 외 리가제 연쇄 반응(LCR)(Wu 및 Wallace, Genomics 4, 560(1989), Landegren 등, Science 241, 1077(1988)), 전사증폭(transcription amplification)(Kwoh 등, Proc. Natl. Acad. Sci. USA 86, 1173(1989)) 및 자가유지 서열 복제(Guatelli 등, Proc. Natl. Acad. Sci. USA 87, 1874(1990)) 및 핵산에 근거한 서열 증폭(NASBA)이 사용될 수 있다.

[0080] 이후, 상기 산물로부터 유전자형을 확인함으로써, 아토피 피부염에 대한 유전자 PTGR2의 유전자형 분석방법을 제공하는 것을 특징으로 할 수 있다. 상기 산물로부터 유전자형을 분석하기 위해서는 당업자에게 알려진 어떠한 방법이든 사용가능하나, 예를 들어 서열 분석, 마이크로어레이(microarray)에 의한 혼성화, 대립유전자 특이적인 PCR(allele specific PCR), 다이내믹 대립유전자 혼성화 기법(dynamic allelespecific hybridization, DASH), PCR 연장 분석, PCR-SSCP, PCR-RFLP 분석 또는 TaqMan 기법, SNPlex 플랫폼(Applied Biosystems), 질량 분석법(예를 들면, Sequenom의 MassARRAY 시스템), 미니-시퀀싱(mini-sequencing) 방법, Bio-Plex 시스템(BioRad), CEQ and SNPstream 시스템(Beckman), Molecular Inversion Probe 어레이 기술(예를 들면, Affymetrix GeneChip), Fluidigm 시스템 및 BeadArray Technologies(예를 들면, Illumina GoldenGate 및 Infinium 분석법)를 포함하나, 이에 한정되지 않는다. 상기 방법들 또는 본 발명이 속하는 기술분야의 당업자에게 이용 가능한 다른 방법에 의해, 마이크로세틀라이트, SNP 또는 다른 종류의 다형성 마커를 포함한, 다형성 마커에서의 하나 이상의 대립유전자가 확인될 수 있다. 이와 같은 다형성 부위의 염기를 결정하는 것은 바람직하게는 SNP 칩을 통해 수행할 수 있다.

[0081] 상기에서, 시퀀싱 분석은 염기서열 결정을 위한 통상적인 방법을 사용할 수 있으며, 자동화된 유전자분석기를 이용하여 수행될 수 있다. 또한, 대립유전자 특이적 PCR은 SNP가 위치하는 염기를 3' 말단으로 하여 고안한 프라이머를 포함한 프라이머 세트에 상기 SNP가 위치하는 DNA 단편을 증폭하는 PCR 방법을 의미한다. 상기 방법의 원리는, 예를 들어, 특정 염기가 C 또는 A로 치환된 경우, 상기 치환되지 않은 원래 서열을 포함하는 프라이머 및 적당한 크기의 DNA 단편을 증폭할 수 있는 반대 방향 프라이머를 고안하여 PCR 반응을 수행할 경우, 상기 SNP 위치의 염기가 치환되지 않은 경우에는 증폭반응이 정상적으로 수행되어 원하는 위치의 밴드가 관찰되고, 상기 염기가 C 또는 A로 치환된 경우에는 프라이머는 주형 DNA에 상보결합할 수 있으나, 3' 말단 쪽이 상보결합을 하지 못함으로써 증폭반응이 제대로 수행되지 않는 점을 이용한 것이다. DASH는 통상적인 방법으로 수행될 수 있고, 바람직하게는 프린스 등에 의한 방법에 의하여 수행될 수 있다.

[0082] 한편, PCR 연장 분석은 먼저 단일염기 다형성이 위치하는 염기를 포함하는 DNA 단편을 프라이머 쌍으로 증폭한 다음, 반응에 첨가된 모든 뉴클레오타이드를 탈인산화시킴으로써 불활성화시키고, 여기에 SNP 특이적 연장 프라이머, dNTP 혼합물, 디디옥시뉴클레오타이드, 반응 완충액 및 DNA 중합효소를 첨가하여 프라이머 연장반응을 수행함으로써 이루어진다. 이때, 연장 프라이머는 SNP가 위치하는 염기의 5' 방향의 바로 인접한 염기를 3'

말단으로 삼으며, dNTP 혼합물에는 디디옥시뉴클레오타이드와 동일한 염기를 갖는 핵산이 제외되고, 상기 디디옥시뉴클레오타이드는 SNP를 나타내는 염기 종류 중 하나에서 선택된다. 예를 들어, A에서 G로의 치환이 있는 경우, dGTP, dCTP 및 TTP 혼합물과 ddATP를 반응에 첨가할 경우, 상기 치환이 일어난 염기에서 프라이머는 DNA 중합효소에 의하여 연장되고, 몇 염기가 지난 후 A 염기가 최초로 나타나는 위치에서 ddATP에 의하여 프라이머 연장반응이 종결된다. 만일 상기 치환이 일어나지 않았다면, 그 위치에서 연장반응이 종결되므로, 상기 연장된 프라이머의 길이를 비교함으로써 SNP를 나타내는 염기 종류를 판별할 수 있게 된다.

[0083] 이때, 검출방법으로는 연장 프라이머 또는 디디옥시뉴클레오타이드를 형광 표지한 경우에는 일반적인 염기서열 결정에 사용되는 유전자 분석기(예를 들어, ABI사의 Model 3700 등)를 사용하여 형광을 검출함으로써 상기 SNP를 검출할 수 있으며, 무-표지된 연장 프라이머 및 디디옥시뉴클레오타이드를 사용할 경우에는 MALDI-TOF(matrix assisted laser desorption ionization-time of flight) 기법을 이용하여 분자량을 측정함으로써 상기 SNP를 검출할 수 있다.

[0084] 또한 상기 산물의 유전자형에서, 서열번호 1의 PTGR2 유전자의 염기서열에 있어서, 73849885 번째 염기가 C이고, 73850297 번째 염기가 C이며, 73851433 번째 염기가 T이고, 73851970 번째 염기가 G이며, 73854619 번째 염기가 G이고, 73858104 번째 염기가 G이고, 73858548 번째 염기가 G이고, 73862199 번째 염기가 G이고, 73862202 번째 염기가 G이고, 73871846 번째 염기가 T이고, 73874094 번째 염기가 A이고, 73874421 번째 염기가 G이고, 73875453 번째 염기가 C이고, 73880710 번째 염기가 A이고, 73880736 번째 염기가 C이고, 73882116 번째 염기가 A이고, 73882524 번째 염기가 G인 해플로타입을 포함하고 있는 경우를 확인하면, 피검자는 아토피 피부염이 발병한 것으로 예측 또는 진단할 수 있다.

발명의 효과

[0085] 본 발명의 특징 및 이점을 요약하면 다음과 같다:

[0086] (i) 본 발명은 PTGR2 유전자 내의 rs1968105#G 또는 rs9646165#A의 단일염기다형성(SNP)를 포함하는 연속된 DNA 서열로 구성된 아토피 피부염 진단을 위한 마커 조성물을 제공한다.

[0087] (ii) 또한, 본 발명은 PTGR2 유전자 내에서의 연관불평형 분석을 통해 해플로타입 마커 조성물을 제공한다.

[0088] (iii) 본 발명은 다양한 원인에 의해 발병되는 아토피 피부염 관련 유전자들의 연관성 분석을 통해 유전적 특성을 확인하였으며, 이를 통해 발굴한 마커 조성물은 아토피 피부염을 정확히 구분할 수 있는 기능을 보이며, 이 기능을 이용하여, 아토피 피부염에 대한 예측 또는 진단 조성물으로써 유용하게 활용될 수 있다.

도면의 간단한 설명

[0089] 도 1은 PTGR2 유전자의 물리적 지도, 해플로타입 및 LD(Linkage Disequilibrium)를 나타낸 것이다. 도 1A는 PTGR2 유전자의 물리적 지도와 본 연구에서 확인된 SNP의 위치를 나타낸 것으로, 염색체 14q24.3(참조 게놈 서열 NM_001146154.1) 상의 PTGR2에서 확인된 다형성을 나타낸다. 이때, ■은 암호화된 엑손이고, □은 5'-UTR(5'-untranslated region, 5'-비번역 부위)과 3'-UTR(3'-비번역 부위)을 의미한다.

도 1B는 PTGR2 유전자의 해플로타입을 나타낸 것이다.

도 1C는 PTGR2 유전자의 LD 플롯이고, 본 실험에서 조사된 SNP는 하나의 LD 블록을 구성하며, 블록 번호는 LD 계수 |D'|의 수치를 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0090] 이하, 실시예를 통하여 본 발명을 더욱 상세히 설명하고자 한다. 이들 실시예는 오로지 본 발명을 보다 구체적으로 설명하기 위한 것으로, 본 발명의 요지에 따라 본 발명의 범위가 이들 실시예에 의해 제한되지 않는다는 것은 당업계에서 통상의 지식을 가진 자에 있어서 자명할 것이다.

[0091] 실시예

[0092] <실시예 1> 피검자의 선정

[0093] 서울 아산 병원 소아 천식 아토피 센터에서 Hanifin and Rajka의 기준에 따라 진단된 아토피 피부염(AD) 환자 420명을 모집하였다. 상기 환자들의 고통 정도는 SCORing AD(SCORAD)를 사용하여 다음과 같이 평가 및 구분하였다: 경미증(mild)<15, 15≤중등도(moderate)<40, 중증(severe)≥40.

[0094] 정상 대조군(n=444)은 서울지역에 위치한 9개 초등학교와 16 개의 유치원으로부터 모집하였다. 정상 대조군은 아토피 피부염(AD), 음식 알레르기, 알레르기 비염, 천식, 알츠하이머 병력이 없거나, 부모님의 아토피 피부염 병력이 없고, 피부 발진 검사 결과가 음성인 구성원들로 모집하였다.

[0095] 모든 피검자들에 대해 말초혈액(peripheral blood)으로부터 총 혈청 IgE 수치를 측정하였고, 이때 총 혈청 IgE 수치는 형광 효소 면역 측정법(ImmunoCAP system; Phadia AB, Uppsala, Sweden)을 사용하였다. 또한 피검자는 16 알레르겐에 대하여 피부단자검사를 통해 민감도를 측정하였다(사상세모무늬먼지진드기(*Dermatophagoides pteronyssinus*), 큰다리먼지진드기(*Dermatophagoides farina*), 개 상피(dog epithelium), 고양이 상피(cat epithelium), 바퀴벌레(cockroach), 잔디, 여러 나무의 꽃가루 혼합물 1 및 2, *Alternaria*(알터나리아), *Aspergillus*(아스페르길루스), 돼지풀(ragweed), 쑥(mugwort), 우유, 달걀 흰자, 땅콩 및 콩).

[0096] 상기 피부단자검에서 하나 이상의 포지티브 반응을 나타내면 아토피로 정의하였고, 포지티브 반응은 음성 대조군을 뺀 후, 직경 3 mm 초과하는 것으로 정의하였다. 본 발명에서 한국판 어린이 천식 및 알레르기에 관한 국제 조사(ISAAC) 설문지를 사용하였고, 모든 피험자는 본 발명에 참가하기 위해 부모로부터 서면 동의를 받아 진행하였고, 프로토콜은 지역윤리위원회로부터 승인받았다.

[0097] 피검자들의 임상학적 특징들을 표 1에 요약하였다. 앞서 살펴본 바와 같이 420명의 아토피 피부염 환자와 444명의 정상 대조군으로 구성된 집단(cohort)을 모집하였고, 아토피 피부염 환자들의 평균연령이 정상 대조군보다 대체적으로 낮았다(7.0 ± 3.7 year vs. 11.9 ± 2.1 year, $P < 0.01$). 총 혈청 IgE 수치는 아토피 피부염 환자가 정상 대조군보다 유의하게 더 높았다(660.8 ± 1040.1 IU/ml vs. 119.9 ± 225.1 IU/ml, $P < 0.01$).

표 1

임상 프로파일 (Clinical profile)	아토피 피부염 환자	정상 대조군	P-value
피검자 수(n)	420	444	-
연령(year)	7.0 ± 3.7	11.9 ± 2.1	< 0.01
성별(male, %)	54.9	41.0	-
SCORAD	69.66 ± 18.28	NA	-
총 IgE(IU/ml)	660.8 ± 1040.1	119.9 ± 225.1	< 0.01

[0099] SCORAD, scoring AD; NA, not applicable.

[0100] 테이터는 평균 $\pm$ 표준편차이다.

[0101] P값은 아토피 피부염 환자와 정상 대조군 사이의 차이에 대한 수치이다.

[0102] <실시예 2> PTGR2의 단일염기다형성(SNP) 유전자형(genotyping), LD, 해플로타입(Haplotype) 분석

[0103] 추적연구 대상으로 GWAS에서 가장 유의한 연관성을 보인 유전자인 PTGR2를 선택하였다. 마이너 대립형질 빈도(minor allele frequency, MAF) 및 연관 불평형(linkage disequilibrium, LD)를 고려하여 유전자로부터 유전자 앞쪽 2 kb 와 뒤쪽 1 kb의 범위를 포함하는 17개의 PTGR2 유전자 SNP를 선택하였다. 구체적으로 MAF 0.05 이상 인이고, 아시아 인종 자료(중국 및 일본 포함)에서 LD를 고려하여 1000 Genomes Project(<http://browser.1000genomes.org/index.html>)로부터 총 17개의 SNP를 선택하였다.

[0104] DNA는 WizPrep™ gDNA Mini Kit(Wizbio solutions, 성남, 한국)를 이용하여 말초 혈액으로부터 분리하였다. 상 기 혈액 시료의 유전형은 Fluidigm SNP Type™ 분석 플랫폼을 갖춘 고효율 플루이다임 EP1 시스템(high-throughput Fluidigm EP1 system)(Fluidigm, South San Francisco, CA)을 사용하여 분석하였다. 메뉴얼에 따라 수행하였고, 특정 표적 증폭 반응을 사용하여 Qiagen 2X Multiplex PCR Master Mix(Qiagen, Hilden, Germany)를 사용하여 표적 게놈 영역의 카피수를 증가시켰다. 동적 어레이 IFC(Integrated Fluidic Circuits, Fluidigm)에서 증폭 반응을 수행하고, 형광 강도를 EP1 판독기로 측정하였다. 유전형은 Fluidigm SNP Genotyping Analysis 프로그램을 사용하여 분석하였다. SNP 결정(determination)은 육안 검사를 통해 수행되었고, 전체 시료 중에서 call rate가 95% 이상인 SNP를 선택하였다.

[0105] 통계분석(statistics)은 다음과 같이 수행되었다. 연관 불평형(LD)은 Broad Institute(<http://www.broadinstitute.org/mpg/haploview>)로부터 얻은 Haploview v4.2 소프트웨어를 사용하여

분석하였다. Lewontin's D'(|D'|) 및 LD 상관계수 r^2 를 검사하여 2대립유전자 위치(biallelic loci)의 모든 쌍 사이의 연관 불균형(LD)를 측정하였다. PHASE 소프트웨어를 사용하여 해플로타입(Haplotypes) 연관성을 평가하였다. 해플로타입 연관성(association)의 경우, 0.05 이상의 빈도를 갖는 일반적인 해플로타입만 분석되었다.

[0106] 로지스틱 분석 및 선형 회귀 분석은 통계 분석 시스템 소프트웨어(SAS, v9.4)을 사용하여 연관 분석을 수행하였다. 다변수 분석을 위해, 연령과 성별을 공변량으로 하여 P 값을 계산하였다. 0.05 미만의 P 값은 통계적으로 유의하다고 간주하였다.

[0107] 도 1은 PTGR2 유전자의 물리적 지도, 해플로타입 및 LD를 나타낸 것이다. 도 1A는 PTGR2 유전자의 물리적 지도와 본 연구에서 확인된 SNP 유전형을 나타낸 것이다. ■은 암호화된 엑손이고, □은 5'-UTR(5'-untranslated region)과 3'-UTR을 의미한다. 도 1B는 PTGR2 유전자의 해플로타입을 나타낸 것이다. 도 1C는 PTGR2 유전자의 LD 플롯이고, 본 실험에서 조사된 SNP는 하나의 LD 블록을 구성하며, 블록 번호는 LD 계수 |D'|의 수치를 나타낸다.

[0108] 도 1에 나타난 바와 같이, 상술한 방법에 따라 염색체 14q24.3(~35 kb)에서 PTGR2(NM_001146154.1)의 총 17개 SNP(단일염기다형성)를 확인 및 선택되었고, 성공적으로 864명 피검자들을 대상으로 유전형을 결정하였음을 확인하였다. 또한, 염기서열의 쌍(Pair-wise) 비교를 통해 하나의 LD 블록을 확인하였고, PHASE 소프트웨어를 사용하여 4 개의 해플로타입(ht1, ht2, ht3, ht4)을 확인하였다.

[0109] <실시예 3> PTGR2 임상 유전형(clinical phenotypes)과 아토피 피부염의 연관성 분석(Association analysis)

[0110] Case-control 분석을 위해 피검자들의 나이와 성별을 공변량으로 하여, 로지스틱 회귀 분석(Logistic regression analysis)을 수행하였고, 이를 표 2, 3에 나타내었다. 표 2는 아토피 피부염과 SNP의 연관성 분석 결과이다. 표 3은 아토피 피부염과 해플로타입의 연관성 분석 결과이다.

표 2

[0111]

SNP	위치 (a.a. change)	Minor allele	MAF		Genetic model					
					Additive		Dominant		Recessive	
			아토피 피 부염 환자 (420명)	정상 대조 군 (444명)	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
rs35423 936	Promoter	G	0.062	0.062	0.98 (0.52-1.6 2)	0.93	0.96 (0.56-1.6 2)	0.86	1.88 (0.12-29.2 0)	0.65
rs17092 737	Promoter	T	0.425	0.431	0.98 (0.77-1.2 6)	0.90	0.96 (0.66-1.4 0)	0.83	1.00 (0.64-1.5 8)	0.99
rs72627 109	Promoter	C	0.431	0.444	0.96 (0.75-1.2 3)	0.74	0.93 (0.63-1.3 5)	0.69	0.97 (0.62-1.5 2)	0.90
rs80206 30	5'UTR	A	0.163	0.166	0.82 (0.57-1.1 8)	0.29	0.82 (0.56-1.2 1)	0.32	0.59 (0.11-3.0 8)	0.53
rs17092 878	Intron	A	0.473	0.498	0.89 (0.69-1.1 4)	0.36	0.86 (0.58-1.2 8)	0.45	0.85 (0.56-1.2 9)	0.45
rs96461 65	Intron	A	0.289	0.264	1.36 (1.01-1.8 3)	0.044	1.45 (1.00-2.1 1)	0.05	1.50 (0.73-3.0 5)	0.27
rs19627 95	Intron	A	0.121	0.111	0.95 (0.64-1.4 3)	0.82	0.94 (0.61-1.4 4)	0.77	1.25 (0.19-8.1 0)	0.82
rs19681 06	Intron	G	0.033	0.048	0.55 (0.29-1.0 4)	0.07	0.58 (0.30-1.1 2)	0.10	.	0.98
rs19681 05	Intron	G	0.037	0.059	0.50 (0.27-0.9 2)	0.026	0.52 (0.28-0.9 8)	0.041	.	0.98
rs17782 222	Intron	A	0.061	0.044	1.33 (0.75-2.3 4)	0.33	1.38 (0.77-2.4 7)	0.28	.	0.99

rs11159042	Exon (Gly76Gly)	A	0.406	0.391	1.15 (0.89-1.48)	0.28	1.25 (0.86-1.81)	0.24	1.14 (0.71-1.83)	0.60
rs45559033	Intron	G	0.467	0.435	1.20 (0.94-1.54)	0.14	1.45 (0.98-2.13)	0.06	1.11 (0.72-1.71)	0.63
rs11621087	Intron	T	0.061	0.052	1.12 (0.64-1.97)	0.70	1.12 (0.64-1.97)	0.70	.	.
rs78870464	Intron	G	0.073	0.071	0.87 (0.53-1.43)	0.59	0.90 (0.54-1.51)	0.69	.	0.98
rs17782239	Intron	T	0.150	0.150	0.78 (0.54-1.14)	0.20	0.76 (0.51-1.14)	0.18	0.85 (0.17-4.37)	0.85
rs6574151	Intron	A	0.342	0.327	1.17 (0.90-1.52)	0.23	1.16 (0.81-1.65)	0.42	1.43 (0.83-2.46)	0.20
rs8005140	Intron	A	0.121	0.113	0.95 (0.64-1.42)	0.81	0.94 (0.61-1.43)	0.76	1.25 (0.19-8.10)	0.82

표 3

Haplotype	빈도		Genetic model					
			Additive		Dominant		Recessive	
	아토피 피부염 환자 (420명)	정상대조군 (444명)	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
ht1	0.424	0.428	0.99 (0.77-1.27)	0.91	0.97 (0.67-1.41)	0.86	1.00 (0.63-1.58)	0.99
ht2	0.286	0.253	1.39 (1.05-1.85)	0.02	1.43 (0.99-2.05)	0.05	1.82 (0.92-3.62)	0.09
ht3	0.061	0.068	0.73 (0.43-1.24)	0.24	0.74 (0.43-1.27)	0.27	0.26 (0.01-10.47)	0.48
ht4	0.062	0.062	0.98 (0.59-1.62)	0.93	0.96 (0.56-1.62)	0.86	1.88 (0.12-29.20)	0.65

표 2 및 3에서 로지스틱 회귀 분석의 p 값은 나이와 성별을 공변량으로 하여 조정하였고, 아토피 피부염 환자와 정상대조군으로부터 편차에 대한 수치이다. 굵은 글씨는 $P < 0.05$ 의 통계적 유의성을 나타내며, MAF는 마이너 대립형질 빈도, a.a는 아미노산, OR은 승산비(odds ratio), CI는 신뢰 구간(confidence interval)을 의미한다.

표 2에 나타난 바와 같이, 2 개의 SNP($P = 0.026$ at rs1968105, $P = 0.044$ at rs9646165)이 아토피 피부염 발병과 상당히 유의한 연관성을 가지고 있음을 확인할 수 있었다. 상기 SNP를 제외한 PTGR2 유전자 내의 다른 유전적 다형성과 아토피 피부염 사이에는 어떠한 연관성도 검출되지 않았다. 또한, 해플로타입 분석에서 해플로타입2(ht2, $P = 0.02$)가 아토피 피부염 발병과 상당히 유의한 연관성을 가지고 있음을 확인할 수 있었다. 즉, 2 개의 SNP($P = 0.026$ at rs1968105, $P = 0.044$ at rs9646165)와 ht2는 아토피 피부염의 위험성에 대해 민감한 효과를 가지고 있음을 알 수 있다.

아토피 피부염 환자와 정상 대조군 사이에서 PTGR2 유전자형과 아토피 피부염의 SCORAD와의 연관성 분석을 위해, 피검자들의 나이와 성별을 공변량으로 하여, 선형 회귀 분석(linear regression analysis)을 수행하였고, 이를 표 4에 나타내었다.

표 4

SNP	위치 (a.a. change)	Minor allele	SCORAD					
			Score			P-value		
			M/M (n)	M/m (n)	m/m (n)	Additive	Dominant	Recessive

rs35423 936	Promoter	G	69.23±5.41 (164)	72.87±14.87 (24)	62.6 (1)	0.39	0.33	0.73
rs17092 737	Promoter	T	69.76±8.79 (63)	69.15±7.25 (91)	70.80±11.97 (35)	0.99	0.84	0.81
rs72627 109	Promoter	C	70.00±8.89 (62)	68.99±7.23 (91)	70.73±11.79 (36)	0.94	0.76	0.82
rs80206 30	5'UTR	A	67.78±5.99 (128)	74.34±9.76 (58)	59.13±34.14 (3)	0.09	0.036	0.30
rs17092 878	Intron	A	69.96±9.70 (52)	69.55±7.21 (93)	69.54±10.48 (44)	0.76	0.77	0.85
rs96461 65	Intron	A	72.90±7.40 (97)	65.92±7.51 (77)	64.15±20.29 (10)	0.020	0.017	0.38
rs19627 95	Intron	A	67.92±5.66 (144)	75.45±11.37 (44)	65.4 (1)	0.021	0.015	0.77
rs19681 06	Intron	G	69.39±5.28 (173)	72.54±18.14 (16)	.	0.87	0.87	.
rs19681 05	Intron	G	69.34±5.30 (171)	72.62±17.12 (18)	.	0.81	0.81	.
rs17782 222	Intron	A	69.28±5.36 (167)	72.53±15.46 (22)	.	0.44	0.44	.
rs11159 042	Exon (Gly76Gly)	A	70.44±8.67 (66)	70.94±7.28 (95)	63.45±11.99 (28)	0.24	0.77	0.07
rs45559 033	Intron	G	70.94±9.74 (53)	69.937.03 (99)	67.10±11.03 (37)	0.43	0.61	0.43
rs11621 087	Intron	T	69.43±5.37 (167)	69.90±15.25 (21)	.	0.90	0.90	.
rs78870 464	Intron	G	68.81±5.41 (162)	74.73±14.38 (27)	.	0.13	0.13	.
rs17782 239	Intron	T	67.24±5.77 (136)	76.21±10.78 (50)	69.93±40.38 (3)	0.006	0.002	0.96
rs65741 51	Intron	A	71.95±7.80 (85)	68.40±7.46 (84)	65.17±14.57 (20)	0.10	0.12	0.32
rs80051 40	Intron	A	67.92±5.66 (144)	75.45±11.37 (44)	65.40 (1)	0.021	0.015	0.77
Haplotype								
ht1			69.78±8.79 (63)	69.15±7.25 (91)	70.80±11.97 (35)	0.99	0.84	0.81
ht2			72.83±7.21 (102)	65.54±7.72 (72)	67.85±17.52 (15)	0.06	0.02	0.79
ht3			68.55±5.34 (165)	77.30±15.78 (24)	.	0.02	0.02	.
ht4			69.23±5.41 (164)	72.87±14.87 (24)	62.60 (1)	0.39	0.33	0.73

[0117] 선형 회귀 분석의 P 값은 공변수로서 연령, 성별을 조정한 공동-우성-열성 유전자형 모델을 사용하여 계산되었다.

[0118] M/M, M/m, 및 m/m은 각각 보편적 대립유전자(major allele)에 대한 동형접합자(homozygote), 이형접합자(heterozygote) 및 희귀 대립유전자(minor allele)에 대한 동형접합자(homozygote)를 나타낸다.

[0119] 굵은 수치는 P<0.05의 통계적 유의성을 나타내며, a.a는 아미노산, n은 피검사자의 수를 의미한다.

[0120] 표 4는 아토피 피부염 환자에서 SCORAD와 총 IgE를 위해 연령과 성별을 공변량으로 하여 조정된 선형 회귀 분석 결과이다. 이에 따르면 아토피 피부염 환자에서 아토피 피부염 평가의 유의한 지침으로써 사용될 수 있다. 구체적으로 4 SNP(rs9646165에서 P = 0.02, rs1962795에서 P = 0.021, rs17782239에서 P = 0.006, rs8005140에서 P = 0.021)와 해플로타입3(haplotype 3, ht3, P = 0.02)은 SCORAD에서 유의한 신호를 갖고 있는 것으로 확인되었다. 게다가 rs8020630(P = 0.036)과 ht2(P = 0.02)는 우성 유전자형 모델에서 SCORAD와 유의한 연관성을 보였다.

[0121] 아토피 피부염 환자와 정상 대조군 사이에서 PTGR2 유전자형과 총 IgE의 연관성 분석을 위해, 피검자들의 나이와 성별을 공변량으로 하여, 선형 회귀 분석(linear regression analysis)을 수행하였고, 이를 표 5에 나타내었다.

다.

표 5

SNP	Position (a.a. change)	Minor allele	Total IgE					
			Score			P-value		
			M/M (n)	M/m (n)	m/m (n)	Additive	Dominant	Recessive
rs35423936	Promoter	G	650.9±34.0 (366)	735.5±106.2 (48)	186.6±131.9 (2)	0.88	0.76	0.47
rs17092737	Promoter	T	591.9±49.5 (143)	724.9±52.4 (191)	625.0±69.4 (81)	0.57	0.31	0.86
rs72627109	Promoter	C	594.9±50.1 (141)	720.9±52.2 (191)	623.2±68.0 (84)	0.63	0.37	0.85
rs8020630	5'UTR	A	661.5±39.1 (286)	668.8±60.1 (124)	298.9±122.0 (6)	0.96	0.88	0.43
rs17092878	Intron	A	555.1±50.7 (120)	719.7±51.1 (198)	661.1±66.8 (98)	0.37	0.18	0.92
rs9646165	Intron	A	631.5±45.1 (196)	654.9±51.0 (165)	563.3±102.8 (30)	0.66	0.83	0.54
rs1962795	Intron	A	684.0±38.3 (319)	587.3±60.9 (93)	272.8±136.4 (4)	0.47	0.56	0.38
rs1968106	Intron	G	665.7±33.8 (388)	557.7±105.4 (28)	.	0.70	0.70	.
rs1968105	Intron	G	669.9±34.1 (385)	516.2±92.7 (31)	.	0.55	0.55	.
rs1778222	Intron	A	672.5±35.2 (365)	557.6±78.1 (51)	.	0.54	0.54	.
rs11159042	Exon (Gly76Gly)	A	667.6±55.1 (147)	659.8±46.5 (201)	634.6±77.0 (68)	0.64	0.75	0.65
rs45559033	Intron	G	716.8±66.0 (118)	631.2±43.8 (208)	644.8±68.0 (90)	0.45	0.39	0.74
rs11621087	Intron	T	677.2±35.5 (364)	536.1±75.1 (51)	.	0.46	0.46	.
rs78870464	Intron	G	678.2±36.0 (355)	543.4±69.6 (61)	.	0.49	0.49	.
rs17782239	Intron	T	687.6±39.9 (297)	606.7±57.3 (112)	286.3±116.9 (6)	0.61	0.75	0.40
rs6574151	Intron	A	694.1±51.2 (184)	608.2±45.2 (181)	708.1±99.2 (51)	0.62	0.48	0.98
rs8005140	Intron	A	684.0±38.3 (319)	587.3±60.9 (93)	272.8±136.4 (4)	0.47	0.56	0.38
Haplotype								
ht1			589.1±49.1 (144)	724.9±52.4 (191)	835.6±69.4 (81)	0.55	0.30	0.87
ht2			654.7±44.7 (215)	665.3±51.8 (165)	648.8±108.1 (36)	0.74	0.81	0.73
ht3			669.5±34.9 (367)	584.0±84.3 (48)	168.0 (1)	0.70	0.74	0.72
ht4			650.9±34.0 (366)	735.5±106.2 (48)	186.6±131.9 (2)	0.88	0.76	0.47

[0123] 선형 회귀 분석의 P 값은 공변수로서 연령, 성별을 조정한 공동-우성-열성 유전자형 모델을 사용하여 계산되었다.

[0124] M/M, M/m, 및 m/m은 각각 보편적 대립유전자(major allele)에 대한 동형접합자(homozygote), 이형접합자(heterozygote) 및 희귀 대립유전자(minor allele)에 대한 동형접합자(homozygote)를 나타낸다.

[0125] 굵은 수치는 P<0.05의 통계적 유의성을 나타내며, a.a는 아미노산, n은 피검사자의 수를 의미한다.

[0126] 총 IgE과의 연관성 분석 결과인 표 5를 살펴본 결과, SNPs와 haplotypes와의 유의성이 있는 연관성은 관찰되지 않았다(P> 0.05, 표 5). 이를 통해 아토피 피부염의 임상지표 중 하나인 인 총 IgE에 대해서는 유전적 연관성이

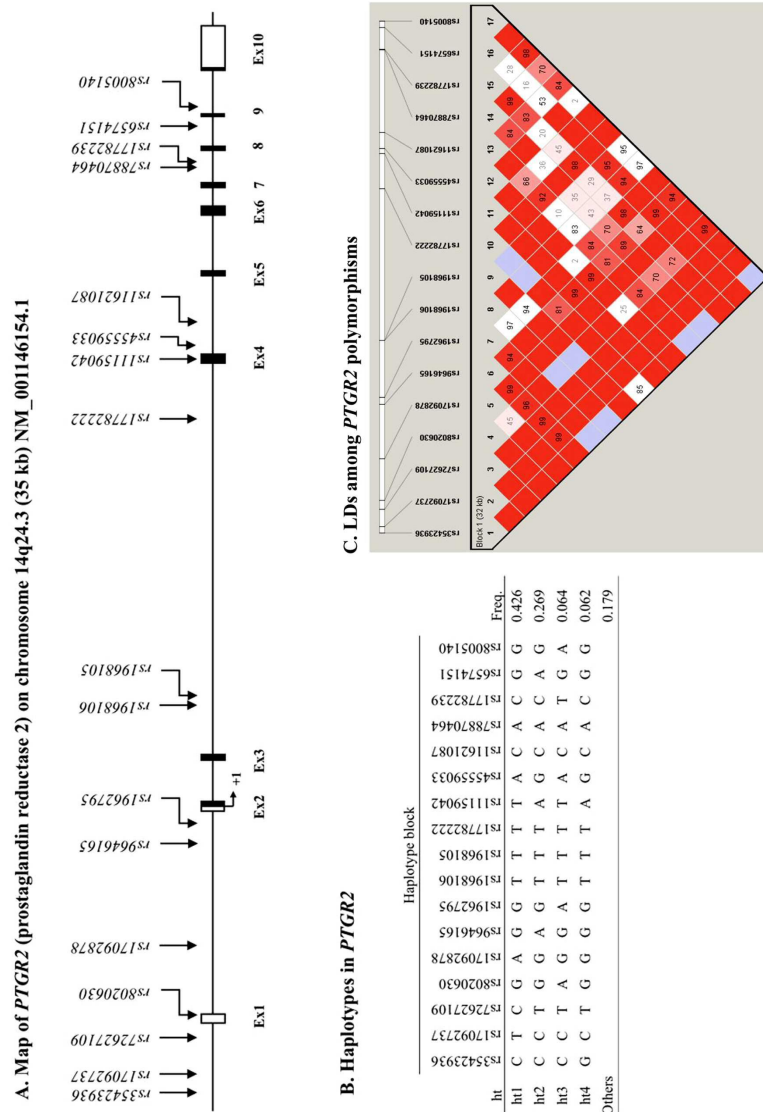
나타나지 않음을 확인하였다.

[0127]

이상으로 본 발명의 특정한 부분을 상세히 기술하였는바, 당업계의 통상의 지식을 가진 자에게 있어서 이러한 구체적인 기술은 단지 바람직한 구현 예일 뿐이며, 이에 본 발명의 범위가 제한되는 것이 아닌 점은 명백하다. 따라서, 본 발명의 실질적인 범위는 첨부된 청구항과 그의 등가물에 의하여 정의된다고 할 것이다.

도면

도면1



서열 목록

- <110> THE ASAN FOUNDATION; University of Ulsan Foundation For Industry Cooperation
- <120> Composition and method for diagnosing atopic dermatitis
- <130> HPC8039
- <160> 1
- <170> KoPatentIn 3.0
- <210> 1
- <211> 35635

<212> DNA

<213> Artificial Sequence

<220><223> Homo sapiens prostaglandin reductase 2 (PTGR2) sequence :

hg38_ncbiRefSeqCurated_NM_001146154.1, chr14:73849831-73885465

<400> 1

tgtggtggca tgcacctgta gttccagcta cttagtaggc tcaattgctt gaacccggga	60
ggcagagggtt gcagtaggcc gagatcgtgc cactgcactc cagcctgggc gacagagaga	120
gactccgtct caaaaaaaaa aaaaaaaaaa aagaaaagaa aaaagaaaa aaagaaatg	180
attctcagga attgcttgaa cgatgagaag gaggcagtac aaagtctgc aaagaggag	240
gagtagggaa ttccagattt atggcaata taacaaaggg tcaggctgctt ggactcaggc	300
aggataactt gtccttttac ttattagtta cttaacctga cttagaactt aactagctta	360
acttcggttt cctcattcgt aaaatggtaa taataataat agggctgttt tgaaaattaa	420
atcatgtagg caaaaaatca tataaggcat tatccaactc aaactacgca ttatgtaaat	480
gatatgaatc cttagtgta gagaaaattt agaagaagag gataaactgg atagaaatgt	540
acctataaat cctctactca tctcttgaat ttttgggaac tcattttggc gagaaaaatc	600
acacttgta caaggagtgc acagacccta gaatcaagag agctaaaact aattaatata	660
catacatata tttaaaattt ttgagacaag gtatcagtct gtcacccagg ctggagtgc	720
gtggtgtgat caaggctgc tgcagcctac tgagcagctg agagactacc agcaagcacc	780
accacaccta gctaatttca ttttattttt tttttttga agaaatgcgg tttcaccacg	840
ttgccaggc tggctctgaa ctctgactc aagcaatcca cctgcctctg cctccaaag	900
gtctgagatt acaggcgtga gccactgcgt ccagcccaa aactaattta atattgcagg	960
gttcttttc tattttgaaa gtgtttcttt taaatcaagt tctagattac ctttatttta	1020
tactactact tttgtctat ttgagcattt ttttttcat catttggttt aagtcattac	1080
aaaaggcatt gaaggctgt ctacctgtcc atgacagata aaagagatta gctcctgtct	1140
gaaaatgttt gcaaagtagt gtaattctcc tagtccttcc tgaagctttc ccaaagaagc	1200
aaaaatatc agacataaag ctaccaggaa ctccaacctg ttgatcact attccttgc	1260
tcaccaacct agtcatttt cagctactcc atgcatgact tgtctgagc tattagctcc	1320
cacatgaata caagcaattc ttatactgc ctaattcaga tgatctttct tagggtttga	1380
agtccttggg gatccgaagc tgtcctaggt atacattct ctcaaccag tccagcacia	1440
aaggagacga ctaacgatga atgaaattga gaaatcagcc cacttggtat atcttgagg	1500
aatttggtat tcttcgtgaa ataaaagagg cgcaaagggg ccctggatta gggtccggag	1560

agaggggttt tatccagca ctgcccttaa ccagctgaaa ggtagggtgt ccaatcacat	1620
cagctctccg agactcagtt tccccatcta taaaacaagg agtaggcttg acggtccctg	1680
tctcactata atttcccaat cgaggttcag ggagaaatcc aaaaaactaa taaaacttct	1740
tcacggactg ctaaattgta actcactggt caaggctgca agatttgctc tttgggtcaa	1800
ggcagtaggg gcggccacaa ggaaccaatc tcagttaagt tattaatgca ctttgcccag	1860
actagtaact ctccccgac gttecgcatt tcccttgctt ctcccggaaa cgtggggcgc	1920
ccaggaatcc taaacttcca atccctaagc ttccaaacc cggccctggg aggggagggg	1980
gcgcggagcc gggcctagtc gcgcgcctgc gtactgcggt cagagggccg aggcctgggg	2040
gcgagctggg gtcgtgcagt acagcctctt tccggcaaat cacgcgagat ttcgttcacc	2100
cgggctccac agggtcagtg acgcgcgagc ggctcccgcg actgtggctc ccgaagctgc	2160
ttcgggcccc agcggtaggg aagaagactc tcaaccaggg tgacttctag gcgtctctcc	2220
ggagctaccg gggacactgg cgttattgcc cgcaacactg agataggga tgctcgagaa	2280
acagaagcag gttttgggt ttttttgaa ttactggggg aaagcaacct gaggtttga	2340
cgtaattgtg tttaggtgc ctctgaaact ttcatttaac acagatttgt tgagcacgta	2400
ccatgcgccg ggcatagggg attggactca catTTTTTTg tgtgtgtgac ggagtctcgc	2460
tctgtctgtc gccaggttg gagtgcagtg gcgcgatctc ggctcactgc agcctccatc	2520
tctcgggttc aagcgattct cctgcctcag cctcccgagt agctgggatt actactgcgc	2580
ccgataatt ttttgtattt ttgtagaga cgagatttca ccgtgttagc caggatggtc	2640
ttgatctcct gacctctga tctgcctgcc tcggcctctc agagtgttg gattacaggc	2700
gtgagccatc gcgcccggcc tggattcaca tcattgaaca aagcagacga gtccttatcc	2760
tcttagcgt gacgagctgg gcagttagtt tgaaattctg ggctagagac gcaaattatt	2820
attaatttcc atataaagtg gttgtagga ggagccacag ggctcaggga ggtcactgag	2880
aaaggaaga gcaagatagg aatctccaca tgcaagactt cgggagggga agggggagac	2940
cagaaaaccg gaaggggagg agaagactgg cagagagatc cagaaagaca cgtggtttgt	3000
ttgtttgttt tttagacaa agtctcgtc tgttgcccag gctggagtgt cctggctgga	3060
tctcggtca ctgcagctc agtctcctgg gttcaggcag ttctccagct tcagcctcct	3120
gtgtagctgg gactatagc gccaccacc acgttaatt tttgtatttt tagtagagcg	3180
gggttttccc atgttgcca gactggtctc gaattcctgg cctcaagtga gcctcccacc	3240
tcagcctccc aaagtgttg aattacaggt gtgagccacg gctctgggcc cagaaagata	3300
cgttttataa caaggcgtgg tcaacagggt tgctagtggc agaataatcg cacgaaataa	3360

gggtattaaa aggtattagg aagctcatgt gatggaaata gcaggtgccg tgttgagta	3420
tgggttctag cgttattaca agttgggcga ctctgggcga attattctct tctgtacct	3480
cttttttctt ctcttttttc ttttttttg agatggagtt tcgctcttgt cgcccatgct	3540
ggagtgcaat ggcgtgatct cggctcacig caacctccac ctcccggttt aaagcgattc	3600
cccagcctca gcctccagag tagctgggat tacaggcgcc cgccaccacg cgggctgat	3660
ttttgtattt ttagtagagg tggggtttcg cgtttttgtc ggagctggtc tcaaactcct	3720
gacctcaagt gatctgccc cctcagcttc ccaaagtgtc gggattacag gcctgagcca	3780
ccatacccg ccccttttct catctataaa attgagagga tgatattacc aacctcacag	3840
tttttttggg gtgaggggtgg gggaggtgggt ttaaacaggc tagtgaatga gtctgtgctc	3900
agttcctggc atttaggaat attaaggaaa caggttgaga gaagggactt tttccaagtt	3960
cccacctggt tagaactagg agaagtactc tgttcttctt acttaaatga cttttccctt	4020
ttaccaggct gcaaggtaaa cactcaacac actctatcat cacatcggag caaatcgtc	4080
tgttattatt ttggaatatt agcaccaagt aggaatagta tcccaggcaa ataggatttg	4140
ctgagttttt ttccctact tgggatgaat ttaacaatct accttcttt cagatcttgg	4200
aacttgatgt gtcttatttt agctgttaat tttatgacta aatctaggta tctatatttt	4260
atttattttt attatttatt ttttttttg agacaaagtt tcgctcttgt cccccaggct	4320
tgagtgcaat ggatcagttt cggctcacig caacctccgc ttcccagggt caagtgattc	4380
tcctgcctca gcctccggag tagctgggat tacaggcacc cgccaccacg cctggttaat	4440
ttttgtattt ttagtagggc agagtttcac catgttggcc aggetggtct caaactcctg	4500
acctcaagtg atccacctgc ctacgcctcc caaagtgcta ggattacagg tgtgagccac	4560
cgtgcccagc caagttaat atatttgaaa tgtgtgtttt tttgtacta ttgagggata	4620
tcacacaatt atcctatatt aattgtagta ttgacactt gaaaagtttc taatttaatg	4680
cagagtgaga gaactttaat agtatatttt tcatgctgat aagcagtgtt tggacaaatc	4740
tataatatac atattaatat gttaaacaaa taaagcatac cattgacagt aattagaggg	4800
caaatatgat atttcatttg gaaagactgc ttgaaaaatt tagccattat ggttctatta	4860
ctacttaagg acttttagac ccacttaatt tggaaaactt tggcttttta atccttatat	4920
aattgacaca tatgggtgatt ctgtagaaaag ttctcttttt aaaacatatg tagccctaca	4980
tgacattatg gacctcttgg tgtgatgaat ggatgtggca ccaccttaa agattcttgc	5040
ctgaaaaatt gaacctgaag taatcaagcc tctagatgta actagcagtc atggataagg	5100

ggaggagata gaagaaacat ttccttcttt ctttaaaagc tagtcattta ccttgtgatg 5160
 glaaatgaca ccatgaggaa gcaatcaaag aacattctac aagcaaagta atgaagaaac 5220
 aaatcaaggg catgaaataa atggaaggga gactcaaact gtagtatgtg gaccttgaat 5280
 cctgatittha atgaaccaac tgtaaaaaga catctttgaa acatttttga aaccaactgt 5340
 aaaaagatat ctttgaacat cctgaaacgt caggaaattt gaatatgcac taggtattag 5400
 atgatgataa ggaattgcc a gtactactat taggtgtaat aatggcatag tgtttatgtt 5460
 taaatctgtg atggctgggt tatagtcatt tgccatataa tgacggttca gacaatgaca 5520

 gacgatgagc ctcccataac attataatgg aattgaaaaa ttcctcttgc ctggtgatgt 5580
 catagtcac agaacgttat agtgcaatta ctttatitit attatttatt tattgagaca 5640
 gggctctact ctgtcaccca gactggagta cagtggcgtg atctcggctc accgcaacct 5700
 ccacctcttg agttcaagca attctcctgc ctgagtcctt gtagtagctg gattacaggc 5760
 gtgcaccacc atgcctggct gatttttgta tgttttagtag agacagggtt tcaccatgtt 5820
 ggtcaggctg gttgtgaact cctgacctta aatgatgcac ctgcctcagc ctcccaaagt 5880
 gctgggatta caggcatgag ccaccacgcc cggagtatta ctttatitit taagaaattt 5940

 agtggggccg ggcgcgggtg ctcatgccig taatcccagc actttgggag gccgaggcgg 6000
 gtggatcatg aggtcaggag tttgagaaca gcctgaccaa catggtgaaa ccccgctctt 6060
 actaaaaata caaaaattag ccgggcatgg tggcacgcac ctgtaatccc agctacttgg 6120
 gaggctgagg caggagaatc gcttgaatct ggaaggcaga ggttgcagtg agctgagatt 6180
 gcgccactgc actccagcct gggcgacaga gcaagactct gtttcaaaaa aaaaagaaag 6240
 aaatttagtg tagtctaagt gcatagtgtt aataaagtct aagtagccta cagtaatgtc 6300
 ctaggctttc acattccctc accactcact cctgactca tccagagcaa cttctaactt 6360

 ggtaggctcc attcatggta agttacccta tataggtata ccattttaaa atcttttttg 6420
 ttgtttttt ttttgaggcg gagtttact cgtcacccag actggagtgc agtggcgcaa 6480
 tcttggtcct ctgcaatctc cactccagg gttcaagcga ttctcctgcc tcagcctccc 6540
 aagtagctgg gactacaggc atgggccacc tgcccaccgg ctattttttt ttgtattttt 6600
 agtagagaca ggatttcac atgttagcca ggctggctc gaactcctga cctcaggatga 6660
 tctgcccccc ttggcctccc aaagtgtggt gattacaagg gtgagccacc gctcccagcc 6720
 catttttaaa tcttttacc catattttta ctgtactttt tctatgttta gatacacaaa 6780

 tacttaccat cgtgttccag ttgtctgcag cattcagtag agcaacatgc catataggtt 6840
 tgtaccgtag aagtataag ctaaacata tcacctaggt gtctagtaat ctgtaccatg 6900
 taggtttgtg taagtatact ctgatgttca cacaatgatt aaatcaccta atgtcacatt 6960

tctcagcatg tatcccagta gtcaagctgt acaaagcact cacaagagtg cttggcacat	7020
agtaacacta agggcttgat tttctaattt tctactgttg ataaagtaag caaagtgatt	7080
ttcagcaagt cagatagtc tagtactgga agttcttttg ctttcttcat tagctcaaat	7140
tctaaatcca gaataaactg gccagatgta gtggctcacg cctgtaatcc cagcactttg	7200
ggaggtcgag atgggaggat cacttgagcc caggaattcg agaccagcct gagcaacata	7260
gcgagacctc atctcaaaaa acaaacaga ataaactgat cctttactta ttggccatat	7320
tgtgatatgg aatttcctat aaggtaagga tttttaatt ccttgatata ttacttttat	7380
tgtcagctaa actataaagg aatttcgtt ttttttttt ttaattcacc actgataaga	7440
aggcaaagga atttccattt tgatgaacga atagaagagc ctttgaaaat ttggttacgt	7500
ggctgggctt gggtgctcat gcctgtaatc ccagcacttt gtgaggccaa ggtgggcaga	7560
tcacttgagg tcaggagtgc aagaccgcc tggccaacat ggtgaaaccc tgtctctact	7620
aaaaatacaa aaaattagca gggcaccgtg gtgcacgcct gtaatcccag ctacttggga	7680
agctgaggca ggagaattgt ttgaaccag gaagcagagg ttgcggtgag ctgagatcgc	7740
gccactgcac tccagcttgg gtgacagagt aagattctgt cccccactcc aaaaaaaaaa	7800
tttttttgat tacatgtcaa ttaagcagtt agtgtttttt ttttttttt ttttttgaga	7860
cagaatctcg ctctgttgc caggctggag tgcagtgggt cgatctcggc tcaactgcaag	7920
ctccgcctcc cgggtccacg ccattctcct gcgtcagcct ccggagtagc tgggactaca	7980
ggctcccgt accacgcctg gcttattttt ttgtattttt agtagagacg gggtttcacc	8040
atgttaggca ggatggtctt gaactcttga ccttgtgatc caccgcctc ggcctcccaa	8100
agtgttggga ttacaggcgt gagccaccgc gccagctgc agttiagcgt tttattttgc	8160
cactagatta actgctagct ctggaggaat atatcatgtt gtttgtttta tgttttttc	8220
ttgtctttat ttaaaaacaa gtgcagtatt gagaaggga aagagtagaa caaggggttt	8280
ggcctgtaac tgactgaaca atcaggtgat ataactcact acctttggac cagccttgtg	8340
ttgttttctt ttaatcgagc cataggagc ttaatcatga tgggtgtcct gaagatgaat	8400
gagcttgaat tatctgagac caaaaacatc ttgcccact agggccgggt gtggtggctc	8460
atgcctgtaa tccaacact gggaagctga ggcgggcaga tcacgaggtc aggagattga	8520
gaccatcctg gctaacacag tgaaccccg tctctactaa aaatacaaaa aaattagctg	8580
ggcatggtgg caggcgctg tcgtcccatc tactcgggag gctgaggcag gagaatggcg	8640
tgaaccagg aggcagagct tgcagtcagc agagatcaca ccactgcact ccagactggg	8700
caacagagcg agactccgtc tttaaaataa ataaataaat aaacaacat ttgcccac	8760
ttaatcttca atgaacttgt tcctattagt taaaaattct aactagattt agtctttctt	8820

catatatttg ctttaggagg caaactctga caggcgagtt tatctttttg gatcctaaat	8880
gigatttttc tccctcttca ttgggaaaca caaatggact tcttttttat ttttaaattg	8940
ttgaatactt caaatcttgt gatttaaaaa tttttttatt tataggagta ttttatacag	9000
aaatcttgtg aaaccactgc ccaaccagag caatgattgt tcaaagagtg gtattgaatt	9060
ctcgacctgg tatgtatttg cgatgaatga acaactctga tctttgataa gtattaatgc	9120
aaatggagga ataaaaacta atgtgatgtt catgttgtaa attaaggtag taagatattg	9180
acactagtaa agatattaga agaaactctt taaacgatgt aataattttc aagtgttagc	9240
ttatcaagac cattttttaa cagtgtgcta gtttttcaag tgataactta tccaggcatt	9300
tgttgatgaa ataaggtatt ttggaatcat aatgtctgaa aggttacatt acaaataat	9360
giggcttttg tttctttaga gtatgttatt attggtcctt aaaatttgct tttctttttt	9420
atgtgaaatg tactaacata tcagtgtgta catgtaaact gcatcataaa aacagatatg	9480
gaaagaacct gagagatttg ggtcactgct gccacttga ggaagtcact taacttttct	9540
gaactcaatt ttagtaaaat gagcagttag attaagatct ttatgtacta ttcccacctt	9600
ataaatttaa caattctatg aagagtaaatt ttgcattaaa catatctgaa tatgagtacc	9660
ttctagtggg aaccttttct tcttacaat ttttctcata cctcaaattt ttctaatttc	9720
ctacttecta ttttttcaga cttcattatt attactttta tattcgttta agtttctata	9780
atgatgacct ttccittaca atttttctca tgcctcaaatt ttttaattt tccacttttt	9840
cagactttat tattattact tttatattca ttttaagtct tatttgcttc ttaaaataag	9900
gaaggaaagt ttagccgcgc tccatcctgg gtgacagagt aagaccttgt ctctaaaaaa	9960
aaaaaagaaa tacatactga gtcattttct tagttacaga agaattaca aacaatatag	10020
atttacttgt gcctgttttt ttttttttga cacggagtct cgctctgtca tgcagtggca	10080
caatctcggc tcagtgaac ttccacctcc cgggttcaag cgattcttct gcctcagcct	10140
cccagtagc tgggactaca ggcgtgcgcc accacgcctg gctagttttt gtatttttgg	10200
tagaatcaag gtttcacat gttagccagg ctggtctgaa actcctgacc tctgatcca	10260
cctgtttcag cctcccaaag tcctgggggtt acaggcatga gccaccacgc ctggccctt	10320
gtgcctgttt taaaaggaac tgttctgaat cctagagtga ccttttaaag aataataagt	10380
atatcccacc actatgggag gccaaagtgg gcagatcgtg aggtcaggaa tttgagacca	10440
gcctggccag catggtaaaa ccccatctct actaaaaata ttaaaaagta gccgggcatg	10500
gigtgcatg cctgtaatcc cagctactcg ggaggctgag tcaggagaat cacttgaacc	10560

caggaggcag aggttgcagt aagccgagat cgcgccactg cattccagcc tgggtgatag	10620
agcaagactc tgtctcaaat aataataata ataatagtat atagcaaagt ggcttttttt	10680
aactttatat ttttgcataa tatttttagga aaaaatggta atccagtggc agagaatttc	10740
cgaatggaag aagtctatit accagataat attaatagaag gacaagtaca agttagaact	10800
ctttatcttt ctgtggatcc ttacatggta agaatacagga tgttgtaact tggtgaaaaa	10860
taactttatt ttattatitit tcatgttgta tctgaatctt attgtgcagt cccttataga	10920
ttggtaaata ttigagcagt tataatattga taccttttaa ttatagattg cctttgtttt	10980
tattcactag tgtaattaag gaatgttgta gagggaacag ataatagtag tacatgtatg	11040
tgattcaggt gccttcaaat gttcagatta ctccccagta gtaataaag ctgaagata	11100
glaaaacagg attgcaaatt cagtaaataa agctttaaga tagtaaaata agattgcaaa	11160
ttctttcaga tccactttat aagagaggct atgtatgaga tctttattta tttctagttg	11220
acctcctttt cttttctttt tctttttttt aagcgatgag gtctcgctgt gttgcccagg	11280
ctggtctcaa ttttctggg ctcaggcaat cctctgactt tggcctcctt gaagtgtgg	11340
cattacaggc atgagccact gtgcttagcc ctgtagtga tctctttttg aaaggaacag	11400
actcacttat gctacttcta gataaataag atttagtgaa aacgtataca ttgctgaaga	11460
ctgcagataa gaagccatta ggcacaggga ctttggctct acgagctggc gtcttgcttt	11520
acttctccca gaggtcatac agtcttttct ttacttctaa ctttttcta tggccatgac	11580
aagcataact ctttatgtca atttaccctt tgctattaag aggtcagct gggcatggtg	11640
gctcttgctt gtaatcttag cttttttgga ggccaaggca ggattgcttg agcccaggag	11700
ttcaagacca gcctggacaa gagagtgaaga cccagtctca acaaaaaatt aaaaatttag	11760
acaggtgtgg tgggtgtgtc ctgtagtcct agctactga gaggtgatg tgggcagatc	11820
gtttgagcct gggaggttga ggctgcattg agccatgatt gcgccactgc acttccgttc	11880
gggtgatgga gtgagaccct gttgcaaaaa aaaaagtctc agtgcttgga atcagattgg	11940
tgcatgagg aaattggccc catgtgtctg tttgttaac tgtttcatag gttgctgaac	12000
gactgtatga gtttctgtt gctgttttga catagtactg taaacttagt ggcttttttt	12060
tttttgggac ggagtctcac tctgttcctt aggtcggagt gcagtgccac tatcttggt	12120
cactgcaacc tctgctcctt gggttcaagt gattctcttg gctcagcttc ccgagtagct	12180
gggattgcag tcttacacca ccacaccag ctaatttttt gtatttttag tagagatggg	12240
gtcccaccag actggccagc ctggtgtcaa actgttcttt ttttttgtt atgtagaggc	12300
tgattatctt ttacattatt gaagacattt ttaaaaaatg agcaacttta ttatttttat	12360
taacttttgt tgttgttgtt tgtttgtttt tttttgagac agagtcttgc tctgtcgccc	12420

aggctggagt gcagtggcgc gatctcggct cactgcaagc tccgcctccc gggttcacgc	12480
cattctcctg cctcagccac ccgagcagcc aggactacag gcaaccgcca ccacgcccgg	12540
ctaatttttt gtatttttag tagagatggg gtttactgt gttagccagg atggcttga	12600
tctcctgacc tcgtgatcca cccgcctcgg cctcccaaag tgctgggatt acaggcgtga	12660
gccactgcgc ctggcctatt tttatcaaca ttctaataag ttgaacaga agtatataaa	12720
gtaggaagtg aaagagtgcc cgtacaaagg atcgcttatt gatctgctgt ctaggtaaca	12780
gcaatgtacc aatgttaatg tctgctttt gatattatat tgcagctatt aaagacgtca	12840
ctactcgggc aagctgagtg aaggactgtg gacttcattt tattaatttt taactttttt	12900
gaaatgagat ctctattgct gaggctgagt gcagtgggtg gaatacagct cactgcagcc	12960
tcgacctcct gggctcaagc agtcctccca tctcagcttc ccaactagca gggaccacag	13020
gtatgcacca ccacaccgg ctaatttttt tgtttttgt agagatgggg tcttgccatg	13080
ttgcccaggt tggctctgaa ctctggact caagcaatac tcttatcttg gcctccaaag	13140
tgctgagatt tcaggtgtga gccactacat ctgtctccat atacttaaaa aatttaacag	13200
ttttattata ttcacacact atacaattta ctcaactaaa gagtacaact aaattatfff	13260
tagtatattc acggaattct gcaactgtca ttacaatcag tttacaaca ttttcatccc	13320
cagaagaaac cctgaaccta tgggtataatt agcaatcatt cccattttcc cctcctcacc	13380
ccacctctg ctgtatacaa ccagtaatct actttttgtt cctatagatt tgcctattct	13440
ggccatttca tataaataaa aacatacaat atgttattct tttttgtgt ttttaagaga	13500
cagagtcttg cactgttgcc caggtctggaa tgcagtgggg caatcatagc tcacggtgc	13560
cttgaccttc tgggctcaag caatcctctc acctcagctg agagctcttg agtagctgag	13620
accataggtg tgtgccacca tgcccggcta attaaaattt tttaaaaaaa cattttaaaa	13680
ttgtttatff atgaatcatt gcaaagaata aactgttatt tattatttta tttagagatg	13740
agttctgttt tatcttttgt catttatact ctcatacttt caggctctct ttttttttt	13800
ttggcagaca gtttactct atccccagg ctggagtga ggggctcaat ctgagctcat	13860
tgcaacctcc acctcctgga ttcaagagat tctcatgcct caacctccg agtagctggg	13920
attaacaggc gtgcaccacc acaccagct aatttttgta ttttagtag agatggggtt	13980
tcaccatgtt ggccaggtg gtctcaagct cctgacctca ggtgatccac ccgcatcagc	14040
ctctcaaagt gctaggatta cagggtgag ccatcacacc tggccgactc atatctaaga	14100
aggcttttcc ccaccaagg tcacctcttg ctatgagaca gccaagctg gtcttgaatt	14160
ccagggatcc tctgacctca gcctcctgag tagctggggg ttaggtggg agctactgtg	14220
cctggcataa tatatgacct tttgtgtctg gcttctttca tttagtatgt gtttacaagg	14280

ctcatcatgt tgiagtttat acttcttttt attgctgaat aatgttacat tgtatggata	14340
cagtatctta tctgttcac agatgatgga catgtgggtt gtttacaca cttggctatt	14400
agcaatactt ctgctatgaa catttgtgta tatgtttttg tgtggactgt gtttttggtt	14460
ctcttgggta tatacataag agtaaaatta ctaggtcata cagtaagtat atgtttaacc	14520
ttttgaggaa ctaccgaact attttccaaa gtggctttac cattttccaa tcccaccatc	14580
agtgtgtgaa ggttccagtt ttccacatcc tcactaacac ttgttattat ctgtcttttt	14640
ggttatagcc attctagtgt gtgtaaagtg atatctcctt gtggttttct tttgtatttc	14700
cttatggcta ataatgttga gcatgttttt aatgtgctta ttggtcattt gcatgtcttt	14760
tttgagaca tgictattca gatttcttgt gcattttaaa agttggatta tctgtgtttt	14820
tattactgag ttgtaagagt tgtttatata ttcttgatat aagcttcttg ctggatgtat	14880
gatctataaa tattttcttc tatgctgtag gttgtctttt tattctcttg atagtatcca	14940
ttgaagcaca aaagttttta attttgataa agtccgattt atctcttttt ttgtttgtca	15000
cgtatacttt caatactttc agactcataf ctaagaaggc ttgacctcac ttaaggfcat	15060
gaagatttag tegtatatit tctgctaaga gttttatagt ttgtctctaa catttaagtc	15120
agtaacacat ttggggttaa actctatgta tgcgtatgag gaaggggtct gtattaatca	15180
agattcttta gagagacaga aacacaaaca cacacatagg catgcacca cagacatgca	15240
cacacatata tatgagaggg gacttataag aggaattggc tcacaagatt atggaagctg	15300
agagttccca tgataggctg tctacaaact ggagaaccag agaagctggt agcatggctc	15360
aatccaagct ggaaggcccg acaaccaagg aagccattgg tatacttgtc agtccaagg	15420
cctgagagcc caggaagcca tgggtgcagt ctgaggtac aaaggctgaa gaacctggag	15480
ttcttatgtc agagggcagg agaagaaggg tgtcctgctt tgcaggagag agagagaatt	15540
tgtccttcac ctgctttttt gttccatctg ggtccttgt tgaatggatg gtgccaccc	15600
acattgaggg cagatcttca ccacatagtc cacagactca cctgctagtt acctctgaaa	15660
acacctcaa ggacatatcc agaaacaatg cttcacaagc catctaggca ttcctctatc	15720
cagtcaagtt gacgcctaaa attatccatc acagggccca actcattctt ttttacaatc	15780
tcaaactccc tgggctcaag tgattctcct gtctcagact cccaagtaac tgggacgata	15840
ggtgcaagcc accacacctg acctaaactc actcttttgc aagtgaatat acagtcgtgt	15900
cagcactttt ttttttttta agagagacca gaaatatttg accacgaagg gacatgaacc	15960
agcactgttt gttgaaaaga taattctttc ctcatgtgtt tgtcttagta ctcttgttga	16020

aaaccaactg accataaata tgagggttta tttctggact ttcaactcta ttctagcaat	16080
ctgtatatct gtcatgctag tacctcgtc tcttaattac tgtagctttg taataagttt	16140
tgaaatgagc aagtgtgagt cttccaatct gtttttcttt tttttttttt gagatggagt	16200
cttgctctgt cgccagctcg gagtgcagtg gcgctatctt ggctcactgc aacctccact	16260
tccccgggtc aagtgggtcc cctgcctcag cctcctgagt agctgggact acaggcaccc	16320
accaccacac ctggctaata ttttgtatit ttagcagaga cggggtttca ccatattggc	16380
caggtgggtc ttgaactcct gaccttcagg tgatccaccc acctcggcct ccgaaagtgc	16440
tggtgattaca ggcatgagct actatgcctg aacaattttt tgtattttag tagagacagg	16500
gtttcaccat gtggccagg atggcttga tctcctaacc ttgtgatcct cccgcctcag	16560
cctcccaaag tgctgggatt acaggtgtga gccacccac ctggcctgtt tttctttttt	16620
aagattgttt tggtctctct gggtcccatg aattttcata tgaactttag aaccagcttg	16680
tcagtttctg cacagaagat ttgataggg atttgatag ggattgcttt gagtctgtag	16740
atcaatttgg ggagtattgt cttctgtgat ggttcatttt atgtgtccac ttaactaggc	16800
catgggatgt ctaaatagct ggttgaacgt tatttctggg tgtgtctgca aggggtgttc	16860
tagaagagat tagcttttgg attggtggac tgaaaaaagc agatgtcctt ctctaagtg	16920
ggtgggcac atccagtcca ttgagggcct aaggagaaca aaatgtagaa gaaggttgct	16980
gggtgcggtg tctcacact gtaatcccag cactttggga ggctgaggcg ggcagatcac	17040
ctgaggtcag acgttcagga ccagcctggc caacgtggtg aaaccccgtc tctactaaga	17100
atacaaaaat tagccgggtg tgggtggcagg cacctgtaat cccagctact caggaagctg	17160
aggcaagaga attgtttgaa cccgggaggc ggagattgca gtaagccaag atcgcggcac	17220
tgcactccag cctgtgtgac agagcaagac tctgtctcaa aaaaaaaaaa aaaaaaaca	17280
aaaaaacaaa aaaaagaaga agaagaaggt tgaactaact ctgactgctt aagctggagc	17340
atcattcttc tctgtcctc agagttcctg gtctcaggc cttcagatct ggactagaat	17400
ctctactgtc agcttacctg ctctcattca ctggaaccac actgtggct tttctgagtc	17460
ttcagcttgc agatggcaga taatgggact tcttagcctc tctaactgca tgagccaata	17520
ccttataata aatctcttca tagatatgta tgtatcctgt tggttctgtt tctctggaga	17580
attctgatga aaacaatate ctaatacact gatccatgaa tgcatgatgt ctttctactt	17640
atttgggtgt tcattaatit ctttctgtct ttaaaaaaaaa aagattttag gggaggaaaa	17700
atgatggaac aactaggatt ataataaaa ttgaattaat gatcttcact tcgaacactc	17760
ccaacttcaa agcaaaaaa aaaacaaaa aacctctttt tctccaggg aagtaaatat	17820
tttgtctact tactcatttc atgctagaaa gcacagtctt gagaattttg cctctaccag	17880

gtcctgtgga tccacacctc taaataccctc tcaattcagt ccatgtctta gtaattccag	17940
ggaagtctac ctagtctgaa tcaccgttgt ttttttcta atggggtagg tacccttgca	18000
tacacacttt gcccccttaa gccatttttc aactgcagc cagaatgcca gagtgttct	18060
tcttataaca gcaactaatt gtgtcacctt gcttaaagtg gtttaatggc ttccattag	18120
ttttggaatt gagaccaatc agccttgaca tggcttaaga agctacataa cctgggctag	18180
gcacggtggc tcacgcttgt aatcctagca ctttgggagg ccgaggtggg tggataacct	18240
gaggtcagga gtttgacacc agcctggcca acatagtga accccgtctc tactaaaaat	18300
acaaaaaaat tagccaggcg tgggtggcaca tgcctgtagt cccagctact tgggaggctg	18360
aggcaggaaa atcacttgag cccgggaggc cgaggttgca gtgagccgag atcacgccac	18420
tgcactccag cctgggcaac agagcaagac tctgtctcaa aaaaaaaaag aagctacata	18480
acctgactcc tttacctgtc caggctcaga gaagcctgcc aattaccctc acactgttgc	18540
ctcagtggat gtctgtcttt tgtaacacat cagtgcctcc tctgtctta tctttgtatg	18600
aatcatgttt ttaatttagt gtattttagg atgttttgac atcttagggg gcttgtgaat	18660
cctagagtaa ctgtccctgg caggatttgc ttatttttag agataacaac ttgcatttga	18720
gtcttgccat tcatatacaa accaaccaat ccaaagccca tatccctaaa ctctcctct	18780
atctaactca cactccctg ccttaaatca cctcagggcc agatgctgca caacttaaga	18840
accacccctg tatatcagag ccctctgaca ttattaaagc tgtcatgtaa tatactcaag	18900
gttacctacc ctgccttgcc cgttctttcc tgtggaaacc aaagatactt ctccctaaa	18960
aaaaaaaaag aagactctgg gccatgctct tccttgctcc tttgccctct tgaccaactc	19020
ttgtgcttcc ccatgcgacc ctctctggca tattgtgccc tcccctctcg gcaactgtaa	19080
gaaattcttc tttaaggca gttgtcttcg tatctatcat ttaccatac ctgggttaaaa	19140
cagagtccca ggtacatatt aaagcaagcc ttcatacatg ttggccctct atctaaaagc	19200
ctcttccac tcctttccct ttacctggta atccctgtta ttccttagat gcctgcttta	19260
aagagatttc ctttggtaaa tcacctgaa cctcagact agtcagacc tctctttgat	19320
attttcctct tgacattcag catttatecc aattgaaagt aataattaca tttgtgtagt	19380
tattagatta tctgtcttcc ttagtaaaaa gtaagcttat gggctgggtg ccatggctca	19440
tacttataat cccagcacac tgggaggctg aggagagg atcacttgac cccaggagtt	19500
tgaaccatc ctgggcaaca cagaaagatg ccatcaatac caaaaaagg aaattaggtg	19560
agtgttaagg tgcaccagcc actctggagg ctgaggcggg aggatcactt gagcccgga	19620
ggtgggagga tcaacttgag ccgggaagtg ggaggatcac ttgagcccag gaggtcgaag	19680
ctgtagttag ctgtgatcat gccactgcac tccagcctgg gcaacagagt gagaccgtgc	19740

ctcaaaaaaa aaaaaaaaag tttatgaaaa caaagttcgt atttatcact gtatctccag	19800
tgactagcat aatgtttatt aaatgttaat tgcataaagg aaggaagaaa aatcactttg	19860
ggatactttc actagacatt ctctatgcct caticcttta aaaaaataa tcaggtctggg	19920
cgaggtggct catgctggta atcccagcat ttggggaggc tgaggcgggc ggatcacttg	19980
aggtcaggag tttagacca gcctggccaa catggtgaaa caccgtctct acgaaaacta	20040
caaaaattag ccaggcgtgg tggcatgtgc ctgttatctc agctactcag gacgtgagg	20100
caggagaatc gcttagccg gggaggcgga ggttgcatg agctgagatt gtgccattcc	20160
actgcactcc agcttagtg acagagactc cgtctcagag agaaaaaaaa aaaagaatca	20220
gctgggcatg gtggcttggg cctgtaattc cagctacttg ggagacttga ggtgggagga	20280
tggcttgaag ccatgagttc aagtcagct tgggcagcaa cgcaagacc caactctaaa	20340
aaatagaaaa atcttaagca accttttttt ttttttttt tgagatagag tcttgcctg	20400
tgcagtggca caatcttggc ttactgccac ctctgcctcc cgggttcaag tgattctct	20460
gcctcagcct tctgaatagc tgggattaca gatgcctgcc accacgccca gctaattttt	20520
gtatttttta gtagagatgg ggtttcgcca tgttggccag gctggtctca aactctgac	20580
ctcaggtgat ctgcccgcct tggcctccca gagtgcctggg attacagaca tgagccacca	20640
caccagacct cttaagcagc tttaaaaggg aaaagggagt ctaacaaaca ttaaataatt	20700
gagattttat atcattccaa attcacaact aaaattaatc attatgtgta cacagtattt	20760
tacatgatta tccttaatgt gcatatacaa ctgggcttcc atattctttt tattcttttt	20820
aaaattttta ttttcatttg acagagagga gtgtcttact gtgttggcca gtctggtctc	20880
aaagtcttgg ccicaagcga tcctctgcc tcagcctccc aaagtgggtg gtttacagat	20940
gtgctttcat attctttaag tgactttata tggaaaaact agttcaaata ttattatttt	21000
agtggttaat aacattctgt tctgttgaca tcatttgttt agctataggg aattacttgt	21060
gtgatattgt gatttataat aggaaatatt tggcttctt gctctttgct ggcacacaat	21120
tcctaaagcc cttagaatct ccaaggtgtc tttttgtatt ctaatgagat gacaagtgg	21180
tgggagctcc tggatagcct caggatgggg gctggttacc aaggcaacca accatatgat	21240
cagagggttg gaacttttat tcctaccacc ttacctccag gtaggggaca gcaggtgaag	21300
gttagtatta tcaccaatgg ccagtgggtt aatcaatcat gtctatataa tgacgccttc	21360
atcaaaacc aaaaggacct gattgaggag ttccaaata gctgaacaca tgggtattcc	21420
tgtagggtgg tgcacctgga gaggtgtgg aaactccacg cttttccca tataacttac	21480

cccatgcac tcttctatct ggctgttcat cttttttttt tttttttttt tttttttaat	21540
ttttgttctt tttttgtctt tttaaaaaat ttttattatt tttatttggc tattcatcta	21600
tatcctttgt aatatacctt ataataaaca gtaaagtga gtaaagtatt tctctgactt	21660
ctgtgaggtg ctctagcaaa ttaaaccaaa ggagagggtg gagtggggac cctgatttat	21720
agccagtttg tcagaagcac agaccataac ctagggcttg agattggcat ctgaagtagg	21780
tgcagctttg tgggactgag cctttaatct gtgggacctg aactatatac caggtagata	21840
gtgtcacagt tgaactgaat ggaattaggg gacaccaggt tgatgtcctc tgcagaattg	21900
cttgggttgt gtaggggaa ataccctcc acatctgttt ttagaagtgt tgagtgtga	21960
gtgaaaggag gaaaaagtag atttttttcc tatatctcag agagcttatt tccatttttc	22020
ctactagctg tgtttaagtg cccttggctc ggccagtgtc tatggctgca cttgctttta	22080
aggaatacgt gggaggtggc acacaagcag aagcaaatgt attcttacag aagcttctgc	22140
atcagttctc tgtttcgtct tttacccgtt ttacctgggt acaccgaacc acatgggacc	22200
ccttaatcag gtgtccttcc atgttttcca tttcctggtc cgtgctgtcg tcctctacat	22260
tttggatttc atgatatagg caagagttag ccacagagga agagagcctt gactgtctacc	22320
aagagaagtt gtagccttc ctcttgagat ggggtgcagt cccattgtga cattctgtct	22380
ttgaggggtg aagagagcac ttatctgagg ttagggggca tgaaataggg agcgtacttg	22440
gtctgaagtt gggatgagca actgtataaa attcacttca cttctttttg ggttatttcc	22500
actggggaga cagaattgct agcccaaaag aaatcaagtt ttatgggtga tatgttgcta	22560
gtttgttttc cataaaaaat agttttataa tatcagcaac atgtgagtgt aacctaatg	22620
gccacaatcc ttigaaacca tcaaaacact attttgtatt atttttgtct ggtataacgt	22680
glataatatt acttcagttt ccttaagtaa ttcttttaaa ttttagctt tattcactag	22740
catacatctc atgcatttaa atgtacaatt cagtgaattt tggaatactc acagagcttt	22800
gcaactgtca cctcagtttt ttaatttttt atttaaaaac aaatttatg acaataaaa	22860
attgttgcca ggcatgtgtg ctcacgctg taatcccagc actttgggag gccaggcca	22920
gcagatcaac gaggtcagga gtctgagacc agcctggcca atatggcgaa actccgtctc	22980
tactaaaaat acaaaaatta gccgggtgtg gtggtgcgtg cctglagtcc cagctatgcg	23040
ggaggctgag gcagaataat cgcttgaacc cagggaagtg aggttgact gagccagat	23100
cataccacta cactccagcc tgggtgacag agactccatc tcaaaaaaaaa aaaaaatttt	23160
ttttatggtg tacaacatga tgttttgaat tgtttacagg ccaggcgtgg tggctcatgc	23220
cagtaatccc agcactttga gaggtgcgg gcagtacacg tgaggttggg agttcgagac	23280
cagcctgacc aacatggaga aaccacatct ctactaaaaa taaaaagta gccaggcatg	23340

ttggcacatg cctctaacc cagctactcg ggaggctgag gcaggagaat cgcttgaacc	23400
caggaggcag aggttgcagt gacccatagat cgcgtcattg tactccagcc tgggcaaaga	23460
gagcgaaact ccatgtcaaa aaaaagaaaa agaaaaaaga aatatgttta tattgtgcag	23520
tggctatatc atgctaacat gctaattaac atggattata tcacatactt attgtttgtg	23580
gtgagaacac ttagaatcta cccctatggc aagtttcttt ttttcttttc ttttcttttt	23640
tttttttgag acggagtttc gctcttgttg cccaggctgg atggagtaca gtggtgcaat	23700
cttggctcac tgcaacctcc acctcccagag ttcaattgat tctcctgcct cagcctcctg	23760
agtagctggg attacaggca cccgccatca tgcctggcta atttttgtat ttttgcagag	23820
atggggtttc accatgttgg ccaggctgtt cttgaactcc tgacctcaag tgatctgccc	23880
gcctcagcct cccaaagtgc taggattata ggcttgagcc accgttccca gccagtgaca	23940
attttcaagt atacaatata ttgttagcca ctgttcccag ccagtggcaa ttttcaagta	24000
tataatacat tgttattaac tatagtcact atgttgtttc aggtcttatg ttttaagtctt	24060
gaatgtatth ttagttgatt tttatgtatt gtgtgcaata tattactcat tatatgcttt	24120
tagtcattgg gtaaattagt agagtggttt gacattattg gataaaatta tcttaatgtt	24180
ttcttttttc agcgtttag aatgaatgaa gacactggca ctgattatat aacaccttgg	24240
cagctatctc aagtcgttga tggaggaggt attggaatta tagaagaaag caaacacaca	24300
aatttgacta aaggcgatth tgtgacttct tttctattggc cctggcaaac caaggttatt	24360
ctggatggaa atagccctga aaaggtgata tatatatgaa catactctgat tttttttccc	24420
cgtataatth ttaactttgtg catttgataat tcagtttgtt ttgtaatcag ggaaaaataa	24480
aagcatatat tttcttttag ctagctgcca aaatgtggaa agcctttgtt ccttttcac	24540
tgttagctgt gtagggaagt ttaaattgtc ctcttgaagg tttgataatt gacccgtcta	24600
aaataaactg acaataaaca gattaacagg agaaaaacca ttttaattca tgtatgtgca	24660
ctggagcctc acagaagaat aggagactca aaggaggac cagatggttg aaggtgttct	24720
ttattttgtt ctttgcctgt ttgtttgttt tagagatggg ttgtcactgt gtgccccgg	24780
ctggtctcaa actcctagge tcaagcaatc ctcccacctc agcctcccaa agtgctgaga	24840
ttacaggtgt gagcttgaag tttttatagc taaaggaatg gggcttgagg tttttggggg	24900
aatggcaac aaaaagggaagg gggagggaagg ttgcatggta aacaaaggtt atcttattgt	24960
agataaagtc tctcaggtaa taaaagtgtt ttccagagcag ctgccaaaag aataggtaat	25020
agcctgactg ggtctgctaa tgactttatt tatttagaga cggcgtctca ctctgtcacc	25080
caggctggag tgatcttggc tcaactgcaac ctctgcctcc tgggttcaag tgatcttccc	25140
acctcagctt cccgagttagc tgggattata ggtgtgcacc accacgcca gctaattttt	25200

tttatTTTT gtagagatgg ggttttgcca tgttgccag gctgttctca aactcctgac	25260
ctcaagtgat ctgtctgtct tggcttccca aagtgtctggg attacaggcg tgagtcactg	25320
caccagcca tttttattaa acttgttgta gaaacaagag tcttaccatg ttgccaggc	25380
tggctttgaa ctcttggcct caaccgatcc tccgcctca gcccacaaa gtgtgggat	25440
tacaggtgtg aaccactgtg ccagcctta atggtacatt ttaagcatta gtaatgagaa	25500
ctgttggagt ttgagttgat ttttttttt ttcaagacag agtcttgctc tgttgccag	25560
gctggagcgc agtggcgca tctcggtca ctcaacctc tgccacctgg gttcaagcga	25620
ttctcctgcc tcagccacct gggtagctgg gattacaggc gtctaccacc atggctggct	25680
aatttttgta ttttagtag agatgggatt tcagcatctt ggccagggtg gtgttgaatg	25740
cctgacctca tgatccacc gccttggcct ccaaagtgc tgggattaca ggcgtgagcc	25800
actgtaccgg gcctgagttg atatttatag gtacatgttt atgttgaat gtctcagaga	25860
cttaagatc ttccattgtt gccaaagtatt ttctgttaa gcacttctt ttttttttt	25920
tttccatac ttggggcatt aacaagaaag aatattttgt ctcttacctg ggaaatttt	25980
taaaagtctc ttttttttt ttttttttg agatggagtc ttgctttgtc gccaggctg	26040
gagtgcagtg gcacaatctt ggctcactgc aacctccacc tcccaggctc aaacaattct	26100
ctgcctcagc ctcttgagta gctgggtttt gggcacccac caccacacct ggctaatttt	26160
ttgtattttt agtagagacg gggtttcacc attttgcca ggctggctct gaactcctga	26220
ccacatgate caccacctc agcctcctta agtactggga ttacaggcgt gagccaccgc	26280
accggcctt aaagtctct taatataaag agaagtaata ttgttttta ctctttaca	26340
ataaagtata gaattttagc aagactgatg aatggtttac tcttctata accatattgg	26400
gttatttatt atactgatgg actgccataa cagagtacta caggctgggt ggtttaaaca	26460
acagaaattt atttctcac agttctggaa gctaaaagtc caagatcaga gttctggaaa	26520
atttggttct tggtagggc cctcctctg gattgcagat ggctgcctc ttgctgtgtc	26580
ctcaggcagt ctttctctg tgggtgtgca gagagataga tctgtttct tcttcttcta	26640
aggacataag tcctttttt ttttttttg agatgaagtc taactctgtc gccaggctg	26700
gagtgcagtg gcatgatctc agctcactgc aacctcctgg gttcaagcga ttctcctgcc	26760
tcagcctctc gagtactgg gattacaggc gcatgccacc atgccttctc aattttttgt	26820
atttttagta gagatgggtt ttcccatgt tggctaggct ggtctcgaa tcctggtctc	26880
aggtgatccg cctgcctcgg tctcccaaag tgctgggatt acagatgtga gccactgtgc	26940

ccggcctgga cacaagtctt gttggactag gacccacct ttatgtctc attcaatcct	27000
aatgtctcc ttaagtgcc tatctccaaa taaagtcaca ctggggctta gggctttact	27060
atataaatgt tgtggggacg caattcagtc cataacacgt tgtctctact ttgttgtctc	27120
aaaacaggaa ttcttagtat ttttaaaggg tatatatatt attttaggta gaccacaaac	27180
ttgtggatgg acacctttca tattttcttg gagctatagg tatgcctggg ttgacttcct	27240
tgattgggat acaggaaaaa ggtcatataa ctgctggatc taataagaca atggttgtca	27300
gtggggccgc aggtgcctgt ggatctgtgg ctgggcagggt aaactttctg agaattattt	27360
gattttctga ttatttgacg attgttataa ttcttaattg aaattccgga tatatgtata	27420
ccattaaaaa tataaaattg gataaatttt tttttcttt tttgagacag agtctccagg	27480
ctggagtgcg gtggcgcaat cttggctcac tgcaacctcc gcctcctggg ttcaagagat	27540
tctcatgcct tagcctctag agtagctggg actacaggtg tgtgccacca cacctggcta	27600
atttttgtat ttttagtaga gatggggttt tgccatgttg gccaggctgt tctcgaactc	27660
ctggcctcaa gagatccacc tgcttcggcc tcccaaagtt ctgggattat aggtgtgagc	27720
taccacgcc agccaaaatt gactaaattt ataatgactt tcacacgttt ctgttagaaa	27780
tggttttttt tttttaccta cttagactct ctactaaca gtatgaaaa tgtagatttt	27840
agtgtcttat tctctgttaa ccaaaaggag atatagtaaa tgctaaagct ttgctttcat	27900
tattataagg aagagaatga tagagaagct gtttgggaatt gcatgaatta aaagtcaaac	27960
tcacatttta aattaatc ctcaataag aaaccactta ttgaattcta gaagtgtaaa	28020
cattttgata gaattttatg aagtcggcct ggcacagtgg ctcatgcctg caatcccagt	28080
gctttgggag gccaaacttg caggattact tgaggaaagg agcacaagac cagcctgggc	28140
aacatagtaa gactcttcag taaataaatt agcccagcgt ggtggcacgc ctgttgtcct	28200
aactacttga gaggctgagg caggaagagt acttcagccc aggagtttga agttacagtg	28260
agctatgatt gcaccactgt actccagcct gggtgacaga cctgtttct ttaaaaaaac	28320
aaaaaaagaa tcttataaaa gtacatcctt ttataaataa tagcctttat aggttgaaac	28380
cattgaaaac aaacatttg tccgagctgt gattttagtt cgccacaagc ttgaaaaata	28440
atgaattgca cgtgatgttt tgggtctgat ttttatcc ttttatgaaa tcagatttaa	28500
tttgcgaat tttattttt taaatttaatt tattattctt tgattttata tttttcaatg	28560
agtgaagcc aaaaatttg agattttaac atttaacta cctatggtcc ccaacttatg	28620
atggttctac tttacagtga tgcaaaagcc ctatgcattc agtgaagcc atacttgaa	28680
ttttgaattt tggcttttt ccaggctagc gatatcagta tgacactctc tcgagatgct	28740
gggcagcggc aacaagctgt agctcctggg cagccaccgt atgtggctga acaactcgt	28800

gaatcatgcc cagctggagt ggatgtttat ttgacaatg accataaggg taaacaactg	28860
glactctacg gtgtactgtg ttgccagatg attttgccca gccataagtg ttctgagcat	28920
gttttaggta ggctaggcta agctatgatg ttcagtaggt taggtatatt aaatgcattt	28980
tttgcttagg atattttcaa ctaatgatgg ttttatggg acatatccct atcataagtt	29040
gaggaacata agtctacctt tgttcaattt taataatttt tttttaggat gacattgtct	29100
gccctcagtt cttttagttt caaattactt ttcattaaaa gcagcgtaat ctaaccttaa	29160
ttgtaatgaa cactgtcata taccgtaat atcttgtata ctggtagatt ttacattta	29220
aatgtgacaa tataaagcca tttaatcttc aaatttcccc cctagattgg ccattttcta	29280
ggttgttcca gagtgggtggg aatttgtgga acacatgaga aatgcatcct cttagacctca	29340
gaactgggct ttgatgctgc aattaattat aaaaaagaca atgtggcaga acagctccgt	29400
gaatcatgcc cagctggagt ggatgtttat ttgacaatg ttggtggtaa catcagtgat	29460
acagtataaa gtcaggttgt ttgctgattt ctataacaaa tttgaatact gcactttata	29520
tgttgtgata tctttttacc taatactgag aaaaaattt ataaaattgg agggattatt	29580
gtgacttaat gtgatgatca gtattgtatt taagtgtac aaaagagctt gctgatacct	29640
acttccattt actgcagat ttgtactatg aacttcaatt tcagttattt gctaccaata	29700
ttatatTTTT cacttttggg ttatttttgt gaaaaattac agttgcattg ttgaaatgct	29760
gaagcactga tcctaactc accactgacc aactcacagc catggcccca gacctatggg	29820
ttcataggc tgcattgatt tatgcataat tactaacaat attataaaat atcaacatga	29880
ggtattaaca ctacatagtc ctttctgca tctaaaatta tttgcagttg agattattaa	29940
attgtttaaa tccaaaaaaa tacagttcta ccactaataa gtagcaacaa caactaatgt	30000
aattcttttg ctgattacca accatttatt tgccttgggt ttcctttgga atgtaattat	30060
ttctagctag ttaaaattct taaatactc taacaggtaa attaaaaaaa aaatcgaact	30120
agttgacaga aaggattaaa tagtcaagga ggcttcatt atggcagcca taaaatcagt	30180
aaacatagga aagggatatt ttatcattat ttttctctgt gcagatgaat gagaacagcc	30240
acatcatcct gtgtgggtcaa atttctcagt acaacaaaga tgtgccttat cctccccgc	30300
tatccctgc tatagaggca atccagaaag aaagaaacat cacaaggtgt gttcttcctc	30360
tttgccctta ttaccaggtt catgggggtt taaatatcat agatgtgtat gaaatctatt	30420
gttcttcata gacacttttt ctttattaaa taaaacttta ggtcgggcat agtggcttat	30480
gcctgtaatc ctagcacttt gggatgccga gacgggtgga ttgcttgagg ccaggagttc	30540
aaaactagcc tggccaacat ggtgaaaccc cttctctact aaaaatacaa aaaaattagc	30600
cgggtgtagt gatgggcgcc tgtaatccca gctacttggg aggcctgaggc aggagaatca	30660

cttgaacctg ggaggtggag gttgcagtga gccaaaatca tgccactgca ctccatccag	30720
cctgggcaat gaaaacgaaa ctccatctca aaaaaacaaa aaaaaaaaaa caaaaaaaaa	30780
cctttaaaat acacattttt taagctgggt caatatcttag gttctggcta actagctacc	30840
tgaaaactga tagttattgg ctttgactgg cctgggttaa gtcacatgag ggtcgagaaa	30900
ggatgctacc cacaaaggaa aaattcacgg tgctattaac aaatgggggt gtagatactg	30960
ggcagacaaa aattattatt atttgagatg gagtctcgct cagccgccca ggctggagta	31020
aagtggcgcg atcgagctc actaaaacca ccgtctccca ggttcaagcg attctcctgt	31080
ctcagcctcc cgagtagctg ggactacagg cgctcactac catgcccggc taattttttt	31140
gtatttttag tagagacggg ggtttcaccg tgttggccag gctggtcttg aaaaattatt	31200
aatagaagtc tacatttaca cctgtaacta tgcataaac tatctgatca gactcaatgt	31260
gttatgtcct aatattttta ggattaacag atattagaat tctggaatgg tttgatcttg	31320
gaataccact cccgttatat tctctgatca ttatcctttt ctctctcact gcagggaaag	31380
atttctggta ttaaattata aagacaaatt tgagcctggc attctacagc tgagtcagtg	31440
gtttaagaa ggaaagctaa aggtagaact tctattcttt atcaatataa ttcttctgc	31500
tacttgatat aattttgcat tattttaatg tactatagca aacatactag ttctattata	31560
aaaattccta caaaagaaaa gtatgctgaa tattataaat taactctacc tcccagggtga	31620
gtataacctc ctgatcaaca aacccttgct tgttattccc tttatctaaa attttttagc	31680
tttactttcc atagattata ttacaaaac attaggtttt ccccaaatg taaaattatg	31740
tttattaatt taacaaatat ttatgggatg ctactatgtg ttagaagctg tgctttctca	31800
aaaataatgt gagcacctcc tgatgcttat accaccatac ttcttccctg tcctccctgg	31860
acattcta at gtatgtctta gttgttgatc gaaaatcagg agttattaat ctgggcttca	31920
gacatactgg cctaggetat tctttttttt tttttttttt tttttttttt tgagacggag	31980
tctcactttg tcaccaggc tgtagtgcag tggcgcaatc tcggctcact gcaacctccg	32040
cctctaggt tcaagcaatt attctgcttc agcctccga gtagctggga ttacaggcac	32100
atgccactcg ctgagctaat ttttgatttt ttagtagaga tggggtttca ccatgttggc	32160
caggctggtc tcaaactcct gacctcaagt gttctgcca tccagcctcc caaagtctg	32220
ggattacagt tgtgaaccag tgcacctgga ctggcctagg ctattcttga aacctgcaaa	32280
tgataagcaa agttgtacct ttgtgtgctt ctgtactgtt ctttttctgt tttagagtgt	32340
atagcttgtg atctgttact aattgaagag attaaacaat gggaaaatag gatattactt	32400

taaaaaatca attgatagga agctttttctg ggctacctac atgaaataaa gaatatctat	32460
ttttagacaa agtgctaaat agtattgctg gctttttaa gtatggaaac tatcatatag	32520
aaacattgaa gtttaaagac tattaaatct agctattttg atttacagat taaagagacg	32580
gtaataaatg ggttggaaaa catgggaggt aagatgaatg tagacttatt atatacacat	32640
gctcagcaca caaagataaa gtctctttat ttttatttat ttatttttga gacggagtct	32700
tgctctgttg cccaggetgg agtgtagtgg catgatctca gctcactgca agctccgcct	32760
tccaggetca cgccattctc ctgcctcagc ctccaagta gctgggacta caggcgcctg	32820
ccaccacgcc tggctaattt ttgtatttt tagtagagac ggggtttcac tgtcttagcc	32880
aggatggtat cgatctctg acctcgtgat ctgcctgcct tggcctcca aagtgtctggg	32940
attacaagcg tgagccacca cgcctggccc ttttttttt ttttttttt ttttttttt	33000
ttttttttga gacggagtct tgctctgtct cccaggetgg agtgtagtgg tatgatcttg	33060
gctcactgca acctccacct cccgagatca agcaattctc ctgcctcagt ctaccaagta	33120
gctgggacta caggcgtgca ctggtatgcc tggctaattt ttgtatttt agtagagaca	33180
gggtttcaac atattggcca ggctgggtctt gaactcctga attcaagtga ttgacctgcc	33240
tccgcctctc aaagtgtctg gattacaggc gtgagccacc gcacaggcca aagtctttta	33300
tcttttaatg ctacctctct ggctccctt tccctgcct tccctccct cactccctt	33360
ttttttttt ttttttttt ttttttttt aaagagatga tcttgctctg ttgccaggc	33420
tggagtgcag tggcatgac atagctaact gtaacttatg aacctatggg ctcaagcaat	33480
cctctacca gtagtagtgg tactacaatt acctgccact atgccaacc ttttttttt	33540
ttttaattt ccatagaac agggctctac tatattgtcc aggcaactcc tggcctcaaa	33600
tgatcctct gcctcagct cccaaagtgc tggaaattac aggtgtgagt gcctagccct	33660
ttttttaatt tttagacag ggccatgctc tgtcaccgg gctggagtgc agtggcgga	33720
tcatggctca ctgcagctc tactccag gttaggtga tgcctccacc ttagccccc	33780
gagtagctgg gactacaggc acacaccacc actccagct gatttttga ttttagtag	33840
agatggggtt tcacatgtt ggtcaggtg gtcttaaact cctgacttca agtgatccac	33900
ccacctggc ctccaaagt gctgggatta cagtcgtgag gccccatgcc cggcctcac	33960
tcacatttta ttccatttt ttttgctc tagtcttcag atactgtaa caatgaacat	34020
tgatgtgcat accagttaga taccaattaa atctttccag tatgggcaaa acatttttt	34080
aagatgttaa ttgggttagt ttgtgtgtaa ctttttactt gagatttctt ttcatgacca	34140
ttttcttcta tgcctaatat tcaaaatcat tgaaaaagt atttaacata attttatatg	34200
ttgtaccttt tggaaacaag tatgatagt acatttatct tttcagataa cctactttct	34260

ctttttccag ctgcattcca gtccatgatg acaggaggta acattggaaa gcagatagtt	34320
tgcatttcag aagaaatctc tttgtaattg ctgtaaatgt catcaaggca atcatagatt	34380
tcttttccat tttgcatatt ttcaaagata tgttaaaaaa tccttagact atacatagct	34440
cttgatttaa atgtgatcat aggtgttatt tttagttgca tagggtatit gatacaatca	34500
ttaatggatc atacacaata ggttttttaa aattaataac ttttagtaat tacttttatt	34560
aattttaa atagaacgctg agaggcactt tgtaaagatt tgtaaactg gaaacgtttt	34620
acatgatctg atacaacat taatgaatca tacacaatag gttttttaa attaatatta	34680
ataactttta ttaattttaa atagaatgct taaaataaaa tagaatgctt gagaggcact	34740
gagtaaagat ttgttgaact ggaaatgttt tacatgattc ttaaaactgaa acttggtgta	34800
aaaatagaat tgagatggcc tttttttcac attgtagact gaaaagagac ttaatggat	34860
gatgtgtaca tagggactgg gggcaggatt gggggtttcg gagcttgtgt aacagttttt	34920
gggataggag accagcggtt ttgggttggg attagtatgg gaagataaat aacttaggtt	34980
ggttaagttg acaagattta ctccagagga tcctctcttt gtatttgcca aataatttac	35040
tgtatagcct aaaaactcca tatatattga gaaaagcata tgtttatttt aggttagcag	35100
gcacatactg tcaagttgta aagattgaga gggcaaacag atgtaaacat cactttagg	35160
tgattaaaaa gattgacagc cgggcctgat gtctcaagcc tgtaatecta gcactttggg	35220
aggcagaggc gggcagatca cttaggttca ggagttcaag accagcctgg ccaacacggt	35280
gaaaccccat ctctctaaaa atacaaaaat tagctgggcg tgggtggcaca cgcctgtaat	35340
cccagctact caggaggtg aggttgagga atcatttgaa cctgggaggg ggaggttgcc	35400
gtgagctgag ctigcaccat tgcactccag cctgggcgac aagagtgaaa ttccatctca	35460
aaaaacaaaa acagattgac aacaaagaca gtttcagaaa atgacaggac tgggcaaatt	35520
aacaaatgtt tgtaaacatg aatgttcagg aactactgat gtacctcaa agtttgtttt	35580
attaattgta ctcaaccctc gcagaacagt aaaactgaag attattgttt ctgaa	35635