(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2013년08월07일
(11) 등록번호 10-1289774
(24) 등록일자 2013년07월19일

(51) 국제특허분류(Int. Cl.)
C12Q 1/68 (2006.01) *G01N 33/48* (2006.01)
G01N 33/53 (2006.01) *G01N 33/574* (2006.01)
(21) 출원번호 10-2006-7022685
(22) 출원일자(국제) 2005년03월31일
심사청구일자 2010년02월04일
(85) 번역문제출일자 2006년10월30일
(65) 공개번호 10-2007-0031894
(43) 공개일자 2007년03월20일
(86) 국제출원번호 PCT/US2005/010645
(87) 국제공개번호 WO 2005/094357
국제공개일자 2005년10월13일
(30) 우선권주장
60/558,218 2004년03월31일 미국(US)
(뒷면에 계속)
(56) 선행기술조사문헌
Cancer Res., Vol. 156, No. 17, pp3859-3861 (1996.09.01)
전체 청구항 수 : 총 16 항

(73) 특허권자
다나-파버 캔서 인스티튜트 인크.
미국 메사추세츠 02215 보스턴 브룩클린 애비뉴 450
더 제너럴 하스피탈 코포레이션
미국, 메사추세츠 02114, 보스톤 프룻트 스트리트 55
(72) 발명자
벨 다프네 위니프레드
미국 02474 매사추세츠주 알링턴 아리조나 테라스 7 아파트 #3
하버 다니엘 에이.
미국 02467 매사추세츠주 체스트넛 힐 모나드녹 로드 34
(뒷면에 계속)
(74) 대리인
문기상, 문두현

심사관 : 이준혁

(54) 발명의 명칭 **표피성장인자 수용체를 표적으로 하는 치료에 대한 암의반응성을 결정하는 방법**

(57) 요 약

본 발명은 표피성장인자 수용체(EGFR) 치료에 대한 암의 반응성을 결정하기 위한 방법에 관한 것이다. 바람직한 실시태양에서는, erbB1 유전자의 키나제 도메인에서의 적어도 하나의 변이의 존재는 티로신 키나제 저해제인 제피티니브에 대한 민감성을 부여한다. 따라서 이러한 돌연변이에 대한 진단분석법은 제피티니브나 에를로티니브 및 다른 티로신 키나제 저해제의 투여를 상기 약제들에 가장 잘 반응할 환자들에게 투여할 수 있도록 할 것이다.

대 표 도

A B

(72) 발명자

잔 파시 안테로

미국 02464 매추세츠주 뉴턴 라운드우드 로드 128

존슨 브루스 이.

미국 02445 매사추세츠주 브룩라인 시버 스트리트 80

린치 토마스 제이.

미국 02465 매사추세츠주 뉴턴 프린스 스트리트 131

메이어슨 매튜

미국 01742 매사추세츠주 콩코드 인디펜던스 로드 307

파에즈 후안 길레모

미국 02115 매사추세츠주 보스톤 리버웨이 352 #22

셀러스 윌리엄 알.

미국 02445 매사추세츠주 브룩라인 시버 스트리트 80

세틀만 제프리 이.

미국 02458 매사추세츠주 뉴턴 웨이벌리 애비뉴 140

소델라 라파엘라

미국 01730 매사추세츠주 베드포드 선셋 로드 38

(30) 우선권주장

60/561,095 2004년04월09일 미국(US)
60/565,753 2004년04월27일 미국(US)
60/565,985 2004년04월27일 미국(US)
60/574,035 2004년05월25일 미국(US)
60/577,916 2004년06월07일 미국(US)
60/592,287 2004년07월29일 미국(US)

특허청구의 범위

청구항 1

삭제

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

삭제

청구항 6

삭제

청구항 7

삭제

청구항 8

삭제

청구항 9

삭제

청구항 10

삭제

청구항 11

삭제

청구항 12

삭제

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

청구항 16

삭제

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

삭제

청구항 66

삭제

청구항 67

삭제

청구항 68

삭제

청구항 69

삭제

청구항 70

삭제

청구항 71

삭제

청구항 72

삭제

청구항 73

삭제

청구항 74

삭제

청구항 75

삭제

청구항 76

삭제

청구항 77

삭제

청구항 78

삭제

청구항 79

삭제

청구항 80

삭제

청구항 81

삭제

청구항 82

삭제

청구항 83

삭제

청구항 84

삭제

청구항 85

삭제

청구항 86

삭제

청구항 87

삭제

청구항 88

삭제

청구항 89

삭제

청구항 90

삭제

청구항 91

삭제

청구항 92

삭제

청구항 93

삭제

청구항 94

삭제

청구항 95

삭제

청구항 96

삭제

청구항 97

삭제

청구항 98

삭제

청구항 99

삭제

청구항 100

삭제

청구항 101

삭제

청구항 102

삭제

청구항 103

삭제

청구항 104

삭제

청구항 105

삭제

청구항 106

삭제

청구항 107

삭제

청구항 108

삭제

청구항 109

삭제

청구항 110

삭제

청구항 111

삭제

청구항 112

삭제

청구항 113

삭제

청구항 114

삭제

청구항 115

삭제

청구항 116

삭제

청구항 117

삭제

청구항 118

삭제

청구항 119

야생형 EGFR 유전자에 대한, 비소세포 폐암에 걸린 환자의 종양 조직의 EGFR 유전자의 엑손 18, 19 또는 21에서의 적어도 하나의 누클레오티드 변이의 존재 또는 부재를 검출하는 단계를 포함하고,

상기 누클레오티드 변이는

a) (i) 서열번호 512의 719 위치의 글리신의 시스테인으로의 치환(G719C);

(ii) 서열번호 512의 719 위치의 글리신의 세린으로의 치환(G719S); 또는

(iii) 서열번호 512의 719 위치의 글리신의 알라닌으로의 치환(G719A);

으로 이루어진 아미노산 변화를 초래하는, 엑손 18에서의 치환,

b) 적어도 서열번호 512의 747, 748, 및 749 위치의 아미노산인 류신, 아르기닌, 및 글루탐산의 결실을 포함하는 아미노산 변화를 초래하는, EGFR 유전자의 엑손 19에서의 인프레임 결실, 또는

c) (i) 서열번호 512의 858 위치의 류신의 아르기닌으로의 치환(L858R); 또는

(ii) 서열번호 512의 861 위치의 류신의 글루타민으로의 치환(L861Q)

으로 이루어진 아미노산 변화를 초래하는, 엑손 21에서의 치환

을 포함하고,

상기 적어도 하나의 누클레오티드 변이의 존재는 제피티니브 또는 에를로티니브가 유효 가능성이 있음을 나타내는, 비소세포 폐암에 걸린 인간 환자에 있어서 제피티니브 또는 에를로티니브의 증가된 유효 가능성 결정 방법.

청구항 120

제119항에 있어서,

상기 적어도 서열번호 512의 747, 748, 및 749 위치의 아미노산인 류신, 아르기닌, 및 글루탐산의 결실이, 서열번호 512의 코돈 746의 아미노산인 글루탐산의 결실을 더 포함하는, 비소세포 폐암에 걸린 인간 환자에 있어서 제피티니브 또는 에를로티니브의 증가된 유효 가능성 결정 방법.

청구항 121

삭제

청구항 122

삭제

청구항 123

삭제

청구항 124

삭제

청구항 125

삭제

청구항 126

삭제

청구항 127

제119항에 있어서,

상기 적어도 하나의 누클레오티드 변이의 존재 또는 부재가 염기서열분석(sequencing), 증폭, 핵산 프로브 혼성화, 미스매치 절단 분석(mismatch cleavage analysis), 단일 가닥 입체구조 다형성(single strand conformation polymorphism), 변성 구배 겔 전기영동법(denaturing gradient gel electrophoresis) 또는 온도 구배 겔 전기영동법 분석에 의해 결정되는, 비소세포 폐암에 걸린 인간 환자에 있어서 제피티니브 또는 에를로티니브의 증가된 유효 가능성 결정 방법.

청구항 128

제119항에 있어서,

상기 엑손 18에서의 적어도 하나의 누클레오티드 변이가, 서열번호 511의 누클레오티드 2155의 구아닌의 티민으로의 치환 또는 구아닌의 아데닌으로의 치환, 또는 서열번호 511의 누클레오티드 2156의 구아닌의 시토신으로의 치환을 포함하는, 비소세포 폐암에 걸린 인간 환자에 있어서 제피티니브 또는 에를로티니브의 증가된 유효 가능성 결정 방법.

청구항 129

제119항에 있어서,

상기 엑손 19에서의 적어도 하나의 누클레오티드 변이가 서열번호 511의 누클레오티드 2240~2247의 결실을 포함하는, 비소세포 폐암에 걸린 인간 환자에 있어서 제피티니브 또는 에를로티니브의 증가된 유효 가능성 결정 방법.

청구항 130

제119항에 있어서,

상기 엑손 19에서의 적어도 하나의 누클레오티드 변이가, 서열번호 511의 누클레오티드 2235~2249의 결실, 누클레오티드 2236~2250의 결실, 누클레오티드 2239~2247의 결실, 누클레오티드 2240~2250의 결실, 누클레오티드 2240~2251의 결실, 누클레오티드 2240~2257의 결실, 누클레오티드 2239~2250의 결실, 누클레오티드 2237~2255의 결실, 누클레오티드 2239~2248의 결실, 누클레오티드 2239~2251의 결실, 또는 누클레오티드 2238~2255의 결실을 포함하는, 비소세포 폐암에 걸린 인간 환자에 있어서 제피티니브 또는 에를로티니브의 증가된 유효 가능성 결정 방법.

청구항 131

제119항에 있어서,

상기 엑손 21에서의 적어도 하나의 누클레오티드 변이가, 서열번호 511의 누클레오티드 2573의 티민의 구아닌으로의 치환, 또는 서열번호 511의 누클레오티드 2582의 티민의 아데닌으로의 치환을 포함하는, 비소세포 폐암에 걸린 인간 환자에 있어서 제피티니브 또는 에를로티니브의 증가된 유효 가능성 결정 방법.

청구항 132

(A) EGFR 유전자의 엑손 18, 19 또는 21에서의 적어도 하나의 누클레오티드 변이를 검출하도록 설계된 적어도 하나의 핵산 프로브로서, 상기 누클레오티드 변이가 제119항에 기재된 누클레오티드 변이이고, 상기 검출이 상기 누클레오티드 변이 염기서열에의 특이적 혼성화에 기반한 것인 핵산 프로브;

(B) 결합 반응(annealing reaction)의 수행에 필요한 산물 및 시약; 및

(C) 사용설명서

를 포함하는, 제119항의 방법에 사용되는 키트.

청구항 133

제132항에 있어서,

상기 핵산 프로브가 EGFR 유전자의 엑손 19 또는 21에서의 적어도 하나의 누클레오티드 변이를 검출하고,

상기 누클레오티드 변이가 적어도 서열번호 512의 747, 748, 및 749 위치의 아미노산인 류신, 아르기닌, 및 글루탐산의 결실을 포함하는 아미노산 변화를 초래하는 코돈 746 내지 753 내의 결실로 이루어진 EGFR 유전자의 엑손 19에서의 인프레임 결실을 초래하는 돌연변이이거나, 또는

상기 누클레오티드 변이가 서열번호 512의 858 위치의 류신의 아르기닌으로의 치환(L858R)으로 이루어진 아미노산 변화를 초래하는 엑손 21에서의 치환인, 키트.

청구항 134

제132항에 있어서,

상기 핵산 프로브가 EGFR 유전자의 엑손 18에서의 적어도 하나의 누클레오티드 변이를 검출하고, 상기 누클레오티드 변이가 서열번호 512의 719 위치의 글리신의 시스테인으로의 치환(G719C)으로 이루어진 아미노산 변화를 초래하는 엑손 18에서의 치환인, 키트.

청구항 135

제132항에 있어서,

상기 핵산 프로브가 EGFR 유전자의 엑손 18에서의 적어도 하나의 누클레오티드 변이를 검출하고,

상기 누클레오티드 변이가,

서열번호 512의 719 위치의 글리신의 세린으로의 치환(G719S)으로 이루어진 아미노산 변화를 초래하는 엑손 18에서의 치환이거나, 또는

서열번호 512의 719 위치의 글리신의 알라닌으로의 치환(G719A)으로 이루어진 아미노산 변화를 초래하는 엑손 18에서의 치환인, 키트.

청구항 136

a) EGFR 유전자의 엑손 18, 19 또는 21의 경계 부분 또는 내의 핵산 영역에 결합(annealing)하도록 설계된 적어도 하나의 축퇴성 프라이머쌍(degenerate primer pair)으로서, 상기 프라이머쌍이 EGFR 유전자의 엑손 18, 19, 또는 21에서의 적어도 하나의 누클레오티드 변이를 포함하는 핵산 서열을 특이적으로 증폭하며, 상기 누클레오티드 변이가 제119항에 기재된 누클레오티드 변이인, 축퇴성 프라이머쌍;

b) PCR 증폭의 수행에 필요한 산물 및 시약; 및

c) 사용설명서

를 포함하는, 제119항의 방법에 사용되는 키트.

청구항 137

제136항에 있어서,

상기 누클레오티드 변이가 서열번호 512의 719 위치의 글리신의 시스테인으로의 치환(G719C)으로 이루어진 아미노산 변화를 초래하는 엑손 18에서의 치환인, 키트.

청구항 138

제136항에 있어서,

상기 누클레오티드 변이가, 서열번호 512의 719 위치의 글리신의 세린으로의 치환(G719S)으로 이루어진 아미노산 변화를 초래하는 엑손 18에서의 치환이거나, 또는 서열번호 512의 719 위치의 글리신의 알라닌으로의 치환(G719A)으로 이루어진 아미노산 변화를 초래하는 엑손 18에서의 치환인, 키트.

청구항 139

제136항에 있어서,

상기 누클레오티드 변이가 적어도 서열번호 512의 747, 748, 및 749 위치의 아미노산인 류신, 아르기닌, 및 글루탐산의 결실을 포함하는 아미노산 변화를 초래하는 코돈 746 내지 753 내의 결실로 이루어진 EGFR 유전자의 엑손 19에서의 인프레임 결실을 초래하는 돌연변이이거나, 또는

상기 누클레오티드 변이가 서열번호 512의 858 위치의 류신의 아르기닌으로의 치환(L858R)으로 이루어진 아미노산 변화를 초래하는 엑손 21에서의 치환인, 키트.

청구항 140

a) EGFR 유전자의 엑손 19 또는 21 내의 핵산 영역에 결합하도록 설계된 적어도 하나의 핵산 프로브;

b) 결합 반응의 수행에 필요한 산물 및 시약; 및

c) 사용설명서

를 포함하고,

상기 프로브가 EGFR 유전자의 엑손 19 또는 21에서의 적어도 하나의 누클레오티드 변이를 포함하는 핵산 서열에 선택적 결합 조건 하에 특이적으로 결합하며,

상기 누클레오티드 변이가 적어도 서열번호 512의 747, 748, 및 749 위치의 아미노산인 류신, 아르기닌, 및 글루탐산의 결실을 포함하는 아미노산 변화를 초래하는 코돈 746 내지 753 내의 결실로 이루어진 EGFR 유전자의 엑손 19에서의 인프레임 결실을 초래하는 돌연변이이거나, 또는 상기 누클레오티드 변이가 서열번호 512의 858 위치의 류신의 아르기닌으로의 치환(L858R)으로 이루어진 아미노산 변화를 초래하는 엑손 21에서의 치환인, 제119항의 방법에 사용되는 키트.

청구항 141

삭제

청구항 142

삭제

청구항 143

삭제

청구항 144

삭제

청구항 145

삭제

청구항 146

삭제

청구항 147

삭제

청구항 148

삭제

청구항 149

삭제

청구항 150

삭제

청구항 151

삭제

청구항 152

삭제

청구항 153

삭제

청구항 154

삭제

청구항 155

삭제

청구항 156

삭제

청구항 157

삭제

청구항 158

삭제

청구항 159

삭제

명 세 서

배 경 기 술

[0001] 전립선암이나, 유방암, 결장암, 폐암, 이자암, 난소암, 지라암, 고환암, 흉선암 등의 상피세포암(epithelial cell cancer)은, 상피세포들의 비정상적인 가속화된 성장을 특징으로 하는 질환들이다. 이러한 가속화된 성장은 초기에는 종양 형성을 유발하게 된다. 결국에는 다른 기관 부위들로의 전이도 일어날 수 있다. 다양한 암들에 대한 진단 및 치료상의 진전이 있어왔지만, 이러한 질환들은 여전히 상당한 사망률로 귀결된다.

[0002] 폐암은 산업화된 국가들에 있어서 암 사망의 주도적 원인으로 남아있다. 폐에서 시작되는 암은 두 가지 주요한 타입으로 나누어지는데, 현미경 상에 그 암세포들이 어떻게 보이느냐에 따라 비소세포(non-small cell) 폐암과 소세포(small cell) 폐암으로 나누어진다. 비소세포 폐암(편평세포암종(squamous cell carcinoma), 선암종(adenocarcinoma), 그리고 대세포암종(large cell carcinoma))은 일반적으로 소세포 폐암에 비해 보다 느리게 다른 기관으로 퍼진다. 약 75%의 폐암 케이스가 비소세포 폐암으로 분류되며(예, 선암종), 나머지 25%가 소세포 폐암이다. 비소세포 폐암(NSCLC)은 미국, 일본 그리고 서부유럽에서의 암 사망의 주도적 원인이다. 병이 진행된 환자들의 경우, 화학요법은 생존과 관련하여 약간의 유익한 점들을 제공하기는 하지만 심각한 독성을 감내해야한다. 이러한 점은 종양의 성장을 지시하는 결정적인 유전적 손상을 특이적으로 표적으로 하는 치료제에 대한 필요성을 강조한다(Schiller JH et al., *N Engl. J. Med.*, 346: 92-98, 2002).

[0003] 표피성장인자 수용체(Epidermal Growth Factor Receptor; 이하 "EGFR")는 질량이 170 킬로달톤(kDa)으로 막에 결합된 단백질인데, 상피세포의 표면에 발현된다. EGFR은 세포주기조절분자들의 부류(class)인 단백질 티로신 키나제의 성장인자수용체 군(family)의 일원이다(W. J. Gullick et al., 1986, *Cancer Res.*, 46:285-292). EGFR은 세포외 도메인에 자신의 리간드(EGF 또는 TGF-α)가 결합하면 활성화되고, 그 결과, 동 수용체의 세포내 티로신 키나제 도메인이 자가인산화반응을 일으키게 된다(S. Cohen et al., 1980, *J. Biol. Chem.*, 255:4834-4842; A. B. Schreiber et al., 1983, *J. Biol. Chem.*, 258:846-853).

[0004] EGFR은, 성장촉진 종양유전자(growth promoting oncogene)인 erbB 또는 ErbB1의 단백질 산물인데, 이 erbB 또는 ErbB1은 많은 인간 암들의 발병과 진행에 있어서 극히 중요한 역할을 한다고 믿어지고 있는, 원발암유전자(protooncogenes)의 ERBB 군의 일원이다. 특히, 아교모세포종(glioblastomas)뿐만 아니라 유방암, 방광암, 폐암, 머리 암, 목 부위 암, 및 위암에서 EGFR의 발현이 증대된다는 것이 관찰되어 왔다. ERBB 종양유전자 군은 구조적으로 연관된 4개의 막횡단(transmembrane) 수용체들, 즉, EGFR, HER-2/neu(erbB2), HER-3(erbB3), 및 HER-4(erbB4)를 코딩하고 있다. 임상적으로, 종양에 있어서 ERBB 종양유전자의 증폭 및/또는 수용체의 과다발현은 치료법에 대한 반응성과 상관관계가 있을 뿐만 아니라 병의 재발과 환자의 좋지 않은 예후와도 상관관계가 있다는 것이 보고되어 있다(L. Harris et al., 1999, Int. *J. Biol. Markers,* 14:8-15; 및 J. Mendelsohn and J. Baselga, 2000, *Oncogene*, 19:6550-6565).

[0005] EGFR은 세 개의 주요 도메인들로 구성되어 있는데, 당화되어 있고(glycosylated), 2개의 시스테인 풍부 영역을 갖는 리간드-결합 포켓(ligand-binding pocket)을 함유하는 세포외 도메인(extracellular domain: ECD); 짧은 막횡단 도메인; 및 고유의 티로신 키나제 활성을 갖는 세포내 도메인이 그것이다. 막횡단 영역은 리간드-결합 도메인을 세포내 도메인과 결합시켜준다. 비당화형태의 EGFR에 대한 연구와 함께 아미노산과 DNA 염기서열분석은 EGFR의 단백질 골격이 질량이 132 kDa이고, 1186개의 아미노산 잔기를 갖는다는 것을 보여준다(A. L. Ulrich et al., 1984, *Nature*, 307:418-425; J. Downward et al., 1984, *Nature*, 307:521-527; C. R. Carlin et al., 1986, *Mol. Cell. Biol.*, 6:257-264; 및 F.L.V. Mayes and M. D. Waterfield, 1984, *The EMBO J.*, 3:531-537).

[0006] EGF나 또는 TGF-α가 EGFR에 결합하면 신호전달경로가 활성화되고, 그 결과, 세포가 증식된다. EGFR 분자의 이량체화(dimerization), 입체구조의 변화 (conformational change), 그리고 내재화(internalization)는 세포내 신호를 전달하는 기능을 하는데, 이는 세포의 성장을 조절하는 것으로 이어진다(G. Carpenter and S. Cohen, 1979, *Ann. Rev. Biochem.*, 48:193-216). 성장인자수용체 기능의 조절에 영향을 주거나 또는 수용체 및/또는 리간드의 과다발현으로 이어지는 유전자적 변경들의 결과, 세포가 증식된다. 거기에 더해서, EGFR은 세포의 분화와, 세포의 운동성(motility)의 향상, 단백질의 분비, 혈관의 신생(neovascularization), 침범(invasion), 암세포의 전이 및 화학요법제나 방사선에 대한 내성에 있어서 어떤 역할을 한다는 결론이 내려져 있다(M.-J. Oh et al., 2000, *Clin. Cancer Res.*, 6:4760-4763).

[0007] EGFR에 대한 다양한 저해제들이 동정되었으며, 그 중에는 이미 다양한 암들에 대한 치료로서 임상 시험이 진행되고 있는 여러 개가 포함되어 있다. 최신의 요약은 de Bono, J. S. and Rowinsky, E.K. (2002), "The ErbB

Receptor Family: A Therapeutic Target For Cancer", *Trends in Molecular Medicine*, 8, S19-26을 참조할 수 있다.

[0008] 암 치료에 있어서 치료적 개입(therapeutic intervention)을 위한 유망한 표적의 세트에는 HER 키나제 축(HER-kinase axis)에 속하는 구성원들이 포함된다. 그것들은, 예를 들면, 전립선이나 폐, 유방의 고형상피세포종양(solid epithelial tumor)들에 있어서, 자주 상향조절(upregulated)되어 있고, 또한 아교모세포종(glioblastoma)들에서도 상향조절되어 있다. 표피성장인자 수용체(EGFR)는 HER 키나제 축의 구성원 중의 하나로서, 몇가지 서로 다른 암 치료법들의 개발을 위한 선택의 표적이 되어 왔다. EGFR 티로신 키나제 저해제(EGFR tyrosine kinase inhibitors: EGFR-TKIs)들은 이러한 치료법들 중의 하나인데, 그 이유는 EGFR 경로의 활성화에는 티로신 잔기들의 가역인산화 과정이 요구되기 때문이다. 달리 말하자면, EGFR-TKI들은 종양 세포들의 성장과 분화를 일으키는 세포의 신호경로를 촉발 및/또는 유지하는 일을 하고 있는 세포 표면의 수용체를 차단한다. 구체적으로는, 이들 저해제들이 HER-1로 명명되는 EGFR 키나제 도메인을 간섭(interfere)한다고 믿어진다. 보다 더 유망한 EGFR-TKI들 중에는 3가지의 화합물 시리즈들이 있는데, 그들은 퀴나졸린(quinazolines)과 피리도피리미딘(pyridopyrimidines), 그리고 피롤로피리미딘(pyrrolopyrimidines)이다.

[0009] 임상적인 개발에 있어서 보다 진보된 화합물들 중의 2가지로서 제피티니브(Gefitinib)(AstraZeneca UK Ltd.에 의해 개발된 화합물 ZD1839; 상표명 IRESSA 로 구입가능; 이하 "IRESSA")와 에를로티니브(Erlotinib)(Genentech, Inc.과 OSI Pharmaceuticals, Inc.에 의해 개발된 화합물 OSI774; 상표명 TARCEVA로 구입가능; 이하 "TARCEVA")가 있는데, 둘 다 고무적인 임상결과들을 낳고 있다. IRESSA나 TARCEVA 를 사용하는 종래의 암 치료법은 각 화합물을 500 mg을 넘지 않는 양으로 매일, 경구 투여하는 것이다. IRESSA는, 2003년 5월에, 진행된 비소세포 폐암 환자들의 치료를 위해 승인되어, 이러한 제품들 중 제일 먼저 미국 시장에 진입하게 되었다.

[0010] IRESSA는 경구적으로 유효한(orally active) 퀴나졸린으로서 EGFR 분자상의 티로신 키나제의 인산화를 직접 저해함으로써 기능한다. 아데노신 삼인산(ATP)의 결합부위에 경쟁적으로 작용하며, 이는 HER-키나제 축의 억제로 이어진다. IRESSA 반응의 정확한 메카니즘은 완전하게 이해되고 있지는 않다. 하지만, EGFR의 존재가 IRESSA의 작용에 필수 전제임을 여러 연구들은 개시한다.

[0011] 이들 화합물들을 사용하는데 있어서의 심대한 한계는 그 수용자들이 그 치료법에 최초로 반응한 이후, 그 치료효과들에 대해 내성을 발달시키는 경우가 있거나, 또는 EGFR-TKI에 대해 측정가능한 정도의 반응을 전혀 보이지 않을 수도 있다는 점이다. 사실, 진행된 비소세포 폐암 환자들 중 단지 10-15%만이 EGFR 키나제 저해제 들에 반응한다. 따라서 IRESSA 와 TARCEVA 에 대한 민감도 밑에 놓인 분자수준의 메카니즘에 대한 보다 나은 이해는 아마도 표적 치료법에 있어서 그러한 치료를 통해 가장 많이 혜택을 보는 인간들에게 극히 유익할 것이다.

[0012] 암, 특히 폐암이나 난소암, 유방암, 뇌암, 대장암, 그리고 전립선암들 같은 상피세포암의 만족할 만한 치료에 대한 당 기술분야의 상당한 요구가 있으며, 여기서 만족할 만한 치료라는 것은 TKI 치료법으로부터의 이점들과 환자들에게 나타난 무반응성을 극복하는 것을 포함한다. 이러한 치료는 개인들의 건강에, 그리고 특히, 암이 특히 흔한 나이가 많은 인간들의 건강에 극적인 효과를 줄 수 있다.

[0013] 제피티니브(IRESSA®) 같은 티로신 키나제 저해제(TKI) 치료법은 상술한 바와 같은 암에 걸린 인간들 중의 대다수에게는 효과적이지 않다. 본 발명자들은 놀랍게도, EGFR의 키나제 도메인에서의 체세포 돌연변이의 존재가, IRESSA나 TARCEVA 같은 TKI에 대한 EGFR의 민감성을 현저하게 증가시킴을 발견하였다. 예를 들면, 그러한 암을 가진 환자들 중 30% 미만이 현재 통용되는 TKI 들에 의한 치료에 민감한 반면, EGFR 키나제 도메인에서의 돌연변이를 가진 환자들 중에서 50%를 넘는 환자들, 보다 바람직하게는 60, 70, 80, 90 %의 환자들이 민감하다. 또, 이러한 돌연변이들은 EGFR의 키나제 활성을 증가시킨다. 따라서 이러한 돌연변이들을 가진 환자들은 현재의 티로신 키나제 저해제(TKI), 예를 들면 제피티니브 같은, 치료법에 반응할 것이다.

[0014] 따라서, 본 발명은 암에 걸린 인간 환자에 있어서 표피성장인자 수용체(EGFR) 표적 치료의 유효 가능성을 결정하는 신규한 방법을 제공한다. 이 방법은 야생종 erbB1 유전자에 대하여, 상기 환자의 erbB1 유전자의 키나제 도메인에서의 적어도 하나의 핵산 변이의 존재 또는 부재를 검출하는 것을 포함한다. 적어도 하나의 변이가 존재하면 그것은 EGFR 표적 치료가 유효할 것이라는 것을 나타낸다. 바람직하게는, 상기 핵산 변이는 EGFR의 키나제 활성을 증대시킨다. 그렇게 되면, 상기 환자는 EGFR 표적 치료를 받을 수 있게 된다. 본 발명의 하나의 태양에서는, EGFR 표적 치료가 티로신 키나제 저해제이다. 바람직한 태양에서는, 상기 티로신 키나제 저해제가 아닐리노퀴나졸린(anilinoquinazoline)이다. 아닐리노퀴나졸린은 합성 아닐리노퀴나졸린이어도 된다. 바람직하게는 합성 아닐리노퀴나졸린은 제피티니브이거나 또는 에를로티니브이다. 또 다른 태양에서는, EGFR 표적 치

료가 4-디메틸아미노-부트-2-엔산[4-(3-클로로-4-플루오로-페닐아미노)-3-시아노-7-에톡시-퀴놀린-6-일]-아미드("EKB-569", "EKI-569"로 칭하기도 함. 예로서 WO/2005/018677과 Torrance et al., *Nature Medicine*, vol.6, No.9, Sept. 2000, p. 1024를 참조) 및/또는 HKI-272 또는 HKI-357(Wyeth; Greenberger et al., Proc. 11th NCI EORTC-AACR Symposium on New Drugs in Cancer Therapy, *Clinical Cancer Res*. Vol.6 Supplement, Nov. 2000, ISSN 1078-0432; Rabindran et al., *Cancer Res*. 64: 3958-3965(2004); Holbro and Hynes, *Ann. Rev. Pharm. Tox*. 44:195-217(2004); Tsou et al., *J. Med. Chem*. 2005, 48, 1107-1131; 및 Tejpar et al., *J. Clin. Oncol*. ASCO Annual Meeting Proc. Vol. 22, No. 14S: 3579(2004) 참조)를 포함하는 비가역적 EGFR 저해제이다.

[0015] 본 발명의 하나의 태양에서는, 암에 걸렸거나 암이 발병할 위험이 있는 환자로부터 얻은 생물학적 시료로부터 EGFR을 얻는다. EGFR(또는 erbB1 유전자)의 키나제 도메인에서의 변이는 ATP-결합 포켓의 입체구조에 영향을 준다. 바람직하게는, EGFR의 키나제 도메인에서의 변이는 엑손 18, 19, 20 또는 21에서의 인프레임 결실(in frame deletion) 또는 치환이다.

[0016] 하나의 태양에서는, 상기 인프레임 결실은 EGFR(erbB1)의 엑손 19에 존재한다. 상기 엑손 19에서의 인프레임 결실은 바람직하게는 적어도 코돈 747, 748, 749, 및 750의 아미노산인 류신, 아르기닌, 글루탐산, 및 알라닌의 결실을 포함한다. 하나의 태양에서, 상기 인프레임 결실은 누클레오티드 2235~2249을 포함하며, 아미노산 746~750(글루탐산, 류신, 아르기닌, 글루탐산, 및 알라닌 서열)이 결실되어 있다(표 2, 표 S2, 도 2B, 도 4A, 도 5, 서열번호 511, 도 6C, 및 도 8C 참조). 또 다른 태양에서, 상기 인프레임 결실은 누클레오티드 2236~2250을 포함하며, 아미노산 746~750이 결실되어 있다(표 S2, 도 5, 서열번호 511, 및 도 6C 참조). 또는, 상기 인프레임 결실은 누클레오티드 2240~2251을 포함하거나(표 2, 도 2C, 도 4A, 도 5, 서열번호 511 참조), 또는 누클레오티드 2240~2257을 포함한다(표 2, 표 S3A, 도 2A, 도 4A, 도 5, 서열번호 511, 도 6C, 및 도 8E 참조). 또는, 상기 인프레임 결실은, 누클레오티드 2239~2247을 포함하고 이와 함께 누클레오티드 2248의 구아닌이 시토신으로 치환된 것이거나(표 S3A 및 도 8D 참조), 또는 누클레오티드 2237의 아데닌이 티민으로 치환된 것과 함께 누클레오티드 2238~2255의 결실을 포함하거나(표 S3A 및 도 8F 참조), 또는 누클레오티드 2254~2277의 결실을 포함한다(표 S2(서열번호 437) 참조). 또는, 상기 인프레임 결실은 표 S3B에 나타낸 바와 같이, 누클레오티드 2239~2250delTTAAGAGAAGCA (서열번호 554); 2251A>C, 또는 2240~2250delTAAGAGAAGCA (서열번호 720), 또는 2257~2271delCCGAAAGCCAACAAG (서열번호 721)를 포함한다.

[0017] 또 다른 태양에서, 상기 치환은 EGFR의 엑손 21에 존재한다. 엑손 21에서의 상기 치환은 적어도 하나의 아미노산을 포함한다. 하나의 태양에서, 엑손 21에서의 상기 치환은 누클레오티드 2573의 티민이 구아닌으로 치환된 것을 포함한다(도 4A 및 도 5, 서열번호 511 참조). 이 치환의 결과, 아미노산이 치환되는데, 아미노산 858의 야생형 류신이 아르기닌으로 대체된다(도 5, 표 2, 표 S2, 표 S3A, 도 2D, 도 6A, 도 8B, 및 서열번호 512 참조). 또는 엑손 21에서의 상기 치환은 누클레오티드 2582의 티민이 아데닌으로 치환된 것을 포함한다(도 4A, 및 도 5, 서열번호 511 참조). 이 치환의 결과, 아미노산이 치환되는데, 아미노산 861의 야생형 류신이 글루타민으로 대체된다(표 5(각각 표시된 순서대로 서열번호 740~762), 표 2(각각 표시된 순서대로 서열번호 730~739), 도 2E, 표 S3B(각각 표시된 순서대로 서열번호 554 및 720~729) 및 서열번호 512 참조).

[0018] 상기 치환은 EGFR의 엑손 18에도 존재할 수 있다. 하나의 태양에서, 엑손 18에서의 치환은, 누클레오티드 2155의 구아닌이 티민으로 치환된 것이다(도 4A 및 도 5, 서열번호 511 참조). 이 치환의 결과, 아미노산이 치환되는데, 코돈 719의 야생형 글리신이 시스테인으로 치환된다(도 5, 서열번호 512 참조). 또 다른 태양에서 엑손 18에서의 치환은 누클레오티드 2155의 구아닌이 아데닌으로 치환된 것이고, 그 결과 아미노산이 치환, 즉, 코돈 719의 야생형 글리신이 세린으로 치환된다(표 S2, 도 6B, 도 8A, 도 5, 서열번호 511 및 512 참조).

[0019] 또 다른 태양에서 상기 치환은 서열번호 511의 누클레오티드 2316 후이면서 누클레오티드 2317 전인 곳에 구아닌, 구아닌, 티민(GGT)이 삽입된 것이다(2316_2317 ins GGT). 이는 아미노산 772에 발린(V)을 삽입(P772_H733 insV)하는 것으로도 기술할 수 있다. 다른 돌연변이들은 표 S3B에 나타나 있으며, 예로서, 서열번호 511의 누클레오티드 2309의 후이면서 누클레오티드 2310의 전인 곳에 CAACCCGG를 삽입하는 것과, 서열번호 511의 누클레오티드 2311의 후이면서 누클레오티드 2312의 전인 곳에 GCGTGGACA를 삽입하는 것을 포함한다. 상기 치환은 또한 엑손 20에 존재할 수도 있으며, 하나의 태양에서는 누클레오티드 2334 및 2335에서 GG가 AA로 치환된 것이다(표 S3B 참조).

[0020] 요약하면, 바람직한 태양들에서, erbB1 유전자의 핵산 변이는, 서열번호 511의 누클레오티드 2155의 구아닌의 티민으로의 치환, 또는 구아닌의 아데닌으로의 치환, 서열번호 511의 누클레오티드 2235~2249, 2240~2251,

2240~2257, 2236~2250, 2254~2277, 또는 2236~2244의 결실, 서열번호 511의 누클레오티드 2316의 후이면서 누클레오티드 2317의 전인 곳에 누클레오티드인 구아닌, 구아닌, 티민(GGT)의 삽입, 서열번호 511의 누클레오티드 2573의 티민의 구아닌으로의 치환 또는 서열번호 511의 누클레오티드 2582의 티민의 아데닌으로의 치환이다.

[0021] 적어도 하나의 핵산 변이의 존재 또는 부재를 검출하는 것은 그 수용체를 코딩하고 있는 핵산의 한 세그먼트를 증폭함으로써 결정될 수 있다. 증폭될 상기 세그먼트의 길이는 1000 누클레오티드 이하, 바람직하게는 500 누클레오티드 이하, 가장 바람직하게는 100 누클레오티드 이하이다. 증폭될 상기 세그먼트는 복수의 변이를 포함할 수 있다.

[0022] 또 다른 태양에서는, 적어도 하나의 변이의 존재 또는 부재를 검출하는 것은 변이 부위를 함유하는 EGFR 핵산을 적어도 하나의 핵산 프로브에 접촉시키는 것을 포함한다. 상기 프로브는, 변이 부위를 포함하고 그 변이 부위에 상보적인 누클레오티드 염기들을 함유하는 핵산 서열과 선택적 혼성화(selective hybridization) 조건하에서 우선적(preferentially)으로 혼성화한다. 혼성화는 검출가능한 표지로써 검출될 수 있다.

[0023] 또 다른 태양에서, 적어도 하나의 변이의 존재 또는 부재를 검출하는 것은 적어도 하나의 핵산 염기서열을 분석하여 기존에 알려진 erbB1 핵산 서열과 상기 분석된 서열을 비교하는 것을 포함한다. 또는, 적어도 하나의 변이의 존재 또는 부재는 적어도 하나의 핵산 서열의 질량분석 결정을 포함한다.

[0024] 바람직한 태양에서, 적어도 하나의 핵산 변이의 존재 또는 부재를 검출하는 것은 중합효소연쇄반응(PCR)을 수행하는 것을 포함한다. 가상(hypothetical)의 변이를 함유하고 있는 erbB1 핵산 서열을 증폭하고 그 증폭된 핵산의 누클레오티드 서열을 결정한다. 증폭된 핵산의 누클레오티드 염기서열을 결정하는 것은 적어도 하나의 핵산 세그먼트의 서열 분석을 포함한다. 또는, 증폭결과물은 그들의 크기에 따라 서로 분리해 낼 수 있는 방법들 중, 어떤 방법에 의해서든지 분석될 수 있는데, 이 방법들에는 자동 또는 수동 겔 전기영동법이나 다른 유사한 방법들이 포함된다.

[0025] 또는, 적어도 하나의 변이의 존재 또는 부재를 검출하는 것은 유전자 내의 복수의 변이의 하플로타입(haplotype)을 결정하는 것을 포함한다.

[0026] 또 다른 태양에서, EGFR 변이의 존재 또는 부재는 erbB1 유전자의 산물(단백질)을 분석함으로써 검출될 수 있다. 이 태양에서는, EGFR 변이체에 특이적으로 결합하는 프로브가 사용된다. 바람직한 태양에서는, 그 프로브는 EGFR 변이체에 우선적으로 결합하는 항체이다. EGFR의 변이체가 존재하면 EGFR 표적 치료가 유효할 수 있을 것으로 예견된다. 또는, 상기 프로브는 항체 단편이나 키메라 항체(chimeric antibody), 인간화된 항체(humanized antibody), 또는 앱테이머(aptamer)이다.

[0027] 본 발명은 또한 적어도 하나의 핵산 변이를 EGFR 유전자(erbB1) 내에 포함하는 핵산 서열에 선택적 결합 조건하에서 특이적으로 결합하는 프로브를 제공한다. 하나의 태양에서, 상기 변이는 erbB1의 키나제 도메인에서의 돌연변이이며, 이 돌연변이는 ATP-결합 포켓에 구조적 변화를 부여한다.

[0028] 본 발명의 프로브는 약 500 누클레오티드 이하의 염기들, 바람직하게는 약 100 누클레오티드 이하의 염기들, 가장 바람직하게는 약 50 누클레오티드 이하 또는 약 25 누클레오티드 이하의 염기들로 된 핵산 서열을 포함할 수 있다. 그 프로브는 DNA, RNA, 또는 펩티드 핵산(PNA)으로 구성될 수도 있다. 나아가서, 그 프로브는 검출가능 표지를 함유할 수도 있는데, 그러한 검출가능 표지로는, 예를 들면, 형광 표지나 효소적 표지가 있다.

[0029] 본 발명은 또한 암에 걸린 환자에 있어서, 표피성장인자 수용체(EGFR)에 대한 표적 치료의 유효 가능성을 결정하는 새로운 방법을 제공한다. 상기 방법은 환자로부터 얻은 생물학적 시료 중의 EGFR의 키나제 활성을 결정하는 것을 포함한다. EGFR 리간드로 자극한 후, 대조군에 비하여 키나제 활성이 증가되면, EGFR 표적 치료가 유효할 가능성이 있음을 나타낸다.

[0030] 본 발명은 또한 암에 걸렸거나 암이 발병할 위험이 있는 환자를 치료하는 신규한 방법을 제공한다. 상기 방법은 환자의 EGFR의 키나제 도메인이 적어도 하나의 변이를 함유하는지의 여부를 결정하는 것을 포함한다. 바람직하게는, 그 EGFR은 종양이나 암의 부위에 위치하고 있고, 핵산의 변이는 체세포 변이이다. 그러한 변이의 존재는 EGFR 표적 치료가 유효할 것임을 나타낸다. 만일 그러한 변이가 존재하면, 티로신 키나제 저해제가 환자에게 투여된다.

[0031] 위와 같이, 확인된 환자에게 투여한 상기 티로신 키나제 저해제는 아닐리노퀴나졸린일 수 있고, 또는, 예를 들면, EKB-569, HKI-272 및/또는 HKI-357(Wyeth)과 같은 비가역적 티로신 키나제 저해제일 수 있다. 바람직하게는, 아닐리노퀴나졸린은 합성 아닐리노퀴나졸린이고, 가장 바람직하게는 상기 합성 아닐리노퀴나졸린은 제피티

니브, 에를로티니브이다.

[0032] 본 발명의 방법으로 치료받을 수 있는 암은 예를 들면 위장관암, 전립선암, 난소암, 유방암, 머리와 목 암, 폐암, 비소세포(non-small cell) 폐암, 신경계암, 신장암, 망막암, 피부암, 간암, 이자암, 생식-비뇨기암, 그리고 방광암 등이 있으나, 이에 한정되는 것은 아니다. 바람직한 태양에서, 그 암은 비소세포 폐암이다.

[0033] 본 발명의 PCR 방법을 수행하기 위한 키트도 제공된다. 상기 키트는 적어도 하나의 축퇴성 프라이머 쌍을 갖는데, 이 프라이머 쌍은 EGFR 키나제 도메인의 ATP-결합 포켓을 코딩하고 있는 유전자들의 경계 부분의 핵산 영역에 결합하도록 설계되어 있다. 이에 더하여, 상기 키트는 PCR 증폭반응을 수행하는데 필요한 산물과 시약, 및 설명서를 포함하고 있다.

[0034] 바람직한 태양에서, 키트에 포함된 프라이머 쌍들은 서열번호 505, 506, 507, 및 508로 이루어진 군으로부터 선택된다. 실시예에서의 표 6과 7에 기재되어 있는 프로브들도 또한 바람직한 프로브들이다.

[0035] 또 다른 태양에서, 본 발명은 표피성장인자 수용체(EGFR) 변이체의 키나제 촉매 활성을 저해하는 화합물을 선별하는 방법을 개시한다. 첫 번째 단계로서, EGFR 변이체를 가능성이 있는 화합물과 접촉시킨다. 그 결과로서 나타나는 EGFR 변이체의 키나제 활성을 검출하고 EGFR 변이체의 키나제 활성을 저해하는 화합물을 선별한다. 하나의 태양에서는, 상기 EGFR 변이체는 세포 내부에 들어 있다. 상기 방법은 키나제 도메인에 2차 돌연변이를 갖는 EGFR 변이체들의 키나제 활성을 저해하는 화합물을 선별하는데에도 또한 사용될 수 있는데, 여기서 상기 2차 돌연변이는 예를 들면 제피티니브나 에를로티니브 같은 TKI에 대해 내성을 부여한다.

[0036] 하나의 태양에서, EGFR의 변이체는 표지되어 있다. 또 다른 태양에서는 EGFR이 고체 지지체에 결합되어 있다. 바람직한 태양에서는 고체 지지체는 단백질 칩이다.

[0037] 본 발명의 또 다른 태양에서는, 표피성장인자 수용체(EGFR) 변이체의 키나제 촉매 활성을 저해하는 약제 조성물이 개시된다. EGFR 변이체의 키나제 촉매 활성을 저해하는 화합물은 항체, 항체 단편, 소분자, 펩티드, 단백질, 안티센스 핵산, 리보자임, PNA, siRNA, 올리고누클레오티드 엡테이머, 및 펩티드 엡테이머로 이루어진 군으로부터 선택된다.

[0038] EGFR 매개 질병을 가진 환자를 치료하는 방법도 개시되어 있다. 이 방법에 의하면, 상기 환자에게 표피성장인자 수용체(EGFR) 변이체의 키나제 촉매 활성을 저해하는 약제 조성물을 투여한다.

[0039] 하나의 태양에서, 상기 EGFR 매개 질병은 암이다. 바람직한 태양에서, 그 암은 상피세포 기원의 것이다. 예로서, 그 암은 위장관암, 전립선암, 난소암, 유방암, 머리 및 목 암, 폐암, 비소세포 폐암, 신경계암, 신장암, 망막암, 피부암, 간암, 이자암, 생식-비뇨기 암, 방광암이다. 바람직한 태양에서, 그 암은 비소세포 폐암이다.

[0040] 또 다른 태양에서는, erbB1 유전자의 키나제 도메인에서의 2차 돌연변이의 획득을 예측하는(또는 돌연변이를 선별하는) 방법이 개시된다. 유효량이지만 준치사량의 티로신 키나제 저해제를 erbB1 유전자의 변이형을 발현하는 세포와 접촉시킨다. 티로신 키나제 저해제의 성장 저지 효과에 내성을 갖는 세포들을 선별하여 erbB1 키나제 도메인 내의 추가적인 돌연변이의 존재에 대하여 erbB1 핵산을 분석한다. 하나의 태양에서, 상기 세포는 시험관내(in vitro)에 있다. 또 다른 태양에서는 상기 세포는 유전자도입 동물로부터 얻어진다. 하나의 태양에서, 상기 유전자도입 동물은 마우스이다. 이 마우스 모델에서, 연구될 세포들은 종양 생검으로부터 얻어진다. 본 발명에 의해 선별된, erbB1 키나제 도메인에 2차 돌연변이를 갖는 세포들은, 위의 방법에서, 키나제 도메인에 2차 돌연변이를 갖는 EGFR 변이체의 키나제 활성을 저해하는 화합물을 선별하는데 사용될 수 있다.

[0041] erbB1 유전자의 키나제 도메인의 2차 돌연변이 획득을 예측하기 위한 하나의 또 다른 태양에서는, 맨처음 erbB1 유전자의 변이형을 발현하는 세포들을 유효량의 돌연변이 유발제(mutagenizing agent)와 접촉시킨다. 상기 돌연변이 유발제는, 예를 들면, 에틸 메탄설포네이트(EMS), N-에틸-N-니트로소우레아(ENU), N-메틸-N-니트로소우레아(MNU), 포카르박신 하이드로클로라이드(Prc), 메틸 메탄설포네이트(MeMS), 클로람뷰실(Chl), 멜팔란(melphalan), 포르카르바진 하이드로클로라이드, 사이클로포스파미드(Cp), 디에틸 설페이트(Et_2SO_4), 아크릴아미드 모노머(AA), 트리에틸렌 멜라민(TEM), 니트로겐 머스타드, 빈크리스틴(vincristine), 디메틸니트로스아민, N-메틸-N'-니트로-니트로소구아니딘(MNNG), 7,12-디메틸벤즈안트라센(DMBA), 에틸렌 옥사이드, 헥사메틸포스포르아미드, 비설판(bisulfan), 또는 에틸 메탄설포레이트(EtMs)이다. 그 후 상기 세포를 유효량이지만 준치사량의 티로신 키나제 저해제와 접촉시킨다. 티로신 키나제 저해제의 성장 저지 효과에 내성을 갖는 세포들을 선별하고, 그 erbB1의 키나제 도메인에서의 추가적인 돌연변이들의 존재 여부에 대하여 그 erbB1 핵산을 분석한다.

실 시 예

[0280] [실시예 1]

[0281] 종양 표본의 누클레오티드 염기서열분석

[0282] IRB에서 승인한 프로토콜 하에, 제피티니브로 후속 치료를 받은 NSCLC를 가진 환자로부터, 초기 진단 또는 외과적 처치로 얻은 종양 표본을 수집(collect)하였다. 4개의 사례에 대해서는, 대응하는 정상 조직과 함께, 동결(frozen) 종양 표본을 사용할 수 있었고, 나머지 표본에 대해서는 파라핀 포매(paraffin-embedded) 재료를 사용하였다. 또한, 원발성 NSCLC의 25개의 비선별 사례(unselected case)(15개는 기관지폐포암, 7개는 선암종, 3개는 대세포 폐암)는, 대응하는 정상 조직과 함께, 메사추세츠 제너럴 하스피탈 종양 은행(Messachusetts General Hospital tumor bank)으로부터 입수하였다. 전체 EGFR 코딩 서열의 돌연변이를 분석하기 위해, 표본으로부터 DNA를 추출하여, 모두 28개의 엑손을 증폭시킨 다음, 클로닝하지않은 PCR 단편(uncloned PCR fragment)을 자동 염기서열분석한 후, 이형접합성 돌연변이의 존재에 대해 센스 및 안티센스 방향 둘다에서의 전기영동도(electropherogram)를 분석하였다. 모든 염기서열 변이체는 다중 독립 PCR 증폭(multiple independent PCR amplification)에 의해 확인하였다. 프라이머 염기서열 및 증폭 조건은 Supplementary Material에서 제공한다. 또한, 엑손 19 및 21에서의 *EGFR* 돌연변이는, 다양한 세포조직(histology)(하기에 기재함)을 나타내는 78개의 암 유래 세포주 패널과 함께, 유방(15개의 사례), 직장(20개의 사례), 신장(16개의 사례), 및 뇌(4개의 사례)의 원발성 종양에서 찾아내었다.

[0283] 돌연변이 EGFR 구조체의 기능적 분석(Functional Analysis of Mutant EGFR Constructs)

[0284] L858R 및 delL747-P753insS 돌연변이를, 지정부위 돌연변이유발법(site-directed mutagenesis)을 이용하여 전장 *EGFR* 코딩 서열에 도입하여, 시스토메갈로바이러스(cytomegalovirus)-유래 발현 구조체(pUSE, Upstate)에 삽입하였다. Cos-7 세포를, 상기 발현 구조체 1㎍을 사용하여 트랜스펙션시킨 다음, 18시간 후, DMEM 결핍 소태아혈청내에서 5×10^4 세포/웰(12웰 플레이트, Coster)로 재평판배양하였다. 16시간의 혈청기아(serum starvation) 후, 10ng/㎖의 EGF(SIGMA)로 세포를 자극하였다. 제피티니브 저해를 확증하기 위해, EGF 첨가(10ng/㎖의 EGF로 30분 자극) 3시간 전에, 상기 약제를 배양 배지에 첨가하였다. 100㎕의 Laemmli 용해 완충액(lysis buffer)을 사용하여 세포용해물(cell lysate)를 제조한 후, 10% SDS-PAGE 상에서 단백질을 확인(resolution)한 다음, PVDF 막에 전달(transfer)한 후, 증강된 화학발광시약(Amersham)을 사용하여 웨스턴 블롯 분석(Western blot analysis)을 행하였다. EGFR의 자가인산화는 포스포티로신 Y-1068에 대한 항체를 사용하여 측정하고, 필적하는 단백질 발현은 항-EGFR 항체(anti-EGFR antibody)를 사용하여 나타내었다(조작 농도 1:1000; Cell Signaling Technology)

[0285] 돌연변이 분석

[0286] 중합효소연쇄반응(polymerase chain reaction:PCR)을 이용하여, 원발성 종양 조직 또는 종양 유래 세포주로부터 단리한 DNA를 사용하여, EGFR 유전자를 포함하는 28개의 엑손을 증폭시켰다. 사용된 프라이머쌍은 다음과 같다: 엑손 1,

CAGATTTGGCTCGACCTGGACATAG (센스) (서열번호 : 513) 및
CAGCTGATCTCAAGGAAACAGG (안티센스) (서열번호 : 514); 엑손 2,
GTATTATCAGTCAC TAAAGCTCAC (센스) (서열번호 : 515) 및
CACACTTCAAGTGGAATTCTGC (서열번호 : 516); 엑손 3, CTCGTG
TGCATTAGGGTTCAACTGG (센스) (서열번호 : 517) 및
CCTTCTCCGAGGTGGAATTGAGTGAC (안티센스) (서열번호 : 518); 엑손 4,
GCTAATTGCGGGACTCTTGTTCGCAC (센스) (서열번호 : 519) 및

[0287]

TACATGC TTTTCTAGTGGTCAG (안티센스) (서열번호 520); 엑손 5, GGTCTCAAGTGATTCTACAAACCAG (센스) (서열번호 : 521) 및 CCTTCACCTACTGGTTCACATCTG (안티센스) (서열번호 : 522); 엑손 6, CATGGT TTGACTTAGTTTGAATGTGG (센스) (서열번호 : 523) 및 GGATACTAAAGATACTTTGTCAC CAGG(안티센스) (서열번호 : 524); 엑손 7, GAACACTAGGCTGCAAAGACAGTAAC (센스) (서열번호 : 525) 및 CCAAGCAAGGCAAACACATCCACC(안티센스) (서열번호 : 526); 엑손 8, GGAGGATGGAGCC TTTCCATCAC (센스) (서열번호 : 527) 및 GAAGAGGAAGATGTGTTCCTTTGG (안티센스) (서열번호 : 528); 엑손 9 및 10, GAATGAAGGATGATGTGGCAGTGG (센스) (서열번호 : 529) 및 CAAAACATCAGCC ATTAACGG (안티센스) (서열번호 : 530) ; 엑손 11, CCACTTACTGTTCATATAATACAGAG (센스) (서열번호 : 531) 및 CATGTGAGATAGCATTTGGGAATGC (안티센스) (서열번호 : 532) ; 엑손 12, CATGACCT ACCATCATTGGAAAGCAG (센스) (서열번호 : 533) 및 GTAATTTCACAGTTAGGAATC (센스) (서열번호 : 534) ; 엑손 13, GTCACCCAAGGTCATGGAGCACAGG (센스) (서열번호 : 535) 및 CAGAATGC CTGTAAAGCTATAAC (안티센스) (서열번호 : 536) ; 엑손 14, GTCCTGGAGTCCCAACTCCTTGAC (센스) (서열번호 : 537) 및 GGAAGTGGCTCTGA TGGCCGTCCTG (안티센스) (서열번호 : 538) ; 엑손 15, CCAC TCACACACACTAAATATTTTAAG (센스) (서열번호 : 539) 및 GACCAAAACACCTTAAGTAA CTGACTC (안티센스) (서열번호 : 540) ; 엑손 16, CCAA TCCAACATCCAGACACATAG (센스) (서열번호 : 541) 및 CCAGAGCCATAGAAACTTGATCAG (안티센스) (서열번호 : 542) ; 엑손 17, GTATGGACTATGGC ACTTCAATTGCATGG (센스) (서열번호 : 543) 및 CCAGAGAACATGGCAACCAGCACAGGAC (안티센스) (서열번호 : 544) ; 엑손 18, CAAATGAGCTGGCAAGTGCCGTGTC (센스) (서열번호 : 545) 및 GAGTTT CCCAAACACTCAGTGAAAC (안티센스) (서열번호 : 546) 또는 CAAGTGCCGTGTCCTGGCACCCAAGC (센스) (서열번호 : 675) 및 CCAAACACTCAGTGAAACAAAGAG (안티센스) (서열번호 : 676); 엑손 19 GCAATATCAGCC TTAGG TGCGGCTC (센스) (서열번호 : 547) 및 CATAGAAAGTGAACATTTAGGATGTG (안티센스) (서열번호 : 548) ; 엑손

[0288]

20, CCATGAGTACGTATTTTGAAACTC (센스) (서열번호 : 549) 및
CATATCC CCATGGC AAACTCTTGC (안티센스) (서열번호 : 550); 엑손 21,
CTAACGTTCGCCAG CCATAAGTCC (센스) (서열번호 : 551) 및
GCTGCGAGCTCACCCAGAATGTCTGG (안티센스) (서열번호 : 552); 엑손 22,
GACGGG TCCTGGGGTGATCTGGCTC (센스) (서열번호 : 553) 및
CTCAGTACAATAGATAGACAGCAATG (안티센스) (서열번호: 684); 엑손
23, CAGGACTACAGAAATGTAGGTTTC (센스) (서열번호 : 555) 및
GTGCCTG CCTTAAGTAATGTGATGAC (안티센스) (서열번호 : 556); 엑손
24, GACTGG AAGTGTCGCA TCACCAATG (센스) (서열번호: 557) 및
GGTTTAATAATGCGATCTGGGACAC (안티센스) (서열번호 : 558); 엑손 25,
GCAGCTATAATTTAGAGAACCAAGG (센스) (서열번호: 559) 및 GGTT
AAAATTGACTTC ATTTCCATG (안티센스) (서열번호 : 560); 엑손 26,
CCTAGTTGCTCTAAA ACTAACG (센스) (서열번호 : 561) 및
CTGTGAGGCGTGACAGCCGTGCAG (안티센스) (서열번호: 562); 엑손 27,
CAACCTACTAATCAG AACCAGCATC (센스) (서열번호 : 563) 및
CCTTCACTGTGTCTGC AAATCTGC (안티센스) (서열번호 : 564); 엑손 28,
CCTGTCATAAGTCTCCTTGTTGAG (센스) (서열번호 : 565) 및
CAGTCTGTGGGTCTAAG AGCTAATG (안티센스) (서열번호 : 566).

[0289]

[0290] 결합(anealing) 온도는 58℃(엑손 1, 3, 4, 7-10, 12-25, 27, 및 28), 56℃(엑손 2, 5, 6, 및 26), 또는 52℃(엑손 11)이었다.

[0291] 보관 종양 조직(archival tumor tissue)으로부터 추출한 DNA의 네스티드(nested) PCR 증폭은 다음과 같이 수행하였다. 엑손 2, 5, 6, 7, 11, 12, 14, 16, 18, 19, 20, 21, 23, 24, 25, 26, 및 27에 대한 초기 PCR은 상기한 프라이머 및 조건을 이용하여 행하였다. 이어서, 이 반응액의 2㎕를, 하기의 안쪽(internal) 프라이머쌍을 사용하여 2차 PCR에 의해 증폭시켰다: 엑손 2,

CAGGAATGGGTGAGTCTCTGTGTG (센스) (서열번호 : 567) 및
GTGGAATTCTGCCCAGGCCTTTC (안티센스) (서열번호 : 568); 엑손 5,
GATTCTACAAACCA GCCAGCCAAAC (센스) (서열번호 : 569) 및
CCTACTGGTTCACATCTGACCCTG (안티센스) (서열번호 : 570); 엑손 6,
GTTTGAATGTGGTTTCGTTGGAAG (센스) (서열번호 : 571) 및
CTTTGTCACCAGG CAGAGG GCAATATC (안티센스) (서열번호 : 572); 엑손

[0292]

7, GACAGTAACTTGGGCTTTCTGAC (센스) (서열번호 : 573) 및
CATCCACCCAAAGACTCTCCAAG (안티센스) (서열번호 : 574); 엑손 11,
CTGTTCATA TAATAC AGAGTCCCTG (센스) (서열번호 : 575) 및
GAGAGATGCAGGAGCTCTGTGC (안티센스) (서열번호 : 576); 엑손 12,
GCAGTTTGTAGTCAATCAAAGGTGG (센스) (서열번호 : 577) 및
GTAATTTAAATGGGAAT AGCCC (안티센스) (서열번호 : 578); 엑손 14,
CAACTCCTTGACCATTACCTCAAG (센스) (서열번호 : 579) 및
GATGGCCGTCCTGCCCACACAGG (안티센스) (서열번호 : 580); 엑손 16,
GAGTAGTTTAGCA TATATTGC (센스) (서열번호 : 581) 및
GACAGTCAGAAATGCAGGAAAGC (안티센스) (서열번호 : 582); 엑손 18,
CAAGTGCCGTGTCCTGGCACCCAAGC (센스) (서열번호 : 583) 및
CCAAACACTCA GTGAAACAAAGAG (안티센스) (서열번호 : 584) 또는
GCACCCAAGCCCATGCCGTGGCTGC (센스) (서열번호 : 677) 및
GAAACAAAGAGTAAAGTAGATGATGG (안티센스) (서열번호 : 678); 엑손
19, CCTTAGGTGCGGCTCCACAGC (센스) (서열번호 : 585) 및
CATTTAGGATGTGGAGATGAGC (안티센스) (서열번호 : 586); 엑손 20,
GAAACTCAAG ATCGCATTCATGC (센스) (서열번호 : 587) 및
GCAAACTCTTGCTATCCCAGGAG (안티센스) (서열번호 : 588); 엑손 21,
CAGCCATAAGTCCTCGACGTGG (센스) (서열번호 : 589) 및
CATCCTCCCCT GCATGTGTTAAAC (안티센스) (서열번호 : 590); 엑손 23,
GTAGGTTTCTAAACATCAAGAAAC (센스) (서열번호 : 591) 및
GTGATGACATTTCTCCAGGGATGC (안티센스) (서열번호 : 592); 엑손 24,
CATCACCA ATGCCTTCTTTAAGC (센스) (서열번호 : 593) 및
GCTGGAGGGTTTAATAATGCGATC (안티센스) (서열번호 : 594); 엑손 25,
GCAAACACACAGGCACCTGCTGGC (센스) (서열번호 : 595) 및 CATTTC
CATGTGAGTTTCACTAGATGG (안티센스) (서열번호 : 596); 엑손 26,
CACCTTCACAATATACCCTCCATG (센스) (서열번호 : 679) 및
GACAGCCGTGCAGGGAAAAACC (안티센스) (서열번호 : 680); 엑손 27,
GAACCAGCATCTCAAGGAGATCTC (센스) (서열번호 : 681) 및
GAGCACCTGGCTTGGACACTGGAG (안티센스) (서열번호 : 682).

[0293]

[0294] 나머지 엑손에 대한 네스티드 PCR 증폭은 하기 프라이머를 사용하여 1차 PCR에 의해 수행되었다: 엑손 1, GACCGGACGACAGGCCACCTCGTC (센스) (서열번호 : 597) 및 GAAGAACGAAACGTCCCGTTCCTCC (안티센스) (서열번호 : 598); 엑손 3, GTTGAGCACT CGTGTGCATTAGG (센스) (서열번호 : 599) 및 CTCAGTGCACGTGTACTGGGTA (안티센스) (서열번호 : 600); 엑손 4, GTTCACTGGGCTAATTGCGGGACTCTTGTTCGCAC (센스) (서열번호 : 601) 및 GGTA AATACATGCTTTTCTAGTGGTCAG (안티센스) (서열번호 : 602); 엑손 8, GGAGGATGGA GCCTTTCCATCAC (센스) (서열번호 : 603) 및 GAAGAGGAAGATGTGTTCCTTTGG (안티센스) (서열번호 : 604); 엑손 9, GAATGAAGGATGATGTGGCAGTGG (센스) (서열번호 : 605) 및 GTATGTGTGAAGGAG TCACTGAAAC (안티센스) (서열번호 : 606); 엑손 10, GGTGAGTCACAGGTTCAGTTGC (센스) (서열번호 : 607) 및 CAAAACATCAGCCATTAACGG (안티센스) (서열번호 : 608); 엑손 13, GTAGCCAGCATGTC TGTGTCAC (센스) (서열번호 : 609) 및 CAGAATGCCTGTAAAGCTATAAC (안티센스) (서열번호 : 610); 엑손 15, CATTTGGCTTTCCCCACTCACAC (센스) (서열번호 : 611) 및 GACCAAAACACCTTAA GTAACTGACTC (안티센스) (서열번호 : 612); 엑손 17, GAAGCTACATAGTGTCTCACTTTCC (센스) (서열번호 : 613) 및 CACAACTGCTAATGGCCCGTTCTCG (안티센스) (서열번호 : 614); 엑손 22, GAGCAGCCCTGAACTCCGTCAGACTG (센스) (서열번호 : 683) 및 CTCAGTACAATAGATAGACAGCAATG (안티센스) (서열번호 : 684); 엑손 28a GCTCC TGCTCCCTGTCATAAGTC (센스) (서열번호 : 615) 및 GAAGTCCTGCTGGTAGTCAGGGTTG (안티센스) (서열번호 : 616); 엑손 28b, CTGCAGTGGGCAACCCCGAGTATC (센스) (서열번호 : 617) 및 CAGTC TGTGGGTCTAAGAGCTAATG (안티센스) (서열번호 : 618).

[0295]

[0296] 2차 PCR 증폭은 하기 프라이머쌍을 사용하여 수행하였다: 엑손 1, GACAGGCCACCTCGTCGGCGTC (센스) (서열번호 : 619) 및 CAGCTGATCTCAAGGAAACAGG (안티센스) (서열번호 : 620); 엑손 3, CTCGTG TGCATTA GGGTTCAACTGG (센스) (서열번호 : 621) 및 CCTTCTCCGAGGTGGAATTGAGTGAC (안티센스) (서열번호 : 622); 엑손 4,

[0297]

GCTAATTGCGGGACTCTTGTTCGCAC (센스) (서열번호 : 623) 및
TACATGCTTT TCTAGTGGTCAG (안티센스) (서열번호 : 624); 엑손 8,
CCTTTCCATCACCCCTCAAGAGG (센스) (서열번호 : 625) 및
GATGTGTTCCTTTGGAGGTGGCATG (안티센스) (서열번호 : 626); 엑손 9,
GATGTGG CAGTGGCGGTTCCGGTG (센스) (서열번호 : 627) 및
GGAGTCACTGAAACAAACAACAGG (안티센스) (서열번호 : 628); 엑손 10,
GGTTCAGTTGCTTGTATAAAG (센스) (서열번호 : 629) 및
CCATTAACGGT AAAATTTCAGAAG (안티센스) (서열번호 : 630); 엑손 13,
CCAAGGTCATGGAGCACAGG (센스) (서열번호 : 631) 및
CTGTAAAGCTATAACAACAACCTGG (안티센스) (서열번호 : 632); 엑손 15,
CCACTCACA CACACTAAATATTTTAAG (센스) (서열번호 : 633) 및
GTAACTGACTCAAATACAAACCAC (안티센스) (서열번호 : 634); 엑손 17,
GAAGCTACATAGTGTCTCACTTTCC (센스) (서열번호 : 635) 및 CACAA
CTGCTAATGGCCCGTTCTCG (안티센스) (서열번호 : 636); 엑손 22,
GACGGGTCCTGGGGTGATCTGGCTC (센스) (서열번호 : 685) 및
CTCAGTACAATAGATAGACAGCAATG (안티센스) (서열번호 : 686); 엑손
28a, CCTGTCATAAG TCTCCTTGTTGAG (센스) (서열번호 : 637) 및
GGTAGTCAGGGTTGTCCAGG (안티센스) (서열번호 : 638); 엑손 28b,
CGAGTATCTCAACACTGTCCAGC (센스) (서열번호 : 639) 및

[0298] CTAAGAGCTAATGCGGGC ATGGCTG (안티센스) (서열번호 : 640).

[0299] 엑손 1의 증폭을 위한 결합 온도는 54℃였다. 1차 및 2차 증폭 둘다에서의 결합온도는 58℃(엑손 3, 4, 7-10, 12-17, 19-25, 27, 및 28), 56℃(엑손 2, 5, 6, 및 26), 또는 52℃(엑손 11 및 18)였다.

[0300] PCR 엠프리콘(amplicon)은 서열분석하기 전에 엑소누클레아제 I(United States Biochemical, Cleveland, OH), 및 쉬림프 알칼린 포스파타제(United States Biochemical, Cleveland, OH)를 사용하여 정제하였다. 그 다음, 정제된 DNA를 희석하여, ABI BigDye Terminator kit v1.1(ABI, Foster City, CA)를 사용하여 제조자의 지시에 따라 사이클-염기서열분석을 행하였다. 염기서열분석 반응액을 ABI3100 유전자 분석기 상에서 전기영동하였다. 이형접합성 위치를 확인하기 위해 Factura와 Sequence Navigator software를 함께 사용하여, 센스 및 안티센스 양방향으로 돌연변이이의 존재에 대하여 전기영동도를 분석하였다. 모든 염기서열 변이는 다중 독립 PCR 증폭 및 염기서열분석에 의해 확인하였다.

[0301] 암 유래 세포주(Cancer-Derived Cell Lines):

[0302] 14개의 폐암 유래 세포주 패널의 *EGFR* 돌연변이에 대해 분석하였다. 이들은 NSCLC(N=5), 소세포 폐암(N=6), 선편평상피 세포암(N=1), 기관지 유암종(N=1), 및 알려지지 않은 세포조직(N=1)의 종양으로부터 유래한 것이다. 특이적 세포주는 NCI-H460, NCI-522, HOP-92, NCI-H841, NCI-H734, NCI-H2228, NCI-H596, NCI-H727, NCI-H446, NCI-H1781, NCI-H209, NCI-H510, NCI-H82, NCI-H865 였다. 또한, 64개의 암 유래 세포주가 엑손 19 및 21에서 돌연변이를 갖는 것으로 스크리닝되었다. 이들은 하기의 세포조직을 나타내었다: 유방암(BT549, BT483, UACC893, HS467T, HS578T, MCF7, MCF7-ADR, MDA-MB-15, MDA-MB-175, MDA-MB-231, MDA-MB-415, MDA-MB-436, MDA-MB-453, MDA-MB-468, T47D), 난소암(ES-2, IGROV-1, MDAH2774, OV1063, OVCAR3, OVCAR4, OVCAR5, SKOV3, SW626), CNS 암류(SF-295, SNB-19, U-251, CCF-STTG1, SW-1088, SW-1783, T98G, M059K, A172, SK-N-DA, SK-N-MC), 백혈병(CCRF-CEM, K562, MOLT-4, RPMI8226, SR), 전립선암(DU-145, PC-3), 직장암(COLO-205, HCT-116, HCT-15, HT-29, SW-620), 신장암(786-O, ACHN, CAKI-1, SN-12C, UO-31), 악성흑색종(LOX-IMVI, M14, SKMEL2, UACC-62), 뼈육종(SAOS-2), 및 머리 및 목암(O11, O13, O19, O28, O22, O29, O12). 상기 머리 및 목암 세포주는 메사추세츠 제너럴 하스파탈/메사추세츠 눈 및 귀 진료소(Massachusetts Eye and Ear Infirmary)의 제임스 로코 박사로부터 제공받았다. 모든 다른 세포주는 ATCC(American Type Culture Collection)(Manassa, VA)를

통해서 얻을 수 있다.

[0303] 게놈 DNA는 스냅-동결(snap-frozen) 종양 표본으로부터 단리하였다. 종양 표본은 먼저 미리차게해서 살균해둔 막자사발을 사용하여 미세 분말로 파쇄하였다. 파쇄한 종양조직을, 즉시 100mM 염화나트륨, 10mM Tris pH7.5, 25mM EDTA(소디움 에틸렌디아민 테트라아세테이트) pH8.0, 및 0.5%(w/v) 소디움도데실설페이트, 및 100㎍/㎖ 프레쉬 프로테나제 K로 구성된 DNA 추출용액에 옮겨넣고, 37℃에서 하룻밤 배양, 또는 50℃에서 3시간 동안 배양하였다. 그 다음, DNA는 표준 페놀-클로로포름법을 사용하여 추출하여, 에탄올에 침전시킨 후, 70% 에탄올로 세정하여, 풍건한 다음, TE 완충액에 재현탁시켰다. 상기 DNA 농도는 분광광도분석에 의해 결정하였다. 인간 EGFR의 엑손 19 및 21은 하기 프라이머쌍을 사용하여 PCR에 의해 증폭시켰다:

[0304] 엑손 19 센스 프라이머, 5'-GCAATATCAGCCTTAGGTGCGGCTC-3'(서열번호 505);

[0305] 엑손 19 안타센스 프라이머, 5'-CATAGAAAGTGAACATTTAGGATGTG-3'(서열번호 506);

[0306] 엑손 21 센스 프라이머, 5'-CTAACGTTCGCCAGCCATAAGTCC-3'(서열번호 507);

[0307] 엑손 21 안티센스 프라이머, 5'-GCTGCGAGCTCACCCAGAATGTCTGG-3'(서열번호 508).

[0308] 각 시료에 대해서, 1X Expand Long Template 완충액 1(Roche, Mannhein Germany), 50μM의 서열분석 등급 dATP(Amersham Biosciences, Cleveland OH), 50μM의 서열분석 등급 dCTP(Amersham Biosciences, Cleveland OH), 50μM의 서열분석 등급 dGTP(Amersham Biosciences, Cleveland OH), 50μM의 서열분석 등급 dTTP(Amersham Biosciences, Cleveland OH), 0.2μM의 센스 프라이머, 0.2μM 안티센스 프라이머, 1/6 용량의 TaqStart 항체(1.1㎍/㎕)와 함께 얼음속에서 5분간 전배양한 1.25 유니트 Expand Long Template 효소 혼합액(Taq DNA 폴리머라제/Tgo DNA 폴리머라제)(Roche, Mannhein Germany), 및 물로 구성된 최종 용량 25㎕의 PCR 반응액에서, 20ng의 게놈 DNA를 증폭시켰다. 또한, 각각의 일련의 증폭반응은 DNA 주형을 뺀 음성 대조군을 포함한다. 양 엑손에 대한 PCR 반응 조건은, MJ-Research PTC-200 또는 PTC-225 thermal-cycler(MJ-Research, Waltham MA) 상에서, 95℃에서 2분간 1주기 반응한 후; 95℃에서 30초, 58℃에서 30초, 및 72℃에서 45초를 1주기로 하여 40주기 반응시킨 다음; 72℃에서 10분간 최종 연장 반응 후, 4℃에서 유지하는 것으로 하였다.

[0309] 환자 재료로부터의 증폭 및 음성 대조군으로부터의 미증폭을 확인하기 위해, PCR 산물을 0.8% 아가로스 겔상에서 전기영동하여 분석하였다. PCR 산물은 염기서열 분석 전에, 각 PCR 엠프리콘 10㎕와 0.5㎕의 엑소누클레아제 I(10U/㎕)(Amersham Biosciences, Cleveland OH), 및 1㎕의 쉬림프 알칼린 포스파타제(1U/㎕)(Amersham Biosciences, Cleveland OH)를 혼합하고, 37℃에서 20분간 배양한 후, thermal-cycler(MJ-Research, Waltham MA) 상에서 80℃에서 15분 동안 불활성화시켜, 정제하였다. 정제된 DNA는, 상기 엠프리콘의 강도에 따라, 물에 희석시킨 다음, ABI BigDye Terminator kit v1.1(Applied Biosystems, Foster City, CA)을 사용하여 제조자의 지시에 따라 사이클-염기서열분석을 수행하였다. 사이클-염기서열분석은 MJ-Research thermal-cycler상에서 하기 사이클링 조건으로 수행하였다: 염기서열 분석용 프라이머는 다음과 같다:

[0310] 엑손 19 센스 프라이머, 5'-GCAATATCAGCCTTAGGTGCGGCTC-3'(서열번호 505);

[0311] 엑손 19 안티센스 프라이머, 5'-CATAGAAAGTGAACATTTAGGATGTG-3'(서열번호 506);

[0312] 엑손 21 센스 프라이머, 5'-CTAACGTTCGCCAGCCATAAGTCC-3'(서열번호 507) 또는 5'-CGTGGAGAGGCTCAGAGCCTGGCATG-3'(서열번호 687);

[0313] 엑손 21 안티센스 프라이머, 5'-GCTGCGAGCTCACCCAGAATGTCTGG-3'(서열번호 508).

[0314] 염기서열분석 반응액은 ABI3100 유전자 분석기(Applied Biosystems, Foster City, CA)상에서 전기영동하였다. Factura 및 Sequence Navigator(Applied Biosystems, Foster City, CA) 소프트웨어 프로그램을 사용하여, 잠정적인 이형접합성 위치를 표시하고 그들을 평가 결과를 표시하였다(display). 2차 피크의 높이가 1차 피크의 높이보다 높거나, 또는 동등하거나, 30%인 누클레오티드 위치는 이형접합성으로 표시하였으며, 센스 및 안티센스 판독 둘다를 분석하여 확인하였다. 돌연변이의 존재를 나타내는 서열을 가진 시료는 다시 증폭시킨 후, 확인용 염기서열분석을 행하였다.

[0315] 엑손 19 및 21에 대한 염기서열분석에 사용된 프라이머의 위치

[0316] 인트론 프라이머는 소문자로 아래에 줄을 그어서 나타내었다.

[0317] 인트론 염기서열은 소분자로 나타내었다.

[0318] 엑손 염기서열은 대문자로 나타내었다.

[0319] EGFR 엑손 19(5'-3')(서열번호 641)

gcaatatcagccttaggtgcggctccacagccccagtgtccctcaccttcggggtgcatcgc
tggtaacatccacccagatcactgggcagcatgtggcaccatctcacaattgccagttaacg
tcttccttctctctctgtcatagGGACTCTGGATCCCAGAAGGTGAGAAAGTTAAAATTCCC
GTCGCTATCAAGGAATTAAGAGAAGCAACATCTCCGAAAGCCAACAAGGAAATCCTCGATgt
gagtttctgctttgctgtgtgggggtccatggctctgaacctcaggcccacctttctcatg
tctggcagctgctctgctctagaccctgctcatctccacatcctaaatgttcactttctatg

[0320]

[0321] EGFR 엑손 21(5'-3')(서열번호 642) 또는 (서열번호 687)

ctaacgttcgccagccataagtcctcgacgtggagaggctcagagcctggcatgaacatgac
cctgaattcggatgcagagcttcttcccatgatgatctgtccctcacagcagggtcttctct
gtttcagGGCATGAACTACTTGGAGGACCGTCGCTTGGTGCACCGCGACCTGGCAGCCAGGA
ACGTACTGGTGAAAACACCGCAGCATGTCAAGATCACAGATTTTGGGCTGGCCAAACTGCTG
GGTGCGGAAGAGAAAGAATACCATGCAGAAGGAGGCAAAgtaaggaggtggctttaggtcag
ccagcattttcctgacaccagggaccaggctgccttcccactagctgtattgtttaacacat
gcaggggaggatgctctccagacattctgggtgagctcgcagc

[0322]

[0323] **결과**

[0324] 제피티니브 반응자의 임상적 특성

[0325] 메사추세츠 제너럴 하스피탈에서 2000년 이후, 진행성, 화학요법-난치성 NSCLC를 가진 환자는 단일 약제 제피티니브로 치료해왔다. 총 275명의 환자가, FDA에 의해 2003년 5월에 승인받기 전, 특례적 이용의 이용범위 확장 프로그램(compassionate use expanded access program)의 일부로서 치료받았다. 후에 그 날짜를 시판공급(commercial supply)일로 사용하였다. 이 기간 동안, 25명의 환자가 유의한 임상 반응을 갖는 것으로 임상의에 의해 확인되었다. 유의한 임상 반응은, RECIST 기준을 이용하여 측정가능한 병변을 가진 환자, 또는 이 기준을 이용하여 측정할 수 없는 종양 존재량을 가진 환자에 대해 부분 반응(partial response)으로서 규정하였으며, 평가가능한 반응은 두명의 의사가 평가하였다. 표 1은 초기 진단 시에 얻어진 이용가능한 종양 표본 9개의 사례에 대한 임상 특성을 나타낸다. 나머지 제피티니브 반응자에 대해서는, 조직(tissue)을 이용할 수 없으며, 그 이유는 대부분 진단 표본이 바늘 흡인액으로부터 얻은 세포검사에 제한되었기 때문이었다. 군(group)으로서, 9명의 환자는 제피티니브로부터 실질적인 이점을 경험하였다. 약물 치료의 개시로부터의 정중 생존율(median survival)은 18개월이 넘었으며, 치료의 정중 지속기간(median duration)은 16개월보다 길었다. 이전 보고서와 마찬가지로, 제피티니브 반응자는 여성, 흡연이력의 부재, 및 기관지폐포 조직을 가진 종양에서 높은 유병율(prevalence)을 갖는다(11, 12). 사례 6은 제피티니브 반응성 코호트(gefitinib-responder cohort)를 나타낸다. 이 환자는 32살 남자로, 흡연이력이 없고, 다발성 뇌 병변 및 오른쪽 폐에 기관지폐포 선암종으로 진단된 병변이 존재하였다. 이 남자는 전체 뇌 방사선요법으로 치료받은 다음, 그 종양이 반응하지 않은 일련의 화학요법(카르보플라틴 및 겜시타빈; 도세탁셀; 비노렐빈)에 의해 치료받았다. 쇠퇴하는 기능 상태 및 진행성 폐암 종양 존재량에 따라서, 1일당 250mg의 제피티니브로 치료를 시작하였다. 그의 숨참이 급속히 개선되었으며, 치료 개시 6주 후, 폐 CT 스캔에서는 도 1에 나타내는 바와 같이 극적인 개선을 나타내었다.

[0326] 제피티니브 반응자에서의 EGFR 돌연변이

[0327] 본 발명자들은, 제피티니브에 현저한 반응성을 가진 NSCLC의 사례들은 이들 종양에서 성장인자 신호전달 경로에 의해 필수적인 역할을 하는 것으로 지적된 EGFR에서의 체세포 돌연변이를 보유한다는 가설을 세웠다. 이러한 돌연변이를 찾기 위해, 본 발명자들은 먼저, 신경아교종(glioma) 특성을 가진 EGFR의 세포외 도메인

(extracellular domain)내에서의 재배열(rearrangement)에 대해 시험하였다: 그 결과, 아무것도 검출되지 않았다. 따라서, 본 발명자들은 개별 엑손의 PCR 증폭을 이용하여 유전자의 전체 코딩 영역의 염기서열을 분석하였다. 9개 중 8개의 사례에서 이형접합성 돌연변이가 확인되었으며, 이들 모두 EGFR의 키나제 도메인내에 군집되어 있었다(표 2 및 도 2). 4개의 종양은 아미노산 746~750(delE746-A750; 사례 1), 747~750(delL747-T751insS; 사례 2), 및 747~752(delL747-P753insS; 사례 3 및 4)를 제거하는 인프레임 결실(in-frame deletion)을 가지고 있었다. 상기 결실 중 후자 2개는 결실 변곡점(deletion breakpoint)에서의 신규 코돈의 생성 결과 일어난, 세린 잔기의 삽입(insertion)과 관련이 있었다. 특히, 이들 4개의 결실은 중첩되어있으며, 모든 사례에서 공유하는 엑손 19내에 4개의 아미노산 결실(류신, 아르기닌, 글루탐산 및 알라닌, 코돈 747~750)을 가지고 있었다. 또 다른 세개의 종양은 엑손 21에 아미노산 치환을 가지고 있었다: 코돈 858의 류신이 아르기닌으로 치환(L858R; 사례 5 및 6), 코돈 861의 류신을 글루타민으로 치환(L861Q; 사례 7). 상기 L861Q 돌연변이는, 마우스 *egfr* 유전자에서의 동일한 아미노산의 변화가 다크 스킨(dsk5) 성향의 원인이 되기 때문에, 변경된 EGFR 신호전달(altered EGFR signaling)과 관련하여, 특히 중요하다(18). 키나제 도메인내의 4번째 미스센스 돌연변이는 엑손 18내의 코돈 719의 글리신이 시스테인으로 치환된 것이다(G719C; 사례 8). 사례 1, 4, 5 및 6에 대해서는 대응하는 정상 조직을 이용할 수 있었으며, 이 정상조직은 단지 야생형 염기서열만을 나타내었으며, 이는 종양을 형성하는 동안, 상기 돌연변이가 체세포적으로 일어남을 의미한다. 제피티니브에 대한 응답에 실패한 NSCLC의 7개의 사례에서는 돌연변이가 확인되지 않았다(P=0.0007; 2-면 피셔 정확검증(2-sided Fisher's exact test)).

[0328] NSCLC 및 기타 암 종류에서의 특이적 EGFR 돌연변이의 유병율

[0329] EGFR 세포외 도메인에 영향을 미치는 재배열에 대해 집중적으로 연구된 신경아교종(gliomas)과 달리(15), NSCLC에서의 *EGFR* 돌연변이 빈도는 밝혀지지 않았다. 따라서, 본 발명자들은, 앞선 임상시험에서는 제피티니브 반응성과 관련있었던, 기관지폐포 세포조직을 가진 15개의 사례를 포함하여, 제피티니브 연구와 관련없는 NSCLC의 25개의 1차 사례들에 대해 유전자의 전체 코딩 영역의 염기서열을 분석하였다(11,12). 이형접합성 돌연변이는 두개의 기관지폐포암에서 검출되었다. 두 사례는 제피티니브 반응자에서 발견되었던 것과 동일한 키나제 도메인에서의 인프레임 결실, 즉 delL747-P753insS 및 delE746-A750(표 2)을 가지고 있었다. *EGFR* 돌연변이의 확실한 군집화(clustering)가 보이면, 총 55개의 원발성 종양 및 78개의 암 유래 세포주의 엑손 19 및 21의 염기서열을 분석하였다(보충 재료 참조). 돌연변이가 검출되지 않은 경우에는, 이들이 암의 서브세트(subset)에서만 일어났음을 암시하며, 이때, EGFR 신호전달은 종양발생에 결정적인 역할을 할 수 있다.

[0330] 돌연변이 EGFR 단백질의 EGF-유도 활성화 및 제피티니브 저해의 증가

[0331] 이 돌연변이들에 의해 코딩되는 기능적 특성을 연구하기 위해, L747-S752insS 결실 및 L858R 미스센스 돌연변이를 배양세포에서 발현시켰다. 야생형 및 돌연변이 구조체(construct)의 Cos-7 세포로의 과도 트랜스펙션(transient transfection)은 등가의 발현 수준을 나타내었으며, 이는 상기 돌연변이가 단백질의 안정성에 영향을 미치지 못함을 나타낸다. EGFR 활성화는 수용체 자가인산화의 표지자로서 통상 사용되는 티로신^{1068}잔기의 인산화를 측정함에 의해 정량화하였다(19). 혈청 및 관련 성장인자의 부재 시, 야생형 EGFR이나 돌연변이 EGFR 어느 것도 자가인산화를 나타내지 않았다(도 3a). 그러나, EGF를 첨가하면, 야생형 수용체와 비교하여, 미스센스 및 결실 *EGFR* 돌연변이 둘다에 대해서는 2~3배 증가된 수용체 활성화가 유도되었다. 또한, 정상 EGFR 활성화는 15분 후에, 수용체 내재화(internalization)와 맞춰, 하향조절(downregulate)된 반면, 상기 두개의 돌연변이 수용체는 3시간까지 활성화가 지속된 것으로 나타났다(도 3a). EGF 첨가 후에 항체로 EGFR 인산화를 측정했을 때 동일한 결과가 얻어졌다(도시하지 않음).

[0332] EGFR 키나제 돌연변이 8개 중 7개는, 제피티니브의 표적으로 되는 ATP 틈(cleft) 부근에 존재하기 때문에, 본 발명자들은 상기 돌연변이 단백질이 상기 저해제에 대한 민감성을 변경시켰는지의 여부를 결정하였다. EGF-유도 수용체 자가인산화가 다양한 농도의 제피티니브로 전처리한 세포에서 측정되었다. 특히, 양 돌연변이 수용체는 제피티니브에 의한 저해에 증가된 민감성을 나타내었다. 야생형 EGFR은 0.1μM의 IC_{50}을 가졌으며, 2μM의 제피티니브에서 자가인산화가 완전히 저해된 반면, 상기 두개의 돌연변이 단백질은 0.015μM의 IC_{50}을 가졌으며, 0.2μM에서 자가인산화가 폐기되었다(도 3b). 약물동태 연구에 의하면, 400~600mg의 제피티니브를 매일 경구투여하면 1.1~1.4μM의 평균 항정상태 최저 혈장농도(mean steady-state trough plasma concentration)로 되

지만, 현재 추천되는 250mg의 1일 투여량은 0.4μM의 평균 최저 농도로 됨이 나타나 있으므로, 약제 민감성에 있어서의 상기한 차이는 임상적으로 의미가 있다(20).

[0333] [실시예 2]

[0334] EGFR의 키나제 도메인내에 돌연변이를 보유하고 있고, 그 결과 제피니티브 치료에 의한 성장 저해에 민감한 종양 세포는 그 키나제 도메인내에서 "두번째 부위(second-site)" 돌연변이가 일어날 수 있으며, 이는 제피티니브에 대한 내성을 부여하는 반면, 야생형 EGFR에 비해 증가된 EGFR 신호전달(signaling)을 나타낸다는 의미에서 여전히 "활성화(activating)" 한다. 이러한 제피티니브 내성 돌연변이는 2개의 산발성(sporadic) 인간 NSCLC 세포주, 즉 NCI-1650 및 NCI-1975로부터 얻는다. 각 세포주는 상기 EGFR의 키나제 도메인을 가진 이형접합성 돌연변이(heterozygous mutation)를 포함하며, 그 결과, 제피티니브에 대해 민감성을 나타낼 것으로 예측된다. NCI-1650에서의 EGFR 돌연변이는 엑손 19내의 2235-2249 위치(delLE746-A750)에서의 15개의 누클레오티드의 인프레임 결실(in-frame deletion)로 이루어지고, 한편 NCI-1975는 엑손 21에 누클레오티드 2573의 T가 G로 치환된 미스센스 돌연변이(L858R)를 갖는다. 여기서 NCI-H1975에서의 L858R 돌연변이는 시험관내에서 활성화 상태이고, 제피티니브에 대한 증가된 민감성을 나타낸다.

[0335] NCI-1650 및 NCI-1975 유래의 제피티니브 내성 세포주는, EMS(ethyl methanesulfonate)를 사용하여 무작위로 화학적 돌연변이를 유발한 다음, 제피티니브 보충 배지에 배양하여 증식한 내성 클론(clone)을 선별하여 분리한다. 개개의 클론을 계대배양한 다음, 상기 EGFR 키나제 도메인에 대응하는 게놈 DNA의 특이적 PCR 매개 증폭(specific PCR-mediated amplification)을 행한 후, 그 EGFR 유전자의 누클레오티드 염기서열을 결정한다.

[0336] 이 방법의 하나의 변화로서, 제피티니브에 대한 내성을 부여하는 EGFR 유전자내에 자연적으로 획득된 돌연변이를 선별하기 위해서, 수주일 또는 수개월에 걸쳐 점차 증가된 농도의 제피티니브의 존재하에, 상기 두 세포주를 연속계대배양하는 것이 포함된다. 선별된 세포(상대적으로 높은 제피티니브 농도에서 지속적으로 증식하는 세포)는 콜로니로서 단리하고, 돌연변이는 상기한 바와 같이 동정하였다.

[0337] [실시예 3]

[0338] 수용체 티로신 키나제의 돌연변이가 NSCLC에서 원인되는 역할(causal role)을 하는지의 여부를 결정하기 위해서, 일본의 나고야 시립대학병원으로부터 얻은 58개의 시료와 메사추세츠주 보스톤의 브리검 여성병원(Brigham and Woman's Hospital)으로부터 얻은 61개의 시료로 이루어지는 119개의 원발성 NSCLC 종양 세트내에서의 체세포 유전자 변화(somatic genetic alteration)을 조사하였다. 상기 종양은, EGFR 키나제 저해제에 의한 치료 기록이 없는 74명의 남성과 45명의 여성으로부터 얻은 70개의 폐선암종(lung adernocarcinoma)과 49개의 다른 NSCLC 종양을 포함하였다.

[0339] 초기 스크린으로서, 41개의 폐선암종을 포함한 58개의 NSCLC 시료의 서브세트 유래의 게놈 DNA로부터 얻은 58개의 인간 수용체 티로신 키나제 유전자(*) 중 47개의, 활성화 루프를 코딩하는 엑손을 증폭시켜 염기서열을 분석하였다. 이 종양들 중 3개는, 모두 폐선암종이며, 동일한 환자의 정상 폐조직으로부터 얻은 DNA에서는 존재하지 않는 이형접합성 미스센스 돌연변이가 *EGFR*에 나타났다(표 S2:S0361, S0388, S0389). 다른 수용체 티로신 키나제 유전자 유래의 엠프리콘(amplicon)에서는 돌연변이가 검출되지 않았다. 세개의 종양 모두 동일한 *EGFR* 돌연변이를 가졌으며, 858 위치의 류신("L")이 아르기닌("R")으로 변화된것으로 예측하였다(도 6A; CTG→CGG;"L858R"). 여기서 모든 숫자는 인간 EGFR에 관련된 것이다.

[0340] 다음에, 119개의 NSCLC 종양의 전체 콜렉션내의 *EGFR*의 엑손 2~25를 시험하였다. 게놈 DNA의 엑손 염기서열분석 결과, 총 16개의 종양에서 *EGFR*의 미스센스 및 결실 돌연변이가 확인되었고, 돌연변이는 모두 키나제 도메인의 엑손 18~21내에 있었다. 이러한 군내에서의 모든 염기서열 변화는 종양 DNA내에서 이형접합성이었으며; 각각의 경우, 동일한 환자로부터 얻은 쌍을 이루는 정상 폐조직은 야생형 염기서열을 나타내었고, 상기 돌연변이는 체세포를 기원으로 함이 확인되었다. 누클레오티드와 단백질 서열 변화의 분포, 및 이들의 비정상성과 관련된 환자의 특성은 표 S2에 요약하여 나타내었다.

[0341] 치환 돌연변이 G719S 및 L858R은 각각 두개 및 세개의 종양에서 검출되었다. "G719S" 돌연변이는 719 위치의 글리신(G)이 세린(S)으로 변화한다(도 6b). 이 돌연변이들은 각각, 누클레오티드 트리포스페이트 결합 도메인 또는 P-루프의 GXGXXG 모티프(motif)내와, 활성화 루프(52)의 상당히 보존된 DFG 모티프에 인접한 부위에 위치

하고 있다(참조, 예를 들면 도 7). 상기 돌연변이가 일어난 잔기들은 모든 단백질 키나제에서 거의 변화하지 않으며, B-Raf 단백질 세린-트레오닌 키나제내의 유사 잔기들(G463 및 L596)은 직장, 난소 및 폐 암종에서 체세포적으로 변이된다(*41, 53*)(도 6A, 6B).

[0342] 또한, EGFR의 키나제 도메인내의 코돈 746~759를 포함하는 영역에 모여있는 다중 결실 돌연변이(multiple deletion mutation)를 검출하였다. 10개의 종양은, 누클레오티드 2235 또는 2236 위치에서 시작하여, EGFR 코돈 746~750을 제거하는 두개의 중첩된 15-누클레오티드 결실 중의 하나를 보유하고 있었다(Del-1; 도 6C 및 8C; 표 S2). 다른 종양으로부터 얻은 *EGFR* DNA는 코돈 752~759의 결실에 의한 이형접합성 24-누클레오티드 갭을 나타내었다(Del-2; 도 6c). 대표적인 크로마토그램(chromatogram)은 도 8a~8f에 나타낸다.

[0343] EGFR 키나제 도메인의 활성 형태의 3차원적 구조(*54*)내에서의 치환 돌연변이 및 Del-1 결실의 위치는 도 7에 나타낸다. 상기 서열 변화는 상기 키나제의 활성 부위 주변에 군집(cluster)되어 있으며, 상기 치환 돌연변이는 활성화 루프와 글리신이 풍부한 P-루프, 즉 많은 단백질 키나제에서의 자동조절에 중요한 것으로 알려진 구조적인 요소 주변에 위치한다(*52*).

[0344] 두개의 다른 종양 종류에서 두개의 부가적인 EGFR 돌연변이가 동정되었다. 즉, 급성 골수성 백혈병(AML)에서는 *EGFR* 돌연변이 G857V가, 전이성 육종에서는 *EGFR* 돌연변이 L883S가 동정되었다. 상기 "G857V" 돌연변이는 857 위치의 글리신(G)이 발린(V)으로 치환된 것이며, "L883S" 돌연변이는 883 위치의 류신(L)이 세린(S)로 치환된 것이다. 이러한 발견은, EGFR에서의 돌연변이가 여러 종류의 종양에서 일어나며, 무엇보다는 중요한 것은, EGFR 저해제가 이러한 돌연변이를 가진 환자를 치료함에 있어서 유효함을 암시한다. 이것은 예를 들면 티로신 키나제 저해제인 제피티니브(상표명 이레사(Iressa™)로 시판됨), 에를티니브(상표명 타르세바(Tarceva™)로 시판됨) 등의 사용, 및 NSCLC 이외의 종양 종류를 치료할 가능성을 확장시킨다.

[0345] 상기 *EGFR* 돌연변이는 일본계 및 미국계 환자 집단에서 기재된 상이한 환자 특성과 두드러진 상관 관계를 나타낸다. 상기한 바와 같이, 임상 시험 결과, 일본계 환자에서는 유럽계 우세 집단(predominantly European-derived population)에서보다 더 높은 반응성을 나타내고(27.5% 대 10.4%, in multi-institutional phase II trial); 부분적인 반응은 미국계의 여성, 비흡연자, 및 선암종을 가진 환자에서 더 빈번하게 보이는(*49-51*) 등의, 티로신 키나제 저해제 제피티니브(상표명 이레사)에 대한 응답에 있어서 유의한 변이성(variability)이 드러났다. *EGFR* 돌연변이는 다른 NSCLC(1/49 또는 2%)에서 보다 선암종(15/70 또는 21%)에서 더 빈번하게 일어나고; 남성(7/74 또는 9%) 보다는 여성(9/45 또는 20%)에서 더 빈번하게 일어나고; 미국계 환자(1/61 또는 2%, 및 선암종 1/29 또는 3%)보다 일본계 환자(15/58 또는 26%, 및 선암종 14/41 또는 32%)에서 더 빈번하게 일어났다. *EGFR* 돌연변이의 가장 높은 분율은 선암종을 가진 일본계 여성(8/14 또는 57%)에서 관찰되었다. 특히, *EGFR* 돌연변이의 존재와 상관관계가 있는 환자의 특성은 제피티니브 치료에 대한 임상적인 응답과 상관관계가 있는 것으로 나타났다.

[0346] *EGFR* 돌연변이가 제피티니브 민감성의 결정인자 인가를 조사하기 위해서, 제피티니브로 치료하는 동안 응답성으로 나타난 환자 5명과 진행성으로 나타난 환자 4명으로부터 사전치료 NSCLC 시료를 얻었으며, 이 환자들은, 제피티니브의 이용범위 확장 프로그램(expanded access program)상 또는 규제 승인(regulatory approval) 후에 다나-파버 암연구소(Dana-Farber Cancer Institute)에서 치료받은 125명 이상의 환자 중에서 선별하였다. 환자들 중의 4명은 부분적인 방사선 반응성(치료 2달 후 CT 스캔에서 50% 이상의 종양이 퇴행함)을 나타내었고, 5번째 환자는 2달 이내에 현저한 증상 개선을 경험하였다. 모든 환자들은 미국계 백인(Caucasian)이었다.

[0347] 상기 키나제 도메인(엑손 18~24)의 염기서열분석 결과, 제피티니브상에서 진행한 종양을 가진 4명의 환자로부터 얻은 종양에서는 돌연변이가 확인되지 않은 반면, 제피티니브 반응성 환자로부터 얻은 5개의 종양은 모두 *EGFR* 키나제 도메인 돌연변이를 보유하고 있었다. 카이제곱검정(Chi-squared test) 결과, 제피티니브 반응자(5/5)와 무반응자(0/4)간의 *EGFR* 돌연변이 빈도에 있어서 통계학적으로 유의한 차이($p=0.0027$)가 확인된 한편, 제피티니브 반응자와 비선별 미국계 NSCLC 환자간에도 $p<10^{-12}$(*)의 유의한 차이가 확인되었다. 한개의 제피티니브 민감성 폐선암종에서는 앞서 비선별 종양에서 관찰된 *EGFR* L858R 돌연변이가 동정되었다(도 6a; 표 S3A, IR3T). 3개의 제피티니브 민감성 종양은 이형접합성 인프레임 결실을 보유하고 있었으며(도 6c 및 표 S3A 및 S3B, 두 사례에서는 Del-3, 한 사례에는 Del-4), 하나는 동형접합성 인프레임 결실을 보유하고 있었다(도 6c 및 표 S3A 및 S3B, Del-5). 이 결실들 각각은, 비선별 종양에서 발견된 경우에도, *EGFR*의 코돈 746~753 영역내에 있었다. 이들 3개의 결실 각각은 아미노산 치환과도 관련이 있다(표 S3A-S3C). 4개의 시료 모두를 정상 조직에 대응시켜 보면, 이들 돌연변이가 체세포 돌연변이인 것으로 확인되었다.

[0348] 실시예 3A : 프라이머 설계

[0349] 수용체 티로신 키나제의 cDNA 염기서열을 유전자은행(GeneBank)(서열등록번호가 표 S1에 나열되어 있음)으로부터 얻은 다음, BLAT 정렬을 이용하여 인간 게놈 어셈블리(http://genome.ucsc.edu)에 적용시켜 엑손/인트론 경계(exon/intro boundaries)를 동정하였다. 프라이머3 프로그램(Primer3 Program)(http://frodo.wi.mit.edu/primer3/primer3_code.html)을 이용하여, 엑손 서열 및 각 측면부에 위치하는 적어도 250bp의 인트론 서열 또는 인접한 엑손 서열을 증폭하도록, 바깥쪽 유전자(external gene) 특이적 프라이머쌍을 설계하였다. 그 결과 예상되는 엠프리콘을 사용하여, 엑손측에 접해 있고(통상 엑손/인트론 경계 부분으로부터 50bp 이상) M13 정방향 또는 역방향 프라이머 꼬리가 첨부되어 있는 안쪽 프라이머(internal primer)를 설계하였다. 이러한 네스티드 프라이머 세트(nested primer set)는 DNA 대조군으로부터 적당한 엠프리콘 사이즈의 고품질의 서열(high-quality sequence)을 얻기 위해 시험하였다. 나고야 시립 대학교 의과대학으로부터 얻은 58개의 원발성 폐암 시료 세트를 이용하여, 47개의 티로신 키나제의 수용체 티로신 키나제 활성화 루프를 코딩하는 엑손을 포함하는 엠프리콘을 증폭시켜 서열분석하였다. 또한, 전장 EGFR(full length EGFR)을 포함하는 엠프리콘도 증폭시켰다.

[0350] 실시예 3B : 게놈 DNA의 PCR 및 염기서열분석 방법

[0351] 384-웰 포멧 네스티드 PCR 기구(384-well format nested PCR setup)를 이용하여, 특이적 프라이머를 사용하여 티로신 키나제 엑손 및 인접한 인트론 염기서열을 증폭시켰다. 각 PCR 반응은 5ng DNA, 1X HotStar 완충액, 0.8mM dNTPs, 1mM $MgCl_2$, 0.2U HotStar 효소(Qiagen, Valencia, CA), 및 0.2μM 정방향 및 역방향 프라이머를 함유하는 10μL 반응 용량으로 시행하였다. PCR 주기 파라미터는 95℃에서 15분간 1주기(cycle); 95℃에서 20분간, 60℃에서 30초간 및 72℃에서 1분간을 1주기로 하여 35주기; 72℃에서 3분간 1주기로 하였다.

[0352] 상기 시험 결과 얻어진 PCR 산물은 고체상 가역 고정화 화학법(solid phase reversible immobilization chemistry)으로 정제한 다음, 범용의 M13 프라이머를 이용하여 양방향 염료-전사종결인자 형광 염기서열분석(bi-directional dye-terminator fluorescent sequencing)을 행하였다. 염기서열분석할 단편은 ABI Prism 3700 DNA Analyzer(Applied Biosystems, Foster City, CA)를 이용한 모세관 전기영동(capillary electrophoresis)을 통해 검출하였다. PCR과 염기서열분석은 주식회사 에이전코트 바이오사이언스(Agencourt Bioscience Coporation)(Beverly, MA)에서 수행하였다.

[0353] 실시예 3B : 염기서열 분석 및 확인(validation)

[0354] 정방향(F) 및 역방향(R) 크로마토그램은 Mutation Surveyor 2.03 (SoftGenetics, State College, PA)에 의해 일괄적으로 분석한 다음, 매뉴얼 재검토를 행하였다. 한 방향 또는 양방향에서 확인된 고품질 염기서열 변이는 후보 돌연변이(candidate mutation)로서 평가하였다. 후보 돌연변이를 가진 엑손은 상기한 바와 같이 오리지날 DNA 시료로부터 다시 증폭시켜 다시 염기서열분석을 행하였다.

[0355] 실시예 3C : 환자

[0356] 폐종양 표본은 임상시험심사위원회(Institutional Review Board)의 연구 승인 하에 나고야 시립 대학병원 및 브리검 여성 병원에서 치료받은 비소세포 폐암에 걸린 환자(비선별 일본계 환자 종양 및 제피티니브-처리 미국계 환자 종양), 및 브리검 여성 병원의 익명의 종양 은행(tumor bank)(비선별 미국계 환자 종양)으로부터 얻었다. 성별, 나이, 및 조직적합성(histology)에 대한 정보는 대부분의 시료로부터 얻을 수 있었다. 또한, 환자 시료는 다나-파버 암연구소에서의 제피티니브의 개방표지 임상시험(open-label clinical trial)이 행해진 환자들로부터도 얻었다(13). 제피티니브에 대한 응답성은 표준 기준을 이용하여 규정하였다(참조, 예, A. B. Miller, B. Hoogstraten, M. Staquet, A. Winkler, 1981 Cancer 47, 207-14). 이 연구에 대해서는 IRB 승인을 얻었다.

[0357] 상기 제피티니브 응답성 환자 중에서, 2명은 사전에 적어도 1주기(cycle)의 화학요법을 받았으며, 1명은 사전에 방사선 요법을 받았고, 1명은 화학요법을 병행하였으며, 1명은 다른 어떠한 치료도 받지 않았다. 제피티니브 무감응성 환자에 대해서, 제피티니브 치료 2개월 후, 기준 CT 스캔과 비교하여, CT 스캔 상에 새로운 종양 병변

이 출현하거나 또는 현존 종양 병변이 증식하는 경우는 치료 실패로 규정하였다.

[0358] 실시예 3D : 환자 시료의 cDNA 염기서열분석

[0359] Trizo TM(Invistrogen, Valencia, CA)을 사용하여 조직 시료로부터 총 RNA(total RNA)를 분리하여 RN easyTM mini-elute cleanup kit(Qiagen, Valencia, CA)를 사용하여 정제하였다. cDNA는 총 RNA 2㎍으로부터 수퍼스크립트 II 역전사효소(Invitrogen Life technologies, Carlsbad, CA)를 제조자의 권장요건에 따라 사용하여 역전사시켜 얻었다. 얻어진 cDNA는 후속 EGFR의 PCR 증폭용 주형(template)으로서 사용된다.

[0360] PCR의 구성 성분은 20mM Tris-HCl(pH 8.4), 50mM KCl, 1.5mM $MgCl_2$, 0.1mM의 dATP, dCTP, dGTP, dTTP 각각, 0.2μM의 각 프라이머, 및 0.05units/㎕ Taq 폴리머라제(Taq Platinum, GIBCO BRL, Gaithersburg, MD)이다. 단편 "a"를 증폭시키기 위해서는 상기 반응에 4%의 DMSO의 첨가가 요구된다. 프라이머 염기서열은 표 S4에 열거하였다. 상기 정방향 및 역방향 프라이머는 각각 18개 염기쌍의 돌출형(overhanging) M13 정방향 및 역방향 염기서열을 이용하여 합성한다. 열변화(thermocycling) 조건은 94℃, 4분간 반응시킨 후; 변성단계 94℃에서 20초간, 연장단계 72℃에서 20초간, 주기1 60℃로부터 주기11 50℃까지 주기당 1℃씩 감소된 결합(anealing) 단계 20초를, 11주기를 행한 다음; 주기 11을 25회 반복하였다. 그 다음, 72℃에서 6분간 배양한 다음 4℃에 넣어 프로그램을 종료한다.

[0361] PCR 반응의 분취표본(aliquot)은 물과 1:50으로 희석시킨다. 희석된 PCR 산물은 M13 Forward Big Dye Primer kit(Perkin-Elmer/Applied Biosystems, Foster City, CA)를 사용하여 제조자의 권장요건에 따라, 염기서열분석을 행하였다. 염기서열분석할 산물은 형광 서열분석기(Applied Biosystems제 모델 3100, Foster City, CA)상에서 분리하였다. 기본값(base call)은 기기의 소프트웨어에 의해 형성하고, 육안으로 관찰하여 재검토한다. 각 염기서열은 Sequencher 4.1 소프트웨어(Gene Code Corp.)를 사용하여 대응하는 정상 염기서열과 비교한다.

[0362] 실시예 3E : EGFR 변이체를 발현하는 종양 종류

[0363] 두개의 다른 종류의 종양에서, EGFR에서의 2개의 부가적인 돌연변이가 확인되었다. 857 위치의 글리신(G)이 발린(V)으로 치환된 EGFR 돌연변이("G857V")가 급성 골수종 백혈병(AML)에서 동정되었으며, 883 위치의 류신(L)이 세린으로 치환된 EGFR 돌연변이("L883S")가 전이성 육종에서 확인되었다.

[0364] 실시예 3F : 세포주(Cell line)

[0365] 시험관내의 NSCLC 세포주에 대한 제피티니브의 효과를 시험하였다. 1개의 세포주, H3255는 제피티니브에 특히 민감하였다(IC_{50} 40nM). 나머지 세포주들은 훨씬 높은 IC_{50}을 가지고 있었다. 예를 들어, 야생형 세포주 H1666은 돌연변이 세포주보다 50배 높은 2μM의 IC_{50}을 가지고 있다. 이 세포주 유래의 EGFR을 서열분석하면, L858R 미스센스 돌연변이를 포함하는 한편, 다른 세포주들은 EGFR에 대해 야생형이었다. EGFR의 저해 또는 EGFR에 의한 AKT 및 ERK의 인산화에는 EGFR 야생형 세포주와 비교하여 훨씬 더 낮은 농도의 제피티니브가 필요하며, EGFR 야생형 세포는 동일한 효과를 얻기 위해서 적어도 100배 이상의 높은 농도의 제피티니브를 필요로 한다. 이러한 발견은 돌연변이 수용체는 제피티니브의 효과에 더 민감함을 암시한다.

[0366] 실시예 3G: 병행 치료

[0367] 에를로티니브를 사용하거나 또는 사용하지 않고 카르보플라틴/파클리탁셀의 무작위 임상시험 치료를 받은 진행성 NSCLC을 가진 환자의 종양 표본을 분석하였다. 이 시험의 임상적 값(portion)은 두개의 치료군에서 동등한 생존율을 나타내었다. 1076명의 환자 중 228명으로부터 얻은 종양 표본을 염기서열분석에 이용할 수 있었다. 이 환자들의 전임상 특성은, 기본 인구학적 자료(baseline demographics), 반응율, 정중 및 전체 생존율 면에서, 상기 치료군 전체적으로 차이가 없었다.

[0368] 티로신 키나제 도메인의 엑손 18~21의 염기서열을 분석하였으며, 12.7퍼센트의 돌연변이 빈도에 해당하는 29개의 돌연변이가 동정되었다.

[0369] 전체적으로 EGFR 돌연변이를 가진 환자는 화학요법 단독 치료를 받았는지 또는 에를로티니브와 병행 치료를 받았는지 여부에 관계없이 더 높은 생존율을 갖는다. 이러한 차이는 p값이 0.001 이하로 통계학적으로 유의하다. 이러한 발견은, EGFR 돌연변이가 제피티니브 및 에를로티니브에 대한 반응성의 예측변수가 되는 것 외에, 향상된 생존율에 대한 예후가 될 수 있는 가능성을 높인다.

[0370] (*)상기 비선별 미국계 환자들에서의 EGFR 돌연변이 빈도(1/16)는 보고된 제피티니브 반응 빈도 10.4%에 비해서 낮게 나타난다. 이러한 차이는 통계학적으로 적정한 유의성을 갖는다(카이제곱검정법에 의한 p=0.025). 따라서 이러한 결과는 여전히 가능성(chance)에 기인한 것이거나, EGFR 돌연변이를 갖지 않은 반응자 부분(fraction)에 기인한 것이거나, 또는 이 종양 콜렉션에서 EGFR 돌연변이를 실험적으로 검출하는데 실패함에 기인한 것일 수 있다. 제피티니브 반응성 미국계 환자의 EGFR 돌연변이 빈도(5/5)를 제피티니브 반응성의 예측빈도(10.4%)와 비교해 보면, 카이제곱확률은 다시 10～12 미만으로 된다.

[0371] [실시예 4]

[0372] **연구 설계:**

[0373] 2004년 8월부터 2005년 1월까지, 메사추세츠 제너럴 하스피탈(MGH), 다나-파버 암연구소(DFCI), 및 브리검 여성병원(BWH)에, 체세포 EGFR 키나제 도메인 서열분석이 의뢰된 NSCLC 환자의 후향성 코오트 연구(retrospective cohort study)를 수행하였다. 이들 세 기관은 1년에 약 1200명의 폐암 환자를 치료하는, 산학 연계 벤처 암센터인, 다나-파버/파트너스 캔서케어(Dana-Farber/Parters CancerCare: DF/PCC)를 구성한다. 2004년 8월에는, *EGFR* 키나제 도메인 염기서열분석이 DF/PCC에서 임상적 용도로 유용하게 사용되었다. 임상의는 시험을 의뢰할 환자를 선별할 수 있으나, 환자들은 충분히 적합한 이용가능한 종양 표본을 가질 것이 요구되었다. 종양 세포는 MGH 및 BWH 병리전문의에 의한 조직접합성 검사에 의거하여 표본의 적어도 50%를 구성해야하며, 그 표본은 원발성 또는 전이성 종양의 절제(resection), 기관지경 생검(bronchoscopic biopsy), 또는 중심부 바늘 생검(core needle biopsy), 또는 흉수 유래의 세포 군집으로부터 얻은 것이어야 한다. 드문 경우에는, 미세 바늘 흡인액 시료가 적당한 것으로 결정되었다. 시료는 파라핀 포매 조직 또는 동결 조직이었다. 편평세포종양(62)에서의 EGFR 돌연변이 발생율이 낮기 때문에, 이것을 가진 것으로 진단된 환자는 시험에 적격하지 않았다.

[0374] 본 발명자들은, 모든 염기서열분석이 수행되고 판독되는 진단 시험 시설인, 하버드 의과대학/파트너스 헬스케어 센터 포 제네틱스 앤드 게노믹스(CLIA# 22D1005307)의 분자의학실험실(LMM)에서 보유하고 있는 EGFR 케이스 로그(EGFR case log)를 사용하여, EGFR 시험을 경험한 환자를 동정(identify)하였다. 그리고, 본 연구기간 동안 NSCLC으로 진단되어 DF/PPC로부터 EGFR 시험을 의뢰받은 모든 환자들을 포함하였다.

[0375] 환자의 나이, 성별, 및 인종은 전자 의학 기록 시스템으로부터 수집하였다. 흡연상태, 암병력, EGFR 키나제 도메인 서열분석 결과, 및 후속 EGFR-TKI 치료 계획은 지정된 의사의 챠트 리뷰를 이용하여 기록하였다. 특히, 흡연상태 및 암병력은 의사와 간호사의 노트로부터 얻었다. 폐암 진단 전 1년이상 흡연을 중단한 사람을 이전 흡연자(former smoker)로 규정하고, 평생동안 100개피 미만을 피운 사람을 비흡연자(never-smokers)로 규정하였다. 진단 1년 이내에 흡연을 중단하였거나 또는 진단 시에 흡연중이었던 흡연자는 현재 흡연자(current smoker)로 분류하였다. 흡연 갑년(Pack-year)은 하루에 피운 갑(pack)의 수와 흡연 년도수를 곱하여 계산하였다. 종양 세포조직(histology) 및 EGFR 키나제 도메인 염기서열분석 결과는 병리 보고서로부터 얻었다. 모든 병리 표본은 MGH 또는 BWH에서 중점적으로 검토하였으며, 세포조직은 세계보건기구(WTO) 분류 시스템(63)을 이용하여 범주를 나누었다. 후속 치료 계획은 의사의 노트로부터 얻었다.

[0376] 완성된 데이터(complete data)로부터 나이, 성별, 종양 세포조직, 및 EGFR 돌연변이 상태(status)를 입수할 수 있었다. 인종(12%), 시험 당시의 종양 단계(stage), 흡연상태(6%), 선행치료(5%), 및 후속 EGFR-TKI 치료계획(11%)에 대해서 누락된 데이터가 있었다. 이 연구는 DF/PCC의 임상시험심사위원회(IRB)에 의해 승인받았다.

[0377]

[0378] EGFR 유전자 염기서열분석:

[0379] 동결되고, 포르말린 고정된, 또는 파라핀 포매된(FFPE) 종양 조직의 연속 절편을 잘라내어 글라스 슬라이드에 올려놓았다. 병리전문의가 적어도 50%의 생존 종양 세포로 이루어진 종양 조직 부분을 동정하였다. FFPE 시료를 크실렌과 에탄올로 추출하여 파라핀을 제거하였다. FFPE 시료 및 동결 조직 시료 둘다를 프로테나제 K로 하룻밤 분해시켰다(digest). 표준 방법을 사용하여 조직 및 말초 전혈(peripheral whole blood)로부터 게놈 DNA

를 추출하였다. 게놈 DNA는 DNA Genetek-Oragene™ 키트를 사용하여 타액 시료로부터 추출하였다.

[0380] 개개의 네스티드 PCR 반응 세트를 이용하여 EGFR의 키나제 도메인(엑손 18~24 및 인접한 인트론 영역)을 증폭시켰다. 상기 네스티드 PCR 증폭에 사용된 프라이머는 표 S1A 및 S1B 및 서열번호 1~424에 기재되어 있으며 이 프라이머의 5' 말단에 범용 서열(5'tgtaaaacgacggccagt)(서열번호 645)이 부가된다. 얻어진 PCR 산물은 염료-전사종결인자 염기서열분석(dye-terminator sequencing)에 의해 직접 양 방향으로 서열분석하였다. PCR은 5ng의 게놈 DNA, 2mM의 $MgCl_2$, 0.75㎕의 DMSO, 1M의 베타인, 0.2mM의 dNTPs, 20pmol의 프라이머들, 0.2㎕의 AmpliTaq Gold™(Applied Biosystems), 1X 완충액(AmpliTaq Gold와 함께 제공됨)을 함유하는 반응용량 15㎕으로 384-웰 플레이트에서 수행하였다. 열변화 조건은 다음과 같다: 95℃에서 10분; 95℃에서 30초, 60℃에서 30초, 72℃에서 1분을 1주기로 하여 30주기; 및 72℃에서 10분. PCR 산물은 Ampure™ 마그네틱 비드(Agencourt)로 정제하였다.

[0381] 염기서열분석할 산물은 Cleanseq™ 마그네틱 비드(Agencourt)를 사용하여 정제한 다음, ABI 3730 DNA Analyzer 상에서의 모세관 전기영동에 의해 분리하였다. 염기서열 분석(sequence analysis)은 2명의 재검토자가 Mutation Surveyor (SoftGenetics, State College, PA)을 이용하여 매뉴얼에 따라 수행하였다. 비동의(nonsynonymous) DNA 염기서열 변이체는 오리지날 게놈 DNA(original genomic DNA)의 3~5회 독립 PCR 반응의 분석에 의해 확인하였다. 비동의 DNA 염기서열 변이체를 가진 개체로부터 얻은 혈액 또는 타액 시료를 분석하여, 그 염기서열 변화가 종양 조직에만 특이한(unique)지의 여부를 결정하였다.

[0382] 통계학적 분석:

[0383] 환자의 인구학적 임상적 특성과 *EGFR* 돌연변이 상태 사이의 일변량 연관성(univariate association)을 평가하기 위해서 로지스틱 회귀 모델(logistic regresstion model)을 구성하였다. 돌연변이 양성 상태의 유의한 예측변수를 확인하기 위해, 앞선 연구에서 돌연변이의 예측변수로 확인되었던 독립 변수, 특히 성별, 인종, 세포조직, 및 흡연 상태를 포함하는 다변량 로지스틱 회귀 모델(multivariable logistic regression model)을 구성하였다. 6명은 환자는 PCR에 실패하여 *EGFR* 돌연변이 데이터가 없기 때문에 이 분석에서 제외하였다. 모든 분석은 통계 소프트웨어 SAS(버젼 8.02, SAS Institute, Cary, NC)를 사용하여 수행하였다.

[0384] **결과:**

[0385] 환자의 특성:

[0386] 본 연구 기간 동안 암의 임상 치료의 한 부분으로서 체세포 EGFR 키나제 돌연변이 염기서열분석이 의뢰된, NSCLC를 가진 환자 100명의 평균 연령은 60.7세이고 여성이 63%였다(표 4). 환자 대부분은 백인(76%) 또는 아시아인(7%)였고, 시험 당시 전이성으로 진단된 병변을 가지고 있었다(67%). *EGFR* 돌연변이에 대해 시험한 거의 모든 환자들(94%)은 선암종, 기관지폐포 암종(bronchiolo-alveolar carcinoma:BAC) 특징을 가진 선암종, 또는 순수 BAC를 가지고 있었다. 환자 중의 대략 1/3이 비흡연자였다. *EGFR* 시험 의뢰에 앞서 처치된 치료법으로는 수술(50%), 가슴 방사선요법(22%), 화학요법(47%), 및 EGFR 직접 표적 치료(11%)를 들 수 있다.

[0387] 동정된 돌연변이:

[0388] 시험 의뢰로부터 사용가능성 있는 결과를 얻는데 걸리는 평균 시간은 12 영업일(business day)이었다. 제출된 표본의 대다수는 파라핀 포매 표본이었다(74%). 74개의 파라핀 포매 표본 중의 6개(7%)는 PCR 증폭을 실패한 반면, 26개의 동결 표본은 모두 성공적으로 증폭되었다. 총 25개의 동정된 돌연변이에 대해서, 해석가능한 결과를 가진 94명의 환자 중, 23명(24%)은 *EGFR* 키나제 도메인내에 적어도 하나의 돌연변이를 가짐이 확인되었으며, 이 환자들 중 2명은 각각 2점 돌연변이를 나타내었다(표 5). 상기 돌연변이를 가진 23명의 환자 중에서, 9명(39%)은 하나 이상의 점 돌연변이를 가지고 있었으며, 12명(52%)은 엑손 19에 인프레임 중첩 결실을 가지고 있었고, 2명의 환자(9%)는 엑손 20에 중복(duplication)을 가지고 있었다. 상기 점 돌연변이는 엑손 18 및 21에 존재하였으며, 5개의 2573T>G(L858R)와, 2126A>T(E709V), 2155G>A(G719S), 2156G>C(G719A), 2327G>A(R776H), 2543C>T(P848L), 및 2582TA(L861Q)를 각각 한개씩 포함하였다. 상기 점 돌연변이 중의 하나

(P848L)는 종양 표본 및 볼점막 면봉으로부터 얻은 단핵세포 둘다에서 검출되었다. 엑손 22, 23, 또는 24에서는 돌연변이가 검출되지 않았다.

[0389] 돌연변이의 예측변수:

[0390] 본 발명의 시료 중에서, *EGFR* 돌연변이 상태와, 나이(p=0.61), 여성(p=0.92), 아시아인종(p=0.08), 또는 의뢰 당시의 전이성 병변(p=0.43, 표 4) 사이에 유의한 연관성은 없었다. 비선종성(non-adenocarcinoma) 종양 세포 조직을 가진 6명의 환자 중, 돌연변이를 가진 것으로 확인된 환자는 없었다. 선암종, BAC 특징을 가진 선암종 및 순수 BAC를 가진 환자 중, BAC/BAC 특징과 *EGFR* 돌연변이 상태 사이에는 연관성이 없었다(p=0.35).

[0391] 17명의 현재 흡연자 중, 돌연변이를 가진 것으로 확인된 환자는 없었다. 비흡연자는 이전 흡연자보다 *EGFR* 돌연변이를 가질 가능성이 유의하게 높았다(교차비(odds ratio)[OR]=3.08, 95% 신뢰구간[CI] 1.09~8.76). 흡연 갑년의 평균수는, EGFR 돌연변이 양성 환자(0.7 갑년)가 EGFR 돌연변이 음성 환자에 비해서(25.0 갑년, $p < 0.001$) 유의하게 낮았다. 각각의 부가적인 흡연 갑년에 대해서는, 돌연변이를 가질 가능성(likelihood)면에서 4% 감소하였다(OR=0.96, 95% CI 0.93-0.99).

[0392] 상기 흡연 갑년 수는, 성별, 인종, 및 종양 조직접합성에 대하여 조절한 후, 돌연변이 상태의 유의한 예측변수였다(OR=0.96, 95% CI 0.93-0.99).

[0393] 시험 정보의 후속 사용:

[0394] *EGFR* 돌연변이 양성 환자는 후속 EGFR-TKI 치료 계획이 보고될 가능성(86%)이 *EGFR* 돌연변이 음성 환자(11%, $p < 0.001$)보다 유의하게 높았다. 임상의는, 38%의 사례에서 *EGFR* 결과가 그들의 우선적인 추천 치료법에 영향을 미친다고 보고하였다. 이러한 사례들로는, 시험결과 음성(negative)으로 진단받기 전에 EGFR-TKI 치료를 추천받은 돌연변이 양성 환자 23명 중 14명(61%)과, EGFR-TKI 치료를 추천받지 않았거나, 또는 시험결과 양성으로 진단을 받은 후에 EGFR-TKI 치료를 추천받은 71명의 돌연변이 음성 환자 중의 24명(34%)이 포함된다.

[0395] EGFR 돌연변이 상태에 의해, 전이성 병변(54%)을 가진 환자에서의 우선 치료 옵션을 변경할 가능성이 국소 또는 국소적으로 진행된 병변을 가진 환자(19%, p=0.003)에서 보다 높았다. 이러한 조사결과를 얻은 경우, 전이성 환자에서의 의사결정 과정(decision-making process)을 더 분석하였다(도 10). 시험 결과가 치료법 추천에 영향을 미친, 전이성 병변을 가진 31명의 환자 중에서, 5명의 돌연변이 양성 환자는 1차(first-line) EGFR-TKI 치료를 받았고 6명의 돌연변이 양성 환자는 화학요법 대신에 2차 EGFR-TKI 치료를 받았다. 20명의 돌연변이 음성 환자는 그들의 음성 *EGFR* 시험 결과에 근거하여 3차 치료 또는 그 이상까지 EGFR-TKI 치료를 연기하도록 권고받았다. 시험 결과가 치료 추천에 영향을 미치지 않은 전이성 병변을 가진 26명의 환자 중에서, 2명의 돌연변이 음성 환자는 그들의 음성 결과에도 불구하고, 1차 EGFR-TKI 치료를 받았고, 4명의 돌연변이 양성 환자를 포함한 9명의 환자는 2차 또는 3차 EGFR-TKI 치료를 받았으며, 2명의 돌연변이 양성 환자를 포함한 15명의 환자는 EGFR-TKI 치료 추천을 받지 않았다. 전이성 병변을 가진 환자들 중 3명은 1차 EGFR-TKI 치료를 평가(evaluate)하는 시험에 참여하였다. 전이성 병변을 가진 환자 중의 9명은 미리 또는 EGFR 시험 시에 EGFR-TKI를 받았다.

[0396] **토의:**

[0397] 본 발명자들은 본 연구에서 NSCLC를 가진 100명의 환자들에게 암의 임상 치료의 일부분으로서 체세포 *EGFR* 돌연변이 스크리닝을 행하는 연구를 수행하여, 이 시험이 실행가능하고 NSCLC 환자의 치료에 유의하게 영향을 미침을 알아내었다. *EGFR* 돌연변이를 보유한 환자는 돌연변이가 없는 환자보다 EGFR-TKI 치료법을 추천받을 가능성이 유의하게 높았다. 의사들은 1/3 이상의 사례에서 시험 결과에 근거하여 치료 추천을 조정하였으며, 전이성 병변을 가진 환자는 조정할 가능성이 더 높았다. 본 연구의 환자 시료에서, 의사들은 양성 EGFR 시험 결과를 이용하여, 몇몇 환자, 특히 1차, 2차 치료 중인 환자에 대해서, 화학요법보다 EGFR-TKI를 우선적으로 시행하도록 하였다. 그러나, 음성 EGFR 시험 결과에 의해, 의사가 선별한 환자에게 EGFR-TKI를 투여하는 것을 막은 것은 아니었다. 상기 시험 결과가 임상적 의사 결정에 영향을 미치지 않은 많은 환자들은 시험당시 초기 단계였거나, 병변을 잘라내었거나 또는 전이성 병변에 대해 이미 EGFR-TKI를 받은 상태였다. 이것은, EGFR-TKI의 보조치료법으로서의 유용성이 알려져 있지 않고, 동정된 EGFR 돌연변이를 갖지 않는 소수의 환자에게도 EGFR-TKI

치료의 이점이 있기 때문에, 타당하다(65, 66-70, 71).

[0398] 또한, 본 연구는, 분자 진단법이 *EGFR* 돌연변이를 가진 환자를 동정하는 임상적 능력을 향상시킬 수 있다는 증거를 제공한다. 많은 종약학자들은, NSCLC를 가진 환자에 대한 의사결정과정을 가이드하기 위해, 현재 *EGFR* 돌연변이와 연관된 임상적 특징과 EGFR-TKI에 대한 반응성을 이용한다. 실제로, *EGFR* 시험이 의뢰된 본 연구의 환자 집단은 이러한 특징이 더 우세한 것으로 나타났다. 예를 들면, 의뢰된 환자의 95%가 선암종 또는 BAC 종양 세포조직을 가졌으나, 이에 비해, 일반 NSCLC 집단에서는 45%의 환자가 선암종 또는 BAC 종양 세포조직을 가졌다(72). 본 연구 집단의 29%가 흡연자인 반면, 일반 NSCLC 집단의 비흡연자 범위는 2~10%로 보고되어 있으며, NSCLC를 가진 여성에서는 27% 만큼 높아질 수 있다(73-75). 마찬가지로, 본 연구 집단은 단지 17%만이 현재 흡연자로 이루어졌으나, 이에 비해 새롭게 진단된 NSCLC 환자 중의 현재 흡연자 비율은 38~75%였다(75, 78-80). 본 연구의 임상적으로 선별된 집단은 결론적으로 EGFR 돌연변이 비율이 24%였으며, 이 비율은 비선별 가용 NSCLC 종양 시료를 시험한 본 연구 또는 다른 미국계 환자 군에 의해 보고된 비율보다 실질적으로 더 높았다(65-66, 81). 하지만, 임상의들이 EGFR 돌연변이로 예측되는 임상적 특징을 가진 시험용 환자를 선별하기 위해서 시도한 것으로 나타났으나, 돌연변이 빈도는 여전히 단지 24%였으며, 분자 진단법이 임상적 결정에 대한 정보의 가용성을 증가시킨다는 사실이 부각됨을 인지하는 것은 중요하다.

[0399] 앞서 기재한 돌연변이 상태의 예측변수를 조절한 이후, 본 연구 환자에서의 *EGFR* 돌연변이 상태의 가장 강력한 예측변수는 흡연 상태이며, 흡연 이력의 증가는 *EGFR* 돌연변이를 보유할 현저히 감소된 가능성과 관련이 있었다. 본 연구 결과는 *EGFR* 돌연변이 가능성면에서 흡연 상태의 중요성을 보고하고 있는 다른 일련의 사례들과 일치한다(66, 69, 70, 81, 82). 편평세포 종양에서는 *EGFR* 돌연변이 유병률이 매우 낮아, 선암성 종양에 대한 시험에 수고가 많이 들어가므로, 흡연 이력이 짧거나 없는 환자에 대해 장차 노력을 집중하는 것이 적합할 수 있다. 하지만, 전형적인 임상적 특징을 보이지 않는 환자에서의 EGFR 돌연변이의 보고는 엄격한 시험 제한에 반하여 권고한다(83). 본 발명자들은, 돌연변이와 관련있는 것으로 생각되는 다른 임상적 특징을 조사한 결과, 아시아인종과 BAC 종양 세포조직은 EGFR 돌연변이 상태를 예측하는데 유의하지 않은 경향을 가짐을 알아내었다. 이러한 연관성면에서의 통계학적 유의성의 부족은 시료 사이즈가 작은 것에 기인한 것일 수 있다.

[0400] 상기 시험은 실행가능하였으며 암의 임상 치료의 시간적 제약면에서 적합하였다. 거의 모든 종양들은 분석을 위해 제공되어 해석가능한 결과가 도출되었다. PCR 증폭에 실패한 6개의 표본은 모두 파라핀 포매 표본이었으며, 이에 반해 PCR 증폭에 실패한 동결 표본은 없었다. 가능한 한, 신선한 동결 표본이 EGFR 돌연변이 시험용으로 바람직한 기재(substrate)이다.

[0401] 부분적인 또는 완전한 EGFR 염기서열 분석을 행한 것으로 보고된 NSCLC 시료는 2500개에 가까웠다. 본 연구의 환자들은, 엑손 19의 9~23개의 염기쌍의 결실과 엑손 18 및 21에서의 단일 아미노산이 치환을 유도한 점 돌연변이가 중첩된, 이전 보고와 유사한 돌연변이를 나타내었다. 본 발명자들은 5개의 점 돌연변이를 찾아내었다(E709V, G719S, G719A, L858R, 및 L861Q). 이 점 돌연변이 중의 한개는 앞서 다른 아미노산 치환이 기재되었던 코돈에서 아미노산 치환(R776H)을 일으킴이 확인되었다. E709V 및 R776H 변이체는 각각 코돈 719를 포함하는 공지의 제피티니브 민감성 돌연변이(geftinib-sensitizing mutation)와 함께 확인되었다. 엑손 21에서의 P848L 돌연변이는 체세포 시료 또는 볼점막 시료 둘다에서 확인되었으며, 이것은 분명하지 않지만 배계열 변이체(germline variant)일 수 있음을 암시한다. 상기 환자는 EGFR 돌연변이 시험 전에, 제피티니브 치료를 받은 15개월 동안 안정성 병변(stable disease)을 가진 선암종을 보유한 비흡연 여성이었다. 상기 P848L 돌연변이가 밝혀졌을 경우, 이 여성은 최근 진행성 병변을 가졌던 것으로 확인되었으며 에를로티니브 치료를 시작하였다. 이때는 에를로티니브에 대한 반응성에 관한 정보를 이용할 수 없다.

[0402] 상기 (2253_2276 del) 결실은 앞서 기재한 엑손 19 결실과 중첩된다. 본 연구의 환자들에서의 결실은 그 범주를 2개의 그룹 중의 하나로 나눌 수 있다: 최소한 코돈 747-759에 걸쳐있는 것(아미노산 서열 LRE); 및 코돈 752-759에 걸쳐있는 것(도 11). 데이터에 기록된 모든 엑손 19 결실의 분석결과는 코돈 747-759에 걸쳐있는 TK 영역으로부터 다양한 아미노산이 결실될 수 있음을 암시한다. 필수 공통 코돈(required common codon)이 결실된 것은 나타나지 않았다; 그러나 모든 결실에서 745 위치의 리신 잔기는 유지되는 것으로 검출되었다.

[0403] 본 연구의 두개의 엑손 20 돌연변이 중의 하나는, *EGFR* 시험을 수행한 후 에를로티니브 치료를 받아서 이때 두달동안 안정성 병변을 가졌던, 재발된 선암종을 가진 비흡연 여성에서 확인된다. 나머지 하나는 EGFR-TKI 치료를 받았으나, 심한 발진 때문에 그것을 견뎌내지 못한, 전이성 암종을 가진 이전 흡연 남성에서 확인되었다. 임상적으로 관련 있는 엑손 20에서의 EGFR 돌연변이의 동정은 EGFR의 TK 영역의 포괄적인 염기서열분석의 중요성을 강조한다.

[0404] 결론적으로, 본 연구는 NSCLC 환자에 있어서, 암의 임상 치료의 일부로서, 체세포 돌연변이에 대한 EGFR 유전자의 TK 영역의 포괄적인 스크리닝의 실행가능성 및 유용성을 나타낸다. 그 시험 결과는 EGFR-TKI 응답의 임상적 예측변수에 관한 유용한 정보를 제공한다. 현재 흡연자는 흡연 이력의 갑년 수가 높은 이전 흡연자 만큼, 돌연변이를 보유할 가능성이 적다.

[0405] [실시예 5]

[0406] **비소세포 폐암에 대한 EGFR 유전자 시험**

[0407] 표준 시행 과정

[0408] 임상적 표시(Clinical Indication):

[0409] 이 시험은 비소세포 폐암을 가진 개체에 대해서 나타낸다.

[0410] 분석 원리:

[0411] *EGFR* 유전자 시험은 *EGFR*의 키나제 도메인내의 돌연변이를 검출하는 유전자 검사이다. 먼저, 종양 생검으로부터 DNA를 얻는다. 그 다음, 직접 양방향 유전자 염기서열분석법(direct bi-directional gene sequencing)을 이용하여, *EGFR*의 7개의 엑손(18, 19, 20, 21, 22, 23, 24)의 DNA 염기서열을 결정한다. 다음, 얻어진 염기서열을 공지의 *EGFR* 염기서열과 비교하여 DNA 염기서열의 변화를 확인한다. 종양 조직에서 DNA 염기서열 변화가 검출되는 경우, 오리지날 조직 시료(original tissue sample)에 대하여 상기 시험을 반복한다. 상기 변화가 제피티니브 또는 에를로티니브 반응자에 있어서 이미 보고된 것이 아닌 경우에는, 돌연변이가 종양조직에서 구성적(constitutive)인지(그 결과 정상적으로 일어나는 다형성(polymorphism)일 수 있음) 또는 체세포적으로 일어난 것인지를 결정하기 위해, 그 개체의 혈액 시료를 이용하여 상기 시험을 또 수행한다.

[0412] 시료 요건:

[0413] 최소한 100ng의 DNA가 조직 시료로부터 얻어질 것이 요구된다.

[0414] 주(note): 매우 적은 양의 DNA가 조직 시료로부터 얻어질 수 있다. 이러한 DNA의 농도는 정확하게 정량화할 수 없다.

[0415] 질 조정(Quality Control):

[0416] 사용된 대조군

[0417] 각 엑손에 대해서 2개의 음성 대조군(물) 및 양성 대조군(인간 DNA)이 PCR 반응에 포함된다. 상기 음성 대조군은 얻어진 서열이 오염된 결과가 아님을 확증하기 위해 전 과정을 통해 처리된다. pGEM 양성 대조군 및 ABI 어레이(array) 대조군은 염기서열분석 단계에 포함된다.

[0418] 대조군 제조 및 저장:

[0419] PCR에 대한 양성 대조군은 클론테크 인간 DNA(Clontech human DNA) 또는 익명의 혈액 시료로부터 얻은 인간 DNA이며 4℃에 저장된다. PCR 반응에 대한 음성 대조군은 실온에서 저장한 고순도 분자생물학 등급(HyPure Molecular Biology Grade)의 물이다. 상기 pGEM 양성 염기서열분석 반응 대조군 및 ABI 어레이 대조군은 -20℃에 저장된다.

[0420] 허용 한계(Tolerance Limit) 및 개체 대조군이 실패한 경우에 취하는 단계:

[0421] 상기 양성 PCR 대조군이 실패(fail)하고 음성 대조군 및 시료가 패스(pass)하는 경우, 그 PCR 결과는 패스로 지정하여 염기서열분석을 수행한다. 음성 대조군이 DNA 증폭의 증거를 나타내는 경우, 환자 DNA의 새로운 분취표본(aliquot)을 이용하여 전체 반응을 반복한다. 상기 pGEM 대조군이 실패하고 시험 반응액이 실패한 경우, PCR 산물의 두번째 분취표본을 이용하여 염기서열분석 반응을 반복한다. 염기서열분석 대조군은 실패하고 시험 반응액은 패스한 경우, 염기서열분석을 반복할 필요는 없다. 주: 파라핀 포매 조직 시료로부터 DNA 추출하는 것은 효율이 낮기 때문에, 바깥쪽 PCR 반응(external PCR 반응)은 종종 가시적인 산물을 생성하지 못한다. 안쪽

(internal) PCR 반응은 가시적 산물을 생성해야 한다. 적당한 PCR 산물이 얻어졌음을 확인하기 위해 겔상에서 검출된 산물의 사이즈를 예상되는 사이즈(하기 참조)와 비교해야한다. 안쪽 PCR 산물이 겔상에 가시적으로 보이지 않으면, 엑손 특이적 PCR 실패가 반복될 것이다.

[0422] 개별 시료에 대한 PCR 증폭이 실패한 경우, 첨가 DNA 시료를 두배로 증가하여 새로운 라운드의 PCR를 시도해야 한다. PCR 증폭이 다시 실패하면, 가능하다면 환자로부터 새로운 DNA 시료를 얻어야 한다. 상기 시료가 파라핀 포매 조직 시료인 경우, 부가 슬라이드를 벗겨내야 한다. 가능하다면, 오리지날 시료 생성에 사용된 것보다 더 많은 슬라이드를 벗겨내고, 프로테이나제 K를 3일밤 동안 반응시켜야 한다.

[0423] 장치와 시약(모든 시약은 특별하게 언급하지 않는 한 1년 동안 안정함):

[0424] PCR 및 염기서열분석(통상, PCR 및 염기서열분석 장치 및 시약은 당업자에게 잘 알려져 있으며, 또한 상기에 기재하였음.)

[0425] · 프라이머: (하기 표 6 및 7 참조)

[0426] <표 6 > 바깥쪽 PCR 프라이머

[0427]

엑손	정방향 프라이머 염기서열 , (5'→3')	서열 번호	역방향 프라이머 염기서열 , (5'→3')	서열 번호
18	TCAGAGCCTGTGTTTCTACCAA	653	TGGTCTCACAGGACCACTGATT	646
19	AAATAATCAGTGTGATTCGTGGAG	654	GAGGCCAGTGCTGTCTCTAAGG	647
20	ACTTCACAGCCCTGCGTAAAC	655	ATGGGACAGGCACTGATTTGT	648
21	GCAGCGGGTTACATCTTCTTTC	656	CAGCTCTGGCTCACACTACCAG	649
22	CCTGAACTCCGTCAGACTGAAA	657	GCAGCTGGACTCGATTTCCT	650
23	CCTTACAGCAATCCTGTGAAACA	658	TGCCCAATGAGTCAAGAAGTGT	651
24	ATGTACAGTGCTGGCATGGTCT	659	CACTCACGGATGCTGCTTAGTT	652

[0428] <표 7> 안쪽 PCR 프라이머

[0429]

엑손	정방향 프라이머 염기서열, (5'→3')	역방향 프라이머 염기서열 , (5'→3')	생성물 길이 (bp)
18	TCCAAATGAGCTGGCAAGTG (서열번호 660)	TCCCAAACACTCAGTGAAACAAA (서열번호 667)	397
19	GTGCATCGCTGGTAACATCC (서열번호 661)	TGTGGAGATGAGCAGGGTCT (서열번호 668)	297
20	ATCGCATTCATGCGTCTTCA (서열번호 662)	ATCCCCATGGCAAACTCTTG (서열번호 669)	378
21	GCTCAGAGCCTGGCATGAA (서열번호 663)	CATCCTCCCCTGCATGTGT (서열번호670)	348
22	TGGCTCGTCTGTGTGTGTCA (서열번호 664)	CGAAAGAAAATACTTGCATGTCAGA (서열번호 671)	287
23	TGAAGCAAATTGCCCAAGAC (서열번호 665)	TGACATTTCTCCAGGGATGC (서열번호672)	383
24	AAGTGTCGCATCACCAATGC (서열번호 666)	ATGCGATCTGGGACACAGG (서열번호673)	302
F 프라이머 링커	tgtaaaacgacggccagt (서열번호 645)	모든 정방향 프라이머의 5' 말단	18
R 프라이머 링커	aacagctatgaccatg (서열번호 674)	모든 역방향 프라이머의 5' 말단	16

[0430] · 주의 사항

[0431] <표 8>

[0432]

작업(task)	지시사항(instructions)	위험요인(risk)
1. PCR 셋업(setup)	-PCR 후드를 사용할 것 -전용 피펫 및 여과 팁을 사용할 것 -PCR 후드내에서만 시약을 개봉할 것	-PCR 반응의 오염
2. PCR 후드(hood)의 사용	-후드내에서 PCR 이후 시료 또는 시약은 사용하지 말 것.	-PCR 반응의 오염

[0433] ·바깥쪽 PCR을 위한 PCR 반응 혼합액(mix) 제조

[0434] 모든 과정은 클린 후드가 아니라, 게놈 DNA용 PCR 후드에서 수행하였다.

[0435] 1. Taq Gold와 dNTP를 얼음속에서 녹인다.

[0436] 2. 튜브(15㎖짜리 에펜도르프 튜브)에서 하기 표9의 조성으로 매스터 혼합액(master mix)을 제조한다. 물, 베타인, 10X 완충액, $MgCl_2$, DMSO, Taq Gold 및 dTNP를 이 순서로 첨가한다. 각 시약을 첨가하는 동안 조심스럽게 피펫 누르기(pipetting up-and-down)를 반복하여 시약들을 혼합하는 것이 매우 중요하다.

[0437] 3. DNA는 분취(aliquoting)하기 전에 매스터 혼합물에 첨가해야 한다. 대용량의 매스터 혼합액를 제조한 다음, 각 환자 또는 대조군에 대한 개개의 튜브에 96㎕씩 분취한다(8 rxn에 대해 충분함). 이 96㎕의 매스터 혼합액에 5ng/㎕의 DNA 8㎕를 첨가한다. 그 다음, 13㎕를, 플레이트의 개개의 웰(well)에 첨가하거나 또는 스트립 튜브(strip tube)에 넣고 멀티채널 피펫터를 사용하여 피페팅한다.

[0438] 4. 전체 384-웰 반응 플레이트용으로, 약 415개 반응에 충분한 매스터 혼합액를 제조한다.

[0439] 5. 기포를 제거하기 위해 매스터 혼합액 플레이트를 회전(spin)시킨다.

[0440] 6. 많은 세트의 프라이머를 사용하는 경우, 개개의 플레이트내에 정방향 프라이머와 역방향 프라이머를 갖도록, 96-웰 플레이트에 프라이머 세트를 첨가해야 한다.

[0441] 7. 멀티채널 피펫을 사용하여 상기 프라이머들을 첨가한다. 조심스럽게 피펫 누르기를 반복하여 확실히 혼합한다.

[0442] 8. 플레이트를 회전시켜 기포를 제거한다.

[0443] 9. 하기 표 10의 주기(cycle)를 이용하여 증폭시킨다.

[0444] 주: PCR은 384-웰 플레이트에서 행한다.

[0445] <표 9>

[0446]

시약	**반응당 용량 *(㎕)***
멸균 ddH_2O	4.90
5M 베타인	3.00
10X 완충액	1.50
염화마그네슘	1.50
DMSO	0.75
Taq	0.20
dNTP	0.15
PCR 정방향 프라이머1 (농도. 20pmol/uL)	1.00
PCR역방향 프라이머2 (농도. 20pmol/uL)	1.00
DNA (농도. 5ng/uL)	1.00
PCR 반응의 총 용량	15.00

[0447] <표 10> PCR 증폭 주기

Taq Gold 활성화	10 분	95 ℃	
변성	30 초	95 ℃	30 주기
결합	30 초	60 ℃	
	1 분	72 ℃	
연장	10 분	72 ℃	
유지	∞	4 ℃	

[0448]

[0449] 주: 바깥쪽 PCR을 수행한 후에는 클린업(cleanup)이 필요없다.

[0450] · 안쪽 PCR을 위한 PCR 반응 혼합액(mix) 제조

[0451] 안쪽 PCR 셋업은 몇가지 예외사항을 제외하고는 바깥쪽 PCR과 거의 동일하다.

[0452] 1. PCR 후드에서 바깥쪽 PCR에서 기재한 바와 같은 대용량의 매스터 혼합액를 제조한다.

[0453] 2. 매스터 혼합액를 7개의 스트립 튜브로 분취하고, 멀티채널 피펫으로 12㎕씩 384-웰 플레이트내로 분취한다.

[0454] 3. 정방향 및 역방향 안쪽 프라이머 각각을 1㎕씩 첨가한다. 일시적으로 플레이트를 밀봉한다.

[0455] 4. 후드로부터 꺼내어, 플레이트를 회전시키고 후속 PCR 셋업 부위로 이동시킨다.

[0456] 5. 전용 피펫을 사용하여 바깥쪽 PCR 산물 1㎕를 각 반응액으로 분취한다.

[0457] 6. 가열 밀봉하고 다시 회전시킨다.

[0458] 7. 바깥쪽 PCR과 동일한 증폭 주기를 행한다.

[0459] · 클린업 이전에 1% 겔상에서 PCR 산물 전개. 반복 PCR을 위한 패스/실패 엑손 결정

[0460] · Ampure 마그네틱 비드 클린업을 이용한 안쪽 PCR의 클린업

[0461] - 클린업

[0462] 1. Ampure 마그네틱 비드 플레이트를 비드 침전물이 없어질 때까지 진탕혼합(vortex)한다.

[0463] 2. Ampure 비드의 온도는 실온인 것이 매우 중요하다.

[0464] 3. 클린업용으로 사용되는 시약으로 조정하기 위해, Biomek사의 384-웰 Ampure 프로토콜을 이용하여, 반응 용량을 12㎕로 변경한다. 이것이 안되면, 오류가 생긴다.

[0465] 4. 상기 프로그램이 완성된 후, 웰 당 20㎕의 멸균한 ddH_2O로 플레이트를 수화시킨다. 물을 첨가하는 동안, 조심스럽게 피펫 누르기를 하여 확실히 혼합한다.

[0466] 5. 플레이트를 회전시켜 기포를 제거한다.

[0467] 6. 상기 플레이트를 마그네트(magnet)상에 올려놓고 비드를 분리한다. 염기서열분석 반응을 셋업하기 위해 DNA 1㎕를 취한다. 나머지는 나중에 사용하기 위해 -20℃에서 저장한다.

[0468] · 염기서열분석 프로토콜

[0469] - 염기서열분석 반응 혼합액 제조

[0470] 1. BigDye 3.1을 암소(dark place)의 얼음 중에서 녹인다.

[0471] 2. 튜브(15㎖ 에펜도르프 튜브)에서 하기 표 11의 조성으로 매스터 혼합액을 제조한다. 물, 완충액, DMSO 및

BigDye는 이 순서로 첨가해야 한다.

[0472] 3. 각 시약을 첨가할 때는 조심스럽게 피펫 조작을 반복하여 시약을 혼합하는 것이 매우 중요하다.

[0473] 4. 염기서열분석용 범용 프라이머를 사용하는 경우, 상기 프라이머는 매스터 혼합액에 동시에 첨가할 수도 있다. 상기 프라이머가 특이적인 경우에는, 상기 매스터 혼합액을 384-웰 플레이트에 넣은 후에 개별적으로 첨가해야 한다.

[0474] 5. 통상 전체 384-웰 플레이트 반응용으로, 약 415개의 반응에 충분한 매스터 혼합액을 제조한다.

[0475] 6. 상기 매스터 혼합액이 셋업되는 즉시, 그 혼합액을 8 웰의 스트립 튜브에 나누어 넣는다(매스터 혼합액을 분취하는데 저장액을 사용하지 말 것. 그것은 폐시약일 수 있음.)

[0476] 7. 멀티채널 피펫을 사용하여 상기 매스터 시약을 384-웰 플레이트로 분취할 수 있다.

[0477] 8. 매스트 혼합액 플레이트를 회전시켜 기포를 제거한다.

[0478] 9. 멀티채널 피펫을 사용하여, 염기서열분석할 PCR 산물을 첨가한다. 피펫 조작을 반복하여 확실히 혼합한다.

[0479] 10. 플레이트를 회전시켜 기포를 제거한다.

[0480] 11. 하기 표 12의 주기를 이용하여 증폭시킨다.

[0481] <표 11>

[0482]

시약	반응당 용적 (μL)
멸균 ddH_2O	4.38
5X ABI 완충액	3.65
DMSO	0.50
ABI BigDye 3.1	0.35
염기서열분석 프라이머 농도	0.12
안쪽 PCR 반응으로부터 얻은 DNA	1.00
반응의 총용량	10.00

[0483] <표 12> 염기서열분석용 증폭 주기

[0484]

변성	10초	96 °C	25 주기
결합	5 초	50 °C	
연장	4 분	60 °C	
유지	∞	4 °C	

[0485] ·Cleanseq 마그네틱 비드 클린업을 통한 클린업

[0486] 1. Cleanseq 마그네틱 비드 플레이트를 비드 침전물이 없어질 때까지 진탕혼합한다.

[0487] 2. Biomek사의 Cleanseq 384-웰 플레이트 프로그램을 이용하여 시료를 클린업한다.

[0488] 3. 상기 프로그램이 수행되는 즉시, 오리지널 플레이트를 -20℃에 보관한다. ABI 3730상에서의 진행을 위해, 상기 클린업 시료를 포함하는 새로운 플레이트를 을 준비한다.

[0489] (주: PCR 산물이 300bp보다 짧은 경우, 그것을 ABI 3730상에 넣기 전에 시료를 희석해야 한다.)

[0490] ·EGFR 시험용 돌연변이 조사 주형(Mutation Surveyor Template)을 제조하여 LMM/Sequencing/Sequences-MS Review/EGFR 상에 보관한다.

[0491] · 반복 결과 기준

[0492] 모든 양성 결과는, 오리지널 조직 시료 유래의 환자 DNA의 두번째 분취표본으로부터 DNA 서열 변화가 검출된 특이적인 엑손을 증폭 및 염기서열분석함에 의해, 반복된다. 또한, 상기 검출된 서열 변화가 제피티니브 또는 에를티니브-반응자에게서 사전에 검출되지 않은 경우, 환자의 혈액 시료로부터 추출한 DNA를 종양 조직과 비교하기 위해 병행하여 증폭시켜야 한다.

[0493] 분명한 염기서열을 생성하지 못한 엑손은, 특정한 기술적 문제점에 근거하여, 추출, PCR 또는 염기서열분석 중 어느 하나를 반복한다.

[0494] · 평가 파라미터

[0495] 시험의 민감성 - 체세포 EGFR 키나제 도메인 돌연변이는 NSCLC를 가진 개체의 대략 13% 정도에서 확인되었다(Paez JG et al., 2004). 또한, 체세포 EGFR 키나제 도메인 돌연변이는 제피티니브 반응성이 있는 NSCLC를 가진 14명의 개체 중 13명(92.8%)에서 확인되었다(Paez JG et al, 2004, Lynch, et al., 2004). 알려진 돌연변이에 대해 100% 민감성을 나타낸 상기 시험의 기술적 민감성의 확인 및 본 발명자 실험실의 염기서열분석대(sequencing platform)의 확인은 100% 민감성을 보인다(하기 "방법의 정확성" 참조).

[0496] 모자이크 시료의 돌연변이 검출을 위한 민감성은 25%인 것으로 결정되었다(즉, 이형접합성 돌연변이는 세포 혼합물 중 50% 존재하는 경우에 검출될 수 있다.). 본 발명자들은 파라핀 포매 조직 중의 20% 정도는 고품질의 DNA를 얻지 못함을 알아내었다. 이들 시료로부터는 서열 정보를 얻을 수 없었다.

[0497] 시험의 특이성 - 공개된 문헌에서는 EGFR에 체세포 돌연변이가 없는 개체는 제피티니브에 대한 응답성이 없음을 기재하고 있다(11/11). 양방향 염기서열분석의 결과에 기인하여 돌연변이를 발견할 가능성(chance)은 0%에 가깝다(하기 "방법의 정확성" 참조). 이처럼, 본 발명의 시험의 특이성은 거의 100%이다.

[0498] 방법의 정확성 - DNA 염기서열분석 방법은 분자적 진단에 있어서의 우수한 전형에 해당한다. 본 발명자의 실험실은 98.5%의 정확도를 갖는 것으로 보고된 ABI 3730 DNA 애널라이저를 사용한다. 본 발명자들은, 이것과 양방향 염기서열분석, 자동 크로마토그램 분석과 돌연변이 조사(Mutation Surveyor), 및 위양성의 매뉴얼 분석을 조합하여, 100%의 정확도를 얻었다. 이것은 10만 염기 이상의 원료 염기서열 분석을 기준으로 한다. 이 측정법의 상세한 내용에 대해서는, 본 발명의 품질 보장 프로그램 매뉴얼을 참고한다.

[0499] 주: 본 발명자들은 이러한 결과들이 본 발명의 실험대의 100% 정확성을 보장한다고는 생각하지 않는다. 염기서열분석 오류는 일어날 수 있으며, 따라서, 대용량 염기서열분석 프로젝트에 의해 99.99%의 정확성이 확인되었음을 보고한다.

[0500] 시험의 재현성 - 시험의 정확성에 기인하여, 결과를 얻은 경우, 그 결과들은 시험의 정확성(99.99%)에 필적하는 재현성이 있다. 그러나, 경우에 따라서, 하기에 열거한 인자들(하기 '방법의 한계' 참조) 또는 설명할 수 없는 기술적 이유에 의한 PCR 또는 염기서열분석의 실패에 의해 시험이 실패할 수 있다. 이러한 경우, 아무런 결과가 얻어지지 않으며, 결과가 얻어질 때까지 측정(assay)을 반복하거나 또는 환자 표본을 적용불가(unacceptable)로 간주한다. 본 발명의 품질 보장 프로그램 매뉴얼(Quality Assurance Program manual)에 따른 확인 보고서에서는 각 측정 단계 및 표본이 특정한 비율로 실패함을 확인할 수 있다.

[0501] 결과들의 정상 범위 - *EGFR* 유전자의 정상 염기서열은 유전자은행 접근법(GenBank accession)을 사용하여 온라인 상에서 확인할 수 있다: NT_033968.5(게놈 염기서열) 및 NM_005228.3(mRNA 염기서열).

[0502] · 방법의 한계

[0503] 하나 이상의 엑손에 걸쳐있는 대량의 결실은, 특히 그것이 이형접합체에 존재하는 경우, 상기 염기서열분석법으로는 검출되지 않는다. 상기 키나제 도메인의 *EGFR* 유전자 바깥쪽(outside)에서의 돌연변이는 이 측정법에 의해 검출되지 않는다. PCR에 의한 증폭을 방해하는 저해제가 DNA 시료 중에 존재할 수 있다. 분해된 DNA(degrated DNA)는 분석가능한 데이터를 생성하지 못할 수 있으며 표본을 다시 제출해야 할 수도 있다. 드문 염기서열 변이(rare sequence variation)나 표적 프라이머 염기서열의 2차 구조는 PCR 증폭에 영향을 미칠 수 있으며 그 결과 돌연변이(들)가 하나의 대립유전자 영역내에서 빠져버릴 수 있다.

[0504] [실시예 6]

[0505] 제피티니브(Iressa™)는 표피성장인자 수용체(EGFR)를 표적으로 하는 티로신 키나제 저해제이며, EGFR 키나제 도메인내에 활성화 돌연변이를 가진 비소세포 폐암(NSCLC)에서 극적인 임상적 반응을 유도한다. 본 발명자들은, 이러한 EGFR 변이체는 세포 생존(cell survival)을 촉진하는 Akt 및 STAT 신호전달 경로를 선택적으로 활성시키지만, 세포 증식(proliferation)을 유도하는 Erk/MAPK 신호전달에 대해서는 효과가 없는 것으로 보고한다. EGFR 변이체를 발현하는 NSCLC는 EGFR 변이체의 siRNA-매개 넉다운(knockdown) 또는 Akt 및 STAT 신호전달의 약리작용 저해제로 치료 후에, 심한 세포사멸(apoptosis)을 경험하였으며, 종래의 화학요법 약제에 의해 유도된 세포사멸에 대해서는 비교적 내성을 나타내었다. 따라서, EGFR 변이체는 생존 신호를 선택적으로 전달(transduce)하며 NSCLC는 이 생존신호에 좌우된다; 결국, 제피티니브에 의한 그들의 신호 저해는 현저한 임상적 반응의 근간이 된다.

[0506] EGFR 군(EGFR family)의 수용체 티로신 키나제는 증식, 생존, 이동, 및 분화를 포함하는 세포의 필수적인 기능을 조절하며, 고형 종양의 병인 및 진행에 있어서 중심적인 역할을 한다(R. N. Jorissen et al. *Exp. Cell. Res*. 284, 21(2003), H. S. Earp, T. L. Dawson, X. Li, H. Yu, *Breast Cancer Res. Treat*. 35, 115(1995)). EGFR은 유방, 폐, 직장, 난소, 및 뇌 종양에서 빈번하게 과다발현되며, 이는 촉매 도메인 내의 ATP 포켓과 결합함으로써 EGFR 키나제 활성을 저해하는 제피티니브 등의 특정 약리작용 저해제의 발달(development)을 촉진한다(A. E. Wakeling et. al., Cancer Res. 62, 5749(2002)). 제피티니브는 화학요법-난치성 NSCLC를 가진 환자의 대략 10%에서 극적인 임상 반응을 유도하였다(J. Baselga et al., *J. Clin. Oncol*. 20, 4292(2002), M. Fukuoka et al., *J. Clin. Oncol*. 21, 2237(2003), G. Giaccone et al., *J Clin Oncol*. 22, 777(2004), M. G. Kris et al., *JAMA* 290, 2149(2003)). 실질적으로 모든 제피티니브 반응성 폐암은 EGFR 키나제 도메인내에 체세포 돌연변이를 보유하는 반면, 비반응성 사례에서는 돌연변이가 확인되지 않았다(T. J. Lynch at al., *N. Engl. J. Med*. 350, 2129(2004), J. G. Paez et al., *Science* 304, 1497(2004)). 이러한 이형접합성 돌연변이들은 ATP-결합 포켓에 군집되어 있는 작은 인프레임 결실 및 미스센스 치환을 포함한다.

[0507] 본 발명자들은, EGFR 돌연변이의 과도 트랜스펙션(transient transfection)을 이용하여, 두 종류의 돌연변이가 증가된 EGF-의존적 수용체 활성화를 유도하며, 이는 EGFR의 두드러진 C-말단 인산화 부위 중의 하나인 Y1068의 자가인산화에 의해 측정되었음을 앞서 보고하였다(T. J. Lynch et al. *N. Engl. J. Med*. 350, 2129 (2004)).

[0508] EGFR 변이체에 의한 신호전달에 있어서 질적 차이 연구를 가능하게 하기 위해서, 본 발명자는 야생형 EGFR 또는 EGFR 변이체를 발현하는 비형질전환 마우스의 유방 표피 세포(NMuMg)의 안정한 세포주를 생성하여, 별개의 하류 영향인자(downstream effector)의 활성화에 관련된 다중 티로신 키나제의 EGF-매개 자동 인산화를 분석하였다(R. N. Jorissen et al. *Exp. Cell Res*. 284, 31(2003)). 제피티니브 반응성 환자 유래의 종양에서 검출된 두 개의 재발 돌연변이 중의 하나 또는 야생형 EGFR 중의 어느 하나를 발현하는 세포주가 생성되었다: 미스센스 돌연변이 L858R 및 18bp 인프레임 결실, delL747-P753insS. 야생형 EGFR과 상기 두개의 EGFR 변이체 사이에, 여러 C-말단 부위에서, 현저하게 다른 티로신 인산화 패턴이 관찰되었다. Y1045와 Y1173의 EGFR-유도 인산화는 실질적으로 EGFR 야생형과 EGFR 변이체 사이에 구별할 수 없는 반면, Y992 및 Y1068의 인산화는 실질적으로 변이체 둘다에서 증가하였다. 흥미롭게도, Y845는 L858R 미스센스 변이체에서 고도로 인산화되었으나 야생형 또는 결실 변이체에서는 인산화되지 않았으며, 따라서, 상기 두 종류의 EGFR 돌연변이 사이를 구별하는데 특이적인 것으로 나타났다. 야생형 및 돌연변이 수용체에서 보이는 상이한 EGF-유도 티로신 인산화 패턴은 과도 트랜스펙트된 COS7 세포에서 재현가능하였으며, 세포종류에 특이적인 잠정적인 효과를 확보하였다.

[0509] 따라서, 제피티니브 민감성 EGFR 변이체는 야생형 EGFR에 의해 매개되는 것들과는 질적으로 다른 신호를 전달한다. 이러한 차이는 기질 특이성에 영향을 미치는 촉매 포켓내에서의 구조적 변화, 또는 EGFR 신호전달을 조절하는 보조 단백질과의 변경된 상호작용으로부터 직접 일어날 수 있다.

[0510] EGFR 변이체에 의해 안정적으로 트랜스펙트된 세포주를 확립함으로써, 공유한(shared) 세포 배경에서의 EGFR의 주요 하류(downstream) 표적의 인산화 상태를 비교할 수 있다. Ras를 통한 Erk1 및 Erk2, PLCγ/PI3K를 통한 Akt, 및 JAK2를 통한 STAT3 및 STAT5의 EGF-유도 활성화는 EGFR의 발암 효과를 매개하는 필수적인 하류(downstream) 경로이다(R. N. Jorissen et al., *Exp. Cell Res.* 284, 31(2003)). EGF-유도 Erk 활성화는, 야생형 EGFR 또는 2개의 활성화 EGFR 변이체 중의 어느 하나를 발현하는 세포 중에서는 본질적으로 구별할 수 없었다. 대조적으로, Akt 및 STAT5 둘다의 인산화는 상기 EGFR 변이체 중 어느 하나를 발현하는 세포에서 실질적으로 증강되었다. STAT3의 증가된 인산화는 EGFR 변이체를 발현하는 세포에서도 마찬가지로 관찰되었다. 상기 EGFR 변이체에 의한 상기 비변경 Erk 활성화(unaltered Erk activation)는, Ras 활성화 및 후속 Erk 인산화를 유도하는 Shc 및 Grb-2 어뎁터(adaptor)에 대한 중요 도킹 부위인, Y1173의 증가된 인산화의 부재와 일치한다(R. N. Jorissen et al., *Exp. Cell Res.* 284, 31(2003)). 상기 EGFR 변이체의 활성화 후에 증가된 Akt 및 STAT 인산화는, 이미 Akt 및 STAT 활성화에 연결되어있는 Y992 및 Y1068 인산화의 증가와 일치하였다(R. N. Jorissen et al., *Exp. Cell Res.* 284, 31(2003)). 따라서, EGFR 변이체내의 C-말단 잔기의 선택적 EGF-유도 자가인산화는 하류 신호전달 경로의 선택적 활성화와 상당히 상관관계가 있다.

[0511] 이러한 관찰을, EGFR 돌연변이가 종양발생을 유도하는 것으로 나타난 폐암세포로 확장시키기 위해서, 본 발명자들은 5개의 NSCLC 종양으로부터 유래하는 세포주를 연구하였다. NCI-H1975는 재발 이형접합성 미스센스 돌연변이 L858R을 보유하며, NCI-H1650은 인프레임 결실 delE746-A750를 갖는 반면, NCI-H358, NCI-H1666, 및 NCI-H1734는 야생형 EGFR를 발현한다. 트랜스펙트된 세포에서와 같이, Y992 및 Y1068의 EGF-유도 자가인산화는 내인성 *EGFR* 돌연변이를 가진 두개의 세포주에서 현저히 증강하였으며, 이는 Erk가 아니라, Akt 및 STAT5의 인산화였다.

[0512] EGFR의 발암 활성은 세포 증식과 세포 생존 모두를 촉진시키는 신호의 활성화를 나타낸다(S. Grant, L. Qiao, P. Dent, *Front. Biosci.* 7, d376(2002)). 이러한 경로가 중첩(overlap)되어 나타나지만, Erk 키나네의 Ras-매개 활성화는 실질적으로 EGFR의 증식 활성에 기여하는 반면, Akt 및 STAT의 활성화는 항 세포사멸 기능(anti-apoptotic function)과 상당히 연관되어 있다(S. Grant, L. Qiao, P. Dent, *Front. Biosci.* 7, d376(2002), F. Chang et al., *Leukemia* 17, 1263(2003), F. Chang et al., *Leukemia* 17, 590(2003), F. Chang et al., *Int. J. Oncol.* 22, 469(2003), V. Calo et al., *J. Cell Physiol.* 197, 157(2003), T. J. Ahonen et al., *J. Biol. Chem.* 278, 27287(2003)). *EGFR* 돌연변이를 보유하고 있는 상기 두개의 폐암 세포주는, 야생형 수용체를 가진 세포에서는 관찰되지 않은, 낮은 혈청(serum) 농도에서의 EGF에 대한 증식 반응을 나타내었다. 그러나, 그들의 증식율 및 세포충실도(confluence)에서의 세포 밀도는 정상 혈청 농도에서와 거의 동등하였다.

[0513] SiRNA

[0514] 반대로, 세포사멸 경로는 EGFR 변이체를 가진 폐암세포에서 현저히 다르다: 이들 세포주내에서의 EGFR 변이체의 siRNA-매개 특이적 불활성화 결과, 빠르게 대량의 세포사멸이 일어난다. delE746-A750-특이적 siRNA로 트랜스펙트된 NCI-H1650 세포와 마찬가지로, L858R-특이적 siRNA로 트랜스펙트된 NCI-H1975 세포의 약 90%가 96시간내에 죽었다. *EGFR* 돌연변이에 특이적인 siRNA는 상기 선택적 돌연변이(alternative mutation)를 발현하는 세포에서는 효과가 없었으며, 야생형과 돌연변이 *EGFR* 둘다를 표적으로 하는 siRNA는 단지 야생형 수용체만을 발현하는 세포의 생육성(viability)에 대해 최소한의 영향을 미치지만, *EGFR* 변이체를 발현하는 세포주에서는 급격한 세포사(cell death)를 유도하였다. 그 대응 *EGFR* 대립 유전자를 특이적으로 표적으로 삼는 siRNA의 능력은, 면역블로팅에 의해, 트랜스펙트된 COS7 세포에서 확인되었다. 따라서, EGFR 변이체의 발현은 이들 돌연변이를 보유하고 있는 폐암에서의 프로-세포사멸(pro-apoptotic) 신호의 억압(suppression)에 필수적인 것으로 나타난다. 또한, 야생형 수용체만을 발현하는 폐암세포는 EGFR 발현 시 유사한 의존성을 나타내지 않는다는 사실은, 야생형 EGFR을 과다발현하는 인간 종양의 상대적인 제피티니브-무감응성에 대해서 설명한다. EGFR 변이체를 보유하는 폐암에서의 제피티니브의 유효성은, 이 세포들이 절대적으로 의존하고 있는 결정적인 항-세포사멸 경로의 저해는 물론, 돌연변이 수용체의 변경된 생화학적 특성을 모두 반영할 수 있다. 본 발명자들은 EGFR 변이체가 야생형 수용체보다 EGF-의존적 자가인산화의 제피티니브 저해에 더 민감함을 앞서 보고하였다(T. J. Lynch et al., N. Engl. J. Med. 350, 2129(2004)). 돌연변이 수용체에 의한 이러한 증가된 약제 민감성은 Erk 및 STAT5 활성화 둘다에서도 관찰되었다. 따라서, 돌연변이 수용체에 의한 EGF-유도 신호전달은 특정 자가인산화 반응을 거쳐 하류 영향인자의 선택적 활성화를 나타내지만, 그들의 제피티니브에 의한 증강된 저해는 균일하며, 변경된 약제의 ATP 포켓 변이체 대한 결합을 반영할 수 있다.

[0515] 본 발명자들은, EGFR-매개 NSCLC 생존에 있어서의 증가된 Akt 및 STAT 신호전달의 관련성을 확립하기 위해서, 특이적인 약리작용 저해제를 이용하여 이 경로들을 표적으로 하였다. *EGFR* 돌연변이를 보유하고 있는 폐암세포는 야생형 수용체를 가진 세포보다 제피티니브에 100배 이상 민감하였다. 또한, EGFR 변이체를 발현하는 세포는 야생형 EGFR만을 발현하는 세포보다 Akt 또는 STAT 신호전달의 약리작용 저해에 더 민감하였다. EGFR 돌연변이 폐암 세포는 Akt/STAT-매개 항-세포사멸 신호의 교란(disruption)에 대한 증가된 민감성을 나타내었지만, 그들은 통상 사용되는 화학요법 약제인 독소루비신(doxorubicin) 및 시스플라틴(cisplatin), 및 프로-세포사멸 Fas-리간드에 의해 유도되는 세포사 신호에 대해 현저히 증가된 내성을 나타내었다.

[0516] EGFR 변이체를 가진 세포에서의 증강된 Akt/STAT 신호전달에 의해, 부가적인 치료 표적을 제공할 수 있는 한편, 종래의 화학요법이 이 종양들에 대해 덜 효과적일 수 있는 가능성을 높였다.

[0517] "종양유전자 중독(oncogene addiction)"은, 그들이 의존하고 있는 증식 신호를 억압한 후 종양 세포의 세포사멸을 설명하기 위해서 제안되었다(I. B. Weinstein, *Science* 297, 63(2002)). 흥미롭게도, 이마티니브(Imatinib)(Gleevec™)는 BCR-ABL 전위(translocation) 산물을 발현하는 만성 골수성 백혈병 및 활성화 c-Kit 돌연변이를 발현하는 위장관 기질 종양에서, 효율적으로 세포사를 야기하며, 둘다 상기 약제에 의해서 효과적으로 저해되는 구성적 STAT를 빈번하게 나타낸다(T. Kindler et al., *Leukemia* 17, 999(2003), G. P. Paner et al., *Anticancer Res*. 23, 2253(2003)). 마찬가지로, EGFR 키나제 돌연변이를 가진 폐암세포에 있어서, 제피티니브 반응성은 수용체 변이체에 의해 전달된 필수적인 항-세포사멸 신호의 효과적인 저해면에서 큰 부분을 차지할 수 있다.

[0518] 재료 및 방법

[0519] 면역블로팅(Immunoblotting)

[0520] 배양 세포로부터 얻은 세포용해물(lysate)은, 프로테아제 저해제를 함유하는, 얼음속에서 차게한 RIPA 용해 용액(ice-cold RIPA lysis solution)(1% 트리톤 X-100, 0.1% SDS, 50mM Tris-HCl, pH 7.4, 150mM NaCl, 1mM EDTA, 1mM EGTA, 10mM β-글리세롤 포스페이트, 10mM NaF, 1mM Na-오르토바나데이트)으로 제조하였다. 조직파편(debris)은 마이크로퓨저에서 4℃에서 12000 xg로 10분간 원심분리하여 제거하였다. 맑아진 용해물을 겔 로딩 완충액(gel loading buffer) 중에서 끓여서 10% SDS-PAGE에 의해 분리하였다. 단백질은 니트로셀룰로스에 전기전달(electrotransfer)시킨 후, 특이적 항체를 이용하여 검출한 다음, 당근과산화효소-접합 2차 염소항체(horseradish peroxidase-conjugated secondary goat antibody)와 배양하여(Cell signaling(Beverly, MA; 1:2000)), 증강된 화학발광(DuPont NEN)으로 현상하여 자가방사기록법(autoradiography)을 행한다. 포스포-EGFR Y845, Y992, Y1045, Y1068, 포스포-STAT5(tyr694), 포스포-AKT(Ser473), 포스포-ERK1/2(Thr202/Tyr204), AKT, STAT5, 및 ERK1/2 항체는 New England Biolab(Beverly, MA)으로부터 얻었다. 총 EGFR Ab-20 항체는 NeoMarkers(Fremont, CA)로부터 얻었다. 포스포-EGFR Y1173 항체는 Upstate Biotechnology(Lake Placid, NY)으로 부터 얻었고 총 포스포티로신 항체 PY-20은 Transduction Laboratories(Lexington, KY)로부터 얻었다. 모든 항체는 1:1000으로 희석하여 사용하였다.

[0521] EGFR 발현 벡터

[0522] 야생형, L858 또는 del L747-P753insS 돌연변이를 코딩하는 전장 EGFR 야생형 발현 구조체를, 표준 방법을 사용하여 플라스미드 pUSEamp내로 서브클로닝하였다. 모든 구조체는 DNA 염기서열 분석에 의해 확인하였다.

[0523] 세포주 및 트랜스펙션

[0524] COS7 세포 및 NMuMg(정상 마우스 유방 상피) 세포를, 2mM L-글루타민 및 50U/㎖ 페니실린/스트렙토마이신의 존재 하에 10%의 소태아혈청을 함유하는 DMEM (Dulbecco's modified Eagle's media)에서 성장시켰다. 상기 NCI-H358, NCI-H1650, NCI-H1734, NCI-H1666, 및 NCI-H1975 인간 폐암 세포주는 ATCC(American Type Culture Collection)의 콜렉션으로부터 얻은 후, 10% 소태아혈청, 2mM L-글루타민, 50U/㎖ 페니실린/스트렙토마이신 및 1mM 피루브산나트륨을 함유하는 RPMI1640에서 성장시켰다. 이들은, 본 명세서내에서, H358, H1650, H1734, H1666, 및 H1975와 같이 각각 약칭한다. COS7 세포의 과도 트랜스펙션은 Lipopectamin 2000(Invitrogen;

Carlsbad, CA)을 사용하여 수행하였다. 플라스미드(1㎍)는 10cm 디쉬상에서 80%의 세포충실도(confluence)로 세포내에 트랜스펙션시켰다. 12시간 후, 상기 세포를 수확하여 혈청이 없는 12-웰 플레이트에 재접종하였다. 다음날, 30ng/㎖의 EGF로 세포를 자극하였다. 안정한 NMuMg 세포주는 EGFR 발현 구조체와 약제-선별가능 플라스미드 pBABE puro를 함께 트랜스펙션시켜 제조한 다음, 3㎍/㎖ 푸로마이신(puromycin) 중에서 선별하였다. 약제 내성 세포의 풀(pool)을 분석용으로 사용하였다. 안정하게 트랜스펙트된 세포내에서의 EGFR 발현은 면역블로팅에 의해 확인하였다.

[0525] EGFR 발현의 SiRNA-매개 "녹다운(knockdown)"

[0526] EGFR L858R에 대한 siRNA는 누클레오티드 염기서열 CACAGATTTTGGGCGGGCCAA (서열번호 688)를 표적으로 하여 설계하였으며, GCTATCAAAACATCTCCGAAA(서열번호 689) 염기서열은 delE745-A750(Qiagen; Valencia, CA)용으로 사용하였다. EGFR의 모든 형태를 표적으로 하기 위해서는, 인간 야생형 EGFR에 대응하는 시판 siRNA를 Dharmcon(Lafayette.co)으로부터 입수하였다. siRNA의 트랜스펙션은 리포펙타민 2000(Lipofactamine 2000)(Invitrogen)을 사용하여 제조자의 지시에 따라 수행하였다. 세포는 MTT 측정법을 사용하여 96시간 후의 생존율에 대하여 측정하였다.

[0527] 세포사멸 측정(Apoptosis assay)

[0528] 10,000 세포를 96-웰 플레이트의 개별 웰에 접종하였다. 6시간 후, 배지를 바꾸어, 증가된 농도의 독소루비신(Sigma; St.Louis, MO), 시스플라틴(Sigma), Fas-리간드(인간 활성화, 클론 CH11; Upstate Biotechnology), Ly294002(Sigma), 또는 AG490(Calbiochem; La Jolla, CA) 존재 하에 세포를 유지하였다. 96시간 후, 세포의 생존율은 MTT 측정법을 이용하여 결정하였다. 카스파제 면역염색(caspase immunostaining)을 위해, 10mm 커버슬립(coverslip)상에 10,000개의 세포를 접종하였다. 다음날 siRNA로 트랜스펙션시켰다(상기에 상세히 기재되어 있음). 72시간 후, 실온에서 10분간 4% 파라포름알데히드 중에서 세포를 고정시켰다. 이어서, 0.5% 트리톤 X-100에서 5분간 침투시킨(permeabilized) 다음, 5% 정상염소혈청(NGS)에서 1시간 동안 블럭(block)시켰다. 그 다음, 상기 커버슬립을, 1:100으로 희석한 1차 항체(Cell Signaling사제의 분할(cleaved) 카스파제-3 Asp175 5A1) 중에서, 4℃에서, 하룻밤 배양하였다. 다음날 상기 커버슬립을 PBS로 3회 세정하고, 5% 정상염소혈청 및 0.5㎍/㎖의 DAPI(4',6-디아미디노-2-페닐인돌) 중에서, 1:250으로 희석한 2차 항체(염소 항-토기 텍사스-레드 접합; Jackson Immunoresearch 사제; West Grove, PA) 와 함께 1시간 동안 배양하였다. 상기 커브슬립을 PBS로 3회 세정한 후, 그 위에 Molecular Probes(Eugene, OR)사제의 ProLong Gold anti-fade 시약을 올려놓았다.

[0529] 세포 생존율 측정

[0530] 5mg/㎖ MTT(티아졸릴블루; Sigma) 용액 10㎕를 96-웰 플레이트의 각 웰에 첨가하였다. 37℃에서 2시간 배양한 후, 상기 배지를 제거하고, 각 웰에 100㎕의 산성 이소프로판올(0.1N HCl)을 첨가하여 상기 MTT를 가용화(solubilize)시켰다. 흡광도는 570nm에서의 분광광도 분석에 의해 결정하였다.

[0531] 성장 곡선

[0532] H-358, H-1650, H-1734, H-1666, 및 H-1975 세포의 성장 곡선은 96-웰 플레이트의 개별 웰에 1000개의 세포를 접종하여 얻었다. 각 세포주는 8개의 별개의 웰에서 평판배양하였다. 수일이 경과한 후, 상기 세포를 4% 포름알데히드 중에서 고정시킨 후 0.1%(w/v)의 크리스탈 바이올렛 용액으로 염색하였다. 상기 크리스탈 용액은 10% 아세트산 100㎕에 용해시켰으며, 상대 세포수를 결정하기 위해 플레이트 리더를 사용하여 570nm에서의 흡광도를 측정하였다.

[0533] 돌연변이 동정

[0534] EGFR 내에 돌연변이를 보유하고 있는 산발성 NSCLC 세포주를 동정하기 위해, 본 발명자들은 15개의 NSCLC 세포

주 패널(panel)내의 엑손 19 및 21을, 상기한 바와 같이, 염기서열분석하였다. 세포주는 흡연이력에 관계없는 기관지폐포 세포조직(bronchoalveolar histology)을 가진 종양(NCI-H358, NCI-H650, NCI-H1650) 또는 비흡연자 중에서 발생한 선암종으로부터 유래한 것에 근거하여 분석용으로 선별하였다(NCI-H1435, NCI-H1563, NCI-H1651, NCI-H1734, NCI-H1793, NCI-H1975, NCI-H2291, NCI-H2342, NCI-H2030, NCI-H1838, NCI-H2347, NCI-H2023). NCI-H1666은 단지 야생형 EGFR만을 보유하는 것으로 보고되었다(상기 실시예 참조). 모든 세포주는 미국의 ATCC으로부터 얻을 수 있다.

[0535] 명세서 전반에 걸쳐 여기에 인용된 참고자료는 전체적으로 참고로서 기재하였다.

[0536] [참 조]

1. Schiller JH, Harrington D, Belani CP, et al. Comparison of four chemotherapy regimens for advanced non-small cell lung cancer. N Engl J Med 2002; 346:92-98.
2. Druker BJ, Talpaz M, Resta DJ et al. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in Chronic Myeloid Leukemia. N Engl J Med 2001;344:1031-1037.
3. Arteaga CL. ErbB-targeted therapeutic approaches in human cancer. Exp Cell Res. 2003; 284:122-30.
4. Jorissen RN, Walker F, Pouliot N, Garrett TP, Ward CW, Burgess AW. Epidermal growth factor receptor: mechanisms of activation and signaling. Exp Cell Res 2003;284:31-53
5. Luetteke NC, Phillips HK, Qui TH, Copeland NG, Earp HS, Jenkins NA, Lee DC. The mouse waved-2 phenotype results from a point mutation in the EGF receptor tyrosine kinase. Genes Dev 1994;8:399-413.
6. Nicholson RI, Gee JMW, Harper ME. EGFR and cancer prognosis. Eur J Cancer. 2001;37:S9-15
7. Wong AJ, Ruppert JM, Bigner SH, et al. Structural alterations of the epidermal growth factor receptor gene in human gliomas. Proc Natl Acad Sci. 1992;89:2965-2969.
8. Ciesielski MJ, Genstermaker RA. Oncogenic epidermal growth factor receptor mutants with tandem duplication: gene structure and effects on receptor function. Oncogene 2000; 19:810-820.
9. Frederick L, Wang W-Y,Eley G, James CD. Diversity and frequency of epidermal growth factor receptor mutations in human glioblastomas. Cancer Res 2000; 60:1383-1387.

[0537]

10. Huang H-JS, Nagane M. Klingbeil CK, et al. The enhanced tumorigenic activity of a mutant epidermal growth factor receptor common in human cancers is mediated by threshold levels of constitutive tyrosine phophorylation and unattenuated signaling. J Biol Chem 1997;272:2927-2935

11. Pegram MD, Konecny G, Slamon DJ. The molecular and cellular biology of HER2/neu gene amplification/overexpression and the clinical development of herceptin (trastuzumab) therapy for breast cancer. Cancer Treat Res 2000;103:57-75.

12. Ciardiello F, Tortora G. A novel approach in the treatment of cancer targeting the epidermal growth factor receptor. Clin Cancer Res. 2001;7:2958-2970

13. Wakeling AE, Guy SP, Woodburn JR et al. ZD1839 (Iressa): An orally active inhibitor of Epidermal Growth Factor signaling with potential for cancer therapy. Cancer Res 2002;62:5749-5754.

14. Moulder SL, Yakes FM, Muthuswamy SK, Bianco R, Simpson JF, Arteaga CL. Epidermal growth factor receptor (HER1) tyrosine kinase inhibitor ZD1839 (Iressa) inhibits HER2/neu (erbB2)-overexpressing breast cancer cells *in vitro* and *in vivo*. Cancer Res 2001;61:8887-8895.

15. Moasser MM, Basso A, Averbuch SD, Rosen N. The tyrosine kinase inhibitor ZD1839 ("Iressa") inhibits HER2-driven signaling and suppresses the growth of HER-2 overexpressing tumor cells. Cancer Res 2001;61:7184-7188.

16. Ranson M, Hammond LA, Ferry D, et al. ZD1839, a selective oral epidermal growth factor receptor-tyrosine kinase inhibitor, is well tolerated and active in patients with solid, malignant tumors: results of a phase I trial. J Clin Oncol. 2002; 20: 2240-2250.

17. Herbst RS, Maddox A-M, Rothernberg ML, et al. Selective oral epidermal growth factor receptor tyrosine kinase inhibitor ZD1839 is generally well tolerated and has activity in non-small cell lung cancer and other solid tumors: results of a phase I trial. J Clin Oncol. 2002;20:3815-3825.

[0538]

18. Baselga J, Rischin JB, Ranson M, et al. Phase I safety, pharmacokinetic and pharmacodynamic trial of ZD1839, a selective oral Epidermal Growth Factor Receptor tyrosine kinase inhibitor, in patients with five selected solid tumor types. J Clin Onc 2002;20:4292-4302.

19. Albanell J, Rojo F, Averbuch S, et al. Pharmacodynamic studies of the epidermal growth factor receptor inhibitor ZD1839 in skin from cancer patients: histopathologic and molecular consequences of receptor inhibition. J Clin Oncol. 2001; 20:110-124.

20. Kris MG, Natale RB, Herbst RS, et al. Efficacy of Gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer: A randomized trial. JAMA 2003;290:2149-2158.

21. Fukuoka M, Yano S, Giaccone G, et al. Multi-institutional randomized phase II trial of gefitinib for previously treated patients with advanced non-small-cell lung cancer. J Clin Oncol 2003;21:2237-2246.

22. Giaccone G, Herbst RS, Manegold C, et al. Gefitinib in combination with gemcitabine and cisplatin in advanced non-small-cell lung cancer: A phase III trial-INTACT 1. J Clin Oncol 2004;22:777-784.

23. Herbst RS, Giaccone G, Schiller JH, et al. Gefitinib in combination with paclitaxel and carboplatin in advanced non-small-cell lung cancer: A phase III trial – INTACT 2. J Clin Oncol 2004;22:785-794.

24. Rich JN, Reardon DA, Peery T, et al. Phase II Trial of Gefitinib in recurrent glioblastoma. J Clin Oncol 2004;22:133-142

25. Cohen MH, Williams GA, Sridhara R, et al. United States Food and Drug Administration Drug Approval Summary: Gefitinib (ZD1839;Iressa) Tablets. Clin Cancer Res. 2004;10:1212-1218.

26. Cappuzzo F, Gregorc V, Rossi E, et al. Gefitinib in pretreated non-small-cell lung cancer (NSCLC): Analysis of efficacy and correlation with HER2 and

[0539]

epidermal growth factor receptor expression in locally advanced or Metastatic NSCLC. J Clin Oncol. 2003;21:2658-2663.

27. Fitch KR, McGowan KA, van Raamsdonk CD, et al. Genetics of Dark Skin in mice. Genes & Dev 2003;17:214-228.

28. Nielsen UB, Cardone MH, Sinskey AJ, MacBeath G, Sorger PK. Profiling receptor tyrosine kinase activation by using Ab microarrays. Proc Natl Acad Sci USA 2003;100:9330-5.

29. Burgess AW, Cho H, Eigenbrot C, et al. An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. Mol Cell 2003;12:541-552.

30. Stamos J, Sliwkowski MX, Eigenbrot C. Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. J Biol Chem. 2002;277:46265-46272.

31. Lorenzato A, Olivero M, Patrane S, et al. Novel somatic mutations of the MET oncogene in human carcinoma metastases activating cell motility and invasion. Cancer Res 2002; 62:7025-30.

32. Davies H, Bignell GR, Cox C, et al. Mutations of the BRAF gene in human cancer. Nature 2002;417:906-7.

33. Bardelli A, Parsons DW, Silliman N, et al. Mutational analysis of the tyrosine kinome in colorectal cancers. Science 2003;300:949.

34. Daley GQ, Van Etten RA, Baltimore D. Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome. Science 1990;247:824-30.

35. Heinrich, MC, Corless CL, Demetri GD, et al. Kinase mutations and imatinib response in patients with metastatic gastrointestinal stromal tumor. J Clin Oncol 2003;21:4342-4349.

36. Li B, Chang C, Yuan M, McKenna WG, Shu HG. Resistance to small molecule inhibitors of epidermal growth factor receptor in malignant gliomas. Cancer Res 2003;63:7443-7450.

[0540]

37. C. L. Sawyers, Genes Dev 17, 2998-3010 (2003).
38. G. D. Demetri et al., N Engl J Med 347, 472-80 (2002).
39. B. J. Druker et al., N Engl J Med 344, 1038-42. (2001).
40. D. J. Slamon et al., N Engl J Med 344, 783-92 (2001).
41. H. Davies et al., Nature 417, 949-54 (2002).
42. Bardelli et al., Science 300, 949 (2003).
43. Y. Samuels et al., Science (2004).
44. Jemal et al., CA Cancer J Clin 54, 8-29 (2004).
45. S. Breathnach et al., J Clin Oncol 19, 1734-1742 (2001).
46. V. Rusch et al., Cancer Res 53, 2379-85 (1993).
47. R. Bailey et al., Lung Cancer 41 S2, 71 (2003).
48. M. Fukuoka et al., J Clin Oncol 21, 2237-46 (2003).
49. P. A. Janne et al., Lung Cancer 44, 221-230 (2004).
50. M. G. Kris et al., Jama 290, 2149-58 (2003).
51. V. A. Miller et al., J Clin Oncol 22, 1103-9 (2004).
52. M. Huse, J. Kuriyan, Cell 109, 275-82 (2002).
53. K. Naoki, T. H. Chen, W. G. Richards, D. J. Sugarbaker, M. Meyerson, Cancer Res 62, 7001-3 (2002).
54. J. Stamos, M. X. Sliwkowski, C. Eigenbrot, J Biol Chem 277, 46265-72 (2002).
55. T. Fujishita et al., Oncology 64, 399-406 (2003).
56. M. Ono et al., Mol Cancer Ther 3, 465-472 (2004).
57. M. C. Heinrich et al., J Clin Oncol 21, 4342-9 (2003).
58. G. Giaccone et al., J Clin Oncol 22, 777-84 (2004).
59. R. S. Herbst et al., J Clin Oncol 22, 785-94 (2004).
60. H. Yamazaki et al., Mol Cell Biol 8, 1816-20 (1988).
61. M. E. Gorre et al., Science 293, 876-80 (2001).
62. Marchetti A, Martella C, Felicioni L, et al: EGFR mutations in non-small-cell lung cancer: analysis of a large series of cases and development of a rapid and sensitive method for diagnostic screening with potential implications on pharmacologic treatment. J Clin Oncol 23:857-65, 2005.

[0541]

63. Franklin WA: Diagnosis of lung cancer: pathology of invasive and preinvasive neoplasia. Chest 117:80S-89S, 2000.

64. Paez JG, Janne PA, Lee JC, et al: EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 304:1497-500, 2004.

65. Lynch TJ, Bell DW, Sordella R, et al: Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 350:2129-39, 2004.

66. Pao W, Miller V, Zakowski M, et al: EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci U S A 101:13306-11, 2004

67. Huang SF, Liu HP, Li LH, et al: High frequency of epidermal growth factor receptor mutations with complex patterns in non-small cell lung cancers related to gefitinib responsiveness in Taiwan. Clin Cancer Res 10:8195-203, 2004.

68. Han SW, Kim TY, Hwang PG, et al: Predictive and Prognostic Impact of Epidermal Growth Factor Receptor Mutation in Non-Small-Cell Lung Cancer Patients Treated With Gefitinib. J Clin Oncol, 2005.

69. Tokumo M, Toyooka S, Kiura K, et al: The relationship between epidermal growth factor receptor mutations and clinicopathologic features in non-small cell lung cancers. Clin Cancer Res 11:1167-73, 2005.

70. Mitsudomi T, Kosaka T, Endoh H, et al: Mutations of the Epidermal Growth Factor Receptor Gene Predict Prolonged Survival After Gefitinib Treatment in Patients with Non-Small-Cell Lung Cancer With Postoperative Recurrence. J Clin Oncol, 2005.

71. Pao W, Wang TY, Riely GJ, et al: KRAS Mutations and Primary Resistance of Lung Adenocarcinomas to Gefitinib or Erlotinib. PLoS Med 2:e17, 2005

72. Read WL, Page NC, Tierney RM, et al: The epidemiology of bronchioloalveolar carcinoma over the past two decades: analysis of the SEER database. Lung Cancer 45:137-42, 2004.

[0542]

73. Sanderson Cox L, Sloan JA, Patten CA, et al: Smoking behavior of 226 patients with diagnosis of stage IIIA/IIIB non-small cell lung cancer. Psychooncology 11:472-8, 2002.
74. Radzikowska E, Glaz P, Roszkowski K: Lung cancer in women: age, smoking, histology, performance status, stage, initial treatment and survival. Population-based study of 20 561 cases. Ann Oncol 13:1087-93, 2002.
75. Tong L, Spitz MR, Fueger JJ, et al: Lung carcinoma in former smokers. Cancer 78:1004-10, 1996.
76. de Perrot M, Licker M, Bouchardy C, et al: Sex differences in presentation, management, and prognosis of patients with non-small cell lung carcinoma. J Thorac Cardiovasc Surg 119:21-6, 2000
77. Capewell S, Sankaran R, Lamb D, et al: Lung cancer in lifelong non-smokers. Edinburgh Lung Cancer Group. Thorax 46:565-8, 1991
78. Gritz ER, Nisenbaum R, Elashoff RE, et al: Smoking behavior following diagnosis in patients with stage I non-small cell lung cancer. Cancer Causes Control 2:105-12, 1991
79. Sridhar KS, Raub WA, Jr.: Present and past smoking history and other predisposing factors in 100 lung cancer patients. Chest 101:19-25, 1992
80. Barbone F, Bovenzi M, Cavallieri F, et al: Cigarette smoking and histologic type of lung cancer in men. Chest 112:1474-9, 1997
81. Shigematsu H, Lin L, Takahashi T, et al: Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst 97:339-46, 2005
82. Kosaka T, Yatabe Y, Endoh H, et al: Mutations of the epidermal growth factor receptor gene in lung cancer: biological and clinical implications. Cancer Res 64:8919-23, 2004
83. Cho D, Kocher O, Tenen DG, et al: Unusual cases in multiple myeloma and a dramatic response in metastatic lung cancer: case 4. Mutation of the epidermal growth factor receptor in an elderly man with advanced, gefitinib-responsive, non-small-cell lung cancer. J Clin Oncol 23:235-7, 2005

[0543]

[0544] <표 1> 비소세포 폐암을 가진 9명의 환자의 특성 및 제피티니브에 대한 반응성

[0545]

환자 번호	성별	제피티니브 치료시작 나이(년)	병리 종류 *	사전 치료 수	흡연 상태 †	치료 기간 (개월)	전체 생존 ‡	*EGFR* 돌연변이 ∫	반응성¶
1	F	70	BAC	3	비흡연	15.6	18.8	있음	주요(major): 개선된 폐 병변
2	M	66	BAC	0	비흡연	>14.0	>14.0	있음	주요: 개선된 양측 폐 병변
3	M	64	Adeno	2	비흡연	9.6	12.9	있음	일부(partial): 개선된 폐병변 및 소프트-조직 덩어리
4	F	81	Adeno	1	이전흡연	>13.3	>21.4	있음	소수(minor): 개선된 흉막병변
5	F	45	Adeno	2	비흡연	>14.7	>14.7	있음	부분: 개선된 간 병변
6	M	32	BAC	3	비흡연	>7.8	>7.8	있음	주요: 개선된 폐 병변
7	F	62	Adeno	1	이전흡연	>4.3	>4.3	있음	부분 : 개선된 간 및 폐 병변

8	F	58	Adeno	1	이전흡연	11.7	17.9	있음	부분 : 개선된 간 병변
9	F	42	BAC	2	비흡연	>33.5	>33.5	없음	부분 : 개선된 폐 결절

[0546] * 기관지폐포 암종(BAC)의 요소를 가진 선암종(Adeno)은 BAC로서 기재한다.

[0547] †흡연상태는 시험개시 12개월내에 담배를 전혀 피우지 않은 환자는 이전흡연으로, 일생동안 100개피 이하로 피운 사람을 비흡연자로 규정하였다.

[0548] ‡전체 생존은 제피티니브 치료시작부터 죽음까지 측정하였다,

[0549] ∫*EGFR*은 표피성장인자 수용체 유전자를 나타낸다.

[0550] ¶부분(partial)반응은 고형 종양에서의 반응 평가 기준을 사용하여 평가하였다; 주요(major) 및 소수(minor) 반응은 이러한 기준을 사용하여 반응이 측정할 수 없었던 환자에서는 두명의 의사에 의해서 평가하였다.

[0551] <표 2> 비소세포 폐암을 가진 환자의 EGFR의 티로신 키나제 도메인에서의 체세포 돌연변이

환자	서열 번호	돌연변이	돌연변이 효과
제피티니브에 대한 반응성을 가진 환자			
환자 1	730	15개의 누클레오티드 결실 (2235-2249)	인프레임 결실(746-750)
환자 2	731	12개의 누클레오티드 결실 (2240-2251)	인프레임 결실(747-751) 및 세린 잔기의 삽입
환자 3	732	18개의 누클레오티드 결실 (2240-2257)	인프레임 결실(747-753) 및 세린 잔기의 삽입
환자 4	733	18개의 누클레오티드 결실 (2240-2257)	인프레임 결실(747-753) 및 세린 잔기의 삽입
환자 5	734	누클레오티드 2573의 T가 G로 치환	아미노산 치환(L858R)
환자 6	735	누클레오티드 2573의 T가 G로 치환	아미노산 치환(L858R)
환자 7	736	누클레오티드 2582의 T가 A로 치환	아미노산 치환(L861Q)
환자 8	737	누클레오티드 2155의 G가 T로 치환	아미노산 치환(G719C)
제피티니브에 노출하지 않은 환자 *			
환자 A	738	18개의 누클레오티드 결실 (2240-2257)	인프레임 결실(747-753) 및 세린 잔기의 삽입
환자 B	739	15개의 누클레오티드 결실 (2235-2249)	인프레임 결실(746-750)

[0552] 삭제

[0553] * 제피티니브에 노출하지 않은 25명의 환자(15명은 기관지폐포암, 7명은 선암종, 및 3명은 대세포 암종) 중에서, 2명(환자 A 및 B : 둘다 기관지폐포암을 가짐)은 *EGFR* 돌연변이를 갖는다. 다양한 조직 타입을 나타내는 14개의 폐암 세포주에서는 돌연변이가 발견되지 않았다: 비소세포 폐암(6개의 표본), 소세포 폐암(6개의 표본), 기관지 카르시노이드(1개의 표본), 및 불명 타입(1개의 표본). *EGFR*내에서 동정된 다형성 변이체(polymorphic variant)는 다음을 포함한다: 티로신 키나제 도메인 내의, 누클레오티드 1562의 G가 A로 치환, 누클레오티드 1887의 T가 A로 치환, 및 기능적 중요성이 알려지지 않은 배계열 변이체, 누클레오티드 2885의 G가 A로 치환.

〈표 4〉 NSCLC 치료의 일부분으로서 *EGFR* 돌연변이에 대한 시험을 행한 100명의 환자중 집단 특성

특성	빈도
평균 연령 (표준 편차)	60.7 (11.0)
여성	63
인종	
백인	76
아시아인	7
기타	5
불명	12
시험시의 단계	
I	15
II	4
III	10
IV	67
불명	4
조직학	
순수 BAC	1
BAC 특징을 가진 선암종	24
선암종	69
NSCLC, 모든 다른 아류	6
흡연 상태	
현재 흡연자	17
이전 흡연자	48
비흡연자	29
불명	6
현재 및 이전 흡연자의 평균 흡연량, 갑년(표준 편차)	39.0 (32.3)
전단부터 *EGFR* 시험까지의 평균 시간, 개월(표준 편차)	18.7 (78.4)
사전 화학 요법 치료	47
사전 EGFR 표적 치료	11

[0554] BAC = 기관지폐포 암종, EGFR = 표피성장인자 수용체

〈표 5〉 등정된 표피성장인자 수용체 체세포 유전자 돌연변이

[0555]

환자	성별	조직학	흡연 갑년	엑손	누클레오티드 변화	아미노산 변화	서열번호
1	F	Adeno	0	18 18	2126A>T 2155G>A	E709V G719S	740
2	F	A+BAC	60	18 20	2156G>C 2327G>A	G719A R776H	741
3	F	A+BAC	0	19	2235_2249 del	K745_A750 del ins K	742
4	M	A+BAC	0	19	2235_2249 del	K745_A750 del ins K	743
5	F	Adeno	5	19	2235_2249 del	K745_A750 del ins K	744
6	M	Adeno	불명	19	2235_2249 del	K745_A750 del ins K	745
7	F	Adeno	0	19	2236_2250 del	E746_A750 del	746
8	M	Adeno	45	19	2236_2250 del	E746_A750 del	747
9	F	Adeno	불명	19	2236_2250 del	E746_A750 del	748
10	M	A+BAC	12	19	2237_2255 del ins T	E746_S752 del ins V	749
11	M	Adeno	1	19	2239_2248 del ins C	L747_A750 del ins P	750
12	M	A+BAC	0	19	2239_2251 del ins C	L747_T751 del ins P	751
13	F	Adeno	30	19	2253_2276 del	T751_I759 del ins T	752
14	F	Adeno	0	19	2254_2277 del	S752_I759 del	753
15	F	Adeno	0	20	2303_2311 dup	D770_N771 ins SVD	754
16	M	Adeno	5	20	2313_2318 dup CCCCCA	P772_H773 dup	755
17	F	Adeno	0	21	2543C>T	P848L*	756
18	M	BAC	0	21	2573T>G	L858R	757
19	F	A+BAC	0	21	2573T>G	L858R	758
20	M	A+BAC	1	21	2573T>G	L858R	759
21	F	Adeno	0	21	2573T>G	L858R	760
22	F	Adeno	15	21	2573T>G	L858R	761
23	F	Adeno	0	21	2582T>A	L861Q	762

Adeno = 선암종, Adeno + BAC = 기관지폐포 암종 특징을 가진 선암종 , BAC = 기관지폐포 암종
* 이 돌연변이는 배계열 변이체로서 동정되었다.

〈표 S1A〉 선별된 EGFR 및 수용체 티로신키나제 엑손 증폭용 프라이머
엑손(서열번호 : 1-212)

유전자	RefSeq	엑손	서열번호	F 네스티드	R 네스티드
ALK	NM_004304	24	1,2	GGAAATATAGGGAAGGGAAGGAA	TTGACAGGGTACCAGCGAGATGA
ALK	NM_004304	25	3,4	CTGAACCGCCAAGGACTCAT	TTTTCCCTCCCTACTAACACACG
AXL	NM_021913	19	5,6	ACTGATGCCCTGACCCTGTT	CCCATGGTTCCCCACTCTT
CSF1R	NM_005211	18	7,8	AGGGACTCCAAAGCCATGTG	CTCTCTGGGGCCATCCACT
CSF1R	NM_005211	19	9,10	CATTGTCAAGGGCAATGTAAGTG	CTCTCACCAACCCTCGCTGT
DDR1	NM_013994	15	11,12	ACATGGGGAGCCAGAGTGAC	TGCAACCCAGAGAAACGTGTG
DDR2	NM_006182	16	13,14	TGAGCTTTCAACCCTAGTTTGTTG	GTTTGCCTCCTGCTGTCTCA
DKFZp761P1010	NM_018423	8	15,16	TGTCCTTGTGTTTTTGAAGATTCC	TGCAGACAGATGACAACACATGAA
EGFR	NM_005228	2	17,18	TGGGTGAGTCTCTGTGTGGAG	CATTGCCATAGCAAAAATAAACACA
EGFR	NM_005228	3	19,20	GGTTCAACTGGGCGTCCTA	CCTTCTCCGAGGTGGAATTG
EGFR	NM_005228	4	21,22	CGCACCATGGCATCTCTTTA	AAAACGATCTCTATGTCCGTGGT
EGFR	NM_005228	5	23,24	CAGCCAGCCAAACAATCAGA	TCTTTGGAGTCTTCAGAGGGAAA
EGFR	NM_005228	6	25,26	TGTGGTTTCGTTGGAAGCAA	AATTGACAGCTCCCCCACAG
EGFR	NM_005228	7	27,28	GGCTTTCTGACGGGAGTCAA	CCACCCAAAGACTCTCCAAGA
EGFR	NM_005228	8	29,30	CCTTTCCATCACCCCTCAAG	AGTGCCTTCCCATTGCCTAA
EGFR	NM_005228	9	31,32	ACCGGAATTCCTTCCTGCTT	CACTGAAACAAACAACCAGGGTGA
EGFR	NM_005228	10	33,34	AGGGGGTGAGTCACAGGTTC	TCAGAAGAAATGTTTTTATTCCAAGG
EGFR	NM_005228	11	35,36	GCAAATCCAATTTTCCCACTT	GCAGGAGCTCTGTGCCCTAT
EGFR	NM_005228	12	37,38	TCCCACAGCATGACCTACCA	TTTGCTTCTTAAGGAACTGAAAA
EGFR	NM_005228	13	39,40	TGTCACCCAAGGTCATGGAG	CAAAAGCCAAGGGCAAAGAA
EGFR	NM_005228	14	41,42	GGAGTCCCAACTCCTTGACC	GTCCTGCCCACACAGGATG
EGFR	NM_005228	15	43,44	GCTTTCCCCACTCACACACA	CAAACCTCGGCAATTTCGTTG
EGFR	NM_005228	16	45,46	CCACCAATCCAACATCCAGA	TGGCCCAGAGCCATAGAAAC
EGFR	NM_005228	17	47,48	TTCCAAGATCATTCTACAAGATGTCA	GCACATTCAGAGATTCTTTCTGC
EGFR	NM_005228	18	49,50	TCCAAATGAGCTGGCAAGTG	TCCCAAACACTCAGTGAAACAAA
EGFR	NM_005228	19	51,52	GTGCATCGCTGGTAACATCC	TGTGGAGATGAGCAGGGTCT
EGFR	NM_005228	20	53,54	ATCGCATTCATGCGTCTTCA	ATCCCCATGGCAAACTCTTG
EGFR	NM_005228	21	55,56	GCTCAGAGCCTGGCATGAA	CATCCTCCCCTGCATGTGT
EGFR	NM_005228	22	57,58	TGGCTCGTCTGTGTGTGTCA	CGAAAGAAAATACTTGCATGTCAGA
EGFR	NM_005228	23	59,60	TGAAGCAAATTGCCCAAGAC	TGACATTTCTCCAGGGATGC
EGFR	NM_005228	24	61,62	AAGTGTCGCATCACCAATGC	ATGCGATCTGGGACACAGG
EGFR	NM_005228	25	63,64	GGCACCTGCTGGCAATAGAC	TGACTTCATATCCATGTGAGTTTCACT
EGFR	NM_005228	26	65,66	TATACCCTCCATGAGGCACA	GGGAAAAACCCACACAGGAA
EGFR	NM_005228	27	67,68	TCAGAACCAGCATCTCAAGGA	GATGCTGGAGGGAGCACCT
EGFR	NM_005228	28_1	69,70	CCTTGTTGAGGACATTCACAGG	ATGTGCCCGAGGTGGAAGTA
EPHA1	NM_005232	14	71,72	GGAGGGCAGAGGACTAGCTG	GTGCCTGGCCAAGTCTTTGT
EPHA1	NM_005232	15	73,74	CTGCAGCCTAGCAACAGAGC	AAGAACCAGAGGAGCCAGGA
EPHA2	NM_004431	13	75,76	CGGGTAAGGATGTGGGTTGT	CAGGTGTTCTGCCTCCTGAA
EPHA2	NM_004431	14	77,78	GCTTCAGGAGGCAGAACACC	GGAGCAAGCCTAAGAAGGTTCA
EPHA3	NM_005233	10	79,80	GCCTTGTATCCATTTGCCACA	TGACAACACGTTTTGGCGTCAT
EPHA3	NM_005233	11	81,82	TGCATATTCCATTTCAGAACAGA	AAACAGTTTCATTGCTGCTAAAT
EPHA4	NM_004438	13	83,84	CCGGATACAGATACCCAAAAAGA	GGAGGCTTCAAGGGATCGAGA
EPHA4	NM_004438	14	85,86	GCTGTTGTCCTGCTTGGCTA	TGGTTGTAATGTTGAACTAGCTTGC
EPHA7	NM_004440	13	87,88	TGGCTGTCAGCTAAATAAGCATGT	TCAATTTGCTTCATTTCTCCTGTT
EPHA7	NM_004440	14	89,90	TGCTGCTGAACTACCAACCAA	TGTGGTAGTAATTGTGGAAAACTG
EPHA8	NM_020526	13	91,92	CAAAGCACCGTCTCAACTCG	CCCGAAACTGCCAACTTCAT
EPHA8	NM_020526	14	93,94	GGAAAACAGGACCCCAGTGT	CCCTCCTCCACAGAGCTGAT
EPHB1	NM_004441	7	95,96	GACAGAAGCTGACAAGCAGCA	AGGTTCCATTCCCTCCCAGT
EPHB1	NM_004441	8	97,98	TGGGAGTGAGAGTTTGGAAGAA	TATGAGGCCGTGAGCTCGAAA
EPHB2	NM_017449	11	99,100	AGGGCCCTGCTCTGGTTT	CCAATTGGGCGTTAGTCGAAA
EPHB2	NM_017449	12	101,	CTCATGAGATTGGGGCATCA	AGGCCCATGATCTCAGAAGC
EPHB3	NM_004443	11	103,	GGTTGCAGGAGAGACGAGGT	AGGCCCTTCACCCTGTGAC
EPHB3	NM_004443	12	105	ATGACCCCTCCGATCCTACC	TAATCCTGCTCCACGGCCATT

[0556]

〈표S1A〉 계속

EPHB4	NM_004444	14	107	GGAAAAAGCAGAGGCAGGTG	TGGTCTCAAGAACCCAGCAG
EPHB6	NM_004445	16	109	GACACCCTCCCCCTCTCAT	ACTATGACACCCCGGCTGAG
EPHB6	NM_004445	17	111	TGCTTGATGTAAAACCCTTGG	GCAATCCAACAGCCATGAGA
ERBB2	NM_004448	21	113	GGAGCAAACCCCTATGTCCA	TCCTCCAACTGTGTGTTGTGG
ERBB3	NM_001982	21	115	TGGGGACCACTGCTGAGAG	TGCAGCCTTCTCTCCTTGAA
FGFR1	NM_000604	14	117	GCAGAGCAGTGTGGCAGAAG	ACAGGTGGGAAGGGACTGG
FGFR1	NM_000604	15	119	AGTGGGGTGGGCTGAGAAC	TCTCTGGGGCAGAAAGAGGA
FGFR2	NM_000141	14	121	ACCCGGCCACACTGTATTTC	CATCCCACCCAGCTCTCAAC
FGFR2	NM_000141	15	123	AGGGCATAGCCCTATTGAGC	CCCAGGAAAAAGCCAGAGAA
FGFR3	NM_000142	13	125	CAGGTGTGGGTGGAGTAGGC	CTCAGGCGCCATCCACTT
FGFR3	NM_000142	14	127	AAGAAGACGACCAACGTGAGC	AGGAGCTCCAGGGCACAG
FGFR4	NM_002011	14	129	CCTCCTCTGTAAAGTGGGTGGA	AGAGGGCCTCAGTGCAGGAGT
FGFR4	NM_002011	15	131	AGATGGGGCAGAACTGGATG	GGGTCCCAGACCAAATCTGA
FLT1	NM_002019	23	133	AGGTGCTCCCTTCACAGCAT	TTCAGGGACTACAGCTGAGGAA
FLT1	NM_002019	24	135	GCCGTATGTTATCTGGGAGGT	TGGGCCCATTACACTTTAAGA
FLT3	NM_004119	20	137	TTCCATCACCGGTACCTCCT	CCATAAATCAAAAATGCACCACA
FLT3	NM_004119	21	139	GAGTGGTCTTAGGAAGATGATGC	AAAGTCATGGGCTGCAATACAA
FLT4	NM_002020	23	141	ATGGTCCCCACTGCTTGG	AGGAGCTCACCTCACCCTGT
IGF1R	NM_000875	18	143	CCTTGCGTCTCTCCACACAT	TGGCAACGGGTAACAATGAA
INSR	NM_000208	18	145	GGCTGAGGTAAGCTGCTTCG	AAAAAGAAGTATCTTGCCCCTTT
INSR	NM_000208	19	147	AACCCCTCTTAGGGCTCTGTG	CAGGAGGATGGCAGGCTTC
KDR	NM_002253	24	149	CGTAGAGAGCTTCAGGACCTGTG	TTCCGAGAAGTTTTGCCTGA
KIT	NM_000222	17	151	TGTGAACATCATTCAAGGCGTA	AAAATGTGTGATATCCCTAGACAGG
KIT	NM_000222	18	153	TCCACATTTCAGCAACAGCA	GGCTGCTTCCTGAGACACAGT
LTK	NM_002344	16	155	TATCTACCGGTGCGGGACTT	AGGTGTAGCCTCCCCTCACA
MERTK	NM_006343	17	157	AGGCTGGTGGTGTCTCTGTG	CAAGCTGCCAACCCTCAGTT
MET	NM_000245	19	159	TGGATTTCAAATACTGAAGCCACT	TGGAATTGGTGGTGTTGAATTT
MUSK	NM_005592	15 1	161	GGGCTTCATATGTTCTGACATGG	CAGAGGACCACGCCATAGG
MUSK	NM_005592	15 2	163	CCGAGATTTAGCCACCAGGA	CCTGGGAAGCAAACAACACA
NTRK1	NM_002529	15	165	AGGTCCCCAGTCTCCTCTCC	AGACCCATGCAGCCATCCTA
NTRK1	NM_002529	16	167	CGTGAACCACCGAGCTTGT	AGAGGGGCAGAAGGGGAAC
NTRK2	NM_006180	15	169	GGTGGGGGTGAGGAGCTTAG	TCGTTTAAGCCACCCAGTCA
NTRK2	NM_006180	16	171	TGCAAATAAGGAAAGCAAACA	TCCTGACATGGTCTTCCAACC
NTRK3	NM_002530	17	173	CAGCATCTTCACACACCTCTGA	GCTGGCTCTAAATCCCACCT
NTRK3	NM_002530	18	175	CTAATCCGGGAAGTTGTTGC	TTCTGTATCAGCAGCTTCTCTGTG
PDGFRA	NM_006206	18	177	CAAGTGCCACCATGGATCA	GGCAGTGTACTGACCCCTTGA
PDGFRA	NM_006206	19	179	GCACAAGTTATTAAGAGCCCAAGG	AGCATACTGGCCTCACACCA
PDGFRB	NM_002609	18	181	GCACATGGGCAGTGTTGTATTT	GAGCCCCACACAGATTTCCT
PDGFRB	NM_002609	19	183	ATGGGACGGAGAAGTGGTTG	TCCCTGTATCAGGGCTCCTC
PTK7	NM_002821	18	185	TTCCTACGCAGCACACCAAT	GCAGGCACTAAACCCTTTCC
PTK7	NM_002821	19	187	GCACGCATGTGACCAATTTC	AGCCCTGAGAGGGAGGTAGG
RET	NM_000323	15	189	CACACACCACCCCTCTGCT	AAAGATTTGGGGTGAGGGCTA
RET	NM_000323	16	191	CTGAAAGCTCAGGGATAGGG	CTGGCCAAGCTGCACAGA
ROR1	NM_005012	09 1	193	TGCAGCCAACGATTTGAAAG	GGAAAGCCCCAAGTCTGAAA
ROR1	NM_005012	09 2	195	TCATCATGAGATCCCCACACT	GCATTTCCCCCTGAAGGAGT
ROR1	NM_005012	09 3	197	TGGATTCAGTAACCAGGAAGTGA	CCCATTCCACCAGGATGATT
ROR1	NM_005012	09 4	199	GTTTCCAGCTGCCCACTACC	GCTCGAAACCACATGTTCCA
RYK	NM_002958	13	201	CTGGATTTGGGGTTCTCTGC	CGGGAACAGCTAGCAGATTTTT
TEK	NM_000459	18	203	GGGAATTTTGGAGGGGAACT	GCTTCAGTCACCACAGAGCA
TEK	NM_000459	19	205	TGAGTCTACCCAGCAATCATTTG	TTCCCGAGAGCTACAGGACA
TIE	NM_005424	18	207	GGTAACAAGGGTACCCACGAA	GTTTGAGGGGCTGAGTGTGG
TIE	NM_005424	19	209	CCTCACCCTTAGGGCTTGTG	AGCCCAGGTCATGCCTTAGA
TYRO3	NM_006293	18	211,212	GGGTAGCTTGGGAGCAAAGA	CCAAACCCCAGAGAGCAGAC

[0557]

〈 표 S1B 〉 선별된 EGFR 및 수용체 티로신키나제 엑손 증폭용 프라이머
엑손 (서열번호 213 - 424)

유전자	RefSeq	엑손	서열번호	F	R
ALK	NM_004304	24	213,214	CATTTCCCCTAATCCTTTTCCA	GTGATCCCAGATTTAGGCCTTC
ALK	NM_004304	25	215	GCCTCTCGTGGTTTGTTTTGTC	CCCAGGGTAGGGTCCAATAATC
AXL	NM_021913	19	217	CTTCCTGGTGGAGGTGACTGAT	CAGGCATAGTGTGTGATGGTCA
CSF1R	NM_005211	18	219	TCACGATACACATTCTCAGATCC	GAAGATCTCCCAGAGGAGGATG
CSF1R	NM_005211	19	221	CGTAACGTGCTGTTGACCAAT	AAACGAGGGAAGAGCCAGAAAG
DDR1	NM_013994	15	223	TGGGGAGCACAATAAAAGAAGA	ACTCTTGGCTCCTGGATTCTTG
DDR2	NM_006182	16	225	GGAAGTCAGTGTGCAGGGAATA	TTTTAGCAGAAATAGGCAAGCA
DKFZp761P1010	NM_018423	8	227	TGGTAATCCTAAACACAATGCAGA	CTGGGCAACACAGTGAGATCCT
EGFR	NM_005228	2	229	TCACAAATTTCTTTGCTGTGTCC	CATGGAACTCCAGATTAGCCTGT
EGFR	NM_005228	3	231	GATTGTTGCAGATCGTGGACAT	CGCTTAAATCTTCCCATTCCAG
EGFR	NM_005228	4	233	CTCCATGGCACCATCATTAACA	CTCAGGACACAAGTGCTCTGCT
EGFR	NM_005228	5	235	GCAGTTCATGGTTCATCTTCTTTT	CAAAATAGCCCACCCTGGATTA
EGFR	NM_005228	6	237	CTTTCTGCATTGCCCAAGATG	CAAGGTCTCAGTGAGTGGTGGA
EGFR	NM_005228	7	239	GAGAAGGGTCTTTCTGACTCTGC	CAGGTGTTTCTCCTGTGAGGTG
EGFR	NM_005228	8	241	CACATTGCGGCCTAGAATGTTA	ACCCCGTCACAACCTTCAGT
EGFR	NM_005228	9	243	GCCGTAGCCCCAAAGTGTACTA	TCAGCTCAAACCTGTGATTTCC
EGFR	NM_005228	10	245	CTCACTCTCCATAAATGCTACGAA	GACTTAACGTGTCCCCTTTTGC
EGFR	NM_005228	11	247	GCCTCTTCGGGGTAATCAGATA	GAAGTCTGTGGTTTAGCGGACA
EGFR	NM_005228	12	249	ATCTTTTGCCTGGAGGAACTTT	CAGGGTAAATTCATCCCATTGA
EGFR	NM_005228	13	251	CAGCAGCCAGCACAACTACTTT	TTGGCTAGATGAACCATTGATGA
EGFR	NM_005228	14	253	TGAATGAAGCTCCTGTGTTTACTC	ATGTTCATCGCAGGCTAATGTG
EGFR	NM_005228	15	255	AAAACAGGGAGAACTTCTAAGCAA	CATGGCAGAGTCATTCCCACT
EGFR	NM_005228	16	257	CAATGCTAGAACAACGCCTGTC	TCCCTCCACTGAGGACAAAGTT
EGFR	NM_005228	17	259	GGGAGAGCTTGAGAAAGTTGGA	ATTTCCTCGGATGGATGTACCA
EGFR	NM_005228	18	261	TCAGAGCCTGTGTTTCTACCAA	TGGTCTCACAGGACCACTGATT
EGFR	NM_005228	19	263	AAATAATCAGTGTGATTCGTGGAG	GAGGCCAGTGCTGTCTCTAAGG
EGFR	NM_005228	20	265	ACTTCACAGCCCTGCGTAAAC	ATGGGACAGGCACTGATTTGT
EGFR	NM_005228	21	267	GCAGCGGGTTACATCTTCTTTC	CAGCTCTGGCTCACACTACCAG
EGFR	NM_005228	22	269	CCTGAACTCCGTCAGACTGAAA	GCAGCTGGACTCGATTTCCT
EGFR	NM_005228	23	271	CCTTACAGCAATCCTGTGAAACA	TGCCCAATGAGTCAAGAAGTGT
EGFR	NM_005228	24	273	ATGTACAGTGCTGGCATGGTCT	CACTCACGGATGCTGCTTAGTT
EGFR	NM_005228	25	275	TAAGGCACCCACATCATGTCA	TGGACCTAAAAGGCTTACAATCA_A
EGFR	NM_005228	26	277	GCCTTTTAGGTCCACTATGGAATG	CCAGGCGATGCTACTACTGGTC
EGFR	NM_005228	27	279	TCATAGCACACCTCCCTCACTG	ACACAACAAAGAGCTTGTGCAG
EGFR	NM_005228	28_1	281	CCATTACTTTGAGAAGGACAGGAA	TATTCTTGCTGGATGCGTTTCT
EPHA1	NM_005232	14	283	AGGAGGGCAGAGGACTAGCTG	GGCAATGTGAATGTGCACTG
EPHA1	NM_005232	15	285	CTTGAACCTGGGAGGTGGAG	ATCAGGGTGGGAGGAGTAAAGA
EPHA2	NM_004431	13	287	CCCACTTACCTCTCACCTGTGC	GTGAACTTCCGGTAGGAAATGG
EPHA2	NM_004431	14	289	AGGGGACCTCAAGGGAGAAG	AGATCATGCCAGTGAACTCCAG
EPHA3	NM_005233	10	291	GGACCAGGAAAGTCCTTGCTTT	GGTGGGGAACATTAAACTGAGG
EPHA3	NM_005233	11	293	GCTTCAGGTTGTTTTGTTGCAG	ACCCTTGCTTGAGGGAAATATG
EPHA4	NM_004438	13	295	CCCAGCTCCTAGGGTACAGTCT	CAGTCAGCTTCAAAATCCCTCTT
EPHA4	NM_004438	14	297	TCACTTCCCTGTGAGTAAAGAAAA	GGCCATTTAATTCTTGTCCTTGA
EPHA7	NM_004440	13	299	TGGACTTGTGCAAACTCAAACTG	TCCCAATATAGGGCAGTCATGTT
EPHA7	NM_004440	14	301	TCTCAATCAGTTGAGTTGCCTTG	AGCTGTGCAAGTGTGGAAACAT
EPHA8	NM_020526	13	303	GCTGTGAGGGTAAATGAGACCA	GTCTCCTGGTGAGTGACTGTGG
EPHA8	NM_020526	14	305	CCTTCCTTCGTCTCCACAGC	GTCCTTGTGCCAACAGTCGAG
EPHB1	NM_004441	7	307	GCTTGGCAAGGAGAAGAGAACA	GCTTGCTTTCTTGCTTGAACAAC
EPHB1	NM_004441	8	309	GCTGGTCACCTTGAGCTTCTCT	CCATGCTGGGCTCTTTGATTA
EPHB2	NM_017449	11	311	CACCACTCTGAAGTTGGCCTCT	ATGGCTCTGCACATTTGTTCC
EPHB2	NM_017449	12	313	CAGAGTGGGAAAAGGCACTTCA	CCAGAGTCCTGTGCAGACATTC
EPHB3	NM_004443	11	315	ATGGGGATTAACTGGGATGTTG	CGTAGCTCCAGACATCACTAGCA
EPHB3	NM_004443	12	317	GCAACCTGGTCTGCAAAGTCTC	ACCCAGCAGTCCAGCATGAG

[0558]

〈 표 S1B 〉 계속

EPHB4	NM_004444	14	319	GAGTTTCAGTGAGCCAAGATCG	TTACAGGCTTGAGCCACTAGGC
EPHB6	NM_004445	16	321	AAGCTTCCAGGAGACGAGGTC	GTCCCTGAAATCCCTCAAACC
EPHB6	NM_004445	17	323	TGCTCCATAAACGTGACTATTGC	GTAAGAGGGTGGGCTGGAATCT
ERBB2	NM_004448	21	325	CTTAGACCATGTCCGGGAAAAC	CACATCACTCTGGTGGGTGAAC
ERBB3	NM_001982	21	327	AAATTTCATCCCAAAACCAACC	CCAGTCCCAAGTTCTTGATCATT
FGFR1	NM_000604	14	329	ACAAGTCGGCTAGTTGCATGG	TCTCAGATGAAACCACCAGCAC
FGFR1	NM_000604	15	331	TTCATCTGAGAAGCAAGGAGTGG	CCAGGGAGAAAGCAGGACTCTA
FGFR2	NM_000141	14	333	TTCTGGCGGTGTTTTGAAATTA	CTCAACATTGACGGCCTTTCTT
FGFR2	NM_000141	15	335	TCAGCTCTTAAACAGGGCATAGC	GAAATGCAGCAGCCACTAAAGA
FGFR3	NM_000142	13	337	CTCACCTTCAAGGACCTGGTGT	CAGGGAGGGGTAGAAACCACA
FGFR3	NM_000142	14	339	GGAGAGGTGGAGAGGCTTCAG	GAGACTCCCAGGACAGACACCT
FGFR4	NM_002011	14	341	CACTCGTTCCTCACCCTTCC	AGGACTCACACGTCACTCTGGT
FGFR4	NM_002011	15	343	GGACAATGTGATGAAGATTGCTG	ATAGCAGGATCCCAAAAGACCA
FLT1	NM_002019	23	345	GGCTTGGGGACCTGTATTTGTA	CAGTGGCCTTTCTGAGCCTTAC
FLT1	NM_002019	24	347	GCACTCTAGCTCCCTCTTTTAGC	TTTTACAGTAGAGGGCAGACATGC
FLT3	NM_004119	20	349	GCCACCATAGCTGCAGAATTAG	CCCAAGGACAGATGTGATGCTA
FLT3	NM_004119	21	351	GCCTTTGTTCGAGAGGAGTTGT	GTTCACGCTCTCAAGCAGGTTA
FLT4	NM_002020	23	353	ATTCCACAAGCTCTCTCCATGA	CTTGCCCCAAGATGCCTAAG
IGF1R	NM_000875	18	355	TGCTTGGTATTTGCTCATCATGT	CCCTTAGCTAGCCCACTGACAA
INSR	NM_000208	18	357	CTCCTGGGAGTGGTGTCCAA	CCTGGGCAACAGACAGAGTAAG
INSR	NM_000208	19	359	CTTCACTTCCCCATGCGTACC	GGGTTCACAATGCCTACAGGA
KDR	NM_002253	24	361	AAAATCTGTGACTTTGGCTTGG	GGGAGGAGACATTCTTTGATTTG
KIT	NM_000222	17	363	GCAGTCCTGAGAAGAAAACAGC	CTTCACATGCCCCAAAATTACA
KIT	NM_000222	18	365	TGAGCCATGTATTTCAGAGGTGA	TACATTTCAGCAGGTGCGTGTT
LTK	NM_002344	16	367	TTGCCTACTCTGTAGGGATATTGC	ATAGGGCATGTAGCCCAGTGA
MERTK	NM_006343	17	369	GCTCTGCTGTTGGTCCTCACT	TTGCAAAGCACACATCTTCTGA
MET	NM_000245	19	371	TGGCAATGTCAATGTCAAGCAT	GTATGTTGCCCCACTCAACAAA
MUSK	NM_005592	15_1	373	TGCATTTCCTAGCTGAGACTCC	TGCCATCTCGCACGTAGTAAAT
MUSK	NM_005592	15_2	375	CTCTCCTGTGCTGAGCAGCTTT	TGTTTCCAATCACTGGCTTTCA
NTRK1	NM_002529	15	377	GAACCATGGGCTGTCTCTGG	ATCTGGGATAGCGAAGGAGACA
NTRK1	NM_002529	16	379	ATTACAGGCCACACGCCATC	AAGGCAAGAATAAGGGAGGAAGA
NTRK2	NM_006180	15	381	GCTCTCAGGACTGCAGAAGTACA	GAGGAACCAATCCCACTCACAC
NTRK2	NM_006180	16	383	TCACTCTTTGCCTTCTGTCTCTG	GCACTGTGCTTTGCTTTCTCAG
NTRK3	NM_002530	17	385	TGTCTCCTTTATCGTAGGTCTCCA	CACCACATTTCCTACAGTTCCA
NTRK3	NM_002530	18	387	CACTGTGCACCAGACAGACAAA	TGTGGTTTTCTGTATCAGCAGCTT
PDGFRA	NM_006206	18	389	CAGGGAGTCTGAAATCATCAGG	TCAAGTATCTAGCCCCAAATCCA
PDGFRA	NM_006206	19	391	GGCAATATTGACCATTCATCATTC	AGGCCAGGAGTAAGACGCAAC
PDGFRB	NM_002609	18	393	AAGAACGTACGTGTGGTGTTGG	CGCTATACTTGCTCCATGCACT
PDGFRB	NM_002609	19	395	AGGAAACAGCCTCTGGTCCTC	GTCAATGCTCAGACAGGGAGAT
PTK7	NM_002821	18	397	CCCAGGAAGGCAGGTACTGTTA	TTTTACAACCACCAAGGGTGTG
PTK7	NM_002821	19	399	TCGTGTGGTTACCTCCAGATTTT	AAATTAGCCAGGGAGTGGAGGT
RET	NM_000323	15	401	CATGCCATGCTATGGCTCAC	AGGCTGAGCGGAGTTCTAATTG
RET	NM_000323	16	403	ATCTCAGCAATCCACAGGAGGT	ATTTGCCTCACGAACACATCAT
ROR1	NM_005012	09_1	405	TGGAAAGTTGTCTATGGCACCTC	ATGGGCAGCAAGGACTTACTCT
ROR1	NM_005012	09_2	407	CACCCCAATATTGTCTGCCTTC	GGCTCGGGAACATGTAATTAGG
ROR1	NM_005012	09_3	409	CCATCATGTATGGCAAATTCTCTT	TGGCGTCTCCTAGTAAAGATGCT
ROR1	NM_005012	09_4	411	GCCAGATTGCTGGTTTCATTG	GGCTAAAACACAAAGCACCATT
RYK	NM_002958	13	413	GGGAAGTCATCCACAAAGACCT	GGTCTGGGTCACAGCTCCTC
TEK	NM_000459	18	415	TTCTTCTGCCAAGATGTGGTGT	TGCAGATGCTGCAATCATGTTA
TEK	NM_000459	19	417	TGGACCCCGAAAGATAAATAGG	TTCTGCACTCCTCTGGAAACTG
TIE	NM_005424	18	419	GGGTGAGAGCCAACACTGATCT	CTGTGCCCTCTCATCTCACACT
TIE	NM_005424	19	421	AGAACCTAGCCTCCAAGATTGC	ACACCTTCCAAGACTCCTTCCA
TYRO3	NM_006293	18	423,424	GACTCGAGGGTGGGAGACAG	GCTGTCACTAGGTGTCCTGAGC

[0559]

[0560]

〈 표S2 〉미처리 폐암에서의 *EGFR* 돌연변이 상태

시료	조직학	기원	성별	엑손	서열 변화	서열 번호	누글레오티드	아미노산
S0514	선암종	미국	F	18	치환	425	2155G>A	G719S
S0377	선암종	일본	F	18	치환	426	2155G>A	G719S
S0418	선암정	일본	F	19	Del-1a	427	2235_2249delGGAATTAAGAGAAGC	E746_A750del
S0363	대세포 암종	일본	F	19	Del-1a	428	2235_2249delGGAATTAAGAGAAGC	E746_A750del
S0380	선암종	일본	M	19	Del-1a	429	2235_2249delGGAATTAAGAGAAGC	E746_A750del
S0399	선암종	일본	F	19	Del-1a	430	2235_2249delGGAATTAAGAGAAGC	E746_A750del
S0353	선암종	일본	F	19	Del-1a	431	2235_2249delGGAATTAAGAGAAGC	E746_A750del
S0385	선암종	일본	M	19	Del-1a	432	2235_2249delGGAATTAAGAGAAGC	E746_A750del
S0301	선암종	일본	M	19	Del-1a	433	2235_2249delGGAATTAAGAGAAGC	E746_A750del
S0412	선암종	일본	M	19	Del-1b	434	2236_2250delGAATTAAGAGAAGCA	E746_A750del
S0335	선암종	일본	M	19	Del-1b	435	2236_2250delGAATTAAGAGAAGCA	E746_A750del
S0405	선암종	일본	F	19	Del-1b	436	2236_2250delGAATTAAGAGAAGCA	E746_A750del
S0439	선암종	일본	M	19	Del-2	437	2254_2277delTCTCCGAAAGCCAACAAGGAAATC	S752_I759del
S0361	선암종	일본	F	21	치환	438	2573T>G	L858R
S0388	선암종	일본	F	21	치환	439	2573T>G	L858R
S0389	선암종	일본	F	21	치환	440	2573T>G	L858R

등록특허 10-1289774

〈표 S3A〉 제피티니브처리 폐암에서의 *EGFR* 돌연변이 상태

[0561]

제피티니브 민감성	시료	조직학	기원	성별	엑손	서열 변화	서열 번호	누클레오티드	아미노산
Y	IR1T	선암종	U.S.	M	19	Del-3	441	2239_2247delTTAAGAGAA, 2248G>C	L747_E749del, A750P
Y	P003	선암종	U.S.	M	19	Del-3	442	2239_2247delTTAAGAGAA, 2248G>C	L747_E749del, A750P
Y	IR4T	기관지폐포성 암종	U.S.	F	19	Del-4	443	2240_2257delTAAGAGAAGCAACATCTC	L747_S752del, P753S
Y	IR2T	선암종	U.S.	F	19	Del-5	444	2238_2255delATTAAGAGAAGCAACATC, 2237A>T	L747_S752del, E746V
Y	IR3T	선암종	U.S.	F	21	치환	445	2573T>G	L858R
Y in vitro	IRG H3255	선암종	U.S.	F	21	치환	446	2573T>G	L858R
N	IR5	선암종	U.S.	F	18-24	미검출		n/a	n/a
N	IR6	선암종	U.S.	M	18-24	미검출		n/a	n/a
N	IR8	선암종	U.S.	F	18-24	미검출		n/a	n/a
N	IR9	NSCLC	U.S.	F	18-24	미검출		n/a	n/a

[0562]

〈 표S3B 〉 표2, 표S2, 또는 표 S3A 에서 나타내지 않은 돌연변이

시료	조직	엑손	서열 변화	누클레오티드	아미노산	서열 번호
Tar4T	폐선암종	19	결실	2239-2250delTTAAGAGAAGCA; 2251A>C	L747_A750del; T751T	554
AD355	폐선암종	19	결실	2240-2250delTAAGAGAAGCA	L747_T751del	720
IR TT	폐선암종	19	결실	2257-2271delCCGAAAGCCAACAAG	P753_K757del	721
AD240	폐선암종	20	결실	2309-2310insCAACCCGG	D770_N771insNPG	722
AD261	폐선암종	20	결실	2311-2312insGCGTGGACA	D770_N771insSVD	723
	폐선암종	20	결실	2316-2317insGGT	P772_H773insV	724
AD356	폐선암종	20	치환	2334-2335GG>AA	G779S	725
SP02-23	급성골수종 백혈병	21	치환	2570G>T	G857V	726
SP08-94	신경아교종	21	치환	2582T>A	L861Q	727
SP06-45	육종	21	치환	2648T>C	L883S	728
AD241	직장 선암종	22	치환	2686G>T	D896Y	729

등록특허 10-1289774

[0563]

〈 표 S3C 〉 EGFR 에서, 이마티니브에 대한 내성을 가진 BCR-ABL 돌연변이 위치 및 유사 위치

내성돌연변이로 되는 Abl1 잔기	유사 EGFR 잔기	동일/유사/비보존적
Met-244	Lys-714	비보존적
Leu-248	Leu-718	동일
Gly-250	Ser-720	비보존적
Gln-252	Ala-722	비보존적
Tyr-253	Phe-723	유사
Glu-255	Thr-725	비보존적
Asp-276	Ala-750	비보존적
Thr-315	Thr-790	동일
Phe-317	Leu-792	유사
Met-351	Met-825	동일
Glu-355	Glu-829	동일
Phe-359	Leu-833	유사
His-396	His-870	동일
Ser-417	Thr-892	유사
Phe-486	Phe-961	동일

[0564]

〈 표 S4 〉 cDNA 염기서열 분석용 프라이머

프라이머 명칭	서열번호	프라이머 염기서열 5'→3'
cDNA_EGFR_aF	447	TGTAAAACGACGGCCAGTCGCCCAGACCGGACGACA
cDNA_EGFR_aR	448	CAGGAAACAGCTATGACCAGGGCAATGAGGACATAACCA
cDNA_EGFR_bF	449	TGTAAAACGACGGCCAGTGGTGGTCCTTGGGAATTTGG
cDNA_EGFR_bR	450	CAGGAAACAGCTATGACCCCATCGACATGTTGCTGAGAAA
cDNA_EGFR_cF	451	TGTAAAACGACGGCCAGTGAAGGAGCTGCCCATGAGAA
cDNA_EGFR_cR	452	CAGGAAACAGCTATGACCCGTGGCTTCGTCTCGGAATT
cDNA_EGFR_dF	453	TGTAAAACGACGGCCAGTGAAACTGACCAAAATCATCTGT
cDNA_EGFR_dR	454	CAGGAAACAGCTATGACCTACCTATTCCGTTACACACTTT
cDNA_EGFR_eF	455	TGTAAAACGACGGCCAGTCCGTAATTATGTGGTGACAGAT
cDNA_EGFR_eR	456	CAGGAAACAGCTATGACCGCGTATGATTTCTAGGTTCTCA
cDNA_EGFR_fF	457	TGTAAAACGACGGCCAGTCTGAAAACCGTAAAGGAAATCAC
cDNA_EGFR_fR	458	CAGGAAACAGCTATGACCCCTGCCTCGGCTGACATTC
cDNA_EGFR_gF	459	TGTAAAACGACGGCCAGTTAAGCAACAGAGGTGAAAACAG
cDNA_EGFR_gR	460	CAGGAAACAGCTATGACCGGTGTTGTTTTCTCCCATGACT
cDNA_EGFR_hF	461	TGTAAAACGACGGCCAGTGGACCAGACAACTGTATCCA
cDNA_EGFR_hR	462	CAGGAAACAGCTATGACCTTCCTTCAAGATCCTCAAGAGA
cDNA_EGFR_iF	463	TGTAAAACGACGGCCAGTGATCGGCCTCTTCATGCGAA
cDNA_EGFR_iR	464	CAGGAAACAGCTATGACCACGGTGGAGGTGAGGCAGAT
cDNA_EGFR_jF	465	TGTAAAACGACGGCCAGTCGAAAGCCAACAAGGAAATCC
cDNA_EGFR_jR	466	CAGGAAACAGCTATGACCATTCCAATGCCATCCACTTGAT
cDNA_EGFR_kF	467	TGTAAAACGACGGCCAGTAACACCGCAGCATGTCAAGAT
cDNA_EGFR_kR	468	CAGGAAACAGCTATGACCCTCGGGCCATTTTGGAGAATT
cDNA_EGFR_lF	469	TGTAAAACGACGGCCAGTTCAGCCACCCATATGTACCAT
cDNA_EGFR_lR	470	CAGGAAACAGCTATGACCGCTTTGCAGCCCATTTCTATC
cDNA_EGFR_mF	471	TGTAAAACGACGGCCAGTACAGCAGGGCTTCTTCAGCA
cDNA_EGFR_mR	472	CAGGAAACAGCTATGACCTGACACAGGTGGGCTGGACA
cDNA_EGFR_nF	473	TGTAAAACGACGGCCAGTGAATCCTGTCTATCACAATCAG
cDNA_EGFR_nR	474	CAGGAAACAGCTATGACCGGTATCGAAAGAGTCTGGATTT
cDNA_EGFR_oF	475	TGTAAAACGACGGCCAGTGCTCCACAGCTGAAAATGCA
cDNA_EGFR_oR	476	CAGGAAACAGCTATGACCACGTTGCAAAACCAGTCTGTG

도면의 간단한 설명

[0042] 도 1A~1B는 난치의 비소세포 폐암(NSCLC)에 있어서의 제피티니브 반응에 대한 대표도이다. 도 1A는 제피티니

브로 처치하기 전의 오른쪽 폐의 커다란 덩어리를 보여주는, 케이스 6(표 1)의 흉부 CT 스캔이며, 도 1B는 제피티니브 처치 개시로부터 6주 후의 현저한 개선을 보여준다.

[0043] 도 2는 제피티니브-반응성 종양에서의 EGFR 돌연변이들을 보여준다.

[0044] 도 2A~2C는 키나제 도메인 내에 이형접합(heterozygous) 인프레임 결실을 갖는 종양 표본 내의 EGFR 유전자의 누클레오티드 서열을 보여준다(이중 피크들)(각각 도시된 순서대로 서열번호 643~644 및 690~699). 상기 결실의 두 절단점을 명시하기 위해 센스와 안티센스 방향에서의 트레이싱(tracings)을 표시하였고, 야생형 누클레오티드 서열은 대문자로, 돌연변이 서열은 소문자로 나타내었다. delL747-T751insS 돌연변이의 5' 절단점 앞에는 코딩된 아미노산을 변경하지 않는 T로부터의 C로의 치환이 선행한다.

[0045] 도 2D와 도 2E는 이형접합 미스센스 돌연변이(heterozygous missense mutation)(화살표)들을 보여주는데, 이들의 결과로 티로신 키나제 도메인 내에서 아미노산이 치환된다(서열번호 701 및 703). 이중 피크들은 이형접합 돌연변이 부위에서의 2개의 누클레오티드를 나타낸다. 비교를 위하여, 대응하는 야생형 서열도 나타내었다(서열번호 700 및 702).

[0046] 도 2F는 EGF 리간드에 의해 결합된 이량체화된 EGFR 분자의 도식도이다. 세포외 도메인(2개의 수용체 리간드[L]-도메인들과 1개의 퓨린유사(furin-like) 도메인을 함유함), 막횡단 영역, 및 세포질 도메인(촉매 키나제 도메인을 함유함)이 강조되어 있다. EGFR의 자가인산화에 의해 활성화되는 다운스트림 효과기(STAT3, MAP 키나제(MAPK), 및 AKT)와 함께, 수용체 활성화의 표지로 사용되는 자가인산화 부위인, 티로신^{1068}의 위치(Y-1068)가 지적되어 있다. 모두 티로신 키나제 도메인 내에 있는, 종양과 관련된 돌연변이들의 위치가 나타나 있다.

[0047] 도 3은 EGFR 돌연변이체의 증진된 EGF-의존성 활성화와 EGFR 돌연변이체의 제피티니브에 대한 증대된 감수성을 보여준다.

[0048] 도 3A는 EGF를 혈청 결핍 세포들(serum starved cells)에 첨가한 후의, delL747-P753insS와 L858R 돌연변이체의 리간드-유도 활성화의 시간 경로(time course)를 야생형 EGFR과 비교하여 보여준다. Cos-7 세포들(대조군; 오른쪽 패널)에서 발현된 EGFR의 전체 수준과 비교하여, EGFR의 자가인산화가 수용체 활성화의 표지로 이용되며, EGFR(왼쪽 패널)의 인산화된 티로신^{1068} 잔기를 특이적으로 인식하는 항체를 이용한 웨스턴 블로팅법이 사용된다. EGFR의 자가인산화는 EGF(10 ng/ml) 첨가 후, 간격을 두고 측정된다.

[0049] 도 3B는 야생형과 돌연변이체 수용체의 인산화의 EGF-유도를 그래프로 나타낸 것이다(패널 A를 참조). 세 개의 독립적인 실험들로부터의 자가방사선상(autoradiograph)을 NIH 이미지 소프트웨어를 사용하여 계량하였으며, EGFR 인산화의 강도는 전체 단백질 발현에 대해 정규화되어 있고, 표준편차와 함께 그 수용체의 퍼센트 활성화로 표현되어 있다.

[0050] 도 3C는 제피티니브에 의한 EGFR 활성화 저해의 투여량 의존성을 보여준다. 야생형 또는 돌연변이형 수용체를 발현하는 Cos-7 세포를 100 ng/ml의 EGF로 30분간 자극하고 웨스턴 블로팅 분석하여 EGFR 티로신^{1068}의 자가인산화를 나타내었다. 세포들은 처리되지 않았거나(U), 또는 도면에 나타나 있는 바와 같이, 제피티니브의 농도를 점차 증가시키면서 3시간 동안 전처리된 것들이다(왼쪽 패널). 발현된 EGFR 단백질의 총량은 대조군(오른쪽 패널)에 나타나 있다.

[0051] 도 3D은 패널 3C에 묘사된 2개의 실험들로부터의 결과를 계량화한 것이다(NIH 이미지 소프트웨어). 인산화된 EGFR의 농도는 단백질 발현 수준에 대해 정규화되었으며, 그 수용체의 퍼센트 활성화로 나타내어져 있다.

[0052] 도 4는 EGFR의 ATP-결합 포켓 내에 있는 주요한 부위들에서의 돌연변이들의 클러스터링을 나타낸다.

[0053] 도 4A는 복수의 경우의 NSCLC(서열번호 495~504(DNA))에 있어서, EGFR 유전자의 엑손 19의 중첩되는 인프레임 결실의 위치와, 엑손 21의 미스센스 돌연변이의 위치를 보여준다. 각 엑손의 부분적인 누클레오티드 서열이 나타나 있는데, 점선에 의해 결실이 표시되어 있고, 미스센스 돌연변이는 밑줄 및 강조되어 있으며, 야생형 EGFR 누클레오티드와 아미노산 서열들이 나타나 있다(서열번호 493 및 494(DNA) 및 509~510(아미노산)).

[0054] 도 4B는 키나제 도메인의 아미노 로브(N)와 카르복시 로브(C)의 사이에 위치한 EGFR ATP 틈의 3차원적 구조를 보여준다(좌표는 PDB 1M14로부터 얻은 것이고, Cn3D 소프트웨어를 이용하여 도시한 것이다). 제피티니브를 나타내는 저해제는 상기 ATP 틈을 차지하고 있는 것으로 그려져 있다. 2개의 미스센스 돌연변이들의 위치가 키나제의 활성화 루프 내에 나타나 있고, 3개의 인프레임 결실들은 모두 다른 루프 내에 존재하며, 상기 ATP 틈의

측면에 위치하고 있다.

[0055] 도 4C는 EGFR 키나제 도메인의 클로즈업 도면인데, ATP나 또는 저해제와의 결합에 참여하는 주요한 아미노산 잔기가 나타나 있다. 구체적으로는, 제피티니브 같은 4-아닐리노퀴나졸린 화합물은 ATP-결합 부위를 차지함으로써 촉매과정을 저해하며, 메티오닌^{793}(M793) 및 시스테인^{775}(C775) 잔기와 수소결합을 형성한다. 반면 그들의 아닐리노 환(ring)은 메티오닌766(M766), 리신745(K745), 및 류신788(L788) 잔기에 가깝다. 돌연변이의 표적으로 되는 상기 루프 내에서의 인 프레임 결실은 저해제에 대한 이들 아미노산의 위치를 변경시키는 것으로 예측된다. 돌연변이된 잔기들은 상기 티로신 키나제의 활성화 루프 내에 나타나 있다.

[0056] 도 5A-5D에는 erbB1 유전자의 누클레오티드와 아미노산 서열이 나타나 있다. 아미노산들은 당업자들에게 알려져 있는 단일 글자들로 표시되어 있다. 키나제 도메인에서의 누클레오티드 변이들은 환자 번호에 의해 강조되어 있다(표 2 참조). 서열번호 511은 누클레오티드 1에서 3633까지를 포함한다. 서열번호 512는 아미노산 1에서 1210까지를 포함한다.

[0057] 도 6A-6C: EGFR과 B-Raf 키나제 도메인들 내의 선택된 영역들의 서열 정렬. 인간 NSCLC에서의 EGFR 돌연변이들의 묘사. NSCLC 종양에서의 EGFR(gb:X00588;) 돌연변이들은 회색으로 강조되어 있다. 복수의 종양형들(5)에서의 B-Raf(gb:M95712) 돌연변이들은 검은색으로 강조되어 있다. 별표 표시들(*)은 EGFR과 B-Raf 사이에 보존된 잔기들을 가리킨다. 도 6A는 활성화 루프 내(서열번호 477-479)의 L858R 돌연변이를 나타낸다. 도 6B는 P-루프(서열번호 480-482) 내의 G719S 돌연변이를 나타낸다. 도 6C는 EGFR 엑손 19(서열번호 483-489) 내의 결실 돌연변이를 나타낸다.

[0058] 도 7: EGFR 키나제 도메인의 3차원적 구조 내의, 미스센스 돌연변이들인 G719S와 L858R의 위치 및 Del-1 결실의 위치. 활성화 루프는 노란색으로, P-루프는 파란색으로 나타나 있고, 그리고 C-로브 및 N-로브는 표시되어 있다. 돌연변이나 또는 결실의 표적이 된 잔기들은 빨간색으로 강조되어 있다. Del-1 돌연변이는 코돈 746~750 내의 ELREA 잔기들을 표적으로 한다. 돌연변이들은 키나제들 내의 고도로 보존된 영역들 내에 위치하며, ATP와 또한 제피티니브와 에를로티니브가 결합할 것으로 예측되는 영역을 둘러싸고 있는, P-루프와 활성화 루프들에서 발견된다.

[0059] 도 8A-8F. 정상조직 및 종양조직들로부터 얻은 EGFR DNA의 대표적인 크로마토그램들. 확인된 돌연변이들의 위치는 다음과 같다. 도 8A는 엑손 18 키나제 도메인 P루프(서열번호 704~705)를 나타낸다. 도 8B는 엑손 21 키나제 도메인 A-루프(서열번호 706~707)를 나타낸다. 도 8C는 엑손 19 키나제 도메인 Del-1(서열번호 708~710)을 나타낸다. 도 8D는 엑손 19 키나제 도메인 Del-3(서열번호 711~713)을 나타낸다. 도 8E는 엑손 19 키나제 도메인 Del-4(서열번호 714~716)을 나타낸다. 도 8F는 엑손 19 키나제 도메인 Del-5(서열번호 717~719)을 나타낸다.

[0060] 도 9: EGFR과 BCR-ABL 폴리펩티드들의 서열 정렬과, 약에 대한 내성을 갖는 표현형을 부여하는 잔기들의 위치. GenBank 고유번호 NM_005228에 개시된 누클레오티드 서열에 의해 코딩되는 EGFR 폴리펩티드(서열번호 492)와, GenBank 고유번호 M14752에 개시된 누클레오티드 서열에 의해 코딩되는 BCR-ABL 폴리펩티드(서열번호 491)가 정렬되어 있고, 보존된 잔기들은 음영처리되어 있다. 티로신 키나제 저해제 이마티니브(imatinib)(STI571, Glivec/Gleevec)에 대한 내성을 부여하는 BCR-ABL 돌연변이들은 별표(*)로 표시되어 있다.

[0061] 도 10은 EGFR 검사가 진행 중인 전이성 NSCLC 환자를 위한 의사결정 프로세스를 보여준다.

[0062] 도 11은 EGFR 엑손 18-24(크기를 반영하지 않은)의 도식도이다. 화살표들은 확인된 돌연변이들의 위치를 가리킨다. 별표(*)는 각 위치에서 돌연변이를 가진 환자의 수를 나타낸다. 확대된 도면은 엑손 19의 결실들의 중첩과, 각 결실(서열번호 511의 누클레오티드 2233~2277 및 서열번호 512의 잔기 745~759)을 갖는 환자 수를 보여 준다. 이 결과는 현재까지의 모든 EGFR 돌연변이를 포함하는 것은 아니다.

[0063] [발명의 상세한 설명]

[0064] 본 발명은 암에 걸린 환자에 있어서 표피성장인자수용체(EGFR) 표적 치료의 유효 가능성을 결정하는 신규한 방법을 제공한다. 상기 방법은 상기 환자의 erbB1 유전자의 키나제 도메인에서의 적어도 하나의 핵산 변이의 존재 또는 부재를 검출하는 것을 포함한다. 적어도 하나의 변이가 존재하는 것은 EGFR 표적 치료가 유효할 것임을 나타낸다. 바람직하게, 상기 핵산 변이는 EGFR 키나제 활성을 증대시킨다. 그렇게 되면, 상기 환자는 EGFR 표적 치료를 받을 수 있게 된다. 본 발명의 하나의 태양에서는, 상기 EGFR 표적 치료는 티로신 키나제 저해제이다. 바람직한 태양에서는, 상기 티로신 키나제 저해제는 아닐리노퀴나졸린이다. 상기 아닐리노퀴나

졸린은 합성 아닐리노퀴나졸린일 수 있다. 바람직하게는 상기 합성 아닐리노퀴나졸린은 제피티니브이거나 또는 에를로티니브이다.

[0065] **정의:**

[0066] "ErbB1", "표피성장인자 수용체", 및 "EGFR"이라는 용어들은 본 원에서 서로 호환적으로 사용되며, 예를 들면 Carpenter et al. Ann. Rev. Biochem. 56:881-914(1987)에 개시되어 있는 천연 서열 EGFR 및 그 변이체들(예, Humphrey et al. PNAS(USA) 87:4207-4211(1990)에 개시된 EGFR의 결실 돌연변이)을 포함한다. erbB1은 EGFR 단백질 산물을 코딩하는 유전자를 가리킨다.

[0067] 본 원에 사용된 "키나제 활성을 증대시키는 핵산 변이"라는 용어는 키나제 활성을 증대시키는 결과를 낳는, 유전자의 누클레오티드 서열에서의 변이(즉, 돌연변이)를 가리킨다. 증대된 키나제 활성은 핵산에서의 변이의 직접적인 결과이고, 그 유전자가 코딩하는 단백질과 관련되어 있다.

[0068] 본 원에 사용된 "약", 또는 "화합물"이라는 용어는 질병이나 상태를 치료 또는 예방 또는 조절하기 위해 환자에게 투여하는 화학물질(chemical entity) 또는 생물학적 산물 또는 이들의 조합을 가리킨다. 화학물질 또는 생물학적 산물은 바람직하게는 저분자량 화합물이지만 반드시 저분자량 화합물이어야 하는 것은 아니고, 더 큰 화합물일 수도 있으며, 그 예에는 핵산 올리고머, 아미노산, 탄수화물, 단백질, 올레고누클레오티드, 리보자임, DNA효소, 당단백질, siRNA, 지질단백질, 엡테이머 및 그들의 변형체(modification)나 조합이 제한 없이 포함된다.

[0069] 본 발명의 의미맥락에서의 "유전형"이라는 용어는 유전자의 특정한 대립형질을 가리키며, 특정 부위에서의 핵산 서열에 존재하는 특정 누클레오티드에 의해서 정의될 수 있다.

[0070] "유전자의 변이형", "유전자형", 또는 "대립유전자"라는 용어는 집단 내에서의 어떤 유전자의 하나의 특정한 형태를 가리키는 용어이며, 그 특정한 형태는 유전자 서열 내부의, 적어도 하나의, 및 흔히는 둘 이상의 변이체 부위의 서열에 있어서 동일 유전자의 다른 형태들과 다른 것이다. 그 유전자의 서로 다른 대립유전자들간에서 서로 다른 이러한 변이체 부위들의 서열들은 "유전자 서열 변이", "변이", 또는 "변이체"라고 명명한다. 비록 돌연변이라는 용어가 유해한 표현형과 관련된 대립유전자를 지칭할 때 흔히 사용되기는 하지만, 당 기술분야에 알려진 동등한 다른 용어 중에는 돌연변이와 다형성이 포함된다. 본 발명의 바람직한 양상에서, 상기 변이들은 본 원 명세서에 포함된 변이표에 게재된 변이들로 구성된 군으로부터 선택된다.

[0071] 본 발명의 의미맥락에서 "프로브"라는 용어는 구조상 서로 다른 표적 분자들을 검출 가능할 정도로 구분할 수 있는 분자를 가리킨다. 검출은 사용되는 프로브의 타입에 따라, 및 표적 분자의 타입에 따라 다양한 서로 다른 방법으로 행하여질 수 있다. 따라서, 예를 들면, 검출은 표적 분자의 활성 수준의 식별에 기초할 수도 있으나, 바람직하게는, 특이적 결합의 검출에 기초한다. 그러한 특이적 결합의 예에는, 항체 결합과 핵산 프로브 혼성화가 포함된다. 따라서, 프로브로는 효소 기질, 항체와 항체 단편, 및 바람직하게는, 핵산 혼성화 프로브를 예로 들 수 있다.

[0072] 본 원에 사용된 "유효한"과 "유효성"이라는 용어는 약리학적인 유효성과 생리학적인 안전성 모두를 포함한다. 약리학적인 유효성은 환자에게 있어서 소망의 생물학적 효과를 거둘 수 있는 치료능력을 가리킨다. 생리학적인 안전성은 치료의 투여의 결과로 나타나는, 독성의 수준, 또는 세포, 조직 및/또는 개체 수준에서의 다른 생리학적인 역효과를 가리킨다(흔히 부작용이라고 불린다). "덜 유효한"의 의미는 그 치료가 치료요법상 약리학적인 유효성이 상당히 낮은 수준을 의미하거나 및/또는 치료요법상 생리학적인 역효과 수준이 더 큰 것을 의미한다.

[0073] 본 원에 사용된 "프라이머"라는 용어는 올리고누클레오티드를 가리키는데, 폴리누클레오티드에 대해 상보적인 프라이머 연장 산물(primer extension product)의 합성이 촉매되는 조건하에 놓여지면 프라이머는 상보 가닥을 따라 폴리누클레오티드 합성의 개시점으로 역할할 수 있다. 이러한 조건들은 4개의 서로 다른 누클레오티드 삼인산 또는 누클레오시드 유사체들과, DNA 중합효소 및/또는 역전사효소 같은 하나 이상의 중합제들이 적합한 온도 하에 적합한 완충액("완충액"이란 보조인자인 대체물 또는 pH나 이온 강도 등에 영향을 미치는 대체물을 포함한다) 중에 존재하는 것이다. 프라이머는 중합효소를 위한 시약의 존재하에 연장 산물의 합성을 시발하기 위해서는 충분히 길어야만 한다. 전형적인 프라이머는 표적 서열에 대해 실질적으로 상보적인 서열 길이를 적어도 약 5 누클레오티드 함유하나, 좀 더 긴 프라이머들이 바람직하다. 대개 프라이머들은 15~26 누클레오티드를 함유하나, 더 긴 프라이머들도 채택될 수 있다.

[0074] 프라이머는 항상 표적 서열, 즉 프라이머가 결합할 수 있는 증폭될 특정 서열에 대해 실질적으로 상보적인 서열을 함유한다. 프라이머는 또한 필요에 따라 프로모터 서열을 포함한다. "프로모터 서열"이라는 용어는 RNA 중

합효소에 의해 특이적으로 인식되는 단일 가닥의 핵산 서열을 뜻한다. 상기 RNA 중합효소는 인식된 서열에 결합하여 전사과정을 개시하는데, 이 전사과정을 통해 RNA 전사체가 만들어진다. 원칙적으로 어떤 프로모터 서열이든, 그 개시서열을 인식할 수 있는 중합효소가 알려져 있고 입수 가능하면 채용할 수 있다. 알려져 있는 유용한 프로모터들은 어떤 박테리오파지 중합효소, 예를 들면, 박테리오파지 T3, T7, 또는 SP6 등에 의해 인식되는 것들이다.

[0075] "마이크로어레이(microarray)"는 고체 지지체의 표면 위에 형성된 각각 한정된 면적을 갖는 불연속적(discrete) 영역들의 선형 또는 2차원 배열이다. 마이크로어레이 위의 불연속적인 영역들의 밀도는 단일 고체상 지지체 표면 위에서 검출되는 표적 폴리누클레오티드의 총 수에 의해 결정되며, 바람직하게는 적어도 약 50/cm^2, 보다 바람직하게는 적어도 약 100/cm^2, 더 바람직하게는 적어도 약 500/cm^2, 한층 더 바람직하게는 적어도 약 1,000/cm^2이다. 본 발명에서 사용되는 DNA 마이크로어레이는 칩, 또는 표적 폴리누클레오티드를 증폭하거나 복제할 때 사용하는 다른 표면 위에 놓여진 올리고누클레오티드 프라이머의 배열이다. 배열 내의 각 특정 프라이머 그룹의 위치가 알려져 있기 때문에, 표적 폴리누클레오티드의 정체는 그것이 마이크로어레이상의 특정 위치에 결합함에 기초하여 결정될 수 있다.

[0076] "표지"라는 용어는 분석 시료에서 표적 폴리누클레오티드의 존재를 가리키는 검출 가능한 신호를 생성할 수 있는 구성물을 가리킨다. 적합한 표지로는 방사성 동위원소, 누클레오티드 발색단, 효소, 기질, 형광 분자, 화학발광부분, 자기 입자(magnetic particle), 생물발광부분 등을 들 수 있다. 이와 같이, 표지는 분광법적인 수단, 광화학적 수단, 생화학적 수단, 면역화학적 수단, 전기적인 수단, 광학적인 수단 또는 화학적인 수단에 의해 검출 가능한 어떤 구성물이라도 좋다.

[0077] "지지체"라는 용어는 비드, 입자, 딥스틱(dipstick), 섬유, 필터, 막, 및 유리 슬라이드와 같은 통상적인 지지체 및 유리 슬라이드 같은 실란 또는 실리케이트 지지체를 가리킨다.

[0078] "증폭"이라는 용어는 광의로 사용되어 증폭 산물을 생성하는 것을 의미하며, 그러한 증폭 산물의 예로서는 추가의 표적 분자, 또는 표적 유사 분자, 또는 시료에 표적 분자가 존재함으로써 생성되는, 표적 분자에 대해 상보적인 분자 등을 들 수 있다. 표적이 핵산인 상황에서, 증폭 산물은 DNA 또는 RNA 중합효소, 또는 역전사효소 등에 의해 효소적으로 생성될 수 있다.

[0079] 본 원에 사용된 "생물학적 시료"란 한 개인으로부터 분리된 조직이나 체액의 시료를 가리키며, 그 예로서는 혈액, 혈장, 혈청, 종양 생검, 오줌, 변, 가래, 척수액, 흉수, 유두흡인액, 림프액, 피부의 외부 절개, 호흡기, 장, 생식비뇨기의 관, 눈물, 침, 젖, 세포(혈액 세포를 포함하지만 이에 한정하지 않음), 종양, 기관, 및 *in vitro* 세포배양액 성분 시료 등을 들 수 있으나, 이에 한정되는 것은 아니다. 바람직한 태양에서, 이러한 시료는 절제나, 기관지경 생검, 또는 원발암(primary tumor) 또는 전이된 암의 중심부바늘생검(core needle biopsy), 또는 흉수로부터의 세포 블럭으로부터 얻어진다. 부가적으로, 미세바늘흡인물 시료(fine needle aspirate sample)가 사용된다. 시료는 파라핀 포매(paraffin-embedded) 조직이어도 좋고 또는 냉동시킨 조직이어도 좋다.

[0080] "항체"라는 용어는 항원에 결합할 수 있는 면역 글로불린 단백질을 의미한다. 본 원에 사용된 항체는, 관심있는 항원 또는 항원성 단편에 결합할 수 있는, 예를 들면 F(ab')2, Fab', Fab와 같은 항체 단편을 포함하는 의미이다. 바람직하게는, 항원에 항체가 결합함으로써 EGFR의 변이형의 활성을 저해한다.

[0081] "인간화된 항체"라는 용어는, 본 원에서, 예를 들면 Fab, Fab', F(ab')2, 및 Fv와 같은 항체 단편들로만 구성되는 항체뿐만 아니라, 온전한 항체 분자, 즉, 2개의 온전한 경쇄과 2개의 온전한 중쇄로 구성된 항체를 기술하는데 사용되며, 그 CDR은 인간이 아닌 소스에서 유래하는 것이고, Ig 분자 또는 그 단편의 나머지 부분은 인간 항체, 바람직하게는 인간의 항체를 코딩하는 핵산 서열로부터 만들어진 인간 항체로부터 유래하는 것이다.

[0082] 본 원에 사용된 "인간 항체", 및 "인간화된 항체"라는 용어는 항체 분자의 모든 부분이 인간의 항체를 코딩하는 핵산 서열로부터 유래된 항체를 기술하기 위해 사용된다. 그러한 인간 항체는 항체 치료에 사용되는 것으로서 가장 바람직한 것인데, 왜냐하면 그러한 항체는 인간 환자에 있어서 면역 반응을 거의 유도하지 않거나 또는 전혀 유도하지 않기 때문이다.

[0083] 본 원에 사용된 "키메라 항체"라는 용어는 위의 "인간화된 항체" 용어의 정의에서 기술한 바와 같이 항체 단편뿐만 아니라 항체 분자를 기술하는데 사용된다. "키메라 항체"라는 용어는 인간화된 항체를 포함한다. 키메라 항체는 제1 포유동물 종으로부터 유래하는 중쇄 또는 경쇄의 아미노산 서열로 된 적어도 하나의 부분과, 제2의

다른 포유동물 종에서 유래하는 중쇄 또는 경쇄의 아미노산 서열로 된 또 다른 부분을 갖는다.

[0084] 바람직하게는, 가변부위는 비-인간 포유동물 종으로부터 유래하는 것이고, 불변부위는 인간 종으로부터 유래된 것이다. 특히, 키메라 항체는 바람직하게는, 항체의 가변부위를 코딩하는 비-인간 포유동물로부터의 9 누클레오티드 서열과 항체의 불변부위를 코딩하는 인간으로부터의 누클레오티드 서열로부터 만들어진다.

[0085] 표 2는 본 발명에 기재되어 있는 방법에 관련된 erbB1 유전자의 키나제 도메인에서의 DNA 서열 변이의 부분적인 리스트이다. 이러한 변이들은 제피티니브에 반응한 NSCLC 환자들로부터 얻은 생물학적 시료와 제피티니브에 노출되지 않은 환자들로부터 얻은 생물학적 시료에 대한 연구중에서 본 발명자들에 의해 확인된 것들이다.

[0086] 핵산 분자들은 당업자에게 주지된 많은 과정들 중 특정 생물학적 시료에 맞는 특정 분리 과정을 선택하여 특정의 생물학적 시료로부터 분리해 낼 수 있다. 예를 들면, 고형의 재료로부터 핵산 분자를 얻음에 있어서는 냉동-해동 및 알칼린 용해 과정이 유용할 수 있고, 오줌으로부터 핵산 분자를 얻는 방법으로서는 가열 및 알칼린 용해 과정이 유용할 수 있고, 혈액으로부터 핵산을 얻는 방법으로서는 단백질분해효소 K 추출법을 사용할 수 있다(Rolff, A et al. PCR: Clinical Diagnostics and Research, Springer(1994)).

[0087] **검출 방법**

[0088] 암에 걸렸거나 암이 발병할 위험이 있는 환자들에 있어서 erbB1 유전자의 키나제 도메인에서의 특정 변이나 복수의 변이들의 존재 또는 부재를 결정하는 것은 다양한 방식으로 수행가능하다. 그러한 시험들은 통상 생물학적 시료들, 즉, 조직 생검, 오줌, 변, 가래, 혈액, 세포, 긁어낸 조직, 유방 흡인액, 또는 다른 세포성 물질들로부터 수집된 DNA나 또는 RNA를 이용하여 수행되며, 또한 다양한 방법들을 통해 수행될 수 있다. 이러한 방법들에는, 다음에 국한하지는 않지만, PCR, 대립유전자-특이적 프로브와의 혼성화, 돌연변이의 효소적 검출, 미스매치에 대한 화학적 절단(cleavage), 질량분석 또는 DNA 서열결정이 있으며, 미량서열결정(minisequencing)도 포함된다. 특정의 태양에서, 대립유전자-특이적 프로브와의 혼성화는 2가지 포맷으로 수행가능하다:(1) 다수의 DNA 칩 응용들에서처럼, 고체상(유리, 실리콘, 나일론 막)에 결합된 대립유전자-특이적 올리고누클레오티드와 용액 중의 표지된 시료, 또는 (2) 결합된 시료(흔히 복제된 DNA 또는 PCR 증폭된 DNA)와 용액 중의 표지된 올리고누클레오티드(대립유전자에 특이적인 것 또는 혼성화에 의해 서열분석할 수 있도록 짧은 것). 진단 시험은 흔히 고체 지지체 위에 변이들의 구획(panel)을 포함하며, 이에 따라 하나 이상의 변이들을 동시에 결정할 수 있다.

[0089] 또 다른 양상에서, erbB1 유전자에 있어서 핵산 변이를 증가시키는 적어도 하나의 키나제 활성의 존재를 검출하는 것은 하플로타입 결정 시험을 수반하기도 한다. 하플로타입 결정 방법은 예로서 WO 00/04194에 개시된 바와 같이, 당업자에게 알려져 있다.

[0090] 바람직하게는, 키나제 활성을 증가시키는 적어도 하나의 핵산 변이의 존재 또는 부재를 결정하는 것은 중합효소 연쇄반응(PCR) 같은 방법에 의해 변이 부위나 또는 부위들의 서열을 결정하는 것을 포함한다. 또는, 키나제 활성을 증가시키는 핵산 변이의 존재 또는 부재를 결정하는 방법에는 사슬 종료 DNA 서열결정(chain terminating DNA sequencing) 또는 미량서열결정, 올리고누클레오티드 혼성화, 또는 질량분석이 포함될 수 있다.

[0091] 본 발명의 방법들은 암에 걸렸거나 암이 발병할 위험이 있는 환자에 있어서 EGFR 표적 치료의 유효 가능성(또는 유효성의 결핍)을 예측하는데 사용될 수 있다. 바람직하게는, 암은 여기에 한정하는 것은 아니지만, 위장관암, 전립선암, 난소암, 유방암, 머리와 목 암, 폐암, 비소세포 폐암, 신경계 암, 신장암, 망막암, 피부암, 간암, 이자암, 생식-비뇨기 암, 및 방광암 같은 상피세포 기원의 암을 포함한다. 바람직한 태양에서 상기 암은 비소세포 폐암이다.

[0092] 본 발명은 일반적으로 암에 걸렸거나 암이 발병할 위험이 있는 환자에 있어서 EGFR 표적 치료의 유효성을 나타내주는 erbB1 유전자의 키나제 도메인에서의 변이들을 동정하는 것과 관계가 있다. 부가적으로, EGFR의 키나제 도메인에서의 특정 변이를 동정하는 것은 진단적 또는 예후적 시험에 이용될 수 있다. 예로서, erbB1의 키나제 도메인에 적어도 하나의 변이가 존재하는 것은, 환자에게, 예를 들면, 티로신 키나제 저해제 같은 EGFR을 표적으로 하는 화합물에 의한 치료가 유익할 수 있을 것임을 나타낸다.

[0093] 진단 시험을 위한 방법은 당업계에 잘 알려져 있으며, 여기에 참조로 편입되는 특허출원 WO 00/04194에 개시되어 있다. 모범이 되는 방법에서, 상기 진단 시험은, erbB1 유전자 서열의 키나제 도메인에서의 하나 또는 그 이상의 알려진 변이상에 걸친 DNA 또는 RNA 세그먼트(일반적으로 RNA를 cDNA로 변환한 후)를 증폭하는 것을 포함한다. 그 후 이 증폭된 세그먼트 중의 핵산 변이들을 동정하기 위해, 이 증폭된 세그먼트를 서열결정하고,

및/또는 폴리아크릴아미드 겔 전기영동에 건다.

[0094] **PCR**

[0095] 하나의 태양에서, 본 발명은 PCR에 의해, 또는 연결 연쇄 반응(ligation chain reaction: LCR)(예로서, Landegran, et al., 1988. *Science* 241: 1077-1080; 및 Nakazawa, et al., 1994. *Proc. Natl. Acad. Sci.* USA 91: 360-364 참조)에 의해 시험용 생물학적 시료 중의 erbB1 유전자의 키나제 도메인에서의 변이체를 탐색하는 방법을 제공한다. LCR은 특히 EGFR 유전자의 점돌연변이를 검출하는데 유용할 수 있다(Abravaya, et al., 1995. *Nucl. Acids Res.* 23: 675-682 참조). 상기 방법은 표적 서열을 증폭하기 위한, 그 유전자의 하나 이상의 보존된 영역에 대응하는 축퇴성 프라이머를 설계하는 단계, 상기 프라이머를 사용하고 시험용 생물학적 시료로부터 얻어진 DNA 또는 cDNA를 주형으로 하여 증폭 반응을 시키는 단계 및 상기 PCR 산물의 분석 단계를 포함한다. 시험용 생물학적 시료의 PCR 산물의 대조군 시료와의 비교로부터 시험용 생물학적 시료의 변이가 드러난다. 그 변화는 시험용 생물학적 시료 중에 핵산 변이의 존재일 수도 있고 부재일 수도 있다.

[0096] 대안적 증폭방법에는 자기 지속성 서열 복제(Guatelli, et al., 1990. *Proc. Natl. Acad. Sci.* USA 87: 1874-1878 참조), 전사 증폭 시스템(Kwoh, et al., 1989. *Proc. Natl. Acad. Sci.* USA 86: 1173-1177 참조), Qb 레플리카제(Lizardi, et al., 1988. *BioTechnology* 6: 1197 참조), 또는 다른 어떤 핵산 증폭 방법이 포함된다. 그 후 당업자에게 잘 알려진 기술을 사용하여 증폭된 분자를 검출한다. 이들 검출 스킴은 핵산 분자들이 매우 적은 수로 존재할 때, 핵산 분자들을 검출하는데 특히 유용하다.

[0097] 본 발명에 따른 유용한 프라이머는 예를 들면 서열번호 493, 494, 509, 및 510을 가이드로 하여 erbB1 유전자의 키나제 도메인의 단백질의 아미노산 서열 또는 핵산 서열을 이용하여 설계한다. 프라이머는 유전자의 상동 영역에서 설계되는데, 여기서 적어도 2개의 상동 영역은 길이나 핵산 서열에 있어서 가변인 서열의 불일치(divergent) 영역에 의해 분리되어 있다.

[0098] 예를 들면, 동일하거나, 또는 고도로 상동인, 바람직하게는 적어도 80%~85%, 더 바람직하게는 적어도 90%~99% 상동인 적어도 약 6개, 바람직하게는 적어도 8~10개의 연속된 아미노산의 서열. 가장 바람직하게는, 상기 아미노산 서열은 100% 동일한 것이다. 정방향과 역방향 프라이머들은, 코돈 축퇴성의 유지 및 알려진 유전자 패밀리 구성원 중 주어진 위치에서 다양한 아미노산들이 나타남에 기초하여 설계된다. 여기서 가리키는 상동성의 정도는 표준 서열비교 소프트웨어인 protein-BLAST 같은 것을 디폴트 세팅으로 하고 이용하여 얻어진 아미노산 서열 분석에 기초한다(http://www.ncbi.nlm.nih.gov/BLAST/).

[0099] 아래의 표 3은 축퇴 코드들의 사용법과 그들의 표준 심볼들을 나타낸다.

[0100] <표 3>

[0101]

	T	C	A	G
T	TTT Phe (F)	TCT Ser (S)	TAT Tyr (Y)	TGT Cys (C)
	TTC "	TCC "	TAC	TGC
	TTA Leu (L)	TCA "	TAA Ter	TGA Ter
	TTG "	TCG "	TAG Ter	TGG Trp (W)
C	CTT Leu (L)	CCT Pro (P)	CAT His (H)	CGT Arg (R)
	CTC "	CCC "	CAC "	CGC "
	CTA "	CCA "	CAA Gln (Q)	CGA "
	CTG "	CCG "	CAG "	CGG "
A	ATT Ile (I)	ACT Thr (T)	AAT Asn (N)	AGT Ser (S)
	ATC "	ACC "	AAC "	AGC "
	ATA "	ACA "	AAA Lys (K)	AGA Arg (R)
	ATG Met (M)	ACG "	AAG "	AGG "
G	GTT Val (V)	GCT Ala (A)	GAT Asp (D)	GGT Gly (G)
	GTC "	GCC "	GAC "	GGC "
	GTA "	GCA "	GAA Glu (E)	GGA "
	GTG "	GCG "	GAG "	GGG "

[0102] 바람직하게는 L, R, 및 S 같은 6배 축퇴 코돈들은 실제에 있어서는 6배 이상의 축퇴를 일으키므로 회피된다. L의 경우에, TTR과 CTN은 YTN(8배 축퇴)으로 되어버리고, R의 경우 CGN과 AGR은 MGN(8배 축퇴)으로 되어버리고

마지막으로 S의 경우, TCN, 및 AGY는 WSN(16배 축퇴)으로 되어버릴 수 있다. 3가지 경우 모두에 있어서 이들 중 6개는 표적 서열에 매치될 것이다. 이러한 특이성의 상실을 피하기 위해, 이러한 영역들을 피하거나, 또는 예를 들면 하나의 집단에는 S에 대하여 TCN을 포함시키고, 다른 집단에는 S에 대하여 AGY를 포함시키는 것과 같이 각각 대안적(alternative) 축퇴 코돈을 갖는 2개의 집단으로 만드는 것이 바람직하다.

[0103] 프라이머들은 많은 입수할 수 있는 컴퓨터 프로그램들을 이용하여 설계될 수 있으며, Oligo Analyzer 3.0; Oligo Calculator; NetPrimer; Methprimer; Primer3; WebPrimer; PrimerFinder; Primer9; Oligo2002; Pride 또는 GenomePride; Oligos; 및 Codehop 등이 포함되지만 이것들에 한정되는 것은 아니다. 이러한 프로그램에 대한 상세한 정보는 예를 들면, www.molbiol.net 로부터 얻을 수 있다.

[0104] 프라이머들은 당업자에게 알려져 있는 표지들을 이용하여 표지될 수 있다. 그러한 표지로는 방사성 표지나 형광 표지, 염색 표지, 및 효소 표지가 포함되나 이에 한정되지 않는다.

[0105] 증폭결과물의 분석은 증폭결과물을 그들의 크기에 따라 분리할 수 있는 방법이면 어느 것이나 사용하여 수행될 수 있으며, 이러한 방법으로는 자동과 수동의 겔 전기영동, 질량분석 등이 있다.

[0106] 또는, 증폭결과물은 서열차이를 이용하여, 예를 들면 핵산 어레이에 혼성화하는 것뿐만 아니라 SSCP, DGGE, TGGE, 화학적 절단, 또는 제한 단편 다형성(restriction fragment polymorphism)을 이용하여 분리될 수 있다.

[0107] 핵산 분리, 증폭, 및 분석 방법은 당업자에게는 일상적인 것이며, 프로토콜예는, 예를 들면, Molecular Cloning: A Laboratory Manual (3-Volume Set) Ed. Joseph Sambrook, David W. Russel, 및 Joe Sambrook, Cold Spring Harbor Laboratory; 3판(2001. 1. 15), ISBN: 0879695773에서 찾아볼 수 있다. PCR 증폭에 사용되는 방법에 대한 특별히 유용한 프로토콜 소스는 PCR (Basics: From Background to Bench) by M.J. McPherson, S. G. Moeller, R. Beynon, C. Howe, Springer Verlag; 1판(2000. 10. 15), ISBN: 0387916008 이다.

[0108] 바람직하게는, 인간 EGFR의 엑손 19 및 21은 다음의 프라이머를 이용하여 중합효소연쇄반응(PCR)에 의해 증폭된다: 엑손 19 센스 프라이머, 5'-GCAATATCAGCCTTAGGTGCGGCTC-3' (서열번호 505); 엑손 19 안티센스 프라이머, 5'-CATAGAAAGTGAACATTTAGGATGTG-3' (서열번호 506); 엑손 21 센스 프라이머, 5'-CTAACGTTCGCCAGCCATAAGTCC-3' (서열번호 507); 및 엑손 21 안티센스 프라이머, 5'-GCTGCGAGCTCACCCAGAATGTCTGG-3' (서열번호 508).

[0109] 대안적 태양에서, 시료 세포로부터의 EGFR 유전자 내에 있는 돌연변이는 제한효소 절단 패턴 상의 변경에 의해 동정될 수 있다. 예를 들면, 시료와 대조군 DNA를 분리하고, 증폭하고(필요에 따라), 1 이상의 제한 엔도누클레아제로 소화(digest)하고, 단편들의 길이를 겔 전기영동에 의해 결정하고, 비교한다. 시료와 대조군 DNA 간의 단편 길이 차이는 시료 DNA에서의 돌연변이를 시사한다. 또한, 서열 특이적 리보자임을 사용하여(예로서, 미국 특허번호 5,493,531 참조), 리보자임 절단 부위의 발달 또는 상실에 의해 특정 돌연변이의 존재를 평가할 수 있다.

[0110] EGFR 유전자에서의 돌연변이를 검출하는 다른 방법은, 절단제로부터의 보호에 의해 RNA/RNA 또는 RNA/DNA의 이형 2본쇄(heteroduplex)에 있어서 미스매치된(mismatched) 염기들을 검출하는 방법을 포함한다(예로서, Myers, et al., 1985. Science 230: 1242 참조). 일반적으로, "미스매치 절단" 기술은 조직 시료로부터 얻은 잠재적 돌연변이 RNA 또는 DNA와 함께 야생형 EGFR 서열을 함유하는 RNA 또는 DNA를 혼성화함으로써 형성된(표지된) 이형 2본쇄를 마련하는 것에 의해 시작된다. 이중 가닥 2본쇄를, 대조군과 시료 가닥 간의 염기쌍 미스매치로 인해 존재하게 되는, 2본쇄 중의 단일 가닥으로 된 영역을 절단하는 물질로 처리한다. 예를 들면, RNA/DNA 2본쇄를 리보누클레아제로 처리할 수 있고, DNA/DNA 혼성물(hybrid)을 S1 누클레아제로 처리하여 그 미스매치된 영역을 효소적으로 소화(digest)할 수 있다. 또 다른 태양에서는, DNA/DNA 또는 RNA/DNA 2본쇄를 하이드록실아민 또는 오스뮴 테트록사이드, 및 피페리딘으로 처리하여 그 미스매치된 영역을 소화(digest)할 수 있다. 미스매치된 영역의 소화 후에 얻어진 물질들은 변성 폴리아크릴아미드 겔상에서 크기에 따라 분리하여 돌연변이 부위를 결정한다(예로서, Cotton, et al., 1988. *Proc. Natl. Acad. Sci.* USA 85: 4397; Saleeba, et al., 1992. *Methods Enzymol.* 217: 286-295 참조). 어떤 태양에서는, 검출을 위해 대조군 DNA 또는 RNA를 표지할 수 있다.

[0111] 또 다른 태양에서는, 미스매치 절단 반응에, 세포 시료로부터 얻어진 EGFR cDNA에서의 점돌연변이를 검출 및 지도작성(mapping)하기 위해 규정된 시스템에서 이중 가닥 DNA 내의 미스매치된 염기쌍을 인식하는 하나 이상의 단백질(소위 "DNA 미스매치 수복" 효소)이 채용된다. 예를 들면 E.coli mutY 효소는 G/A 미스매치에서 A를 절단하며 HeLa 세포로부터의 DNA 글리코실라제는 G/T 미스매치에서 T를 절단한다(예로서, Hsu, et al., 1994.

Calcinogenesis 15: 1657-1662 참조). 모범적인 태양에 따르면, 예를 들면 DEL-1부터 DEL-5, G719S, G857V, L883S 또는 L858R EGFR 서열과 같은 돌연변이 EGFR 서열에 기초한 프로브를 시험 세포(들)로부터의 cDNA 또는 다른 DNA 산물에 혼성화시킨다. 그 2본쇄를 DNA 미스매치 수복효소로 처리하고 만약 그 절단 산물이 있다면 전기영동 등에 의해 검출할 수 있다(예로서 미국 특허번호 5,459,039 참조).

[0112] 또 다른 태양에서는, EGFR 유전자에서의 돌연변이를 확인하는데 있어서 전기영동 이동성(electrophoretic mobility)의 변경을 이용할 것이다. 예를 들면, 단일 가닥 입체구조 다형성(SSCP)을 이용하여 돌연변이와 야생형 핵산 간의 전기영동 이동성 차이를 검출할 수 있다(Orita, et al., 1989. *Proc. Natl. Acad. Sci.* USA: 86: 2766; Cotton, 1993. *Mutat. Res.* 285: 125-144; Hayashi, 1992. *Genet. Anal. Tech. Appl.* 9: 73-79 참조). 시료와 대조군 EGFR 핵산의 단일 가닥 DNA 단편들은 변성되고(denatured) 재결합(renature)될 것이다. 단일-가닥 핵산의 2차 구조는 서열에 따라 변하는데, 그에 따른 전기영동 이동성의 변경에 의해 심지어 단일 염기 변화까지도 검출 가능하게 된다. DNA 단편들은 표지되거나 표지된 프로브로 검출될 수 있다. 분석의 감도는 그 2차 구조가 서열 변화에 보다 민감한 RNA(DNA를 사용하는 것보다 민감)를 사용함으로써 향상될 수 있다. 하나의 태양에서, 본 방법은 이형 2본쇄 분석을 이용하여, 이중가닥 이형 2본쇄 분자를, 그 전기영동 이동성의 변화에 기초하여 분리한다(Keen, et al., 1991. *Trends Genet.* 7: 5 참조).

[0113] 또 다른 태양에서는, 변성제(denaturant)의 농도 구배를 갖는 폴리아크릴아미드 겔 내의 돌연변이 또는 야생형의 단편들의 이동을, 변성 구배 겔 전기영동법(denaturing gradient gel electrophoresis: DGGE)을 사용하여 분석한다(Myers, et al., 1985. *Nature* 313: 495 참조). 분석의 방법으로서 DGGE가 사용될 때, DNA는 완전히 변성되지 않도록 하기 위해 변형(modify)될 것인데, 이는, 예를 들면, PCR에 의해, 대략 40bp의 고융점 GC-풍부 DNA인 GC 클램프를 부가하는 것이다. 또 다른 태양에서는, 대조군과 시료 DNA의 이동성 차이를 확인하기 위해 변성 구배 대신에 온도 구배가 이용된다(Rosenbaum and Reissner, 1987. *Biophys. Chem.* 265: 12753).

[0114] 점돌연변이를 검출하는 다른 기술들의 예는, 선택적 올리고누클레오티드 혼성화, 선택적 증폭, 또는 선택적 프라이머 연장을 포함하나, 이에 한정되는 것은 아니다. 예를 들면, 올리고누클레오티드 프라이머를 그 중심부에 알려진 돌연변이가 위치하도록 제조한 후, 오직 완벽한 매치일 때만 혼성화를 허용하는 조건하에서, 표적 DNA와 혼성화시킬 수 있다(Saiki, et al., 1986. *Nature* 324: 163; Saiki, et al., 1989. *Proc. Natl. Acad. Sci.* USA 86: 6230 참조). 그러한 대립유전자 특이적 올리고누클레오티드들은 그 올리고누클레오티드들이 혼성화 막에 부착하여 표지된 표적 DNA와 혼성화 될 때, PCR 증폭된 표적 DNA나 또는 다수의 서로 다른 돌연변이들과 혼성화 된다.

[0115] 또는, 선택적 PCR 증폭에 의존하는, 대립유전자 특이적 증폭 기술을 본 발명과 함께 사용할 수 있다. 특이적 증폭을 위한 프라이머로서 사용되는 올리고누클레오티드는 관심 대상의 돌연변이를 그 분자의 중심부에 가질 수 있으며(증폭이 분별 혼성화(differential hybridization)에 의존하도록. 예로서, Gibbs, et al., 1989. *Nucl. Acids Res.* 17: 2437-2448 참조), 또는 프라이머의 극단의 3'-말단(extreme 3'-terminus)에 가질 수도 있는데, 이렇게 할 경우 적합한 조건하에서라면, 미스매치, 또는 중합효소 연장(polymerase extension)이 방지 또는 감소된다(예로서, Prossner, 1993. Tibtech. 11: 238 참조). 부가적으로, 절단에 기초한 검출이 가능하도록 돌연변이 영역 내에 신규한 제한 부위를 도입하는 것이 바람직할 수도 있다(Gasparini, et al., 1992. *Mol. Cell Probes* 6: 1 참조). 어떤 태양에서는, 증폭용 Taq 연결효소(ligase)를 사용하여 증폭을 행할 수도 있을 것이다(Barany, 1991. *Proc. Natl. Acd. Sci.* USA 88: 189 참조). 그러한 경우에는, 5' 서열의 3'-말단에서 완벽한 매치가 있을 때만 연결(ligation)이 될 것이므로, 이로써 증폭의 존재 또는 부재를 찾음으로써, 알려진 돌연변이가 특정 부위에 존재하는 것을 검출할 수 있게 될 것이다.

[0116] **고형의 지지체와 프로브**

[0117] 대안적 태양에서는, 적어도 하나의 핵산 변이의 존재 또는 부재를 검출하는 것은, 상술한 바와 같이 하여 확인된, erbB1 유전자의 소망 영역에 대응하는 핵산 서열을 프로브와 접촉시키는 것을 포함한다. 상기 프로브는, 예를 들면, 분별 결합 또는 혼성화에 의해, 특정 형태의 유전자, 또는 특정 변이 또는 변이들의 존재를 구별할 수 있다. 따라서, 본보기로 되는 프로브로는 핵산 혼성화 프로브, 펩티드 핵산 프로브, 적어도 하나의 누클레오티드 유사체를 또한 함유하는 누클레오티드-함유 프로브, 예를 들면 단일클론 항체와 같은 항체, 및 여기서 논의한 다른 프로브들을 들 수 있다. 당업자들은 특정 특이성을 가진 프로브를 준비하는데 익숙하다. 당업자들은 유전자의 2개의 변이형들 간의 구분을 최적화하기 위해 다양한 변수를 조정할 수 있음을 인식할 것이며, 이러한 변수의 조정에는 염 농도, 온도, pH의 변화, 및 테트라메틸 암모늄 클로라이드 같은, GC *vs.* AT 염기쌍의 분별 친화도(differential affinity)에 영향을 주는 다양한 화합물들의 첨가 등이 포함된다(F. M. Ausubel,

R. Brent, R. E. Kingston, D. D. Moore, J.G. Seidman, K. Struhl 및 V. B. Chanda(편집자), John Wiley & Sons에 의한 「Current Protocols in Molecular Biology(분자생물학에서의 최근 프로토콜)」 참조).

[0118] 따라서, 바람직한 태양에서는, 적어도 하나의 변이의 존재 또는 부재의 검출은, 적어도 하나의 변이부위를 포함하는 핵산 서열을, 프로브와, 바람직하게는 핵산 프로브와 접촉시키는 것을 포함하며, 여기서 상기 프로브는, 혼성화가 선택적 혼성화 조건하에서 수행될 때, 변이 부위에 비-상보적 염기를 갖는 핵산 서열 형태에 혼성화되는 것과 비교해서, 변이 부위에 상보적인 염기를 함유하는 핵산 서열 형태에 우선적으로 혼성화한다. 그러한 핵산 혼성화 프로브는 2개 또는 그 이상의 변이 부위에 걸쳐 있다. 달리 구체적으로 적시하지 않는다면, 핵산 프로브는 염기 쌍 짝짓기 기능이 유지되는 한, 하나 이상의 핵산 유사체, 표지 또는 다른 대체물 또는 부분들을 포함할 수 있다.

[0119] 상기 프로브는, 예를 들면, 서열번호 495, 497, 또는 499의 결실된 영역의 양 쪽에, 적어도 3개의 연속적인 누클레오티드에 결합하도록 설계될 수 있다. 그러한 프로브는, 적합한 조건하에서 혼성화 될 때, EGFR의 변이형에는 결합할 것이나, 야생형 EGFR에는 결합하지 않을 것이다.

[0120] 그러한 혼성화 프로브는 당업계에 잘 알려져 있다(예로서, Sambrook et al., Eds.,(최신판), Molecular Cloning: A Laboratory Manual, (3판, 2001), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. 참조). 엄격한 혼성화 조건들은 전형적으로 약 1 M 미만의 농도로 염을 포함할 것이며, 더 보편적으로는 약 500 mM 미만, 바람직하게는 약 200 mM 미만의 농도로 포함한다. 혼성화 온도는 5℃ 까지 낮아질 수 있으나, 전형적으로는 22℃ 보다 높고, 더 전형적으로는 약 30℃보다 높고, 바람직하게는 약 37℃를 넘는다. 보다 긴 단편은 특이적 혼성화를 위해 보다 높은 혼성화 온도를 요구한다. 다른 요소들도 혼성화의 엄격성에 영향을 주기도 하는데, 여기에는, 염의 구성과 상보적 가닥의 길이, 유기 용매의 존재, 및 염기 미스매칭의 정도가 포함되며, 사용되는 변수들의 조합은 어느 변수 하나의 절대 척도(absolute measure)보다 더 중요하다. 조절될 수 있는 다른 혼성화 조건들로는 완충액 타입 및 농도, 용액 pH, 배경 결합(background binding)을 감소시키기 위한 차단제(blocking reagent)(예로서, 반복서열, Cot1 DNA, 차단 단백질 용액)의 존재와 농도, 세제(detergent)의 타입과 농도, 폴리누클레오티드의 상대농도를 증가시키는 폴리머 같은 분자, 금속 이온과 그 농도, 킬레이터와 그 농도, 및 당 기술분야에 알려져 있거나 발견할 수 있는 다른 조건들을 들 수 있다. 주어진 프로브에 대해 완벽하게 상보적인 서열에 대한 최적의 융해 온도(melting temperature)를 예측하기 위해 공식을 사용할 수도 있으나, 혼성화 조건의 어떤 세트하에서 프로브에 대한 실제 융해 온도는 경험적으로 결정되어야만 한다. 또한, 주어진 조건 세트 하에서의 프로브의 정확한 융해 온도를 결정하기 위해, Sambrook et al., "Molecular Cloning," 3판, Cold Spring Harbor Laboratory Press, 2001에 기재된 바와 같이, 그 프로브의 정확한 상보성을 시험할 수도 있다. 혼성화 온도는, 표적 폴리누클레오티드와 결합된 지지체를 사용하는 주어진 혼성화 용액에 대해, 소망하는 엄격성의 수준에서 검출 가능한 프로브의 결합이 검출되는 온도 범위가 확인될 때까지, 체계적으로 변경할 수 있다. 상기 엄격성은, 오직 고도의 상보성을 갖는 표적 폴리누클레오티드만이 혼성화하는 높은 엄격성, 또는 프로브에 상보적인 영역을 갖는 추가적인 표적 폴리누클레오티드가, 비상보적 표적 폴리누클레오티드와 상기 지지체에, 비특이적 결합으로부터 비롯되는 배경 수준 이상으로 검출 가능하게 혼성화하는 낮은 엄격성을 말한다. 지지체 상의 잠재적 표적 폴리누클레오티드에 의해 주어진 조건 세트하에서 혼성화를 수행한 후, 그 지지체를, 엄격성을 점진적으로 상향하는 조건하에(전형적으로는 염 농도를 낮추거나 및/또는 온도를 증가시키지만 다른 조건을 변경할 수도 있음), 뚜렷한 포지티브 신호가 보일만큼 배경 결합이 낮아질 때까지, 세정한다. 이것은 그 진행상황을 모니터할 수 있는데, 여기에는 프로브가 방사선사진법적으로 방사성 동위원소로 표지되는 가이거 계수기를 이용하는 방법과, 형광 촬상 장치(fluorescent imager)를 이용하는 방법, 및 프로브 결합을 검출할 수 있는 다른 수단들을 이용하는 방법들이 있다. 상기 지지체는 그러한 과정 중 말라서는 안되며, 마르면 상기 프로브는 심지어 배경 지점들에까지 비가역적으로 결합되어버릴 수도 있다. 프로브가 바람직하지 않은 배경 또는 가짜 포지티브 신호를 내보내면, 포지티브 신호가 배경과 구분될 때까지, 차단제를 채용하거나 프로브의 다른 영역을 사용하거나 또는 다른 프로브를 사용한다. 배경을 뛰어넘는 만족할 만한 신호를 제공하는 조건을 일단 찾으면, 포지티브 신호를 제공하는 당해 표적 누클레오티드를 분리하여 더욱 특징을 규명한다. 상기 분리된 폴리누클레오티드를 서열결정할 수 있으며, 그 서열을 데이터뱅크 엔트리나 또는 알려진 서열과 비교할 수 있고, 필요하다면, 당업계에 알려진 기술에 의해 전장 클론을 얻을 수 있으며, 그 폴리누클레오티드를 적합한 벡터 및 숙주를 사용하여 발현하여, 동정된 폴리누클레오티드가, 그 프로브 폴리누클레오티드의 기원에 유사한 활성을 갖는 단백질을 코딩하고 있는지 여부를 결정할 수 있다. 상기 프로브는 10~50 누클레오티드에서부터 시작하나, 예를 들면 약 50~500 누클레오티드 또는 그 이상의 보다 큰 프로브도 채택가능하다.

[0121] **고체상 지지체**

[0122] 본 발명의 고체상 지지체는 핵산 혼성화와 합성을 지지하는데 적합한 고형의 물질과 구조면 어느 것이나 가능하다. 바람직하게는, 상기 고체상 지지체는 적어도 하나의 실질적으로 단단한 표면을 포함하는데, 그 표면 위에 올리고누클레오티드 또는 올리고누클레오티드 프라이머들이 움직일 수 없게 고정될 수 있다. 상기 고체상 지지체는, 예를 들면, 유리, 합성 폴리머, 플라스틱, 메쉬(mesh) 구조가 아닌 단단한 나일론 또는 세라믹 등으로 만들어질 수 있다. 다른 적합한 고체 지지체 물질들은 당업자에게 알려져 있고, 곧 바로 입수가능하다. 상기 고형 지지체의 크기는 DNA 마이크로어레이 기술에 유용한 표준형의 마이크로어레이 크기면 어느 크기라도 좋으며, 그 크기는 본 발명의 반응을 수행하는데 사용되는 특정 기계에 맞도록 맞출 수도 있다. 올리고누클레오티드를 고정할 목적에 맞는 고체상 지지체를 얻기 위한 방법과 물질은 당업자에게 알려져 있고, 예를 들면, 미국 특허번호 5,919,523에 기재되어 있는데, 여기에 그 개시 내용을 참고로 편입한다.

[0123] 상기 고체 지지체는 용기를 갖는 유체의 내부에 마련되거나 그 부분일 수도 있다. 예를 들면, 고체 지지체는 챔버 내에 놓여질 수 있는데, 중합효소연쇄반응(PCR)이 상기 지지체 위에서 행하여질 수 있도록, 상기 챔버의 측면이 고체 지지체의 가장자리를 밀봉한다. 하나의 구체적인 예에서, 상기 챔버는 직각형의 지지체 양변 위에 벽을 가질 수 있는데, 이는 PCR 혼합물이 확실하게 지지체 위에 머물도록 하며, 또한 전체 표면이 프라이머들을 제공하는데 유용하도록 만든다.

[0124] 본 발명의 올리고누클레오티드 또는 올리고누클레오티드 프라이머들은 올리고누클레오티드를 고체 지지체 위의 특정 위치에 고착(fix), 고정(immobilize), 마련(provide), 및/또는 도포(apply)할 수 있는 사용 가능한 어떤 수단에 의해 고체 지지체의 표면에 고착되고, 고정되고, 마련되고, 및/또는 도포된다. 예를 들면, 포토리소그래피 기술(Affymetrix, Santa Clara, Calif.)을 사용하여, 여기에 참조형태로 편입되어 있는 미국 특허번호 5,919,523, 5,837,832, 5,831,070, 및 5,770,722에 기재되어 있는 바와 같이, 올리고누클레오티드 프라이머를 칩이나 고체 지지체 위의 특정 위치에 도포할 수 있다. 상기 올리고누클레오티드 프라이머는 고형의 지지체에 Brown과 Shalon에 의한 미국 특허번호 5,807,522(1998)에 기재된 바와 같이 도포될 수 있다. 부가적으로, 상기 프라이머는 Genetic MicroSystems사제(Woburn, Mass), GeneMachines사제(San Carlos, Calif.) 또는 Cartesian Technologies사제(Irvine, Calif.)의 로봇 시스템을 이용하여 고체 지지체에 도포될 수 있다.

[0125] 본 발명의 하나의 측면에서는, 생물학적인 시료로부터의 표적 폴리누클레오티드의 고체상 증폭이 수행되는데, 여기서는 복수의 올리고누클레오티드 프라이머 그룹들이 고체상 지지체 위에 고정된다. 바람직한 태양에서, 한 그룹내의 프라이머들은 적어도 프라이머들의 제1 세트를 포함하는데, 이 제1 세트의 프라이머들은 그 서열이 동일하며, 표적 폴리누클레오티드의 규정된 서열에 상보적이고, 또 적합한 조건하에서 표적 폴리누클레오티드에 혼성화할 수 있으며, 핵산 합성(즉, 체인 연장 또는 신장)에 대한 개시 프라이머로서 적합한 것이다. 기준 서열의 특정 영역을 커버하는 선별된 프라이머들은 그룹으로서 고체 지지체 위에 불연속적 위치에 고정된다. 바람직하게는, 그룹간의 거리는 증폭결과물을 검출하는데 사용되는 검출수단의 해상도 보다 크다. 바람직한 태양에서, 프라이머들은 고정되어 마이크로어레이 또는 칩의 형태를 이루고 있으며, 이들은 자동화된 프로세싱을 통해 처리되어 분석될 수 있다. 고정된 프라이머들은 핵산 증폭 수단에 적합한 조건 하에서, 표적 폴리누클레오티드의 고체상 증폭에 사용된다. 이러한 방식으로, erbB1 유전자의 키나제 도메인에서의 다양한 잠재적 변이들의 존재 또는 부재를 일회의 분석을 통해 결정할 수 있다.

[0126] 어떤 생물학적 소스가 erbB1 유전자의 키나제 도메인 위에 적어도 하나의 키나제 활성을 증가시키는 변이를 갖는지를 결정함에 있어서, 건강한 개체로부터 분리된 표적 폴리누클레오티드의 집단을 대조군으로 사용할 수 있다. 또는, 동일 개체의 건강한 조직으로부터 분리된 표적 폴리누클레오티드를 마찬가지로 대조군으로 사용할 수도 있다.

[0127] 마이크로어레이 상의 in situ-타입 PCR 반응은 Embretson et al., *Nature* 362: 359-362 (1993); Gosden et al., *BioTechniques* 15(1): 78-80 (1993); Heniford et al., Nuc. Acid Res. 21(14): 3159-3166 (1993); Long et al., *Histochemistry* 99: 151-162 (1993); Nuovo et al., *PCR Methods and Applications* 2(4): 305-312 (1993); Patterson et al., *Science* 260: 976-979 (1993)에 기재된 바와 같이 하여 수행할 수 있다.

[0128] 또는, erbB1의 키나제 도메인에서의 변이들은 지지체 위에서 PCR을 수행하지 않고도 고체상 기술에 의해 결정될 수 있다. 각자 erbB1의 키나제 도메인에 특징적인 변이를 이중, 삼중, 또는 사중으로 갖고 있는 복수의 올리고누클레오티드 프로브들을 고체상의 지지체에 결합할 수 있다. 당업자에게 알려져 있으며 상술한 바와 같은 선택적 혼성화 기술에 의해 상기 시험용 생물학적 시료에서의 변이의 존재 또는 부재를 검출할 수도 있다.

[0129] **질량분석**

[0130] 또 다른 태양에서, erbB1 유전자의 키나제 도메인에서의 키나제 활성을 증가시키는 핵산 변이의 존재 또는 부재는 질량분석법을 사용하여 결정할 수 있다. 질량분석을 수행할 핵산 분자의 적합한 양을 얻기 위해, 증폭은 필수적일 것이다. 본 발명에 사용하기에 적합한 증폭 과정의 예로는 클로닝(Sambrook et al., Molecular Cloning: A Laboratory Manual, 3판, Cold Spring Harbor Laboratory Press, 2001), 중합효소연쇄반응(PCR)(C.R. Newton 및 A. Graham, PCR, BIOS Publishers, 1994), 연결효소연쇄반응(LCR)(Wiedmann, M., et al., (1994) *PCR Methods Appl.* Vol. 3, pp. 57-64; F. Barnay *Proc. Natl. Acad. Sci.* USA 88, 189-93 (1991), 가닥치환증폭(strand displacement amplification: SDA)(G. Terrance Walker et al., *Nucleic Acids Res.* 22, 2670-77 (1994)) 및 RT-PCR(Higuchi, et al., *Bio/Technology* 11: 1026-1030 (1993)), 대립유전자-특이적 증폭(ASA), 및 전사 기반 처리법(transcription based process) 과 같은 변형들을 들 수 있다.

[0131] 질량분석을 용이하게 하기 위해, 검출될 핵산 서열을 함유하고 있는 핵산 분자는 고체 지지체에 고정될 수 있다. 적합한 고체 지지체의 예로서는 비드(예로서, 실리카 겔, 조절된 다공 유리(controlled pore glass), 자석, 세파덱스/세파로스(Sephadex/Sepharose), 셀룰로스), 편평한 표면 또는 칩(예로서, 유리섬유 필터, 유리 표면, 금속 표면(철, 금, 은, 알루미늄, 동, 및 실리콘), 모세관, 플라스틱(예로서, 폴리에틸렌, 폴리프로필렌, 폴리아미드, 폴리비닐리덴디플루오라이드 막 또는 마이크로타이터 판)); 또는, 비드 또는 편평한 표면을 구성하는 것과 유사한 물질로 만들어진 핀이나 빗(comb), 또는 웨이퍼(예로서, 실리콘 웨이퍼)와 같은 편평한 표면에 있는 구멍(pit)에 들어있는 비드를 들 수 있다.

[0132] 고정은, 예를 들면, 혼성화에 기초하여 달성할 수가 있으며, 그 혼성화는 지지체에 이미 고정된 캡쳐 핵산 서열과, 역시 검출될 핵산 서열을 함유하는 핵산 분자 내에 함유되어 있는 상보적 핵산 서열과의 사이에서 일어난다. 상기 상보적 핵산 분자들 간의 혼성화가 지지체에 의해 방해받지 않도록, 캡쳐 핵산은 고체 지지체와 캡쳐 핵산 서열 사이에 적어도 약 5 누클레오티드의 길이를 갖는 스페이서 영역(spacer region)을 포함할 수 있다. 그 형성된 2본쇄(duplex)는 레이저 펄스의 영향 하에 절단될 것이며, 탈착이 개시될 수 있다. 그 고체 지지체에 결합된 염기 서열은 자연(natural) 올리고리보누클레오티드 또는 올리고디옥시리보누클레오티드뿐만 아니라 그 유사체(예, 티오-변형 포스포디에스테르나 또는 포스포트리에스테르 골격)를 통해 제공될 수 있고, 또는, 상기 염기 서열을 효소적인 분해에 덜 민감하게 해서, 고체 지지체에 결합되어 있는 캡쳐 염기 서열의 전체적인 안정성을 증가시키는, PNA 유사체(예로서, Nielsen et al., Science, 254, 1497 (1991) 참조)와 같은 올리고누클레오티드 모방체들을 사용함으로써 제공될 수 있다.

[0133] 질량분석법 분석 이전에, 예를 들면, 증발(volatilization)에 요구되는 레이저 에너지를 낮추기 위해, 및/또는 분절화(fragmentation)를 최소화하기 위해, 핵산 분자를 "콘디쇼닝(conditioning)" 하는 것이 유용할 수 있다. 콘디쇼닝은 바람직하게는 표적 검출 부위가 고정되어 있는 동안에 수행한다. 콘디쇼닝의 한 예는, 핵산 분자의 포스포디에스테르 골격의 변형(modification)(예로서, 양이온 교환)인데, 이것은 누클레오티드 단위마다 결합된 양이온들의 이질성(heterogeneity)에 의한 피크 넓어짐(peak broadening)을 제거하는데 유용할 수 있다. 핵산 분자를 알킬아이오다이드, 아이오도아세트아미드, β-아이오도에탄올, 2,3-에폭시-1-프로판올과 같은 알킬화 물질과 접촉시킴으로써, 핵산 분자의 모노티오 포스포디에스테르 결합을 포스포트리에스테르 결합으로 변환시킬 수 있다. 마찬가지로, 포스포디에스테르 결합을, 트리알킬실릴 클로라이드를 사용하여, 전하를 갖지 않는 유도체로 변환시킬 수도 있다. 콘디쇼닝을 더 하는 것은, N7-또는 N9-디아자퓨린 누클레오티드 또는 RNA 빌딩 블럭과 같은 퓨린 제거(질량분석 도중의 분절화)에 대한 감도를 감소시키는 누클레오티드의 편입(incorporation), 또는 올리고누클레오티드 트리에스테르의 사용, 또는 알킬화된 포스포로티오에이트기의 편입, 또는 PNA와 같은 올리고누클레오티드 모방체의 사용을 포함한다.

[0134] 어떤 응용에 있어서는, 특정의 포획된(captured) 핵산 단편 위의 2개 이상의 (돌연변이된) 위치(loci)(어레이의 2개 이상의 지점(spot))를 동시에 검출하는 것이 유용할 수 있으며, 또는 다양한 고체 지지체 위의 올리고누클레오티드 또는 올리고누클레오티드 모방체 어레이를 사용하여, 병렬 처리(parallel processing)를 수행하는 것이 유용할 수도 있다. "다중처리(multiplexing)"는 몇 가지 서로 다른 방법들에 의해 달성될 수 있다. 예를 들면, 대응하는 검출자(detector)(프로브) 분자(예로서, 올리고누클레오티드 또는 올리고누클레오티드 모방체)를 사용함으로써, 몇 가지 돌연변이를 하나의 표적 서열 상에서 동시에 검출할 수 있다. 그러나, 검출자 올리고누클레오티드 D1, D2, 및 D3 간의 분자량 차이가 동시 검출(multiplexing)이 가능하도록 충분히 커야 한다. 이것은 서열 자체(조성 또는 길이) 또는 검출자 올리고누클레오티드에 질량-변경 작용기 M1-M3 을 도입함으로써 달성될 수 있다.

[0135] 본 발명에서의 사용을 위한 바람직한 질량분석의 형식은 매트릭스 지원 레이저 탈착 이온화(MALDI), 전기 분사(electrospray: ES), 이온 사이클로트론 공명(ICR), 및 푸리에 변환이다. 질량분석을 수행하는 방법들은 당업자에게 알려져 있으며, *Methods of Enzymology*, Vol. 193: "Mass Spectrometry"(J.A. McCloskey, 편집자), 1990, Academic Press, New York에 더 기술되어 있다.

[0136] **서열분석**

[0137] 다른 바람직한 태양에서, 적어도 하나의, 키나제 활성을 증가시키는 핵산 변이의 존재 또는 부재를 결정하는 것은 적어도 하나의 핵산서열을 서열분석하는 것을 포함한다. 상기 서열분석은, 적어도 하나의 변이 부위를 포함하고 또한 복수의 변이 부위를 포함할 수도 있는 erbB1의 키나제 도메인의 한 부분 또는 여러 부분을 서열 분석하는 것을 포함한다. 바람직하게는 상기 부분의 길이는 500 누클레오티드 이하이고, 보다 바람직하게는 100 누클레오티드 이하이며, 가장 바람직하게는 45 누클레오티드 이하이다. 그러한 서열분석은 당업자들에게 알려져 있는 다양한 방법들에 의해 수행될 수 있는데, 그러한 방법에는 디데옥시 종결법(예로서, 염료로 표지된 디데옥시 누클레오티드를 사용), 미세서열분석법, 및 질량분석법이 포함된다.

[0138] **면역검출법**

[0139] 하나의 태양에서, 적어도 하나의 키나제의 활성을 증가시키는 핵산 변이의 존재 또는 부재를 결정하는 것은 EGFR의 다운스트림 표적의 활성화 상태를 결정하는 것을 포함한다.

[0140] 본원의 발명자들은 EGFR의 주요한 다운스트림 표적의 인산화 상태를 비교했었다. 예를 들면, Ras를 통한 Erk1과 Erk2의 EGF-유도 활성화, PLCγ/PI3K를 통한 Akt의 EGF-유도 활성화, 및 JAK2를 통한 STAT3와 STAT5의 EGF-유도 활성화가 검사되었다. Ras를 통한 Erk1과 Erk2, PLCγ/PI3K를 통한 Akt, JAK2를 통한 STAT3와 STAT5는 EGFR의 발암성 효과를 매개하는 필수적인 다운스트림 경로이다(R.N. Jorissen et al., *Exp. Cell Res.* **284**, 31(2003)).

[0141] 본원 발명자들은, EGF로 유도된 Erk의 활성화는, 야생형 EGFR 또는 2개의 활성화하는 EGFR 돌연변이체 중 어느 것을 발현하는 세포들 사이에서 분별불가능하다는 것을 입증하였다.

[0142] 대조적으로, Akt와 STAT5 양자의 인산화는, 돌연변이 EGFR들 중의 어느 것을 발현하는 세포에 있어서 실질적으로 증대되었다. STAT3의 증대된 인산화는 돌연변이 EGFR을 발현하는 세포들에서 유사하게 관찰되었다. 따라서, EGFR 돌연변이체 내의 C-말단 티로신 잔기의 선택적 EGF-유도 자가인산화반응은 다운스트림 신호전달경로의 선택적 활성화와 매우 상관관계가 높다.

[0143] 본원의 하나의 태양에서는 EGFR 돌연변이체의 존재를 당 업계에 잘 알려진 면역학적인 기술들을 사용하여 결정할 수 있으며, 그러한 기술에는 예를 들면, 면역조직화학법, 면역세포화학법, FACS 스캐닝, 면역블로팅법, 방사면역측정법, 웨스턴 블로팅법, 면역침전법, 효소결합 면역흡착법(ELISA)과 같은 항체 기술 및 EGFR의 활성화된 다운스트림 표적들에 대한 항체를 사용하는 파생 기술들이 있다. 그러한 표적의 예로서는, 인산화된 STAT3, 인산화된 STAT5, 및 인산화된 Akt을 들 수 있다. 인산기-특이적 항체를 사용하여, STAT3, STAT5, 및 Akt의 활성화 상태를 결정할 수 있다. STAT3, STAT5, 및 Akt의 활성화는 EGFR 돌연변이체의 활성화의 진단 지표로서 유용하다.

[0144] 본 발명의 하나의 태양에서, 활성화된(인산화된) STAT3, STAT5, 및 Akt의 존재는 EGFR 표적치료가 유효할 수 있음을 나타낸다.

[0145] 본 발명은 면역조직화학법 또는 면역세포화학법에 의해 시험용 생물학적 시료 중의 erbB1 유전자의 키나제 도메인에서의 변이체들을 탐색하는 방법을 제공한다.

[0146] 예를 들면 면역조직화학법(“IHC”)과 면역세포화학법(“ICC”)을 사용할 수도 있다. IHC는 면역화학법을 조직 절편에 적용하는 것이고, 반면에 ICC는 면역화학법을, 예를 들면, 액체계 표본작업(liquid-based preparation)과 같은 특정의 세포학적 표본작업을 거친 후의 세포나 조직 임프린트(imprint)에 적용하는 것이다. 면역화학법은 특이적 항체의 사용에 기반한 기술들의 집합으로서, 여기서는 항체를 사용하여 세포 내부 또는 세포 표면의 분자를 특이적으로 표적으로 한다. 상기 항체는 전형적으로 표지를 함유하는데, 이 표지는 생화학적 반응을 하며, 이에 따라 표적 분자를 만나게 되면, 변색한다. 어떤 경우에 있어서는 신호 증폭을 특정의 프로토콜에 통합시킬 수 있으며, 이때 1차 특이적 항체의 적용 후 표지 착색(marker stain)을 포함하는 2차 항체가 적용된다.

[0147] 면역조직화학법적 분석은 당업자에게 알려져 있다(예로서, Jalkanen, et al., *J. Cell. Biol.* 101: 976-985

(1985); Jalkanen, et al., *J. Cell. Biol.* 105: 3087-3096 (1987) 참조).

[0148] 항체는, 다클론항체이든 또는 단일클론항체이든, 다양한 상업적 공급자들로부터 구매할 수 있거나, 또는 예를 들면, Harlow et al., Antibodies: A Laboratory Mannual, 2판; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1988)에 기재되어 있는 것과 같은 잘 알려진 방법들을 통해 제조할 수도 있다. 일반적으로, 본 발명에 유용한 항체의 예로는 안티-포스포-STAT3, 안티-포스포-STAT5, 및 안티-포스포-Akt 항체를 들 수 있다. 이러한 항체들은, 예를 들면, Upstate Biotechnology사(Lake Placid, NY), New England Biolabs사(Beverly, MA), NeoMarkers사(Fremont, CA)로부터 구입 가능하다.

[0149] 전형적으로, 면역조직화학법을 위해서는, 환자로부터 조직 절편을 얻어 이것을 알콜, 아세톤, 파라포름알데히드 같은 적합한 고착제에 의해 고착시키며, 이것이 항체와 반응한다. 면역조직화학법의 통상의 방법들은 Harlow 및 Lane (eds) (1988) "Antibodies A Laboratory Mannual", Cold Spring Harbor Press, Cold Spring Harbor, New York; Ausbel et al. (eds) (1987), in Current Protocols In Molecular Biology, John Wiley and Sons (New York, NY)에 기재되어 있다. 그러한 검출 분석에 적합한 생물학적 시료들은, 세포, 조직 생검, 전혈(whole blood), 혈장, 혈청, 가래, 뇌척수 액, 유방 흡인액, 흉수, 소변 등을 포함하나 이것들에 한정되는 것은 아니다.

[0150] 직접 표지술에서는 표지된 항체를 활용한다. 간접 표지술에서는 시료를 다시 표지된 물질과 반응시킨다.

[0151] 또는, 면역세포화학법을 활용할 수 있다. 일반적으로 환자로부터 세포를 얻어, 알콜, 아세톤, 파라포름알데하이드 같은 적합한 고착제에 의해 고착하며, 그것이 항체와 반응한다. 인간 시료의 면역세포학적 염색은 당업자에게 알려져 있고, 예를 들면 Brauer et al., 2001 (FASEB J, 15, 2689-2701), Smith-Swintosky et al., 1997에 기재되어 있다.

[0152] 본 발명의 면역학적 방법은 단지 적은 양의 생물학적 물질을 요구하기 때문에 유익하다. 그 방법은 세포 수준에서 행해질 수 있고, 최소 한 개의 세포가 요구된다. 바람직하게는, 암에 걸렸거나 암이 발병할 위험이 있는 환자로부터 몇 개의(several) 세포들을 얻어, 본 발명의 방법에 따라 분석한다.

[0153] **다른 진단 방법들**

[0154] 돌연변이 EGFR 단백질을 검출하는 제제는 돌연변이 EGFR 단백질에 결합할 수 있는 항체이며, 바람직하게는 검출 가능하게 표지된 항체이다. 항체는 다클론 항체일 수도 있고, 더 바람직하게는 단일클론항체일 수도 있다. 완전한 항체 또는 예를 들면 F_{ab}, $F_{(ab)2}$와 같은 항체 단편이 사용될 수도 있다. 프로브 또는 항체에서 "표지된"이라는 용어의 의미는, 검출가능한 물질을 프로브 또는 항체에 커플링(즉, 물리적 결합)을 통해 직접 표지한 경우뿐만 아니라, 직접 표지된 다른 시약과의 반응성에 의해 프로브 또는 항체를 간접 표지한 경우 모두를 포괄한다. 간접 표지의 예로는, 형광 표지된 2차 항체를 사용하여 1차 항체를 검출하는 것과, 형광 표지된 스트렙타비딘(streptavidin)에 의해 검출될 수 있도록 비오틴으로 DNA 프로브의 말단을 표지하는 것을 들 수 있다. "생물학적 시료"란 용어는 피험자(subject)의 내부에 존재하는 조직, 세포, 생물학적 액뿐만 아니라, 피험자로부터 분리된 조직, 세포, 생물학적 액을 포함한다. 즉, 본 발명의 검출 방법은, 생체내(*in vivo*) 뿐만 아니라 생체외(*in vitro*)에서 생물학적 시료 중의 돌연변이 EGFR mRNA, 단백질, 또는 게놈 DNA를 검출하는데 사용할 수 있다. 예를 들면, 돌연변이 EGFR mRNA의 검출을 위한 생체외 기술로는 노던 혼성화법(Northen hybridization)과 *in situ* 혼성화법을 들 수 있다. 돌연변이 EGFR 단백질의 검출을 위한 생체외 기술로는 효소결합 면역흡착법(ELISA), 웨스턴 블로팅법, 면역침전법, 및 면역형광법을 들 수 있다. 돌연변이 EGFR 게놈 DNA의 검출을 위한 생체외 기술로는 서던 혼성화법(Southern hybridization)이 있다. 또한 돌연변이 EGFR 단백질의 검출을 위한 *in vivo* 기술로는 피험자에게 표지된 항-돌연변이 EGFR 단백질 항체를 도입하는 방법이 있다. 예를 들면, 그 항체는, 피험자 내에서의 그 존재와 위치를 표준 영상 기술(standard imaging techniques)에 의해 검출 가능한 방사성 표지로 표지될 수 있다.

[0155] 하나의 태양에서는, 생물학적 시료는 피험자로부터의 단백질 분자를 함유한다. 대안적으로, 생물학적 시료는 피험자로부터의 mRNA 분자 또는 게놈 DNA 분자를 포함할 수도 있다.

[0156] 또 다른 태양에서, 상기 방법들은 대조군 피험자로부터 대조군 생물학적 시료를 얻고, 그 생물학적 시료 내의 돌연변이 EGFR 단백질, mRNA, 또는 게놈 DNA의 존재가 검출될 수 있도록, 그 대조군 시료를, 돌연변이 EGFR 단백질, mRNA, 또는 게놈 DNA를 검출할 수 있는 화합물 또는 시약과 접촉시키고, 그 대조군 시료 내의 돌연변이 EGFR 단백질, mRNA, 또는 게놈 DNA의 존재를 상기 시험용 시료 내의 돌연변이 EGFR 단백질, mRNA, 또는 게놈

DNA의 존재와 비교하는 것을 포함한다.

[0157] 또 다른 태양에서, 상기 진단 분석은 돌연변이 EGFR의 활성에 대한 것이다. 구체적인 태양에서, 상기 돌연변이 EGFR 활성은 티로신 키나제 활성이다. 그러한 진단 분석은, 적어도 하나의 EGFR 기질의, EGFR-매개 인산화를 검출하기 위한 것이다. EGFR 활성의 수준은 예를 들면 다양한 돌연변이 EGFR 폴리펩티드, 돌연변이 EGFR을 함유하는 다양한 조직, 적어도 하나의 돌연변이 EGFR을 가지고 있을 것으로 의심되는 암 조직으로부터의 생검물 등에서 분석될 수 있다. 필요에 따라, 이들 다양한 세포, 조직 또는 그 추출물에서의 EGFR 활성 수준을 비교할 수 있다. 하나의 태양에서, 암 조직에서의 높은 수준의 EGFR 활성은, 하나 이상의 티로신 키나제 저해제에 의한 치료에 민감한 암의 진단 지표이다. 관련된 태양에서는, 처리된 또는 처리되지 않은 생검 시료, 세포주, 형질전환 동물 또는 이들 중 어느 것으로부터의 추출물 사이에서 EGFR 활성 수준을 결정하여, 주어진 치료가 돌연변이 EGFR 활성에 미치는 영향을 처리되지 않은 대조군과 비교하여 결정할 수 있다.

[0158] **환자 치료 방법**

[0159] 하나의 태양에서, 본 발명은 erbB1 유전자의 키나제 도메인에서의, 적어도 하나의, 키나제 활성을 증가시키는 핵산 변이의 존재 또는 부재를 결정함으로써, 암에 걸렸거나 암이 발병할 위험이 있는 환자에 대한 치료를 선택하는 방법을 제공한다. 또 다른 태양에서, 상기 변이는 복수의 변이이며, 여기서 복수란 하나, 둘, 셋 또는 그 이상의 유전자 위치(loci)의 변이를 포함할 수 있다.

[0160] 어떤 태양에서, 적어도 하나의 변이의 존재는 환자의 치료가 효과적이거나 또는 유익함(beneficial)(또는 유익할 가능성이 있음)을 가리킨다. "치료가 효과적"이라는 것은, 유익한 치료 효과의 가능성이, erbB1 유전자의 키나제 도메인에 특정 키나제 활성을 증가시키는 적합한 핵산 변이가 없는 인간보다 크다는 것을 의미한다.

[0161] 이 치료는 티로신 키나제 저해제의 투여를 포함한다. 이 치료는 치료들의 조합을 포함할 수 있는데, 티로신 키나제 저해제와, 다른 티로신 키나제 저해제, 화학요법, 방사선요법 등과의 조합을 포함하나 이에 한정되지는 않는다.

[0162] 따라서, 티로신 키나제 저해제의 투여와 관련하여, 암에 "대하여 유효한" 약(drug)이란, 임상적으로 적합한 방법으로 투여한 결과 적어도 통계적으로 유의미한 수의 환자에게서 증상의 개선, 치유, 질병 부담의 경감, 종양의 크기 또는 세포수 감소, 수명의 연장, 삶의 질의 향상 또는 특정 유형의 질병 또는 상태의 치료에 익숙한 의사에 의해 일반적으로 긍정적으로 인식될 수 있는 다른 효과 등과 같은 유익한 효과가 있음을 가리킨다.

[0163] 바람직한 태양에서, 이 화합물은 아닐리노퀴나졸린 또는 합성 아닐리노퀴나졸린이다. 유럽 특허공개번호 EP 0566226에는 EGFR 티로신 키나제에 대해 활성을 갖는 아닐리노퀴나졸린이 개시되어 있다. 유럽 특허출원 EP 0520722와 EP 0566226에는 특정한 4-아닐리노퀴나졸린 유도체가 수용체 티로신 키나제 저해제로서 유용하다는 사실이 밝혀져 있다. 이들 화합물들에 의해 밝혀진 구조와 활성간의 긴밀한 관계는 명확하게 한정된 결합 양식(binding mode)을 제안한다. 즉 퀴나졸린환은 아데닌 포켓에 결합하며, 아닐리노환은 이웃한 고유의(unique) 친지질성 포켓에 결합한다. 세 개의 4-아닐리노퀴나졸린 유사체(두개는 가역적이며 다른 한 개는 비가역적인 저해제임)는 임상적으로 항암제로 평가되었다(Denny, Farmaco January-February 2001; 56(1-2):51-6 참조). 또는, 상기 화합물은 EKB-569로서, EGFR 키나제 저해제이다(Torrance et al., *Nature Medicine*, Vol. 6, No. 9, Sept. 2000, p.1024 참조). 가장 바람직한 태양에서, 이 화합물은 제피티니브(IRESSA®) 또는 에를로티니브(TARCEVA®)이다.

[0164] 여기에 기술된 적어도 하나의 돌연변이 EGFR를 함유하는 암세포를 표적으로 하는 치료는, 단독으로 투여되거나 다른 적합한 항암 치료 및/또는 당업자에게 알려져 있는 치료제와 조합하여 투여된다. 하나의 태양에서, 암과 같은 병변의 치료는, 치료를 요하는 피치료자에게, 제피티니브, 에를로티니브와 같은 EGFR 키나제 활성을 저해하는 화합물의 치료적 유효량을 단독 투여하거나 또는 적어도 하나의 다른 항암제 또는 치료와 조합하여 투여하는 것을 포함한다. 표적으로 된 소분자 의약(small molecule drugs) 또는 항체 기반 전략을 통한 활성화된 단백질 키나제의 저해는 암 치료에 효과적인 접근방법으로 판명되었다(G. D. Demetri et al., *N. Eng. J. Med.* 347, 472(2002); B. J. Druker et al., *N. Engl. J. Med.* 344, 1038(2001); D. J. Slamon et al., *N. Engl. J. Med.* 344, 783 (2001) 참조).

[0165] 하나의 태양에서, 항암제는 적어도 하나의 화학요법제이다. 관련된 태양에서, 항암제는 적어도 하나의 방사선 치료이다. 변형된 태양에서, 항암치료는 항혈관생성 치료이다(예, 엔도스타틴, 안지오스타틴, TNP-470, 카플로스타틴(Caplostatin)(Stachi-Fainaro et al., *Cander Cell* 7(3), 251 (2005)).

[0166] 치료제들은 서로 동일해도 달라도 좋으며, 예를 들면, 치료용 방사성 핵종, 약(drug), 호르몬, 호르몬 길항제,

수용체 길항제, 다른 시약에 의해 활성화된 효소 또는 전효소, 오토크린(autocrine), 사이토카인, 또는 당업자에게 알려진 적합한 항암제 중의 어느 것이다. 하나의 태양에서 항암제는, 항-VEGF 항체로서, 고형암과 혈액악성종양에 대한 항 혈관생성 치료에 있어서 성공적임이 입증된 아바스틴(Avastin)이다(Ribatti et al., 2003, J. Hematother *Stem Cell Res.* 12(1), 11-12 참조). 독소(Toxin) 역시 본 발명의 방법에 이용될 수 있다. 본 발명에 이용되는 또 다른 치료제로는 항DNA제, 항RNA제, 안티센스 올리고누클레오티드와 같은 방사성 표지된 올리고누클레오티드, 항단백질제, 항 염색질 세포독성제 또는 항미생물제를 들 수 있다. 당업자에게 알려져 있는 다른 치료제, 및 그러한 다른 치료제의 본 발명에 따른 사용이 특이적으로 고려된다.

[0167] 항종양제는 알킬화제, 대사저해제, 호르몬제, 항생제, 항체, 생물학적 항암제(anti-cancer biological), 글리벡(gleevec), 콜히친(colchicine), 빈카 알칼로이드(vinka alkaloid), L-아스파라기나제, 프로카르바진(procarbazine), 히드록시우레아, 미토탄(mitotane), 니트로소우레아, 또는 이미다졸 카르복스아미드(imidazole carboxamide)와 같은 다수의 화학요법제 중의 하나일 수 있다. 적합한 제제들은 튜뷸린(tubulin)의 탈분극을 촉진하거나 또는 종양세포 증식을 방해하는 것들이다. 본 발명의 범위 내에서 고려되는 화학요법제에는 미국 식품의약국(Food and Drug Administration)과 미국 보건사회복지성(Department of Health and Human Services)에 의한, 치료동등성이 평가된 승인된 의약품의 오렌지북(Orange Book of Approved Drug Products with Therapeutic Equivalence Evaluations)에 등재된 항암제들이 포함되나, 이것에 한정되는 것은 아니다. 화학요법제의 예로 카르보플라틴(carboplatin), 파클리탁셀을 들 수 있으나 이에 한정되는 것은 아니다. EGFR 키나제 활성을 표적으로 하는 치료 역시 방사선치료와 함께 투여될 수 있다. 당 기술분야에 알려져 있는 다른 항암치료들도 본 발명의 범위 내에서 고려된다.

[0168] 치료제는 화학요법제일 수도 있다. 화학요법제들은 당 기술분야에 알려져 있으며 적어도 탁산, 니트로겐 머스타드, 에틸렌이민 유도체, 알킬 설포네이트, 니트로소우레아, 트리아젠, 폴산(folic acid) 유사체, 피리미딘 유사체, 퓨린 유사체, 빈카 알칼로이드, 항생제, 효소, 백금착화합물, 우레아 치환체, 메틸 히드라진 유도체, 부신피질억제제 또는 길항제를 포함한다. 더 구체적으로, 화학요법제는 스테로이드, 프로게스틴, 에스트로겐, 안티에스트로겐, 안드로겐 등의, 그러나 이것에 한정되지 않는 군으로부터 선택된 하나 이상일 수 있다. 더 구체적으로, 화학요법제는 아자리빈(azaribine), 블레오마이신(bleomycin), 브리오스타틴-1(bryostatin-1), 부설판(busulfan), 카르무스틴(carmustine), 클로람부실(chlorambucil), 카르보플라틴(carboplatin), 시스플라틴(cisplatin), CPT-11, 시클로포스파미드(cyclophosphamide), 시타라빈(cytarabine), 다카르바진(dacarbazine), 다크티노마이신(dactinomycin), 다우노루비신(daunorubicin), 덱사메타손(dexamethasone), 디에틸스틸베스트롤(diethylstilbestrol), 독소루비신(doxorubicin), 에티닐 에스트라디올(ethinyl estradiol), 에토포시드(etoposide), 플루오로우라실(fluorouracil), 플루옥시메스테론(fluoxymesterone), 겜시타빈(gemcitabine), 히드록시프로게스테론 카프로에이트(hydroxyprogesterone caproate), 히드록시우레아, L-아스파라기나제, 류코보린(leucovorin), 로무스틴(lomustine), 메클로레타민(mechlorethamine), 메드로프로게스테론 아세테이트(medroprogesterone acetate), 메게스트롤 아세테이트(megestrol acetate), 멜팔란, 머캅토퓨린, 메토트렉세이트(methotrexate), 미트라마이신(mithramycin), 미토마이신(mitomycin), 미토탄(mitotane), 파클리탁셀(paclitaxel), 페닐 부티레이트, 프레드니손(prednisone), 프로카르바진(procarbazine), 세무스틴 스트렙토조신(semustine streptozocin), 타목시펜(tamoxifen), 탁산(taxane), 탁솔, 테스토스테론 프로피오네이트, 탈리도마이드, 티오구아닌, 티오테파(thiotepa), 우라실 머스타드(uracil mustard), 빈블라스틴(vinblastine) 또는 빈크리스틴(vincristine)일 수 있다. 화학요법제의 어떠한 조합이든 그 이용이 고려된다. 화학요법제의 투여는 EGFR 활성을 표적으로 하는 치료의 전, 중 또는 후일 수 있다.

[0169] 다른 적합한 치료제는 방사성 동위원소, 붕소 부가물(boron addend), 면역조절자(immunomodulator), 독소(toxin), 광활성제 또는 광활성염료, 암 화학치료약, 항바이러스약, 항진균약, 항박테리아약, 항원충약, 화학민감화약(chemosensitizing drug)으로 이루어진 군으로부터 선택된다(미국 특허번호 4,925,648 및 4,932,412 참조). 적합한 화학요법제는 Remington's Pharmaceutical Sciences, 19판 (Mack Publishing Co. 1995)과 Goodman and Gilman's The Pharmacological Basis of Therapeutics (Goodman et al., Eds. Macmillan Publishing Co., New York, 1980년 및 2001년판)에 기술되어 있다. 실험 의약들과 같은 다른 적합한 화학요법제들은 당업자에게 알려져 있다. 또, 적합한 치료적 방사성 동위원소는 α-방사체, β-방사체, γ-방사체, 오제 전자 방사체, α-입자를 방사하는 중성자 포획제와 전자 포획에 의해 감쇠되는 방사성 동위원소로 구성되는 군으로부터 선택된다. 바람직하게는 방사성 동위원소는 225Ac, 198Au, 32P, 125I, 131I, 90Y, 186Re, 188Re, 67Cu, 177Lu, 213Bi, 10B 및 211At로 이루어진 군으로부터 선택된다.

[0170] 둘 이상의 치료제를 사용함에 있어서는, 그들은 동일할 수도 있고 다른 것일 수도 있다. 예를 들면, 치료제는

서로 다른 방사성 핵종들(radionuclides), 또는 약품과 방사성 핵종을 포함할 수 있다. 바람직한 태양에서, EGFR 활성을 표적으로 하는 치료는 돌연변이 EGFR 키나제 활성을 저해한다.

[0171] 또 다른 태양에서는, 개별적인 에너지 방출의 결과 서로 다른 거리(distance)에 걸쳐 유효한 서로 다른 동위원소들을 제1 및 제2 치료제로 사용한다. 이러한 시약들은 더욱 효과적인 종양 치료를 달성하기 위하여 이용될 수 있으며, 정상적인 임상 상황에서 서로 다른 크기의 다발성 종양을 보이는 환자를 치료하는데 유용하다.

[0172] 입수가능한 동위원소 중 극히 작은 종양 퇴적물과 단일 세포를 치료하는데 유용한 것은 매우 적다. 이러한 상황에서는 약(drug) 또는 독소(toxin)가 더욱 유용한 치료제로 될 수 있다. 따라서 본 발명의 바람직한 태양에서는, 동위원소를, 약, 독소, 중성자 포획제와 같은 비 동위원소 종과 조합하여 사용한다. 본 발명과 연관되어 이용될 수 있는 많은 약과 독소는 세포에 대하여 세포독성(cytotoxic) 효과가 있는 것으로 알려져 있다. 이들은 Merck Index, Goodman and Gilman 등과 같은 약과 독소의 개론서 및 위에서 언급한 참고문헌들에서 발견된다.

[0173] 세포내의 단백질 합성을 간섭하는 약 역시 본 발명의 방법에 사용될 수 있는데, 이러한 약들은 당업자에게 알려져 있으며, 푸로마이신(puromycin), 시클로헥시미드(cycloheximide),리보누클레아제를 포함한다.

[0174] 본 발명의 치료 방법은 암 치료에 사용될 수 있다. 방사성 동위원소, 약, 독소가 암세포와 관련되어 또는 암세포에 의해 제조된 표지에 특이적으로 결합하는 항체 및 항체 단편에 접합(conjugate)될 수 있다는 사실은 잘 알려져 있다. 또한 이러한 항체 접합체(antibody conjugates)를 사용하여 방사성 동위원소, 약, 독소를 종양 위치에 도달시켜, 치료 효과를 향상시키고 부작용을 최소화할 수 있다. 이러한 시약과 방법에 대한 예들은 Wawrzynczak 및 Thorpe(in Introduction to the Cellular and Molecular Biology of cancer, L. M. Franks and N. M. Teich, eds, Chapter 18, pp.378-410, Oxford University Press. Oxford, 1986), Immunoconjugates: Antibody Conjugates in Radioimaging and Therapy of cancer(C. W. Vogel, ed., 3-300, Oxford University Press, N.Y., 1987), Dillman, R. O. (CRC Critical Reviews in *Oncology/Hematology* 1:357, CRC Press, Inc., 1984), Pastan et al.(*Cell* 47: 641, 1986), Vitetta et al. (*Science* 238:1098-1104, 1987) 및 Brady et al.(*Int. J. Rad. Oncol. Biol. Phys*. 13:1535-1544, 1987)에 검토되어 있다. 암에 대한 면역접합체의 이용 및 다른 형태의 치료에 대한 다른 예들은 특히 미국 특허번호 4,331,647, 4,348,376, 4,361,544, 4,468,457, 4,444,744, 4,460,459, 4,460,561, 4,624,846, 4,818,709, 4,046,722, 4,671,958, 4,046,784, 5,332,567, 5,443,953, 5,541,297, 5,601,825, 5,635,603, 5,637,288, 5,677,427, 5,686,578, 5,698,178, 5,789,554, 5,922,302, 6,187,287호, 및 6,319,500에 개시되어 있다.

[0175] 부가적으로, 본 발명의 치료 방법은, 어떤 세포 독성 제제의 부작용을 방해, 완화, 역전시키는 다른 화합물 또는 기술과 조합하여 이용할 수 있다. 이러한 조합의 예로는 미국 특허번호 4,624,846에 기술된 바와 같은, 빠른 제거를 위해 IL-1을 항체와 함께 투여하는 것을 들 수 있다. 이러한 투여는, EGFR 활성을 표적으로 하는 1차 치료로부터 3시간~72시간 후에, 항암제(세포독성 성분으로서의 방사성 동위원소, 약 또는 독소)와 조합하여 수행할 수 있다. 이것은 순환계로부터의 접합체, 약, 독소의 제거를 증진하며 치료제에 의하여 야기되는 척수 및 다른 조혈독성(hematopoietic toxcity)을 완화 또는 역전시키는데 이용될 수 있다.

[0176] 본 발명의 다른 양상에서, 암 치료는, 예를 들면 약과 방사성 동위원소; 또는 방사성 동위원소와, 중성자 활성 치료(neutron-activated therapy)를 위한 붕소-10; 또는 약과 생물학적 반응 조절자(biological response modifier); 또는 융합분자접합체(fusion molecule conjugate)와 생물학적 반응 조절자 등과 같이, 둘 이상의 살종양제(tumoricidal agent)의 조합을 포함할 수 있다. 사이토카인을 치료적 투여 계획(therapeutic regimen)에 포함시킴으로써 각 성분의 효율성을 극대화할 수 있다.

[0177] 유사하게, β 또는 α 방사체인 방사성 동위원소와 접합된 어떤 항백혈병 및 항림프종 항체들은, 이들이 순전히 종양세포에로만 향하여지지 않을 경우, 골수 및 다른 조혈 부작용을 야기할 것이다. 이것은 특히 종양 세포가 순환계 또는 혈액 형성 기관에 있을 경우 발견된다. 적어도 하나의 조혈 사이토카인(예, 콜로니 자극 인자, G-CSF, GM-CSF와 같은 성장 인자)의 동시 및/또는 순차 투여는, 항암 효과를 증가시키는 동시에 조혈 부작용(hematopoietic side effect)을 감소 또는 개선시킴에 있어서 바람직하다.

[0178] Harvert, "Nuclear Medicine Therapy", New York, Thieme Medical Publishers, 1087, pp.1-340에 기술된 바와 같이, 다양한 방사성핵종 치료법을 암과 또 다른 병리적 상태의 치료에 이용할 수 있음은 당 기술분야에 잘 알려져 있다. 이러한 과정을 경험한 임상의사는, 그 조혈 부작용을 경감하기 위한 과정에, 여기에 기술된 사이토카인 보조 치료를 용이하게 적용할 수 있을 것이다. 이와 유사하게, EGFR 활성을 표적으로 하는 치료와 함께

투여되는 세포독성 약에 의한 치료는, 예를 들면 암 또는 다른 세포 증식 질환을 위한 치료에 사용될 수 있다. 그러한 치료는 동위원소 또는 방사성 표지된 항체에 의한 방사성 동위원소 치료법과 유사한 원칙에 의해 지배된다. 통상의 숙련된 임상의사는 1차 항암 치료의 전, 중 및/또는 후에 부가적인 항암 치료 투여를 적용할 수 있을 것이다.

[0179] **키트**

[0180] 본 발명은 또한, erbB1 유전자의 키나제 도메인에서의 표적 핵산을 증폭하기 위한 축퇴성 프라이머, 및 증폭 프로토콜과 그 결과의 분석을 포함하는 설명서를 포함하는 예측, 진단, 예후 키트를 제공한다. 이 키트는 필요에 따라 완충액, 효소, 및 증폭의 수행과 증폭 결과물의 분석을 위한 용기들을 포함할 수 있다. 이 키트는 DNA 마이크로어레이와 같은 다른 도구(tool)를 포함하는 탐색, 진단, 또는 예후 키트의 일 구성요소일 수도 있다. 바람직하게는, 이 키트는 정상적인 조직 시료로부터 분리된 핵산 및/또는 erbB1 유전자의 키나제 도메인에서의 서로 다른 변이들을 나타내는 일련의 시료들로부터 분리된 핵산 등과 같은 하나 이상의 대조군 주형(template)도 제공한다.

[0181] 하나의 태양에서, 이 키트는 둘 이상의 프라이머쌍을 제공하며, 각 프라이머쌍은 erbB1 유전자의 서로 다른 영역(각 영역은 잠재적 변이 부위임)을 증폭할 수 있으며, 이에 의해 하나의 반응 또는 몇 개의 병렬적 반응(parallel reaction)으로 생물학적 시료 중의 몇 가지 유전자 변이의 발현을 분석하는 키트를 제공한다.

[0182] 이 키트 내의 프라이머는 예를 들면 형광으로 표지되어 있어서 증폭 결과물과 이어지는 핵산 변이들의 분석을 용이하게 할 수 있다.

[0183] 하나의 태양에서는, 2 이상의 변이를 한번의 분석으로 검출할 수 있다. 그러므로 조합 키트는 erbB1 유전자의 키나제 도메인의 서로 다른 세그먼트들을 증폭할 수 있는 프라이머들을 포함할 것이다. 그 프라이머들은, 변이들을 분별할 수 있도록, 서로 다르게 표지될 수도 있다(예를 들어 다른 색의 형광 표지를 사용).

[0184] 상기 키트 내에 함유되어 있는 프라이머들은 다음 프라이머들을 포함할 수 있다: 엑손 19 센스 프라이머, 5'-GCAATATCAGCCTTAGGTGCGGCTC-3'(서열번호: 505); 엑손 19 안티센스 프라이머, 5'-CATAGAAAGTGAACATTTAGGATGTG-3'(서열번호: 506); 엑손 21 센스 프라이머, 5'-CTAACGTTCGCCAGCCATAAGTCC-'3(서열번호: 507); 엑손 21 안티센스 프라이머, 5'-GCTGCGAGCTCACCCAGAATGTCTGG-3'(서열번호: 508).

[0185] 바람직한 태양에서, 프라이머들은 서열번호 646~673(표 5 와 6 참조)로 구성되는 군에서 선택된다. 이들 프라이머들은 정방향 프라이머의 5' 말단에 서열번호 645를 가지며, 역방향 프라이머의 5' 말단에 서열번호 674를 가진다.

[0186] **면역검출 키트**

[0187] 또 다른 태양에서, 본 발명은 다운스트림 EGFR 표적(즉, STAT3, STAT5, 및 Akt)의 활성화 수준을 검출하는데 사용하기 위한 면역학적 키트를 제공한다. 이러한 키트는 일반적으로 STAT3, STAT5, 또는 Akt의 인산화된 형태에 면역특이성을 갖는 하나 이상의 항체들을 포함할 것이다.

[0188] 본 발명은, 인산화된 Akt, STAT3, 및 STAT5 단백질로 구성되는 군으로부터 선택되는 포유동물 세포 내의 인산화된 단백질에 면역특이적으로(immunospecifically) 결합할 수 있는 항체와, 이 항체를 이용하여 포유동물 세포에서 Akt, STAT3, 또는 STAT5 경로 활성화(pathway activation)를 조사할 수 있는 설명서를 포함하는 키트를 제공한다. 바람직한 방법에 있어서는, 이 키트는 인산화된 Akt, STAT3, 또는 STAT5 단백질로 구성되는 군으로부터 선택되는 포유동물 세포 내의 인산화된 단백질에 각각 면역특이적으로 결합할 수 있는 서로 다른 항체들을 포함한다.

[0189] 상기 키트는 일반적으로 a) 약학적으로 허용가능한 담체; b) 적합한 용기에 담겨 있는, 인산화된 STAT3, STAT5, 또는 Akt에 대한 항체; 및 c) 면역검출제를 포함한다. 단클론 또는 다클론 항체들은 상업적으로 입수가능하며, 예를 들면 Current Protoclos in Immunology, John Wiley & Sons, Edited by: John E. Coligan, Ada M. Krusibeek, David H. Margulies, Ethan M. Shevach, Warren Strober, 2001에 기술된 바와 같은 당업자에게 알려져 있는 방법으로 제조할 수도 있다.

[0190] 어떤 태양들에서, 항원과 항체는 칼럼 매트릭스 또는 마이크로타이터 플레이트의 웰(well)과 같은 고체 지지체에 결합되어 있을 수 있다. 이 키트의 면역검출제는 주어진 항체 또는 항원 그 자체와 결합되거나(associated) 연결된(linked) 검출가능한 표지를 포함하는 다양한 형태들 중에서 하나를 택할 수 있다. 2차 결합 배위자에 결합(associated) 또는 부착되어(attached) 있는 검출가능한 표지도 고려된다. 2차 배위자의 예로는 1차 항체

또는 항원에 결합친화성(binding affinity)을 갖는 2차 항체들이 있다.

[0191] 당 기술분야에 알려져 있는 적합한 분석 표지로는 글루코스 산화효소와 같은 효소 표지(enzyme label); 요오드(^{131}I, ^{125}I, ^{123}I, ^{121}I), 탄소(^{14}C), 황(^{35}S), 트리튬(^{3}H), 인듐(^{115m}In, ^{113m}In, ^{112}In, ^{111}In), 및 테크네튬(technetium)(^{99}Tc, ^{99m}Tc), 탈륨(^{201}Ti), 갈륨(^{68}Ga, ^{67}Ga), 팔라듐(^{103}Pd), 몰리브데늄(^{99}Mo), 제논(^{133}Xe), 불소(^{18}F), ^{153}Sm, ^{177}Lu, ^{159}Gd, ^{149}Pm, ^{140}La, ^{175}Yb, ^{166}Ho, ^{90}Y, ^{47}Sc, ^{186}Re, ^{188}Re, ^{142}Pr, ^{105}Rh, ^{97}Ru와 같은 방사성 동위원소들; 루미놀(luminol)과 같은 발광 표지들; 플루오레세인(fluorescein), 및 로다민(rhodamine)과 같은 형광 표지들, 및 비오틴(biotin)을 들 수 있다.

[0192] 본 키트에 사용되는 더 적합한 면역검출제는 2성분제를 포함하며, 이 2성분제는 1차 항체나 1차 항원에 결합친화성을 갖는 2차 항체와 함께, 상기 2차 항체에 결합친화성을 갖는, 검출 표지에 연결되어 있는 3차 항체를 포함한다.

[0193] 다수의 표지의 예가 당 기술분야에 알려져 있으며 그러한 모든 표지들은 본원발명에 관하여 사용될 수 있다. 방사성 표지, 핵자기 스핀 공명 동위원소(nuclear magnetic spin-resonance isotopes), 형광 표지, 및 적합한 기질과 접촉하여 색을 띠는 결과물을 생성할 수 있는 효소 표지(emzyme tag)가 그 적합한 예이다.

[0194] 상기 키트는 완전히 접합된 형태, 중간체 형태, 또는 키트의 사용자에 의해 접합될, 분리되어 있는 부분들, 중의 어느 것으로 된 항체-표지 접합체(antibody-label conjugate)를 함유할 수 있다.

[0195] 상기 키트는, 검출 분석을 위한 표준 곡선(standard curve)을 제작하기 위해 사용되거나 또는 양성 대조군으로서 사용될 수 있는, 표지된 또는 표지되지 않은 항원의 적절히 등분된(aliquored) 조성물을 더 포함할 수 있다.

[0196] 본 발명의 키트는, 그 형태를 불문하고, 일반적으로, 생물제가 들어 있는, 및 바람직하게는 생물제가 적절히 등분되어 있는 하나 이상의 용기를 포함할 것이다. 이 키트의 성분들은 수성 매체 중에 또는 동결건조된 형태로 포장될 수 있다.

[0197] 본 발명의 면역검출 키트는 필요에 따라 부가적으로 하나 이상의 다양한 암(cancer) 표지 항체 또는 항원을 함유할 수 있다. 이러한 키트는 다양한 환자를 시험하는데 더 유용한 암 표지의 패널(panel)을 제공할 수 있다. 이러한 부가적인 표지의 예로는 p53, 사이클린 D1, p16, 티로시나제, MAGE, BAGE, PAGE, MUC18, CEA, p27,[bgr]HCG 뿐만 아니라 PAS, SeLe(X), HCG와 같은 다른 종양 표지들 또는 당업자에게 알려져 있는 다른 표지들을 들 수 있다.

[0198] 키트의 용기들은, 일반적으로, 항체 또는 항원이 들어 있거나 바람직하게는 적절하게 등분되어 있는 적어도 하나의 바이알, 시험관, 플라스크, 병, 또는 주사기 또는 다른 용기를 포함할 것이다. 2차 또는 3차 결합 배위자(binding ligand) 또는 부가적인 성분이 제공될 경우, 이 키트는 일반적으로 이 배위자 또는 성분을 담을 수 있는 제2의, 제3의 또는 다른 부가적인 용기들을 또한 함유할 것이다.

[0199] 본 발명의 키트는, 전형적으로, 상업적 판매를 위해 기밀된, 항체, 항원 및 , 다른 시약 용기들을 담기 위한 수단을 포함할 것이다. 이러한 용기들로는 소망 바이알이 보유되는, 사출 또는 블로우 성형된 플라스틱 용기들을 들 수 있다.

[0200] 본 발명의 방법은 또한 EGFR 변이형의 키나제 활성을 간섭하는 화합물의 동정을 포괄한다. 변이체 EGFR은 그 키나제 도메인에 적어도 하나의 변이를 포함한다. 예를 들어 이러한 화합물은 티로신 키나제 저해제이다. 수용체의 키나제 활성을 간섭하는 화합물을 동정하는 방법들은 당업자에게 일반적으로 알려져 있으며, 예를 들면 본 명세서에 참고문헌으로서 편입되는 Dhanabal et al., *Cancer Res.* 59:189-197(1999); Xin et al., *J. Biol. Chem.* 274:9116-9121(1999); Sheu et al., *AntiCancer Res.* 18:4435-4441; Ausprunk et al., *Dev. Biol.* 38:237-248(1974); Gimbrone et al., *J. Natl. cancer Inst.* 52:413-427; Nicosia et al., *In vitro* 18:538-549에 더 기술되어 있다. 일반적으로, erbB1 유전자의 키나제 도메인에서의 적어도 하나의 변이의 증가된 키나제 활성 특성을 간섭하는 화합물은 여기 개시된 방법을 이용하여 동정된다.

[0201] **고체 지지체(Solid Support)**

[0202] 또 다른 태양에서, 본 발명은 본 발명의 방법을 실행하기 위해 키트를 제공한다. 하나의 태양에서는, 고체 지지체 상의 erbB1 유전자의 키나제 도메인내에서의 변이를 검출하기 위한 키트를 기재한다. 상기 키트는, 예를 들면, 하나의 측정법(assay)으로 복수의 변이를 검출하기 위한 재료 및 시약을 포함한다. 상기 키트는, 예를 들면, 고체 지지체, 특정세트의 표적 폴리누클레오티드용 올리고누클레오티드 프라이머, 중합효소연쇄반응(PCR)

시약 및 구성성분, 예를 들면, DNA 합성 효소, 표지 물질, 및 기타 완충액 및 세정용 시약을 포함할 수 있다. 또한, 상기 키트는 고체 지지체 상에서 특이적 표적을 증폭시키기 위한 키트 사용 설명서를 포함할 수 있다. 상기 키트가, 예를 들면, 표적 폴리누클레오티드의 특정 세트를 증폭시키기 위해서, 고체 지지체상에 이미 고정된 프라이머 세트를 갖도록 제조된 고체 지지체를 포함하는 경우에는, 이러한 제조된 고체 지지체의 설계 및 구성은 상기에 기재한 바와 같다. 또한, 상기 키트는, 예를 들면, 상기 지지체가 *in-situ* 타입 PCR 기기를 사용하여 PCR 증폭을 할 수 있는 경우에 *in-situ* 타입 또는 고체상 타입 PCR 공정을 이용하여, PCR을 수행하는데 필요한 시약을 포함한다. 상기 키트에 포함된 PCR 시약은 범용 PCR 완충액, 열안정성 폴리머라제(예, Taq 폴리머라제), 누클레오티드(예, dNTPs), 및 다른 구성성분 및 표지 분자(예를 들면, 상기한 바와 같은 직접 또는 간접 표지용)를 포함한다. 상기 키트는 고정된 프라이머(immobilized primer)를 단독으로 또는, 필요에 따라서, 용액상 프라이머(solution phase primer)와 함께 사용하는 PCR 증폭법의 실행을 지지하기 위해 조립(assemble)될 수 있다.

[0203] 필요에 따라서, 상기 키트는 많은 EGFR 변이에 특이적인 올리고누클레오티드가 부착된(affixed) 고체 지지체를 포함할 수 있으며, 이는 도 4A-4C 및 도 7 및 8에 나타내었다. 시험의 생물학적 시료는, erbB1의 키나제 도메인내의 변이의 존재 또는 부재를 결정하기 위해, 선택적 혼성화 조건 하에서, 고체 지지체에 적용될 수 있다.

[0204] 또한, 본 발명의 방법은 변이형 EGFR의 키타제 활성을 저해하는 화합물의 동정을 포함한다. 상기 EGFR 변이체는 그 키나제 도메인내에 적어도 하나의 변이를 포함한다. 그러나, 다른 태양에서는, 상기 EGFR 변이체는 제1 TKI, 예를 들면, 제피티니브 또는 에를로티니브에 저항성을 부여하는 2차 돌연변이를 포함한다. 이러한 화합물은, 예를 들면, 티로신 키나제 저해제일 수 있다. 수용체의 키나제 활성을 저해하는 화합물을 동정하는 방법은 당업자에게 일반적으로 알려져 있으며, 또한 예를 들면, 본 명세서 중에 참고로 기재한 Dhanabal et al., *Cancer Res*. 59:189-197(1999); Xin et al., *J. Bio. Chem*. 274:9116-9121(1999); Sheu et al., *Anticancer Res*. 18:4435-4441; Ausprunk et al., *Dev. Biol*. 38:237-248(1974); Gimbrone et al., *J. Natl. Cancer Inst*. 52:413-427; Nicosia et al., *In vitro* 18:537-549에 기재되어 있다. erbB1 유전자의 키나제 도메인내에 존재하는 적어도 하나의 변이의 특징적인 증강된 키나제 활성을 저해하는 화합물은, 통상, 본 명세서 중에 개시한 방법을 사용하여, 동정된다.

[0205] 동정 즉시, 이 화합물들은 EGFR 표적 치료를 필요로 하는 환자, 예를 들면 암에 걸렸거나 암이 발병할 위험이 있는 환자에게 투여된다.

[0206] 상기 투여 경로는 정맥내(I.V.), 근육내(I.M.), 피하(S.C.), 진피내(I.D.), 복막내(I.P.), 경막내(I.T.), 흉막내, 자궁내, 직장, 질내, 국소, 종양내 등일 수 있다. 본 발명의 화합물은 주사 또는 여러시간에 걸친 점진적인 주입에 의해 비경구적으로 투여할 수 있으며, 연동운동 수단(peristaltic means)에 의해서 전달할 수도 있다.

[0207] 투여는 경점막 또는 경피 수단에 의해서도 가능하다. 경점막 또는 경피 투여하기 위해서는, 침투할 장벽에 적합한 침투제가 제제(formulation)에 사용된다. 이러한 침투제는 당업자에게 일반적으로 알려져 있으며, 예를 들면, 경점막 투여용 담즙염 및 푸시딘산 유도체를 들 수 있다. 또한, 계면활성제(detergent)는 침투를 촉진시키기 위해서 사용할 수 있다. 경점막 투여는 비강 스프레이, 또는, 예를 들면, 좌약을 사용하여 행할 수 있다. 경구투여를 위해, 본 발명의 화합물은 캡슐, 정제(tablet) 및 강장제(tonic) 등의 종래의 경구 투여 형태로 제제화된다.

[0208] 국소 투여를 위해, 본 발명의 약제학적 조성물(키나제 활성 저해제)은 일반적으로 당업계에서 공지된 바와 같이, 연고, 고약, 겔, 또는 크림으로 제제화된다.

[0209] 본 발명의 치료 조성물은, 예를 들면, 단위 투여량의 주입에 의해, 통상 정맥내 투여된다. 본 발명의 치료 조성물에 관련하여 사용된 경우 "단위 투여량(unit dose)"은 피험자에 대한 단위 투여(unitary dosage)로서 적합한, 물리적으로 구별되는 단위를 말하며, 각 단위는 필요한 희석제(즉, 담체 또는 수송체(vehicle))와 함께, 원하는 치료 효과를 얻을 수 있도록 계산된 소정량의 활성 물질을 포함한다.

[0210] 상기 조성물은 상기 투여 제제와 양립할 수 있는 방식으로, 치료 유효량으로 투여된다. 투여할 양과 시점은 치료받을 피험자, 활성 성분을 이용할 수 있는 피험자의 시스템의 능력, 및 원하는 치료 효과의 정도에 따라 다르다. 투여에 필요한 활성 성분의 정확한 양은 진료의의 판단에 따라 다르며 각 개체마다 상이하다.

[0211] 본 발명의 방법을 실행하는데 유용한 티로신 키나제 저해제는 본 명세서 중에 기재되어 있다. 의도한 용도에 적합한, 활성 성분을 함유하는 어떠한 제제 또는 약제 전달 시스템은, 당업자에게 일반적으로 공지되어 있는 바

에 따라, 사용할 수 있다. 경구투여, 직장투여, 국소투여 또는 비경구 투여용(흡입, 피하, 복막내, 근육내 및 정맥내 투여를 포함함)으로 적합한 약제학적으로 허용가능한 담체는 당업자에게 공지되어 있다. 상기 담체는 제제의 다른 성분과 양립되고 피험자에게 해를 끼치지 않는다는 의미에서 약제학적으로 허용가능한 것이어야 한다.

[0212] 본원에 사용된 "약제학적으로 허용가능한(pharmaceutically acceptable)", "생리적으로 허용가능한(physiologically tolerable)" 및 그들의 문법적인 변화 어구는 서로 호환하여 사용하며, 그 재료가 바람직하지 못한 생리학적 효과를 유발하지 않고 포유동물에 투여할 수 있음을 나타낸다.

[0213] 비경구적인 투여에 적합한 제제로는, 적합하게는, 피험자의 혈액과 바람직하게는 등장인, 활성 화합물의 살균된 수성 약제를 들 수 있다. 따라서, 이러한 제제는, 적합하게는, 멸균수, 5% 덱스트로스 멸균수 또는 식염수(saline)를 포함할 수 있다. 또한, 유용한 제제로는, 적당한 용제로 희석시 비경구 투여용으로 적합한 용액으로 되는 화합물을 함유하는 농축 용액 또는 고체를 들 수 있다.

[0214] 소화관내(enteral) 투여를 위해서는, 화합물을 캡슐, 교갑(cachet), 정제(tablet) 또는 마름모꼴 정제(lozenge) 등의 개별 단위의 불활성 담체에 삽입할 수 있으며, 각 담체는 소정량의 활성 화합물을 분말 또는 과립, 또는 수성액 또는 비수성액의 현탁액 또는 용액(예, 시럽, 엘릭시르(elixir), 에멀션 또는 드래프트(draught))으로서 함유한다. 적합한 담체로는 전분 또는 당류를 들 수 있으며, 윤활제, 향료, 바인더, 및 기타 천연 물질을 포함할 수 있다.

[0215] 정제는, 필요에 따라, 하나 이상의 보조 성분을 포함하여, 압축 또는 성형에 의해 제조할 수 있다. 압축 정제는 자유-유동 형태(예, 분말 또는 과립)의 활성 화합물을, 필요에 따라, 보조 성분(예, 바인더, 윤활제, 불활성 희석제, 표면 활성 또는 분산제)과 혼합하여, 적합한 기계로 압축하여 제조할 수 있다. 성형 정제는 분말상 활성 화합물과 적합한 담체의 혼합물을 적합한 기계로 성형하여 제조할 수 있다.

[0216] 시럽 또는 현탁액은 당(예, 슈크로스)의 농축된 수성 용액에 활성 화합물을 첨가하여 제조할 수 있으며, 보조 성분을 더 첨가해도 좋다. 이러한 보조성분으로는 향료, 당 결정화 방지제, 또는 다른 성분의 용해도를 증가시키는 약제(예, 글리세롤 또는 소르비톨 등의 다가알콜)를 들 수 있다.

[0217] 직장 투여용 제제는 좌약의 주재료(base)로서 종래의 담체, 예를 들면, 코코아 버터 또는 Witepsol S55™(Dynamite Nobel Chemical, Germany)를 함유한, 좌약 형태로서 존재할 수 있다.

[0218] 경구 투여용 제제는 증강제를 함유해도 좋다. 경구적으로 허용가능한 흡수 증강제로는 소디움라우릴설페이트, 팔미토일 카르니틴, 라우레스-9, 포스파티딜콜린, 시클로덱스트린 및 그 유도체 등의 계면활성제; 소디움 데옥시콜레이트, 소디움 타우로콜레이트, 및 소디움 푸시데이트 등의 담즙염; EDTA, 시트르산 및 살리실레이트 등의 킬레이트제; 및 올레산, 라우르산, 아실카르니틴, 모노 및 디글리세라이드 등의 지방산을 들 수 있다. 기타 경구 흡수 증강제로는 염화벤잘코늄, 염화벤즈에토늄, CHAPS(3-(3-콜아미도프로필)-디메틸암모니오-1-프로판설포네이트), Big-CHAPS(N,N-비스(3-D-글루콘아미도프로필)-콜아미드), 클로로부탄올, 옥톡시놀-9, 벤질알콜, 페놀, 크레졸, 및 알킬알콜을 들 수 있다. 본 발명의 특히 바람직한 경구 흡수 증강제는 소디움라우릴설페이트이다.

[0219] 필요에 따라, 상기 화합물은 리포솜(liposome) 또는 미소구체(microsphere) (또는 미세입자(microparticle)) 형태로 투여해도 좋다. 환자 투여용 리포솜 및 미소구체를 제조하는 방법은 당업자에게 잘 알려져 있으며, 미국 특허번호 4,789,734에 리포솜내에 생물학적 물질을 캡슐화하는 방법이 참고로 기재되어 있다. 특히, 상기 물질은 수성 용액에 용해되어 있으며, 필요에 따라, 계면활성제와 함께 적합한 인지질 및 지질이 첨가되며, 상기 물질은, 필요에 따라, 투석 또는 초음파 처리된다. 공지의 방법의 재검토는 G. Gregoriadis, Chapter 14, "리포솜" *Drug Carries in Biology and Madicine*, pp. 287-341(Academic Press, 1979)에 의해 제공된다.

[0220] 폴리머 또는 단백질로 형성된 미소구체는 당업자에게 잘 알려져 있으며, 위장관 경로를 통해 직접 혈류로 통과하도록 맞춰질 수 있다. 필요에 따라, 상기 화합물은 미소구체, 또는 미소구체 복합체에 삽입할 수 있으며, 수일에서 수개월간에 걸쳐 천천히 방출되도록 이식(implant)할 수 있다(참고, 미국 특허번호 4,906,474, 4,926,673 및 3,625,214, 및 Jein, TIPS 19:155-157(1998)).

[0221] 하나의 태양에서, 본 발명의 티로신 키나제 저해제는, 정맥내 투여 후 모세혈관상에서 머무르기에 적합한 사이즈로 한 리포솜 또는 미세입자내로 제제화시킬 수 있다. 상기 리포솜 또는 미세입자가 허혈조직(ischemic tissue) 주변의 모세혈관상에 머무르면, 상기 약제는 그들이 가장 유효할 수 있는 부위에 국소적으로 투여될 수 있다. 허혈조직을 표적으로 하는 적합한 리포솜은 통상 약 200나노미터보다 작고, 또한 일반적으로 단층소포체

(unilamella vesicles)이다(참고, 예를 들면, Baldeschweiler에 의한 "Liposomal targeting of ischemic tissue"의 제목의 미국 특허번호 5,593,688에 기재).

[0222] 바람직한 미세입자는 폴리글리콜라이드, 폴리락티드 및 그들의 공중합체 등의 생분해성 폴리머로 제조한 것이다. 당업자는 원하는 약제 방출 속도 및 원하는 투여량을 포함한 각종 인자에 따라 적합한 담체 시스템을 쉽게 결정할 수 있다.

[0223] 하나의 태양에서, 상기 제제는 카테터(catheter)를 통해서 직접 혈관 내부에 투여된다. 이 투여는, 예를 들면, 카테터내의 구멍을 통해서 일어날 수 있다. 활성 화합물이 비교적 긴 반감기(약 1일~1주일)를 가진 태양에서는, 상기 제제는 미국 특허번호 5,410,016에 Hubbell et al.이 개시한 바와 같은, 생분해성 고분자 히드로겔에 포함될 수 있다. 이 고분자 히드로겔은 조직 루멘의 내부에 전달되어, 상기 폴리머의 분해에 따라 경시적으로 상기 활성 화합물을 방출된다. 원한다면, 상기 고분자 히드로겔은 그 내부에 활성 화합물이 분산된 미세입자 또는 리포솜을 가짐으로써, 활성 화합물의 제어 방출에 대한 또다른 메카니즘을 제공할 수 있다.

[0224] 상기 제제는, 적합하게는, 단위 투여 형태로 존재할 수 있으며, 제약학 분야의 당업자에게 공지된 방법 중 어느 하나로 제조할 수 있다. 모든 방법은 하나 이상의 보조성분을 구성하는 담체와 활성 화합물을 혼합하는 단계를 포함한다. 일반적으로, 상기 제제는 활성 화합물과 액체 담체 또는 미세하게 쪼개진 고체 담체를 균일하게 잘 혼합함에 의해 제조하고, 그 후, 필요에 따라, 생성물을 원하는 단위 투여 형태로 형상화한다.

[0225] 상기 제제는 종래의 약제학적 제제에 이용되던 하나 이상의 부가적인 보조성분(예, 희석제, 완충액, 향료, 바인더, 표면 활성화제, 점증제, 윤활제, 현탁용제, 보존제(항산화제 포함) 등)을 더 포함할 수 있다.

[0226] 본 발명의 방법의 화합물은, 호흡기도 투여용으로, 스너프(snuff) 또는 분무기용 에어로졸 또는 용액으로서, 또는 흡입용 미세 분말로서, 단독으로 또는 락토스 등의 불활성 담체와 조합하여 제공될 수 있다. 이러한 경우에 활성 화합물의 입자는 적합하게는 50마이크론 미만의 직경을 갖고, 바람직하게는 10마이크론 미만, 더 바람직하게는 2~5마이크론의 직경을 갖는다.

[0227] 통상, 비강 투여용으로는 약산의 pH가 바람직하다. 본 발명의 조성물의 pH로는 약 2~5가 바람직하고, 약 pH 3.5~3.9가 더 바람직하고, pH 3.7이 가장 바람직하다. pH의 조정은 염산 등의 적당한 산을 첨가하여 행한다.

[0228] 그 내부에 용해 또는 분산된 활성 성분을 함유하는 약리학적 조성물의 제조는 당업계에 잘 알려져 있으며, 제제화에 의거하여 제한할 필요는 없다. 일반적으로 이러한 조성물은 액체 용액 또는 현탁액의 어느 하나로서 주입할 수 있도록 제조되지만, 사용하기 전에 용액 또는 현탁액으로 하기에 적합한 고체 형태로 제조할 수도 있다. 상기 제제는 유상으로 제조할 수도 있다.

[0229] 상기 활성 성분은, 약제학적으로 허용가능하고 활성 성분과 양립할 수 있는 부형제(excipient)와, 본 발명의 치료법에 사용하기 적합한 양으로 혼합할 수 있다. 적합한 부형제로는, 예를 들면, 물, 식염수, 덱스트로스, 글리세롤, 에탄올 등 및 그들의 조합물이 있다. 또한, 필요에 따라, 상기 조성물은 흡습제 또는 유화제, pH 완충제 등 활성성분의 유효성을 증강시키는 소량의 보조성분(auxiliary substance)을 함유할 수 있다.

[0230] 본 발명의 키나제 저해제는 그 안에 약제학적으로 허용가능한 구성성분의 염을 포함할 수 있다. 약제학적으로 허용가능한 염으로는, 예를 들면, 염산 또는 인산 등의 무기산, 또는 아세트산, 타르타르산, 만델릭산 등의 유기산으로 형성되는 산 부가염(acid addition salt)(폴리펩티드의 유리 아미노기로 형성됨)을 들 수 있다. 또한, 유리 카르복실기로 형성된 염은, 예를 들면, 나트륨, 칼륨, 암모늄, 칼슘 또는 수산화제2철 등의 무기 염기, 및 이소프로필아민, 트리메틸아민, 2-에틸아미노에탄올, 히스티딘, 프로카인 등의 유기 염기로부터 유도될 수 있다.

[0231] 생리학적으로 허용가능한 담체는 당업계에 잘 알려져 있다. 액체 담체의 예로는, 상기 활성 성분 및 물 외에는 어떠한 재료도 함유하지 않는 멸균 수성 용액, 또는 생리학적 pH 값의 인산나트륨, 인산염-완충 식염수 등의 생리학적 식염수 또는 둘다 등 완충액을 함유하는 멸균 수성 용액이 있다. 또한, 수성 담체는 하나 이상의 완충액 염은 물론, 염화나트륨 및 염화칼륨 등의 염류, 덱스트로스, 폴리에틸렌글리콜 및 기타 용질(solute)을 함유할 수 있다.

[0232] 또한, 액체 조성물은 물을 제외하고도 액상(liquid phase)을 더 함유할 수 있다. 이러한 부가적인 액상의 예로는 글리세린, 면실유 등의 식물성 오일, 및 물-오일 에멀션이 있다.

[0233] **돌연변이 예측**

[0234] 또 다른 태양에서, 본 발명은 티로신 키나제 저해제로 치료한 후 erbB1 유전자에 존재하는 변이를 예측하는 방법을 개시한다. 티로신 키나제에 의한 항암치료에 대한 반응 결과, 종종 그 화합물 또는 다른 유사 화합물에 대한 내성이 생긴다는 것은 일반적으로 공지된 사실이다. 이러한 내성은 약제 표적, 예를 들면, EGFR 내에서의 돌연변이의 획득을 통해 생기는 것으로 생각된다. 이러한 돌연변이 예측(및 선별) 능력은 더 나은 치료 옵션 및 더 낮은 재발을 참작해야한다.

[0235] 본 발명의 하나의 태양에서는, 제피티니브(또는 유사한 EGFR 표적 치료제)에 반응하였으나 뒤이어 재발한 암환자의 종양 시료로부터, EGFR 키나제 도메인을 코딩하는 DNA를 단리하여 염기서열을 분석한다. 이러한 환자들에서의 재발은 EGFR 키나제 도메인내에서의 2차 돌연변이의 획득을 내포하는 것으로 예상된다. 그 다음, 본 발명에 개시한 방법을 사용하여, 이러한 새롭게 규정된 돌연변이를 표적으로 하며 그 키나제 활성을 저해하는 화합물을 동정한다. 이러한 화합물은 단독으로 또는 다른 공지의 EGFR 표적 치료제와 조합하여, EGFR 키나제 도메인에 1차 또는 2차(상기와 같은) 돌연변이를 가진 환자의 치료에 사용할 수 있다.

[0236] 하나의 태양에서는, 시험관내(*in vitro*)에서, EGFR 키나제(촉매) 도메인내에 존재하는 변이를 예측한다. 이 방법에서는, 인간 암 세포주에서 동정된 키나제 도메인 돌연변이를 포함하는 cDNA로, 세포(예, 섬유아세포)를 안정적으로 트랜스펙션시킨다. 예를 들면, 상기 세포는 도 4A에 기재된 바와 같은 서열번호 495 등의 돌연변이를 보유하고 있는 EGFR에 의해, 또는 다수의 동정 또는 미동정 키나제 도메인-변이 EGFR에 의해 트랜스펙션시킬 수 있다. 키나제 도메인-변이 EGFR로 세포를 트랜스펙션시키면, 배양시 세포의 비정상적인 증식이 일어날 것이다. 안정한 트랜스펙션 방법은 당업자에게 공지되어 있으며, 또한 본 명세서에 참고로 기재한, Current Protocols in Molecular Biololgy(F. M. Ausubel, R Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, K. Struhl 및 V. B. Chanda(편집), John Wiley & Sons., 2004)에 규정되어 있다. 그 다음, 상기 트랜스펙트된 세포에, 세포 증식을 저해할 것으로 예상되는 약제, 바람직하게는 티로신 키나제 저해제를 준치사량의 유효 투여량으로 투여한다. 바람직한 태양에서는, 상기 약제는 아닐리노퀴나졸린, 합성 아닐리노퀴나졸린, 제피티니브 또는 에를로티니브이다. 상기 세포를 약제의 존재 하에 연속적으로 배양하여(passage), 생존한 서브클론을 선별한다. 많은 세대를 거쳐서 생존한 세포(즉, 화합물에 내성이 있는 세포)에 대해서, erbB1 유전자에 존재하는 변이에 대한 선별 및 분석을 행한다. 그 결과, 2차 돌연변이는, 생체내(*in vivo*)에서 티로신 키나제 저해제로 반복 치료한 후에 일어나는 것으로 예측할 수 있다.

[0237] 필요에 따라서, 세포를, 인간 NSCLC 세포주(예, NCI-1650 및 NCI-1975) 유래의 제피티니브 내성 돌연변이 cDNA로 트랜스펙션시킨다. 각 세포주는 EGFR의 키나제 도메인내에 이형접합성 돌연변이를 가지고 있으므로, 제피티니브에 대해 민감성일 것으로 예측된다. 상기 NCI-1650에서의 EGFR 돌연변이는 엑손 19의 2235-2249 위치에서의 15개의 누클레오티드의 인프레임 결실(in-frame deletion)(delLE746-A750)에 의해 이루어지는 한편, NCI-1975는 엑손 21내의 누클레오티드 2573의 T를 G로 치환한 미스센스 돌연변이(L858R)를 갖는다. 여기서 나타낸 바와 같이, NCI-H1975에서의 상기 L858R 돌연변이는 활성화하여, 시험관내에서 제피티니브에 대한 증가된 민감성을 부여한다. EGFR 키나제 도메인 돌연변이를 보유하고 있는 다른 암 세포주를 이용할 수도 있다. 상기 암 세포주로는 폐암은 물론 이러한 돌연변이를 보유하는 것으로 알려진 다른 암을 들 수 있다.

[0238] 상기 세포는 세포의 2차 돌연변이 획득 빈도를 증가시키기 위해서 돌연변이원(mutagen)으로 처리해도 좋다. 돌연변이원은 투여량 처방, 전달 모드, 및 돌연변이원 투여 시의 유기체 또는 세포의 발육 단계에 따라 다른 빈도로 돌연변이를 유발할 수 있으며, 이 모든 파라미터는 다른 돌연변이원 또는 돌연변이유발 기술에 대한 종래 기술에 개시되어 있다. 상기 돌연변이원은 에틸메탄설포네이트(EMS), N-에틸-N-니트로소우레아(MNU) 등의 알킬화제일 수 있다. 필요에 따라서, 상기 돌연변이원은, 예를 들면, 포카르박신 히드로클로라이드(Prc), 메틸메탄설포네이트(MeMS), 클로르암부실(Chl), 멜팔란, 포르카르바진 히드로클로라이드, 시클로포스파미드(Cp), 디에틸설페이트(Et_2SO_4), 아크릴아미드 모노머(AA), 트리에틸렌 멜라민(TEM), 니트로겐 머스타드, 빈크리스틴, 디메틸니트로소아민, N-메틸-N'-니트로-니트로소구아니딘(MNNG), 7,12-디메틸벤즈안트라센(DMBA), 에틸렌옥사이드, 헥사메틸포스포아미드, 비설판, 및 에틸메탄설포레이트(EtMs)일 수 있다.

[0239] 돌연변이원으로 세포를 처치하는 방법은, 본 명세서에 참고로 기재한, 예를 들면, 미국특허 6,015,670에 기재되어 있다. 돌연변이유발 후, 세포(즉, EGFR 변이체로 트랜스펙트된 세포, 또는 인간 암 세포주 유래의 세포)를 제피티니브 보충 배지에서 배양하여, 성장한 내성 클론을 선별한다. 개개의 클론을 계대배양한 후, 예를 들면, EGFR 키나제 도메인에 대응하는 게놈 DNA를 특이적 PCR-매개 증폭하여, 그 EGFR 유전자의 누클레오티드 염기서열을 결정한다.

[0240] 또 다른 태양에서는, 제피티니브 내성을 부여하는, EGFR 유전자내의 자발적으로 획득된 돌연변이를 선별하기 위

해서, 세포(EGFR 변이를 가진 세포)를 제피티니브(또는 유사한 티로신 키나제 저해제)의 농도를 점차 증가시킨 배지에서 수주 또는 수개월에 걸쳐 연속 배양한다. 선별된 세포(상대적으로 높은 제피티니브 농도에서 지속적으로 증식한 세포)는 콜로니로서 단리할 수 있으며, 돌연변이는 상기와 같이 동정할 수 있다. 따라서, 이러한 변이는 생체내에서 티로신 키나제 저해제로 반복하여 처리한 후에 발생하는 것으로 예상할 수 있다(참조, Scappini et al., *Cancer*, April 1, 2004, Vol. 100, pg.1459.)

[0241] 또 다른 태양에서, EGFR 유전자의 변이형은 그것을 안정하게 선별된 세포주에 재도입하기 전에 DNA 수복-결핍 박테리아 균주(DNA repair-deficient bacterial strain)내에서 증식시킬 수 있다. 이러한 박테리아에서의 복제는 돌연변이유발 빈도를 증가시킨다. 필요에 따라서, "오류가 일어나기 쉬운(error-prone)" PCR을 이용하여, 당업자에게 공지된 표준 방법을 사용하여, 시험관내에서 상기 클로닝된 EGFR DNA의 돌연변이 빈도를 증가시킬 수 있다.

[0242] 또 다른 태양에서는, 생체내(*in vivo*)에서, erbB1 유전자의 키나제 도메인내에 존재하는 변이를 예측한다. 예를 들면, 키나제 활성이 증가한 erbB1 유전자의 변이형을, 암 모델이 되는 동물(즉, 마우스)에 트랜스펙션시킨다. 그 다음, 상기 동물에, 화합물, 바람직하게는 아닐리노퀴나졸린, 합성 아닐리노퀴나졸린, 제피티니브 또는 에를로티니브를 유효 투여량으로 투여한다. 상기 화합물에 반복 노출시켰을 때, 상기 암은 처음에는 저해된다. 이러한 화합물로 치료한 인간에서와 마찬가지로, 동물에서의 종양 세포도 이러한 치료에 내성을 부여하는 돌연변이를 획득한다. 본 발명의 방법은 이러한 내성 종양의 erbB1 유전자의 단리 및 특성화를 고려한다. 이러한 새롭게 특성화된 변이를 특이적으로 표적으로 하는 화합물은 이러한 변이된 erbB1 유전자를 보유할 것으로 짐작되는 환자의 치료에 유용하다. 이러한 환자들로는, 예를 들면, 티로신 키나제 저해제에 의한 치료에 초기에는 반응하지만, 그 후에는 동일 또는 유사한 화합물에 대한 반응에 실패한 환자를 들 수 있다.

[0243] 동물 모델을 만드는 방법은 당업자에게 공지되어 있으며, 또한 예를 들면, 본 명세서에 참고로 기재한 Ohashi et al., *Cell*, 65:305-317(1991); Adams et al., *Nature*, 325;223-228(1987); 및 Roman et al., *Cell*, 61:383-396(1990)에 기재되어 있다. 수정된 난모세포의 경우, 바람직한 유전자 도입 방법은 미세주입(microinjection)에 의한 것인 반면(참조, Leder et al., 미국 특허번호 4,736,866 및 5,175,383), 배아줄기(ES)세포 경우의 바람직한 방법은 전기천공법(electroporation)이다. 하지만, 레트로바이러스 감염 등의 바이러스 전달 시스템, 또는 리포솜 융합을 포함한 다른 방법들도 사용할 수 있다. 상기 핵산의 단리 및 특성화는 상기 및 실시예에 기재되어 있다.

[0244] 상기에 기재한 erbB1 유전자에 존재하는 키나제 활성 증가 변이는, 본 발명의 방법을 사용하여, 환자에서 스크리닝(진단에 의해 또는 예후에 의해)할 수 있다. 또한, 이러한 돌연변이의 존재 또는 부재는 EGFR을 표적으로 하는 화합물(예, 티로신 키나제 저해제)에 의한 치료에 내성을 가진 환자를 결정하기 위한 기준으로서 사용될 수 있다.

[0245] 생체내 또는 시험관내에서 검출된, 이러한 새롭게 규정된 변이를 특이적으로 표적으로 하는 화합물은 당업자에게 공지된 방법 및 본 명세서에 기재한 방법을 사용하여 선별할 수 있다. 후보 약제 스크리닝법은 EGFR 변이형의 활성을 저해하는 생활성(bioactive) 후보 약제를 동정하기 위해 사용될 수 있다. 인간 세포에 대한 독성이 낮은 약제에 대한 스크리닝법이 특히 중요하다. 이러한 목적으로, 시험관내에서 표지된 단백질-단백질 결합 측정법, 전기영동 이동성 시프트 측정법, 효소활성 측정법, 단백질 결합에 대한 면역측정법 등을 포함한 다양한 측정법이 사용될 수 있다. 상기 정제된 돌연변이 EGFR 단백질은, 분자간 상호작용, 전달체 기능 등을 모델링하기 위해 사용될 수 있는 3차원적 결정(crystal) 구조를 결정하기 위해 사용될 수도 있다. 이러한 화합물은, 예를 들면, EGFR의 키나제 활성을 저해하는 티로신 키나제 저해제, 항체, 엡테이머(aptamer), siRNA, 및 벡터일 수 있다.

[0246] 또 다른 태양에서, 본 발명의 방법에 유용한 화합물은, 단일클론 항체, 사람 키메라 항체, 및 재조합 항체를 포함한, EGFR 변이체를 통하여 키나제 신호전달을 방해하는 항체 및 그들의 단편이며, 이들은 EGFR의 키나제 활성을 저해하는 능력에 의해 특징지어지며 독성이 낮다.

[0247] 항체의 중화(neutralizing)는, 토끼 또는 마우스 등의 동물에서는, 키나제 도메인에 적어도 하나의 핵산 변이가 존재하는 EGFR로 면역접종(immunization)함으로써 쉽게 일어난다. 면역접종한 마우스는, 대량의 항-EGFR 단일클론 항체를 생성하기 위해 배양되는, 하이브리도마(hybridoma) 제조용 B세포의 근원(source)을 제공하는데 특히 유용하다. 키메라 항체(chimeric antibody)는 다른 동물 종으로부터 유래하는 2개 이상의 세그먼트 또는 부분에 의해 특징지어지는 면역글로빈 분자이다. 일반적으로, 상기 키메라 항체의 가변부위(variable region)는 생쥐의 단일클론 항체 등, 인간을 제외한 포유동물의 항체로부터 유래하고, 면역글로빈의 불변부위(constant

region)는 인간 면역글로빈 분자에서 유래한다. 바람직하게는, 양 영역 및 상기 조합은 통상적으로 결정되는 바와 같이 낮은 면역원성(immunogenicity)을 갖는다. 인간화 항체(humanized antibody)는 생쥐의 항원 결합 부위는 그대로 유지한 채, 생쥐의 불변부위를 인간의 불변부위로 대체하는 유전공학 기술에 의해 만들어진 면역글로빈 분자이다. 그 결과 얻어진 마우스-인간 키메라 항체는 인간에 있어서 감소된 면역원성과 향상된 약물동태를 가져야 한다. 본 발명의 방법에 유용한, 고친화성 단일클론 항체 및 그들의 키메라 유도체의 바람직한 예는 유럽특허출원 EP 186,833; PCR 특허출원 WO 92/16553; 및 미국특허번호 6,090,923에 기재되어 있다.

[0248] 상기한 바와 같은, 기존에 존재하는 또는 새롭게 동정된 화합물은 1차 및/또또는 2차 EGFR 돌연변이를 보유하는 환자의 치료에 유용하다.

[0249] 바람직한 태양에서, 상기 화합물은 키나제 도메인에 적어도 하나의 변이가 존재하는 EGFR의 티로신 키나제 활성 저해제, 특히, 다른 티로신 키나제에 비해서 "변이된(mutated)" EGFR에서 선택적으로 작용하는 소분자 저해제이다.

[0250] EGFR의 저해제로는, 특히 제한은 없지만,

[0251] PID 153035, 4-(3-클로로아닐리노)퀴나졸린 등의 퀴나졸린류, 피리도피리미딘류, 피리미도피리미딘류, CGP 59326, CGP 60261 및 CGP 62706 등의 피롤로피리미딘류, 피라졸로피리미딘류, 4-(페닐아미노)-7H-피롤로[2,3-d]피리미딘류(Traxler et al.,(1996) *J. Med Chem* 39:2285-2292), 쿠르쿠민(디페룰로일메탄)(Laxmin arayana, et al., (1995), *Carcinogen* 16:1741-1745), 4,5-비스(4-플루오로아닐리노)프탈이미드(Buchdunger et al.(1995) *Clin. Cancer Res.* 1:813-821; Dinney et al.(1997) *Clin. Cancer Res.* 3:161-168) 등의 티로신 키나제 저해제;

[0252] 니트로티오펜 부위(moiety)를 함유하는 타이르포스틴류(tyrphostins) (Brunton et al. (1996) *Anti Cancer Drug Design* 11:265-296);

[0253] 프로테인 키나제 저해제 ZD-1839(AstraZeneca);

[0254] CP-358774(Pfizer, Inc.);

[0255] PD-01 83805(Warner-Lambert), EKB-569(Torrance et al., *Nature Medicine*, Vol.6, No.9, Sept. 2000, p.1024), HKI-272 및 357(Wyeth); 또는

[0256] 국제특허출원 WO99/09016(American Cyanamid), WO98/43960(American Cyanamid), WO97/38983(Warender Labert), WO99/06378(Warender Labert), WO99/06396(Warender Labert), WO96/30347(Pfizer, Inc.), WO96/33978(Zeneca), WO96/33977(Zeneca), 및 WO96/33980(Zeneca) 등을 들 수 있다.

[0257] 또 다른 태양에서는, 안티센스 전략을 사용하여 EGFR 변이체의 키나제 활성을 저해할 수 있다. 이러한 접근법은, 예를 들면, 안티센스 핵산으로 특이적 mRNA를 매스킹(masking)하거나 또는 리보자임으로 절단(cleaving)함으로써, 그 특이적 mRNA의 번역(translation)을 블로킹하는 안티센스 핵산 또는 리보자임을 이용할 수 있다. 안티센스 기술의 일반적인 내용에 대해서는, 예를 들면, Antisense DNA 및 RNA(Cold Spring Harbor Laboratory, D. Melton, ed., 1988)를 참고한다.

[0258] 변이 EGFR 유전자 전사의 가역적 단시간 저해(reversible short inhibition)도 유용할 수 있다. 이러한 저해는 siRNA를 사용하여 성취할 수 있다. RNA 간섭(RNA interference: RNAi) 기술은 소간섭 RNA(small interfering RNA: siRNA) 등의 소 RNA 분자를 사용하여 유전자 발현을 방해한다. 이러한 기술은, 이번에는, RNAi가 식물에서 곤충, 포유동물에 이르기까지 많은 살아있는 생물의 대부분의 세포에서 유전자 발현을 저해(silencing)하는 자연적인 생물학적 메카니즘이라는 사실을 이용한다(McManus et al., *Nature Reviews Genetics*, 2002, 3(10) p.737). RNAi는, 유전자의 분자 중간체인 mRNA 카피가 파괴되는 것을 보장함으로써, 유전자가 기능적 단백질을 생성하는 것을 방해한다. siRNA는, 하기한 바와 같이, 그 형태 그대로(naked form) 또는 벡터에 도입된 형태로 사용될 수 있다. 또한, 돌연변이 EGFR 유전자 전사를 특이적으로 저해하는 엡테이머(aptamer)를 이용할 수 있다(참고, 예를 들면, 미국특허 6,699,843). 본 발명에 유용한 엡테이머는 SELEX 프로세스를 사용하여 동정할 수 있다(참고, 예를 들면, 미국특허출원 5,707,796, 5,763,177, 6,011,577, 5,580,737, 5,567,588, 및 5,660,985).

[0259] "안티센스 핵산" 또는 "안티센스 올리고누클레오티드"는 단일 가닥(single stranded) 핵산 분자로서, 세포질내 조건(cytoplasmic conditions)하에서 RNA 또는 DNA 분자 중의 상보적인 염기와 혼성화 시에 그 상보적인 염기의 역할을 저해한다. 상기 RNA가 mRNA 전사체인 경우, 상기 안티센스 핵산은 상대전사체 (countertranscript)이거

나 또는 mRNA를 간섭하는 상보적인 핵산(mRNA-interfering complementary nucleic acid)이다. 본원에서 사용되는 바와 같이, "안티센스"는, 광범위하게, RNA-RNA 상호작용, RNA-DNA 상호작용, 리보자임, RNAi, 엡테이머 및 Rnase-H 매개 정지(Rnase-H mediated arrest)를 포함한다.

[0260] 리보자임은 DNA 제한 엔도뉴클리아제(restriction endonuclease)와 다소 유사한 방식으로 다른 단일가닥 RNA 분자를 특이적으로 절단할 수 있는 능력을 가진 RNA 분자이다. 리보자임은, 그들 자신의 인트론을 잘라내는 능력을 가진 특정한 mRNA의 관찰에서 발견되었다. 이러한 리보자임의 누클레오티드 염기서열을 수식(modify)함으로써, 연구자들은 RNA 분자 중의 특이적 누클레오티드 염기서열을 인식하여 그들을 절단하는 분자를 설계할 수 있었다(Cech, 1989, *Science* 245 (4915) p.276). 이들은 염기서열-특이적이기 때문에, 특정 염기서열을 가진 mRNA만을 불활성화시킨다.

[0261] 안티센스 핵산 분자는 세포내 발현용 재조합 유전자에 의해 코딩될 수 있고(예, 미국 특허번호 5,814,500; U.S.5,811,234), 또는 필요에 따라서 합성하여 제조할 수도 있다(예, 미국 특허번호 5,780,607).

[0262] 또한, 본 발명은 암에 걸린 환자, 특히, EGFR의 키나제 도메인내에 적어도 하나의 핵산 변이가 존재하는 환자의 치료 방법을 제공한다. 상기 치료법은 적합한 시간창(time window)내에 siRNA-함유 조성물을 환자에게 투여하는 단계를 포함한다. 상기 siRNA는 화학적으로 합성해도 좋으며, 시험관내 전사(in vitro transcription) 등을 이용하여 생성할 수 있다. 또한, 상기 siRNA 분자는 환자의 종양에서 동정된 돌연변이에 정확하게 대응하는 방식으로 개개의 환자에 맞춰 제작할 수 있다. siRNA는 단지 단일 누클레오티드에 차이가 있는, 누클레오티드 염기서열 간을 구별할 수 있기 때문에, 단일 누클레오티드 치환 또는 수개의 누클레오티드의 소결실 중 어느 하나와 관련된 EGFR 유전자의 변이형(둘다 여기서 기재한 바와 같이 종양에서 동정된 것임)을 특이적으로 표적으로 하는 siRNA를 설계할 수 있다. siRNA는 Brummelkamp et al., *Science* 296;550-553, 2002; Jaque et al., *Nature* 418;435-438, 2002; Elbashir S.M. et al.(2001), *Nature*, 411:494-498; MaCaffrey et al.,(2002), *Nature*, 418:38-39; Xia H. et al.(2001), *Nat. Biotech.* 20:1006-1010; Novia et al. (2002), Nat. Med. 8:681-686; 및 미국 특허출원번호 20030198627에 기재되어 있다.

[0263] 변이된 수용체 및 정상 수용체 둘다를 저해하는 제피티니브 등의 약제의 사용과 관련된, 이러한 치료 전략의 중요한 이점은, 상기 변이된 EGFR에 대해 특이적으로 검출된 siRNA가 야생형 EGFR을 저해하지 않는다는 점이다. 이것은, 설사 및 피부염을 포함한 제피티니브 치료의 "부작용"이, EGFR의 기능이 필요한 정상조직에서의 EGFR 저해 결과로서 일어난다고 일반적으로 알려져 있기 때문에, 중요하다.

[0264] siRNA의 종양으로 전달은 현재 이용할 수 있는 유전자 전달 "운반체 (vehicle)" 중 어느 하나를 통해 성취할 수 있다. 이들로는 아데노바이러스, 렌티바이러스, 단순포진 바이러스, 우두 바이러스, 및 레트로바이러스 등의 바이러스 벡터 뿐만 아니라, 화학적 매개 유전자 전달 시스템(예, 리포솜), 또는 기계적인 DNA 전달 시스템(DNA guns)을 들 수 있다. 이러한 siRNA-매개 유전자 발현 저해용으로 발현되는 올리고누클레오티드는 길이가 18~28개인 누클레오티드이다.

[0265] 또 다른 태양에서, 상기 화합물은 키나제 도메인에 적어도 하나의 변이가 존재하는 EGFR을 코딩하는 인간 염기서열에 특이적인 안티센스 분자이다. 상기 투여 치료제는 안티센스 올리고누클레오티드, 특히, 원래의 핵산으로부터 화학적 수식이 있는 합성 올리고누클레오티드, 또는 그러한 안티센스 분자를 RNA로서 발현하는 핵산 구조체일 수 있다. 상기 안티센스 염기서열은 표적 EGFR 유전자의 mRNA에 상보적이며, 표적 유전자 산물의 발현을 저해한다(참고, Nyce et al. (1997) *Nature* 385:720). 안티센스 분자는, Rnase H의 활성화 또는 입체장애(steric hinderance)를 통해, 번역(translation)에 유용한 mRNA의 양을 감소시킴으로써, 유전자 발현을 저해한다. 안티센스 분자 한개 또는 조합물이 투여될 수 있으며, 이 안티센스 분자 조합물의 경우, 단일 표적 유전자와는 상이한 다중 염기서열, 또는 여러 상이한 유전자에 상보적인 염기서열을 포함할 수 있다.

[0266] 바람직한 표적 유전자는 그 키나제 도메인에 적어도 하나의 핵산 변이가 존재하는 EGFR이다. 상기 유전자 염기서열은, 예를 들면, 도 5A-5D 등에 기재되어 있다. 일반적으로, 상기 안티센스 염기서열은 동물 숙주와 동일한 종 기원을 갖는다.

[0267] 안티센스 분자는 적합한 벡터 중의 표적 유전자 염기서열 전부 또는 일부의 발현에 의해 생성할 수 있으며, 이때, 상기 벡터는 표적 세포에 도입되어 발현된다. 전사 개시는 안티센스 가닥이 RNA 분자로서 생성되도록 맞춰진다.

[0268] 상기 안티센스 RNA는 내인성 센스 가닥 mRNA와 혼성화하여, 표적 유전자의 발현을 막는다. 상기 원래의 전사 개시 영역, 또는 외인성 전사 개시 영역이 사용될 수 있다. 프로모터(promoter)는 시험관내에서 재조합법에 의

해 도입되거나, 또는 상기 염기서열의 염색체로의 상동 삽입(homologous integration)의 결과로서 도입될 수 있다. 근육세포에서 활성이 있는 다수의 강력한 프로모터는 당업자에게 공지되어 있으며, 이-엑틴 프로모터, SV40 초기 및 후기 프로모터(early and late promoter), 인간 사이토르네갈로바이러스(cytornegalovirus) 프로모터, 레트로바이러스 LTR 등을 들 수 있다. 전사 벡터는, 일반적으로, 핵산 염기서열의 삽입용으로 제공되는, 프로모터 서열 부근에 위치한 알맞은 제한 부위(restriction site)를 갖는다. 전사 카세트(cassettes)는 전사개시 부위, 표적 유전자 또는 그들의 단편, 및 전사종결 부위를 포함하도록 제조될 수 있다. 상기 전사 카세트는 플라스미드, 레트로바이러스(예, 렌티바이러스), 아데노바이러스 등의 다양한 벡터에 도입될 수 있으며, 이 벡터들은 세포내에서 일반적으로 적어도 거의 하루동안, 더 일반적으로는 거의 수일동안, 일시적으로 또는 안정적으로 유지될 수 있다.

[0269] 또한, 엡테이머도 유용하다. 엡테이머는 유망한 새로운 군의 치료 올리고누클레오티드 또는 펩티드이며, 시험관내에서 예를 들면 리간드 수용체 등의 높은 친화성을 가진 소정의 표적과 특이적으로 결합시켜 선별한다. 그들의 결합 특성은 분자내에서 핵염기 쌍을 이룸에 의해 함께 결합되어 3차원 구조를 형성하는 올리고누클레오티드의 능력을 반영한 것이다. 엡테이머는, 정상 및 수식될 수 있는 합성 DNA, RNA 또는 펩티드 서열(예, 펩티드 핵산(PNA), 티오인산화 DNA 등)로서, 표적 단백질 및 리간드(리피드, 탄수화물, 대사산물 등)와 상호작용한다. 다른 태양에서는, EGFR 변이체에 특이적인 RNA 엡테이머를 치료제로서 세포에 도입하여 발현시킬 수 있다.

[0270] 펩티드 핵산(PNA)은, 어떤 면에서는, 올리고누클레오티드 및 그 유사체와 유사한 화합물이며 따라서 모방(mimic) DNA 및 RNA일 수 있다. PNA에서는, 올리고누클레오티드의 데옥시리보스 골격(backbone)이 슈도-펩티드 골격으로 대체되어 있다(Nielsen et al. 1991 Science 254, 1457-1500). 각 서브유니트, 또는 모노머는 이 골격에 부착된 자연스럽게 발생한 또는 부자연스럽게 발생한 핵염기를 갖는다. 하나의 이러한 골격은 아미드 결합을 통하여 결합된 N-(2-아미노에틸)글리신의 반복 단위로 구성된다. PNA는 왓슨과 크릭의 염기쌍 및 나선구조 형성을 통해 상보적인 핵산과 혼성화한다. 상기 슈도-펩티드 골격은 더 뛰어난 혼성화 특성(Egholm et al., *Biochem. Pharmacol.*(1993)365, 566-568), 효소 분해에 대한 저항성(Demmidov et al. *Biochem. Pharmacol.*(1994) 48, 1310-1313) 및 다양한 화학적 수식에 대한 접근로(Nielsen and Haaima Chemical Society Reviews(1997) 73-78)를 제공한다. EGFR 변이체에 특이적인 PNA는 치료제로서 세포에 도입되어 발현될 수 있다.

[0271] EGFR 변이체를 표적으로 하는 화합물로 치료할 환자들로는, 예를 들면, EGFR에 1차 또는 2차 돌연변이가 존재하는 것으로 진단된 환자, 티로신 키나제 저해제에 의한 치료에 대해 처음에는 반응하지만 동일 또는 유사 화합물에 의한 후속 치료에 대해서는 반응에 실패한 환자를 들 수 있다. 필요에 따라서, 2차 EGFR 돌연변이를 표적으로 하는 화합물은, 병용치료로서, 1차 EGFR 돌연변이를 표적으로 하는 화합물(예를 들면, 제피티니브)과 함께 암환자에게 투여할 수도 있다. 1차 및 2차 EGFR 돌연변이 둘다를 표적으로 하는 화합물을 병용함으로써, 내성 가능성을 줄일 수 있다.

[0272] EGFR 티로신 키나제 저해제인 제피티니브, 에를로티니브 등을 포함한, 현재 공지된 항암치료제에 내성을 부여하는 부가적인 EGFR 돌연변이는 본 발명의 범위에 포함된다.

[0273] 내성 EGFR 변이체는, 이 키나제 영역에 높은 상동성을 가진 단백질을 포함한, 관련 티로신 키나제 도메인의 키나제 도메인에서 동정된 돌연변이와 유사한 돌연변이를 갖는 것으로 예상된다. 유사 단백질에서의 돌연변이를 기재하고 있는 논문으로는 BCR-ABL에 관련된 공지의 논문들을 들 수 있다(참조, 예를 들면, Bradford et al. *Blood*. 2003 Jul 1;102(1):276-83, Epub 2003 Mar 06; Hochhaus et al., *Leukemia*. 2002 Nov;16(11):2190-6; 및 Al-Ali et al., *Hematol J.* 2004;5(1):55-60).

[0274] 공지된 EGFR 티로신 키나제 저해제에 내성이 있는 EGFR 변이체로는, 이미티니브 내성 표현형을 나타내는(confirm) c-able(BCR-ABL) 잔기와 유사한 비야생형 잔기를 하나 이상의 위치에 가진, 어떠한 하나 이상의 EGFR 폴리펩티드 또는 그것을 코딩하는 누클레오티드를 들 수 있다. EGFR에 돌연변이가 일어났을 때 약제 내성을 부여하는 잔기로는 특히, 제한은 없지만, 예를 들면, P-루프 및 활성화 루프 등을 포함한, 키나제 도메인 유래의 잔기를 들 수 있으며, 상기 EGFR 폴리펩티드내의 변이된 잔기는 c-able 잔기와 유사하다. 고려된 내성 EGFR 변이체는 BCR-ABL 중의 적어도 Met 244, Leu 248, Gly 250, Gln 252, Tyr 253, Glu 255, Asp 276, Thr 315, Phe 317, Met 351, Glu 355, Phe 359, His 396, Ser 417, 및 Phe 486의 위치와 유사한 아미노산 위치에 비야생형 잔기를 갖는다(참조, 예를 들면 표 S3C 및 도 9). 이 BCL-ABL 잔기들은 EGFR의 Lys 714, Leu 718, Ser 720, Ala 722, Phe 723, Thr 725, Ala 750, Thr 790, Leu 792, Met 825, Glu 829, Leu 833, His 870, Thr 892, Phe 961 잔기에 각각 대응한다(참조, 예를 들면, 표 S3C, 도 9).

[0275] **예후 시험(Prognostic Testing)**

[0276] 본 발명의 방법은 암 발병의 예후 지표(prognostic indicator)로 사용된다. 필요에 따라, 본 발명은 존재하지만 아직 진단되지 않은 암 또는 검출할 수 없는 단계의 암을 검출하기 위해 사용된다. 암이 발병할 위험이 있는 환자는, 본 발명의 방법을 이용하여, 키나제 활성이 증가하는 erbB1 유전자내의 핵산 변이의 존재에 대해 스크리닝된다. erbB1 유전자의 키나제 도메인 내에서의 변이의 존재는 암이 존재하거나 또는 곧 생길 것을 나타낸다. 따라서, erbB1 유전자의 키나제 도메인 내에서의 변이의 존재는 환자가 EGFR 표적 치료에 유리함을 나타낸다. 여기서 기재한 바와 같이, EGFR 표적 치료법은 티로신 키나제 저해제에 의한 바람직한 치료법이다.

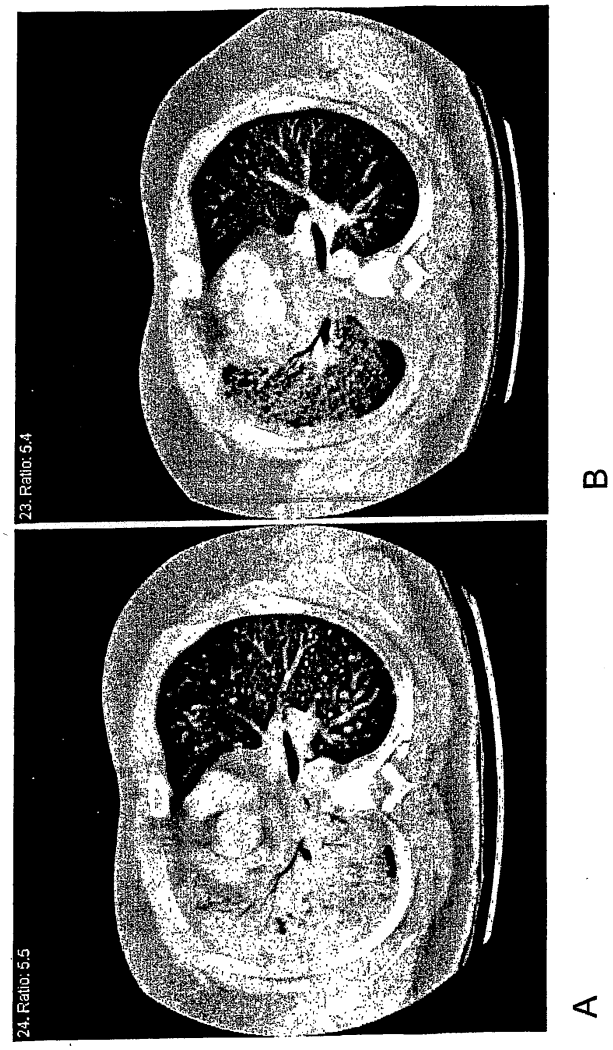
[0277] 본 발명의 바람직한 태양에서는, 환자로부터 생물학적 시료를 얻어, erbB1 유전자의 키나제 도메인에서의 핵산 변이의 존재 또는 부재에 대해 스크리닝한다. 상기 시료는, 예를 들면, 혀, 입, 뺨, 기관(trachea), 기관지튜브, 폐 등에서 얻은 조직, 또는 가래나 폐흡인액으로부터 얻은 체액을 포함한, 환자로부터 얻은 어떠한 시료일 수 있다. 이러한 생물학적 표본을 얻는 방법은 당업자에게 잘 알려져 있다.

[0278] 따라서, 본 발명은, 피험자로부터 시험시료를 얻어 EGFR 변이체 단백질 또는 핵산(mRNA, 게놈 DNA)을 검출하여, EGFR 변이체의 비정상적인 발현 및 활성과 관련된 질병 또는 장애를 동정하는 방법을 제공하며, 이 방법에서, EGFR 변이체 단백질 또는 핵산의 존재는 EGFR 변이체의 발현 또는 활성과 관련된 질병 또는 장애를 가졌거나 또는 발병할 위험이 있는 피험자에 대한 진단이다. 여기서 사용되는 "시험시료"는 관심있는 피험자로부터 얻은 생물학적 시료이다. 예를 들면, 시험 시료는 생체액(예, 혈청), 세포 시료, 또는 조직, 특히 조직생검 시료일 수 있다.

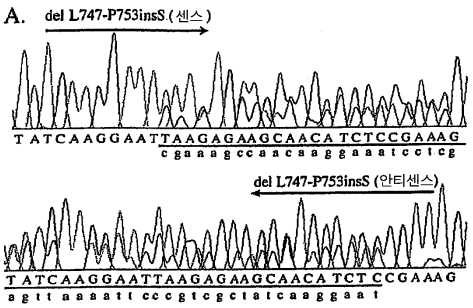
[0279] 또한, 본원에서 기재하는 예후 측정법은, EGFR 변이체의 비정상적인 발현 또는 활성과 관련된 질병 또는 장애를 치료하기 위해, 피험자에게 약제(예를 들면, 작용제, 길항제, 펩티도미메틱(peptidomimetic), 단백질, 펩티드, 핵산, 소분자, 또는 기타 약제 후보)를 투여해야하는지 여부를 결정하는 데 사용할 수 있다. 예를 들면, 이러한 방법은 피험자가 장애에 대해 약제로 효과적으로 치료할 수 있는지의 여부를 결정하는 데 사용된다. 따라서, 본 발명은, 피험자로부터 시험 시료을 얻어서 돌연변이 EGFR 단백질 또는 핵산을 검출함으로써, EGFR 변이체의 비정상적인 발현 및 활성과 관련된 질병에 대하여 약제를 사용하여 효과적으로 치료할 수 있는지 여부를 결정하는 방법을 제공한다(예, 돌연변이 EGFR 단백질 또는 핵산이 존재하는 경우, EGFR 변이체의 발현 또는 활성과 관련된 장애를 치료하기 위해 약제를 투여할 수 있는 환자로 진단한다.).

도면

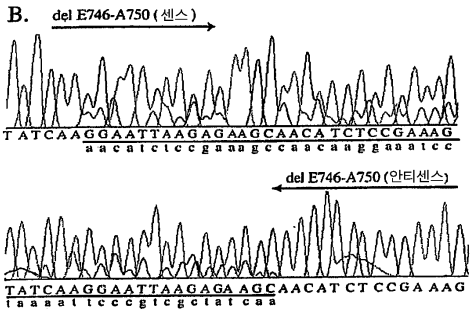
도면1


23. Ratio: 5.4
24. Ratio: 5.5
B
A

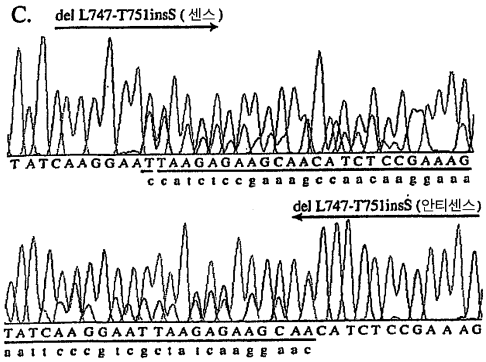
도면2A


A.
del L747-P753insS (센스)
TATCAAGGAATTAAGAGAAGCAACA TCTCCGAAAG
cgaaagccaacaaggaaatcctcg
del L747-P753insS (안티센스)
TATCAAGGAATTAAGAGAAGCAACA TCT CCGAAAG
agtt aaaatt cccgtcgctatcaaggaat

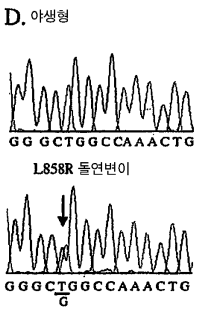
도면2B



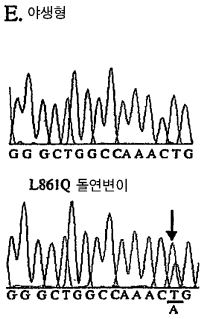
도면2C



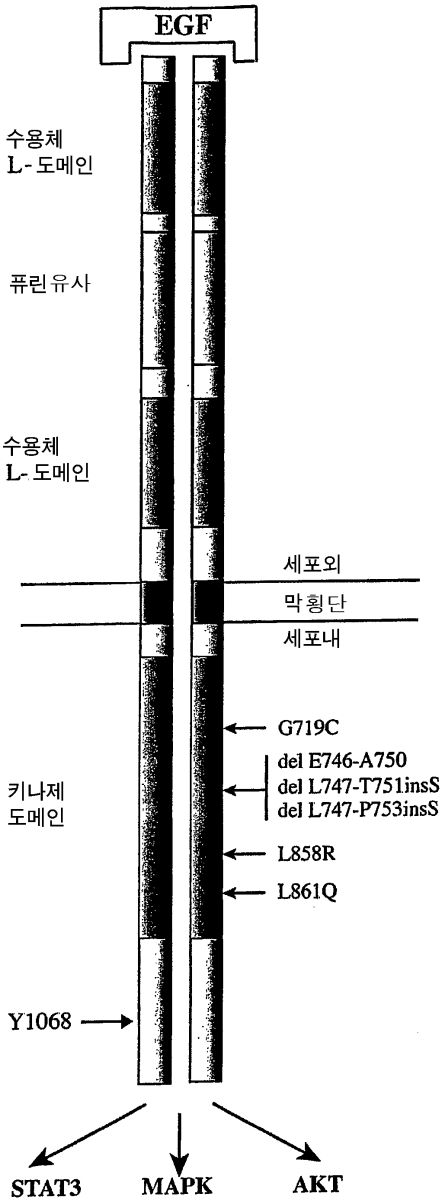
도면2D



도면2E

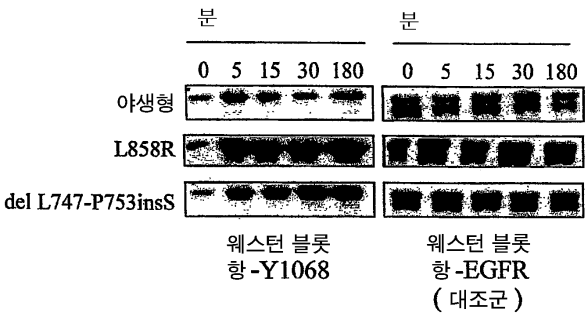

E. 야생형
GG GCTGGCCAAACTG
L861Q 돌연변이
GG GCTGGCCAAACTG
A

도면2F


EGF
수용체
L-도메인
퓨린유사
수용체
L-도메인
세포외
막횡단
세포내
G719C
del E746-A750
del L747-T751insS
del L747-P753insS
L858R
L861Q
키나제
도메인
Y1068
STAT3
MAPK
AKT

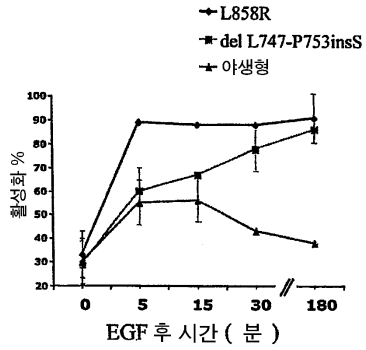
도면3A

A



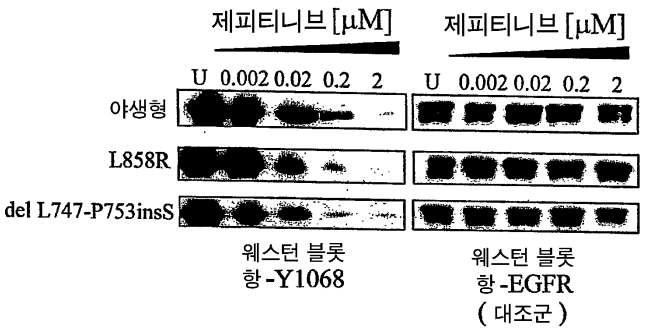
도면3B

B

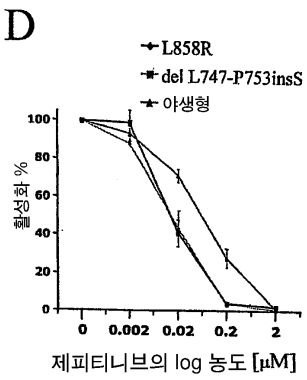


도면3C

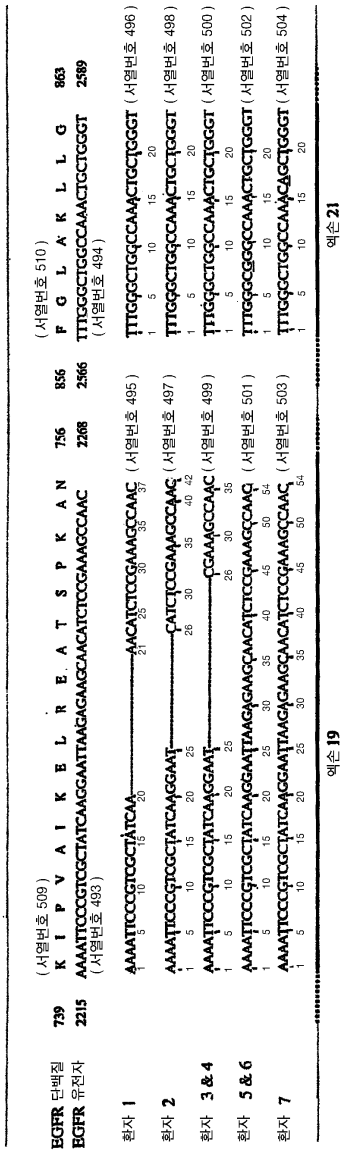
C



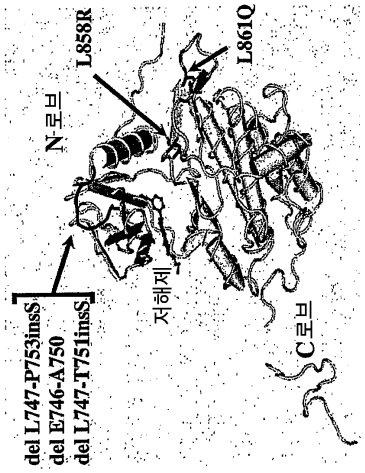
도면3D


D
L858R
del L747-P753insS
야생형
활성화 %
100
80
60
40
20
0
0
0.002
0.02
0.2
2
제피티니브의 log 농도 [μM]

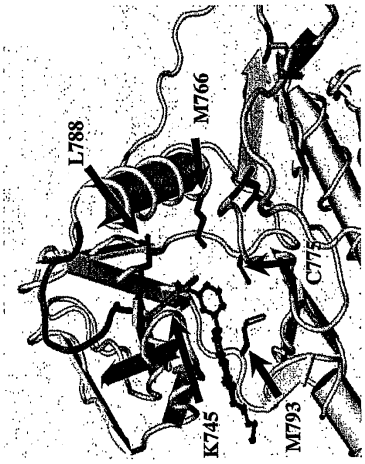
도면4A


EGFR 단백질 739 (서열번호 509) K I P V A I K E L R E A T S P K A N 756 856 (서열번호 510) F G L A K L L G 863
EGFR 유전자 2215 AAAATTCCCGTCGCTATCAAGGAATTAAGAGAAGCAACATCTCCGAAAGCCAAC 2268 2566 TTTGGCCTGGCCAAACTGCTGGGT 2589
(서열번호 493) (서열번호 494)
환자 1 AAAATTCCCGTCGCTATCAA----------AACATCTCCGAAAGCCAAC (서열번호 495) TTTGGCCTGGCCAAACTGCTGGGT (서열번호 496)
환자 2 AAAATTCCCGTCGCTATCAAGGAAT----------CATCTCCGAAAGCCAAC (서열번호 497) TTTGGCCTGGCCAAACTGCTGGGT (서열번호 498)
환자 3 & 4 AAAATTCCCGTCGCTATCAAGGAAT----------CGAAAGCCAAC (서열번호 499) TTTGGCCTGGCCAAACTGCTGGGT (서열번호 500)
환자 5 & 6 AAAATTCCCGTCGCTATCAAGGAATTAAGAGAAGCAACATCTCCGAAAGCCAAC (서열번호 501) TTTGGCCTGGCCAAACTGCTGGGT (서열번호 502)
환자 7 AAAATTCCCGTCGCTATCAAGGAATTAAGAGAAGCAACATCTCCGAAAGCCAAC (서열번호 503) TTTGGCCTGGCCAAACTGCTGGGT (서열번호 504)
엑손 19 엑손 21

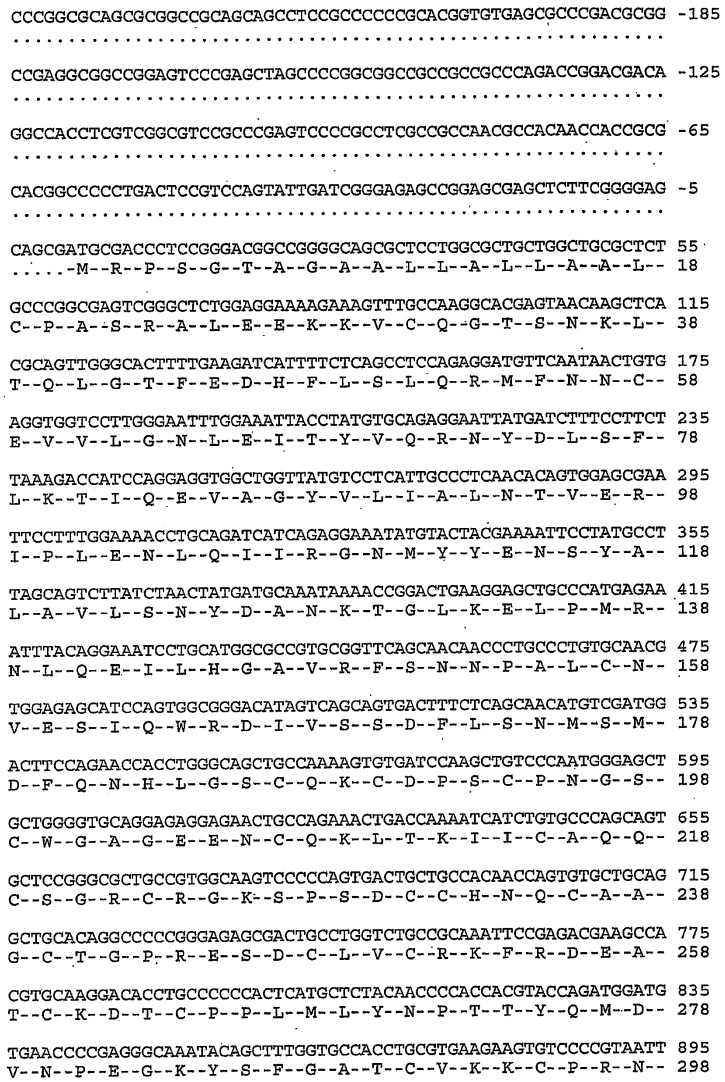
도면4B


L858R
L861Q
N-로브
del L747-P753insS
del E746-A750
del L747-T751insS
저해제
C-로브

도면4C


L788
M766
K745
M793

도면5A


CCCGGCGCAGCGCGGCCGCAGCAGCCTCCGCCCCCCGCACGGTGTGAGCGCCCGACGCGG -185
CCGAGGCGGCCGGAGTCCCGAGCTAGCCCCGGCGGCCGCCGCCGCCCAGACCGGACGACA -125
GGCCACCTCGTCGGCGTCCGCCCGAGTCCCCGCCTCGCCGCCAACGCCACAACCACCGCG -65
CACGGCCCCCTGACTCCGTCCAGTATTGATCGGGAGAGCCGGAGCGAGCTCTTCGGGGAG -5
CAGCGATGCGACCCTCCGGGACGGCCGGGGCAGCGCTCCTGGCGCTGCTGGCTGCGCTCT 55
.....-M--R--P--S--G--T--A--G--A--A--L--L--A--L--L--A--A--L-- 18
GCCCGGCGAGTCGGGCTCTGGAGGAAAAGAAAGTTTGCCAAGGCACGAGTAACAAGCTCA 115
C--P--A--S--R--A--L--E--E--K--K--V--C--Q--G--T--S--N--K--L-- 38
CGCAGTTGGGCACTTTTGAAGATCATTTTCTCAGCCTCCAGAGGATGTTCAATAACTGTG 175
T--Q--L--G--T--F--E--D--H--F--L--S--L--Q--R--M--F--N--N--C-- 58
AGGTGGTCCTTGGGAATTTGGAAATTACCTATGTGCAGAGGAATTATGATCTTTCCTTCT 235
E--V--V--L--G--N--L--E--I--T--Y--V--Q--R--N--Y--D--L--S--F-- 78
TAAAGACCATCCAGGAGGTGGCTGGTTATGTCCTCATTGCCCTCAACACAGTGGAGCGAA 295
L--K--T--I--Q--E--V--A--G--Y--V--L--I--A--L--N--T--V--E--R-- 98
TTCCTTTGGAAAACCTGCAGATCATCAGAGGAAATATGTACTACGAAAATTCCTATGCCT 355
I--P--L--E--N--L--Q--I--I--R--G--N--M--Y--Y--E--N--S--Y--A-- 118
TAGCAGTCTTATCTAACTATGATGCAAATAAAACCGGACTGAAGGAGCTGCCCATGAGAA 415
L--A--V--L--S--N--Y--D--A--N--K--T--G--L--K--E--L--P--M--R-- 138
ATTTACAGGAAATCCTGCATGGCGCCGTGCGGTTCAGCAACAACCCTGCCCTGTGCAACG 475
N--L--Q--E--I--L--H--G--A--V--R--F--S--N--N--P--A--L--C--N-- 158
TGGAGAGCATCCAGTGGCGGGACATAGTCAGCAGTGACTTTCTCAGCAACATGTCGATGG 535
V--E--S--I--Q--W--R--D--I--V--S--S--D--F--L--S--N--M--S--M-- 178
ACTTCCAGAACCACCTGGGCAGCTGCCAAAAGTGTGATCCAAGCTGTCCCAATGGGAGCT 595
D--F--Q--N--H--L--G--S--C--Q--K--C--D--P--S--C--P--N--G--S-- 198
GCTGGGGTGCAGGAGAGGAGAACTGCCAGAAACTGACCAAAATCATCTGTGCCCAGCAGT 655
C--W--G--A--G--E--E--N--C--Q--K--L--T--K--I--I--C--A--Q--Q-- 218
GCTCCGGGCGCTGCCGTGGCAAGTCCCCCAGTGACTGCTGCCACAACCAGTGTGCTGCAG 715
C--S--G--R--C--R--G--K--S--P--S--D--C--C--H--N--Q--C--A--A-- 238
GCTGCACAGGCCCCCGGGAGAGCGACTGCCTGGTCTGCCGCAAATTCCGAGACGAAGCCA 775
G--C--T--G--P--R--E--S--D--C--L--V--C--R--K--F--R--D--E--A-- 258
CGTGCAAGGACACCTGCCCCCCCACTCATGCTCTACAACCCCACCACGTACCAGATGGATG 835
T--C--K--D--T--C--P--P--L--M--L--Y--N--P--T--T--Y--Q--M--D-- 278
TGAACCCCGAGGGCAAATACAGCTTTGGTGCCACCTGCGTGAAGAAGTGTCCCCGTAATT 895
V--N--P--E--G--K--Y--S--F--G--A--T--C--V--K--K--C--P--R--N-- 298

도면5B

```
ATGTGGTGACAGATCACGGCTCGTGCGTCCGAGCCTGTGGGGCCGACAGCTATGAGATGG 955
Y--V--V--T--D--H--G--S--C--V--R--A--C--G--A--D--S--Y--E--M-- 318

AGGAAGACGGCGTCCGCAAGTGTAAGAAGTGCGAAGGGCCTTGCCGCAAAGTGTGTAACG 1015
E--E--D--G--V--R--K--C--K--K--C--E--G--P--C--R--K--V--C--N-- 338

GAATAGGTATTGGTGAATTTAAAGACTCACTCTCCATAAATGCTACGAATATTAAACACT 1075
G--I--G--I--G--E--F--K--D--S--L--S--I--N--A--T--N--I--K--H-- 358

TCAAAAACTGCACCTCCATCAGTGGCGATCTCCACATCCTGCCGGTGGCATTTAGGGGTG 1135
F--K--N--C--T--S--I--S--G--D--L--H--I--L--P--V--A--F--R--G-- 378

ACTCCTTCACACATACTCCTCCTCTGGATCCACAGGAACTGGATATTCTGAAAACCGTAA 1195
D--S--F--T--H--T--P--P--L--D--P--Q--E--L--D--I--L--K--T--V-- 398

AGGAAATCACAGGGTTTTTGCTGATTCAGGCTTGGCCTGAAAACAGGACGGACCTCCATG 1255
K--E--I--T--G--F--L--L--I--Q--A--W--P--E--N--R--T--D--L--H-- 418

CCTTTGAGAACCTAGAAATCATACGCGGCAGGACCAAGCAACATGGTCAGTTTTCTCTTG 1315
A--F--E--N--L--E--I--I--R--G--R--T--K--Q--H--G--Q--F--S--L-- 438

CAGTCGTCAGCCTGAACATAACATCCTTGGGATTACGCTCCCTCAAGGAGATAAGTGATG 1375
A--V--V--S--L--N--I--T--S--L--G--L--R--S--L--K--E--I--S--D-- 458

GAGATGTGATAATTTCAGGAAACAAAAATTTGTGCTATGCAAATACAATAAACTGGAAAA 1435
G--D--V--I--I--S--G--N--K--N--L--C--Y--A--N--T--I--N--W--K-- 478

AACTGTTTGGGACCTCCGGTCAGAAAACCAAAATTATAAGCAACAGAGGTGAAAACAGCT 1495
K--L--F--G--T--S--G--Q--K--T--K--I--I--S--N--R--G--E--N--S-- 498

GCAAGGCCACAGGCCAGGTCTGCCATGCCTTGTGCTCCCCCGAGGGCTGCTGGGGCCCGG 1555
C--K--A--T--G--Q--V--C--H--A--L--C--S--P--E--G--C--W--G--P-- 518

AGCCCAGGGACTGCGTCTCTTGCCGGAATGTCAGCCGAGGCAGGGAATGCGTGGACAAGT 1615
E--P--R--D--C--V--S--C--R--N--V--S--R--G--R--E--C--V--D--K-- 538

GCAACCTTCTGGAGGGTGAGCCAAGGGAGTTTGTGGAGAACTCTGAGTGCATACAGTGCC 1675
C--N--L--L--E--G--E--P--R--E--F--V--E--N--S--E--C--I--Q--C-- 558

ACCCAGAGTGCCTGCCTCAGGCCATGAACATCACCTGCACAGGACGGGGACCAGACAACT 1735
H--P--E--C--L--P--Q--A--M--N--I--T--C--T--G--R--G--P--D--N-- 578

GTATCCAGTGTGCCCACTACATTGACGGCCCCCACTGCGTCAAGACCTGCCCGGCAGGAG 1795
C--I--Q--C--A--H--Y--I--D--G--P--H--C--V--K--T--C--P--A--G-- 598

TCATGGGAGAAAACAACACCCTGGTCTGGAAGTACGCAGACGCCGGCCATGTGTGCCACC 1855
V--M--G--E--N--N--T--L--V--W--K--Y--A--D--A--G--H--V--C--H-- 618

TGTGCCATCCAAACTGCACCTACGGATGCACTGGGCCAGGTCTTGAAGGCTGTCCAACGA 1915
L--C--H--P--N--C--T--Y--G--C--T--G--P--G--L--E--G--C--P--T-- 638

ATGGGCCTAAGATCCCGTCCATCGCCACTGGGATGGTGGGGGCCCTCCTCTTGCTGCTGG 1975
N--G--P--K--I--P--S--I--A--T--G--M--V--G--A--L--L--L--L--L-- 658

TGGTGGCCCTGGGGATCGGCCTCTTCATGCGAAGGCGCCACATCGTTCGGAAGCGCACGC 2035
V--V--A--L--G--I--G--L--F--M--R--R--R--H--I--V--R--K--R--T-- 678
```

도면5C

```
TGCGGAGGCTGCTGCAGGAGAGGGAGCTTGTGGAGCCTCTTACACCCAGTGGAGAAGCTC 2095
L--R--R--L--L--Q--E--R--E--L--V--E--P--L--T--P--S--G--E--A-- 698

CCAACCAAGCTCTCTTGAGGATCTTGAAGGAAACTGAATTCAAAAAGATCAAAGTGCTGG 2155
P--N--Q--A--L--L--R--I--L--K--E--T--E--F--K--K--I--K--V--L-- 718

GCTCCGGTGCGTTCGGCACGGTGTATAAGGGACTCTGGATCCCAGAAGGTGAGAAAGTTA 2215
G--S--G--A--F--G--T--V--Y--K--G--L--W--I--P--E--G--E--K--V-- 738

AAATTCCCGTCGCTATCAAGGAATTAAGAGAAGCAACATCTCCGAAAGCCAACAAGGAAA 2275
K--I--P--V--A--I--K--E--L--R--E--A--T--S--P--K--A--N--K--E-- 758

TCCTCGATGAAGCCTACGTGATGGCCAGCGTGGACAACCCCCACGTGTGCCGCCTGCTGG 2335
I--L--D--E--A--Y--V--M--A--S--V--D--N--P--H--V--C--R--L--L-- 778

GCATCTGCCTCACCTCCACCGTGCAGCTCATCACGCAGCTCATGCCCTTCGGCTGCCTCC 2395
G--I--C--L--T--S--T--V--Q--L--I--T--Q--L--M--P--F--G--C--L-- 798

TGGACTATGTCCGGGAACACAAAGACAATATTGGCTCCCAGTACCTGCTCAACTGGTGTG 2455
L--D--Y--V--R--E--H--K--D--N--I--G--S--Q--Y--L--L--N--W--C-- 818

TGCAGATCGCAAAGGGCATGAACTACTTGGAGGACCGTCGCTTGGTGCACCGCGACCTGG 2515
V--Q--I--A--K--G--M--N--Y--L--E--D--R--R--L--V--H--R--D--L-- 838

CAGCCAGGAACGTACTGGTGAAAACACCGCAGCATGTCAAGATCACAGATTTTGGGCTGG 2575
A--A--R--N--V--L--V--K--T--P--Q--H--V--K--I--T--D--F--G--L-- 858

CCAAACTGCTGGGTGCGGAAGAGAAAGAATACCATGCAGAAGGAGGCAAAGTGCCTATCA 2635
A--K--L--L--G--A--E--E--K--E--Y--H--A--E--G--G--K--V--P--I-- 878

AGTGGATGGCATTGGAATCAATTTTACACAGAATCTATACCCACCAGAGTGATGTCTGGA 2695
K--W--M--A--L--E--S--I--L--H--R--I--Y--T--H--Q--S--D--V--W-- 898

GCTACGGGGTGACTGTTTGGGAGTTGATGACCTTTGGATCCAAGCCATATGACGGAATCC 2755
S--Y--G--V--T--V--W--E--L--M--T--F--G--S--K--P--Y--D--G--I-- 918

CTGCCAGCGAGATCTCCTCCATCCTGGAGAAAGGAGAACGCCTCCCTCAGCCACCCATAT 2815
P--A--S--E--I--S--S--I--L--E--K--G--E--R--L--P--Q--P--P--I-- 938

GTACCATCGATGTCTACATGATCATGGTCAAGTGCTGGATGATAGACGCAGATAGTCGCC 2875
C--T--I--D--V--Y--M--I--M--V--K--C--W--M--I--D--A--D--S--R-- 958

CAAAGTTCCGTGAGTTGATCATCGAATTCTCCAAAATGGCCCGAGACCCCCAGCGCTACC 2935
P--K--F--R--E--L--I--I--E--F--S--K--M--A--R--D--P--Q--R--Y-- 978

TTGTCATTCAGGGGGATGAAAGAATGCATTTGCCAAGTCCTACAGACTCCAACTTCTACC 2995
L--V--I--Q--G--D--E--R--M--H--L--P--S--P--T--D--S--N--F--Y-- 998

GTGCCCTGATGGATGAAGAAGACATGGACGACGTGGTGGATGCCGACGAGTACCTCATCC 3055
R--A--L--M--D--E--E--D--M--D--D--V--V--D--A--D--E--Y--L--I-- 1018

CACAGCAGGGCTTCTTCAGCAGCCCCTCCACGTCACGGACTCCCCTCCTGAGCTCTCTGA 3115
P--Q--Q--G--F--F--S--S--P--S--T--S--R--T--P--L--L--S--S--L-- 1038

GTGCAACCAGCAACAATTCCACCGTGGCTTGCATTGATAGAAATGGGCTGCAAAGCTGTC 3175
S--A--T--S--N--N--S--T--V--A--C--I--D--R--N--G--L--Q--S--C-- 1058
```

도면5D

```
CCATCAAGGAAGACAGCTTCTTGCAGCGATACAGCTCAGACCCCACAGGCGCCTTGACTG 3235
P--I--K--E--D--S--F--L--Q--R--Y--S--S--D--P--T--G--A--L--T-- 1078

AGGACAGCATAGACGACACCTTCCTCCCAGTGCCTGAATACATAAACCAGTCCGTTCCCA 3295
E--D--S--I--D--D--T--F--L--P--V--P--E--Y--I--N--Q--S--V--P-- 1098

AAAGGCCCGCTGGCTCTGTGCAGAATCCTGTCTATCACAATCAGCCTCTGAACCCCGCGC 3355
K--R--P--A--G--S--V--Q--N--P--V--Y--H--N--Q--P--L--N--P--A-- 1118

CCAGCAGAGACCCACACTACCAGGACCCCCACAGCACTGCAGTGGGCAACCCCGAGTATC 3415
P--S--R--D--P--H--Y--Q--D--P--H--S--T--A--V--G--N--P--E--Y-- 1138

TCAACACTGTCCAGCCCACCTGTGTCAACAGCACATTCGACAGCCCTGCCCACTGGGCCC 3475
L--N--T--V--Q--P--T--C--V--N--S--T--F--D--S--P--A--H--W--A-- 1158

AGAAAGGCAGCCACCAAATTAGCCTGGACAACCCTGACTACCAGCAGGACTTCTTTCCCA 3535
Q--K--G--S--H--Q--I--S--L--D--N--P--D--Y--Q--Q--D--F--F--P-- 1178

AGGAAGCCAAGCCAAATGGCATCTTTAAGGGCTCCACAGCTGAAAATGCAGAATACCTAA 3595
K--E--A--K--P--N--G--I--F--K--G--S--T--A--E--N--A--E--Y--L-- 1198

GGGTCGCGCCACAAAGCAGTGAATTTATTGGAGCATGA 3633 (서열번호511)
R--V--A--P--Q--S--S--E--F--I--G--A--*- 1210 (서열번호512)
```

도면6

A

		서열 번호	
	활성화 루프		
L858R	KTPQHVKITDFGRAKLLGAEEKEYH	870	477
EGFR	KTPQHVKITDFGLAKLLGAEEKEYH	870	478
BRAF	HEDLTVKIGDFGLATVKSRWSGSHQ	608	479
	*** *****		

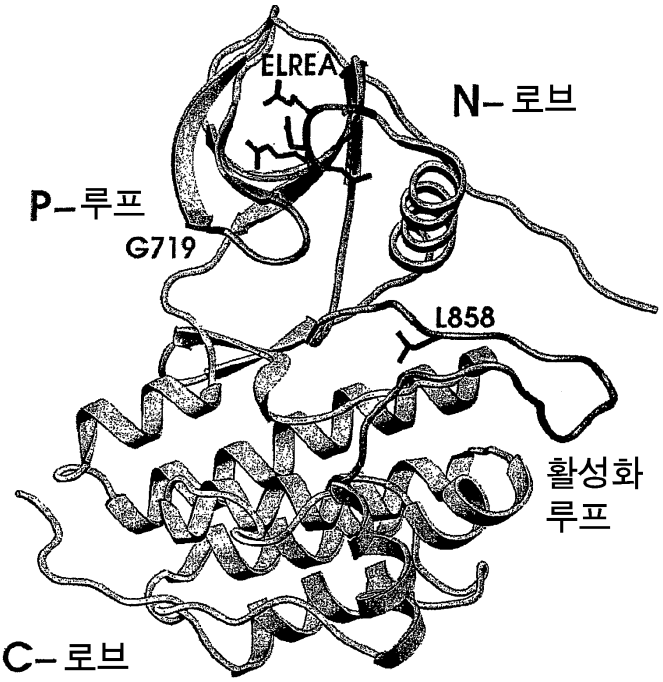
B

	P-루프		
G719S	ETEFKKIKVLSSGAFGTVYKGLWIP	733	480
EGFR	ETEFKKIKVLGSGAFGTVYKGLWIP	733	481
BRAF	DGQITVGQRIGSGSFGTVYKGKWHG	477	482
	*** ******* *		

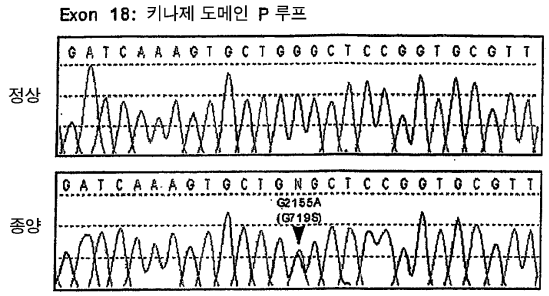
C

	742 |750|752		
Del-1	VAIK-----T-SPKANKEILDEAYV	765	483
Del-2	VAIKELREAT-----LDEAYV	765	484
Del-3	VAIKE---T-SPKANKEILDEAYV	765	485
Del-4	VAIKE-----SKANKEILDEAYV	765	486
Del-5	VAIKV-----PKANKEILDEAYV	765	487
EGFR	VAIKELREAT-SPKANKEILDEAYV	765	488
BRAF	VAVKMLNVTAPTPQQLQAFKNEVGV	503	489
	** * * *		

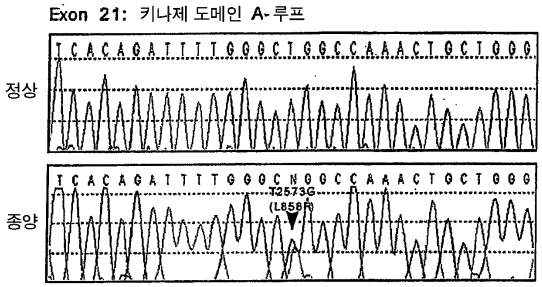
도면7



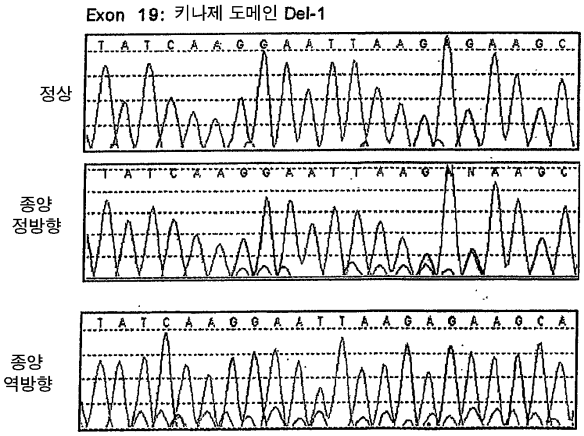
도면8A



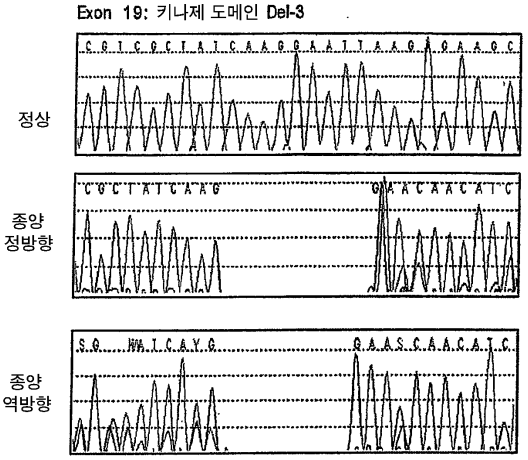
도면8B



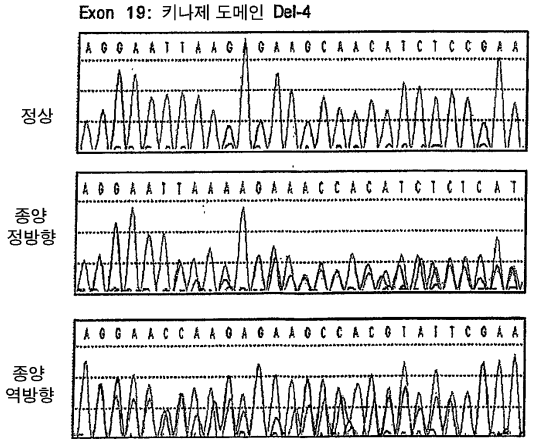
도면8C



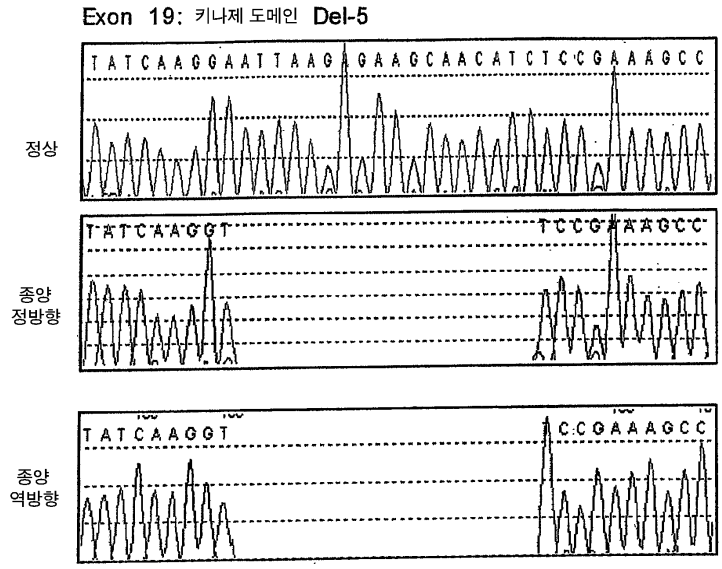
도면8D


Exon 19: 키나제 도메인 Del-3
정상
종양
정방향
종양
역방향

도면8E


Exon 19: 키나제 도메인 Del-4
정상
종양
정방향
종양
역방향

도면8F



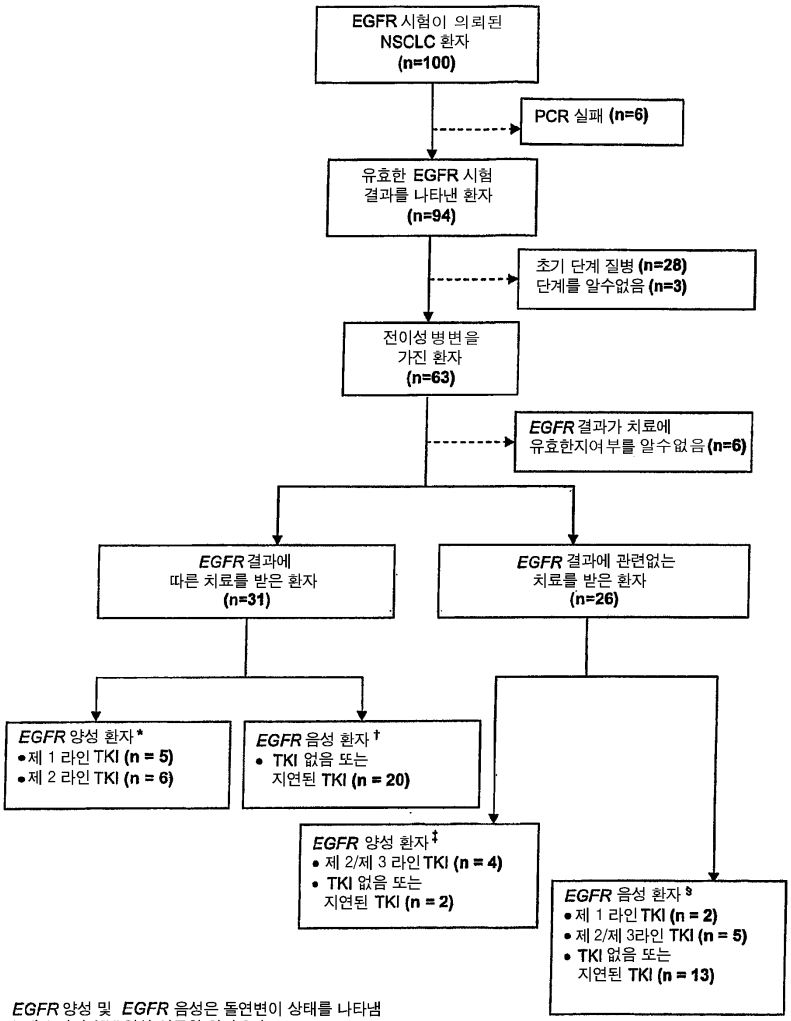
도면9

EGFR에 대한 Abl1의 정렬

```
                712                                                 790
Abl1(242)  ITMKHKLGGGQYGEVYEGVWK----KYSLTVAVKTLKEDT--MEVEEFLKEAAVMKEIKHPNLVQLLGVCTREPPFYIIT
EGFR(712)  FKKIKVLGSGAFGTVYKGLWIPEGEKVKIPVAIKELREATSPKANKEILDEAYVMASVDNPHVCRLLGICLTS-TVQLIT
보존서열          LG G FG VY GLW     K I VAIK LKE T     E L EA VM I P L  LLGIC         IIT
                791                                                 869
Abl1(316)  EFMTYGNLLDYLRECNREVNAVVLLYMATQISSAMEYLEKKNFIHRDLAARNCLVGENHLVKVADFGLSRLMTGDTYTA
EGFR(791)  QLMPFGCLLDYVREHKDN-IGSQYLLNWCVQIAKGMNYLEDRRLVHRDLAARNVLVKTPQHVKITDFGLAKLLGAEEKEY
보존서열          M FG LLDYLRE   N I A  LL    QIA AM YLE K  IHRDLAARN LV     VKI DFGLAKLL AD
                870                                                 949
Abl1(396)  HA-GAKFPIKWTAPESLAYNKFSIKSDVWAFGVLLWEIATYGMSPYPGIDLSQVYELLEKDYRMERPEGCPEKVYELMRA
EGFR(870)  HAEGGKVPIKWMALESILHRIYTHQSDVWSYGVTVWELMTFGSKPYDGIPASEISSILEKGERLPQPPICTIDVYMIMVK
보존서열          HA GAK PIKW A ESI H  FS  SDVWAFGV LWEI TFG  PY GI  S I  ILEK  RL  P  C   VY IM
                950                                                 979
Abl1(475)  CWQWNPSDRPSFAEIHQAF---------- (서열 번호 491)
EGFR(950)  CWMIDADSRPKFRELIIEFSKMARDPQRYL  (서열 번호 492)
보존서열          CW     RP F EI   F
```

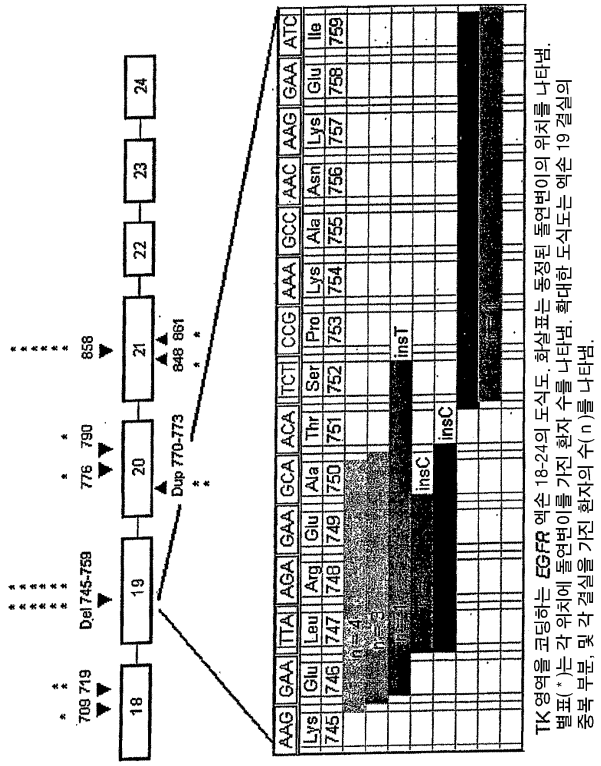
도면10

EGFR 시험을 경험한 전이성 NSCLC를 가진 환자에 대한 의사결정과정



EGFR 양성 및 *EGFR* 음성은 돌연변이 상태를 나타냄
* 제 1 라인 TKI 임상 시도한 환자 2명
† 시험 전에 TKI 를 받은 환자 2명
‡ 시험 전에 TKI 를 받은 환자 3명
§ 제 1 라인 TKI 임상 시도한 환자 1명과 시험전에 TKI를 받은 환자 4명

도면11



TK 영역을 코딩하는 *EGFR* 엑손 18-24의 도식도. 화살표는 동정된 돌연변이의 위치를 나타냄. 별표(*)는 각 위치에 돌연변이를 가진 환자 수를 나타냄. 확대한 도식도는 엑손 19 결실의 중복 부분, 및 각 결실을 가진 환자의 수(n)를 나타냄.

서 열 목 록

<110> THE GENERAL HOSPITAL CORPORATION
DANA-FARBER CANCER INSTITUTE, INC.

<120> METHOD TO DETERMINE RESPONSIVENESS OF CANCER TO EPIDERMAL GROWTH FACTOR RECEPTOR TARGETING TREATMENTS

<130> 030258-057952-KR

<140> KR 1020067022685
<141> 2006-10-30

<150> 60/558,218
<151> 2004-03-31

<150> 60/561,095
<151> 2004-04-09

<150> 60/565,753
<151> 2004-04-27

<150> 60/565,985
<151> 2004-04-27

<150> 60/574,035
<151> 2004-05-25

<150> 60/577,916
<151> 2004-06-07

<150> 60/592,287
<151> 2004-07-29

<160> 762

<170> PatentIn version 3.5

<210> 1
<211> 23
<212> DNA
<213> Homo sapiens

<400> 1
ggaaatatag ggaagggaag gaa 23

<210> 2
<211> 22
<212> DNA
<213> Homo sapiens

<400> 2
ttgacagggt accaggagat ga 22

<210> 3
<211> 20
<212> DNA
<213> Homo sapiens

<400> 3
ctgaaccgcc aaggactcat 20

<210> 4

<211> 23
<212> DNA
<213> Homo sapiens

<400> 4
ttttccctcc ctactaacac acg 23

<210> 5
<211> 20
<212> DNA
<213> Homo sapiens

<400> 5
actgatgccc tgaccctgtt 20

<210> 6
<211> 19
<212> DNA
<213> Homo sapiens

<400> 6
cccatggttc cccactctt 19

<210> 7
<211> 20
<212> DNA
<213> Homo sapiens

<400> 7
agggactcca aagccatgtg 20

<210> 8
<211> 19
<212> DNA
<213> Homo sapiens

<400> 8
ctctctgggg ccatccact 19

<210> 9
<211> 23

<212> DNA
<213> Homo sapiens

<400> 9
cattgtcaag ggcaatgtaa gtg 23

<210> 10
<211> 20
<212> DNA
<213> Homo sapiens

<400> 10
ctctcaccaa ccctcgctgt 20

<210> 11
<211> 20
<212> DNA
<213> Homo sapiens

<400> 11
acatggggag ccagagtgac 20

<210> 12
<211> 20
<212> DNA
<213> Homo sapiens

<400> 12
tgcaacccag agaaagtgtg 20

<210> 13
<211> 24
<212> DNA
<213> Homo sapiens

<400> 13
tgagctttca accctagttt gttg 24

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

<213> Homo sapiens

<400> 14
gtttgcctcc tgctgtctca 20

<210> 15
<211> 24
<212> DNA
<213> Homo sapiens

<400> 15
tgtccttgtg tttttgaaga ttcc 24

<210> 16
<211> 23
<212> DNA
<213> Homo sapiens

<400> 16
tgcagacaga tgacaaacat gaa 23

<210> 17
<211> 21
<212> DNA
<213> Homo sapiens

<400> 17
tgggtgagtc tctgtgtgga g 21

<210> 18
<211> 25
<212> DNA
<213> Homo sapiens

<400> 18
cattgccata gcaaaaataa acaca 25

<210> 19
<211> 19
<212> DNA

<213> Homo sapiens

<400> 19
ggttcaactg ggcgtccta 19

<210> 20
<211> 20
<212> DNA
<213> Homo sapiens

<400> 20
ccttctccga ggtggaattg 20

<210> 21
<211> 20
<212> DNA
<213> Homo sapiens

<400> 21
cgcaccatgg catctcttta 20

<210> 22
<211> 23
<212> DNA
<213> Homo sapiens

<400> 22
aaaacgatct ctatgtccgt ggt 23

<210> 23
<211> 20
<212> DNA
<213> Homo sapiens

<400> 23
cagccagcca aacaatcaga 20

<210> 24
<211> 23
<212> DNA

<213> Homo sapiens

<400> 24
tctttggagt cttcagaggg aaa 23

<210> 25
<211> 20
<212> DNA
<213> Homo sapiens

<400> 25
tgtggtttcg ttggaagcaa 20

<210> 26
<211> 20
<212> DNA
<213> Homo sapiens

<400> 26
aattgacagc tcccccacag 20

<210> 27
<211> 20
<212> DNA
<213> Homo sapiens

<400> 27
ggctttctga cgggagtcaa 20

<210> 28
<211> 21
<212> DNA
<213> Homo sapiens

<400> 28
ccacccaaag actctccaag a 21

<210> 29
<211> 20
<212> DNA

<213> Homo sapiens

<400> 29
cctttccatc acccctcaag 20

<210> 30
<211> 20
<212> DNA
<213> Homo sapiens

<400> 30
agtgccttcc cattgcctaa 20

<210> 31
<211> 20
<212> DNA
<213> Homo sapiens

<400> 31
accggaattc cttcctgctt 20

<210> 32
<211> 23
<212> DNA
<213> Homo sapiens

<400> 32
cactgaaaca aacaacaggg tga 23

<210> 33
<211> 20
<212> DNA
<213> Homo sapiens

<400> 33
agggggtgag tcacaggttc 20

<210> 34
<211> 26
<212> DNA

<213> Homo sapiens

<400> 34
tcagaagaaa tgtttttatt ccaagg 26

<210> 35
<211> 21
<212> DNA
<213> Homo sapiens

<400> 35
gcaaatccaa ttttcccact t 21

<210> 36
<211> 20
<212> DNA
<213> Homo sapiens

<400> 36
gcaggagctc tgtgccctat 20

<210> 37
<211> 20
<212> DNA
<213> Homo sapiens

<400> 37
tcccacagca tgacctacca 20

<210> 38
<211> 23
<212> DNA
<213> Homo sapiens

<400> 38
tttgcttctt aaggaactga aaa 23

<210> 39
<211> 20
<212> DNA

<213> Homo sapiens

<400> 39
tgtcacccaa ggtcatggag 20

<210> 40
<211> 20
<212> DNA
<213> Homo sapiens

<400> 40
caaaagccaa gggcaaagaa 20

<210> 41
<211> 20
<212> DNA
<213> Homo sapiens

<400> 41
ggagtcccaa ctccttgacc 20

<210> 42
<211> 19
<212> DNA
<213> Homo sapiens

<400> 42
gtcctgccca cacaggatg 19

<210> 43
<211> 20
<212> DNA
<213> Homo sapiens

<400> 43
gctttcccca ctcacacaca 20

<210> 44
<211> 20
<212> DNA

<213> Homo sapiens

<400> 44
caaacctcgg caatttgttg 20

<210> 45
<211> 20
<212> DNA
<213> Homo sapiens

<400> 45
ccaccaatcc aacatccaga 20

<210> 46
<211> 20
<212> DNA
<213> Homo sapiens

<400> 46
tggcccagag ccatagaaac 20

<210> 47
<211> 26
<212> DNA
<213> Homo sapiens

<400> 47
ttccaagatc attctacaag atgtca 26

<210> 48
<211> 23
<212> DNA
<213> Homo sapiens

<400> 48
gcacattcag agattctttc tgc 23

<210> 49
<211> 20
<212> DNA

<213> Homo sapiens

<400> 49
tccaaatgag ctggcaagtg 20

<210> 50
<211> 23
<212> DNA
<213> Homo sapiens

<400> 50
tcccaaacac tcagtgaaac aaa 23

<210> 51
<211> 20
<212> DNA
<213> Homo sapiens

<400> 51
gtgcatcgct ggtaacatcc 20

<210> 52
<211> 20
<212> DNA
<213> Homo sapiens

<400> 52
tgtggagatg agcagggtct 20

<210> 53
<211> 20
<212> DNA
<213> Homo sapiens

<400> 53
atcgcattca tgcgtcttca 20

<210> 54
<211> 20
<212> DNA

<213> Homo sapiens

<400> 54
atccccatgg caaactcttg 20

<210> 55
<211> 19
<212> DNA
<213> Homo sapiens

<400> 55
gctcagagcc tggcatgaa 19

<210> 56
<211> 19
<212> DNA
<213> Homo sapiens

<400> 56
catcctcccc tgcatgtgt 19

<210> 57
<211> 20
<212> DNA
<213> Homo sapiens

<400> 57
tggctcgtct gtgtgtgtca 20

<210> 58
<211> 25
<212> DNA
<213> Homo sapiens

<400> 58
cgaaagaaaa tacttgcatg tcaga 25

<210> 59
<211> 20
<212> DNA

<213> Homo sapiens

<400> 59
tgaagcaaat tgcccaagac 20

<210> 60
<211> 20
<212> DNA
<213> Homo sapiens

<400> 60
tgacatttct ccagggatgc 20

<210> 61
<211> 20
<212> DNA
<213> Homo sapiens

<400> 61
aagtgtcgca tcaccaatgc 20

<210> 62
<211> 19
<212> DNA
<213> Homo sapiens

<400> 62
atgcgatctg ggacacagg 19

<210> 63
<211> 20
<212> DNA
<213> Homo sapiens

<400> 63
ggcacctgct ggcaatagac 20

<210> 64
<211> 27
<212> DNA

<213> Homo sapiens

<400> 64
tgacttcata tccatgtgag tttcact 27

<210> 65
<211> 20
<212> DNA
<213> Homo sapiens

<400> 65
tataccctcc atgaggcaca 20

<210> 66
<211> 20
<212> DNA
<213> Homo sapiens

<400> 66
gggaaaaacc cacacaggaa 20

<210> 67
<211> 21
<212> DNA
<213> Homo sapiens

<400> 67
tcagaaccag catctcaagg a 21

<210> 68
<211> 19
<212> DNA
<213> Homo sapiens

<400> 68
gatgctggag ggagcacct 19

<210> 69
<211> 22
<212> DNA

<213> Homo sapiens

<400> 69
ccttgttgag gacattcaca gg 22

<210> 70
<211> 20
<212> DNA
<213> Homo sapiens

<400> 70
atgtgcccga ggtggaagta 20

<210> 71
<211> 20
<212> DNA
<213> Homo sapiens

<400> 71
ggagggcaga ggactagctg 20

<210> 72
<211> 20
<212> DNA
<213> Homo sapiens

<400> 72
gtgcctggcc aagtctttgt 20

<210> 73
<211> 20
<212> DNA
<213> Homo sapiens

<400> 73
ctgcagccta gcaacagagc 20

<210> 74
<211> 20
<212> DNA

<213> Homo sapiens

<400> 74
aagaaccaga ggagccagga 20

<210> 75
<211> 20
<212> DNA
<213> Homo sapiens

<400> 75
cgggtaagga tgtgggttgt 20

<210> 76
<211> 20
<212> DNA
<213> Homo sapiens

<400> 76
caggtgttct gcctcctgaa 20

<210> 77
<211> 20
<212> DNA
<213> Homo sapiens

<400> 77
gcttcaggag gcagaacacc 20

<210> 78
<211> 22
<212> DNA
<213> Homo sapiens

<400> 78
ggagcaagcc taagaaggtt ca 22

<210> 79
<211> 21
<212> DNA

<213> Homo sapiens

<400> 79
gccttgtatc catttgccac a 21

<210> 80
<211> 21
<212> DNA
<213> Homo sapiens

<400> 80
tgacaacacg ttttgggtca t 21

<210> 81
<211> 23
<212> DNA
<213> Homo sapiens

<400> 81
tgcatattcc atttcagaac aga 23

<210> 82
<211> 23
<212> DNA
<213> Homo sapiens

<400> 82
aaacagtttc attgctgcta aat 23

<210> 83
<211> 23
<212> DNA
<213> Homo sapiens

<400> 83
ccggatacag atacccaaaa aga 23

<210> 84
<211> 20
<212> DNA

<213> Homo sapiens

<400> 84
ggaggcttca agggatgaga 20

<210> 85
<211> 20
<212> DNA
<213> Homo sapiens

<400> 85
gctgttgtcc tgcttggcta 20

<210> 86
<211> 25
<212> DNA
<213> Homo sapiens

<400> 86
tggttgtaat gttgaactag cttgc 25

<210> 87
<211> 24
<212> DNA
<213> Homo sapiens

<400> 87
tggctgtcag ctaaataagc atgt 24

<210> 88
<211> 24
<212> DNA
<213> Homo sapiens

<400> 88
tcaatttgct tcatttctcc tgtt 24

<210> 89
<211> 21
<212> DNA

<213> Homo sapiens

<400> 89
tgctgctgaa ctaccaacca a 21

<210> 90
<211> 24
<212> DNA
<213> Homo sapiens

<400> 90
tgtggtagta attgtggaaa actg 24

<210> 91
<211> 20
<212> DNA
<213> Homo sapiens

<400> 91
caaagcaccg tctcaactcg 20

<210> 92
<211> 20
<212> DNA
<213> Homo sapiens

<400> 92
cccgaaactg ccaacttcat 20

<210> 93
<211> 20
<212> DNA
<213> Homo sapiens

<400> 93
ggaaaacagg accccagtgt 20

<210> 94
<211> 20
<212> DNA

<213> Homo sapiens

<400> 94
ccctcctcca cagagctgat 20

<210> 95
<211> 21
<212> DNA
<213> Homo sapiens

<400> 95
gacagaagct gacaagcagc a 21

<210> 96
<211> 20
<212> DNA
<213> Homo sapiens

<400> 96
aggttccatt ccctcccagt 20

<210> 97
<211> 22
<212> DNA
<213> Homo sapiens

<400> 97
tgggagtgag agtttggaag aa 22

<210> 98
<211> 20
<212> DNA
<213> Homo sapiens

<400> 98
tatgaggccg tgagctgaaa 20

<210> 99
<211> 18
<212> DNA

<213> Homo sapiens

<400> 99
agggccctgc tctggttt 18

<210> 100
<211> 20
<212> DNA
<213> Homo sapiens

<400> 100
ccaattgggc gttagtgaaa 20

<210> 101
<211> 20
<212> DNA
<213> Homo sapiens

<400> 101
ctcatgagat tggggcatca 20

<210> 102
<211> 20
<212> DNA
<213> Homo sapiens

<400> 102
aggcccatga tctcagaagc 20

<210> 103
<211> 20
<212> DNA
<213> Homo sapiens

<400> 103
ggttgcagga gagacgaggt 20

<210> 104
<211> 19
<212> DNA

<213> Homo sapiens

<400> 104
aggcccttca ccctgtgac 19

<210> 105
<211> 20
<212> DNA
<213> Homo sapiens

<400> 105
atgacccctc cgatcctacc 20

<210> 106
<211> 20
<212> DNA
<213> Homo sapiens

<400> 106
taatcctgct ccacggcatt 20

<210> 107
<211> 20
<212> DNA
<213> Homo sapiens

<400> 107
ggaaaaagca gaggcaggtg 20

<210> 108
<211> 20
<212> DNA
<213> Homo sapiens

<400> 108
tggtctcaag aacccagcag 20

<210> 109
<211> 19
<212> DNA

<213> Homo sapiens

<400> 109
gacaccctcc ccctctcat 19

<210> 110
<211> 20
<212> DNA
<213> Homo sapiens

<400> 110
actatgacac cccggctgag 20

<210> 111
<211> 21
<212> DNA
<213> Homo sapiens

<400> 111
tgcttgatgt aaaaacccttg g 21

<210> 112
<211> 20
<212> DNA
<213> Homo sapiens

<400> 112
gcaatccaac agccatgaga 20

<210> 113
<211> 20
<212> DNA
<213> Homo sapiens

<400> 113
ggagcaaacc cctatgtcca 20

<210> 114
<211> 21
<212> DNA

<213> Homo sapiens

<400> 114
tcctccaact gtgtgttgtg g 21

<210> 115
<211> 19
<212> DNA
<213> Homo sapiens

<400> 115
tggggaccac tgctgagag 19

<210> 116
<211> 20
<212> DNA
<213> Homo sapiens

<400> 116
tgcagccttc tctccttgaa 20

<210> 117
<211> 20
<212> DNA
<213> Homo sapiens

<400> 117
gcagagcagt gtggcagaag 20

<210> 118
<211> 19
<212> DNA
<213> Homo sapiens

<400> 118
acaggtggga agggactgg 19

<210> 119
<211> 19
<212> DNA

<213> Homo sapiens

<400> 119
agtggggtgg gctgagaac 19

<210> 120
<211> 20
<212> DNA
<213> Homo sapiens

<400> 120
tctctggggc agaaagagga 20

<210> 121
<211> 20
<212> DNA
<213> Homo sapiens

<400> 121
acccggccac actgtatttc 20

<210> 122
<211> 20
<212> DNA
<213> Homo sapiens

<400> 122
catcccaccc agctctcaac 20

<210> 123
<211> 20
<212> DNA
<213> Homo sapiens

<400> 123
agggcatagc cctattgagc 20

<210> 124
<211> 20
<212> DNA

<213> Homo sapiens

<400> 124
cccaggaaaa agccagagaa 20

<210> 125
<211> 20
<212> DNA
<213> Homo sapiens

<400> 125
caggtgtggg tggagtaggc 20

<210> 126
<211> 18
<212> DNA
<213> Homo sapiens

<400> 126
ctcaggcgcc atccactt 18

<210> 127
<211> 21
<212> DNA
<213> Homo sapiens

<400> 127
aagaagacga ccaacgtgag c 21

<210> 128
<211> 18
<212> DNA
<213> Homo sapiens

<400> 128
aggagctcca gggcacag 18

<210> 129
<211> 22
<212> DNA

<213> Homo sapiens

<400> 129
cctcctctgt aaagtgggtg ga 22

<210> 130
<211> 20
<212> DNA
<213> Homo sapiens

<400> 130
agagggcctc agtgcagagt 20

<210> 131
<211> 20
<212> DNA
<213> Homo sapiens

<400> 131
agatggggca gaactggatg 20

<210> 132
<211> 20
<212> DNA
<213> Homo sapiens

<400> 132
gggtcccaga ccaaatctga 20

<210> 133
<211> 20
<212> DNA
<213> Homo sapiens

<400> 133
aggtgctccc ttcacagcat 20

<210> 134
<211> 22
<212> DNA

<213> Homo sapiens

<400> 134
ttcagggact acagctgagg aa 22

<210> 135
<211> 21
<212> DNA
<213> Homo sapiens

<400> 135
gccgtatgtt atctgggagg t 21

<210> 136
<211> 21
<212> DNA
<213> Homo sapiens

<400> 136
tgggcccatt acactttaag a 21

<210> 137
<211> 20
<212> DNA
<213> Homo sapiens

<400> 137
ttccatcacc ggtacctcct 20

<210> 138
<211> 23
<212> DNA
<213> Homo sapiens

<400> 138
ccataaatca aaaatgcacc aca 23

<210> 139
<211> 23
<212> DNA

<213> Homo sapiens

<400> 139
gagtggtctt aggaagatga tgc 23

<210> 140
<211> 22
<212> DNA
<213> Homo sapiens

<400> 140
aaagtcatgg gctgcaatac aa 22

<210> 141
<211> 18
<212> DNA
<213> Homo sapiens

<400> 141
atggtcccca ctgcttgg 18

<210> 142
<211> 20
<212> DNA
<213> Homo sapiens

<400> 142
aggagctcac ctcaccctgt 20

<210> 143
<211> 20
<212> DNA
<213> Homo sapiens

<400> 143
ccttgcgtct ctccacacat 20

<210> 144
<211> 20
<212> DNA

<213> Homo sapiens

<400> 144
tggcaacggg taacaatgaa 20

<210> 145
<211> 20
<212> DNA
<213> Homo sapiens

<400> 145
ggctgaggta agctgcttcg 20

<210> 146
<211> 23
<212> DNA
<213> Homo sapiens

<400> 146
aaaaagaagt atcttgcccc ttt 23

<210> 147
<211> 21
<212> DNA
<213> Homo sapiens

<400> 147
aacccctctt agggctctgt g 21

<210> 148
<211> 19
<212> DNA
<213> Homo sapiens

<400> 148
caggaggatg gcaggcttc 19

<210> 149
<211> 23
<212> DNA

<213> Homo sapiens

<400> 149
cgtagagagc ttcaggacct gtg 23

<210> 150
<211> 20
<212> DNA
<213> Homo sapiens

<400> 150
ttccgagaag ttttgcctga 20

<210> 151
<211> 22
<212> DNA
<213> Homo sapiens

<400> 151
tgtgaacatc attcaaggcg ta 22

<210> 152
<211> 25
<212> DNA
<213> Homo sapiens

<400> 152
aaaatgtgtg atatccctag acagg 25

<210> 153
<211> 20
<212> DNA
<213> Homo sapiens

<400> 153
tccacatttc agcaacagca 20

<210> 154
<211> 21
<212> DNA

<213> Homo sapiens

<400> 154
ggctgcttcc tgagacacag t 21

<210> 155
<211> 20
<212> DNA
<213> Homo sapiens

<400> 155
tatctaccgg tgcgggactt 20

<210> 156
<211> 20
<212> DNA
<213> Homo sapiens

<400> 156
aggtgtagcc tcccctcaca 20

<210> 157
<211> 20
<212> DNA
<213> Homo sapiens

<400> 157
aggctggtgg tgtctctgtg 20

<210> 158
<211> 20
<212> DNA
<213> Homo sapiens

<400> 158
caagctgcca accctcagtt 20

<210> 159
<211> 24
<212> DNA

<213> Homo sapiens

<400> 159
tggatttcaa atactgaagc cact 24

<210> 160
<211> 22
<212> DNA
<213> Homo sapiens

<400> 160
tggaattggt ggtgttgaat tt 22

<210> 161
<211> 23
<212> DNA
<213> Homo sapiens

<400> 161
gggcttcata tgttctgaca tgg 23

<210> 162
<211> 19
<212> DNA
<213> Homo sapiens

<400> 162
cagaggacca cgccatagg 19

<210> 163
<211> 20
<212> DNA
<213> Homo sapiens

<400> 163
ccgagattta gccaccagga 20

<210> 164
<211> 20
<212> DNA

<213> Homo sapiens

<400> 164
cctgggaagc aaacaacaca 20

<210> 165
<211> 20
<212> DNA
<213> Homo sapiens

<400> 165
aggtccccag tctcctctcc 20

<210> 166
<211> 20
<212> DNA
<213> Homo sapiens

<400> 166
agacccatgc agccatccta 20

<210> 167
<211> 19
<212> DNA
<213> Homo sapiens

<400> 167
cgtgaaccac cgagcttgt 19

<210> 168
<211> 19
<212> DNA
<213> Homo sapiens

<400> 168
agaggggcag aaggggaac 19

<210> 169
<211> 20
<212> DNA

<213> Homo sapiens

<400> 169
ggtgggggtg aggagcttag 20

<210> 170
<211> 20
<212> DNA
<213> Homo sapiens

<400> 170
tcgtttaagc cacccagtca 20

<210> 171
<211> 21
<212> DNA
<213> Homo sapiens

<400> 171
tgcaaataag gaaagcaaac a 21

<210> 172
<211> 21
<212> DNA
<213> Homo sapiens

<400> 172
tcctgacatg gtcttccaac c 21

<210> 173
<211> 22
<212> DNA
<213> Homo sapiens

<400> 173
cagcatcttc acacacctct ga 22

<210> 174
<211> 20
<212> DNA

<213> Homo sapiens

<400> 174
gctggctcta aatcccacct 20

<210> 175
<211> 20
<212> DNA
<213> Homo sapiens

<400> 175
ctaatccggg aagttgttgc 20

<210> 176
<211> 24
<212> DNA
<213> Homo sapiens

<400> 176
ttctgtatca gcagcttctc tgtg 24

<210> 177
<211> 19
<212> DNA
<213> Homo sapiens

<400> 177
caagtgccac catggatca 19

<210> 178
<211> 21
<212> DNA
<213> Homo sapiens

<400> 178
ggcagtgtac tgacccctg a 21

<210> 179
<211> 24
<212> DNA

<213> Homo sapiens

<400> 179
gcacaagtta ttaagagccc aagg 24

<210> 180
<211> 20
<212> DNA
<213> Homo sapiens

<400> 180
agcatactgg cctcacacca 20

<210> 181
<211> 22
<212> DNA
<213> Homo sapiens

<400> 181
gcacatgggc agtgttgtat tt 22

<210> 182
<211> 20
<212> DNA
<213> Homo sapiens

<400> 182
gagccccaca cagatttcct 20

<210> 183
<211> 20
<212> DNA
<213> Homo sapiens

<400> 183
atgggacgga gaagtggttg 20

<210> 184
<211> 20
<212> DNA

<213> Homo sapiens

<400> 184
tccctgtatc agggctcgtc 20

<210> 185
<211> 20
<212> DNA
<213> Homo sapiens

<400> 185
ttcctacgca gcacaccaat 20

<210> 186
<211> 20
<212> DNA
<213> Homo sapiens

<400> 186
gcaggcacta aacccttcc 20

<210> 187
<211> 20
<212> DNA
<213> Homo sapiens

<400> 187
gcacgcatgt gaccaatttc 20

<210> 188
<211> 20
<212> DNA
<213> Homo sapiens

<400> 188
agccctgaga gggaggtagg 20

<210> 189
<211> 19
<212> DNA

<213> Homo sapiens

<400> 189
cacacaccac ccctctgct 19

<210> 190
<211> 21
<212> DNA
<213> Homo sapiens

<400> 190
aaagatttgg ggtgagggct a 21

<210> 191
<211> 20
<212> DNA
<213> Homo sapiens

<400> 191
ctgaaagctc agggataggg 20

<210> 192
<211> 18
<212> DNA
<213> Homo sapiens

<400> 192
ctggccaagc tgcacaga 18

<210> 193
<211> 20
<212> DNA
<213> Homo sapiens

<400> 193
tgcagccaac gatttgaaag 20

<210> 194
<211> 20
<212> DNA

<213> Homo sapiens

<400> 194
ggaaagcccc aagtctgaaa 20

<210> 195
<211> 21
<212> DNA
<213> Homo sapiens

<400> 195
tcatcatgag atccccacac t 21

<210> 196
<211> 20
<212> DNA
<213> Homo sapiens

<400> 196
gcatttcccc ctgaaggagt 20

<210> 197
<211> 23
<212> DNA
<213> Homo sapiens

<400> 197
tggattcagt aaccaggaag tga 23

<210> 198
<211> 20
<212> DNA
<213> Homo sapiens

<400> 198
cccattccac caggatgatt 20

<210> 199
<211> 20
<212> DNA

<213> Homo sapiens

<400> 199
gtttccagct gcccactacc 20

<210> 200
<211> 20
<212> DNA
<213> Homo sapiens

<400> 200
gctcgaaacc acatgttcca 20

<210> 201
<211> 20
<212> DNA
<213> Homo sapiens

<400> 201
ctggatttgg ggttctctgc 20

<210> 202
<211> 22
<212> DNA
<213> Homo sapiens

<400> 202
cgggaacagc tagcagattt tt 22

<210> 203
<211> 20
<212> DNA
<213> Homo sapiens

<400> 203
gggaattttg gaggggaact 20

<210> 204
<211> 20
<212> DNA

<213> Homo sapiens

<400> 204
gcttcagtca ccacagagca 20

<210> 205
<211> 23
<212> DNA
<213> Homo sapiens

<400> 205
tgagtctacc cagcaatcat ttg 23

<210> 206
<211> 20
<212> DNA
<213> Homo sapiens

<400> 206
ttcccgagag ctacaggaca 20

<210> 207
<211> 21
<212> DNA
<213> Homo sapiens

<400> 207
ggtaacaagg gtacccacga a 21

<210> 208
<211> 20
<212> DNA
<213> Homo sapiens

<400> 208
gtttgagggg ctgagtgtgg 20

<210> 209
<211> 20
<212> DNA

<213> Homo sapiens

<400> 209
cctcaccctt agggcttgtg 20

<210> 210
<211> 20
<212> DNA
<213> Homo sapiens

<400> 210
agcccaggtc atgccttaga 20

<210> 211
<211> 20
<212> DNA
<213> Homo sapiens

<400> 211
gggtagcttg ggagcaaaga 20

<210> 212
<211> 20
<212> DNA
<213> Homo sapiens

<400> 212
ccaaacccca gagagcagac 20

<210> 213
<211> 22
<212> DNA
<213> Homo sapiens

<400> 213
catttcccct aatccttttc ca 22

<210> 214
<211> 22
<212> DNA

<213> Homo sapiens

<400> 214
gtgatcccag atttaggcct tc 22

<210> 215
<211> 22
<212> DNA
<213> Homo sapiens

<400> 215
gcctctcgtg gtttgttttg tc 22

<210> 216
<211> 22
<212> DNA
<213> Homo sapiens

<400> 216
cccagggtag ggtccaataa tc 22

<210> 217
<211> 22
<212> DNA
<213> Homo sapiens

<400> 217
cttcctggtg gaggtgactg at 22

<210> 218
<211> 22
<212> DNA
<213> Homo sapiens

<400> 218
caggcatagt gtgtgatggt ca 22

<210> 219
<211> 23
<212> DNA

<213> Homo sapiens

<400> 219
tcacgataca cattctcaga tcc 23

<210> 220
<211> 22
<212> DNA
<213> Homo sapiens

<400> 220
gaagatctcc cagaggagga tg 22

<210> 221
<211> 21
<212> DNA
<213> Homo sapiens

<400> 221
cgtaacgtgc tgttgaccaa t 21

<210> 222
<211> 22
<212> DNA
<213> Homo sapiens

<400> 222
aaacgaggga agagccagaa ag 22

<210> 223
<211> 22
<212> DNA
<213> Homo sapiens

<400> 223
tggggagcac aataaaagaa ga 22

<210> 224
<211> 22
<212> DNA

<213> Homo sapiens

<400> 224
actcttggct cctggattct tg 22

<210> 225
<211> 22
<212> DNA
<213> Homo sapiens

<400> 225
ggaagtcagt gtgcagggaa ta 22

<210> 226
<211> 22
<212> DNA
<213> Homo sapiens

<400> 226
ttttagcaga aataggcaag ca 22

<210> 227
<211> 24
<212> DNA
<213> Homo sapiens

<400> 227
tggtaatcct aaacacaatg caga 24

<210> 228
<211> 22
<212> DNA
<213> Homo sapiens

<400> 228
ctgggcaaca cagtgagatc ct 22

<210> 229
<211> 23
<212> DNA

<213> Homo sapiens

<400> 229
tcacaaattt ctttgctgtg tcc 23

<210> 230
<211> 23
<212> DNA
<213> Homo sapiens

<400> 230
catggaactc cagattagcc tgt 23

<210> 231
<211> 22
<212> DNA
<213> Homo sapiens

<400> 231
gattgttgca gatcgtggac at 22

<210> 232
<211> 22
<212> DNA
<213> Homo sapiens

<400> 232
cgcttaaatc ttcccattcc ag 22

<210> 233
<211> 22
<212> DNA
<213> Homo sapiens

<400> 233
ctccatggca ccatcattaa ca 22

<210> 234
<211> 22
<212> DNA

<213> Homo sapiens

<400> 234
ctcaggacac aagtgctctg ct 22

<210> 235
<211> 24
<212> DNA
<213> Homo sapiens

<400> 235
gcagttcatg gttcatcttc tttt 24

<210> 236
<211> 22
<212> DNA
<213> Homo sapiens

<400> 236
caaaatagcc caccctggat ta 22

<210> 237
<211> 21
<212> DNA
<213> Homo sapiens

<400> 237
ctttctgcat tgcccaagat g 21

<210> 238
<211> 22
<212> DNA
<213> Homo sapiens

<400> 238
caaggtctca gtgagtggtg ga 22

<210> 239
<211> 23
<212> DNA

<213> Homo sapiens

<400> 239
gagaagggtc tttctgactc tgc 23

<210> 240
<211> 22
<212> DNA
<213> Homo sapiens

<400> 240
caggtgtttc tcctgtgagg tg 22

<210> 241
<211> 22
<212> DNA
<213> Homo sapiens

<400> 241
cacattgcgg cctagaatgt ta 22

<210> 242
<211> 20
<212> DNA
<213> Homo sapiens

<400> 242
accccgtcac aaccttcagt 20

<210> 243
<211> 22
<212> DNA
<213> Homo sapiens

<400> 243
gccgtagccc caaagtgtac ta 22

<210> 244
<211> 22
<212> DNA

<213> Homo sapiens

<400> 244
tcagctcaaa cctgtgattt cc 22

<210> 245
<211> 24
<212> DNA
<213> Homo sapiens

<400> 245
ctcactctcc ataaatgcta cgaa 24

<210> 246
<211> 22
<212> DNA
<213> Homo sapiens

<400> 246
gacttaacgt gtccccttt gc 22

<210> 247
<211> 22
<212> DNA
<213> Homo sapiens

<400> 247
gcctcttcgg ggtaatcaga ta 22

<210> 248
<211> 22
<212> DNA
<213> Homo sapiens

<400> 248
gaagtctgtg gtttagcgga ca 22

<210> 249
<211> 22
<212> DNA

<213> Homo sapiens

<400> 249
atcttttgcc tggaggaact tt 22

<210> 250
<211> 22
<212> DNA
<213> Homo sapiens

<400> 250
cagggtaaat tcatcccatt ga 22

<210> 251
<211> 22
<212> DNA
<213> Homo sapiens

<400> 251
cagcagccag cacaactact tt 22

<210> 252
<211> 23
<212> DNA
<213> Homo sapiens

<400> 252
ttggctagat gaaccattga tga 23

<210> 253
<211> 24
<212> DNA
<213> Homo sapiens

<400> 253
tgaatgaagc tcctgtgttt actc 24

<210> 254
<211> 22
<212> DNA

<213> Homo sapiens

<400> 254
atgttcatcg caggctaatg tg 22

<210> 255
<211> 24
<212> DNA
<213> Homo sapiens

<400> 255
aaaacaggga gaacttctaa gcaa 24

<210> 256
<211> 21
<212> DNA
<213> Homo sapiens

<400> 256
catggcagag tcattcccac t 21

<210> 257
<211> 22
<212> DNA
<213> Homo sapiens

<400> 257
caatgctaga acaacgcctg tc 22

<210> 258
<211> 22
<212> DNA
<213> Homo sapiens

<400> 258
tccctccact gaggacaaag tt 22

<210> 259
<211> 22
<212> DNA

<213> Homo sapiens

<400> 259
gggagagctt gagaaagttg ga 22

<210> 260
<211> 22
<212> DNA
<213> Homo sapiens

<400> 260
atttcctcgg atggatgtac ca 22

<210> 261
<211> 22
<212> DNA
<213> Homo sapiens

<400> 261
tcagagcctg tgtttctacc aa 22

<210> 262
<211> 22
<212> DNA
<213> Homo sapiens

<400> 262
tggtctcaca ggaccactga tt 22

<210> 263
<211> 24
<212> DNA
<213> Homo sapiens

<400> 263
aaataatcag tgtgattcgt ggag 24

<210> 264
<211> 22
<212> DNA

<213> Homo sapiens

<400> 264
gaggccagtg ctgtctctaa gg 22

<210> 265
<211> 21
<212> DNA
<213> Homo sapiens

<400> 265
acttcacagc cctgcgtaaa c 21

<210> 266
<211> 21
<212> DNA
<213> Homo sapiens

<400> 266
atgggacagg cactgatttg t 21

<210> 267
<211> 22
<212> DNA
<213> Homo sapiens

<400> 267
gcagcgggtt acatcttctt tc 22

<210> 268
<211> 22
<212> DNA
<213> Homo sapiens

<400> 268
cagctctggc tcacactacc ag 22

<210> 269
<211> 22
<212> DNA

<213> Homo sapiens

<400> 269
cctgaactcc gtcagactga aa 22

<210> 270
<211> 20
<212> DNA
<213> Homo sapiens

<400> 270
gcagctggac tcgatttcct 20

<210> 271
<211> 23
<212> DNA
<213> Homo sapiens

<400> 271
ccttacagca atcctgtgaa aca 23

<210> 272
<211> 22
<212> DNA
<213> Homo sapiens

<400> 272
tgcccaatga gtcaagaagt gt 22

<210> 273
<211> 22
<212> DNA
<213> Homo sapiens

<400> 273
atgtacagtg ctggcatggt ct 22

<210> 274
<211> 22
<212> DNA

<213> Homo sapiens

<400> 274
cactcacgga tgctgcttag tt 22

<210> 275
<211> 21
<212> DNA
<213> Homo sapiens

<400> 275
taaggcaccc acatcatgtc a 21

<210> 276
<211> 24
<212> DNA
<213> Homo sapiens

<400> 276
tggacctaaa aggcttacaa tcaa 24

<210> 277
<211> 24
<212> DNA
<213> Homo sapiens

<400> 277
gcctttagg tccactatgg aatg 24

<210> 278
<211> 22
<212> DNA
<213> Homo sapiens

<400> 278
ccaggcgatg ctactactgg tc 22

<210> 279
<211> 22
<212> DNA

<213> Homo sapiens

<400> 279
tcatagcaca cctccctcac tg 22

<210> 280
<211> 22
<212> DNA
<213> Homo sapiens

<400> 280
acacaacaaa gagcttgtgc ag 22

<210> 281
<211> 24
<212> DNA
<213> Homo sapiens

<400> 281
ccattacttt gagaaggaca ggaa 24

<210> 282
<211> 22
<212> DNA
<213> Homo sapiens

<400> 282
tattcttgct ggatgcgttt ct 22

<210> 283
<211> 21
<212> DNA
<213> Homo sapiens

<400> 283
aggagggcag aggactagct g 21

<210> 284
<211> 20
<212> DNA

<213> Homo sapiens

<400> 284
ggcaatgtga atgtgcactg 20

<210> 285
<211> 20
<212> DNA
<213> Homo sapiens

<400> 285
cttgaacctg ggaggtggag 20

<210> 286
<211> 22
<212> DNA
<213> Homo sapiens

<400> 286
atcagggtgg gaggagtaaa ga 22

<210> 287
<211> 22
<212> DNA
<213> Homo sapiens

<400> 287
cccacttacc tctcacctgt gc 22

<210> 288
<211> 22
<212> DNA
<213> Homo sapiens

<400> 288
gtgaacttcc ggtaggaaat gg 22

<210> 289
<211> 20
<212> DNA

<213> Homo sapiens

<400> 289
aggggacctc aagggagaag 20

<210> 290
<211> 22
<212> DNA
<213> Homo sapiens

<400> 290
agatcatgcc agtgaactcc ag 22

<210> 291
<211> 22
<212> DNA
<213> Homo sapiens

<400> 291
ggaccaggaa agtccttgct tt 22

<210> 292
<211> 22
<212> DNA
<213> Homo sapiens

<400> 292
ggtggggaac attaaactga gg 22

<210> 293
<211> 22
<212> DNA
<213> Homo sapiens

<400> 293
gcttcaggtt gttttgttgc ag 22

<210> 294
<211> 22
<212> DNA

<213> Homo sapiens

<400> 294
accettgctt gagggaaata tg 22

<210> 295
<211> 22
<212> DNA
<213> Homo sapiens

<400> 295
cccagctcct agggtacagt ct 22

<210> 296
<211> 23
<212> DNA
<213> Homo sapiens

<400> 296
cagtcagctt caaaatccct ctt 23

<210> 297
<211> 24
<212> DNA
<213> Homo sapiens

<400> 297
tcacttccct gtgagtaaag aaaa 24

<210> 298
<211> 23
<212> DNA
<213> Homo sapiens

<400> 298
ggccatttaa ttcttgtcct tga 23

<210> 299
<211> 23
<212> DNA

<213> Homo sapiens

<400> 299
tggacttgtg caaactcaaa ctg 23

<210> 300
<211> 23
<212> DNA
<213> Homo sapiens

<400> 300
tcccaatata gggcagtcat gtt 23

<210> 301
<211> 23
<212> DNA
<213> Homo sapiens

<400> 301
tctcaatcag ttgagttgcc ttg 23

<210> 302
<211> 22
<212> DNA
<213> Homo sapiens

<400> 302
agctgtgcaa gtgtggaaac at 22

<210> 303
<211> 22
<212> DNA
<213> Homo sapiens

<400> 303
gctgtgaggg taaatgagac ca 22

<210> 304
<211> 22
<212> DNA

<213> Homo sapiens

<400> 304
gtctcctggt gagtgactgt gg 22

<210> 305
<211> 20
<212> DNA
<213> Homo sapiens

<400> 305
ccttccttcg tctccacagc 20

<210> 306
<211> 21
<212> DNA
<213> Homo sapiens

<400> 306
gtccttgtgc caacagtcga g 21

<210> 307
<211> 22
<212> DNA
<213> Homo sapiens

<400> 307
gcttggcaag gagaagagaa ca 22

<210> 308
<211> 23
<212> DNA
<213> Homo sapiens

<400> 308
gcttgctttc ttgcttgaac aac 23

<210> 309
<211> 22
<212> DNA

<213> Homo sapiens

<400> 309
gctggtcacc ttgagcttct ct 22

<210> 310
<211> 21
<212> DNA
<213> Homo sapiens

<400> 310
ccatgctggg ctctttgatt a 21

<210> 311
<211> 22
<212> DNA
<213> Homo sapiens

<400> 311
caccactctg aagttggcct ct 22

<210> 312
<211> 21
<212> DNA
<213> Homo sapiens

<400> 312
atggctctgc acatttgttc c 21

<210> 313
<211> 22
<212> DNA
<213> Homo sapiens

<400> 313
cagagtggga aaaggcactt ca 22

<210> 314
<211> 22
<212> DNA

<213> Homo sapiens

<400> 314
ccagagtcct gtgcagacat tc 22

<210> 315
<211> 22
<212> DNA
<213> Homo sapiens

<400> 315
atggggatta actgggatgt tg 22

<210> 316
<211> 23
<212> DNA
<213> Homo sapiens

<400> 316
cgtagctcca gacatcacta gca 23

<210> 317
<211> 22
<212> DNA
<213> Homo sapiens

<400> 317
gcaacctggt ctgcaaagtc tc 22

<210> 318
<211> 20
<212> DNA
<213> Homo sapiens

<400> 318
acccagcagt ccagcatgag 20

<210> 319
<211> 22
<212> DNA

<213> Homo sapiens

<400> 319
gagtttcagt gagccaagat cg 22

<210> 320
<211> 22
<212> DNA
<213> Homo sapiens

<400> 320
ttacaggctt gagccactag gc 22

<210> 321
<211> 21
<212> DNA
<213> Homo sapiens

<400> 321
aagcttccag gagacgaggt c 21

<210> 322
<211> 21
<212> DNA
<213> Homo sapiens

<400> 322
gtccctgaaa tccctcaaac c 21

<210> 323
<211> 23
<212> DNA
<213> Homo sapiens

<400> 323
tgctccataa acgtgactat tgc 23

<210> 324
<211> 22
<212> DNA

<213> Homo sapiens

<400> 324
gtaagagggt gggctggaat ct 22

<210> 325
<211> 22
<212> DNA
<213> Homo sapiens

<400> 325
cttagaccat gtccgggaaa ac 22

<210> 326
<211> 22
<212> DNA
<213> Homo sapiens

<400> 326
cacatcactc tggtgggtga ac 22

<210> 327
<211> 22
<212> DNA
<213> Homo sapiens

<400> 327
aaatttcatc ccaaaaccaa cc 22

<210> 328
<211> 23
<212> DNA
<213> Homo sapiens

<400> 328
ccagtcccaa gttcttgatc att 23

<210> 329
<211> 21
<212> DNA

<213> Homo sapiens

<400> 329
acaagtcggc tagttgcatg g 21

<210> 330
<211> 22
<212> DNA
<213> Homo sapiens

<400> 330
tctcagatga aaccaccagc ac 22

<210> 331
<211> 23
<212> DNA
<213> Homo sapiens

<400> 331
ttcatctgag aagcaaggag tgg 23

<210> 332
<211> 22
<212> DNA
<213> Homo sapiens

<400> 332
ccagggagaa agcaggactc ta 22

<210> 333
<211> 22
<212> DNA
<213> Homo sapiens

<400> 333
ttctggcggt gttttgaaat ta 22

<210> 334
<211> 22
<212> DNA

<213> Homo sapiens

<400> 334
ctcaacattg acggcctttc tt 22

<210> 335
<211> 23
<212> DNA
<213> Homo sapiens

<400> 335
tcagctctta aacagggcat agc 23

<210> 336
<211> 22
<212> DNA
<213> Homo sapiens

<400> 336
gaaatgcagc agccactaaa ga 22

<210> 337
<211> 22
<212> DNA
<213> Homo sapiens

<400> 337
ctcaccttca aggacctggt gt 22

<210> 338
<211> 21
<212> DNA
<213> Homo sapiens

<400> 338
cagggagggg tagaaaccac a 21

<210> 339
<211> 21
<212> DNA

<213> Homo sapiens

<400> 339
ggagaggtgg agaggcttca g 21

<210> 340
<211> 22
<212> DNA
<213> Homo sapiens

<400> 340
gagactccca ggacagacac ct 22

<210> 341
<211> 20
<212> DNA
<213> Homo sapiens

<400> 341
cactcgttcc tcacccttcc 20

<210> 342
<211> 22
<212> DNA
<213> Homo sapiens

<400> 342
aggactcaca cgtcactctg gt 22

<210> 343
<211> 23
<212> DNA
<213> Homo sapiens

<400> 343
ggacaatgtg atgaagattg ctg 23

<210> 344
<211> 22
<212> DNA

<213> Homo sapiens

<400> 344
atagcaggat cccaaaagac ca 22

<210> 345
<211> 22
<212> DNA
<213> Homo sapiens

<400> 345
ggcttgggga cctgtatttg ta 22

<210> 346
<211> 22
<212> DNA
<213> Homo sapiens

<400> 346
cagtggcctt tctgagcctt ac 22

<210> 347
<211> 23
<212> DNA
<213> Homo sapiens

<400> 347
gcactctagc tccctctttt agc 23

<210> 348
<211> 24
<212> DNA
<213> Homo sapiens

<400> 348
ttttacagta gagggcagac atgc 24

<210> 349
<211> 22
<212> DNA

<213> Homo sapiens

<400> 349
gccaccatag ctgcagaatt ag 22

<210> 350
<211> 22
<212> DNA
<213> Homo sapiens

<400> 350
cccaaggaca gatgtgatgc ta 22

<210> 351
<211> 22
<212> DNA
<213> Homo sapiens

<400> 351
gcctttgttc gagaggagtt gt 22

<210> 352
<211> 22
<212> DNA
<213> Homo sapiens

<400> 352
gttcacgctc tcaagcaggt ta 22

<210> 353
<211> 22
<212> DNA
<213> Homo sapiens

<400> 353
attccacaag ctctctccat ga 22

<210> 354
<211> 20
<212> DNA

<213> Homo sapiens

<400> 354
cttgccccaa gatgcctaag 20

<210> 355
<211> 23
<212> DNA
<213> Homo sapiens

<400> 355
tgcttggtat ttgctcatca tgt 23

<210> 356
<211> 22
<212> DNA
<213> Homo sapiens

<400> 356
cccttagcta gcccactgac aa 22

<210> 357
<211> 20
<212> DNA
<213> Homo sapiens

<400> 357
ctcctgggag tggtgtccaa 20

<210> 358
<211> 22
<212> DNA
<213> Homo sapiens

<400> 358
cctgggcaac agacagagta ag 22

<210> 359
<211> 21
<212> DNA

<213> Homo sapiens

<400> 359
cttcacttcc ccatgcgtac c 21

<210> 360
<211> 21
<212> DNA
<213> Homo sapiens

<400> 360
gggttcacaa tgcctacagg a 21

<210> 361
<211> 22
<212> DNA
<213> Homo sapiens

<400> 361
aaaatctgtg actttggctt gg 22

<210> 362
<211> 23
<212> DNA
<213> Homo sapiens

<400> 362
gggaggagac attctttgat ttg 23

<210> 363
<211> 22
<212> DNA
<213> Homo sapiens

<400> 363
gcagtcctga gaagaaaaca gc 22

<210> 364
<211> 22
<212> DNA

<213> Homo sapiens

<400> 364
cttcacatgc cccaaaatta ca 22

<210> 365
<211> 23
<212> DNA
<213> Homo sapiens

<400> 365
tgagccatgt atttcagagg tga 23

<210> 366
<211> 22
<212> DNA
<213> Homo sapiens

<400> 366
tacatttcag caggtgcgtg tt 22

<210> 367
<211> 24
<212> DNA
<213> Homo sapiens

<400> 367
ttgcctactc tgtagggata ttgc 24

<210> 368
<211> 21
<212> DNA
<213> Homo sapiens

<400> 368
atagggcatg tagcccagtg a 21

<210> 369
<211> 21
<212> DNA

<213> Homo sapiens

<400> 369
gctctgctgt tggtcctcac t 21

<210> 370
<211> 22
<212> DNA
<213> Homo sapiens

<400> 370
ttgcaaagca cacatcttct ga 22

<210> 371
<211> 22
<212> DNA
<213> Homo sapiens

<400> 371
tggcaatgtc aatgtcaagc at 22

<210> 372
<211> 22
<212> DNA
<213> Homo sapiens

<400> 372
gtatgttgcc ccactcaaca aa 22

<210> 373
<211> 22
<212> DNA
<213> Homo sapiens

<400> 373
tgcatttcct agctgagact cc 22

<210> 374
<211> 22
<212> DNA

<213> Homo sapiens

<400> 374
tgccatctcg cacgtagtaa at 22

<210> 375
<211> 22
<212> DNA
<213> Homo sapiens

<400> 375
ctctcctgtg ctgagcagct tt 22

<210> 376
<211> 22
<212> DNA
<213> Homo sapiens

<400> 376
tgtttccaat cactggcttt ca 22

<210> 377
<211> 20
<212> DNA
<213> Homo sapiens

<400> 377
gaaccatggg ctgtctctgg 20

<210> 378
<211> 22
<212> DNA
<213> Homo sapiens

<400> 378
atctgggata gcgaaggaga ca 22

<210> 379
<211> 20
<212> DNA

<213> Homo sapiens

<400> 379
attacaggcc acacgccatc 20

<210> 380
<211> 23
<212> DNA
<213> Homo sapiens

<400> 380
aaggcaagaa taagggagga aga 23

<210> 381
<211> 23
<212> DNA
<213> Homo sapiens

<400> 381
gctctcagga ctgcagaagt aca 23

<210> 382
<211> 22
<212> DNA
<213> Homo sapiens

<400> 382
gaggaaccaa tcccactcac ac 22

<210> 383
<211> 23
<212> DNA
<213> Homo sapiens

<400> 383
tcactctttg ccttctgtct ctg 23

<210> 384
<211> 22
<212> DNA

<213> Homo sapiens

<400> 384
gcactgtgct ttgctttctc ag 22

<210> 385
<211> 24
<212> DNA
<213> Homo sapiens

<400> 385
tgtctccttt atcgtaggtc tcca 24

<210> 386
<211> 22
<212> DNA
<213> Homo sapiens

<400> 386
caccacattt cctacagttc ca 22

<210> 387
<211> 22
<212> DNA
<213> Homo sapiens

<400> 387
cactgtgcac cagacagaca aa 22

<210> 388
<211> 24
<212> DNA
<213> Homo sapiens

<400> 388
tgtggttttc tgtatcagca gctt 24

<210> 389
<211> 22
<212> DNA

<213> Homo sapiens

<400> 389
cagggagtct gaaatcatca gg 22

<210> 390
<211> 23
<212> DNA
<213> Homo sapiens

<400> 390
tcaagtatct agccccaaat cca 23

<210> 391
<211> 24
<212> DNA
<213> Homo sapiens

<400> 391
ggcaatattg accattcatc attc 24

<210> 392
<211> 21
<212> DNA
<213> Homo sapiens

<400> 392
aggccaggag taagacgcaa c 21

<210> 393
<211> 22
<212> DNA
<213> Homo sapiens

<400> 393
aagaacgtac gtgtggtgtt gg 22

<210> 394
<211> 22
<212> DNA

<213> Homo sapiens

<400> 394
cgctatactt gctccatgca ct 22

<210> 395
<211> 21
<212> DNA
<213> Homo sapiens

<400> 395
aggaaacagc ctctggtcct c 21

<210> 396
<211> 22
<212> DNA
<213> Homo sapiens

<400> 396
gtcaatgctc agacagggag at 22

<210> 397
<211> 22
<212> DNA
<213> Homo sapiens

<400> 397
cccaggaagg caggtactgt ta 22

<210> 398
<211> 22
<212> DNA
<213> Homo sapiens

<400> 398
ttttacaacc accaagggtg tg 22

<210> 399
<211> 23
<212> DNA

<213> Homo sapiens

<400> 399
tcgtgtggtt acctccagat ttt 23

<210> 400
<211> 22
<212> DNA
<213> Homo sapiens

<400> 400
aaattagcca gggagtggag gt 22

<210> 401
<211> 20
<212> DNA
<213> Homo sapiens

<400> 401
catgccatgc tatggctcac 20

<210> 402
<211> 22
<212> DNA
<213> Homo sapiens

<400> 402
aggctgagcg gagttctaat tg 22

<210> 403
<211> 22
<212> DNA
<213> Homo sapiens

<400> 403
atctcagcaa tccacaggag gt 22

<210> 404
<211> 22
<212> DNA

<213> Homo sapiens

<400> 404
atttgcctca cgaacacatc at 22

<210> 405
<211> 23
<212> DNA
<213> Homo sapiens

<400> 405
tggaaagttg tctatggcac ctc 23

<210> 406
<211> 22
<212> DNA
<213> Homo sapiens

<400> 406
atgggcagca aggacttact ct 22

<210> 407
<211> 22
<212> DNA
<213> Homo sapiens

<400> 407
caccccaata ttgtctgcct tc 22

<210> 408
<211> 22
<212> DNA
<213> Homo sapiens

<400> 408
ggctcgggaa catgtaatta gg 22

<210> 409
<211> 24
<212> DNA

<213> Homo sapiens

<400> 409
ccatcatgta tggcaaattc tctt 24

<210> 410
<211> 23
<212> DNA
<213> Homo sapiens

<400> 410
tggcgtctcc tagtaaagat gct 23

<210> 411
<211> 21
<212> DNA
<213> Homo sapiens

<400> 411
gccagattgc tggtttcatt g 21

<210> 412
<211> 22
<212> DNA
<213> Homo sapiens

<400> 412
ggctaaaaca caaagcacca tt 22

<210> 413
<211> 22
<212> DNA
<213> Homo sapiens

<400> 413
gggaagtcat ccacaaagac ct 22

<210> 414
<211> 20
<212> DNA

<213> Homo sapiens

<400> 414
ggtctgggtc acagctcctc 20

<210> 415
<211> 22
<212> DNA
<213> Homo sapiens

<400> 415
ttcttctgcc aagatgtggt gt 22

<210> 416
<211> 22
<212> DNA
<213> Homo sapiens

<400> 416
tgcagatgct gcaatcatgt ta 22

<210> 417
<211> 22
<212> DNA
<213> Homo sapiens

<400> 417
tggaccccga aagataaata gg 22

<210> 418
<211> 22
<212> DNA
<213> Homo sapiens

<400> 418
ttctgcactc ctctggaaac tg 22

<210> 419
<211> 22
<212> DNA

<213> Homo sapiens

<400> 419
gggtgagagc caacactgat ct 22

<210> 420
<211> 22
<212> DNA
<213> Homo sapiens

<400> 420
ctgtgccctc tcatctcaca ct 22

<210> 421
<211> 22
<212> DNA
<213> Homo sapiens

<400> 421
agaacctagc ctccaagatt gc 22

<210> 422
<211> 22
<212> DNA
<213> Homo sapiens

<400> 422
acaccttcca agactccttc ca 22

<210> 423
<211> 20
<212> DNA
<213> Homo sapiens

<400> 423
gactcgaggg tgggagacag 20

<210> 424
<211> 22
<212> DNA

<213> Homo sapiens

<400> 424
gctgtcacta ggtgtcctga gc 22

<210> 425
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 425
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctga 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 426
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 426
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctga 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 427
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 427
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 428
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 428
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 429
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 429
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg ccccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 430
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 430
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 431
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 431
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 432
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 432
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 433
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 433
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 434
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 434
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga ccccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 435
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 435
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 436
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 436
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga ccccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 437
<211> 3854
<212> DNA
<213> Homo sapiens

<400> 437
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacact cgatgaagcc tacgtgatgg 2520

ccagcgtgga caacccccac gtgtgccgcc tgctgggcat ctgcctcacc tccaccgtgc 2580

agctcatcac gcagctcatg cccttcggct gcctcctgga ctatgtccgg gaacacaaag 2640

acaatattgg ctcccagtac ctgctcaact ggtgtgtgca gatcgcaaag ggcatgaact 2700

acttggagga ccgtcgcttg gtgcaccgcg acctggcagc caggaacgta ctggtgaaaa 2760

caccgcagca tgtcaagatc acagattttg ggctggccaa actgctgggt gcggaagaga 2820

aagaatacca tgcagaagga ggcaaagtgc ctatcaagtg gatggcattg gaatcaattt 2880

tacacagaat ctatacccac cagagtgatg tctggagcta cggggtgact gtttgggagt 2940

tgatgacctt tggatccaag ccatatgacg gaatccctgc cagcgagatc tcctccatcc 3000

tggagaaagg agaacgcctc cctcagccac ccatatgtac catcgatgtc tacatgatca 3060

tggtcaagtg ctggatgata gacgcagata gtcgcccaaa gttccgtgag ttgatcatcg 3120

aattctccaa aatggcccga gacccccagc gctaccttgt cattcagggg gatgaaagaa 3180

tgcatttgcc aagtcctaca gactccaact tctaccgtgc cctgatggat gaagaagaca 3240

tggacgacgt ggtggatgcc gacgagtacc tcatcccaca gcagggcttc ttcagcagcc 3300

cctccacgtc acggactccc ctcctgagct ctctgagtgc aaccagcaac aattccaccg 3360

tggcttgcat tgatagaaat gggctgcaaa gctgtcccat caaggaagac agcttcttgc 3420

agcgatacag ctcagacccc acaggcgcct tgactgagga cagcatagac gacaccttcc 3480

tcccagtgcc tgaatacata aaccagtccg ttcccaaaag gcccgctggc tctgtgcaga 3540

atcctgtcta tcacaatcag cctctgaacc ccgcgcccag cagagaccca cactaccagg 3600

acccccacag cactgcagtg ggcaaccccg agtatctcaa cactgtccag cccacctgtg 3660

tcaacagcac attcgacagc cctgcccact gggcccagaa aggcagccac caaattagcc 3720

tggacaaccc tgactaccag caggacttct ttcccaagga agccaagcca aatggcatct 3780

ttaagggctc cacagctgaa aatgcagaat acctaagggt cgcgccacaa agcagtgaat 3840

ttattggagc atga 3854

<210> 438
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 438
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcggg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 439
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 439
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcggg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 440
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 440
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcggg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 441
<211> 3869
<212> DNA
<213> Homo sapiens

<400> 441
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaaccaacat ctccgaaagc caacaaggaa atcctcgatg 2520

aagcctacgt gatggccagc gtggacaacc cccacgtgtg ccgcctgctg ggcatctgcc 2580

tcacctccac cgtgcagctc atcacgcagc tcatgcccttcggctgcctc ctggactatg 2640

tccgggaaca caaagacaat attggctccc agtacctgct caactggtgt gtgcagatcg 2700

caaagggcat gaactacttg gaggaccgtc gcttggtgca ccgcgacctg gcagccagga 2760

acgtactggt gaaaacaccg cagcatgtca agatcacaga ttttgggctg gccaaactgc 2820

tgggtgcgga agagaaagaa taccatgcag aaggaggcaa agtgcctatc aagtggatgg 2880

cattggaatc aattttacac agaatctata cccaccagag tgatgtctgg agctacgggg 2940

tgactgtttg ggagttgatg acctttggat ccaagccata tgacggaatc cctgccagcg 3000

agatctcctc catcctggag aaaggagaac gcctccctca gccacccata tgtaccatcg 3060

atgtctacat gatcatggtc aagtgctgga tgatagacgc agatagtcgc ccaaagttcc 3120

gtgagttgat catcgaattc tccaaaatgg cccgagaccc ccagcgctac cttgtcattc 3180

agggggatga aagaatgcat ttgccaagtc ctacagactc caacttctac cgtgccctga 3240

tggatgaaga agacatggac gacgtggtgg atgccgacga gtacctcatc ccacagcagg 3300

gcttcttcag cagcccctcc acgtcacgga ctcccctcct gagctctctg agtgcaacca 3360

gcaacaattc caccgtggct tgcattgata gaaatgggct gcaaagctgt cccatcaagg 3420

aagacagctt cttgcagcga tacagctcag accccacagg cgccttgact gaggacagca 3480

tagacgacac cttcctccca gtgcctgaat acataaacca gtccgttccc aaaaggcccg 3540

ctggctctgt gcagaatcct gtctatcaca atcagcctct gaaccccgcg cccagcagag 3600

acccacacta ccaggacccc cacagcactg cagtgggcaa ccccgagtat ctcaacactg 3660

tccagcccac ctgtgtcaac agcacattcg acagccctgc ccactgggcc cagaaaggca 3720

gccaccaaat tagcctggac aaccctgact accagcagga cttctttccc aaggaagcca 3780

agccaaatgg catctttaag ggctccacag ctgaaaatgc agaataccta agggtcgcgc 3840

cacaaagcag tgaatttatt ggagcatga 3869

<210> 442
<211> 3869
<212> DNA
<213> Homo sapiens

<400> 442
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaaccaacat ctccgaaagc caacaaggaa atcctcgatg 2520

aagcctacgt gatggccagc gtggacaacc cccacgtgtg ccgcctgctg ggcatctgcc 2580

tcacctccac cgtgcagctc atcacgcagc tcatgcccct cggctgcctc ctggactatg 2640

tccgggaaca caaagacaat attggctccc agtacctgct caactggtgt gtgcagatcg 2700

caaagggcat gaactacttg gaggaccgtc gcttggtgca ccgcgacctg gcagccagga 2760

acgtactggt gaaaacaccg cagcatgtca agatcacaga ttttgggctg gccaaactgc 2820

tgggtgcgga agagaaagaa taccatgcag aaggaggcaa agtgcctatc aagtggatgg 2880

cattggaatc aattttacac agaatctata cccaccagag tgatgtctgg agctacgggg 2940

tgactgtttg ggagttgatg acctttggat ccaagccata tgacggaatc cctgccagcg 3000

agatctcctc catcctggag aaaggagaac gcctccctca gccacccata tgtaccatcg 3060

atgtctacat gatcatggtc aagtgctgga tgatagacgc agatagtcgc ccaaagttcc 3120

gtgagttgat catcgaattc tccaaaatgg cccgagaccc ccagcgctac cttgtcattc 3180

agggggatga aagaatgcat ttgccaagtc ctacagactc caacttctac cgtgccctga 3240

tggatgaaga agacatggac gacgtggtgg atgccgacga gtacctcatc ccacagcagg 3300

gcttcttcag cagcccctcc acgtcacgga ctcccctcct gagctctctg agtgcaacca 3360

gcaacaattc caccgtggct tgcattgata gaaatgggct gcaaagctgt cccatcaagg 3420

aagacagctt cttgcagcga tacagctcag accccacagg cgccttgact gaggacagca 3480

tagacgacac cttcctccca gtgcctgaat acataaacca gtccgttccc aaaaggcccg 3540

ctggctctgt gcagaatcct gtctatcaca atcagcctct gaaccccgcg cccagcagag 3600

acccacacta ccaggacccc cacagcactg cagtgggcaa ccccgagtat ctcaacactg 3660

tccagcccac ctgtgtcaac agcacattcg acagccctgc ccactgggcc cagaaaggca 3720

gccaccaaat tagcctggac aaccctgact accagcagga cttctttccc aaggaagcca 3780

agccaaatgg catctttaag ggctccacag ctgaaaatgc agaataccta agggtcgcgc 3840

cacaaagcag tgaatttatt ggagcatga 3869

<210> 443
<211> 3860
<212> DNA
<213> Homo sapiens

<400> 443
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaatcgaaag ccaacaagga aatcctcgat gaagcctacg 2520

tgatggccag cgtggacaac ccccacgtgt gccgcctgct gggcatctgc ctcacctcca 2580

ccgtgcagct catcacgcag ctcatgccct tcggctgcct cctggactat gtccgggaac 2640

acaaagacaa tattggctcc cagtacctgc tcaactggtg tgtgcagatc gcaaagggca 2700

tgaactactt ggaggaccgt cgcttggtgc accgcgacct ggcagccagg aacgtactgg 2760

tgaaaacacc gcagcatgtc aagatcacag attttgggct ggccaaactg ctgggtgcgg 2820

aagagaaaga ataccatgca gaaggaggca aagtgcctat caagtggatg gcattggaat 2880

caattttaca cagaatctat acccaccaga gtgatgtctg gagctacggg gtgactgttt 2940

gggagttgat gacctttgga tccaagccat atgacggaat ccctgccagc gagatctcct 3000

ccatcctgga gaaaggagaa cgcctccctc agccacccat atgtaccatc gatgtctaca 3060

tgatcatggt caagtgctgg atgatagacg cagatagtcg cccaaagttc cgtgagttga 3120

tcatcgaatt ctccaaaatg gcccgagacc cccagcgcta ccttgtcatt caggggggatg 3180

aaagaatgca tttgccaagt cctacagact ccaacttcta ccgtgccctg atggatgaag 3240

aagacatgga cgacgtggtg gatgccgacg agtacctcat cccacagcag ggcttcttca 3300

gcagcccctc cacgtcacgg actcccctcc tgagctctct gagtgcaacc agcaacaatt 3360

ccaccgtggc ttgcattgat agaaatgggc tgcaaagctg tcccatcaag gaagacagct 3420

tcttgcagcg atacagctca gaccccacag gcgccttgac tgaggacagc atagacgaca 3480

ccttcctccc agtgcctgaa tacataaacc agtccgttcc caaaaggccc gctggctctg 3540

tgcagaatcc tgtctatcac aatcagcctc tgaaccccgc gcccagcaga gacccacact 3600

accaggaccc ccacagcact gcagtgggca accccgagta tctcaacact gtccagccca 3660

cctgtgtcaa cagcacattc gacagccctg cccactgggc ccagaaaggc agccaccaaa 3720

ttagcctgga caaccctgac taccagcagg acttctttcc caaggaagcc aagccaaatg 3780

gcatctttaa gggctccaca gctgaaaatg cagaatacct aagggtcgcg ccacaaagca 3840

gtgaatttat tggagcatga 3860

<210> 444
<211> 3860
<212> DNA
<213> Homo sapiens

<400> 444
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gttccgaaag ccaacaagga aatcctcgat gaagcctacg 2520

tgatggccag cgtggacaac ccccacgtgt gccgcctgct gggcatctgc ctcacctcca 2580

ccgtgcagct catcacgcag ctcatgccct tcggctgcct cctggactat gtccgggaac 2640

acaaagacaa tattggctcc cagtacctgc tcaactggtg tgtgcagatc gcaaagggca 2700

tgaactactt ggaggaccgt cgcttggtgc accgcgacct ggcagccagg aacgtactgg 2760

tgaaaacacc gcagcatgtc aagatcacag attttgggct ggccaaactg ctgggtgcgg 2820

aagagaaaga ataccatgca gaaggaggca aagtgcctat caagtggatg gcattggaat 2880

caattttaca cagaatctat acccaccaga gtgatgtctg gagctacggg gtgactgttt 2940

gggagttgat gacctttgga tccaagccat atgacggaat ccctgccagc gagatctcct 3000

ccatcctgga gaaaggagaa cgcctccctc agccacccat atgtaccatc gatgtctaca 3060

tgatcatggt caagtgctgg atgatagacg cagatagtcg cccaaagttc cgtgagttga 3120

tcatcgaatt ctccaaaatg gcccgagacc cccagcgcta ccttgtcatt caggggggatg 3180

aaagaatgca tttgccaagt cctacagact ccaacttcta ccgtgccctg atggatgaag 3240

aagacatgga cgacgtggtg gatgccgacg agtacctcat cccacagcag ggcttcttca 3300

gcagcccctc cacgtcacgg actcccctcc tgagctctct gagtgcaacc agcaacaatt 3360

ccaccgtggc ttgcattgat agaaatgggc tgcaaagctg tcccatcaag gaagacagct 3420

tcttgcagcg atacagctca gaccccacag gcgccttgac tgaggacagc atagacgaca 3480

ccttcctccc agtgcctgaa tacataaacc agtccgttcc caaaaggccc gctggctctg 3540

tgcagaatcc tgtctatcac aatcagcctc tgaaccccgc gcccagcaga gacccacact 3600

accaggaccc ccacagcact gcagtgggca accccgagta tctcaacact gtccagccca 3660

cctgtgtcaa cagcacattc gacagccctg cccactgggc ccagaaaggc agccaccaaa 3720

ttagcctgga caaccctgac taccagcagg acttctttcc caaggaagcc aagccaaatg 3780

gcatctttaa gggctccaca gctgaaaatg cagaatacct aagggtcgcg ccacaaagca 3840

gtgaatttat tggagcatga 3860

<210> 445
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 445
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcggg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 446
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 446
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcggg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 447
<211> 36
<212> DNA
<213> Homo sapiens

<400> 447
tgtaaaacga cggccagtcg cccagaccgg acgaca 36

<210> 448
<211> 39
<212> DNA
<213> Homo sapiens

<400> 448
caggaaacag ctatgaccag ggcaatgagg acataacca 39

<210> 449
<211> 38
<212> DNA
<213> Homo sapiens

<400> 449
tgtaaaacga cggccagtgg tggtccttgg gaatttgg 38

<210> 450
<211> 40
<212> DNA
<213> Homo sapiens

<400> 450
caggaaacag ctatgacccc atcgacatgt tgctgagaaa 40

<210> 451
<211> 38
<212> DNA
<213> Homo sapiens

<400> 451
tgtaaaacga cggccagtga aggagctgcc catgagaa 38

<210> 452
<211> 38
<212> DNA
<213> Homo sapiens

<400> 452
caggaaacag ctatgacccg tggcttcgtc tcggaatt 38

<210> 453
<211> 40
<212> DNA
<213> Homo sapiens

<400> 453
tgtaaaacga cggccagtga aactgaccaa aatcatctgt 40

<210> 454
<211> 40
<212> DNA
<213> Homo sapiens

<400> 454
caggaaacag ctatgaccta cctattccgt tacacacttt 40

<210> 455
<211> 40
<212> DNA
<213> Homo sapiens

<400> 455
tgtaaaacga cggccagtcc gtaattatgt ggtgacagat 40

<210> 456
<211> 40
<212> DNA
<213> Homo sapiens

<400> 456
caggaaacag ctatgaccgc gtatgatttc taggttctca 40

<210> 457
<211> 41
<212> DNA
<213> Homo sapiens

<400> 457
tgtaaaacga cggccagtct gaaaaccgta aaggaaatca c 41

<210> 458
<211> 37
<212> DNA
<213> Homo sapiens

<400> 458
caggaaacag ctatgacccc tgcctcggct gacattc 37

<210> 459
<211> 40
<212> DNA
<213> Homo sapiens

<400> 459
tgtaaaacga cggccagtta agcaacagag gtgaaaacag 40

<210> 460
<211> 40
<212> DNA
<213> Homo sapiens

<400> 460
caggaaacag ctatgaccgg tgttgttttc tcccatgact 40

<210> 461
<211> 38
<212> DNA
<213> Homo sapiens

<400> 461
tgtaaaacga cggccagtgg accagacaac tgtatcca 38

<210> 462
<211> 40
<212> DNA
<213> Homo sapiens

<400> 462
caggaaacag ctatgacctt ccttcaagat cctcaagaga 40

<210> 463
<211> 38
<212> DNA
<213> Homo sapiens

<400> 463
tgtaaaacga cggccagtga tcggcctctt catgcgaa 38

<210> 464
<211> 38
<212> DNA
<213> Homo sapiens

<400> 464
caggaaacag ctatgaccac ggtggaggtg aggcagat 38

<210> 465
<211> 39
<212> DNA
<213> Homo sapiens

<400> 465
tgtaaaacga cggccagtcg aaagccaaca aggaaatcc 39

<210> 466
<211> 40
<212> DNA
<213> Homo sapiens

<400> 466
caggaaacag ctatgaccat tccaatgcca tccacttgat 40

<210> 467
<211> 39
<212> DNA
<213> Homo sapiens

<400> 467
tgtaaaacga cggccagtaa caccgcagca tgtcaagat 39

<210> 468
<211> 39
<212> DNA
<213> Homo sapiens

<400> 468
caggaaacag ctatgaccct cgggccattt tggagaatt 39

<210> 469
<211> 39
<212> DNA
<213> Homo sapiens

<400> 469
tgtaaaacga cggccagttc agccacccat atgtaccat 39

<210> 470
<211> 39
<212> DNA
<213> Homo sapiens

<400> 470
caggaaacag ctatgaccgc tttgcagccc atttctatc 39

<210> 471
<211> 38
<212> DNA
<213> Homo sapiens

<400> 471
tgtaaaacga cggccagtac agcagggctt cttcagca 38

<210> 472
<211> 38
<212> DNA
<213> Homo sapiens

<400> 472
caggaaacag ctatgacctg acacaggtgg gctggaca 38

<210> 473
<211> 40
<212> DNA
<213> Homo sapiens

<400> 473
tgtaaaacga cggccagtga atcctgtcta tcacaatcag 40

<210> 474
<211> 40
<212> DNA
<213> Homo sapiens

<400> 474
caggaaacag ctatgaccgg tatcgaaaga gtctggattt 40

<210> 475
<211> 38
<212> DNA
<213> Homo sapiens

<400> 475
tgtaaaacga cggccagtgc tccacagctg aaaatgca 38

<210> 476
<211> 39
<212> DNA
<213> Homo sapiens

<400> 476
caggaaacag ctatgaccac gttgcaaaac cagtctgtg 39

<210> 477
<211> 25
<212> PRT
<213> Homo sapiens

<400> 477
Lys Thr Pro Gln His Val Lys Ile Thr Asp Phe Gly Arg Ala Lys Leu
1 5 10 15

Leu Gly Ala Glu Glu Lys Glu Tyr His
20 25

<210> 478
<211> 25
<212> PRT
<213> Homo sapiens

<400> 478
Lys Thr Pro Gln His Val Lys Ile Thr Asp Phe Gly Leu Ala Lys Leu
1 5 10 15

Leu Gly Ala Glu Glu Lys Glu Tyr His
20 25

<210> 479
<211> 25
<212> PRT
<213> Homo sapiens

<400> 479
His Glu Asp Leu Thr Val Lys Ile Gly Asp Phe Gly Leu Ala Thr Val
1 5 10 15

Lys Ser Arg Trp Ser Gly Ser His Gln
20 25

<210> 480
<211> 25
<212> PRT

<213> Homo sapiens

<400> 480
Glu Thr Glu Phe Lys Lys Ile Lys Val Leu Ser Ser Gly Ala Phe Gly
1 5 10 15

Thr Val Tyr Lys Gly Leu Trp Ile Pro
20 25

<210> 481
<211> 25
<212> PRT
<213> Homo sapiens

<400> 481
Glu Thr Glu Phe Lys Lys Ile Lys Val Leu Gly Ser Gly Ala Phe Gly
1 5 10 15

Thr Val Tyr Lys Gly Leu Trp Ile Pro
20 25

<210> 482
<211> 25
<212> PRT
<213> Homo sapiens

<400> 482
Asp Gly Gln Ile Thr Val Gly Gln Arg Ile Gly Ser Gly Ser Phe Gly
1 5 10 15

Thr Val Tyr Lys Gly Lys Trp His Gly
20 25

<210> 483
<211> 19
<212> PRT
<213> Homo sapiens

<400> 483
Val Ala Ile Lys Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu Asp Glu

1 5 10 15

Ala Tyr Val

<210> 484
<211> 16
<212> PRT
<213> Homo sapiens

<400> 484
Val Ala Ile Lys Glu Leu Arg Glu Ala Thr Leu Asp Glu Ala Tyr Val
1 5 10 15

<210> 485
<211> 21
<212> PRT
<213> Homo sapiens

<400> 485
Val Ala Ile Lys Glu Pro Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu
1 5 10 15

Asp Glu Ala Tyr Val
20

<210> 486
<211> 18
<212> PRT
<213> Homo sapiens

<400> 486
Val Ala Ile Lys Glu Ser Lys Ala Asn Lys Glu Ile Leu Asp Glu Ala
1 5 10 15

Tyr Val

<210> 487
<211> 18

<212> PRT
<213> Homo sapiens

<400> 487
Val Ala Ile Lys Val Pro Lys Ala Asn Lys Glu Ile Leu Asp Glu Ala
1 5 10 15

Tyr Val

<210> 488
<211> 24
<212> PRT
<213> Homo sapiens

<400> 488
Val Ala Ile Lys Glu Leu Arg Glu Ala Thr Ser Pro Lys Ala Asn Lys
1 5 10 15

Glu Ile Leu Asp Glu Ala Tyr Val
20

<210> 489
<211> 25
<212> PRT
<213> Homo sapiens

<400> 489
Val Ala Val Lys Met Leu Asn Val Thr Ala Pro Thr Pro Gln Gln Leu
1 5 10 15

Gln Ala Phe Lys Asn Glu Val Gly Val
20 25

<210> 490
<211> 6
<212> PRT
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
illustrative mutation motif

<220>
<221> MOD_RES
<222> (2)..(2)
<223> Any amino acid

<220>
<221> MOD_RES
<222> (4)..(5)
<223> Any amino acid

<400> 490
Gly Xaa Gly Xaa Xaa Gly
1 5

<210> 491
<211> 252
<212> PRT
<213> Homo sapiens

<400> 491
Ile Thr Met Lys His Lys Leu Gly Gly Gly Gln Tyr Gly Glu Val Tyr
1 5 10 15

Glu Gly Val Trp Lys Lys Tyr Ser Leu Thr Val Ala Val Lys Thr Leu
20 25 30

Lys Glu Asp Thr Met Glu Val Glu Glu Phe Leu Lys Glu Ala Ala Val
35 40 45

Met Lys Glu Ile Lys His Pro Asn Leu Val Gln Leu Leu Gly Val Cys
50 55 60

Thr Arg Glu Pro Pro Phe Tyr Ile Ile Thr Glu Phe Met Thr Tyr Gly
65 70 75 80

Asn Leu Leu Asp Tyr Leu Arg Glu Cys Asn Arg Gln Glu Val Asn Ala

85 90 95

Val Val Leu Leu Tyr Met Ala Thr Gln Ile Ser Ser Ala Met Glu Tyr
100 105 110

Leu Glu Lys Lys Asn Phe Ile His Arg Asp Leu Ala Ala Arg Asn Cys
115 120 125

Leu Val Gly Glu Asn His Leu Val Lys Val Ala Asp Phe Gly Leu Ser
130 135 140

Arg Leu Met Thr Gly Asp Thr Tyr Thr Ala His Ala Gly Ala Lys Phe
145 150 155 160

Pro Ile Lys Trp Thr Ala Pro Glu Ser Leu Ala Tyr Asn Lys Phe Ser
165 170 175

Ile Lys Ser Asp Val Trp Ala Phe Gly Val Leu Leu Trp Glu Ile Ala
180 185 190

Thr Tyr Gly Met Ser Pro Tyr Pro Gly Ile Asp Leu Ser Gln Val Tyr
195 200 205

Glu Leu Leu Glu Lys Asp Tyr Arg Met Glu Arg Pro Glu Gly Cys Pro
210 215 220

Glu Lys Val Tyr Glu Leu Met Arg Ala Cys Trp Gln Trp Asn Pro Ser
225 230 235 240

Asp Arg Pro Ser Phe Ala Glu Ile His Gln Ala Phe
245 250

<210> 492
<211> 268
<212> PRT
<213> Homo sapiens

<400> 492

```
Phe Lys Lys Ile Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr
1               5                  10                  15

Lys Gly Leu Trp Ile Pro Glu Gly Glu Lys Val Lys Ile Pro Val Ala
            20                  25                  30

Ile Lys Glu Leu Arg Glu Ala Thr Ser Pro Lys Ala Asn Lys Glu Ile
        35                  40                  45

Leu Asp Glu Ala Tyr Val Met Ala Ser Val Asp Asn Pro His Val Cys
    50                  55                  60

Arg Leu Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Ile Thr Gln
65                  70                  75                  80

Leu Met Pro Phe Gly Cys Leu Leu Asp Tyr Val Arg Glu His Lys Asp
                85                  90                  95

Asn Ile Gly Ser Gln Tyr Leu Leu Asn Trp Cys Val Gln Ile Ala Lys
            100                 105                 110

Gly Met Asn Tyr Leu Glu Asp Arg Arg Leu Val His Arg Asp Leu Ala
        115                 120                 125

Ala Arg Asn Val Leu Val Lys Thr Pro Gln His Val Lys Ile Thr Asp
    130                 135                 140

Phe Gly Leu Ala Lys Leu Leu Gly Ala Glu Glu Lys Glu Tyr His Ala
145                 150                 155                 160

Glu Gly Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu
                165                 170                 175

His Arg Ile Tyr Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr
            180                 185                 190
```

Val Trp Glu Leu Met Thr Phe Gly Ser Lys Pro Tyr Asp Gly Ile Pro
195 200 205

Ala Ser Glu Ile Ser Ser Ile Leu Glu Lys Gly Glu Arg Leu Pro Gln
210 215 220

Pro Pro Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp
225 230 235 240

Met Ile Asp Ala Asp Ser Arg Pro Lys Phe Arg Glu Leu Ile Ile Glu
245 250 255

Phe Ser Lys Met Ala Arg Asp Pro Gln Arg Tyr Leu
260 265

<210> 493
<211> 54
<212> DNA
<213> Homo sapiens

<400> 493
aaaattcccg tcgctatcaa ggaattaaga gaagcaacat ctccgaaagc caac 54

<210> 494
<211> 24
<212> DNA
<213> Homo sapiens

<400> 494
tttgggctgg ccaaactgct gggt 24

<210> 495
<211> 39
<212> DNA
<213> Homo sapiens

<400> 495
aaaattcccg tcgctatcaa aacatctccg aaagccaac 39

<210> 496
<211> 24
<212> DNA
<213> Homo sapiens

<400> 496
tttgggctgg ccaaactgct gggt 24

<210> 497
<211> 42
<212> DNA
<213> Homo sapiens

<400> 497
aaaattcccg tcgctatcaa ggaatcatct ccgaaagcca ac 42

<210> 498
<211> 24
<212> DNA
<213> Homo sapiens

<400> 498
tttgggctgg ccaaactgct gggt 24

<210> 499
<211> 36
<212> DNA
<213> Homo sapiens

<400> 499
aaaattcccg tcgctatcaa ggaatcgaaa gccaac 36

<210> 500
<211> 24
<212> DNA
<213> Homo sapiens

<400> 500
tttgggctgg ccaaactgct gggt 24

<210> 501

<211> 54
<212> DNA
<213> Homo sapiens

<400> 501
aaaattcccg tcgctatcaa ggaattaaga gaagcaacat ctccgaaagc caac 54

<210> 502
<211> 24
<212> DNA
<213> Homo sapiens

<400> 502
tttgggcggg ccaaactgct gggt 24

<210> 503
<211> 54
<212> DNA
<213> Homo sapiens

<400> 503
aaaattcccg tcgctatcaa ggaattaaga gaagcaacat ctccgaaagc caac 54

<210> 504
<211> 24
<212> DNA
<213> Homo sapiens

<400> 504
tttgggctgg ccaaacagct gggt 24

<210> 505
<211> 25
<212> DNA
<213> Homo sapiens

<400> 505
gcaatatcag ccttaggtgc ggctc 25

<210> 506
<211> 26

<212> DNA
<213> Homo sapiens

<400> 506
catagaaagt gaacatttag gatgtg 26

<210> 507
<211> 24
<212> DNA
<213> Homo sapiens

<400> 507
ctaacgttcg ccagccataa gtcc 24

<210> 508
<211> 26
<212> DNA
<213> Homo sapiens

<400> 508
gctgcgagct cacccagaat gtctgg 26

<210> 509
<211> 18
<212> PRT
<213> Homo sapiens

<400> 509
Lys Ile Pro Val Ala Ile Lys Glu Leu Arg Glu Ala Thr Ser Pro Lys
1 5 10 15

Ala Asn

<210> 510
<211> 8
<212> PRT
<213> Homo sapiens

<400> 510
Phe Gly Leu Ala Lys Leu Leu Gly

1 5

<210> 511
<211> 3633
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(3630)

<400> 511
atg cga ccc tcc ggg acg gcc ggg gca gcg ctc ctg gcg ctg ctg gct 48
Met Arg Pro Ser Gly Thr Ala Gly Ala Ala Leu Leu Ala Leu Leu Ala
1 5 10 15

gcg ctc tgc ccg gcg agt cgg gct ctg gag gaa aag aaa gtt tgc caa 96
Ala Leu Cys Pro Ala Ser Arg Ala Leu Glu Glu Lys Lys Val Cys Gln
20 25 30

ggc acg agt aac aag ctc acg cag ttg ggc act ttt gaa gat cat ttt 144
Gly Thr Ser Asn Lys Leu Thr Gln Leu Gly Thr Phe Glu Asp His Phe
35 40 45

ctc agc ctc cag agg atg ttc aat aac tgt gag gtg gtc ctt ggg aat 192
Leu Ser Leu Gln Arg Met Phe Asn Asn Cys Glu Val Val Leu Gly Asn
50 55 60

ttg gaa att acc tat gtg cag agg aat tat gat ctt tcc ttc tta aag 240
Leu Glu Ile Thr Tyr Val Gln Arg Asn Tyr Asp Leu Ser Phe Leu Lys
65 70 75 80

acc atc cag gag gtg gct ggt tat gtc ctc att gcc ctc aac aca gtg 288
Thr Ile Gln Glu Val Ala Gly Tyr Val Leu Ile Ala Leu Asn Thr Val
85 90 95

gag cga att cct ttg gaa aac ctg cag atc atc aga gga aat atg tac 336
Glu Arg Ile Pro Leu Glu Asn Leu Gln Ile Ile Arg Gly Asn Met Tyr
100 105 110

tac gaa aat tcc tat gcc tta gca gtc tta tct aac tat gat gca aat 384
Tyr Glu Asn Ser Tyr Ala Leu Ala Val Leu Ser Asn Tyr Asp Ala Asn
115 120 125

```
aaa acc gga ctg aag gag ctg ccc atg aga aat tta cag gaa atc ctg        432
Lys Thr Gly Leu Lys Glu Leu Pro Met Arg Asn Leu Gln Glu Ile Leu
    130                 135                 140

cat ggc gcc gtg cgg ttc agc aac aac cct gcc ctg tgc aac gtg gag        480
His Gly Ala Val Arg Phe Ser Asn Asn Pro Ala Leu Cys Asn Val Glu
145                 150                 155                 160

agc atc cag tgg cgg gac ata gtc agc agt gac ttt ctc agc aac atg        528
Ser Ile Gln Trp Arg Asp Ile Val Ser Ser Asp Phe Leu Ser Asn Met
                165                 170                 175

tcg atg gac ttc cag aac cac ctg ggc agc tgc caa aag tgt gat cca        576
Ser Met Asp Phe Gln Asn His Leu Gly Ser Cys Gln Lys Cys Asp Pro
            180                 185                 190

agc tgt ccc aat ggg agc tgc tgg ggt gca gga gag gag aac tgc cag        624
Ser Cys Pro Asn Gly Ser Cys Trp Gly Ala Gly Glu Glu Asn Cys Gln
        195                 200                 205

aaa ctg acc aaa atc atc tgt gcc cag cag tgc tcc ggg cgc tgc cgt        672
Lys Leu Thr Lys Ile Ile Cys Ala Gln Gln Cys Ser Gly Arg Cys Arg
    210                 215                 220

ggc aag tcc ccc agt gac tgc tgc cac aac cag tgt gct gca ggc tgc        720
Gly Lys Ser Pro Ser Asp Cys Cys His Asn Gln Cys Ala Ala Gly Cys
225                 230                 235                 240

aca ggc ccc cgg gag agc gac tgc ctg gtc tgc cgc aaa ttc cga gac        768
Thr Gly Pro Arg Glu Ser Asp Cys Leu Val Cys Arg Lys Phe Arg Asp
                245                 250                 255

gaa gcc acg tgc aag gac acc tgc ccc cca ctc atg ctc tac aac ccc        816
Glu Ala Thr Cys Lys Asp Thr Cys Pro Pro Leu Met Leu Tyr Asn Pro
            260                 265                 270

acc acg tac cag atg gat gtg aac ccc gag ggc aaa tac agc ttt ggt        864
Thr Thr Tyr Gln Met Asp Val Asn Pro Glu Gly Lys Tyr Ser Phe Gly
        275                 280                 285

gcc acc tgc gtg aag aag tgt ccc cgt aat tat gtg gtg aca gat cac        912
Ala Thr Cys Val Lys Lys Cys Pro Arg Asn Tyr Val Val Thr Asp His
    290                 295                 300
```

```
ggc tcg tgc gtc cga gcc tgt ggg gcc gac agc tat gag atg gag gaa        960
Gly Ser Cys Val Arg Ala Cys Gly Ala Asp Ser Tyr Glu Met Glu Glu
305                 310                 315                 320

gac ggc gtc cgc aag tgt aag aag tgc gaa ggg cct tgc cgc aaa gtg       1008
Asp Gly Val Arg Lys Cys Lys Lys Cys Glu Gly Pro Cys Arg Lys Val
                325                 330                 335

tgt aac gga ata ggt att ggt gaa ttt aaa gac tca ctc tcc ata aat       1056
Cys Asn Gly Ile Gly Ile Gly Glu Phe Lys Asp Ser Leu Ser Ile Asn
            340                 345                 350

gct acg aat att aaa cac ttc aaa aac tgc acc tcc atc agt ggc gat       1104
Ala Thr Asn Ile Lys His Phe Lys Asn Cys Thr Ser Ile Ser Gly Asp
        355                 360                 365

ctc cac atc ctg ccg gtg gca ttt agg ggt gac tcc ttc aca cat act       1152
Leu His Ile Leu Pro Val Ala Phe Arg Gly Asp Ser Phe Thr His Thr
    370                 375                 380

cct cct ctg gat cca cag gaa ctg gat att ctg aaa acc gta aag gaa       1200
Pro Pro Leu Asp Pro Gln Glu Leu Asp Ile Leu Lys Thr Val Lys Glu
385                 390                 395                 400

atc aca ggg ttt ttg ctg att cag gct tgg cct gaa aac agg acg gac       1248
Ile Thr Gly Phe Leu Leu Ile Gln Ala Trp Pro Glu Asn Arg Thr Asp
                405                 410                 415

ctc cat gcc ttt gag aac cta gaa atc ata cgc ggc agg acc aag caa       1296
Leu His Ala Phe Glu Asn Leu Glu Ile Ile Arg Gly Arg Thr Lys Gln
            420                 425                 430

cat ggt cag ttt tct ctt gca gtc gtc agc ctg aac ata aca tcc ttg       1344
His Gly Gln Phe Ser Leu Ala Val Val Ser Leu Asn Ile Thr Ser Leu
        435                 440                 445

gga tta cgc tcc ctc aag gag ata agt gat gga gat gtg ata att tca       1392
Gly Leu Arg Ser Leu Lys Glu Ile Ser Asp Gly Asp Val Ile Ile Ser
    450                 455                 460

gga aac aaa aat ttg tgc tat gca aat aca ata aac tgg aaa aaa ctg       1440
Gly Asn Lys Asn Leu Cys Tyr Ala Asn Thr Ile Asn Trp Lys Lys Leu
465                 470                 475                 480
```

```
ttt ggg acc tcc ggt cag aaa acc aaa att ata agc aac aga ggt gaa       1488
Phe Gly Thr Ser Gly Gln Lys Thr Lys Ile Ile Ser Asn Arg Gly Glu
                485                 490                 495

aac agc tgc aag gcc aca ggc cag gtc tgc cat gcc ttg tgc tcc ccc       1536
Asn Ser Cys Lys Ala Thr Gly Gln Val Cys His Ala Leu Cys Ser Pro
            500                 505                 510

gag ggc tgc tgg ggc ccg gag ccc agg gac tgc gtc tct tgc cgg aat       1584
Glu Gly Cys Trp Gly Pro Glu Pro Arg Asp Cys Val Ser Cys Arg Asn
        515                 520                 525

gtc agc cga ggc agg gaa tgc gtg gac aag tgc aac ctt ctg gag ggt       1632
Val Ser Arg Gly Arg Glu Cys Val Asp Lys Cys Asn Leu Leu Glu Gly
    530                 535                 540

gag cca agg gag ttt gtg gag aac tct gag tgc ata cag tgc cac cca       1680
Glu Pro Arg Glu Phe Val Glu Asn Ser Glu Cys Ile Gln Cys His Pro
545                 550                 555                 560

gag tgc ctg cct cag gcc atg aac atc acc tgc aca gga cgg gga cca       1728
Glu Cys Leu Pro Gln Ala Met Asn Ile Thr Cys Thr Gly Arg Gly Pro
                565                 570                 575

gac aac tgt atc cag tgt gcc cac tac att gac ggc ccc cac tgc gtc       1776
Asp Asn Cys Ile Gln Cys Ala His Tyr Ile Asp Gly Pro His Cys Val
            580                 585                 590

aag acc tgc ccg gca gga gtc atg gga gaa aac aac acc ctg gtc tgg       1824
Lys Thr Cys Pro Ala Gly Val Met Gly Glu Asn Asn Thr Leu Val Trp
        595                 600                 605

aag tac gca gac gcc ggc cat gtg tgc cac ctg tgc cat cca aac tgc       1872
Lys Tyr Ala Asp Ala Gly His Val Cys His Leu Cys His Pro Asn Cys
    610                 615                 620

acc tac gga tgc act ggg cca ggt ctt gaa ggc tgt cca acg aat ggg       1920
Thr Tyr Gly Cys Thr Gly Pro Gly Leu Glu Gly Cys Pro Thr Asn Gly
625                 630                 635                 640

cct aag atc ccg tcc atc gcc act ggg atg gtg ggg gcc ctc ctc ttg       1968
Pro Lys Ile Pro Ser Ile Ala Thr Gly Met Val Gly Ala Leu Leu Leu
                645                 650                 655
```

```
ctg ctg gtg gtg gcc ctg ggg atc ggc ctc ttc atg cga agg cgc cac       2016
Leu Leu Val Val Ala Leu Gly Ile Gly Leu Phe Met Arg Arg Arg His
            660                 665                 670

atc gtt cgg aag cgc acg ctg cgg agg ctg ctg cag gag agg gag ctt       2064
Ile Val Arg Lys Arg Thr Leu Arg Arg Leu Leu Gln Glu Arg Glu Leu
        675                 680                 685

gtg gag cct ctt aca ccc agt gga gaa gct ccc aac caa gct ctc ttg       2112
Val Glu Pro Leu Thr Pro Ser Gly Glu Ala Pro Asn Gln Ala Leu Leu
    690                 695                 700

agg atc ttg aag gaa act gaa ttc aaa aag atc aaa gtg ctg ggc tcc       2160
Arg Ile Leu Lys Glu Thr Glu Phe Lys Lys Ile Lys Val Leu Gly Ser
705                 710                 715                 720

ggt gcg ttc ggc acg gtg tat aag gga ctc tgg atc cca gaa ggt gag       2208
Gly Ala Phe Gly Thr Val Tyr Lys Gly Leu Trp Ile Pro Glu Gly Glu
                725                 730                 735

aaa gtt aaa att ccc gtc gct atc aag gaa tta aga gaa gca aca tct       2256
Lys Val Lys Ile Pro Val Ala Ile Lys Glu Leu Arg Glu Ala Thr Ser
            740                 745                 750

ccg aaa gcc aac aag gaa atc ctc gat gaa gcc tac gtg atg gcc agc       2304
Pro Lys Ala Asn Lys Glu Ile Leu Asp Glu Ala Tyr Val Met Ala Ser
        755                 760                 765

gtg gac aac ccc cac gtg tgc cgc ctg ctg ggc atc tgc ctc acc tcc       2352
Val Asp Asn Pro His Val Cys Arg Leu Leu Gly Ile Cys Leu Thr Ser
    770                 775                 780

acc gtg cag ctc atc acg cag ctc atg ccc ttc ggc tgc ctc ctg gac       2400
Thr Val Gln Leu Ile Thr Gln Leu Met Pro Phe Gly Cys Leu Leu Asp
785                 790                 795                 800

tat gtc cgg gaa cac aaa gac aat att ggc tcc cag tac ctg ctc aac       2448
Tyr Val Arg Glu His Lys Asp Asn Ile Gly Ser Gln Tyr Leu Leu Asn
                805                 810                 815

tgg tgt gtg cag atc gca aag ggc atg aac tac ttg gag gac cgt cgc       2496
Trp Cys Val Gln Ile Ala Lys Gly Met Asn Tyr Leu Glu Asp Arg Arg
            820                 825                 830
```

```
ttg gtg cac cgc gac ctg gca gcc agg aac gta ctg gtg aaa aca ccg       2544
Leu Val His Arg Asp Leu Ala Ala Arg Asn Val Leu Val Lys Thr Pro
        835                 840                 845

cag cat gtc aag atc aca gat ttt ggg ctg gcc aaa ctg ctg ggt gcg       2592
Gln His Val Lys Ile Thr Asp Phe Gly Leu Ala Lys Leu Leu Gly Ala
    850                 855                 860

gaa gag aaa gaa tac cat gca gaa gga ggc aaa gtg cct atc aag tgg       2640
Glu Glu Lys Glu Tyr His Ala Glu Gly Gly Lys Val Pro Ile Lys Trp
865                 870                 875                 880

atg gca ttg gaa tca att tta cac aga atc tat acc cac cag agt gat       2688
Met Ala Leu Glu Ser Ile Leu His Arg Ile Tyr Thr His Gln Ser Asp
                885                 890                 895

gtc tgg agc tac ggg gtg act gtt tgg gag ttg atg acc ttt gga tcc       2736
Val Trp Ser Tyr Gly Val Thr Val Trp Glu Leu Met Thr Phe Gly Ser
            900                 905                 910

aag cca tat gac gga atc cct gcc agc gag atc tcc tcc atc ctg gag       2784
Lys Pro Tyr Asp Gly Ile Pro Ala Ser Glu Ile Ser Ser Ile Leu Glu
        915                 920                 925

aaa gga gaa cgc ctc cct cag cca ccc ata tgt acc atc gat gtc tac       2832
Lys Gly Glu Arg Leu Pro Gln Pro Pro Ile Cys Thr Ile Asp Val Tyr
    930                 935                 940

atg atc atg gtc aag tgc tgg atg ata gac gca gat agt cgc cca aag       2880
Met Ile Met Val Lys Cys Trp Met Ile Asp Ala Asp Ser Arg Pro Lys
945                 950                 955                 960

ttc cgt gag ttg atc atc gaa ttc tcc aaa atg gcc cga gac ccc cag       2928
Phe Arg Glu Leu Ile Ile Glu Phe Ser Lys Met Ala Arg Asp Pro Gln
                965                 970                 975

cgc tac ctt gtc att cag ggg gat gaa aga atg cat ttg cca agt cct       2976
Arg Tyr Leu Val Ile Gln Gly Asp Glu Arg Met His Leu Pro Ser Pro
            980                 985                 990

aca gac tcc aac ttc tac cgt gcc  ctg atg gat gaa gaa  gac atg gac      3024
Thr Asp Ser Asn Phe Tyr Arg Ala  Leu Met Asp Glu Glu  Asp Met Asp
        995                 1000                1005
```

```
gac gtg  gtg gat gcc gac gag  tac ctc atc cca cag  cag ggc ttc        3069
Asp Val  Val Asp Ala Asp Glu  Tyr Leu Ile Pro Gln  Gln Gly Phe
    1010                1015                1020

ttc agc  agc ccc tcc acg tca  cgg act ccc ctc ctg  agc tct ctg        3114
Phe Ser  Ser Pro Ser Thr Ser  Arg Thr Pro Leu Leu  Ser Ser Leu
    1025                1030                1035

agt gca  acc agc aac aat tcc  acc gtg gct tgc att  gat aga aat        3159
Ser Ala  Thr Ser Asn Asn Ser  Thr Val Ala Cys Ile  Asp Arg Asn
    1040                1045                1050

ggg ctg  caa agc tgt ccc atc  aag gaa gac agc ttc  ttg cag cga        3204
Gly Leu  Gln Ser Cys Pro Ile  Lys Glu Asp Ser Phe  Leu Gln Arg
    1055                1060                1065

tac agc  tca gac ccc aca ggc  gcc ttg act gag gac  agc ata gac        3249
Tyr Ser  Ser Asp Pro Thr Gly  Ala Leu Thr Glu Asp  Ser Ile Asp
    1070                1075                1080

gac acc  ttc ctc cca gtg cct  gaa tac ata aac cag  tcc gtt ccc        3294
Asp Thr  Phe Leu Pro Val Pro  Glu Tyr Ile Asn Gln  Ser Val Pro
    1085                1090                1095

aaa agg  ccc gct ggc tct gtg  cag aat cct gtc tat  cac aat cag        3339
Lys Arg  Pro Ala Gly Ser Val  Gln Asn Pro Val Tyr  His Asn Gln
    1100                1105                1110

cct ctg  aac ccc gcg ccc agc  aga gac cca cac tac  cag gac ccc        3384
Pro Leu  Asn Pro Ala Pro Ser  Arg Asp Pro His Tyr  Gln Asp Pro
    1115                1120                1125

cac agc  act gca gtg ggc aac  ccc gag tat ctc aac  act gtc cag        3429
His Ser  Thr Ala Val Gly Asn  Pro Glu Tyr Leu Asn  Thr Val Gln
    1130                1135                1140

ccc acc  tgt gtc aac agc aca  ttc gac agc cct gcc  cac tgg gcc        3474
Pro Thr  Cys Val Asn Ser Thr  Phe Asp Ser Pro Ala  His Trp Ala
    1145                1150                1155

cag aaa  ggc agc cac caa att  agc ctg gac aac cct  gac tac cag        3519
Gln Lys  Gly Ser His Gln Ile  Ser Leu Asp Asn Pro  Asp Tyr Gln
    1160                1165                1170
```

```
cag gac  ttc ttt ccc aag gaa  gcc aag cca aat ggc  atc ttt aag        3564
Gln Asp  Phe Phe Pro Lys Glu  Ala Lys Pro Asn Gly  Ile Phe Lys
    1175                 1180                 1185

ggc tcc  aca gct gaa aat gca  gaa tac cta agg gtc  gcg cca caa        3609
Gly Ser  Thr Ala Glu Asn Ala  Glu Tyr Leu Arg Val  Ala Pro Gln
    1190                 1195                 1200

agc agt  gaa ttt att gga gca  tga                                   3633
Ser Ser  Glu Phe Ile Gly Ala
    1205                 1210
```

<210> 512
<211> 1210
<212> PRT
<213> Homo sapiens

<400> 512

```
Met Arg Pro Ser Gly Thr Ala Gly Ala Ala Leu Leu Ala Leu Leu Ala
1               5                  10                  15

Ala Leu Cys Pro Ala Ser Arg Ala Leu Glu Glu Lys Lys Val Cys Gln
            20                  25                  30

Gly Thr Ser Asn Lys Leu Thr Gln Leu Gly Thr Phe Glu Asp His Phe
        35                  40                  45

Leu Ser Leu Gln Arg Met Phe Asn Asn Cys Glu Val Val Leu Gly Asn
    50                  55                  60

Leu Glu Ile Thr Tyr Val Gln Arg Asn Tyr Asp Leu Ser Phe Leu Lys
65                  70                  75                  80

Thr Ile Gln Glu Val Ala Gly Tyr Val Leu Ile Ala Leu Asn Thr Val
                85                  90                  95

Glu Arg Ile Pro Leu Glu Asn Leu Gln Ile Ile Arg Gly Asn Met Tyr
```

100 105 110

Tyr Glu Asn Ser Tyr Ala Leu Ala Val Leu Ser Asn Tyr Asp Ala Asn
115 120 125

Lys Thr Gly Leu Lys Glu Leu Pro Met Arg Asn Leu Gln Glu Ile Leu
130 135 140

His Gly Ala Val Arg Phe Ser Asn Asn Pro Ala Leu Cys Asn Val Glu
145 150 155 160

Ser Ile Gln Trp Arg Asp Ile Val Ser Ser Asp Phe Leu Ser Asn Met
165 170 175

Ser Met Asp Phe Gln Asn His Leu Gly Ser Cys Gln Lys Cys Asp Pro
180 185 190

Ser Cys Pro Asn Gly Ser Cys Trp Gly Ala Gly Glu Glu Asn Cys Gln
195 200 205

Lys Leu Thr Lys Ile Ile Cys Ala Gln Gln Cys Ser Gly Arg Cys Arg
210 215 220

Gly Lys Ser Pro Ser Asp Cys Cys His Asn Gln Cys Ala Ala Gly Cys
225 230 235 240

Thr Gly Pro Arg Glu Ser Asp Cys Leu Val Cys Arg Lys Phe Arg Asp
245 250 255

Glu Ala Thr Cys Lys Asp Thr Cys Pro Pro Leu Met Leu Tyr Asn Pro
260 265 270

Thr Thr Tyr Gln Met Asp Val Asn Pro Glu Gly Lys Tyr Ser Phe Gly
275 280 285

Ala Thr Cys Val Lys Lys Cys Pro Arg Asn Tyr Val Val Thr Asp His

290 295 300

Gly Ser Cys Val Arg Ala Cys Gly Ala Asp Ser Tyr Glu Met Glu Glu
305 310 315 320

Asp Gly Val Arg Lys Cys Lys Lys Cys Glu Gly Pro Cys Arg Lys Val
325 330 335

Cys Asn Gly Ile Gly Ile Gly Glu Phe Lys Asp Ser Leu Ser Ile Asn
340 345 350

Ala Thr Asn Ile Lys His Phe Lys Asn Cys Thr Ser Ile Ser Gly Asp
355 360 365

Leu His Ile Leu Pro Val Ala Phe Arg Gly Asp Ser Phe Thr His Thr
370 375 380

Pro Pro Leu Asp Pro Gln Glu Leu Asp Ile Leu Lys Thr Val Lys Glu
385 390 395 400

Ile Thr Gly Phe Leu Leu Ile Gln Ala Trp Pro Glu Asn Arg Thr Asp
405 410 415

Leu His Ala Phe Glu Asn Leu Glu Ile Ile Arg Gly Arg Thr Lys Gln
420 425 430

His Gly Gln Phe Ser Leu Ala Val Val Ser Leu Asn Ile Thr Ser Leu
435 440 445

Gly Leu Arg Ser Leu Lys Glu Ile Ser Asp Gly Asp Val Ile Ile Ser
450 455 460

Gly Asn Lys Asn Leu Cys Tyr Ala Asn Thr Ile Asn Trp Lys Lys Leu
465 470 475 480

Phe Gly Thr Ser Gly Gln Lys Thr Lys Ile Ile Ser Asn Arg Gly Glu
485 490 495

```
Asn Ser Cys Lys Ala Thr Gly Gln Val Cys His Ala Leu Cys Ser Pro
            500                 505                 510

Glu Gly Cys Trp Gly Pro Glu Pro Arg Asp Cys Val Ser Cys Arg Asn
        515                 520                 525

Val Ser Arg Gly Arg Glu Cys Val Asp Lys Cys Asn Leu Leu Glu Gly
    530                 535                 540

Glu Pro Arg Glu Phe Val Glu Asn Ser Glu Cys Ile Gln Cys His Pro
545                 550                 555                 560

Glu Cys Leu Pro Gln Ala Met Asn Ile Thr Cys Thr Gly Arg Gly Pro
                565                 570                 575

Asp Asn Cys Ile Gln Cys Ala His Tyr Ile Asp Gly Pro His Cys Val
            580                 585                 590

Lys Thr Cys Pro Ala Gly Val Met Gly Glu Asn Asn Thr Leu Val Trp
        595                 600                 605

Lys Tyr Ala Asp Ala Gly His Val Cys His Leu Cys His Pro Asn Cys
    610                 615                 620

Thr Tyr Gly Cys Thr Gly Pro Gly Leu Glu Gly Cys Pro Thr Asn Gly
625                 630                 635                 640

Pro Lys Ile Pro Ser Ile Ala Thr Gly Met Val Gly Ala Leu Leu Leu
                645                 650                 655

Leu Leu Val Val Ala Leu Gly Ile Gly Leu Phe Met Arg Arg Arg His
            660                 665                 670

Ile Val Arg Lys Arg Thr Leu Arg Arg Leu Leu Gln Glu Arg Glu Leu
        675                 680                 685
```

Val Glu Pro Leu Thr Pro Ser Gly Glu Ala Pro Asn Gln Ala Leu Leu
690 695 700

Arg Ile Leu Lys Glu Thr Glu Phe Lys Lys Ile Lys Val Leu Gly Ser
705 710 715 720

Gly Ala Phe Gly Thr Val Tyr Lys Gly Leu Trp Ile Pro Glu Gly Glu
725 730 735

Lys Val Lys Ile Pro Val Ala Ile Lys Glu Leu Arg Glu Ala Thr Ser
740 745 750

Pro Lys Ala Asn Lys Glu Ile Leu Asp Glu Ala Tyr Val Met Ala Ser
755 760 765

Val Asp Asn Pro His Val Cys Arg Leu Leu Gly Ile Cys Leu Thr Ser
770 775 780

Thr Val Gln Leu Ile Thr Gln Leu Met Pro Phe Gly Cys Leu Leu Asp
785 790 795 800

Tyr Val Arg Glu His Lys Asp Asn Ile Gly Ser Gln Tyr Leu Leu Asn
805 810 815

Trp Cys Val Gln Ile Ala Lys Gly Met Asn Tyr Leu Glu Asp Arg Arg
820 825 830

Leu Val His Arg Asp Leu Ala Ala Arg Asn Val Leu Val Lys Thr Pro
835 840 845

Gln His Val Lys Ile Thr Asp Phe Gly Leu Ala Lys Leu Leu Gly Ala
850 855 860

Glu Glu Lys Glu Tyr His Ala Glu Gly Gly Lys Val Pro Ile Lys Trp
865 870 875 880

```
Met Ala Leu Glu Ser Ile Leu His Arg Ile Tyr Thr His Gln Ser Asp
                885                 890                 895

Val Trp Ser Tyr Gly Val Thr Val Trp Glu Leu Met Thr Phe Gly Ser
            900                 905                 910

Lys Pro Tyr Asp Gly Ile Pro Ala Ser Glu Ile Ser Ser Ile Leu Glu
        915                 920                 925

Lys Gly Glu Arg Leu Pro Gln Pro Pro Ile Cys Thr Ile Asp Val Tyr
    930                 935                 940

Met Ile Met Val Lys Cys Trp Met Ile Asp Ala Asp Ser Arg Pro Lys
945                 950                 955                 960

Phe Arg Glu Leu Ile Ile Glu Phe Ser Lys Met Ala Arg Asp Pro Gln
                965                 970                 975

Arg Tyr Leu Val Ile Gln Gly Asp Glu Arg Met His Leu Pro Ser Pro
            980                 985                 990

Thr Asp Ser Asn Phe Tyr Arg Ala  Leu Met Asp Glu Glu  Asp Met Asp
        995                 1000                1005

Asp Val  Val Asp Ala Asp Glu  Tyr Leu Ile Pro Gln  Gln Gly Phe
    1010                1015                1020

Phe Ser  Ser Pro Ser Thr Ser  Arg Thr Pro Leu Leu  Ser Ser Leu
    1025                1030                1035

Ser Ala  Thr Ser Asn Asn Ser  Thr Val Ala Cys Ile  Asp Arg Asn
    1040                1045                1050

Gly Leu  Gln Ser Cys Pro Ile  Lys Glu Asp Ser Phe  Leu Gln Arg
    1055                1060                1065
```

```
Tyr Ser  Ser Asp Pro Thr Gly  Ala Leu Thr Glu Asp  Ser Ile Asp
    1070                1075                1080

Asp Thr  Phe Leu Pro Val Pro  Glu Tyr Ile Asn Gln  Ser Val Pro
    1085                1090                1095

Lys Arg  Pro Ala Gly Ser Val  Gln Asn Pro Val Tyr  His Asn Gln
    1100                1105                1110

Pro Leu  Asn Pro Ala Pro Ser  Arg Asp Pro His Tyr  Gln Asp Pro
    1115                1120                1125

His Ser  Thr Ala Val Gly Asn  Pro Glu Tyr Leu Asn  Thr Val Gln
    1130                1135                1140

Pro Thr  Cys Val Asn Ser Thr  Phe Asp Ser Pro Ala  His Trp Ala
    1145                1150                1155

Gln Lys  Gly Ser His Gln Ile  Ser Leu Asp Asn Pro  Asp Tyr Gln
    1160                1165                1170

Gln Asp  Phe Phe Pro Lys Glu  Ala Lys Pro Asn Gly  Ile Phe Lys
    1175                1180                1185

Gly Ser  Thr Ala Glu Asn Ala  Glu Tyr Leu Arg Val  Ala Pro Gln
    1190                1195                1200

Ser Ser  Glu Phe Ile Gly Ala
    1205                1210
```

<210> 513
<211> 25
<212> DNA
<213> Homo sapiens

<400> 513
cagatttggc tcgacctgga catag 25

<210> 514
<211> 22
<212> DNA
<213> Homo sapiens

<400> 514
cagctgatct caaggaaaca gg 22

<210> 515
<211> 24
<212> DNA
<213> Homo sapiens

<400> 515
gtattatcag tcactaaagc tcac 24

<210> 516
<211> 22
<212> DNA
<213> Homo sapiens

<400> 516
cacacttcaa gtggaattct gc 22

<210> 517
<211> 19
<212> DNA
<213> Homo sapiens

<400> 517
tgcattaggg ttcaactgg 19

<210> 518
<211> 26
<212> DNA
<213> Homo sapiens

<400> 518

ccttctccga ggtggaattg agtgac 26

<210> 519
<211> 26
<212> DNA
<213> Homo sapiens

<400> 519
gctaattgcg ggactcttgt tcgcac 26

<210> 520
<211> 22
<212> DNA
<213> Homo sapiens

<400> 520
tacatgcttt tctagtggtc ag 22

<210> 521
<211> 25
<212> DNA
<213> Homo sapiens

<400> 521
ggtctcaagt gattctacaa accag 25

<210> 522
<211> 24
<212> DNA
<213> Homo sapiens

<400> 522
ccttcaccta ctggttcaca tctg 24

<210> 523
<211> 26
<212> DNA
<213> Homo sapiens

<400> 523

catggtttga cttagtttga atgtgg 26

<210> 524
<211> 27
<212> DNA
<213> Homo sapiens

<400> 524
ggatactaaa gatactttgt caccagg 27

<210> 525
<211> 26
<212> DNA
<213> Homo sapiens

<400> 525
gaacactagg ctgcaaagac agtaac 26

<210> 526
<211> 24
<212> DNA
<213> Homo sapiens

<400> 526
ccaagcaagg caaacacatc cacc 24

<210> 527
<211> 23
<212> DNA
<213> Homo sapiens

<400> 527
ggaggatgga gcctttccat cac 23

<210> 528
<211> 24
<212> DNA
<213> Homo sapiens

<400> 528

gaagaggaag atgtgttcct ttgg 24

<210> 529
<211> 24
<212> DNA
<213> Homo sapiens

<400> 529
gaatgaagga tgatgtggca gtgg 24

<210> 530
<211> 21
<212> DNA
<213> Homo sapiens

<400> 530
caaaacatca gccattaacg g 21

<210> 531
<211> 26
<212> DNA
<213> Homo sapiens

<400> 531
ccacttactg ttcatataat acagag 26

<210> 532
<211> 25
<212> DNA
<213> Homo sapiens

<400> 532
catgtgagat agcatttggg aatgc 25

<210> 533
<211> 26
<212> DNA
<213> Homo sapiens

<400> 533

catgacctac catcattgga aagcag 26

<210> 534
<211> 21
<212> DNA
<213> Homo sapiens

<400> 534
gtaatttcac agttaggaat c 21

<210> 535
<211> 25
<212> DNA
<213> Homo sapiens

<400> 535
gtcacccaag gtcatggagc acagg 25

<210> 536
<211> 23
<212> DNA
<213> Homo sapiens

<400> 536
cagaatgcct gtaaagctat aac 23

<210> 537
<211> 24
<212> DNA
<213> Homo sapiens

<400> 537
gtcctggagt cccaactcct tgac 24

<210> 538
<211> 25
<212> DNA
<213> Homo sapiens

<400> 538

ggaagtggct ctgatggccg tcctg 25

<210> 539
<211> 27
<212> DNA
<213> Homo sapiens

<400> 539
ccactcacac acactaaata ttttaag 27

<210> 540
<211> 27
<212> DNA
<213> Homo sapiens

<400> 540
gaccaaaaca ccttaagtaa ctgactc 27

<210> 541
<211> 24
<212> DNA
<213> Homo sapiens

<400> 541
ccaatccaac atccagacac atag 24

<210> 542
<211> 24
<212> DNA
<213> Homo sapiens

<400> 542
ccagagccat agaaacttga tcag 24

<210> 543
<211> 29
<212> DNA
<213> Homo sapiens

<400> 543

gtatggacta tggcacttca attgcatgg 29

<210> 544
<211> 28
<212> DNA
<213> Homo sapiens

<400> 544
ccagagaaca tggcaaccag cacaggac 28

<210> 545
<211> 25
<212> DNA
<213> Homo sapiens

<400> 545
caaatgagct ggcaagtgcc gtgtc 25

<210> 546
<211> 25
<212> DNA
<213> Homo sapiens

<400> 546
gagtttccca aacactcagt gaaac 25

<210> 547
<211> 25
<212> DNA
<213> Homo sapiens

<400> 547
gcaatatcag ccttaggtgc ggctc 25

<210> 548
<211> 26
<212> DNA
<213> Homo sapiens

<400> 548

catagaaagt gaacatttag gatgtg 26

<210> 549
<211> 24
<212> DNA
<213> Homo sapiens

<400> 549
ccatgagtac gtattttgaa actc 24

<210> 550
<211> 24
<212> DNA
<213> Homo sapiens

<400> 550
catatcccca tggcaaactc ttgc 24

<210> 551
<211> 24
<212> DNA
<213> Homo sapiens

<400> 551
ctaacgttcg ccagccataa gtcc 24

<210> 552
<211> 26
<212> DNA
<213> Homo sapiens

<400> 552
gctgcgagct cacccagaat gtctgg 26

<210> 553
<211> 25
<212> DNA
<213> Homo sapiens

<400> 553

gacgggtcct ggggtgatct ggctc 25

<210> 554
<211> 3866
<212> DNA
<213> Homo sapiens

<400> 554
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaaccatctc cgaaagccaa caaggaaatc ctcgatgaag 2520

cctacgtgat ggccagcgtg gacaaccccc acgtgtgccg cctgctgggc atctgcctca 2580

cctccaccgt gcagctcatc acgcagctca tgcccttcgg ctgcctcctg gactatgtcc 2640

gggaacacaa agacaatatt ggctcccagt acctgctcaa ctggtgtgtg cagatcgcaa 2700

agggcatgaa ctacttggag gaccgtcgct tggtgcaccg cgacctggca gccaggaacg 2760

tactggtgaa aacaccgcag catgtcaaga tcacagattt tgggctggcc aaactgctgg 2820

gtgcggaaga gaaagaatac catgcagaag gaggcaaagt gcctatcaag tggatggcat 2880

tggaatcaat tttacacaga atctataccc accagagtga tgtctggagc tacggggtga 2940

ctgtttggga gttgatgacc tttggatcca agccatatga cggaatccct gccagcgaga 3000

tctcctccat cctggagaaa ggagaacgcc tccctcagcc acccatatgt accatcgatg 3060

tctacatgat catggtcaag tgctggatga tagacgcaga tagtcgccca aagttccgtg 3120

agttgatcat cgaattctcc aaaatggccc gagaccccca gcgctacctt gtcattcagg 3180

gggatgaaag aatgcatttg ccaagtccta cagactccaa cttctaccgt gccctgatgg 3240

atgaagaaga catggacgac gtggtggatg ccgacgagta cctcatccca cagcagggct 3300

tcttcagcag cccctccacg tcacggactc ccctcctgag ctctctgagt gcaaccagca 3360

acaattccac cgtggcttgc attgatagaa atgggctgca aagctgtccc atcaaggaag 3420

acagcttctt gcagcgatac agctcagacc ccacaggcgc cttgactgag gacagcatag 3480

acgacacctt cctcccagtg cctgaataca taaaccagtc cgttcccaaa aggcccgctg 3540

gctctgtgca gaatcctgtc tatcacaatc agcctctgaa ccccgcgccc agcagagacc 3600

cacactacca ggacccccac agcactgcag tgggcaaccc cgagtatctc aacactgtcc 3660

agcccacctg tgtcaacagc acattcgaca gccctgccca ctgggcccag aaaggcagcc 3720

accaaattag cctggacaac cctgactacc agcaggactt ctttcccaag gaagccaagc 3780

caaatggcat ctttaagggc tccacagctg aaaatgcaga atacctaagg gtcgcgccac 3840

aaagcagtga atttattgga gcatga 3866

<210> 555
<211> 24
<212> DNA
<213> Homo sapiens

<400> 555
caggactaca gaaatgtagg tttc 24

<210> 556
<211> 26
<212> DNA
<213> Homo sapiens

<400> 556
gtgcctgcct taagtaatgt gatgac 26

<210> 557
<211> 25
<212> DNA

<213> Homo sapiens

<400> 557
gactggaagt gtcgcatcac caatg 25

<210> 558
<211> 25
<212> DNA
<213> Homo sapiens

<400> 558
ggtttaataa tgcgatctgg gacac 25

<210> 559
<211> 25
<212> DNA
<213> Homo sapiens

<400> 559
gcagctataa tttagagaac caagg 25

<210> 560
<211> 21
<212> DNA
<213> Homo sapiens

<400> 560
aaaattgact tcatttccat g 21

<210> 561
<211> 22
<212> DNA
<213> Homo sapiens

<400> 561
cctagttgct ctaaaactaa cg 22

<210> 562
<211> 24
<212> DNA

<213> Homo sapiens

<400> 562
ctgtgaggcg tgacagccgt gcag 24

<210> 563
<211> 25
<212> DNA
<213> Homo sapiens

<400> 563
caacctacta atcagaacca gcatc 25

<210> 564
<211> 24
<212> DNA
<213> Homo sapiens

<400> 564
ccttcactgt gtctgcaaat ctgc 24

<210> 565
<211> 24
<212> DNA
<213> Homo sapiens

<400> 565
cctgtcataa gtctccttgt tgag 24

<210> 566
<211> 25
<212> DNA
<213> Homo sapiens

<400> 566
cagtctgtgg gtctaagagc taatg 25

<210> 567
<211> 24
<212> DNA

<213> Homo sapiens

<400> 567
caggaatggg tgagtctctg tgtg 24

<210> 568
<211> 23
<212> DNA
<213> Homo sapiens

<400> 568
gtggaattct gcccaggcct ttc 23

<210> 569
<211> 25
<212> DNA
<213> Homo sapiens

<400> 569
gattctacaa accagccagc caaac 25

<210> 570
<211> 24
<212> DNA
<213> Homo sapiens

<400> 570
cctactggtt cacatctgac cctg 24

<210> 571
<211> 24
<212> DNA
<213> Homo sapiens

<400> 571
gtttgaatgt ggtttcgttg gaag 24

<210> 572
<211> 27
<212> DNA

<213> Homo sapiens

<400> 572
ctttgtcacc aggcagaggg caatatc 27

<210> 573
<211> 23
<212> DNA
<213> Homo sapiens

<400> 573
gacagtaact tgggctttct gac 23

<210> 574
<211> 23
<212> DNA
<213> Homo sapiens

<400> 574
catccaccca aagactctcc aag 23

<210> 575
<211> 25
<212> DNA
<213> Homo sapiens

<400> 575
ctgttcatat aatacagagt ccctg 25

<210> 576
<211> 22
<212> DNA
<213> Homo sapiens

<400> 576
gagagatgca ggagctctgt gc 22

<210> 577
<211> 25
<212> DNA

<213> Homo sapiens

<400> 577
gcagtttgta gtcaatcaaa ggtgg 25

<210> 578
<211> 22
<212> DNA
<213> Homo sapiens

<400> 578
gtaatttaaa tgggaatagc cc 22

<210> 579
<211> 24
<212> DNA
<213> Homo sapiens

<400> 579
caactccttg accattacct caag 24

<210> 580
<211> 23
<212> DNA
<213> Homo sapiens

<400> 580
gatggccgtc ctgcccacac agg 23

<210> 581
<211> 21
<212> DNA
<213> Homo sapiens

<400> 581
gagtagttta gcatatattg c 21

<210> 582
<211> 23
<212> DNA

<213> Homo sapiens

<400> 582
gacagtcaga aatgcaggaa agc 23

<210> 583
<211> 26
<212> DNA
<213> Homo sapiens

<400> 583
caagtgccgt gtcctggcac ccaagc 26

<210> 584
<211> 24
<212> DNA
<213> Homo sapiens

<400> 584
ccaaacactc agtgaaacaa agag 24

<210> 585
<211> 21
<212> DNA
<213> Homo sapiens

<400> 585
ccttaggtgc ggctccacag c 21

<210> 586
<211> 22
<212> DNA
<213> Homo sapiens

<400> 586
catttaggat gtggagatga gc 22

<210> 587
<211> 23
<212> DNA

<213> Homo sapiens

<400> 587
gaaactcaag atcgcattca tgc 23

<210> 588
<211> 23
<212> DNA
<213> Homo sapiens

<400> 588
gcaaactctt gctatcccag gag 23

<210> 589
<211> 22
<212> DNA
<213> Homo sapiens

<400> 589
cagccataag tcctcgacgt gg 22

<210> 590
<211> 24
<212> DNA
<213> Homo sapiens

<400> 590
catcctcccc tgcatgtgtt aaac 24

<210> 591
<211> 24
<212> DNA
<213> Homo sapiens

<400> 591
gtaggtttct aaacatcaag aaac 24

<210> 592
<211> 24
<212> DNA

<213> Homo sapiens

<400> 592
gtgatgacat ttctccaggg atgc 24

<210> 593
<211> 23
<212> DNA
<213> Homo sapiens

<400> 593
catcaccaat gccttcttta agc 23

<210> 594
<211> 24
<212> DNA
<213> Homo sapiens

<400> 594
gctggagggt ttaataatgc gatc 24

<210> 595
<211> 24
<212> DNA
<213> Homo sapiens

<400> 595
gcaaacacac aggcacctgc tggc 24

<210> 596
<211> 21
<212> DNA
<213> Homo sapiens

<400> 596
catgtgagtt tcactagatg g 21

<210> 597
<211> 24
<212> DNA

<213> Homo sapiens

<400> 597
gaccggacga caggccacct cgtc 24

<210> 598
<211> 25
<212> DNA
<213> Homo sapiens

<400> 598
gaagaacgaa acgtcccgtt cctcc 25

<210> 599
<211> 23
<212> DNA
<213> Homo sapiens

<400> 599
gttgagcact cgtgtgcatt agg 23

<210> 600
<211> 22
<212> DNA
<213> Homo sapiens

<400> 600
ctcagtgcac gtgtactggg ta 22

<210> 601
<211> 35
<212> DNA
<213> Homo sapiens

<400> 601
gttcactggg ctaattgcgg gactcttgtt cgcac 35

<210> 602
<211> 28
<212> DNA

<213> Homo sapiens

<400> 602
ggtaaataca tgcttttcta gtggtcag 28

<210> 603
<211> 23
<212> DNA
<213> Homo sapiens

<400> 603
ggaggatgga gcctttccat cac 23

<210> 604
<211> 24
<212> DNA
<213> Homo sapiens

<400> 604
gaagaggaag atgtgttcct ttgg 24

<210> 605
<211> 24
<212> DNA
<213> Homo sapiens

<400> 605
gaatgaagga tgatgtggca gtgg 24

<210> 606
<211> 25
<212> DNA
<213> Homo sapiens

<400> 606
gtatgtgtga aggagtcact gaaac 25

<210> 607
<211> 22
<212> DNA

<213> Homo sapiens

<400> 607
ggtgagtcac aggttcagtt gc 22

<210> 608
<211> 21
<212> DNA
<213> Homo sapiens

<400> 608
caaaacatca gccattaacg g 21

<210> 609
<211> 22
<212> DNA
<213> Homo sapiens

<400> 609
gtagccagca tgtctgtgtc ac 22

<210> 610
<211> 23
<212> DNA
<213> Homo sapiens

<400> 610
cagaatgcct gtaaagctat aac 23

<210> 611
<211> 23
<212> DNA
<213> Homo sapiens

<400> 611
catttggctt tccccactca cac 23

<210> 612
<211> 27
<212> DNA

<213> Homo sapiens

<400> 612
gaccaaaaca ccttaagtaa ctgactc 27

<210> 613
<211> 25
<212> DNA
<213> Homo sapiens

<400> 613
gaagctacat agtgtctcac tttcc 25

<210> 614
<211> 25
<212> DNA
<213> Homo sapiens

<400> 614
cacaactgct aatggcccgt tctcg 25

<210> 615
<211> 23
<212> DNA
<213> Homo sapiens

<400> 615
gctcctgctc cctgtcataa gtc 23

<210> 616
<211> 25
<212> DNA
<213> Homo sapiens

<400> 616
gaagtcctgc tggtagtcag ggttg 25

<210> 617
<211> 24
<212> DNA

<213> Homo sapiens

<400> 617
ctgcagtggg caaccccgag tatc 24

<210> 618
<211> 20
<212> DNA
<213> Homo sapiens

<400> 618
tgtgggtcta agagctaatg 20

<210> 619
<211> 22
<212> DNA
<213> Homo sapiens

<400> 619
gacaggccac ctcgtcggcg tc 22

<210> 620
<211> 22
<212> DNA
<213> Homo sapiens

<400> 620
cagctgatct caaggaaaca gg 22

<210> 621
<211> 25
<212> DNA
<213> Homo sapiens

<400> 621
ctcgtgtgca ttagggttca actgg 25

<210> 622
<211> 26
<212> DNA

<213> Homo sapiens

<400> 622
ccttctccga ggtggaattg agtgac 26

<210> 623
<211> 26
<212> DNA
<213> Homo sapiens

<400> 623
gctaattgcg ggactcttgt tcgcac 26

<210> 624
<211> 22
<212> DNA
<213> Homo sapiens

<400> 624
tacatgcttt tctagtggtc ag 22

<210> 625
<211> 23
<212> DNA
<213> Homo sapiens

<400> 625
cctttccatc acccctcaag agg 23

<210> 626
<211> 25
<212> DNA
<213> Homo sapiens

<400> 626
gatgtgttcc tttggaggtg gcatg 25

<210> 627
<211> 24
<212> DNA

<213> Homo sapiens

<400> 627
gatgtggcag tggcggttcc ggtg 24

<210> 628
<211> 24
<212> DNA
<213> Homo sapiens

<400> 628
ggagtcactg aaacaaacaa cagg 24

<210> 629
<211> 21
<212> DNA
<213> Homo sapiens

<400> 629
ggttcagttg cttgtataaa g 21

<210> 630
<211> 24
<212> DNA
<213> Homo sapiens

<400> 630
ccattaacgg taaaatttca gaag 24

<210> 631
<211> 20
<212> DNA
<213> Homo sapiens

<400> 631
ccaaggtcat ggagcacagg 20

<210> 632
<211> 25
<212> DNA

<213> Homo sapiens

<400> 632
ctgtaaagct ataacaacaa cctgg 25

<210> 633
<211> 27
<212> DNA
<213> Homo sapiens

<400> 633
ccactcacac acactaaata ttttaag 27

<210> 634
<211> 24
<212> DNA
<213> Homo sapiens

<400> 634
gtaactgact caaatacaaa ccac 24

<210> 635
<211> 25
<212> DNA
<213> Homo sapiens

<400> 635
gaagctacat agtgtctcac tttcc 25

<210> 636
<211> 20
<212> DNA
<213> Homo sapiens

<400> 636
ctgctaatgg cccgttctcg 20

<210> 637
<211> 24
<212> DNA

<213> Homo sapiens

<400> 637
cctgtcataa gtctccttgt tgag 24

<210> 638
<211> 20
<212> DNA
<213> Homo sapiens

<400> 638
ggtagtcagg gttgtccagg 20

<210> 639
<211> 23
<212> DNA
<213> Homo sapiens

<400> 639
cgagtatctc aacactgtcc agc 23

<210> 640
<211> 25
<212> DNA
<213> Homo sapiens

<400> 640
ctaagagcta atgcgggcat ggctg 25

<210> 641
<211> 372
<212> DNA
<213> Homo sapiens

<400> 641
gcaatatcag ccttaggtgc ggctccacag ccccagtgtc cctcaccttc ggggtgcatc 60

gctggtaaca tccacccaga tcactgggca gcatgtggca ccatctcaca attgccagtt 120

aacgtcttcc ttctctctct gtcataggga ctctggatcc cagaaggtga gaaagttaaa 180

attcccgtcg ctatcaagga attaagagaa gcaacatctc cgaaagccaa caaggaaatc 240

ctcgatgtga gtttctgctt tgctgtgtgg gggtccatgg ctctgaacct caggcccacc 300

ttttctcatg tctggcagct gctctgctct agaccctgct catctccaca tcctaaatgt 360

tcactttcta tg 372

<210> 642
<211> 415
<212> DNA
<213> Homo sapiens

<400> 642
ctaacgttcg ccagccataa gtcctcgacg tggagaggct cagagcctgg catgaacatg 60

accctgaatt cggatgcaga gcttcttccc atgatgatct gtccctcaca gcagggtctt 120

ctctgtttca gggcatgaac tacttggagg accgtcgctt ggtgcaccgc gacctggcag 180

ccaggaacgt actggtgaaa acaccgcagc atgtcaagat cacagatttt gggctggcca 240

aactgctggg tgcggaagag aaagaatacc atgcagaagg aggcaaagta aggaggtggc 300

tttaggtcag ccagcatttt cctgacacca gggaccaggc tgccttccca ctagctgtat 360

tgtttaacac atgcagggga ggatgctctc cagacattct gggtgagctc gcagc 415

<210> 643
<211> 35
<212> DNA
<213> Homo sapiens

<400> 643
tatcaaggaa ttaagagaag caacatctcc gaaag 35

<210> 644
<211> 24
<212> DNA

<213> Homo sapiens

<400> 644
cgaaagccaa caaggaaatc ctcg 24

<210> 645
<211> 18
<212> DNA
<213> Homo sapiens

<400> 645
tgtaaaacga cggccagt 18

<210> 646
<211> 22
<212> DNA
<213> Homo sapiens

<400> 646
tggtctcaca ggaccactga tt 22

<210> 647
<211> 22
<212> DNA
<213> Homo sapiens

<400> 647
gaggccagtg ctgtctctaa gg 22

<210> 648
<211> 21
<212> DNA
<213> Homo sapiens

<400> 648
atgggacagg cactgatttg t 21

<210> 649
<211> 22
<212> DNA

<213> Homo sapiens

<400> 649
cagctctggc tcacactacc ag 22

<210> 650
<211> 20
<212> DNA
<213> Homo sapiens

<400> 650
gcagctggac tcgatttcct 20

<210> 651
<211> 22
<212> DNA
<213> Homo sapiens

<400> 651
tgcccaatga gtcaagaagt gt 22

<210> 652
<211> 22
<212> DNA
<213> Homo sapiens

<400> 652
cactcacgga tgctgcttag tt 22

<210> 653
<211> 22
<212> DNA
<213> Homo sapiens

<400> 653
tcagagcctg tgtttctacc aa 22

<210> 654
<211> 24
<212> DNA

<213> Homo sapiens

<400> 654
aaataatcag tgtgattcgt ggag 24

<210> 655
<211> 21
<212> DNA
<213> Homo sapiens

<400> 655
acttcacagc cctgcgtaaa c 21

<210> 656
<211> 22
<212> DNA
<213> Homo sapiens

<400> 656
gcagcgggtt acatcttctt tc 22

<210> 657
<211> 22
<212> DNA
<213> Homo sapiens

<400> 657
cctgaactcc gtcagactga aa 22

<210> 658
<211> 23
<212> DNA
<213> Homo sapiens

<400> 658
ccttacagca atcctgtgaa aca 23

<210> 659
<211> 22
<212> DNA

<213> Homo sapiens

<400> 659
atgtacagtg ctggcatggt ct 22

<210> 660
<211> 20
<212> DNA
<213> Homo sapiens

<400> 660
tccaaatgag ctggcaagtg 20

<210> 661
<211> 20
<212> DNA
<213> Homo sapiens

<400> 661
gtgcatcgct ggtaacatcc 20

<210> 662
<211> 20
<212> DNA
<213> Homo sapiens

<400> 662
atcgcattca tgcgtcttca 20

<210> 663
<211> 19
<212> DNA
<213> Homo sapiens

<400> 663
gctcagagcc tggcatgaa 19

<210> 664
<211> 20
<212> DNA

<213> Homo sapiens

<400> 664
tggctcgtct gtgtgtgtca 20

<210> 665
<211> 20
<212> DNA
<213> Homo sapiens

<400> 665
tgaagcaaat tgcccaagac 20

<210> 666
<211> 20
<212> DNA
<213> Homo sapiens

<400> 666
aagtgtcgca tcaccaatgc 20

<210> 667
<211> 23
<212> DNA
<213> Homo sapiens

<400> 667
tcccaaacac tcagtgaaac aaa 23

<210> 668
<211> 20
<212> DNA
<213> Homo sapiens

<400> 668
tgtggagatg agcagggtct 20

<210> 669
<211> 20
<212> DNA

<213> Homo sapiens

<400> 669
atccccatgg caaactcttg 20

<210> 670
<211> 19
<212> DNA
<213> Homo sapiens

<400> 670
catcctcccc tgcatgtgt 19

<210> 671
<211> 25
<212> DNA
<213> Homo sapiens

<400> 671
cgaaagaaaa tacttgcatg tcaga 25

<210> 672
<211> 20
<212> DNA
<213> Homo sapiens

<400> 672
tgacatttct ccagggatgc 20

<210> 673
<211> 19
<212> DNA
<213> Homo sapiens

<400> 673
atgcgatctg ggacacagg 19

<210> 674
<211> 16
<212> DNA

<213> Homo sapiens

<400> 674
aacagctatg accatg 16

<210> 675
<211> 26
<212> DNA
<213> Homo sapiens

<400> 675
caagtgccgt gtcctggcac ccaagc 26

<210> 676
<211> 24
<212> DNA
<213> Homo sapiens

<400> 676
ccaaacactc agtgaaacaa agag 24

<210> 677
<211> 25
<212> DNA
<213> Homo sapiens

<400> 677
gcacccaagc ccatgccgtg gctgc 25

<210> 678
<211> 26
<212> DNA
<213> Homo sapiens

<400> 678
gaaacaaaga gtaaagtaga tgatgg 26

<210> 679
<211> 24
<212> DNA

<213> Homo sapiens

<400> 679
caccttcaca atataccctc catg 24

<210> 680
<211> 22
<212> DNA
<213> Homo sapiens

<400> 680
gacagccgtg cagggaaaaa cc 22

<210> 681
<211> 24
<212> DNA
<213> Homo sapiens

<400> 681
gaaccagcat ctcaaggaga tctc 24

<210> 682
<211> 24
<212> DNA
<213> Homo sapiens

<400> 682
gagcacctgg cttggacact ggag 24

<210> 683
<211> 26
<212> DNA
<213> Homo sapiens

<400> 683
gagcagccct gaactccgtc agactg 26

<210> 684
<211> 26
<212> DNA

<213> Homo sapiens

<400> 684
ctcagtacaa tagatagaca gcaatg 26

<210> 685
<211> 25
<212> DNA
<213> Homo sapiens

<400> 685
gacgggtcct ggggtgatct ggctc 25

<210> 686
<211> 26
<212> DNA
<213> Homo sapiens

<400> 686
ctcagtacaa tagatagaca gcaatg 26

<210> 687
<211> 26
<212> DNA
<213> Homo sapiens

<400> 687
cgtggagagg ctcagagcct ggcatg 26

<210> 688
<211> 21
<212> DNA
<213> Homo sapiens

<400> 688
cacagatttt gggcgggcca a 21

<210> 689
<211> 21
<212> DNA

<213> Homo sapiens

<400> 689
gctatcaaaa catctccgaa a 21

<210> 690
<211> 35
<212> DNA
<213> Homo sapiens

<400> 690
tatcaaggaa ttaagagaag caacatctcc gaaag 35

<210> 691
<211> 29
<212> DNA
<213> Homo sapiens

<400> 691
agttaaaatt cccgtcgcta tcaaggaat 29

<210> 692
<211> 35
<212> DNA
<213> Homo sapiens

<400> 692
tatcaaggaa ttaagagaag caacatctcc gaaag 35

<210> 693
<211> 29
<212> DNA
<213> Homo sapiens

<400> 693
aacatctccg aaagccaaca aggaaatcc 29

<210> 694
<211> 35
<212> DNA

<213> Homo sapiens

<400> 694
tatcaaggaa ttaagagaag caacatctcc gaaag 35

<210> 695
<211> 21
<212> DNA
<213> Homo sapiens

<400> 695
taaaattccc gtcgctatca a 21

<210> 696
<211> 35
<212> DNA
<213> Homo sapiens

<400> 696
tatcaaggaa ttaagagaag caacatctcc gaaag 35

<210> 697
<211> 25
<212> DNA
<213> Homo sapiens

<400> 697
ccatctccga aagccaacaa ggaaa 25

<210> 698
<211> 35
<212> DNA
<213> Homo sapiens

<400> 698
tatcaaggaa ttaagagaag caacatctcc gaaag 35

<210> 699
<211> 23
<212> DNA

<213> Homo sapiens

<400> 699
aattcccgtc gctatcaagg aac 23

<210> 700
<211> 15
<212> DNA
<213> Homo sapiens

<400> 700
gggctggcca aactg 15

<210> 701
<211> 15
<212> DNA
<213> Homo sapiens

<400> 701
gggcgggcca aactg 15

<210> 702
<211> 15
<212> DNA
<213> Homo sapiens

<400> 702
gggctggcca aactg 15

<210> 703
<211> 15
<212> DNA
<213> Homo sapiens

<400> 703
gggctggcca aacag 15

<210> 704
<211> 27
<212> DNA

<213> Homo sapiens

<400> 704
gatcaaagtg ctgggctccg gtgcgtt 27

<210> 705
<211> 27
<212> DNA
<213> Homo sapiens

<220>
<221> modified_base
<222> (14)..(14)
<223> a, c, t, g, unknown or other

<400> 705
gatcaaagtg ctgngctccg gtgcgtt 27

<210> 706
<211> 31
<212> DNA
<213> Homo sapiens

<400> 706
tcacagattt tgggctggcc aaactgctgg g 31

<210> 707
<211> 31
<212> DNA
<213> Homo sapiens

<220>
<221> modified_base
<222> (16)..(16)
<223> a, c, t, g, unknown or other

<400> 707
tcacagattt tgggcnggcc aaactgctgg g 31

<210> 708

<211> 21
<212> DNA
<213> Homo sapiens

<400> 708
tatcaaggaa ttaagagaag c 21

<210> 709
<211> 21
<212> DNA
<213> Homo sapiens

<220>
<221> modified_base
<222> (17)..(17)
<223> a, c, t, g, unknown or other

<400> 709
tatcaaggaa ttaaganaag c 21

<210> 710
<211> 22
<212> DNA
<213> Homo sapiens

<400> 710
tatcaaggaa ttaagagaag ca 22

<210> 711
<211> 27
<212> DNA
<213> Homo sapiens

<400> 711
cgtcgctatc aaggaattaa gagaagc 27

<210> 712
<211> 20
<212> DNA
<213> Homo sapiens

<400> 712
cgctatcaag gaacaacatc 20

<210> 713
<211> 19
<212> DNA
<213> Homo sapiens

<400> 713
sgatcaygga ascaacatc 19

<210> 714
<211> 28
<212> DNA
<213> Homo sapiens

<400> 714
aggaattaag agaagcaaca tctccgaa 28

<210> 715
<211> 28
<212> DNA
<213> Homo sapiens

<400> 715
aggaattaaa agaaaccaca tctctcat 28

<210> 716
<211> 28
<212> DNA
<213> Homo sapiens

<400> 716
aggaaccaag agaagccacg tattcgaa 28

<210> 717
<211> 37
<212> DNA
<213> Homo sapiens

<400> 717

tatcaaggaa ttaagagaag caacatctcc gaaagcc 37

<210> 718
<211> 19
<212> DNA
<213> Homo sapiens

<400> 718
tatcaaggtt ccgaaagcc 19

<210> 719
<211> 19
<212> DNA
<213> Homo sapiens

<400> 719
tatcaaggtt ccgaaagcc 19

<210> 720
<211> 3867
<212> DNA
<213> Homo sapiens

<400> 720
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaatacatct ccgaaagcca acaaggaaat cctcgatgaa 2520

gcctacgtga tggccagcgt ggacaacccc cacgtgtgcc gcctgctggg catctgcctc 2580

acctccaccg tgcagctcat cacgcagctc atgcccttcg gctgcctcct ggactatgtc 2640

cgggaacaca aagacaatat tggctcccag tacctgctca actggtgtgt gcagatcgca 2700

aagggcatga actacttgga ggaccgtcgc ttggtgcacc gcgacctggc agccaggaac 2760

gtactggtga aaacaccgca gcatgtcaag atcacagatt ttgggctggc caaactgctg 2820

ggtgcggaag agaaagaata ccatgcagaa ggaggcaaag tgcctatcaa gtggatggca 2880

ttggaatcaa ttttacacag aatctatacc caccagagtg atgtctggag ctacggggtg 2940

actgtttggg agttgatgac ctttggatcc aagccatatg acggaatccc tgccagcgag 3000

atctcctcca tcctggagaa aggagaacgc ctccctcagc cacccatatg taccatcgat 3060

gtctacatga tcatggtcaa gtgctggatg atagacgcag atagtcgccc aaagttccgt 3120

gagttgatca tcgaattctc caaaatggcc cgagaccccc agcgctacct tgtcattcag 3180

ggggatgaaa gaatgcattt gccaagtcct acagactcca acttctaccg tgccctgatg 3240

gatgaagaag acatggacga cgtggtggat gccgacgagt acctcatccc acagcagggc 3300

ttcttcagca gcccctccac gtcacggact cccctcctga gctctctgag tgcaaccagc 3360

aacaattcca ccgtggcttg cattgataga aatgggctgc aaagctgtcc catcaaggaa 3420

gacagcttct tgcagcgata cagctcagac cccacaggcg ccttgactga ggacagcata 3480

gacgacacct tcctcccagt gcctgaatac ataaaccagt ccgttcccaa aaggcccgct 3540

ggctctgtgc agaatcctgt ctatcacaat cagcctctga accccgcgcc cagcagagac 3600

ccacactacc aggacccccа cagcactgca gtgggcaacc ccgagtatct caacactgtc 3660

cagcccacct gtgtcaacag cacattcgac agccctgccc actgggccca gaaaggcagc 3720

caccaaatta gcctggacaa ccctgactac cagcaggact tctttcccaa ggaagccaag 3780

ccaaatggca tctttaaggg ctccacagct gaaaatgcag aatacctaag ggtcgcgcca 3840

caaagcagtg aatttattgg agcatga 3867

<210> 721
<211> 3863
<212> DNA

<213> Homo sapiens

<400> 721
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tgaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 722
<211> 3886
<212> DNA
<213> Homo sapiens

<400> 722
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ggcaacccca acgtgtgccg 2580

cctgctgggc atctgcctca cctccaccgt gcagctcatc acgcagctca tgcccttcgg 2640

ctgcctcctg gactatgtcc gggaacacaa agacaatatt ggctcccagt acctgctcaa 2700

ctggtgtgtg cagatcgcaa agggcatgaa ctacttggag gaccgtcgct tggtgcaccg 2760

cgacctggca gccaggaacg tactggtgaa aacaccgcag catgtcaaga tcacagattt 2820

tgggctggcc aaactgctgg gtgcggaaga gaaagaatac catgcagaag gaggcaaagt 2880

gcctatcaag tggatggcat tggaatcaat tttacacaga atctataccc accagagtga 2940

tgtctggagc tacggggtga ctgtttggga gttgatgacc tttggatcca agccatatga 3000

cggaatccct gccagcgaga tctcctccat cctggagaaa ggagaacgcc tccctcagcc 3060

acccatatgt accatcgatg tctacatgat catggtcaag tgctggatga tagacgcaga 3120

tagtcgccca aagttccgtg agttgatcat cgaattctcc aaaatggccc gagacccca 3180

gcgctacctt gtcattcagg gggatgaaag aatgcatttg ccaagtccta cagactccaa 3240

cttctaccgt gccctgatgg atgaagaaga catggacgac gtggtggatg ccgacgagta 3300

cctcatccca cagcagggct tcttcagcag cccctccacg tcacggactc ccctcctgag 3360

ctctctgagt gcaaccagca acaattccac cgtggcttgc attgatagaa atgggctgca 3420

aagctgtccc atcaaggaag acagcttctt gcagcgatac agctcagacc ccacaggcgc 3480

cttgactgag gacagcatag acgacacctt cctcccagtg cctgaataca taaaccagtc 3540

cgttcccaaa aggcccgctg gctctgtgca gaatcctgtc tatcacaatc agcctctgaa 3600

ccccgcgccc agcagagacc cacactacca ggacccccac agcactgcag tgggcaaccc 3660

cgagtatctc aacactgtcc agcccacctg tgtcaacagc acattcgaca gccctgccca 3720

ctgggcccag aaaggcagcc accaaattag cctggacaac cctgactacc agcaggactt 3780

ctttcccaag gaagccaagc caaatggcat ctttaagggc tccacagctg aaaatgcaga 3840

atacctaagg gtcgcgccac aaagcagtga atttattgga gcatga 3886

<210> 723
<211> 3887
<212> DNA
<213> Homo sapiens

<400> 723
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacagcgt ggacaacccc cacgtgtgcc 2580

gcctgctggg catctgcctc acctccaccg tgcagctcat cacgcagctc atgcccttcg 2640

gctgcctcct ggactatgtc cgggaacaca aagacaatat tggctcccag tacctgctca 2700

actggtgtgt gcagatcgca aagggcatga actacttgga ggaccgtcgc ttggtgcacc 2760

gcgacctggc agccaggaac gtactggtga aaacaccgca gcatgtcaag atcacagatt 2820

ttgggctggc caaactgctg ggtgcggaag agaaagaata ccatgcagaa ggaggcaaag 2880

tgcctatcaa gtggatggca ttggaatcaa ttttacacag aatctatacc caccagagtg 2940

atgtctggag ctacggggtg actgtttggg agttgatgac ctttggatcc aagccatatg 3000

acggaatccc tgccagcgag atctcctcca tcctggagaa aggagaacgc ctccctcagc 3060

cacccatatg taccatcgat gtctacatga tcatggtcaa gtgctggatg atagacgcag 3120

atagtcgccc aaagttccgt gagttgatca tcgaattctc caaaatggcc cgagaccccc 3180

agcgctacct tgtcattcag ggggatgaaa gaatgcattt gccaagtcct acagactcca 3240

acttctaccg tgccctgatg gatgaagaag acatggacga cgtggtggat gccgacgagt 3300

acctcatccc acagcagggc ttcttcagca gcccctccac gtcacggact cccctcctga 3360

gctctctgag tgcaaccagc aacaattcca ccgtggcttg cattgataga aatgggctgc 3420

aaagctgtcc catcaaggaa gacagcttct tgcagcgata cagctcagac cccacaggcg 3480

ccttgactga ggacagcata gacgacacct tcctcccagt gcctgaatac ataaaccagt 3540

ccgttcccaa aaggcccgct ggctctgtgc agaatcctgt ctatcacaat cagcctctga 3600

accccgcgcc cagcagagac ccacactacc aggaccccca cagcactgca gtgggcaacc 3660

ccgagtatct caacactgtc cagcccacct gtgtcaacag cacattcgac agccctgccc 3720

actgggccca gaaaggcagc caccaaatta gcctggacaa ccctgactac cagcaggact 3780

tctttcccaa ggaagccaag ccaaatggca tctttaaggg ctccacagct gaaaatgcag 3840

aatacctaag ggtcgcgcca caaagcagtg aatttattgg agcatga 3887

<210> 724
<211> 3881
<212> DNA
<213> Homo sapiens

<400> 724
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc cggtcacgtg tgccgcctgc 2580

tgggcatctg cctcacctcc accgtgcagc tcatcacgca gctcatgccc ttcggctgcc 2640

tcctggacta tgtccgggaa cacaaagaca atattggctc ccagtacctg ctcaactggt 2700

gtgtgcagat cgcaaagggc atgaactact tggaggaccg tcgcttggtg caccgcgacc 2760

tggcagccag gaacgtactg gtgaaaacac cgcagcatgt caagatcaca gattttgggc 2820

tggccaaact gctgggtgcg gaagagaaag aataccatgc agaaggaggc aaagtgccta 2880

tcaagtggat ggcattggaa tcaattttac acagaatcta tacccaccag agtgatgtct 2940

ggagctacgg ggtgactgtt tgggagttga tgacctttgg atccaagcca tatgacggaa 3000

tccctgccag cgagatctcc tccatcctgg agaaaggaga acgcctccct cagccaccca 3060

tatgtaccat cgatgtctac atgatcatgg tcaagtgctg gatgatagac gcagatagtc 3120

gcccaaagtt ccgtgagttg atcatcgaat tctccaaaat ggcccgagac ccccagcgct 3180

accttgtcat tcaggggat gaaagaatgc atttgccaag tcctacagac tccaacttct 3240

accgtgccct gatggatgaa gaagacatgg acgacgtggt ggatgccgac gagtacctca 3300

tcccacagca gggcttcttc agcagcccct ccacgtcacg gactcccctc ctgagctctc 3360

tgagtgcaac cagcaacaat tccaccgtgg cttgcattga tagaaatggg ctgcaaagct 3420

gtcccatcaa ggaagacagc ttcttgcagc gatacagctc agaccccaca ggcgccttga 3480

ctgaggacag catagacgac accttcctcc cagtgcctga atacataaac cagtccgttc 3540

ccaaaaggcc cgctggctct gtgcagaatc ctgtctatca caatcagcct ctgaaccccg 3600

cgcccagcag agacccacac taccaggacc cccacagcac tgcagtgggc aaccccgagt 3660

atctcaacac tgtccagccc acctgtgtca acagcacatt cgacagccct gcccactggg 3720

cccagaaagg cagccaccaa attagcctgg acaaccctga ctaccagcag gacttctttc 3780

ccaaggaagc caagccaaat ggcatcttta agggctccac agctgaaaat gcagaatacc 3840

taagggtcgc gccacaaagc agtgaattta ttggagcatg a 3881

<210> 725
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 725

cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctaa 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 726
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 726
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgtgctgg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 727
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 727

cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820

ccaaacagct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 728
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 728
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc atcggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 729
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 729

cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt tatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 730
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 730
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaacccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 731
<211> 3866
<212> DNA
<213> Homo sapiens

<400> 731

cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaatcatctc cgaaagccaa caaggaaatc ctcgatgaag 2520

cctacgtgat ggccagcgtg gacaaccccc acgtgtgccg cctgctgggc atctgcctca 2580

cctccaccgt gcagctcatc acgcagctca tgcccttcgg ctgcctcctg gactatgtcc 2640

gggaacacaa agacaatatt ggctcccagt acctgctcaa ctggtgtgtg cagatcgcaa 2700

agggcatgaa ctacttggag gaccgtcgct tggtgcaccg cgacctggca gccaggaacg 2760

tactggtgaa aacaccgcag catgtcaaga tcacagattt tgggctggcc aaactgctgg 2820

gtgcggaaga gaaagaatac catgcagaag gaggcaaagt gcctatcaag tggatggcat 2880

tggaatcaat tttacacaga atctataccc accagagtga tgtctggagc tacggggtga 2940

ctgtttggga gttgatgacc tttggatcca agccatatga cggaatccct gccagcgaga 3000

tctcctccat cctggagaaa ggagaacgcc tccctcagcc acccatatgt accatcgatg 3060

tctacatgat catggtcaag tgctggatga tagacgcaga tagtcgccca aagttccgtg 3120

agttgatcat cgaattctcc aaaatggccc gagaccccca gcgctacctt gtcattcagg 3180

gggatgaaag aatgcatttg ccaagtccta cagactccaa cttctaccgt gccctgatgg 3240

atgaagaaga catggacgac gtggtggatg ccgacgagta cctcatccca cagcagggct 3300

tcttcagcag cccctccacg tcacggactc ccctcctgag ctctctgagt gcaaccagca 3360

acaattccac cgtggcttgc attgatagaa atgggctgca aagctgtccc atcaaggaag 3420

acagcttctt gcagcgatac agctcagacc ccacaggcgc cttgactgag gacagcatag 3480

acgacacctt cctcccagtg cctgaataca taaaccagtc cgttcccaaa aggcccgctg 3540

gctctgtgca gaatcctgtc tatcacaatc agcctctgaa ccccgcgccc agcagagacc 3600

cacactacca ggacccccac agcactgcag tgggcaaccc cgagtatctc aacactgtcc 3660

agcccacctg tgtcaacagc acattcgaca gccctgccca ctgggcccag aaaggcagcc 3720

accaaattag cctggacaac cctgactacc agcaggactt ctttcccaag gaagccaagc 3780

caaatggcat ctttaagggc tccacagctg aaaatgcaga atacctaagg gtcgcgccac 3840

aaagcagtga atttattgga gcatga 3866

<210> 732
<211> 3860
<212> DNA
<213> Homo sapiens

<400> 732
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaatcgaaag ccaacaagga aatcctcgat gaagcctacg 2520

tgatggccag cgtggacaac ccccacgtgt gccgcctgct gggcatctgc ctcacctcca 2580

ccgtgcagct catcacgcag ctcatgccct tcggctgcct cctggactat gtccgggaac 2640

acaaagacaa tattggctcc cagtacctgc tcaactggtg tgtgcagatc gcaaagggca 2700

tgaactactt ggaggaccgt cgcttggtgc accgcgacct ggcagccagg aacgtactgg 2760

tgaaaacacc gcagcatgtc aagatcacag attttgggct ggccaaactg ctgggtgcgg 2820

aagagaaaga ataccatgca gaaggaggca aagtgcctat caagtggatg gcattggaat 2880

caattttaca cagaatctat acccaccaga gtgatgtctg gagctacggg gtgactgttt 2940

gggagttgat gacctttgga tccaagccat atgacggaat ccctgccagc gagatctcct 3000

ccatcctgga gaaaggagaa cgcctccctc agccacccat atgtaccatc gatgtctaca 3060

tgatcatggt caagtgctgg atgatagacg cagatagtcg cccaaagttc cgtgagttga 3120

tcatcgaatt ctccaaaatg gcccgagacc cccagcgcta ccttgtcatt caggggggatg 3180

aaagaatgca tttgccaagt cctacagact ccaacttcta ccgtgccctg atggatgaag 3240

aagacatgga cgacgtggtg gatgccgacg agtacctcat cccacagcag ggcttcttca 3300

gcagcccctc cacgtcacgg actcccctcc tgagctctct gagtgcaacc agcaacaatt 3360

ccaccgtggc ttgcattgat agaaatgggc tgcaaagctg tcccatcaag gaagacagct 3420

tcttgcagcg atacagctca gaccccacag gcgccttgac tgaggacagc atagacgaca 3480

ccttcctccc agtgcctgaa tacataaacc agtccgttcc caaaaggccc gctggctctg 3540

tgcagaatcc tgtctatcac aatcagcctc tgaaccccgc gcccagcaga gacccacact 3600

accaggaccc ccacagcact gcagtgggca accccgagta tctcaacact gtccagccca 3660

cctgtgtcaa cagcacattc gacagccctg cccactgggc ccagaaaggc agccaccaaa 3720

ttagcctgga caaccctgac taccagcagg acttctttcc caaggaagcc aagccaaatg 3780

gcatctttaa gggctccaca gctgaaaatg cagaatacct aagggtcgcg ccacaaagca 3840

gtgaatttat tggagcatga 3860

<210> 733
<211> 3860
<212> DNA
<213> Homo sapiens

<400> 733
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaatcgaaag ccaacaagga aatcctcgat gaagcctacg 2520

tgatggccag cgtggacaac ccccacgtgt gccgcctgct gggcatctgc ctcacctcca 2580

ccgtgcagct catcacgcag ctcatgccct tcggctgcct cctggactat gtccgggaac 2640

acaaagacaa tattggctcc cagtacctgc tcaactggtg tgtgcagatc gcaaagggca 2700

tgaactactt ggaggaccgt cgcttggtgc accgcgacct ggcagccagg aacgtactgg 2760

tgaaaacacc gcagcatgtc aagatcacag attttgggct ggccaaactg ctgggtgcgg 2820

aagagaaaga ataccatgca gaaggaggca aagtgcctat caagtggatg gcattggaat 2880

caattttaca cagaatctat acccaccaga gtgatgtctg gagctacggg gtgactgttt 2940

gggagttgat gacctttgga tccaagccat atgacggaat ccctgccagc gagatctcct 3000

ccatcctgga gaaaggagaa cgcctccctc agccacccat atgtaccatc gatgtctaca 3060

tgatcatggt caagtgctgg atgatagacg cagatagtcg cccaaagttc cgtgagttga 3120

tcatcgaatt ctccaaaatg gcccgagacc cccagcgcta ccttgtcatt caggggggatg 3180

aaagaatgca tttgccaagt cctacagact ccaacttcta ccgtgccctg atggatgaag 3240

aagacatgga cgacgtggtg gatgccgacg agtacctcat cccacagcag ggcttcttca 3300

gcagcccctc cacgtcacgg actcccctcc tgagctctct gagtgcaacc agcaacaatt 3360

ccaccgtggc ttgcattgat agaaatgggc tgcaaagctg tcccatcaag gaagacagct 3420

tcttgcagcg atacagctca gaccccacag gcgccttgac tgaggacagc atagacgaca 3480

ccttcctccc agtgcctgaa tacataaacc agtccgttcc caaaaggccc gctggctctg 3540

tgcagaatcc tgtctatcac aatcagcctc tgaaccccgc gcccagcaga gacccacact 3600

accaggaccc ccacagcact gcagtgggca accccgagta tctcaacact gtccagccca 3660

cctgtgtcaa cagcacattc gacagccctg cccactgggc ccagaaaggc agccaccaaa 3720

ttagcctgga caaccctgac taccagcagg acttctttcc caaggaagcc aagccaaatg 3780

gcatctttaa gggctccaca gctgaaaatg cagaatacct aagggtcgcg ccacaaagca 3840

gtgaatttat tggagcatga 3860

<210> 734
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 734
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcggg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 735
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 735

cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcggg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 736
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 736
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820

ccaaacagct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 737
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 737

cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgt 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 738
<211> 3860
<212> DNA
<213> Homo sapiens

<400> 738
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaatcgaaag ccaacaagga aatcctcgat gaagcctacg 2520

tgatggccag cgtggacaac ccccacgtgt gccgcctgct gggcatctgc ctcacctcca 2580

ccgtgcagct catcacgcag ctcatgccct tcggctgcct cctggactat gtccgggaac 2640

acaaagacaa tattggctcc cagtacctgc tcaactggtg tgtgcagatc gcaaagggca 2700

tgaactactt ggaggaccgt cgcttggtgc accgcgacct ggcagccagg aacgtactgg 2760

tgaaaacacc gcagcatgtc aagatcacag attttgggct ggccaaactg ctgggtgcgg 2820

aagagaaaga ataccatgca gaaggaggca aagtgcctat caagtggatg gcattggaat 2880

caattttaca cagaatctat acccaccaga gtgatgtctg gagctacggg gtgactgttt 2940

gggagttgat gacctttgga tccaagccat atgacggaat ccctgccagc gagatctcct 3000

ccatcctgga gaaaggagaa cgcctccctc agccacccat atgtaccatc gatgtctaca 3060

tgatcatggt caagtgctgg atgatagacg cagatagtcg cccaaagttc cgtgagttga 3120

tcatcgaatt ctccaaaatg gcccgagacc cccagcgcta ccttgtcatt caggggggatg 3180

aaagaatgca tttgccaagt cctacagact ccaacttcta ccgtgccctg atggatgaag 3240

aagacatgga cgacgtggtg gatgccgacg agtacctcat cccacagcag ggcttcttca 3300

gcagcccctc cacgtcacgg actcccctcc tgagctctct gagtgcaacc agcaacaatt 3360

ccaccgtggc ttgcattgat agaaatgggc tgcaaagctg tcccatcaag gaagacagct 3420

tcttgcagcg atacagctca gaccccacag gcgccttgac tgaggacagc atagacgaca 3480

ccttcctccc agtgcctgaa tacataaacc agtccgttcc caaaaggccc gctggctctg 3540

tgcagaatcc tgtctatcac aatcagcctc tgaaccccgc gcccagcaga gacccacact 3600

accaggaccc ccacagcact gcagtgggca accccgagta tctcaacact gtccagccca 3660

cctgtgtcaa cagcacattc gacagccctg cccactgggc ccagaaaggc agccaccaaa 3720

ttagcctgga caaccctgac taccagcagg acttctttcc caaggaagcc aagccaaatg 3780

gcatctttaa gggctccaca gctgaaaatg cagaatacct aagggtcgcg ccacaaagca 3840

gtgaatttat tggagcatga 3860

<210> 739
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 739

cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 740
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 740
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg taactgaatt caaaaagatc aaagtgctga 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 741
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 741

cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

cctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cacctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 742
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 742
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaacccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 743
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 743

cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 744
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 744
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaacccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 745
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 745
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 746
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 746
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaacccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 747
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 747

cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 748
<211> 3863
<212> DNA
<213> Homo sapiens

<400> 748
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520

acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc tgcctcacct 2580

ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640

aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700

gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760

tggtgaaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880

aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940

tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000

cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120

tgatcatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180

atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300

tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca accagcaaca 3360

attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac agcatagacg 3480

acaccttcct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540

ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagacccac 3600

actaccagga cccccacagc actgcagtgg gcaacccccga gtatctcaac actgtccagc 3660

ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780

atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840

gcagtgaatt tattggagca tga 3863

<210> 749
<211> 3860
<212> DNA
<213> Homo sapiens

<400> 749

cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gttccgaaag ccaacaagga aatcctcgat gaagcctacg 2520

tgatggccag cgtggacaac ccccacgtgt gccgcctgct gggcatctgc ctcacctcca 2580

ccgtgcagct catcacgcag ctcatgccct tcggctgcct cctggactat gtccgggaac 2640

acaaagacaa tattggctcc cagtacctgc tcaactggtg tgtgcagatc gcaaagggca 2700

tgaactactt ggaggaccgt cgcttggtgc accgcgacct ggcagccagg aacgtactgg 2760

tgaaaacacc gcagcatgtc aagatcacag attttgggct ggccaaactg ctgggtgcgg 2820

aagagaaaga ataccatgca gaaggaggca aagtgcctat caagtggatg gcattggaat 2880

caattttaca cagaatctat acccaccaga gtgatgtctg gagctacggg gtgactgttt 2940

gggagttgat gacctttgga tccaagccat atgacggaat ccctgccagc gagatctcct 3000

ccatcctgga gaaaggagaa cgcctccctc agccacccat atgtaccatc gatgtctaca 3060

tgatcatggt caagtgctgg atgatagacg cagatagtcg cccaaagttc cgtgagttga 3120

tcatcgaatt ctccaaaatg gcccgagacc cccagcgcta ccttgtcatt caggggggatg 3180

aaagaatgca tttgccaagt cctacagact ccaacttcta ccgtgccctg atggatgaag 3240

aagacatgga cgacgtggtg gatgccgacg agtacctcat cccacagcag ggcttcttca 3300

gcagcccctc cacgtcacgg actcccctcc tgagctctct gagtgcaacc agcaacaatt 3360

ccaccgtggc ttgcattgat agaaatgggc tgcaaagctg tcccatcaag gaagacagct 3420

tcttgcagcg atacagctca gaccccacag gcgccttgac tgaggacagc atagacgaca 3480

ccttcctccc agtgcctgaa tacataaacc agtccgttcc caaaaggccc gctggctctg 3540

tgcagaatcc tgtctatcac aatcagcctc tgaaccccgc gcccagcaga gacccacact 3600

accaggaccc ccacagcact gcagtgggca accccgagta tctcaacact gtccagccca 3660

cctgtgtcaa cagcacattc gacagccctg cccactgggc ccagaaaggc agccaccaaa 3720

ttagcctgga caaccctgac taccagcagg acttctttcc caaggaagcc aagccaaatg 3780

gcatctttaa gggctccaca gctgaaaatg cagaatacct aagggtcgcg ccacaaagca 3840

gtgaatttat tggagcatga 3860

<210> 750
<211> 3869
<212> DNA
<213> Homo sapiens

<400> 750
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaaccaacat ctccgaaagc caacaaggaa atcctcgatg 2520

aagcctacgt gatggccagc gtggacaacc cccacgtgtg ccgcctgctg ggcatctgcc 2580

tcacctccac cgtgcagctc atcacgcagc tcatgcccтt cggctgcctc ctggactatg 2640

tccgggaaca caaagacaat attggctccc agtacctgct caactggtgt gtgcagatcg 2700

caaagggcat gaactacttg gaggaccgtc gcttggtgca ccgcgacctg gcagccagga 2760

acgtactggt gaaaacaccg cagcatgtca agatcacaga ttttgggctg gccaaactgc 2820

tgggtgcgga agagaaagaa taccatgcag aaggaggcaa agtgcctatc aagtggatgg 2880

cattggaatc aattttacac agaatctata cccaccagag tgatgtctgg agctacgggg 2940

tgactgtttg ggagttgatg acctttggat ccaagccata tgacggaatc cctgccagcg 3000

agatctcctc catcctggag aaaggagaac gcctccctca gccacccata tgtaccatcg 3060

atgtctacat gatcatggtc aagtgctgga tgatagacgc agatagtcgc ccaaagttcc 3120

gtgagttgat catcgaattc tccaaaatgg cccgagaccc ccagcgctac cttgtcattc 3180

agggggatga aagaatgcat ttgccaagtc ctacagactc caacttctac cgtgccctga 3240

tggatgaaga agacatggac gacgtggtgg atgccgacga gtacctcatc ccacagcagg 3300

gcttcttcag cagcccctcc acgtcacgga ctcccctcct gagctctctg agtgcaacca 3360

gcaacaattc caccgtggct tgcattgata gaaatgggct gcaaagctgt cccatcaagg 3420

aagacagctt cttgcagcga tacagctcag accccacagg cgccttgact gaggacagca 3480

tagacgacac cttcctccca gtgcctgaat acataaacca gtccgttccc aaaaggcccg 3540

ctggctctgt gcagaatcct gtctatcaca atcagcctct gaaccccgcg cccagcagag 3600

acccacacta ccaggacccc cacagcactg cagtgggcaa ccccgagtat ctcaacactg 3660

tccagcccac ctgtgtcaac agcacattcg acagccctgc ccactgggcc cagaaaggca 3720

gccaccaaat tagcctggac aaccctgact accagcagga cttctttccc aaggaagcca 3780

agccaaatgg catctttaag ggctccacag ctgaaaatgc agaataccta agggtcgcgc 3840

cacaaagcag tgaatttatt ggagcatga 3869

<210> 751
<211> 3866
<212> DNA
<213> Homo sapiens

<400> 751

cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaaccatctc cgaaagccaa caaggaaatc ctcgatgaag 2520

cctacgtgat ggccagcgtg gacaaccccc acgtgtgccg cctgctgggc atctgcctca 2580

cctccaccgt gcagctcatc acgcagctca tgcccttcgg ctgcctcctg gactatgtcc 2640

gggaacacaa agacaatatt ggctcccagt acctgctcaa ctggtgtgtg cagatcgcaa 2700

agggcatgaa ctacttggag gaccgtcgct tggtgcaccg cgacctggca gccaggaacg 2760

tactggtgaa aacaccgcag catgtcaaga tcacagattt tgggctggcc aaactgctgg 2820

gtgcggaaga gaaagaatac catgcagaag gaggcaaagt gcctatcaag tggatggcat 2880

tggaatcaat tttacacaga atctataccc accagagtga tgtctggagc tacggggtga 2940

ctgtttggga gttgatgacc tttggatcca agccatatga cggaatccct gccagcgaga 3000

tctcctccat cctggagaaa ggagaacgcc tccctcagcc acccatatgt accatcgatg 3060

tctacatgat catggtcaag tgctggatga tagacgcaga tagtcgccca aagttccgtg 3120

agttgatcat cgaattctcc aaaatggccc gagaccccca gcgctacctt gtcattcagg 3180

gggatgaaag aatgcatttg ccaagtccta cagactccaa cttctaccgt gccctgatgg 3240

atgaagaaga catggacgac gtggtggatg ccgacgagta cctcatccca cagcagggct 3300

tcttcagcag cccctccacg tcacggactc ccctcctgag ctctctgagt gcaaccagca 3360

acaattccac cgtggcttgc attgatagaa atgggctgca aagctgtccc atcaaggaag 3420

acagcttctt gcagcgatac agctcagacc ccacaggcgc cttgactgag gacagcatag 3480

acgacacctt cctcccagtg cctgaataca taaaccagtc cgttcccaaa aggcccgctg 3540

gctctgtgca gaatcctgtc tatcacaatc agcctctgaa ccccgcgccc agcagagacc 3600

cacactacca ggacccccac agcactgcag tgggcaaccc cgagtatctc aacactgtcc 3660

agcccacctg tgtcaacagc acattcgaca gccctgccca ctgggcccag aaaggcagcc 3720

accaaattag cctggacaac cctgactacc agcaggactt ctttcccaag gaagccaagc 3780

caaatggcat ctttaagggc tccacagctg aaaatgcaga atacctaagg gtcgcgccac 3840

aaagcagtga atttattgga gcatga 3866

<210> 752
<211> 3854
<212> DNA
<213> Homo sapiens

<400> 752
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaaccct cgatgaagcc tacgtgatgg 2520

ccagcgtgga caacccccac gtgtgccgcc tgctgggcat ctgcctcacc tccaccgtgc 2580

agctcatcac gcagctcatg cccttcggct gcctcctgga ctatgtccgg gaacacaaag 2640

acaatattgg ctcccagtac ctgctcaact ggtgtgtgca gatcgcaaag ggcatgaact 2700

acttggagga ccgtcgcttg gtgcaccgcg acctggcagc caggaacgta ctggtgaaaa 2760

caccgcagca tgtcaagatc acagattttg ggctggccaa actgctgggt gcggaagaga 2820

aagaatacca tgcagaagga ggcaaagtgc ctatcaagtg gatggcattg gaatcaattt 2880

tacacagaat ctatacccac cagagtgatg tctggagcta cggggtgact gtttgggagt 2940

tgatgacctt tggatccaag ccatatgacg gaatccctgc cagcgagatc tcctccatcc 3000

tggagaaagg agaacgcctc cctcagccac ccatatgtac catcgatgtc tacatgatca 3060

tggtcaagtg ctggatgata gacgcagata gtcgcccaaa gttccgtgag ttgatcatcg 3120

aattctccaa aatggcccga gacccccagc gctaccttgt cattcagggg gatgaaagaa 3180

tgcatttgcc aagtcctaca gactccaact tctaccgtgc cctgatggat gaagaagaca 3240

tggacgacgt ggtggatgcc gacgagtacc tcatcccaca gcagggcttc ttcagcagcc 3300

cctccacgtc acggactccc ctcctgagct ctctgagtgc aaccagcaac aattccaccg 3360

tggcttgcat tgatagaaat gggctgcaaa gctgtcccat caaggaagac agcttcttgc 3420

agcgatacag ctcagacccc acaggcgcct tgactgagga cagcatagac gacaccttcc 3480

tcccagtgcc tgaatacata aaccagtccg ttcccaaaag gcccgctggc tctgtgcaga 3540

atcctgtcta tcacaatcag cctctgaacc ccgcgcccag cagagaccca cactaccagg 3600

acccccacag cactgcagtg ggcaaccccg agtatctcaa cactgtccag cccacctgtg 3660

tcaacagcac attcgacagc cctgcccact gggcccagaa aggcagccac caaattagcc 3720

tggacaaccc tgactaccag caggacttct ttcccaagga agccaagcca aatggcatct 3780

ttaagggctc cacagctgaa aatgcagaat acctaagggt cgcgccacaa agcagtgaat 3840

ttattggagc atga 3854

<210> 753
<211> 3854
<212> DNA
<213> Homo sapiens

<400> 753

cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacact cgatgaagcc tacgtgatgg 2520

ccagcgtgga caacccccac gtgtgccgcc tgctgggcat ctgcctcacc tccaccgtgc 2580

agctcatcac gcagctcatg cccttcggct gcctcctgga ctatgtccgg gaacacaaag 2640

acaatattgg ctcccagtac ctgctcaact ggtgtgtgca gatcgcaaag ggcatgaact 2700

acttggagga ccgtcgcttg gtgcaccgcg acctggcagc caggaacgta ctggtgaaaa 2760

caccgcagca tgtcaagatc acagattttg ggctggccaa actgctgggt gcggaagaga 2820

aagaatacca tgcagaagga ggcaaagtgc ctatcaagtg gatggcattg gaatcaattt 2880

tacacagaat ctatacccac cagagtgatg tctggagcta cggggtgact gtttgggagt 2940

tgatgacctt tggatccaag ccatatgacg gaatccctgc cagcgagatc tcctccatcc 3000

tggagaaagg agaacgcctc cctcagccac ccatatgtac catcgatgtc tacatgatca 3060

tggtcaagtg ctggatgata gacgcagata gtcgcccaaa gttccgtgag ttgatcatcg 3120

aattctccaa aatggcccga gacccccagc gctaccttgt cattcagggg gatgaaagaa 3180

tgcatttgcc aagtcctaca gactccaact tctaccgtgc cctgatggat gaagaagaca 3240

tggacgacgt ggtggatgcc gacgagtacc tcatcccaca gcagggcttc ttcagcagcc 3300

cctccacgtc acggactccc ctcctgagct ctctgagtgc aaccagcaac aattccaccg 3360

tggcttgcat tgatagaaat gggctgcaaa gctgtcccat caaggaagac agcttcttgc 3420

agcgatacag ctcagacccc acaggcgcct tgactgagga cagcatagac gacaccttcc 3480

tcccagtgcc tgaatacata aaccagtccg ttcccaaaag gcccgctggc tctgtgcaga 3540

atcctgtcta tcacaatcag cctctgaacc ccgcgcccag cagagaccca cactaccagg 3600

acccccacag cactgcagtg ggcaaccccg agtatctcaa cactgtccag cccacctgtg 3660

tcaacagcac attcgacagc cctgcccact gggcccagaa aggcagccac caaattagcc 3720

tggacaaccc tgactaccag caggacttct ttcccaagga agccaagcca aatggcatct 3780

ttaagggctc cacagctgaa aatgcagaat acctaagggt cgcgccacaa agcagtgaat 3840

ttattggagc atga 3854

<210> 754
<211> 3887
<212> DNA
<213> Homo sapiens

<400> 754
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacagcgt ggacaacccc cacgtgtgcc 2580

gcctgctggg catctgcctc acctccaccg tgcagctcat cacgcagctc atgcccttcg 2640

gctgcctcct ggactatgtc cgggaacaca aagacaatat tggctcccag tacctgctca 2700

actggtgtgt gcagatcgca aagggcatga actacttgga ggaccgtcgc ttggtgcacc 2760

gcgacctggc agccaggaac gtactggtga aaacaccgca gcatgtcaag atcacagatt 2820

ttgggctggc caaactgctg ggtgcggaag agaaagaata ccatgcagaa ggaggcaaag 2880

tgcctatcaa gtggatggca ttggaatcaa ttttacacag aatctatacc caccagagtg 2940

atgtctggag ctacggggtg actgtttggg agttgatgac ctttggatcc aagccatatg 3000

acggaatccc tgccagcgag atctcctcca tcctggagaa aggagaacgc ctccctcagc 3060

cacccatatg taccatcgat gtctacatga tcatggtcaa gtgctggatg atagacgcag 3120

atagtcgccc aaagttccgt gagttgatca tcgaattctc caaaatggcc cgagaccccc 3180

agcgctacct tgtcattcag ggggatgaaa gaatgcattt gccaagtcct acagactcca 3240

acttctaccg tgccctgatg gatgaagaag acatggacga cgtggtggat gccgacgagt 3300

acctcatccc acagcagggc ttcttcagca gcccctccac gtcacggact cccctcctga 3360

gctctctgag tgcaaccagc aacaattcca ccgtggcttg cattgataga aatgggctgc 3420

aaagctgtcc catcaaggaa gacagcttct tgcagcgata cagctcagac cccacaggcg 3480

ccttgactga ggacagcata gacgacacct tcctcccagt gcctgaatac ataaaccagt 3540

ccgttcccaa aaggcccgct ggctctgtgc agaatcctgt ctatcacaat cagcctctga 3600

accccgcgcc cagcagagac ccacactacc aggacccccа cagcactgca gtgggcaacc 3660

ccgagtatct caacactgtc cagcccacct gtgtcaacag cacattcgac agccctgccc 3720

actgggccca gaaaggcagc caccaaatta gcctggacaa ccctgactac cagcaggact 3780

tctttcccaa ggaagccaag ccaaatggca tctttaaggg ctccacagct gaaaatgcag 3840

aatacctaag ggtcgcgcca caaagcagtg aatttattgg agcatga 3887

<210> 755
<211> 3884
<212> DNA
<213> Homo sapiens

<400> 755
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacccccac gtgtgccgcc 2580

tgctgggcat ctgcctcacc tccaccgtgc agctcatcac gcagctcatg cccttcggct 2640

gcctcctgga ctatgtccgg gaacacaaag acaatattgg ctcccagtac ctgctcaact 2700

ggtgtgtgca gatcgcaaag ggcatgaact acttggagga ccgtcgcttg gtgcaccgcg 2760

acctggcagc caggaacgta ctggtgaaaa caccgcagca tgtcaagatc acagattttg 2820

ggctggccaa actgctgggt gcggaagaga aagaatacca tgcagaagga ggcaaagtgc 2880

ctatcaagtg gatggcattg gaatcaattt tacacagaat ctatacccac cagagtgatg 2940

tctggagcta cggggtgact gtttgggagt tgatgacctt tggatccaag ccatatgacg 3000

gaatccctgc cagcgagatc tcctccatcc tggagaaagg agaacgcctc cctcagccac 3060

ccatatgtac catcgatgtc tacatgatca tggtcaagtg ctggatgata gacgcagata 3120

gtcgcccaaa gttccgtgag ttgatcatcg aattctccaa aatggcccga gacccccagc 3180

gctaccttgt cattcagggg gatgaaagaa tgcatttgcc aagtcctaca gactccaact 3240

tctaccgtgc cctgatggat gaagaagaca tggacgacgt ggtggatgcc gacgagtacc 3300

tcatcccaca gcagggcttc ttcagcagcc cctccacgtc acggactccc ctcctgagct 3360

ctctgagtgc aaccagcaac aattccaccg tggcttgcat tgatagaaat gggctgcaaa 3420

gctgtcccat caaggaagac agcttcttgc agcgatacag ctcagacccc acaggcgcct 3480

tgactgagga cagcatagac gacaccttcc tcccagtgcc tgaatacata aaccagtccg 3540

ttcccaaaag gcccgctggc tctgtgcaga atcctgtcta tcacaatcag cctctgaacc 3600

ccgcgcccag cagagaccca cactaccagg acccccacag cactgcagtg ggcaaccccg 3660

agtatctcaa cactgtccag cccacctgtg tcaacagcac attcgacagc cctgcccact 3720

gggcccagaa aggcagccac caaattagcc tggacaaccc tgactaccag caggacttct 3780

ttcccaagga agccaagcca aatggcatct ttaagggctc cacagctgaa aatgcagaat 3840

acctaagggt cgcgccacaa agcagtgaat ttattggagc atga 3884

<210> 756
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 756
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacactgc agcatgtcaa gatcacagat tttgggctgg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 757
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 757

cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcggg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 758
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 758
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcggg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 759
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 759

cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcggg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 760
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 760
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcggg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 761
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 761
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcggg 2820

ccaaactgct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

<210> 762
<211> 3878
<212> DNA
<213> Homo sapiens

<400> 762
cccggcgcag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgaggcggc cggagtcccg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcgtccg cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600

tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900

gctccgggcg ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

gctgcacagg cccccgggag agcgactgcc tggtctgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080

tgaaccccga gggcaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560

cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620

gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740

gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccgg 1800

agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860

gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920

acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040

tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160

atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220

tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280

tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460

aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520

tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580

gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640

tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760

cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820

ccaaacagct gggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000

ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacccatat 3060

gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtcgcc 3120

caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180

ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240

gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300

cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360

gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga ccccacaggc gccttgactg 3480

aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540

aaaggcccgc tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcgc 3600

ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660

tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720

agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780

aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840

gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878