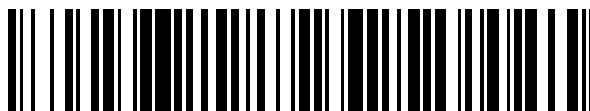


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 741 574**

51 Int. Cl.:

C12Q 1/68 (2008.01)

C07H 21/04 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **31.03.2005** **E 11178202 (5)**

97 Fecha y número de publicación de la concesión europea: **08.05.2019** **EP 2439285**

54 Título: **Método para determinar la respuesta del cáncer a tratamientos dirigidos al receptor del factor de crecimiento epidérmico**

30 Prioridad:

31.03.2004 US 558218 P
09.04.2004 US 561095 P
27.04.2004 US 565753 P
27.04.2004 US 565985 P
25.05.2004 US 574035 P
07.06.2004 US 577916 P
29.07.2004 US 592287 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
11.02.2020

73 Titular/es:

THE GENERAL HOSPITAL CORPORATION
(50.0%)
55 Fruit Street
Boston, MA 02114, US y
DANA-FARBER CANCER INSTITUTE, INC.
(50.0%)

72 Inventor/es:

BELL, DAPHNE WINIFRED;
HABER, DANIEL A.;
JANNE, PASI ANTERO;
JOHNSON, BRUCE E.;
LYNCH, THOMAS J.;
MEYERSON, MATTHEW;
PAEZ, JUAN GUILLERMO;
SELLERS, WILLIAM R.;
SETTLEMAN, JEFFREY E. y
SORDELLA, RAFFAELLA

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 741 574 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Método para determinar la respuesta del cáncer a tratamientos dirigidos al receptor del factor de crecimiento epidérmico.

Antecedentes

- 5 Los cánceres de células epiteliales, por ejemplo, cáncer de próstata, cáncer de mama, cáncer de colon, cáncer de pulmón, cáncer de páncreas, cáncer de ovario, cáncer del bazo, cáncer testicular, cáncer del timo, etc., son enfermedades caracterizadas por el crecimiento acelerado anormal de las células epiteliales. Este crecimiento acelerado causa inicialmente que se forme un tumor. Eventualmente, también se puede producir metástasis a diferentes órganos. Aunque se han logrado avances en el diagnóstico y el tratamiento de diferentes tipos de cáncer, estas enfermedades siguen provocando una mortalidad significativa.

- 10 El cáncer de pulmón sigue siendo la principal causa de muerte por cáncer en los países industrializados. Los cánceres que comienzan en los pulmones se dividen en dos tipos principales, el cáncer de pulmón de células no pequeñas y el cáncer de pulmón de células pequeñas, dependiendo de cómo lucen las células bajo un microscopio. El cáncer de pulmón de células no pequeñas (carcinoma de células escamosas, adenocarcinoma y carcinoma de células grandes) generalmente se propaga a otros órganos más lentamente de lo que lo hace el cáncer de pulmón de células pequeñas. Alrededor del 75 por ciento de los casos de cáncer de pulmón se clasifican como cáncer de pulmón de células no pequeñas (por ejemplo, los adenocarcinomas), y el otro 25 por ciento son cáncer de pulmón de células pequeñas. El cáncer de pulmón de células no pequeñas (NSCLC) es la principal causa de muerte por cáncer en los Estados Unidos, Japón y Europa Occidental. Para los pacientes con enfermedad avanzada, la quimioterapia proporciona un modesto beneficio para la supervivencia, pero a costa de una toxicidad significativa, lo que subraya la necesidad de agentes terapéuticos que estén específicamente dirigidos a las lesiones genéticas críticas que dirigen el crecimiento del tumor (Schiller J. H. et al., *N Engl J Med*, 346: 92 - 98, 2002).

- 25 El receptor del factor de crecimiento epidérmico (EGFR) es una proteína enlazada a la membrana de 170 kilodaltons (kDa), expresada en la superficie de las células epiteliales. EGFR es un miembro de la familia de receptores del factor de crecimiento de proteínas tirosina quinasas, una clase de moléculas reguladoras del ciclo celular. (W. J. Gullick et al., 1986, *Cancer Res.*, 46: 285 - 292). EGFR se activa cuando su ligando (ya sea EGF o TGF- α) se enlaza al dominio extracelular, lo que resulta en la autofosforilación del dominio intracelular de la tirosina quinasa del receptor (S. Cohen et al., 1980, *J. Biol. Chem.*, 255: 4834 - 4842; A. B. Schreiber et al., 1983, *J. Biol. Chem.*, 258: 846 - 853).

- 30 EGFR es el producto proteico de un oncogén que promueve el crecimiento, *erbB* o *ErbB1*, que no es más que uno de los miembros de una familia, es decir, la familia ERBB de protooncogenes, que se cree que desempeñan un papel fundamental en el desarrollo y la progresión de muchos cánceres humanos. En particular, se ha observado un aumento de expresión del EGFR en cáncer de mama, vejiga, pulmón, cabeza, cuello y estómago, así como de glioblastomas. La familia ERBB de oncogenes codifica cuatro receptores transmembrana estructuralmente relacionados, a saber, EGFR, HER-2/*neu* (*erbB2*), HER-3 (*erbB3*) y HER-4 (*erbB4*). Clínicamente, se ha reportado que la amplificación del oncogén ERBB y/o la sobreexpresión del receptor en tumores se correlacionan con la recurrencia de la enfermedad y con un pobre pronóstico para el paciente, así como con la capacidad de respuesta a la terapia. (L. Harris et al., 1999, *Int. J. Biol. Markers*, 14: 8 - 15, y J. Mendelsohn y J. Baselga, 2000, *Oncogene*, 19: 6550 - 6565).

- 40 El EGFR se compone de tres dominios principales, a saber, el dominio extracelular (ECD), que está glicosilado y contiene el bolsillo que enlaza al ligando con dos regiones ricas en cisteína; un dominio transmembrana corto, y un dominio intracelular que tiene actividad intrínseca de tirosina quinasa. La región transmembrana une el dominio de enlazamiento del ligando con el dominio intracelular. El análisis de secuencia de aminoácidos y de ADN, así como los estudios de las formas no glicosiladas del EGFR, indican que el esqueleto de la proteína del EGFR tiene una masa de 132 kDa, con 1186 residuos de aminoácidos (A. L. Ullrich 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; y F. L. V. Mayes y M. D. Waterfield, 1984, *The EMBO J.*, 3: 531 - 537).

- 50 El enlazamiento de EGF o TGF- α con EGFR activa una ruta de transducción de la señal y trae como resultado la proliferación celular. La dimerización, los cambios conformacionales y la internalización de las moléculas del EGFR funcionan para transmitir señales intracelulares que conducen a la regulación del crecimiento celular (G. Carpenter y S. Cohen, 1979, *Ann. Rev. Biochem.*, 48: 193 - 216). Las alteraciones genéticas que afectan la regulación de la función del receptor del factor de crecimiento, o que conducen a la sobreexpresión del receptor y/o del ligando, dan como resultado la proliferación celular. Además, se ha determinado que el EGFR desempeña un papel en la diferenciación celular, la mejora de la movilidad celular, la secreción de proteína, la neovascularización, invasión, metástasis y resistencia de las células cancerosas a los agentes quimioterapéuticos y de radiación. (M. -J. Oh et al., 2000, *Clin. Cancer Res.*, 6: 4760 - 4763).

Se han identificado una variedad de inhibidores del EGFR, incluyendo una cantidad que ya ha sometida a ensayos clínicos para el tratamiento de varios tipos de cáncer. Para ver un resumen reciente, véase de Bono, J. S. y

Rowinsky, E. K. (2002), "The ErbB receptor Family: A Therapeutic Targer for Cancer", Trends in Molecular Medicine, 8, S19-26.

Un conjunto prometedor de objetivos para la intervención terapéutica en el tratamiento del cáncer incluye los miembros del eje de quinasa HER. Se favorece frecuentemente su expresión en tumores epiteliales sólidos, por ejemplo, de la próstata, de pulmón y de mama, y también se favorece su expresión en los tumores de glioblastoma. El receptor del factor de crecimiento epidérmico (EGFR) es un miembro del eje de la quinasa HER, y ha sido el objetivo de elección para el desarrollo de varias terapias diferentes contra el cáncer. Los inhibidores de la tirosina quinasa del EGFR (EGFR-TKI) se encuentran entre estas terapias, ya que se requiere la fosforilación reversible de los residuos de tirosina para la activación de la ruta del EGFR. En otras palabras, los EGFR-TKI bloquean un receptor de la superficie celular responsable de la activación y/o el mantenimiento de la ruta de señalización celular que induce el crecimiento de células tumorales y la división. Específicamente, se cree que estos inhibidores interfieren con el dominio quinasa del EGFR, denominado como HER-1. Entre los EGFR-TKI más prometedores están tres series de compuestos: quinazolininas, piridopirimidinas y pirrolopirimidinas.

Dos de los compuestos más avanzados en el desarrollo clínico incluyen el Gefitinib (compuesto ZD1839 desarrollado por AstraZeneca UK Ltd.; disponible bajo el nombre comercial IRESSA; en lo sucesivo, "IRESSA") y Erlotinib (compuesto OSI-774 desarrollado por Genentech, Inc. y OSI Pharmaceuticals, Inc.; disponible bajo el nombre comercial de TARCEVA; en lo sucesivo, "TARCEVA"), ambos han generado resultados clínicos alentadores. El tratamiento convencional contra el cáncer tanto con IRESSA como con TARCEVA supone la administración oral diaria de no más de 500 mg de los compuestos respectivos. En mayo de 2003, IRESSA se convirtió en el primero de estos productos en llegar al mercado de los Estados Unidos, cuando fue aprobado para el tratamiento de pacientes con cáncer de pulmón avanzado de células no pequeñas.

IRESSA es una quinazolina oralmente activa que funciona inhibiendo en forma directa la fosforilación de la tirosina quinasa en la molécula del EGFR. Compite por el sitio de enlazamiento del trifosfato de adenosina (ATP), conduciendo a la supresión del eje de la quinasa HER. El mecanismo exacto de la respuesta de la IRESSA no es completamente entendido, sin embargo, los estudios sugieren que la presencia del EGFR es un prerrequisito necesario para su acción.

Una limitación significativa en el uso de estos compuestos es que los receptores de los mismos pueden desarrollar una resistencia a sus efectos terapéuticos después de que inicialmente responden a la terapia, o pueden no responder a los EGFR-TKI en ningún grado medible en absoluto. De hecho, sólo el 10 - 15 por ciento de los pacientes con cáncer de pulmón avanzado de células no pequeñas responde a los inhibidores de la quinasa del EGFR. Por lo tanto, sería extremadamente benéfico una mejor comprensión de los mecanismos moleculares subyacentes de la sensibilidad a IRESSA y TARCEVA en la selección de la terapia para aquellos individuos que tienen más probabilidades de beneficiarse de dicha terapia.

Existe una necesidad significativa en la técnica por un tratamiento satisfactorio del cáncer, y específicamente de cánceres de células epiteliales tales como los cánceres de pulmón, de ovario, de mama, de cerebro, de colon y de próstata, que incorporan los beneficios de la terapia del TKI y vencen la falta de sensibilidad mostrada por los pacientes. Dicho tratamiento podría tener un impacto dramático en la salud de las personas, y especialmente en las personas mayores, entre los que el cáncer es muy común.

Resumen

La presente invención se define en y por las reivindicaciones adjuntas. En la siguiente descripción y en las reivindicaciones, una referencia a una posición de nucleótido de la SEQ ID NO: 511 es una referencia a la numeración de acuerdo con la Figura 5.

La terapia con el inhibidor de tirosina quinasa (TKI) tal como con gefitinib (IRESSA®) no es efectiva en la gran mayoría de las personas que se ven afectadas con los cánceres mencionados anteriormente. Los presentes inventores han descubierto sorprendentemente que la presencia de mutaciones somáticas en el dominio quinasa del EGFR aumenta sustancialmente la sensibilidad del EGFR con TKI tal como IRESSA, TARCEVA. Por ejemplo menos del 30% de los pacientes que tienen tal cáncer son susceptibles al tratamiento por los TKI actuales, mientras que más del 50%, más preferiblemente del 60, 70, 80, 90% de los pacientes que tienen una mutación en el dominio quinasa del EGFR son susceptibles. Además, estas mutaciones confieren aumento de la actividad de la quinasa del EGFR. Por lo tanto, los pacientes que tienen estas mutaciones probablemente serán sensibles a la terapia actual con inhibidor de tirosina quinasa (TKI), por ejemplo, gefitinib.

En consecuencia, la presente invención permite un nuevo método para determinar la probabilidad de la efectividad de un tratamiento dirigido al receptor del factor de crecimiento epidérmico (EGFR) en un paciente humano afectado con cáncer. El método comprende la detección de la presencia o ausencia de al menos una diferencia del ácido nucleico en el dominio quinasa del gen erbB1 de dicho paciente en relación con el gen de tipo silvestre erbB1. La presencia de al menos una diferencia indica que el tratamiento dirigido al EGFR es probable que sea efectivo. Preferiblemente, la diferencia de ácido nucleico aumenta la actividad de la quinasa del EGFR. El paciente puede ser tratado entonces con un tratamiento dirigido al EGFR. El tratamiento dirigido al EGFR puede ser un inhibidor de la

tirosina quinasa. Preferiblemente, el inhibidor de la tirosina quinasa es una anilinoquinazolina. La anilinoquinazolina puede ser una anilinoquinazolina sintética. Preferiblemente, la anilinoquinazolina sintética es ya sea gefitinib o erlotinib. El tratamiento dirigido al EGFR puede ser un inhibidor irreversible del EGFR, que incluye al ácido 4-dimetilamino-but-2-enoico [4 - (3-cloro-4-fluorofenilamino)-3-ciano-7-etoxi-quinolin-6-il]amida ("EKB-569", a veces también denominado como "EKI-569", véase, por ejemplo el documento WO/2005/018677 y Torrance et al., *Nature Medicine*, vol. 6, No. 9, septiembre de 2000, p. 1024) y/o HKI-272 o HKI-357 (Wyeth; véase Greenberger et al., *Proc 11avo Simposio NCI EORTC-AACR sobre nuevos medicamentos en la terapia contra el cáncer*, *Clinical Cancer Res.* Vol. 6, Suplemento, noviembre de 2000, ISSN 1078 - 0432; en Rabindran et al., *Cancer Res.* 64: 3958 - 3965 (2004); Holbro y Hynes, *Ann. Rev. Pharm. Tox.* 44: 195 - 217 (2004); Tsou et al., *J. Med. Chem.* 2005, 48, 1107 - 1131, y Tejpar et al., *J. Clin. Oncol. ASCO Annual Meeting Proc.* Vol., 22, No. 14S: 3579 (2004)).

El EGFR se puede obtener de una muestra biológica de un paciente con o en riesgo de desarrollar cáncer. La variación en el dominio quinasa del EGFR (o el gen *erbB1*) efectúa la estructura conformacional del bolsillo de enlazamiento del ATP. Preferiblemente, la variación en el dominio quinasa del EGFR es una supresión en el marco o una sustitución en el exón 18, 19, 20 o 21.

La supresión en el marco puede ser en el exón 19 del EGFR (*erbB1*). La supresión en el marco en el exón 19 comprende preferiblemente una supresión de al menos los aminoácidos leucina, arginina, ácido glutámico y alanina, en los codones 747 748, 749, y 750. La supresión en el marco puede comprender los nucleótidos 2235 a 2249 y elimina los aminoácidos 746 a 750 (la secuencia de ácido glutámico, leucina, arginina, ácido glutámico, y alanina), véase la Tabla 2, la Tabla S2, la Figura 2B, la Figura 4A, la Figura 5, la SEQ ID No: 511, la Figura 6C y la Figura 8C. En otra forma de realización, la supresión en el marco comprende los nucleótidos 2236 a 2250 y elimina los aminoácidos 746 a 750, véase la Tabla S2, la Figura 5, la SEQ ID No: 511, y la Figura 6C. Alternativamente, la supresión en el marco comprende los nucleótidos 2240 a 2251, véase la Tabla 2, la Figura 2C, la Figura 4A, la Figura 5, la SEQ ID No: 511, o los nucleótidos 2240 a 2257, véase la Tabla 2, la Tabla S3A, la Figura 2A, la Figura 4A, la Figura 5, la SEQ ID No: 511, la Figura 6C, y la Figura 8E. Alternativamente, la supresión en el marco comprende los nucleótidos 2239 a 2247 junto con una sustitución de citosina por guanina en el nucleótido 2248, véase la Tabla S3A y la Figura 8D, o una supresión de los nucleótidos 2238 a 2255 junto con una sustitución de timina por adenina en el nucleótido 2237, véase la Tabla S3A y la Figura 8F, o una supresión de los nucleótidos 2254 a 2277, véase la Tabla S2. Alternativamente, la supresión en el marco comprende los nucleótidos 2239 - 2250 de 1TTAAGAGAAGCA; 2251A> C, o 2240 - 2254 de 1TAAGAGAAGCA, o 2257 - 2271 de 1CCGAAAGCCAACAAG, como se muestra en la Tabla S3B.

La sustitución puede ser en el exón 21 del EGFR. La sustitución en el exón 21 comprende al menos un aminoácido. La sustitución en el exón 21 puede comprender una sustitución de una guanina por una timina en el nucleótido 2573, véase la Figura 4A y la Figura 5, la SEQ ID No: 511. Esta sustitución da como resultado una sustitución de aminoácidos, donde la Leucina de tipo silvestre se reemplaza con una arginina en el aminoácido 858, véase la Figura 5, la Tabla 2, la Tabla S2, la Tabla S2A, la Figura 2D, la Figura 6A, la figura 8B, y la SEQ ID No: 512. Alternativamente, la sustitución en el exón 21 comprende una sustitución de una adenina por una timina en el nucleótido 2582, véase la Figura 4A y la Figura 5, la SEQ ID No: 511. Esta sustitución da como resultado una sustitución de aminoácidos, donde la Leucina de tipo silvestre se reemplaza con una glutamina en el aminoácido 861, véase la Figura 5, la Tabla 2, la Figura 2E, la Tabla S3B, y la SEQ ID NO: 512.

La sustitución también puede ser en el exón 18 del EGFR. La sustitución puede ser en el exón 18 y puede ser una timina por una guanina en el nucleótido 2155, véase la Figura 4A y la Figura 5, la SEQ ID No: 511. Esta sustitución da como resultado una sustitución de aminoácidos, en donde la glicina de tipo silvestre se sustituye con una cisteína en el codón 719, véase la Figura 5, la SEQ ID No: 512. La sustitución en el exón 18 puede ser una adenina por una guanina en el nucleótido 2155 que resulta en una sustitución de aminoácidos, en donde la glicina de tipo silvestre se sustituye por una serina en el codón 719, véase la Tabla S2, la figura 6B, la figura 8A, la figura 5, la SEQ ID No: 511 y 512.

La sustitución puede ser una inserción de guanina, guanina y timina (GGT) después del nucleótido 2316 y antes del nucleótido 2317 de la SEQ ID No: 511 (236_2317 ins GGT). Esto también puede ser descrito como una inserción de valina (V) en el aminoácido 772 (P772_H733 INSV). Otras mutaciones se muestran en la Tabla S3B e incluyen, por ejemplo, y la inserción de CAACCCGG después del nucleótido 2309 y antes del nucleótido 2310 de la SEQ ID No: 511 y una inserción de GCGTGGACA después del nucleótido 2311 y antes del nucleótido 2312 de la SEQ ID No: 511. La sustitución también puede ser en el exón 20 y en una forma de realización es una sustitución de AA por GG en los nucleótidos 2334 y 2335, véase la Tabla S3B.

En resumen, en formas de realización preferidas, la variación de ácidos nucleicos del gen *erbB1* es una sustitución de una timina por una guanina o una adenina por una guanina en el nucleótido 2155 de la SEQ ID No: 511, una supresión de los nucleótidos 2235 a 2249, 2240 a 2251, 2240 a 2257, 2236 a 2250, 2254 a 2277, o 2236 a 2244 de la SEQ ID No: 511, una inserción de los nucleótidos guanina, guanina y timina (GGT) después del nucleótido 2316 y antes del nucleótido 2317 de la SEQ ID No: 511, y una sustitución de una guanina por una timina en el nucleótido 2573 o una adenina por una timina en el nucleótido 2582 de la SEQ ID No: 511.

La detección de la presencia o ausencia de al menos una variación de ácido nucleico se puede determinar mediante

la amplificación de un segmento de ácido nucleico que codifica al receptor. El segmento a amplificar es de 1000 nucleótidos de longitud, preferiblemente, 500 nucleótidos de longitud, y lo más preferiblemente 100 nucleótidos de longitud o menos. El segmento que se amplifica puede incluir una pluralidad de variaciones.

5 La detección de la presencia o ausencia de al menos una variación puede permitir el contacto del ácido nucleico del EGFR que contiene un sitio de variación con al menos una sonda de ácido nucleico. La sonda preferentemente hibrida con una secuencia de ácido nucleico que incluye un sitio de variación y que contiene bases de nucleótidos complementarios en el sitio de la variación en condiciones de hibridación selectiva. La hibridación se puede detectar con un marcador detectable.

10 La detección de la presencia o ausencia de al menos una variación puede comprender la secuenciación de al menos una secuencia de ácido nucleico y la comparación de la secuencia obtenida con la secuencia conocida del ácido nucleico erbB1. Alternativamente, la presencia o ausencia de al menos una variación comprende la determinación espectrométrica de masa de al menos una secuencia de ácido nucleico.

15 La detección de la presencia o ausencia de al menos una variación del ácido nucleico puede comprender realizar una reacción en cadena de la polimerasa (PCR). La secuencia del ácido nucleico erbB1 que contiene la variación hipotética se amplifica y se determina la secuencia de nucleótidos del ácido nucleico amplificado. La determinación de la secuencia de nucleótidos del ácido nucleico amplificado comprende la secuenciación de al menos un segmento de ácido nucleico. Alternativamente, los productos de amplificación se pueden analizar mediante el uso de cualquier método capaz de separar los productos de amplificación según su tamaño, incluyendo la electroforesis en gel automatizada y manual y similares.

20 Alternativamente, la detección de la presencia o ausencia de al menos una variación comprende la determinación del haplotipo de una pluralidad de variaciones en un gen.

25 En otra forma de realización, la presencia o ausencia de una variación del EGFR puede ser detectada mediante el análisis del producto del gen erbB1 (proteína). En esta forma de realización, se utiliza una sonda que se enlaza específicamente a una variante del EGFR. En una forma de realización preferida, la sonda es un anticuerpo que se enlaza preferentemente a una variante del EGFR. La presencia de una variante del EGFR predice la probabilidad de la eficacia de un tratamiento dirigido al EGFR. Alternativamente, la sonda puede ser un fragmento de anticuerpo, anticuerpo quimérico, anticuerpo humanizado o un aptámero.

30 La presente invención proporciona además una sonda que se enlaza específicamente en condiciones de enlazamiento selectivo a una secuencia de ácido nucleico que comprende al menos una variación del ácido nucleico en el gen EGFR (erbB1). En una forma de realización, la variación es una mutación en el dominio quinasa de erbB1 que confiere un cambio estructural en el bolsillo de enlazamiento de ATP.

35 La sonda de la presente invención puede comprender una secuencia de ácido nucleico de aproximadamente 500 bases de nucleótidos, preferiblemente de aproximadamente 100 bases de nucleótidos, y más preferiblemente de aproximadamente 50 o aproximadamente 25 bases de nucleótidos o menos de longitud. La sonda puede estar compuesta de ADN, ARN, o ácido nucleico peptídico (PNA). Además, la sonda puede contener un marcador detectable, tal como, por ejemplo, un marcador fluorescente o enzimático.

40 La presente invención proporciona adicionalmente un nuevo método para determinar la probabilidad de la efectividad de un tratamiento dirigido al receptor del factor de crecimiento epidérmico (EGFR) en un paciente afectado con cáncer. El método comprende la determinación de la actividad de la quinasa del EGFR en una muestra biológica de un paciente. Un aumento en la actividad de la quinasa después de la estimulación con un ligando del EGFR, en comparación con un control normal, indica que el tratamiento dirigido al EGFR es probable que sea efectivo.

45 La presente invención proporciona además un método novedoso para el tratamiento de un paciente afectado con o en riesgo de desarrollar cáncer. El método implica la determinación de si el dominio quinasa del EGFR de un paciente contiene al menos una variación del ácido nucleico. Preferiblemente, el EGFR se encuentra en el sitio del tumor o cáncer y la variación del ácido nucleico es somática. La presencia de tal variación indica que un tratamiento dirigido al EGFR será efectivo. Si la variación está presente, se administra el inhibidor de la tirosina quinasa al paciente.

50 Como anteriormente, el inhibidor de la tirosina quinasa administrado a un paciente identificado puede ser una anilinoquinazolina o un inhibidor de tirosina quinasa irreversible, tal como, por ejemplo, EKB-569, HKI-272 y/o HKI-357 (Wyeth). Preferiblemente, la anilinoquinazolina es una anilinoquinazolina sintética y lo más preferiblemente la anilinoquinazolina sintética es gefitinib y erlotinib.

55 El cáncer a ser tratado por los métodos de la presente invención incluye, por ejemplo, pero no se limitan a, cáncer gastrointestinal, cáncer de próstata, cáncer de ovario, cáncer de mama, cáncer de cabeza y cuello, cáncer de pulmón, cáncer de pulmón de células no pequeñas, cáncer del sistema nervioso, cáncer de riñón, cáncer de retina, cáncer de piel, cáncer de hígado, cáncer de páncreas, cáncer genitourinario y cáncer de vejiga. En una forma de realización preferida, el cáncer es cáncer de pulmón de células no pequeñas.

También está comprendido un kit para la implementación de los métodos de PCR de la presente invención. El kit incluye al menos un par de cebadores degenerados diseñados para hibridar con regiones de ácido nucleico que bordean los genes que codifican para el bolsillo de enlazamiento de ATP del dominio quinasa del EGFR. Además, el kit contiene los productos y los reactivos necesarios para llevar a cabo la amplificación por PCR, y las instrucciones.

- 5 En una forma de realización preferida, los pares de cebadores contenidos en el kit se seleccionan entre el grupo que consiste de la SEQ ID N o: 505, la SEQ ID No: 506, la SEQ ID No: 507, y la SEQ ID No: 508. También se prefieren los cebadores enumerados en la Tabla 6 y 7 en los ejemplos.

- En aún otra forma de realización, la presente invención da a conocer un método para seleccionar un compuesto que inhibe la actividad catalítica de quinasa de una variante del receptor del factor de crecimiento epidérmico (EGFR).
10 Como una primera etapa, una variante del EGFR se pone en contacto con un compuesto potencial. La actividad de la quinasa resultante de la variante del EGFR se detecta a continuación, y se selecciona un compuesto que inhibe la actividad de la quinasa de la variante del EGFR. En una forma de realización, la variante del EGFR está contenida dentro de una célula. El método también se puede utilizar para seleccionar un compuesto que inhibe la actividad de la quinasa de la variante del EGFR que tiene una mutación secundaria en el dominio quinasa que confiere
15 resistencia a un TKI, por ejemplo, gefitinib o erlotinib.

En una forma de realización, se marca la variante del EGFR. En otra forma de realización, el EGFR se enlaza a un soporte sólido. En una forma de realización preferida, el soporte sólido es un chip de proteína.

- En aún otra forma de realización de la presente invención se divulga una composición farmacéutica que inhibe la actividad catalítica de la quinasa de una variante del receptor del factor de crecimiento epidérmico (EGFR). El
20 compuesto que inhibe la actividad catalítica de la quinasa de una variante del EGFR se selecciona del grupo que consiste de un anticuerpo, un fragmento de anticuerpo, una molécula pequeña, un péptido, una proteína, un ácido nucleico antisentido, una ribozima, un ANP, un ARNpi, un aptámero de oligonucleótido, y el aptámero peptídico.

- También se da a conocer un método para tratar un paciente que tiene una enfermedad mediada por EGFR. De acuerdo con el método, se administra al paciente la composición farmacéutica que inhibe la actividad catalítica de la
25 quinasa de una variante del receptor del factor de crecimiento epidérmico (EGFR).

- En una forma de realización, la enfermedad mediada por EGFR es cáncer. En una forma de realización preferida, el cáncer es de origen epitelial. Por ejemplo, el cáncer es cáncer gastrointestinal, cáncer de próstata, cáncer de ovario, cáncer de mama, cáncer de cabeza y cuello, cáncer de pulmón, cáncer de pulmón de células no pequeñas, cáncer
30 del sistema nervioso, cáncer de riñón, cáncer de retina, cáncer de piel, cáncer de hígado, cáncer pancreático, cáncer genitourinario y cáncer de vejiga. En una forma de realización preferida, el cáncer es cáncer de pulmón de células no pequeñas.

- En otra forma de realización, se divulga un método para predecir la adquisición de mutaciones secundarias (o la selección de mutaciones) en el dominio quinasa del gen erbB1. Se pone en contacto una célula que expresa una forma variante del gen erbB1 con una dosis efectiva incluso sub-letal de un inhibidor de la tirosina quinasa. Se
35 seleccionan las células que son resistentes a un efecto de la detención del crecimiento del inhibidor de la tirosina quinasa y se analiza el ácido nucleico de erbB1 por la presencia de mutaciones adicionales en el dominio quinasa de erbB1. En una forma de realización, la célula está in vitro. En otra forma de realización, se obtiene la célula de un animal transgénico. En una forma de realización, el animal transgénico es un ratón. En este modelo de ratón, las células a ser estudiadas se obtienen a partir de una biopsia del tumor. Las células que contienen una mutación
40 secundaria en el dominio quinasa de erbB1 seleccionadas por la presente invención se pueden utilizar en los métodos anteriores para seleccionar un compuesto que inhibe la actividad de la quinasa de la variante del EGFR que tiene una mutación secundaria en el dominio quinasa.

- En una forma de realización alternativa para la predicción de la adquisición de mutaciones secundarias en el dominio quinasa del gen erbB1, se ponen primero en contacto células que expresan una forma variante del gen erbB1 con
45 una cantidad efectiva de un agente mutagénico. La mutagénesis es, por ejemplo, metanosulfonato de etilo (EMS), N-etil-N-nitrosourea (ENU), N-metil-N-nitrosourea (MNU), clorhidrato de focarbaxina (Prc), metanosulfonato de metilo (MeMS), clorambucil (Chl), melfalán, clorhidrato de porcarbaxina, ciclofosfamida (Cp), sulfato de dietilo (Et₂SO₄), monómero de acrilamida (AA), trietilen-melamina (TEM), mostaza nitrogenada, vincristina, dimetilnitrosamina, N-metil-N'-nitro-nitrosoguanidina (MNNG), 7,12 dimetilbenz(a)antraceno (DMBA), óxido de etileno,
50 hexametilfosforamida, bisulfán, o etil metanosulfonato (ETMs). Se pone luego la célula en contacto con una dosis efectiva incluso sub-letal de un inhibidor de la tirosina quinasa. Se seleccionan las células que son resistentes a un efecto de la detención del crecimiento del inhibidor de la tirosina quinasa y se analiza el ácido nucleico de erbB1 por la presencia de mutaciones adicionales en el dominio quinasa erbB1.

Breve descripción de las figuras

- 55 Las figuras 1A-1B muestran una ilustración representativa de la respuesta al gefitinib en un cáncer difícil de pulmón de células no pequeñas (NSCLC). Tomografía computarizada de tórax del caso 6 (Tabla 1), que demuestra (Figura 1A) una gran masa en el pulmón derecho antes del tratamiento con gefitinib y (Figura 1B) una marcada mejoría seis semanas después de iniciado el tratamiento con Gefitinib.

La Figura 2 muestra las mutaciones del EGFR en tumores sensibles al gefitinib.

Las Figuras 2A-2C muestran la secuencia de nucleótidos del gen *EGFR* en muestras de tumores con supresiones en el marco de heterocigotos dentro del dominio quinasa (picos dobles) (SEQ ID Nos: 643 a 654, respectivamente, en orden de aparición). Se muestran los seguimientos en ambas direcciones tanto sentido como antisentido para demostrar los dos puntos de rompimiento de la supresión; se muestra la secuencia de nucleótidos de tipo silvestre en letras mayúsculas, y la secuencia mutante en letras minúsculas. El punto de rompimiento 5' de la mutación del L747-T751insS está precedida por una sustitución de T por C que no altera el aminoácido codificado.

La Figura 2D y la Figura 2E muestran mutaciones heterocigotas en sentido erróneo (flechas) que resultan en sustituciones de aminoácidos dentro del dominio de la tirosina quinasa (SEQ ID Nos: 656 y 658). Los picos dobles representan dos nucleótidos en el sitio de mutaciones heterocigotas. Para la comparación, también se muestra la secuencia de tipo silvestre correspondiente (SEQ ID Nos: 655 y 657).

La figura 2F es una representación esquemática de moléculas del EGFR dimerizado enlazadas por medio del ligando EGF. Se resaltan el dominio extracelular (que contiene dos dominios del ligando del receptor [L] y un dominio como el de la furina), la región transmembrana y el dominio citoplasmático (que contiene el dominio catalítico de la quinasa). Se indica la posición de la tirosina¹⁰⁶⁸ (Y-1068), un sitio de autofosforilación utilizado como un marcador de la activación del receptor, junto con los efectores secuencia abajo activados por la autofosforilación del EGFR (STAT3, MAP quinasa (MAPK), y AKT). Se muestra la ubicación de las mutaciones asociadas a tumores, todos dentro del dominio de la tirosina quinasa.

La Figura 3 demuestra la activación mejorada que depende de EGF del EGFR mutante y aumento de la sensibilidad del EGFR mutante al gefitinib.

La figura 3A muestra una evolución en el tiempo de la activación inducida por ligandos de los mutantes del L747-P753insS y L858R, en comparación con el EGFR de tipo silvestre, después de la adición de EGF a las células privadas de suero. Se utiliza la autofosforilación del EGFR como un marcador de la activación del receptor, utilizando una transferencia tipo Western con un anticuerpo que reconoce específicamente el residuo de tirosina¹⁰⁶⁸ fosforilada del EGFR (panel de la izquierda), en comparación con los niveles totales del EGFR expresado en células Cos-7 (control; panel de la derecha). Se mide la autofosforilación del EGFR en intervalos después de la adición de EGF (10 ng/ml).

La Figura 3B es una representación gráfica de la inducción de EGF de la fosforilación del receptor de tipo silvestre y mutante (véase el panel A). Se cuantificaron las autorradiografías de tres experimentos independientes usando el software de imágenes NIH; se normalizó la intensidad de la fosforilación del EGFR con la expresión de la proteína total, y se muestra como la activación del receptor en porcentaje, con una desviación estándar.

La figura 3C muestra una inhibición dependiente de la dosis de la activación del EGFR por gefitinib. La autofosforilación de la tirosina¹⁰⁶⁸ del EGFR se demuestra por medio de análisis de transferencias tipo Western de células Cos-7 que expresan receptores de tipo silvestre o mutante, y se estimularon con 100 ng/ml de EGF durante 30 min. Las células se dejaron sin tratar (U) o se las trató previamente durante 3 horas con concentraciones crecientes de gefitinib como se muestra (panel de la izquierda). Las cantidades totales de proteína del EGFR expresadas se muestran como control (panel de la derecha).

La Figura 3D muestra la cuantificación de los resultados de dos experimentos descritos para el panel 3C (software de imágenes NIH). Las concentraciones del EGFR fosforilado se normalizaron a los niveles de expresión de proteína y se expresaron como la activación del receptor en porcentaje.

La Figura 4 demuestra la agrupación de mutaciones en los sitios críticos dentro del bolsillo de enlazamiento de ATP del EGFR.

La Figura 4A muestra la posición de la superposición de supresiones en el marco en el exón 19 y mutaciones en sentido erróneo en el exón 21 del gen *EGFR*, en múltiples casos de NSCLC (SEQ ID Nos: 495 a 504 (ADN)). Se muestra la secuencia parcial de nucleótidos para cada exón, con supresiones marcadas por líneas punteadas y las mutaciones de sentido erróneo resaltadas y subrayadas; se muestran las secuencias de nucleótidos y aminoácidos del EGFR de tipo silvestre (SEQ ID Nos: 493 y 494 (ADN) y 509-510 (amino ácido)).

La Figura 4B muestra la estructura tridimensional de la hendidura de ATP del EGFR flanqueada por los lóbulos amino (N) y carboxilo (C) del dominio quinasa (coordenadas derivadas de PDB 1M14, y mostradas mediante el uso de software Cn3D). El inhibidor, que representa Gefitinib, se describe ocupando la hendidura de ATP. Se muestran las ubicaciones de las dos mutaciones de sentido erróneo, dentro del bucle de activación de la quinasa; las tres supresiones en el marco están todas presentes dentro de otro bucle, que flanquea la hendidura de ATP.

La Figura 4C es un primer plano del dominio quinasa del EGFR, que muestra los residuos de aminoácidos críticos implicados en el enlazamiento ya sea a ATP o al inhibidor. Específicamente, los compuestos de 4-anilinoquinazolina tales como gefitinib inhiben la catálisis ocupando el sitio de enlazamiento de ATP, donde forman enlaces de hidrógeno con residuos de metionina⁷⁹³ (M793) y cisteína⁷⁷⁵ (C775), mientras que su anillo de anilina está cerca de

residuos de metionina⁷⁶⁶ (M766), lisina⁷⁴⁵ (K745) y leucine⁷⁸⁸ (L788). Se predice que las supresiones en el marco dentro del bucle que están dirigidas por mutaciones alteran la posición de estos aminoácidos en relación con el inhibidor. Se muestran residuos mutados en el bucle de activación de la tirosina quinasa.

La Figura 5 muestra la secuencia de nucleótidos y de aminoácidos del gen erbB1. Los aminoácidos se representan como letras individuales, que son conocidas por aquellos capacitados en la técnica. Las variaciones de nucleótidos en el dominio quinasa se destacan por medio del número del paciente, véase la Tabla 2. La SEQ ID No: 511 incluye los nucleótidos 1 a 3633. La SEQ ID No: 512 incluye los aminoácidos 1 a 1210.

Figuras 6A - 6C: Alineación de secuencias de regiones seleccionadas dentro de los dominios quinasa del EGFR y de B-Raf. Representación de las mutaciones del EGFR en NSCLC humano. Las mutaciones del EGFR (gb: X00588) en los tumores de NSCLC se resaltan en gris. Las mutaciones de B-Raf (gb: M95712) en múltiples tipos de tumores (5) se resaltan en negro. Los asteriscos indican los residuos conservados entre EGFR y B-Raf. La Figura 6A representa mutaciones del L858R en el bucle de activación (SEQ ID Nos 477 - 479). La Figura 6B representa al mutante G719S en el bucle de P (SEQ ID Nos 480 - 482). La Figura 6C representa mutantes de supresión del exón 19 del EGFR (SEQ ID Nos 483 - 489).

Figura 7: Posiciones de las mutaciones de sentido erróneo G719S y L858R y la supresión Del-1 en la estructura tridimensional del dominio quinasa del EGFR. El bucle de activación se muestra en amarillo, el bucle P es de color azul y el lóbulo C y el lóbulo N son como se indica. Los residuos dirigidos por mutación o supresión se resaltan en rojo. La mutación Del-1 dirige los residuos ELREA en los codones 746 a 750. Las mutaciones se encuentran en regiones altamente conservadas dentro de quinasas y se encuentran en bucle p y en el bucle de activación, que rodean la región donde se predice que ATP y también gefitinib y erlotinib se enlazan.

Figuras 8A-8F. Cromatogramas representativos de ADN del EGFR a partir de tejido normal y de los tejidos tumorales. Las ubicaciones de las mutaciones identificadas son las siguientes. La Figura 8A representa el exón 18 del bucle P del dominio quinasa (SEQ ID Nos. 659 - 660). La Figura 8B representa el exón 21 del bucle A del dominio quinasa (SEQ ID Nos. 661 - 662). La Figura 8C representa el exón 19 de Del-1 del dominio quinasa (SEQ ID Nos. 663 - 665). La Figura 8D representa el exón 19 de Del-3 del dominio quinasa (SEQ ID Nos. 666 - 668). La Figura 8E representa el exón 19 de Del-4 del dominio quinasa (SEQ ID No. 669 - 671). La Figura 8F representa el exón 19 de Del-5 del dominio quinasa (SEQ ID No. 672 - 674).

Figura 9: La alineación de la secuencia de los polipéptidos BCR-ABL y EGFR y la ubicación de los residuos que confieren un fenotipo resistente a fármacos. El polipéptido EGFR (SEQ ID No: 492) codificado por la secuencia de nucleótidos descrita en el GenBank, número de acceso NM005228 y el polipéptido BCR-ABL (SEQ ID No: 491) codificado por la secuencia de nucleótidos descrita en el GenBank número de acceso M14752 son residuos alineados y conservados, se presentan sombreados. Las mutaciones de BCR-ABL que confieren resistencia al inhibidor de tirosina quinasa imatinib (STI571, Glivec/Gleevec) se indican mediante asteriscos.

La Figura 10 muestra el proceso de toma de decisión para un paciente con NSCLC metastásico en fase de pruebas del EGFR.

La Figura 11 muestra un diagrama de los exones 18-24 del EGFR (no a escala). Las flechas describen la ubicación de las mutaciones identificadas. Los asteriscos denotan el número de pacientes con mutaciones en cada ubicación. El diagrama ampliado representa la superposición de las supresiones del exón 19, y el número de pacientes (n) con cada supresión. Obsérvese que estos resultados no son inclusivos de todas las mutaciones del EGFR hasta la fecha.

Descripción detallada

La presente invención proporciona un nuevo método para determinar la probabilidad de la eficacia de un tratamiento dirigido al receptor del factor de crecimiento epidérmico (EGFR) en un paciente afectado con cáncer. El método comprende la detección de la presencia o ausencia de al menos una variación del ácido nucleico en el dominio quinasa del gen erbB de dicho paciente. La presencia de al menos una variación indica que tratamiento dirigido al EGFR probablemente sea efectivo. Preferiblemente, la variación de ácido nucleico aumenta la actividad quinasa del EGFR. El paciente puede ser tratado luego con un tratamiento dirigido al EGFR. En una forma de realización de la presente invención, el tratamiento dirigido al EGFR es un inhibidor de tirosina quinasa. En una forma de realización preferida, el inhibidor de tirosina quinasa es una anilinoquinazolina. El anilinoquinazolina puede ser una anilinoquinazolina sintética. Preferiblemente, la anilinoquinazolina sintético es o bien gefitinib o erlotinib.

Definiciones:

Los términos "ErbB1", "receptor del factor de crecimiento epidérmico", y "EGFR" se utilizan aquí de forma intercambiable y se refieren al EGFR de la secuencia nativa divulgada, por ejemplo, en Carpenter et al. Ann. Rev. Biochem. 56: 881-914 (1987), que incluye variantes del mismo (por ejemplo, un EGFR mutante por supresión como en Humphrey et al. PNAS (EUA) 87: 4.207-4.211 (1990)). ErbB1 se refiere al gen que codifica al producto de la proteína del EGFR.

El término "variación de ácidos nucleicos que incrementa la actividad de la quinasa" como se utiliza aquí se refiere a una variación (es decir, una mutación) en la secuencia de nucleótidos de un gen que se traduce en un aumento de la actividad de la quinasa. El aumento de la actividad de la quinasa es un resultado directo de la variación en el ácido nucleico y está asociado con la proteína codificada por el gen.

- 5 El término "fármaco" o "compuesto" como se utiliza aquí se refiere a una entidad química o producto biológico, o combinación de entidades químicas o de productos biológicos, administrado a una persona para tratar o prevenir o controlar una enfermedad o condición. La entidad química o producto biológico es preferiblemente, pero no necesariamente, un compuesto de bajo peso molecular, sino que también puede ser un compuesto más grande, por ejemplo, un oligómero de ácidos nucleicos de aminoácidos o de carbohidratos que incluye sin limitación proteínas, oligonucleótidos, ribozimas, ADNzimas, glicoproteínas, los ARNpi, lipoproteínas, aptámeros, y modificaciones y combinaciones de los mismos.

El término "genotipo" en el contexto de esta invención se refiere a la forma alélica particular de un gen, que se puede definir por medio del (de los) nucleótido(s) particular(es) presente(s) en una secuencia de ácido nucleico en un (unos) sitio(s) particular(es).

- 15 Los términos "forma variante de un gen", "forma de un gen", o "alelo" se refieren a una forma específica de un gen en una población, siendo diferente la forma específica de otras formas del mismo gen en la secuencia de al menos un, y frecuentemente más de un, sitios de variantes dentro de la secuencia del gen. Las secuencias en estos sitios de variantes que se diferencian entre los diferentes alelos de un gen se llaman "variaciones de la secuencia del gen" o "variaciones" o "variantes". Otros términos conocidos en la técnica que son equivalentes incluyen mutación y polimorfismo, aunque se utiliza a menudo mutación para referirse a un alelo asociado con un fenotipo perjudicial. En aspectos preferidos de esta invención, las variaciones se seleccionan del grupo que consiste de las variaciones enumeradas en las tablas de variación presentadas aquí.

- 25 En el contexto de esta invención, el término "sonda" se refiere a una molécula que puede distinguir en forma detectable entre moléculas objetivo que difieren en su estructura. La detección puede llevarse a cabo en una variedad de diferentes maneras, dependiendo del tipo de sonda utilizada y del tipo de molécula objetivo. Así, por ejemplo, la detección puede basarse en la discriminación de niveles de actividad de la molécula objetivo, pero preferiblemente se basa en la detección de enlazamientos específicos. Los ejemplos de tales enlazamientos específicos incluyen enlazamiento de anticuerpos e hibridación de sondas de ácido nucleico. Por lo tanto, por ejemplo, las sondas pueden incluir sustratos enzimáticos, anticuerpos y fragmentos de anticuerpos, y preferiblemente sondas de hibridación de ácido nucleico.

- 30 Como se utilizan aquí, los términos "efectivo" y "efectividad" incluyen tanto efectividad farmacológica como seguridad fisiológica. La efectividad farmacológica se refiere a la capacidad del tratamiento que da como resultado un efecto biológico deseable en el paciente. La seguridad fisiológica se refiere al nivel de toxicidad, u otros efectos fisiológicos adversos a nivel celular, de órganos y/o de organismos (a menudo denominados como efectos secundarios) que resultan de la administración del tratamiento. "Menor efectividad" significa que el tratamiento da como resultado un nivel más bajo terapéuticamente de la efectividad farmacológica y/o un mayor nivel terapéutico de efectos fisiológicos adversos.

- 40 El término "cebador", como se utiliza aquí, se refiere a un oligonucleótido que es capaz de actuar como un punto de iniciación de la síntesis de polinucleótidos a lo largo de una hebra complementaria cuando se coloca bajo condiciones en las cuales se cataliza la síntesis de un producto de extensión del cebador que es complementario con un polinucleótido. Tales condiciones incluyen la presencia de cuatro diferentes trifosfatos de nucleótidos o análogos de nucleósidos y uno o más agentes de polimerización tales como ADN polimerasa y/o transcriptasa inversa, en un amortiguador apropiado ("amortiguador" incluye sustituyentes que son cofactores, o que afectan el pH, la fuerza iónica, etc.), y una temperatura adecuada. Un cebador debe ser suficientemente largo para cebar la síntesis de los productos de extensión en presencia de un agente para la polimerasa. Un cebador típico contiene al menos aproximadamente 5 nucleótidos de longitud de una secuencia sustancialmente complementaria a la secuencia objetivo, pero se prefieren cebadores algo más largos. Por lo general, los cebadores contienen aproximadamente 15-26 nucleótidos, pero se pueden emplear también cebadores más largos.

- 50 Un cebador siempre contendrá una secuencia sustancialmente complementaria con la secuencia objetivo, que es la secuencia específica que va a ser amplificada, con la cual puede hibridarse. Un cebador puede, opcionalmente, comprender también una secuencia promotora. El término "secuencia promotora" define una sola hebra de una secuencia de ácido nucleico que es reconocida específicamente por una ARN polimerasa que se enlaza con una secuencia reconocida e inicia el proceso de transcripción por medio del cual se produce un transcrito de ARN. En principio, se puede emplear cualquier secuencia promotora para la cual existe una polimerasa conocida y disponible que es capaz de reconocer la secuencia de iniciación. Los promotores conocidos y útiles son aquellos que son reconocidos por ciertas polimerasas de bacteriófagos, tales como los bacteriófagos T3, T7 o SP6.

Un "microarreglo" es un arreglo lineal o bidimensional de regiones preferiblemente discretas, teniendo cada una un área definida, formada sobre la superficie de un soporte sólido. La densidad de las regiones discretas sobre un microarreglo se determina por medio del número total de polinucleótidos objetivo que van a ser detectados sobre la

superficie de un único soporte en fase sólida, preferiblemente al menos aproximadamente 50/cm², más preferiblemente al menos aproximadamente 100/cm², incluso más preferiblemente al menos aproximadamente 500/cm², y aún más preferiblemente al menos aproximadamente 1.000/cm². Como se utiliza aquí, un microarreglo es un arreglo de cebadores de oligonucleótidos colocados sobre un chip u otras superficies utilizadas para amplificar o clonar polinucleótidos objetivo. Ya que la posición de cada grupo particular de cebadores en el arreglo es conocida, se pueden determinar las identidades de los polinucleótidos objetivo con base en su enlazamiento con una posición particular en el microarreglo.

El término "etiqueta" se refiere a una composición capaz de producir una señal detectable indicativa de la presencia del polinucleótido objetivo en una muestra del ensayo. Los marcadores adecuados incluyen radioisótopos, cromóforos de nucleótidos, enzimas, sustratos, moléculas fluorescentes, fracciones quimioluminiscentes, partículas magnéticas, fracciones bioluminiscentes, y similares. Como tal, una etiqueta es cualquier composición detectable por medios espectroscópicos, fotoquímicos, bioquímicos, inmunoquímicos, eléctricos, ópticos o químicos.

El término "soporte" se refiere a soportes convencionales tales como perlas, partículas, tiras reactivas, fibras, filtros, membranas y soportes de silano o de silicato tales como placas de vidrio.

El término "amplificación" se utiliza en sentido amplio para significar la creación de un producto de amplificación que puede incluir, por ejemplo, moléculas objetivo adicionales, o moléculas tipo objetivo o moléculas complementarias a la molécula objetivo, las cuales se crean en virtud de la presencia de la molécula objetivo en la muestra. En la situación en donde el objetivo es un ácido nucleico, se puede elaborar enzimáticamente un producto de amplificación con ADN o ARN polimerasa o transcriptasas inversas.

Como se utiliza aquí, una "muestra biológica" se refiere a una muestra de tejido o fluido aislada de un individuo, que incluye, pero no se limita a, por ejemplo, sangre, plasma, suero, biopsias de tumor, orina, heces, esputo, líquido cefalorraquídeo, líquido pleural, aspirados de pezón, fluido linfático, secciones externas de la piel, del tracto respiratorio, intestinal, y genitourinario, lágrimas, saliva, leche, células (incluyendo, pero sin limitarse a, células de la sangre), tumores, órganos, y también muestras de constituyentes de cultivo celular in vitro. En una forma de realización preferida, la muestra es de una resección, biopsia broncoscópica, o biopsia con aguja gruesa de un tumor primario o metastásico, o un bloque de células de líquido pleural. Además, se utilizan muestras de aspirado con aguja fina. Las muestras pueden ser ya sea de tejido embebido en parafina o congelado.

El término "anticuerpo" significa que se trata de una proteína de inmunoglobulina que es capaz de enlazar un antígeno. Anticuerpo como se utiliza se entiende que incluye fragmentos de anticuerpo, por ejemplo, F(ab')₂, Fab', Fab, capaces de enlazar al antígeno o un fragmento antigénico de interés. Preferiblemente, el enlazamiento del anticuerpo con el antígeno inhibe la actividad de una forma variante del EGFR.

El término "anticuerpo humanizado" se utiliza aquí para describir moléculas de anticuerpo completas, es decir, compuesta de dos cadenas livianas completas y dos cadenas pesadas completas, así como anticuerpos que consisten únicamente de fragmentos de anticuerpos, por ejemplo, Fab, Fab', F(ab)₂, y Fv, en los que las CDR proceden de una fuente no humana y la porción restante de la molécula de Ig o fragmento de la misma se deriva de un anticuerpo humano, preferiblemente producido a partir de una secuencia de ácido nucleico que codifica un anticuerpo humano.

Los términos "anticuerpo humano" y "anticuerpo humanizado" se utilizan aquí para describir un anticuerpo a partir del cual se derivan todas las porciones de la molécula de anticuerpo de una secuencia de ácido nucleico que codifica un anticuerpo humano. Tales anticuerpos humanos son los más deseables para uso en terapias de anticuerpos, ya que tales anticuerpos podrían provocar poca o ninguna respuesta inmune en el paciente humano.

El término "anticuerpo quimérico" se utiliza aquí para describir una molécula de anticuerpo, así como fragmentos de anticuerpos, como se describió anteriormente en la definición del término "anticuerpo humanizado". El término "anticuerpo quimérico" incluye anticuerpos humanizados. Los anticuerpos quiméricos tienen al menos una porción de una secuencia de aminoácidos de cadena pesada o liviana que se obtiene a partir de una primera especie de mamífero y otra parte de la secuencia de aminoácidos de cadena pesada o liviana que se obtiene a partir de una segunda especie diferente de mamífero.

Preferiblemente, la región variable se deriva de una especie de mamífero no humano y la región constante se deriva de una especie humana. Específicamente, el anticuerpo quimérico se produce preferiblemente a partir de una secuencia de 9 nucleótidos de un mamífero no humano que codifica una región variable y una secuencia de nucleótidos de un humano que codifica una región constante de un anticuerpo.

La Tabla 2 es una lista parcial de variaciones de la secuencia de ADN en el dominio quinasa de erbB1 relevante para los métodos descritos en la presente invención. Estas variaciones fueron identificadas por los inventores en estudios de muestras biológicas de pacientes con NSCLC que respondieron a gefitinib y pacientes sin exposición a gefitinib.

Las moléculas de ácido nucleico se pueden aislar de una muestra biológica particular utilizando cualquiera de una cantidad de procedimientos, que son conocidos en el arte, siendo el procedimiento particular de aislamiento elegido

el apropiado para la muestra biológica particular. Por ejemplo, pueden ser útiles los procedimientos de lisis alcalina y de congelación-descongelación para obtener moléculas de ácido nucleico a partir de materiales sólidos; pueden ser útiles los procedimientos de lisis alcalina y calor para obtener moléculas de ácido nucleico de la orina; y se puede utilizar extracción de la proteinasa K para obtener ácido nucleico de la sangre (Rolff, A et al. PCR: Clinial Diagnostics and Research, Springer (1994)).

Métodos de detección

La determinación de la presencia o ausencia de una variación particular o de una pluralidad de variaciones en el dominio quinasa del gen *erbB1* en un paciente con, o en riesgo de desarrollar cáncer, se puede llevar a cabo en una variedad de formas. Tales pruebas se llevan a cabo comúnmente utilizando ADN o ARN recolectado de muestras biológicas, por ejemplo, biopsias de tejidos, orina, heces, esputo, sangre, células, raspados de tejido, aspirados de mama u otros materiales celulares, y se pueden realizar por medio de una variedad de métodos que incluyen, pero no se limitan a, PCR, hibridación con sondas específicas de alelos, detección de mutación enzimática, escisión química de faltas de correlación, espectrometría de masas o secuenciación de ADN, incluyendo minisequenciación. En formas particulares de realización, se pueden llevar a cabo hibridaciones con sondas específicas de alelos en dos formatos: (1) oligonucleótidos específicos de alelos enlazados a una fase sólida (vidrio, silicio, membranas de nailon) y las muestras marcadas en solución, como en muchas aplicaciones de chips de ADN, o (2) muestra enlazada (a menudo ADN clonado o ADN amplificado por PCR) y oligonucleótidos marcados en solución (ya sea específico del alelo o corto para permitir la secuenciación por medio de hibridación). Los ensayos de diagnóstico pueden involucrar un panel de variaciones, a menudo sobre un soporte sólido, que permite la determinación simultánea de más de una variación.

En otro aspecto, la determinación de la presencia de al menos una actividad de quinasa que incrementa la variación de ácidos nucleicos en el gen *erbB1* puede implicar una prueba de determinación del haplotipo. Los métodos para determinar los haplotipos son conocidos por aquellos capacitados en el arte, como por ejemplo, en el documento WO 00/04194.

Preferiblemente, la determinación de la presencia o ausencia de una actividad de quinasa que incrementa la variación de ácidos nucleicos involucra la determinación de la secuencia del sitio o sitios de la variación por medio de métodos tales como la reacción en cadena de la polimerasa (PCR). Alternativamente, la determinación de la presencia o ausencia de una actividad de quinasa que incrementa la variación de ácidos nucleicos puede abarcar secuenciación o minisequenciación del ADN de terminación de la cadena, hibridación de oligonucleótidos o espectrometría de masas.

Los métodos de la presente invención se pueden utilizar para predecir la probabilidad de la efectividad (o falta de efectividad) de un tratamiento dirigido al EGFR en un paciente afectado con o en riesgo de desarrollar cáncer. Preferiblemente, los cánceres incluyen cáncer de origen epitelial, incluyendo, pero sin limitarse a, cáncer gastrointestinal, cáncer de próstata, cáncer de ovario, cáncer de mama, cáncer de cabeza y cuello, cáncer de pulmón, cáncer de pulmón de células no pequeñas, cáncer del sistema nervioso, cáncer de riñón, cáncer de retina, cáncer de piel, cáncer de hígado, cáncer de páncreas, cáncer genitourinario y cáncer de vejiga. En una forma de realización preferida, el cáncer es cáncer de pulmón de células no pequeñas.

La presente invención se refiere generalmente a la identificación de las variaciones en el dominio quinasa del gen *erbB1* que son indicativas de la eficacia de un tratamiento dirigido al EGFR en un paciente con o en riesgo de desarrollar cáncer. Adicionalmente, la identificación de variaciones específicas en el dominio quinasa del EGFR, en efecto, puede ser utilizada como una prueba de diagnóstico o pronóstico. Por ejemplo, la presencia de al menos una variación en el dominio quinasa de *erbB1* indica que un paciente probablemente se beneficiará del tratamiento con un compuesto dirigido al EGFR, tal como, por ejemplo, un inhibidor de tirosina quinasa.

Los métodos para las pruebas de diagnóstico son bien conocidas en el arte y se divulgan en la solicitud de patente WO 00/04194. En un método de ejemplo, la prueba de diagnóstico comprende la amplificación de un segmento de ADN o ARN (generalmente después de convertir el ARN en ADNc) que abarca una o más variaciones conocidas en el dominio quinasa de la secuencia del gen *erbB1*. Este segmento amplificado es luego secuenciado, y/o ometido a electroforesis en gel de poli(acrilamida) con el fin de identificar las variaciones de ácido nucleico en el segmento amplificado.

PCR

En una forma de realización, la invención proporciona un método de cribado de las variantes en el dominio quinasa del gen *erbB1* en una muestra biológica de ensayo por medio de PCR o, alternativamente, en una reacción en cadena de ligación (LCR) (véase, por ejemplo, Landegran, et al., 1988. Science 241: 1077-1080; y Nakazawa, et al., 1994. Proc. Natl. Acad. Sci. EUA. 91: 360-364), el último de los cuales puede ser particularmente útil para detectar mutaciones puntuales en el gen para EGFR (véase, Abravaya, et al., 1995. Nucl. Acids Res. 23: 675-682). El método comprende las etapas de diseño de cebadores degenerados para amplificar la secuencia objetivo, correspondiendo los cebadores a una o más regiones conservadas del gen, amplificando la reacción con los cebadores utilizando, como molde, un ADN o ADNc obtenido a partir de una muestra biológica de prueba y analizando los productos de

PCR. La comparación de los productos de la PCR de la muestra biológica de prueba con una muestra de control indica las variaciones en la muestra biológica de prueba. El cambio puede ser o bien en ausencia o en presencia de una variación de ácido nucleico en la muestra biológica de prueba.

Los métodos alternativos de amplificación incluyen: replicación de secuencia autosostenida (véase, Guatelli, et al., 1990., Proc. Natl. Acad. Sci. EUA 87: 1874 - 1878), sistema de amplificación transcripcional (véase, Kwok, et al., 1989, Proc. Natl. Acad. Sci. EUA 86: 1173 - 1177); Qb replicasa (véase, Lizardi, et al., 1988. BioTechnology 6: 1197), o cualquier otro método de amplificación de ácido nucleico, seguido por la detección de las moléculas amplificadas usando técnicas bien conocidas por aquellos capacitados en la técnica. Estos esquemas de detección son especialmente útiles para la detección de moléculas de ácido nucleico si tales moléculas están presentes en cantidades muy bajas.

Los cebadores se pueden diseñar utilizando secuencias de aminoácidos de las secuencias de ácido nucleico o proteína del dominio quinasa del gen erbB1 como guía, por ejemplo, la SEQ ID No: 493, la SEQ ID No: 494, la SEQ ID No: 509, y la SEQ ID NO: 510. Los cebadores se diseñan en las regiones homólogas del gen en el que al menos dos regiones de homología están separadas por una región divergente de secuencia variable, siendo la secuencia variable, ya sea en longitud o en la secuencia de ácidos nucleicos.

Por ejemplo, la secuencia idéntica o altamente homóloga, preferiblemente al menos 80% - 85%, más preferiblemente al menos 90 - 99% de la secuencia homóloga de aminoácidos de al menos aproximadamente 6, preferiblemente al menos 8 - 10 aminoácidos consecutivos. Lo más preferiblemente, la secuencia de aminoácidos es 100% idéntica. Los cebadores directos e inversos se diseñan con base en el mantenimiento de la degeneración del codón y la representación de los diferentes aminoácidos en una posición dada entre los miembros conocidos de la familia de genes. El grado de homología como se menciona en el presente documento se basa en el análisis de una secuencia de aminoácidos usando un software de comparación de secuencia estándar, como proteína-BLAST utilizando la configuración predeterminada (<http://www.ncbi.nlm.nih.gov/BLAST/>) .

La Tabla 3 a continuación representa el uso de los códigos de degenerados y sus símbolos estándar:

	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)
	CCTC"	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	AGA Arg (R)
	ATG Met	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"

Preferiblemente se evitan cualquiera de los codones degenerados 6 veces tales como L, R y S ya que en la práctica introducirán una degeneración 6 veces más alta. En el caso de L, TTR y CTN son YTN comprometidos (degeneración de 8 veces), en el caso de R, CGN y AGR se comprometen en MGN (degeneración de 8 veces), y, finalmente, S, TCN y AGY que pueden verse comprometidas con WSN (degeneración de 16 veces). En los tres casos en 6 de ellos coincidirá con la secuencia objetivo. Para evitar esta pérdida de especificidad, es preferible evitar

estas regiones, o para hacer dos poblaciones, cada una con el codón degenerado alternativo, por ejemplo, para S incluyen TCN en una reserva, y AGY en la otra.

Los cebadores pueden diseñarse utilizando una cantidad de programas de ordenador disponibles, incluyendo, pero sin limitarse a Oligo Analyzer 3.0; Oligo Calculator; NetPrimer; Methprimer; Primer3; WebPrimer; PrimerFinder; Primer9; Oligo2002; Pride o GenomePride; Oligos y Codehop. La información detallada acerca de estos programas se puede obtener, por ejemplo, de www.molbiol.net.

Los cebadores se pueden marcar usando las etiquetas conocidas por una persona capacitada en la técnica. Tales marcadores incluyen, pero no se limitan a marcadores radioactivos, fluorescentes, colorantes, y enzimáticos.

El análisis de los productos de amplificación se puede realizar utilizando cualquier método capaz de separar los productos de amplificación de acuerdo con su tamaño, incluyendo electroforesis automatizada y manual en gel, espectrometría de masas, y similares.

Alternativamente, los productos de amplificación se pueden separar utilizando las diferencias de secuencia, utilizando SSCP, DGGE, TGGE, escisión química o polimorfismos de fragmentos de restricción, así como la hibridación, por ejemplo, con matrices de ácidos nucleicos.

Los métodos de aislamiento de ácidos nucleicos, amplificación y análisis son de rutina para una persona capacitada en la técnica y se pueden encontrar ejemplos de protocolos, por ejemplo, en el Molecular Cloning: A Laboratory Manual (conjunto de 3 volúmenes) Ed. Joseph Sambrook, David W. Russell y Joe Sambrook, Cold Spring Harbor Laboratory, tercera edición (Enero 15, 2001), ISBN: 0879695773. Una fuente de protocolos particularmente útil para los métodos utilizados en la amplificación por PCR es PCR (Basics: From Background to Bench) por M. J. McPherson, S. G. Müller, R. Beynon, C. Howe, Springer Verlag, primera edición (octubre 15, 2000), ISBN: 0387916008.

Preferiblemente, se amplifican los exones 19 y 21 del EGFR humano por medio de la reacción en cadena de la polimerasa (PCR) utilizando los siguientes cebadores: cebador sentido del Exón 19, 5'-GCAATATCAGCCTTAGGTG CGGCTC-3' (SEQ ID NO: 505); cebador antisentido del Exón 19, 5'-CATAGAAAGTGAACATTTAGGATGTG-3' (SEQ ID NO: 506); cebador sentido del Exón 21, 5'-CTAACGTTTCG CCAGCCATAAGTCC-3' (SEQ ID NO: 507), y cebador antisentido del Exón 21, 5'-GCTGCGAGCTCACCCAG AATGTCTGG-3' (SEQ ID No: 508).

En una realización alternativa, se pueden identificar las mutaciones en un gen *EGFR* de una célula de muestra por medio de las alteraciones en los patrones de escisión de la enzima de restricción. Por ejemplo, se aísla el ADN de la muestra y del control, se amplifica (opcionalmente), se digiere con una o más endonucleasas de restricción, y se determina la longitud de los fragmentos por electroforesis en gel y se comparan. Las diferencias en longitud de los fragmentos entre el ADN de la muestra y del control indican mutaciones en el ADN de la muestra. Por otra parte, el uso de ribozimas específicos de la secuencia (véase, por ejemplo, la patente de los Estados Unidos No. 5.493.531) se puede utilizar para cuantificar la presencia de mutaciones específicas por medio del desarrollo o la pérdida de un sitio de escisión del ribozima.

Otros métodos para detectar mutaciones en el gen *EGFR* incluyen métodos en los cuales se usa protección de los agentes de escisión para detectar bases mal emparejadas en heterodúplex de ARN/ARN o ARN/ADN. Véase, por ejemplo, Myers, et al., 1985. Science 230: 1242. En general, la técnica del arte de "escisión de falta de correlación" comienza proporcionando heterodúplex formados por hibridación (etiquetado) de ARN o ADN que contiene la secuencia del EGFR de tipo silvestre con ARN o ADN potencialmente mutante obtenido a partir de una muestra de tejido. Los dúplex bicatenarios se tratan con un agente que escinde regiones monocatenarias del dúplex tales como aquellas que existirán debido a una falta de correlación de pares de bases entre las cadenas del control y de la muestra. Por ejemplo, se pueden tratar dúplex de ARN/ADN con ARNasa e híbridos de ADN/ADN tratados con nucleasa S1 para digerir enzimáticamente las regiones no coincidentes. En otras formas de realización, se pueden tratar ya sea los dúplex de ADN/ADN o ARN/ADN con hidroxilamina o tetróxido de osmio y con piperidina con el fin de digerir las regiones no coincidentes. Después de la digestión de las regiones no coincidentes, se separa luego el material resultante por tamaño en geles de poliacrilamida desnaturalizantes para determinar el sitio de la mutación. Véase, por ejemplo Cotton, et al., 1988. Proc. Natl. Acad. Sci. EUA 85: 4397; Saleeba, et al., 1992. Methods Enzymol. 2 17: 286 - 295. En una forma de realización, se puede marcar el ADN o ARN de control para la detección.

En aún otra forma de realización, la reacción de escisión de emparejamientos erróneos emplea una o más proteínas que reconocen pares de bases desapareadas en ADN bicatenario (llamadas enzimas de "reparación de emparejamientos erróneos del ADN") en sistemas definidos para detectar y mapear mutaciones puntuales en los ADNc del EGFR obtenidos a partir de muestras de células. Por ejemplo, la enzima mutY de *E. coli* escinde A en emparejamientos erróneos de G/A y la timidina ADN glicosilasa de células HeLa escinde T en emparejamientos erróneos de G/T. Véase, por ejemplo, Hsu, et al., 1994. Carcinogenesis 15: 1657 - 1662. De acuerdo con un ejemplo de una forma de realización, se hibrida una sonda basada en una secuencia mutante del EGFR, por ejemplo, una secuencia del EGFR DEL-1 hasta DEL-5, G719S, G857V, L883S o L858R, con un ADNc u otro producto de ADN de una(s) célula(s) de prueba. Se trata el dúplex con una enzima de reparación de emparejamientos erróneos de ADN, y los productos de escisión, si los hay, se pueden detectar a partir de protocolos de electroforesis o similares. Véase,

por ejemplo, la patente de los Estados Unidos No. 5.459.039.

En otras formas de realizaciones, se usarán las alteraciones en la movilidad electroforética para identificar mutaciones en genes *EGFR*. Por ejemplo, se puede utilizar un polimorfismo de conformación monocatenaria (SSCP) para detectar diferencias en la movilidad electroforética entre ácidos nucleicos mutantes y de tipo silvestre. Véase, por ejemplo, Orita, et al., 1989. Proc. Natl. Acad. Sci. EUA: 86: 2766; Cotton, 1993. Mutat. Res. 285: 125 - 144; Hayashi, 1992. Genet. Anal. Tech. Appl. 9: 73 - 79. Se desnaturalizarán fragmentos de ADN monocatenarios de ácidos nucleicos del *EGFR* del control y de la muestra y se permitirá que se renaturalicen. La estructura secundaria de ácidos nucleicos monocatenarios varía de acuerdo con la secuencia, la alteración resultante en la movilidad electroforética permite la detección incluso de un solo cambio de base. Los fragmentos de ADN pueden marcarse o detectarse con sondas marcadas. La sensibilidad del ensayo puede ser mejorada mediante el uso de ARN (en lugar de ADN), en donde la estructura secundaria es más sensible a un cambio en la secuencia. En una forma de realización, el método objetivo utiliza análisis de heterodúplex para separar moléculas de heterodúplex bicatenarias con base en los cambios en la movilidad electroforética. Véase, por ejemplo, Keen, et al., 1991. Trends Genet. 7: 5.

En aún otra forma de realización, el movimiento de fragmentos mutantes o de tipo silvestre en geles de poliacrilamida que contienen un gradiente de desnaturizante se ensaya usando electroforesis en gel de gradiente desnaturizante (DGGE). Véase, por ejemplo, Myers, et al., 1985. Nature 313: 495. Cuando se utiliza DGGE como el método de análisis, se modificará el ADN para asegurar que no se desnaturaliza completamente, por ejemplo, añadiendo una abrazadera de GC de aproximadamente 40 pb de ADN rico en GC de alto punto de fusión por medio de PCR. En una forma de realización adicional, se usa un gradiente de temperatura en lugar de un gradiente desnaturizante para identificar diferencias en la movilidad del ADN de control y de muestra. Véase, por ejemplo, Rosenbaum y Reissner, 1987. Biophys. Chem., 265: 12753.

Los ejemplos de otras técnicas para detectar mutaciones puntuales incluyen, pero no se limitan a, hibridación selectiva de oligonucleótidos, amplificación selectiva, o extensión selectiva de cebadores. Por ejemplo, se pueden preparar cebadores de oligonucleótidos en los que se coloca la mutación conocida en el centro y luego se hibrida con ADN objetivo bajo condiciones que permiten la hibridación sólo si se encuentra una coincidencia perfecta. Véase, por ejemplo, Saiki, et al., 1986. Nature 324: 163; Saiki, et al., 1989. Proc. Natl. Acad. Sci. EUA. 86: 6230. Tales oligonucleótidos específicos de alelo se hibridan con ADN objetivo amplificado por PCR o una cantidad de mutaciones diferentes cuando los oligonucleótidos están unidos a la membrana de hibridación y se hibridan con el ADN objetivo marcado.

Alternativamente, puede utilizarse la tecnología de amplificación específica de alelo que depende de la amplificación selectiva por PCR conjuntamente con la presente invención. Los oligonucleótidos utilizados como cebadores para la amplificación específica pueden portar la mutación de interés en el centro de la molécula (por lo que la amplificación depende de hibridación diferencial; véase, por ejemplo, Gibbs, et al., 1989., Nucl. Acids Res. 17: 2437 - 2448) o en el terminal 3' del extremo de un cebador donde, en condiciones apropiadas, falta de coincidencia puede prevenir, o reducir la extensión de la polimerasa (véase, por ejemplo, Prossner, 1993. Tibtech. 11: 238). Además, puede ser deseable introducir un nuevo sitio de restricción en la región de la mutación para crear una detección basada en la escisión. Véase, por ejemplo, Gasparini, et al., 1992. Mol. Cell Probes 6: 1. Se prevé que en ciertas formas de realizaciones también puede realizar la amplificación usando Taq ligasa para la amplificación. Véase, por ejemplo, Barany, 1991. Proc. Natl. Acad. Sci. EUA. 88: 189. En tales casos, la ligación se producirá sólo si existe una correspondencia perfecta en el terminal 3' de la secuencia 5', haciendo posible detectar la presencia de una mutación conocida en un sitio específico mediante la búsqueda de la presencia o ausencia de amplificación.

Soporte sólido y sonda

En una forma de realización alternativa, la detección de la presencia o ausencia de la al menos una variación de ácido nucleico que implica poner en contacto una secuencia de ácido nucleico correspondiente a la región deseada del gen *erbB1*, identificado anteriormente, con una sonda. La sonda es capaz de distinguir una forma particular del gen o la presencia o una variación o variaciones particulares, por ejemplo, mediante enlazamiento o hibridación diferencial. Por lo tanto, los ejemplos de sondas incluyen sondas de hibridación de ácido nucleico, sondas de ácido nucleico peptídico, sondas que contienen nucleótidos que también contienen al menos un análogo de nucleótido, y anticuerpos, por ejemplo, anticuerpos monoclonales, y otras sondas como se discute en este documento. Aquellos capacitados en la técnica están familiarizados con la preparación de sondas con especificidades particulares. Aquellos capacitados en la técnica reconocerán que se pueden ajustar una variedad de variables para optimizar la discriminación entre dos formas variantes de un gen, incluyendo cambios en la concentración salina, temperatura, pH y la adición de diferentes compuestos que afectan la afinidad diferencial de GC versus los pares de bases AT, tales como cloruro de tetrametil amonio. (Véase Current Protocols in Molecular Biology por F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, K. Struhl y V. B. Chanda (editores), John Wiley & Sons.)

Por lo tanto, en formas de realización preferidas, la detección de la presencia o ausencia de al menos una variación implica poner en contacto una secuencia de ácido nucleico que incluye al menos un sitio de variación con una sonda, preferiblemente una sonda de ácido nucleico, donde la sonda preferentemente hibrida con una forma de la secuencia de ácido nucleico que contiene una base complementaria en el sitio de la variación en comparación con la hibridación con una forma de la secuencia de ácido nucleico que tiene una base no complementaria en el sitio de la

variación, donde se lleva a cabo la hibridación en condiciones de hibridación selectiva. Tal sonda de hibridación de ácido nucleico puede abarcar dos o más sitios de variación. A menos que se especifique otra cosa, una sonda de ácido nucleico puede incluir uno o más análogos de ácido nucleico, etiquetas u otros sustituyentes o fracciones siempre que se conserva la función de apareamiento de bases.

- 5 La sonda se puede diseñar para enlazarse, por ejemplo, al menos con tres nucleótidos continuos en ambos lados de la región suprimida de la SEQ ID No: 495, la SEQ ID NO: 497, o la SEQ ID NO: 499. Tales sondas, cuando se hibrida en las condiciones apropiadas, se enlazarán con la forma variante del EGFR, pero no se enlazará con el EGFR de tipo silvestre.

- 10 Tales sondas de hibridación son bien conocidas en la técnica (véase, por ejemplo, Sambrook et al., Eds, (la edición más reciente), Molecular Cloning: A Laboratory Manual, (tercera edición, 2001), vol. 1 - 3, Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York). Las condiciones de hibridación rigurosa incluirán típicamente concentraciones de sal de menos de aproximadamente 1 M, más usualmente aproximadamente menos de 500 mM y preferiblemente aproximadamente menos de 200 mM. Las temperaturas de hibridación pueden ser tan bajas como 5° C, pero son típicamente mayores a 22° C, más típicamente mayores aproximadamente a 30° C, y preferiblemente por encima de aproximadamente 37° C. Los fragmentos más largos pueden requerir temperaturas de hibridación mayores para hibridación específica. Otros factores pueden afectar la rigurosidad de la hibridación, incluyendo la composición de bases y la longitud de las cadenas complementarias, la presencia de disolventes orgánicos y el grado de falta de correlación de las bases; la combinación de parámetros usada es más importante que la medida absoluta de cualquiera de ellos solo. Otras condiciones de hibridación que se pueden controlar incluyen el tipo y la
- 20 concentración del amortiguador, el pH de la solución, la presencia y concentración de los reactivos de bloqueo (por ejemplo, las secuencias de repetición, ADN CotI, el bloqueo de las soluciones de proteína) para disminuir el enlazamiento de fondo, el(los) tipo(s) de detergente(s) y las concentraciones, moléculas tales como polímeros que aumentan la concentración relativa de los polinucleótidos, ión(es) metálicos(s) y su(s) concentración(es), el(los) agente(s) quelante(s) y su(s) concentración(es), y otras condiciones conocidas o que pueden ser descubiertas en la técnica. Se pueden utilizar fórmulas para predecir la temperatura de fusión óptima para una secuencia perfectamente complementaria para una sonda dada, pero las temperaturas de fusión verdaderas para una sonda bajo un conjunto de condiciones de hibridación deben ser determinadas empíricamente. Además, se puede probar una sonda contra su complemento exacto para determinar una temperatura de fusión precisa bajo un conjunto dado de condiciones como se describe en Sambrook et al., "Molecular Cloning", 3 edición, Cold Spring Harbor Laboratory
- 30 Press, 2001. Se pueden alterar sistemáticamente las temperaturas de hibridación para una solución de hibridación dada usando un soporte asociado con polinucleótidos objetivo hasta identificar un rango de temperatura que permita la detección del enlazamiento de una sonda detectable al nivel de rigurosidad deseado, ya sea con alta rigurosidad donde sólo hibridan polinucleótidos objetivo con una alto grado de complementariedad, o con una menor rigurosidad donde polinucleótidos objetivo adicionales que tienen regiones de complementariedad con la sonda detectable hibridan por encima del nivel de fondo proporcionado a partir del enlazamiento no específico con polinucleótidos objetivo no complementarios y para el soporte. Cuando se lleva a cabo la hibridación con polinucleótidos objetivo potenciales sobre un soporte bajo un conjunto dado de condiciones, se lava el soporte a continuación, bajo condiciones crecientes de rigurosidad (típicamente menor concentración de sal y/o aumento de la temperatura, pero se pueden alterar otras condiciones) hasta disminuir el enlazamiento de fondo hasta el punto donde se pueden
- 40 observar señales positivas distintas. Esto puede ser monitoreado en progreso usando un contador Geiger en donde se marcó radiactivamente la sonda, radiográficamente, utilizando un formador de imágenes fluorescentes, o por otros medios de detección de enlazamiento de la sonda. No se permite que el soporte se seque durante tales procedimientos, o se puede enlazar en forma irreversible la sonda incluso con ubicaciones de fondo. Cuando una sonda produce un fondo indeseable o falsos positivos, se emplean reactivos de bloqueo, o se usan diferentes regiones de la sonda o sondas diferentes hasta que se puedan distinguir las señales positivas de las de fondo. Una vez que se encuentran las condiciones que proporcionan una señal satisfactoria por encima del fondo, los polinucleótidos objetivo que proporcionan una señal positiva se aíslan y caracterizan adicionalmente. Los polinucleótidos aislados pueden ser secuenciados; la secuencia se puede comparar con las entradas del banco de datos o con secuencias conocidas; en caso necesario, se pueden obtener clones de longitud completa por medio de
- 50 técnicas conocidas en la técnica; y se pueden expresar los polinucleótidos utilizando vectores y huéspedes adecuados para determinar si el polinucleótido identificado codifica una proteína que tiene actividad similar a aquella de la cual se deriva el polinucleótido sonda. Las sondas pueden ser de 10 - 50 nucleótidos. Sin embargo, también se pueden emplear sondas más largas, por ejemplo, de 50 - 500 nucleótidos o más grandes.

Soporte de fase sólida

- 55 El soporte de fase sólida de la presente invención puede ser de cualquier material y estructura sólidos adecuados para dar soporte a la hibridación y síntesis de nucleótidos. Preferiblemente, el soporte de fase sólida comprende al menos una superficie sustancialmente rígida sobre la cual se pueden inmovilizar oligonucleótidos o cebadores de oligonucleótidos. El soporte de fase sólida puede ser elaborado, por ejemplo, de vidrio, polímero sintético, plástico, nailon de malla no dura o de cerámica. Se conocen otros materiales adecuados de soporte sólido y se encuentran fácilmente disponibles para aquellos capacitados en la técnica. El tamaño del soporte sólido puede ser cualquiera de los tamaños estándar de microarreglos, útiles para la tecnología de microarreglos de ADN, y se puede adaptar el tamaño a la máquina particular que está siendo utilizada para llevar a cabo una reacción de la invención. Los métodos y materiales para la formación de derivados de soportes de fase sólida con el fin de inmovilizar
- 60

oligonucleótidos son conocidos por aquellos capacitados en la técnica y se describen, por ejemplo, en la patente de los Estados Unidos No. 5.919.523.

El soporte sólido puede ser proporcionado en o formar parte de un recipiente que contiene fluido. Por ejemplo, el soporte sólido se puede colocar en una cámara con lados que crean un sello a lo largo del borde del soporte sólido de manera que contenga la reacción en cadena de la polimerasa (PCR) sobre el soporte. En un ejemplo específico, la cámara puede tener paredes a cada lado de un soporte rectangular para garantizar que la mezcla de la PCR se mantenga en el soporte y también para volver toda la superficie útil para proporcionar los cebadores.

Los oligonucleótidos o cebadores de oligonucleótidos de la invención se fijan, inmovilizan, proporcionan y/o aplican a la superficie del soporte sólido utilizando cualquier medio disponible para fijar, inmovilizar, proporcionar y/o aplicar los oligonucleótidos en una ubicación particular en el soporte sólido. Por ejemplo, se puede utilizar fotolitografía (Affymetrix, Santa Clara, California) para aplicar los cebadores de oligonucleótidos en una posición particular en un chip o soporte sólido, como se describe en las patentes de los Estados Unidos Nos. 5.919.523, 5.837.832, 5.831.070, y 5.770.722. Los cebadores de oligonucleótidos también se pueden aplicar a un soporte sólido tal como se describe en Brown y Shalon, patente de los Estados Unidos No. 5.807.522 (1998). Adicionalmente, se pueden aplicar los cebadores a un soporte sólido utilizando un sistema robótico, tal como uno fabricado por Genetic Microsystems (Woburn, Massachusetts), GeneMachines (San Carlos, California) o Cartesian Technologies (Irvine, California).

En un aspecto de la invención, se lleva a cabo la amplificación en fase sólida de polinucleótidos objetivo a partir de una muestra biológica, en donde se inmovilizan múltiples grupos de cebadores de oligonucleótidos sobre un soporte de fase sólida. En una forma de realización preferida, los cebadores dentro de un grupo comprenden al menos un primer conjunto de cebadores que son idénticos en secuencia y son complementarios a una secuencia definida del polinucleótido objetivo, capaz de hibridar con el polinucleótido objetivo bajo condiciones apropiadas, y son adecuados como cebadores iniciales para la síntesis de ácido nucleico (es decir, extensión o alargamiento de la cadena). Los cebadores seleccionados que cubren una región particular de la secuencia de referencia se inmovilizan, como grupo, sobre un soporte sólido en una ubicación discreta. Preferiblemente, la distancia entre los grupos es mayor que la resolución de los medios de detección que se utilizan para la detección de los productos amplificados. En una forma de realización preferida, los cebadores se inmovilizan para formar un microarreglo o chip que puede ser procesado y analizado a través de procesamiento automatizado. Se utilizan los cebadores inmovilizados para la amplificación en fase sólida de polinucleótidos objetivo bajo condiciones adecuadas para un medio de amplificación de ácidos nucleicos. De esta forma, se pueden determinar en un ensayo la presencia o ausencia de una variedad de variaciones potenciales en el dominio quinasa del gen *erbB1*.

Se puede utilizar una población de polinucleótidos objetivo aislada de un individuo sano como control para determinar si una fuente biológica tiene al menos una variación que incrementa la actividad de quinasa en el dominio quinasa del gen *erb1*. Alternativamente, se pueden utilizar los polinucleótidos objetivo aislados de tejido sano del mismo individuo como control como anteriormente.

Se pueden llevar a cabo reacciones PCR del tipo *in situ* en los microarreglos esencialmente como se describe, por ejemplo en 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): 3.159-3.166 (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).

Alternativamente, se pueden determinar las variaciones en el dominio quinasa de *erbB1* mediante técnicas de fase sólida sin llevar a cabo una PCR sobre el soporte. Una pluralidad de sondas de oligonucleótido, conteniendo cada una una variación distinta en el dominio quinasa de *erbB1*, en duplicado, triplicado o cuadruplicado, se pueden enlazar con el soporte de fase sólida. Se puede detectar la presencia o ausencia de variaciones en la muestra biológica de prueba por medio de técnicas de hibridación selectiva, conocidas por aquellos capacitados en la técnica y descritas anteriormente.

Espectrometría de masas

En otra forma de realización, se determinan la presencia o ausencia de variaciones de ácido nucleico que incrementan la actividad de quinasa en el dominio quinasa del gen *erbB1* mediante espectrometría de masas. Para obtener una cantidad apropiada de moléculas de ácido nucleico sobre las cuales llevar a cabo la espectrometría de masas, puede ser necesaria una amplificación. Los ejemplos de procedimientos apropiados de amplificación para uso en la invención incluyen: clonación (Sambrook et al., *Molecular Cloning: A Laboratory Manual*, tercera Edición, Cold Spring Harbor Laboratory Press, 2001), reacción en cadena de la polimerasa (PCR) (C. R. Newton y A. Graham, PCR, BIOS Publishers, 1994), reacción en cadena de la ligasa (LCR) (Wiedmam, M., et al., (1994) *PCR Methods Appl.* Vol. 3, páginas. 57-64; F. Barnay Proc. Natl. Acad. Sci. EUA. 88,189-93 (1991)), amplificación por desplazamiento de cadena (SDA) (G. Terrance Walker et al., *Nucleic Acids Res.* 22, 2670-77 (1994)) y variaciones tales como RT-PCR (Higuchi, et al., *Bio/Technology* 11: 1026-1030 (1993)), amplificación específica de alelos (ASA) y procesos basados en la transcripción.

Para facilitar el análisis espectrométrico de masas, se puede inmovilizar una molécula de ácido nucleico que

contiene una secuencia de ácido nucleico que va a ser detectada a un soporte sólido. Los ejemplos de soportes sólidos adecuados incluyen perlas (por ejemplo, de gel de sílice, de vidrio de poro controlado, magnéticas, de Sefadex/Sefarosa, de celulosa), superficies planas o chips (por ejemplo, de filtros de fibra de vidrio, de superficies de vidrio, de superficies metálicas (de acero, de oro, de plata, de aluminio, de cobre y de silicio), capilares, plástico (por ejemplo, membranas de polietileno, polipropileno, poliamida, polivinilidendifloruro o placas de microtitulación)); o alfileres o peines hechos a partir de materiales similares que comprenden perlas o superficies planas o perlas colocadas en pozos en superficies planas tales como obleas (por ejemplo, obleas de silicio).

La inmovilización se puede lograr, por ejemplo, con base en la hibridación entre una secuencia de ácido nucleico de captura, que ya ha sido inmovilizada al soporte y una secuencia de ácido nucleico complementaria, que también está contenida dentro de la molécula de ácido nucleico que contiene la secuencia de ácido nucleico que va a ser detectada. Por lo tanto esa hibridación entre las moléculas de ácido nucleico complementarias no se ve obstaculizada por el soporte, el ácido nucleico de captura puede incluir una región espaciadora de al menos aproximadamente cinco nucleótidos de longitud entre el soporte sólido y la secuencia de ácido nucleico de captura. El dúplex formado se escindiría bajo la influencia del pulso láser y se puede iniciar la desorción. La secuencia base enlazada al soporte sólido se puede presentar a través de oligorribonucleótidos u oligodesoxirribonucleótidos naturales, así como análogos (por ejemplo, una estructura de fosfodiéster o fosfotriéster tio-modificado) o empleando imitaciones de oligonucleótidos tales como análogos de PNA (véase, por ejemplo Nielsen et al., Science, 254,1497 (1991)) que tornan la secuencia base menos susceptible a degradación enzimática y por lo tanto incrementan la estabilidad total de la secuencia base de captura enlazada al soporte sólido.

Antes del análisis espectrométrico de masas, puede ser útil "acondicionar" las moléculas de ácido nucleico, por ejemplo, para disminuir la energía láser requerida para la volatilización y/o para minimizar la fragmentación. El acondicionamiento se lleva a cabo preferiblemente mientras se inmoviliza un sitio de detección objetivo. Un ejemplo de acondicionamiento es la modificación de la estructura de fosfodiéster de la molécula de ácido nucleico (por ejemplo, intercambio catiónico), que puede ser útil para la eliminación del ensanchamiento de pico debido a una heterogeneidad en los cationes enlazados por unidad de nucleótido. El contacto de una molécula de ácido nucleico con un agente de alquilación tal como yoduro de alquilo, yodoacetamida, β-yodoetanol, 2,3-epoxi-1-propanol, los enlaces monotio fosfodiéster de una molécula de ácido nucleico se pueden transformar en un enlace fosfotriéster. Del mismo modo, los enlaces fosfodiéster se pueden transformar en derivados no cargados que emplean cloruros de trialkylsililo. El acondicionamiento adicional involucra la incorporación de nucleótidos que reducen la sensibilidad por despurinización (fragmentación durante la MS) tal como nucleótidos de N7 o N9-deazapurina, o bloques de construcción de ARN o utilizando triésteres de oligonucleótido o incorporando funciones fosforotioato que están alquiladas o empleando imitaciones de oligonucleótidos tales como PNA.

Para ciertas aplicaciones, puede ser útil detectar simultáneamente más de un loci (mutado) sobre un fragmento particular de ácido nucleico capturado (sobre un punto de un arreglo) o puede ser útil llevar a cabo un procesamiento paralelo por medio del uso de oligonucleótidos o arreglos miméticos de oligonucleótido sobre diferentes soportes sólidos. Se puede lograr una "multiplexación" por medio de diferentes metodologías. Por ejemplo, se pueden detectar simultáneamente diferentes mutaciones sobre una secuencia objetivo empleando las moléculas detectoras correspondientes (sonda) (por ejemplo, oligonucleótidos o miméticos de oligonucleótidos). Sin embargo, las diferencias de peso molecular entre los oligonucleótidos de detección D1, D2 y D3 deben ser lo suficientemente grandes para que sea posible la detección simultánea (multiplexación). Esto se puede lograr ya sea por medio de la secuencia misma (composición o longitud) o por medio de la introducción de grupos funcionales que modifican la masa M1-M3 dentro del oligonucleótido de detección.

Los formatos preferidos del espectrómetro de masas para uso en la invención son ionización de desorción láser asistida por matriz (MALDI), electroaspersión (ES), resonancia ciclotrón de iones (ICR) y transformadas de Fourier. Los métodos para realizar la espectrometría de masas son conocidos por aquellos capacitados en la técnica y se describen adicionalmente en Methods of Enzymology, vol. 193: "Mass Spectrometry" (JA McCloskey, editor), 1990, Academic Press, Nueva York.

Secuenciación

En otras formas de realización preferidas, la determinación de la presencia o ausencia de al menos una actividad de quinasas que incrementa la variación de ácidos nucleicos involucra la secuenciación de al menos una secuencia de ácido nucleico. La secuenciación involucra la secuenciación de una porción o porciones del dominio quinasas de erbB1 que incluye al menos un sitio de variación, y puede incluir una pluralidad de tales sitios. Preferiblemente, la porción es de 500 nucleótidos o menos de longitud, más preferiblemente 100 nucleótidos o menos, y lo más preferiblemente 45 nucleótidos o menos de longitud. Tal secuenciación se puede llevar a cabo por medio de diferentes métodos reconocidos por aquellos capacitados en la técnica, que incluye el uso de métodos de terminación didesoxi (por ejemplo, utilizando nucleótidos didesoxi marcados con colorantes), minisequenciación, y el uso de métodos espectrométricos de masas.

Inmunodetección

En una forma de realización, la determinación de la presencia o ausencia de al menos una actividad de quinasas que

incrementa la variación de ácidos nucleicos involucra la determinación del estado de activación de objetivos secuencia abajo del EGFR.

Los inventores de la presente solicitud han comparado el estado de fosforilación de la mayoría de los objetivos secuencia abajo del EGFR. Por ejemplo, se ha examinado la activación inducida por EGF de Erk1 y Erk2, a través de Ras, de Akt a través de PLC γ /PI3K, y de STAT3 y STAT5 a través de JAK2. Erk1 y Erk2, a través de Ras, Akt a través de PLC γ /PI3K y STAT3 y STAT5 a través de JAK2 son rutas esenciales secuencia abajo que median los efectos oncogénicos del EGFR (R. N. Jorissen et al., Exp. Cell Res. 284, 31 (2003)).

Los inventores de la presente solicitud han mostrado que la activación de Erk inducida por EGF es indistinguible entre las células que expresan EGFR de tipo silvestre o cualquiera de los dos que activan los mutantes del EGFR.

10 En contraste, la fosforilación tanto de Akt como de STAT5 fue sustancialmente elevada en células que expresan cualquiera de los EGFR mutantes. Se observó en forma similar una mayor fosforilación de STAT3 en células que expresan los EGFR mutantes. Por lo tanto, la autofosforilación selectiva inducida por EGF de los residuos de tirosina del terminal C dentro de los mutantes del EGFR está bien correlacionada con la activación selectiva de las rutas de señalización secuencia abajo.

15 En una forma de realización de la presente solicitud, la presencia de mutaciones del EGFR se puede determinar utilizando técnicas inmunológicas bien conocidas en la técnica, por ejemplo, técnicas de anticuerpos tales como inmunohistoquímica, inmunocitoquímica, barrido para FACS, inmunotransferencias, radioinmunoensayos, transferencia tipo Western, inmunoprecipitación, ensayos por inmunoabsorción ligados a enzimas (ELISA), y técnicas de derivación que hacen uso de anticuerpos dirigidos contra objetivos activados secuencia abajo del EGFR. Los ejemplos de tales objetivos incluyen, por ejemplo, STAT3 fosforilado, STAT5 fosforilado, y Akt fosforilado. Utilizando anticuerpos fosfo-específicos, se puede determinar el estado de activación de STAT3, STAT5, y Akt. La activación de STAT3, STAT5, y Akt es útil como un indicador de diagnóstico de la activación de mutaciones del EGFR.

25 En una forma de realización de la presente invención, la presencia de STAT5, STAT3, o Akt activados (fosforilados) indica que un tratamiento dirigido al EGFR es probable que sea efectivo.

La invención proporciona un método de cribado para variantes en el dominio quinasa del gen *erbB1* en una muestra biológica de prueba por medio de métodos inmunohistoquímicos o inmunocitoquímicos.

30 Se pueden utilizar, por ejemplo, técnicas inmunohistoquímicas ("IHC") e inmunocitoquímicas ("ICC"). IHC es la aplicación de inmunocitoquímica a secciones de tejido, mientras que ICC es la aplicación de inmunocitoquímica a impresiones de células o tejidos después de que han sido objeto de preparaciones citológicas específicas, tales como, por ejemplo, preparaciones en base líquida. La inmunocitoquímica es una familia de técnicas basadas en el uso de un anticuerpo específico, en donde se utilizan anticuerpos para moléculas específicamente objetivo dentro o sobre la superficie de las células. El anticuerpo contiene típicamente un marcador que experimentará una reacción bioquímica, y por lo tanto experimenta un cambio de color, por el encuentro con las moléculas objetivo. En algunos casos, la amplificación de la señal puede estar integrada en el protocolo particular, en donde un anticuerpo secundario, que incluye la mancha del marcador, sigue a la aplicación de un anticuerpo específico primario.

Los ensayos de inmunohistoquímica son conocidos por aquellos capacitados en la técnica (por ejemplo, véase Jalkanen, et al., J. Cell. Biol. 101: 976-985 (1985); Jalkanen, et al., J. Cell. Biol. 105: 3087-3096 (1987)).

40 Los anticuerpos, policlonales o monoclonales, pueden ser adquiridos a una variedad de proveedores comerciales, o pueden ser fabricados utilizando métodos bien conocidos, por ejemplo, como se describe en Harlow et al., Antibodies: A Laboratory Manual, 2^a edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York (1988). En general, los ejemplos de anticuerpos útiles en la presente invención incluyen anticuerpos anti-fosfo-STAT3, anti-fosfo-STAT5, y anti-fosfo-Akt. Tales anticuerpos pueden ser adquiridos, por ejemplo, a través de Upstate Biotechnology (Lake Placid, NY), New England Biolabs (Beverly, MA), NeoMarkers (Fremont, CA).

45 Por lo general, para inmunohistoquímica, se obtienen secciones de tejido de un paciente y se fijan por medio de un agente de fijación adecuado, tal como alcohol, acetona, y paraformaldehído, con el cual se hace reaccionar un anticuerpo. Los métodos convencionales para inmunohistoquímica se describen en Harlow y Lane (eds) (1988) En "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor, Nueva York; Ausbel et al., (eds) (1987), en Current Protocols in Molecular Biology, John Wiley and Sons (Nueva York, Nueva York). Las muestras biológicas apropiadas para tales ensayos de detección incluyen, pero no se limitan a, células, biopsias de tejido, sangre entera, plasma, suero, esputo, fluido cerebro espinal, aspirados de mama, líquido pleural, orina y similares.

Para las técnicas de marcación directa, se utiliza un anticuerpo marcado. Para las técnicas de marcación indirecta, se hace reaccionar adicionalmente la muestra con una sustancia marcada.

55 Alternativamente, se puede utilizar inmunocitoquímica. En general, las células se obtienen de un paciente y se fijan por medio de un agente de fijación adecuado, tal como alcohol, acetona, y paraformaldehído, con el cual se hace reaccionar un anticuerpo. Los métodos de coloración inmunocitológica de muestras humanas son conocidos por

aquellos capacitados en el arte y están descritos, por ejemplo, en Brauer et al., 2001 (FASEB J, 15, 2689-2701), Smith-Swintosky et al., 1997.

Los métodos inmunológicos de la presente invención son ventajosos ya que requieren únicamente pequeñas cantidades de material biológico. Tales métodos pueden ser desarrollados a nivel celular y por lo tanto requieren como mínimo una célula. Preferiblemente, se obtienen diferentes células de un paciente afectado con o en riesgo de desarrollar cáncer y se evalúan de acuerdo con los métodos de la presente invención.

Otros métodos de diagnóstico

Un agente para detectar la proteína del EGFR mutante es un anticuerpo capaz de enlazarse a la proteína del EGFR mutante, preferiblemente un anticuerpo con un marcador detectable. Los anticuerpos pueden ser policlonales, o más preferiblemente, monoclonales. Se puede utilizar un anticuerpo intacto, o un fragmento del mismo (por ejemplo, F_{ab} o $F_{(ab)2}$). El término "marcado", con respecto a la sonda o anticuerpo, pretende abarcar el etiquetado directo de la sonda o anticuerpo por acoplamiento (es decir, enlazamiento físico) de una sustancia detectable con la sonda o anticuerpo, así como marcación indirecta de la sonda o anticuerpo por reacción con otro reactivo que está marcado directamente. Ejemplos de marcación indirecta incluyen la detección de un anticuerpo primario usando un anticuerpo secundario marcado en forma fluorescente y marcación final de una sonda de ADN con biotina de manera que pueda ser detectada con estreptavidina marcada en forma fluorescente. El término "muestra biológica" pretende incluir tejidos, células y fluidos biológicos aislados de un individuo, así como tejidos, células y fluidos presentes en un sujeto. Es decir, se puede utilizar el método de detección de la invención para detectar ARNm mutante del EGFR, proteína, o ADN genómico en una muestra biológica in vitro así como in vivo. Por ejemplo, las técnicas in vitro para la detección de ARNm mutante del EGFR incluyen hibridaciones tipo Northern e hibridaciones in situ. Las técnicas in vitro para la detección de proteína del EGFR mutante incluyen ensayos por inmunoabsorción ligados a enzimas (los ELISA), transferencias tipo Western, inmunoprecipitaciones, e inmunofluorescencia. Las técnicas in vitro para detección de ADN genómico mutante del EGFR incluyen hibridaciones tipo Southern. Además, las técnicas in vivo para detección de proteína del EGFR mutante incluyen la introducción en un individuo de un anticuerpo marcado de proteína anti-EGFR mutante. Por ejemplo, el anticuerpo puede ser marcado con un marcador radiactivo cuya presencia y localización en un individuo se puede detectar mediante técnicas estándar de formación de imágenes.

En una forma de realización, la muestra biológica contiene moléculas de proteína del individuo de prueba. Alternativamente, la muestra biológica puede contener moléculas de ARNm del individuo de prueba o moléculas de ADN genómico del individuo de prueba.

En otra forma de realización, los métodos involucran además la obtención de una muestra biológica de control de un individuo de control, poniendo en contacto la muestra de control con un compuesto o agente capaz de detectar proteína del EGFR mutante, ARNm, o ADN genómico, de tal manera que se detecta la presencia de proteína del EGFR mutante, ARNm o ADN genómico en la muestra biológica, y se compara la presencia de proteína del EGFR mutante, ARNm o ADN genómico en la muestra de control con la presencia de proteína del EGFR mutante, ARNm o ADN genómico en la muestra de prueba.

En una forma diferente de realización, el ensayo diagnóstico es para actividad del EGFR mutante. En una forma de realización específica, la actividad del EGFR mutante es una actividad de tirosina quinasa. El ensayo diagnóstico es para detectar fosforilación mediada por EGFR de al menos un sustrato del EGFR. Los niveles de actividad del EGFR se pueden evaluar, por ejemplo, para diferentes polipéptidos del EGFR mutante, diferentes tejidos que contienen EGFR mutante, biopsias de tejidos cancerosos que se sospecha que tengan al menos un EGFR mutante, y similares. Pueden hacerse opcionalmente comparaciones de los niveles de actividad del EGFR en las diferentes células, tejidos o extractos de la misma. En una forma de realización, los altos niveles de actividad del EGFR en tejido canceroso sirven de diagnóstico para cánceres que pueden ser susceptibles de tratamientos con uno o más inhibidores de tirosina quinasa. En formas de realización relacionadas, se pueden determinar los niveles de actividad del EGFR entre muestras de biopsias tratadas y no tratadas, líneas celulares, animales transgénicos, o extractos de cualquiera de estos, para determinar el efecto de un tratamiento dado sobre la actividad del EGFR mutante en comparación con un control no tratado.

Método de tratamiento de un paciente

En una forma de realización, la invención proporciona un método para seleccionar un tratamiento para un paciente afectado por o en riesgo de desarrollar cáncer por medio de la determinación de la presencia o ausencia de al menos una actividad de quinasa que incrementa la variación de ácidos nucleicos en el dominio quinasa del gen *erbB1*. En otra forma de realización, la variación es una pluralidad de variaciones, mediante el cual una pluralidad puede incluir variaciones de uno, dos, tres o más loci de genes.

En ciertas formas de realización, la presencia de al menos una variación es indicativa de que el tratamiento será efectivo o bien beneficioso (o más probablemente que sea beneficioso) en el paciente. La manifestación de que el tratamiento será efectivo significa que la probabilidad un efecto terapéutico beneficioso será mayor que en una persona que no tenga la presencia adecuada de la(s) variación(ones) de ácido nucleico que incrementan la actividad de la quinasa particular en el dominio quinasa del gen *erbB1*.

El tratamiento involucrará la administración de un inhibidor de tirosina quinasa. El tratamiento puede involucrar una combinación de tratamientos, que incluyen, pero no se limitan a, un inhibidor de tirosina quinasa en combinación con otros inhibidores de tirosina quinasa, quimioterapia, radiación, etc.

Por lo tanto, en relación con la administración de un inhibidor de tirosina quinasa, un fármaco que sea "efectivo contra" un cáncer indica que la administración en una forma clínicamente apropiada da como resultado un efecto beneficioso para al menos una fracción de pacientes estadísticamente significativa, tal como una mejoría de los síntomas, una cura, una reducción en la carga de la enfermedad, una reducción en la masa o en el número de células del tumor, prolongación de la vida, una mejoría en la calidad de vida, u otro efecto generalmente reconocido como positivo por parte de los médicos familiarizados con el tratamiento del tipo particular de enfermedad o condición.

En una forma de realización preferida, el compuesto es una anilinoquinazolina o anilinoquinazolina sintética. La publicación de la patente europea No. 0566226 divulga anilinoquinazolinas que tienen actividad contra la tirosina quinasa del receptor del factor de crecimiento epidérmico (EGF). También se sabe a partir de las solicitudes de patente europeas números 0520722 y 0566226 que ciertos derivados de 4-anilinoquinazolina son útiles como inhibidores de tirosina quinasa del receptor. Las relaciones muy estrechas de la actividad con la estructura mostradas por estos compuestos sugieren un modo de enlace claramente definido, en donde el anillo de quinazolina se enlaza en el bolsillo de adenina y el anillo anilino se enlaza en un bolsillo lipofílico único adyacente. Se han evaluado clínicamente tres análogos de 4-anilinoquinazolina (dos inhibidores reversibles y uno irreversible) como compuestos farmacéuticos contra el cáncer. Denny, *Farmaco* enero - febrero 2001; 56 (1 - 2): 51 - 6. Alternativamente, el compuesto es EKB-569, un inhibidor de quinasa del receptor de EGF (Torrance et al., *Nature Medicine*, vol. 6, No. 9, septiembre de 2000, página 1024). En una forma de realización más preferida, el compuesto es gefitinib (IRESSA®) o erlotinib (TARCEVA®).

El tratamiento dirigido a las células cancerosas que contienen al menos un EGFR mutante descrito aquí puede ser administrado solo o en combinación con cualquier otro tratamiento apropiado contra el cáncer y/o agente terapéutico conocido por alguien capacitado en el arte. En una forma de realización, existe un tratamiento de una patología, tal como un cáncer, que comprende administrar a un individuo que requiera del mismo cantidades terapéuticamente efectivas de un compuesto que inhibe la actividad de quinasa del EGFR, tales como gefitinib, erlotinib, etc., administrados solos o en combinación con al menos otro agente o terapia contra el cáncer. La inhibición de proteína quinasas activadas a través del uso de compuestos farmacéuticos de moléculas pequeñas específicas o de estrategias basadas en anticuerpos se ha convertido en un enfoque efectivo como terapia contra el cáncer. Véase, por ejemplo, G. D. Demetri et al., *N. Engl. 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).

En una forma de realización, el agente contra el cáncer es al menos un agente quimioterapéutico. En una forma de realización relacionada, el agente contra el cáncer es al menos una radioterapia. En una variante de una forma de realización, la terapia contra el cáncer es una terapia antiangiogénica (por ejemplo, endostatina, angiostatina, TNP-470, caplostatina (Stachi-Fainaro et al., *Cancer Cell* 7 (3), 251 (2005)).

Los agentes terapéuticos pueden ser iguales o diferentes, y pueden ser, por ejemplo, radionucleidos terapéuticos, fármacos, hormonas, antagonistas de hormonas, antagonistas de receptores, enzimas o proenzimas activadas por otro agente, autocrinas, citoquinas o cualquier agente adecuado contra el cáncer conocido por aquellos capacitados en la técnica. En una forma de realización, el agente contra el cáncer es avastina, un anticuerpo anti-VEGF probado exitosamente en terapia antiangiogénica del cáncer tanto contra cánceres sólidos y como contra neoplasias hematológicas. Véase, por ejemplo, Ribatti et al., 2003 *J Hematother Stem Cel Res.* 12 (1), 11-22. También se pueden utilizar toxinas en los métodos de la presente invención. Otros agentes terapéuticos útiles en la presente invención incluyen anti-ADN, anti-ARN, oligonucleótidos marcados radiactivamente, tales como oligonucleótidos antisentido, agentes citotóxicos o antimicrobianos anti-proteína y anti-cromatina. Otros agentes terapéuticos son conocidos por aquellos capacitados en el arte, y se contempla específicamente el uso de tales otros agentes terapéuticos de acuerdo con la presente invención.

El agente antitumoral puede ser uno entre numerosos agentes quimioterapéuticos tales como un agente alquilante, un antimetabolito, un agente hormonal, un antibiótico, un anticuerpo, un compuesto biológico contra el cáncer, Gleeevec, colchicina, un alcaloide de la vinca, L-asparaginasa, procarbazona, hidroxiurea, mitotano, nitrosoureas o una imidazol carboxamida. Agentes adecuados son aquellos agentes que promueven la despolarización de tubulina o prohíben la proliferación de células tumorales. Los agentes quimioterapéuticos contemplados dentro del alcance de la invención incluyen, pero no se limitan a, agentes contra el cáncer enumeradas en el Orange Book of Approved Drug Products With Therapeutic Equivalence Evaluations, como los recopilados por la Administración de Alimentos y Medicamentos y el Departamento de Salud y Servicios Humanos de los Estados Unidos. Ejemplos no limitantes de agentes quimioterapéuticos incluyen, por ejemplo, carboplatino y paclitaxel. Los tratamientos dirigidos a la actividad de la quinasa del EGFR también se pueden administrar con un tratamiento de terapia por radiación. Se contemplan tratamientos adicionales contra el cáncer conocidos en el arte dentro del alcance de la invención.

El agente terapéutico puede ser un agente quimioterapéutico. Los agentes quimioterapéuticos son conocidos en la técnica e incluyen al menos los taxanos, mostazas de nitrógeno, derivados de etilenimina, sulfonatos de alquilo,

nitrosoureas, triazenos, análogos de ácido fólico, análogos de pirimidina, análogos de purina, alcaloides de la vinca, antibióticos, enzimas, complejos de coordinación de platino, urea sustituida, derivados de metil hidracina, supresores adrenocorticales, o antagonistas. Más específicamente, los agentes quimioterapéuticos pueden ser uno o más agentes escogidos del grupo no limitante de esteroides, progestinas, estrógenos, antiestrógenos, o andrógenos. Aún más específicamente, los agentes quimioterapéuticos pueden ser azaribina, bleomicina, briostatina-1, busulfán, carmustina, clorambucilo, carboplatino, cisplatino, CPT-11, ciclofosfamida, citarabina, dacarbazina, dactinomycin, daunorubicina, dexametasona, dietilestilbestrol, doxorubicina, etinil estradiol, etopósido, fluorouracilo, flouximesterona, gemcitabina, caproato de hidroxiprogesterona, hidroxiaurea, L-asparaginasa, leucovorina, lomustina, mecloretamina, acetato de medroprogesterona, acetato de megestrol, melfalán, mercaptopurina, metotrexato, mitramicina, mitomicina, mitotano, paclitaxel, butirato de fenilo, prednisona, procarbazona, semustina, estreptozocina, tamoxifeno, taxanos, taxol, propionato de testosterona, talidomida, tioguanina, tiotepa, mostaza de uracilo, vinblastina, o vincristina. También se contempla el uso de cualquiera de la combinación de agentes quimioterapéuticos. La administración del agente quimioterapéutico puede hacerse antes, durante o después de la administración de un tratamiento dirigido a la actividad del EGFR.

Otros agentes terapéuticos adecuados se seleccionan del grupo que consiste de radioisótopo, adición de boro, inmunomodulador, toxina, agente fotoactivo o colorante, compuesto farmacéutico quimioterapéutico contra el cáncer, compuesto farmacéutico antiviral, compuesto farmacéutico contra hongos, compuesto farmacéutico antibacteriano, compuesto farmacéutico contra protozoarios y agente quimiosensibilizador (véanse, las patentes de los Estados Unidos Nos. 4.925.648 y 4.932.412). Se describen agentes quimioterapéuticos adecuados en Remington's Pharmaceutical Sciences, 19ª edición. (Mack Publishing Co., 1995), y en The Pharmacological Basis of Therapeutics de Goodman y Gilman (Goodman et al., Eds. Macmillan Publishing Co., New York, ediciones 1980 y 2001). Otros agentes quimioterapéuticos adecuados, tales como compuestos farmacéuticos experimentales, son conocidos por aquellos capacitados en el arte. Por otra parte, se selecciona un radioisótopo terapéutico adecuado del grupo que consiste de emisores α , emisores β , emisores γ , emisores de electrones Auger, agentes de captura de neutrones que emiten partículas α y radioisótopos que decaen por la captura de electrones. Preferiblemente, se selecciona el radioisótopo del grupo que consiste de 225Ac, 198Au, 32P, 125I, 131I, 90Y, 186Re, 188Re, 67Cu, 177Lu, 213Bi, 10B, y 211At.

Cuando se utiliza más de un agente terapéutico, pueden ser iguales o diferentes. Por ejemplo, los agentes terapéuticos pueden incluir diferentes radionucleidos, o un compuesto farmacéutico y un radionucleido. En una forma de realización preferida, el tratamiento dirigido a la actividad del EGFR inhibe la actividad de quinasa del EGFR mutante.

En otra forma de realización, se utilizan diferentes isótopos que son efectivos en diferentes distancias como resultado de sus emisiones individuales de energía como primero y segundo agentes terapéuticos. Tales agentes pueden ser utilizados para lograr un tratamiento más efectivo de los tumores, y son útiles en pacientes que presentan múltiples tumores de diferentes tamaños, como en circunstancias clínicas normales.

Pocos de los isótopos disponibles son útiles para el tratamiento de depósitos tumorales muy pequeños y células individuales. En estas situaciones, un compuesto farmacéutico o toxina pueden ser un agente terapéutico más útil. En consecuencia, en formas de realización preferidas de la presente invención, se utilizan isótopos en combinación con especies no isotópicas tales como compuestos farmacéuticos, toxinas, y agentes de captura de neutrones. Se conocen muchos compuestos farmacéuticos y toxinas que tienen efectos citotóxicos sobre las células, y pueden ser utilizados en conexión con la presente invención. Se encuentran en compendios de compuestos farmacéuticos y toxinas, tales como el Merck Index, Goodman y Gilman, y similares, y en las referencias citadas anteriormente.

Los compuestos farmacéuticos que interfieren con la síntesis de proteína intracelular también pueden ser utilizados en los métodos de la presente invención; tales compuestos farmacéuticos son conocidos por aquellos capacitados en el arte e incluyen puomicina, cicloheximida y ribonucleasa.

Los métodos terapéuticos de la invención pueden ser utilizados para terapia contra el cáncer. Se sabe bien que los radioisótopos, los compuestos farmacéuticos y las toxinas se pueden conjugar con anticuerpos o con fragmentos de anticuerpos que se enlazan específicamente con marcadores que son producidos por o están asociados con células cancerosas, y que tales conjugados de anticuerpo pueden ser usados para dirigir los radioisótopos, compuestos farmacéuticos o toxinas a los sitios tumorales para mejorar su eficacia terapéutica y minimizar los efectos secundarios. Ejemplos de estos agentes y métodos son revisados en un Wawrzynczak y Thorpe (en Introduction to the Cellular and Molecular Biology of Cancer, L. M. Franks y N. M. Teich, eds. capítulo 18, páginas 378 - 410, Oxford University Press. Oxford, 1986), en Immunoconjugates: Antibody Conjugates in Radioimaging and Therapy of Cancer (C. W. Vogel, ed., 3 - 300, Oxford University Press, N.Y., 1987), en Dillman, R. O. (CRC Critical Reviews in Oncology/Hematology 1: 357, CRC Press, Inc., 1984), en Pastan et al., (Cell 47: 641, 1986). En Vitetta et al., (Science 238: 1098 - 1104, 1987) y en Brady et al., (Int. J. Rad. Oncol. Biol. Phys. 13: 1535 - 1544, 1987). Otros ejemplos del uso de immunoconjugados para el cáncer y otras formas de terapia han sido divulgados, entre otros, en las patentes de los Estados Unidos Nos. 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, 6.187.287, y 6.319.500.

Adicionalmente, se pueden utilizar los métodos de tratamiento de la invención en combinación con otros compuestos o técnicas para prevenir, mitigar o revertir los efectos secundarios de ciertos agentes citotóxicos. Los ejemplos de tales combinaciones incluyen, por ejemplo, la administración de IL-1 junto con un anticuerpo para rápida eliminación, como se describe, por ejemplo en la Patente de los Estados Unidos No. 4.624.846. Tal administración se puede

- 5 realizar de 3 a 72 horas después de la administración de un tratamiento terapéutico primario dirigido a la actividad del EGFR en combinación con un agente contra el cáncer (por ejemplo, con un radioisótopo, compuesto farmacéutico o toxina como el componente citotóxico). Esto se puede utilizar para mejorar la eliminación del conjugado, el compuesto farmacéutico o la toxina del torrente sanguíneo y para mitigar o revertir la toxicidad mieloide y otra toxicidad hematopoyético causada por el agente terapéutico.
- 10 En otro aspecto de la invención, la terapia contra el cáncer puede involucrar una combinación de más de un agente tumoricida, por ejemplo, un compuesto farmacéutico y un radioisótopo, o un radioisótopo y un agente de boro 10 para terapia activada por neutrones, o un compuesto farmacéutico y un modificador de respuesta biológica, o un conjugado de moléculas de fusión y un modificador de respuesta biológica. Se puede integrar la citoquina en tal régimen terapéutico para maximizar la eficacia de cada componente de la misma.
- 15 En forma similar, ciertos anticuerpos contra la leucemia y contra linfoma conjugados con radioisótopos que son emisores β o α pueden inducir efectos secundarios mieloides y otros efectos secundarios hematopoyéticos cuando estos agentes no están dirigidos únicamente a las células tumorales. Esto se observa particularmente cuando las células tumorales están en el torrente sanguíneo y en los órganos que forma la sangre. Se prefiere la administración concomitante y/o posterior de al menos una citoquina hematopoyética (por ejemplo, factores de crecimiento, tales
- 20 como factores estimulantes de colonias, tales como G-CSF y GM-CSF) para reducir o mejorar los efectos secundarios hematopoyéticos, aumentando al mismo tiempo los efectos contra el cáncer.

Se conoce bien en el arte que se pueden utilizar diferentes métodos de terapia con radionucleidos para el tratamiento del cáncer y otras condiciones patológicas, como se describe, por ejemplo, en Harbert, "Nuclear Medicine Therapy", Nueva York, Thieme Medical Publishers, 1087, páginas 1-340. Un médico experimentado en

25 estos procedimientos será capaz de adaptar fácilmente la terapia adyuvante de citoquina descrita aquí a tales procedimientos para mitigar cualquiera de los efectos secundarios hematopoyéticos de los mismos. En forma similar, se puede utilizar una terapia con compuestos farmacéuticos citotóxicos, administrada con el tratamiento dirigido a la actividad del EGFR, por ejemplo, para el tratamiento del cáncer o de otras enfermedades proliferativas celulares. Tal tratamiento está gobernado por medio de principios análogos a los de la terapia de radioisótopos con isótopos o anticuerpos marcados en forma radioactiva. El médico ordinariamente capacitado será capaz de adaptar la

30 administración de la terapia adicional contra el cáncer durante y/o después de la terapia primaria contra el cáncer.

Kits

Por consiguiente, la presente invención proporciona kits predictivos, de diagnóstico, y de pronóstico que comprenden cebadores degenerados para amplificar un ácido nucleico objetivo en el dominio quinasa del gen erbB1

35 e instrucciones que incluyen la amplificación del protocolo y el análisis de los resultados. El kit también puede incluir alternativamente amortiguadores, enzimas, y contenedores para llevar a cabo la amplificación y el análisis de los productos de la amplificación. El kit también puede ser un componente de un kit de cribado, diagnóstico o pronóstico que incluye otras herramientas tales como microarreglos de ADN. Preferiblemente, el kit también proporciona una o más plantillas de control, tales como ácidos nucleicos aislados de una muestra de tejido normal, y/o una serie de

40 muestras que representan diferentes variaciones en el dominio quinasa del gen erbB1.

En una forma de realización, el kit proporciona dos o más pares de cebadores, cada par capaz de amplificar una región diferente del gen erbB1 (cada región un sitio de variación potencial) proporcionando por lo tanto un kit para análisis de la expresión de diferentes variaciones de los genes en una muestra biológica en una reacción o en varias reacciones paralelas.

- 45 Los cebadores en los kits pueden ser marcados, por ejemplo marcados en forma fluorescente, para facilitar la detección de los productos de la amplificación y el análisis consecuente de las variaciones de ácido nucleico.

En una forma de realización, se puede detectar más de una variación en un análisis. Un kit de combinación incluirá por lo tanto cebadores capaces de amplificar diferentes segmentos del dominio quinasa del gen erbB1. Los cebadores pueden ser marcados en forma diferencial, por ejemplo, utilizando diferentes marcadores fluorescentes,

50 para diferenciar entre las variaciones.

Los cebadores contenidos en el kit pueden incluir los siguientes cebadores: cebador sentido del Exón 19, 5'-GCAATATCAGCCTTAGSTGCGGCTC-3' (SEQ ID No: 505), cebador antisentido del exón 19, 5'-CATAGAAAGTGAACATTTAGGATGTG-3' (SEQ ID No: 506); cebador sentido del exón 21, 5'-CTAACGTTCCGCCAGCCATAAGTC

55 C-3' (SEQ ID No: 507) y cebador antisentido del exón 21, 5'-GCTGCGAGCTCACCCAGAATGTCTGG-3' (SEQ ID No: 508).

En una forma de realización preferida, se seleccionan los cebadores del grupo que consiste de las SEQ ID Nos. 646 a 673 (véanse las Tablas 5 y 6). Estos cebadores tienen la SEQ ID No. 645 sobre el extremo 5' del cebador hacia adelante y la SEQ ID No. 674 en el extremo 5' de los cebadores inversos.

Kits de inmunodetección

En formas de realización adicionales, la invención proporciona kits inmunológicos para uso en la detección de los niveles de activación de los objetivos del EGFR secuencia abajo (es decir, STAT3, STAT5, y Akt). Tales kits generalmente contendrán uno o más anticuerpos que tienen inmunoespecificidad para la forma fosforilada de STAT3, STAT5, o Akt.

Un kit que comprende un anticuerpo capaz de enlazarse en forma inmunoespecífica a una proteína fosforilada en una célula de mamífero seleccionada del grupo que consiste de proteínas Akt, STAT3, y STAT5 fosforiladas e instrucciones para el uso del anticuerpo para examinar la célula de mamífero para la activación de la ruta de Akt, STAT3 o STAT5 es provista en la presente invención. En métodos preferidos, el kit comprende diferentes anticuerpos, cada uno de los cuales es capaz de enlazarse inmunoespecíficamente con proteínas fosforiladas en una célula de mamífero seleccionada del grupo que consiste de proteínas Akt, STAT3, STAT5 fosforiladas.

El kit comprende, en general, a) un vehículo farmacéuticamente aceptable; b) un anticuerpo dirigido contra STAT3, STAT5, o Akt fosforilados, en un medio contenedor adecuado; y c) un reactivo de inmunodetección. Los anticuerpos (monoclonales o policlonales) se encuentran disponibles comercialmente y también se pueden preparar por medio de métodos conocidos por aquellos capacitados en la técnica, por ejemplo, en Current Protocols in Immunology, John Wiley & Sons, Editado por: John E. Coligan, Ada M. Kruisbeek, David H. Margulies, Ethan M. Shevach, Warren Strober, 2001.

En ciertas formas de realización, el antígeno o el anticuerpo puede estar unido a un soporte sólido, tal como una matriz de columna o pozo de una placa de microtitulación. Los reactivos de inmunodetección del kit pueden adoptar cualquiera de una variedad de formas, incluyendo aquellas etiquetas detectables que se asocian con, o que se enlazan con, el anticuerpo dado o antígeno mismo. También se contemplan las etiquetas detectables que se asocian con o que se unen a un ligando de enlazamiento secundario. Ejemplos de ligandos secundarios son aquellos anticuerpos secundarios que tienen afinidad de enlazamiento por el primer anticuerpo o antígeno.

Los marcadores de ensayo adecuados son conocidos en la técnica e incluyen marcadores enzimáticos, tales como, glucosa oxidasa, radioisótopos, tales como yodo (^{131}I , ^{125}I , ^{123}I , ^{121}I), carbono (^{14}C), azufre (^{35}S), tritio (^3H), indio ($^{115\text{m}}\text{In}$, $^{113\text{m}}\text{In}$, ^{112}In , ^{111}In), y tecnecio (^{99}Tc , $^{99\text{m}}\text{Tc}$), talio (^{201}Tl), galio (^{68}Ga , ^{67}Ga), paladio (^{103}Pd), molibdeno (^{99}Mo), xenón (^{133}Xe), flúor (^{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}Ru ; etiquetas luminiscentes tales como luminol; y etiquetas fluorescentes, tales como fluoresceína y rodamina, y biotina.

Otros reactivos de inmunodetección adecuados para uso en los presentes kits incluyen el reactivo de dos componentes que comprende un anticuerpo secundario que tiene afinidad de enlazamiento por el primer anticuerpo o antígeno, junto con un tercer anticuerpo que tiene afinidad de enlazamiento por el segundo anticuerpo, en donde el tercer anticuerpo se enlaza a un marcador detectable.

Se conocen una cantidad de ejemplos de etiquetas o marcadores en la técnica y todas estas etiquetas se puede emplear en conexión con la presente invención. Los marcadores radiactivos, los isótopos de resonancia de espín magnético nuclear, etiquetas fluorescentes y etiquetas de enzimas capaces de generar un producto coloreado tras el contacto con un sustrato apropiado son ejemplos adecuados.

Los kits pueden contener conjugados de etiqueta-anticuerpo, ya sea en forma completamente conjugada, en forma de compuestos intermedios, o como fracciones separadas para ser conjugados por el usuario del kit.

Los kits pueden incluir además una composición adecuada en alícuotas de un antígeno ya sea marcado o no marcado, tal como se puede utilizar para preparar una curva estándar para un ensayo de detección o como un control positivo.

Los kits de la invención, independientemente del tipo, generalmente incluirán uno o más contenedores en los cuales se colocan los agentes biológicos y, preferiblemente, en alícuotas adecuadas. Los componentes de los kits pueden empacarse o bien en medio acuoso o en forma liofilizada.

Los kits de inmunodetección de la invención pueden contener adicionalmente uno o más de una variedad de otros anticuerpos o antígenos marcadores de cáncer, si así se desea. Tales kits podrían por lo tanto proporcionar un panel de marcadores de cáncer, que pueden ser utilizados mejor en las pruebas de una variedad de pacientes. A modo de ejemplo, dichos marcadores adicionales podrían incluir, otros marcadores tumorales, tales como PSA, SeLe (X), HCG, así como p53, ciclina D1, p16, tirosinasa, MAGE, BAGE, PAGE, MUC18, CEA, p27, [bgr]HCG u otros marcadores conocidos por aquellos capacitados en la técnica.

El medio contenedor de los kits, generalmente, incluirá al menos un vial, un tubo de ensayo, un matraz, una botella, o incluso una jeringa u otro medio contenedor, en el que se pueden colocar el anticuerpo o el antígeno, y preferiblemente, en forma de alícuotas adecuadas. Cuando se proporciona un segundo o un tercer ligando de enlazamiento o un componente adicional, el kit también contendrá generalmente un segundo, un tercer u otro contenedor adicional en el que se puede colocar este ligando o componente.

Los kits de la presente invención también incluirán típicamente un medio para contener el anticuerpo, el antígeno, y cualesquiera otros contenedores de reactivos en un confinamiento cerrado para la venta comercial. Estos contenedores pueden incluir contenedores plásticos moldeados por inyección o por soplado en los que se mantienen los viales deseados.

- 5 Los métodos de la presente invención también abarcan la identificación de compuestos que interfieren con la actividad de la quinasa de una forma variante del EGFR. El EGFR variante comprende al menos una variación en su dominio quinasa. Tales compuestos pueden, por ejemplo, ser inhibidores de la tirosina quinasa. Los métodos para la identificación de compuestos que interfieren con la actividad de la quinasa de un receptor son generalmente conocidos por aquellos capacitados en la técnica y se describen adicionalmente, por ejemplo, en Dhanabal et al.,
- 10 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 - 549. En general, se identifican los compuestos, usando los métodos divulgados en este documento, que interfieren con la actividad mejorada característica de la quinasa de al menos una variación en el dominio quinasa del gen erbB1.

15 Soporte sólido

- En otra forma de realización, la invención proporciona un kit para la práctica de los métodos de la invención. En una forma de realización, se describe un kit para la detección de variaciones en el dominio quinasa del gen erbB1 en un soporte sólido. El kit puede incluir, por ejemplo, los materiales y reactivos para la detección de una pluralidad de variaciones en un ensayo. El kit puede incluir por ejemplo, un soporte sólido, cebadores de oligonucleótidos para un
- 20 conjunto específico de polinucleótidos objetivo, reactivos para la reacción en cadena de la polimerasa y componentes, por ejemplo, enzimas para la síntesis de ADN, materiales marcadores, y otros amortiguadores y reactivos para el lavado. El kit también puede incluir instrucciones para uso del kit para amplificar objetivos específicos sobre un soporte sólido. Cuando el kit contiene un soporte sólido preparado que tiene un conjunto de cebadores ya fijados sobre el soporte sólido, por ejemplo, para la amplificación de un conjunto particular de
- 25 polinucleótidos objetivo, el diseño y la construcción de tal soporte sólido preparado se describió anteriormente. El kit también incluye reactivos necesarios para la realización de una PCR sobre un soporte sólido, por ejemplo utilizando un procedimiento de PCR de tipo de fase sólida o del tipo in situ, donde el soporte es capaz de amplificación por PCR usando un máquina de PCR de tipo in situ. Los reactivos de PCR, incluidos en el kit, incluyen los amortiguadores de PCR habituales, una polimerasa termoestable (por ejemplo, Taq ADN polimerasa), nucleótidos
- 30 (por ejemplo, los dNTP), y otros componentes y moléculas de marcación (por ejemplo, para marcación directa o indirecta tal como se describe más arriba). Los kits se pueden ensamblar para apoyar la práctica del método de amplificación por PCR utilizando cebadores inmovilizados solos o, alternativamente, junto con los cebadores en fase de solución.

- Alternativamente, el kit puede incluir un soporte sólido con oligonucleótidos específicos fijados a cualquier cantidad de variaciones del EGFR, que se definen adicionalmente en las Figuras 4A - 4C y en las Figuras 7 y 8. Se puede aplicar una muestra biológica de prueba al soporte sólido, bajo condiciones de hibridación selectivas, para la
- 35 determinación de la presencia o ausencia de variaciones en el dominio quinasa de erbB1.

- Los métodos de la presente invención también abarcan la identificación de compuestos que interfieren con la actividad de la quinasa de una forma variante del EGFR. El EGFR variante comprende al menos una variación en su dominio quinasa. Sin embargo, en una forma de realización alternativa, el EGFR variante comprende una mutación secundaria que confiere resistencia a una primera TKI por ejemplo, gefitinib o erlotinib. Tales compuestos pueden, por ejemplo, ser inhibidores de la tirosina quinasa. Los métodos para la identificación de compuestos que interfieren con la actividad de la quinasa de un receptor son generalmente conocidos por aquellos capacitados en la técnica y se describen adicionalmente, por ejemplo, en Dhanabal et al., Cancer Res. 59: 189 - 197 (1999); Xin et al., J. Biol.
- 40 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 - 549. En general, se identifican compuestos, usando los métodos divulgados en este documento, que interfieren con la actividad mejorada característica de la quinasa de al menos una variación en el dominio quinasa del gen erbB1. Tales variaciones conocidas se describen en las Figuras 4, 7, 8 y en la Tabla 2.

- 50 Una vez identificados, tales compuestos se administran a los pacientes que requieran de tratamiento dirigido al EGFR, por ejemplo, los pacientes afectados con o en riesgo de desarrollar cáncer.

- La vía de administración puede ser intravenosa (I.V.), intramuscular (IM), subcutánea (SC), intradérmica (ID), intraperitoneal (IP), intratecal (IT), intrapleurales, intrauterina, rectal, vaginal, tópica, intratumoral y similares. Los compuestos de la invención se pueden administrar por vía parenteral mediante inyección o por infusión gradual en el
- 55 tiempo y pueden ser suministrados por medio peristáltico.

La administración puede ser por vía transmucosa o transdérmica. Para la administración transmucosa o transdérmica, se usan penetrantes apropiados para permear la barrera en la formulación. Tales penetrantes son generalmente conocidos en la técnica, e incluyen, por ejemplo, para la administración transmucosa sales biliares y derivados de ácido fusídico. Además, pueden usarse detergentes para facilitar la permeación. La administración

transmucosa puede ser a través de aerosoles nasales, por ejemplo, o usando supositorios. Para la administración oral, los compuestos de la invención se formulan en formas de administración oral convencionales tales como cápsulas, comprimidos y tónicos.

5 Para la administración tópica, la composición farmacéutica (inhibidor de la actividad de la quinasa) se formula en pomadas, ungüentos, geles, o cremas, como se conoce generalmente en la técnica.

10 Las composiciones terapéuticas de esta invención se administran convencionalmente por vía intravenosa, como mediante inyección de una dosis unitaria, por ejemplo. El término "dosis unitaria" cuando se usa en referencia a una composición terapéutica de la presente invención se refiere a unidades físicamente discretas adecuadas como dosificación unitaria para el individuo, conteniendo cada unidad una cantidad predeterminada de material activo calculada para producir el efecto terapéutico deseado junto con el diluyente requerido, es decir, el portador, o vehículo.

15 Las composiciones se administran de una manera compatible con la formulación de la dosificación, y en una cantidad terapéuticamente efectiva. La cantidad a administrar y el momento dependen del sujeto a tratar, de la capacidad del sistema del individuo para utilizar el ingrediente activo, y del grado de efecto terapéutico deseado. Las cantidades precisas de ingrediente activo requeridas para ser administradas dependen del juicio del médico tratante y son particulares para cada individuo.

20 Los inhibidores de tirosina quinasa útiles para la práctica de los métodos de la presente invención se describen en este documento. Se puede utilizar cualquier formulación o sistema de suministro de fármacos que contengan los ingredientes activos, que sea adecuada para el uso previsto, como generalmente las conocen aquellos capacitados en la técnica. Los vehículos farmacéuticamente aceptables adecuados para administración oral, rectal, tópica o parenteral (incluyendo inhalación, subcutánea, intraperitoneal, intramuscular e intravenosa) son conocidos por aquellos capacitados en la técnica. El vehículo debe ser farmacéuticamente aceptable en el sentido de ser compatible con los otros ingredientes de la formulación y no perjudicial para el receptor de los mismos.

25 Como se usa en este documento, los términos, "farmacéuticamente aceptable", "fisiológicamente tolerable" y variaciones gramaticales de los mismos, como se refieren a composiciones, vehículos, diluyentes y reactivos, se utilizan de manera intercambiable y significa que los materiales pueden ser administrados a o sobre un mamífero sin la producción de efectos fisiológicos indeseables.

30 Las formulaciones adecuadas para administración parenteral incluyen convenientemente una preparación acuosa estéril del compuesto activo que es preferiblemente isotónica con la sangre del receptor. Por lo tanto, tales formulaciones pueden contener convenientemente agua destilada, dextrosa al 5% en agua destilada o solución salina. Las formulaciones útiles también incluyen soluciones concentradas o sólidos que contienen el compuesto que tras la dilución con un disolvente apropiado producen una solución adecuada para administración parental como se mencionó más arriba.

35 Para la administración enteral, se puede incorporar un compuesto en un vehículo inerte en unidades discretas tales como cápsulas, tabletas, comprimidos o pastillas, conteniendo cada una una cantidad predeterminada del compuesto activo; como un polvo o gránulos; o una suspensión o solución en un líquido acuoso o en un líquido no acuoso, por ejemplo, un jarabe, un elixir, una emulsión o una poción. Los vehículos adecuados pueden ser almidones o azúcares e incluyen lubricantes, saborizantes, aglutinantes y otros materiales de la misma naturaleza.

40 Un comprimido puede fabricarse por compresión o moldeado, opcionalmente con uno o más ingredientes accesorios. Las tabletas comprimidas se pueden preparar comprimiendo en una máquina adecuada el compuesto activo en una forma de flujo libre, por ejemplo, un polvo o gránulos, opcionalmente mezclados con ingredientes accesorios, por ejemplo, aglutinantes, lubricantes, diluyentes inertes, agentes dispersantes o tensoactivos. Los comprimidos moldeados se pueden preparar moldeando en una máquina adecuada una mezcla del compuesto activo en polvo con cualquier vehículo adecuado.

45 Un jarabe o suspensión se puede fabricar añadiendo el compuesto activo a una solución acuosa concentrada de un azúcar, por ejemplo, sacarosa, a la que también se puede añadir cualquier ingrediente accesorio. Tales ingredientes accesorios pueden incluir saborizantes, un agente para retardar la cristalización del azúcar o un agente para aumentar la solubilidad de cualquier otro ingrediente, por ejemplo, como un alcohol polihídrico, por ejemplo, glicerol o sorbitol.

50 Las formulaciones para administración rectal pueden presentarse como un supositorio con un vehículo convencional, por ejemplo, manteca de cacao o Witepsol S55 (marca registrada de Dynamite Nobel Chemical, Alemania), para una base de supositorio.

55 Las formulaciones para administración oral se pueden presentar con un potenciador. Potenciadores de la absorción por vía oral aceptables incluyen agentes tensoactivos tales como lauril sulfato de sodio, palmitoil carnitina, Laureth-9, fosfatidilcolina, ciclodextrina y derivados de los mismos; sales biliares tales como desoxicolato de sodio, taurocolato de sodio, glicolato de sodio, y fusidato de sodio, agentes quelantes como EDTA, cítrico ácido y salicilatos; y ácidos grasos (por ejemplo, ácido oleico, ácido láurico, acilcarnitinas, mono y diglicéridos). Otros potenciadores de la

absorción oral incluyen cloruro de benzalconio, cloruro de bencetonio, CHAPS (3-(3-colamidopropil)-dimetilamonio-1-propansulfonato), Big-CHAPS (N,N-bis(3-D-gluconamidopropil)-colamida), clorobutanol, octoxinol-9, alcohol bencílico, fenoles, cresoles, y alcoholes de alquilo. Un potenciador de la absorción oral especialmente preferido para la presente invención es lauril sulfato de sodio.

- 5 Alternativamente, el compuesto puede ser administrado en liposomas o microesferas (o micropartículas). Los métodos para preparar liposomas y microesferas para administración a un paciente son bien conocidos por aquellos capacitados en la técnica. La patente de los Estados Unidos No. 4.789.734, describe procedimientos para encapsular materiales biológicos en liposomas. Esencialmente, el material se disuelve en una solución acuosa, se añaden los fosfolípidos y lípidos apropiados, junto con tensoactivos si se requiere, y el material dializado o sonificado, como, según se requiera. Una revisión de los métodos conocidos es proporcionada por G. Gregoriadis, Capítulo 14, "Liposomes", Drug Carriers in Biology and Medicine, páginas 287 - 341 (Academic Press, 1979).

- 10 Las microesferas formadas a partir de polímeros o proteínas son bien conocidas para aquellos capacitados en la técnica, y pueden ser adaptadas para pasar a través del tracto gastrointestinal directamente al torrente sanguíneo. Alternativamente, se puede incorporar el compuesto y las microesferas, o se puede implantar la composición de microesferas, para liberación lenta durante un periodo de tiempo que varía de días a meses. Véanse, por ejemplo, las patentes de los Estados Unidos Nos. 4.906.474, 4.925.673 y 3.625.214, y Jein, TIPS 19: 155 - 157 (1998).

- 15 En una forma de realización, el inhibidor de la tirosina quinasa de la presente invención se pueden formular en un liposoma o micropartícula, que es de tamaño adecuado para alojarse en lechos capilares después de la administración intravenosa. Cuando el liposoma o la micropartícula se aloja en los lechos capilares que rodean el tejido isquémico, los agentes se pueden administrar localmente en el sitio en el que pueden ser más eficaces. Liposomas adecuados para ser dirigidos a tejido isquémico son por lo general aproximadamente menores de 200 nanómetros y también son típicamente vesículas unilamelares, como se divulga, por ejemplo, en la patente de los Estados Unidos No. 5.593.688 de Baldeschweiler, titulada "Liposomal targeting of ischemic tissue".

- 20 Las micropartículas preferidas son aquellas preparadas a partir de polímeros biodegradables, tales como poliglicólido, poliláctida y copolímeros de los mismos. Aquellos capacitados en la técnica puede determinar fácilmente un sistema portador apropiado dependiendo de diversos factores, incluyendo la velocidad deseada de liberación del fármaco y la dosificación deseada.

- 25 En una forma de realización, las formulaciones se administran a través de catéter directamente en el interior de los vasos sanguíneos. La administración puede ocurrir, por ejemplo, a través de los agujeros en el catéter. En aquellas formas de realizaciones en las que los compuestos activos tienen una vida media relativamente larga (del orden de 1 día a una semana o más), las formulaciones se pueden incluir en hidrogeles poliméricos biodegradables, tales como los divulgados en la patente de los Estados Unidos No. 5.410.016 de Hubbell et al. Estos hidrogeles poliméricos pueden ser suministrados en el interior de un lumen de tejido y los compuestos activos liberados con el tiempo a media que el polímero se degrada. Si se desea, los hidrogeles poliméricos pueden tener micropartículas o liposomas que incluyen el compuesto activo dispersado en los mismos, proporcionando otro mecanismo para la liberación controlada de los compuestos activos.

- 30 Las formulaciones pueden presentarse convenientemente en forma de dosificación unitaria y pueden prepararse por cualquiera de los métodos bien conocidos en la técnica del campo farmacéutico. Todos los métodos incluyen la etapa de poner el compuesto activo en asociación con un vehículo que constituye uno o más ingredientes accesorios. En general, las formulaciones se preparan poniendo en contacto en forma uniforme e íntima el compuesto activo junto con un vehículo líquido o un vehículo sólido finamente dividido y, a continuación, si es necesario, se conforma el producto en la forma de dosificación unitaria deseada.

- 35 Las formulaciones pueden incluir además uno o más ingredientes accesorios opcionales utilizados en la técnica de las formulaciones farmacéuticas, por ejemplo, diluyentes, amortiguadores, agentes saborizantes, aglutinantes, agentes activos de superficie, espesantes, lubricantes, agentes de suspensión, conservantes (incluyendo antioxidantes) y similares.

- 40 Los compuestos de los presentes métodos se pueden presentar para la administración al tracto respiratorio como una sustancia para aspirar o un aerosol o solución para un nebulizador, o como un polvo microfino para insuflación, solos o en combinación con un vehículo inerte tal como lactosa. En tal caso, las partículas de compuesto activo tienen adecuadamente diámetros de menos de 50 micras, preferiblemente menos de 10 micras, más preferiblemente entre 2 y 5 micras.

- 45 En general, para administración nasal se prefiere un pH ligeramente ácido. Preferiblemente, las composiciones de la invención tienen un pH de aproximadamente 3 a 5, más preferiblemente de aproximadamente 3,5 hasta aproximadamente 3,9 y lo más preferiblemente de 3,7. El ajuste del pH se consigue mediante la adición de un ácido apropiado, tal como ácido clorhídrico.

50 La preparación de una composición farmacológica que contiene ingredientes activos disueltos o dispersados en ella se entiende bien en la técnica y no necesita estar limitada con base en la formulación. Típicamente, tales composiciones se preparan como inyectables, ya sea como soluciones o suspensiones líquidas, sin embargo,

también se pueden preparar formas sólidas adecuadas para solución, o suspensiones en líquido antes de ser usadas. La preparación también se puede emulsionar.

El ingrediente activo se puede mezclar con excipientes que sean farmacéuticamente aceptables y compatibles con el ingrediente activo y en cantidades adecuadas para uso en los métodos terapéuticos descritos en este documento.

- 5 Los excipientes adecuados son, por ejemplo, agua, solución salina, dextrosa, glicerol, etanol o similares y combinaciones de los mismos. Además, si se desea, la composición puede contener cantidades menores de sustancias auxiliares tales como agentes humectantes o emulsionantes, agentes amortiguadores del pH y similares que mejoran la eficacia del ingrediente activo.

- 10 El inhibidor de la quinasa de la presente invención puede incluir sales farmacéuticamente aceptables de los componentes en el mismo. Las sales farmacéuticamente aceptables incluyen las sales de adición ácida (formadas con los grupos amino libres del polipéptido) que se forman con ácidos inorgánicos tales como, por ejemplo, ácidos clorhídrico o fosfórico, o ácidos orgánicos tales como acético, tartárico, mandélico y similares. Las sales formadas con los grupos carboxilo libres también se pueden derivar de bases inorgánicas tales como, por ejemplo, hidróxido de sodio, potasio, amonio, calcio o férrico, y bases orgánicas tales como isopropilamina, trimetilamina, 2 - etilamino
- 15 etanol, histidina, procaína y similares.

- Los vehículos fisiológicamente tolerables son bien conocidos en la técnica. Ejemplos de vehículos líquidos son soluciones acuosas estériles que no contienen materiales además de los ingredientes activos y agua, o contienen un amortiguador tal como fosfato de sodio a un valor de pH fisiológico, solución salina fisiológica o ambos, tal como solución salina amortiguada con fosfato. Aún más, los vehículos acuosos pueden contener más de una sal
- 20 amortiguadora, así como sales tales como cloruros de sodio y de potasio, dextrosa, polietilenglicol y otros solutos.

Las composiciones líquidas también pueden contener fases líquidas además de y con la exclusión del agua. Ejemplos de tales fases líquidas adicionales son glicerina, aceites vegetales tales como aceite de semilla de algodón, y emulsiones agua-aceite.

La predicción de mutaciones

- 25 En otra forma de realización, la presente invención divulga un método para predecir las variaciones en el gen *erbB1* después del tratamiento con un inhibidor de la tirosina quinasa. Es generalmente conocido que la respuesta al tratamiento del cáncer con un inhibidor de la tirosina quinasa es seguido a menudo por la resistencia a ese u otros compuestos similares. Se cree que tal resistencia surge a través de la adquisición de mutaciones en el objetivo del fármaco, por ejemplo, en el EGFR. La capacidad para predecir (y seleccionar) tales mutaciones permitirán obtener mejores opciones de tratamiento y menos recaídas.
- 30

- En una forma de realización de la presente invención, se aísla y secuencia el ADN que codifica el dominio quinasa del EGFR a partir de una muestra tumoral de pacientes con cáncer que han respondido al gefitinib (o un tratamiento similar dirigido al EGFR) pero han recaído posteriormente. Se espera que la recaída en tales pacientes involucre la adquisición de mutaciones secundarias dentro del dominio quinasa del EGFR. Los compuestos que dirigen, e
- 35 inhiben la actividad de la quinasa de estas mutaciones recién definidas se identifican luego usando los métodos divulgados en este documento. Tales compuestos se pueden usar solos, o en combinación con otros tratamientos conocidos dirigidos al EGFR, para el tratamiento de pacientes con cáncer con mutaciones primarias o secundarias (como anteriormente) en el dominio quinasa del EGFR.

- En una forma de realización, la predicción de las variaciones en el dominio quinasa (catalítico) del EGFR (gen *erbB1*) se realiza *in vitro*. En este método, las células, por ejemplo células fibroblásticas, se transfectaron de forma estable con los ADNc que contienen mutaciones en el dominio quinasa que se han identificado en líneas celulares de cáncer humano. Por ejemplo, las células pueden ser transfectadas con un EGFR que porta una mutación tal como la SEQ ID NO: 495, descrita adicionalmente en la Figura 4A, o con cualquier cantidad de los EGFR mutados del dominio quinasa identificados o aún no identificados. La transfección de los EGFR mutados del dominio quinasa
- 40 en las células dará lugar a la proliferación aberrante de las células en cultivo. Los métodos de transfección estable son conocidos por aquellos capacitados en la técnica y se definen adicionalmente en Current Protocols in Molecular Biology por F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, K. Struhl y V. B. Chanda (Editores), John Wiley & Sons, 2004. Las células transfectadas reciben a continuación, una dosis aún no letal efectiva de una composición farmacéutica, preferiblemente un inhibidor de la tirosina quinasa, que se predice que inhibe la proliferación celular. En una forma de realización preferida, el fármaco es un anilinoquinazolina, anilinoquinazolina sintética, gefitinib o erlotinib. Las células se pasan en serie en presencia de fármacos y se seleccionan los subclones que sobreviven. Durante muchas generaciones, se seleccionan las células que sobreviven (es decir, que son resistentes al compuesto), y se analizaron las variaciones en el gen *erbB1*. Se pueden predecir luego las variaciones secundarias que se presentan después del tratamiento repetido con un inhibidor de la tirosina quinasa *in vivo*.
- 50

- 55 Alternativamente, las células se transfectan con ADNc mutante resistente a gefitinib derivado de líneas celulares de NSCLC humano, por ejemplo, NCI-1650 y NCI-1975. Cada línea celular tiene una mutación heterocigota con el dominio quinasa del EGFR, y se espera, por lo tanto, que sea sensible a gefitinib. La mutación del EGFR en NCI-1650 consiste en una supresión en el marco de 15 nucleótidos en la posición 2235 - 2249 (delLE746-A750) en el

exón 19, mientras que NCI-1975 tiene una mutación en sentido erróneo en el exón 21 que sustituye una G por T en el nucleótido 2573 (L858R). Como se muestra en este documento, la mutación L858R en NCI-H1975 se activa y confiere mayor sensibilidad al gefitinib *in vitro*. Se pueden utilizar otras líneas celulares de cáncer que albergan mutaciones del dominio quinasa del EGFR. Las líneas celulares de cáncer pueden incluir cáncer de pulmón, así como otros tipos de cáncer que se encuentra que albergan tales mutaciones.

Las células pueden ser tratadas con un mutágeno con el fin de aumentar la frecuencia con la que las células adquieren mutaciones secundarias. Un mutágeno puede inducir mutaciones en diferentes frecuencias dependiendo del régimen de dosificación, modo de administración, y la etapa de desarrollo del organismo o célula después de la administración del mutágeno, donde todos los parámetros se dan a conocer en la técnica anterior para diferentes mutágenos o técnicas de mutagénesis. El mutágeno puede ser un agente alquilante, tal como metanosulfonato de etilo (EMS), N-etil-N-nitrosourea (ENU) o N-metil-N-nitrosourea (MNU). Alternativamente, el mutágeno puede ser, por ejemplo, clorhidrato de focarbaxina (Prc), metanosulfonato de metilo (MeMS), clorambucilo (Chl), melfalán, clorhidrato de porcarbazona, ciclofosfamida (Cp), sulfato de dietilo (Et₂SO₄), monómero de acrilamida (AA), trietilén melamina (TEM), mostaza nitrogenada, vincristina, dimetilnitrosamina, N-metil-N'-nitro-nitrosoguanidina (MNNG), 7,12-dimetilbenz(a)antraceno (DMBA), óxido de etileno, hexametilfosforamida, bisulfán, y metanosulfonato de etilo (EtMs). Los métodos de tratamiento de las células con mutágenos se describen, por ejemplo, en la patente de los Estados Unidos No. 6.015.670. Después de la mutagénesis, las células (es decir, transfectadas con EGFR variante o derivado de una línea celular de cáncer humano) pueden ser cultivadas en un medio suplementado con gefitinib para seleccionar el resultado de los clones resistentes. El subcultivo de los clones individuales puede ser seguido, por ejemplo, por la determinación de la secuencia de nucleótidos del gen *EGFR* después de la amplificación específica mediada por PCR del ADN genómico correspondiente al dominio quinasa del EGFR.

En otra forma de realización, las células (con una variación del EGFR) se pasan en forma serial en presencia de concentraciones gradualmente crecientes de gefitinib (o un inhibidor similar de la tirosina quinasa) durante el transcurso de varias semanas o meses con el fin de seleccionar la adquisición espontánea de mutaciones en el gen *EGFR* que confiere resistencia a gefitinib. Las células seleccionadas (que continúan proliferando a una concentración relativamente alta de gefitinib) se pueden aislar en forma de colonias, y se identificarán las mutaciones como se describió anteriormente. Se puede predecir por lo tanto que ocurrirán tales variaciones tras el tratamiento repetido con un inhibidor de tirosina quinasa *in vivo*. Véase, por ejemplo, Scappini et al., Cancer, abril 1, 2004, vol. 100, página 1459.

En aún otra forma de realización, una forma variante del gen EGFR se puede propagar en una cepa bacteriana deficiente para la reparación del ADN antes de volver a introducirlo en líneas celulares seleccionadas de forma estable. La replicación en tales bacterias mejorará la frecuencia de mutagénesis. Alternativamente, se puede utilizar PCR "propenso a errores" para mejorar la frecuencia de mutaciones en el ADN clonado del EGFR *in vitro*, utilizando métodos estándar, conocidos por aquellos capacitados en la técnica.

En otra forma de realización, la predicción de las variaciones en el dominio quinasa del gen *erbB1* se realiza *in vivo*. Por ejemplo, se transfecta una forma variante que incrementa la actividad de la quinasa del gen *erbB1* en un animal, es decir, un ratón, generando un modelo de cáncer. El animal es tratado luego con una dosis efectiva de un compuesto, preferiblemente una anilinoquinazolina, una anilinoquinazolina sintética, gefitinib o erlotinib. Tras la exposición repetida al compuesto, se inhibe inicialmente el cáncer. Al igual que en los seres humanos tratados con tales compuestos, las células tumorales en el animal adquieren mutaciones que las hacen resistentes a tal tratamiento. Los métodos de la presente invención permiten el aislamiento y la caracterización del gen *erbB1* en tales tumores resistentes. Los compuestos que dirigen específicamente estas variaciones recién caracterizadas son útiles en el tratamiento de pacientes de los que se sospecha que portan tal gen *erbB1* mutado. Tales pacientes incluyen, por ejemplo, a los pacientes que inicialmente responden a la terapia con un inhibidor de la tirosina quinasa, pero que posteriormente no responden al mismo o a un compuesto similar.

Los métodos de creación de un modelo animal son conocidos por aquellos capacitados en la técnica y se definen con más detalle, por ejemplo, en Ohashi et al., Cell, 65: 305 - 317 (1991); Adams et al., Nature, 325: 223 - 228 (1987); Romano et al., Cell, 61: 383 - 396 (1990). En el caso de oocitos fertilizados, el método preferido para la introducción de transgenes es por microinyección, véase, por ejemplo, Leder et al., patentes de los Estados Unidos Nos. 4.736.866 y 5.175.383, mientras que en el caso de las células madre embrionarias (ES), el método preferido es la electroporación. Sin embargo, se pueden utilizar otros métodos, incluyendo los sistemas de suministro virales tales como infección retroviral, o de fusión liposomal. El aislamiento y la caracterización de ácidos nucleicos se describe anteriormente y en los ejemplos.

Las variaciones que incrementan la actividad de la quinasa identificadas anteriormente en el gen *erbB1* pueden ser cribadas en pacientes (en diagnóstico o pronóstico), utilizando los métodos de la presente invención. La presencia o ausencia de tales mutaciones puede ser utilizada entonces como un criterio para determinar la sensibilidad de aquellas al tratamiento con un compuesto dirigido al EGFR, tal como, por ejemplo, un inhibidor de tirosina quinasa.

Los compuestos que se dirigen específicamente a estas variaciones recién definidas, ya sea detectados *in vivo* o *in vitro*, se pueden seleccionar mediante técnicas conocidas en la técnica y se discuten en este documento. Se pueden utilizar ensayos de cribado de fármacos candidatos para identificar agentes candidatos bioactivos que inhiban la

actividad de las formas variantes del EGFR. De particular interés son los ensayos de cribado para agentes que tengan una baja toxicidad para las células humanas. Se puede utilizar una amplia variedad de ensayos para este propósito, incluyendo ensayos de enlazamiento proteína - proteína marcados in vitro, ensayos de cambio de movilidad electroforética, ensayos de actividad enzimática, inmunoensayos para enlazamiento de proteína, y similares. La proteína purificada del EGFR mutante también puede usarse para la determinación de la estructura tridimensional del cristal, que puede ser utilizada para el modelado de las interacciones intermoleculares, la función del transportador, etc. Tales compuestos pueden ser, por ejemplo, inhibidores de la tirosina quinasa, anticuerpos, aptámeros, los ARNpi, y vectores que inhiben la actividad de quinasa del EGFR.

En otra forma de realización, los compuestos útiles en el método de la presente invención son anticuerpos que interfieren con la señalización de la quinasa a través del EGFR mutante, incluyendo los anticuerpos monoclonales quiméricos, humanizados y recombinantes y fragmentos de los mismos que se caracterizan por su capacidad para inhibir la actividad de quinasa del EGFR y que tienen baja toxicidad.

Los anticuerpos neutralizantes se elevan rápidamente en animales tales como conejos o ratones mediante la inmunización con un EGFR con al menos una variación de ácido nucleico en su dominio quinasa. Los ratones inmunizados son particularmente útiles para proporcionar fuentes de células B para la fabricación de hibridomas, que a su vez se cultivan para producir grandes cantidades de anticuerpos monoclonales anti-EGFR. Los anticuerpos quiméricos son moléculas de inmunoglobulina caracterizadas por dos o más segmentos o porciones derivadas de diferentes especies animales. En general, la región variable del anticuerpo quimérico se deriva de un anticuerpo de mamífero no humano, tal como un anticuerpo monoclonal de murino, y la región constante de la inmunoglobulina se deriva de una molécula de inmunoglobulina humana. Preferiblemente, ambas regiones y la combinación tienen baja inmunogenicidad, determinada de forma rutinaria. Los anticuerpos humanizados son moléculas de inmunoglobulina creadas por técnicas de ingeniería genética en las que las regiones constantes del murino son reemplazadas con sus contrapartes humanas al tiempo que retienen las regiones de enlazamiento de antígeno de murino. El anticuerpo quimérico de ratón-humano resultante debería haber reducido la inmunogenicidad y mejorado la farmacocinética en humanos. Los ejemplos preferidos de anticuerpos monoclonales de alta afinidad y derivados quiméricos de los mismos, útiles en los métodos de la presente invención, se describen en la solicitud de patente europea EP 186.833; la solicitud de patente PCT WO 92/16553 y la patente de los Estados Unidos No. 6.090.923.

Los compuestos existentes o los recientemente identificados como se describió anteriormente son útiles en el tratamiento de pacientes portadores de mutaciones primarias y/o secundarias del EGFR.

En una forma de realización preferida, el compuesto es un inhibidor de la actividad de tirosina quinasa de un EGFR con un al menos una variación en su dominio quinasa, particularmente inhibidores de moléculas pequeñas que tienen acción selectiva sobre los EGFR "mutados" en comparación con otras tirosina quinasas. Los inhibidores del EGFR incluyen, pero no se limitan a, inhibidores de tirosina quinasa tales como quinazolininas, tales como PID 153035, 4-(3-cloroanilino) quinazolina o CP - 358.774, piridopirimidinas, pirimidopirimidinas, pirrolopirimidinas, tales como CGP 59326, CGP 60261 y CGP 62706 y pirazolopirimidinas, 4-(fenilamino)-7H-pirrolo[2,3-d]pirimidinas (Traxler et al., (1996) J. Med. Chem. 39: 2285 - 2292), curcumina (diferuloil metano) (Laxmin arayana, et al., (1995), Carcinogen 16: 1741 - 1745), 4,5-bis(4-fluoroanilino)ftalimida (Buchdunger et al. (1995) Clin. Cancer Res. 1: 813 - 821; Dinney et al. (1997) Clin. Cancer Res. 3: 161 - 168); tirfostinas que contienen fracciones de nitrotofeno (Brunton et al. (1996) Anti Cancer Drug Design 11: 265 - 295), el inhibidor de proteína quinasa ZD-1 839 (AstraZeneca), CP - 358.774 (Pfizer, Inc.); PD-01 83805 (Warner-Lambert), EKB-569 (Torrance et al., Nature Medicine, Vol. 6, No. 9, septiembre de 2000, página 1024), HKI-272 y HKI-357 (Wyeth); o como se describe en la solicitud internacional de patente WO 99/09016 (American Cyanamid); WO 98/43960 (American Cyanamid); WO 97/38983 (Warener Labert); WO 99/06378 (Warner Lambert); WO 99/06396 (Warner Lambert); WO 96/30347 (Pfizer, Inc.); WO 96/33978 (Zeneca); WO 96/33977 (Zeneca), y WO 96/33980 (Zeneca).

En otra forma de realización, se puede utilizar una estrategia antisentido para interferir con la actividad de quinasa de un EGFR variante. Este enfoque puede, por ejemplo, utilizar ácidos nucleicos antisentido o ribozimas que bloquean la traducción de un ARNm específico, ya sea mediante el enmascaramiento de ese ARNm con un ácido nucleico antisentido o escindiéndolo con una ribozima. Para una discusión general de la tecnología antisentido, véase, por ejemplo, Antisense DNA y RNA, (Cold Spring Harbor Laboratory, D. Melton, ed., 1988).

La inhibición corta reversible de la transcripción del gen EGFR de la variante, también puede ser útil. Tal inhibición se puede lograr mediante el uso de los ARNpi. La tecnología de ARN de interferencia (ARNi) evita la expresión de genes mediante el uso de moléculas pequeñas de ARN, tales como los ARN pequeños de interferencia (ARNpi). Esta tecnología a su vez toma ventaja del hecho de que el ARNi es un mecanismo biológico natural para silenciar genes en la mayoría de las células de muchos organismos vivos, de plantas a insectos a mamíferos (McManus et al., Nature Reviews Genetics, 20023, (10) página 737). El ARNi evita que un gen produzca una proteína funcional al garantizar que la molécula intermedia, la copia del ARN mensajero del gen se destruye. Los ARNpi pueden ser utilizados en una forma simple e incorporarse en un vector, tal como se describe a continuación. Se puede hacer uso además de aptámeros para inhibir específicamente la transcripción del gen EGFR de la variante, véase, por ejemplo, la patente de los Estados Unidos No. 6.699.843. Los aptámeros útiles en la presente invención pueden ser identificados usando el proceso SELEX. Los métodos de SELEX han sido descritos, por ejemplo, en las patentes de los Estados Unidos Nos. 5.707.796, 5.763.177, 6.011.577, 5.580.737, 5.567.588, y 5.660.985.

Un "ácido nucleico antisentido" u "oligonucleótido antisentido" es una molécula de ácido nucleico monocatenario, que, por hibridación en condiciones citoplasmáticas con bases complementarias en una molécula de ARN o ADN, inhibe la función de esta última. Si el ARN es un transcrito de ARN mensajero, el ácido nucleico antisentido es un contratranscrito o un ácido nucleico complementario que interfiere con el ARNm. Tal como se utiliza en la actualidad, "antisentido" incluye ampliamente las interacciones ARN-ARN, las interacciones ARN-ADN, ribozimas, ARNi, aptámeros y a captura mediada por ARNasa-H.

Los ribozimas son moléculas de ARN que poseen la capacidad de escindir específicamente otras moléculas de ARN monocatenarias de una manera algo análoga a las endonucleasas de restricción de ADN. Los ribozimas se descubrieron a partir de la observación de que ciertos ARNm tienen la capacidad para escindir sus propios intrones. Mediante la modificación de la secuencia de nucleótidos de estos ribozimas, los investigadores han sido capaces de modificar por ingeniería genética moléculas que reconocen secuencias de nucleótidos específicos en una molécula de ARN y escindir las (Cech, 1989, Science 245 (4915) página 276). Debido a que son específicos de la secuencia, sólo los ARNm con secuencias particulares se inactivan.

Las moléculas de ácido nucleico antisentido puede ser codificadas por un gen recombinante para la expresión en una célula (por ejemplo, las patentes de los Estados Unidos Nos. 5.814.500; 5.811.234), o, alternativamente, se pueden preparar sintéticamente (por ejemplo, la patente de los Estados Unidos No. 5.780.607).

La presente invención proporciona además métodos de tratamiento de pacientes con cáncer. En particular, los pacientes con al menos una variación de ácido nucleico en el dominio quinasa del EGFR. El método de tratamiento comprende la administración de una composición que contiene ARNpi a un paciente dentro de una ventana de tiempo apropiada. Los ARNpi pueden ser sintetizados químicamente, producidos usando transcripción in vitro, etc. Además, la molécula de ARNpi se puede personalizar para pacientes individuales en una forma tal que corresponda en forma precisa a la mutación identificada en su tumor. Ya que ARNpi puede discriminar entre secuencias de nucleótidos que difieren por sólo un único nucleótido, es posible diseñar ARNpi que dirijan únicamente una forma mutante del gen EGFR que está asociada, ya sea con una única sustitución de nucleótidos o una pequeña supresión de varios nucleótidos, ambos han sido identificados en tumores como se describe en este documento. Los ARNpi han sido descritos en 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, McCaffrey et al. (2002), Nature, 418: 38 - 39; Xia H. et al. (2002), Nat. Biotech. 20: 1006 - 1010, Novina et al. (2002), Nat. Med. 8: 681 - 686, y la solicitud de patente de los Estados Unidos No. 20030198627.

Una ventaja importante de tal estrategia terapéutica en relación con el uso de fármacos tales como gefitinib, que inhiben tanto el receptor mutado como el receptor normal, es que ARNpi dirigido específicamente contra el EGFR mutado no debe inhibir el EGFR de tipo silvestre. Esto es significativo porque en general se cree que los "efectos secundarios" del tratamiento con gefitinib, que incluyen diarrea y dermatitis, son una consecuencia de la inhibición del EGFR en tejidos normales que requieren de su función.

El suministro de ARNpi a los tumores potencialmente se puede lograr a través de cualquiera de los diferentes "vehículos" de suministro de genes que están actualmente disponibles. Estos incluyen vectores virales, tales como adenovirus, lentivirus, virus del herpes simple, virus vacuna, y retrovirus, así como sistemas de suministro de genes mediada por productos químicos (por ejemplo, liposomas), o sistemas de suministro mecánico de ADN (pistolas de ADN). Los oligonucleótidos que van a ser expresados por tal inhibición mediada por ARNpi de la expresión génica estarían entre 18 y 28 nucleótidos de longitud.

En otra forma de realización, los compuestos son moléculas antisentido específicas para secuencias humanas que codifican para un EGFR que tiene al menos una variación en su dominio quinasa. El agente terapéutico administrado pueden ser oligonucleótidos antisentido, particularmente oligonucleótidos sintéticos, que tienen modificaciones químicas de los ácidos nucleicos nativos, o constructos de ácido nucleico que expresan tales moléculas antisentido como ARN. La secuencia antisentido es complementaria al ARNm de los genes EGFR objetivo, e inhibe la expresión de los productos génicos objetivo (véase, por ejemplo Nyce et al. (1997) Nature 385: 720). Las moléculas antisentido inhiben la expresión génica mediante la reducción de la cantidad de ARNm disponible para la traducción, a través de la activación de ARNasa H o impedimento estérico. Se puede administrar una o una combinación de moléculas antisentido, donde una combinación puede comprender múltiples secuencias diferentes de un solo gen objetivo, o secuencias que complementan varios genes diferentes.

Un gen objetivo preferido es un EGFR con al menos una variación de ácido nucleico en su dominio quinasa. La secuencia del gen se incorpora en este documento, tal como, por ejemplo, en la Figura 5. En general, la secuencia antisentido tendrá la misma especie de origen que el huésped animal.

Las moléculas antisentido pueden ser producidas por expresión de la totalidad o de una parte de la secuencia del gen objetivo en un vector apropiado, donde se introduce el vector y se expresa en las células objetivo. La iniciación de la transcripción se orientará de tal manera que la cadena antisentido se produce como una molécula de ARN.

El ARN antisentido hibrida con el ARNm de cadena sentido endógena, bloqueando de este modo la expresión del gen objetivo. Se pueden emplear la región de iniciación transcripcional nativa, o una región de iniciación

transcripcional exógena. El promotor puede introducirse por métodos recombinantes in vitro, o como el resultado de la integración homóloga de la secuencia en un cromosoma. Se conocen en el arte muchos promotores fuertes que son activos en células musculares, incluyendo el promotor O-actina, promotores tardíos y tempranos de SV40, un promotor de citomegalovirus humano, LTR retrovirales, etc. Los vectores de transcripción generalmente tienen sitios de restricción convenientes localizados cerca de la secuencia promotora para prever la inserción de secuencias de ácido nucleico. Los casetes de transcripción se pueden preparar para que comprendan una región de iniciación de la transcripción, el gen objetivo o un fragmento del mismo, y una región de terminación transcripcional. Los casetes de transcripción se pueden introducir en una variedad de vectores, por ejemplo un plásmido; un retrovirus, por ejemplo lentivirus; adenovirus, y similares, donde los vectores son capaces de mantenerse en forma transitoria o estable en las células, por lo general por un período de al menos aproximadamente un día, más usualmente durante un período de al menos aproximadamente varios días.

Los aptámeros son también útiles. Los aptámeros son una nueva clase prometedora de oligonucleótidos o péptidos terapéuticos y se seleccionan in vitro para enlazarse específicamente a un objetivo determinado con alta afinidad, tales como por ejemplo receptores del ligando. Sus características de enlazamiento son probablemente un reflejo de la capacidad de los oligonucleótidos para formar estructuras tridimensionales que se mantienen unidas por el apareamiento intramolecular de nucleobases. Los aptámeros son secuencias sintéticas de ADN, ARN o de péptidos que pueden ser normales y modificadas (por ejemplo, ácido péptido nucleico (PNA), ADN tiofosforilado, etc.) que interactúan con una proteína, ligando objetivo (lípidos, carbohidratos, metabolitos, etc.) En una forma de realización adicional, se pueden introducir o expresar en aptámeros de ARN específicos para un EGFR variante en una célula como agente terapéutico.

Los ácidos péptido nucleicos (PNA) son compuestos que en algunos aspectos son similares a oligonucleótidos y sus análogos y por lo tanto pueden imitar ADN y ARN. En PNA, la cadena principal de desoxirribosa de oligonucleótidos ha sido reemplazada por una cadena principal pseudopeptídica (Nielsen et al. 1991 Science 254, 1457 - 1500). Cada subunidad, o monómero, tiene una nucleobase de origen natural o de origen no natural unida a esta cadena principal. Una de tales cadenas principales está construida de unidades de repetición de N-(2-aminoetil)glicina enlazada a través de enlaces de amida. PNA se hibrida con ácidos nucleicos complementarios a través del apareamiento de bases de Watson y Crick y la formación de la hélice. La cadena principal de pseudopéptido proporciona propiedades de hibridación superiores (Egholm et al. Nature (1993) 365, 566 - 568), resistencia a la degradación enzimática (Demidov et al. Biochem. Pharmacol. (1994) 48, 1310 - 1313) y acceso a una variedad de modificaciones químicas (Nielsen y Haaima Chemical Society Reviews (1997) 73 - 78). Los PNA específicos para un EGFR variante pueden ser introducidos o expresados en una célula como agente terapéutico. Los PNA han sido descritos, por ejemplo, en la solicitud de patente de los Estados Unidos No. 20040063906.

Los pacientes a ser tratados con un compuesto que se dirige a una variante del EGFR incluyen, por ejemplo, los pacientes diagnosticados con una mutación primaria o secundaria en su EGFR, los pacientes que inicialmente responden a la terapia con inhibidor de tirosina quinasa, pero que posteriormente no responden al mismo compuesto o a uno similar. Alternativamente, los compuestos que se dirigen a las mutaciones del EGFR secundarios pueden ser administrados a pacientes con cáncer en combinación con compuestos que se dirigen a las mutaciones primarias del EGFR, por ejemplo, gefitinib, como una terapia de combinación. Mediante la combinación de compuestos que se dirigen tanto a mutaciones primarias como secundarias del EGFR, se reducirá la probabilidad de resistencia.

Mutaciones adicionales del EGFR que confieren resistencia a terapéuticos contra el cáncer actualmente conocidos, incluyendo pero sin limitarse a inhibidores de tirosina quinasa del EGFR gefitinib, erlotinib y similares, están dentro del alcance de la invención. Se predice que mutantes resistentes del EGFR tienen mutantes análogos a los mutantes identificados en los dominios quinasa de proteínas que contienen un dominio relacionado de tirosina quinasa que tienen alta homología en esta región quinasa. Los artículos que describen mutaciones en proteínas análogas incluyen aquellos conocidos en la técnica para BCR-ABL. Véase, por ejemplo, Bradford et al. Blood. 2003 1 de julio; 102 (1): 276 - 83, Epub 2003, 06 de marzo; Hochhaus et al., Leukemia. 2002 Nov. 16 (11): 2190 - 6; y Al-Ali et al., J. Hematol 2004; 5 (1): 55 - 60.

Un EGFR mutante resistente a los inhibidores conocidos de la tirosina quinasa del EGFR incluye uno o más polipéptidos del EGFR, o un nucleótido que los codifica, con un residuo de tipo no silvestre en una o más posiciones análogas a los residuos c-abl (BCR-ABL) que confirman un fenotipo resistente al imatinib. Los residuos que cuando mutan en EGFR confieren resistencia a los fármacos incluyen especialmente aquellos residuos del dominio quinasa, incluyendo pero sin limitarse, por ejemplo, al bucle P y el bucle de activación, en donde los residuos mutados en el polipéptido del EGFR son análogos a los residuos c-abl. Los mutantes contemplados resistentes del EGFR tienen residuos de tipo no silvestre en las posiciones de aminoácidos que son análogos a al menos las posiciones Met 244, Leu 248, Gly 250, Gln 252, Tyr 253, Glu 255, Asp 276, Thr 315, Phe 317, Met 351, Glu 355, Phe 359, Su 396, Ser 417, y Phe 486 de BCR-ABL, véase, por ejemplo, la Tabla S3C y la FIG. 9. Estos residuos BCL-ABL corresponden a los residuos de 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, respectivamente, en EGFR. Véase, por ejemplo, la Tabla S3C, FIG. 9.

Pruebas de pronóstico

Los métodos de la presente invención se utilizan como un indicador de pronóstico del desarrollo de cáncer. Alternativamente, los métodos se utilizan para detectar el cáncer que está presente, pero aún no ha sido diagnosticado o está en una etapa que es indetectable. Los pacientes en situación de riesgo de desarrollar cáncer son seleccionados, utilizando los métodos de la presente invención, por la presencia de actividad de la quinasa que incrementa la variación de ácido nucleico en el gen *erbB1*. La presencia de una variación o variaciones en el dominio quinasa del gen *erbB1* indica la presencia o la presencia inminente de cáncer. Por lo tanto, la presencia de variaciones en el dominio quinasa del gen *erbB1* sugiere que un paciente se beneficiaría de un tratamiento dirigido al EGFR. Como se describe en este documento, un tratamiento dirigido al EGFR es preferentemente un tratamiento con un inhibidor de la tirosina quinasa.

En una forma de realización preferida de la presente invención, se selecciona un paciente por la presencia o ausencia de variaciones de ácido nucleico en el dominio quinasa del gen *erbB1* mediante la obtención de una muestra biológica. La muestra puede ser cualquier muestra del paciente, incluyendo el tejido, por ejemplo, de la lengua, boca, mejilla, tráquea, tubo bronquial, pulmones, etc. o fluido, por ejemplo, a partir de esputo o aspirado de pulmón. Los métodos de obtención de estos especímenes biológicos son bien conocidos por aquellos capacitados en la técnica.

Por lo tanto, la invención proporciona un método para identificar una enfermedad o trastorno asociado con la expresión o actividad del EGFR mutante aberrante en la cual se obtiene una muestra de ensayo de un individuo y se detecta la proteína del EGFR mutante o el ácido nucleico (por ejemplo, ARNm, ADN genómico), en donde la presencia de la proteína del EGFR mutante o el ácido nucleico es diagnóstica para un individuo que tiene o está en riesgo de desarrollar una enfermedad o trastorno asociado con la expresión o actividad del EGFR mutante aberrante. Como se usa en este documento, una "muestra de ensayo" se refiere a una muestra biológica obtenida de un individuo de interés. Por ejemplo, una muestra de ensayo puede ser un fluido biológico (por ejemplo, suero), muestra celular, o tejido, especialmente una muestra de biopsia de tejido.

Por otra parte, los ensayos de pronóstico descritos en este documento pueden ser utilizados para determinar si se le puede administrar a un individuo un agente (por ejemplo, un agonista, antagonista, peptidomimético, proteína, péptido, ácido nucleico, molécula pequeña, u otro fármaco candidato) para tratar una enfermedad o trastorno asociado con la expresión o actividad del EGFR mutante aberrante. Por ejemplo, se pueden utilizar tales métodos para determinar si un individuo puede ser tratado efectivamente con un agente para un trastorno. Por lo tanto, la invención proporciona métodos para determinar si un individuo puede ser tratado efectivamente con un agente para un trastorno asociado con la expresión o actividad del EGFR mutante aberrante en donde se obtiene una muestra de ensayo y se detecta la proteína o el ácido nucleico del EGFR mutante (por ejemplo, en donde la presencia de la proteína o el ácido nucleico del EGFR mutante es diagnóstica para un individuo al que se le puede administrar el agente para tratar un trastorno asociado con la expresión o actividad del EGFR mutante).

Ejemplos

Ejemplo 1

Análisis de la secuencia de nucleótidos de especímenes tumorales

Los especímenes tumorales de procedimientos quirúrgicos o de diagnóstico inicial se obtuvieron de pacientes con NSCLC que fueron posteriormente tratados con Gefitinib, en virtud de un protocolo aprobado por el IRB. Especímenes tumorales congelados, junto con el tejido normal correspondiente, estaban disponibles para cuatro casos, y se utilizó material embebido en parafina para los especímenes restantes. Además, se obtuvieron 25 casos no seleccionados de NSCLC primario (15 broncoalveolares, 7 de adenocarcinoma, y 3 de cánceres de pulmón de células grandes), con los tejidos normales correspondientes, del banco de tumores del Hospital General de Massachusetts. Para el análisis mutacional de toda la secuencia de codificación del EGFR, se extrajo ADN de los especímenes, seguido por la amplificación de todos los 28 exones, secuenciación automatizada de los fragmentos de PCR no clonados, y el análisis de electroferogramas, tanto en dirección sentido como antisentido para la presencia de mutaciones heterocigotas. Se confirmaron todas las variantes de la secuencia por medio de amplificaciones independientes múltiples de la PCR. Las secuencias del cebador y las condiciones de amplificación se suministran en la sección Material complementario. Se buscaron también mutaciones del EGFR en los exones 19 y 21 en tumores primarios de mama (15 casos), de colon (20 casos), de riñón (16 casos), y de cerebro (4 casos), junto con un panel de 78 líneas celulares de cáncer que representan diversas histologías (enumeradas a continuación).

Análisis funcional de constructos del EGFR mutante

Se introdujeron las mutaciones L858R y delL747-P753insS en la secuencia de codificación de longitud completa del EGFR usando mutagénesis dirigida al sitio y se insertaron en un constructo de expresión conducido por citomegalovirus (pUSE, Upstate). Se transfectaron células Cos-7 (Lipofectamine 2000, Invitrogen), utilizando 1 µg de los constructos de expresión, seguido después de 18 horas por una resiembra a razón de 5 x10⁴ células/pozo (placas de 12 pozos, Costar) en DMEM que carecen de suero de ternera fetal. Después de 16 horas de privación de suero, se estimularon las células con 10 ng/ml de EGF (SIGMA). Para demostrar la inhibición de Gefitinib, se añadió

el fármaco al medio de cultivo 3 horas antes de la adición del EGF (30 min de estimulación con 100 ng/ml del EGF). Se prepararon los lisados celulares en 100 µL de amortiguador de lisis de Laemmli, seguido por la resolución de proteínas en SDS-PAGE al 10%, transferencia a membranas de PVDF, y análisis de transferencia tipo Western usando reactivo de quimioluminiscencia mejorado (Amersham). Se midió la autofosforilación del EGFR utilizando anticuerpo para fosfotirosina Y-1068, y se mostró la expresión comparable de proteína utilizando anticuerpo anti-EGFR (concentración de trabajo de 1:100; Cell Signaling Technology).

Análisis mutacional

Se utilizó la reacción en cadena de la polimerasa para amplificar los 28 exones que contenían el gen EGFR utilizando ADN aislado de tejido tumoral primario o líneas de células derivadas de tumores. Los pares de cebadores utilizados fueron: Exón 1, CAGATTTGGCTCGACCTGGACATAG (sentido) (SEQ ID NO: 513) y CAGCTGATCTCAAGGAAACAGG (antisentido) (SEQ ID NO: 514); Exón 2, GTATTATCAGTCAC TAAAGCTCAC (sentido) (SEQ ID NO: 515) y CACACTTCAAGTGAATTCTGC (SEQ ID NO: 516); Exón 3, CTCGTG TGCATTAGGGTTCAACTGG (sentido) (SEQ ID NO: 517) y CCTTCTCCGAGGTGGAATTGAGTGAC (antisentido) (SEQ ID NO: 518); Exón 4, GCTAATTGCGGACTCTTGTTCCGAC (sentido) (SEQ ID NO: 519) y TACATGC TTTTCTAGTGGTCAG (antisentido) (SEQ ID NO: 520); Exón 5, GGTCTCAAGTGATTCTACAAACCAG (sentido) (SEQ ID NO: 521) y CCTTCACCTACTGGTTCACATCTG (antisentido) (SEQ ID NO: 522); Exón 6, CATGGT TTGACTTAGTTGAATGTGG (sentido) (SEQ ID NO: 523) y GGATACTAAAGATACTTTGTCAC CAGG (antisentido) (SEQ ID NO: 524); Exón 7, GAACACTAGGCTGCAAAGACAGTAAC (sentido) (SEQ ID NO: 525) y CCAAGCAAGGCAAACACATCCACC (antisentido) (SEQ ID NO: 526); Exón 8, GGAGGATGGAGCC TTTCCATCAC (sentido) (SEQ ID NO: 527) y GAAGAGGAAGATGTGTTCTTTGG (antisentido) (SEQ ID NO: 528); Exones 9 y 10, GAATGAAGGATGATGTGGCAGTGG (sentido) (SEQ ID NO: 529) y CAAAACATCAGCC ATTAACGG (antisentido) (SEQ ID NO: 530); Exón 11, CCACTTACTGTTCATATAATACAGAG (sentido) (SEQ ID NO: 531) y CATGTGAGATAGCATTTGGGAATGC (antisentido) (SEQ ID NO: 532); Exón 12, CATGACCT ACCATCATTGGAAAGCAG (sentido) (SEQ ID NO: 533) y GTAATTTACAGTTAGGAATC (sentido) (SEQ ID NO: 534); Exón 13, GTCACCCAAGGTCATGGAGCACAGG (sentido) (ID SEC NO: 535) y CAGAATGC CTGTAAAGCTATAAC (antisentido) (SEQ ID NO: 536); Exón 14, GTCCTGGAGTCCCAACTCCTTGAC (sentido) (SEQ ID NO: 537) y GGAAGTGGCTCTGA TGGCCGTCCTG (antisentido) (SEQ ID NO: 538); Exón: 15, CCAC TCACACACACTAAATATTTTAAG (sentido) (SEQ ID NO: 539) y GACCAAAACACCTTAAGTAA CTGACTC (antisentido) (SEQ ID NO: 540); Exón 16, CCAA TCCAACATCCAGACACATAG (sentido) (SEQ ID NO: 541) y CCAGAGCCATAGAACTTGATCAG (antisentido) (SEQ ID NO: 542); Exón 17, GTATGGACTATGGC ACTTCAATTGCATGG (sentido) (SEQ ID NO: 543) y CCAGAGAACATGGCAACCAGCACAGGAC (antisentido) (SEQ ID NO: 544); Exón 18, CAAATGAGCTGGCAAGTGCCGTGTC (sentido) (SEQ ID NO: 545) y GAGTTT CCCAAACACTCAGTGAAAC (antisentido) (SEQ ID NO: 546) o CAAGTGCCGTGTCCTGGCACCCAAGC (sentido) (SEQ ID NO: 675) y CCAAACACTCAGTGAAACAAAGAG (antisentido) (SEQ ID NO: 676); Exón 19, GCAATATCAGCC TTAGG TGC GGCTC (sentido) (SEQ ID NO: 547) y CATAGAAAGTGAACATTTAGGATGTG (antisentido) (SEQ ID NO: 548); Exón 20, CCATGAGTACGTATTTTGAAGCTC (sentido) (SEQ ID NO: 549) y CATATCC CCATGGC AAACCTTTGC (antisentido) (SEQ ID NO: 550); Exón 21, CTAACGTTCCGCCAGCCATAAGTCC (sentido) (SEQ ID NO: 551) y GCTGCGAGCTACCCAGAATGTCTGG (antisentido) (SEQ ID NO: 552); Exón 22, GACGGG TCCTGGGGTGATCTGGCTC (sentido) (SEQ ID NO: 553) y CTCAGTACAATAGATAGACAGCAATG (antisentido) (SEQ ID NO: 684); Exón 23, CAGGACTACAGAAATGTAGGTTTC (sentido) (SEQ ID NO: 555) y GTGCCTG CCTTAAGTAATGTGATGAC (antisentido) (SEQ ID NO: 556); Exón 24, GACTGG AAGTGTGCGA TCACCAATG (sentido) (SEQ ID NO: 557) y GGTTTAATAATGCGATCTGGGACAC (antisentido) (SEQ ID NO: 558); Exón 25, GCAGCTATAATTTAGAGAACCAAGG (sentido) (ID SEC N °: 559) y GGTT AAAATTGACTTC ATTTCCATG (antisentido) (SEQ ID NO: 560); Exón 26, CTTAGTTGCTCTAAA ACTAAGC (sentido) (SEQ ID NO: 561) y CTGTGAGGCGTGACAGCCGTGCAG (antisentido) (SEQ ID NO: 562); Exón 27, CAACCTACTAATCAG AACCAGCATC (sentido) (SEQ ID NO: 563) y CCTTCACTGTGTCTGC AAATCTGC (antisentido) (SEQ ID NO: 564); Exón 28, CCTGTCATAAGTCTCCTTGTGAG (sentido) (SEQ ID NO: 565) y CAGTCTGTGGGTCTAAG AGCTAATG (antisentido) (SEQ ID NO: 566). Las temperaturas de hibridación fueron 58 °C (exones 1, 3, 4, 7 - 10, 12 - 25, 27, y 28), 56 °C (exones 2, 5, 6, y 26), o 52 °C (exón 11).

La amplificación por PCR anidada del ADN extraído de tejido tumoral de archivo se llevo a cabo de la siguiente manera. Se generó un PCR inicial para los exones 2, 5, 6, 7, 11, 12, 14, 16, 18, 19, 20, 21, 23, 24, 25, 26, y 27, utilizando los cebadores y condiciones descritas anteriormente. Posteriormente, se amplificaron 2 µl de esta reacción en una PCR secundaria utilizando los siguientes pares de cebadores internos: Exón 2, CAGGAATGGGTGAGTCTCTGTGTG (sentido) (SEQ ID NO: 567) y GTGGAATTCTGCCAGGCCTTTC (antisentido) (SEQ ID NO: 568); Exón 5, GATTCTACAAACCA GCCAGCCAAAC (sentido) (SEQ ID NO: 569) y CCTACTGGTTCACATCTGACCTG (antisentido) (SEQ ID NO: 570); Exón 6, GTTTGAATGTGGTTTCGTTGGAAG (sentido) (SEQ ID NO: 571) y CTTTGTACACAGG CAGAGG GCAATATC (antisentido) (SEQ ID NO: 572); Exón 7, GACAGTAACTTGGGCTTTCTGAC (sentido) (SEQ ID NO: 573) y CATCCACCCAAAGACTCTCCAAG (antisentido) (SEQ ID NO: 574); Exón 11, CTGTTTATA TAATAC AGAGTCCCTG (sentido) (SEQ ID NO: 575) y GAGAGATGCAGGAGCTCTGTGC (antisentido) (SEQ ID NO: 576); Exón 12, GCAGTTTGTAGTCAATCAAAGGTGG (sentido) (SEQ ID NO: 577) y GTAATTTAAATGGGAAT AGCCC (antisentido) (SEQ ID NO: 578); Exón 14, CAACTCCTTGACATTACCTCAAG (sentido) (SEQ ID NO: 579) y GATGGCCGTCTGCCACACAGG

(antisentido) (SEQ ID NO: 580); Exón 16, GAGTAGTTTAGCA TATATTGC (sentido) (SEQ ID NO: 581) y GACAGTCAGAAATGCAGGAAAGC (antisentido) (SEQ ID NO: 582); Exón 18, CAAGTGCCGTGTCCTGGCACCCAAGC (sentido) (SEQ ID NO: 583) y CCAAACACTCA GTGAAACAAAGAG (antisentido) (SEQ ID NO: 584) o GCACCCAAGCCCATGCCGTGGCTGC (sentido) (SEQ ID NO: 677) y
 5 GAAACAAAGAGTAAAGTAGATGATGG (antisentido) (SEQ ID NO: 678); Exón 19, CCTTAGGTGCGGCTCCACAGC (sentido) (SEQ ID NO: 585) y CATTAGGATGTGGAGATGAGC (antisentido) (SEQ ID NO: 586); Exón 20, GAAACTCAAG ATCGCATTATGC (sentido) (SEQ ID NO: 587) y GCAAACCTTTGCTATCCCAGGAG (antisentido) (SEQ ID NO: 588); Exón 21, CAGCCATAAGTCCTCGACGTGG (sentido) (SEQ ID NO: 589) y CATCCTCCCCT GCATGTGTAAAC (antisentido) (SEQ ID NO: 590); Exón 23, GTAGGTTTCTAAACATCAAGAAAC (sentido) (SEQ
 10 ID NO: 591) y GTGATGACATTTCTCCAGGGATGC (antisentido) (SEQ ID NO: 592); Exón 24, CATCACCA ATGCCTTCTTTAAGC (sentido) (SEQ ID NO: 593) y GCTGGAGGGTTTAATAATGCGATC (antisentido) (SEQ ID NO: 594); Exón 25, GCAAACACACAGGCACCTGCTGGC (sentido) (SEQ ID NO: 595) y CATTTC CATGTGAGTTTCACTAGATGG (antisentido) (SEQ ID NO: 596); Exón 26, CACCTTCACAATATACCCTCCATG (sentido) (SEQ ID NO: 679) y GACAGCCGTGCAGGAAAAACC (antisentido) (SEQ ID NO: 680); Exón 27,
 15 GAACCAGCATCTCAAGGAGATCTC (sentido) (SEQ ID NO: 681) y GAGCACCTGGCTTGGACACTGGAG (antisentido) (SEQ ID NO: 682).

Las amplificaciones por PCR anidada para el resto de los exones consistió de PCR primaria utilizando los siguientes cebadores. Exón 1, GACCGACGACAGGCCACCTCGTC (sentido) (SEQ ID NO: 597) y GAAGAACGAAACGTCCCGTTCTCC (antisentido) (SEQ ID NO: 598); Exón 3, GTTGAGCACT
 20 CGTGTGCATTAGG (sentido) (SEQ ID NO: 599) y CTCAGTGACGTGTACTGGGTA (antisentido) (SEQ ID NO: 600); Exón 4, GTTCACTGGGCTAATTGCGGACTCTTGTTCGCAC (sentido) (SEQ ID NO: 601) y GGTA AATACATGCTTTTCTAGTGGTCAG (antisentido) (SEQ ID NO: 602); Exón 8, GGAGGATGGA GCCTTTCCATCAC (sentido) (SEQ ID NO: 603) y GAAGAGGAAGATGTGTTCTTTGG (antisentido) (SEQ ID NO: 604); Exón 9, GAATGAAGGATGATGTGGCAGTGG (sentido) (SEQ ID NO: 605) y GTATGTGTGAAGGAG TCACTGAAAC
 25 (antisentido) (SEQ ID NO: 606); Exón 10, GGTGAGTCACAGGTTTCAGTTGC (sentido) (SEQ ID NO: 607) y CAAAACATCAGCCATTAACGG (antisentido) (SEQ ID NO: 608); Exón 13, GTAGCCAGCATGTC TGTGTAC (sentido) (SEQ ID NO: 609) y CAGAATGCCTGTAAAGCTATAAC (antisentido) (SEQ ID NO: 610); Exón 15, CATTTGGCTTTCCCCACTCACAC (sentido) (SEQ ID NO: 611) y GACCAAAACACCTTAA GTAAGTACTC
 30 (antisentido) (SEQ ID NO: 612); Exón 17, GAAGCTACATAGTGTCTCACTTTCC (sentido) (SEQ ID NO: 613) y CACAACCTGCTAATGGCCCGTTCTCG (antisentido) (SEQ ID NO: 614); Exón 22, GAGCAGCCCTGAACTCCGTCAGACTG (sentido) (SEQ ID NO: 683) y CTCAGTACAATAGATAGACAGCAATG (antisentido) (SEQ ID NO: 684); Exón 28a, GCTCC TGCTCCCTGTCATAAGTC (sentido) (SEQ ID NO: 615) y
 GAAGTCTGCTGGTAGTCAGGGTTG (antisentido) (SEQ ID NO: 616); Exón 28b, CTGCAGTGGCAACCCGAGTATC (sentido) (SEQ ID NO: 617) y CAGTC TGTGGGTCTAAGAGCTAATG
 35 (antisentido) (SEQ ID NO: 618). La amplificación secundaria por PCR se realizó utilizando pares de cebadores: Exón 1, GACAGGCCACCTCGTCGCGTC (sentido) (SEQ ID NO: 619) y CAGCTGATCTCAAGGAAACAGG (antisentido) (SEQ ID NO: 620); Exón 3, CTCGTG TGCATTA GGGTTCACTGG (sentido) (SEQ ID NO : 621) y CCTTCTCCGAGGTGGAATTGAGTGAC (antisentido) (SEQ ID NO: 622); Exón 4, GCTAATTGCGGGACTCTTGTTCGCAC (sentido) (SEQ ID NO: 623) y TACATGCTTT TCTAGTGGTCAG
 40 (antisentido) (SEQ ID NO: 624); Exón 8, CCTTTCCATCACCCCTCAAGAGG (sentido) (SEQ ID NO: 625) y 'GATGTGTTCCCTTTGGAGGTGGCATG (antisentido) (SEQ ID NO: 626); Exón 9, GATGTGG CAGTGGCGGTTCCGGTG (sentido) (SEQ ID NO: 627) y GGAGTCACTGAAACAAACAACAGG (antisentido) (SEQ ID NO: 628); Exón 10, GGTTCAAGTTGCTGTATAAAG (sentido) (SEQ ID NO: 629) y CCATTAACGGT AAAATTTTCAAG (antisentido) (SEQ ID NO: 630); Exón 13, CCAAGGTCATGGAGCACAGG (sentido) (SEQ ID
 45 NO: 631) y CTGTAAAGCTATAACAACAACCTGG (antisentido) (SEQ ID NO: 632); Exón 15, CCACTCACA CACTAAATATTTTAAG (sentido) (SEQ ID NO: 633) y GTAAGTACTCAAATACAAACCAC (antisentido) (SEQ ID NO: 634); Exón 17, GAAGCTACATAGTGTCTCACTTTCC (sentido) (SEQ ID NO: 635) y CCAA CTGCTAATGGCCCGTTCTCG (antisentido) (SEQ ID NO: 636); Exón 22, GACGGGTCCTGGGGTGATCTGGCTC
 50 (sentido) (SEQ ID NO: 685) y CTCAGTACAATAGATAGACAGCAATG (antisentido) (SEQ ID NO: 686); Exón 28a, CCTGTCATAAG TCTCCTTGTGAG (sentido) (SEQ ID NO: 637) y GGTAGTCAGGGTTGTCCAGG (antisentido) (SEQ ID NO: 638); Exón 28b, CGAGTATCTCAACACTGTCCAGC (sentido) (SEQ ID NO: 639) y CTAAGAGCTAATGCGGGC ATGGCTG (antisentido) (SEQ ID NO: 640). La temperatura de hibridación para las
 55 amplificaciones del exón 1 fue de 54° C. Las temperaturas de hibridación tanto para las amplificaciones primaria como secundaria fueron de 58 °C (exones 3, 4, 7 - 10, 12 - 17, 19 - 25, 27, y 28), 56 °C (exones 2, 5, 6, y 26), o 52 °C (exones 11 y 18).

Los amplicones de la PCR se purificaron usando exonucleasa I (United States Biochemical, Cleveland, OH), y fosfatasa alcalina de camarón (United States Biochemical, Cleveland, OH) antes de la secuenciación. El ADN purificado se diluyó y se realizó el ciclo de secuenciación utilizando el kit v1.1 BigDye Terminator de ABI (ABI, Foster City, CA) de acuerdo con las instrucciones del fabricante. Las reacciones de secuenciación se sometieron a
 60 electroforesis en un analizador genético ABI 3100. Se analizaron los electroferogramas tanto en dirección sentido como antisentido por la presencia de mutaciones, utilizando el software Sequence Navigator en combinación con Factura para marcar las posiciones de los heterocigotos. Todas las variantes de la secuencia se confirmaron en múltiples amplificaciones independientes de PCR y reacciones de secuenciación.

Líneas celulares derivadas de cáncer:

Se analizó un panel de 14 líneas celulares derivadas de cáncer de pulmón se analizó por mutaciones del EGFR. Estos se derivaron de tumores de NSCLC (N = 5), cáncer de pulmón de células pequeñas (N = 6), carcinoma adenoescamoso (N = 1), carcinoide bronquial (N = 1), y de histología desconocida (N = 1). Las líneas celulares específicas fueron: NCI-H460, NCI-522, HOP-92, NCIH841, NCIH734, NCIH2228, NCIH596, NCIH727, NCIH446, NCIH1781, NCIH209, NCIH510, NCIH82, NCIH865. Además, se seleccionaron 64 líneas celulares derivadas de cáncer para las mutaciones en los exones 19 y 21. Estas representaron las siguientes histologías: cáncer de mama (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), cáncer de ovario (ES-2, IGROV-1, MDAH2774, OV1063, OVCAR3, OVCAR4, OVCAR5, SKOV3, SW626), cánceres del SNC (SF-295, SNB-19, U-251, CCF-STTG1, SW-1088, SW-1783, T98G, M059K, A172, SK-N-DZ, SK-N-MC), leucemia (CCRF-CEM, K562, MOLT-4, RPMI8226, SR), cáncer de próstata (DU-145, PC-3), cáncer de colon (COLO-205, HCT-116, HCT-15, HT-29, SW-620), cáncer renal (786-O, ACHN, CAKI-1, SN-12C, UO-31), melanoma (LOX-IMVI, M14, SKMEL2, UACC-62), osteosarcoma (SAOS-2), y cánceres de cabeza y cuello (O11, O13, O19, O28, O22, O29, O12). Las líneas de células de cáncer de cabeza y cuello fueron proporcionadas por el Dr. James Rocco, Massachusetts General Hospital/Massachusetts Eye and Ear Infirmary. El resto de las líneas celulares están disponibles a través de la American Type Culture Collection (Manassas, VA).

El ADN genómico se aisló a partir de especímenes tumorales congelados en forma instantánea. Se molieron primero especímenes de tumor hasta un polvo fino usando un mortero y mano de mortero previamente enfriados y esterilizados. Se transfirió inmediatamente el tejido tumoral a una solución de extracción de ADN que consiste de cloruro de sodio 100 mM, Tris 10 mM pH 7,5, EDTA (ácido etilendiamina tetra acético disódico) 25 mM pH 8,0, y dodecil sulfato de sodio al 0,5% (p/v), y 100 µg/ml de proteinasa K recientemente preparada y se incubó durante la noche a 37° C o durante 3 horas a 50° C. Se extrajo luego el ADN, utilizando los métodos estándar de fenol-cloroformo, se precipitó con etanol, se lavó con etanol al 70%, se secó al aire y se resuspendió en amortiguador TE. Se determinó la concentración de ADN espectrofotométricamente. Se amplificaron los Exones 19 y 21 del EGFR humano mediante la reacción en cadena de la polimerasa usando los siguientes pares de cebadores: cebador sentido del Exon19, 5'-GCAATATCAGCCTTAGGTGCGGCTC-3' (SEQ ID NO: 505); cebador antisentido del Exón 19, 5'-CATAGAA AGTGAACATTTAGGATGTG-3' (SEQ ID NO: 506); cebador sentido del Exón 21, 5'-CTAACGTTCCG CCAGCCATAAGTCC-3' (SEQ ID NO: 507); cebador antisentido del Exón 21, 5'-GCTGCGAGCTCACCCAG AATGTCTGG-3' (SEQ ID NO: 508). Para cada muestra, se amplificaron 20 ng de ADN genómico en una reacción PCR que consiste de 1X amortiguador Expand Long Template 1 (Roche, Mannheim Alemania), dATP grado secuenciación 50 M (Amersham Biosciences, Cleveland, OH), dCTP grado secuenciación 50 µM (Amersham Biosciences, Cleveland OH), dGTP grado secuenciación 50 µM (Amersham Biosciences, Cleveland OH), dTTP grado secuenciación 50 µM (Amersham Biosciences, Cleveland OH), cebador sentido 0,2 µM, cebador antisentido 0,2 µM, 1,25 unidades de mezcla de la enzima Expand Long Template (Taq ADN polimerasa/Tgo ADN polimerasa) (Roche, Mannheim, Alemania) que ha sido previamente incubada durante 5 minutos en hielo con 1/6 de volumen de Anticuerpo TaqStart (1,1 µg/µl) (Clontech, Palo Alto, CA) y agua hasta un volumen final de 25 µl. Cada serie de amplificaciones también incluye un control negativo para el que se omite el molde de ADN. Las condiciones de los ciclos de la PCR para ambos exones fueron 95° C durante 2 min seguido de 40 ciclos de 95° C durante 30 s, 58° C durante 30 s y 72° C durante 45 s, y una extensión final de 72° C durante 10 min seguido de mantenimiento a 4° C en un Termociclador PTC-200 o PTC-225 de MJ-Research (MJ-Research, Waltham, MA).

Los productos de la PCR se resolvieron por electroforesis a través de un gel de agarosa al 0,8% para asegurar la amplificación del material del paciente y que no hubiera amplificación en el control negativo. Se purificaron los productos de la PCR antes de la secuenciación mediante la mezcla de 10 µl de cada amplicón de la PCR con 0,5 µl de exonucleasa I (10 U/µl) (United States Biochemical, Cleveland, OH), y 1 µl de fosfatasa alcalina de camarón (1 U/µl) (United States Biochemical, Cleveland, OH) e incubando a 37° C durante 20 minutos, seguido de inactivación a 80° C durante 15 minutos en un termociclador (MJ-Research, Waltham, MA). Se diluyó ADN purificado en agua, de acuerdo con la intensidad del amplicón, y se llevó a cabo el ciclo de secuenciación utilizando el kit v1.1 BigDye Terminator de ABI (Applied Biosystems, Foster City, CA) de acuerdo con las instrucciones del fabricante. El ciclo de secuenciación se realizó en un termociclador de MJ-Research usando las siguientes condiciones para el ciclo: los cebadores utilizados para la secuenciación fueron: cebador sentido del Exón 19, 5'-GCAATATCAGCCTTAGGTGCGGCTC-3' (SEQ ID NO: 505); cebador antisentido del Exón 19, 5'-CATAG AAAGTGAACATTTAGGATGTG-3' (SEQ ID NO: 506); cebador sentido del Exón 21, 5'-CTAACGTTCCGCCAG CCATAAGTCC-3' (SEQ ID NO: 507) o 5'-CGTGGAGAGGCTCAGAGCCTGGCATG-3' (SEQ ID NO: 687); cebador antisentido del Exón 21, 5'-GCTGCGAGCTCACCCAGAAATGTCTGG-3' (SEQ ID NO: 508). Las reacciones de secuenciación se sometieron a electroforesis en un analizador genético ABI 3100 (Applied Biosystems, Foster City, CA). Se utilizaron los programas de software Factura y Sequence Navigator (Applied Biosystems, Foster City, CA) para marcar las posiciones posibles de los heterocigotos y los muestra para su evaluación. Las posiciones de los nucleótidos en las cuales la altura del pico secundario fue mayor que, o igual a, 30% de la altura del pico principal se marcaron como heterocigotos y se confirmaron mediante análisis tanto de lecturas sentido como antisentido. Las muestras con una secuencia indicativa de la presencia de una mutación fueron amplificadas nuevamente y secuenciadas para la confirmación.

Posición de los cebadores utilizados en el análisis de secuencia con relación a los Exones 19 y 21

Los cebadores intrónicos se muestran en minúsculas y subrayados.

La secuencia intrónica se muestra en minúsculas.

La secuencia exónica se muestra en mayúsculas.

Exón 19 del EGFR (5' - 3') (SEQ ID NO: 641)

gcaatatcagccttaggtgaggctccacagccccagtggtccctcaccttcgggggtgcacgc
tggttaacatccacccagatcactgggcagcatgtggcaccatctcacaattgccagttaacg
tcttccttctctctctgtcatagGGACTCTGGATCCCAGAAGGTGAGAAAGTTAAAATCCCC
GTCGCTATCAAGGAATTAAGAGAAGCAACATCTCCGAAAGCCAACAAGGAAATCCTCGATgt
gagtttctgctttgctgtgtgggggtccatggctctgaacctcaggccccaccttttctcatg
tctggcagctgctctgctctagaccctgctcatctccacatcctaataatgttcactttctatg

5

Exón 21 del EGFR (5' - 3') (SEQ ID NO: 642) o (SEQ ID NO: 687)

ctaacggttcgccagccataagtccctcgacgtggagaggctcagagcctggcatgaacatgac
cctgaattcggatgcagagcttcttcccatgatgatctgtccctcacagcagggtcttctct
gtttcagGGCATGAACTACTTGGAGGACCGTCGCTTGGTGACCGCGACCTGGCAGCCAGGA
ACGTA CTGGTGAAAA CACCGCAGCATGTCAAGATCACAGATTTTGGGCTGGCCAAACTGCTG
GGTGCGGAAGAGAAA GAATACCATGCAGAAGGAGGCAAAGtaaggaggtggctttaggtcag
ccagcattttcctga caccagggaccaggctgccttcccactagctgtattgtttaacacat
gcagggggaggatgctctccagacatctctgggtgagctcgcagc

10 Resultados

Características clínicas de las respuestas al gefitinib

Los pacientes con NSCLC avanzado reacio a la quimioterapia han sido tratados con gefitinib como único agente desde el año 2000 en el Hospital General de Massachusetts. Un total de 275 pacientes fueron tratados, tanto antes de su aprobación en mayo de 2003 por la FDA, como parte de un programa de acceso más amplio de uso compasivo, y después de esa fecha a través del uso del producto obtenido en el comercio. Durante este período, se identificó por parte de los médicos que 25 pacientes tenían respuestas clínicas significativas. Una respuesta clínica significativa se define ya sea como una respuesta parcial según criterios RECIST para los pacientes con enfermedad medible, o para los pacientes cuya carga tumoral no podía ser cuantificada usando estos criterios, dos médicos evaluaron una respuesta que pudiera ser medida. La Tabla 1 muestra las características clínicas de 9 casos para los cuales las muestras de tumor obtenidas en el momento del diagnóstico inicial estaban disponibles. Para los otros que responden al gefitinib, el tejido no estaba disponible, más comúnmente, porque las muestras de diagnóstico se limitaron a una citología de los aspirados con aguja. Como grupo, los 9 pacientes experimentaron beneficio sustancial con el gefitinib. La supervivencia media desde el inicio del tratamiento con el fármaco es de más de 18 meses, y la duración media de la terapia es mayor a 16 meses. De acuerdo con informes previos, aquellos que responden al gefitinib tienen una alta prevalencia de sexo femenino, ausencia de antecedentes de tabaquismo, y tumores con histología broncoalveolar (11, 12). El caso 6 es representativo de la cohorte sensible al Gefitinib. Este paciente es un hombre de 32 años de edad, sin antecedentes de tabaquismo, que presentaba múltiples lesiones cerebrales y la enfermedad en el pulmón derecho se diagnosticó como carcinoma broncoalveolar. El paciente fue tratado con radioterapia total del cerebro, seguido por una serie de regímenes de quimioterapia a la que no respondió su tumor (carboplatino y gemcitabina; docetaxel; vinorelbina). Con un estado funcional tendiente a empeorar y un agobio progresivo del tumor del pulmón, el paciente empezó la terapia con 250 mg por día de gefitinib. Su falta de aliento mejoró rápidamente y un escáner TC del pulmón 6 semanas después del inicio del tratamiento reveló la mejoría dramática que se muestra en la Figura 1.

30

Mutaciones del EGFR en aquellos que responden al gefitinib

Se formuló la hipótesis de que los casos de NSCLC con respuestas sorprendentes al gefitinib podrían albergar mutaciones somáticas en el *EGFR*, lo que indica el papel esencial desempeñado por esta ruta de señalización del factor de crecimiento en estos tumores. Para buscar tales mutaciones, primero se analizaron los reordenamientos dentro del dominio extracelular del EGFR que son característicos de los gliomas (15): no se detectó ninguno. Por lo tanto, se secuenció toda la región de codificación del gen usando amplificación por PCR de los exones individuales. No se observaron mutaciones heterocigóticas en 8 de los 9 casos, todos los cuales fueron agrupados dentro del dominio quinasa del EGFR (Tabla 2 y Figura 2). Cuatro tumores tenían supresiones en el marco por la remoción de los aminoácidos 746 - 750 (delE756-A750; caso 1), 747 a 750 (delL747-T751insS; caso 2), y 747 a 752 (delL747-P753insS; casos 3 y 4). Las dos últimas supresiones se asociaron con la inserción de un residuo de serina, resultante de la generación de un nuevo codón en el rompimiento de la supresión. Sorprendentemente, estas cuatro supresiones se superpusieron, con la supresión de cuatro aminoácidos (leucina, arginina, ácido glutámico y alanina, en los codones 747 a 750) dentro del exón 19 compartido por todos los casos (véase la Figura 4a). Otros tres tumores tenían sustituciones de aminoácidos en el exón 21: leucina por arginina en el codón 858 (L858R; casos 5 y 6), y leucina por glutamina en el codón 861 (L861 Q; caso 7). La mutación L861Q es de particular interés, ya que el mismo cambio de aminoácidos en el gen *egfr* de ratón es responsable por el rasgo de piel oscura (*dsk5*), asociado con la señalización del EGFR alterado (18). Una cuarta mutación de sentido erróneo en el dominio quinasa dio lugar a una sustitución de glicina por cisteína en el codón 719 dentro del exón 18 (G719C; caso 8). Estaba disponible tejido normal emparejado para los casos de 1, 4, 5 y 6, y mostró sólo la secuencia de tipo silvestre, lo que indica que las mutaciones habían surgido somáticamente, durante la formación del tumor. No se observaron mutaciones en siete casos de NSCLC que no lograron responder al gefitinib ($P = 0,0007$; prueba exacta de Fisher bilateral).

Prevalencia de mutaciones específicas del EGFR en NSCLC y otros tipos de cáncer

A diferencia de los gliomas, en los que los reordenamientos que afectan al dominio extracelular del EGFR han sido estudiados extensamente (15), la frecuencia de las mutaciones del *EGFR* en NSCLC no ha sido definida. Por lo tanto, se secuenció toda la región de codificación del gen en 25 casos primarios de NSCLC no relacionados con el estudio del gefitinib, incluyendo 15 con histología broncoalveolar, que se ha asociado con sensibilidad al gefitinib en ensayos clínicos previos (11, 12). Se detectaron mutaciones heterocigotas en dos tipos de cáncer broncoalveolar. Ambos casos tenían supresiones en el marco en el dominio quinasa idénticas a aquellas encontrados en aquellos que responden al gefitinib, a saber delL747-P753insS y delE746-A750 (Tabla 2). Dada la aparente agrupación de mutaciones del EGFR, se secuenciaron los exones 19 y 21 en un total de 55 tumores primarios y 78 líneas celulares derivadas de cáncer, que representan diversos tipos de tumores (véase el Material complementario). No se detectaron mutaciones, lo que sugiere que estas surgen sólo en un subconjunto de cánceres, en los que la señalización del EGFR pueden jugar un papel crítico en la tumorigénesis.

Aumento en la activación inducida por EGF y la inhibición por gefitinib de las proteínas del EGFR mutante

Para estudiar las propiedades funcionales codificadas por estas mutaciones, se expresaron la supresión de L747-S752insS y los mutantes de sentido erróneo L858R en células cultivadas. La transfección transitoria de los constructos mutantes y de tipo silvestre en células Cos-7 demostró niveles de expresión equivalentes, lo que indica que las mutaciones no afectan la estabilidad de la proteína. Se cuantificó la activación del EGFR mediante la medición de la fosforilación del residuo de tirosina ¹⁰⁶⁸, comúnmente usado como un marcador de la autofosforilación del receptor (19). En la ausencia de factores de crecimiento asociados y en suero, ni el EGFR de tipo silvestre ni el mutante demostraron autofosforilación (Figura 3a). Sin embargo, la adición de EGF condujo a un aumento de 2-3 veces en la activación del receptor tanto para los mutantes del *EGFR* de supresión y de sentido erróneo, en comparación con el receptor de tipo silvestre. Por otra parte, mientras que la activación del EGFR normal estaba subregulada después de 15 min, en forma consistente con la internalización del receptor, los dos receptores mutantes demostraron activación continua hasta por 3 horas (Figura 3a). Se obtuvieron resultados similares con anticuerpos que miden el total de la fosforilación del EGFR después de la adición de EGF (no se muestra).

Ya que 7 de 8 mutaciones de quinasa del EGFR residen cerca de la hendidura del ATP, que está dirigida por el gefitinib, se determinó si las proteínas mutantes habían alterado la sensibilidad con el inhibidor. Se midió la autofosforilación del receptor inducida por EGF en las células pretratadas con concentraciones variables de gefitinib. Sorprendentemente, ambos receptores mutantes mostraron un aumento de la sensibilidad con la inhibición por gefitinib. El EGFR de tipo silvestre tenía una IC_{50} de 0,1 μM y mostró una inhibición completa de la autofosforilación con 2 μM gefitinib, mientras que las dos proteínas mutantes tenían una IC_{50} de 0,015 μM y la anulación de la autofosforilación con 0,2 μM (Figura 3b). Esta diferencia en la sensibilidad al fármaco puede ser clínicamente relevante, ya que los estudios farmacocinéticos indican que la administración oral diaria de 400 - 600 mg de gefitinib da como resultado una concentración estacionaria mínima en plasma de 1,1 - 1,4 μM , mientras que la dosis diaria recomendada actualmente de 250 mg conduce a una concentración mínima de 0,4 μM (20).

Ejemplo 2

Las células tumorales que albergan mutaciones dentro del dominio quinasa del EGFR, y por lo tanto son sensibles a la inhibición del crecimiento por el tratamiento con gefitinib, pueden experimentar mutaciones en un "segundo sitio", también dentro del dominio quinasa, que confieren resistencia al gefitinib pero que aún se "activan" en el sentido de que presentan un aumento de señalización del EGFR con respecto al EGFR de tipo silvestre. Tales mutantes

resistentes al gefitinib se generan a partir de dos líneas celulares de NSCLC humano esporádico a saber NCI-1650 y NCI-1975. Cada línea celular contiene una mutación heterocigota con el dominio quinasa del EGFR, y, por lo tanto, se espera que sea sensible al gefitinib. La mutación del EGFR en NCI-1650 consiste en una supresión en el marco de 15 nucleótidos en la posición 2235 - 2249 (delLE746-A750) en el exón 19, mientras que NCI-1975 tiene una mutación en sentido erróneo en el exón 21 que sustituye una G por T en el nucleótido 2573 (L858R). La mutación L858R en NCI-H1975 se ha demostrado en este documento que se activa y que confiere un aumento de sensibilidad al gefitinib in vitro.

Las líneas de células resistentes al gefitinib, derivadas tanto de NCI-1650 como de NCI-1975 se aíslan, después de mutagénesis química aleatoria usando EMS (metanosulfonato de etilo) seguido por el cultivo en medio suplementado con gefitinib para seleccionar el resultado de los clones resistentes. El subcultivo de los clones individuales es seguido por la determinación de la secuencia de nucleótidos del gen EGFR después de la amplificación específica mediada por PCR del ADN genómico correspondiente al dominio quinasa del EGFR.

Una variación de esta estrategia implica el paso en serie de estas dos líneas celulares en presencia de concentraciones gradualmente crecientes de gefitinib en el transcurso de varias semanas o meses con el fin de seleccionar la adquisición espontánea de mutaciones en el gen EGFR que confieren resistencia al gefitinib. Las células seleccionadas (que continuarán proliferando a una concentración relativamente alta de gefitinib) se aíslan en forma de colonias, y se identifican las mutaciones como se ha descrito anteriormente.

Ejemplo 3

Para determinar si la mutación de tirosina quinasas del receptor desempeña un papel causal en el NSCLC, se buscaron alteraciones genéticas somáticas en una serie de 119 tumores primarios de NSCLC, que consisten en 58 muestras del Hospital Universitario de la ciudad de Nagoya en Japón y 61 del Hospital Brigham y de Mujeres en Boston, Massachusetts. Los tumores incluyeron 70 adenocarcinomas de pulmón y otros 49 tumores de NSCLC de 74 pacientes varones y 45 pacientes de sexo femenino, ninguno de los cuales había documentado tratamiento con inhibidores de quinasa del EGFR.

Como una selección inicial, se amplifican y secuencian los exones que codifican los bucles de activación de 47 de los 58 genes de tirosina quinasa receptora humana (*) (Tabla S1) a partir de ADN genómico de un subconjunto de 58 muestras de NSCLC que incluye 41 adenocarcinomas de pulmón. Tres de los tumores, todos adenocarcinomas de pulmón, mostraron mutaciones de sentido erróneo heterocigotas en *EGFR* que no están presentes en el ADN de tejido pulmonar normal de los mismos pacientes (Tabla S2; S0361, S0388, S0389). No se detectaron mutaciones en amplicones de otros genes de la tirosina quinasa receptora. Todos los tres tumores tenían la misma mutación del EGFR, que se predijo que cambia la leucina ("L") en la posición 858 por arginina ("R") (Figura 6A; CTG → CGG; "L858R"), en donde toda la numeración se refiere a EGFR humano.

A continuación se examinaron los exones 2 al 25 del *EGFR* en la colección completa de 119 tumores de NSCLC. La secuenciación del exón de ADN genómico reveló mutaciones de sentido erróneo y la supresión de mutaciones del *EGFR* en un total de 16 tumores, todo dentro de los exones 18 a 21 del dominio quinasa. Todas las alteraciones de secuencia en este grupo fueron heterocigotas en el ADN del tumor; en cada caso, el tejido pulmonar normal aparejado del mismo paciente mostró una secuencia de tipo silvestre, lo que confirma que las mutaciones son de origen somático. La distribución de las alteraciones de las secuencias de nucleótidos y de proteínas, y las características del paciente asociadas con estas anomalías, se resumen en la Tabla S2.

Las mutaciones de sustitución G719S y L858R se detectaron en dos y tres tumores, respectivamente. La mutación "G719S" cambia la glicina (G) en la posición 719 por serina (S) (Figura 6B). Estas mutaciones se localizan en el motivo GXGXXG (SEQ ID NO: 490) del dominio de enlazamiento de trifosfato del nucleótido o bucle P y adyacente al motivo DFG altamente conservado en el bucle de activación (52), respectivamente. Véase por ejemplo, la Figura 7. Los residuos mutados son casi invariables en todas las proteínas quinasas y los residuos análogos (G463 y L596) en la proteína serina-treonina quinasa B-Raf son somáticamente mutados en carcinoma colorrectal, de ovario y de pulmón (41, 53) (Figura 6A, 6B).

También se detectaron múltiples mutaciones de supresión agrupadas en la región que abarca los codones 746 a 759 dentro del dominio quinasa del EGFR. Diez tumores transportaban uno de las dos supresiones de 15 nucleótidos que se superponen eliminando los codones del EGFR 746 - 750, a partir ya sea del nucleótido 2235 o 2236 (Del-1; Figuras 6C y 8C, Tabla S2). El ADN del EGFR de otro tumor mostró una brecha de 24 nucleótidos heterocigotos que conduce a la supresión de los codones 752 - 759 (Del-2; Figura 6C). Los cromatogramas representativos se muestran en las Figuras 8A - 8F.

Las posiciones de las mutaciones de sustitución y la supresión Del-1 en la estructura tridimensional de la forma activa del dominio quinasa del EGFR (54) se muestran en la Figura 7. Obsérvese que las alteraciones de la secuencia se agrupan alrededor del sitio activo de la quinasa, y que las mutaciones de sustitución se encuentran en el bucle de activación y el bucle P rico en glicina, elementos estructurales que se sabe que son importantes para la autorregulación en muchas proteína quinasas (52).

Se han identificado dos mutaciones adicionales del EGFR en dos tipos de tumores diferentes. Es decir, se ha

identificado la mutación G857V del *EGFR* en la leucemia mielógena aguda (AML) y la mutación L883S del *EGFR* en un sarcoma metastásico. La mutación "G857V" tiene la glicina (G) en la posición 857 sustituida con una valina (V), mientras que la mutación "L883S" tiene la leucina (L) en la posición 883 sustituido con una serina (S). Estos hallazgos sugieren que se producen mutaciones en *EGFR* en varios tipos de tumores y, lo más importante, que los inhibidores del *EGFR* serían eficaces en el tratamiento de los pacientes portadores de tales mutaciones. Esto expande el uso de inhibidores de quinasa tales como, por ejemplo, los inhibidores de la tirosina quinasa gefitinib (comercializada como IressaTM), erlotinib (comercializado como TarcevaTM), y similares, en el tratamiento de otros tipos de tumores diferentes al NSCLC.

Las mutaciones del *EGFR* muestran una notable correlación con las características de los pacientes diferenciales descritas en las poblaciones de pacientes japonesas y de los Estados Unidos. Como se señaló anteriormente, los ensayos clínicos revelan una variabilidad significativa en la respuesta al inhibidor de tirosina quinasa gefitinib (IressaTM), con mayores respuestas observadas en los pacientes japoneses que en una población predominantemente europea (27,5% versus 10,4%, en un ensayo de fase II multi-institucional) (48); y con respuestas parciales observadas con más frecuencia en mujeres no fumadoras en los Estados Unidos y en pacientes con adenocarcinomas (49 - 51). Se demuestra que las mutaciones del *EGFR* fueron más frecuentes en los adenocarcinomas (15/70 o 21%) que en otros NSCLC (1/49 o 2%); más frecuente en mujeres (9/45 o 20%) que en hombres (7/74 o 9%), y más frecuentemente en los pacientes de Japón (15/58 o 26%, y 14/41 con adenocarcinomas o 32%) que en aquellos de los Estados Unidos (1/61 o 2%, y 1/29 con adenocarcinomas o 3%). Se observó la fracción más alta de mutaciones del *EGFR* en las mujeres japonesas con adenocarcinoma (8/14 o 57%). En particular, las características de los pacientes que se correlacionan con la presencia de mutaciones del *EGFR* parecen ser aquellas que se correlacionan con la respuesta clínica al tratamiento con gefitinib.

Para investigar si las mutaciones del *EGFR* pueden ser un factor determinante de la sensibilidad al gefitinib, se obtuvieron muestras de NSCLC antes del tratamiento de 5 pacientes que respondieron y 4 pacientes que progresaron durante el tratamiento con gefitinib, donde estos pacientes se identificaron de un total de más de 125 pacientes tratados en el Instituto del Cáncer Dana-Farber, ya sea en un programa de acceso expandido o después de la aprobación reglamentaria del gefitinib (49). Cuatro de los pacientes tuvieron respuestas radiográficas parciales ($\geq 50\%$ de regresión tumoral en una tomografía computarizada después de 2 meses de tratamiento), mientras que el quinto paciente experimentó dramática mejoría de los síntomas en menos de dos meses. Todos los pacientes eran de los Estados Unidos y caucásicos.

Aunque la secuenciación del dominio quinasa (exones 18 a 24) no reveló mutaciones en los tumores de los cuatro pacientes cuyos tumores progresaron con gefitinib, los cinco tumores de pacientes sensibles a gefitinib albergaban mutaciones del dominio quinasa del *EGFR*. La prueba de Chi cuadrado reveló la diferencia en la frecuencia de la mutación del *EGFR* entre aquellos que respondieron al gefitinib (5/5) y los que no respondieron (0/4) que es estadísticamente significativa con $p = 0,0027$, mientras que la diferencia entre quienes respondieron al gefitinib y pacientes estadounidenses no seleccionados con NSCLC (5/5 frente a 1/61) también fue significativa con $p < 10^{-12}$ (*). La mutación L858R del *EGFR*, observado previamente en los tumores no seleccionados, se identificó en un adenocarcinoma de pulmón sensible a gefitinib (Figura 6A; Tabla S3A, IR3T). Tres tumores sensibles a gefitinib contenían supresiones heterocigotas en el marco (Fig. 6C y Tablas S3A y S3B, Del-3 en dos casos y Del-4 en uno) y uno contenía una supresión homocigota en el marco (Fig. 6C y Tablas S3A y S3B, Del-5). Cada una de estas supresiones estaba dentro de la región del codón 746 - 753 del *EGFR*, donde también se encontraron supresiones en tumores no seleccionados. Cada una de estas tres supresiones también se asocia con una sustitución de aminoácidos (Tablas S3A - S3C). En las cuatro muestras donde el tejido normal coincidió, se confirmaron estas mutaciones como somáticas.

Ejemplo 3A: Diseño del cebador

Las secuencias de ADNc de tirosina quinasas receptoras se obtuvieron a partir del GenBank (números de acceso enumerados en la Tabla S1), y fueron a la formación del genoma humano (<http://genome.ucsc.edu>) utilizando la alineación BLAT para identificar los límites del exón/intrón. Se diseñaron pares de cebadores específicos del gen externo para amplificar secuencias de exón y al menos 250 pb de secuencia intrónica de flanco o secuencia exónica adyacente en cada lado utilizando el programa Primer3 (http://frodo.wi.mit.edu/primer3/primer3_code.html). Los amplicones resultantes predichos se utilizaron entonces para diseñar cebadores internos que flanquean el exón (generalmente mayor de 50 pb desde el límite del exón/intrón) y que contienen anexas colas de cebador hacia adelante e inversas M13. Estos conjuntos de cebadores anidados se pusieron a prueba para el tamaño adecuado del amplicón y la secuencia de alta calidad a partir del ADN de control. Los amplicones que abarcan exones que codifican el bucle de activación de tirosina quinasa receptora de 47 tirosina quinasas se amplificaron y secuenciaron en un conjunto de 58 muestras de cáncer de pulmón primario de la escuela de medicina de la Universidad de la Ciudad de Nagoya. Además, los amplicones que cubren toda la longitud del *EGFR* también fueron amplificados.

Ejemplo 3B: Métodos PCR y de secuenciación de ADN genómico

Los exones de la tirosina quinasa y las secuencias intrónicas de flanco fueron amplificados utilizando cebadores específicos en una disposición de PCR anidado en formato de 384 pozos. Cada reacción PCR contenía 5 ng de ADN, 1X de amortiguador HotStar, los dNTP de 0,8 mM, MgCl₂ 1 mM, 0,2 U de Enzima HotStar (Qiagen, Valencia,

CA), y cebadores hacia delante e inversos 0,2 μ M en un volumen de reacción de 10 μ l. Los parámetros de los ciclos de PCR fueron: un ciclo de 95° C durante 15 min, 35 ciclos de 95° C durante 20 s, 60° C durante 30 s y 72° C durante 1 min, seguido de un ciclo de 72° C durante 3 min.

- 5 Los productos PCR resultantes se purificaron por medio de química de inmovilización reversible en fase sólida seguida por secuenciación fluorescente terminadora de coloración bidireccional con cebadores universales M13. Los fragmentos de secuenciación se detectaron a través de electroforesis capilar utilizando un analizador de ADN ABI Prism 3700 (Applied Biosystems, Foster City, CA). Se llevaron a cabo una PCR y secuenciación por parte de Agencourt Bioscience Corporation (Beverly, MA).

Ejemplo 3B: Análisis de secuencias y validación

- 10 Se analizaron Cromatogramas hacia adelante (F) e inversos (R) fueron analizadas por lotes por medio del Mutation Surveyor 2.03 (SoftGenetics, State College, PA), seguido por una revisión manual. Las variaciones en la secuencia de alta calidad que se encuentran en una o ambas direcciones se anotaron como mutaciones candidatas. Se volvieron a amplificar los exones que albergan mutaciones candidatas a partir de la muestra original de ADN y se volvió a secuenciar como anteriormente.

Ejemplo 3C: Los pacientes

- 20 Los especímenes de tumor de pulmón se obtuvieron de pacientes con cáncer de pulmón de células no pequeñas tratados en el Hospital de la Universidad de la Ciudad de Nagoya y el Hospital Brigham y de Mujeres (tumores de japoneses no seleccionados y tumores de estadounidenses tratados con gefitinib, respectivamente) y del banco de tumores anónimos del Hospital Brigham y de Mujeres (muestras de estadounidenses no seleccionadas) bajo estudios aprobados de la Junta de Revisión Institucional. La información sobre el género, la edad, y la histología estaba disponible para la mayoría de las muestras. Las muestras de pacientes también fueron obtenidas de pacientes tratados en un ensayo clínico de etiqueta abierta de gefitinib en el Instituto del Cáncer Dana-Farber (13). Las respuestas al gefitinib se definieron mediante criterios estándar (Véase, por ejemplo, AB Miller, B. Hoogstraten, M. Staquet, A. Winkler, 1981 Cancer 47, 207 - 14). Se obtuvo la aprobación del IRB para estos estudios.

- 25 De los pacientes que responden al gefitinib, hubo dos pacientes que habían sido tratados previamente con al menos un ciclo de quimioterapia, un paciente previamente tratado con terapia de radiación, un paciente tratado al mismo tiempo con quimioterapia, y un paciente que no recibió ningún otro tratamiento. Para los pacientes no sensibles al gefitinib, el fracaso del tratamiento se definió como la aparición de nuevas lesiones tumorales o el crecimiento de lesiones tumorales existentes en una tomografía computarizada después de 2 meses de tratamiento con gefitinib en comparación con una tomografía computarizada de línea de base.

Ejemplo 3D: Secuenciación de ADNc de muestras de los pacientes

- 35 Se aisló ARN total a partir de muestras de tejido utilizando Trizol™ (Invitrogen, Carlsbad, CA) y se purificó usando un kit de limpieza de minieluato ARNeasy™ (Qiagen, Valencia, CA). Se transcribe ADNc a partir de 2 μ g de ARN total con transcriptasa inversa Superscript II (Invitrogen Life Technologies, Carlsbad, CA), de acuerdo con las recomendaciones del fabricante. El ADNc se usa como molde para amplificaciones posteriores por PCR del EGFR.

- 40 Los componentes de la PCR son: Tris-HCl 20 mM (pH 8,4), KCl 50 mM, MgCl₂ 1,5 mM, dATP, dCTP, dGTP, dTTP cada uno 0,2 μ M, 0,2 μ M de cada cebador, y 0,05 unidades/ μ l de Taq polimerasa (Taq Platinum, GIBCO BRL, Gaithersburg, MD). La amplificación del fragmento "a" requiere la adición de DMSO al 4% a la reacción. Las secuencias de cebadores se enumeran en la Tabla S4. Los cebadores hacia adelante e inversos se sintetizan con 18 pares de bases de secuencias hacia delante e inversas de M13 que sobresalen, respectivamente. Las condiciones de los ciclos térmicos son: 94° C, 4 min, seguido por 11 ciclos, con una etapa de desnaturalización a 94° C durante 20 s, una etapa de extensión a 72° C durante 20 s, y con una etapa de hibridación de 20 s que disminuyó 1° C/ciclo, desde 60° C en un ciclo hasta 50° C en el ciclo de 11; se repitió luego el ciclo 11, 25 veces. Una incubación de 6 minutos a 72° C seguida de un remojo a 4° C completan el programa.

- 45 Una parte alícuota de la reacción de PCR se diluye en proporción 1:50 con agua. El producto de PCR diluido se secuenciar usando un kit cebador M13 Forward Big Dye (Perkin-Elmer/Applied Biosystems, Foster City, CA), de acuerdo con las recomendaciones del fabricante. Los productos de secuenciación se separan en un secuenciador fluorescente (modelo 3100 de Applied Biosystems, Foster City, CA). Las llamadas de bases se realizan mediante el software del instrumento, y se revisan por medio de inspección visual. Cada secuencia se compara con la secuencia normal correspondiente utilizando el software Sequencher 4.1 (Gene Codes Corp.).

Ejemplo 3E: Tipos de tumores que expresan EGFR mutante

- 55 Se encontraron dos mutaciones adicionales en EGFR en dos tipos de tumores diferentes. Se identificó una mutación del EGFR que sustituye una glicina (G) para una valina (V) en posición 857 ("G857V") en leucemia mielógena aguda (AML). Una mutación EGFR que sustituye una leucina (L) con una serina (S) en posición 883 ("L883S") en un sarcoma metastásico.

Ejemplo 3F: Líneas celulares

Se examinaron los efectos del gefitinib sobre líneas de células de NSCLC. Una línea celular, H3255, fue particularmente sensible al gefitinib, con una IC_{50} de 40 nM. Otras líneas celulares tenían IC_{50} mucho más altos. Por ejemplo, una línea celular de tipo silvestre H1666 tiene un IC_{50} de 2 μ M, que es 50 veces mayor que para la línea celular mutante. Cuando se secuenció el EGFR de esta línea celular, contenía la mutación de sentido erróneo L858R, mientras que las otras líneas celulares eran de tipo silvestre para EGFR. Se requirieron concentraciones mucho menores de gefitinib para desactivar la fosforilación del EGFR y también AKT y ERK por medio del EGFR en comparación con las células de tipo silvestre para EGFR, que requieren al menos concentraciones 100 veces más altas de gefitinib para lograr el mismo efecto. Estos hallazgos sugieren que el receptor mutante es más sensible a los efectos del gefitinib.

También téngase en cuenta aquí,

Ejemplo 3G: Terapias de combinación

Se analizaron especímenes tumorales de pacientes con NSCLC avanzado tratados en un ensayo aleatorio de carboplatino/paclitaxel con o sin erlotinib. La porción clínica de este ensayo demostró una supervivencia equivalente en los dos brazos del tratamiento. Se encontraron disponibles especímenes tumorales para secuenciación de 228 de los 1076 pacientes. Las características clínicas preliminares de estos pacientes no son diferentes de las del grupo como un todo con respecto a los datos demográficos de referencia, la velocidad de respuesta, la supervivencia media y total.

Se secuenciaron los exones 18 - 21 del dominio tirosina quinasa y se identificaron 29 mutaciones, para una frecuencia de mutación de 12,7 por ciento.

En su conjunto, los pacientes con mutaciones del EGFR tienen una mejor supervivencia, independientemente de si recibieron tratamiento con quimioterapia únicamente o en combinación con erlotinib. Estas diferencias son estadísticamente significativas con un valor p de menos de 0,001. Estos resultados plantean la posibilidad de que las mutaciones del EGFR, además de ser predictores de la respuesta al gefitinib y al erlotinib, también pueden servir de pronóstico para una mejor supervivencia.

(*) Obsérvese que la frecuencia de mutación del EGFR en los pacientes estadounidenses no seleccionados, 1 de 61, parece ser menor en comparación con la frecuencia de la respuesta reportada al gefitinib de 10,4%. Esta diferencia tiene una importancia estadística moderada ($p = 0,025$ por medio del ensayo de chi-cuadrado). Por lo tanto este resultado podría aún deberse al azar, podría ser debido a una fracción de los que respondieron que no tienen mutaciones del EGFR, o podría deberse a una falla para detectar experimentalmente mutaciones del EGFR en esta colección de tumores. Si se compara la frecuencia de la mutación del EGFR en pacientes estadounidenses sensibles al gefitinib (5/5) con la frecuencia esperada de respuesta al gefitinib (10,4%), la probabilidad de chi-cuadrado es nuevamente menor a 10 - 12.

Ejemplo 4

Diseño del estudio:

Se realizó un estudio retrospectivo de una cohorte de pacientes con NSCLC que se refiere a la secuenciación del dominio quinasa del *EGFR* somático desde agosto de 2004 hasta enero de 2005 en el Hospital General de Massachusetts (MGH), el instituto del cáncer Dana-Farber (DFCI) y el Hospital Brigham y de Mujeres (BWH). Estas tres instituciones comprenden Dana-Farber/Partners CancerCare (DF/PCC), un centro de cáncer académico asociado que atiende aproximadamente 1200 pacientes con cáncer de pulmón al año. En agosto de 2004, se encontraba disponible la secuenciación del dominio quinasa del EGFR para uso clínico en DF/PCC. Los médicos podrían seleccionar que pacientes remitir a las pruebas, sin embargo, los pacientes necesitan tener especímenes tumorales apropiados y suficientes disponibles. Las células tumorales tenían que comprender al menos 50% del espécimen con base en el examen histológico por parte de los patólogos de referencia del MGH y BWH, y el espécimen tenía que ser de una resección, biopsia broncoscópica, o biopsia con aguja gruesa de un tumor primario o metastásico, o un bloque de células de fluido pleural. En raros casos, se consideraron adecuadas las muestras aspiradas con aguja fina. Las muestras podrían ser o bien tejido embebido en parafina o congelado. Debido a la baja incidencia de mutaciones del *EGFR* en pacientes con tumores de células escamosas (62) con este diagnóstico, no fueron elegibles para el ensayo.

Se identificaron los pacientes sometidos a los ensayos del *EGFR* utilizando el registro del caso del EGFR mantenido en el Laboratorio de Medicina Molecular (LMM), de la Escuela de Medicina de Harvard/Partners HealthCare Center for Genetics y Genomics (CLIA # 22D 1005307), el centro de análisis de diagnóstico en donde se realizaron e interpretaron todas las secuenciaciones. Se incluyeron todos los pacientes referidos para análisis del *EGFR* de DF/PCC con un diagnóstico de NSCLC durante el período del estudio.

Se recolectaron datos del paciente como edad, género y raza del sistema electrónico de registro médico. Se documentaron datos como el consumo de tabaco, antecedentes de cáncer, resultados de la secuenciación del

dominio quinasa del *EGFR*, y los posteriores planes de tratamiento del *EGFR*-TKI utilizando una revisión estructurada del médico. Específicamente, se obtuvieron la condición de fumador y antecedentes de cáncer de las notas médicas y de enfermería. Los ex fumadores se definieron como pacientes que habían dejado de fumar al menos un año antes del diagnóstico de cáncer de pulmón y los no fumadores se definieron como pacientes que habían fumado menos de 100 cigarrillos en su vida. Los fumadores que habían dejado de fumar al año de su diagnóstico o que estaban fumando en el momento del diagnóstico fueron clasificados como fumadores actuales. Se calcularon los paquetes fumados al año multiplicando el número de paquetes fumados por día por el número de años como fumador. Se obtuvieron los resultados de la histología tumoral y de la secuenciación del dominio quinasa del *EGFR* a partir de los reportes de patología. Todas las muestras de patología fueron revisadas en un solo sitio, ya sea en el MGH o en el BWH y se categorizó la histología utilizando el sistema de clasificación de la Organización Mundial de la Salud (OMS), (63). Los planes posteriores de tratamiento se obtuvieron de las notas del médico.

Los datos completos estaban disponibles para edad, género, histología del tumor, y el estado de mutación del *EGFR*. Se perdieron datos para la raza (12%), estado del tumor en el momento del análisis (4%), estatus de fumador (6%), tratamientos previos (5%), y planes de tratamiento posteriores del *EGFR*-TKI (11%). Este estudio fue aprobado por la Junta de Revisión Institucional de DF/PCC.

Secuenciación del gen *EGFR*:

Se cortaron secciones seriales ya sea de tejido tumoral embebido en parafina (FFPE), fijado en formalina o congelado y se colocaron sobre un portaobjetos de vidrio. Se identificó una región de tejido tumoral que consistía de al menos 50% de células tumorales viables por parte de un patólogo. Se extrajeron muestras de FFPE con etanol y xileno para remover la parafina. Se digirieron tanto las muestras de tejido congelados como las FFPE con proteinasa K durante la noche. Se extrajo el ácido desoxirribonucleico genómico (ADN) del tejido y de sangre entera periférica utilizando procedimientos estándar. Se extrajo el ADN genómico de muestras de saliva utilizando el kit para saliva de ADN Genotek-Oragene™.

Se amplificó el dominio quinasa del *EGFR* (exones 18 - 24 y las regiones intrónicas de flanqueo) en un conjunto de reacciones de la reacción en cadena de la polimerasa anidada (PCR). Los cebadores utilizados en las ampliaciones de PCR anidada se describen en la Tabla S1A y B y la SEQ ID 1 - 424 con la adición de secuencias universales a los extremos 5' de los cebadores (5'-tgtaaacgacgagccagt) (SEQ ID NO: 645). Se secuenciaron directamente los productos PCR en forma bidireccionalmente por medio de la secuenciación terminadora de coloración. Se realizó la PCR en una placa de 384 pozos en un volumen de 15 µl que contenía 5 ng de ADN genómico, MgCl₂ 2 mM, 0,75 µl de DMSO, Betaína 1 M, los dNTP 0,2 mM, 20 pmoles de cebadores, 0,2 µl de AmpliTaq Gold® (Applied Biosystems), 1X amortiguador (suministrado con AmpliTaq Gold). Las condiciones de los ciclos térmicos fueron las siguientes: 95° C durante 10 minutos; 95° C durante 30 segundos, 60° C durante 30 segundos, 72° C durante 1 minuto durante 30 ciclos; y 72° C durante 10 minutos. Se purificaron los productos PCR con perlas magnéticas Ampure® (Agencourt).

Los productos de la secuenciación se purificaron utilizando perlas magnéticas Cleanseq™ (Agencourt) y se separaron por medio de electroforesis capilar en un analizador de ADN ABI 3730 (Applied Biosystems). El análisis de secuencia se realizó por medio del Mutación Surveyor (SoftGenetics, State College, PA) y manualmente por medio de dos revisores. Se confirmaron las variantes de secuencia de ADN no sinónimas por medio del análisis de 3-5 reacciones PCR independientes de la muestra original de ADN genómico. Se analizaron muestras de sangre o saliva de individuos con variantes de secuencia de ADN no sinónimas para determinar si los cambios de secuencia eran exclusivos del tejido tumoral.

Análisis estadístico:

Se construyeron modelos de regresión logística para evaluar la asociación univariada entre características demográficas y clínicas del paciente y el estado de mutación del *EGFR*. Para identificar predictores significativos del status positivo de la mutación, se construyó un modelo de regresión logística multivariable que incluía variables independientes identificadas en estudios previos como predictores de mutaciones, específicamente género, raza, histología y condición de fumador. Se excluyeron seis pacientes de estos análisis debido a la falta de datos de mutaciones del *EGFR*, como resultado de un daño del PCR. Todos los análisis se realizaron utilizando un software estadístico SAS (versión 8.02, SAS Institute, Cary, NC).

Resultados:

Características de los pacientes:

Entre los 100 pacientes con NSCLC referidos para secuenciación del dominio quinasa del *EGFR* somático como parte del cuidado clínico del cáncer durante el periodo de estudio, la edad promedio fue de 60,7 años y el 63% eran mujeres (Tabla 4). La mayoría de los pacientes eran de raza blanca (76%) o asiáticos (7%), y tenían enfermedad metastática al momento de ordenar los análisis (67%). Casi todos los pacientes (94%) analizados para las mutaciones del *EGFR* tenían características de adenocarcinoma, adenocarcinoma con carcinoma broncoalveolar (BAC), o BAC puro. Aproximadamente un tercio de los pacientes nunca fueron fumadores. La terapia administrada antes de remitir a los pacientes para ensayos del *EGFR* incluían cirugía (50%), radioterapia torácica (22%),

quimioterapia (47%), y terapia destinada dirigida al EGFR (11%).

Mutaciones identificadas:

El tiempo promedio desde la remisión para los análisis hasta la disponibilidad de los resultados fue de 12 días hábiles. La mayoría de los especímenes remitidos estaban embebidos en parafina (74%). Seis de los 74 especímenes embebidos en parafina (8%) fallaron en las amplificaciones de la PCR, mientras que todos los 26 especímenes congelados fueron amplificados con éxito. Entre los 94 pacientes con resultados interpretables, se encontró que 23 (24%) tenían al menos una mutación en el dominio quinasa del *EGFR*, con dos de estos pacientes demostrando dos mutaciones puntuales cada uno, para un total de 25 mutaciones identificadas (Tabla 5). Entre los 23 pacientes con mutaciones, 9 (39%) tenían una o más mutaciones puntuales, 12 (52%) tenían supresiones que se superponen en el marco en el exón 19 y dos pacientes (9%) tenían duplicaciones en el exón 20. Las mutaciones puntuales fueron en los exones 18 y 21, e incluían cinco 2573T> G (L858R), y uno de cada uno de 2126A> T (E709V), 2155G> A (G719S), 2156G> C (G719A), 2327G> A (R776H), 2543C> T (P848L) y 2582T> A (L861Q). Una de las mutaciones puntuales (P848L) fue detectada tanto en el espécimen tumoral como en las células mononucleares obtenidas de un hisopo bucal. No se detectaron mutaciones en los exones 22, 23, o 24.

Predictores de mutaciones:

En nuestro ejemplo, no hubo una asociación significativa entre el estado de mutación del *EGFR* y la edad ($p = 0,61$), género femenino ($p = 0,92$), raza asiática ($p = 0,08$), o enfermedad metastática al momento de la remisión ($p = 0,43$, Tabla 4). Se encontró que ninguno de los 6 pacientes con histología tumoral diferente a adenocarcinoma tiene mutaciones. Entre los pacientes con características de adenocarcinoma, adenocarcinoma con BAC y BAC puro, no hubo asociación entre las características BAC/BAC y el estado de mutación del EGFR ($p = 0,35$).

Se encontró que ninguno de los 17 fumadores actuales tiene una mutación. Los no fumadores eran más propensos a tener una mutación del EGFR que los ex fumadores (*odds ratio* [OR] = 3,08, intervalo de confianza del 95% [IC]: 1,09 a 8,76). El número promedio paquetes fumados al año fue significativamente menor entre los pacientes con mutación positiva del *EGFR* (0,7 paquetes al año) frente a los pacientes con mutación negativa del *EGFR* (25,0 paquetes al año, $p < 0,001$). Por cada paquete adicional fumado al año, hubo una disminución del 4% en la probabilidad de tener una mutación (OR = 0,96, IC del 95% 0,93 - 0,99).

El número de paquetes fumados al año sigue siendo un predictor significativo del status de mutación después de controlar el género, la raza y la histología del tumor (OR = 0,96, IC del 95%: 0,93 - 0,99).

Uso posterior de la información de la prueba:

Los pacientes con mutación positiva del *EGFR* fueron significativamente más propensos a tener planes documentados para recibir un tratamiento posterior del EGFR-TKI (86%) que los pacientes con mutación negativa del *EGFR* (11%, $p < 0,001$). Los médicos documentaron que los resultados del *EGFR* afectaron su priorización de las terapias recomendadas en un 38% de los casos. Estos casos incluyeron 14 (61%) de los 23 pacientes con mutación positiva para quienes se recomendó terapia del EGFR-TKI antes de lo que hubiera tenido que hacerse para pacientes con mutación negativa, y 24 (34%) de los 71 pacientes con mutación negativa para quienes la terapia del EGFR-TKI no fue recomendada o fue recomendada más tarde de lo que hubiera sido para pacientes con mutación positiva.

Era más probable que el estado de mutación del EGFR cambiara la priorización de las opciones de tratamiento en pacientes con enfermedad metastática (54%) que en pacientes con enfermedad local o localmente avanzada (19%, $p = 0,003$). Teniendo en cuenta este resultado, se analizó el proceso de toma de decisión en pacientes con metástasis (Figura 10). Entre los 31 pacientes con enfermedad metastática cuyos resultados del análisis afectaron las recomendaciones de tratamiento, se les ofreció a cinco pacientes de primera línea con mutación positiva el tratamiento del EGFR-TKI y se les ofreció a seis pacientes de segunda línea con mutación positiva el tratamiento del EGFR-TKI en lugar de quimioterapia. Se alentó a veinte pacientes con mutación negativa a aplazar el tratamiento del EGFR-TKI hasta un tratamiento de tercera línea o superior con base en sus resultados negativos de las pruebas del EGFR. Entre los 26 pacientes con enfermedad metastática cuyos resultados de la prueba no afectaron las recomendaciones del tratamiento, dos pacientes con mutación negativa recibieron tratamiento de primera línea del EGFR-TKI a pesar de sus resultados negativos, nueve pacientes que incluían cuatro pacientes con mutación positiva recibieron tratamiento de segunda o tercera línea del EGFR-TKI y 15 pacientes, incluyendo dos pacientes con mutación positiva no recibieron una recomendación para un tratamiento del EGFR-TKI. Tres de los pacientes con enfermedad metastática participaron en ensayos para evaluar la terapia de primera línea del EGFR-TKI. Nueve de los pacientes con enfermedad metastática habían recibido previamente o estaban recibiendo los EGFR-TKI al momento del análisis del EGFR.

Discusión:

Se estudiaron los primeros 100 pacientes con NSCLC para experimentar selección por las mutaciones somáticas del *EGFR* como parte del cuidado clínico del cáncer en nuestra institución y se encontró que el análisis era factible e impactó significativamente el tratamiento de los pacientes con NSCLC. Era significativamente más probable que los

pacientes con mutaciones del *EGFR* recibieran recomendaciones para terapia del EGFR-TKI que los pacientes sin mutaciones. Los médicos ajustaron sus recomendaciones de tratamiento con base en los resultados del ensayo en más de un tercio de los casos, y era más probable que lo hicieran en pacientes con enfermedad metastática. En nuestra muestra de pacientes, los médicos utilizaron resultados positivos de la prueba del EGFR para ayudarlos a tomar la decisión para priorizar los EGFR-TKI sobre la quimioterapia para algunos pacientes, especialmente para tratamiento de primera o segunda línea. Sin embargo, los resultados negativos de las pruebas del EGFR no impidieron que los médicos administraran los EGFR-TKI a pacientes seleccionados. Muchos de los pacientes en quienes los resultados de la prueba no impactaron en la toma de decisiones clínicas tenían enfermedad resecada en etapa temprana o ya habían recibido una EGFR-TKI para enfermedad metastática al momento de la prueba. Esto es razonable ya que la utilidad de las EGFR-TKI como terapia adyuvante no se conoce y existe un beneficio con la terapia del EGFR-TKI en una cantidad pequeña de pacientes sin mutaciones identificadas del EGFR (65, 66 - 70, 71).

Nuestro estudio también proporciona evidencia de que el diagnóstico molecular puede mejorar la capacidad clínica para identificar pacientes con mutaciones del *EGFR*. Muchos oncólogos utilizan actualmente las características clínicas asociadas con mutaciones del *EGFR* y la respuesta a los EGFR-TKI para guiar el proceso de toma de decisiones para los pacientes con NSCLC. En efecto, nuestra población de pacientes remitidos para el ensayo del *EGFR* demostró un aumento de la prevalencia de tales características. Por ejemplo, 95% de los pacientes remitidos tenían adenocarcinoma o histología tumoral BAC, en comparación con 45% en la población general de NSCLC (72). Mientras que los no fumadores comprendían el 29% de nuestra población, la incidencia de los no fumadores en la población general de NSCLC se ha reportado como el 2 - 10%, y puede ser tan alta como 27% en mujeres con NSCLC (73 - 75). Del mismo modo, nuestra población consistía en únicamente de 17% de fumadores actuales, en comparación con la tasa de 38 - 75% de fumadores actuales entre los pacientes con NSCLC recientemente diagnosticados (75, 78 - 80). Nuestra población clínicamente seleccionada tuvo como consecuencia una tasa de mutación del *EGFR* de 24%, que es sustancialmente superior las tasas documentadas por nosotros y por otros grupos estadounidenses que analizaron muestras tumorales de NSCLC disponibles no seleccionadas (65 - 66, 81). Sin embargo, es importante señalar que mientras que los médicos parecían estar tratando de seleccionar pacientes para las pruebas que tenían las características clínicas de predicción de mutaciones del *EGFR*, la frecuencia de mutación era aún únicamente del 24%, destacando el hecho de que el diagnóstico molecular aumenta la información disponible para la toma de decisiones clínicas.

El estatus de fumador fue el predictor más fuerte del estado de mutación del EGFR en nuestros pacientes, con un aumento del historial de tabaquismo asociado a una significativa disminución de la probabilidad de albergar una mutación del EGFR, después de controlar predictores previamente descritos del estado de mutación. Nuestros resultados son consistentes con otra serie de casos que documentan la importancia de la condición de fumador en la probabilidad de mutaciones del *EGFR* (66, 69, 70, 81, 82). Al igual que la extremadamente baja prevalencia de mutaciones del EGFR en tumores de células escamosas (62) se han desplazado los esfuerzos de los análisis hacia los tumores de adenocarcinoma, puede ser apropiado enfocar los esfuerzos futuros sobre los pacientes con un historial bajo o nulo de tabaquismo. Sin embargo, los informes de mutaciones del EGFR en pacientes sin características clínicas típicas desaconsejan las limitaciones estrictas de los análisis (83). Cuando se examinan las otras características clínicas que se cree que están asociadas con las mutaciones, se encontró que la histología tumoral BAC y de raza asiática no tienen tendencias significativas para predecir el estado de mutación del *EGFR*. La falta de significancia estadística en estas asociaciones puede ser debido al pequeño tamaño de la muestra.

La prueba era factible y se adapta a las limitaciones de tiempo de cuidado clínico del cáncer. Casi todos los tumores enviados para análisis produjeron resultados interpretables. Las seis muestras que fallaron la amplificación por PCR estaban todas embebidas en parafina, mientras que ninguno de los especímenes congelados falló en la amplificación por PCR. Cuando está disponible, el tejido fresco congelado es el sustrato preferible para pruebas de mutación del *EGFR*.

Se han reportado cerca de 2.500 muestras de NSCLC que hasta el momento han sido sometidos a análisis de secuencia del EGFR parcial o completo. Nuestros pacientes demostraron mutaciones similares a los informes anteriores, con la superposición de las supresiones del exón 19 de 9 - 23 pares de bases y mutaciones puntuales que conducen a sustituciones de aminoácidos individuales en los exones 18 y 21. Se encontró que cinco de las mutaciones puntuales han sido descritas anteriormente (E709V, G719S, G719A, L858R, y L861Q). Se encontró que una de las mutaciones puntuales provoca una sustitución de aminoácido en un codón donde se ha descrito previamente una sustitución de aminoácidos diferente (R776H). Se encontró que cada una de las variantes E709V y R776H en combinación con una mutación conocida que sensibiliza al gefitinib que involucra al codón 719. Se encontró que la mutación P848L en el exón 21 tanto en muestras somáticas como bucales, lo que sugiere que puede ser una variante de línea germinal de significado incierto. La paciente era una mujer no fumadora con adenocarcinoma que tenía enfermedad estable durante 15 meses en tratamiento con gefitinib, antes de la prueba de mutación del EGFR. Cuando se reveló la mutación P848L, se le había encontrado recientemente que tiene una enfermedad progresiva y se inició el tratamiento con erlotinib. No existe información disponible acerca de la respuesta al erlotinib hasta el momento.

Los solapamientos de la supresión (2253_2276 del) describieron previamente las supresiones del exón 19. Las supresiones en nuestros pacientes se pueden clasificar en uno de dos grupos: aquellos codones que abarcan 747-

749 como mínimo (secuencia de aminoácidos LRE), y aquellos codones que abarcan 752 - 759 (Figura 11). El análisis de todas las supresiones del exón 19 reportados hasta la fecha sugiere que una amplia variedad de aminoácidos se puede eliminar de la región TK que abarca los codones 747 - 759. No parece ser un codón común requerida suprimido; sin embargo, todas las supresiones detectadas mantenían un residuo de lisina en la posición 745.

Una de las dos mutaciones del exón 20 están en una mujer no fumadora con adenocarcinoma recurrente que fue tratado con erlotinib después de que se realizó una prueba del EGFR y ha tenido enfermedad estable durante dos meses en este momento. El otro es un hombre ex-fumador con adenocarcinoma metastásico que fue tratado con un EGFR-TKI, pero no pudo tolerarlo debido a una erupción cutánea severa. La identificación de mutaciones del EGFR clínicamente relevantes en el exón 20 subraya la importancia de la secuenciación completa de la región TK del EGFR.

En conclusión, este estudio demuestra la viabilidad y la utilidad de la detección integral de la región TK del gen EGFR para mutaciones somáticas en pacientes con NSCLC como parte de la atención clínica del cáncer. El resultado de la prueba proporciona información útil con respecto a los predictores clínicos de respuesta a EGFR-TKI. Los fumadores actuales son menos propensos a albergar una mutación, al igual que los ex fumadores con un alto número de paquetes al año de historia de tabaquismo.

Ejemplo 5

Análisis del gen EGFR para cáncer de pulmón de células no pequeñas, un procedimiento operativo estándar.

Indicaciones clínicas:

Esta prueba está indicada para las personas con cáncer pulmonar de células no pequeñas.

Principio analítico

La prueba de gen *EGFR* es una prueba genética que detecta mutaciones en el dominio quinasa del *EGFR*. Se obtiene primero ADN a partir de una biopsia de tumor. La secuencia de ADN de 7 exones (18, 19, 20, 21, 22, 23, 24) del *EGFR* se determina entonces por medio de secuenciación génica bidireccional directa. Se compara luego la secuencia obtenida con la secuencia conocida del *EGFR* para identificar cambios de la secuencia de ADN. Si se detecta un cambio de secuencia de ADN en el tejido tumoral, la prueba se repetirá en la muestra de tejido original. Si no ha sido previamente reportado el cambio en alguien que responde al gefitinib o erlotinib, la prueba también se llevará a cabo con una muestra de sangre del individuo para determinar si la mutación es constitutiva (y por lo tanto probablemente ocurra normalmente un polimorfismo) o se ha producido de forma somática en el tejido tumoral.

Requerimientos del espécimen:

Se requiere un mínimo de 100 ng de ADN de la muestra de tejido. Nota: se pueden extraer cantidades extremadamente pequeñas de ADN de muestras de tejido. La concentración de este ADN no se puede cuantificar con precisión.

Control de calidad:

Controles utilizados

Se incluyen dos controles negativos (agua) y un control positivo (ADN humano) para cada exón en las reacciones PCR. El control negativo debe proceder a través de todo el procedimiento para asegurar que la secuencia obtenida no es el resultado de una contaminación. Un control positivo pGEM y un control de la matriz ABI están incluidos en la etapa de secuenciación.

Preparación y almacenamiento del control:

El control positivo para PCR es o bien ADN humano Clontech o ADN humano a partir de una muestra de sangre anónima y se almacena a 4° C. El control negativo para la reacción PCR es agua grado biológico de HyPure Molecular almacenada a temperatura ambiente. El control de la reacción de secuenciación positivo pGEM y el control de la matriz ABI se almacenan a 20° C.

Límites de tolerancia y pasos a seguir si se produce un error de control individual:

Si el control positivo de PCR falla, pero los controles negativos y las muestras pasan, los resultados de la PCR se considerarán que pasan y se llevara a cabo la secuenciación. Si un control negativo muestra evidencia de amplificación de ADN, se repetirá toda la reacción con una nueva parte alícuota de ADN del paciente. Si el control pGEM falla y las reacciones de las pruebas fallan, la ejecución de la secuenciación se repetirá con una segunda alícuota del producto de PCR. Si los controles de secuenciación fallan, pero las reacciones de prueba pasan la no necesita ser repetida la secuenciación. NOTA: Debido al baja rendimiento de la extracción de ADN a partir de muestras de tejidos embebidos en parafina, las reacciones de PCR externos a menudo no producen productos

visibles. Las reacciones internas PCR deben producir productos visibles. El tamaño del producto detectado en el gel debe ser comparado con los tamaños esperados (véase más adelante) para asegurar que el producto de PCR apropiada se ha obtenido. Si un producto de PCR interno no es visible en el gel, se deben repetir las fallas de la PCR específicas del exón.

- 5 Si la amplificación de la PCR de una muestra individual falla, una nueva ronda de PCR se debe intentar con un aumento de dos veces en el molde de entrada de ADN. Si la amplificación de la PCR falla nuevamente, se debe adquirir una nueva muestra de ADN para ese paciente si está disponible. Si la muestra era una muestra de tejido embebido en parafina, se deben obtener muestras adicionales. Si se encuentran disponibles, se deben obtener más muestras que las utilizadas para generar la muestra original y se debe permitir la digestión en proteinasa K durante tres noches.

Equipos y reactivos (todos los reactivos estables durante un año a menos que se indique lo contrario)

PCR y secuenciación (en general, el equipo y los reactivos para PCR y secuenciación son conocidos por aquellos capacitados en la técnica y pueden ser utilizados en este documento, como se señaló anteriormente).

Cebadores: (véanse las Tablas 6 y 7 a continuación)

Tabla 6: cebadores externos de PCR:

Exón	Secuencia del cebador hacia adelante, (5'→3')	SEQ ID NOS	Secuencia inversa del cebador, (5'→3')	SEQ ID NOS
18	TCAGAGCCTGTGTTTCTACAA	653	TGGTCTCACAGGACCACTGATT	646
19	AAATAATCAGTGTGATTCGTGGAG	654	GAGGCCAGTGCTGTCTCTAAGG	647
20	ACTTCACAGCCCTGCGTAAAC	655	ATGGGACAGGCACTGATTTGT	648
21	GCAGCGGGTTACATCTTCTTTC	656	CAGCTCTGGCTCACACTACCAG	649
22	CCTGAAGTCCGTCAGACTGAAA	657	GCAGCTGGACTCGATTTCT	650
23	CCTTACAGCAATCCTGTGAAACA	658	TGCCCAATGAGTCAAGAAGTGT	651
24	ATGTACAGTGCTGGCATGGTCT	659	CACTCACGGATGCTGCTTAGTT	652

TABLA 7: Cebadores internos de PCR:

Exón	Secuencia del cebador hacia adelante, (5'→3')	Secuencia inversa del cebador, (5'→3')	Longitud del producto (pb)
18	TCCAAATGAGCTGGCAAGTG (SEQ ID NO 660)	TCCCAAACACTCAGTGAAACAAA (SEQ ID NO 667)	397
19	GTGCATCGCTGGTAACATCC (SEQ ID NO 661)	TGTGGAGATGAGCAGGGTCT (SEQ ID NO 668)	297
20	ATCGCATTTCATGCGTCTTCA (SEQ ID NO 662)	ATCCCCATGGCAAACCTCTTG (SEQ ID NO 669)	378
21	GCTCAGAGCCTGGCATGAA (SEQ ID NO 663)	CATCCTCCCCTGCATGTGT (SEQ ID NO 670)	348
22	TGGC T CGTCTGTGTGTGTCA (SEQ ID NO 664)	CGAAAGAAAATACTTGCATGTCAGA (SEQ ID NO 671)	287
23	TGAAGCAAATTGCCAAGAC (SEQ ID NO 665)	TGACATTTCTCCAGGGATGC (SEQ ID NO 672)	383
24	AAGTGTGCGATCACCAATGC (SEQ ID NO 666)	ATGCGATCTGGGACACAGG (SEQ ID NO 673)	302
Enlazador del cebador F	tgtaaacgacgagccagt (SEQ ID NO 645)	Extremo 5' de todos los cebadores hacia adelante	18
Enlazador del cebador R	aacagctatgaccatg (SEQ ID NO 674)	Extremo 5' de todos los cebadores inversos	16

Precauciones

Tabla 8

Tarea		Instrucción(es)	Riesgo
1.	Montaje de la PCR	Uso de campana para PCR Uso dedicado de pipetas y puntas filtradas Únicamente abrir reactivos en la campana	Contaminación de la reacción de la PCR
2.	Uso de campana para PCR	No utilizar ninguna muestra o reactivo después de la PCR en la campana	Contaminación de la reacción de la PCR

Preparar de la mezcla de reacción PCR para PCR externo

- 5 Todos los procedimientos se realizan en una campana para PCR para ADN genómico, no la campana limpia.
1. Descongelar la Taq Gold y dNTP sobre hielo.
2. Preparar la mezcla maestra en un tubo (tubos Eppendorf o de 15 mL) usando la tabla presentada más abajo. Deben añadirse agua, betaina, amortiguador 10X, MgCl₂, DMSO, Taq Gold y dNTP en el orden indicado. Es muy importante mezclar los reactivos pipeteando hacia arriba y hacia abajo suavemente mientras que añade cada reactivo.
- 10 3. El ADN se debe añadir a la mezcla maestra antes de formar las alícuotas. Después de elaborar un gran volumen de mezcla maestra, se toma una alícuota de 96 µl (suficiente para 8 rxns) a un tubo separado para cada paciente o control. Se añaden 8 µl de ADN a razón de 5 ng/µl a los 96 µl de la mezcla maestra. Se puede añadir luego 13 µl a los pozos individuales de la placa o colocarlos en tubos de tiras y pipetearlos con una pipeta multicanal.
- 15 4. Para una placa de 384 pozos llena de reactivos, se elabora mezcla maestra suficiente para aproximadamente 415 reacciones.
5. Girar la placa de la mezcla maestra para deshacerse de las burbujas de aire.
6. Si se utiliza un grupo grande de cebadores, ayudaría a tenerlos en placas de 96 pozos con cebadores hacia adelante y cebadores inversos en placas separadas.
- 20 7. Añadir los cebadores utilizando una pipeta multicanal. Asegurarse de mezclar pipeteando suavemente hacia arriba y hacia abajo.
8. Girar la placa para eliminar cualquier burbuja de aire.
9. Utilizar el ciclo siguiente para amplificación.
- Nota: la PCR se realiza en placas de 384 pozos.

25

Tabla 9

Reactivo	Volumen por reacción (µL)
ddH ₂ O autoclavada	4,90
Betaina 5M	3,00
Amortiguador 10X	1,50
Cloruro de magnesio	1,50
DMSO	0,75
Taq	0,20
dNTP	0,15
Cebador 1 hacia adelante para la PCR (conc. 20 pmol/µL)	1,00
Cebador 2 inverso para la PCR (conc. 20 pmol/µL)	1,00
ADN (conc. 5 ng/µL)	1,00

Volumen total de reacción PCR	15,00
-------------------------------	-------

Tabla 10: Ciclo de amplificación de la PCR

Activación de Taq Gold	10 minutos	95° C	
Desnaturalización	30 segundos	95° C	30 ciclos
Hibridación	30 segundos	60° C	
	1 minuto	72° C	
Extensión	10 minutos	72° C	
Sostenimiento	∞	4° C	
Nota: No es necesaria una limpieza después de realizar la PCR externa.			

Preparación de la mezcla de reacción PCR para PCR interna

- 5 El montaje de la PCR interna es casi el mismo que el de la PCR externa con unas pocas excepciones.
 1. Elaborar el gran volumen de mezcla maestra como se describe para la PCR externa en la campana de PCR.
 2. Alícuota MM para tubos de 7 tiras y pipeta multicanal de 12 ul en la placa de 384 pozos.
 3. Añadir 1 ul de cada uno de los cebadores internos hacia delante e inversos. Sellar temporalmente la placa.
 4. Remover de la campana, desacelerar la placa y proceder a la zona de montaje posterior de la PCR.
- 10 5. Utilizar pipetas dedicadas para alícuotas de 1 ul de producto PCR externo en cada reacción.
 6. Sellar con calor y girar nuevamente.
 7. Correr el mismo ciclo de amplificación como en el caso externo.

Correr los productos PCR sobre un gel al 1% antes de la limpieza. Determinar los exones que pasan/fallan para repetir la PCR.

- 15 Limpieza de la PCR interna utilizando perlas magnéticas Ampure

Limpieza

1. Someter a agitación tipo vórtice la placa de perlas magnéticas Ampure hasta que no haya depósito de perlas.
2. Es muy importante que la temperatura de las perlas Ampure esté a temperatura ambiente.
3. Uso del protocolo Ampure de 384 pozos en el Biomek y cambio del volumen de reacción hasta 12uL para acomodar los reactivos utilizados para la limpieza. Si esto no se hace, se puede generar un error.
- 20 4. Después de que se completa el programa, se hidrata la placa 20 uL de ddH₂O autoclavada por pozo. Mientras se añade el agua, asegúrese de mezclar pipeteando suavemente hacia arriba y hacia abajo.
5. Girar la placa para eliminar cualquier burbuja de aire.
6. Colocar la placa en un imán para separar las perlas. Ahora se puede tomar 1 uL del ADN para configurar las reacciones de secuenciación. Guardar el resto a -20° C para un uso posterior.
- 25

Protocolo de secuenciación

Preparación de la mezcla de reacción para secuenciación

1. Descongelar BigDye 3.1 en un lugar oscuro, sobre hielo.
2. Preparar la mezcla maestra en un tubo (tubos Eppendorf o de 15 ml) utilizando la tabla que se muestra más abajo. Se deben añadir agua, amortiguador, DMSO y BigDye en el orden indicado.
- 30 3. Es muy importante mezclar los reactivos pipeteando suavemente hacia arriba y hacia abajo mientras se añade

cada reactivo.

4. Cuando se utiliza un cebador universal para secuenciación, también se puede añadir el cebador a la mezcla maestra en este momento. Si el cebador es único se debe añadir en forma individual después de que la mezcla maestra está en la placa de 384 pozos.

- 5 5. Usualmente, para una placa completa de 384 pozos de reacciones, elaborar suficiente mezcla maestra para aproximadamente 415 reacciones.

6. Una vez que se configura la mezcla maestra, se divide la mezcla en 8 pozos de tubos de tiras. (No utilizar depósitos para tomar alícuotas de la mezcla maestra. Eso sería un desperdicio de reactivos)

7. Ahora se puede utilizar una pipeta multicanal para tomar alícuotas de la mezcla maestra en la placa de 384 pozos.

- 10 8. Girar la placa de la mezcla maestra para deshacerse de las burbujas de aire.

9. Añadir el producto PCR para ser secuenciado utilizando una pipeta multicanal. Asegurarse de mezclar pipeteando hacia arriba y hacia abajo.

10. Girar la placa para eliminar cualquier burbuja de aire.

11. Utilizar el ciclo que se presenta a continuación para la amplificación.

15

Tabla 11

Reactivo	Volumen por reacción (µL)
ddH ₂ O autoclavada	4,38
Amortiguador ABI 5X	3,65
DMSO	0,50
BigDye 3.1 ABI	0,35
Concentración del cebador de secuenciación	0,12
ADN de la reacción PCR interna	1,00
Volumen total de la reacción	10,00

Tabla 12: Ciclo de amplificación para secuenciación

Desnaturalización	10 segundos	96°C	25 ciclos
Hibridación	5 segundos	50°C	
Extensión	4 minutos	60°C	
Sostenimiento	∞	4°C	

Limpieza a través de las perlas magnéticas Cleanseq

- 20 1. Someter a agitación tipo vórtice la placa de perlas magnéticas CleanSeq hasta que no haya depósito de perlas.

2. Utilizar el programa para placas de 384 pozos en el Biomek para limpiar las muestras.

3. Una vez corrido el programa, guardar la placa original a -20° C. La nueva placa con las muestras limpias está lista para ir al ABI 3730.

(Nota: Si los productos PCR son más cortos que 300 pb se debe diluir la muestra antes de colocarla en el ABI 3730)

- 25 Crear moldes Mutation Surveyor para la prueba del EGFR y guardarlos en LMM/secuenciación/secuencias - MS Review/EGFR.

Repetición de los criterios de los resultados

Todos los resultados positivos se repiten por medio de la amplificación y secuenciación del (de los) exón(es) específico(s) en la cual se haya detectado un cambio en la secuencia de ADN de una segunda alícuota del ADN del paciente derivado de la muestra de tejido original. Además, el ADN extraído de una muestra de la sangre del paciente se debe correr en forma paralela para comparar con el tejido del tumor si el cambio detectado de la secuencia no ha sido previamente detectado en un paciente que responde al gefitinib o al erlotinib.

Cualquier exón que no produjera una secuencia clara se repetirá ya sea a partir de la extracción, la PCR o la secuenciación, con base en los problemas técnicos específicos.

Parámetros del ensayo

Sensibilidad del ensayo: se han encontrado mutaciones somáticas del dominio quinasa del EGFR aproximadamente en 13% de los individuos con NSCLC (Paez J. G. et al., 2004). Además, se han encontrado mutaciones somáticas en el dominio quinasa del *EGFR* en 13/14 individuos (92,8%) con NSCLC que fueron sensibles al gefitinib (Paez JG et al., 2004, Lynch, et al., 2004). La validación de la sensibilidad técnica del ensayo demostró 100% de sensibilidad a las mutaciones conocidas y la validación de la plataforma de secuenciación en nuestro laboratorio muestra una sensibilidad del 100% (véase "precisión de la técnica" más abajo). Se determinó que la sensibilidad para la detección de la mutación de muestras mosaico es del 25% (es decir, se pueden detectar mutaciones heterocigotos cuando están presentes en un 50% de una mezcla de células). Se ha encontrado que hasta un 20% del tejido embebido en parafina no produce un ADN de alta calidad. No se puede obtener información de la secuencia a partir de estas muestras.

Especificidad del ensayo: hasta la fecha, la literatura publicada indica que ningún individuo con una mutación somática en EGFR fuera sensibles al gefitinib (11/11). La posibilidad de encontrar una mutación debido a un artefacto de secuenciación bidireccional es cercana al 0% (véase "precisión de la técnica" a continuación). Como tal, la especificidad del ensayo es aproximadamente del 100%.

Precisión de la técnica: la técnica de secuenciación de ADN es el estándar dorado en los diagnósticos moleculares. Este laboratorio utiliza el analizador de ADN ABI 3730 que ha reportado una precisión del 98,5%. Combinando esto con una secuenciación bidireccional, el análisis automatizado de un cromatograma con Mutation Surveyor, y el análisis manual de los falsos positivos, se ha logrado una tasa de precisión del 100%. Esto con base en el análisis de más de 100.000 bases de secuencias sin procesar. Para más información sobre esta evaluación, consulte nuestro manual de Programa de Aseguramiento de Calidad.

Nota: no se asume que estos resultados garantizan el 100% de precisión de esta plataforma. Se sabe que pueden presentarse errores en la secuenciación y por lo tanto, se reporta una precisión que es del 99,99% que ha sido encontrada por proyectos de secuenciación a gran escala (Hill et al., 2000).

Reproducibilidad del ensayo: debido a la precisión del ensayo, cuando se obtienen los resultados, tienen una reproducibilidad igual a la precisión del ensayo (99,99%). Sin embargo, a veces, el ensayo puede fallar debido a factores como los enumerados más adelante (véase Limitaciones del método) o debido a fallas en la secuenciación o de la PCR debido a razones técnicas sin explicación. En estos casos, no se obtienen resultados y se repite el ensayo hasta obtener un resultado o el espécimen del paciente se considera inaceptable. Las tasas específicas de fracaso de cada etapa del ensayo y de los especímenes pueden encontrarse en nuestro manual de Programa de Aseguramiento de Calidad.

Rango normal de los resultados: la secuencia normal del gen *EGFR* se puede encontrar en línea utilizando los números de acceso del GenBank: Nt_033968.5 (secuencia genómica) y NM_005228.3 (secuencia de ARNm).

Limitaciones del Método:

Grandes supresiones que abarcan uno o más exones no serán detectadas por el método de secuenciamiento, particularmente si están presentes en mutaciones de heterocigotos en el gen *EGFR* fuera del dominio quinasa, no serán detectados por el ensayo. Pueden estar presentes inhibidores en la muestra de ADN impidiendo la amplificación por PCR. El ADN degradado puede no producir datos analizables y puede requerirse reenviar espécimen. Las variaciones raras de la secuencia o estructuras secundarias de las secuencias del cebador objetivo podrían afectar la amplificación por PCR y por lo tanto la(s) mutación(es) podría(n) perderse en aquella región de un alelo.

Ejemplo 6

El gefitinib (Iressa) es un inhibidor de la tirosina quinasa que dirige al receptor del factor de crecimiento epidérmico (EGFR), e induce respuestas clínicas dramáticas en los cánceres de pulmón de células no pequeñas (NSCLC) con mutaciones de activación dentro del dominio quinasa del EGFR. Reportamos que estos EGFR mutantes activan selectivamente las rutas de señalización Akt y STAT, que promueven la supervivencia celular, pero no tienen ningún efecto sobre la señalización de Erk/MAPK, que induce la proliferación. Los NSCLC que expresan los EGFR mutantes sufrieron una amplia apoptosis después de la desactivación mediada por ARNpi del EGFR mutante o el tratamiento con inhibidores farmacológicos de la señalización de Akt y STAT, y eran relativamente resistentes a la

apoptosis inducida por fármacos quimioterapéuticos convencionales. Por lo tanto, los EGFR mutantes transducen selectivamente las señales de supervivencia de las cuales se vuelven dependientes los NSCLC; en consecuencia, la inhibición de esas señales por gefitinib puede ser la verdadera razón de respuestas clínicas notables.

Las tirosina quinasas receptoras de la familia del EGFR regulan funciones celulares esenciales incluyendo la proliferación, supervivencia, migración, y diferenciación, y parecen jugar un papel central en la etiología y progreso de los tumores sólidos (R. N. Jorissen et al., Exp. Cell Res. 284, 31 (2003), H. S. Earp, T. L. Dawson, X. Li, H. Yu, Breast Cancer Res. Treat. 35, 115 (1995)). El EGFR se sobreexpresa frecuentemente en cáncer de mama, de pulmón, de colon, de ovario, y en tumores cerebrales, lo que provocó el desarrollo de inhibidores farmacológicos específicos, tales como gefitinib, que afectan la actividad de quinasa del EGFR mediante el enlazamiento del bolsillo de ATP en el dominio catalítico (A- E. Wakeling et al., Cancer Res. 62, 5749 (2002)). Gefitinib ha inducido respuestas clínicas dramáticas en aproximadamente el 10% de los pacientes con NSCLC refractario a la quimioterapia (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)). Prácticamente todos los cánceres de pulmón sensibles a gefitinib albergan mutaciones somáticas en el dominio quinasa del EGFR, mientras que no se han observado mutaciones en los casos no sensibles (T. J. Lynch et al., N. Engl. J. Med. 350, 2129 (2004), J. G. Paez et al., Science 304, 1497 (2004)). Estas mutaciones heterocigotas incluyen supresiones pequeñas en el marco y sustituciones en sentido erróneo agrupadas dentro del bolsillo de enlazamiento del ATP.

Utilizando transfecciones transitorias de los EGFR mutantes, se mostró previamente que ambos tipos de mutaciones conducen a una mayor activación del receptor de depende de EGF, tal como se mide por la autofosforilación de Y1068, uno de los sitios de fosforilación prominentes del terminal C del EGFR. (T. J. Lynch et al., N. Engl. J. Med. 350, 2129 (2004)).

Para permitir los estudios de diferencias cualitativas en la señalización por los EGFR mutantes, hemos generado líneas estables de células epiteliales mamarias de ratón no transformadas (NMuMg) que expresan EGFR de tipo silvestre o mutante, y se analizó la autofosforilación mediada por EGF de múltiples residuos de tirosina vinculados a la activación de distintos efectores secuencia abajo (R. N. Jorissen et al., Exp. Cell Res. 284, 31 (2003)). Se generaron líneas celulares que expresan, ya sea EGFR de tipo silvestre o una de las dos mutaciones recurrentes detectadas en los tumores de los pacientes que responden a gefitinib: la mutación en sentido erróneo L858R y la supresión en el marco de 18 pb, delL747-P753insS. No se observaron patrones de fosforilación de tirosina significativamente diferentes entre los EGFR de tipo silvestre y los dos mutantes en diferentes sitios del terminal C. La fosforilación inducida por EGF de Y1045 y Y1173 era virtualmente indistinguible entre los EGFR de tipo silvestre y mutante, mientras que la fosforilación de Y992 y Y1068 se incrementó sustancialmente en ambos mutantes. Curiosamente, Y845 fue altamente fosforilado en el mutante de sentido erróneo L858R, pero no en el de tipo silvestre o el mutante de supresión, y por lo tanto parece ser único para distinguir entre los dos tipos de mutaciones del EGFR. El patrón de fosforilación de la tirosina inducida por EGF diferencial observado con los receptores de tipo silvestre y mutante fue reproducible en células COS7 transfectadas transitoriamente, asegurando contra efectos potenciales específicos del tipo de célula.

Por lo tanto, los EGFR mutantes sensibles gefitinib transducen señales que son cualitativamente distintas de aquellas mediadas por EGFR de tipo silvestre. Estas diferencias pueden resultar directamente de las alteraciones estructurales en el bolsillo catalítico afectando la especificidad del sustrato, o de las interacciones alteradas con las proteínas accesorias que modulan la señalización del EGFR.

El establecimiento de líneas celulares transfectadas establemente con los EGFR mutantes hizo posible comparar el estado de fosforilación de los principales objetivos secuencia abajo del EGFR en un fondo celular compartido. La activación inducida por EGF de Erk1 y Erk2, a través de Ras, Akt a través de PLC γ /PI3K y de STAT3 y STAT5 a través de JAK2, son rutas esenciales secuencia abajo que median los efectos oncogénicos del EGFR (R. N. Jorissen et al., Exp. Cell Res. 284, 31 (2003)). La activación de Erk inducida por EGF era esencialmente indistinguible entre las células que expresan EGFR de tipo silvestre o cualquiera de los dos que activan mutantes del EGFR. Por el contrario, la fosforilación de Akt y STAT5 se elevó sustancialmente en las células que expresan cualquiera de los EGFR mutantes. El aumento de la fosforilación de STAT3 se observó también en forma similar en células que expresan EGFR mutantes. La activación de Erk inalterada por los EGFR mutantes es coherente con la ausencia de aumento de la fosforilación de Y1173, un importante sitio de acoplamiento para los adaptadores Shc y grb-2 que conduce a la activación de Ras y la posterior fosforilación de Erk (R. N. Jorissen et al., Exp. Cell Res. 284, 31 (2003)). El aumento de fosforilación de Akt y STAT después de la activación de los EGFR mutantes es consistente con el aumento en fosforilación de Y992 y Y1068, ambos de los cuales han sido previamente vinculados a la activación de Akt y STAT (R. N. Jorissen et al., Exp. Cell Res. 284, 31 (2003)). Por lo tanto, la autofosforilación selectiva inducida por EGF de los residuos de tirosina del terminal C dentro de los mutantes del EGFR está bien correlacionada con la activación selectiva de las rutas de señalización secuencia abajo.

Para extender estas observaciones a las células de cáncer de pulmón en las cuales las mutaciones del EGFR parecen dirigir la tumorigénesis, hemos estudiado las líneas derivadas de cinco tumores de NSCL. NCI-H1975 transporta la mutación de sentido erróneo heterocigota recurrente L858R y NCI-H1650 tiene la supresión en el marco delE746-A750, mientras que el NCI-358, NCI-H1666 y NCI-H1734 expresan *EGFR* de tipo silvestre. Al igual que en las células transfectadas, la autofosforilación inducida por EGF de Y992 y Y1068 se elevó marcadamente en

las dos líneas con mutaciones del EGFR endógeno, al igual que la fosforilación de Akt y STAT5, pero no Erk.

La actividad oncogénica del EGFR refleja la activación de señales que promueven tanto la proliferación celular y como la supervivencia celular (S. Grant, L. Qiao, P. Dent, *Front. Biosci.* 7, d376 (2002)). Mientras que estas rutas exhiben solapamiento, la activación mediada por Ras de las quinasas Erk contribuye sustancialmente a la actividad proliferativa del EGFR, mientras que la activación de Akt y STAT está vinculada en gran medida a una función anti-apoptótica (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)). Las dos líneas celulares de cáncer de pulmón que albergan mutaciones del EGFR mostraron una respuesta proliferativa a EGF a bajas concentraciones de suero que no se observó en las células con receptores de tipo silvestre. Sin embargo, su tasa de proliferación y la densidad de las células en confluencia fueron comparables a las concentraciones séricas normales.

ARNpi

En contraste, las rutas apoptóticas eran marcadamente diferentes en las células de cáncer de pulmón con EGFR mutantes: la inactivación específica mediada por ARNpi del EGFR mutante en estas líneas celulares resultó en una rápida y masiva apoptosis. Alrededor del 90% de las células NCI-H1975 transfectadas con ARNpi específico de L858R murieron en un lapso de 96 horas, como lo hicieron las células NCI-H1650 transfectadas con ARNpi específico de delE746-A750. El ARNpi específico, para cualquier mutación del EGFR no tuvo ningún efecto sobre las células que expresan la mutación alternativa, y ARNpi que dirige tanto a *EGFR* de tipo silvestre como mutante tuvo un efecto mínimo sobre la viabilidad de las células que expresan sólo los receptores de tipo silvestre, pero indujo una rápida muerte celular en líneas que expresan mutantes del EGFR. La capacidad de los ARNpi para dirigir específicamente a los alelos correspondientes del EGFR fue confirmada en las células transfectadas COS7 por medio de inmunotransferencia. Por lo tanto, parece esencial la expresión de los EGFR mutantes para la supresión de las señales pro-apoptóticas en los cánceres de pulmón que albergan estas mutaciones. El hecho de que las células de cáncer de pulmón que expresan sólo los receptores de tipo silvestre no muestran una dependencia similar sobre la expresión del EGFR también puede dar cuenta de la relativa insensibilidad al gefitinib de tumores humanos que sobreexpresan EGFR de tipo silvestre.

La efectividad de gefitinib en cánceres de pulmón que albergan EGFR mutantes puede reflejar tanto su inhibición de las rutas anti-apoptóticas críticas sobre la cuales estas células se han vuelto estrictamente dependientes, así como las propiedades bioquímicas alteradas de los receptores mutantes. Hemos informado anteriormente que los EGFR mutantes son más sensibles a la inhibición de gefitinib de autofosforilación dependiente de EGF que los receptores de tipo silvestre (T. J. Lynch et al., *N. Engl. J. Med.* 350, 2129 (2004)). También se observó este aumento de sensibilidad al fármaco por parte de los receptores mutantes, tanto para la activación de Erk como de STAT5. Por lo tanto, mientras que la señalización inducida por EGF por parte de los receptores mutantes demuestra activación selectiva de los efectores secuencia abajo a través de eventos de autofosforilación diferenciales, su mayor inhibición por gefitinib es uniforme, y puede reflejar enlazamiento alterado del fármaco con el bolsillo del ATP mutante.

Para establecer la relevancia del aumento de señalización de Akt y STAT en superviviente del NSCL mediado por EGFR, nos centramos en estas rutas con inhibidores farmacológicos específicos. Las células de cáncer de pulmón que albergan mutaciones del EGFR eran 100 veces más sensibles al gefitinib que las células con receptores de tipo silvestre. Las células que expresan EGFR mutantes también eran más sensibles a la inhibición farmacológica de la señalización de Akt o STAT que las células que expresan sólo EGFR de tipo silvestre. Mientras que las células de cáncer de pulmón mutantes del EGFR mostraron aumento de la sensibilidad a la interrupción de las señales anti-apoptóticas mediadas por Akt/STAT, demostraron notable aumento de la resistencia a las señales de muerte celular inducidas por los agentes quimioterapéuticos comúnmente utilizados doxorubicina y cisplatino, y el ligando de FAS pro-apoptótico.

La señalización mejorada de Akt/STAT en las células con EGFR mutante puede proporcionar por lo tanto, un objetivo terapéutico adicional, mientras aumenta la posibilidad de que la quimioterapia convencional pueda ser menos efectiva contra estos tumores.

Se ha propuesto "adicción oncogénica" para explicar la apoptosis de las células cancerosas después de la supresión de una señal proliferativa de la cual se han vuelto dependientes (I. B. Weinstein, *Science* 297, 63 (2002)). Curiosamente, imatinib (Gleevec) activa eficientemente la muerte celular en leucemias mieloides crónicas que expresan el producto de la translocación BCR-ABL y en los tumores del estroma gastrointestinal que expresan la activación de mutaciones de c-Kit, los cuales exhiben frecuentemente activación constitutiva de STAT que es inhibida eficazmente por el fármaco (T. Kindler et al., *Leukemia* 17, 999 (2003), G. P. Paner et al., *Anticancer Res.* 23, 2253 (2003)). Del mismo modo, en las células de cáncer de pulmón con mutaciones de la quinasa del EGFR, la sensibilidad al gefitinib puede resultar en gran parte de su inhibición efectiva de señales anti-apoptóticas esenciales transducidas por el receptor mutante.

Materiales y métodos

Inmunotransferencia

Los lisados de células cultivadas se prepararon en solución de lisis RIPA enfriada con hielo (Triton X-100 al 1%, SDS al 0,1%, Tris-HCl 50 mM, pH 7,4, NaCl 150 mM, EDTA 1mM, EGTA 1 mM, β -glicerol-fosfato, NaF 10 mM, ortovanadato de sodio 1 mM, que contienen inhibidores de proteasa. Se removieron los residuos por centrifugación en una microcentrífuga a 12.000 xg durante 10 min a 4° C. Se calentaron a ebullición los lisados clarificados en gel
 5 que carga amortiguador y se separaron por medio de SDS-PAGE al 10%. Se electrotransfirieron las proteínas a nitrocelulosa y se detectaron con anticuerpos específicos, seguido por incubación con anticuerpo de cabra secundario conjugado con peroxidasa de rábano (señalización celular (Beverly, MA; 1: 2000) y desarrollo con quimioluminiscencia mejorada (DuPont NEN), seguido por autorradiografía). Se obtuvieron los anticuerpos fosfo-EGFR Y845, Y992, Y1045, Y1068, fosfo-STAT5 (tyr694), fosfo-Akt (Ser473), fosfo-Erk1/2 (Thr202/Tyr204), Akt,
 10 STAT5 y ERK1/2 a través de New England Biolabs (Beverly, MA). Se obtuvo anticuerpo Ab-20 del EGFR total a través de NeoMarkers (Fremont, CA). El anticuerpo Y1173 de fosfo-EGFR a través de Upstate Biotechnology (Lake Placid, NY) y el anticuerpo total de fosfotirosina PY-20 a través de Transduction Laboratories (Lexington, KY). Todos los anticuerpos se utilizaron en una dilución de 1:1000.

Vectores de expresión del EGFR

- 15 Los constructos de expresión de longitud completa del EGFR que codifican las mutaciones L858 o delL747-P753insS de tipo silvestre fueron sub-clonados utilizando métodos estándar en el plásmido pUSEamp. Se confirmaron todos los constructos por medio de análisis de secuencia de ADN.

Líneas celulares y transfecciones

- 20 Se cultivaron células COS7 y células NMuMg (epiteliales mamarias de ratón normales) en DMEM (medio de Eagle modificado de Dulbecco) con suero de ternera fetal al 10% en presencia de L-glutamina 2mM y 50 U/ml de penicilina/estreptomicina. Se obtuvieron las líneas celulares de cáncer de pulmón humano NCI-H358, NCI-H1650, NCI-H1734, NCI-H1666 y NCI-H1975 a través de la colección American Type Culture Collection y se cultivaron en RPMI1640 con suero bovino fetal al 10%, L-glutamina 2 mM, 50 U/ml de penicilina/estreptomicina y piruvato de sodio 1mM. Se denominan en el texto, en forma abreviada, como H358, H1650, H1734, H1666 y H1975,
 25 respectivamente. La transfección transitoria de células COS7 se realizó utilizando Lipofectamina 2000 (Invitrogen, Carlsbad, CA). Se transfectó el plásmido (1 μ g) en células con un 80% de confluencia en una placa de 10 cm. Después de 12 horas, se recogieron las células y se volvieron a sembrar en placas de 12 pozos en ausencia de suero. Al día siguiente, se estimularon las células con 30 ng/ml de EGF. Se prepararon las líneas celulares estables NMuMg mediante cotransfección de los constructos de expresión del EGFR con el plásmido seleccionable por el fármaco pBABE puro, seguido por selección en 3 μ g/ml de puomicina. Se utilizaron reservorios de células
 30 resistentes al fármaco para el análisis. La expresión del EGFR en las células transfectadas de forma estable se confirmó mediante inmunotransferencia.

"Desactivación" mediada por ARNpi de la expresión del EGFR

- 35 Se diseñó ARNpi para L858R del EGFR para dirigir la secuencia de nucleótidos CACAGATTTTGGGCGGGCCAA (SEQ ID NO. 688), mientras que se utilizó la secuencia GCTATCAAAACATCTCCGAAA (SEQ ID NO: 689) para la delE745-A750 (Qiagen, Valencia, CA). Para dirigir todas las formas del EGFR, se obtuvo ARNpi preparado comercialmente correspondiente a EGFR de tipo silvestre humano a través de Dharmacon (Lafayette, CO). Se llevó a cabo la transfección de los ARNpi con Lipofectamina 2000 (Invitrogen) según las instrucciones del fabricante. Las células se ensayaron para determinar la viabilidad después de 96 horas usando el ensayo de MTT.

40 Ensayo de apoptosis

- Se sembraron 10.000 células en pozos individuales de una placa de 96 pozos. Después de 6 horas, se cambió el medio y se mantuvieron las células en presencia de concentraciones crecientes de doxorrubicina (Sigma, St. Louis, MO), cisplatino (Sigma), ligando Fas (activador humano, clon CH11; Upstate Biotechnology), Ly294002 (Sigma), o AG490 (Calbiochem, La Jolla, CA). Después de 96 horas, se determinó la viabilidad de las células utilizando el
 45 ensayo de MTT. Para la inmunocoloración de la caspasa, se sembraron 10.000 células sobre cubreobjetos de 10 mm. Al día siguiente fueron transfectadas con ARNpi (véase la sección anterior para más detalles). Después de 72 horas, se fijaron las células en paraformaldehído al 4% a temperatura ambiente durante 10 min. Posteriormente, se permeabilizaron durante 5 min en Triton X-100 al 0,5% y se bloquearon durante 1 h en suero normal de cabra al 5% (NGS). Se incubaron los cubreobjetos durante la noche a 4° C en anticuerpo primario (caspasa-3 Asp 175 5A1
 50 escindida a través de Cell Signaling) con una dilución de 1:100. Al día siguiente, se lavaron los cubreobjetos 3 veces en PBS y se incubaron durante 1 hora con el anticuerpo secundario (conjugado de Texas-red anti-conejo de cabra; de Jackson ImmunoResearch, West Grove, PA) con una dilución 1:250 en suero de cabra normal al 5% y 0,5 μ g/ml de DAPI (4',6-diamidino-2-fenilindol). Después de 3 lavadas en PBS, se montaron los cubreobjetos con reactivo contra la decoloración ProLong Gold de Molecular Probes (Eugene, OR).

55 Ensayo de viabilidad celular

Se añadieron 10 μ l de solución de 5 mg/ml de MTT (azul de tiazolilo; Sigma) a cada pozo de una placa de 96 pozos. Después de 2 horas de incubación a 37° C, se removió el medio y se solubilizó la MTT mediante la adición de 100 μ l de isopropanol ácido (HCl 0,1 normal) a cada pozo. Se determinó la absorbancia espectrofotométricamente a 570

nm.

Curva de crecimiento

- 5 Se obtuvieron curvas de crecimiento para las células H-358, H-1650, H-1734, y H-1975 sembrando 1.000 células en pozos individuales de placas de 96 pozos. Se sembró en placa cada línea celular en 8 pozos separados. Los siguientes días, se fijaron las células en formaldehído al 4% y se tiñeron con solución de cristal violeta al 0,1% (p/v). Se solubilizó luego el cristal violeta 100 µl de ácido acético al 10%, y se midió la absorbancia a 570 nm utilizando un lector de placas para determinar el número relativo de células.

Identificación de la mutación

- 10 Para identificar líneas de células esporádicas de NSCLC que albergan mutaciones dentro del EGFR, se secuenciaron los exones 19 y 21 dentro de un panel de 15 líneas celulares de NSCLC, como se describió anteriormente. Se seleccionaron líneas celulares para análisis con base en su derivación a partir de tumores de histología broncoalveolar, independientemente de los antecedentes de tabaquismo (NCI-H358, NCI-H650, NCI-H1650), o de adenocarcinomas que surgen dentro de los no fumadores (NCI-H1435, NCI-H1563, NCI-H1651, NCI-H1734, NCI-H1793, NCI-H1975, NCI-H2291, NCI-H2342, NCI-H2030, NCI-H1838, NCI-H2347, NCI-H2023). Se ha reportado que NCI-H1666 albergar únicamente EGFR de tipo silvestre (véanse los ejemplos anteriores). Todas las
- 15 líneas celulares están disponibles a través de la American Type Culture Collection.

Referencias

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.
- 20 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
- 25 5. Luetke 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.
- 30 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.
- 35 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 phosphorylation 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.
- 40 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.
- 45 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.
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 Oncol* 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 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.
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).
- 5 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).
- 10 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).
- 15 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).
- 20 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).
- 25 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.
63. Franklin WA: Diagnosis of lung cancer: pathology of invasive and preinvasive neoplasia. *Chest* 117:80S-89S, 2000.
- 30 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
- 35 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.
- 40 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.

5 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.

10 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.

15 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

20 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

25 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

30 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

Tabla 1. Características de 9 pacientes con cáncer de pulmón de células no pequeñas y sensibilidad al gefitinib

Paciente No.	Sexo	Edad al inicio de la terapia de gefitinib	Tipo patológico*	No. de regímenes anteriores	Estatus de fumador†	Duración de la terapia	Supervivencia total‡	Mutación del EGFR§	Respuesta¶
		años				meses			
1	F	70	BAC	3	Nunca	15,6	18,8	Sí	Mayor; lesiones pulmonares mejoradas
2	M	66	BAC	0	Nunca	>14,0	>14,0	Sí	Mayor; lesiones pulmonares bilaterales mejoradas
3	M	64	Adeno	2	Nunca	,6	12,9	Sí	Parcial; lesiones pulmonares mejoradas
4	F	81	Adeno	1	Ex-fumador	>13,3	>21,4	Sí	Menor; enfermedad pleural mejorada
5	F	45	Adeno	2	Nunca	>14,7	>14,7	Sí	Parcial; lesiones hepáticas mejoradas
6	M	32	BAC	3	Nunca	>7,8	>7,8	Sí	Mayor; lesiones pulmonares mejoradas
7	F	62	Adeno	1	Ex-fumador	>4,3	>4,3	Sí	Parcial; lesiones pulmonares y hepáticas mejoradas

8	F	58	Adeno	1	Ex-fumador	11,7	17,9	Sí	Parcial; lesiones hepáticas mejoradas
9	F	42	BAC	2	Nunca	>33,5	>33,5	No	Parcial; nódulos pulmonares mejorados

* Adenocarcinoma (Adeno) con cualquier elemento del carcinoma broncoalveolar (BAC) se enumera como BAC

† El estatus de fumador se definió como de ex-fumador si el paciente no había fumado ningún cigarrillo en un lapso de 12 meses antes de entrar y nunca si el paciente había fumado menos de 100 cigarrillos en toda su vida.

‡ Se midió la supervivencia total desde el comienzo del tratamiento con gefitinib hasta la muerte.

§ EGFR denota el gen para el receptor del factor de crecimiento epidérmico

¶ Se evaluó una respuesta parcial con el uso de criterios de evaluación de respuesta en tumores sólidos; se evaluaron las respuestas mayor y menor por parte de dos médicos en pacientes en quienes no se podía medir la respuesta con el uso de estos criterios.

Tabla 2. Mutaciones somáticas en el dominio tirosina quinasa de EGR en pacientes con cáncer de pulmón de células no pequeñas

Paciente	Mutación	Efecto de la mutación
Pacientes con sensibilidad al gefitinib		
Paciente 1	Supresión de 15 nucleótidos (2235 - 2249)	Supresión en el marco (746 - 750)
Paciente 2	Supresión de 12 nucleótidos (2240 - 2251)	Supresión en el marco (747 - 751) e inserción de un residuo de serina
Paciente 3	Supresión de 18 nucleótidos (2240 - 2257)	Supresión en el marco (747 - 753) e inserción de un residuo de serina
Paciente 4	Supresión de 18 nucleótidos (2240 - 2257)	Supresión en el marco (747 - 753) e inserción de un residuo de serina
Paciente 5	Sustitución de G por T en el nucleótidos 2573	Sustitución de aminoácido (L858R)
Paciente 6	Sustitución de G por T en el nucleótidos 2573	Sustitución de aminoácido (L858R)
Paciente 7	Sustitución de A por T en el nucleótidos 2582	Sustitución de aminoácido (L861Q)
Paciente 8	Sustitución de T por G en el nucleótidos 2155	Sustitución de aminoácido (G719C)
Pacientes sin exposición al gefitinib*		
Paciente A	Supresión de 18 nucleótidos (2240 - 2257)	Supresión en el marco (747 - 753) e inserción de un residuo de serina
Paciente B	Supresión de 15 nucleótidos (2235 - 2249)	Supresión en el marco (746 - 750)

* Entre los 25 pacientes sin exposición al gefitinib (15 con cáncer broncoalveolar, 7 con adenocarcinoma, y 3 con carcinoma de células grandes), 2 (Pacientes A y B) – quienes tenían cáncer broncoalveolar – tenían mutaciones del *EGFR*. No se encontraron mutaciones en 14 líneas celulares de cáncer de pulmón que representan tipos histológicos diversos: cáncer pulmonar de células no pequeñas (6 especímenes), cáncer pulmonar de células pequeñas (6 especímenes), carcinoma bronquial (1 espécimen), y un tipo desconocido (un espécimen). Las variantes polimórficas identificadas dentro del *EGFR* incluyen las siguientes: la sustitución de A por G en el nucleótido 1562, la sustitución de A por T en el nucleótido 1887, y una variante de línea germinal de significado funcional desconocido, la sustitución de A por G en el nucleótido 2885 dentro del dominio de tirosina quinasa.

Tabla 4: Características de la población entre 100 pacientes analizados por las mutaciones del EGFR como parte del cuidado del NSCLC

Características	Frecuencia
Edad promedio, años (desviación estándar)	60,7 (11,0)
Mujer	63
Raza	
Blanca	76
Asiática	7
Otra	5
Desconocida	12
Etapa al momento del ensayo	
I	15
II	4

III	10
IV	67
Desconocida	4
Histología	
BAC puro	1
Adenocarcinoma con características de BAC	24
Adenocarcinoma	69
NSCLC, todos los otros subtipos	6
Estatus de fumador	
Actual	17
Ex-fumador	48
Nunca	29
Desconocido	6
Cantidad promedio fumada por fumadores actuales y ex-fumadores, paquetes por año (desviación estándar)	39,0 (32,3)
Tiempo promedio de diagnóstico para la prueba del <i>EGFR</i> , meses (desviación estándar)	18,7 (78,4)
Tratamiento previo de quimioterapia	47
Tratamiento previo dirigido al <i>EGFR</i>	11

BAC = carcinoma broncoalveolar, EGFR = receptor del factor de crecimiento epidérmico

Tabla 5: Mutaciones identificadas del gen somático para el receptor del factor de crecimiento epidérmico

Paciente	Género	Histología	Paquetes fumados por año	Exón	Cambio de nucleótido	Cambio de aminoácido
1	F	Adeno	0	18 18	2126A>T 2155G>A	E709V G719S
2	F	A+BAC	60	18 20	2156G>C 2327G>A	G719A R776H
3	F	A+BAC	0	19	2235_2249 del	K745_A750 del ins K
4	M	A+BAC	0	19	2235_2249 del	K745_A750 del ins K
5	F	Adeno	5	19	2235_2249 del	K745_A750 del ins K
6	M	Adeno	Desconocido	19	2235_2249 del	K745_A750 del ins K
7	F	Adeno	0	19	2236_2250 del	E746_A750 del
8	M	Adeno	45	19	2236_2250 del	E746_A750 del
9	F	Adeno	Desconocido	19	2236_2250 del	E746_A750 del
10	M	A+BAC	12	19	2237_2255 del ins T	E746_S752 del ins V
11	M	Adeno	1	19	2239_2248 del ins C	L747_A750 del ins P

(continuación)

Paciente	Género	Histología	Paquetes fumados por año	Exón	Cambio de nucleótido	Cambio de aminoácido
12	M	A+BAC	0	19	2239_2251 del ins C	L747_T751 del ins P
13	F	Adeno	30	19	2253_2276 del	T751_1759 del ins T
14	F	Adeno	0	19	2254_2277 del	S752_I759 del
15	F	Adeno	0	20	2303_2311 dup	D770_N771 ins SVD
16	M	Adeno	5	20	2313_2318 dup CCCCCGA	P772_H773 dup
17	F	Adeno	0	21	2543C>T	P848L*
18	M	BAC	0	21	2573T>G	L858R
19	F	A+BAC	0	21	2573T>G	L858R
20	M	A+BAC	1	21	2573T>G	L858R
21	F	Adeno	0	21	2573T>G	L858R
22	F	Adeno	15	21	2573T>G	L858R
23	F	Adeno	0	21	2582T>A	L861Q

Adeno = Adenocarcinoma, Adeno + BAC = Adenocarcinoma con características de carcinoma broncoalveolar, BAC = Carcinoma broncoalveolar puro

* Esta mutación fue identificada como una variante de línea germinal

Tabla S1A: Cebadores para amplificación del EGFR seleccionado y exones de tirosina quinasa receptora (SEQ ID NO: 1 - 212)

Gen	RefSeq	Exón	SEQ ID NO	F Anidado	R Anidado
ALK	NM 004304	24	1,2	GGAAATATAGGGAAGGGAAGGAA	TGACAGGGTACCAGGAGATGA
ALK	NM 004304	25	3,4	CTGAACCGCCAAGGACTCAT	TTTCCCTCCCTACTAAACACACG
AXL	NM 021913	19	5,6	ACTGATGCCCTGACCCCTGTT	CCCATGGTTCCTCCACTCTT
CSF1R	NM 005211	18	7,8	AGGGACTCCAAGCCATGTG	CTCTCTGGGGCCATCCACT
CSF1R	NM 005211	19	9,10	CATTGTCAAGGCAATGTAAGTG	CTCTCACCACCCCTCGCTGT
DDR1	NM 013994	15	11,12	ACATGGGGAGCCAGAGTGAC	TGCAACCCAGAGAAAAGTG
DDR2	NM 006182	16	13,14	TGAGCTTTCACCCCTAGTTGTG	GTTCCTCTCTGCTGCTCA
DKFZp761P1010	NM 018423	3	15,16	TGTCTTGTGTTTGTGAAGATTCC	TGCAGACAGATGACAAACATGAA
EGFR	NM 005228	2	17,18	TGGGTGAGTCTCTGTGTGGAG	CATTGCCATAGCAAAAATAAACACA
EGFR	NM 005228	3	19,20	GTTTCAACTGGGCGTCTA	CCTTCTCCGAGGTGGAATTG
EGFR	NM 005228	4	21,22	CGCACCATGGCATCTCTTIA	AAAAAGATCTCTATGTCCGTGGT
EGFR	NM 005228	5	23,24	CAGCCAGCCAAACAATCAGA	TCTTTGGAGTCTTCAGAGGGAAA
EGFR	NM 005228	6	25,26	TGTGGTTTCGTGGAAACAA	AATTGACAGCTCCCCACAG
EGFR	NM 005228	7	27,28	GGCTTCTGACGGGAGTCAA	CCACCCAAAGACTCTCAAGA
EGFR	NM 005228	8	29,30	CCTTCCATCACCCCTCAAG	AGTGCTTCCCATTTGCTTAA
EGFR	NM 005228	9	31,32	ACCGGAATTCCTTCTCTGTT	CACTGAAACAACAACAGGGTGA
EGFR	NM 005228	10	33,34	AGGGGGTGAGTCACAGGTTT	TCAGAAGAAATGTTTATTCCAAGG
EGFR	NM 005228	11	35,36	GCAAAATCAATTTTCCACTT	GCAGGAGCTCTGTGCCCTAT
EGFR	NM 005228	12	37,38	TCCACAGCATGACCTACCA	TTTGCTTCTTAAGGAACGTGAAA
EGFR	NM 005228	13	39,40	TGTCACCAAGGTCTATGGAG	CAAAAGCCAAAGGCAAGAA
EGFR	NM 005228	14	41,42	GGAGTCCCAACTCCTTGACC	GTCTGCCACACAGGATG
EGFR	NM 005228	15	43,44	GCTTTCCTCACTCACACACA	CAAACTCGCAATTTGTTG
EGFR	NM 005228	16	45,46	CCACCAATCCAACATCCAGA	TGGCCAGAGCCATAGAAAC
EGFR	NM 005228	17	47,48	TTCCAAGATCATTTACAAGATGTCA	GCACATTCAGAGATTCTTCTGTC
EGFR	NM 005228	18	49,50	TCCAAATGAGCTGGCAAGTG	TCCAAACACTCAGTGAAACAAA
EGFR	NM 005228	19	51,52	GTGCATCGCTGGTAACATCC	TGTGGAGATGAGCAGGCTCT
EGFR	NM 005228	20	53,54	ATCGCATTCATGCGTCTTCA	ATCCCCATGGCAAACTCTTG
EGFR	NM 005228	21	55,56	GCTCAGAGCCTGGCATGAA	CATCTCCCTGCTATGTTGT
EGFR	NM 005228	22	57,58	TGGCTCGTCTGTGTGTGCA	CGAAAGAAAATACTTGATGTGAGA
EGFR	NM 005228	23	59,60	TGAAGCAAAATTGCCCAAGAC	TGACATTTCTCCAGGATGTC
EGFR	NM 005228	24	61,62	AAGTGTCGATCACCAATGC	ATGCGATCTGGGACACAGG
EGFR	NM 005228	25	63,64	GGCACTGCTGGCAATAGAC	TGACTTCATATCCATGTGAGTTTCACT
EGFR	NM 005228	26	65,66	TATACCTCCATGAGGCACA	GGGAAAAACCCACACAAGAA
EGFR	NM 005228	27	67,68	TCAGAACAGCATCTCAAGGA	GATGCTGGAGGGAGCACTT
EGFR	NM 005228	28	69,70	CCTTGTGTGAGGACATTCACAGG	ATGTGCCGAGGTGGAGTA
EPHA1	NM 005232	14	71,72	GGAGGGCAGAGGACTAGCTG	GTGCTGGCCAAAGTCTTGT
EPHA1	NM 005232	15	73,74	CTGCAGCTAGCAACAGAGC	AAGAACCAGAGGAGCCAGGA
EPHA2	NM 004431	13	75,76	CGGGTAAGGATGTGGGTTGT	CAGGTGTCTGCTCCTGAA
EPHA2	NM 004431	14	77,78	GCTTCAGGAGGCAGAACACC	GGAGCAAGCCTAAGAAAGTTCA
EPHA3	NM 005233	10	79,80	GCCTTGTATCCATTTOCCACA	TGACAACACGTTTGGGTCAT
EPHA3	NM 005233	11	81,82	TGCATATTCCATTTCAGAACAGA	AAACAGTTTCATTGCTGCTAAAT
EPHA4	NM 004438	13	83,84	CCGGATACAGATACCCAAAAGA	GGAGGCTTCAAGGATCAGA
EPHA4	NM 004438	14	85,86	GCTGTTGCTGCTTGGCTA	TGGTGTAAATGTTGAAGTAGCTTGC
EPHA7	NM 004440	13	87,88	TGGCTGTCAGCTAAATAAGCATGT	TCAATTGCTTCATTCTTCTGTT
EPHA7	NM 004440	14	89,90	TGCTGCTGAACCTACCAACCAA	TGTGTAGTAATTGTGAAAACCTG
EPHA8	NM 020526	13	91,92	CAAAAGCACCGTCTCACTCG	CCCGAACTGCCAACTTCAT
EPHA8	NM 020526	14	93,94	GGAAAACAGGACCCAGTGT	CCCTCTCCACAGAGCTGAT
EPHB1	NM 004441	7	95,96	GACAGAGCTGACAAGCAGCA	AGGTTCCATTCCCTCCAGT
EPHB1	NM 004441	8	97,98	TGGGAGTGAGAGTTTGAAGAA	TATGAGGCCGTGAGCTGAAA
EPHB2	NM 017449	11	99,100	AGGGCCCTGCTCTGGTTT	CCAATTGGGGCTTAGTGAAA
EPHB2	NM 017449	12	101	CTCATGAGATTGGGGCATCA	AGGCCATGATCTCAGAACGC
EPHB3	NM 004443	11	103	GTTTGCAGGAGAGACGAGGT	AGGCCCTTCCCTGTGAC
EPHB3	NM 004443	12	105	ATGACCCCTCCGATCCTACC	TATCTGCTGCACGGCATT

EPHB4	NM 004444	14	107	GGAAAAAGCAGAGGCAGGTG	TGGTCTCAAGAAGCCAGCAG
EPHB6	NM 004445	16	109	GACACCTCCCCCTCTCAT	ACTATGACACCCCGCTCAG
EPHB6	NM 004445	17	111	TGCTTGATGTAAACCCCTGG	GCAATCCAACAGCCATGAGA
ERBB2	NM 004448	21	113	GGAGCAAAACCCCTATGTCCA	TCCTCCAACGTGTGTCTGTGG
ERBB3	NM 001982	21	115	TGGGGACCACTGCTGAGAG	TGCAGCCTTCTCTCCTTGA
FGFR1	NM 000604	14	117	GCAGAGCAGTGTGGCAGAAG	ACAGGTGGGAAGGGACTGG
FGFR1	NM 000604	15	119	AGTGGGGTGGGCTGAGAAC	TCTCTGGGGCAGAAAGAGGA
FGFR2	NM 000141	14	121	ACCCGGCCACACTGTATTTC	CATCCACCCAGCTCTCAAC
FGFR2	NM 000141	15	123	AGGGCATAGCCCTATTGAGC	CCCAGGAAAAAGCCAGAGAA
FGFR3	NM 000142	13	125	CAGGTGTGGGTGGAGTAGGC	CTCAGGCGCCATCCACTT
FGFR3	NM 000142	14	127	AAGAAGACGACCAACGTGAGC	AGGAGCTCCAGGGCACAAG
FGFR4	NM 002011	14	129	CCTCCTCTGTAAAGTGGGTGGA	AGAGGGCCTCAGTGCAGAGT
FGFR4	NM 002011	15	131	AGATGGGGCAGAACTGGATG	GGGTCCCAGACCAAATCTGA
FLT1	NM 002019	23	133	AGGTGCTCCCTTCACAGCAT	TTCAGGGACTACAGCTGAGGAA
FLT1	NM 002019	24	135	GCCGTATGTTATCTGGGAGGT	TGGGCCCATTAACCTTTAAGA
FLT3	NM 004119	20	137	ITCCATCACCCGTACCTCCT	CCATAAATCAAAAAATGCACCACA
FLT3	NM 004119	21	139	GAGTGGTCTTAGGAAGATGATGC	AAAGTCATGGGTGCAATACAA
FLT4	NM 002020	23	141	ATGGTCCCCACTGCTTGG	AGGAGCTCACTCACTTGT
IGF1R	NM 000875	18	143	CCTTGGCTCTCTCCACACAT	TGGCAACGGGTAAACAATGAA
INSR	NM 000208	18	145	GGCTGAGGTAAGCTGCTTCG	AAAAAGAAGTATCTTGCCTTT
INSR	NM 000208	19	147	AACCCCTCTTAGGGCTCTGTG	CAGGAGGATGGCAGGCTTC
KDR	NM 002253	24	149	CGTAGAGAGCTTCAGGACCTGTG	ITCCGAGAAGTTTGGCTTGA
KIT	NM 000222	17	151	TGTGAACATCATTCAGGCGTA	AAAATGTGTGATATCCCAGACAGG
KIT	NM 000222	18	153	TCCACATTTAGCAACAGCA	GGCTGCTTCTGAGACACAGT
LTK	NM 002344	16	155	TATCTACCGGTGCGGACTT	AGGTGTAGCCTCCCTCACA
MERTK	NM 006343	17	157	AGGCTGGTGGTGTCTCTGTG	CAAGCTGCCAACCCCTCAGTT
MET	NM 000245	19	159	TGGATTTCAAATACTGAAGCCACT	TGGAAATGGTGGTGTGAAATTT
MUSK	NM 005592	15	161	GGGCTTCATATGTTCTGACATGG	CAGAGGACCACGCCATAAGG
MUSK	NM 005592	15	163	CCGAGATTTAGCCACCAGGA	CCTGGGAAGCAAAACAACA
NTRK1	NM 002529	15	165	AGGTCCCCAGTCTCCTCTCC	AGACCCATGCAGCCATCCTA
NTRK1	NM 002529	16	167	CGTGAACCACCGAGCTTGT	AGAGGGGCAGAGGGGAAC
NTRK2	NM 006180	15	169	GGTGGGGGTGAGGAGCTTAG	TCGTTTAAGCCACCCAGTCA
NTRK2	NM 006180	16	171	TGCAATAAGGAAAGCAAACA	TCCTGACATGGTCTTCCAACC
NTRK3	NM 002530	17	173	CAGCATCTTCACACCTCTGA	GCTGGCTCTAAATCCCACT
NTRK3	NM 002530	18	175	CTAATCCGGGAAGTTGTTGC	TTCTGTATCAGCAGTTCTCTGTG
PDGFRA	NM 006206	18	177	CAAGTGCCACCATTGGATCA	GGCAGTGTACTGACCCCTGA
PDGFRA	NM 006206	19	179	GCACAAGTTATTAAGAGCCCAAGG	AGCATACTGGCCTCACAACA
PDGFRB	NM 002609	18	181	GCACATGGGCAGTGTGTATT	GAGCCCCACACAGATTCTCT
PDGFRB	NM 002609	19	183	ATGGGACGGGAGAGTGGTTG	TCCCTGTATCAGGGCTCGTC
PTK7	NM 002821	18	185	TTCTACGCAGCACACCAAT	GCAGGCCTAAACCCCTTCC
PTK7	NM 002821	19	187	GCACGCATGTGACCAATTC	AGCCCTGAGAGGGAGGTAGG
RET	NM 000323	15	189	CACACACCACCCCTCTGCT	AAAGATTGGGGTGAGGCTA
RET	NM 000323	16	191	CTGAAAGCTCAGGGATAGGG	CTGGCCAAGCTGCACAGAA
ROR1	NM 005012	09	193	TGCAGCCAACGATTTGAAAG	GGAAAGCCCCAAGTCTGAAA
ROR1	NM 005012	09	195	TCATCATGAGATCCCCACACT	GCATTTCCCCCTGAAGGAGT
ROR1	NM 005012	09	197	TGGATTCACTAACAGGAAGTGA	CCCATTCACAGGATGATT
ROR1	NM 005012	09	199	GTTTCCAGTGTCCCACTACC	GCTCGAAACCATGTTCCA
RYK	NM 002958	13	201	CTGGATTTGGGGTCTCTGC	CGGGAACAGCTAGCAGATTTT
TEK	NM 000459	18	203	GGGAATTTTGGAGGGGAAGT	GCTTCAGTCACCACAGAGCA
TEK	NM 000459	19	205	TGAGTCTACCCAGCAATCATTTG	TTCCCGAGAGCTACAGGACA
TIE	NM 005424	18	207	GGTAACAAGGGTACCCACGAA	GTTTGAGGGGCTGAGTGTGG
TIE	NM 005424	19	209	CCTCACCTTAGGGCTTGTG	AGCCAGGTCAATGCCCTAGA
TYRO3	NM 006293	18	211,212	GGGTAGCTTGGGAGCAAAGA	CCAAACCCAGAGAGCAAGAC

Tabla S1B: Cebadores para amplificación del EGFR seleccionado y exones de la tirosina quinasa receptora (SEQ ID NOS: 213 – 424)

Gen	RefSeq	Exón	SEQ ID NO	F Externo	R Externo
ALK	NM_004304	24	213,214	CATTTCCCTAATCCTTTTCCA	GTGATCCCAGATTTAAGCCTTC
ALK	NM_004304	25	215	GCCTCTCGTGGTTTGTTTGTCT	CCCAGGGTAGGGTCCAATAATC
AXL	NM_021913	19	217	CTTCTCGTGGAGGTGACTGAT	CAGGCATAGTGTGTATGGTCA
CSF1R	NM_005211	18	219	TCACGATACACATTCTCAGATCC	GAAGATCTCCAGAGGAGGATG
CSF1R	NM_005211	19	221	CGTAACGTGCTGTGACCAAT	AAACGAGGGAAGAGCCAGAAAG
DDR1	NM_013994	15	223	TGGGGAGCACATAAAAGAAGA	ACTCTGGCTCCTGGATTCTTG
DDR2	NM_006182	16	225	GGAAAGTCAGTGTGCAGGAATA	TTTATAGCAGAAATAGGCAAGCA
DKFZp761P1010	NM_018423	8	227	TGGTAATCCTAAACACAATGCAGA	CTGGGCAACACAGTGAGATCCT
EGFR	NM_005228	2	229	TCACAAATTTCTTTGCTGTGTCC	CATGGAACTCCAGATTAGCCTGT
EGFR	NM_005228	3	231	GATTGTGTGAGATCGTGGACAT	CGCTTAAATCTTCCCATTCAG
EGFR	NM_005228	4	233	CTCCATGGCACCATCATTAAACA	CTCAGGACACAAGTCTCTGCT
EGFR	NM_005228	5	235	GCAGTTCATGGTTCATCTTCTTTT	CAAAATAGCCACCTCGGATTA
EGFR	NM_005228	6	237	CTTCTGCATTGCCCAAGATG	CAAGGTCTCAGTGAGTGGTGA
EGFR	NM_005228	7	239	GAGAAGGGTCTTCTGACTCTGC	CAGGTGTTTCTCCTGTGAGGTG
EGFR	NM_005228	8	241	CACATTGCGGCCTAGAATGTTA	ACCCCGTCACAACCTTCAGT
EGFR	NM_005228	9	243	GCCGTAGCCCCAAAGTGTACTA	TCAGCTCAAACTGTGATTTCC
EGFR	NM_005228	10	245	CTCACTCTCCATAAATGCTACGAA	GACTTAACGTGTCCCTTTTGC
EGFR	NM_005228	11	247	GCCTCTTCGGGGTAATCAGATA	GAAGTCTGTGGTTAGCGGACA
EGFR	NM_005228	12	249	ATCTTTTGCCTGGAGGAACCTT	CAGGGTAAATTCATCCCATTTGA
EGFR	NM_005228	13	251	CAGCAGCCAGCACAACTACTTT	TTGGCTAGATGAACCATTTGATGA
EGFR	NM_005228	14	253	TGAATGAAGCTCCTGTGTTACTC	ATGTTTCATCGCAGGCTAATGTG
EGFR	NM_005228	15	255	AAAACAGGGAGAACTTCTAAGCAA	CATGCGCAGAGTCATTCCCACT
EGFR	NM_005228	16	257	CAATGCTAGAACACGCTGTC	TCCCTCCACTGAGGACAAAGTT
EGFR	NM_005228	17	259	GGGAGAGCTTGAGAAAGTTGGA	ATTCTCTCGGATGGATGTACCA
EGFR	NM_005228	18	261	TCAGAGCCTGTGTTTCTACCAA	TGGTCTCACAGGACCACTGATT
EGFR	NM_005228	19	263	AAATAATCAGTGTGATTCGTGGAG	GAGGCCAGTGTCTCTAAGG
EGFR	NM_005228	20	265	ACTTCACAGCCCTGCGTAAAC	ATGGGACAGGCAGTGAATTTGT
EGFR	NM_005228	21	267	GCAGCGGGTTACATCTTCTTTC	CAGCTCTGGCTCACACTACCAG
EGFR	NM_005228	22	269	CCTGAATCCGTGAGACTGAAA	GCAGCTGGACTCGATTTCCT
EGFR	NM_005228	23	271	CCTTACAGCAATCCTGTGAAACA	TGCCCAATGAGTCAAGAAGTGT
EGFR	NM_005228	24	273	ATGTACAGTGTGCGATGGTCT	CACCTACGGATGCTGCTTAGTT
EGFR	NM_005228	25	275	TAAGGCACCCACATCATGTCA	TGGACCTAAAAGGCTTACAATCA
EGFR	NM_005228	26	277	GCCTTTTAGGTCCACTATGGAATG	CCAGGCGATGCTACTACTGGTC
EGFR	NM_005228	27	279	TCATAGCACACCTCCCTCACTG	ACACAACAAAGAGCTTGTGCAAG
EGFR	NM_005228	28	281	CCATTACTTTGAGAAGGACAGGAA	TATTCTTGCTGGATGCGTTTCT
EPHA1	NM_005232	14	283	AGGAGGGCAGAGGACTAGCTG	GGCAATGTGAATGTGCACTG
EPHA1	NM_005232	15	285	CTTGAACCTGGGAGGTGGAG	ATCAGGGTGGGAGGAGTAAAGA
EPHA2	NM_004431	13	287	CCCACCTACCTCTCACCTGTGC	GTGAACCTCCGATAGGAAATGG
EPHA2	NM_004431	14	289	AGGGGAOCTCAAGGGAGAAAG	AGATCATGCCAGTGAACCTCAG
EPHA3	NM_005233	10	291	GGACCAGGAAAGTCTTGTCTTT	GGTGGGGAACATTAAACTGAGG
EPHA3	NM_005233	11	293	GCTTCAGGTTGTTTGTGTGAG	AOCCTTGCTTGAGGGAAATATG
EPHA4	NM_004438	13	295	CCCAGCTCCTAGGGTACAGTCT	CAGTCAGCTTCAAAATCCCTCTT
EPHA4	NM_004438	14	297	TCACTTCCCTGTGAGTAAAGAAAA	GGCCATTAAATCTTGTCTTGA
EPHA7	NM_004440	13	299	TGGACTTGTGCAAACTCAAACTG	TCCCAATATAGGGCAGTCATGTT
EPHA7	NM_004440	14	301	TCTCAATCAGTTGAGTTGCCCTTG	AGCTGTGCAAGTGTGGAACAT
EPHA8	NM_020526	13	303	GCTGTGAGGGTAAATGAGACCA	GTCTCCTGGTGAGTGACTGTGG
EPHA8	NM_020526	14	305	CCTTCCTCTGCTCCACAGC	GTCCTTGTGCCAACAGTCGAG
EPHB1	NM_004441	7	307	GCTTGGCAAGGAGAAGAGAACA	GCTTGGCTTCTTGTCTGAACAAC
EPHB1	NM_004441	8	309	GCTGGTCACCTTGAGCTTCTCT	CCATGCTGGGCTCTTTGATTA
EPHB2	NM_017449	11	311	CACCACTCTGAAGTTGGCCTCT	ATGGCTCTGCACATTTGTTC
EPHB2	NM_017449	12	313	CAGAGTGGGAAAAAGGCACCTCA	CCAGAGTCTGTGTCAGACATTC
EPHB3	NM_004443	11	315	ATGGGGATTAACTGGGATGTTG	CGTAGCTCCAGACATCACTAGCA
EPHB3	NM_004443	12	317	GCAACCTGGTCTGCAAAGTCTC	ACCCAGCAGTCCAGCATGAG

Tabla S1B (continuación)

EPHB4	NM 004444	14	319	GAGTTTCAGTGAGCCAAGATCG	TTACAGGCTTGAGCCACTAGGC
EPHB5	NM 004445	16	321	AAGCTTCCAGGAGACGAGGTC	GTCCCTGAAATCCCTCAAACC
EPHB6	NM 004445	17	323	TGCTCCATAAACGTGACTATTGC	GTAAGAGGGTGGGCTGGAATCT
ERBB2	NM 004448	21	325	CTTAGAACATGTCCGGGAAAAC	CACATCACTCTGGTGGGTGAAC
ERBB3	NM 001982	21	327	AAATTTTCATCCAAAACCAACC	CCAGTCCCAAGTCTTGATCATT
FGFR1	NM 000604	14	329	ACAAGTCGGCTAGTTGCATGG	TCTCAGATGAAACCACCAGCAC
FGFR1	NM 000604	15	331	TTCACTGTGAGAAGCAAGGATGG	CCAGGGAGAAAAGCAGGACTCTA
FGFR2	NM 000141	14	333	TTCTGGCGGTGTTTGAATTA	CTCAACATTGACGGCCTTCTT
FGFR2	NM 000141	15	335	TCAGCTCTTAAACAGGGCATAGC	GAAATGCAGCAGCCACTAAAGA
FGFR3	NM 000142	13	337	CTCACCTTCAAGGACCTGGTGT	CAGGGAGGGGTAGAAACCACA
FGFR3	NM 000142	14	339	GGAGAGGTGGAGAGGCTTCAG	GAGACTCCAGGACAGACACCT
FGFR4	NM 002011	14	341	CACTCGTCTCTCACCTTCC	AGGACTCACACGTCACTCTGGT
FGFR4	NM 002011	15	343	GGACAATGTGATGAAGATTGCTG	ATAGCAGGATCCCAAAAGACCA
FLT1	NM 002019	23	345	GGCTTGGGACCTGTATTTGTA	CAGTGGCCTTCTGAGCCTTAC
FLT1	NM 002019	24	347	GCACCTAGCTCCCTCTTTAGC	TTTACAGTAGAGGGCAGACATGC
FLT3	NM 004119	20	349	GCCACCATAGCTGCAGAATTAG	CCCAAGGACAGATGTGATGCTA
FLT3	NM 004119	21	351	GCCTTGTGTGAGAGGAGTTGT	GTTACAGCTCTCAAGCAGGTTA
FLT4	NM 002020	23	353	ATTCCACAAGCTCTCTCCATGA	CTTGCCCAAGATGCCTAAG
IGFIR	NM 000875	18	355	TGCTTGGTATTGTCTCATCATGT	CCCTTAGCTAGCCCACTGACAA
INSR	NM 000208	18	357	CTCTGGGAGTGGTGTCCAA	CCTGGGCAACAGACAGAGTAAG
INSR	NM 000208	19	359	CTTCACTTCCCATGCGTACC	GGGTTCAACAATGCCTACAGGA
KDR	NM 002253	24	361	AAAATCTGTGACTTTGGCTTGG	GGGAGGAGACATTTCTTGATTTG
KIT	NM 000222	17	363	GCAGTCTGAGAAGAAAACAGC	CTTCACATGCCCCAAAATTACA
KIT	NM 000222	18	365	TGAGCCATGTATTTCAGAGGTGA	TACATTTACGAGGTGCGTGT
LTK	NM 002344	16	367	TTGCCTACTCTGTAGGGATATTGC	ATAGGGCATGTAGCCCATGTA
MERTK	NM 006343	17	369	GCCTGCTGTTGGTCTCACT	TTGCAAAGCACACATCTCTGA
MET	NM 000243	19	371	TGGCAATGTCAATGTCAAGCAT	GTATGTTGCCCACTCAACAAA
MUSK	NM 005592	15 1	373	TGCATTTCTAGCTGAGACTCC	TGCCATCTCGCAGTAGTAAAT
MUSK	NM 005592	15 2	375	CTCTCCTGTGCTGAGCAGCTTT	TGTTTCCAATCACTGGCTTTCA
NTRK1	NM 002529	15	377	GAACCATGGGCTGTCTCTGG	ATCTGGGATAGCGAAGGAGACA
NTRK1	NM 002529	16	379	ATTACAGGCCACAGCCATC	AAGGCAAGAATAAGGGAGGAAGA
NTRK2	NM 006180	15	381	GCTCTCAGGACTGCAGAAGTACA	GAGGAACCAATCCCACTCACAC
NTRK2	NM 006180	16	383	TCACTCTTGGCTTCTGTCTCTG	GCACTGTGCTTTGCTTCTCAG
NTRK3	NM 002530	17	385	TGTCTCCTTATCGTAGGCTCCA	CACCACATTTCTACAGTTCCA
NTRK3	NM 002530	18	387	CACGTGTGACCAGACAGACAAA	TGTGGTTTTCTGTATCAGCAGCTT
PDGFRA	NM 006206	18	389	CAGGGAGTCTGAAATCATCAGG	TCAAGTATCTAGCCCCAAATCCA
PDGFRA	NM 006206	19	391	GGCAATATTGACCATTTCATTC	AGGCCAGGAGTAAGACGCAAC
PDGFRB	NM 002609	18	393	AAGAACGTACGTGTGGTGTGG	CGCTATACTTGCTCCATGCACT
PDGFRB	NM 002609	19	395	AGGAAACAGCCTCTGGTCTC	GTCAATGCTCAGACAGGGAGAT
PTK7	NM 002821	18	397	CCCAGGAAGCAGGTAAGTTA	TTTACAACCAAGGGGTGTG
PTK7	NM 002821	19	399	TCGTGTGTTACCTCCAGATTTT	AAATTAGCCAGGGAGTGGAGGT
RET	NM 000323	15	401	CATGCCATGCTATGGCTCAC	AGGCTGAGCGGAGTTCTAATTG
RET	NM 000323	16	403	ATCTCAGCAATCCACAGGAGGT	ATTGCTTCACGAACACATCAT
ROR1	NM 005012	09 1	405	TGGAAAGTTGTCTATGGCACTC	ATGGGCAGCAAGGACTTACTCT
ROR1	NM 005012	09 2	407	CACCCCAATATTGTCTGCCTTC	GGCTCGGGAACATGTAATTAGG
ROR1	NM 005012	09 3	409	CCATCATGTATGGCAAAATCTCTT	TGGCGTCTCTAGTAAAGATGCT
ROR1	NM 005012	09 4	411	GCCAGATTGCTGTTTCATTG	GGCTAAACACAAAGCACCATT
RYK	NM 002958	13	413	GGGAAATCATCCACAAAGACCT	GGTCTGGGTACAGCTCCTC
TEK	NM 000459	18	415	TTCTTCTGCCAAGATGTGGTGT	TGCAGATGTGCAATCATGTGTA
TEK	NM 000459	19	417	TGGACCCCGAAAGATAAATAGG	TTCTGCACTCCTCTGGAAACTG
TIE	NM 005424	18	419	GGGTGAGAGCCAACACTGATCT	CTGTGCCCTCTCATCTCACT
TIE	NM 005424	19	421	AGAACCTAGCTCCAAGATTGC	ACACCTTCCAAGACTCCTTCCA
TYRO3	NM 006293	18	423,424	GACTCGAGGGTGGGAGACAG	GC7GTCACTAGGTGTCTGAGC

Tabla S2: estatus de la mutación del EGFR en cáncer de pulmón no tratado

Muestra	Histología	Fuente	Género	Exón	Alteración de la secuencia	SEQ ID NO	Nucleótido	Aminoácido
S0514	adenocarcinoma	EUA	F	18	Sustitución	425	2155G>A	G719S
S0377	adenocarcinoma	Japón	F	18	Sustitución	426	2155G>A	G719S
S0418	adenocarcinoma	Japón	F	19	Del-1a	427	2235_2249delGGAAATTAAGAGAAAGC	E746_A750del
S0363	cáncer de células grandes	Japón	F	19	Del-1a	428	2235_2249delGGAAATTAAGAGAAAGC	E746_A750del
S0380	adenocarcinoma	Japón	M	19	Del-1a	429	2235_2249delGGAAATTAAGAGAAAGC	E746_A750del
S0399	adenocarcinoma	Japón	F	19	Del-1a	430	2235_2249delGGAAATTAAGAGAAAGC	E746_A750del
S0353	adenocarcinoma	Japón	F	19	Del-1a	431	2235_2249delGGAAATTAAGAGAAAGC	E746_A750del
S0385	adenocarcinoma	Japón	M	19	Del-1a	432	2235_2249delGGAAATTAAGAGAAAGC	E746_A750del
S0301	adenocarcinoma	Japón	M	19	Del-1a	433	2235_2249delGGAAATTAAGAGAAAGC	E746_A750del
S0412	adenocarcinoma	Japón	M	19	Del-1b	434	2236_2250delGAATTAAGAGAGAGCA	E746_A750del
S0335	adenocarcinoma	Japón	M	19	Del-1b	435	2236_2250delGAATTAAGAGAGAGCA	E746_A750del
S0405	adenocarcinoma	Japón	F	19	Del-1b	436	2236_2250delGAATTAAGAGAGAGCA	E746_A750del
S0439	adenocarcinoma	Japón	M	19	Del-2	437	2254_2277delTCTCGCAAAAGCCAAACAGGAAATC	S752_1759del
S0361	adenocarcinoma	Japón	F	21	Sustitución	438	2573T>G	L858R
S0388	adenocarcinoma	Japón	F	21	Sustitución	439	2573T>G	L858R
S0389	adenocarcinoma	Japón	F	21	Sustitución	440	2573T>G	L858R

Tabla S3A: Estado de la mutación del EGFR en cáncer de pulmón tratado con gefitinib

Sensibilidad al Gefitinib	Muestra	Histología	Fuente	Género	Exones	Alteración de la secuencia	SEQ ID NO	Nucleótido	Aminoácido
Y	IR1T	adenocarcinoma	EUA	M	19	Del-3	441	2239_2247delTTAAGAGAA, 2248G>C	L747_E749del, A750P
Y	P003	adenocarcinoma	EUA	M	19	Del-3	442	2239_2247delTTAAGAGAA, 2248G>C	L747_E749del, A750P
Y	IR4T	carcinoma broncoalveolar	EUA	F	19	Del-4	443	2240_2257delTTAAGAGAAGCAACATCTC	L747_S752del, P753S
Y	IR2T	adenocarcinoma	EUA	F	19	Del-5	444	2238_2255delATTAAAGAGAGCAACATC, 2237A>T	L747_S752del, E746V
Y	IR3T	adenocarcinoma	EUA	F	21	Sustitución	445	2573T>G	L858R
Y	IRG	adenocarcinoma	EUA	F	21	Sustitución	446	2573T>G	L858R
<i>in vitro</i>	H3255								
N	IR5	adenocarcinoma	EUA	F	18 - 24	No detectada		n/a	n/a
N	IR6	adenocarcinoma	EUA	M	18 - 24	No detectada		n/a	n/a
N	IR8	adenocarcinoma	EUA	F	18 - 24	No detectada		n/a	n/a
N	IR9	NSCLC	EUA	F	18 - 24	No detectada		n/a	n/a

Tabla S3B: Mutaciones del EGFR no mostradas en la Tabla 2, Tabla S2 o Tabla S3A

Muestra	Tejido	Exón	Alteración de la secuencia	Nucleótido	Aminoácido
Tar4T	Adenocarcinoma de pulmón	19	Supresión	2239-2250delTTAAGAGAAGCA; 2251A>C	L747_A750del; T751T
AD355	Adenocarcinoma de pulmón	19	Supresión	2240-2254delTAAGAGAAGCA	L747_T751del
IR TT	Adenocarcinoma de pulmón	19	Supresión	2257-2271delCCGAAAGCCAACAAG	P753_K757del
AD240	Adenocarcinoma de pulmón	20	Inserción	2309-2310insCAACCCGG	D770_N771ins NPG
AD261	Adenocarcinoma de pulmón	20	Inserción	2311-2312insGCGTGGACA	D770_N771ins SVD
	Adenocarcinoma de pulmón	20	Inserción	2316-2317insGGT	P772_H773ins V
AD356	Adenocarcinoma de pulmón	20	Sustitución	2334-2335GG>AA	G779S
SP02-23	Leucemia mieloide aguda	21	Sustitución	2570G>T	G857V
SP08-94	Glioma	21	Sustitución	2582T>A	L861Q
SP06-45	Sarcoma	21	Sustitución	2648T>C	L883S
AD241	Adenocarcinoma de colon	22	Sustitución	2686G>T	D896Y

Tabla S3C: Posición de los mutantes BCR-ABL resistentes a imatinib y posiciones análogas en EGFR

Residuo de Ab11 sometido a mutación de resistencia	Residuo del EGFR análogo	Idéntico/similar/no conservado
Met-244	Lys-714	No conservado
Leu-248	Leu-718	Idéntico
Gly-250	Ser-720	No conservado
Gln-252	Ala-722	No conservado
Tyr-253	Phe-723	Similar
Glu-255	Thr-725	No conservado
Asp-276	Ala-750	No conservado
Thr-315	Thr-790	Idéntico
Phe-317	Leu-792	Similar
Met-351	Met-825	Idéntico
Glu-355	Glu-829	Idéntico
Phe-359	Leu-833	Similar
His-396	His-870	Idéntico
Ser-417	Thr-892	Similar
Phe-486	Phe-961	Idéntico

Tabla S4: Cebadores utilizados para la secuenciación de ADNc

Nombre del cebador	SEQ ID NO	Secuencia 5' a 3' del cebador
ADNc_EGFR_aF	447	TGTAAAACGACGGCCAGTCGCCCAGACCGGACGACA
ADNc_EGFR_aR	448	CAGGAAACAGCTATGACCAGGGCAATGAGGACATAACCA
ADNc_EGFR_bF	449	TGTAAAACGACGGCCAGTGGTGGTCCTTGGGAATTTGG
ADNc_EGFR_bR	450	CAGGAAACAGCTATGACCCCATCGACATGTTGCTGAGAAA
ADNc_EGFR_cF	451	TGTAAAACGACGGCCAGTGAAGGAGCTGCCCATGAGAA
ADNc_EGFR_cR	452	CAGGAAACAGCTATGACCCGTGGCTTCGTCTCGGAATT
ADNc_EGFR_dF	453	TGTAAAACGACGGCCAGTGAAACTGACCAAAATCATCTGT
ADNc_EGFR_dR	454	CAGGAAACAGCTATGACCTACCTATTCCGTTACACACTTT
ADNc_EGFR_eF	455	TGTAAAACGACGGCCAGTCCGTAATTATGTGGTGACAGAT
ADNc_EGFR_eR	456	CAGGAAACAGCTATGACCGCGTATGATTTCTAGGTTCTCA
ADNc_EGFR_fF	457	TGTAAAACGACGGCCAGTCTGAAAACCGTAAAGGAAATCAC
ADNc_EGFR_fR	458	CAGGAAACAGCTATGACCCCTGCCTCGGCTGACATTC
ADNc_EGFR_gF	459	TGTAAAACGACGGCCAGTTAAGCAACAGAGGTGAAAACAG
ADNc_EGFR_gR	460	CAGGAAACAGCTATGACCGGTGTTGTTTTCTCCCATGACT
ADNc_EGFR_hF	461	TGTAAAACGACGGCCAGTGGACCAGACAACTGTATCCA
ADNc_EGFR_hR	462	CAGGAAACAGCTATGACCTTCCTCAAGATCCTCAAGAGA
ADNc_EGFR_iF	463	TGTAAAACGACGGCCAGTGATCGGCCTCTTCATGCGAA
ADNc_EGFR_iR	464	CAGGAAACAGCTATGACCACGGTGGAGGTGAGGCAGAT
ADNc_EGFR_jF	465	TGTAAAACGACGGCCAGTCGAAAGCCAACAAGGAAATCC
ADNc_EGFR_jR	466	CAGGAAACAGCTATGACCATTCCAATGCCATCCACTTGAT
ADNc_EGFR_kF	467	TGTAAAACGACGGCCAGTAACACCGCAGCATGTCAAGAT
ADNc_EGFR_kR	468	CAGGAAACAGCTATGACCCTCGGGCCATTTTGGAGAATT
ADNc_EGFR_lF	469	TGTAAAACGACGGCCAGTTCAGCCACCCATATGTACCAT
ADNc_EGFR_lR	470	CAGGAAACAGCTATGACCGCTTTGCAGCCCATTCTATC
ADNc_EGFR_mF	471	TGTAAAACGACGGCCAGTACAGCAGGGCTTCTTCAGCA

Nombre del cebador	SEQ ID NO	Secuencia 5' a 3' del cebador
ADNc_EGFR_mR	472	CAGGAAACAGCTATGACCTGACACAGGTGGGCTGGACA
ADNc_EGFR_nF	473	TGTAACACGACGGCCAGTGAATCCTGTCTATCACAATCAG
ADNc_EGFR_nR	474	CAGGAAACAGCTATGACCGGTATCGAAAGAGTCTGGATTT
ADNc_EGFR_oF	475	TGTAACACGACGGCCAGTGCTCCACAGCTGAAAATGCA
ADNc_EGFR_oR	476	CAGGAAACAGCTATGACCACGTTGCAAAACCAGTCTGTG

Los siguientes párrafos numerados definen realizaciones particulares de la presente invención:

1. Un método para determinar la probabilidad de la efectividad de un tratamiento dirigido al receptor del factor de crecimiento epidérmico (EGFR) en un paciente humano afectado con o en riesgo de desarrollar cáncer, que comprende: la detección de la presencia o ausencia de al menos una variación de ácido nucleico en el dominio quinasa del gen erbB1 de dicho paciente en relación con el gen erbB1 de tipo silvestre, en donde la presencia de al menos una variación de ácido nucleico indica que el tratamiento dirigido al EGFR es probable que sea efectivo.
2. El método del párrafo 1, en donde la variación de ácido nucleico aumenta la actividad de quinasa.
3. El método del párrafo 1, en donde el gen erbB1 se obtiene partir de una muestra biológica de dicho paciente.
4. El método del párrafo 1, en donde la variación en el dominio quinasa del gen erbB1 elabora la estructura conformacional del bolsillo de enlazamiento de ATP.
5. El método del párrafo 1, en donde la variación en el dominio quinasa de erbB1 está en un exón del gen erbB1 seleccionado del grupo que consiste del exón 18, 19, 20 o 21.
6. El método del párrafo 5, en donde la variación es en el exón 18, 19 o 21.
7. El método del párrafo 1, en donde la variación en el dominio quinasa del gen erbB1 es una supresión, sustitución, o inserción en el marco.
8. El método del párrafo 7, en donde la supresión en el marco es en el exón 19 del gen erbB1.
9. El método del párrafo 8, en donde la supresión en el marco en el exón 19 del gen erbB1 comprende una supresión de al menos los aminoácidos leucina, arginina, ácido glutámico y alanina, en los codones 747, 748, 749 y 750 de la SEQ ID NO: 512.
10. El método del párrafo 8, en donde la supresión en el marco en el exón 19 del gen erbB1 comprende una supresión de al menos los aminoácidos leucina, arginina y ácido glutámico en el codón 747, 748 y 749 de la SEQ ID NO: 512.
11. El método del párrafo 8, en donde la supresión en el marco comprende los nucleótidos seleccionados del grupo que consiste de 2235 - 2249, 2240 - 2251 y 2240 - 2257 de la SEQ ID NO: 511.
12. El método del párrafo 7, en donde la sustitución es en el exón 21 del gen erbB1.
13. El método del párrafo 12, en donde la sustitución en el exón 21 comprende al menos un aminoácido.
14. El método del párrafo 12, en donde la sustitución en el exón 21 comprende una sustitución del grupo que consiste de una guanina por una timina en el nucleótido 2573 de la SEQ ID NO: 511, y una adenina para una timina en el nucleótido 2582 de la SEQ ID NO: 511.
15. El método del párrafo 7, en donde la sustitución es en el exón 18 del gen erbB1.
16. El método del párrafo 15, en donde la sustitución en el exón 18 es una timina por una guanina o una serina por una guanina en el nucleótido 2155 de la SEQ ID NO: 511.
17. El método del párrafo 1, en donde la detección de la presencia o ausencia de dicha al menos una variación comprende la amplificación de un segmento de ácido nucleico.

18. El método del párrafo 17, en donde el segmento que va a ser amplificado es de 1000 nucleótidos de longitud o menos.
19. El método del párrafo 17, en donde el segmento que va a ser amplificado incluye una pluralidad de variaciones.
20. El método del párrafo 1, en donde la detección de la presencia o ausencia de dicha al menos una variación comprende poner en contacto el ácido nucleico de erbB1 con al menos una sonda de ácido nucleico, en donde dicha al menos una sonda preferentemente se hibrida con una secuencia de ácido nucleico que comprende dicha variación en condiciones de hibridación selectivas.
21. El método del párrafo 1, en donde la detección de la presencia o ausencia de dicha al menos una variación comprende la secuenciación de al menos una secuencia de ácido nucleico.
22. El método del párrafo 1, en donde la detección de la presencia o ausencia de dicha al menos una variación comprende la determinación por espectrometría de masas de al menos una secuencia de ácido nucleico.
23. El método del párrafo 1, en donde la detección de la presencia o ausencia de dicha al menos una variación comprende llevar a cabo una reacción en cadena de la polimerasa (PCR) para amplificar ácido nucleico que comprende la secuencia de codificación de erbB1, y la determinación de la secuencia de nucleótidos del ácido nucleico amplificado.
24. El método del párrafo 23, en donde la determinación de la secuencia de nucleótidos del ácido nucleico amplificado comprende la secuenciación de al menos un segmento de ácido nucleico.
25. El método del párrafo 23, en donde la determinación de la secuencia de nucleótidos del ácido nucleico amplificado comprende correr el segmento de ácido nucleico amplificado sobre un gel y la determinación del tamaño de los segmentos.
26. El método del párrafo 1, en donde la detección de la presencia o ausencia de dicha al menos una variación comprende determinar el haplotipo de una pluralidad de variaciones en un gen.
27. Una sonda que se enlaza específicamente en condiciones de enlazamiento selectivas a una secuencia de ácido nucleico que comprende por lo menos una variación en el gen erbB1, en donde la variación es una mutación en el dominio quinasa de erbB1 que confiere un cambio estructural en el bolsillo de enlazamiento de ATP.
28. La sonda del párrafo 27, en donde dicha variación es en un exón del gen erbB1 seleccionado del grupo que consiste del exón 18, 19, 20 o 21.
29. La sonda del párrafo 27, en donde dicha sonda comprende una secuencia de ácido nucleico de 500 bases de nucleótidos o menos de longitud.
30. La sonda del párrafo 27, en donde dicha sonda comprende ADN.
31. La sonda del párrafo 27, en donde dicha sonda comprende ADN y al menos un análogo de ácido nucleico.
32. La sonda del párrafo 27, en donde dicha sonda comprende ácido nucleico peptídico (PNA).
33. La sonda del párrafo 27, que comprende además un marcador detectable.
34. La sonda del párrafo 27, en donde el marcador detectable es un marcador fluorescente.
35. La sonda del párrafo 27, en donde dicha sonda comprende al menos 10 ácidos nucleicos consecutivos que consisten de al menos los ácidos nucleicos 15 - 25 de la SEQ ID NO 495, o sus adiciones.
36. La sonda del párrafo 27, en donde dicha sonda comprende al menos 10 ácidos nucleicos consecutivos que consisten de al menos los ácidos nucleicos 20 - 30 de la SEQ ID NO 497, o sus adiciones.
37. La sonda del párrafo 27, en donde dicha sonda comprende al menos 10 ácidos nucleicos consecutivos que consisten de al menos los ácidos nucleicos 20 - 30 de la SEQ ID NO 499, o sus adiciones.
38. Un método para determinar la probabilidad de la efectividad de un tratamiento dirigido al EGFR en un paciente que comprende:
la determinación de la actividad de la quinasa del gen erbB1 en una muestra biológica de dicho paciente, en donde un incremento en actividad de la quinasa después de la estimulación con un ligando del EGFR, en comparación con un control, indica que el tratamiento dirigido al EGFR es probable que sea efectivo.
39. El método de los párrafos 1 y 38, en donde el tratamiento dirigido al EGFR es un inhibidor de la tirosina quinasa.
40. El método del párrafo 39, en donde el inhibidor de la tirosina quinasa es una anilinoquinazolina.

41. El método del párrafo 40, en donde la anilinoquinazolina es una anilinoquinazolina sintética.
42. El método del párrafo 41, en donde la anilinoquinazolina sintética se selecciona del grupo que consiste de gefitinib y erlotinib.
- 5 43. Un método para determinar la probabilidad de la efectividad de un tratamiento dirigido al receptor del factor de crecimiento epidérmico (EGFR) en un paciente humano afectado con o en riesgo de desarrollar cáncer, que comprende:
 - a. la detección de la presencia o ausencia de al menos una variación de ácido nucleico en el exón 18, 19, 20 o 21 mediante la realización de una reacción en cadena de la polimerasa (PCR) para amplificar una porción del exón 18, 19, 20, o 21; y
 - 10 b. la determinación de la secuencia de nucleótidos del ácido nucleico amplificado por secuenciación de al menos una porción del exón amplificado 18, 19, 20, o 21, en donde la presencia de al menos una variación de nucleótidos en el exón 18, 19, 20 o 21 en comparación con un control de erbB1 de tipo silvestre indica que el tratamiento dirigido al EGFR es probable que sea efectivo.
- 15 44. Un método de tratamiento de un paciente afectado con o en riesgo de desarrollar cáncer, que comprende detectar la presencia o ausencia de al menos una variación de ácido nucleico en el dominio quinasa del gen erbB1 del paciente, en donde se le administra al paciente un tratamiento dirigido al EGFR si se detecta la presencia de dicha al menos una variación de ácido nucleico.
45. Un método de tratamiento de un paciente afectado con o en riesgo de desarrollar cáncer, que comprende:
 - a. detectar la presencia o ausencia de al menos una variación de ácido nucleico en el exón 18, 19, 20 o 21 mediante la realización de una reacción en cadena de la polimerasa (PCR) para amplificar una porción del exón 18, 19, 20, o 21,
 - 20 b. determinar la secuencia de nucleótidos del ácido nucleico amplificado por medio de la secuenciación de al menos una porción del exón amplificado 18, 19, 20 o 21; y
 - 25 c. administrar un tratamiento dirigido al EGFR al paciente si se detecta la presencia de dicha al menos una variación de ácido nucleico.
46. El método de los párrafos 44 y 45, en donde el tratamiento dirigido al EGFR es un inhibidor de la tirosina quinasa.
47. El método del párrafo 46, en donde el inhibidor de la tirosina quinasa es una anilinoquinazolina.
48. El método del párrafo 47, en donde la anilinoquinazolina es una anilinoquinazolina sintética.
- 30 49. El método del párrafo 48, en donde la anilinoquinazolina sintética se selecciona del grupo que consiste de gefitinib y erlotinib.
50. El método de los párrafos 44 y 45, en donde se obtiene el gen erbB1 a partir de una muestra biológica de dicho paciente.
- 35 51. El método de los párrafos 44 y 45, en donde dicho cáncer se selecciona del grupo que consiste de cáncer gastrointestinal, cáncer de próstata, cáncer de ovario, cáncer de mama, cáncer de cabeza y cuello, cáncer de pulmón, cáncer de pulmón de células no pequeñas, cáncer del sistema nervioso, cáncer de riñón, cáncer de retina, cáncer de piel, cáncer de hígado, cáncer de páncreas, cáncer genitourinario y cáncer de vejiga.
52. El método del párrafo 51, en donde dicho cáncer es cáncer de pulmón de células no pequeñas.
53. El método de los párrafos 44 y 45, en donde la variación de ácido nucleico aumenta la actividad de la quinasa.
- 40 54. El método de los párrafos 44 y 45, en donde la variación en el dominio quinasa del gen erbB1 efectúa la conformación del bolsillo de enlazamiento de ATP.
55. El método de los párrafos 44 y 45, en donde la variación en el dominio quinasa de erbB1 es en un exón del gen erbB1 seleccionado de entre el grupo que consiste del exón 18, 19, 20 o 21.
56. El método del párrafo 55, en donde la variación es en el exón 18, 19 o 21.
- 45 57. El método de los párrafos 44 y 45, en donde la variación en el dominio quinasa del gen erbB1 es una supresión o una sustitución en el marco.
58. El método del párrafo 57, en donde la supresión en el marco es en el exón 19 del gen erbB1.

59. El método del párrafo 58, en donde la supresión en el marco en el exón 19 del gen erbB1 comprende una supresión de por lo menos los aminoácidos leucina, arginina, ácido glutámico y alanina, en los codones 747, 748, 749 y 750.
- 5 60. El método del párrafo 58, en donde la supresión en el marco en el exón 19 del gen erbB1 comprende una supresión de por lo menos los aminoácidos leucina, arginina, glutamina y ácido en el codón 747, 748 y 749 de la SEQ ID NO: 512.
61. El método del párrafo 59, en donde la supresión en el marco comprende los nucleótidos seleccionados de entre el grupo que consiste de 2235 - 2249, 2240 - 2251, y 2240 - 2257 de la SEQ ID NO: 511.
62. El método del párrafo 57, en donde la sustitución es en el exón 21 del gen erbB1.
- 10 63. El método del párrafo 62, en donde la sustitución en el exón 21 comprende una sustitución del grupo que consiste de una guanina por una timina en el nucleótido 2573 de la SEQ ID NO: 511, y una adenina para una timina en el nucleótido 2582 nucleótidos de la SEQ ID NO: 511.
64. El método del párrafo 62, en donde la sustitución en el exón 21 comprende al menos un aminoácido.
65. El método del párrafo 57, en donde la sustitución es en el exón 18 del gen erbB1.
- 15 66. El método del párrafo 65, en donde la sustitución en el exón 18 es una timina por una guanina en el nucleótido 2155 de la SEQ ID NO: 511.
67. Un kit que comprende:
- a. al menos un par cebador degenerado diseñado para hibridar con regiones de ácido nucleico limítrofes o dentro del exón 18, 19, 20 o 21 del dominio quinasa del EGFR;
- 20 b. productos y reactivos necesarios para llevar a cabo la amplificación por PCR, e
- c. instrucciones.
68. Un kit de acuerdo con el párrafo 67, en donde el cebador comprende la secuencia de los cebadores seleccionados de entre el grupo consiste de las SEQ ID NOS: 505 - 508, y las SEQ ID NOS: 646 - 673 con la SEQ ID NO: 645 en el extremo 5' de todos los cebadores hacia delante y la SEQ ID NO: 674 en el extremo 5' de todos los
- 25 cebadores inversos.
69. Un kit que comprende:
- a. al menos una sonda diseñada para hibridar con regiones de ácido nucleico dentro de los exones 18, 19, 20 o 21 del dominio quinasa del EGFR;
- b. productos y reactivos necesarios para llevar a cabo la reacción de hibridación; e
- 30 c. instrucciones.
70. El kit del párrafo 69, en donde al menos una sonda está unida a un soporte sólido.
71. Un kit que comprende:
- a. al menos una sonda diseñada para enlazarse al bolsillo de enlazamiento de ATP de la proteína del dominio quinasa del EGFR;
- 35 b. los productos y los reactivos necesarios para llevar a cabo la reacción de enlazamiento; e
- c. instrucciones.
72. El kit de acuerdo con el párrafo 71, en donde la sonda es un anticuerpo, un fragmento de anticuerpo, o un anticuerpo quimérico.
73. El kit de acuerdo con el párrafo 72, en donde la sonda comprende además un marcador detectable.
- 40 74. Un método para seleccionar un compuesto que inhibe la actividad catalítica de la quinasa de una variante del receptor del factor de crecimiento epidérmico (EGFR), que comprende:
- a. poner en contacto el compuesto con una variante del EGFR; y
- b. la detección de la actividad de la quinasa resultante de la variante del EGFR, en donde se selecciona un compuesto que inhibe la actividad de la quinasa de la variante del EGFR.

75. El método del párrafo 74, en donde se marca la variante del EGFR.
76. El método del párrafo 74, en donde la variante del EGFR se une a un soporte sólido.
77. El método del párrafo 74, en donde dicho soporte sólido es un chip de proteína.
78. Un compuesto que inhibe la actividad catalítica de la quinasa de una variante del EGFR como se identifica en el párrafo 74, en donde el compuesto se selecciona del grupo que consiste de un anticuerpo, un fragmento de anticuerpo, una molécula pequeña, un péptido, una proteína, un ácido nucleico antisentido, una ribozima, PNA, ARNpi, un aptámero de oligonucleótido, y un aptámero de péptido.
79. El método del párrafo 74, en donde el EGFR variante comprende una mutación secundaria en el dominio quinasa del gen erbB1.
80. Una composición farmacéutica que comprende el inhibidor de la actividad quinasa del EGFR identificado en los párrafos 74 - 79.
81. Un método de tratamiento de un paciente que tiene una enfermedad mediada por EGFR, que comprende administrar a dicho paciente la composición farmacéutica del párrafo 80.
82. El método del párrafo 81, en donde la enfermedad mediada por EGFR es cáncer.
83. El método del párrafo 82, en donde el cáncer se selecciona de entre el grupo que consiste en cáncer gastrointestinal, cáncer de próstata, cáncer de ovario, cáncer de mama, cáncer de cabeza y cuello, cáncer de pulmón, cáncer de pulmón de células no pequeñas, cáncer del sistema nervioso, cáncer de riñón, cáncer de retina, cáncer de piel, cáncer de hígado, cáncer de páncreas, cáncer genitourinario y cáncer de vejiga.
84. El método del párrafo 83, en donde el cáncer es cáncer de pulmón de células no pequeñas.
85. Un método para predecir la adquisición de mutaciones secundarias en el dominio quinasa del gen erbB1 que comprende:
 - a. poner en contacto una célula que tiene una forma variante del gen erbB1 con una dosis subletal de un inhibidor de la tirosina quinasa,
 - b. seleccionar células que sean resistentes a un efecto de detención del crecimiento del inhibidor de la tirosina quinasa; y
 - c. analizar el ácido nucleico de erbB1 de dichas células resistentes por la presencia de mutaciones secundarias en el dominio quinasa de erbB1.
86. El método del párrafo 85, en donde la célula está in vitro.
87. El método del párrafo 85, en donde se obtiene la célula de un animal transgénico.
88. El método del párrafo 87, en donde el animal transgénico es un ratón.
89. El método del párrafo 87, en donde se obtiene la célula a partir de una biopsia de tumor.
90. El método del párrafo 85, que comprende además poner en contacto primero las células con una cantidad efectiva de un agente mutagénico.
91. El método del párrafo 90, en el que el agente mutagénico se selecciona de entre el grupo que consiste de metanosulfonato de etilo (EMS), N-etil-N-nitrosourea (ENU), N-metil-N-nitrosourea (MNU), clorhidrato de focarbaxina (Prc), metanosulfonato de metilo (MEMS), clorambucilo (Chl), melfalán, clorhidrato procarbazona, ciclofosfamida (Cp), sulfato de dietilo (Et₂SO₄), monómero de acrilamida (AA), trietilén melamina (TEM), mostaza de nitrógeno, vincristina, dimetilnitrosamina, N-metil-N'-nitro-nitrosoguanidina (MNNG), 7,12-dimetilbenz(a)antraceno (DMBA), óxido de etileno, hexametilfosforamida, bisulfán, y metanosulfonato de etilo (EtMs).
92. El método del párrafo 85, que comprende además la propagación de una forma variante del gen EGFR en una cepa bacteriana deficiente en la reparación del ADN antes de introducirlo en una célula.
93. Un método para seleccionar un compuesto que comprende poner en contacto el compuesto con una variante del receptor del factor de crecimiento epidérmico (EGFR) que tiene una mutación secundaria en el dominio quinasa y la detección de la actividad resultante de la quinasa, en donde se selecciona un compuesto que inhibe la actividad de la quinasa del EGFR variante.
94. El método del párrafo 93, en donde la mutación secundaria da como resultado una resistencia a gefitinib o erlotinib.

95. Un ácido nucleico aislado que comprende la SEQ ID NO: 495.
96. Un ácido nucleico aislado que comprende la SEQ ID NO: 511, en donde se suprimen los nucleótidos 2235 a 2249.
97. Un ácido nucleico aislado que comprende la SEQ ID NO: 497.
- 5 98. Un ácido nucleico aislado que comprende la SEQ ID NO: 511, en donde se suprimen los nucleótidos 2240 a 2251.
99. Un ácido nucleico aislado que comprende la SEQ ID NO: 499.
100. Un ácido nucleico aislado que comprende la SEQ ID NO: 511, en donde se suprimen los nucleótidos 2240 a 2257.
- 10 101. Un ácido nucleico aislado que comprende la SEQ ID NO: 502.
102. Un ácido nucleico aislado que comprende la SEQ ID NO: 511, en donde se sustituye la guanina en el nucleótido 2573 por una timina.
103. Un ácido nucleico aislado que comprende la SEQ ID NO: 504.
104. Un ácido nucleico aislado que comprende la SEQ ID NO: 511, en donde se sustituye la adenina en el nucleótido 2582 por una timina.
- 15 105. Un ácido nucleico aislado de la SEQ ID NO: 511, en donde se sustituye una timina en el nucleótido 2155 por una guanina.
106. Una proteína aislada que tiene la secuencia de aminoácidos de la SEQ ID NO: 512, en donde se suprimen los aminoácidos 746 a 750.
- 20 107. Una proteína aislada que tiene la secuencia de aminoácidos de la SEQ ID NO: 512, en donde se suprimen los aminoácidos 747 - 751.
108. Una proteína aislada que tiene la secuencia de aminoácidos de la SEQ ID NO: 512, en donde se suprimen los aminoácidos 747 a 753.
- 25 109. Una proteína aislada que tiene la secuencia de aminoácidos de la SEQ ID NO: 512, en donde se sustituye la leucina en el aminoácido 858 con una arginina.
110. Una proteína aislada que tiene la secuencia de aminoácidos de la SEQ ID NO: 512, en donde se sustituye la leucina en el aminoácido 861 con una glutamina.
111. Una proteína aislada que tiene la secuencia de aminoácidos de la SEQ ID NO: 512, en donde se sustituye la glicina en el aminoácido 719 con una cisteína.
- 30 112. El método del párrafo 1, en donde la variación de ácido nucleico del gen erbB1 se selecciona de entre el grupo que consiste de una sustitución de una timina por una guanina o una serina por una guanina en el nucleótido 2155 de la SEQ ID NO 511, una supresión de nucleótidos 2235 a 2249, 2240 a 2251, 2240 a 2257, 2236 a 2250, 2254 a 2277 o 2236 a 2244 de la SEQ ID NO: 511, y una sustitución de una guanina por una timina en el nucleótido 2573 o una adenina para una timina en el nucleótido 2582 de la SEQ ID NO: 511.
- 35 113. Un método para determinar la probabilidad de efectividad de un tratamiento dirigido al receptor del factor de crecimiento epidérmico (EGFR) en un paciente afectado con o en riesgo de desarrollar cáncer, que comprende:
 - a. obtener una muestra biológica de dicho paciente; y
 - b. determinar si Akt, STAT5 o STAT3 se activan en dicho paciente, en donde Akt, STAT5, o STAT3 activados indican que dicho tratamiento dirigido al EGFR es probable que sea efectivo.
- 40 114. El método del párrafo 113, en donde la muestra biológica es una biopsia o un aspirado.
115. El método del párrafo 113, en donde Akt, STAT3, o STAT5 activado está fosforilado.
116. El método del párrafo 113, en donde el Akt, STAT5, o STAT3 activado se determina inmunológicamente.
117. El método del párrafo 116, en donde los métodos de detección inmunológica se seleccionan del grupo que consiste de inmunohistoquímica, inmunocitoquímica, barrido por FACS, inmunotransferencia, radioinmunoensayos, transferencia tipo Western, inmunoprecipitación, o ensayos por inmunoabsorción ligados a enzimas (ELISA).
- 45

118. El método del párrafo 116, en donde el método de detección inmunológica es inmunohistoquímica o inmunocitoquímica utilizando anticuerpos anti-fosfo Akt, anti-fosfo STAT3 o anti-fosfo STAT5.

Listado de secuencias

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

5 <120> MÉTODO PARA DETERMINAR LA RESPUESTA DEL CÁNCER A TRATAMIENTOS DIRIGIDOS AL RECEPTOR DEL FACTOR DE CRECIMIENTO EPIDÉRMICO

<130> 030258-55145

<140> PCT/US05/10645

<141> 2005-03-31

10 <150> 60/558,218

<151> 2004-03-31

<150> 60/561,095

<151> 2004-04-09

<150> 60/565,753

15 <151> 2004-04-27

<150> 60/565,985

<151> 2004-04-27

<150> 60/574,035

<151> 2004-05-25

20 <150> 60/577,916

<151> 2004-06-07

<150> 60/592,287

<151> 2004-07-29

<160> 762

25 <170> PatentIn Ver. 3.3

<210> 1

<211> 23

<212> ADN

<213> Homo sapiens

30 <400> 1

ggaaatatag ggaagggaag gaa 23

<210> 2

<211> 22

<212> ADN

35 <213> Homo sapiens

<400> 2

ttgacagggt accaggagat ga 22

<210> 3
 <211> 20
 <212> ADN
 <213> Homo sapiens
 5 <400> 3
 ctgaaccgcc aaggactcat 20
 <210> 4
 <211> 23
 <212> ADN
 10 <213> Homo sapiens
 <400> 4
 tttccctcc ctactaacac acg 23
 <210> 5
 <211> 20
 15 <212> ADN
 <213> Homo sapiens
 <400> 5
 actgatgccc tgaccctgtt 20
 <210> 6
 20 <211> 19
 <212> ADN
 <213> Homo sapiens
 <400> 6
 cccatggttc ccactctt 19
 25 <210> 7
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 7
 30 agggactcca aagccatgtg 20
 <210> 8
 <211> 19
 <212> ADN
 <213> Homo sapiens
 35 <400> 8
 ctctctgggg ccatccact 19
 <210> 9

<211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 9
 5 cattgtcaag ggcaatgtaa gtg 23
 <210> 10
 <211> 20
 <212> ADN
 <213> Homo sapiens
 10 <400> 10
 ctctcaccaa ccctcgctgt 20
 <210> 11
 <211> 20
 <212> ADN
 15 <213> Homo sapiens
 <400> 11
 acatggggag ccagagtgac 20
 <210> 12
 <211> 20
 20 <212> ADN
 <213> Homo sapiens
 <400> 12
 tgcaaccag agaaagtggtg 20
 <210> 13
 25 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 13
 tgagcttca accctagttt gttg 24
 30 <210> 14
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 14
 35 gttgcctcc tgctgtctca 20
 <210> 15
 <211> 24

<212> ADN
 <213> Homo sapiens
 <400> 15
 tgtcctgtg ttttgaaga ttcc 24
 5 <210> 16
 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 16
 10 tgcagacaga tgacaaacat gaa 23
 <210> 17
 <211> 21
 <212> ADN
 <213> Homo sapiens
 15 <400> 17
 tgggtgagtc tctgtgtgga g 21
 <210> 18
 <211> 25
 <212> ADN
 20 <213> Homo sapiens
 <400> 18
 cattgccata gcaaaaataa acaca 25
 <210> 19
 <211> 19
 25 <212> ADN
 <213> Homo sapiens
 <400> 19
 ggtcaactg ggcgtccta 19
 <210> 20
 30 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 20
 ccttctccga ggtggaattg 20
 35 <210> 21
 <211> 20
 <212> ADN

<213> Homo sapiens
 <400> 21
 cgcaccatgg catctcttta 20
 <210> 22
 5 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 22
 aaaacgatct ctatgtccgt ggt 23
 10 <210> 23
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 23
 15 cagccagcca aacaatcaga 20
 <210> 24
 <211> 23
 <212> ADN
 <213> Homo sapiens
 20 <400> 24
 tctttggagt cttcagaggg aaa 23
 <210> 25
 <211> 20
 <212> ADN
 25 <213> Homo sapiens
 <400> 25
 tgtggttcg ttgaagcaa 20
 <210> 26
 <211> 20
 30 <212> ADN
 <213> Homo sapiens
 <400> 26
 aattgacagc tccccacag 20
 <210> 27
 35 <211> 20
 <212> ADN
 <213> Homo sapiens

<400> 27
 ggctttctga cgggagtaa 20
 <210> 28
 <211> 21
 5 <212> ADN
 <213> Homo sapiens
 <400> 28
 ccacccaaag acttccaag a 21
 <210> 29
 10 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 29
 cctttccatc acccctcaag 20
 15 <210> 30
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 30
 20 agtgccttc cattgcctaa 20
 <210> 31
 <211> 20
 <212> ADN
 <213> Homo sapiens
 25 <400> 31
 accggaattc cttctgctt 20
 <210> 32
 <211> 23
 <212> ADN
 30 <213> Homo sapiens
 <400> 32
 cactgaaaca aacaacagg tga 23
 <210> 33
 <211> 20
 35 <212> ADN
 <213> Homo sapiens
 <400> 33

agggggtgag tcacaggttc 20
 <210> 34
 <211> 26
 <212> ADN
 5 <213> Homo sapiens
 <400> 34
 tcagaagaaa tgttttatt ccaagg 26
 <210> 35
 <211> 21
 10 <212> ADN
 <213> Homo sapiens
 <400> 35
 gcaaatccaa tttcccact t 21
 <210> 36
 15 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 36
 gcaggagctc tgtgccctat 20
 20 <210> 37
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 37
 25 tcccacagca tgacctacca 20
 <210> 38
 <211> 23
 <212> ADN
 <213> Homo sapiens
 30 <400> 38
 ttgcttctt aaggaactga aaa 23
 <210> 39
 <211> 20
 <212> ADN
 35 <213> Homo sapiens
 <400> 39
 tgtcacccaa ggatcatggag 20

<210> 40
 <211> 20
 <212> ADN
 <213> Homo sapiens
 5 <400> 40
 caaaagccaa gggcaaagaa 20
 <210> 41
 <211> 20
 <212> ADN
 10 <213> Homo sapiens
 <400> 41
 ggagtcccaa ctccttgacc 20
 <210> 42
 <211> 19
 15 <212> ADN
 <213> Homo sapiens
 <400> 42
 gtcctgccca cacaggatg 19
 <210> 43
 20 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 43
 gctttcccca ctcacacaca 20
 25 <210> 44
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 44
 30 caaacctcgg caattgttg 20
 <210> 45
 <211> 20
 <212> ADN
 <213> Homo sapiens
 35 <400> 45
 ccaccaatcc aacatccaga 20
 <210> 46

<211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 46
 5 tggcccagag ccatagaaac 20
 <210> 47
 <211> 26
 <212> ADN
 <213> Homo sapiens
 10 <400> 47
 ttccaagatc attctacaag atgtca 26
 <210> 48
 <211> 23
 <212> ADN
 15 <213> Homo sapiens
 <400> 48
 gcacattcag agattctttc tgc 23
 <210> 49
 <211> 20
 20 <212> ADN
 <213> Homo sapiens
 <400> 49
 tccaaatgag ctggcaagtg 20
 <210> 50
 25 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 50
 tcccaaacac tcagtgaaac aaa 23
 30 <210> 51
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 51
 35 gtgcatcgct ggtaacatcc 20
 <210> 52
 <211> 20

<212> ADN
 <213> Homo sapiens
 <400> 52
 tgtggagatg agcaggggtct 20
 5 <210> 53
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 53
 10 atcgcatcca tgcgtcttca 20
 <210> 54
 <211> 20
 <212> ADN
 <213> Homo sapiens
 15 <400> 54
 atcccatgg caaactcttg 20
 <210> 55
 <211> 19
 <212> ADN
 20 <213> Homo sapiens
 <400> 55
 gctcagagcc tggcatgaa 19
 <210> 56
 <211> 19
 25 <212> ADN
 <213> Homo sapiens
 <400> 56
 catcctcccc tgcattgt 19
 <210> 57
 30 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 57
 tggctcgtct gtgtgtgtca 20
 35 <210> 58
 <211> 25
 <212> ADN

<213> Homo sapiens
 <400> 58
 cgaaagaaaa tacttgcattg tcaga 25
 <210> 59
 5 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 59
 tgaagcaaat tgccaagac 20
 10 <210> 60
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 60
 15 tgacatttct ccagggatgc 20
 <210> 61
 <211> 20
 <212> ADN
 <213> Homo sapiens
 20 <400> 61
 aagtgtcgca tcaccaatgc 20
 <210> 62
 <211> 19
 <212> ADN
 25 <213> Homo sapiens
 <400> 62
 atgcatctg ggacacagg 19
 <210> 63
 <211> 20
 30 <212> ADN
 <213> Homo sapiens
 <400> 63
 ggcacctgct ggcaatagac 20
 <210> 64
 35 <211> 27
 <212> ADN
 <213> Homo sapiens

<400> 64
 tgacttcata tccatgtgag ttctact 27
 <210> 65
 <211> 20
 5 <212> ADN
 <213> Homo sapiens
 <400> 65
 tataccctcc atgaggcaca 20
 <210> 66
 10 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 66
 gggaaaaacc cacacaggaa 20
 15 <210> 67
 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 67
 20 tcagaaccag catctcaagg a 21
 <210> 68
 <211> 19
 <212> ADN
 <213> Homo sapiens
 25 <400> 68
 gatgctggag ggagcacct 19
 <210> 69
 <211> 22
 <212> ADN
 30 <213> Homo sapiens
 <400> 69
 ccttggtgag gacattcaca gg 22
 <210> 70
 <211> 20
 35 <212> ADN
 <213> Homo sapiens
 <400> 70

atgtgcccgga ggtggaagta 20
 <210> 71
 <211> 20
 <212> ADN
 5 <213> Homo sapiens
 <400> 71
 ggagggcaga ggactagctg 20
 <210> 72
 <211> 20
 10 <212> ADN
 <213> Homo sapiens
 <400> 72
 gtgcctggcc aagtcttgt 20
 <210> 73
 15 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 73
 ctgcagccta gcaacagagc 20
 20 <210> 74
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 74
 25 aagaaccaga ggagccagga 20
 <210> 75
 <211> 20
 <212> ADN
 <213> Homo sapiens
 30 <400> 75
 cgggtaagga tgtgggtgt 20
 <210> 76
 <211> 20
 <212> ADN
 35 <213> Homo sapiens
 <400> 76
 cagggtgttct gcctcctgaa 20

<210> 77
 <211> 20
 <212> ADN
 <213> Homo sapiens
 5 <400> 77
 gcttcaggag gcagaacacc 20
 <210> 78
 <211> 22
 <212> ADN
 10 <213> Homo sapiens
 <400> 78
 ggagcaagcc taagaagggt ca 22
 <210> 79
 <211> 21
 15 <212> ADN
 <213> Homo sapiens
 <400> 79
 gccttgatc cattgccac a 21
 <210> 80
 20 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 80
 tgacaacacg ttttgggtca t 21
 25 <210> 81
 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 81
 30 tgcattatcc attcagaac aga 23
 <210> 82
 <211> 23
 <212> ADN
 <213> Homo sapiens
 35 <400> 82
 aaacagtttc attgctgcta aat 23
 <210> 83

<211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 83
 5 ccggatacag ataccaaaa aga 23
 <210> 84
 <211> 20
 <212> ADN
 <213> Homo sapiens
 10 <400> 84
 ggaggcttca agggatgaga 20
 <210> 85
 <211> 20
 <212> ADN
 15 <213> Homo sapiens
 <400> 85
 gctgtgtcc tgcttgcta 20
 <210> 86
 <211> 25
 20 <212> ADN
 <213> Homo sapiens
 <400> 86
 tgggtgtaat gttgaactag ctgac 25
 <210> 87
 25 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 87
 tggctgtcag ctaaataagc atgt 24
 30 <210> 88
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 88
 35 tcaatttgct tcatttctcc tggt 24
 <210> 89
 <211> 21

<212> ADN
 <213> Homo sapiens
 <400> 89
 tgctgctgaa ctaccaacca a 21
 5 <210> 90
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 90
 10 tgtggtagta attgtggaaa actg 24
 <210> 91
 <211> 20
 <212> ADN
 <213> Homo sapiens
 15 <400> 91
 caaagcacccg tctcaactcg 20
 <210> 92
 <211> 20
 <212> ADN
 20 <213> Homo sapiens
 <400> 92
 cccgaaactg ccaactcat 20
 <210> 93
 <211> 20
 25 <212> ADN
 <213> Homo sapiens
 <400> 93
 ggaaaacagg accccagtg 20
 <210> 94
 30 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 94
 ccctcctcca cagagctgat 20
 35 <210> 95
 <211> 21
 <212> ADN

<213> Homo sapiens
 <400> 95
 gacagaagct gacaagcagc a 21
 <210> 96
 5 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 96
 aggttcatt ccctcccagt 20
 10 <210> 97
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 97
 15 tgggagtgag agttggaag aa 22
 <210> 98
 <211> 20
 <212> ADN
 <213> Homo sapiens
 20 <400> 98
 tatgaggccg tgagctgaaa 20
 <210> 99
 <211> 18
 <212> ADN
 25 <213> Homo sapiens
 <400> 99
 agggccctgc tctggtt 18
 <210> 100
 <211> 20
 30 <212> ADN
 <213> Homo sapiens
 <400> 100
 ccaattgggc gttagtgaaa 20
 <210> 101
 35 <211> 20
 <212> ADN
 <213> Homo sapiens

<400> 101
 ctcatgagat tggggcatca 20
 <210> 102
 <211> 20
 5 <212> ADN
 <213> Homo sapiens
 <400> 102
 aggcccatga tctcagaagc 20
 <210> 103
 10 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 103
 ggttgagga gagacgaggt 20
 15 <210> 104
 <211> 19
 <212> ADN
 <213> Homo sapiens
 <400> 104
 20 aggccctca ccctgtgac 19
 <210> 105
 <211> 20
 <212> ADN
 <213> Homo sapiens
 25 <400> 105
 atgaccctc cgatcctacc 20
 <210> 106
 <211> 20
 <212> ADN
 30 <213> Homo sapiens
 <400> 106
 taatcctgct ccacggcatt 20
 <210> 107
 <211> 20
 35 <212> ADN
 <213> Homo sapiens
 <400> 107

ggaaaaagca gaggcaggtg 20
 <210> 108
 <211> 20
 <212> ADN
 5 <213> Homo sapiens
 <400> 108
 tgggtctcaag aaccagcag 20
 <210> 109
 <211> 19
 10 <212> ADN
 <213> Homo sapiens
 <400> 109
 gacaccctcc ccctctcat 19
 <210> 110
 15 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 110
 actatgacac cccggctgag 20
 20 <210> 111
 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 111
 25 tgcttgatgt aaaacccttg g 21
 <210> 112
 <211> 20
 <212> ADN
 <213> Homo sapiens
 30 <400> 112
 gcaatccaac agccatgaga 20
 <210> 113
 <211> 20
 <212> ADN
 35 <213> Homo sapiens
 <400> 113
 ggagcaaacc cctatgtcca 20

<210> 114
 <211> 21
 <212> ADN
 <213> Homo sapiens
 5 <400> 114
 tctccaact gtgtgttg g 21
 <210> 115
 <211> 19
 <212> ADN
 10 <213> Homo sapiens
 <400> 115
 tggggaccac tgctgagag 19
 <210> 116
 <211> 20
 15 <212> ADN
 <213> Homo sapiens
 <400> 116
 tgcagccttc tctcctgaa 20
 <210> 117
 20 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 117
 gcagagcagt gtggcagaag 20
 25 <210> 118
 <211> 19
 <212> ADN
 <213> Homo sapiens
 <400> 118
 30 acaggtggga agggactgg 19
 <210> 119
 <211> 19
 <212> ADN
 <213> Homo sapiens
 35 <400> 119
 agtggggtgg gctgagaac 19
 <210> 120

<211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 120
 5 tctctggggc agaaagagga 20
 <210> 121
 <211> 20
 <212> ADN
 <213> Homo sapiens
 10 <400> 121
 acccggccac actgtatttc 20
 <210> 122
 <211> 20
 <212> ADN
 15 <213> Homo sapiens
 <400> 122
 catcccaccc agctctcaac 20
 <210> 123
 <211> 20
 20 <212> ADN
 <213> Homo sapiens
 <400> 123
 agggcatagc cctattgagc 20
 <210> 124
 25 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 124
 cccaggaaaa agccagagaa 20
 30 <210> 125
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 125
 35 caggtgtggg tggagtaggc 20
 <210> 126
 <211> 18

<212> ADN
 <213> Homo sapiens
 <400> 126
 ctcaggcgcc atccactt 18
 5 <210> 127
 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 127
 10 aagaagacga ccaacgtgag c 21
 <210> 128
 <211> 18
 <212> ADN
 <213> Homo sapiens
 15 <400> 128
 aggagctcca gggcacag 18
 <210> 129
 <211> 22
 <212> ADN
 20 <213> Homo sapiens
 <400> 129
 cctcctctgt aaagtgggtg ga 22
 <210> 130
 <211> 20
 25 <212> ADN
 <213> Homo sapiens
 <400> 130
 agagggcctc agtcagagt 20
 <210> 131
 30 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 131
 agatggggca gaactggatg 20
 35 <210> 132
 <211> 20
 <212> ADN

<213> Homo sapiens
 <400> 132
 ggggtcccaga ccaaattctga 20
 <210> 133
 5 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 133
 aggtgctccc ttacagcat 20
 10 <210> 134
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 134
 15 ttcagggact acagctgagg aa 22
 <210> 135
 <211> 21
 <212> ADN
 <213> Homo sapiens
 20 <400> 135
 gccgtatgtt atctgggagg t 21
 <210> 136
 <211> 21
 <212> ADN
 25 <213> Homo sapiens
 <400> 136
 tgggcccatt acactttaag a 21
 <210> 137
 <211> 20
 30 <212> ADN
 <213> Homo sapiens
 <400> 137
 ttccatcacc ggtacctcct 20
 <210> 138
 35 <211> 23
 <212> ADN
 <213> Homo sapiens

<400> 138
 ccataaatca aaaatgcacc aca 23
 <210> 139
 <211> 23
 5 <212> ADN
 <213> Homo sapiens
 <400> 139
 gagtggctct aggaagatga tgc 23
 <210> 140
 10 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 140
 aaagtcattg gctgcaatac aa 22
 15 <210> 141
 <211> 18
 <212> ADN
 <213> Homo sapiens
 <400> 141
 20 atggccccca ctgcttgg 18
 <210> 142
 <211> 20
 <212> ADN
 <213> Homo sapiens
 25 <400> 142
 aggagctcac ctcaccctgt 20
 <210> 143
 <211> 20
 <212> ADN
 30 <213> Homo sapiens
 <400> 143
 ccttgctgt ctccacacat 20
 <210> 144
 <211> 20
 35 <212> ADN
 <213> Homo sapiens
 <400> 144

tggcaacggg taacaatgaa 20
 <210> 145
 <211> 20
 <212> ADN
 5 <213> Homo sapiens
 <400> 145
 ggctgaggta agctgcttcg 20
 <210> 146
 <211> 23
 10 <212> ADN
 <213> Homo sapiens
 <400> 146
 aaaaagaagt atctgcccc ttt 23
 <210> 147
 15 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 147
 aaccctctt agggctctgt g 21
 20 <210> 148
 <211> 19
 <212> ADN
 <213> Homo sapiens
 <400> 148
 25 caggaggatg gcaggcttc 19
 <210> 149
 <211> 23
 <212> ADN
 <213> Homo sapiens
 30 <400> 149
 cgtagagagc ttcaggacct gtg 23
 <210> 150
 <211> 20
 <212> ADN
 35 <213> Homo sapiens
 <400> 150
 ttccgagaag tttgcctga 20

<210> 151
 <211> 22
 <212> ADN
 <213> Homo sapiens
 5 <400> 151
 tgtgaacatc attcaaggcg ta 22
 <210> 152
 <211> 25
 <212> ADN
 10 <213> Homo sapiens
 <400> 152
 aaaatgtgtg atatccctag acagg 25
 <210> 153
 <211> 20
 15 <212> ADN
 <213> Homo sapiens
 <400> 153
 tccacatttc agcaacagca 20
 <210> 154
 20 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 154
 ggctgcttcc tgagacacag t 21
 25 <210> 155
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 155
 30 tatctaccgg tgcgggactt 20
 <210> 156
 <211> 20
 <212> ADN
 <213> Homo sapiens
 35 <400> 156
 aggtgtagcc tcccctcaca 20
 <210> 157

<211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 157
 5 aggctggtgg tgtctctgtg 20
 <210> 158
 <211> 20
 <212> ADN
 <213> Homo sapiens
 10 <400> 158
 caagctgccca accctcagtt 20
 <210> 159
 <211> 24
 <212> ADN
 15 <213> Homo sapiens
 <400> 159
 tggatttcaa atactgaagc cact 24
 <210> 160
 <211> 22
 20 <212> ADN
 <213> Homo sapiens
 <400> 160
 tggaattggt ggtgttgaat tt 22
 <210> 161
 25 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 161
 gggcttcata tgtctgaca tgg 23
 30 <210> 162
 <211> 19
 <212> ADN
 <213> Homo sapiens
 <400> 162
 35 cagaggacca cgccatagg 19
 <210> 163
 <211> 20

<212> ADN
 <213> Homo sapiens
 <400> 163
 ccgagattta gccaccagga 20
 5 <210> 164
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 164
 10 cctgggaagc aaacaacaca 20
 <210> 165
 <211> 20
 <212> ADN
 <213> Homo sapiens
 15 <400> 165
 aggtccccag tctcctctcc 20
 <210> 166
 <211> 20
 <212> ADN
 20 <213> Homo sapiens
 <400> 166
 agaccatgc agccatccta 20
 <210> 167
 <211> 19
 25 <212> ADN
 <213> Homo sapiens
 <400> 167
 cgtgaaccac cgagcttgt 19
 <210> 168
 30 <211> 19
 <212> ADN
 <213> Homo sapiens
 <400> 168
 agaggggcag aaggggaac 19
 35 <210> 169
 <211> 20
 <212> ADN

<213> Homo sapiens
 <400> 169
 ggtgggggtg aggagcttag 20
 <210> 170
 5 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 170
 tcgtttaagc caccagtc 20
 10 <210> 171
 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 171
 15 tgcaaataag gaaagcaaac a 21
 <210> 172
 <211> 21
 <212> ADN
 <213> Homo sapiens
 20 <400> 172
 tcctgacatg gtctccaac c 21
 <210> 173
 <211> 22
 <212> ADN
 25 <213> Homo sapiens
 <400> 173
 cagcatcttc acacacctct ga 22
 <210> 174
 <211> 20
 30 <212> ADN
 <213> Homo sapiens
 <400> 174
 gctggctcta aatcccacct 20
 <210> 175
 35 <211> 20
 <212> ADN
 <213> Homo sapiens

<400> 175
 ctaatccggg aagttgtgc 20
 <210> 176
 <211> 24
 5 <212> ADN
 <213> Homo sapiens
 <400> 176
 ttctgtatca gcagcttctc tgtg 24
 <210> 177
 10 <211> 19
 <212> ADN
 <213> Homo sapiens
 <400> 177
 caagtgccac catggatca 19
 15 <210> 178
 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 178
 20 ggcagtgtac tgacccttg a 21
 <210> 179
 <211> 24
 <212> ADN
 <213> Homo sapiens
 25 <400> 179
 gcacaagtta ttaagagccc aagg 24
 <210> 180
 <211> 20
 <212> ADN
 30 <213> Homo sapiens
 <400> 180
 agcatactgg cctcacacca 20
 <210> 181
 <211> 22
 35 <212> ADN
 <213> Homo sapiens
 <400> 181

	gcacatgggc agtgtgtat tt	22
	<210> 182	
	<211> 20	
	<212> ADN	
5	<213> Homo sapiens	
	<400> 182	
	gagccccaca cagatttcct	20
	<210> 183	
	<211> 20	
10	<212> ADN	
	<213> Homo sapiens	
	<400> 183	
	atgggacgga gaagtgttg	20
	<210> 184	
15	<211> 20	
	<212> ADN	
	<213> Homo sapiens	
	<400> 184	
	tccctgtatc agggctcgtc	20
20	<210> 185	
	<211> 20	
	<212> ADN	
	<213> Homo sapiens	
	<400> 185	
25	ttcctacgca gcacaccaat	20
	<210> 186	
	<211> 20	
	<212> ADN	
	<213> Homo sapiens	
30	<400> 186	
	gcaggcacta aaccctttcc	20
	<210> 187	
	<211> 20	
	<212> ADN	
35	<213> Homo sapiens	
	<400> 187	
	gcacgcatgt gaccaatttc	20

<210> 188
 <211> 20
 <212> ADN
 <213> Homo sapiens
 5 <400> 188
 agccctgaga gggaggtagg 20
 <210> 189
 <211> 19
 <212> ADN
 10 <213> Homo sapiens
 <400> 189
 cacacaccac ccctctgct 19
 <210> 190
 <211> 21
 15 <212> ADN
 <213> Homo sapiens
 <400> 190
 aaagattgg ggtgagggt a 21
 <210> 191
 20 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 191
 ctgaaagctc agggataggg 20
 25 <210> 192
 <211> 18
 <212> ADN
 <213> Homo sapiens
 <400> 192
 30 ctggccaagc tgcacaga 18
 <210> 193
 <211> 20
 <212> ADN
 <213> Homo sapiens
 35 <400> 193
 tgcagccaac gattgaaag 20
 <210> 194

<211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 194
 5 ggaaagcccc aagtctgaaa 20
 <210> 195
 <211> 21
 <212> ADN
 <213> Homo sapiens
 10 <400> 195
 tcatcatgag atccccacac t 21
 <210> 196
 <211> 20
 <212> ADN
 15 <213> Homo sapiens
 <400> 196
 gcatttcccc ctgaaggagt 20
 <210> 197
 <211> 23
 20 <212> ADN
 <213> Homo sapiens
 <400> 197
 tggattcagt aaccaggaag tga 23
 <210> 198
 25 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 198
 cccattccac caggatgatt 20
 30 <210> 199
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 199
 35 gtttcagct gccactacc 20
 <210> 200
 <211> 20

<212> ADN
 <213> Homo sapiens
 <400> 200
 gctcgaaacc acatgtcca 20
 5 <210> 201
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 201
 10 ctggatttgg ggtctctgc 20
 <210> 202
 <211> 22
 <212> ADN
 <213> Homo sapiens
 15 <400> 202
 cggaacagc tagcagattt tt 22
 <210> 203
 <211> 20
 <212> ADN
 20 <213> Homo sapiens
 <400> 203
 gggaatttg gaggggaact 20
 <210> 204
 <211> 20
 25 <212> ADN
 <213> Homo sapiens
 <400> 204
 gcttcagtca ccacagagca 20
 <210> 205
 30 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 205
 tgagtctacc cagcaatcat ttg 23
 35 <210> 206
 <211> 20
 <212> ADN

<213> Homo sapiens
 <400> 206
 ttcccgagag ctacaggaca 20
 <210> 207
 5 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 207
 ggtaacaagg gtaccacga a 21
 10 <210> 208
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 208
 15 gtttgagggg ctgagtgtg 20
 <210> 209
 <211> 20
 <212> ADN
 <213> Homo sapiens
 20 <400> 209
 cctcacctt agggctgtg 20
 <210> 210
 <211> 20
 <212> ADN
 25 <213> Homo sapiens
 <400> 210
 agcccaggtc atgcctaga 20
 <210> 211
 <211> 20
 30 <212> ADN
 <213> Homo sapiens
 <400> 211
 gggtagcttg ggagcaaaga 20
 <210> 212
 35 <211> 20
 <212> ADN
 <213> Homo sapiens

<400> 212
 ccaaacccca gagagcagac 20
 <210> 213
 <211> 22
 5 <212> ADN
 <213> Homo sapiens
 <400> 213
 catttcccct aatcctttc ca 22
 <210> 214
 10 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 214
 gtgatcccag atttaggcct tc 22
 15 <210> 215
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 215
 20 gcctctcgtg gttgttttg tc 22
 <210> 216
 <211> 22
 <212> ADN
 <213> Homo sapiens
 25 <400> 216
 cccagggtag ggtccaataa tc 22
 <210> 217
 <211> 22
 <212> ADN
 30 <213> Homo sapiens
 <400> 217
 cttcctggtg gaggtgactg at 22
 <210> 218
 <211> 22
 35 <212> ADN
 <213> Homo sapiens
 <400> 218

caggcatagt gtgtgatggt ca 22
 <210> 219
 <211> 23
 <212> ADN
 5 <213> Homo sapiens
 <400> 219
 tcacgataca cattctcaga tcc 23
 <210> 220
 <211> 22
 10 <212> ADN
 <213> Homo sapiens
 <400> 220
 gaagatctcc cagaggagga tg 22
 <210> 221
 15 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 221
 cgtaacgtgc tgttgaccaa t 21
 20 <210> 222
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 222
 25 aaacgaggga agagccagaa ag 22
 <210> 223
 <211> 22
 <212> ADN
 <213> Homo sapiens
 30 <400> 223
 tggggagcac aataaaagaa ga 22
 <210> 224
 <211> 22
 <212> ADN
 35 <213> Homo sapiens
 <400> 224
 actcttggct cctggattct tg 22

<210> 225
 <211> 22
 <212> ADN
 <213> Homo sapiens
 5 <400> 225
 ggaagtcagt gtcagggaa ta 22
 <210> 226
 <211> 22
 <212> ADN
 10 <213> Homo sapiens
 <400> 226
 ttttagcaga aataggcaag ca 22
 <210> 227
 <211> 24
 15 <212> ADN
 <213> Homo sapiens
 <400> 227
 tggtaatcct aaacacaatg caga 24
 <210> 228
 20 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 228
 ctgggcaaca cagtgagatc ct 22
 25 <210> 229
 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 229
 30 tcacaaattt ctttgctgtg tcc 23
 <210> 230
 <211> 23
 <212> ADN
 <213> Homo sapiens
 35 <400> 230
 catggaactc cagattagcc tgt 23
 <210> 231

<211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 231
 5 gattgttgca gatcgtggac at 22
 <210> 232
 <211> 22
 <212> ADN
 <213> Homo sapiens
 10 <400> 232
 cgcttaaadc ttccattcc ag 22
 <210> 233
 <211> 22
 <212> ADN
 15 <213> Homo sapiens
 <400> 233
 ctccatggca ccatcattaa ca 22
 <210> 234
 <211> 22
 20 <212> ADN
 <213> Homo sapiens
 <400> 234
 ctcaggacac aagtgctctg ct 22
 <210> 235
 25 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 235
 gcagttcatg gttcatcttc tttt 24
 30 <210> 236
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 236
 35 caaaatagcc caccctggat ta 22
 <210> 237
 <211> 21

<212> ADN
 <213> Homo sapiens
 <400> 237
 ctttctgcat tgccaagat g 21
 5 <210> 238
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 238
 10 caaggtctca gtgagtgggtg ga 22
 <210> 239
 <211> 23
 <212> ADN
 <213> Homo sapiens
 15 <400> 239
 gagaaggggtc ttctgactc tgc 23
 <210> 240
 <211> 22
 <212> ADN
 20 <213> Homo sapiens
 <400> 240
 caggtgtttc tctgtgagg tg 22
 <210> 241
 <211> 22
 25 <212> ADN
 <213> Homo sapiens
 <400> 241
 cacattgcgg cctagaatgt ta 22
 <210> 242
 30 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 242
 accccgtcac aaccttcagt 20
 35 <210> 243
 <211> 22
 <212> ADN

<213> Homo sapiens
 <400> 243
 gccgtagccc caaagtgta ta 22
 <210> 244
 5 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 244
 tcagctcaaa cctgtgattt cc 22
 10 <210> 245
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 245
 15 ctcaactctcc ataaatgcta cgaa 24
 <210> 246
 <211> 22
 <212> ADN
 <213> Homo sapiens
 20 <400> 246
 gacttaacgt gtcccctttt gc 22
 <210> 247
 <211> 22
 <212> ADN
 25 <213> Homo sapiens
 <400> 247
 gcctcttcgg ggtaatcaga ta 22
 <210> 248
 <211> 22
 30 <212> ADN
 <213> Homo sapiens
 <400> 248
 gaagtctgtg gtttagcgga ca 22
 <210> 249
 35 <211> 22
 <212> ADN
 <213> Homo sapiens

<400> 249
 atcttttgcc tggaggaact tt 22
 <210> 250
 <211> 22
 5 <212> ADN
 <213> Homo sapiens
 <400> 250
 cagggtaaat tcatccatt ga 22
 <210> 251
 10 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 251
 cagcagccag cacaactact tt 22
 15 <210> 252
 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 252
 20 ttggctagat gaaccattga tga 23
 <210> 253
 <211> 24
 <212> ADN
 <213> Homo sapiens
 25 <400> 253
 tgaatgaagc tcctgtgtt actc 24
 <210> 254
 <211> 22
 <212> ADN
 30 <213> Homo sapiens
 <400> 254
 atgttcacg caggctaag tg 22
 <210> 255
 <211> 24
 35 <212> ADN
 <213> Homo sapiens
 <400> 255

aaaacaggga gaacttctaa gcaa 24
 <210> 256
 <211> 21
 <212> ADN
 5 <213> Homo sapiens
 <400> 256
 catggcagag tcattccac t 21
 <210> 257
 <211> 22
 10 <212> ADN
 <213> Homo sapiens
 <400> 257
 caatgctaga acaacgcctg tc 22
 <210> 258
 15 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 258
 tccctccact gaggacaaag tt 22
 20 <210> 259
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 259
 25 gggagagctt gagaaagtg ga 22
 <210> 260
 <211> 22
 <212> ADN
 <213> Homo sapiens
 30 <400> 260
 atttctcgg atggatgtac ca 22
 <210> 261
 <211> 22
 <212> ADN
 35 <213> Homo sapiens
 <400> 261
 tcagagcctg tgtttctacc aa 22

<210> 262
 <211> 22
 <212> ADN
 <213> Homo sapiens
 5 <400> 262
 tggctcacaccgacctga tt 22
 <210> 263
 <211> 24
 <212> ADN
 10 <213> Homo sapiens
 <400> 263
 aaataatcag tggattcgt ggag 24
 <210> 264
 <211> 22
 15 <212> ADN
 <213> Homo sapiens
 <400> 264
 gaggccagtg ctgtctctaa gg 22
 <210> 265
 20 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 265
 acttcacagc cctgcgtaaa c 21
 25 <210> 266
 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 266
 30 atgggacagg cactgattg t 21
 <210> 267
 <211> 22
 <212> ADN
 <213> Homo sapiens
 35 <400> 267
 gcagcgggtt acatcttctt tc 22
 <210> 268

<211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 268
 5 cagctctggc tcacactacc ag 22
 <210> 269
 <211> 22
 <212> ADN
 <213> Homo sapiens
 10 <400> 269
 cctgaactcc gtcagactga aa 22
 <210> 270
 <211> 20
 <212> ADN
 15 <213> Homo sapiens
 <400> 270
 gcagctggac tcgatttcct 20
 <210> 271
 <211> 23
 20 <212> ADN
 <213> Homo sapiens
 <400> 271
 ccttacagca atcctgtgaa aca 23
 <210> 272
 25 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 272
 tgcccaatga gtcaagaagt gt 22
 30 <210> 273
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 273
 35 atgtacagtg ctggcatggt ct 22
 <210> 274
 <211> 22

<212> ADN
 <213> Homo sapiens
 <400> 274
 cactcacgga tgctgcttag tt 22
 5 <210> 275
 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 275
 10 taaggcaccc acatcatgtc a 21
 <210> 276
 <211> 24
 <212> ADN
 <213> Homo sapiens
 15 <400> 276
 tggacctaaa aggcttaca tcaa 24
 <210> 277
 <211> 24
 <212> ADN
 20 <213> Homo sapiens
 <400> 277
 gccttttagg tcactatgg aatg 24
 <210> 278
 <211> 22
 25 <212> ADN
 <213> Homo sapiens
 <400> 278
 ccaggcgatg ctactactgg tc 22
 <210> 279
 30 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 279
 tcatagcaca cctccctcac tg 22
 35 <210> 280
 <211> 22
 <212> ADN

<213> Homo sapiens
 <400> 280
 acacaacaaa gagcttgtag ag 22
 <210> 281
 5 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 281
 ccattacttt gagaaggaca ggaa 24
 10 <210> 282
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 282
 15 tattcttgct ggatgcgttt ct 22
 <210> 283
 <211> 21
 <212> ADN
 <213> Homo sapiens
 20 <400> 283
 aggagggcag aggactagct g 21
 <210> 284
 <211> 20
 <212> ADN
 25 <213> Homo sapiens
 <400> 284
 ggcaatgtga atgtgcactg 20
 <210> 285
 <211> 20
 30 <212> ADN
 <213> Homo sapiens
 <400> 285
 ctggaacctg ggaggtggag 20
 <210> 286
 35 <211> 22
 <212> ADN
 <213> Homo sapiens

<400> 286
 atcagggtgg gaggagtaaa ga 22
 <210> 287
 <211> 22
 5 <212> ADN
 <213> Homo sapiens
 <400> 287
 ccacttacc tctcacctgt gc 22
 <210> 288
 10 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 288
 gtgaactcc ggtaggaaat gg 22
 15 <210> 289
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 289
 20 aggggacctc aagggagaag 20
 <210> 290
 <211> 22
 <212> ADN
 <213> Homo sapiens
 25 <400> 290
 agatcatgcc agtgaactcc ag 22
 <210> 291
 <211> 22
 <212> ADN
 30 <213> Homo sapiens
 <400> 291
 ggaccaggaa agtccttgct tt 22
 <210> 292
 <211> 22
 35 <212> ADN
 <213> Homo sapiens
 <400> 292

ggtggggaac attaaactga gg 22
 <210> 293
 <211> 22
 <212> ADN
 5 <213> Homo sapiens
 <400> 293
 gcttcagggt gttttgtgc ag 22
 <210> 294
 <211> 22
 10 <212> ADN
 <213> Homo sapiens
 <400> 294
 acccttgctt gagggaaata tg 22
 <210> 295
 15 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 295
 cccagctcct aggggtacagt ct 22
 20 <210> 296
 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 296
 25 cagtcagctt caaaatccct ctt 23
 <210> 297
 <211> 24
 <212> ADN
 <213> Homo sapiens
 30 <400> 297
 tcacttcctt gtgagtaaag aaaa 24
 <210> 298
 <211> 23
 <212> ADN
 35 <213> Homo sapiens
 <400> 298
 ggccatttaa ttctgtcct tga 23

<210> 299
 <211> 23
 <212> ADN
 <213> Homo sapiens
 5 <400> 299
 tggactgtg caaactcaaa ctg 23
 <210> 300
 <211> 23
 <212> ADN
 10 <213> Homo sapiens
 <400> 300
 tccaatata gggcagtcac gtt 23
 <210> 301
 <211> 23
 15 <212> ADN
 <213> Homo sapiens
 <400> 301
 tctcaatcag ttgagtgcc ttg 23
 <210> 302
 20 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 302
 agctgtgcaa gtgtggaac at 22
 25 <210> 303
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 303
 30 gctgtgagg taaatgagac ca 22
 <210> 304
 <211> 22
 <212> ADN
 <213> Homo sapiens
 35 <400> 304
 gtctctggt gagtgactgt gg 22
 <210> 305

<211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 305
 5 ccttccttcg tctccacagc 20
 <210> 306
 <211> 21
 <212> ADN
 <213> Homo sapiens
 10 <400> 306
 gtccttgtag caacagtcga g 21
 <210> 307
 <211> 22
 <212> ADN
 15 <213> Homo sapiens
 <400> 307
 gcttggcaag gagaagagaa ca 22
 <210> 308
 <211> 23
 20 <212> ADN
 <213> Homo sapiens
 <400> 308
 gcttgctttc ttgctgaac aac 23
 <210> 309
 25 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 309
 gctggtcacc ttgagcttct ct 22
 30 <210> 310
 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 310
 35 ccatgctggg ctcttgatt a 21
 <210> 311
 <211> 22

<212> ADN
 <213> Homo sapiens
 <400> 311
 caccactctg aagtggcct ct 22
 5 <210> 312
 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 312
 10 atggctctgc acattgttc c 21
 <210> 313
 <211> 22
 <212> ADN
 <213> Homo sapiens
 15 <400> 313
 cagagtggga aaaggcactt ca 22
 <210> 314
 <211> 22
 <212> ADN
 20 <213> Homo sapiens
 <400> 314
 ccagagtctt gtgcagacat tc 22
 <210> 315
 <211> 22
 25 <212> ADN
 <213> Homo sapiens
 <400> 315
 atggggatta actgggatgt tg 22
 <210> 316
 30 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 316
 cgtagctcca gacatcacta gca 23
 35 <210> 317
 <211> 22
 <212> ADN

<213> Homo sapiens
 <400> 317
 gcaacctggt ctgcaaagtc tc 22
 <210> 318
 5 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 318
 acccagcagt ccagcatgag 20
 10 <210> 319
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 319
 15 gagtttcagt gagccaagat cg 22
 <210> 320
 <211> 22
 <212> ADN
 <213> Homo sapiens
 20 <400> 320
 ttacaggctt gagccactag gc 22
 <210> 321
 <211> 21
 <212> ADN
 25 <213> Homo sapiens
 <400> 321
 aagcttcag gagacgaggt c 21
 <210> 322
 <211> 21
 30 <212> ADN
 <213> Homo sapiens
 <400> 322
 gtccctgaaa tccctcaaac c 21
 <210> 323
 35 <211> 23
 <212> ADN
 <213> Homo sapiens

<400> 323
 tgctccataa acgtgactat tgc 23
 <210> 324
 <211> 22
 5 <212> ADN
 <213> Homo sapiens
 <400> 324
 gtaagagggt gggctggaat ct 22
 <210> 325
 10 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 325
 cttagaccat gtccgggaaa ac 22
 15 <210> 326
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 326
 20 cacatcactc tgggtgggtga ac 22
 <210> 327
 <211> 22
 <212> ADN
 <213> Homo sapiens
 25 <400> 327
 aaatttcac ccaaaaccaa cc 22
 <210> 328
 <211> 23
 <212> ADN
 30 <213> Homo sapiens
 <400> 328
 ccagtcccaa gttcttgatc att 23
 <210> 329
 <211> 21
 35 <212> ADN
 <213> Homo sapiens
 <400> 329

acaagtcggc tagttgcatg g 21
 <210> 330
 <211> 22
 <212> ADN
 5 <213> Homo sapiens
 <400> 330
 tctcagatga aaccaccagc ac 22
 <210> 331
 <211> 23
 10 <212> ADN
 <213> Homo sapiens
 <400> 331
 ttcatctgag aagcaaggag tgg 23
 <210> 332
 15 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 332
 ccaggagagaa agcaggactc ta 22
 20 <210> 333
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 333
 25 ttctggcggg gtttgaaat ta 22
 <210> 334
 <211> 22
 <212> ADN
 <213> Homo sapiens
 30 <400> 334
 ctcaacattg acggccttgc tt 22
 <210> 335
 <211> 23
 <212> ADN
 35 <213> Homo sapiens
 <400> 335
 tcagctctta aacaggcat agc 23

<210> 336
 <211> 22
 <212> ADN
 <213> Homo sapiens
 5 <400> 336
 gaaatgcagc agccactaaa ga 22
 <210> 337
 <211> 22
 <212> ADN
 10 <213> Homo sapiens
 <400> 337
 ctcaccttca aggacctggt gt 22
 <210> 338
 <211> 21
 15 <212> ADN
 <213> Homo sapiens
 <400> 338
 cagggagggg tagaaaccac a 21
 <210> 339
 20 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 339
 ggagaggtgg agaggcttca g 21
 25 <210> 340
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 340
 30 gagactcca ggacagacac ct 22
 <210> 341
 <211> 20
 <212> ADN
 <213> Homo sapiens
 35 <400> 341
 cactcgttcc tcacccttcc 20
 <210> 342

<211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 342
 5 aggactcaca cgctactctg gt 22
 <210> 343
 <211> 23
 <212> ADN
 <213> Homo sapiens
 10 <400> 343
 ggacaatgtg atgaagattg ctg 23
 <210> 344
 <211> 22
 <212> ADN
 15 <213> Homo sapiens
 <400> 344
 atagcaggat cccaaaagac ca 22
 <210> 345
 <211> 22
 20 <212> ADN
 <213> Homo sapiens
 <400> 345
 ggcttgggga cctgtatttg ta 22
 <210> 346
 25 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 346
 cagtggcctt tctgagcctt ac 22
 30 <210> 347
 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 347
 35 gcactctagc tccctctttt agc 23
 <210> 348
 <211> 24

<212> ADN
 <213> Homo sapiens
 <400> 348
 ttctacagta gagggcagac atgc 24
 5 <210> 349
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 349
 10 gccaccatag ctgcagaatt ag 22
 <210> 350
 <211> 22
 <212> ADN
 <213> Homo sapiens
 15 <400> 350
 cccaaggaca gatgtgatgc ta 22
 <210> 351
 <211> 22
 <212> ADN
 20 <213> Homo sapiens
 <400> 351
 gcctttgttc gagaggagtt gt 22
 <210> 352
 <211> 22
 25 <212> ADN
 <213> Homo sapiens
 <400> 352
 gttcacgctc tcaagcaggt ta 22
 <210> 353
 30 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 353
 attccacaag ctcttccat ga 22
 35 <210> 354
 <211> 20
 <212> ADN

<213> Homo sapiens
 <400> 354
 ctgccccaa gatgcctaag 20
 <210> 355
 5 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 355
 tgcttggtat ttgctcatca tgt 23
 10 <210> 356
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 356
 15 cccttagcta gccactgac aa 22
 <210> 357
 <211> 20
 <212> ADN
 <213> Homo sapiens
 20 <400> 357
 ctctctgggag tgggtgtccaa 20
 <210> 358
 <211> 22
 <212> ADN
 25 <213> Homo sapiens
 <400> 358
 cctgggcaac agacagagta ag 22
 <210> 359
 <211> 21
 30 <212> ADN
 <213> Homo sapiens
 <400> 359
 cttcacttcc ccatgcgtac c 21
 <210> 360
 35 <211> 21
 <212> ADN
 <213> Homo sapiens

<400> 360
 gggttcaca tgcctacagg a 21
 <210> 361
 <211> 22
 5 <212> ADN
 <213> Homo sapiens
 <400> 361
 aaaatctgtg actttggctt gg 22
 <210> 362
 10 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 362
 gggaggagac attcttggat ttg 23
 15 <210> 363
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 363
 20 gcagtcctga gaagaaaaca gc 22
 <210> 364
 <211> 22
 <212> ADN
 <213> Homo sapiens
 25 <400> 364
 cttcacatgc cccaaaatta ca 22
 <210> 365
 <211> 23
 <212> ADN
 30 <213> Homo sapiens
 <400> 365
 tgagccatgt atttcagagg tga 23
 <210> 366
 <211> 22
 35 <212> ADN
 <213> Homo sapiens
 <400> 366

tacatttcag caggtgcgtg tt 22
 <210> 367
 <211> 24
 <212> ADN
 5 <213> Homo sapiens
 <400> 367
 ttgcctactc ttagggata ttgc 24
 <210> 368
 <211> 21
 10 <212> ADN
 <213> Homo sapiens
 <400> 368
 atagggcatg tagcccagtg a 21
 <210> 369
 15 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 369
 gctctgctgt tggctctcac t 21
 20 <210> 370
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 370
 25 ttgcaaagca cacatcttct ga 22
 <210> 371
 <211> 22
 <212> ADN
 <213> Homo sapiens
 30 <400> 371
 tggcaatgtc aatgtcaagc at 22
 <210> 372
 <211> 22
 <212> ADN
 35 <213> Homo sapiens
 <400> 372
 gtatgttgcc ccactcaaca aa 22

<210> 373
 <211> 22
 <212> ADN
 <213> Homo sapiens
 5 <400> 373
 tgcatttcct agctgagact cc 22
 <210> 374
 <211> 22
 <212> ADN
 10 <213> Homo sapiens
 <400> 374
 tgccatctcg cacgtagtaa at 22
 <210> 375
 <211> 22
 15 <212> ADN
 <213> Homo sapiens
 <400> 375
 ctctcctgtg ctgagcagct tt 22
 <210> 376
 20 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 376
 tgtttccaat cactggcttt ca 22
 25 <210> 377
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 377
 30 gaaccatggg ctgtctctgg 20
 <210> 378
 <211> 22
 <212> ADN
 <213> Homo sapiens
 35 <400> 378
 atctgggata gcgaaggaga ca 22
 <210> 379

<211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 379
 5 attacaggcc acacgccatc 20
 <210> 380
 <211> 23
 <212> ADN
 <213> Homo sapiens
 10 <400> 380
 aaggcaagaa taagggagga aga 23
 <210> 381
 <211> 23
 <212> ADN
 15 <213> Homo sapiens
 <400> 381
 gctctcagga ctgcagaagt aca 23
 <210> 382
 <211> 22
 20 <212> ADN
 <213> Homo sapiens
 <400> 382
 gaggaaccaa tcccactcac ac 22
 <210> 383
 25 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 383
 tcactctttg ccttctgtct ctg 23
 30 <210> 384
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 384
 35 gcactgtgct ttgctttctc ag 22
 <210> 385
 <211> 24

<212> ADN
 <213> Homo sapiens
 <400> 385
 tgtctccttt atcgtaggtc tcca 24
 5 <210> 386
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 386
 10 caccacattt cctacagttc ca 22
 <210> 387
 <211> 22
 <212> ADN
 <213> Homo sapiens
 15 <400> 387
 cactgtgcac cagacagaca aa 22
 <210> 388
 <211> 24
 <212> ADN
 20 <213> Homo sapiens
 <400> 388
 tgtggttttc tgtatcagca gctt 24
 <210> 389
 <211> 22
 25 <212> ADN
 <213> Homo sapiens
 <400> 389
 cagggagtct gaaatcatca gg 22
 <210> 390
 30 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 390
 tcaagtatct agccccaaat cca 23
 35 <210> 391
 <211> 24
 <212> ADN

<213> Homo sapiens
 <400> 391
 ggcaatattg accatcattc attc 24
 <210> 392
 5 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 392
 aggccaggag taagacgcaa c 21
 10 <210> 393
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 393
 15 aagaacgtac gtgtggtgtt gg 22
 <210> 394
 <211> 22
 <212> ADN
 <213> Homo sapiens
 20 <400> 394
 cgctatactt gctccatgca ct 22
 <210> 395
 <211> 21
 <212> ADN
 25 <213> Homo sapiens
 <400> 395
 aggaaacagc ctctggtcct c 21
 <210> 396
 <211> 22
 30 <212> ADN
 <213> Homo sapiens
 <400> 396
 gtcaatgctc agacagggag at 22
 <210> 397
 35 <211> 22
 <212> ADN
 <213> Homo sapiens

<400> 397
 cccaggaagg caggtactgt ta 22
 <210> 398
 <211> 22
 5 <212> ADN
 <213> Homo sapiens
 <400> 398
 tttacaacc accaaggggtg tg 22
 <210> 399
 10 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 399
 tcgtgtgggt acctccagat ttt 23
 15 <210> 400
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 400
 20 aaattagcca gggagtggag gt 22
 <210> 401
 <211> 20
 <212> ADN
 <213> Homo sapiens
 25 <400> 401
 catgccatgc tatggctcac 20
 <210> 402
 <211> 22
 <212> ADN
 30 <213> Homo sapiens
 <400> 402
 aggctgagcg gaggcttaat tg 22
 <210> 403
 <211> 22
 35 <212> ADN
 <213> Homo sapiens
 <400> 403

atctcagcaa tccacaggag gt 22
 <210> 404
 <211> 22
 <212> ADN
 5 <213> Homo sapiens
 <400> 404
 attgcctca cgaacacatc at 22
 <210> 405
 <211> 23
 10 <212> ADN
 <213> Homo sapiens
 <400> 405
 tggaaagttg tctatggcac ctc 23
 <210> 406
 15 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 406
 atgggcagca aggacttact ct 22
 20 <210> 407
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 407
 25 caccccaata ttgtctgcct tc 22
 <210> 408
 <211> 22
 <212> ADN
 <213> Homo sapiens
 30 <400> 408
 ggctcgggaa catgtaatta gg 22
 <210> 409
 <211> 24
 <212> ADN
 35 <213> Homo sapiens
 <400> 409
 ccatcatgta tggcaaattc tctt 24

<210> 410
 <211> 23
 <212> ADN
 <213> Homo sapiens
 5 <400> 410
 tggcgtctcc tagtaaagat gct 23
 <210> 411
 <211> 21
 <212> ADN
 10 <213> Homo sapiens
 <400> 411
 gccagattgc tggtttcatt g 21
 <210> 412
 <211> 22
 15 <212> ADN
 <213> Homo sapiens
 <400> 412
 ggctaaaaca caaagcacca tt 22
 <210> 413
 20 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 413
 gggaagtcac ccacaaagac ct 22
 25 <210> 414
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 414
 30 ggtctgggtc acagctcctc 20
 <210> 415
 <211> 22
 <212> ADN
 <213> Homo sapiens
 35 <400> 415
 ttctctgcc aagatgtggt gt 22
 <210> 416

<211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 416
 5 tgcagatgct gcaatcatgt ta 22
 <210> 417
 <211> 22
 <212> ADN
 <213> Homo sapiens
 10 <400> 417
 tggaccccga aagataaata gg 22
 <210> 418
 <211> 22
 <212> ADN
 15 <213> Homo sapiens
 <400> 418
 ttctgcactc ctctggaac tg 22
 <210> 419
 <211> 22
 20 <212> ADN
 <213> Homo sapiens
 <400> 419
 gggtagagc caaactgat ct 22
 <210> 420
 25 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 420
 ctgtgccctc tcatctcaca ct 22
 30 <210> 421
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 421
 35 agaacctagc ctccaagatt gc 22
 <210> 422
 <211> 22

<212> ADN

<213> Homo sapiens

<400> 422

acacctcca agactccttc ca 22

5 <210> 423

<211> 20

<212> ADN

<213> Homo sapiens

<400> 423

10 gactcgaggg tgggagacag 20

<210> 424

<211> 22

<212> ADN

<213> Homo sapiens

15 <400> 424

gctgtcacta ggtgtcctga gc 22

<210> 425

<211> 3878

<212> ADN

20 <213> Homo sapiens

<400> 425

cccgggcgag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60

ccgagggcggc cggaggtccc agctagcccc ggcgggcgcc gccgcccaga ccggacgaca 120

ggccacctcg tcggcggtccg cccgagtcgc cgcctcgccg ccaacgccac aaccaccgcg 180

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300

gcccggcgag tcggggtctg 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

tagcagtcctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660

atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720

tggagagcat ccagtggcgg gacatagtca gcagtgaact tctcagcaac atgtcgatgg 780

acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccgacagt 900

gctccggggc ctgccgtggc aagtcccca gtgactgctg ccacaaccag tgtgtgacag 960

gctgcacagg ccccggggag agcgactgcc tggctcgccg caaattccga gacgaagcca 1020

cgtgcaagga cacctgcccc ccactcatgc tctacaacct caccacgtac cagatggatg 1080

tgaaccccca gggcaaatat agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140

atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tgggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380

```

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatcttcagg aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgaggggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggaggggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggcgggga ccagacaact 1980
gtatccagtg tgccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctgggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgtcc atcgccactg ggaatggagg ggcctcctc ttgctgctgg 2220
tgggtggcct ggggatcgcc ctcttcatgc gaaggcgcca catcgttcgg aagcgacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccgt cgctatcaag gaattaaag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgctgctgg 2580
gcatctgcct cactccacc gtgcagctca tcacgcagct catgccctc ggctgcctcc 2640
tggactatgt ccgggaacac aaagacaata ttggtccca gtacctgtc aactggtgtg 2700
tgcatagcgc aaagggcact aactacttgg aggacgctc cttggtgcac cgcgacctg 2760
cagccaggaa cgtactgggt gagaaaagaat accatgcaga aggaggcaaa gtgcctatca 2880
ccaaactgct ggtgcgga atttgaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940
agtggatggc gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000
ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccaccatat 3060
gtaccatcga tgtctacatg atcatggtca agtctggat gatagacgca gatagtgcgc 3120
caaagtccg tgagttgatc atcgaattct ccaaatggc ccgagacccc cagcgtacc 3180
ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240
gtgcctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300
cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360
gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420
ccatcaagga agacagcttc ttgcagcgat acagctcaga cccacaggc gccttgactg 3480
aggacagcat agacgacacc ttcctcccag tgctgaata cataaaccag tccgttccca 3540
aaagcccgcc tggtctctgt cagaatcctg tctatcaca tcagcctctg aaccccgcc 3600
ccagcagaga ccacactac caggacccc acgacactgc agtgggcaac ccgagtatc 3660
tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720
agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttcttccca 3780
aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840
gggtcgcgcc acaagcagtg gaatttattg gagcatga 3878

```

<210> 426

<211> 3878

<212> ADN

5 <213> Homo sapiens

<400> 426

```

cccggcgag cgcggccgca gcagcctccg cccccgcac ggtgtgagcg cccgacgagg 60
ccgagggcgg cggagtcccg agctagcccc ggcgcccgcc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctccg cccgagtcct cgctcgccg ccaacgccac aaccaccgag 180
cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgtgctg gctgcgctct 300
gcccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
agggtgctct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttcttct 480
taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagtcct atctaactat gatgcaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcac ggcccgctgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tgagagcat ccagtggcgg gacatagtca cagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840

```

```

gctgggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccagcagt 900
gctccggggcg ctgccgtggc aagtccccc gtgactgctg ccacaaccag tgtgtgtcag 960
gctgcacagg ccccccggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc cactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccgga gggcaaatat agcttttggt ccacctgcgt gaagaagtgt ccccgtaat 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggatg ttgtgaatth aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccgggtggc tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggttttt ctgattcagg ctggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtgcagc cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaagggcac aggccaggtc tgccatgcct tgtgtcccc cgagggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggatgggtgg ggccctcctc ttgctgtctg 2220
tggtggccct ggggatcggc ctcttcacgc gaaggcgcca catcgctcgg aagcgcacgc 2280
tgccggaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctga 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat ccagaagggt gagaaagtta 2460
aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580
gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640
tggaactatg ccgggaacac aaagacaata ttggctccca gtacctgtc aactggtgtg 2700
tgcatatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760
cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820
ccaaactgct ggggtcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880
agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940
gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000
ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccaccatat 3060
gtaccatcga tgtctacatg atcatggtca agtctggat gatagacgca gatagtccgc 3120
caaagtcccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180
ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240
gtgcctgat ggatgaagaa gacatggacg acgtgggtgga tgccgacgag tacctcatcc 3300
cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360
gtgcaaccag caacaattcc accgtggctt gcatgtagag aaatgggctg caaagctgtc 3420
ccatcaagga agacagcttc ttgcagcgat acagctcaga cccacaggc gccttgactg 3480
aggacagcat agacgacacc ttctcccgag tgctgaata cataaaccag tccgttccca 3540
aaaggcccg cgtgctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgcg 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 acaagcagtg gaatttattg gagcatga 3878

```

<210> 427

<211> 3863

<212> ADN

5 <213> Homo sapiens

<400> 427

```

cccggcgcgag cgcggcgcgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgagg 60
ccgagggcggc cggaggtccc agctagcccc ggcgggcgcc gccgcccaga ccggacgaca 120
ggcgccactcg tcggcgctcc cccgagtcgc cgctcgccg ccaacgccac aaccaccgag 180
cagcgccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccgcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

```

```

cgagttggg 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
tagcagtcct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcccggtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tgagagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccagcagc 900
gctccggggcg ctgccgtggc aagtcctcca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccccggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccgga gggcaaatac agcttttggt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac aggtgtttttg tctgtcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagagcgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tggtaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaatcac aggtgtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactgaaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaaacagc 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agccagggga ctgctctctc tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgcc 1920
accagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggatgggtgg ggccctctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtgg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatctc gatgaagcct 2520
acgtgatggc cagcgtggac aacccccacg tgtgcgcct gctgggcatc tgcctcacct 2580
ccaccgtgca gctcatcacg cagctcatgc cctcggctg cctcctggac tatgtccggg 2640
aacacaaaga caatattggc tcccagtagc 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 ataccacc agagtgatgt ctggagctac ggggtgactg 2940
tttgggagtt gatgacctt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000
cctccatcct ggagaaagga gaacgcctcc ctacgccacc catatgtacc atcgatgtct 3060
acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgccaaaag ttccgtgagt 3120
tgatcatcga attctccaaa atggcccgag acccccagc ctacctgtc attcaggggg 3180
atgaaagaat gcatattgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240
aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300
tcagcagccc ctccacgtca cggactcccc tctgagctc tctgagtgc accagcaaca 3360
attccaccgt ggcttgcat gatagaaatg ggctgcaaa ctgtcccatc aaggaagaca 3420
gcttcttgca gcgatacagc tcagacccca caggcgctt gactgaggac agcatagacg 3480
acaccttctt cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgtggct 3540
ctgtgcagaa tctgtctat cacaatcagc ctctgaaccc cgcgccagc agagaccac 3600
actaccagga cccccacagc actgcagtgg gcaaccccg gtatctcaac actgtccagc 3660
ccacctgtgt caacagcaca ttcgacagcc ttgcccactg ggccagaaa ggcagccacc 3720
aaattagcct ggacaaccct gactaccagc aggaacttct tcccaaggaa gccaaagccaa 3780
atggcatctt taagggtcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840
gcagtgaatt tattggagca tga 3863

```

<210> 428

<211> 3863

<212> ADN

5 <213> Homo sapiens

<400> 428

```

ccccggcgag cggcgccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggcgcc cggagtcgcc agctagcccc ggccggccgcc gccgcccaga cccgacgaca 120
ggccacccctcg tcggcgctccg cccgagtcgc cgcctcgccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagttattgat cgggagagcc ggagcgagct ctcggggag 240
cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccggcgag tcgggctctg gaggaagaaga aagtttgcca aggcacgagt aacaagctca 360
cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtcggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat cttccttct 480
taaagaccat ccaggaggtg gctggttatg tectcattgc cctcaacaca gtggagcgaa 540
ttccttttga aaacctgcag atcatcagag gaaatatgta ctacgaaat tcctatgcct 600
tagcagtcct atctaactat gatgcaataa aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcccgctgc ggttcagcaa caacctgcc ctgtgcaacg 720
tgagagagcat ccagtgccgg gcacatagtc gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgcaaaa agtgatgatc aagctgtccc aatgggagct 840
gctgggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccgagct 900
gctccggcgcg aatccgtggc agtgcctgcc gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcgactgcc tggctgccc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccgga gggcaataac agctttgggt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtcggcgcg aatccagcag tctgtcgctc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaatagggtat tgggtaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactc cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggctag tttctctctg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgtg 1620
gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgttttg gacctccggg cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggcaggttc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agcccgaggga ctgctctctc tgccggaatg tcagccgagg caggggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
acccagagtg cctgctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctgggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtccctcc gatcccgctc atcgccactg ctgggccaag tcttgaaggc tgtccaacga 2160
atgggcctaa gateccgctc atcgccactg ggatgggtgg ggccctcctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcattg gaaggcgcca catcggttcgg aagcgcacgc 2280
tgccgagggt gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccgggtc gttccggcac gtgtataagg gactctggat ccagaaaggt gagaaagtta 2460
aaattcccg cgtatcaaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520
acgtgatggc cagcgtggac aacccccacg tgtgccgctg gctgggcatc tgccctacct 2580
ccaccgtgca gctcatcac cagctcatgc cctcggctg cctcctggac tatgtccggg 2640
aacacaaaga caatatggc tcccagttac gtgtcaactg gtgtgtgcag atcgcaagg 2700
gcatgaaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760
tggtgaaaac accgcagcat gtcaagatca cagatttttg gctggccaaa ctgctgggtg 2820
cggaagagaa agaataccat gcagaaggag gcaaaagtgc tatcaagtgg atggcattgg 2880
aatcaatttt acacagaatc tatacccacc agagtgtatg ctggagctac ggggtgactg 2940
tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000
cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060
acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccagg ttccgtgagt 3120
tgatcatcga attctccaaa atggcccgag acccccagcg ctacctgtc attcaggggg 3180
atgaagaagt gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240
aagaagacat ggacgagctg gtggatgccg acgagtacct catcccacag cagggttct 3300
tcagcagccc ctccacgtca cggactcccc tctgagctc tctgagtga accagcaaca 3360
attccaccgt ggcttgcat gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420
gcttcttgca gcgatacagc tcagacccca caggcgctt gactgaggac agcatagacg 3480
acaccttct cccagtgcct gaatacataa accagtcctg tcccaaaagg cccgtggt 3540
ctgtgcagaa tcctgtctat cacaatcagc ctctgaacc cgcgccagc agagaccac 3600
actaccagga cccccacagc actgcagtgg gcaaccccca gtatctcac actgtccagc 3660
ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720
aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaaagcaa 3780

atggcatctt taagggtcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840
gcagtgaatt tattggagca tga 3863

```

<210> 429

<211> 3863

5 <212> ADN

<213> Homo sapiens

<400> 429

ccccggcgag	cgcgggccgca	gcagcctccg	ccccccgcac	ggtgtgagcg	cccgcgcggg	60
ccgagggcgcc	cggagtcccc	agctagcccc	ggcgggccgc	gcccggccaga	ccggacgcga	120
ggccacctcg	tcggcggtccg	cccaggtccc	cgctcgcccg	ccaacgccac	aaccaccgcg	180
cacggccccc	tgactccgtc	cagtattgat	cgggagagcc	ggagcgagct	cttcggggag	240
cagcgatgag	accctccggg	acggccgggg	cagcgctcct	ggcgctgctg	gctgcgctct	300
gccccggcgag	tcggggtctg	gaggaaaaga	aaagtttgcca	aggcacgagt	aacaagctca	360
cgcagttggg	cacttttgaa	gatcattttc	tcagcctcca	gaggatgttc	aataactgtg	420
aggtgggtcct	tggaattttg	gaaattacct	atgtgcagag	gaattatgat	ctttccttct	480
taaagaccat	ccaggaggtg	gctggttatg	tcctcattgc	cctcaacaca	gtggagcgaa	540
ttcctttgga	aaacctgcag	atcatcagag	gaaatatgta	ctacgaaaat	tcctatgcct	600
tagcagtcct	atctaactat	gatgcaataa	aaaccggact	gaaggagctg	cccatgagaa	660
atttacagga	aatcctgcat	ggcgccgtgc	ggttcagcaa	caaccctgcc	ctgtgcaacg	720
tgagagagcat	ccagtggcgg	gacatagtca	gcagtgactt	tctcagcaac	atgtcgatgg	780
acttccagaa	ccacctgggc	agctgccaaa	agtgatgatcc	aaagctgtccc	aatgggagct	840
gctgggggtgc	aggagaggag	aactgccaga	aactgaccaa	aatcatctgt	gcccagcagt	900
gctccggggcg	ctgccgtggc	aaagtcctcc	gtgactgctg	ccacaaccag	tgtgtctgcag	960
gctgcacagg	ccccggggag	agcgactgcc	tggtctgccc	caaattccga	gacgaagcca	1020
cgtgcaagga	cacctgcccc	ccactcatgc	tctacaaccc	caccacgtac	cagatggatg	1080
tgaaccccgga	gggcaataac	agctttgggtg	ccacctgcgt	gaagaagtgt	ccccgtaatt	1140
atgtggtgac	agatcacggc	tcgtgcgtcc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaagacgg	cgtccgcgaag	tgtaagaagt	gcgaagggcc	ttgccgcaaa	gtgtgtaacg	1260
gaatagggtat	tggtgaattt	aaagactcac	tctccataaa	tgctacgaat	attaaacact	1320
tcaaaaactg	cacctccatc	agtggcgatc	tccacatcct	gccgggtggca	tttaggggtg	1380
actccttcac	acatactcct	cctctggatc	cacaggaact	ggatattctg	aaaaccgtaa	1440
aggaaatcac	agggtttttg	ctgattcagg	cttggcctga	aaacaggacg	gacctccatg	1500
cctttgagaa	cctagaaatc	atacgcgcca	ggaccaagca	acatggtcag	ttttctcttg	1560
cagtcgtcag	cctgaacata	acatccttgg	gattacgctc	cctcaaggag	ataagtgtatg	1620
gagatgtgat	aatttcagga	aaacaaaatt	tggtctatgc	aaatacaata	aactggaaaa	1680
aactgtttgg	gacctccggt	cagaaaacca	aaattataag	caacagaggt	gaaaaacagct	1740
gcaagggccac	aggccaggctc	tgccatgcct	tggtctcccc	cgaggggctgc	tgggggcccg	1800
agcccaggga	ctgcgtctct	tgccgggaatg	tcagccgagg	cagggaatgc	gtggacaagt	1860
gcaaccttct	ggaggggtgag	ccaagggagt	ttgtggagaa	ctctgagtgc	atacagtgcc	1920
accagagtg	cctgcctcag	gccatgaaca	tcacctgcac	aggacgggga	ccagacaact	1980
gtatccagtg	tgcccactac	attgacggcc	cccactgcgt	caagacctgc	ccggcaggag	2040
tcattgggaga	aaacaacacc	ctgggtctgga	agtagcgaga	cgccggccat	gtgtgccacc	2100
tgtgccatcc	aaactgcacc	tacggatgca	ctgggcccag	tcttgaaggc	tgcccaacga	2160
atgggacctaa	gatcccgctc	atcgccactg	ggatgggtgg	ggccctcctc	ttgctgctgg	2220
tggtggccct	ggggatcggc	ctcttcctgc	gaaggcgcca	catcggtcgg	aagcgacgc	2280
tgccgagggt	gctgcaggag	agggagcttg	tggaagcctc	tacacccagt	ggagaagctc	2340
ccaaccaagc	tctcttgagg	atcttgaagg	aaactgaatt	caaaaagatc	aaagtgtctg	2400
gctccggtgc	gttcggcacg	gtgtataagg	gactctggat	cccagaaggt	gagaaagtta	2460
aaattcccgt	cgtctatcaa	acatctccga	aagccaacaa	ggaaatcctc	gatgaagcct	2520
acgtgatggc	cagcgtggac	aacccccacg	tggtccgcct	gctgggcatc	tgccctcacct	2580
ccaccgtgca	gctcatcacg	cagctcatgc	ccttcgggtg	cctcctggac	tatgtccggg	2640
aacacaaaga	caatatgtgc	tcacagtacc	tgctcaactg	gtgtgtgcag	atcgcaaagg	2700
gcattgaacta	cttggaggag	cgtcgcttgg	tgaccgcgca	cctggcagcc	aggaacgtac	2760
tggtgaaaac	accgcagcat	gtcaagatca	cagattttgg	gctggccaaa	ctgctgggtg	2820
cggaagagaa	agaataccat	gcagaaggag	gcaaagtgcc	tatcaagtgg	atggcatttg	2880
aatcaatttt	acacagaatc	tataccacc	agagtgtatg	ctggagctac	ggggtgactg	2940
tttgggagtt	gatgaccttt	ggatccaagc	catatgaccc	aatccctgcc	agcgagatct	3000
cctccatcct	ggagaaagga	gaacgcctcc	ctcagccacc	catatgtacc	atcgatgtct	3060
acatgatcat	ggtcaagtgc	tggtgatag	acgcagatag	tcgcccacaa	ttccgtgagt	3120
tgatcatoga	attctccaaa	atggcccag	acccccagcg	ctacctgttc	attcaggggg	3180
atgaaagaat	gcatttgcca	agtcctacag	actccaactt	ctaccgtgcc	ctgatggatg	3240
aagaagacat	ggacgacgtg	gtggatgccg	acgagtacct	catcccacag	cagggtcttct	3300
tcagcagccc	ctccacgtca	cggactcccc	tcttgagctc	tctgagtgc	accagcaaca	3360
attccaccgt	ggcttgcat	gatagaaatg	ggctgcaaa	ctgtcccac	aagggaagaca	3420
gcttcttgca	gcgatacagc	tcagacccca	caggcgctct	gactgaggac	agcatagacg	3480
acaccttctc	cccagtgctc	gaatacataa	accagtcctg	tcccaaaaagg	cccgtgtggc	3540
ctgtgcagaa	tcctgtctat	cacaatcagc	ctctgaaccc	cgcgcccagc	agagaccac	3600
actaccagga	ccccacagc	actgcagtgg	gcaaccccga	gtatctcaac	actgtccagc	3660
ccacctgtgt	caacagcaca	ttcgacagcc	ctgcccactg	ggcccagaaa	ggcagccacc	3720
aaattagcct	ggacaaccct	gactaccagc	aggacttctt	tcccaaggaa	gccaaagccaa	3780
atggcatctt	taagggtctc	acagctgaaa	atgcagaata	cctaagggtc	gcgccacaaa	3840
gcagtgaatt	tattggagca	tga				3863

<210> 430

5 <211> 3863

<212> ADN

<213> Homo sapiens

<400> 430

```

ccccggcgag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggcggc cggaggtcccg agctagcccc ggcgcccgcc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctccg cccgagtcgc cgctcgccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccggcag 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
tagcagtcct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcac ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tggagagcat ccagtggcgg gacatagtca gcagtgaact tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccacgagct 900
gctccggggc ctgccgtggc aagtccccc gtgactgctg ccacaaccag tgtgtgcag 960
gctgcacagg ccccggggag agcgactgac tggctcgccg caaattccga gacgaagcca 1020
cgtgcaagga cactgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccca gggcaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaa gtgtgtaacg 1260
gaataggtat tgggaattt aaagactcac tctcataaaa tgctacgaat attaaacact 1320
tcaaaaactg cactccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
ccttttgaaa cctagaaata atacgcggca ggaccaagca acatggctag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaaacagt 1740
gcaagggcac agggccaggtc tgccatgcct tgtgtcctcc caggggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggaggttgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcattgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccag tcttgaaggc tgtccaacga 2160
atggccctaa gatcccgctc atcgccactg ggtggtggg ggccctcctc ttgctgctgg 2220
tggtggccct ggggatcgcc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctcctgggt gcctggcagc gtgtataagg gactctggat ccagaaaggt gaaaaagtta 2460
aaattcccg cgtatcaaaa acatctcaga aagccaacaa ggaaatcctc gatgaagcct 2520
acgtgatggc cagcgtggac aacccccacg tgtgccgct gctgggcatc tgcctcacct 2580
ccaccgtgca gctcatcacg cagctcatgc cctcgggtg cctcctggac tatgtccggg 2640
aacacaaaga caattattgg tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700
gcattgaacta cttggaggag cgtcgcttgg tcaccgcga cctggcagcc aggaacgtac 2760
tggtgaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820

cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880
aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940
tttgggagtt gatgacctt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000
cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc atcgatgtct 3060
acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccacaa ttccgtgagt 3120
tgatcatcga attctccaaa atggcccagc acccccagcg ctacctgttc attcaggggg 3180
atgaaaagaa gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240
aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300
tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgc accagcaaca 3360
attccaccgt ggcttgcat gatagaaatg ggctgcaaag ctgtcccac aaggaagaca 3420
gcttcttgca gcgatacagc tcagaccccc caggcgctt gactgaggac agcatagacg 3480
acaccttct cccactgcat gaatacataa accagtccgt tcccacaaag cccgctggct 3540
ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagaccac 3600
actaccagga ccccccacgc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660
ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720
aaattagcct ggacaacct gactaccagc aggaacttct tcccagga gccaagccaa 3780
atggcatctt taagggtccc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840
gcagtgaatt tattggagca tga 3863

```

<210> 431

<211> 3863

<212> ADN

<213> Homo sapiens

5 <400> 431

```

ccccggcgag cgcgggccga gcagcctccg cccccgcac ggtgtgagcg cccgacgagg 60
ccgaggcggc cgaggtcccc agctagcccc ggcgggccgc gccgcccaga ccggacgaca 120
ggccacctcg tcggcggtccg cccgagtcgc cgctcgccg ccaacgccac aaccaccgag 180
cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccgcgag tcggggtctg gaggaaaaga aagtttgcca aggcacgagt aacaagtca 360
cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggagggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagtcctt atctaactat gatgcaataa 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 gccagcagct 900
gctccggggc ctgccgtggc aagtccccc gtgactgctg ccacaaccag tgtgtgcag 960
gctgcacagg cccccgggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccga gggcaaatat agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tggatgaattt aaagactcac tctccataaa tgctacgaat attaaacct 1320
tcaaaaactg cactccatc agtggcgatc tccacatcct gccggtggca tttagggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagg 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 ggaggggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggcggggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atgcacctg ggatgggtgg ggccctctc ttgctgctgg 2220
tgggtggcct ggggatcggc ctcttcatgc gaaggcgcca catcgctcgg aagcgacgc 2280
tgccggaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340

```

```

ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtgtg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520
acgtgatggc cagcgtggac aacccccacg tgtgcgcct gctgggcac tgcctcacct 2580
ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640
aacacaaaga caatattggc tcccagtagc tgcctcaactg gtgtgtgcag atcgcaaagg 2700
gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760
tggtgaaaac accgcagcat gtcaagatca cagatcttgg gctggccaaa ctgctgggtg 2820
cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880
aatcaatctt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940
tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000
cctccatcct ggagaaagga gaacgcctcc ctacgccacc catatgtacc atcgatgtct 3060
acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgccaaaag ttccgtgagt 3120
tgcacatcga attctccaaa atggcccgag acccccagcg ctaccttgtc attcaggggg 3180
atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240
aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggtcttct 3300
tcagcagccc ctccacgtca cggactcccc tcttgagctc tctgagtgc accagcaaca 3360
attccaccgt ggcttgcatc gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420
gcttcttgca gcgatacagc tcagacccca caggcgctt gactgaggac agcatagacg 3480
acaccttctt cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgtggct 3540
ctgtgcagaa tctgtcttat cacaatcagc ctctgaacct cgcgccagc agagaccac 3600
actaccagga cccccacagc actgcagtgg gcaaccccg gtatctcaac actgtccagc 3660
ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720
aaattagcct gcacaaccct gactaccagc aggaactctt tcccaaggaa gccaaagcaa 3780
atggcatctt taagggtctc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840
gcagtgaatt tattggagca tga 3863

```

<210> 432

<211> 3863

<212> ADN

5 <213> Homo sapiens

<400> 432

```

cccgccgcag cgcggccgca gcagcctccg cccccgcac ggtgtgagcg cccgacgcgg 60
ccgaggcggc cggagtcccc agctagcccc ggccgcccga gccgcccaga cccgacgaca 120
ggccacctcg tcggcgctccg cccgagctcc cgctcgccg ccaacgccac aaccaccgcg 180
cagggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccgcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgcagtgggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggagggt gctggttatg tctcatttgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgct 600
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggccgctgac ggttcagcaa caaccctgcc ctgtgcaacg 720
tgagagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtg aggagaggag aactgccaga aactgaccaa aatcatctgt gcccgagct 900
gctccggggc ctgcccgtggc aagtccccca gtgactgtg ccacaaccag tgtgtgcag 960
gctgcacagg cccccgggag agcgactgcc tggctcgccg caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaacct caccacgtac cagatggatg 1080
tgaaccccg gggcaaatat agctttgggt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtagaagt gcgaagggcc ttgccgcaa gtgtgtaacg 1260
gaataggtat tgggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccgggtgga tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggttttt ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtgcgcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttccagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccgtg cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagc cagggaatgc gtggacaagt 1860

```

```

gcaaccttct ggaggggtgag ccaaggggagt ttgtggagaa ctctgagtgc atacagtgcc 1920
acctcagagt cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcacctac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctgggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atggggcctaa gatcccgctc atcgccactg ggatgggtggg ggccctctctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcctgc gaaggcgcca catcggttcgg aagcgcacgc 2280
tgcggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccgggtgc gttcggcacg gtgtataagg gactctggat ccagaagggt gagaaagtta 2460
aaattcccgct cgctatcaaa acatctccga aagccaacaa ggaaatctctc gatgaagcct 2520
acgtgatggc cagcgtggac aacccccacg tgtccgcctc gctgggcac c tgccctcacct 2580
ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640
aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700
gcatgaacta cttggaggag cgctcgcttg tgcaccgca cctggcagcc aggaacgtac 2760
tggtgaaac accgcagcat gtcaagatca cagatttttg gctggccaaa ctgctgggtg 2820
cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880
aatcaatctt acacagaatc tataccacc agagtgtgt ctggagctac ggggtgactg 2940
tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000
cctccatcct ggagaaagga gaacgcctcc ctgaccacc catatgtacc atcgatgtct 3060
acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccagg ttccgtgagt 3120
tgatcatcga attctccaaa atggcccag agccccagcg ctacctgtgc attcaggggg 3180
atgaaagaat gcatttgcca agtctacag actccaact ctaccgtgcc ctgatggatg 3240
aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggtctct 3300
tcagcagccc ctccacgtca cggactcccc tctgagctc tctgagtgc accagcaaca 3360
attccaccgt ggcttgcat gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420
gcttcttgca gcgatacagc tcagacccca caggcgctt gactgaggac agcatagacg 3480
acaccttctc ccagtgcct gaatacataa accagtccgt tcccaaaagg cccgtggct 3540
ctgtgcagaa tctgtctat cacaatcagc ctctgaacc cgcgccacg agagaccac 3600
actaccagga cccccacagc actgcagtgg gcaacccga gtatctcaac actgtccagc 3660
ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggagccacc 3720
aaattagcct ggacaacctc gactaccagc aggaacttct tcccaaggaa gccaaagcaa 3780
atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840
gcagtgaatt tattggagca tga 3863

```

<210> 433

<211> 3863

<212> ADN

5 <213> Homo sapiens

<400> 433

```

cccgccgag cgccggcgca gcagcctccg cccccgcac ggtgtgagcg cccgacgcg 60
ccgagcgggc cggagtcccc agctagcccc ggccggcgcc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctccg cccgagctcc cgccctcgcc ccaacgccac aaccaccgag 180
cacggccccc tgaactcgtc cagtatgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgtctct ggcgctgctg gctgcgtctc 300
gccggcgag tcgggctctg gaggaaga aagtttgcca aggcacgagt aacaagctca 360
cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtgtctct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 600
atttacagga aatcctgcat ggcccgctgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tgagagcat ccagtggcgg gacatagtca gcagtgaact tctcagcaac atgtcgatgg 780
acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtg aggagaggag aactgccaga aactgaccaa aatcatctgt gccagcagct 900
gctccgggag cgtccgtggc aagtcctcca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcgactgcc tggctgccc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccaactcatg tctacaacc caccacgtac cagatggatg 1080
tgaacccgga gggcaaatat agctttgtg ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgcccgaag gtgtgtaacg 1260
gaataggtat tggtgaaatt 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
ccttttgagaa	cctagaaatc	atacgcgga	ggaccaagca	acatggtcag	ttttctcttg	1560
cagtcgtcag	cctgaacata	acatccttgg	gattacgctc	cctcaaggag	ataagtgatg	1620
gagatgtgat	aatcttcagg	aacaaaaatt	tgtgctatgc	aaatacaata	aactggaaaa	1680
aactgtttgg	gacctccggt	cagaaaacca	aaattataag	caacagaggt	gaaaacagct	1740
gcaaggccac	aggccaggtc	tgccatgcct	tgtgctcccc	cgagggctgc	tggggcccg	1800
agcccaggga	ctgcgtctct	tgccggaatg	tcagccgagg	cagggaatgc	gtggacaagt	1860
gcaaccttct	ggaggggtgag	ccaaggaggt	ttgtggagaa	ctctgagtgc	atacagtgcc	1920
acccagagtg	cctgcctcag	gccatgaaca	tcacctgcac	aggacgggga	ccagacaact	1980
gtatccagtg	tgcccactac	attgacggcc	cccactgcgt	caagacctgc	ccggcaggag	2040
tcattgggaga	aaacaacacc	ctgggtctgga	agtacgcaga	cgccggccat	gtgtgccacc	2100
tgtgccatcc	aaactgcacc	tacggatgca	ctgggccagg	tcttgaaggc	tgtccaacga	2160
atgggacctaa	gatcccgtcc	atcgccactg	ggatgggtgg	ggccctcctc	ttgctgctgg	2220
tggtggccct	ggggatcgcc	ctcttcatgc	gaaggcgcca	catcgttcgg	aagcgacgc	2280
tgcgagggt	gctgcaggag	agggagcttg	tgagcctct	tacaccaggt	ggagaagctc	2340
ccaaccaagc	tcctctgagg	atcttgaagg	aaactgaatt	caaaaagatc	aaagtgtctg	2400
gctccggtgc	gttcggcacg	gtgtataagg	gactctggat	cccagaaggt	gagaaagtta	2460
aaattcccggt	cgctatcaaa	acatctccga	aagccaacaa	ggaaatcctc	gatgaagcct	2520
acgtgatggc	cagcgtggac	aacccccacg	tgtgccgcct	gctgggcatc	tgctcacct	2580
ccaccgtgca	gctcatcacg	cagctcatgc	ccttcggctg	cctcctggac	tatgtccggg	2640
aacacaaaga	caatattggc	tcacagtagc	tgctcaactg	gtgtgtgcag	atcgcaagg	2700
gcataaacta	cttggaggag	cgctcgcttg	tgacccgcga	cctggcagcc	aggaacgtac	2760
tggtgaaaaac	accgcagcat	gtcaagatca	cagatcttgg	gctggccaaa	ctgctgggtg	2820
cggaagagaa	agaataccat	gcagaaggag	gcaaagtgcc	tatcaagtgg	atggcatttg	2880
aatcaatttt	acacagaatc	tatacccacc	agagtgtatg	ctggagctac	gggtgactg	2940
tttgggagtt	gatgaccttt	ggatccaagc	catatgacgg	aatccctgcc	agcgagatct	3000
cctccatcct	ggagaaagga	gaacgcctcc	ctcagccacc	catatgtacc	atcgatgtct	3060
acatgatcat	ggtcaagtgc	tggtgatag	acgcagatag	tcgcccagg	ttccgtgagt	3120
tgatcatoga	attctccaaa	atggcccag	acccccagcg	ctacctgtc	attcagggg	3180
atgaaagaat	gcatttgcca	agtcctacag	actccaactt	ctaccgtgcc	ctgatggatg	3240
aagaagacat	ggacgacgtg	gtggatgccg	acgagtacct	catcccacag	cagggttct	3300
tcagcagccc	ctccacgtca	cgactcccc	tcctgagctc	tctgagtga	accagcaaca	3360
attccaccgt	ggcttgcat	gatagaaatg	ggctgcaaag	ctgtcccatc	aaggagaca	3420
gcttcttgca	gcgatacagc	tcagacccca	caggcgctt	gactgaggac	agcatagacg	3480
acaccttct	ccagtgccct	gaatacataa	accagtccgt	tcccaaaagg	cccgtgggt	3540
ctgtgcagaa	tcctgtctat	cacaatcagc	ctctgaacct	cgcgcccagc	agagaccac	3600
actaccaga	ccccacagc	actgcagtgg	gcaaccccca	gtatctcaac	actgtccagc	3660
ccacctgtgt	caacagcaca	ttcgacagcc	ctgcccactg	ggcccagaaa	ggcagccacc	3720
aaattagcct	ggacaaccct	gactaccagc	aggacttctt	tcccaaggaa	gccaaagcaa	3780
atggcatctt	taagggtcc	acagctgaaa	atgcagaata	cctaagggtc	gcgccacaaa	3840
gcagtgaatt	tattggagca	tga				3863

<210> 434

<211> 3863

<212> ADN

5 <213> Homo sapiens

<400> 434

cccggcgag	cgcgccgca	gcagcctccg	ccccccgac	ggtgtgagcg	cccgcgcgg	60
ccgagggcg	cggagtcgg	agctagcccc	ggcgccgccc	gccgcccaga	ccggacgaca	120
ggccacctcg	tcggcgctcg	cccaggtccc	cgctcgccg	ccaacgccac	aaccaccg	180
cacggccccc	tgactccgtc	cagtattgat	cgggagagcc	ggagcgagct	cttcggggag	240
cagcgatgcg	accctccggg	acggccgggg	cagcgctcct	ggcgctgctg	gctgcgctct	300
gcccggcgag	tcgggctctg	gaggaaaaga	aagtgtgcca	aggcacgagt	aacaagctca	360
cgagtggtgg	cacttttgaa	gatcattttc	tcagcctcca	gaggatgttc	aataactgtg	420
aggtggctct	tgggaatttg	gaaattacct	atgtgcagag	gaattatgat	ctttccttct	480
taaagaccat	ccaggaggtg	gctggttatg	tcctcattgc	cctcaacaca	gtggagcgaa	540
ttcctttgga	aaacctgcag	atcatcagag	gaaatatgta	ctacgaaaat	tcctatgcct	600
tagcagtcct	atctaactat	gatgcaataa	aaaccggact	gaaggagctg	cccatgagaa	660
atttacagga	aatcctgcac	ggcgccgtgc	ggttcagcaa	caaccctgcc	ctgtgcaacg	720
tgagagagcat	ccagtggcgg	gacatagtca	cagtgacttc	tctcagcaac	atgtcgatgg	780
acttcagaa	ccacctgggc	agctgcaaaa	agtgatgatc	aagctgtccc	aatgggagct	840
gctgggggtgc	aggagaggag	aactgccaga	aactgaccaa	aatcatctgt	gccagcagct	900

```

gctccgggag ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccccggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatgggat 1080
tgaaccccca gggcaaatat agcttttggt ccacctgctg gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tegtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tgggtgaatt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cactccatc agtggcgatc tccacatcct gccgggtggca ttttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactgaaaaa 1680
aactgtttgg gacctccggg cagaaaaacca aaattataag caacagaggt gaaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggtgctc tggggcccg 1800
agcccgaggc ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgctg caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctgggtctga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggatgggtgg ggccctctct ttgctgctgg 2220
tgggtggccct ggggatcgcc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtggt 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccgt cggtatcaag acatctccga aagccaacaa ggaaatctct gatgaagcct 2520
acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcac tgcctcacct 2580
ccaccgtgca gctcatcac cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640
aacacaaaga caatattggc tcccagtaac 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 ctacgccacc catatgtacc atcgatgtct 3060
acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgccccaaag ttccgtgagt 3120
tgatcatcga attctccaaa atggcccag agccccagcg ctaccttgct attcaggggg 3180
atgaaagaat gcattttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240
aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggtctct 3300
tcagcagccc ctccacgtca cggactcccc tctgagctc tctgagtgc accagcaaca 3360
attccaccgt ggcttgcatg gatagaaatg gggtgcaaaag ctgtcccatc aaggaagaca 3420
gcttcttgca gcgatacagg tcagaccca caggcgctct gactgaggac agcatagacg 3480
acaccttctc ccagtgctc gaatacataa accagtcctg tcccaaaaagg cccgtggct 3540
ctgtgcagaa tctctgtctat cacaatcagc ctctgaaccc cgggccagc agagaccac 3600
actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660
ccacctgtgt caacagcaca ttgcacagcc ctgcccactg ggcccagaaa ggcagccacc 3720
aaattagcct ggacaaccc gactaccagc aggaattctt tcccaaggaa gccaaagcaa 3780
atggcatctt taagggtctc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840
gcagtgaatt tattggagca tga 3863

```

<210> 435

<211> 3863

<212> ADN

5 <213> Homo sapiens

<400> 435

```

cccggcgacg cgcggccgca gcagcctccg cccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggcgg cggagtcccc agctagcccc ggcgcccgcc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctcc cccgagtccc cgcctcgccg ccaacgccac aaccaccgcg 180
cacggccccc tgaactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccggcgagc tcgggctctg gaggaagaag aagtttgcca aggcacgagt aacaagctca 360
cgcagtgggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420

```

```

aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggagggtg gctgggtatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcac ggcgccgtgc gggtcagcaa caaccctgcc ctgtgcaacg 720
tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccacgagct 900
gctccggggcg ctgccgtggc aagtccccc gtgactgctg ccacaaccag tgtgtctgcag 960
gctgcacagg cccccgggag agcgactgcc tgggtctgccg caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccaactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccgga gggcaaatc agcttttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaa gtgtgtaacg 1260
gaatagggtat tggggaattt aaagactcac tctcataaaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggctcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgttttg gacctccggt cagaaaacca aaattataag caacagaggt gaaaaacagct 1740
gcaaggccac aggccagggtc tgccatgcct tgtgctcccc cgagggtgctg tggggcccg 1800
agcccaggga ctgcgtctct tgccggaaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgc 1920
accagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctgggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggatgggtgg ggccctctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat ccagaaggt gagaaagtta 2460
aaattcccg tctctatcaag acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520
acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcac tgcctcacct 2580
ccaccgtgca gctcatcacg cagctcatgc ccttcgggtg cctcctggac tatgtccggg 2640
aacacaaaga caatattggc tccagtagc tgctcaactg gtgtgtgcag atcgcaaagg 2700
gcatgaacta ctggaggag cgtcgcttgg tgcaccgca 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 ctacgccacc catatgtacc atcgatgtct 3060
acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccagg ttccgtgagt 3120
tgatcatcga attctccaaa atggcccag agccccagcg ctaccttgtc attcaggggg 3180
atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240
aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggtcttct 3300
tcagcagccc ctccacgtca cggactcccc tctgagctc tctgagtgea accagcaaca 3360
attccaccgt ggcttgcatg gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420
gcttcttgca gcgatacagc tcagacccca caggcgctt gactgaggac agcatagacg 3480
acaccttctt cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540
ctgtgcagaa tcctgtctat cactgaaccc ctctgaaccc cgcccccagc agagaccac 3600
actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660
ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggagccacc 3720
aaattagcct ggacaacctt gactaccagc aggacttctt tcccaggaa gccaaacca 3780
atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840
gcagtgaatt tattggagca tga 3863

```

<210> 436

<211> 3863

<212> ADN

5 <213> Homo sapiens

<400> 436

```

ccccggcgag cgcggcgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60
ccgaggcgcc cggagtcctc agctagcccc ggcgcccgcc gccgcccaga cccgacgcaca 120
ggccacctcg tcggcgctcc cccgagtcct cgctcgcccg ccaacgccac aaccaccgcg 180
cacggcccc ctagctccgtc cagtagtgat cgggagagcc ggagcgagct ctcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gcccgcgag tcgggctctg gagaaaaa aaagtttgcca aggcacgagt aacaagctca 360
cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggctct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
attacagga aatcctgcat ggcccgctgc ggttcagcaa caacctgcc ctgtgcaacg 720
tgagagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtg agggagagg aactgccaga aactgaccaa aatcatctgt gccagcagt 900
gctccgggct ctagctgggc aagtcctcca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcgactgcc tggctgccc ccaattccga gacgaagcca 1020
cgtgcaaggg cactgcccc ccactcatgc tctacaacct caccacgtac cagatggatg 1080
tgaaccccc gggaacatac agctttggtg ccacctgctg gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggcgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tgggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cactccatc agtgcgcatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacggggca ggaccaagca acatggctag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggcacaggtg tgccatgcct tgtgctccc cgagggtgc tggggcccgg 1800
agccagggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggaggggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgc 1920
accagagatg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccaactc attgacggcc ccaactgctg caagacctgc ccggcaggag 2040
tcattgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atgcgccact ggatgggtgg ggcctcctc ttgtgctg 2220
tggtggccct ggggatccgc ctctcatgc gaaggcgcca catcgttcgg aagcgacgc 2280
tgccggaggt cctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccggctg gttcggcagc gtgtataagg gactctggat cccagaaggt gaaaaagtta 2460
aaattccctg cgtatcaag acatctccga aagccaacaa ggaatcctc gatgaagcct 2520
acgtgatggc cagcgtggac aacccccacg tgtccgcct gctgggcatc tgccctacct 2580
ccaccgtgca gctcatcagc cagctcatgc cctcggtg cctcctggac tatgtccggg 2640
aacacaaaga caatattggc tcccagtagc tgctcaactg gtgtgtgcag atcgcaagg 2700
gcatgaacta cttggaggag cgtcgcttgg tgaccgcga 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 gatgacctt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000
cctccatcct ggagaaagga gaacgcctcc ctacgccacc catatgtacc atcgatgtct 3060
acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccagg ttccgtgagt 3120
tgatcatcga attctccaaa atggcccag acccccagc ctacctgtc attcaggggg 3180
atgaagaat gcatttgcca agtcctacag actccaact ctaccgtgcc ctgatggatg 3240
aagaagacat ggacgacgtg gtggatgcc acgagtacct catcccacag cagggttct 3300
tcagcagccc ctccacgtca cggactcccc tctgagctg tctgagtga accagcaaca 3360
attccaccgt ggcttgcat gatagaaat ggctgcaaac ctgtccatc aaggaagaca 3420
gcttcttgca gcgatacagc tcagacccca caggcgctt gactgaggac agcatagacg 3480
acaccttct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540
ctgtgcagaa tctgtctcat cacaatcagc ctctgaaccc cgcgccagc agagaccac 3600
actaccagga ccccacagc actgcagtgg gcaaccccg gtatctcaac actgtccagc 3660
ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720
aaattagcct ggacaaccct gactaccagc aggacttctt tcccaggaa gccaaagcaa 3780
atggcatctt taagggtctc acagctgaaa atgcagaata cctaagggtc ggcacacaaa 3840

```

gcagtgaatt tattggagca tga

3863

<210> 437

<211> 3854

5 <212> ADN

<213> Homo sapiens

<400> 437

```

ccccggcgag cgcgggccgca gcagcctccg ccccccgac ggtgtgagcg cccgacgcgg 60
ccgagggcggc cggagtcccc agctagcccc ggcgcccgcc gccgcccaga cccgacgaca 120
ggccacctcg tcggcgctccg cccgagtccc cgctcgccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagtgatgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggctct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggagggt gctgggtatg tcctcattgc cctcaacaca gtggagcgaa 540
ttccttttga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaa ccacctgggc agctgcaaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtg aggagaggag aactgccaaa aactgaccaa aatcatctgt gccacgagt 900
gctccggggc ctgcgctggc aagtccccc gtgactgctg ccacaaccag tgtgtgcag 960
gctgcacagg cccccgggag agcgactgcc tggctgccc ccaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaacct caccacgtac cagatggatg 1080
tgaaccccg gggcaaatat agcttttggt ccactgctg gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tgggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccact agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc ccaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggttttt ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg cactccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccactac attgaccgce ccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggctctga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaagg tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggatgggtgg ggccctctc ttgctgctgg 2220
tgggtggcct ggggatcggc ctcttcctgc gaaggcgcca catcgctcgg aagcgcaacg 2280
tgcggaaggc gctgcaggag agggagcttg tggagcctc tacacccagt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaaag aaactgaatt caaaaagatc aaagtgtgg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccg cgtatcaag gaattaagag aagcaacact cgatgaagcc tacgtgatgg 2520
ccagcgtgga caacccccc gtgtgccgce tgctgggcat ctgcctcacc tccaccgtgc 2580
agctcatcac gcagctcatg cccttcggct gcctcctgga ctatgtccgg gaacacaaag 2640
acaatattgg ctcccagtac ctgctcaact ggtgtgtgca gatcgcaaag ggcatgaact 2700
acttgaggga ccgtcgcttg gtgcaccgcg acctggcagc caggaacgta ctggtgaaaa 2760
caccgcagca tgtcaagatc acagattttg ggctggccaa actgctgggt gcggaagaga 2820
aagaatacca tgcagaagga ggcaaaagt ctatcaagtg gatggcattg gaatcaattt 2880
tacacagaat ctataccac cagagtgatg tctggagcta cggggtgact gtttgggagt 2940
tgatgacctt tggatccaag ccatatgacg gaatccctgc cagcgagatc tcctccatcc 3000
tggagaaagg agaacgcct cctcagccac ccatatgtac catcgatgtc tacatgatca 3060
tggctcaagg ctgggatgta gacgcccaga gtcgcccata gttccgtgag ttgatcatcg 3120
aattctccaa aatggcccga gaccccagc gctaccttgt cattcagggg gatgaaagaa 3180
tgcatttgcc aagtcctaca gactccaact tctaccgtgc cctgatggat gaagaagaca 3240
tggacgacgt ggtggatgcc gacgagtacc tcatcccaca gcagggcttc ttcagcagcc 3300
cctccacgtc acggactccc ctctgagct ctctgagtg aaccagcaac aattccaccg 3360

tggcttgcat tgatagaaat gggctgcaaa gctgtcccat caaggaagac agcttcttgc 3420
agcgatacag ctcagacccc acaggcgctt tgactgagga cagcatagac gacaccttcc 3480
tcccagtgcc tgaatacata aaccagtcct tcccaaaaag gcccgctggc tctgtgcaga 3540
atcctgtcta tcacaatata cctctgaacc ccgcgccag cagagacca cactaccagg 3600
accccacag 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

5 <212> ADN

<213> Homo sapiens

<400> 438

```

ccccggcgag cgcgggcgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgagg 60
ccgagggggc cggagtcccc agctagcccc ggcgggcgcc gccgcccaga cgggacgaca 120
ggccacctcg tcggcgctcc cccgagtccc cgctcgcccg ccaacgccac aaccaccgag 180
cacggccccc tgaactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccgcgag tcgggctctg gagggaaaaga 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
ttccttttga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caacctgcc ctgtgcaacg 720
tggaagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccacgagc 900
gctccggggc ctgccgtggc aagtccccc gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcgactgcc tggctgccc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccg gggcaaatat agcttttggg ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tggatgaatt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactggttgg gacctccgtg cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccagggt tgccatgcct tgtgctcccc cgagggtgct tggggcccg 1800
agcccgagg ctcgctctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggcgggga ccagacaact 1980
gtatccagtg tgcccaactac attgacggcc ccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggtggtggg ggccctctc tttgtgctgg 2220
tggtggccct ggggatcggc ctcttcctgc gaaggcgcca catcgctcgg aagcgcacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacacccaag ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtggt 2400
gctccggtgc gtcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgt atggccagcg tggacaaccc ccactgtgct cgcctgctgg 2580
gcactgcctc cacctccacc gtgcagctca tcacgcagct catgccttcc ggctgcctcc 2640
tggaactatg ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700
tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760
cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttggggcgg 2820
ccaaactgct ggtgcggaag gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggtggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940
gctacggggt gactgttttg gagttgatga cctttggatc caagccatat gacggaatcc 3000
ctgcagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccaccatat 3060
gtaccatcga tgtctacatg atcatggtca agtgcgtgat gatagacgca gatagtcgcc 3120
caaagtccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180
ttgtcattca ggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240
gtgcctgat ggtgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300
cacagcagg cttcttcagc agcccctcca cgtcacggac tcccctctg agctctctga 3360
gtgcaaccag caacaattcc accgtggctt gcatgtagag aaatgggctg caaagctgtc 3420
ccatcaagga agacagcttc ttgcagcgat acagctcaga cccacaggc gccttgactg 3480
aggacagcat agacgacacc ttcctcccag tgctgaata cataaaccag tccgttccca 3540
aaaggcccg tggtctgtg cagaatcctg tctatcaca tcagcctctg aaccccgccg 3600
ccagcagaga cccacactac caggaccccc acgacactgc 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 acaagcagtg gaatttattg gagcatga 3878

```

<211> 3878

<212> ADN

<213> Homo sapiens

<400> 439

```

ccccggcgag cgcggcgcca gcagcctccg ccccccgcac ggtgtgagcg cccgacggcg 60
ccgagggcggc cggaggtccc agctagcccc ggcgggcgcc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctccg cccgagtcgc cgcctcgccg ccaacgccac aaccaccgag 180
cacggccccc tgactccgtc cagttattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gcccgcgag tcgggctctg gaggaaga 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 ctacgaaaaa tcctatgcct 600
tagcagtctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caacctgcc ctgtgcaacg 720
tgagagcagc ccagtgccg gacatagtc gacgtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgcaaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtg aggagaggag aactgccaga aactgaccaa aatcatctgt gccagcagt 900
gctccggggc ctgccgtggc aagtcctcca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaacct caccacgtac cagatggatg 1080
tgaaccccca gggcaaatat agcttttggt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaaagagg cgtccgcaag tgaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tgggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaatcac aggggtttttg 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
aactgttttg gacctccggt cagaaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggtgct tggggcccg 1800
agccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggcgggga ccagacaact 1980
gtatccagtg tgccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggatggtggg ggccctctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgacgc 2280
tgcgagggtg gctgcaggag agggagcttg tggagcctc tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400

```

```

gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgtatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgctcgtcgtg 2580
gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640
tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700
tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760
cagccaggaa cgtactgggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcggg 2820
ccaaactgct ggtgcggaag gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880
agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940
gtgccctgat ggatgaagaa gagttgatga cctttggatc caagccatat gacggaatcc 3000
ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccaccatat 3060
gtaccatcga tgtctacatg atcatggtca agtgcgtgat gatagacgca gatagtgcgc 3120
caaagtcccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgtacc 3180
ttgtcattca ggggatgaa agaattgcat tgccaagtcc tacagactcc aacttctacc 3240
gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300
cacagcaggg cttcttcacg agccctcca cgtcacggac tcccctcctg agctctctga 3360
gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420
ccatcaagga agacagcttc ttgcagcgat acagctcaga cccacagggc gccttgactg 3480
aggacagcat agacgacacc ttctctccca tgctgaata cataaaccag tccgttccca 3540
aaaggcccg tggtctgtg cagaatcctg tctatcaca tcagcctctg aaccccgcg 3600
ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660
tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720
agaaaggcag ccaccaaatt agcctggaca acctgacta ccagcaggac ttcttccca 3780
aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840
gggtcgcgccc acaaagcagt gaatttattg gagcatga 3878

```

<210> 440

<211> 3878

<212> ADN

5 <213> Homo sapiens

<400> 440

```

cccgccgcag cgcggccgcga gcagcctccg cccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggcggc cggagtcccc agctagcccc ggccgcccgc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctccg cccgagctcc cgctcgcgcg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatcgc accctccggg acggccgggg cagcgtcctt ggcgctgtgt gctcgcctct 300
gccccggcag tcgggctctg gaggaaga aaagtttcca aggcacgagt aacaagctca 360
cgcagtggg cacttttgaa gatcatttcc 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
tagcagctct atctaactat gatgcaaaa aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caacctgcc ctgtgcaacg 720
tgagagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccacgagct 900
gctccgggcg ctcctgtggc agtccccca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccgggag agcactgcc ttggtctgcg caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc cactcatgc tctacaacc caccacgtac cagatggatg 1080
tgaacccoga gggcaaatat agctttgtgt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaa gtgtgtaacg 1260
gaatagggtat tggatgaatt 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 gaaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggtgctc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920

```

accagagtg	cctgcctcag	gccatgaaca	tcacctgcac	aggacgggga	ccagacaact	1980
gtatccagt	tgcccactac	attgacggcc	cccactgcgt	caagacctgc	ccggcaggag	2040
tcattgggaga	aaacaacacc	ctgggtctgga	agtacgcaga	cgccggccat	gtgtgccacc	2100
tgtgccatcc	aaactgcacc	tacggatgca	ctgggccagg	tcttgaaggc	tgtccaacga	2160
atgggcctaa	gatcccgctc	atcgccactg	ggatgggtggg	ggccctcctc	ttgctgctgg	2220
tggtggccct	ggggatcggc	ctcttcatgc	gaaggcgcca	catcgttcgg	aagcgcacgc	2280
tgcggaggct	gctgcaggag	agggagcttg	tggagcctct	tacaccagct	ggagaagctc	2340
ccaaccaagc	tctcttgagg	atcttgaagg	aaactgaatt	caaaaagatc	aaagtgtctg	2400
gctccggtgc	gttcgggcagc	gtgtataagg	gactctggat	cccagaaggt	gagaaagtta	2460
aaattcccg	cgctatcaag	gaattaaag	aagcaacatc	tccgaaagcc	aacaaggaaa	2520
tcctcgatga	agcctacgtg	atggccagcg	tggacaaccc	ccacgtgtgc	cgctgctgg	2580
gcatctgcct	caactccacc	gtgcagctca	tcacgcagct	catgcccttc	ggctgcctcc	2640
tggactatgt	ccgggaacac	aaagacaata	ttggctccca	gtacctgtct	aactggtgtg	2700
tgcagatcgc	aaagggcatg	aactacttgg	aggaccgtcg	cttgggtgcac	cgcgacctgg	2760
cagccaggaa	cgtactggtg	aaaacaccgc	agcatgtcaa	gatcacagat	tttgggctgg	2820
ccaaactgct	gggtgcggaa	gagaaagaat	acatgcaga	aggaggcaaa	gtgcctatca	2880
agtggatggc	attggaatca	attttacaca	gaatctatac	ccaccagagt	gatgtctgga	2940
gctacggggt	gactgttttg	gagttgatga	cctttggatc	caagccatat	gacggaatcc	3000
ctgcacgcgc	gatctcctcc	atcctggaga	aaggagaacg	cctccctcag	ccaccatat	3060
gtaccatcga	tgtctacatg	atcatgggtca	agtgtggat	gatagacgca	gatagtgcgc	3120
caaagtctcg	tgagttgatc	atcgaattct	ccaaaatggc	ccgagacccc	cagcgctacc	3180
ttgtcattca	gggggatgaa	agaatgcatt	tgccaagtcc	tacagactcc	aacttctacc	3240
gtgcctgat	ggatgaagaa	gacatggacg	acgtgggtgga	tgccgacgag	tacctcatcc	3300
cacagcagg	cttcttcagc	agcccctcca	cgtcacggac	tcccctcctg	agctctctga	3360
gtgcaaccag	caacaattcc	accgtggctt	gcattgatag	aatgggctg	caaagctgtc	3420
ccatcaagga	agacagcttc	ttgcagcgat	acagctcaga	ccccacaggc	gccttgactg	3480
aggacagcat	agacgacacc	ttcctcccag	tgctgaata	cataaaccag	tccgttccca	3540
aaaggccgcg	tggtctctgt	cagaatcctg	tctatcacaa	tcagcctctg	aaccccgccg	3600
ccagcagaga	cccacactac	caggaccccc	acagcactgc	agtgggcaac	cccagtatc	3660
tcaacactgt	ccagcccacc	tgtgtcaaca	gcacattcga	cagccctgcc	cactgggccc	3720
agaaaggcag	ccaccaaatt	agcctggaca	accctgacta	ccagcaggac	ttctttccca	3780
aggaagccaa	gccaaatggc	atctttaagg	gctccacagc	tgaaaatgca	gaatacctaa	3840
gggtcgcgcc	acaagcagct	gaatttattg	gagcatga			3878

<210> 441

<211> 3869

<212> ADN

5 <213> Homo sapiens

<400> 441

cccggcgccag	cgcgggccgca	gcagcctccg	ccccccgcac	ggtgtgagcg	cccgcgcggg	60
ccgagggcgcc	cggagtcccc	agctagcccc	ggcgcccgcc	gccgcccaga	ccggacgaca	120
ggccacctcg	tcggcgctccg	cccaggtccc	cgctcgcccg	ccaacgccac	aaccaccgcg	180
cacggccccc	tgaactccgtc	cagtattgat	cgggagagcc	ggagcgagct	cttcggggag	240
cagcgatgcg	accctccggg	acggccgggg	cagcgctcct	ggcgctgctg	gctgcgctct	300
gccccggcgag	tcgggctctg	gaggaaaaga	aagtgtgcca	aggcacgagt	aacaagctca	360
cgcagttggg	cacttttgaa	gatcattttc	tcagcctcca	gaggatgttc	aataactgtg	420
aggtggtcct	tggaattttg	gaaattacct	atgtgcagag	gaattatgat	ctttccttct	480
taaagagcat	ccaggaggtg	gctggttatg	tcctcattgc	cctcaacaca	gtggagcgaa	540
ttcctttgga	aaacctgcag	atcatcagag	gaaatatgta	ctacgaaaat	tcctatgcct	600
tagcagtcct	atctaactat	gatgcaata	aaaccggact	gaaggagctg	cccatgagaa	660
atttacagga	aatcctgcac	ggcgccgtgc	ggttcagcaa	caaccctgcc	ctgtgcaacg	720
tgagagcagc	ccagtggcgg	gacatagtca	gcagtgaatt	tctcagcaac	atgtcgatgg	780
acttcagaa	ccacctgggc	agctgccaaa	agtgtgatcc	aagctgtccc	aatgggagct	840
gctgggggtgc	aggagaggag	aactgccaga	aactgaccaa	aatcatctgt	gccagcagct	900
gctccggggc	ctgccgtggc	aagtccccc	gtgactgctg	ccacaaccag	tgtgtgcag	960
gctgcacagg	ccccggggag	agcgactgcc	tggtctgccc	caaattccga	gacgaagcca	1020
cgtgcaagg	cacctgcccc	ccactcatgc	tctacaaccc	caccacgtac	cagatggatg	1080
tgaaccccga	gggcaaatat	agctttggtg	ccacctgcgt	gaagaagtgt	ccccgtaatt	1140
atgtggtgac	agatcacggc	tcgtgcgtcc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaagacgg	cgtccgcgaag	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

```

aggaatcac aggggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatgggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aattttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggtctg tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctgggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggatgggtgg ggcctcctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcctgc gaaggcgcca catcggtcgg aagcgacgc 2280
tgccgagggt ctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgcctg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat ccagaagggt gagaagtta 2460
aaattcccg tgcctatcaag gaaccaacat ctccgaaagc caacaaggaa atcctcgatg 2520
aagcctacgt gatggccagc gtggacaacc cccacgtgtg ccgctgctg ggcattctgc 2580
tcacctccac ctgcagctc atcacgcagc tcatgccctt cggtgcctc ctggactatg 2640
tccgggaaca caaagacaat attggctccc agtacctgct caactggtgt gtgcagatcg 2700
caaaggcgat gaactacttg gaggaccgtc gcttgggtgca ccgcgacctg gcagccagga 2760
acgtactggt gaaaacaccg cagcatgtca agatcacaga ttttgggctg gccaaactgc 2820
tggtgctcga agagaaagaa taccatgcag aaggaggcaa agtgctatc aagtggatgg 2880
cattggaatc aattttacac agaattctata cccaccagag tgatgtctgg agctacgggg 2940
tgactgtttg ggagttgatg acctttggat ccaagccata tgacggaatc cctgccagcg 3000
agatctcctc catcctggag aaaggagaac gcctccctca gccaccata tgtaccatcg 3060
atgtctacat gatcatggtc aagtgtctgga tgatagacgc agatagtcgc ccaaagtctc 3120
gtgagttgat catcgaattc tccaaaatgg cccgagacct cttgtcattc 3180
aggggggatga aagaatgcat ttgccaaagtc ctacagactc caacttctac cgtgccctga 3240
tggtatgaaga agacatggac gacgtggtgg atgcccagca gtacctcatc ccacagcagg 3300
gcttcttcag cagccccctc acgtcacgga ctccccctc gagctctctg agtgaacca 3360
gcaacaattc caccgtggct tgcatgtata gaaatgggct gcaaagctgt cccatcaagg 3420
aagacagctt ctgcagcgca tacagctcag accccacagg cgccttgact gaggacagca 3480
tagacgacac ctctctccca gtgcctgaat acataaacca gtccgttccc aaaaggcccc 3540
ctggctctgt gcagaatcct gtctatcaca atcagcctct gaaccccgcg cccagcagag 3600
acccacacta ccaggacccc cacagcactg cagtgggcaa ccccgagtat ctcaacactg 3660
tccagcccat ctgtgtcaac agcacattcg acagccctgc cactggggc cagaaaggca 3720
gccaccaaat tagcctggac aacctgact accagcagga cttctttccc aaggaaagca 3780
agccaaatgg catctttaag ggctccacag ctgaaaatgc agaataccta agggctcgcc 3840
caciaaagcag tgaatttatt ggagcatga 3869

```

<210> 442

<211> 3869

<212> ADN

5 <213> Homo sapiens

<400> 442

```

cccgccgcag cgcggccgca gcagcctccg ccccccgac ggtgtgagcg cccgacggcg 60
ccgagggcgg cggagtcccg agctagcccc ggcgcccgcc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctccg cccgagtcct cgctcgccg ccaacgccac aaccaccgg 180
cagggccccc tgactccgct cagttattgat cggagagacc ggagcgagct cttcggggag 240
cagcgatgag accctccggt acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccgcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgaggtggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttctctct 480
taaagaccat ccaggagggt gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgect 600
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgcctgct gggtcagcaa caacctgcc ctgtgcaacg 720
tgagagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtg aactgccaga aactgaccaa aatcatctgt gccacgagc 900
gctccggggc ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960

```

```

gctgcacagg cccccgggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccca gggcaaatat agcttttggtg ccacttgctg gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tegtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tggatgaatt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttagggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtgcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aattttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggg cagaaaaacca aaattataag caacagaggt gaaaaacagct 1740
gcaaggccac aggccagggt tgccatgcct tegtctcccc cgagggtctg tggggcccg 1800
agcccaggga ctgcgtctct tgcgggaatg tcacgcaggg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 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 ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggacctaa gatcccgctc atcgccactg ggatggtggg ggccctcctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcggtcgg aagcgcacgc 2280
tgccggaggt gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgctgg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccggt cgctatcaag gaaccaacat ctccgaaagc caacaaggaa atcctcgatg 2520
aagctacgtg gatggccagc gtggacaacc cccacgtgtg ccgcctgtg ggcatctgcc 2580
tcacctccac cgtgcagctc atcacgcagc tcatgccctt cggtgcctc ctggactatg 2640
tcccggaaaca caaagacaat attggtccc agtacctgct caactggtgt gtgcagatcg 2700
caaagggcat gaactacttg gaggaccgtc gcttgggtgca ccgcgacctg gcagccaggga 2760
acgtactggt gaaaacaccg cagcatgtca agatcacaga ttttgggctg gccaaactgc 2820
tggtgctcga agagaaagaa taccatgcag aaggaggcaa agtgcctatc aagtggatgg 2880
cattggaatc aattttacac agaattctata cccaccagag tgatgtctgg agctacgggg 2940
tgactgtttg ggagttgatg acctttggat ccaagccata tgacggaatc cctgccagcg 3000
agatctcctc catcctggag aaaggagaac gcctccctca gccaccata tgtaccatcg 3060
atgtctacat gatcatggtc aagtgtctgga tgatagacgc agatagtgc ccaaagtcc 3120
gtgagttgat catcgaattc tccaaaatgg cccgagaccc ccagcgctac cttgtcatte 3180
aggggggatga aagaatgcat ttgccaaagtc ctacagactc caacttctac cgtgccctga 3240
tggtatgaaga agacatggac gacgtggtgg atgccgaga gtacctcatc ccacagcagg 3300
gcttcttcag cagcccctcc acgtcacgga ctcccctcct gagctctctg agtgcaacca 3360
gcaacaattc caccgtggct tgcatgata gaaatgggct gcaaagctgt cccatcaagg 3420
aagacagctt cttgcagcga tacagctcag accccacagg cgccttgact gaggacagca 3480
tagacgacac ctctctccca gtgcctgaat acataaacca gtccgttccc aaaaggccc 3540
ctggctctgt gcagaatccc gtctatcaca atcagcctct gaaccccgcg cccagcagag 3600
acccacacta ccaggacccc cacagcactg cagtgggcaa ccccgagtat ctcaacactg 3660
tccagcccac ctgtgtcaac agcacattcg acagccctgc cactgggccc cagaaaggca 3720
gccaccaaata tagcctggac aaccctgact accagcagga cttctttccc aaggaaagcca 3780
agccaaatgg catcttttaag ggctccacag ctgaaaatgc agaataccta agggtcgcgc 3840
cacaagcag tgaatttatt ggagcatga 3869

```

<210> 443

<211> 3860

<212> ADN

5 <213> Homo sapiens

<400> 443

```

cccgccgcag cgcggccgca gcagcctccg cccccgcac ggtgtgagcg cccgacgcgg 60
ccgaggcggc cggagtcctc agctagcccc ggccgcccgc gccgcccaga cccgacgaca 120
ggccacctcg tcggcgctcc cccgagtcct cgctctcgcc ccaacgccac aaccaccg 180
cacggcccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgtct 300
gccccggcag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgcagttggg cacttttgaa gatcatttcc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480

```

taaagacccat	ccaggaggtg	gctgggttatg	tcctcattgc	cctcaacaca	gtggagcgaa	540
ttccttttga	aaacctgcag	atcatcagag	gaaatatgta	ctacgaaaat	tcctatgcct	600
tagcagtcct	atctaactat	gatgcaaata	aaaccggact	gaaggagctg	cccatgagaa	660
atttacagga	aatcctgcac	ggcgccgtgc	ggttcagcaa	caaccctgcc	ctgtgcaacg	720
tggagagcat	ccagtgggcg	gacatagtca	gcagtgactt	tctcagcaac	atgtcgtatg	780
acttccagaa	ccacctgggc	agctgccaaa	agtgatgatcc	aagctgtccc	aatgggagct	840
gctgggggtgc	aggagaggag	aactgccaga	aactgaccaa	aatcatctgt	gcccagcagt	900
gctccggggcg	ctgccgtggc	aagtccccc	gtgactgctg	ccacaaccag	tgtgctgcag	960
gctgcacagg	ccccggggag	agcgactgcc	tggctctgcc	caaattccga	gacgaagcca	1020
cgtgcaaggga	cacctgcccc	ccactcatgc	tctacaaccc	caccacgtac	cagatggatg	1080
tgaaccccgga	gggcaaatac	agcttttggg	ccacctgcgt	gaagaagtgt	ccccgtaat	1140
atgtggtgac	agatcacggc	tcgtgcgtcc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaagacgg	cgtccgcgaag	tgtgaagaag	gcgaagggcc	ttgccgcaaa	gtgtgtaacg	1260
gaataggtat	tggtgaattt	aaagactcac	tctccataaa	tgtacgaat	attaaacact	1320
tcaaaaactg	cacctccatc	agtgccgatc	tccacatcct	gccgggtggca	tttaggggtg	1380
actccttcac	acatactcct	cctctggatc	cacaggaact	ggatattctg	aaaaccgtaa	1440
aggaaatcac	agggtttttg	ctgattcagg	cttggcctga	aaacaggacg	gacctccatg	1500
ccttttgagaa	cctagaaatc	atacgccgga	ggaccaagca	acatggtcag	ttttctcttg	1560
cagtcgtcag	cctgaacata	acatccttgg	gattacgctc	cctcaaggag	ataagtgatg	1620
gagatgtgat	aattttcagga	aacaaaaatt	tgtgctatgc	aaatacaata	aactggaaaa	1680
aactgtttgg	gacctccggg	cagaaaacca	aaattataag	caacagaggt	gaaaacagct	1740
gcaaggccac	aggccaggtc	tgccatgcct	tgtgctcccc	cgaagggtgc	tggggcccg	1800
agcccaggga	ctgcgtctct	tgccggaatg	tcagccgagg	cagggaatgc	gtggacaagt	1860
gcaacccttct	gaagggtgag	ccaaggaggt	ttgtggagaa	ctctgagtgc	atacagtgcc	1920
accagagtg	cctgcctcag	gccatgaaca	tcacctgcac	aggacgggga	ccagacaact	1980
gtatccagt	tgcccactac	attgacggcc	cccactgcgt	caagacctgc	ccggcaggag	2040
tcattgggaga	aaacaacacc	ctgggtctga	agtacgcaga	cgccggccat	gtgtgccacc	2100
tgtgccatcc	aaactgcacc	tacggatgca	ctgggcccag	tcttgaaggc	tgtccaacga	2160
atgggcctaa	gatcccgtcc	atcgccactg	ggatgggtgg	ggccctcctc	ttgctgctgg	2220
tgggtggccct	ggggatcgcc	ctcttcatgc	gaaggcgcca	catcgttcgg	aagcgcacgc	2280
tgccgaggct	gctgcaggag	agggagcttg	tggagcctct	tacacccagt	ggagaagctc	2340
ccaaccaagc	tctcttgagg	atcttgaagg	aaactgaatt	caaaaagatc	aaagtgtctg	2400
gctccgggtg	gttcgggcac	gtgtataaag	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	tcaactgggt	tgtgcagatc	gcaaagggca	2700
tgaactactt	ggaggaccgt	cgcttgggtc	accgcgacct	ggcagccagg	aacgtactgg	2760
tgaaaacacc	gcagcatgtc	aagatcacag	atcttgggct	ggccaaactg	ctgggtgcgg	2820
aagagaaaga	ataccatgca	gaaggaggca	aagtgcctat	caagtggatg	gcattggaat	2880
caattttaca	cagaatctat	accaccaga	gtgatgtctg	gagctacggg	gtgactgttt	2940
gggagttgat	gacctttgga	tccaagccat	atgacggaat	ccctgccagc	gagatctcct	3000
ccatcctgga	gaaaggagaa	cgctccctc	agccaccat	atgtaccatc	gatgtctaca	3060
tgatcatggt	caagtgtctg	atgatagacg	cagatagtcg	cccaaagtcc	cgtgagttga	3120
tcacgaatt	ctccaaaatg	gcccgagacc	cccagcgcta	ccttgtcatt	caggggggatg	3180
aaagaatgca	tttgccaagt	cctacagact	ccaacttcta	cgtgcccctg	atggatgaag	3240
aagacatgga	cgacgtgggt	gatgccgacg	agtacctcat	cccacagcag	ggcttcttca	3300
gcagcccctc	cacgtcacgg	actcccctcc	tgagctctct	gagtgcaacc	agcaacaatt	3360
ccaccgtggc	ttgcattgat	agaaatgggc	tgcaaaagctg	tcccatcaag	gaagacagct	3420
tcttgacgcg	atcacagctca	gacccacacg	gcgccttgac	tgaggacagc	atagacgaca	3480
ccttctctcc	agtgcctgaa	tacataaacc	agtcctgtcc	caaaaggccc	gctggctctg	3540
tgcagaaatc	tgctatcac	aatcagcctc	tgaacccgca	gcccagcaga	gaccacact	3600
accaggaccc	ccacagcact	gcagtgggca	accccgagta	tctcaacact	gtccagccca	3660
cctgtgtcaa	cagcacattc	gacagccctg	ccactgggca	ccagaaaggc	agccaccaa	3720
ttagcctgga	caaccctgac	taccagcagg	acttctttcc	caagggaagc	aagccaaatg	3780
gcattcttaa	gggctccaca	gctgaaaatg	cagaatacct	aagggtcgcg	ccacaaagca	3840
gtgaatttat	tggagcatga					3860

<210> 444

<211> 3860

<212> ADN

5 <213> Homo sapiens

<400> 444

```

ccccggcgag cgcgggcgca gcagcctccg ccccccgca ggtgtgagcg cccgacgagg 60
ccgaggcggc cggagtcctcg agctagcccc ggcgggcgcc gccgcccaga cgggacgaca 120
ggccacctcg tggcgctccg cccgagtcct cgccctcgcc ccaacgccac aaccaccgag 180
cacggcccc ctagctccgtc cagtattgat cgggagagcc ggagcgagct ctccggggag 240
cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gcccgggcag tccgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggctct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggaggtg gctgggttat tcctcatgct cctcaacaca gtggagcgaa 540
ttccttttga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctct atctaactat gatgcaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcccgctgc ggttcagcaa caacctgcc ctgtgcaacg 720
tgagagagcat ccagtgggcg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaa ccacctgggc agctgccaaa agtgatgctc aagctgtccc aatgggagct 840
gctgggggtg agggagggag aactgccaga aatcatctgt gccagcagt 900
gctccgggag ctgcccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccgggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
cgtgcaagga cactgcccc ccactcatgc tctacaacc caccacgtac cagatggatg 1080
tgaacccga gggcaaatc agctttgggt ccacctgctt gaagaagtgt ccccgtaatt 1140
atgtgggtgac agatcacggc tctgctgctc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tctaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggat tggtagaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaatatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactgaaaaa 1680
aactgtttgg gacctccggt cagaaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac agccagagtc tgccatgcct tgtgctcccc cgagggtgct gggggcccg 1800
agccagggga ctgctctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgc 1920
accagagtg cctgcctcag gccatgaaca tcacctgcac aggcagggga ccagacaact 1980
gtatccagtg tggccactac attgacggcc cccactgctt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaagg tgtccaacga 2160
atgggctaa gatcccgctc atcgccactg ggatgggtgg ggccctctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccggagct gctgcaggag agggagcttg tggagcctct tacaccagtg ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aagtgctgg 2400
gctccgggtc gttcggcagc gtgtataagg gactctggat ccagaaagg gagaaagtta 2460
aaattcccgt cgtatcaag gttccgaaag ccaacaagga aatcctcgat gaagcctacg 2520
tgatggccag cgtggacaac cccacagtg gccgctgct gggcatctgc ctcacctcca 2580
ccgtgcagct catcacgag ctcatgccct cggctgctc cctggactat gtcggggaac 2640
acaaagacaa tattggctcc cagtacctgc tcaactggtg tgtgcagatc gcaagggca 2700
tgaaactact ggaggaccgt cgcttggtgc accgcgacct ggcagccagg aacgtactgg 2760
tgaaaacacc gcagcatgtc aagatcacag attttgggct ggccaaactg ctgggtgagg 2820
aagagaaaga ataccatgca gaaggaggca aagtgcctat caagtggatg gcattggaat 2880
caattttaca cagaatctat acccaccaga gtgatgtctg gagctacggg gtgactgttt 2940
gggagttgat gacctttgga tccaagccat atgacggaat cctgccagc gagatctcct 3000
ccatcctgga gaaggagaa cgctccctc agccaccat atgtaccatc gatgtctaca 3060
tgatcatggt caagtgcctg atgatagacg cagatagtcg cccaaagttc cgtgagtga 3120
tcatcgaatt ctccaaaatg gcccgagacc cccagcgcta ccttgtcatt cagggggatg 3180
aaagaatgca tttgccaaat cctacagact ccaacttcta ccgtgccctg atggatgaag 3240
aagacatgga cgacgtggtg gatgcccagc agtacctcat cccacagcag ggcttcttca 3300
gcagccctc cacgtcacgg actccctcc tgagctctct gagtgcacac agcaacaatt 3360
ccaccgtggc ttgcattgat agaaatgggc tgcaaaagct tccatcaag gaagacagct 3420
tcttgacgag atacagctca gacccacag gcgccttgac tgaggacagc atagacgaca 3480
ccttccctcc agtgccctgaa tacataaacc agtcctgtcc caaaaggccc gctggctctg 3540
tgcagaatcc tgtctatcac aatcagctc tgaacccgc gccagcaga gaccacact 3600
accaggacc ccacagcat gcagtgggca accccagta tctcaacact gtcacgcca 3660
cctgtgtcaa cagcacattc gacagccctg cccactgggc ccagaaaggc agccacaaa 3720
ttagcctgga caacctgac taccagcagg acttctttcc caaggagcc aagccaaatg 3780
gcattcttaa gggctccaca gctgaaaatg cagaataoct aagggtcgcg ccacaaagca 3840
gtgaatttat tggagcatga 3860

```

<210> 445

<211> 3878

<212> ADN

5 <213> Homo sapiens

<400> 445

```

ccccggcgag cgcggcgcca gcagcctccg cccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggcgg cggagtcccc agctagcccc ggcgcccgcc gccgcccaga cccgacgaca 120
ggccacctcg tcggcgctccg cccgagtcgg cgctcgccgg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagttattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccgcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggagggt gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccactgggg agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccacgagct 900
gctccggggc ctgcgctggc aagtccccc gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg cccccgggag agcgactgcc tggctcgccg caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccg gggcaaatat agcttttggt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaatagggtat tgggtgaattt 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
cagtctgtag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gcacctcggg cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg caggggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgcc 1920
accagagtg tgccctacg gccatgaaca tcacctgcac aggcgggga agacacagct 1980
gtatcagtg tgcctctac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaagg ggcctcctc 2160
atgggcctaa gatcccgctc atcgccactg ggatgggtgg ggcctcctc 2220
tggtggccct ggggatcgcc ctcttcacgc gaaggcgcca catcggtcgg aagcgcaacg 2280
tgcggaaggc gctgcaggag agggagcttg tggagcctc tacacccagt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgagg aaactgaatt caaaaagatc aaagtgtgtg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccg cgtatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580
gcatctgcct cacctccacc gtgcagctca tcacgcagct catgccctc ggctgcctcc 2640
tggactatgt ccgggaacac aaagacaata ttggctcca gtacctgtc aactggtgtg 2700
tgcagatgc aaagggcatg aactacttgg aggcagctc cttggtgcac cgcgacctgg 2760
cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcggg 2820
ccaaactgct ggtgcgggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880
agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940
gctacggggg gactgttttg gagttgatga ctttggtatc caagccatat gacggaatcc 3000
ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccaccatat 3060
gtaccatgca tgcctacatg atcaggttca atgctggat gatagacga gatagtccg 3120
caaagtccg tgagttagtc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180
ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240
gtgccctgat ggaatgaaga gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300
cacagcaggg cttcttcacg agccctcca cgtcacggac tcccctctg agctctctga 3360
gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420

ccatcaagga agacagcttc ttgcagcgat acagctcaga cccacaggc gccttgactg 3480
aggacagcat agacgacacc ttcctcccag tgctgaata cataaaccag tccgttccca 3540
aaaggcccg tggtctgtg cagaatcctg tctatcacaa tcagcctctg aaccgccg 3600
ccagcagaga ccacactac caggaccccc aggcactgc agtgggcaac cccagtatc 3660
tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggcc 3720
agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780
aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840
gggtcgcgcc acaagcaggt gaatttatgg gagcatga 3878

```

<210> 446

<211> 3878

5 <212> ADN

<213> Homo sapiens

<400> 446

```

ccccggcgag cgccggcgca gcagcctccg ccccccgac ggtgtgagcg cccgacgagg 60
ccgagggggc cggagtcccg agctagcccc ggccggcgcc gccgcccaga cccgacgaca 120
ggccacctcg tcggcgctccg cccgagtcgc cgctcgcccg ccaacgccac aaccaccgag 180
cacggccccc tgaactccgtc cagttattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccggcag tcgggctctg gagaaaaaga 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
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggccgctgct gggtcagcaa caacctgcc ctgtgcaacg 720
tgagagcagc caagtggcgg gacatagtcg cagtgactt tctcagcaac atgtcgatgg 780
acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggaaggag aactgccaga aactgaccaa aatcatctgt gccagcag 900
gctccgggag ctgcccgtggc aagtcccca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcgactgcc tggctgccc ccaattccga gacgaagcca 1020
cgtgcaagga cactgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccg gggcaaatat agcttttggt ccacctgctg gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcctcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagcagg cgtccgcaag tgtaagaagt gcgaagggcc ttgcccgaac gtgtgtaacg 1260
gaataggtat tggtaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cactccatc agtgccgcat tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggttttt ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactgaaaaa 1680
aactgtttgg gacctccggc cagaaaaacca aaattataag caacagaggt gaaaaacagct 1740
gcaaggccac aggccaggtc tgctgctccc cgagggtgct tggggcccg 1800
agcccgaggc ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggcagggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgctg caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggtggtggg ggccctctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgcggagggt gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgagg aaactgaatt caaaaagatc aaagtgtgg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattccggt cgctatcaag gaattaaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgctgctgg 2580
gcactctgct cactccacc gtgcagctca tcacgcagct catgccttc ggctgctcc 2640
tggaactatg ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700
tgcagatcgc aaagggcag aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760
cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggagg 2820
ccaaactgct ggtgcggaag gagaaagaat accatgcaga aggaggcaaa gtgcctatga 2880
agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940

gctacggggt gactgttttg gagttgatga cctttggatc caagccatat gacggaatcc 3000
ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccaccatat 3060
gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtccgc 3120
caaagttccg tgagttgatc atcgaattct ccaaatggc ccgagacccc cagcgtacc 3180
ttgtcattca ggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240
gtgccctgat ggtgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300
cacagcaggc cttcttcagc agcccccca cgtcacggac tccccctctg agctctctga 3360
gtgcaaccag caacaatttc accgtggctt cattgatag aaatggctg caaagctgtc 3420
ccatcaagga agacagcttc ttgcagcgat acagctcaga cccacaggc gccttgactg 3480
aggacagcat agacgacacc ttctcccag tgctgaata cataaaccag tccgttccca 3540
aaaggcccg tggctctgtg cagaatcctg tctatcacia tcagcctctg aaccccgcg 3600
ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtac 3660
tcaacactgt cagcccacc tgtgtcaaca gcacattcga cagccctgcc cactggggcc 3720
agaaaggcag ccaccaaatt agcctggaca acctgacta ccagcaggac ttctttccca 3780
aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840
gggtcgcgcc acaagcaggt gaatttatgg gagcatga 3878

```

<211> 36
 <212> ADN
 <213> Homo sapiens
 <400> 447
 5 tgtaaaacga cggccagtcg cccagaccgg acgaca 36
 <210> 448
 <211> 39
 <212> ADN
 <213> Homo sapiens
 10 <400> 448
 caggaaacag ctatgaccag ggcaatgagg acataacca 39
 <210> 449
 <211> 38
 <212> ADN
 15 <213> Homo sapiens
 <400> 449
 tgtaaaacga cggccagtgg tggctcttgg gaatttgg 38
 <210> 450
 <211> 40
 20 <212> ADN
 <213> Homo sapiens
 <400> 450
 caggaaacag ctatgacccc atcgacatgt tgctgagaaa 40
 <210> 451
 25 <211> 38
 <212> ADN
 <213> Homo sapiens
 <400> 451
 tgtaaaacga cggccagtga aggagctgcc catgagaa 38
 30 <210> 452
 <211> 38
 <212> ADN
 <213> Homo sapiens
 <400> 452
 35 caggaaacag ctatgaccgg tggcttcgtc tcggaatt 38
 <210> 453
 <211> 40

<212> ADN
 <213> Homo sapiens
 <400> 453
 tgtaaacga cgccagtga aactgaccaa aatcatctgt 40
 5 <210> 454
 <211> 40
 <212> ADN
 <213> Homo sapiens
 <400> 454
 10 caggaaacag ctatgacctt cctattccgt tacacacttt 40
 <210> 455
 <211> 40
 <212> ADN
 <213> Homo sapiens
 15 <400> 455
 tgtaaacga cgccagtcc gtaattatgt ggtgacagat 40
 <210> 456
 <211> 40
 <212> ADN
 20 <213> Homo sapiens
 <400> 456
 caggaaacag ctatgaccgc gtatgatttc taggttctca 40
 <210> 457
 <211> 41
 25 <212> ADN
 <213> Homo sapiens
 <400> 457
 tgtaaacga cgccagtct gaaaaccgta aaggaaatca c 41
 <210> 458
 30 <211> 37
 <212> ADN
 <213> Homo sapiens
 <400> 458
 caggaaacag ctatgacccc tgctcggt gacattc 37
 35 <210> 459
 <211> 40
 <212> ADN

<213> Homo sapiens
 <400> 459
 tgtaaaacga cggccagtta agcaacagag gtgaaaacag 40
 <210> 460
 5 <211> 40
 <212> ADN
 <213> Homo sapiens
 <400> 460
 caggaaacag ctatgaccgg tgtgttttc tccatgact 40
 10 <210> 461
 <211> 38
 <212> ADN
 <213> Homo sapiens
 <400> 461
 15 tgtaaaacga cggccagtgg accagacaac tgtatcca 38
 <210> 462
 <211> 40
 <212> ADN
 <213> Homo sapiens
 20 <400> 462
 caggaaacag ctatgacctt cctcaagat cctcaagaga 40
 <210> 463
 <211> 38
 <212> ADN
 25 <213> Homo sapiens
 <400> 463
 tgtaaaacga cggccagtga tcggcctctt catgcgaa 38
 <210> 464
 <211> 38
 30 <212> ADN
 <213> Homo sapiens
 <400> 464
 caggaaacag ctatgaccac ggtggagggtg aggagat 38
 <210> 465
 35 <211> 39
 <212> ADN
 <213> Homo sapiens

<400> 465
 tgtaaaacga cggccagtcg aaagccaaca aggaaatcc 39
 <210> 466
 <211> 40
 5 <212> ADN
 <213> Homo sapiens
 <400> 466
 caggaaacag ctatgaccat tccaatgcca tccactgat 40
 <210> 467
 10 <211> 39
 <212> ADN
 <213> Homo sapiens
 <400> 467
 tgtaaaacga cggccagtaa caccgcagca tgtaagat 39
 15 <210> 468
 <211> 39
 <212> ADN
 <213> Homo sapiens
 <400> 468
 20 caggaaacag ctatgaccct cgggccattt tggagaatt 39
 <210> 469
 <211> 39
 <212> ADN
 <213> Homo sapiens
 25 <400> 469
 tgtaaaacga cggccagttc agccacccat atgtaccat 39
 <210> 470
 <211> 39
 <212> ADN
 30 <213> Homo sapiens
 <400> 470
 caggaaacag ctatgaccgc ttgcagccc atttctatc 39
 <210> 471
 <211> 38
 35 <212> ADN
 <213> Homo sapiens
 <400> 471

tgtaaaacga cggccagtac agcagggctt cttcagca 38
 <210> 472
 <211> 38
 <212> ADN
 5 <213> Homo sapiens
 <400> 472
 caggaaacag ctatgacctg acacaggtgg gctggaca 38
 <210> 473
 <211> 40
 10 <212> ADN
 <213> Homo sapiens
 <400> 473
 tgtaaaacga cggccagtga atcctgtcta tcacaatcag 40
 <210> 474
 15 <211> 40
 <212> ADN
 <213> Homo sapiens
 <400> 474
 caggaaacag ctatgaccgg tatcgaaaga gtctggatt 40
 20 <210> 475
 <211> 38
 <212> ADN
 <213> Homo sapiens
 <400> 475
 25 tgtaaaacga cggccagtgc tccacagctg aaaatgca 38
 <210> 476
 <211> 39
 <212> ADN
 <213> Homo sapiens
 30 <400> 476
 caggaaacag ctatgaccac gttgcaaaac cagtctgtg 39
 <210> 477
 <211> 25
 <212> PRT
 35 <213> Homo sapiens
 <400> 477

ES 2 741 574 T3

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

5 <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

10 <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

15 <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

20

<210> 481

<211> 25

<212> PRT

<213> Homo sapiens

25 <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
5           20           25

```

<210> 483

<211> 19

<212> PRT

<213> Homo sapiens

10 <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

15 <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

20 <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

25 <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

5 <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

10 <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

15 <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

20 <211> 6

<212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: motivo de mutación ilustrativo

25 <220>

<221> MOD_RES

<222> (2)

<223> residuo variable

<220>

30 <221> MOD_RES

<222> (4)..(5)

<223> residuo variable

<400> 490

Gly Xaa Gly Xaa Xaa Gly
 1 5

<210> 491

5 <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

ES 2 741 574 T3

<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

5 <211> 54

<212> ADN

<213> Homo sapiens

<400> 493

aaaattcccg tcgctatcaa ggaattaaga gaagcaacat ctccgaaagc caac 54

10 <210> 494

<211> 24

<212> ADN

<213> Homo sapiens

<400> 494
 ttgggctgg ccaaactgct gggt 24
 <210> 495
 <211> 39
 5 <212> ADN
 <213> Homo sapiens
 <400> 495
 aaaattcccg tcgctatcaa aacatctccg aaagccaac 39
 <210> 496
 10 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 496
 ttgggctgg ccaaactgct gggt 24
 15 <210> 497
 <211> 42
 <212> ADN
 <213> Homo sapiens
 <400> 497
 20 aaaattcccg tcgctatcaa ggaatcatct ccgaaagcca ac 42
 <210> 498
 <211> 24
 <212> ADN
 <213> Homo sapiens
 25 <400> 498
 ttgggctgg ccaaactgct gggt 24
 <210> 499
 <211> 36
 <212> ADN
 30 <213> Homo sapiens
 <400> 499
 aaaattcccg tcgctatcaa ggaatcgaaa gccaac 36
 <210> 500
 <211> 24
 35 <212> ADN
 <213> Homo sapiens
 <400> 500

ES 2 741 574 T3

tttgggctgg ccaaactgct gggt 24
 <210> 501
 <211> 54
 <212> ADN
 5 <213> Homo sapiens
 <400> 501
 aaaattcccg tcgctatcaa ggaattaaga gaagcaacat ctccgaaagc caac 54
 <210> 502
 <211> 24
 10 <212> ADN
 <213> Homo sapiens
 <400> 502
 tttgggcggg ccaaactgct gggt 24
 <210> 503
 15 <211> 54
 <212> ADN
 <213> Homo sapiens
 <400> 503
 aaaattcccg tcgctatcaa ggaattaaga gaagcaacat ctccgaaagc caac 54
 20 <210> 504
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 504
 25 tttgggctgg ccaaacagct 24
 <210> 505
 <211> 25
 <212> ADN
 <213> Homo sapiens
 30 <400> 505
 gcaatatcag ccttaggtgc ggctc 25
 <210> 506
 <211> 26
 <212> ADN
 35 <213> Homo sapiens
 <400> 506
 catagaaagt gaacatttag gatgtg 26

<210> 507
 <211> 24
 <212> ADN
 <213> Homo sapiens
 5 <400> 507
 ctaacgttcg ccagccataa gtcc 24
 <210> 508
 <211> 26
 <212> ADN
 10 <213> Homo sapiens
 <400> 508
 gctgcgagct caccagaat gtctgg 26
 <210> 509
 <211> 18
 15 <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
 20 <211> 8
 <212> PRT
 <213> Homo sapiens
 <400> 510
 Phe Gly Leu Ala Lys Leu Leu Gly
 1 5
 25 <210> 511
 <211> 3878
 <212> ADN
 <213> Homo sapiens
 <220>
 30 <221> CDS
 <222> (246)..(3875)
 <400> 511

cccggcgcag cgcggccgca gcagcctccg cccccgcac ggtgtgagcg cccgacgcgg 60
 ccgagggcggc cggagtcctcg agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120
 ggccacctcg tcggcgctccg cccgagtcctc cgcctcgccg ccaacgccac aaccaccgcg 180
 cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct ctcgggggag 240
 cagcg atg cga ccc tcc ggg acg gcc ggg gca gcg ctc ctg gcg ctg ctg 290

Met	Arg	Pro	Ser	Gly	Thr	Ala	Gly	Ala	Ala	Leu	Leu	Ala	Leu	Leu		
1					5					10				15		
gct	gcg	ctc	tgc	ccg	gcg	agt	cgg	gct	ctg	gag	gaa	aag	aaa	ggt	tgc	338
Ala	Ala	Leu	Cys	Pro	Ala	Ser	Arg	Ala	Leu	Glu	Glu	Lys	Lys	Val	Cys	
			20						25					30		
caa	ggc	acg	agt	aac	aag	ctc	acg	cag	ttg	ggc	act	ttt	gaa	gat	cat	386
Gln	Gly	Thr	Ser	Asn	Lys	Leu	Thr	Gln	Leu	Gly	Thr	Phe	Glu	Asp	His	
			35					40					45			
ttt	ctc	agc	ctc	cag	agg	atg	ttc	aat	aac	tgt	gag	gtg	gtc	ctt	ggg	434
Phe	Leu	Ser	Leu	Gln	Arg	Met	Phe	Asn	Asn	Cys	Glu	Val	Val	Leu	Gly	
		50					55					60				
aat	ttg	gaa	att	acc	tat	gtg	cag	agg	aat	tat	gat	ctt	tcc	ttc	tta	482
Asn	Leu	Glu	Ile	Thr	Tyr	Val	Gln	Arg	Asn	Tyr	Asp	Leu	Ser	Phe	Leu	
	65					70					75					
aag	acc	atc	cag	gag	gtg	gct	ggt	tat	gtc	ctc	att	gcc	ctc	aac	aca	530
Lys	Thr	Ile	Gln	Glu	Val	Ala	Gly	Tyr	Val	Leu	Ile	Ala	Leu	Asn	Thr	
	80				85				90					95		
gtg	gag	cga	att	cct	ttg	gaa	aac	ctg	cag	atc	atc	aga	gga	aat	atg	578
Val	Glu	Arg	Ile	Pro	Leu	Glu	Asn	Leu	Gln	Ile	Ile	Arg	Gly	Asn	Met	
				100					105					110		
tac	tac	gaa	aat	tcc	tat	gcc	tta	gca	gtc	tta	tct	aac	tat	gat	gca	626
Tyr	Tyr	Glu	Asn	Ser	Tyr	Ala	Leu	Ala	Val	Leu	Ser	Asn	Tyr	Asp	Ala	
			115					120					125			
aat	aaa	acc	gga	ctg	aag	gag	ctg	ccc	atg	aga	aat	tta	cag	gaa	atc	674
Asn	Lys	Thr	Gly	Leu	Lys	Glu	Leu	Pro	Met	Arg	Asn	Leu	Gln	Glu	Ile	
		130					135					140				
ctg	cat	ggc	gcc	gtg	cgg	ttc	agc	aac	aac	cct	gcc	ctg	tgc	aac	gtg	722
Leu	His	Gly	Ala	Val	Arg	Phe	Ser	Asn	Asn	Pro	Ala	Leu	Cys	Asn	Val	
	145					150					155					
gag	agc	atc	cag	tgg	cgg	gac	ata	gtc	agc	agt	gac	ttt	ctc	agc	aac	770
Glu	Ser	Ile	Gln	Trp	Arg	Asp	Ile	Val	Ser	Ser	Asp	Phe	Leu	Ser	Asn	
	160				165				170					175		
atg	tcg	atg	gac	ttc	cag	aac	cac	ctg	ggc	agc	tgc	caa	aag	tgt	gat	818
Met	Ser	Met	Asp	Phe	Gln	Asn	His	Leu	Gly	Ser	Cys	Gln	Lys	Cys	Asp	
			180						185					190		
cca	agc	tgt	ccc	aat	ggg	agc	tgc	tgg	ggt	gca	gga	gag	gag	aac	tgc	866
Pro	Ser	Cys	Pro	Asn	Gly	Ser	Cys	Trp	Gly	Ala	Gly	Glu	Glu	Asn	Cys	
			195					200					205			
cag	aaa	ctg	acc	aaa	atc	atc	tgt	gcc	cag	cag	tgc	tcc	ggg	cgc	tgc	914
Gln	Lys	Leu	Thr	Lys	Ile	Ile	Cys	Ala	Gln	Gln	Cys	Ser	Gly	Arg	Cys	
		210				215						220				
cgt	ggc	aag	tcc	ccc	agt	gac	tgc	tgc	cac	aac	cag	tgt	gct	gca	ggc	962
Arg	Gly	Lys	Ser	Pro	Ser	Asp	Cys	Cys	His	Asn	Gln	Cys	Ala	Ala	Gly	
	225					230						235				
tgc	aca	ggc	ccc	cgg	gag	agc	gac	tgc	ctg	gtc	tgc	cgc	aaa	ttc	cga	1010
Cys	Thr	Gly	Pro	Arg	Glu	Ser	Asp	Cys	Leu	Val	Cys	Arg	Lys	Phe	Arg	
	240				245					250				255		
gac	gaa	gcc	acg	tgc	aag	gac	acc	tgc	ccc	cca	ctc	atg	ctc	tac	aac	1058

ES 2 741 574 T3

Asp	Glu	Ala	Thr	Cys	Lys	Asp	Thr	Cys	Pro	Pro	Leu	Met	Leu	Tyr	Asn	
				260					265					270		
ccc	acc	acg	tac	cag	atg	gat	gtg	aac	ccc	gag	ggc	aaa	tac	agc	ttt	1106
Pro	Thr	Thr	Tyr	Gln	Met	Asp	Val	Asn	Pro	Glu	Gly	Lys	Tyr	Ser	Phe	
			275					280					285			
ggt	gcc	acc	tgc	gtg	aag	aag	tgt	ccc	cgt	aat	tat	gtg	gtg	aca	gat	1154
Gly	Ala	Thr	Cys	Val	Lys	Lys	Cys	Pro	Arg	Asn	Tyr	Val	Val	Thr	Asp	
			290				295					300				
cac	ggc	tgc	tgc	gtc	cga	gcc	tgt	ggg	gcc	gac	agc	tat	gag	atg	gag	1202
His	Gly	Ser	Cys	Val	Arg	Ala	Cys	Gly	Ala	Asp	Ser	Tyr	Glu	Met	Glu	
	305					310					315					
gaa	gac	ggc	gtc	cgc	aag	tgt	aag	aag	tgc	gaa	ggg	cct	tgc	cgc	aaa	1250
Glu	Asp	Gly	Val	Arg	Lys	Cys	Lys	Lys	Cys	Glu	Gly	Pro	Cys	Arg	Lys	
320					325					330					335	
gtg	tgt	aac	gga	ata	ggt	att	ggt	gaa	ttt	aaa	gac	tca	ctc	tcc	ata	1298
Val	Cys	Asn	Gly	Ile	Gly	Ile	Gly	Glu	Phe	Lys	Asp	Ser	Leu	Ser	Ile	
			340						345					350		
aat	gct	acg	aat	att	aaa	cac	ttc	aaa	aac	tgc	acc	tcc	atc	agt	ggc	1346
Asn	Ala	Thr	Asn	Ile	Lys	His	Phe	Lys	Asn	Cys	Thr	Ser	Ile	Ser	Gly	
			355					360					365			
gat	ctc	cac	atc	ctg	ccg	gtg	gca	ttt	agg	ggt	gac	tcc	ttc	aca	cat	1394
Asp	Leu	His	Ile	Leu	Pro	Val	Ala	Phe	Arg	Gly	Asp	Ser	Phe	Thr	His	
			370				375					380				
act	cct	cct	ctg	gat	cca	cag	gaa	ctg	gat	att	ctg	aaa	acc	gta	aag	1442
Thr	Pro	Pro	Leu	Asp	Pro	Gln	Glu	Leu	Asp	Ile	Leu	Lys	Thr	Val	Lys	
	385					390					395					
gaa	atc	aca	ggg	ttt	ttg	ctg	att	cag	gct	tgg	cct	gaa	aac	agg	acg	1490
Glu	Ile	Thr	Gly	Phe	Leu	Leu	Ile	Gln	Ala	Trp	Pro	Glu	Asn	Arg	Thr	
400					405					410					415	
gac	ctc	cat	gcc	ttt	gag	aac	cta	gaa	atc	ata	cgc	ggc	agg	acc	aag	1538
Asp	Leu	His	Ala	Phe	Glu	Asn	Leu	Glu	Ile	Ile	Arg	Gly	Arg	Thr	Lys	
			420						425					430		
caa	cat	ggt	cag	ttt	tct	ctt	gca	gtc	gtc	agc	ctg	aac	ata	aca	tcc	1586
Gln	His	Gly	Gln	Phe	Ser	Leu	Ala	Val	Val	Ser	Leu	Asn	Ile	Thr	Ser	
			435					440					445			
ttg	gga	tta	cgc	tcc	ctc	aag	gag	ata	agt	gat	gga	gat	gtg	ata	att	1634
Leu	Gly	Leu	Arg	Ser	Leu	Lys	Glu	Ile	Ser	Asp	Gly	Asp	Val	Ile	Ile	
			450				455					460				
tca	gga	aac	aaa	aat	ttg	tgc	tat	gca	aat	aca	ata	aac	tgg	aaa	aaa	1682
Ser	Gly	Asn	Lys	Asn	Leu	Cys	Tyr	Ala	Asn	Thr	Ile	Asn	Trp	Lys	Lys	
	465					470					475					
ctg	ttt	ggg	acc	tcc	ggt	cag	aaa	acc	aaa	att	ata	agc	aac	aga	ggt	1730
Leu	Phe	Gly	Thr	Ser	Gly	Gln	Lys	Thr	Lys	Ile	Ile	Ser	Asn	Arg	Gly	
480					485					490					495	
gaa	aac	agc	tgc	aag	gcc	aca	ggc	cag	gtc	tgc	cat	gcc	ttg	tgc	tcc	1778
Glu	Asn	Ser	Cys	Lys	Ala	Thr	Gly	Gln	Val	Cys	His	Ala	Leu	Cys	Ser	
				500					505				510			
ccc	gag	ggc	tgc	tgg	ggc	ccg	gag	ccc	agg	gac	tgc	gtc	tct	tgc	cgg	1826
Pro	Glu	Gly	Cys	Trp	Gly	Pro	Glu	Pro	Arg	Asp	Cys	Val	Ser	Cys	Arg	

ES 2 741 574 T3

515										520										525										
aat	gtc	agc	cga	ggc	agg	gaa	tgc	gtg	gac	aag	tgc	aac	ctt	ctg	gag	1874														
Asn	Val	Ser	Arg	Gly	Arg	Glu	Cys	Val	Asp	Lys	Cys	Asn	Leu	Leu	Glu															
530																535	540													
ggt	gag	cca	agg	gag	ttt	gtg	gag	aac	tct	gag	tgc	ata	cag	tgc	cac	1922														
Gly	Glu	Pro	Arg	Glu	Phe	Val	Glu	Asn	Ser	Glu	Cys	Ile	Gln	Cys	His															
545																550	555													
cca	gag	tgc	ctg	cct	cag	gcc	atg	aac	atc	acc	tgc	aca	gga	cgg	gga	1970														
Pro	Glu	Cys	Leu	Pro	Gln	Ala	Met	Asn	Ile	Thr	Cys	Thr	Gly	Arg	Gly															
560																565	570										575			
cca	gac	aac	tgt	atc	cag	tgt	gcc	cac	tac	att	gac	ggc	ccc	cac	tgc	2018														
Pro	Asp	Asn	Cys	Ile	Gln	Cys	Ala	His	Tyr	Ile	Asp	Gly	Pro	His	Cys															
580																585	590													
gtc	aag	acc	tgc	ccg	gca	gga	gtc	atg	gga	gaa	aac	aac	acc	ctg	gtc	2066														
Val	Lys	Thr	Cys	Pro	Ala	Gly	Val	Met	Gly	Glu	Asn	Asn	Thr	Leu	Val															
595																600	605													
tgg	aag	tac	gca	gac	gcc	ggc	cat	gtg	tgc	cac	ctg	tgc	cat	cca	aac	2114														
Trp	Lys	Tyr	Ala	Asp	Ala	Gly	His	Val	Cys	His	Leu	Cys	His	Pro	Asn															
610																615	620													
tgc	acc	tac	gga	tgc	act	ggg	cca	ggt	ctt	gaa	ggc	tgt	cca	acg	aat	2162														
Cys	Thr	Tyr	Gly	Cys	Thr	Gly	Pro	Gly	Leu	Glu	Gly	Cys	Pro	Thr	Asn															
625																630	635													
ggg	cct	aag	atc	ccg	tcc	atc	gcc	act	ggg	atg	gtg	ggg	gcc	ctc	ctc	2210														
Gly	Pro	Lys	Ile	Pro	Ser	Ile	Ala	Thr	Gly	Met	Val	Gly	Ala	Leu	Leu															
640																645	650										655			
ttg	ctg	ctg	gtg	gtg	gcc	ctg	ggg	atc	ggc	ctc	ttc	atg	cga	agg	cgc	2258														
Leu	Leu	Leu	Val	Val	Ala	Leu	Gly	Ile	Gly	Leu	Phe	Met	Arg	Arg	Arg															
660																665	670													
cac	atc	gtt	cgg	aag	cgc	acg	ctg	cgg	agg	ctg	ctg	cag	gag	agg	gag	2306														
His	Ile	Val	Arg	Lys	Arg	Thr	Leu	Arg	Arg	Leu	Leu	Gln	Glu	Arg	Glu															
675																680	685													
ctt	gtg	gag	cct	ctt	aca	ccc	agt	gga	gaa	gct	ccc	aac	caa	gct	ctc	2354														
Leu	Val	Glu	Pro	Leu	Thr	Pro	Ser	Gly	Glu	Ala	Pro	Asn	Gln	Ala	Leu															
690																695	700													
ttg	agg	atc	ttg	aag	gaa	act	gaa	ttc	aaa	aag	atc	aaa	gtg	ctg	ggc	2402														
Leu	Arg	Ile	Leu	Lys	Glu	Thr	Glu	Phe	Lys	Lys	Ile	Lys	Val	Leu	Gly															
705																710	715													
tcc	ggt	gcg	ttc	ggc	acg	gtg	tat	aag	gga	ctc	tgg	atc	cca	gaa	ggt	2450														
Ser	Gly	Ala	Phe	Gly	Thr	Val	Tyr	Lys	Gly	Leu	Trp	Ile	Pro	Glu	Gly															
720																725	730										735			
gag	aaa	gtt	aaa	att	ccc	gtc	gct	atc	aag	gaa	tta	aga	gaa	gca	aca	2498														
Glu	Lys	Val	Lys	Ile	Pro	Val	Ala	Ile	Lys	Glu	Leu	Arg	Glu	Ala	Thr															
740																745	750													
tct	ccg	aaa	gcc	aac	aag	gaa	atc	ctc	gat	gaa	gcc	tac	gtg	atg	gcc	2546														
Ser	Pro	Lys	Ala	Asn	Lys	Glu	Ile	Leu	Asp	Glu	Ala	Tyr	Val	Met	Ala															
755																760	765													
agc	gtg	gac	aac	ccc	cac	gtg	tgc	cgc	ctg	ctg	ggc	atc	tgc	ctc	acc	2594														
Ser	Val	Asp	Asn	Pro	His	Val	Cys	Arg	Leu	Leu	Gly	Ile	Cys	Leu	Thr															
770																775	780													

ES 2 741 574 T3

tcc acc gtg cag ctc atc acg cag ctc atg ccc ttc ggc tgc ctc ctg Ser Thr Val Gln Leu Ile Thr Gln Leu Met Pro Phe Gly Cys Leu Leu 785 790 795	2642
gac tat gtc cgg gaa cac aaa gac aat att ggc tcc cag tac ctg ctc Asp Tyr Val Arg Glu His Lys Asp Asn Ile Gly Ser Gln Tyr Leu Leu 800 805 810 815	2690
aac tgg tgt gtg cag atc gca aag ggc atg aac tac ttg gag gac cgt Asn Trp Cys Val Gln Ile Ala Lys Gly Met Asn Tyr Leu Glu Asp Arg 820 825 830	2738
cgc ttg gtg cac cgc gac ctg gca gcc agg aac gta ctg gtg aaa aca Arg Leu Val His Arg Asp Leu Ala Ala Arg Asn Val Leu Val Lys Thr 835 840 845	2786
ccg cag cat gtc aag atc aca gat ttt ggg ctg gcc aaa ctg ctg ggt Pro Gln His Val Lys Ile Thr Asp Phe Gly Leu Ala Lys Leu Leu Gly 850 855 860	2834
gcg gaa gag aaa gaa tac cat gca gaa gga ggc aaa gtg cct atc aag Ala Glu Glu Lys Glu Tyr His Ala Glu Gly Gly Lys Val Pro Ile Lys 865 870 875	2882
tgg atg gca ttg gaa tca att tta cac aga atc tat acc cac cag agt Trp Met Ala Leu Glu Ser Ile Leu His Arg Ile Tyr Thr His Gln Ser 880 885 890 895	2930
gat gtc tgg agc tac ggg gtg act gtt tgg gag ttg atg acc ttt gga Asp Val Trp Ser Tyr Gly Val Thr Val Trp Glu Leu Met Thr Phe Gly 900 905 910	2978
tcc aag cca tat gac gga atc cct gcc agc gag atc tcc tcc atc ctg Ser Lys Pro Tyr Asp Gly Ile Pro Ala Ser Glu Ile Ser Ser Ile Leu 915 920 925	3026
gag aaa gga gaa cgc ctc cct cag cca ccc ata tgt acc atc gat gtc Glu Lys Gly Glu Arg Leu Pro Gln Pro Pro Ile Cys Thr Ile Asp Val 930 935 940	3074
tac atg atc atg gtc aag tgc tgg atg ata gac gca gat agt cgc cca Tyr Met Ile Met Val Lys Cys Trp Met Ile Asp Ala Asp Ser Arg Pro 945 950 955	3122
aag ttc cgt gag ttg atc atc gaa ttc tcc aaa atg gcc cga gac ccc Lys Phe Arg Glu Leu Ile Ile Glu Phe Ser Lys Met Ala Arg Asp Pro 960 965 970 975	3170
cag cgc tac ctt gtc att cag ggg gat gaa aga atg cat ttg cca agt Gln Arg Tyr Leu Val Ile Gln Gly Asp Glu Arg Met His Leu Pro Ser 980 985 990	3218
cct aca gac tcc aac ttc tac cgt gcc ctg atg gat gaa gaa gac atg Pro Thr Asp Ser Asn Phe Tyr Arg Ala Leu Met Asp Glu Glu Asp Met 995 1000 1005	3266
gac gac gtg gtg gat gcc gac gag tac ctc atc cca cag cag ggc ttc Asp Asp Val Val Asp Ala Asp Glu Tyr Leu Ile Pro Gln Gln Gly Phe 1010 1015 1020	3314
ttc agc agc ccc tcc acg tca cgg act ccc ctc ctg agc tct ctg agt Phe Ser Ser Pro Ser Thr Ser Arg Thr Pro Leu Leu Ser Ser Leu Ser 1025 1030 1035	3362

ES 2 741 574 T3

gca acc agc aac aat tcc acc gtg gct tgc att gat aga aat ggg ctg 3410
Ala Thr Ser Asn Asn Ser Thr Val Ala Cys Ile Asp Arg Asn Gly Leu
1040 1045 1050 1055

caa agc tgt ccc atc aag gaa gac agc ttc ttg cag cga tac agc tca 3458
Gln Ser Cys Pro Ile Lys Glu Asp Ser Phe Leu Gln Arg Tyr Ser Ser
1060 1065 1070

gac ccc aca ggc gcc ttg act gag gac agc ata gac gac acc ttc ctg 3506
Asp Pro Thr Gly Ala Leu Thr Glu Asp Ser Ile Asp Asp Thr Phe Leu
1075 1080 1085

cca gtg cct gaa tac ata aac cag tcc gtt ccc aaa agg ccc gct ggc 3554
Pro Val Pro Glu Tyr Ile Asn Ser Val Pro Lys Arg Pro Ala Gly
1090 1095 1100

tct gtg cag aat cct gtc tat cac aat cag cct ctg aac ccc gcg ccc 3602
Ser Val Gln Asn Pro Val Tyr His Asn Gln Pro Leu Asn Pro Ala Pro
1105 1110 1115

agc aga gac cca cac tac cag gac ccc cac agc act gca gtg ggc aac 3650
Ser Arg Asp Pro His Tyr Gln Asp Pro His Ser Thr Ala Val Gly Asn
1120 1125 1130 1135

ccc gag tat ctg aac act gtc cag ccc acc tgt gtc aac agc aca ttc 3698
Pro Glu Tyr Leu Asn Thr Val Gln Pro Thr Cys Val Asn Ser Thr Phe
1140 1145 1150

gac agc cct gcc cac tgg gcc cag aaa ggc agc cac caa att agc ctg 3746
Asp Ser Pro Ala His Trp Ala Gln Lys Gly Ser His Gln Ile Ser Leu
1155 1160 1165

gac aac cct gac tac cag cag gac ttc ttt ccc aag gaa gcc aag cca 3794
Asp Asn Pro Asp Tyr Gln Gln Asp Phe Phe Pro Lys Glu Ala Lys Pro
1170 1175 1180

aat ggc atc ttt aag ggc tcc aca gct gaa aat gca gaa tac cta agg 3842
Asn Gly Ile Phe Lys Gly Ser Thr Ala Glu Asn Ala Glu Tyr Leu Arg
1185 1190 1195

gtc gcg cca caa agc agt gaa ttt att gga gca tga 3878
Val Ala Pro Gln Ser Ser Glu Phe Ile Gly Ala
1200 1205 1210

<210> 512

<211> 1210

<212> PRT

5 <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

ES 2 741 574 T3

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

ES 2 741 574 T3

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							

ES 2 741 574 T3

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	Phe		
	1010					1015					1020						
Ser	Ser	Pro	Ser	Thr	Ser	Arg	Thr	Pro	Leu	Leu	Ser	Ser	Leu	Ser	Ala		
1025					1030					1035					1040		
Thr	Ser	Asn	Asn	Ser	Thr	Val	Ala	Cys	Ile	Asp	Arg	Asn	Gly	Leu	Gln		
				1045					1050					1055			
Ser	Cys	Pro	Ile	Lys	Glu	Asp	Ser	Phe	Leu	Gln	Arg	Tyr	Ser	Ser	Asp		
			1060					1065						1070			
Pro	Thr	Gly	Ala	Leu	Thr	Glu	Asp	Ser	Ile	Asp	Asp	Thr	Phe	Leu	Pro		
		1075					1080						1085				
Val	Pro	Glu	Tyr	Ile	Asn	Gln	Ser	Val	Pro	Lys	Arg	Pro	Ala	Gly	Ser		
	1090					1095					1100						
Val	Gln	Asn	Pro	Val	Tyr	His	Asn	Gln	Pro	Leu	Asn	Pro	Ala	Pro	Ser		
1105					1110					1115					1120		

ES 2 741 574 T3

Arg Asp Pro His Tyr Gln Asp Pro His Ser Thr Ala Val Gly Asn Pro
1125 1130 1135

Glu Tyr Leu Asn Thr Val Gln Pro Thr Cys Val Asn Ser Thr Phe Asp
1140 1145 1150

Ser Pro Ala His Trp Ala Gln Lys Gly Ser His Gln Ile Ser Leu Asp
1155 1160 1165

Asn Pro Asp Tyr Gln Gln Asp Phe Phe Pro Lys Glu Ala Lys Pro Asn
1170 1175 1180

Gly Ile Phe Lys Gly Ser Thr Ala Glu Asn Ala Glu Tyr Leu Arg Val
1185 1190 1195 1200

Ala Pro Gln Ser Ser Glu Phe Ile Gly Ala
1205 1210

<210> 513

<211> 25

<212> ADN

5 <213> Homo sapiens

<400> 513

cagattggc tcgacctgga catag 25

<210> 514

<211> 22

10 <212> ADN

<213> Homo sapiens

<400> 514

cagctgatct caaggaaaca gg 22

<210> 515

15 <211> 24

<212> ADN

<213> Homo sapiens

<400> 515

gtattatcag tcactaaagc tcac 24

20 <210> 516

<211> 22

<212> ADN

<213> Homo sapiens

<400> 516

25 cacactcaa gtggaattct gc 22

<210> 517

<211> 19

<212> ADN

<213> Homo sapiens
 <400> 517
 tgcattaggg ttcaactgg 19
 <210> 518
 5 <211> 26
 <212> ADN
 <213> Homo sapiens
 <400> 518
 ccttctccga ggtggaattg agtgac 26
 10 <210> 519
 <211> 26
 <212> ADN
 <213> Homo sapiens
 <400> 519
 15 gctaattgcg ggactcttgt tcgcac 26
 <210> 520
 <211> 22
 <212> ADN
 <213> Homo sapiens
 20 <400> 520
 tacatgcttt tctagtggc ag 22
 <210> 521
 <211> 25
 <212> ADN
 25 <213> Homo sapiens
 <400> 521
 ggtctcaagt gattctacaa accag 25
 <210> 522
 <211> 24
 30 <212> ADN
 <213> Homo sapiens
 <400> 522
 ccttcaccta ctggttcaca tctg 24
 <210> 523
 35 <211> 26
 <212> ADN
 <213> Homo sapiens

<400> 523
 catggttga ctagttga atgtgg 26
 <210> 524
 <211> 27
 5 <212> ADN
 <213> Homo sapiens
 <400> 524
 ggatactaaa gatacttgt caccagg 27
 <210> 525
 10 <211> 26
 <212> ADN
 <213> Homo sapiens
 <400> 525
 gaacactagg ctgcaaagac agtaac 26
 15 <210> 526
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 526
 20 ccaagcaagg caaacacatc cacc 24
 <210> 527
 <211> 23
 <212> ADN
 <213> Homo sapiens
 25 <400> 527
 ggaggatgga gccttccat cac 23
 <210> 528
 <211> 24
 <212> ADN
 30 <213> Homo sapiens
 <400> 528
 gaagaggaag atgtgttcct ttgg 24
 <210> 529
 <211> 24
 35 <212> ADN
 <213> Homo sapiens
 <400> 529

gaatgaagga tgatgtggca gtgg 24
 <210> 530
 <211> 21
 <212> ADN
 5 <213> Homo sapiens
 <400> 530
 caaaacatca gccattaacg g 21
 <210> 531
 <211> 26
 10 <212> ADN
 <213> Homo sapiens
 <400> 531
 ccacttactg ttcataat acagag 26
 <210> 532
 15 <211> 25
 <212> ADN
 <213> Homo sapiens
 <400> 532
 catgtgagat agcatttggg aatgc 25
 20 <210> 533
 <211> 26
 <212> ADN
 <213> Homo sapiens
 <400> 533
 25 catgacctac catcattgga aagcag 26
 <210> 534
 <211> 21
 <212> ADN
 <213> Homo sapiens
 30 <400> 534
 gtaatttcac agttaggaat c 21
 <210> 535
 <211> 25
 <212> ADN
 35 <213> Homo sapiens
 <400> 535
 gtcaccaag gtcattggagc acagg 25

<210> 536
 <211> 23
 <212> ADN
 <213> Homo sapiens
 5 <400> 536
 cagaatgcct gtaaagctat aac 23
 <210> 537
 <211> 24
 <212> ADN
 10 <213> Homo sapiens
 <400> 537
 gtcctggagt cccaactcct tgac 24
 <210> 538
 <211> 25
 15 <212> ADN
 <213> Homo sapiens
 <400> 538
 ggaagtggt ctgatggccg tcctg 25
 <210> 539
 20 <211> 27
 <212> ADN
 <213> Homo sapiens
 <400> 539
 ccactcacac acactaaata tttaag 27
 25 <210> 540
 <211> 27
 <212> ADN
 <213> Homo sapiens
 <400> 540
 30 gaccaaaca ccttaagtaa ctgactc 27
 <210> 541
 <211> 24
 <212> ADN
 <213> Homo sapiens
 35 <400> 541
 ccaatccaac atccagacac atag 24
 <210> 542

<211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 542
 5 ccagagccat agaaactga tcag 24
 <210> 543
 <211> 29
 <212> ADN
 <213> Homo sapiens
 10 <400> 543
 gtatggacta tggcacttca attgcatgg 29
 <210> 544
 <211> 28
 <212> ADN
 15 <213> Homo sapiens
 <400> 544
 ccagagaaca tggcaaccag cacaggac 28
 <210> 545
 <211> 25
 20 <212> ADN
 <213> Homo sapiens
 <400> 545
 caaatgagct ggcaagtgcc gtgtc 25
 <210> 546
 25 <211> 25
 <212> ADN
 <213> Homo sapiens
 <400> 546
 gagtttcca aacactcagt gaaac 25
 30 <210> 547
 <211> 25
 <212> ADN
 <213> Homo sapiens
 <400> 547
 35 gcaatatcag ccttaggtgc ggctc 25
 <210> 548
 <211> 26

<212> ADN
 <213> Homo sapiens
 <400> 548
 catagaaagt gaacatttag gatgtg 26
 5 <210> 549
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 549
 10 ccatgagtag gtatttgaa actc 24
 <210> 550
 <211> 24
 <212> ADN
 <213> Homo sapiens
 15 <400> 550
 catatcccca tggcaaactc ttgc 24
 <210> 551
 <211> 24
 <212> ADN
 20 <213> Homo sapiens
 <400> 551
 ctaacgttcg ccagccataa gtcc 24
 <210> 552
 <211> 26
 25 <212> ADN
 <213> Homo sapiens
 <400> 552
 gctgcgagct caccagaat gtctgg 26
 <210> 553
 30 <211> 25
 <212> ADN
 <213> Homo sapiens
 <400> 553
 gacgggtcct ggggtgatct ggctc 25
 35 <210> 554
 <211> 3866
 <212> ADN

<213> Homo sapiens

<400> 554

```

ccccggcgag cgcggcgcca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggcggc cggagtcccc agctagcccc ggcgcccgcc gccgcccaga cccgacgaca 120
ggccacctcg tcggcgctccg cccgagtccc cgctcgcccg ccaacgccac aaccaccgag 180
cacggccccc tgactccgtc cagtgatgat cgggagagcc ggagcgagct ctcgggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccgcgag tcgggctctg gagaaaaga 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
tagcagtcct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tggaagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccacgagt 900
gctccgggag ctgccgtggc aagtccccc gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcgactgcc tggctgccc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccg gggcaaatat agcttttggt ccacctgctg gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggatg ttgtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactctc cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaatcac aggggttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac agcccgaggt tgccatgcct tgtgctcccc cgagggtgct 1800
agcccgagg gtcgctctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgcc 1920
accagagtg cctgcctcag gccatgaaca tcacctgcac aggcgggga ccagacaact 1980
gtatccagtg ttcccactac attgacggcc cccactgctc caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctgggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggtggtggg ggccctctc tttgctgctg 2220
tggtggccct ggggatcggc ctcttcctgc gaaggcgcca catcggtcgg aagcgcacgc 2280
tgcggaagg cgtgcaggag agggagcttg tggagcctct tacacccaag ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgagg aaactgaatt caaaaagatc aaagtgtgg 2400
gctccggtgc gtcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccg cgtatcaag gaacctctc cgaaagccaa caaggaaatc ctcgatgaag 2520
cctcagtgat ggcagcgtg gacaaccccc acgtgtgccc cctgctgggc atctgcctca 2580
cctccacgt gcaagctcac acgcagctca cgccttcgg ctgectctg gactatgtcc 2640
gggaacacaa agacaatatt ggctcccagt acctgctcaa ctggtgtgtg cagatcgcaa 2700
agggcatgaa ctacttgag gaccgtcgct tgggtgcacc cgacctggca gccaggaaacg 2760
tactggtgaa aacaccgcag catgtcaaga tcacagattt tgggctggcc aaactgctgg 2820
gtgcggaaga gaaagacata catgcagaag gagcaaagt gcctatcaag tggatggcat 2880
tggaatcaat ttacacata atctataccc accagagtga tgtctggagc tacggggtga 2940
ctgtttggga gttgatgacc tttggatcca agccatatga cggaatccct gccagcgaga 3000
tctcctccat cctggagaaa ggagaacgcc tccctcagcc acccatatgt accatcgatg 3060
tctacatgat catggtcaga tgctggatga tagacgcaga tagtcgccc aagttccgtg 3120
agttgatcat cgaattctcc aaaatggccc gagaccccc gcgctacctt gtcattcagg 3180
gggatgaaa aatgcatttg ccaagtccta cagactccaa cttctaccgt gccctgatgg 3240
atgaagaaga catggacgac gtggtggatg ccgacgagta cctcatccca cagcagggct 3300
tcttcagcag ccctccacg tcacggactc cctcctgag ctctctgagt gcaaccagca 3360
acaattccac cgtggcttgc attgatagaa atgggctgca aagctgtccc atcaaggag 3420

acagcttctt gcagcgatac agctcagacc ccacaggcgc cttgactgag gacagcatag 3480
acgacacctt cctcccagtg cctgaatata taaaccagtc cgttcccaaa aggcccgtg 3540
gctctgtgca gaatcctgtc tatcacaatc agcctctgaa ccccgcgccc agcagagacc 3600
cacactacca ggacccccag agcactgcag tgggcaaccc cgagtatctc aacactgtcc 3660
agcccacctg cgtgcacagc acattcgaca gccctgccc ctgggcccag aaaggcagcc 3720
accaaattag cctggacaac cctgactacc agcaggactt ctttcccaag gaagccaagc 3780
caaatggcat ctttaagggc tccacagctg aaaatgcaga atacctagg gtcgcgccac 3840
aaagcagtg atttattgga gcatga 3866

```

<211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 555
 5 caggactaca gaaatgtagg tttc 24
 <210> 556
 <211> 26
 <212> ADN
 <213> Homo sapiens
 10 <400> 556
 gtgcctgcct taagtaatgt gatgac 26
 <210> 557
 <211> 25
 <212> ADN
 15 <213> Homo sapiens
 <400> 557
 gactggaagt gtcgcatcac caatg 25
 <210> 558
 <211> 25
 20 <212> ADN
 <213> Homo sapiens
 <400> 558
 ggtttaataa tgcgatctgg gacac 25
 <210> 559
 25 <211> 25
 <212> ADN
 <213> Homo sapiens
 <400> 559
 gcagctataa ttagagaac caagg 25
 30 <210> 560
 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 560
 35 aaaattgact tcatttccat g 21
 <210> 561
 <211> 22

<212> ADN
 <213> Homo sapiens
 <400> 561
 cctagttgct ctaaaactaa cg 22
 5 <210> 562
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 562
 10 ctgtgaggcg tgacagccgt gcag 24
 <210> 563
 <211> 25
 <212> ADN
 <213> Homo sapiens
 15 <400> 563
 caacctacta atcagaacca gcatc 25
 <210> 564
 <211> 24
 <212> ADN
 20 <213> Homo sapiens
 <400> 564
 ccttcactgt gtctgcaa at ctgc 24
 <210> 565
 <211> 24
 25 <212> ADN
 <213> Homo sapiens
 <400> 565
 cctgtcataa gtctcctgt tgag 24
 <210> 566
 30 <211> 25
 <212> ADN
 <213> Homo sapiens
 <400> 566
 cagtctgtgg gtctaagagc taatg 25
 35 <210> 567
 <211> 24
 <212> ADN

<213> Homo sapiens
 <400> 567
 caggaatggg tgagtctctg tgtg 24
 <210> 568
 5 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 568
 gtggaattct gcccaggcct ttc 23
 10 <210> 569
 <211> 25
 <212> ADN
 <213> Homo sapiens
 <400> 569
 15 gattctacaa accagccagc caaac 25
 <210> 570
 <211> 24
 <212> ADN
 <213> Homo sapiens
 20 <400> 570
 cctactggtt cacatctgac cctg 24
 <210> 571
 <211> 24
 <212> ADN
 25 <213> Homo sapiens
 <400> 571
 gtttgaatgt ggttcgtg gaag 24
 <210> 572
 <211> 27
 30 <212> ADN
 <213> Homo sapiens
 <400> 572
 cttgtcacc aggcagaggg caatac 27
 <210> 573
 35 <211> 23
 <212> ADN
 <213> Homo sapiens

<400> 573
 gacagtaact tgggcttct gac 23
 <210> 574
 <211> 23
 5 <212> ADN
 <213> Homo sapiens
 <400> 574
 catccacca aagactctcc aag 23
 <210> 575
 10 <211> 25
 <212> ADN
 <213> Homo sapiens
 <400> 575
 ctgttcatat aatacagagt ccctg 25
 15 <210> 576
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 576
 20 gagagatgca ggagctctgt gc 22
 <210> 577
 <211> 25
 <212> ADN
 <213> Homo sapiens
 25 <400> 577
 gcagtttgta gtcaatcaaa ggtgg 25
 <210> 578
 <211> 22
 <212> ADN
 30 <213> Homo sapiens
 <400> 578
 gtaatttaaa tgggaatagc cc 22
 <210> 579
 <211> 24
 35 <212> ADN
 <213> Homo sapiens
 <400> 579

caactccttg accattacct caag 24
 <210> 580
 <211> 23
 <212> ADN
 5 <213> Homo sapiens
 <400> 580
 gatggccgtc ctgccacac agg 23
 <210> 581
 <211> 21
 10 <212> ADN
 <213> Homo sapiens
 <400> 581
 gagtagtta gcatatattg c 21
 <210> 582
 15 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 582
 gacagtcaga aatgcaggaa agc 23
 20 <210> 583
 <211> 26
 <212> ADN
 <213> Homo sapiens
 <400> 583
 25 caagtgccgt gtctggcac ccaagc 26
 <210> 584
 <211> 24
 <212> ADN
 <213> Homo sapiens
 30 <400> 584
 ccaaacactc agtgaacaa agag 24
 <210> 585
 <211> 21
 <212> ADN
 35 <213> Homo sapiens
 <400> 585
 ccttaggtgc ggctccacag c 21

<210> 586
 <211> 22
 <212> ADN
 <213> Homo sapiens
 5 <400> 586
 catttaggat gtggagatga gc 22
 <210> 587
 <211> 23
 <212> ADN
 10 <213> Homo sapiens
 <400> 587
 gaaactcaag atcgattca tgc 23
 <210> 588
 <211> 23
 15 <212> ADN
 <213> Homo sapiens
 <400> 588
 gcaaactctt gctatcccag gag 23
 <210> 589
 20 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 589
 cagccataag tcctcgacgt gg 22
 25 <210> 590
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 590
 30 catcctcccc tgcattgtt aaac 24
 <210> 591
 <211> 24
 <212> ADN
 <213> Homo sapiens
 35 <400> 591
 gtaggtttct aaacatcaag aaac 24
 <210> 592

<211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 592
 5 gtgatgacat ttctccaggg atgc 24
 <210> 593
 <211> 23
 <212> ADN
 <213> Homo sapiens
 10 <400> 593
 catcaccaat gccttctta agc 23
 <210> 594
 <211> 24
 <212> ADN
 15 <213> Homo sapiens
 <400> 594
 gctggagggt ttaataatgc gatc 24
 <210> 595
 <211> 24
 20 <212> ADN
 <213> Homo sapiens
 <400> 595
 gcaaacacac aggcacctgc tggc 24
 <210> 596
 25 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 596
 catgtgagtt tcactagatg g 21
 30 <210> 597
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 597
 35 gaccggacga caggccacct cgtc 24
 <210> 598
 <211> 25

<212> ADN
 <213> Homo sapiens
 <400> 598
 gaagaacgaa acgtcccgtt cctcc 25
 5 <210> 599
 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 599
 10 gttgagcact cgtgtgcatt agg 23
 <210> 600
 <211> 22
 <212> ADN
 <213> Homo sapiens
 15 <400> 600
 ctcaagtgcac gtgtactggg ta 22
 <210> 601
 <211> 35
 <212> ADN
 20 <213> Homo sapiens
 <400> 601
 gttcactggg ctaattgcgg gactcttgtt cgcac 35
 <210> 602
 <211> 28
 25 <212> ADN
 <213> Homo sapiens
 <400> 602
 ggtaaataca tgctttcta gtggtcag 28
 <210> 603
 30 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 603
 ggaggatgga gcctttccat cac 23
 35 <210> 604
 <211> 24
 <212> ADN

<213> Homo sapiens
 <400> 604
 gaagaggaag atgtgttcct ttgg 24
 <210> 605
 5 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 605
 gaatgaagga tgatgtggca gtgg 24
 10 <210> 606
 <211> 25
 <212> ADN
 <213> Homo sapiens
 <400> 606
 15 gtatgtgtga aggagtcact gaaac 25
 <210> 607
 <211> 22
 <212> ADN
 <213> Homo sapiens
 20 <400> 607
 ggtgagtcac aggttcagtt gc 22
 <210> 608
 <211> 21
 <212> ADN
 25 <213> Homo sapiens
 <400> 608
 caaaacatca gccattaacg g 21
 <210> 609
 <211> 22
 30 <212> ADN
 <213> Homo sapiens
 <400> 609
 gtagccagca tgtctgtgtc ac 22
 <210> 610
 35 <211> 23
 <212> ADN
 <213> Homo sapiens

<400> 610
 cagaatgcct gtaaagctat aac 23
 <210> 611
 <211> 23
 5 <212> ADN
 <213> Homo sapiens
 <400> 611
 catttggtt tccccactca cac 23
 <210> 612
 10 <211> 27
 <212> ADN
 <213> Homo sapiens
 <400> 612
 gaccaaaca ccttaagtaa ctgactc 27
 15 <210> 613
 <211> 25
 <212> ADN
 <213> Homo sapiens
 <400> 613
 20 gaagctacat agtgtctcac ttcc 25
 <210> 614
 <211> 25
 <212> ADN
 <213> Homo sapiens
 25 <400> 614
 cacaactgct aatggcccg tctcg 25
 <210> 615
 <211> 23
 <212> ADN
 30 <213> Homo sapiens
 <400> 615
 gctcctgctc cctgtcataa gtc 23
 <210> 616
 <211> 25
 35 <212> ADN
 <213> Homo sapiens
 <400> 616

gaagtcctgc tggtagtcag ggttg 25
 <210> 617
 <211> 24
 <212> ADN
 5 <213> Homo sapiens
 <400> 617
 ctgcagtggg caaccccgag tatc 24
 <210> 618
 <211> 20
 10 <212> ADN
 <213> Homo sapiens
 <400> 618
 tgtgggtcta agagctaatg 20
 <210> 619
 15 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 619
 gacaggccac ctgctcggcg tc 22
 20 <210> 620
 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 620
 25 cagctgatct caaggaaaca gg 22
 <210> 621
 <211> 25
 <212> ADN
 <213> Homo sapiens
 30 <400> 621
 ctggtgtgca ttagggttca actgg 25
 <210> 622
 <211> 26
 <212> ADN
 35 <213> Homo sapiens
 <400> 622
 ccttctccga ggtggaattg agtgac 26

<210> 623
 <211> 26
 <212> ADN
 <213> Homo sapiens
 5 <400> 623
 gctaattgcg ggactcttgt tcgcac 26
 <210> 624
 <211> 22
 <212> ADN
 10 <213> Homo sapiens
 <400> 624
 tacatgcttt tctagtggc ag 22
 <210> 625
 <211> 23
 15 <212> ADN
 <213> Homo sapiens
 <400> 625
 cctttccatc acccctcaag agg 23
 <210> 626
 20 <211> 25
 <212> ADN
 <213> Homo sapiens
 <400> 626
 gatgtgtcc ttggagggtg gcatg 25
 25 <210> 627
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 627
 30 gatgtggcag tggcggttcc ggtg 24
 <210> 628
 <211> 24
 <212> ADN
 <213> Homo sapiens
 35 <400> 628
 ggagtcactg aaacaaaca cagg 24
 <210> 629

<211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 629
 5 gggtcagttg ctgtataaa g 21
 <210> 630
 <211> 24
 <212> ADN
 <213> Homo sapiens
 10 <400> 630
 ccattaacgg taaaattca gaag 24
 <210> 631
 <211> 20
 <212> ADN
 15 <213> Homo sapiens
 <400> 631
 ccaaggtcat ggagcacagg 20
 <210> 632
 <211> 25
 20 <212> ADN
 <213> Homo sapiens
 <400> 632
 ctgtaaagct ataacaaca cctgg 25
 <210> 633
 25 <211> 27
 <212> ADN
 <213> Homo sapiens
 <400> 633
 ccactcacac acactaaata ttttaag 27
 30 <210> 634
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 634
 35 gtaactgact caaatacaaa ccac 24
 <210> 635
 <211> 25

<212> ADN
 <213> Homo sapiens
 <400> 635
 gaagctacat agtgtctcac ttcc 25
 5 <210> 636
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 636
 10 ctgctaattgg cccgttctcg 20
 <210> 637
 <211> 24
 <212> ADN
 <213> Homo sapiens
 15 <400> 637
 cctgtcataa gtctccttgt tgag 24
 <210> 638
 <211> 20
 <212> ADN
 20 <213> Homo sapiens
 <400> 638
 ggtagtcagg gttgtccagg 20
 <210> 639
 <211> 23
 25 <212> ADN
 <213> Homo sapiens
 <400> 639
 cgagtatctc aacactgtcc agc 23
 <210> 640
 30 <211> 25
 <212> ADN
 <213> Homo sapiens
 <400> 640
 ctaagagcta atgcgggcat ggctg 25
 35 <210> 641
 <211> 372
 <212> ADN

<213> Homo sapiens

<400> 641

```
gcaatatcag ccttaggtgc ggctccacag cccagtggtc cctcaccttc ggggtgcac 60
gctggtaaca tccacccaga tcaactggca gcatgtggca ccatctcaca attgccagtt 120
aacgtcttcc ttctctctct gtcataaggga ctctggatcc cagaaggtga gaaagttaa 180
attcccgtcg ctatcaagga attaagagaa gcaacatctc cgaaagccaa caaggaaatc 240
ctcgatgtga gtttctgctt tgctgtgtgg ggggtccatgg ctctgaacct caggcccacc 300
ttttctcatg tctggcagct gctctgctct agaccctgct catctccaca tcctaaatgt 360
tcactttcta tg 372
```

<210> 642

5 <211> 415

<212> ADN

<213> Homo sapiens

<400> 642

```
ctaacgttcg ccagccataa gtcctcgacg tggagaggct cagagcctgg catgaacatg 60
accctgaatt cggatgcaga gcttcttccc atgatgatct gtccctcaca gcagggtctt 120
ctctgtttca gggcatgaac tacttgaggg 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 ggggtgagctc gcagc 415
```

10 <210> 643

<211> 35

<212> ADN

<213> Homo sapiens

<400> 643

15 tatcaaggaa ttaagagaag caacatctcc gaaag 35

<210> 644

<211> 24

<212> ADN

<213> Homo sapiens

20 <400> 644

cgaaagccaa caaggaaatc ctcg 24

<210> 645

<211> 18

<212> ADN

25 <213> Homo sapiens

<400> 645

tgtaaaacga cggccagt 18

<210> 646

<211> 22

30 <212> ADN

<213> Homo sapiens
 <400> 646
 tggcttcaca ggaccactga tt 22
 <210> 647
 5 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 647
 gaggccagtg ctgtctctaa gg 22
 10 <210> 648
 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 648
 15 atgggacagg cactgatttg t 21
 <210> 649
 <211> 22
 <212> ADN
 <213> Homo sapiens
 20 <400> 649
 cagctctggc tcacactacc ag 22
 <210> 650
 <211> 20
 <212> ADN
 25 <213> Homo sapiens
 <400> 650
 gcagctggac tcgatttcct 20
 <210> 651
 <211> 22
 30 <212> ADN
 <213> Homo sapiens
 <400> 651
 tgcccaatga gtcaagaagt gt 22
 <210> 652
 35 <211> 22
 <212> ADN
 <213> Homo sapiens

<400> 652
 cactcacgga tgctgcttag tt 22
 <210> 653
 <211> 22
 5 <212> ADN
 <213> Homo sapiens
 <400> 653
 tcagagcctg tgtttctacc aa 22
 <210> 654
 10 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 654
 aaataatcag tgtgattcgt ggag 24
 15 <210> 655
 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 655
 20 acttcacagc cctgcgtaaa c 21
 <210> 656
 <211> 22
 <212> ADN
 <213> Homo sapiens
 25 <400> 656
 gcagcgggtt acatcttctt tc 22
 <210> 657
 <211> 22
 <212> ADN
 30 <213> Homo sapiens
 <400> 657
 cctgaactcc gtcagactga aa 22
 <210> 658
 <211> 23
 35 <212> ADN
 <213> Homo sapiens
 <400> 658

ccttacagca atcctgtgaa aca 23
 <210> 659
 <211> 22
 <212> ADN
 5 <213> Homo sapiens
 <400> 659
 atgtacagtg ctggcatggt ct 22
 <210> 660
 <211> 20
 10 <212> ADN
 <213> Homo sapiens
 <400> 660
 tccaaatgag ctggcaagtg 20
 <210> 661
 15 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 661
 gtgcatcgct ggtaacatcc 20
 20 <210> 662
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 662
 25 atcgattca tgcgtcttca 20
 <210> 663
 <211> 19
 <212> ADN
 <213> Homo sapiens
 30 <400> 663
 gctcagagcc tggcatgaa 19
 <210> 664
 <211> 20
 <212> ADN
 35 <213> Homo sapiens
 <400> 664
 tggctcgtct gtgtgtgtca 20

<210> 665
 <211> 20
 <212> ADN
 <213> Homo sapiens
 5 <400> 665
 tgaagcaaat tgccaagac 20
 <210> 666
 <211> 20
 <212> ADN
 10 <213> Homo sapiens
 <400> 666
 aagtgtcgca tcaccaatgc 20
 <210> 667
 <211> 23
 15 <212> ADN
 <213> Homo sapiens
 <400> 667
 tcccaaacac tcagtgaac aaa 23
 <210> 668
 20 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 668
 tgtggagatg agcagggct 20
 25 <210> 669
 <211> 20
 <212> ADN
 <213> Homo sapiens
 <400> 669
 30 atcccatgg caaactctg 20
 <210> 670
 <211> 19
 <212> ADN
 <213> Homo sapiens
 35 <400> 670
 catcctccc tgcattgt 19
 <210> 671

<211> 25
 <212> ADN
 <213> Homo sapiens
 <400> 671
 5 cgaaagaaaa tacttgcattg tcaga 25
 <210> 672
 <211> 20
 <212> ADN
 <213> Homo sapiens
 10 <400> 672
 tgacatttct ccagggatgc 20
 <210> 673
 <211> 19
 <212> ADN
 15 <213> Homo sapiens
 <400> 673
 atgcgatctg ggacacagg 19
 <210> 674
 <211> 16
 20 <212> ADN
 <213> Homo sapiens
 <400> 674
 aacagctatg accatg 16
 <210> 675
 25 <211> 26
 <212> ADN
 <213> Homo sapiens
 <400> 675
 caagtgccgt gtcctggcac ccaagc 26
 30 <210> 676
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 676
 35 ccaaacactc agtgaacaa agag 24
 <210> 677
 <211> 25

<212> ADN
 <213> Homo sapiens
 <400> 677
 gcaccaagc ccatgccgtg gctgc 25
 5 <210> 678
 <211> 26
 <212> ADN
 <213> Homo sapiens
 <400> 678
 10 gaaacaaaga gtaaagtaga tgatgg 26
 <210> 679
 <211> 24
 <212> ADN
 <213> Homo sapiens
 15 <400> 679
 caccttcaca atataccctc catg 24
 <210> 680
 <211> 22
 <212> ADN
 20 <213> Homo sapiens
 <400> 680
 gacagccgtg cagggaaaaa cc 22
 <210> 681
 <211> 24
 25 <212> ADN
 <213> Homo sapiens
 <400> 681
 gaaccagcat ctcaaggaga tctc 24
 <210> 682
 30 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 682
 gagcacctgg cttggacact ggag 24
 35 <210> 683
 <211> 26
 <212> ADN

<213> Homo sapiens
 <400> 683
 gagcagccct gaactccgtc agactg 26
 <210> 684
 5 <211> 26
 <212> ADN
 <213> Homo sapiens
 <400> 684
 ctcaagtacaa tagatagaca gcaatg 26
 10 <210> 685
 <211> 25
 <212> ADN
 <213> Homo sapiens
 <400> 685
 15 gacgggtcct ggggtgatct ggctc 25
 <210> 686
 <211> 26
 <212> ADN
 <213> Homo sapiens
 20 <400> 686
 ctcaagtacaa tagatagaca gcaatg 26
 <210> 687
 <211> 26
 <212> ADN
 25 <213> Homo sapiens
 <400> 687
 cgtggagagg ctcaagcct ggcattg 26
 <210> 688
 <211> 21
 30 <212> ADN
 <213> Homo sapiens
 <400> 688
 cacagatttt gggcgggcca a 21
 <210> 689
 35 <211> 21
 <212> ADN
 <213> Homo sapiens

<400> 689
gctatcaaaa catctccgaa a 21
<210> 690
<211> 35
5 <212> ADN
<213> Homo sapiens
<400> 690
tatcaaggaa ttaagagaag caacatctcc gaaag 35
<210> 691
10 <211> 29
<212> ADN
<213> Homo sapiens
<400> 691
agttaaaatt cccgtcgcta tcaaggaat 29
15 <210> 692
<211> 35
<212> ADN
<213> Homo sapiens
<400> 692
20 tatcaaggaa ttaagagaag caacatctcc gaaag 35
<210> 693
<211> 29
<212> ADN
<213> Homo sapiens
25 <400> 693
aacatctccg aaagccaaca aggaaatcc 29
<210> 694
<211> 35
<212> ADN
30 <213> Homo sapiens
<400> 694
tatcaaggaa ttaagagaag caacatctcc gaaag 35
<210> 695
<211> 21
35 <212> ADN
<213> Homo sapiens
<400> 695

taaaattccc gtcgctatca a 21
 <210> 696
 <211> 35
 <212> ADN
 5 <213> Homo sapiens
 <400> 696
 tatcaaggaa ttaagagaag caacatctcc gaaag 35
 <210> 697
 <211> 25
 10 <212> ADN
 <213> Homo sapiens
 <400> 697
 ccattctcga aagccaacaa ggaaa 25
 <210> 698
 15 <211> 35
 <212> ADN
 <213> Homo sapiens
 <400> 698
 tatcaaggaa ttaagagaag caacatctcc gaaag 35
 20 <210> 699
 <211> 23
 <212> ADN
 <213> Homo sapiens
 <400> 699
 25 aattcccgctc gctatcaagg aac 23
 <210> 700
 <211> 15
 <212> ADN
 <213> Homo sapiens
 30 <400> 700
 gggctggcca aactg 15
 <210> 701
 <211> 15
 <212> ADN
 35 <213> Homo sapiens
 <400> 701
 gggcgggcca aactg 15

<210> 702
 <211> 15
 <212> ADN
 <213> Homo sapiens
 5 <400> 702
 gggctggcca aactg 15
 <210> 703
 <211> 15
 <212> ADN
 10 <213> Homo sapiens
 <400> 703
 gggctggcca aacag 15
 <210> 704
 <211> 27
 15 <212> ADN
 <213> Homo sapiens
 <400> 704
 gatcaaagtg ctgggctccg gtgcgtt 27
 <210> 705
 20 <211> 27
 <212> ADN
 <213> Homo sapiens
 <220>
 <221> base_modificada
 25 <222> (14)
 <223> a, t, c, g, otro o desconocido
 <400> 705
 gatcaaagtg ctgngctccg gtgcgtt 27
 <210> 706
 30 <211> 31
 <212> ADN
 <213> Homo sapiens
 <400> 706
 tcacagattt tgggctggcc aaactgctgg g 31
 35 <210> 707
 <211> 31
 <212> ADN

<213> Homo sapiens
 <220>
 <221> base_modificada
 <222> (16)
 5 <223> a, t, c, g, otro o desconocido
 <400> 707
 tcacagattt tgggcnggcc aaactgctgg g 31
 <210> 708
 <211> 21
 10 <212> ADN
 <213> Homo sapiens
 <400> 708
 tatcaaggaa ttaagagaag c 21
 <210> 709
 15 <211> 21
 <212> ADN
 <213> Homo sapiens
 <220>
 <221> base_modificada
 20 <222> (17)
 <223> a, t, c, g, otro o desconocido
 <400> 709
 tatcaaggaa ttaaganaag c 21
 <210> 710
 25 <211> 22
 <212> ADN
 <213> Homo sapiens
 <400> 710
 tatcaaggaa ttaagagaag ca 22
 30 <210> 711
 <211> 27
 <212> ADN
 <213> Homo sapiens
 <400> 711
 35 cgtcgctatc aaggaattaa gagaagc 27
 <210> 712
 <211> 20

<212> ADN
 <213> Homo sapiens
 <400> 712
 cgctatcaag gaacaacatc 20
 5 <210> 713
 <211> 19
 <212> ADN
 <213> Homo sapiens
 <400> 713
 10 sgatcaygga ascaacatc 19
 <210> 714
 <211> 28
 <212> ADN
 <213> Homo sapiens
 15 <400> 714
 aggaattaag agaagcaaca tctccgaa 28
 <210> 715
 <211> 28
 <212> ADN
 20 <213> Homo sapiens
 <400> 715
 aggaattaataa agaaccaca tcttcat 28
 <210> 716
 <211> 28
 25 <212> ADN
 <213> Homo sapiens
 <400> 716
 aggaaccaag agaagccacg tattcgaa 28
 <210> 717
 30 <211> 37
 <212> ADN
 <213> Homo sapiens
 <400> 717
 tatcaaggaa ttaagagaag caacatctcc gaaagcc 37
 35 <210> 718
 <211> 19
 <212> ADN

<213> Homo sapiens
 <400> 718
 tatcaagggtt ccgaaagcc 19
 <210> 719
 5 <211> 19
 <212> ADN
 <213> Homo sapiens
 <400> 719
 tatcaagggtt ccgaaagcc 19
 10 <210> 720
 <211> 3867
 <212> ADN
 <213> Homo sapiens
 <400> 720
 cccggcgagc cgccggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60
 ccgagggcggc cggaggtccc agctagcccc ggccggccgcc gccgcccaga ccggacgcaca 120
 ggccacctcg tcggcggtccg cccgagtcgc cgctcgccg ccaacgccac aaccaccgcg 180
 cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
 cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
 gcccgccgag tcgggctctg gaggaaga aagtttgcca aggcacgag aacaagctca 360
 cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
 aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
 taaagaccat ccaggaggtg gctgggttatg tcctcattgc cctcaacaca gtggagcgaa 540
 ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
 tagcagtcct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
 atttacagga aatcctgcat ggccgctgac ggttcagcaa caaccctgcc ctgtgcaacg 720
 tggagagcat ccagtgccgg gacatagtc gacgtgactt tctcagcaac atgtcgatgg 780
 acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
 gctgggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccagcagct 900
 gctccggggc ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960
 gctgcacagg ccccggggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
 cgtgcaagga cacctgcccc cactcatgac tctacaaccc caccacgtac cagatggatg 1080
 tgaaccccg gggcaatac agctttggtg ccacctgctg gaagaagtgt ccccgtaatt 1140
 atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
 aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
 gaataggat tgggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
 tcaaaaactg cacctccatc agtggcgatc tccacatcct gccgggtggca tttaggggtg 1380
 actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
 15 aggaaatcac aggggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500

```

cctttgagaa cctagaaatc atacgcggca ggaccaagca acatgggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aattttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaaacca aaattataag caacagaggt gaaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggtctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgctg caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctgggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgtcc atcgccactg ggatggtggg ggccctctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccgagggt gctgcaggag agggagcttg tggagcctct tacaccagt ggagaagctc 2340
ccaaccaagc tctcttgagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat ccgaaagcca acaaggaaat cctcgatgaa 2460
aaattccgtg cgctatcaag gaatacatct ccgaaagcca acaggtgctg 2520
gcctacgtga tggccagcgt ggacaacccc cactgtgctc gcctgctggg catctgcctc 2580
acctccaccg tgcagctcat cagcagctc atgcccctcg gctgcctcct ggactatgtc 2640
cggggaacaca aagacaatat tggctccag tacctgctca actgggtgtg gcagatcgca 2700
aagggcatag actacttgga ggaccgtcgc ttggtgcacc gcgacctggc agccaggaac 2760
gtactggtga aaacaccgca gcatgtcaag atcacagatt ttgggctggc caaactgctg 2820
ggtgcggaag agaagaata ccatgcagaa ggaggcaaa tgcctatcaa gtggatggca 2880
ttggaatcaa ttttacacag aatctatacc caccagagt atgtctggag ctacggggtg 2940
actgtttggg agttgatgac ctttggatcc aagccatatg acggaatccc tgccagcgag 3000
atctcctcca tctggagaa aggagaacgc ctccctcagc caccatattg taccatgat 3060
gtctacatga tcatggtcaa gtgctggatg atagacgcag atagtgcgcc aaagtccgt 3120
gagttgatca tcgaattctc caaaatggcc cgagaccccc agcgctacct tgtcattcag 3180
ggggatgaaa gaatgcattt gccaaagtct acagactcca acttctaccg tgccctgatg 3240
gatgaagaag acatggacga cgtggtggat gccgacgagt acctcatccc acagcagggc 3300
ttcttcagca gccctccac gtcacggact cccctcctga gctctctgag tgcaaccagc 3360
aacaattcca ccgtggcttg cattgataga aatgggctgc aaagctgtcc catcaaggaa 3420
gacagcttct tgcagcgata cagctcagac ccacagggcg ccttgactga ggacagcata 3480
gacgacacct tctccagct gcctgaatac ataaaccagt ccgttcccaa aaggcccgct 3540
ggctctgtgc agaactctgt ctatcacaat cagcctctga accccgcgcc cagcagagac 3600
ccacactacc aggaccccca cagcactgca gtgggcaacc ccgagtatct caacactgtc 3660
cagcccacct gtgtcaacag cacattcgac agccctgccc actgggcca gaaaggcagc 3720
caccaaatta gcctggacaa ccctgactac cagcaggact tctttcccaa ggaagccaag 3780
ccaaatggca tctttaaggc ctccacagct gaaaatgcag aatacctaag ggtcgcgcca 3840
caaagcagtg aatttattgg agcatga 3867

```

<210> 721

<211> 3863

<212> ADN

5 <213> Homo sapiens

<400> 721

```

cccggcgag cgccggcgca gcagcctccg cccccgcac ggtgtgagcg cccgacgagg 60
ccgaggcgcc cggagtcccg agctagcccc ggccggccgc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctcc cccgagtccc cgctcgccg ccaacgccac aaccaccgag 180
cacggcccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gcccgcgag tcgggctctg gaggaaaaga aagtgtgcca aggcacgagt aacaagctca 360
cgcatgtggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttctctct 480
taaagaccat ccaggagggt gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctct atctaactat gatgcaataa aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgcgctgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tgagagagcat ccagtggcgg gacatagtca gcagtactt tctcagcaac atgtcgatgg 780
acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aatcatctgt gccacgagct 900
gctccggggc ctgccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccgaggag agcagctgcc tggctctgcc caaattccga gacgaagcca 1020

```

cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
 tgaaccccgga gggcaaatat agctttggtg ccacctgctg gaagaagtgt ccccgtaatt 1140
 atgtggtgac agatcacggc tcgtgctgct gagcctgtgg ggccgacagc tatgagatgg 1200
 aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaa gtgtgtaacg 1260
 gaatagggtat tgggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
 tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
 actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
 aggaaatcac aggggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
 cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggctag ttttctcttg 1560
 cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
 gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
 aactgttttg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740
 gcaagggcac agggcagggtc tgccatgcct tgtgctcccc cgagggtgctg tggggcccg 1800
 agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
 gcaaccttct ggaagggtgag ccaagggtg tttgtggaga ctctgagtgc atacagtgc 1920
 acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
 gtatccagtg tgcccactac attgacggcc cccactgctg caagacctgc ccggcaggag 2040
 tcatgggaga aaacaacacc ctgggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
 tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
 atgggacctaa gatcccgctc atcgccactg ggatgggtgg ggccctctc ttgctgctgg 2220
 tgggtggcct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgacgc 2280
 tgccgagggt gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340
 ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
 gctccggtgc gttcggcacg gtgtataagg gactctggat ccagaaggt gagaaagtta 2460
 aaattcccg cgtatcaag gaattaagag aagcaacatc tgaaatctc gatgaagcct 2520
 acgtgatggc cagcgtggac aacccccacg tgtgccgct gctgggcac tgccctcacct 2580
 ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640
 aacacaaaga caatattggc tccagtagc tgctcaactg gtgtgtgcag atcgcaagg 2700
 gcataacta cttggaggac cgtcgcttgg tgcaccgca cctggcagcc aggaacgtac 2760
 tgggtgaaac accgcagcat gtcaagatca cagatttttg gctggccaaa ctgctgggtg 2820
 cggaagagaa agaataccat gcagaaggag gcaaagtgc tatcaagtgg atggcattgg 2880
 aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940
 tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000
 cctccatcct ggagaaagga gaacgcctcc ctacgccacc catatgtacc atcgatgtct 3060
 acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccaaag ttccgtgagt 3120
 tgatcatcga attctccaaa atggcccag agccccagcg ctacctgtc attcaggggg 3180
 atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240
 aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggtctct 3300
 tcagcagccc ctccacgtca cggactcccc tctgagctc tctgagtgc accagcaaca 3360
 attccaccgt ggcttgcat gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420
 gcttcttgca gcgatacagc tcagacccca caggcgctt gactgaggac agcatagacg 3480
 acaccttct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgtgggt 3540
 ctgtgcagaa tctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagaccac 3600
 actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac actgtccagc 3660
 ccactgtgt caacagcaca ttcgacagcc ctgcccactg ggccagaaa ggagccacc 3720
 aaattagcct ggacaaccct gactaccagc aggacttct tcccaaggaa gccaaagcaa 3780
 atggcatct taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840
 gcagtgaatt tattggagca tga 3863

<210> 722

<211> 3886

<212> ADN

5 <213> Homo sapiens

<400> 722

cccggcgag cgcggccgca gcagcctccg cccccgcac ggtgtgagcg cccgacggg 60
 ccgaggcggc cggagtcccg agctagcccc ggcgcccgcc gccgccaga cccgacgaca 120
 ggccacctcg tcggcgctcc cccgagtcct cgctcgccg ccaacgccac aaccaccgag 180
 cacggcccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
 cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
 gccggcgag tcgggctctg gaggaaaaga agtttgcca aggcacgagt aacaagctca 360
 cgcagtggg cacttttgaa gatcatttct tcagcctcca gaggatgttc aataactgtg 420
 aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttcttct 480
 taaagaccat ccaggagggt gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540

ttccttttga	aaacctgcag	atcatcagag	gaaatatgta	ctacgaaaat	tcctatgcct	600
tagcagtcct	atctaactat	gatgcaaata	aaaccggact	gaaggagctg	cccatgagaa	660
atttacagga	aatcctgcat	ggcgccgtgc	ggttcagcaa	caaccctgcc	ctgtgcaacg	720
tggagagcat	ccagtggcgg	gacatagtc	gcagtgactt	tctcagcaac	atgtcgatgg	780
acttccagaa	ccacctgggc	agctgccaaa	agtgatgatcc	aaagctgtccc	aatgggagct	840
gctgggggtgc	aggagaggag	aactgccaga	aactgaccaa	aatcatctgt	gcccagcagt	900
gctccggggcg	ctgccgtggc	aagtccccc	gtgactgctg	ccacaaccag	tgtgtgtgcag	960
gctgcacagg	ccccggggag	agcgactgcc	tggctgtgccg	caaattccga	gacgaagcca	1020
cgtgcaagga	cacctgcccc	ccactcatgc	tctacaaccc	caccacgtac	cagatggatg	1080

tgaaccccca	gggcaaatat	agcttttggg	ccacctgcgt	gaagaagtgt	ccccgtaatt	1140
atgtggtgac	agatcacggc	tcgtgcgtcc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaagacgg	cgtccgcgaag	tgtgaagaat	gcgaagggcc	ttgccgcaaa	gtgtgtaacg	1260
gaataggtat	tggtgaattt	aaagactcac	tctccataaa	tgtacgaat	attaaacact	1320
tcaaaaactg	cacctccatc	agtggcgatc	tccacatcct	gccggtggca	tttaggggtg	1380
actccttcac	acatactcct	cctctggatc	cacaggaact	ggatatctctg	aaaaccgtaa	1440
aggaaatcac	agggtttttg	ctgattcagg	cttggcctga	aaacaggacg	gacctccatg	1500
ccttttgagaa	cctagaaatc	atacgcgga	ggaccaagca	acatggtcag	ttttctcttg	1560
cagtgcgtcag	cctgaacata	acatccttgg	gattacgctc	cctcaaggag	ataagtgatg	1620
gagatgtgat	aattttcagga	aacaaaaatt	tgtgctatgc	aaatacaata	aactggaaaa	1680
aactgtttgg	gacctccggg	cagaaaacca	aaattataag	caacagaggt	gaaaacagct	1740
gcaagggccac	aggccagggtc	tgccatgcct	tgtgctcccc	cgaggggtgc	tggggcccg	1800
agcccaggga	ctgcgtctct	tgccggaatg	tcagccgagg	caggggaatgc	gtggacaagt	1860
gcaacccttct	ggaggggtgag	ccaagggagt	ttgtggagaa	ctctgagtgc	atacagtgcc	1920
accagagtg	cctgcctcag	gccatgaaca	tcacctgcac	aggacgggga	ccagacaact	1980
gtatccagt	tgccactac	attgacggcc	cccactgcgt	caagacctgc	ccggcaggag	2040
tcattgggaga	aaacaacacc	ctggcttgga	agtacgcaga	cgccggccat	gtgtgccacc	2100
tgtgcccatt	aaactgcacc	tacggatgca	ctgggcccag	tcttgaaggc	tgtccaacga	2160
atgggcctaa	gatcccgtcc	atcgccactg	ggatgggtgg	ggccctcctc	ttgctgctgg	2220
tggtggccct	ggggatcgcc	ctcttcatgc	gaaggcgcca	catcgttcgg	aagcgcacgc	2280
tgccgaggct	gctgcaggag	agggagcttg	tggagcctct	tacaccagct	ggagaagctc	2340
ccaaccaagc	tctcttgagg	atcttgaagg	aaactgaatt	caaaaagatc	aaagtgtctg	2400
gctccgggtg	gttcggcacg	gtgtataagg	gactctggat	cccagaaggt	gagaaagtta	2460
aaattcccg	cgctatcaag	gaattaaag	aagcaacatc	tccgaaagcc	aacaaggaaa	2520
tcctcgatga	agcctacgtg	atggccagcg	tggacaaccc	ggcaaccccc	acgtgtgccg	2580
cctgtctggg	atctgcctca	cctccaccgt	gcagctcatc	acgcagctca	tgcccttcgg	2640
ctgctcctg	gactatgtcc	gggaacacaa	agacaatatt	ggctcccagt	acctgctcaa	2700
ctggtgtgtg	gccagcgaga	agggcatgaa	ctacttggag	gaccgtcgct	tgggtgcaccg	2760
cgacctggga	gccaggaacg	tactggtgaa	aacaccgcag	catgtcaaga	tcacagattt	2820
tgggctggcc	aaactgctgg	gtgcggaaga	gaaagaatac	catgcagaag	gaggcaaggt	2880
gcctatcaag	tggatggcat	tggaaatcaat	tttacacaga	atctataccc	accagagtga	2940
tgtctggagc	tacggggtga	ctgtttggga	gttgatgacc	tttggatcca	agccatatga	3000
cggaatccct	gccagcgaga	tctcctccat	cctggagaaa	ggagaacgcc	tccctcagcc	3060
accatcatgt	accatcgatg	tctacatgat	catggtcaag	tgtggtgatga	tagacgcaga	3120
tagtcgccc	aagttccgtg	agttgatcat	cgaattctcc	aaaatggccc	gagaccccc	3180
gcgtacctt	gtcattcagg	gggatgaaag	aatgcatttg	ccaagtccta	cagactccaa	3240
cttctaccgt	gccctgatgg	atgaagaaga	catggacgac	gtggtggatg	ccgacgagta	3300
cctcatccca	cagcagggtc	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	tatcacatc	agcctctgaa	3600
ccccgcggcc	agcagagacc	cacactacca	ggacccccac	agcactgcag	tgggcaaccc	3660
cgagtatctc	aacactgtcc	agcccacctg	tgtcaacagc	acattcgaca	gccctgcccc	3720
ctgggcccag	aaaggcagcc	accaaattag	cctggacaac	cctgactacc	agcaggactt	3780
ctttcccaag	gaagccaagc	caaatggcat	ctttaagggc	tccacagctg	aaaatgcaga	3840
atacctaagg	gtcgcgccac	aaagcagtg	atttattgga	gcatga		3886

<210> 723

<211> 3887

<212> ADN

5 <213> Homo sapiens

<400> 723

```

ccccggcgag cgcgggcgca gcagcctccg ccccccgca ggtgtgagcg cccgacgagg 60
ccgaggcggc cggagtcctcg agctagcccc ggcgggcgcc gccgcccaga cgggacgaca 120
ggccacctcg tggcgctccg cccgagtcct cgccctcgcc ccaacgccac aaccaccgag 180
cacggcccc ctagctccgc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gcccgcgag tggggtctcg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgagttggg cacttttgaa gatcatttct tcagcctcca gaggatgttc aataactgtg 420
aggtggctct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggaggtg gctgggttat tcctcatgct cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctct atctaactat gatgcaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcccgctgc ggttcagcaa caacctgcc ctgtgcaacg 720
tgagagagcat ccagtgggcg gacatagtcg cagtgactt tctcagcaac atgtcgatgg 780
acttcagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aatcatctgt gccagcagt 900
gctccgggag ctgcccgggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccgggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
cgtgcaagga cagctcccc ccactcatgc tctacaacct caccacgtac cagatggatg 1080
tgaacccga gggcaaatc agctttgggt ccacctgctg gaagaagtgt ccccgtaatt 1140
atgtgggtgac agatcacggc tctgctgctc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tctaagaagt gcaagggcc ttgcccgaac gtgtgtaacg 1260
gaataggat tggtagaatt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaaact ggatattctg aaaaccgtaa 1440
aggaatcac agggtttttg ctgattcagg cttggcctga aaacaggagc gacctccatg 1500
cctttgagaa cctagaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactgaaaaa 1680
aactgtttgg gacctccggt cagaaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac agccaggtc tgccatgcct tgtgctcccc cgagggtgc gggggcccg 1800
agccagggga ctgctctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgc 1920
accagagtg cctgcctcag gccatgaaca tcacctgcac aggcagggga ccagacaact 1980
gtatccagtg tgccactac attgacggcc cccactgctg caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggacctaa gatccccctc atcgccactg ggatgggtgg ggccctctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccggagct gctgcaggag agggagcttg tggagcctct tacaccagc ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aagtgctgg 2400
gctccggtgc gttcgccagc gtgtataagg gactctggat ccagaaagt gagaaagtta 2460
aaattccccg cgtatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tctcgtatga agcctacgtg atggccagcg tggacagcgt ggacaacccc cacgtgtgcc 2580
gectgctggg catctgcctc acctccacc tgccagctcat cagcgagctc atgcccttcg 2640
gctgcctcct ggactatgtc cggaacaca aagacaatat tggtcccag tacctgctca 2700
actggtgtgt gcagatcgca aagggcatga actacttggg ggaccgtcgc ttggtgcacc 2760
gcgacctggc agccaggaac gtaactgtga aaacaccgca gcatgtcaag atcacagatt 2820
ttgggctggc caaactgctg ggtgcggaag agaaaagaata ccatgcagaa ggaggcaaa 2880
tgccatcaaa gtggatggca ttggaatcaa ttttacacag aatctatacc caccagagt 2940
atgtctggag ctacgggtg actgtttggg agttgatgac ctttggatcc aagccatag 3000
acggaatccc tgcccgctg atctcctcca tctggagaa aggagaacgc ctccctcagc 3060
cacccatag taccatgat gctcatatga tcatggtcaa gtgctggatg atagacgag 3120
atagtcgccc aaagtccgt gagttgatca tcgaattctc caaaatggcc cgagaccccc 3180
agcgctacct tgtcattcag ggggatgaaa gaatgcatt gccaaagctc acagactcca 3240
actctaccg tgccctgat gatgaagaag acatggacga cgtgggtgat gccagcag 3300
acctcatccc acagcagggc ttcttcagca gcccctccac gtcacggact cccctcctga 3360
gctctctgag tgcaaccaga aacaattcca ccgtggcttg cattgataga aatgggctgc 3420
aaagctgtcc catcaaggaa gacagcttct tgcagcgata cagctcagac cccacaggcg 3480
ccttgactga ggacagcata gacgacacct tcctcccagt gctgaatac ataaaccagt 3540
ccgttcccaa aaggccgct ggctctgtgc agaactcgt ctatcacaat cagcctctga 3600
accccgcgcc cagcagagac ccacactacc aggaccccca cagcactgca gttggcaacc 3660
ccgagtatct caacagtgc cagcccacct gtgtcaacag cacattcgac agccctgccc 3720
actgggccc gaaaggcagc caccaaaata gcctggacaa cctgactac cagcaggact 3780
tctttcccaa ggaagccaag ccaaatggca tctttaagg ctccacagct gaaaatgcag 3840
aatacctaag ggtcgcgcca caaagcagtg aatttatgg agcatga 3887

```

<210> 724

<211> 3881

<212> ADN

5 <213> Homo sapiens

<400> 724

```

ccccggcgag cgcggcgcca gcagcctccg ccccccgac ggtgtgagcg cccgacgcgg 60
ccgagggcgg cggagtcccc agctagcccc ggcgcccgcc gccgcccaga cccgacgcga 120
ggccacctcg tcggcgctcc cccgagtcct cgctcgcccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagttattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccgcgag tcgggctctg gaggaagaaga aagtttgcca aggcacgagt aacaagctca 360
cgagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggagggt gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tggagagcat ccagtgccgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggg agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtg aggagaggag aactgcccaa aactgaccaa aatcatctgt gccacgagct 900
gctccggggc ctgcctgggc aagtccccc gtgactgctg ccacaaccag tgtgtgcag 960
gctgcacagg cccccgggag agcgactgcc tggctgccc ccaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccocga gggcaaatat agctttgggt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260

gaataggtat tgggtaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cactccatc agtgccgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggctag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcaggg aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggg cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccagggt tgccatgcct tgtgctcccc cgaggggctg tgggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggaggggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggcagggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc ccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgtcc atcgccactg gtaggtgggg ggcctcctc ttgctgctgg 2220
tgggtggcct ggggatcgcc ctcttcctgc gaagcgcca catcgctcgg aagcgacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccgggtg gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccgt cgctatcaag gaattaaag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgtg atggccagcg tggacaaccc cggtcacgtg tgccgctgc 2580
tgggcatctg cctcacctcc accgtgcagc tcatcacgca gctcatgcc ttcggtgcc 2640
tcctggacta tgtccgggaa cacaaagaca atattggctc ccagtacctg ctcaactggt 2700
gtgtgcagat cgcaaaggcg atgaactact tggaggaccg tcgcttgggt caccgcgacc 2760
tggcagccag gaacgtact gtgaaaacac cgcagcatgt caagatcaca gattttgggc 2820
tggccaaact gctgggtgag gaagagaaag aataccatgc agaaggaggc aaagtgcct 2880
tcaagtggat ggcattggaa tcaattttac acagaatcta taccaccag agtgatgtct 2940
ggagctacgg ggtgactggt tgggagttga tgaccttgg atccaagcca tatgacggaa 3000
tccctgccag cgagatctcc tccatcctgg aaaaaggaga acgctcctc cagccacca 3060
tatgtcccat gcatgtctac atgatcatgg tcaagtgtcg gatgatagac gcagatagtc 3120
gccccaaagt ccgtgagttg atcatcgaat tctccaaaat ggcccgagac cccagcgct 3180
acctgtcat tcagggggat gaaagaatgc atttgccaag tcctacagac tccaacttct 3240
accgtgccct gatggatgaa gaagacatgg acgacgtgg ggatgccgac gactacctca 3300
tcccacagca gggcttcttc agcagccctc ccacgtcac gactccctc ctgagctctc 3360

tgagtgaac cagcaacaat tccaccgtgg cttgcattga tagaaatgg ctgcaaagct 3420
gtcccatcaa ggaagacagc ttcttgacgc gatcacgtc agacccaca ggcgccttga 3480
ctgaggacag catagacgac accttctctc cagtgcctga atacataaac cagtccgttc 3540
ccaaaaggcc cgctggctct gtgcagaate ctgtctatca caatcagcct ctgaacccc 3600
cgccagcag agaccacac taccaggacc cccacagcac tgagtgggc aaccccgagt 3660
atctcaacac tbtccagccc acctgtgtca acagcacatt cgacagccct gccactggg 3720
cccagaaagg cagccaccaa attagcctgg acaaccctga ctaccagcag gacttctttc 3780
ccaaggaagc caagccaaat ggcattctta agggctccac agctgaaaat gcagaatacc 3840
taagggtcgc gccacaaagc agtgaattta ttggagcatg a 3881

```

<210> 725

<211> 3878

<212> ADN

<213> Homo sapiens

<400> 725

```

ccccgcccag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggcggc cggaggtcccg agctagcccc ggccggccgcc gccgcccaga ccggacgaca 120
ggccacctcg tcggcggtccg cccgagtcgc cgctcgcccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccggcag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtgggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagtcct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcac ggccgctgac gggtcagcaa caaccctgcc ctgtgcaacg 720
tggagagcat ccagtggcgg gacatagtca gcagtgaact tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtg aggagaggag aactgccaga aactgaccaa aatcatctgt gccacgagct 900
gctccggggc ctgccgtggc aagtccccc gtgactgctg ccacaaccag tgtgtgcag 960
gctgcacagg ccccggggag agcgactgac tggctcgccg caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaacct caccacgtac cagatggatg 1080
tgaaccccca gggcaaatag agctttgggt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcgaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tgggtgaatt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cactccatc agtggcgatc tccacatcct gccgggtgga tttagggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac agggttttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggctcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaaacagt 1740
gcaagcccac aggccaggtc tgccatgcct tgtgtcctcc caggggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctgggtctgga agtgcgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccag tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggtgggtggg ggccctctc ttgctgctgg 2220
tgggtggcct ggggatcgcc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccggtgc gttcggcagc gtgtataagg gactctggat ccagaagggt gagaaagtta 2460
aaattcccg cgtctatcaag gaattaaagc aagcaacatc tccgaaagcc aacaaggaaa 2520
tctctgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctaa 2580
gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640
tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgtc aactggtgtg 2700
tgcatagcgc aaagggcagc aactacttgc aggacgtcgc cttggtgcac cgcgacctgg 2760
cagccaggaa cgtactgggtg aaaacaccgg agcatgtcaa gatcacagat tttgggctgg 2820
ccaaactgct ggggtcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940
gctacggggg gactgttttg gagttgatga cctttggatc caagccatat gacggaatcc 3000
ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccaccatat 3060
gtaccatcga tgtctacatg atcatggtca agtgcctgat gatagacgca gatagtgcgc 3120
caaagtcccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180
ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240
gtgccctgat ggatgaagaa gacatggacg acgtgggtga tgccgacgag tacctcatcc 3300
cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctcctg agctctctga 3360
gtgcaaccag caacaatttc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420
ccatcaagga agacagcttc ttgcagcgat acagctcaga cccacagggc gccttgactg 3480
aggacagcat ttcctccacc ttccctgaata cataaaccag tccgttccca 3540
aaaggcccg tggctctgtg cagaatcctg tctatcacia tcagcctctg aaccccgccg 3600
ccagcagaga ccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660
tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720
agaaaggcag ccaccaaatc agcctggaca accctgacta ccagcaggac ttctttccca 3780
aggaagccaa gccaaatggc atctttaagg cctccacagc tgaaaatgca gaatacctaa 3840
gggtcgcgcc acaaaagcag gaatttatgg gagcatga 3878

```

<210> 726

<211> 3878

<212> ADN

<213> Homo sapiens

5 <400> 726

```

cccgggcgcag cgcggccgcga gcagcctccg cccccgcac ggtgtgagcg cccgacgcgg 60
ccgaggcggc cggagtcctcg agctagcccc ggcgcccgcc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctccg cccgagtcctc cgctcgcccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccgcgag 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
tagcagtcct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgcctgct gggtcagcaa caaccctgcc ctgtgcaacg 720
tggagagcat ccagtgccgg gacatagtca gcagtgaact tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccagcagct 900
gctccgggcg ctgccgtggc aagtccccc gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg cccccgggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaacccga gggcaaatc agctttggtg ccacctgctg gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tggatgaatt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440

aggaaatcac aggggttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgttttg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgaggggctg tggggcccg 1800
agcccaggga ctgctctctc tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggcgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgctg caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggtgggtggg ggccctcctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccagtg ggagaagctc 2340

```

ccaaccaagc	tctcttgagg	atcttgaagg	aaactgaatt	caaaaagatc	aaagtgtg	2400
gctccggtgc	gttcggcacg	gtgtataagg	gactctggat	cccagaaggt	gagaaagtta	2460
aaattcccgt	cgctatcaag	gaattaaagag	aagcaacatc	tccgaaagcc	aacaaggaaa	2520
tctctgatga	agcctacgtg	atggccagcg	tggacaaccc	ccacgtgtgc	cgctgtgtg	2580
gcattctgcct	cacctccacc	gtgcagctca	tcacgcagct	catgcccttc	ggctgcctcc	2640
tggactatgt	ccgggaacac	aaagacaata	ttggctccca	gtacctgtctc	aactgggtgtg	2700
tgcagatcgc	aaagggcatg	aactacttgg	aggaccgtcg	cttggtgcac	cgcgacctgg	2760
cagccaggaa	cgtactgggtg	aaaacaccgc	agcatgtcaa	gatcacagat	tttgtgtgtg	2820
ccaaactgct	gggtgcggaa	gagaaagaat	accatgcaga	aggaggcaaa	gtgcctatca	2880
agtggatggc	attggaatca	attttacaca	gaatctatac	ccaccagagt	gatgtctgga	2940
gctacggggt	gactgttttg	gagttgatga	cctttggatc	caagccatat	gacggaatcc	3000
ctgccagcga	gatctcctcc	atcctggaga	aaggagaacg	cctccctcag	ccaccatat	3060
gtaccatcga	tgtctacatg	atcatggtca	agtgtgtgat	gatagacgca	gatagtgcgc	3120
caaagtccg	tgagttgatc	atcgaattct	ccaaaatggc	cggagacccc	cagcgctacc	3180
ttgtcattca	gggggatgaa	agaatgcatt	tgccaaagtcc	tacagactcc	aacttctacc	3240
gtgccctgat	ggatgaagaa	gacatggacg	acgtggtgga	tgccgacgag	tacctcatcc	3300
cacagcagg	cttcttcacg	agccccctca	cgtcacggac	tccccctctg	agctctctga	3360
gtgcaaccag	caacaattcc	accgtggctt	gcattgatag	aaatgggctg	caaagctgtc	3420
ccatcaagga	agacagcttc	ttgcagcgat	acagctcaga	ccccacaggc	gccttgactg	3480
aggacagcat	agacgacacc	ttcctcccag	tgcctgaata	cataaaccag	tccgttccca	3540
aaaggcccg	tggtctctgtg	cagaatcctg	tctatcacaa	tcagcctctg	aaccccgcg	3600
ccagcagaga	ccacactac	caggaccccc	acagcactgc	agtgggcaac	cccaggtatc	3660
tcaacactgt	ccagcccacc	tgtgtcaaca	gcacattcga	cagccctgcc	caactgggcc	3720
agaaaaggcag	ccaccaaat	agcctggaca	accctgacta	ccagcaggac	ttctttccca	3780
aggaagccaa	gccaaatggc	atctttaagg	gctccacagc	tgaaaatgca	gaatacctaa	3840
gggtcgcgcc	acaaagcagt	gaatttattg	gagcatga			3878

<210> 727

<211> 3878

<212> ADN

5 <213> Homo sapiens

<400> 727

cccgccgcag	cgccggccgca	gcagcctccg	ccccccgcac	ggtgtgagcg	cccagcgcgg	60
ccgaggcggc	cggagtcccc	agctagcccc	ggcgcccgcc	gccgcccaga	ccggacgaca	120
ggccacctcg	tccggcgtccg	cccagatccc	cgctcgcgcg	ccaacgccac	aaccaccgcg	180
cacggccccc	tgactccgtc	cagtattgat	cgggagagcc	ggagcgagct	cttcggggag	240
cagcgatcgc	accctccggg	acggcccggg	cagcgctcct	ggcgctgctg	gctgcgctct	300
gccccggcag	tccgggctctg	gaggaaaaga	aagtttgcca	aggcacgagt	aacaagctca	360
cgcatgtggg	cacttttgaa	gatcattttc	tcagcctcca	gaggatgttc	aataactgtg	420
aggtggtcct	tgggaatttg	gaaattacct	atgtgcagag	gaattatgat	ctttccttct	480
taaagaccat	ccaggagggtg	gctggttatg	tcctcattgc	cctcaacaca	gtggagcgaa	540
ttccttttga	aaacctgcag	atcatcagag	gaaatatgta	ctacgaaaat	tcctatgcct	600
tagcagctct	atctaactat	gatgcaaata	aaaccggact	gaaggagctg	cccatgagaa	660
atttacagga	aatcctgcat	ggcgccgtgc	ggttcagcaa	caaccctgcc	ctgtgcaacg	720
tggagagcat	ccagtggcgg	gacatagtca	gcagtgactt	tctcagcaac	atgtcgatgg	780
acttcagaaa	ccacctgggc	agctgccaaa	agtgtgatcc	aagctgtccc	aatgggagct	840
gctgggggtg	aggagaggag	aactgccaga	aactgaccaa	aatcatctgt	gcccagcagt	900
gctccggggc	ctgcctgtgc	aagtccccca	gtgactgctg	ccacaaccag	tgtgtctgcag	960
gctgcacagg	cccccgggag	agcgactgcc	tggctctgcc	caaattccga	gacgaagcca	1020
cgtgcaagga	cacctgcccc	ccactcatgc	tctacaaccc	caccacgtac	cagatggatg	1080
tgaaccccg	gggcaaatat	agctttgggtg	ccacctgcgt	gaagaagtgt	ccccgtaatt	1140
atgtggtgac	agatcacggc	tcgtgcgtcc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaagacgg	cgtccgcaag	tgtagaaggt	gcgaagggcc	ttgccgcaaa	gtgtgtaacg	1260
gaatagggtat	tgggtgaattt	aaagactcac	tctccataaa	tgtacgaat	attaaacact	1320
tcaaaaactg	cacctccatc	agtggcgatc	tccacatcct	gccggtggca	tttaggggtg	1380
actccttcac	acatactcct	cctctggatc	cacaggaact	ggatattctg	aaaaccgtaa	1440
aggaaatcac	agggtttttg	ctgattcagg	cctggcctga	aaacaggacg	gacctccatg	1500
cctttgagaa	cctagaaatc	atacgcgcca	ggaccaagca	acatggtcag	ttttctcttg	1560
cagtctgcag	cctgaacata	acatccttgg	gattacgctc	cctcaaggag	ataagtgatg	1620
gagatgtgat	aatttccagg	aacaaaaaat	tgtgctatgc	aaatacaata	aactggaaaa	1680
aaactgtttg	gactcccggt	cagaaaacca	aaattataag	caacagaggt	gaaaacagct	1740
gcaaggccac	aggccaggtc	tgccatgcct	tgtgctcccc	cgagggctgc	tggggcccg	1800

```

agccagggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggaggggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggctctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atggggctaa gatcccgctc atcgccactg ggatgggtgg ggccctctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcctgc gaagcgccca catcgctcgg aagcgcacgc 2280
tgccgagggt gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgcctg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattccgtg cgtatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgctgctgg 2580
gcatctgcct caactccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640
tgactatgt ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700
tgcagatgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760
cagccaggaa cgtactgggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820
ccaaacagct ggggtcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880
agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940
gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000
ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccaccatat 3060
gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtgcgc 3120
caaagttccg tgagttgatc atcgaattct ccaaatggc ccgagacccc cagcgtacc 3180
ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240
gtgcctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300
cacagcaggg cttcttcagc agcccctcca cgtcacggac tcccctctg agctctctga 3360
gtgcaaccag caacaattcc accgtggtt gcattgatag aaatgggtg caaagctgtc 3420
ccatcaagga agacagcttc ttgcagcat acagctcaga cccacaggc gccttgactg 3480
aggacagcat agacgacacc ttcctccag tgctgaata cataaaccag tccgttccca 3540
aaaggcccg tggctctgtg cagaatcctg tctatcacia tcagcctctg aaccccggc 3600
ccagcagaga ccacactac caggaccccc acagactgc agtgggcaac cccagtatc 3660
tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagcctgcc 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> ADN

5 <213> Homo sapiens

<400> 728

```

cccgccgag cgccggccgca gcagcctccg ccccccgac ggtgtgagcg cccgacgagg 60
ccgagcgggc cggagtccc agctagcccc ggccggccgccc gccgcccaga ccgacgaca 120
ggccacctcg tcggcgctccg cccgagtccc cgctcgccg ccaacgccac aaccaccgg 180
cacggccccc tgactccgtc cagttatgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgtcctt ggcgtgctg gctgcgctct 300
gcccggcgag tcgggctctg gaggaaga aaagttgcca aggcacgagt aacaagctca 360
cgagttggg cacttttgaa gatcattttc tcagcctcca gaggatgtt aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttctctct 480
taaagaccat ccaggaggtg gctggttatg tctcattg cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgcgtgc ggttcagcaa caacctgcc ctgtgcaacg 720
tgagagagcat ccagtggcgg gacatagtca gcagtactt tctcagcaac atgtcgatgg 780
acttcagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aactgccaga aactgaccaa aatcatctgt gccacgagc 900
gctccggggc ctgcctggc agtccccca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcgactgcc tggctgccc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc cactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccg gggcaaatag agctttgggt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagcgg cgtccgcaag tgtaagaagt gcgaagggcc ttgcccgaag gtgtgtaacg 1260
gaataggtat tggatgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320

```

tcaaaaaactg	cacctccatc	agtggcgatc	tccacatcct	gccgggtggca	tttaggggtg	1380
actccttcac	acatactcct	cctctggatc	cacaggaact	ggatattctg	aaaaccgtaa	1440
aggaaatcac	aggggtttttg	ctgattcagg	cttggcctga	aaacaggacg	gacctccatg	1500
cctttgagaa	cctagaaatc	atacgcgga	ggaccaagca	acatggtcag	ttttctcttg	1560
cagtcgctcag	cctgaacata	acatccttgg	gattacgctc	cctcaaggag	ataagtgatg	1620

gagatgtgat	aatttcagga	aacaaaaatt	tgtgctatgc	aaatacaata	aactggaaaa	1680
aactgtttgg	gacctccggt	cagaaaaacca	aaattataag	caacagaggt	gaaaaacagct	1740
gcaaggccac	aggccagggtc	tgccatgcct	tgtgctcccc	cgagggtctg	tggggcccg	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	cggcaggag	2040
tcattgggaga	aaacaacacc	ctggtctgga	agtacgcaga	cgccggccat	gtgtgccacc	2100
tgtgccatcc	aaactgcacc	tacggatgca	ctgggccagg	tcttgaaggc	tgtccaacga	2160
atgggacctaa	gatcccgctc	atcgccactg	ggatgggtggg	ggccctcctc	ttgctgctgg	2220
tgggtggccct	ggggatcggc	ctcttcatgc	gaaggcgcca	catcggttcg	aagcgcacgc	2280
tgcggaggct	gctgcaggag	agggagcttg	tggagcctct	tacacccagt	ggagaagctc	2340
ccaaccaagc	tctcttgagg	atcctgaagg	aaactgaatt	caaaaagatc	aaagtgcctg	2400
gctccggtgc	gttcggcacg	gtgtataagg	gactctggat	cccagaaggt	gagaaagtta	2460
aaattcccgt	cgctatcaag	gaattaagag	aagcaacatc	tccgaaagcc	aacaaggaaa	2520
tcctcgatga	agcctacgtg	atggccacg	tggacaaccc	ccacgtgtgc	cgccctgctg	2580
gcactgcctc	cacctccacc	gtgcagctca	tcacgcagct	catgcccttc	ggctgcctcc	2640
tggactatgt	cggggaacac	aaagacaata	ttggtcccca	gtacctgctc	aactggtgtg	2700
tgcagatcgc	aaagggcatg	aactacttgg	aggaccgtcg	cttgggtgcac	cgcgacctgg	2760
cagccaggaa	cgtactgggtg	aaaacaccgc	agcatgtcaa	gatcacagat	tttgggctgg	2820
ccaaactgct	gggtgcggaa	gagaaagaat	accatgcaga	aggaggcaaa	gtgcctatca	2880
agtggatggc	atcggaatca	attttacaca	gaatctatac	ccaccagagt	gatgtctgga	2940
gctacggggg	gactgttttg	gagttgatga	cctttggatc	caagccatat	gacggaatcc	3000
ctgccagcga	gatctcctcc	atcctggaga	aaggagaacg	cctccctcag	ccacccatat	3060
gtaccatcga	tgtctacatg	atcatggtca	agtctggat	gatagacgca	gatagtgcgc	3120
caaagtctcg	tgagttgatc	atcgaattct	ccaaaatggc	ccgagacccc	cagcgctacc	3180
ttgtcattca	gggggatgaa	agaatgcatt	tgccaagtcc	tacagactcc	aacttctacc	3240
gtgccctgat	ggatgaagaa	gacatggacg	acgtggtgga	tgcgcagcag	tacctcatcc	3300
cacagcaggg	cttcttcacg	agcccctcca	cgtcacggac	tcccctcctg	agctctctga	3360
gtgcaaccag	caacaatttc	accgtggctt	gcattgatag	aatgggtctg	caaagctgtc	3420
ccatcaaggga	agacagcttc	ttgcagcgat	acagctcaga	ccccacaggc	gccttgactg	3480
aggacagcat	agacgacacc	ttcctccag	tgccctgaata	cataaaccag	tccgttccca	3540
aaaggcccg	tggctctgtg	cagaatcctg	tctatcacaa	tcagcctctg	aaccccgccg	3600
ccagcagaga	cccacactac	caggaccccc	acagcaactgc	agtgggcaac	cccaggtatc	3660
tcaacactgt	ccagcccacc	tgtgtcaaca	gcacattcga	cagccctgcc	cactgggccc	3720
agaaaaggcag	ccaccaaatt	agcctggaca	accctgacta	ccagcaggac	ttctttccca	3780
aggaagccaa	gccaaatggc	atccttaagg	gctccacagc	tgaaaatgca	gaatacctaa	3840
gggtcgcgcc	acaagcagct	gaattttattg	gagcatga			3878

<210> 729

<211> 3878

<212> ADN

5 <213> Homo sapiens

<400> 729

cccggcgacg	cgcgggcgca	gcagcctccg	ccccccgcac	ggtgtgagcg	cccgcgcg	60
ccgaggcggc	cggagtcccc	agctagcccc	ggcgcccgcc	gccgcccaga	ccggacgaca	120
ggccacctcg	tcggcgctccg	cccagatccc	cgctcgccg	ccaacgccac	aaccaccg	180
cacggccccc	tgactccgtc	cagtattgat	cgggagagcc	ggagcgagct	cttcggggag	240
cagcgatg	accctccggg	acggccgggg	cagcgctcct	ggcgctgctg	gctgcgctct	300
gcccggcgag	tcgggctctg	gaggaaga	aagtttgcca	aggcacgagt	aacaagctca	360
cgcagttggg	cacttttgaa	gatcattttc	tcagcctcca	gaggatgttc	aataactgtg	420
aggtggtcct	tgggaatttg	gaaattacct	atgtgcagag	gaattatgat	ctttccttct	480
taaagaccat	ccaggagggtg	gctggttatg	tcctcattgc	cctcaacaca	gtggagcgaa	540
ttcctttgga	aaacctgcag	atcatcagag	gaaatatgta	ctacgaaaa	tcctatgcct	600
tagcagctctt	atctaactat	gatgcaata	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
gctgggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccacgagct 900
gctccggggcg ctgcgctggc agtcccccgt gtgactgctg ccacaaccag tgtgtctgcag 960
gctgcacagg ccccggggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccaactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccgga gggcaaatc agctttgggt ccacctgctg gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgctgct gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaatagggtat tgggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccgggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac agggtttttt ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggctag ttttctcttg 1560
cagtctgcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaaacagct 1740
gcaaggccac aggccaggctc tgccatgcct tgtgctcccc cgagggtctg tggggcccg 1800
agcccaggga ctgcgtctct tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
accagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgctg caagacctgc ccggcaggag 2040
tcatgggaga aaacaacac ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggacctaa gatcccgctc atcgccactg ggatgggtgg ggccctctc ttgctgctgg 2220
tggtggccct ggggatcgcc ctcttcctgc gaaggcgcca catcgttcgg aagcgacgc 2280
tgccgaggct cctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtg 2400
gctccgggtgc gttcggcacg gtgtataagg gactctggat ccagaagggt gagaagtta 2460
aaattcccg cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgctgctgg 2580
gcactgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640
tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgtc aactgggtg 2700
tgcatgcgc aaaggcgatg aactacttgg aggaccgtcg cttggtgcac cgcgacctg 2760
cagccaggaa cgtactgggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctg 2820
ccaaactgct ggtgcgga gagaaagaat accatgcaga agggaggcaa gtgcctatca 2880
agtggatggc attggaatca attttacaca gaatctatac ccaccagagt tatgtctgga 2940
gtacgggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000
ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacctatat 3060
gtaccatcga tgtctacatg atcatggtca agtgtggat gatagacgca gatagtgc 3120
caaagtcccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180
ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240
gtgccctgat ggatgaagaa gacatggacg acgtgggtgga tgccgacgag tacctcatcc 3300
cacagcaggg cttcttcagc agccctcca cgtcacggac tcccctctg agctctctga 3360
gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420
ccatcaagga agacagcttc ttgcagcag acagctcaga cccacaggc gccttgactg 3480
aggacagcat agacgacacc ttctcccag tgccctgaata cataaaccag tccgttcca 3540
aaaggcccg tggtctgtg cagaatcctg tctatcacia tcagcctctg aaccccgcg 3600
ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660
tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720
agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780
aggaagccaa gccaaatggc gctccacagc tgaaaatgca gaatacctaa 3840
gggtcgccgc acaaagcagt gaatttatgg gagcatga 3878

```

<210> 730

<211> 3863

<212> ADN

5 <213> Homo sapiens

<400> 730

```

cccgccgcag cgcggccgca gcagcctccg cccccgcac ggtgtgagcg cccgacgcg 60
ccgagggcgc cggagtcctg agctagcccc ggccgcccgc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctcc cccgagtcct cgctcgccg ccaacgccac aaccaccg 180
cacggccccc tgactccgct cagtattgat cgggagagcc ggagcgagct cttcggggag 240

```

```

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccggcag tcgggctctg gaggaaga aagtttgcca aggcacgagt aacaagctca 360
cgagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggaggtg gctgggtatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctctt atctaactat gatgcaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgcctgct ggttcagcaa caaccctgcc ctgtgcaacg 720
tggagagcat ccagtcggcg gacatagtca gcagtgaact tctcagcaac atgtcgatgg 780
acttcagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtg aggagaggag aactgccaga aactgaccaa aatcatctgt gccagcagct 900
gctccggggc ctgcccgtggc aagtcctccc gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg cccccgggag agcgactgcc tggctgtgcc ccaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc cactcatgct tctacaaccc caccacgtac cagatggatg 1080
tgaaccccca gggcaaatat agcttttggtg ccacctgctg gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgctgct gagcctgtgg ggcgacagc tatgagatgg 1200
aggaaacagg cgggttcttg tctaagaaat gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tgggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tcacatcctt gccggtggca tttaggggtg 1380
actccttcac cataactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggttttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactgaaaaa 1680
aactgttttg gacctccggg cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac agggccagtc tgccatgcct tgtgctcccc cgagggtgct tggggcccg 1800

agccagggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaacctctc ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgc 1920
accagagtg cctgcctcag gccatgaaca tcacctgcac aggcggggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc ccaactgcgt caagacctgc cggcaggag 2040
tcattgggaga aaacaacacc ctgggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccag tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggatggtggg ggcctctctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctc tacaccagc ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccgggtg gctccgcaag gtgtataagg gactctggat ccagaaggat gagaaagtta 2460
aaattcccgat cgctatcaaa acatctccga aagccaacaa ggaaatctc gatgaagcct 2520
acgtgatggc cagcgtggac aacccccacg tgtgccgctc gctgggcatc tgcctcacct 2580
ccaccgtgca gctcatcagc cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640
aacacaaagg caatattggc tcaccatacc tgctcaactg gtgtgtgcag atcgcaagg 2700
gcatgaacta cttggaggac cgtcgcttgg tgaccgcga 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 ctacgccacc catatgtacc atcgatgtct 3060
acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccagg ttccgtgagt 3120
tgatcatcga attctccaaa atggcccag acccccagc ctaccttgct attcagggg 3180
atgaagaagt gcatttgcca agtccctacg actccactt ctaccgtgcc ctgatggatg 3240
aagaagacat ggacgacgtg gtggatgccg acgagtacct catccacag cagggttct 3300
tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtga accagcaaca 3360
attccactgt ggcttgcatc gatagaaatg gctgcccac ctgtcccac aaggaagaca 3420
gcttcttgca gcgatacagg tcagacccca caggcgctt gactgaggac agcatagacg 3480
acaccttctt cccagtgctt gaatacataa accagtccgt tccaaaagg cccgtgggt 3540
ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcccccagc agagaccac 3600
actaccagga cccccacagc actgcagtgg gcaaccccca gtatctcaac actgtccagc 3660
ccacctgtgt caacagcaca ttcgacagcc ctgcccactg gcccagaaa ggagccacc 3720
aaattagcct ggacaacct gactaccagc aggacttctt tcccaaggaa gccaagccaa 3780
atggcatctt taagggtccc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840
gcagtgaatt tattggagca tga 3863

```

<210> 731

<211> 3866

<212> ADN

5 <213> Homo sapiens

<400> 731

```

ccccggcgag cgcggcgcgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggcggc cggagtcccc agctagcccc ggcgcccgcc gccgcccaga cccgacgcga 120
ggccacctcg tcggcgctccg cccgagtccc cgctcgcccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagttattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgcg accctccggg acggccgctct cagcgctcct ggcgctgctg gctgcgctct 300
gccccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagacccat ccaggagggt gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccgagct 900
gctccggggc ctgcccgtggc aagtccccc gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg cccccgggag agcgactgcc tggctgccc ccaattccga gacgaagcca 1020
cgtgcaagga cactgcccc ccaactcatgc tctacaacct caccacgtac cagatggatg 1080
tgaaccocga ggcaaaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgcccgaac gtgtgtaacg 1260
gaataggtat tggtgaaatt aaagactcac tctccataaa tgctacgaat attaaacct 1320
tcaaaaaact cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc ccacggaact 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 acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
acccgagtg cctgcctcag gccatgaaca tcacctgcac aggcgggga ccagacaact 1980
gtatccagtg tgccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaagg tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggtggtggg ggccctcctc ttgctgctgg 2220
tgggtgccct gggtatcgcc ctcttcctgc gaaggcgcca catcgttcgg aagcgcaacg 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccgt aactatcaag gaatcatctc caaaggcaaa caaggaaatc ctcgatgaag 2520
cctacgtgat ggcagcgctg gacaaccccc acgtgtgccg cctgctgggc atctgctca 2580
cctccaccgt gcagctcatc acgcagctca tgcccttcgg ctgctcctg gactatgtcc 2640
gggaacacaa agacaatatt ggctcccagt acctgctcaa ctggtgtgtg cagatcgcaa 2700
agggcatgaa ctacttgagg gaccgtcgct tgggtgcacc cgacctggca gccaggaacg 2760
tactggtgaa aacaccgcag catgtcaaga tcacagattt tgggctggcc aaactgctgg 2820
gtgcggaaga gaaagaatac catgcagaag gaggcaaggt gcctatcaag tggatggcat 2880
tggaatcaat ttacacaga atctataccc accagagtga tgtctggagc tacggggtga 2940
ctgtttggga gttgatgacc tttggatcca agccatatga cggaatccct gccagcgaga 3000
tctcctccat cctggagaaa ggagaacgcc tcctcagcc acccatatgt accatcgatg 3060
tctacatgat catggtcaag tgctggatga tagcgcaga tagtcgcca aagttccgtg 3120
agttgatcat cgaattctcc aaaatggccc gagaccccca gcgctacctt gtcattcagg 3180
gggatgaaag aatgcatttg ccaagtccca cagactccaa cttctaccgt gccctgatgg 3240
atgaagaaga catggacgac gtggtggatg ccgacgagta cctcatccca cagcagggtc 3300
tcttcagcac ccctccacg tcacggactc cctcctgag ctctctgagt gcaaccagca 3360
acaattccac cgtggcttgc attgatgaa atgggtgca aagctgtccc atcaaggag 3420
acagcttctt gcagcgatac agctcagacc ccacaggcgc cttgactgag gacagcatag 3480
acgacacctt cctcccagtg cctgaatata taaaccagtc cgttcccaa aggcccgctg 3540
gctctgtgca gaatcctgtc tatcacaatc agcctctgaa ccccgcgccc agcagagacc 3600

cacactacca ggacccccac agcactgcag tgggcaaccc cgagtatctc aacactgtcc 3660
agcccacctg tgtcaacagc acattcgaca gccctgccc ctgggcccag aaaggcagcc 3720
accaaattag cctggacaac cctgactacc agcaggactt ctttcccaag gaagccaagc 3780
caaatggcat ctttaagggc tccacagctg aaaatgcaga atacctaagg gtcgcgccac 3840
aaagcagtga atttattgga gcatga 3860

```

<210> 732

<211> 3860

5 <212> ADN

<213> Homo sapiens

<400> 732

```

ccccggcgag cgcgggcgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggggc cggagtcccg agctagcccc ggcgggcgcc gccgcccaga cgggacgaca 120
ggccacctcg tcggcgctcc cccgagtcct cgctcgcccg ccaacgccac aaccaccgcg 180
cacggcccc tgaactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccggcag tcgggctctg gagggaaaaga 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
tagcagtcct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tgagagcagc ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccacgagt 900
gctccggggc ctgccgtggc aagtcctcca gtgactgctg ccacaaccag tgtgtgcag 960
gctgcacagg ccccggggag agcgactgcc tggctgccc ccaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccg gggcaaatat agcttttggt ccacctgctg gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tggatgaatt 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 acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccaggct tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agccagggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggaagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
accagagtg cctgcctcag gccatgaaca tcacctgcac aggcggggga ccagacaact 1980

gtatccagtg tgcccactac attgacggcc cccactgctg caagacctgc cggcaggag 2040
tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggtggtggg ggccctcctc ttgctgctgg 2220
tggtggccct ggggatcgcc ctcttcacgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaage tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtt 2460
aaattcccg cgtatcaagg gaatcgaaag ccaacaagga aatcctcgat gaagcctacg 2520
tgatggccag cgtggacaac cccacgtgt gccgctgct gggtatctgc ctacacctca 2580
ccgtgcagct catcacgcag ctcatgccct tcggctgcct cctggactat gtccgggaac 2640
acaaagacaa tattggctcc cagtacctgc tcaactgggt tgtgcagatc gcaaaaggca 2700
tgaactact ggaggaccgt cgcttgggtg accgcgacct ggcagccagg aacgtactgg 2760
tgaaaacacc gcagctgtc aagatcacag attttgggt ggccaaactg ctgggtgcgg 2820
aagagaaaga ataccatgca gaaggaggca aagtgcctat caagtggatg gcattggaat 2880
caattttaca cagaatctat accaccaga gtgatgtctg gagctacggg gtgactgttt 2940
gggagttgat gacctttgga tccaagccat atgacggaat ccctgccagc gagatctcct 3000
ccatcctgga gaaaggagaa cgctccctc agccacccat atgtaccatc gatgtctaca 3060

tgatcatggt caagtgtctg atgatagacg cagatagtcg cccaaagtcc cgtgagttga 3120
tcatcgaatt ctccaaaatg gcccgagacc cccagcgcta ccttgtcatt cagggggatg 3180
aaagaatgca tttgccaagt cctacagact ccaacttcta ccgtgccttg atggatgaag 3240
aagacatgga cgaactgggt gatgccagc agtacctcat cccacagcag ggcttcttca 3300
gcagccctc cacgtcacgg actccctcc tgagctctct gagtgcaacc agcaacaatt 3360
ccaccgtggc ttgcattgat agaaatgggc tgcaaagctg tcccatcaag gaagacagct 3420
tcttgacgag atacagctca gacccacag gcgccttgac tgaggacagc atagacgaca 3480
ccttcctccc agtgcctgaa tacataaacc agtccgttcc caaaaggccc gctggctctg 3540
tgacagatcc tgtctatcac aatcagcctc tgaaccccg gccagcaga gaccacact 3600
accaggaccc ccacagcact gcagtgggca accccgagta tctcaacact gtccagccca 3660
cctgtgtcaa cagcacattc gacagccctg cccactgggc ccagaaaggc agccaccaa 3720
ttagcctgga caacctgac taccagcagg acttctttcc caagggaagg aagccaaatg 3780
gcatctttaa gggctccaca gctgaaaaatg cagaatacct aagggtcgcg ccacaaagca 3840
gtgaatttat tggagcatga 3860

```

<210> 733

<211> 3860

<212> ADN

<213> Homo sapiens

5 <400> 733

```

ccccggcgag cgcgggccga gcagcctccg cccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggcggc cggagtcgcc agctagcccc ggcgggccgc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctcc cccgagtcgc cgctcgccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccgcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagtca 360
cgcgattggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggagggt gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagtcct atctaactat gatgcaaata aaacgggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttccagaa ccacctgggc agctgcaaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccagcagct 900
gctccggggcg ctgccgtggc aagtccccc gtgactgctg ccacaaccag tgtgtgcag 960
gctgcacagg cccccgggag agcgactgcc tggctcgccg caaattccga gacgaagcca 1020
cgtgcaagga cactgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccca gggcaaatat 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 cactccatc agtggcgatc tccacatcct gccggtggca tttagggggtg 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 tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggaggggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggcgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggtggtgggg ggccctctc ttgctgctgg 2220
tggtggccct ggggatcgcc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgcgagggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtgtg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattccgct cgctatcaag gaatcgaaag ccaacaagga aatcctcgat gaagcctacg 2520
tgatggccag cgtggacaac cccacgtgt gccgcctgct gggcatctgc ctcacctcca 2580

```

```

ccgtgcagct catcacgcag ctcattgccct tcggtgcct 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 accaccaga gtgatgtctg gagctacggg gtgactgttt 2940
gggagttgat gacctttgga tccaagccat atgacggaat cctgccagc gagatctcct 3000
ccatcctgga gaaaggagaa cgcctccctc agccacccat atgtaccatc gatgtctaca 3060
tgatcatggt caagtgtctg atgatagacg cagatagtcg cccaaagttc cgtgagttga 3120
tcacgcaatt ctccaaaatg gcccgagacc cccagcgcta cctgttcatt caggggggatg 3180
aaagaatgca tttgccaaagt cctacagact ccaacttcta ccgtgccctg atggatgaag 3240
aagacatgga cgacgtggtg gatgccgacg agtacctcat cccacagcag ggcttcttca 3300
gcagccctc cagctcacgg actccctcc tgagctctct gagtgcaacc agcaacaatt 3360
ccaccgtggc ttgcattgat agaaatgggc tgcaaagctg tcccatcaag gaagacagct 3420
tcttgacgg atacagctca gacccacag gcgccttgac tgaggacagc atagacgaca 3480
ccttccctcc agtgctgaa tacataaacc agtccgttcc caaaaggccc gctggctctg 3540
tgcagaatcc tgtctatcac aatcagcctc tgaaccccg cccagcaga gaccacact 3600
accaggaccc ccacagcact gcagtgggca accccgagta tctcaacact gtccagccca 3660
cctgtgtcaa cagcacctg gacagccctg cccactgggc ccagaaaggc agccaccaa 3720
ttagcctgga caaccctgac taccagcagg acttctttcc caaggaaagg aagccaaatg 3780
gcattcttaa gggctccaca gctgaaaatg cagaatacct aagggtcgcg ccacaaagca 3840
gtgaatttat tggagcatga 3860

```

<210> 734

<211> 3878

<212> ADN

5 <213> Homo sapiens

<400> 734

```

cccgggcgag cgcggccgca gcagcctccg cccccgcac ggtgtgagcg cccgacgagg 60
ccgaggcggc cggagtcccc agctagcccc ggcgcccgcc gccgcccaga cccgacgaca 120
ggccacctcg tcggcgctccg cccgagtcct cgctcgcccg ccaacgccac aaccaccgag 180
cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct ctcgggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctcgctctc 300
gccggcgag tcgggctctg gaggaaga aagtttgcca aggcacgagt aacaagctca 360
cgagttggg 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
tagcagtcct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcac ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tggagagcat ccagtggcgg gacatagtaa gcagtgaact tctcagcaac atgtcgatgg 780
acttcagaaa caacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccacgagc 900
gctccggggc ctgccgtggc aagtccccca gtactgctg ccacaaccag tgtgtgcag 960
gctgcacagg ccccggggag agcgactgcc tggctcgccg caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatg tctacaacc caccacgtac cagatggatg 1080
tgaacccgga gggaataatc agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tggatgaatt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cactccatc 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 tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggcaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
accagagtg cctgcctcag gcatgaaca tcacctgcac aggcagggga ccagacaact 1980
gtatccagtg tgccactac attgacggcc cccactgcgt caagacctgc cggcgaggag 2040
tcattgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100

```

tgtgccatcc	aaactgcacc	tacggatgca	ctgggccagg	tcttgaaggc	tgtccaacga	2160
atgggcctaa	gatcccgtcc	atcgccactg	ggatgggtggg	ggccctcctc	ttgctgctgg	2220
tggtggccct	ggggatcggc	ctcttcatgc	gaaggcgcca	catcggttcg	aagcgcacgc	2280
tgcggaggct	gctgcaggag	agggagcttg	tggaacctct	tacacccagt	ggagaagctc	2340
ccaaccaagc	tctcttgagg	atcttgaagg	aaactgaatt	caaaaagatc	aaagtgctgg	2400
gctccggtgc	gttcggcacg	gtgtataaag	gactctggat	cccagaaggt	gagaaagtta	2460
aaattcccgt	cgctatcaag	gaattaagag	aagcaacatc	tccgaaagcc	aacaaggaaa	2520
tctctgatga	agcctacgtg	atggccagcg	tggaacaacc	ccacgtgtgc	cgctgctggg	2580
gcattctgct	cacctccacc	gtgcagctca	tcacgcagct	catgcccttc	ggctgcctcc	2640
tggaactatg	ccgggaacac	aaagacaata	ttggctccca	gtacctgctc	aactgggtgtg	2700
tgcagatcgc	aaagggcatg	aactacttgg	aggaccgtcg	cttgggtcac	cgcgacctgg	2760
cagccaggaa	cgctactggg	aaaacaccgc	agcatgtcaa	gatcacagat	tttggggcgg	2820
ccaaactgct	gggtgcggaa	gagaaagaat	accatgcaga	aggaggcaaa	gtgcctatca	2880
agtggatggc	atttgaatca	atcttacaca	gaatctatac	ccaccagagt	gatgtctgga	2940
gctacggggg	gactgttttg	gagttgatga	cctttggatc	caagccatat	gacggaatcc	3000
ctgccagcga	gatctcctcc	atcctggaga	aaggagaacg	cctccctcag	ccaccatat	3060
gtaccatcga	tgtctacatg	atcatggtca	agtgtgggat	gatagacgca	gatagtgcgc	3120
caaagtcccg	tgagttgatc	atcgaattct	ccaaaatggc	ccgagacccc	cagcgctacc	3180
ttgtcattca	gggggatgaa	agaatgcatt	tgccaagtcc	tacagactcc	aacttctacc	3240
gtgcctgat	ggatgaagaa	gacatggacg	acgtggtgga	tgccgacgag	tacctcatcc	3300
cacagcaggg	cttcttcacg	agccccctca	cgtcacggac	tccccctctg	agctctctga	3360
gtgcaaccag	caacaattcc	accgtggctt	gcattgatag	aatggggctg	caaagctgtc	3420
ccatcaagg	agacagcttc	ttgcagcgat	acagctcaga	ccccacaggc	gccttgactg	3480
aggacagcat	agacgacacc	ttcctcccag	tgctgaata	cataaaccag	tccgttccca	3540
aaaggcccg	tggtctctgt	cagaatcctg	tctatcacia	tcagcctctg	aaccccgcg	3600
ccagcagaga	ccacactac	caggaccccc	acagcactgc	agtgggcaac	cccaggtatc	3660
tcaacactgt	ccagcccacc	tgtgtcaaca	gcacattcga	cagccctgcc	caactgggccc	3720
agaaaggcag	ccaccaaatt	agcctggaca	accctgacta	ccagcaggac	ttctttccca	3780
aggaaagcca	gccaaatggc	atctttaagg	gctccacagc	tgaaaatgca	gaatacctaa	3840
gggtcgcgcc	acaagcagct	gaatttattg	gagcatga			3878

<210> 735

<211> 3878

<212> ADN

5 <213> Homo sapiens

<400> 735

cccgggcgag	cgcggccgca	gcagcctccg	ccccccgcac	ggtgtgagcg	cccgcgcggg	60
ccgaggcggc	cggagtccc	agctagcccc	ggcgcccgcc	gccgcccaga	ccggacgaca	120
ggccacctcg	tggcgctccg	cccagatccc	cgctcgcccg	ccaacgccac	aaccaccgcg	180
cacggccccc	tgactccgtc	cagtattgat	cgggagagcc	ggagcgagct	cttcggggag	240
cagcgatcgc	accctccggg	acggccgggg	cagcgctcct	ggcgctgctg	gctgcgctct	300
gcccggcgag	tcgggctctg	gaggaaaaga	aagtttgcca	aggcacgagt	aacaagctca	360
cgcagttggg	cacttttgaa	gatcattttc	tcagcctcca	gaggatgttc	aataactgtg	420
aggtggtcct	tgggaaattt	gaaattacct	atgtgcagag	gaattatgat	ctttccttct	480
taaagaccat	ccaggagggtg	gctggttatg	tcctcattgc	cctcaacaca	gtggagcgaa	540
ttccttttga	aaacctgcag	atcatcagag	gaaatatgta	ctacgaaaat	tcctatgect	600
tagcagctct	atctaactat	gatgcaaata	aaaccggact	gaaggagctg	cccatgagaa	660
atttacagga	aatcctgcat	ggcgccgtgc	ggttcagcaa	caacctgccc	ctgtgcaacg	720
tggagagcat	ccagtggcgg	gacatagtca	gcagtgactt	tctcagcaac	atgtcgatgg	780
acttcagaaa	ccacctgggc	agctgccaaa	agtgtgatcc	aagctgtccc	aatgggagct	840
gctgggggtg	aggagaggag	aactgccaga	aactgaccaa	aatcatctgt	gcccagcagt	900
gctccggggc	ctgcgctggc	aagtccccc	gtgactgctg	ccacaaccag	tgtgctgcag	960
gctgcacagg	ccccggggag	agcgactgcc	tggtctgccc	caaattccga	gacgaagcca	1020
cgtgcaagga	cacctgcccc	ccactcatgc	tctacaaccc	caccacgtac	cagatgggat	1080
tgaacccoga	gggcaaatat	agctttgggt	ccacctgcgt	gaagaagtgt	ccccgtaatt	1140
atgtggtgac	agatcacggc	tcgtgcgtcc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaaagacg	cgtccgcaag	tgtgaagaag	gcgaagggcc	ttgccgcaaa	gtgtgtaacg	1260
gaatagggtat	tgttgaaatt	aaagactcac	tctccataaa	tgtacgaat	attaacact	1320
tcaaaaactg	cacctccatc	agtggcgatc	tccacatcct	gccgggtgga	tttaggggtg	1380
actccttcac	acatactcct	cctctggatc	cacaggaact	ggatatcttg	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 gaaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggtgctc tggggcccg 1800
agcccaggga ctgctctctc tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgctg caagacctgc ccggcaggag 2040
tcattgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggacctaa gatcccgctc atcgccactg ggatggtggg ggccctctctc ttgctgctgg 2220
tggtggccct ggggatcgcc ctcttcatgc gaaggcgcca catcggttcgg aagcgcacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat ccagaagggt gagaaggtta 2460
aaattcccg tgcctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgtg atggccagcg tggacaacct ccacgtgtgc cgctgctgg 2580
gcattgtcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640
tgactatgt ccgggaacac ttggtctcca gtacctgtc aactggtgtg 2700
tgcatatgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760
cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcggg 2820
ccaaactgct ggggtgcggaa gagaaagaat accatgcaga agggaggcaa gtgcctatca 2880
agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940
gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000
ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccaccatat 3060
gtaccatcga tgtctacatg atcatggtca agtgtggat gatagacgca gatagtgcgc 3120
caaagtcccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180
ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240
gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300
cacagcaggg cttcttcagc agccctcca cgtcacggac tcccctctctg agctctctga 3360
gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420
ccatcaagga agacagcttc ttgcagcga acagctcaga cccacaggc gccttgactg 3480
aggacagcat agacgacacc ttctctccga tgccgaata cataaaccag tccgttccca 3540
aaaggcccg tggctctgtg cagaatcctg tctatcacaa tcagcctctg aaccccgccg 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 gaatttatgg gagcatga 3878

```

<210> 736

<211> 3878

<212> ADN

5 <213> Homo sapiens

<400> 736

```

ccccggcgag cgcgcccgca gcagcctccg cccccgcac ggtgtgagcg cccgacgagg 60
ccgagggcggc cggagtcctcg agctagcccc ggcgcccgcc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctccg cccgagtcctc cgctcgcccg ccaacgccac aaccaccggg 180
cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gcccggcgag tcgggctctg gagggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgagttggg 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
tagcagtcct atctaactat gatgcaaata aaacccgact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caacctgcc ctgtgcaacg 720
tgagagcat ccagtggcgg gacatagtca gcagtactt tctcagcaac atgtcgatgg 780
acttcagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccagcagct 900
gctccgggag ctgcccgtgc aagtccccc gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcgactgcc tggctgccc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc cactcatgc tctacaacct caccacgtac cagatggatg 1080

```

tgaaccccg	gggcaaatc	agctttggtg	ccacctgcgt	gaagaagtgt	ccccgtaatt	1140
atgtggtgac	agatcacggc	tcgtgcgtcc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaagacgg	cgtccgcaag	tgtgaagaag	gcgaagggcc	ttgccgcaaa	gtgtgtaacg	1260
gaataggtat	tggtgaattt	aaagactcac	tctccataaa	tgctacgaat	attaaacact	1320
tcaaaaactg	cacctccatc	agtggcgatc	tccacatcct	gccgggtggc	tttaggggtg	1380
actccttcac	acatactcct	cctctggatc	cacaggaact	ggatattctg	aaaaccgtaa	1440
aggaaatcac	agggtttttg	ctgattcagg	cttggcctga	aaacaggacg	gacctccatg	1500
cctttgagaa	cctagaaatc	atacgcgga	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	ccaaggagtg	ttgtggagaa	ctctgagtgc	atacagtgcc	1920
acccagagtg	cctgcctcag	gccatgaaca	tcacctgcac	aggacgggga	ccagacaact	1980
gtatccagtg	tgcccactac	attgacggcc	cccactgcgt	caagacctgc	ccggcaggag	2040
tcattgggaga	aaacaacacc	ctggctctgga	agtacgcaga	cgccggccat	gtgtgccacc	2100
tgtgccatcc	aaactgcacc	tacggatgca	ctgggcccagg	tcttgaaggc	tgtccaacga	2160
atgggcccata	gatcccgtcc	atcgccactg	ggatgggtggg	ggccctcctc	ttgctgctgg	2220
tggtggccct	ggggatcgcc	ctcttcatgc	gaaggcgcca	catcgttcgg	aagcgcacgc	2280
tgccggaggct	gctgcaggag	agggagcttg	tggagcctct	tacaccagct	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	tggaacaacc	ccacgtgtgc	cgcctgctgg	2580
gcattctgct	cacctccacc	gtgcagctca	tcacgcagct	catgcccttc	ggctgctctc	2640
tggaactatg	ccgggaacac	aaagacaata	ttggctccca	gtacctgctc	aactgggtgtg	2700
tgcatatgcg	aaagggcatg	aactacttgg	aggaccgtcg	cttgggtcac	cgcgacctgg	2760
cagccaggaa	cgtactgggt	aaaacaccgc	agcatgtcaa	gatcacagat	tttgggctgg	2820
ccaaacagct	gggtgcggaa	gagaaagaat	accatgcaga	aggaggcaaa	gtgcctatca	2880
agtggatggc	attggaatca	attttacaca	gaatctatac	ccaccagagt	gatgtctgga	2940
gctacggggg	gactgtttgg	gagttgatga	cctttggatc	caagccatat	gacggaatcc	3000
ctgccagcga	gatctcctcc	atcctggaga	aaggagaacg	cctccctcag	ccaccatat	3060
gtaccatcga	tgtctacatg	atcatggtca	agtgtgggat	gatagacgca	gatagtgcgc	3120
caaagtcccg	tgagttgatc	atcgaattct	ccaaaatggc	ccgagacccc	cagcgctacc	3180
ttgtcattca	gggggatgaa	agaatgcatt	tgccaagtcc	tacagactcc	aacttctacc	3240
gtgccctgat	ggatgaagaa	gacatggacg	acgtgggtgga	tgccgacgag	tacctcatcc	3300
cacagcaggg	cttcttcacg	agcccctcca	cgtcacggac	tcccctcctg	agctctctga	3360
gtgcaaccag	caacaattcc	accgtggctt	gcattgatag	aaatgggctg	caaagctgtc	3420
ccatcaaggga	agacagcttc	ttgcagcgat	acagctcaga	ccccacaggc	gccttgactg	3480
aggacagcat	agacgacacc	ttcctcccag	tgctgaata	cataaaccag	tccgttccca	3540
aaaggcccgc	tggtctctgtg	cagaatcctg	tctatcacaa	tcagcctctg	aaccccgccg	3600
ccagcagaga	cccacactac	caggaccccc	acagcactgc	agtgggcaac	cccaggtatc	3660
tcaacactgt	ccagcccacc	tgtgtcaaca	gcacattcga	cagccctgcc	cactgggccc	3720
agaaaggcag	ccaccaaatt	agcctggaca	acctgacta	ccagcaggac	ttcttcccca	3780
aggaaagcaa	gccaaatggc	atctttaagg	gtccacagac	tgaaaatgca	gaatacctaa	3840
gggtcgcgcc	acaaagcagt	gaatttattg	gagcatga			3878

<210> 737

<211> 3878

<212> ADN

5 <213> Homo sapiens

<400> 737

cccggcgag	cggggccgca	gcagcctccg	ccccccgcac	ggtgtgagcg	cccgcgcggg	60
ccgaggcggc	cggagtcccc	agctagcccc	ggcgcccgcc	gccgcccaga	ccggacgaca	120
ggccacctcg	tcggcgctccg	cccagatccc	cgctcgccg	ccaacgccac	aaccaccgcg	180
cacggccccc	tgactccgtc	cagtattgat	cgggagagcc	ggagcgagct	cttcggggag	240
cagcgatgcg	accctccggg	acggcccggg	cagcgctcct	ggcgctgctg	gctgcgctct	300
gcccggcgag	tcgggctctg	gaggaaaaga	aagtttgcca	aggcacgagt	aacaagctca	360
cgcagttggg	cacttttgaa	gatcattttc	tcagcctcca	gaggatgttc	aataactgtg	420
agggtggtcct	tgggaatttg	gaaattacct	atgtgcagag	gaattatgat	ctttccttct	480
taaagaccat	ccaggagggtg	gctggttatg	tcctcattgc	cctcaacaca	gtggagcgaa	540

```

ttccttttga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagtcct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tggagagcat ccagtgccgg gacatagtc gacgtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccagcagct 900
gctccggggcg ctgccgtggc aagtccccc gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg cccccgggag agcgactgcc tggctcgccg caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc cactcatgc tctacaacc caccacgtac cagatggatg 1080
tgaaccccgga gggcaaatat agctttgggt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tctgtcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tggatgaatt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggttttt ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggcgggga ccagacaact 1980
gtatccagtg tgccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggcttggg agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaagg tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggatgggtgg ggccctctct ttgctgctgg 2220
tggtggccct gggtgctcgg cctctcatgc gaaggcgcca catcggtcgg aagcgcacgc 2280
tgccggaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaaagg aaactgaatt caaaaagatc aaagtgtgtg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccg tgcctatcaag gaattaaag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580
gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640
tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgtc aactggtgtg 2700
tgcatatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760
cagccaggaa cgtactgggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820
ccaaactgct ggggtcggaa gagaaagaat accatgcaga agggagcaaa gtgcctatca 2880
agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940
gctacggggg gactgttttg gagttgatga ccttgggatc caagccatat gacggaatcc 3000
ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccaccatat 3060
gtaccatcga tgtctacatg atcatggtca agtctggat gatagacgca gatagtgcgc 3120
caaagttccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgtacc 3180
ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240
gtgccctgat ggtgaagaa gacatggacg acgtgggtgga tgccgacgag tacctcatcc 3300
cacagcaggg cttcttcagc agccctcca cgtcacggac tcccctctg agctctctga 3360
gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420
ccatcaagga agacagcttc ttgcagcgat acagctcaga cccacaggc gccttgactg 3480
aggacagcat agacgacacc ttctcccag tgcctgaata cataaaccag tccgttccca 3540
aaagcccgcc tggctctgtg cagaatcctg tctatcacia tcagcctctg aaccccgccg 3600
ccagcagaga ccacactac caggaccccc acagactgc 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 acaagcaggt gaatttatgg gagcatga 3878

```

<210> 738

<211> 3860

<212> ADN

5 <213> Homo sapiens

<400> 738

```

ccccggcgag cgccggccgca gcagcctccg cccccgcac ggtgtgagcg cccgacgagg 60

```

```

ccgaggcgccg cggagtcctcg agctagcccc ggcggcccgcc gccgcccaga ccggacgaca 120
ggccacccctcg tcggcgctccg cccgagtcctc cgcctcgccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gcccgccgag tcgggctctg gaggaaga aagtttgcca aggcacgagt aacaagctca 360
cgcagttggg cacttttgaa gatcatttcc tcagcctcca gaggatgttc aataactgtg 420
aggtggctcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttccttttga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagtcct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caacctgcc ctgtgcaacg 720
tgagagagcat ccagtggcgg gcatagtcga gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaa ccacctgggc agctgccaaa agtgatgacc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccacgagct 900
gctccggcg ctgcctgggc aagtccccc gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcgactgcc tggctctgcg caaattccga gacgaagcca 1020
cgtgcaagga cactgcccc cactcatgc tctacaacc caccacgtac cagatggatg 1080
tgaacccca gggcaaatat agctttgggt ccacctgcgt gaagaagtggt ccccgtaatt 1140
atgtgggtgac agatccatgc tcgtgcgtgc gacgctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgaagaagt gcgaagggcc ttgccgcaa gtgtgtaacg 1260
gaataggtat tgggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cactccatc agtgccgatc tccacatcct gccggtggca tttagggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaatcac aggttttttg ctgattcagg cttggcctga aaacaggagc gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtctgac cctgaacata acatccttg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaaacgct 1740
gcaaggccac agggcaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agcccaatg ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtag ccaagggagt ttgtggagaa ctctgagtg atacagtgcc 1920
accagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagt tcgccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcattgggaga aaacaacacc ctggcttgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaagg tgctcaacga 2160
atgggacctaa gatcccgtcc atcgccactg ggatgggtgg ggccctcctc ttgctgctgg 2220
tggtggccct gggtatcggc ctctcatgc gaaggcgcca catcgctcgg aagcgacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaaag aaactgaatt caaaaagatc aaagtgtctg 2400
gctccgggac gttccggcac gtgtataagg gactctggat ccagaaaggt gagaaagtta 2460
aaattccgct cgtatcaag gaatcgaaag ccaacaagga aatcctcgat gaagcctacg 2520

tgatggccag cgtggacaac cccacgtgt gccgctgct gggcatctgc ctcacctcca 2580
ccgtgcagct catcacgag cctatgccct tcggctgcct cctggactat gtccgggaa 2640
acaaagacaa tattggctcc cagtacctgc tcaactggtg tgtgcagatc gcaaagggca 2700
tgaactactt ggaggaccgt cgcttggtgc accgcgacct ggcagccagg aacgtactgg 2760
tgaacacacc gcagcatgtc aagatcacag attttggtgt ggccaaactg ctgggtgccc 2820
aagagaaaga ataccatgca gaaggaggca aagtgcctat caagtggatg gcattggaat 2880
caattttaca cagaatctat acccaccaga gtgatgtctg gagctacggg gtgactgttt 2940
gggagttgat gacctttgga tccaagccat atgacggaat cctgccagc gagatctcct 3000
ccatcctgga gaaaggagaa gcctccctc atgtaccatc atgtaccatc gatgtctaca 3060
tgatcatggt caagtgtctg atgatagacg cagatagtcg cccaaagttc cgtgagtga 3120
tcacgaatt ctccaaaatg gcccgagacc cccagcgcta ccttgtcatt cagggggatg 3180
aaagaatgca tttgccaaat cctacagact ccaacttcta ccgtgccctg atggatgaag 3240
aagacatgga cgagtggtg gatgccgacg agtacctcat cccacagcag ggcttcttca 3300
gcagccctc cacgtcacgg actccctcc tgagctctct gagtgcacac agcaacaatt 3360
ccaccgtggc ttgcattgat agaaatgggc tgcaaaagctg tccatcaag gaagacagct 3420
tcttgacgag atacagctca gacccacag gcgccttgac tgaggacagc atagacgaca 3480
ccttccctcc agtgccctgaa tacataaacc agtccgttcc caaaaggccc gctggctctg 3540
tgacgaatcc tgtctatcac aatcagctc tgaaccccg gcacagcaga gacccacact 3600
accaggaccc ccacagcact gcagtgggca accccagta tctcaacact gtccagccca 3660
cctgtgtcaa cagcacattc gacagccctg cccactgggc ccagaaaggc agccaccaa 3720
ttagcctgga caacctgac taccagcagg acttcttcc caagggaagc aagccaaatg 3780
gcattctttaa gggctccaca gctgaaaatg cagaatacct aagggtcgcg ccacaaagca 3840
gtgaatttat tggagcatga 3860

```

<210> 739

<211> 3863

<212> ADN

5 <213> Homo sapiens

<400> 739

```

ccccggcgag cgcggcgccga gcagcctccg ccccccgac ggtgtgagcg cccgacgagg 60
ccgagggcgg cggagtcccc agctagcccc ggcgcccgcc gccgcccaga cccgacgaca 120
ggccacctcg tcggcgctccg cccgagtcgg cgctcgccgg ccaacgccac aaccaccgag 180
cacggccccc tgactccgtc cagttattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccggcag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggagggt gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccactggggc agtgcgcaaa agtgcgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccacgagct 900
gctccggggc ctgcgctggc aagtccccc gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg cccccgggag agcgactgcc tggctcgccg caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccg gggcaaatac agcttttggt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaatagggtat tgggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtgcgcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tggctatgc aaatacaata aactggaaaa 1680
aactgtttgg gcacctcggg cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
accagagtg tgccctcag gccatgaaca tcacctgcac aggcgggga ccagacact 1980
gtatcagtg tgcacctac attgacggcc cccactgcgt caagacctgc cggcaggag 2040
tcatgggaga aaacaacacc ctggctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaagg ggcctcctc 2160
atgggcctaa gatcccgctc atcgccactg ggatgggtgg ggcctcctc 2220
tggtgccct ggggatcgcc ctcttcacgc gaaggcgcca catcgctcgg aagcgcacgc 2280
tgccggaggc gctgcaggag agggagcttg tgggcctct tacacccagt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgagg aaactgaatt caaaaagatc aaagtgcagg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccg cgtatcaaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520
acgtgatggc cagcgtggac aaccccacg tgtgcgcct gctgggcatc tgccctcacct 2580
ccaccgtgca gctcatcacg cagctcatgc cctcggctg cctcctggac tatgtccggg 2640
aacacaaaga caatattggc tcccagtag 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
tttgaggagt gatgacctt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000
cctccatcct ggagaaagga gaacgcctcc ctacgccacc catatgtacc atcgatgtct 3060
acatgatcat ggtcaagtag tggatgatag acccgatag tcgcccagg ttccgtgatg 3120
tgatcatcga attctccaaa atggcccag ccccccagcg ctacctgtc attcaggggg 3180
atgaaagaat gcatttgcca agtcctacag actccaact ctaccgtgcc ctgatggatg 3240
aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300
tcagcagccc ctccacgtca cggactcccc tctgagctc tctgagtga accagcaaca 3360
attccaccgt ggcttgcatg gatagaaatg ggctgcaaa ctgtcccatc aaggaagaca 3420

gcttcttgca gcgatacagc tcagacccca caggcgctt gactgaggac agcatagacg 3480
acaccttct cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgctggct 3540
ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcccccagc agagacccac 3600
actaccagga cccccacagc actgcagtgg gcaacccgga gtatctcaac actgtccagc 3660
ccacctgtgt caacagcaca ttgcacagcc ctgcccactg ggcccagaaa ggacgccacc 3720
aaattagcct ggacaaccct gactaccagc aggacttct tcccaggaa gccaaagcaa 3780
atggcatct taagggctcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840
gcagtgaatt tattggagca tga 3863

```

<210> 740

<211> 3878

5 <212> ADN

<213> Homo sapiens

<400> 740

```

ccccggcgag cgcgggcgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggcgc cggagtcccg agctagcccc ggcgggcgcc gccgcccaga cccgacgaca 120
ggccacctcg tcggcgctccg cccgagtcgc cgctcgcccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagttattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccggcag tcgggctctg gagaaaaaga aagtttgcca aggcacagct 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
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tgagagcagc caagtggcgg gacatagtcg cagtgactt tctcagcaac atgtcgatgg 780
acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggtgag agtagaggag aactgccaga aactgaccaa aatcatctgt gccagcagt 900
gctccgggag agtcccggtg aagtcccca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcgactgcc tggctgccc ccaattccga gacgaagcca 1020
cgtgcaagga cactgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccgga gggcaaatat agcttttggt ccacctgctg gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggg tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagcagg cgtccgcaag tgtaagaagt gcgaagggcc ttgcccgaag gtgtgtaacg 1260
gaataggtat tggtagaatt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cactccatc agtgccgcat tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactgaaaaa 1680
aactgtttgg gacctccggg cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggcacgggt tgctgctccc cgagggtgct tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggcagggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgctg caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggtggtggg ggccctctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccgagggt gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg taactgaatt caaaaagatc aaagtgtga 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat ccagaagggt gagaaagtta 2460
aaattccggt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgctgctgg 2580
gcactcgctc cactccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640
tggactatgt ccgggaacac aaagacaata ttggtccca gtacctgctc aactggtgtg 2700

tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760
cagccaggaa cgtactgggt aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820
ccaaactgct ggggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880

agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940
gctacggggt gactgtttgg gagttgatga ctttggatc caagccatat gacggaatcc 3000
ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccaccatata 3060
gtaccatcga tgtctacatg atcatggtca agtgctggat gatagacgca gatagtgcgc 3120
caaagtcccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180
ttgtcattca gggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240
gtgcctgat ggatgaagaa gacatggacg acgtgggtga tgccgacgag tacctcatcc 3300
cacagcaggg cttcttcagc agcccccca cgtcacggac tccccctctg agctctctga 3360
gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420
ccatcaagga agacagcttc ttgcagcgat acagctcaga cccacaggc gccttgactg 3480
aggacagcat agacgacacc ttcctcccag tgctgaata cataaaccag tccgttccca 3540
aaaggcccg cgtgctgtgt cagaatcctg tctatcaca tcagcctctg aaccccgcg 3600
ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccagtatc 3660
tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720
agaaaggcag ccaccaaatt agcctggaca acctgacta ccagcaggac ttctttccca 3780
aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840
gggtcgcgcc acaagcagtg gaatttattg gagcatga 3878

```

<210> 741

<211> 3878

<212> ADN

<213> Homo sapiens

5 <400> 741

```

cccgggcgcag cgcggccgcga gcagcctccg cccccgcac ggtgtgagcg cccgacgcgg 60
ccgaggcggc cggagtcgcc agctagcccc ggcgcccgcc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctccg cccgagtcgc cgcctcgccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccgcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgag aacaagctca 360
cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggagggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagtcct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tgagagagct ccagtggcgg gacatagtca gcagtgaact tctcagcaac atgtcgatgg 780
acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccagcagct 900
gctccgggcg ctgccgtggc aagtcccca gtactgctg ccacaaccag tgtgtgcag 960
gctgcacagg cccccgggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaacccga gggcaaatc agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaa gtgtgtaacg 1260
gaataggtat tggatgaatt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cactccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggttttg 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 tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggaggggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggcggggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggatgggtgg ggccctcctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcattg gaaggcgcca catcgctcgg aagcgcacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400

```

```

cctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccgt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cacctgctgg 2580
gcactgcctt cactccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640
tggactatgt ccgggaacac aaagacaata ttggctccca gtacctgtc aactggtgtg 2700
tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760
cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820
ccaaactgct ggtgcggaag gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880
agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940
gtgccctgat ggatgaagaa gagttgatga cctttggatc caagccatat gacggaatcc 3000
ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccaccatat 3060
gtaccatcga tgtctacatg atcatggtca agtctggat gatagacgca gatagtgcgc 3120
caaagtcccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgtacc 3180
ttgtcattca ggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240
gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300
cacagcaggg cttcttcagc agccctcca cgtcacggac tcccctctg agctctctga 3360
gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420
ccatcaagga agacagcttc ttgcagcgat acagctcaga cccacagggc gccttgactg 3480
aggacagcat agacgacacc ttctctccca tgctgaata cataaaccag tccgttccca 3540
aaaggcccg tggtctgtg cagaatcctg tctatcaca tcagcctctg aaccccgcg 3600
ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660
tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720
agaaaggcag ccaccaaatt agcctggaca acctgacta ccagcaggac ttcttccca 3780
aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840
gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

```

<210> 742

<211> 3863

<212> ADN

5 <213> Homo sapiens

<400> 742

```

cccgccgcag cgcggccgcga gcagcctccg cccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggcgg cggagtcccc agctagcccc ggccgcccgc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctccg cccgagctcc cgctcgccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatcgc accctccggg acggccgggg cagcgtcctt ggcgctgtg gctgcgtctt 300
gccccgcgag tcgggctctg gaggaaga aaagtttcca aggcacgagt aacaagctca 360
cgcagttggg cacttttgaa gatcatttcc 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
tagcagctct atctaactat gatgcaaaa aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caacctgcc ctgtgcaacg 720
tgagagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtg aggagaggag aactgccaga aactgaccaa aatcatctgt gccagcagct 900
gctccgggcg ctcctgtggc agtccccca gtgactgtc ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcactgcc ttggtctgcg caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc cactcatgc tctacaacc caccacgtac cagatggatg 1080
tgaacccoga gggcaaatat agctttgttg ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaa gtgtgtaacg 1260
gaatagggtat tggatgaatt 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 cttagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggtgct tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920

```

accagagtg	cctgcctcag	gccatgaaca	tcacctgcac	aggacgggga	ccagacaact	1980
gtatccagt	tgcccactac	attgacggcc	cccactgcgt	caagacctgc	ccggcaggag	2040
tcattgggaga	aaacaacacc	ctgggtctgga	agtacgcaga	cgccggccat	gtgtgccacc	2100
tgtgccatcc	aaactgcacc	tacggatgca	ctgggcccagg	tcttgaaggc	tgtccaacga	2160
atgggcctaa	gatcccgtcc	atcgccactg	ggatgggtggg	ggccctcctc	ttgctgctgg	2220
tggtggccct	ggggatcggc	ctcttcatgc	gaaggcgcca	catcgttcgg	aagcgcacgc	2280
tgcggaggct	gctgcaggag	agggagcttg	tggagcctct	tacacccagt	ggagaagctc	2340
ccaaccaagc	tctcttgagg	atcttgaagg	aaactgaatt	caaaaagatc	aaagtgtctg	2400
gctccggtgc	gttcggcacg	gtgtataagg	gactctggat	cccagaaggt	gagaaagtta	2460
aaattcccgt	cgctatcaaa	acatctccga	aagccaacaa	ggaaatcctc	gatgaagcct	2520
acgtgatggc	cagcgtggac	aacccccacg	tgtgccgcct	gctgggcatc	tgccctcacct	2580
ccaccgtgca	gctcatcacg	cagctcatgc	ccttcggctg	cctcctggac	tatgtccggg	2640
aacacaaaga	caatattggc	tcccagtacc	tgctcaactg	gtgtgtgcag	atcgcaaagg	2700
gcatgaacta	cttggaggac	cgctcgcttg	tgcaccgcga	cctggcagcc	aggaacgtac	2760
tggtgaaaac	accgcagcat	gtcaagatca	cagatttttg	gctggccaaa	ctgctgggtg	2820
cggaagagaa	agaataccat	gcagaaggag	gcaaagtgcc	tatcaagtgg	atggcatttg	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	tgcgccaaag	ttccgtgagt	3120
tgatcatcga	attctccaaa	atggcccggag	acccccagcg	ctacctgttc	attcaggggg	3180
atgaaagaat	gcattttgcca	agtccctacg	actccaactt	ctaccgtgcc	ctgatggatg	3240
aagaagacat	ggacgacgtg	gtggatgccg	acgagtacct	catcccacag	cagggcttct	3300
tcagcagccc	ctccacgtca	cggactcccc	tctgagcttc	tctgagtgca	accagcaaca	3360
attccaccgt	ggcttgcat	gatagaaatg	ggctgcaaag	ctgtcccatc	aaggaagaca	3420
gcttcttgca	gcgatacagc	tcagacccca	caggcgccct	gactgaggac	agcatagacg	3480
acaccttctc	cccagtgctc	gaatacataa	acagtcctgt	tcccaaaagg	cccgtgggtc	3540
ctgtgcagaa	tctgtcttat	cacaatcagc	ctctgaaccc	cgcgcccagc	agagaccac	3600
actaccagga	ccccacagc	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
gcagtggaatt	tattggagca	tga				3863

<210> 743

<211> 3863

<212> ADN

5 <213> Homo sapiens

<400> 743

cccggcgag	cgcggcgca	gcagcctccg	ccccccgcac	ggtgtgagcg	cccagcgcg	60
ccgagggcg	cggagtccc	agctagcccc	ggcgccgcgc	gccgcccaga	ccggacgaca	120
ggccacctcg	tcggcgctcc	cccagatccc	cgcctcgccg	ccaacgccac	aaccaccg	180
cacggccccc	tgaactcgtc	cagtatgat	cgggagagcc	ggagcgagct	cttcggggag	240
cagcgatg	accctccggg	acggccgggg	cagcgctcct	ggcgctgctg	gctgcgctct	300
gcccggcgag	tcgggctctg	gaggaaga	aagtgtgcca	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
gctgggggtg	aggagaggag	aactgccaga	aactgaccaa	aatcatctgt	gcccagcagt	900
gctccggggc	gtgccgtggc	aagtcccca	gtgactgctg	ccacaaccag	tgtgtgacg	960
gctgcacagg	ccccggggag	agcgactgcc	tggctctgcc	caaattccga	gacgaagcca	1020
cgtgcaagga	cacctgcccc	ccactcatgc	tctacaaccc	caccacgtac	cagatggatg	1080
tgaaccccga	gggcaaatat	agctttgtgt	ccacctgcgt	gaagaagtgt	ccccgtaatt	1140
atgtggtgac	agatcacggc	tcgtgcgtcc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaagacgg	cgtccgcaag	tgtagaaggt	gcgaagggcc	ttgccgcaaa	gtgtgtaacg	1260
gaataggtat	tgggtgaattt	aaagactcac	tctccataaa	tgtacgaat	attaaacact	1320
tcaaaaactg	cacctccatc	agtggcgatc	tccacatcct	gccgggtggca	tttaggggtg	1380

```

actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcaggaa aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgaggggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggaggggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgccactac attgacggcc cccactgctg caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctgggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgtcc atcgccactg ggatgggtgg ggcctcctc ttgctgctgg 2220
tgggtggcct ggggatcgcc ctcttcatgc gaaggcgcca catcgttcgg aagcgacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccgt cgctatcaaa acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520
acgtgatggc cagcgtggac aacccccacg tgtgccgct gctgggcac tgcctcacct 2580
ccaccgtgca gctcatcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640
aacacaaaga caatattggc tcccagtagc tgctcaactg gtgtgtgcag atcgcaagg 2700
gcataaacta cttggaggag cgtcgcttgg tgaccgcga cctggcagcc aggaacgtac 2760
tggtgaaaaa 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 tcgcccagg ttccgtgagt 3120
tgatcatoga attctccaaa atggcccag agccccagcg ctacctgtc attcaggggg 3180
atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240
aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggttct 3300
tcagcagccc ctccacgtca cggactcccc tctgagctc tctgagtga accagcaaca 3360
attccaccgt ggcttgcatc gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420
gcttcttgca gcgatacagc tcagacccca caggcgctt gactgaggac agcatagacg 3480
acaccttctt cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgtgggt 3540
ctgtgcagaa tcctgtctat cacaatcagc ctctgaacc cgcgccagc agagaccac 3600
actaccagga cccccacagc actgcagtgg gcaacccga gtatctcaac actgtccagc 3660
ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggccagaaa ggcagccacc 3720
aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaaagcaa 3780
atggcatctt taagggtcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840
gcagtgaatt tattggagca tga 3863

```

<210> 744

<211> 3863

<212> ADN

5 <213> Homo sapiens

<400> 744

```

ccccggcgag cgcggccgca gcagcctccg cccccgcac ggtgtgagcg cccgacgagg 60
ccgagggcgg cggagtcctc agctagcccc ggcgcccgcc gccgccaga ccggacgaca 120
ggccacctcg tcggcgctcg cccgagtcct cgctcgccg ccaacgccac aaccaccgag 180
cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgctctct 300
gccccggcag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggctct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttcttct 480
taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagtcct atctaactat gatgcaataa aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcac ggcccgctgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tgagagagcat ccagtggcgg gacatagtca cagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgcaaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtg aggagaggag aactgcagaa aactgaccaa aatcatctgt gccacgagct 900

```

```

gctccgggag ctgccgtggc aagtccecca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccccggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
cgtgcaagga cacttgcccc ccactcatgc tctacaaccc caccacgtac cagatgggat 1080
tgaaccccca gggcaaatat agcttttggtg ccacctgctg gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tegtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tggtagaatt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cactccatc agtggcgatc tccacatcct gccgggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgttttg gacctccgtt cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccaggct tggcatgcct tgtgctcccc cgagggtctg tggggcccg 1800
agccagggga ctgcgtctct tggcggatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 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 gatcccgctc atcgccactg ggatgggtgg ggccctctct ttgctgctgg 2220
tgggtggcct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtgg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat ccagaaaggt gagaaagtta 2460
aaattcccg tgcctatcaa acatctccga aagccaacaa ggaaatctc gatgaagcct 2520
acgtgatggc cagcgtggac aaccccacg tgtgcgcct gctgggcac tgccctcacct 2580
ccaccgtgca accctgcacg cagctcatgc ccttcggctg cctcctggac tatgtccggg 2640
aacacaaaga caatattggc tcccagtagc tgctcaactg gtgtgtgcag atcgcaaagg 2700
gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760
tgggtgaaac accgcagcat gtcaagatca cagattttgg gctggccaaa ctgctgggtg 2820
cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880
aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac ggggtgactg 2940
tttgggagtt gatgacctt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000
cctccatcct ggagaaagga gaacgcctcc ctacgccacc catatgtacc atcgatgtct 3060

```

```

acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccagg ttccgtgagt 3120
tgatcatcga attctccaaa atggcccag agccccagcg ctacctgtgc attcaggggg 3180
atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240
aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggtcttct 3300
tcagcagccc ctccacgtca cggactcccc tctgagctc tctgagtcca accagcaaca 3360
attccaccgt ggcttgcatg gatagaaatg ggctgcaaag ctgtcccatc aaggaaagaca 3420
gcttcttgca gcgatacagc tcagacccca caggcgctt gactgaggac agcatagacg 3480
acaccttctt ccagtgctc gaatacataa accagtccgt tcccaaaaagg cccgctgggt 3540
ctgtgcagaa tcctgtctat cacaatcagc ctctgaaccc cgcccccagc agagaccac 3600
actaccagga cccccacagc actgcagtgg gcaaccocga gtatctcaac actgtccagc 3660
ccacctgtgt caacagcaca ttgcagagcc ctgcccactg ggccagaaa ggagccacc 3720
aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa gccaaagcaa 3780
atggcatctt taagggtcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840
gcagtgaatt tattggagca tga 3863

```

<210> 745

<211> 3863

<212> ADN

5 <213> Homo sapiens

<400> 745

```

cccggcgag cgcggcgca gcagctccg ccccccgcac ggtgtgagcg cccgacggg 60
ccgagggcgg cggagtccc agctagcccc ggccggcgcc gccggccaga ccggacgaca 120
ggccacctcg tcggcgctcc cccgagtcg cgctcgccg ccaacgccac aaccaccgag 180
caccgcccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccgcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360

```

ES 2 741 574 T3

```

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
tagcagtcct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcccggtgc ggttcagcaa caacctgcc ctgtgcaacg 720
tgagagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccagcagc 900
gctccggggcg ctgccgtggc aagtcctcca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccccggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccgga gggcaaatat agcttttggt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac aggtgtttttg tctgtcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tggtaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggtgtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactgaaaaa 1680
aactgttttg gacctccggt cagaaaaacca aaattataag caacagaggt gaaaaacagc 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agccagggga ctgctctctc tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgcc 1920
accagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggctctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggatgggtgg ggccctctc ttgctgctgg 2220
tgggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtgg 2400
gctccggtgc gttcggcacg gtgtataaag gactctggat ccagaaaggt gagaaagtta 2460
aaattcccgt cgttatcaaa acatctccga aagccaacaa ggaaatctc gatgaagcct 2520
acgtgatggc cagcgtggac aacccccacg tgtgcgcct gctgggcatc tgcctcacct 2580
ccaccgtgca gctcatcacg cagctcatgc cctcggctg cctcctggac tatgtccggg 2640
aacacaaaga caatattggc tcccagtagc tgctcaactg gtgtgtgcag atcgcaaagg 2700
gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc aggaacgtac 2760
tggtgaaaac accgcagcat gtcaagatca cagatttttg gctggccaaa ctgctgggtg 2820
cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg atggcattgg 2880
aatcaatttt acacagaatc agagtgatgt ctggagctac ggggtgactg 2940
tttgggagtt gatgacctt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000
cctccatcct ggagaaagga gaacgcctcc ctacgccacc catatgtacc atcgatgtct 3060
acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccagg ttccgtgagt 3120
tgatcatcga attctccaaa atggcccgag acccccagcg ctacctgtc attcaggggg 3180
atgaaagaat gcatattgcca agtcctacag actccaactt ctaccgtgcc ctgatggatg 3240
aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag cagggcttct 3300
tcagcagccc ctccacgtca cggactcccc tctgagctc tctgagtgc accagcaaca 3360
attccaccgt ggcttgcat gatagaaatg ggctgcaaa ctgtcccatc aaggaagaca 3420
gcttcttgca gcgatacagc tcagacccca caggcgctt gactgaggac agcatagacg 3480
acaccttctt cccagtgcct gaatacataa accagtccgt tcccaaaagg cccgtggct 3540
ctgtgcagaa tctgtctat cacaatcagc ctctgaaccc cgcgccagc agagaccac 3600
actaccagga cccccacagc actgcagtgg gcaaccccg gtatctcaac actgtccagc 3660
ccacctgtgt caacagcaca ttcgacagcc ttgcccactg ggccagaaa ggcagccacc 3720
aaattagcct ggacaaccct gactaccagc aggaacttct tcccaaggaa gccaaagcca 3780
atggcatctt taagggtcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840
gcagtgaatt tattggagca tga 3863

```

<210> 746

<211> 3863

<212> ADN

5 <213> Homo sapiens

<400> 746

```

ccccggcgag cggcgccgca gcagcctccg ccccccgca cgtgtgagcg cccgacgcgg 60
ccgagggcgcc cggagtcgcc agctagcccc ggccggccgccc cccgcccaga cccgacgaca 120
ggccacccctcg tcggcgctccg cccgagctccc cgcctcgcccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagttattgat cgggagagccc ggagcgagct ctcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccggcgag tcgggctctg gaggaagaaga aagtttgcga aggcacgagt aacaagctca 360
cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtcggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat cttccttct 480
taaagaccat ccaggaggtg gctggttatg tctcattg cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaat tcctatgcct 600
tagcagctct atctaactat gatgcaataa aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcccgctgc ggttcagcaa caacctgcc ctgtgcaacg 720
tggagagcat ccagtgccgg gcacatagtc gcagtgactt tctcagcaac atgtcgatgg 780
acttccagaa ccacctgggc agctgcaaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccgagct 900
gctcggtggc aatcctggc gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg cccccgggag agcgactgcc tggctgccc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaacccga gggcaatac agctttgggt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtcggtgat aatctcagga aactccccca gtgactgctg ggcgacagc tatgagatgg 1200
aggagagcgg cgtccgcaag tgaagaagt gcgaaggccc ttgccgcaa gtgtgtaacg 1260
gaataggat tggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaact cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggctag tttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgtg 1620
gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggg cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggcaggttc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agccagggga ctgctctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggaggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
acccagagt cctgctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagt tgccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctgggtctgga agtacgcaga cgcgggccat gtgtgccacc 2100
tgtccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gateccgtcc atcgccactg ggatgggtgg ggccctcctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcggttcgg aagcgcacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaag tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccg cgctatcaag acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520
acgtgatggc cagcgtggac aacccccacg tgtgccctgc gctgggcatc tgccctacct 2580
ccaccttgc gctcatcac cagctcatgc cctcggctg cctcctggac tatgtccggg 2640
aacacaaaga caatatggc tcccagttac tgcctcaact gtgtgtgcag atcgcaaagg 2700
gcatgaacta cttggaggac cgtcgtttgg tgcaccgcga cctggcagcc aggaacgtac 2760
tggtgaaaac accgcagcat gtcaagatca cagatttttg gctggccaaa ctgctgggtg 2820
cggaagagaa agaataccat gcagaaggag gcaaaagtgc tatcaagtgg atggcattgg 2880
aatcaatttt acacagaatc tataccacc agagtgtagt ctggagctac ggggtgactg 2940
tttgggagtt gatgacctt ggatccaagc catatgacgg aatccctgcc agcgagatct 3000
cctccatcct ggagaaagga gaacgcctc ctcagccacc catatgtacc atcgatgtct 3060
acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgcccagg ttccgtgagt 3120
tgatcatcga attctccaaa atggcccag acccccagc ctacctgtc attcaggggg 3180
atgaagaat gcatttgcca agtcctacag actccaact ctacctgcc ctgatggatg 3240

aagaagacat ggacgacgtg gtggatgccg acgagtacct catccacag cagggcttct 3300
tcagcagccc ctccacgtca cggactcccc tctgagctc tctgagtga accagcaaca 3360
attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc aaggaagaca 3420
gcttcttgca gcgatacagc tcagacccca caggcgctt gactgaggac agcatagacg 3480
acaccttct cccagtgcct gaatacataa accagtccgt tccccaaaag cccgtggtc 3540
ctgtgcagaa tctgtctat cacaatcagc ctctgaaccc cgcgcccagc agagaccac 3600
actaccagga cccccacagc actgcagtgg gcaaccccg gtatctcaac actgtccagc 3660
ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa ggcagccacc 3720

aaattagcct ggacaacct gactaccagc aggaattctt tccaaggaa gccaagccaa 3780
atggcatctt taagggtcc acagctgaaa atgcagaata cctaagggtc gcgccacaaa 3840
gcagtgaatt tattggagca tga 3863

```

<210> 747

<211> 3863

5 <212> ADN

<213> Homo sapiens

<400> 747

ccccggcgag	cgcgggccgca	gcagcctccg	ccccccgcac	ggtgtgagcg	cccgcgcggg	60
ccgagggcgcc	cggagtcgcc	agctagcccc	ggcgggccgcc	gcccggccaga	ccggacgcaca	120
ggccacctcg	tcggcggtccg	cccaggtccc	cgctcgcccg	ccaacgccac	aaccaccgag	180
cacggccccc	tgactccgtc	cagtattgat	cgggagagcc	ggagcgagct	cttcggggag	240
cagcgatgcg	accctccggg	acggccgggg	cagcgctcct	ggcgctgctg	gctgcgctct	300
gccccggcgag	tcggggtctg	gaggaaaaga	aagtttgcca	aggcacgagt	aacaagctca	360
cgcagttggg	cacttttgaa	gatcattttc	tcagcctcca	gaggatgttc	aataactgtg	420
aggtggtcct	tggaattttg	gaaattacct	atgtgcagag	gaattatgat	ctttccttct	480
taaagaccat	ccaggaggtg	gctggttatg	tcctcattgc	cctcaacaca	gtggagcgaa	540
ttcctttgga	aaacctgcag	atcatcagag	gaaatatgta	ctacgaaaat	tcctatgcct	600
tagcagtcct	atctaactat	gatgcaataa	aaaccggact	gaaggagctg	cccatgagaa	660
atttcacagga	aaacccgggag	ggcgccgtgc	ggttcagcaa	caaccctgcc	ctgtgcaacg	720
tgagagagcat	ccagtgccgg	gacatagtcg	gcagtgactt	tctcagcaac	atgtcgatgg	780
acttcacagaa	ccacctgggc	agctgccaaa	agtgatgatcc	aagctgtccc	aatgggagct	840
gctgggggtgc	aggagaggag	aactgccaga	aactgaccaa	aatcatctgt	gcccagcagt	900
gctccggggcg	ctgccgtggc	aagtccccc	gtgactgctg	ccacaaccag	tgtgtgcag	960
gctgcacagg	ccccggggag	agcgactgcc	tggtctgccg	caaattccga	gacgaagcca	1020
cgtgcaagga	cacctgcccc	ccactcatgc	tctacaacc	caccacgtac	cagatggatg	1080
tgaaccccgga	gggcaatac	agctttggtg	ccacctgcgt	gaagaagtgt	ccccgtaatt	1140
atgtggtgac	agatcacggc	tcgtgcgtcc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaagacgg	cgtccgcgaag	tgtaagaagt	gcgaagggcc	ttgccgcaaa	gtgtgtaacg	1260
gaatagggtat	ttggtgaattt	aaagactcac	tctccataaa	tgctacgaat	attaaacact	1320
tcaaaaactg	cacctccatc	agtggcgatc	tccacatcct	gcccgtggca	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	gattacgtcc	cctcaaggag	ataagtgatg	1620
gagatgtgat	aatttcagga	aacaaaaatt	tggtctatgc	aaatacaata	aactggaaaa	1680
aactgttttg	gacctccggt	cagaaaacca	aaattataag	caacagaggt	gaaaaacagct	1740
gcaaggccac	aggccaggtc	tgccatgcct	tggtctcccc	cgagggctgc	tgggggcccg	1800
agccacggga	ctgcgtctct	tgccggaatg	tcagccgagg	cagggaatgc	gtggacaagt	1860
gcaaccttct	ggagggtgag	ccaaggaggt	ttgtggagaa	ctctgagtgc	atacagtgc	1920
accagagtg	cctgcctcag	gccatgaaca	tcacctgcac	aggacgggga	ccagacaact	1980
gtatccagt	tgcccactac	attgacggcc	cccactgcgt	caagacctgc	ccggcaggag	2040
tcattgggaga	aaacaacacc	ctggtctgga	agtagcgaga	cgccggccat	gtgtgccacc	2100
tgtgccatcc	aaactgcacc	tacggatgca	ctgggccagg	tcttgaaggc	tggtccaacga	2160
atgggacctaa	gatcccgctc	atcgccactg	ggatgggtgg	ggccctctc	ttgtctgtgg	2220
tggtggccct	ggggatcgcc	ctcttcattg	gaaggcgcca	catcggtcgg	aagcgacgc	2280
tgccgagggt	gctgcaggag	agggagcttg	tggaacctct	tacacccagt	ggagaagctc	2340
ccaaccaagc	tctcttgagg	atcttgaagg	aaactgaatt	caaaaagatc	aaagtgtctg	2400
gctccgggtgc	gttcggcacg	gtgtataagg	gactctggat	cccagaaggt	gagaaagtta	2460
aaattcccgt	cgtctatcaag	acatctccga	aagccaacaa	ggaaatcctc	gatgaagcct	2520
acgtgatggc	cagcgtggac	aacccccacg	tggtccgcct	gctgggcatc	tgctcacct	2580
ccaccgtgca	gctcatcacg	cagctcatgc	ccttcggctg	cctcctggac	tatgtccggg	2640
aacacaaaga	caatattggc	tcccagtacc	tgctcaactg	gtgtgtgcag	atcgcaaagg	2700
gcattgaacta	cttgagggac	cgctcgcttg	tgaccgcga	cctggcagcc	aggaacgtac	2760
tggtgaaaac	accgcagcat	gtcaagatca	cagattttgg	gctggccaaa	ctgctgggtg	2820
cggaagagaa	agaataccat	gcagaaggag	gcaaagtgcc	tatcaagtgg	atggcatttg	2880
aatcaatttt	acacagaatc	tataccacc	agagtgatgt	ctggagctac	ggggtgactg	2940
tttgggagtt	gatgaccttt	ggatccaagc	catatgacgc	aatccctgcc	agcgagatct	3000
cctccatcct	ggagaaagga	gaacgcctcc	ctcagccacc	catatgtacc	atcgatgtct	3060
acatgatcat	ggtcaagtgc	tggtgatag	acgcagatag	tcgcccacaa	ttccgtgagt	3120
tgatcatcga	attctccaaa	atggcccag	acccccagcg	ctacctgtgc	attcaggggg	3180
atgaaagaat	gcatttgcca	agtccctacg	actccaactt	ctaccgtgcc	ctgatggatg	3240
aagaagacat	ggacgacgtg	gtggatgccg	acgagtacct	catcccacag	cagggcttct	3300
tcagcagccc	ctccacgtca	cggactcccc	tcctgagctc	tctgagtgcg	accagcaaca	3360
attccaccgt	ggcttgcat	gatagaaatg	ggctgcaaa	ctgtcccatc	aagggaagaca	3420
gcttcttgca	gcgatacagc	tcagacccca	caggcgctct	gactgaggac	agcatagacg	3480
acaccttctc	ccagtgccct	gaatacataa	accagtcctg	tcccaaaaagg	cccgtgtggc	3540
ctgtgcagaa	tcctgtctat	cacaatcagc	ctctgaaccc	cgcgcccagc	agagaccac	3600
actaccagga	ccccacagc	actgcagtgg	gcaaccccga	gtatctcaac	actgtccagc	3660
ccacctgtgt	caacagcaca	ttcgacagcc	ctgcccactg	ggcccagaaa	ggcagccacc	3720
aaattagcct	ggacaaccct	gactaccagc	aggacttctt	tcccaaggaa	gccaaagcaa	3780
atggcatctt	taagggtctc	acagctgaaa	atgcagaata	cctaagggtc	gcgccacaaa	3840
gcagtgaatt	tattggagca	tga				3863

<210> 748

5 <211> 3863

<212> ADN

<213> Homo sapiens

<400> 748

```

cccggcgcgag cgcggcgcgca gcagcctccg cccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggcggc cggaggtccc agctagcccc ggcgggcgcc gccgcccaga ccggacgaca 120
ggccacctcg tcggcggtccg cccgagtcgc cgctcgccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccgcgag 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
tagcagtcct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcac ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tgagagcagc ccagtggcgg gacatagtca gcagtgaact tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccacgagct 900
gctccggggc ctgccgtggc aagtccccc gtgactgctg ccacaaccag tgtgtgcag 960
gctgcacagg ccccggggag agcgactgcc tggctcgccg caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccca gggcaatac agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tggggaattt aaagactcac tctcataaaa tgctacgaat attaaacact 1320
tcaaaaactg cactccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggctag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaaacagct 1740
gcaagggcac aggccaggtc tgccatgcct tgtgtcctcc caggggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggcggggga ccagacaact 1980
gtatccagtg tgccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcattgggaga aaacaacacc ctgggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccag tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggtgggtggg ggccctcctc ttgtgctgg 2220
tggtggccct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat ccagaaaggt gagaaagtta 2460
aaattcccgt cgtatcaag acatctccga aagccaacaa ggaaatcctc gatgaagcct 2520
acgtgatggc cagcgtggac aacccccacg tgtgcgcct gctgggcatc tgcctcacct 2580
ccaccgtgca gctcatcacg cagctcatgc cctcggctg cctcctggac tatgtccggg 2640
aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag atcgcaaagg 2700
gcatgaacta cttggaggac cgtcgcttgg tgaccgcga cctggcagcc aggaacgtac 2760

```

tggtgaaaac	accgcagcat	gtcaagatca	cagatTTTTg	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	ggagaaaagg	gaacgcctcc	ctcagccacc	catatgtacc	atcgatgtct	3060
acatgatcat	ggtcaagtgc	tggatgatag	acgcagatag	tcgccc aaag	ttccgtgagt	3120
tgatcatcga	attctccaaa	atggcccag	accccagcg	ctacctgtc	attcaggggg	3180
atgaaagaat	gcatttgcca	agtcctacag	actccaactt	ctaccgtgcc	ctgatggatg	3240
aagaagacat	ggacgacgtg	gtggatgccg	acgagtacct	catcccacag	cagggcttct	3300
tcagcagccc	ctccacgtca	cggactcccc	tctgagctc	tctgagtgc	accagcaaca	3360
attccaccgt	ggcttgcat	gatagaaatg	ggctgcaaag	ctgtcccatc	aaggaagaca	3420
gcttcttgca	gcgatacagc	tcagacccca	caggcgcctt	gactgaggac	agcatagacg	3480
acaccttctt	cccagtgcct	gaatacataa	accagtccgt	tcccaaaagg	cccgtgggt	3540
ctgtgcagaa	tcctgtctat	cacaatcagc	ctctgaaccc	cgcgccacg	agagaccac	3600
actaccagga	ccccacagc	actgcagtgg	gcaaccccg	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> ADN

5 <213> Homo sapiens

<400> 749

ccccggcgag	cgcgggcgca	gcagcctccg	ccccccgcac	ggtgtgagcg	cccgcgcggg	60
ccgagggcgg	cggagtcccc	agctagcccc	ggcgcccgcc	gccgcccaga	ccggacgaca	120
ggccacctcg	tcggcgctcc	cccagatccc	cgctcgccg	ccaacgccac	aaccaccgcg	180
cacggccccc	tgactccgtc	cagtattgat	cgggagagcc	ggagcgagct	cttcggggag	240
cagcgatgcg	accctccggg	acggcccggg	cagcgctcct	ggcgctgctg	gctgcgtctt	300
gccccggcag	tcgggctctg	gaggaaaaga	aagtttgcca	aggcacgagt	aacaagctca	360
cgcagttggg	cacttttgaa	gatcattttc	tcagcctcca	gaggatgttc	aataactgtg	420
aggtggtcct	tgggaatttg	gaaattacct	atgtgcagag	gaattatgat	ctttccttct	480
taaagaccat	caaggaggtg	gctggttatg	tcctcattgc	cctcaacaca	gtggagcgaa	540
ttcctttgga	aaacctgcag	atcatcagag	gaaatatgta	ctacgaaaat	tcctatgcct	600
tagcagctct	atctaactat	gatgcaaata	aaaccggact	gaaggagctg	cccatgagaa	660
atttacagga	aatcctgcat	ggcgccgtgc	ggttcagcaa	caaccctgcc	ctgtgcaacg	720
tggagagcat	ccagtggcgg	gacatagtca	gcagtgactt	tctcagcaac	atgtcgatgg	780
acttcagaaa	ccacctgggc	agctgccaaa	agtgtgatcc	aagctgtccc	aatgggagct	840
gctgggggtg	aggagaggag	aactgccaga	aactgaccaa	aatcatctgt	gccagcagct	900
gctccggggc	gtccgctggc	aagtccccc	gtgactgctg	ccacaaccag	tgtgctgcag	960
gctgcacagg	ccccggggag	agcgactgcc	tggctctgcc	caaattccga	gacgaagcca	1020
cgtgcaagga	cacctgcccc	ccactcatgc	tctacaaccc	caccacgtac	cagatggatg	1080
tgaaccccga	gggcaaatat	agctttggtg	ccacctgcgt	gaagaagtgt	ccccgtaatt	1140
atgtggtgac	agatcacggc	tcgtgcgtcc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaagacgg	cgtccgcaag	tgtaaagaat	gcgaagggcc	ttgccgcaaa	gtgtgtaacg	1260
gaataggtat	tgggtgaattt	aaagactcac	tctccataaa	tgtacgaat	attaaacact	1320
tcaaaaactg	cacctccatc	agtggcgatc	tccacatcct	gccgggtggc	tttaggggtg	1380
actccttcac	acatactcct	cctctggatc	cacaggaact	ggatattctg	aaaaccgtaa	1440
aggaatcac	agggtttttg	ctgattcagg	cttggcctga	aaacaggacg	gacctccatg	1500
cctttgagaa	cctagaaatc	atacgcgga	ggaccaagca	acatggtcag	ttttctcttg	1560
cagtcgtcag	cctgaacata	acatccttgg	gattacgctc	cctcaaggag	ataagtgatg	1620
gagatgtgat	aatttcagga	aacaaaaatt	tgtgctatgc	aaatacaata	aactggaaaa	1680
aactgttttg	gacctccggg	cagaaaacca	aaattataag	caacagaggt	gaaaaacagct	1740
gcaaggccac	aggccaggtc	tgccatgcct	tgtgctcccc	caggggctgc	tggggccccg	1800
agcccaggga	ctgcgtctct	tgccggaatg	tcagccgagg	caggggaatg	gtggacaagt	1860
gcaaccttct	ggagggtgag	ccaaggaggt	ttgtggagaa	ctctgagtgc	atacagtgcc	1920
acccagagtg	cctgcctcag	gccatgaaca	tcacctgcac	aggacgggga	ccagacaact	1980
gtatccagtg	tgcccactac	attgacggcc	ccactgcgt	caagacctgc	ccggcaggag	2040
tcattgggga	aaacaacacc	ctggtctgga	agtacgcaga	cgcggcccat	gtgtgccacc	2100
tgtgccatcc	aaactgcacc	tacggatgca	ctgggccagg	tcttgaaggc	tgtccaacga	2160
atgggcctaa	gatcccgtcc	atcgccactg	ggatggtggg	ggccctcctc	ttgctgctgg	2220

tgggtggccct	ggggatcggc	ctcttcatgc	gaaggcgcca	catcggttcgg	aagcgcacgc	2280
tgcggaggct	gctgcaggag	agggagcttg	tgagcctct	tacaccagct	ggagaagctc	2340
ccaaccaagc	tctcttgagg	atcttgaagg	aaactgaatt	caaaaagatc	aaagtgcctg	2400
gctccggtgc	gttcggcacg	gtgtataagg	gactctggat	cccagaaggt	gagaaagtta	2460
aaattcccg	cgctatcaag	gttccgaaag	ccaacaagga	aatcctcgat	gaagcctacg	2520
tgatggccag	cgtggacaac	ccccacgtgt	gccgcctgct	gggcatctgc	ctcacctcca	2580
ccgtgcagct	catcacgcag	ctcatgccct	tcggctgcct	cctggactat	gtccgggaac	2640
acaaagacaa	tattggctcc	cagtacctgc	tcaactgggtg	tgtgcagatc	gcaaagggca	2700
tgaactactt	ggaggaccgt	cgcttgggtg	accgcgacct	ggcagccagg	aacgtactgg	2760
tgaaaacacc	gcagcatgtc	aagatcacag	attttgggct	ggccaaactg	ctgggtgcgg	2820
aagagaaaga	ataccatgca	gaaggaggca	aagtgcctat	caagtggatg	gcattggaat	2880
caattttaca	cagaatctat	accaccaga	gtgatgtctg	gagctacggg	gtgactgttt	2940
gggagttgat	gacctttgga	tccaagccat	atgacggaat	ccctgccagc	gagatctcct	3000
ccatcctgga	gaaaggagaa	cgcctccctc	agccacccat	atgtaccatc	gatgtctaca	3060
tgatcatggg	caagtgcctg	atgatagacg	cagatagtcg	cccaaagttc	cgtgagttga	3120
tcatcgaatt	ctccaaaatg	gcccgagacc	cccagcgcta	ccttgtcatt	caggggggatg	3180
aaagaatgca	tttgccaagt	cctacagact	ccaacttcta	ccgtgccctg	atggatgaag	3240
aagacatgga	cgacgtgggt	gatgccgacg	agtacctcat	cccacagcag	ggcttcttca	3300
gcagccctcc	caagtcacgg	actccctccc	tgagctctct	gagtgcaccc	agcaacaatt	3360
ccaccgtggc	ttgcattgat	agaaatgggc	tgcaaagctg	tcccatcaag	gaagacagct	3420
tcttgcagcg	atacagctca	gacccacacg	gcgccttgac	tgaggacagc	atagacgaca	3480
ccttccctccc	agtgcctgaa	tacataaacc	agtcogttcc	caaaaggccc	gctggctctg	3540
tgcagaatcc	tgtctatcac	aatcagcctc	tgaaccccg	gcccagcaga	gacccacact	3600
accaggaccc	ccacagcact	gcagtgggca	accccgagta	tctcaacact	gtccagccca	3660
cctgtgtcaa	cagcacattc	gacagccctg	cccactgggc	ccagaaaggc	agccaccaaa	3720
ttagcctgga	caaccctgac	taccagcagg	acttctttcc	caaggaagcc	aagccaaatg	3780
gcacctttta	gggctccaca	gctgaaaatg	cagaatacct	aagggtcgcg	ccacaaagca	3840
gtgaatttat	tggagcatga					3860

<210> 750

<211> 3869

<212> ADN

5 <213> Homo sapiens

<400> 750

cccggcgag	cgcggccgca	gcagcctccg	ccccccgcac	ggtgtgagcg	cccagcgagg	60
ccgagggcgg	cggagtcctc	agctagcccc	ggcgcccgcc	gccgcccaga	ccggacgaca	120
ggccaccccg	tgcgcgtccg	cccaggtccc	cgctctcgccg	ccaacgccac	aaccaccgag	180
cacggccccc	tgactccgtc	cagtattgat	cgggagagcc	ggagcgagct	cttcggggag	240
cagcgatgcg	acctccggg	acggccgggg	cagcgctcct	ggcgctgctg	gctgcgctct	300
gcccggcgag	tcgggctctg	gaggaaaaga	aagtttgcca	aggcacgagt	aacaagctca	360
gcgagttggg	cacttttgaa	gatcattttc	tcagcctcca	gaggatgttc	aataactgtg	420
aggtggtcct	tggaattttg	gaaattacct	atgtgcagag	gaattatgat	ctttccttct	480
taaagaccat	ccaggaggtg	gctggttatg	tcctcattgc	cctcaacaca	gtggagcgaa	540
ttccttttga	aaacctgcag	atcatcagag	gaaatatgta	ctacgaaaat	tcctatgcct	600
tagcagtcct	atctaactat	gatgcaaaata	aaaccggact	gaaggagctg	cccatgagaa	660
atttacagga	aatcctgcat	ggcgccgtgc	ggttcagcaa	caaccctgcc	ctgtgcaacg	720
tggagagcat	ccagtgccgg	gacatagtca	gcagtgactt	tctcagcaac	atgtcgatgg	780
acttccagaa	ccacctgggc	agctgccaaa	agtgatgatc	aagctgtccc	aatgggagct	840
gctgggggtg	aggagaggag	aactgccaga	aactgaccaa	aatcatctgt	gcccagcagt	900
gctccggggg	ctgccgtggc	aagtccccca	gtgactgctg	ccacaaccag	tgtgctgcag	960
gctgcacagg	ccccgggag	agcgactgcc	tggtctgccc	caaattccga	gacgaagcca	1020
cgtgcaagga	cacctgcccc	ccactcatgc	tctacaaccc	caccacgtac	cagatggatg	1080
tgaacccoga	gggcaaatat	agctttgggtg	ccacctgcgt	gaagaagtgt	ccccgtaatt	1140
atgtggtgac	agatcacggc	tcgtgcgtcc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaagacgg	cgtccgcgaag	tgtaagaagt	gcgaagggcc	ttgccgcaaa	gtgtgtaacg	1260
gaataggtat	tgggtgaattt	aaagactcac	tctccataaa	tgctacgaat	attaacacct	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	ataagtgtatg	1620
gagatgtgat	aatttcagga	aacaaaaatt	tgtgctatgc	aaatacaata	aactggaaaa	1680
aactgttttg	gacctccggt	cagaaaaacca	aaattataag	caacagaggt	gaaaaacagct	1740

gcaagggccac aggccagggtc tgccatgcct tgtgctcccc cgaggggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgc 1920
accagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctgggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgtcc atcgccactg ggatgggtgg ggccctctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcctgc gaaggcgcca catcgttcgg aagcgacgc 2280
tgccggaggct gctgcaggag agggagcttg tggagcctct tacacccagt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgcctg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat ccagaaaggt gagaaagtta 2460
aaattcccgt cgctatcaag gaaccaacat ctccgaaagc caacaaggaa atcctcgatg 2520
aagcctacgt agatggcagc gtggacaacc cccacgtgtg ccgcctgctg ggcatctgcc 2580
tcacctccac cgtgcagctc atcacgcagc tcatgccctt cggctgcctc ctggactatg 2640
tccgggaaca caaagacaat attggctccc agtacctgct caactgggtg gtgcagatcg 2700
caaagggcat gaactacttg gaggaccgtc gcttgggtgca ccgcgacctg gcagccaggga 2760
acgtactggt gaaaacaccg cagcatgtca agatcacaga ttttgggctg gccaaactgc 2820
tggtgctcga agagaaagaa taccatgcag aaggaggcaa agtgcctatc aagtggatgg 2880
cattggaatc aattttacac agaattctata cccaccagag tgatgtctgg agctacgggg 2940
tgactgtttg ggagttgatg acctttggat ccaagccata tgacggaatc cctgccagcg 3000
agatctcctc catcctggag aaaggagaac gcctccctca gccacccata tgtaccatcg 3060
atgtctacat gatcatggtc aagtgcctgga tgatagacgc agatagtcgc ccaaagtctc 3120
gtgagttgat cctcgaattc tccaaaatgg cccgagaccc ccagcgctac cttgtcattc 3180
agggggatga aagaatgcac ttgccaaagc ctacagactc caacttctac cgtgccctga 3240
tggatgaaga agacatggac gacgtggtgg atgccgacga gtacctcatc ccacagcagg 3300
gcttcttcag cagcccctcc acgtcacgga ctcccctcct gagctctctg agtgcaacca 3360
gcaacaattc caccgtggct tgcatgtgata gaaatgggct gcaaagctgt cccatcaagg 3420
aagacagctt cttgcagcag tacagctcag accccacagg cgccttgact gaggacagca 3480
tagacgacac cttcctccca gtgcctgaat acataaacca gtccgttccc aaaaggcccg 3540
ctggctctgt gcagaatcct gtctatcaca atcagcctct gaaccccgcg cccagcagag 3600

accacacta 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
caciaagcag tgaatttatt ggagcatga 3869

<210> 751

<211> 3866

<212> ADN

5 <213> Homo sapiens

<400> 751

cccgccgag cgccggccga gcagcctccg cccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggcgg cggagtcccc agctagcccc ggccggccgc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctcc cccgagtcgc cgctcgccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgct cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gcccggcgag tcgggctctg gaggaaaaga aagtgtgcca aggcacgagt aacaagctca 360
cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggctct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggagggt gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggccgctgac ggttcagcaa caaccctgcc ctgtgcaacg 720
tgagagcat ccagtggcgg gacatagtca cagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtg aggagaggag aactgccaga aactgaccaa aatcatctgt gccagcagct 900
gctccggggc ctgcccgtggc aagtccccca gtgactgctg ccacaaccag tgtgtgcag 960
gctgcacagg ccccccggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
cgtgcaaggga cactgcctcc cactcatgc tctacaacct caccacgtac cagatggatg 1080
tgaaccccga gggcaaatat agctttgggt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtgggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

```

aggaagacgg cgcccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tgggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cactctccatc agtggcgatc tccacatcct gccgggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggtttttg 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 gacctccggg cagaaaacca aaattataag caacagaggt gaaaaacagct 1740
gcaaggccac agggcagggt tgccatgcct tgtgctcccc cgagggtgctg tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccagggt attgacggcc cccactgctg caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggctggga agtgcgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgtcc atcgccactg ggatgggtgg ggcctctctc ttgctgctgg 2220
tgggtggcct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccgagggt cctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat ccagaagggt gagaaagtta 2460
aaattcccgt cgttatcaag gaacctctc cgaaagccaa caaggaaatc ctcgatgaag 2520
octacgtgat ggcacgctg gacaaccccc acgtgtgccc cctgctgggc atctgctca 2580
cctccaccgt gcaagctcac acgcagctca tgcccttcgg ctgcctcctg gactatgtcc 2640
gggaacacaa agacaatatt ggctcccagt acctgctcaa ctggtgtgtg cagatcgcaa 2700
agggcatgaa ctacttgagg gaccgtcgct tgggtgcacc cgacctggca gccaggaaac 2760
tactggtgaa aacaccgcag catgtcaaga tcacagattt tgggctggcc aaactgctgg 2820
gtgcggaaga gaaagaatac catgcagaag gaggcaaagt gcctatcaag tggatggcat 2880
tggaatcaat ttacacaga atctatcccc accagagtga tgtctggagc tacggggtga 2940
ctgtttggga gttgatgacc tttggatcca agccatatga cggaatccct gccagcgaga 3000
tctcctccat cctggagaaa ggagaacgcc tccctcagcc acccatatgt accatcgatg 3060
tctacatgat catggtcaag tgctggatga tagacgcaga tagtcgcca aagttccgtg 3120
agttgatcat cgaattctcc aaaatggccc gagaccccca gcgctacctt gtcattcagg 3180
gggatgaaag aatgcatttg ccaagtccta cagactccaa cttctaccgt gccctgatgg 3240
atgaagaaga catggacgac gtgggtgatg ccgacgagta cctcatccca cagcagggtc 3300
tcttcagcag ccctccacg tcacggactc cctcctgag ctctctgagt gcaaccagca 3360
acaattccac cgtggcttgc attgatagaa atgggctgca aagctgtccc atcaaggaa 3420
acagcttctt gcagcgatac agctcagacc ccacaggcgc cttgactgag gacagcatag 3480
acgacacctt cctcccagtg cctgaataca taaaccagtc cgttcccaaa aggcccgtg 3540
gctctgtgca gaatcctgtc tatcacaatc agcctctgaa ccccgcgccc agcagagacc 3600
cacactacca ggacccccac agcactgcag tgggcaaccc cgagtatctc aacctgtcc 3660
agcccacctg tgtcaacagc acattcgaca gcctgcccc ctgggcccag aaaggcagcc 3720
accaaattag cctggacaac cctgactacc agcaggactt ctttcccaag gaagccaagc 3780
caaattggcat ctttaagggc tccacagctg aaaatgcaga atacctaagg gtcgcgccac 3840
aaagcagtga atttattgga gcatga 3866

```

<210> 752

<211> 3854

<212> ADN

5 <213> Homo sapiens

<400> 752

```

cccggcgacg cgccggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacggcg 60
ccgaggcggc cggagtcccc agctagcccc ggccggccgc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctccg cccgagtcct cgctcgccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagtattgat cggagagacc ggagcgagct cttcggggag 240
cagcgatgcg accctccggg acggcctcct cagcgctcct ggcgctgctg gctcgctct 300
gccccggcag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttctctct 480
taaagaccat ccaggagggt gctggttatg tctcatttgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgcgctgc ggttcagcaa caaccctgcc ctgtgcaacg 720

```

tggagagcat	ccagtggcgg	gacatagtca	gcagtgactt	tctcagcaac	atgtcgatgg	780
acttccagaa	ccacctgggc	agctgccaaa	agtgatgatcc	aagctgtccc	aatgggagct	840
gctgggggtgc	aggagaggag	aactgccaga	aactgaccaa	aatcatctgt	gcccagcagt	900
gctccggggcg	ctgcggtggc	aagtccccc	gtgactgctg	ccacaaccag	tgtgctgcag	960
gctgcacagg	ccccgggag	agcgactgcc	tggtctgccg	caaattccga	gacgaagcca	1020
cgtgcaagga	cacctgcccc	ccactcatgc	tctacaaccc	caccacgtac	cagatggatg	1080
tgaaccccg	gggcaaatac	agctttggtg	ccacctgctg	gaagaagtgt	ccccgtaatt	1140
atgtggtgac	agatcacggc	tcgtgctgcc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaagacgg	cgctccgcaag	tgtaagaagt	gcgaagggcc	ttgccgcaa	gtgtgtaacg	1260
gaatagggtat	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	atacgcgga	ggaccaagca	acatggtcag	ttttctcttg	1560
cagtcgtcag	cctgaacata	acatccttgg	gattacgctc	cctcaaggag	ataagtgatg	1620
gagatgtgat	aatttcagga	aacaaaaatt	tgtgctatgc	aaatacaata	aactggaaaa	1680
aactgttttg	gacctccggt	cagaaaacca	aaattataag	caacagaggt	gaaaaacagt	1740
gcaaggccac	aggccaggtc	tgccatgcct	tgtgctcccc	cgagggtctg	tggggcccg	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
tcattgggaga	aaacaacacc	ctggtctgga	agtacgcaga	cgccggccat	gtgtgccacc	2100
tgtgccatcc	aaactgcacc	tacggatgca	ctgggcccag	tcttgaaggc	tgtccaacga	2160
atgggacctaa	gatcccgctc	atcgccactg	ggatgggtgg	ggccctcctc	ttgctgctgg	2220
tggtggccct	ggggatcggc	ctcttcctgc	gaaggcgcca	catcgttcgg	aagcgacgc	2280
tgccgaggct	gctgcaggag	agggagcttg	tggagcctct	tacacccagt	ggagaagctc	2340
ccaaccaagc	tctcttgagg	atcttgaagg	aaactgaatt	caaaaagatc	aaagtgtctg	2400
gctccggtgc	gttcggcacg	gtgtataagg	gactctggat	cccagaaggt	gagaaagtta	2460
aaattcccg	cgctatcaag	gaattaagag	aagcaaccct	cgatgaagcc	tacgtgatgg	2520
ccagcgtgga	caacccccac	gtgtgccgcc	tgctgggcat	ctgcctcacc	tccaccgtgc	2580
agctcatcac	gcagctcatg	cccttcggct	gcctcctgga	ctatgtccgg	gaacacaaag	2640
acaattttg	ctcccagtat	ctgctcaact	ggtgtgtgca	gatcgcaaag	ggcatgaact	2700
acttgaggga	ccgtcgcttg	gtgcaccgcg	acctggcagc	caggaacgta	ctggtgaaaa	2760
caccgcagca	tgtcaagatc	acagattttg	ggctggccaa	actgctgggt	gcggaagaga	2820
aagaatacca	tgcaagaagg	ggcaaaagtgc	ctatcaagtgc	gatggcattg	gaatcaattt	2880
tacacagaat	ctatacccac	cagagtgtatg	tctggagcta	cggggtgact	gtttgggagt	2940
tgatgacctt	tggtaccaag	ccatatgacg	gaatccctgc	cagcgagatc	tcctccatcc	3000
tggagaaagg	agaacgcctc	cctcagccac	ccatatgtac	catcgatgtc	tacatgatca	3060
tggtcaagtgc	ctggatgata	gacgcagata	gtcgcccaaa	gttcctgtgag	ttgatcatcg	3120
aattctccaa	aatggcccga	gacccccagc	gctaccttgt	cattcagggg	gatgaaagaa	3180
tgcatttgcc	aagtcctaca	gactccaact	tctaccgtgc	cctgatggat	gaagaagaca	3240
tggaacgacgt	ggtggatgcc	gacgagtacc	tcatcccaca	gcagggtctc	ttcagcagcc	3300
cctccacgtc	acggactccc	ctcctgagct	ctctgagtgc	aaccagcaac	aattccaccg	3360
tggtctgcat	tgatagaaat	gggctgcaaa	gctgtcccat	caaggaagac	agcttcttgc	3420
agcgatacac	ctcagacccc	acaggcgccct	tgactgagga	cagcatagac	gacaccttcc	3480
tcccagtgc	tgaatacata	aaccagtccg	tcccaaaaag	gcccgtggc	tctgtgcaga	3540
atcctgtcta	tcacaatcag	cctctgaacc	ccgcggccag	cagagaccga	cactaccagg	3600
acccccacag	cactgcagtgc	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> ADN

5 <213> Homo sapiens

<400> 753

ccccggcgag	cgcgggcgca	gcagcctccg	ccccccgcac	ggtgtgagcg	cccagcgcg	60
ccgagggcgg	cggagtcccc	agctagcccc	ggcgcccgcc	gccgcccaga	ccggacgaca	120
ggccacctcg	tcggcgctcc	cccagatccc	cgctcgcccg	ccaacgccac	aaccaccg	180

```

cacggccccc tgactccgtc cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gcccggcgag tcgggctctg gaggaagaag aagtttgcca aggcacgagt aacaagctca 360
cgagttggg 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 aatcctcgcat ggcgccgtgc ggttcagcaa caacctgccc ctgtgcaacg 720
tgagagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc agggagaggag aactgccaga aactgaccaa aatcatctgt gccagcagct 900
gctccggggc agtcccggtg cctctgcctc gactgctgct ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcgactgcc tggctgctcg caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccaactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccg gggcaaatac agcttttggtg ccacctgctg gaagaagtgt ccccgtaatt 1140
atgtggtgac aatttcaggg tcgtgcctgc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaatagggtat tgggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac aactactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaatac atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtgcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga aacaaaaatt tgtgctatgc aaatacaata aactgaaaaa 1680
aactgtttgg gacctccggt cagaaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac agggcagggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agccagggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaacctctc ggaggggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggcggggga ccagacaaact 1980
gtatccagtg tgcccactac attgacggcc cccactgctg caagacctgc cggcgaggag 2040
tcatgggaga aaacaacacc ctgggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccag tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggatgggtgg ggccctcctc ttgctgctgg 2220
tgggtggcct ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgccgaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccgggtg gttccgcacg gtgtataagg gactctggat ccagaaaggt gagaaagtta 2460
aaattcccg cgtatcaag gaattaagag aagcaaacct cgatgaagcc tacgtgatgg 2520
ccagcgtgga caacccccac gtgtgccgcc tggctgggcat ctgcctcacc tccaccgtgc 2580
agctcatcac gcagctcatg cccttcggct gccctcctgga ctatgtccgg gaacacaaa 2640
acaatattgg ctcccagtc ctgctcaact ggtgtgtgca gatcgcaaa ggcatgaact 2700
acttgaggga ccgtcgcttg gtgcaccgag acctggcagc caggaaagta ctggtgaaa 2760
caccgcagca tgtcaagatc acagattttg ggctggccaa actgctgggt gccgaagaga 2820
aagaatacca tgcagaagga ggcaagtgc ctatcaagtg gatggcattg gaatcaattt 2880
tacacagaat ctataccac cagagtgtg tctggagcta cggggtgact gtttgggagt 2940
tgatgacctt tggatccaag ccatatgacg gaatccctgc cagcgagatc tcctccatcc 3000
tggaagaaag agaacgcctc cctcagccac ccatatgtac catcgatgtc tacatgatca 3060
tggtcaagtg ctggatgata gacgcagata gtcgccccaa gttccgtgag ttgatcatcg 3120
aattctccaa aatggcccg gacccccagc gctacctgt cattcagggg gatgaaagaa 3180
tgcatttgac aagtcctaga gactccaact tctaccgtgc cctgatggat gaagaagaca 3240
tggaagacgt ggtggatgcc gacgagtacc tcatcccaca gcagggcttc ttcagcagcc 3300
cctccacgtc acggactccc ctctgagct ctctgagtgc aaccagcaac aattccaccg 3360
tggcttgcat tgatagaaat gggctgcaaa gctgtcccat caaggaagac agcttcttgc 3420
agcgatacat ctacagcccc acaggcgctt gactgagga 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 attcgacag cctgccact gggcccagaa aggcagccac caaattagcc 3720
tggaacaacc tgactaccag caggacttct ttcccaagga agccaagcca aatggcatct 3780
ttaagggtc cacagctgaa aatgcagaat acctaagggt cgcgcacaaa agcagtgaat 3840
ttattggagc atga 3854

```

<210> 754

<211> 3887

<212> ADN

5 <213> Homo sapiens

<400> 754

```

ccccggcgag cgcggcgccga gcagcctccg ccccccgac ggtgtgagcg cccgacgcgg 60
ccgagggcgcc cggagtcccc agctagcccc ggcgcccgcc gccgcccaga cccgacgcga 120
ggccacctcg tcggcgctccg cccgagtccc cgctcgccg ccaacgccac aaccaccgag 180
cacggccccc tgactccgtc cagtgattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgcg accctccggg acggccgctct cagcgctcct ggcgctgctg gctgcgctct 300
gccccggcgag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgagtggtgg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggagggt gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttctcttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caacctgcc ctgtgcaacg 720
tggagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttccagaa ccacctgggc agtgcctaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccgagct 900
gctccggggc ctgcccgtggc aagtccccca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg cccccgggag agcgactgcc tggctgccc ccaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaacct caccacgtac cagatggatg 1080
tgaaccccgga ggcaaatatc agctttggtg ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgcccgaac gtgtgtaacg 1260
gaataggtat tgggtgaattt aaagactcac tctccataaa tgctacgaat attaaacct 1320
tcaaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc ccaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtgcgcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaagggcac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgc 1920
accgcagtg cctgcctcag gccatgaaca tccactgcac aggcgggga ccagacaact 1980
gtatccagtg tgccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaagg tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggtggtggg ggccctcctc ttgctgctgg 2220
tgggtggcct gggtcctcgg ctcttcctgc gaaggcgcca catcgttcgg aagcgcaacg 2280
tgccggaggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccgt agctatcaag gaattaaagg aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgtg atggccagcg tggacagcgt ggacaacccc cacgtgtgcc 2580
gcctgctggg catctgcctc acctccaccg tgcagctcat cagcgagctc atgcccttcg 2640
gctgcctcct ggactatgtc cggaacaca aagacaatat cggctcccag tacctgctca 2700
actggtgtgt gcagatcgca aagggcatga actacttggg ggaccgtcgc ttggtgcacc 2760
gcgacctggc agccaggaac gtactggtga aaacaccgca gcatgtcaag atcacagatt 2820
ttgggctggc caaactgctg ggtgcggaag agaaagaata ccatgcagaa ggaggcaaa 2880
tgcctatcaa gtggatggca ttggaatcaa ttttacacag aatctatacc caccagagt 2940
atgtctggag ctacgggggtg actgtttggg agttgatgac ctttggatcc aagccatatg 3000
acggaatccc tgccagcgag atctcctcca tcctggagaa aggagaacgc ctccctcagc 3060
caccatgatg taccatcgat gtctacatga tcatgggtcaa gtgctggatg atagacgcag 3120
atagtcgccc aaagtccgtg gaggttgatc tcgaattctc caaaatggcc cgagaccccc 3180
agcgctacct tgtcattcag ggggatgaaa gaatgcattt gccaaagtcct acagactcca 3240
acttctaccg tgccctgatg gatgaagaag acatggacga cgtggtggat gccgacgagt 3300
acctcatccc acagcagggc ttcttcagca gccctccac gtcacggact cccctcctga 3360
gctctctgag tgcaaccgac aacaattcca ccgtggcttg cattgataga aatgggctgc 3420
aaagtgtccc catcaaggaa gacagcttct tgcagcgata cagctcagac cccacaggcg 3480
ccttgactga ggacagcata gacgacacct tcctcccagt gcctgaatac ataaaccagt 3540
ccgttcccaa aaggcccgtg ggctctgtgc agaactcctg ctatcacaat cagcctctga 3600

accccgcgcc cagcagagac ccacactacc aggaccccca cagcactgca gtgggcaacc 3660
ccgagtatct caacactgtc cagcccacct gtgtcaacag cacattcgac agccctgccc 3720
actgggcccc gaaaggcagc caccaaatta gcctggacaa ccctgactac cagcaggact 3780
tctttcccaa ggaagccaag ccaaattggc tctttaaggg ctccacagct gaaaatgcag 3840
aatacctaag ggtcgcgcca caaagcagtg aatttattgg agcatga 3887

```

<210> 755

<211> 3884

5 <212> ADN

<213> Homo sapiens

<400> 755

```

ccccggcgag cgcggcgcca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggcggc cggagtcccc agctagcccc ggcgcccgcc gccgcccaga cccgacgaca 120
ggccacctcg tcggcgctccg cccgagtccc cgctcgcccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagttattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccggcgag 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
ttccttttga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagtcctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tgagagagcat ccagtggcgg gacatagtca gcagtgactt tctcagcaac atgtcgatgg 780
acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aactgccaga aactgaccaa aatcatctgt gccacgagc 900
gctccggggc ctgccgtggc aagtccccc gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcgactgcc tggctgccc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccg gggcaaata agcttttggg ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaa gtgtgtaacg 1260
gaataggatg ttgtgaattt 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 acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgttttg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac aggccaggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agccagggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgc 1920
accagagtg cctgcctcag gccatgaaca tcacctgcac aggcgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc cggcaggag 2040
tcatgggaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaagg tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggatgggtgg ggccctctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcctatg gaaggcgcca catcgctcgg aagcgcaacg 2280
tgcgaggcgt gctgcaggag agggagcttg tggagcctc tacacccagt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgagg aaactgaatt caaaaagatc aaagtgtgg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccg cgtatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgt atggccagcg tgacaacccc ccaccccac gtgtgcggcc 2580
tgctgggcat ctgcctcacc tccaccgtgc agctcatcac gcagctcatg cccttcggct 2640
gcctcctgga ctatgtccgg gaacacaaag acaatattgg ctccagtag cctgcaccgc 2700
ggtgtgtgca gatcgcaaag ggcataaact actggagga ccgtcgcttg gtgcaccgcg 2760
acctggcagc caggaacgta ctggtgaaaa caccgcagca tgtcaagatc acagattttg 2820
ggctggccaa actgctgggt gcggaagaga agaataacca tgcagaagga ggcaaagtgc 2880
ctatcaagtg gatggcattg gaatcaattt tacacagaat ctataccac cagagtgtatg 2940
tctggagcta cggggtgact gtttgggagt tgatgacctt tggatccaag ccatatgacg 3000
gaatccctgc cagcgagatc tcctccatcc tggagaaagg agaacgcctc cctcagccac 3060
ccatagtac catcgatgtc tacatgatca tggcaagtgc ctggatgata gacgcagata 3120

gtcgcacaaa gttccgtgag ttgatcatcg aattctccaa aatggcccg gacccccagc 3180
gctaccttgt cattcagggg gatgaaagaa tgcatttgcc aagtcctaca gactccaact 3240
tctaccgtgc cctgatggat gaagaagaca tggacgacgt ggtggatgcc gacgagtacc 3300
tcattcccac gcagggcttc ttcagagccc cctccacgtc acggactccc ctctgagct 3360
ctctgagtgc aaccagcaac aattccaccg tggcttgcat tgatagaaat gggctgcaaa 3420
gctgtcccat caaggaagac agcttcttgc agcgatacag ctacagcccc acaggcgct 3480
tgactgagga cagcatagac gacaccttcc tcccagtgcc tgaatacata aaccagtccg 3540
ttcccaaaag gcccgctggc tctgtgcaga atcctgtcta tcacaatcag cctctgaacc 3600
ccgcggccag cagagaccca cactaccag caccocacag cactgcagtg ggcaaccccg 3660
agtatctcaa cactgtccag cccacctgtg tcaacagcac attcgacagc cctgcccact 3720
gggcccagaa aggcagccac caaattagcc tggacaaccc tgactaccag caggacttct 3780
ttcccaagga agccaagcca aatggcatct ttaagggtc cagagctgaa aatgcagaat 3840
acctaagggt cgcgcacaaa agcagtgaat ttattggagc atga 3884

```

<211> 3878

<212> ADN

<213> Homo sapiens

<400> 756

```

ccccggcgag cgcgggcgca gcagcctccg ccccccgcac ggtgtgagcg cccgacgcgg 60
ccgagggcggc cggagtcccc agctagcccc ggcgggcgcc gccgcccaga cgggacgaca 120
ggccacctcg tcggcgctccg cccgagtcgc cgcctcgccg ccaacgccac aaccaccgcg 180
cacggccccc tgaactccgtc cagcagtcgc cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg agggccgggg cagcgtcctt ggcgctgctg gctgcgctct 300
gcccggcgag tcgggctctg gaggaaga aagtttgcca aggcacgagt aacaagctca 360
cgagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaagaccat ccaggagggt gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagtcctt atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caacctgcc ctgtgcaacg 720
tgagagcagc ccagtgccgg gacatagtcg gcagtgactt tctcagcaac atgtcgatgg 780
acttcagaaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctgggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccagcagc 900
gctccggggc ctgcccgtggc aagtcctcca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcgactgcc tggctgccc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaaccccca gggcaaatat agcttttggt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgcccgaag gtgtgtaacg 1260
gaataggtat tggtaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca ttttaggggtg 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 acaaaaaatt tgtctatgc aaatacaata aactgaaaaa 1680
aactgtttgg gacctccggt cagaaaaacca aaattataag caacagaggt gaaaaacagc 1740
gcaaggccac agggcaggtc tgccatgcct tgtgctcccc cgagggctgc tggggccccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggcgggga ccagacaact 1980
gtatccagtg tgccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctggctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgtcc atcgccactg ggtgggtgg ggccctctc ttgctgctgg 2220
tgggtggcctt ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgcgagggt gctgcaggag agggagcttg tggagcctct tacaccagc ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtgg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat ccagaagggt gagaaagtta 2460
aaattcccgt cgctatcaag gaattaaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgcctgctgg 2580
gcacatgcct cacctccacc gtgcagctca tcacgcagct catgccttc ggctgcctcc 2640

```

5

tggactatgt	ccgggaacac	aaagacaata	ttggctccca	gtacctgctc	aactggtgtg	2700
tgcagatcgc	aaagggcatg	aactacttgg	aggaccgtcg	cttggtgcac	cgcgacctgg	2760
cagccaggaa	cgtactggtg	aaaacactgc	agcatgtcaa	gatcacagat	tttgggctgg	2820
cctaaactgct	gggtgcgga	gagaaagaat	accatgcaga	aggaggcaaa	gtgcctatca	2880
agtggatggc	attggaatca	attttacaca	gaatctatac	ccaccagagt	gatgtctgga	2940
gctacggggt	gactgtttgg	gagttgatga	cctttggatc	caagccatat	gacggaatcc	3000
ctgccagcga	gatctcctcc	atcctggaga	aaggagaacg	cctccctcag	ccaccatat	3060
gtaccatcga	tgtctacatg	atcatggtca	agtgtgggat	gatagacgca	gatagtgcgc	3120
caaagtcccg	tgagttgatc	atcgaattct	cctaaatggc	ccgagacccc	cagcgctacc	3180
ttgtcattca	gggggatgaa	agaatgcatt	tgccaagtcc	tacagactcc	aacttctacc	3240
gtgcctgat	ggatgaagaa	gacatggacg	acgtggtgga	tgcgacgag	tacctcatcc	3300
cacagcaggg	cttcttcagc	agcccctcca	cgtcacggac	tcccctcctg	agctctctga	3360
gtgcaaccag	caacaattcc	accgtggtct	gcattgatag	aatgggctg	caaagctgtc	3420
ccatcaagga	agacagcttc	ttgcagcgat	acagctcaga	ccccacaggc	gccttgactg	3480
aggacagcat	agacgacacc	ttcctccag	tgctgaata	cataaaccag	tccgttccca	3540
aaaggcccg	tggtctgtg	cagaatcctg	tctatcacaa	tcagcctctg	aaccccgcc	3600
ccagcagaga	ccacactac	caggaccccc	acagcactgc	agtgggcaac	cccaggtatc	3660
tcaacactgt	ccagcccacc	tgtgtcaaca	gcacattcga	cagccctgcc	cactgggccc	3720
agaaaggcag	ccaccaaatt	agcctggaca	accctgaacta	ccagcaggac	ttctttccca	3780
aggaaagccaa	gccaaatggc	atctttaagg	gctccacagc	tgaaaatgca	gaatacctaa	3840
gggtcgcgcc	acaaagcagt	gaatttattg	gagcatga			3878

<210> 757

<211> 3878

<212> ADN

5 <213> Homo sapiens

<400> 757

cccggcgag	cgccggcgca	gcagcctccg	ccccccgcac	ggtgtgagcg	cccagcgagg	60
ccgagggggc	cggagtcccc	agctagcccc	ggcgcccgcc	gcccggcaga	ccggacgaca	120
ggccacctcg	tcggcgctccg	cccaggtccc	cgctcgcccg	ccaacgccac	aaccaccggg	180
cacggccccc	tgaactcgtc	cagtattgat	cgggagagcc	ggagcgagct	cttcggggag	240
cagcgatcgc	accctccggg	cagcccgggg	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
ttaaagaccat	ccaggagggtg	gctggttatg	tcctcattgc	cctcaacaca	gtggagcgaa	540
ttcctttgga	aaacctgcag	atcatcagag	gaaatatgta	ctacgaaaat	tcctatgcct	600
tagcagtcct	atctaactat	gatgcaaata	aaaccggact	gaaggagctg	cccatgagaa	660
atttacagga	aatcctgcat	ggcgccgtgc	ggttcagcaa	caaccctgcc	ctgtgcaacg	720
tggagagcat	ccagtgccgg	gacatagtca	gcagtgactt	tctcagcaac	atgtcgatgg	780
acttccagaa	ccacctgggc	agctgccaaa	agtgtgatcc	aagctgtccc	aatgggagct	840
gctgggggtgc	agcgaggag	aactgccaga	aactgaccaa	aatcatctgt	gcccagcagt	900
gctccggggc	ctgccgtggc	aagtccccca	gtgactgctg	ccacaaccag	tgtgtctcag	960
gctgcacagg	cccccgagg	agcgactgcc	tggtctgccc	caaattccga	gacgaagcca	1020
cgtgcaagga	cacctgcccc	ccactcatgc	tctacaaccc	caccacgtac	cagatggatg	1080
tgaaccccga	gggcaaatac	agctttgggtg	ccacctgcgt	gaagaagtgt	ccccgtaatt	1140
atgtggtgac	agatcacggc	tcgtgcgtcc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaagacgg	cgtccgcgaag	tgtagaaggt	gcgaagggcc	ttgccgcaaa	gtgtgtaacg	1260
gaataggtat	tggtgaattt	aaagactcac	tctccataaa	tgctacgaat	attaaacact	1320
tcaaaaactg	cacctccatc	agtggcgatc	tccacatcct	gccggtggca	tttaggggtg	1380
actccttcac	acataactcct	cctctggatc	cacaggaaact	ggatattctg	aaaaccgtaa	1440
aggaaatcac	agggtttttg	ctgattcagg	cttggcctga	aaacaggacg	gacctccatg	1500
cctttgagaa	cctagaaatc	atacgcgga	ggaccaagca	acatggtcag	ttttctcttg	1560
cagtcgtcag	cctgaacata	acatccttgg	gattacgctc	cctcaaggag	ataagtgatg	1620
gagatgtgat	aatttcaggag	aacaaaaatt	tgtgctatgc	aaatacaata	aactggaaaa	1680
aactgtttgg	gacctccggt	cagaaaaacca	aaattataag	caacagaggt	gaaaacagct	1740
gcaaggccac	aggccaggtc	tgccatgcct	tgtgctcccc	cgagggctgc	tggggcccg	1800
agcccaggga	ctgcgtctct	tgccggaatg	tcagccgagg	cagggaatgc	gtggacaagt	1860
gcaaccttct	ggagggtgag	ccaaggaggt	ttgtggagaa	ctctgagtgc	atacagtgcc	1920
accagagtg	cctgcctcag	gccatgaaca	tcacctgcac	aggacgggga	ccagacaact	1980
gtatccagt	tgccactac	attgacggcc	ccactgcgt	caagacctgc	ccggcaggag	2040
tcatgggaga	aaacaacacc	ctgggtctgga	agtacgcaga	cgccggccat	gtgtgccacc	2100
tgtgccatcc	aaactgcacc	tacggatgca	ctgggccagg	tcttgaaggc	tgtccaacga	2160

atggggcctaa	gatcccgctcc	atcgccactg	ggatgggtggg	ggccctcctc	ttgctgctgg	2220
tggtggccct	ggggatcggc	ctcttcacgc	gaaggcgcca	catcggtcgg	aagcgcaagc	2280
tgcggaggct	gctgcaggag	agggagcttg	tggagcctct	tacacccagt	ggagaagctc	2340
ccaaccaagc	tctcttgagg	atcttgaagg	aaactgaatt	caaaaagatc	aaagtgctgg	2400
gctccggtgc	gttcggcacg	gtgtataagg	gactctggat	cccagaaggt	gagaaagtta	2460
aaattcccgt	cgtatcaag	gaattaagag	aagcaacatc	tccgaaagcc	aacaaggaaa	2520
tcctcgatga	agcctacgtg	atggccagcg	tggacaaccc	ccacgtgtgc	cgcctgctgg	2580
gcatctgcct	cacctccacc	gtgcagctca	tcacgcagct	catgcccttc	ggctgcctcc	2640
tggactatgt	ccgggaacac	aaagacaata	ttggctccca	gtacctgtc	aactgggtgtg	2700
tgcagatcgc	aaagggcag	aactacttgg	aggaccgtcg	cttggtgcac	cgcgacctgg	2760
cagccaggaa	cgtactgggt	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	ccaccatat	3060
gtaccatcga	tgtctacatg	atcatggtca	agtgtctggat	gatagacgca	gatagtctgc	3120
caaagtcccg	tgaattgatc	atcgaattct	ccaaaatggc	ccgagacccc	cagcgtacc	3180
ttgtcattca	gggggatgaa	agaatgcatt	tgccaaagtcc	tacagactcc	aacttctacc	3240
gtgcctgat	ggatgaagaa	gacatggacg	acgtgggtgga	tgccgacgag	tacctcatcc	3300
cacagcaggg	cttcttcagc	agcccccca	cgtcacggac	tccccctctg	agctctctga	3360
gtgcaaccag	caacaattcc	accgtggctt	gcattgatag	aaatgggctg	caaagctgtc	3420
ccatcaaggga	agacagcttc	ttgcagcgat	acagctcaga	ccccacaggc	gccttgactg	3480
aggacagcat	agacgacacc	ttcctcccga	tgcctgaata	cataaaccag	tccgttccca	3540
aaaggcccg	tggtctgtgt	cagaatcctg	tctatcacaa	tcagcctctg	aaccccgccg	3600
ccagcagaga	ccacactac	caggaccccc	acagcactgc	agtgggcaac	cccaggtatc	3660
tcaacactgt	ccagcccacc	tgtgtcaaca	gcacattcga	cagccctgcc	cactgggccc	3720
agaaaggcag	ccaccaaatt	agcctggaca	acctgacta	ccagcaggac	ttctttccca	3780
aggaagccaa	gccaaatggc	atctttaagg	gtccacagc	tgaaaatgca	gaatacctaa	3840
gggtcgcgcc	acaaagcagt	gaatttatgt	gagcatga			3878

<210> 758

<211> 3878

<212> ADN

5 <213> Homo sapiens

<400> 758

ccccggcgag	cgcggccgca	gcagcctccg	ccccccgcac	ggtgtgagcg	cccgaagcgg	60
ccgagggcgg	cggagtcgcc	agctagcccc	ggcgcccgcc	gccgcccaga	ccggacgaca	120
ggccacctcg	tcggcgctcc	cccagatccc	cgctcgcccg	ccaacgccac	aaccaccgcg	180
cacggccccc	tgaactccgt	cagtatgat	cgggagagcc	ggagcgagct	cttcggggag	240
cagcgatgcg	accctccggg	acggccgggg	cagcgctcct	ggcgctgctg	gctgcgctct	300
gccccggcag	tcgggctctg	gaggaaaaa	aaagtttgcca	aggcacgagt	aacaagctca	360
cgcagttggg	cacttttgaa	gatcattttc	tcagcctcca	gaggatgttc	aataactgtg	420
aggtggctct	tgggaatttg	gaaattacct	atgtgcagag	gaattatgat	ctttccttct	480
taaagaccat	ccaggaggtg	gctggttatg	tcctcattgc	cctcaacaca	gtggagcgaa	540
ttccttttga	aaacctgcag	atcatcagag	gaaatatgta	ctacgaaaat	tcctatgcct	600
tagcagctct	atctaactat	gatgcaaata	aaaccggact	gaaggagctg	cccatgagaa	660
atttacagga	aatcctgcat	ggcgccgtgc	ggttcagcaa	caacctgcc	ctgtgcaacg	720
tggagagcat	ccagtggcgg	gacatagtca	gcagtgactt	tctcagcaac	atgtcgatgg	780
acttccagaa	ccacctgggc	agctgccaaa	agtgtgatcc	aaagctgtccc	aatgggagct	840
gctgggggtg	aggagaggag	aactgccaga	aactgaccaa	aatcatctgt	gccagcagct	900
gctccggggc	ctgcccgtgg	aagtccccc	gtgactgctg	ccacaaccag	tgtgctgcag	960
gctgcacagg	ccccggggag	agcgactgcc	tggctctgcc	caaattccga	gacgaagcca	1020
cgtgcaaggga	cacctgcccc	ccactcatgc	tctacaaccc	caccacgtac	cagatggatg	1080
tgaaccccga	gggcaaatat	agctttgtgt	ccacctgcgt	gaagaagtgt	ccccgtaat	1140
atgtggtgac	agatcacggc	tcgtgcgtcc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaagcagg	cgtccgcaag	tgaagaagt	gcgaagggcc	ttgccgcaaa	gtgtgtaacg	1260
gaataggtat	tgggtgaattt	aaagactcac	tctccataaa	tgtacgaat	attaaacact	1320
tcaaaaactg	cacctccatc	agtggcgatc	tccacatcct	gccgggtggc	tttaggggtg	1380
actccttcac	acatactcct	cctctggatc	cacaggaact	ggatattctg	aaaaccgtaa	1440
aggaaatcac	agggtttttg	ctgattcac	cttggcctga	aaacaggacg	gacctccatg	1500
cctttgagaa	cctagaaatc	atacgcggca	gcaccaagca	acatggtcag	ttttctcttg	1560
cagtcgtcag	cctgaacata	acatccttgg	gattacgctc	cctcaaggag	ataagtgatg	1620
gagatgtgat	aatttcagga	aacaaaaatt	tgtgctatgc	aaatacaata	aactggaaaa	1680

```

aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaagggccac aggccagggtc tgccatgcct tgtgctcccc cgaggggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggaggggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcc cccactgcgt caagacctgc ccggcaggag 2040
tcatgggaga aaacaacacc ctgggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggatgggtgg ggccctcctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcctgc gaaggcgcca catcgctcgg aagcgcacgc 2280
tgccggaggct gctgcaggag agggagcttg tggagcctct tacaccagt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat ccagaagggt gagaaagtta 2460
aaattcccggt cgctatcaag gaattaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgtg atggccagcg tggacaaccc ccacgtgtgc cgctgctgg 2580
gcatctgcct cacctccacc gtgcagctca tcacgcagct catgcccttc ggctgcctcc 2640
tggtactatgt ccgggaacac aaagacaata ttggctccca gtacctgtc aactggtgtg 2700
tgcatatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760
cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggctgg 2820
ccaaactgct ggggtcgga gagaaagaat accatgcaga agggaggcaa gtgcctatca 2880
agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940
gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000
ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccaccatat 3060
gtaccatcga tgtctacatg atcatggtca agtctggat gatagacga gatagtctgc 3120
caaagtctcg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgtacc 3180
ttgtcattca ggggatgaa agaatgcatt tgccaagtcc tacagactcc aacttctacc 3240
gtgccctgat ggtgaagaa gacatggacg acgtgggtgga tgccgacgag tacctcatcc 3300
cacagcaggg cttcttcagc agccctcca cgtcacggac tcccctcctg agctctctga 3360
gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420
ccatcaagga agacagcttc ttgcagcgat acagctcaga cccacaggc gccttgactg 3480
aggacagcat agacgacacc ttctccag tgctgaata cataaaccag tccgttccca 3540
aaagcccgcc tggtctctgt cagaatcctg tctatcaca tcagcctctg aaccccgcg 3600
ccagcagaga cccacactgc caggaccccc acagcactgc agtgggcaac cccgagtatc 3660
tcaacactgt ccagccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720
agaaaggcag ccaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780
aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840
gggtcgcgcc acaagcaggt gaatttatgg gagcatga 3878

```

<210> 759

<211> 3878

<212> ADN

5 <213> Homo sapiens

<400> 759

```

ccccggcgag cgcgggcgca gcagcctccg cccccgcac ggtgtgagcg cccgacgagg 60
ccgagggcggc cggagtcccc agctagcccc ggcgggcgcc gccgcccaga ccgagcagca 120
ggccacctcg tcggcgctcc cccagatccc cgctcgcccg ccaacgccac aaccaccgag 180
cacggccccc tgactccgct cagtattgat cgggagagcc ggagcgagct cttcggggag 240
cagcgatgag accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccggcag tcggggtctg gaggaaga aaagtgtgcca aggcacgagt aacaagctca 360
cgagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggctct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttcttct 480
taaagaccat ccaggaggtg gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgct 600
tagcagtctt atctaactat gatgcaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgc atgacatgca ggttcagcaa caaccctgcc ctgtgcaacg 720
tgagagcat ccagtgccgg gacatagtca cagtgactt tctcagcaac atgtcgatgg 780
acttcagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gccagcag 900
gctccggggc ctgccgtggc aagtccccc gtgactgctg ccacaaccag tgtgtgcag 960
gctgcacagg ccccggggag agcgactgcc tggctctgcc caaattccga gacgaagcca 1020
cgtgcaagga ccactgcccc ccactcatgc tctacaacc caccagctac cagatggatg 1080
tgaaccccg gggcaaatc agctttgggt ccacctgcgt gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tcgtgcgtcc gagcctgtgg ggccgacagc tatgagatgg 1200

```

```

aggaagacgg cgtccgcaag tgtaagaagt gcgaagggcc ttgccgcaaa gtgtgtaacg 1260
gaataggtat tgggtgaattt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cactctccatc agtggcgatc tccacatcct gccgggtggca tttaggggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaaatcac aggggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaaatc atacgcggca ggaccaagca acatggtcag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactggttgg gacctccggg cagaaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccac agggcagggtc tgccatgcct tgtgctcccc cgagggctgc tggggcccg 1800
agcccaggga ctgcgtctct tgccggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaagggagt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgcccactac attgacggcg cccactgcgt caagacctgc ccggcaggag 2040
tcatggtgaga aaacaacacc ctggtctgga agtacgcaga cgccggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggcccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgtcc atcgccactg ggatgggtgg ggcctctctc ttgctgctgg 2220
tgggtggcctt ggggatcggc ctcttcatgc gaaggcgcca catcgttcgg aagcgcacgc 2280
tgcgagggct gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtgg 2400
gctccggtgc gttcggcacg gtgtataagg gactctggat cccagaaggt gagaaagtta 2460
aaattcccgt cgttatcaag gaattaaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgtg atggccagcg tggacaaccc caactgtgct cgctgtctgg 2580
gcactctgct cactccacc gtgcagctca tcacgcagct catgcccttc ggctgctcc 2640
tggaactatg ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700
tgcaagatgc aaagggcag aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760
cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcccgg 2820
ccaaactgct ggggtgcggaa gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880
agtggatggc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940
gctacggggt gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000
ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccaccatat 3060
gtaccatcga tgtctacatg atcatggtca agtgcctggat gatagacgca gatagtgcgc 3120
caaagtccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180
ttgtcattca gggggtgaa agaattgcat tgccaagtcc tacagactcc aacttctacc 3240
gtgccctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300
cacagcaggg cttcttcacg agccccctcca cgtcacggac tccccctctg agctctctga 3360
gtgcaaccag caacaatttc accgtggtct gcattgatag aaatgggctg caaagctgtc 3420
ccatcaagga agacagcttc ttgcagcag acagctcaga ccccacaggc gccttgactg 3480
aggacagcat agacgacacc ttctctccag tgctgaata cataaaccag tccgttccca 3540
aaaggccgc tggtctctgt cagaatcctg tctatcacia tcagcctctg aaccccgcg 3600
ccagcagaga ccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660
tcaacactgt ccagcccacc tgtgtcaaca gcacattcga cagccctgcc cactgggccc 3720
agaaaaggcag caaccaaatt agcctggaca accctgacta ccagcaggac ttctttccca 3780
aggaagccaa gccaaatggc atctttaagg gctccacagc tgaaaatgca gaatacctaa 3840
gggtcgcgcc acaaagcagt gaatttattg gagcatga 3878

```

<210> 760

<211> 3878

<212> ADN

5 <213> Homo sapiens

<400> 760

```

ccccggcgag cgcggccgca gcagcctccg ccccccgcac ggtgtgagcg cccgacggcg 60
ccgaggcggc cggagtcccc agctagcccc ggcggccgcc gccgcccaga ccggacgaca 120
ggccacctcg tcggcgctccg cccgagtccc cgctcgccg ccaacgccac aaccaccgcg 180
cacggccccc tgactccgtc cagtattgat cggagagacc ggagcgagct cttcggggag 240
cagcgatcgc accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gccccggcag tcgggctctg gaggaaaaga aagtttgcca aggcacgagt aacaagctca 360
cgcagtgggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggtcct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttctctct 480
taaagaccat ccaggagggt gctggttatg tcctcattgc cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagctct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgcgctgc ggttcagcaa caaccctgcc ctgtgcaacg 720

```

tggagagcat	ccagtggcgg	gacatagtc	gcagtgactt	tctcagcaac	atgtcgatgg	780
acttccagaa	ccacctgggc	agctgccaaa	agtgatgatcc	aagctgtccc	aatgggagct	840
gctgggggtgc	aggagaggag	aactgccaga	aactgaccaa	aatcatctgt	gcccagcagt	900
gctccggggcg	ctgcggtggc	aagtccccc	gtgactgctg	ccacaaccag	tgtgctgcag	960
gctgcacagg	cccccgggag	agcgactgcc	tggctgtgccg	caaattccga	gacgaagcca	1020
cgtgcaaggga	cacctgcccc	ccactcatgc	tctacaaccc	caccacgtac	cagatggatg	1080
tgaaccccga	gggcaaatac	agctttgggtg	ccacctgctg	gaagaagtgt	ccccgtaatt	1140
atgtggtgac	agatcacggc	tcgtgctgctc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaagacgg	cgtccgcgaag	tgtagaaggt	gcgaagggcc	ttgccgcaaa	gtgtgtaacg	1260
gaatagggtat	tgggtgaattt	aaagactcac	tctccataaa	tgctacgaat	attaaacact	1320
tcaaaaactg	cacctccatc	agtggcgatc	tccacatcct	gccgggtggca	tttaggggtg	1380
actccttcac	acatactcct	cctctggatc	cacaggaact	ggatattctg	aaaaccgtaa	1440
aggaaatcac	agggtttttg	ctgattcagg	cttggcctga	aaacaggacg	gacctccatg	1500
cctttgagaa	cctagaaatc	atcacgcggca	ggaccaagca	acatggtcag	ttttctcttg	1560
cagtgcgtcag	cctgaacata	acatccttgg	gattacgctc	cctcaaggag	ataagtgatg	1620
gagatgtgat	aatttcagga	aacaaaaatt	tgtgctatgc	aaatacaata	aactggaaaa	1680
aactgttttg	gacctccggt	cagaaaacca	aaattataag	caacagaggt	gaaaaacagt	1740
gcaaggccac	aggccaggtc	tgccatgcct	tgtgctcccc	cgagggtctg	tggggcccg	1800
agcccaggga	ctgcgtctct	tgccggaatg	tcagccgagg	cagggaatgc	gtggacaagt	1860
gcaaccttct	ggagggtgag	ccaaggaggt	ttgtggagaa	ctctgagtgc	atacagtgc	1920
accagagtg	cctgcctcag	gccatgaaca	tcacctgcac	aggacgggga	ccagacaact	1980
gtatccagt	tgcccactac	attgacggcc	cccactgctg	caagacctgc	ccggcaggag	2040
tcatgggaga	aaacaacac	ctggtctgga	agtagcgaga	cgccggccat	gtgtgccacc	2100
tgtgccatcc	aaactgcacc	tacggatgca	ctgggccagg	tcttgaaggc	tgtccaacga	2160
atgggacctaa	gatcccgctc	atcgccactg	ggatgggtggg	ggccctcctc	ttgctgctgg	2220
tgggtggccct	ggggatcggc	ctcttcctgc	gaaggcgcca	catcggtcgg	aagcgacgc	2280
tgccgagggt	gctgcaggag	agggagcttg	tggagcctct	tacacccagt	ggagaagctc	2340
ccaaccaagc	tctcttgagg	atcttgaagg	aaactgaatt	caaaaagatc	aaagtgtctg	2400
gctccgggtgc	gttcggcaag	gtgtataagg	gactctggat	cccagaaggt	gagaaagtta	2460
aaattcccg	cgctatcaag	gaattaaag	aagcaacatc	tccgaaagcc	aacaaggaaa	2520
tcctcgatga	agcctacgtg	atggccagcg	tggacaaccc	ccacgtgtgc	cgctgctgg	2580
gcactgcct	cacctccacc	gtgcagctca	tcacgcagct	catgcccttc	ggctgcctcc	2640
tggactatgt	ccgggaacac	aaagacaata	ttggctccca	gtacctgtc	aactgggtgtg	2700
tgcatgcgc	aaagggcatg	aactacttgg	aggaccgtcg	cttgggtgac	cgcgacctgg	2760
cagccaggaa	cgtactgggtg	aaaacaccgc	agcatgtcaa	gatcacagat	tttgggcggg	2820
ccaaactgct	gggtgcggaa	gagaaagaat	accatgcaga	aggaggcaaa	gtgcctatca	2880
agtggatggc	attggaatca	attttacaca	gaatctatac	ccaccagagt	gatgtctgga	2940
gctacgggg	gactgttttg	gagttgatga	cctttggatc	caagccatat	gacggaatcc	3000
ctgccagcga	gatctcctcc	atcctggaga	aaggagaacg	cctccctcag	ccaccatat	3060
gtaccatcga	tgtctacatg	atcatggtca	agtgtggat	gatagacgca	gatagtgcgc	3120
caaagtctcg	tgaattgatc	atcgaattct	ccaaaatggc	ccgagacccc	cagcgctacc	3180
ttgtcattca	gggggatgaa	agaatgcatt	tgccaagtcc	tacagactcc	aacttctacc	3240
gtgccctgat	ggatgaagaa	gacatggacg	acgtgggtgga	tgccgacgag	tacctcatcc	3300
cacagcagg	cttcttcagc	agccctccca	cgtcacggac	tcccctcctg	agctctctga	3360
gtgcaaccag	caacaattcc	accgtggctt	gcattgatag	aatgggctg	caaagctgtc	3420
ccatcaaggga	agacagcttc	ttgcagcgat	acagctcaga	ccccacaggc	gccttgactg	3480
aggacagcat	agacgacacc	ttcctcccag	tgctgaata	cataaaccag	tccgttccca	3540
aaaggcccg	tggtctctgtg	cagaatcctg	tctatcacia	tcagcctctg	aaccccgccg	3600
ccagcagaga	cccacactac	caggaccccc	acagcactgc	agtgggcaac	cccaggtatc	3660
tcaacactgt	ccagcccacc	tgtgtcaaca	gcacattcga	cagccctgcc	cactgggccc	3720
agaaaggcag	ccaccaaatt	agcctggaca	accctgacta	ccagcaggac	ttctttccca	3780
aggaagccaa	gccaaaatgg	gtccacagc	gtccacagc	tgaatatgca	gaatacctaa	3840
gggtcgccgc	acaaagcagt	gaatttatgg	gagcatga			3878

<210> 761

<211> 3878

<212> ADN

5 <213> Homo sapiens

<400> 761

cccggcgag	cgccggcgca	gcagcctccg	ccccccgcac	ggtgtgagcg	cccgcgcggg	60
ccgaggcggc	cggagtcggc	agctagcccc	ggcgcccgcc	gccgcccaga	ccggacgaca	120
ggccacctcg	tcggcgctccg	cccagtcgcc	cgctcgccg	ccaacgccac	aaccaccggg	180
cacggccccc	tgactccgtc	cagtattgat	cgggagagcc	ggagcgagct	cttcggggag	240

cagcgatgcg accctccggg acggccgggg cagcgctcct ggcgctgctg gctgcgctct 300
gcccgccgag tccggctctg gaggaagaaga aagtttgcca aggcacgagt aacaagctca 360
cgcagttggg cacttttgaa gatcattttc tcagcctcca gaggatgttc aataactgtg 420
aggtggctct tgggaatttg gaaattacct atgtgcagag gaattatgat ctttccttct 480
taaaagccat ccaggaggtg gctggttatg tctcattg cctcaacaca gtggagcgaa 540
ttcctttgga aaacctgcag atcatcagag gaaatatgta ctacgaaaat tcctatgcct 600
tagcagtcct atctaactat gatgcaaata aaaccggact gaaggagctg cccatgagaa 660
atttacagga aatcctgcat ggcgccgtgc ggttcagcaa caaccctgcc ctgtgcaacg 720
tggagagcat ccagtggcgg gacatagtc gcagtgactt tctcagcaac atgtcgatgg 780
acttccagaa ccacctgggc agctgccaaa agtgtgatcc aagctgtccc aatgggagct 840
gctggggtgc aggagaggag aactgccaga aactgaccaa aatcatctgt gcccgagct 900
gctccggggc ctgccgtggc aagtcaccca gtgactgctg ccacaaccag tgtgctgcag 960
gctgcacagg ccccggggag agcagctgcc tggctctgccc caaattccga gacgaagcca 1020
cgtgcaagga cacctgcccc ccactcatgc tctacaaccc caccacgtac cagatggatg 1080
tgaacccoga gggcaaatat agctttggtg ccacctgctg gaagaagtgt ccccgtaatt 1140
atgtggtgac agatcacggc tctgctgctc gagcctgtgg ggcgacagc tatgagatgg 1200
aggaagacgg cgtccgcaag tgtagaagt gccaaggccc ttgccgcaaa gtgtgtaacg 1260
gaatgagtat tggtagaatt aaagactcac tctccataaa tgctacgaat attaaacact 1320
tcaaaaactg cacctccatc agtggcgatc tccacatcct gccggtggca tttagggtg 1380
actccttcac acatactcct cctctggatc cacaggaact ggatattctg aaaaccgtaa 1440
aggaatcac agggtttttg ctgattcagg cttggcctga aaacaggacg gacctccatg 1500
cctttgagaa cctagaatat atacgaggca ggaccaagca acatggctag ttttctcttg 1560
cagtcgtcag cctgaacata acatccttgg gattacgctc cctcaaggag ataagtgatg 1620
gagatgtgat aatttcagga acaaaaaatt tgtgctatgc aaatacaata aactggaaaa 1680
aactgtttgg gacctccggt cagaaaacca aaattataag caacagaggt gaaaacagct 1740
gcaaggccca aggcaggctc tgccatgcct tgtgctcccc cgagggtgc tggggcccgg 1800
agccaggga ctgcgtctct tgcgggaatg tcagccgagg cagggaatgc gtggacaagt 1860
gcaaccttct ggagggtgag ccaaggaggt ttgtggagaa ctctgagtgc atacagtgcc 1920
acccagagtg cctgcctcag gccatgaaca tcacctgcac aggacgggga ccagacaact 1980
gtatccagtg tgccactac attgacggcc cccactgctg caagacctgc ccggcaggag 2040
tcattgggga aaacaacacc ctggtctgga agtacgcaga cgcgggccat gtgtgccacc 2100
tgtgccatcc aaactgcacc tacggatgca ctgggccagg tcttgaaggc tgtccaacga 2160
atgggcctaa gatcccgctc atcgccactg ggatggtggg ggcctcctc ttgctgctgg 2220
tggtggccct ggggatcggc ctcttcattg gaaggcgcca catcgctcgg aagcgcacgc 2280
tgccgagggt gctgcaggag agggagcttg tggagcctct tacaccaggt ggagaagctc 2340
ccaaccaagc tctcttgagg atcttgaagg aaactgaatt caaaaagatc aaagtgtctg 2400
gctccggtgc gttcggaacg gtgtataagg gactctggat ccagaaaggt gagaaagtta 2460
aaattcccg tgcctatcaag gaattaaagag aagcaacatc tccgaaagcc aacaaggaaa 2520
tcctcgatga agcctacgtg atggccagcg tggacaaccc ccactgtgct cgcctgctgg 2580
gcactctgct caactccacc gtgcagctca tcacgcagct catgcccctc ggctgcctcc 2640
tggaactatg ccgggaacac aaagacaata ttggctccca gtacctgctc aactggtgtg 2700
tgcagatcgc aaagggcatg aactacttgg aggaccgtcg cttggtgcac cgcgacctgg 2760
cagccaggaa cgtactggtg aaaacaccgc agcatgtcaa gatcacagat tttgggcccg 2820
ccaactcgtg ggggtgcgga gagaaagaat accatgcaga aggaggcaaa gtgcctatca 2880
agtggtatgc attggaatca attttacaca gaatctatac ccaccagagt gatgtctgga 2940
gctacggggg gactgtttgg gagttgatga cctttggatc caagccatat gacggaatcc 3000
ctgccagcga gatctcctcc atcctggaga aaggagaacg cctccctcag ccacctatat 3060
gtaccatcga tgtctacatg atcatggtca agtgcctggat gatagacgca gatagtcgct 3120
caaaagtccg tgagttgatc atcgaattct ccaaaatggc ccgagacccc cagcgctacc 3180
ttgtcattca gggggtatga agaattgcatt tgccaagtcc tacagactcc aacttctacc 3240
gtgcctgat ggatgaagaa gacatggacg acgtggtgga tgccgacgag tacctcatcc 3300
cacagcaggg cttcttcagc agccctcca cgtcacggac tcccctcctg agctctctga 3360
gtgcaaccag caacaattcc accgtggctt gcattgatag aaatgggctg caaagctgtc 3420
ccatcaagga agacagcttc ttgcagcgat acagctcaga cccacagggc gccttgactg 3480
aggacagcat agacgacacc ttcctcccag tgcctgaata cataaaccag tccgttccca 3540
aaaggcccg tggctctgtg cagaatcctg tctatcacia tcagcctctg aaccccgctc 3600
ccagcagaga cccacactac caggaccccc acagcactgc agtgggcaac cccgagtatc 3660
tcaaacactg 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> ADN

5 <213> Homo sapiens

<400> 762

cccggcgcag	cgcgcccgca	gcagcctccg	ccccccgcac	gggtgtgagcg	cccgacgcgg	60
ccgagggcggc	cggagtcgccg	agctagcccc	ggcgggccgcc	gccgcccaga	ccggacgaca	120
ggccacctcg	tcggcgctccg	cccaggtccc	cgccctcgccg	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	cactttttgaa	gatcattttc	tcagcctcca	gaggatgttc	aataactgtg	420
aggtggctct	tgggaatttg	gaaattacct	atgtgcagag	gaattatgat	ctttccttct	480
taaagaccat	ccaggaggtg	gctggttatg	tcctcattgc	cctcaacaca	gtggagcgaa	540
ttcctttgga	aaacctgcag	atcatcagag	gaaatatgta	ctacgaaaat	tcctatgcct	600
tagcagtcct	atctaactat	gatgcaaata	aaaccggact	gaaggagctg	cccatgagaa	660
atttacagga	aatcctgcat	ggcgccgtgc	ggttcagcaa	caaccctgcc	ctgtgcaacg	720
tggagagcat	ccagtggcgg	gacatagtca	gcagtgactt	tctcagcaac	atgtcgatgg	780
acttcagaa	ccacctggcg	agctgccaaa	agtgtgatcc	aagctgtccc	aatgggagct	840
gctgggggtgc	aggagaggag	aactgccaga	aactgaccaa	aatcatctgt	gcccgagcgt	900
gctccggggcg	ctgccgtggc	aagtccccca	gtgactgtctg	ccacaaccag	tgtgctgcag	960
gctgcacagg	ccccggggag	agcgactgcc	tggctctgccg	caaattccga	gacgaagcca	1020
cgtgcaaggga	cacctgcccc	ccactcatgc	tctacaaccc	caccacgtac	cagatggatg	1080
tgaaccccca	gggcaaatat	agctttgggtg	ccacctgcgt	gaagaagtgt	ccccgtaatt	1140
atgtggtgac	agatcacggc	tcgtgcgtcc	gagcctgtgg	ggccgacagc	tatgagatgg	1200
aggaagacgg	cgtccgcgaag	tgtgaagaagt	gcgaagggcc	ttgccgcaaa	gtgtgtaacg	1260
gaaataggtat	tgggtgaattt	aaagactcac	tctccataaa	tgctacgaat	attaacact	1320
tcaaaaactg	cacctccatc	agtggcgatc	tccacatcct	gccggtggca	tttaggggtg	1380
actccttcac	acatactcct	cctctggatc	cacaggaact	ggatattctg	aaaaccgtaa	1440
aggaaatcac	agggtttttg	ctgattcagg	cttgccctga	aaacaggacg	gacctccatg	1500
cctttgagaa	cctagaaatc	atacggcgca	ggaccaagca	acatggtcag	ttttctcttg	1560
cagtcgctcag	cctgaacata	acatccttgg	gattacgctc	cctcaaggag	ataagtgtatg	1620
gagatgtgat	aatttcagga	aacaaaaatt	tgtgctatgc	aaatacaata	aactggaaaa	1680
aactgttttg	gacctccggt	cagaaaacca	aaattataag	caacagaggt	gaaaacagct	1740
gcaaggccac	aggccaggtc	tgccatgcct	tgtgctcccc	cgagggtctg	tggggcccg	1800
agcccaggga	ctgcgtctct	tgccggaatg	tcagccgagg	cagggaatgc	gtggacaagt	1860
gcaaccttct	ggagggtgag	ccaagggagt	ttgtggagaa	ctctgagtgc	atacagtgcc	1920
accagagtg	cctgcctcag	gccatgaaca	tcacctgcac	aggacgggga	ccagacaact	1980
gtatccagtg	tgcccactac	attgacggcc	cccactgcgt	caagacctgc	ccggcaggag	2040
tcatggggaga	aaacaacacc	ctggtctgga	agtacgcaga	cgccggccat	gtgtgccacc	2100
tgtgccatcc	aaactgcacc	tacggatgca	ctgggcccagg	tcttgaaagg	tgcccaacga	2160
atgggacctaa	gaccccgctc	atcgccactg	ggatgggtggg	ggccctcctc	ttgctgctgg	2220
tgggtggccct	ggggatcggc	ctcttcacgc	gaaggcgcca	catcgctcgg	aagcgcacgc	2280
tgccggaggct	gctgcaggag	agggagcttg	tggagcctct	tacacccagt	ggagaagctc	2340
ccaaccaagc	tctcttgagg	atcttgaagg	aaactgaatt	caaaaagatc	aaagtgtctgg	2400
gctccggtgc	gttcggcacg	gtgtataagg	gactctggat	cccagaaggt	gagaaagtta	2460
aaattcccgt	cgctatcaag	gaattaaagag	aagcaacatc	tccgaaagcc	aacaaggaaa	2520
tcctcgatga	agcctacgtg	atggccagcg	tggacaaccc	ccactgtgtc	cgctgtctgg	2580
gcatctgcct	cacctccacc	gtgcagctca	tcacgcagct	catgcccttc	ggctgcctcc	2640
tggaactatgt	ccgggaacac	aaagacaata	ttggctccca	gtacctgtct	aactggtgtg	2700
tgcagatcgc	aaagggcatg	aactacttgg	aggaccgtcg	cttgggtgcac	cgcgacctgg	2760
cagccaggaa	cgtactggtg	aaaacaccgc	agcatgtcaa	gatcacagat	tttgggctgg	2820
ccaaacagct	gggtgcggaa	gagaaagaat	accatgcaga	aggaggcaaa	gtgcctatca	2880
agtggatggc	attggaatca	attttacaca	gaatctatac	ccaccagagt	gatgtctgga	2940
gctacgggggt	gactgttttg	gagttgatga	cctttggatc	caagccatat	gacggaatcc	3000
ctgccagcga	gatctcctcc	atcctggaga	aaggagaacg	cctccctcag	ccacccatat	3060
gtaccatcga	tgtctacatg	atcatgggtca	agtgtgtgat	gatagacgca	gatagtgcgc	3120
caaagtcccg	tgagttgatc	atcgaattct	ccaaaatggc	ccgagacccc	cagcgctacc	3180
ttgtcattca	gggggatgaa	agaatgcatt	tgccaagtcc	tacagactcc	aacttctacc	3240
gtgccctgat	ggatgaagaa	gacatggacg	acgtggtgga	tgccgacgag	tacctcatcc	3300
cacagcaggg	cttcttcacg	agcccccca	cgtcacggac	tccccctcctg	agctctctga	3360
gtgcaaccag	caacaatttc	accgtggctt	gcattgatag	aaatgggctg	caaagctgtc	3420
ccatcaagga	agacagcttc	ttgcagcgat	acagctcaga	ccccacaggc	gccttgactg	3480
aggacagcat	agacgacacc	ttcctcccag	tgccctgaata	cataaaccag	tccgttccca	3540
aaaggcccg	tggctctgtg	cagaatcctg	tctatcacaa	tcagcctctg	aaccccgcg	3600
ccagcagaga	cccacactac	caggaccccc	acagcactgc	agtgggcaac	cccgagtatc	3660

ES 2 741 574 T3

```
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
```

REIVINDICACIONES

1. Una proteína del EGFR que comprende una o más variaciones o mutaciones que confieren capacidad de respuesta a la terapia con inhibidores de la tirosina quinasa cuando está presente en un individuo afectado de cáncer, o que confiere una mayor actividad de la quinasa en el EGFR, seleccionada de:
 - 5 a. una sustitución de cisteína por glicina en la posición 719 de la SEQ ID NO: 763 (G719C) o en una sustitución de serina por glicina en el codón 719 de la SEQ ID NO: 511 (G719S);
 - b. una supresión de al menos los aminoácidos leucina, arginina, ácido glutámico y alanina en los codones 747, 748, 749 y 750 de la SEQ ID NO: 511; o
 - 10 c. una sustitución de aminoácido de arginina por leucina en el codón 858 de la SEQ ID NO: 511 (L858R) o de glutamina por leucina en el codón 861 de la SEQ ID NO: 511 (L861Q).
2. La proteína del EGFR de acuerdo con la reivindicación 1, en donde la variación o mutación es una supresión de los aminoácidos 746 a 750 de la SEQ ID NO: 511 que son la secuencia del ácido glutámico, leucina, arginina, ácido glutámico y alanina.
3. Una proteína del EGFR de acuerdo con la reivindicación 1 o la reivindicación 2, que comprende una o más variaciones o mutaciones en el dominio de la quinasa de la proteína del EGFR, que confiere capacidad de respuesta a la terapia con inhibidores de la tirosina quinasa cuando está presente en un individuo afectado de cáncer, o que confiere un aumento de la actividad de la quinasa en el EGFR, en donde la variación o mutación en el EGFR, en cada caso usando la numeración mostrada en la SEQ ID NO: 511, se selecciona de:
 - 15 (i) una supresión en el marco en el exón 19 que comprende una supresión del aminoácido 746 a 750 de la SEQ ID NO: 511;
 - (ii) una supresión en el marco que comprende los nucleótidos 2235 a 2249 de la SEQ ID NO: 511 de acuerdo con la numeración utilizada en la Figura 5 que elimina los aminoácidos 746 a 750, que son la secuencia del ácido glutámico, leucina, arginina, ácido glutámico y alanina de la SEQ ID NO: 511;
 - 20 (iii) una supresión en el marco que comprende los nucleótidos 2236 a 2250 de la SEQ ID NO: 511 de acuerdo con la numeración utilizada en la Figura 5 que elimina los aminoácidos 746 a 750 de la SEQ ID NO: 511;
 - (iv) una supresión en el marco que comprende los nucleótidos 2240 a 2251 de la SEQ ID NO: 511 de acuerdo con la numeración utilizada en la Figura 5;
 - (v) una supresión en el marco que comprende los nucleótidos 2240 a 2257 de la SEQ ID NO: 511 de acuerdo con la numeración utilizada en la Figura 5;
 - 30 (vi) una supresión en el marco que comprende los nucleótidos 2239 a 2247 junto con una sustitución de citosina por guanina en el nucleótido 2248 de la SEQ ID NO: 511 de acuerdo con la numeración utilizada en la Figura 5;
 - (vii) una supresión en el marco que elimina los nucleótidos 2238 a 2255 junto con una sustitución de timina por adenina en el nucleótido 2237 de la SEQ ID NO: 511 de acuerdo con la numeración utilizada en la Figura 5;
 - 35 (viii) una supresión en el marco que elimina los nucleótidos 2239 a 2250 junto con una sustitución de citosina por adenina en el nucleótido 2251 de la SEQ ID NO: 511 de acuerdo con la numeración utilizada en la Figura 5;
 - (ix) una supresión en el marco que comprende los nucleótidos 2239 a 2250 del TTAAGAGAAGCA de la SEQ ID NO: 511 de acuerdo con la numeración utilizada en la Figura 5;
 - (x) una supresión en el marco que comprende los nucleótidos 2240 a 2254 del TAAGAGAAGCA de la SEQ ID NO: 511 de acuerdo con la numeración utilizada en la Figura 5;
 - 40 (xi) una sustitución en el exón 21 del EGFR que comprende una sustitución de guanina por una timina en el nucleótido 2573 de la SEQ ID NO: 511 de acuerdo con la numeración utilizada en la Figura 5 que da como resultado una sustitución de aminoácidos donde la leucina de tipo silvestre se reemplaza con una arginina en el aminoácido 858 de la SEQ ID NO: 511;
 - 45 (xii) una sustitución en el exón 21 del EGFR que comprende una sustitución de una adenina por una timina en el nucleótido 2582 de la SEQ ID NO: 511 de acuerdo con la numeración utilizada en la Figura 5 que da como resultado una sustitución de aminoácidos donde la leucina de tipo silvestre se reemplaza con una glutamina en el aminoácido 861 de la SEQ ID NO: 511;
 - (xiii) una sustitución en el exón 18 del EGFR que comprende una sustitución de una timina por una guanina en el nucleótido 2155 de la SEQ ID NO: 511 de acuerdo con la numeración utilizada en la Figura 5 que da como resultado

una sustitución de aminoácidos donde la glicina de tipo silvestre está sustituida por una cisteína en el codón 719 de la SEQ ID NO: 511; y

(xiv) una sustitución en el exón 18 del EGFR que comprende una sustitución de una adenina por una guanina en el nucleótido 2155 de la SEQ ID NO: 511 de acuerdo con la numeración utilizada en la Figura 5 que da como resultado una sustitución de aminoácidos donde la glicina de tipo silvestre está sustituida por una serina en el codón 719 de la SEQ ID NO: 511.

4. Una proteína de acuerdo con la reivindicación 1 seleccionada de:

A. una proteína que comprende la secuencia de una cualquiera de las SEQ ID NOS: 477, 480 y 483 a 487 y la secuencia de aminoácidos mostrada en la SEQ ID NO: 512 siempre que esté presente una mutación como se define en la reivindicación 1;

B. una proteína aislada que tiene la secuencia de aminoácidos de la SEQ ID NO: 512, en donde se suprimen los aminoácidos 746 a 750;

C. una proteína aislada que tiene la secuencia de aminoácidos de la SEQ ID NO: 512, en donde se suprimen los aminoácidos 746 a 751;

D. una proteína aislada que tiene la secuencia de aminoácidos de la SEQ ID NO: 512, en donde se suprimen los aminoácidos 747 a 753;

E. una proteína aislada que tiene la secuencia de aminoácidos de la SEQ ID NO: 512, en donde la Leucina en el aminoácido 858 está sustituida con una Arginina;

F. una proteína aislada que tiene la secuencia de aminoácidos de la SEQ ID NO: 512, en donde la Leucina en el aminoácido 861 está sustituida con una Glutamina; y

G. una proteína aislada que tiene la secuencia de aminoácidos de la SEQ ID NO: 512, en donde la glicina en el aminoácido 719 está sustituida con una Cisteína.

5. Un anticuerpo que es una sonda capaz de unirse a una proteína del EGFR mutante de acuerdo con una cualquiera de las reivindicaciones 1 a 4.

6. Uso de una proteína del EGFR de acuerdo con una cualquiera de las reivindicaciones 1 a 4, en un ensayo de selección de fármacos candidatos para identificar agentes candidatos bioactivos que inhiben la actividad de formas variantes del EGFR.

7. Un ácido nucleico del EGFR que comprende una o más variaciones o mutaciones que confieren capacidad de respuesta a la terapia con inhibidores de la tirosina quinasa cuando está presente en un individuo afectado de cáncer, o que confiere una mayor actividad de la quinasa en el EGFR, seleccionado de:

a. una mutación en el exón 18 que da como resultado una sustitución de la cisteína por glicina en el codón 719 de la SEQ ID NO: 511 (G719C) o en una sustitución de serina por glicina en el codón 719 de la SEQ ID NO: 511 (G719S);

b. una supresión en el marco en el exón 19 que da como resultado una supresión de al menos los aminoácidos leucina, arginina, ácido glutámico y alanina en los codones 747, 748, 749 y 750 de la SEQ ID NO: 511; o

c. una mutación en el exón 21 que da como resultado una sustitución de aminoácido de arginina por leucina en el codón 858 de la SEQ ID NO: 511 (L858R) o de glutamina por leucina en el codón 861 de la SEQ ID NO: 511 (L861Q).

8. El ácido nucleico del EGFR de acuerdo con la reivindicación 7, en donde la variación es como se define en una cualquiera de las reivindicaciones 2 a 4.

9. Un ácido nucleico aislado de acuerdo con la reivindicación 7, seleccionado de:

A. un ácido nucleico aislado que comprende la SEQ ID NO: 495;

B. un ácido nucleico aislado que comprende la secuencia de la SEQ ID NO: 511, en donde se suprimen los nucleótidos 2235 a 2249 de acuerdo con la numeración utilizada en la Figura 5;

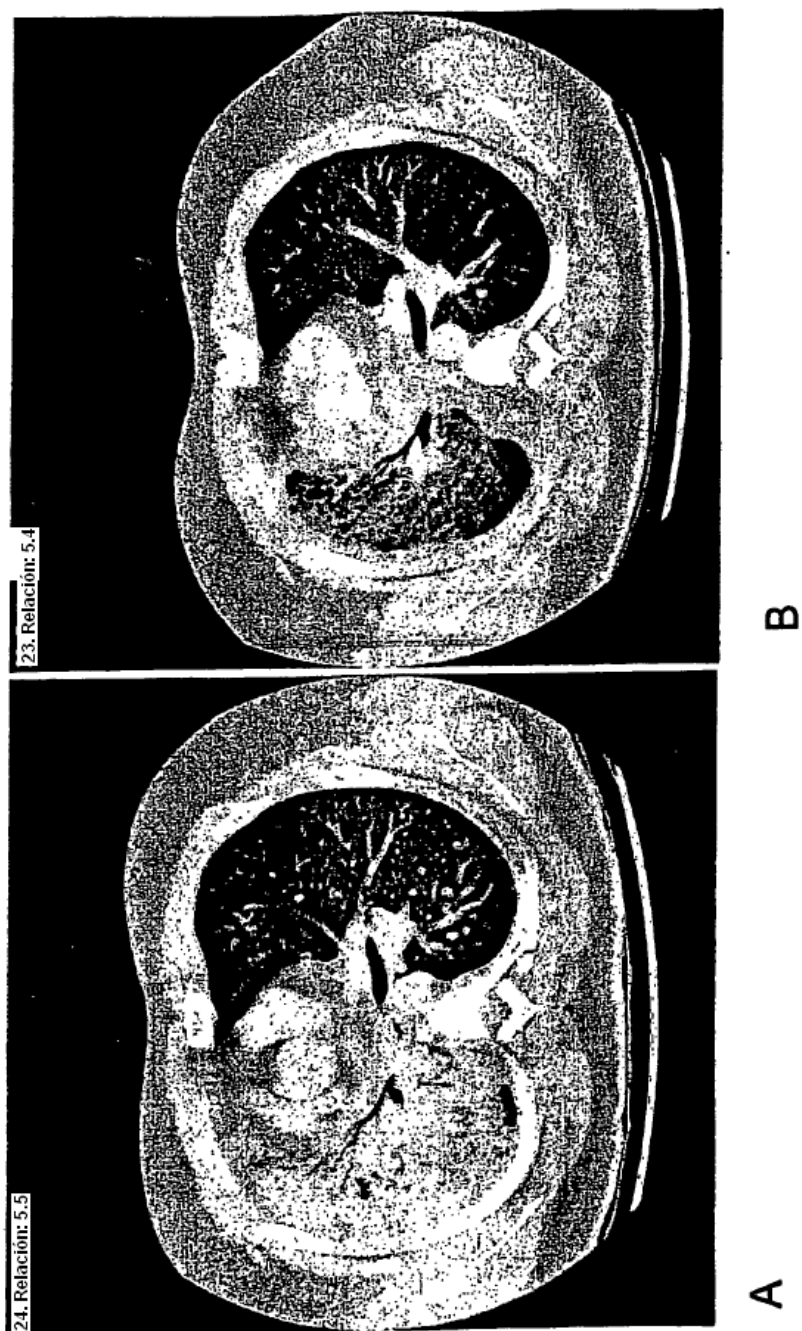
C. un ácido nucleico aislado que comprende la SEQ ID NO: 497;

D. un ácido nucleico aislado que comprende la secuencia de la SEQ ID NO: 511, en donde se suprimen los nucleótidos 2240 a 2251 de acuerdo con la numeración utilizada en la Figura 5;

E. un ácido nucleico aislado que comprende la SEQ ID NO: 499;

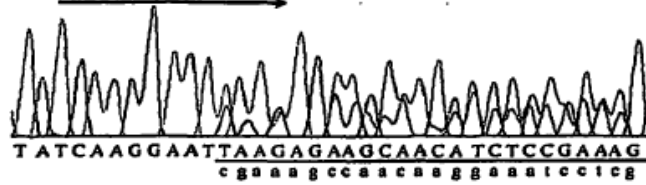
- F. un ácido nucleico aislado que comprende la secuencia de la SEQ ID NO: 511, en donde se suprimen los nucleótidos 2240 a 2257 de acuerdo con la numeración utilizada en la Figura 5;
- G. un ácido nucleico aislado que comprende la SEQ ID NO: 502;
- 5 H. un ácido nucleico aislado que comprende la secuencia de la SEQ ID NO: 511, en donde una guanina se sustituye por la timina en el nucleótido 2573 de acuerdo con la numeración utilizada en la Figura 5;
- I. un ácido nucleico aislado que comprende la SEQ ID NO: 504;
- J. un ácido nucleico aislado que comprende la secuencia de la SEQ ID NO: 511, en donde una adenina se sustituye por la timina en el nucleótido 2582 de acuerdo con la numeración utilizada en la Figura 5;
- 10 K. un ácido nucleico aislado que comprende la secuencia de la SEQ ID NO: 511, en donde una timina se sustituye por la guanina en el nucleótido 2155 de acuerdo con la numeración utilizada en la Figura 5;
- L. un ácido nucleico que codifica una proteína del EGFR de acuerdo con una cualquiera de las reivindicaciones 1 a 4;
- M. un ácido nucleico que codifica un anticuerpo de acuerdo con la reivindicación 5;
- 15 N. ARNpi capaz de mediar específicamente la desactivación de la expresión del EGFR mutante, donde el EGFR mutante comprende una o más variaciones como se define en la reivindicación 9;
- O. un ácido nucleico que comprende la secuencia de cualquiera de las SEQ ID NO: 425, 427, 434, 438, 441, 443, 444, 445, 495, 497, 499, 502, 504, 554, 720, 727, 730 a 739, 741 a 751, 757 y 762;
- 20 P. un ácido nucleico complementario a un ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 7, 8 o 9A a 9O, en donde dicho ácido nucleico complementario se hibrida con el ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 7, 8 o 9A a 9O. en comparación con la hibridación de dicho ácido nucleico complementario con la secuencia de ácido nucleico correspondiente que tiene una base en el sitio de variación que no es complementaria con dicho ácido nucleico complementario bajo condiciones de hibridación selectiva; y
- Q. Un ácido nucleico antisentido a un ácido nucleico de acuerdo con la reivindicación 9P.
- 25 10. Células que se han transfectado con un ADNc de tal manera que comprenden una proteína de acuerdo con una cualquiera de las reivindicaciones 1 a 4 o un ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 7 a 9.
11. Un animal transgénico no humano que contiene células que comprenden una proteína de acuerdo con una cualquiera de las reivindicaciones 1 a 4 o un ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 7 a 9.
- 30 12. Vectores que comprenden un ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 7 a 9.
13. Un kit que comprende:
- 35 (a) al menos una sonda capaz de distinguir la presencia de una variación o variaciones genéticas (i) como se define en una cualquiera de las reivindicaciones 7 a 9 o (ii) que puede causar una variación o variaciones proteicas como se define en una cualquiera de las reivindicaciones 1 a 4, en donde la sonda está diseñada para hibridarse con regiones de ácido nucleico dentro de los exones 18, 19 o 21 del dominio de quinasa del EGFR e hibridar preferentemente con un ácido nucleico que tiene una variación o variaciones genéticas (i) como se define en una cualquiera de las reivindicaciones 7 a 9 o (ii) que puede causar una variación o variaciones de la proteína como se define en cualquiera de las reivindicaciones 1 a 4, en comparación con la hibridación de la sonda con la correspondiente secuencia de ácidos nucleicos que tiene una base en el sitio de la variación que no es
- 40 complementario con la sonda bajo condiciones de hibridación selectiva;
- (b) productos y reactivos requeridos para llevar a cabo la reacción de hibridación; y
- (c) instrucciones.

Figura 1A y B



Figuras 2A - 2E

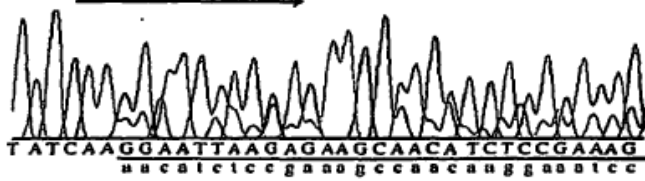
A. del L747-P753insS (sentido)



del L747-P753insS (antisentido)



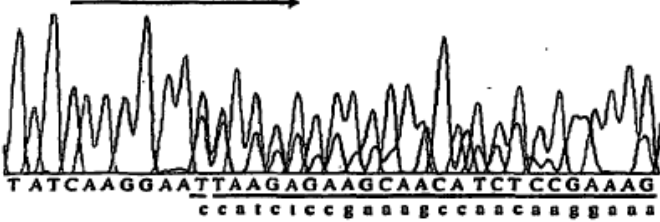
B. del E746-A750 (sentido)



del E746-A750 (antisentido)



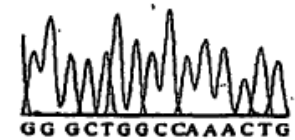
C. del L747-T751insS (sentido)



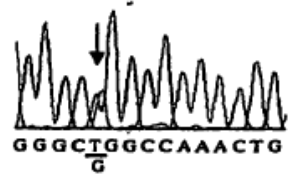
del L747-T751insS (antisentido)



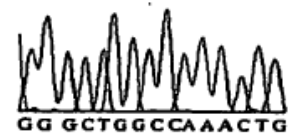
D. Tipo silvestre



Mutación L858R



E. Tipo silvestre



Mutación L861Q

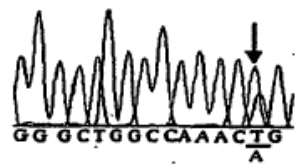
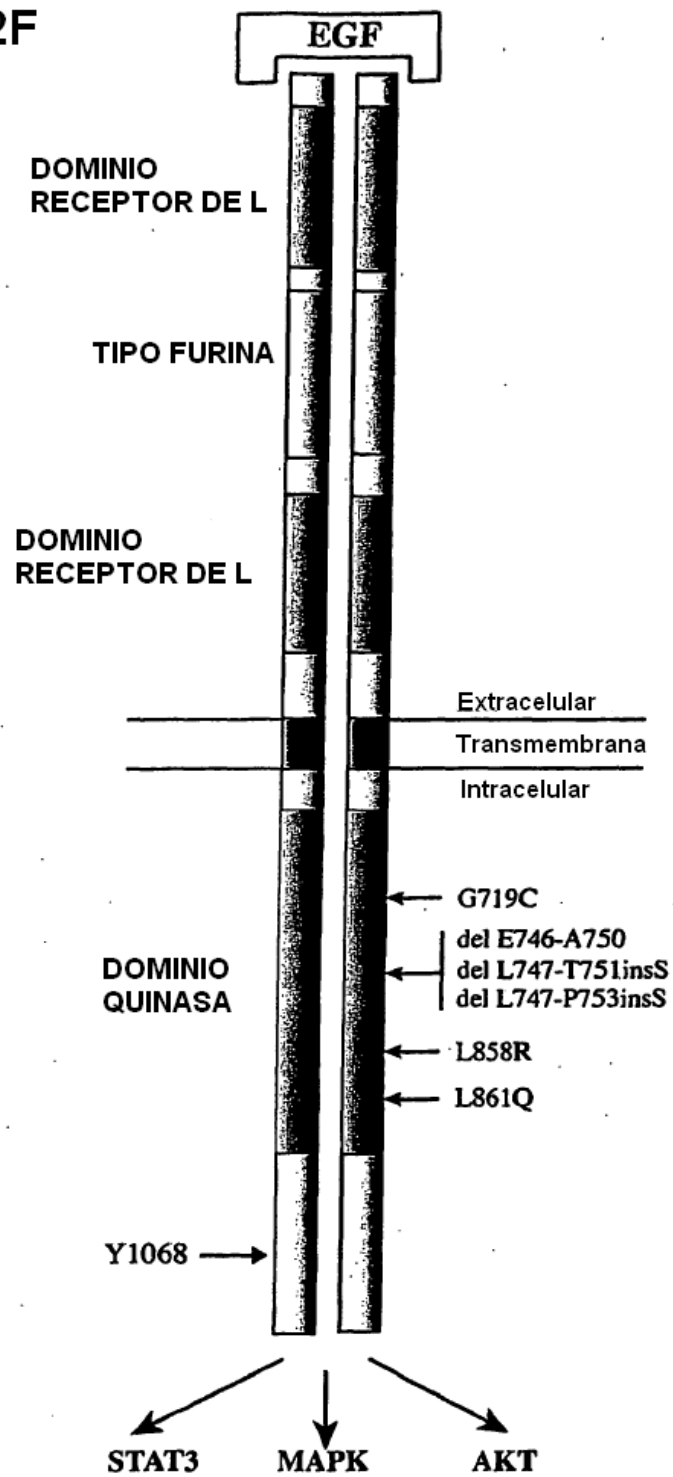
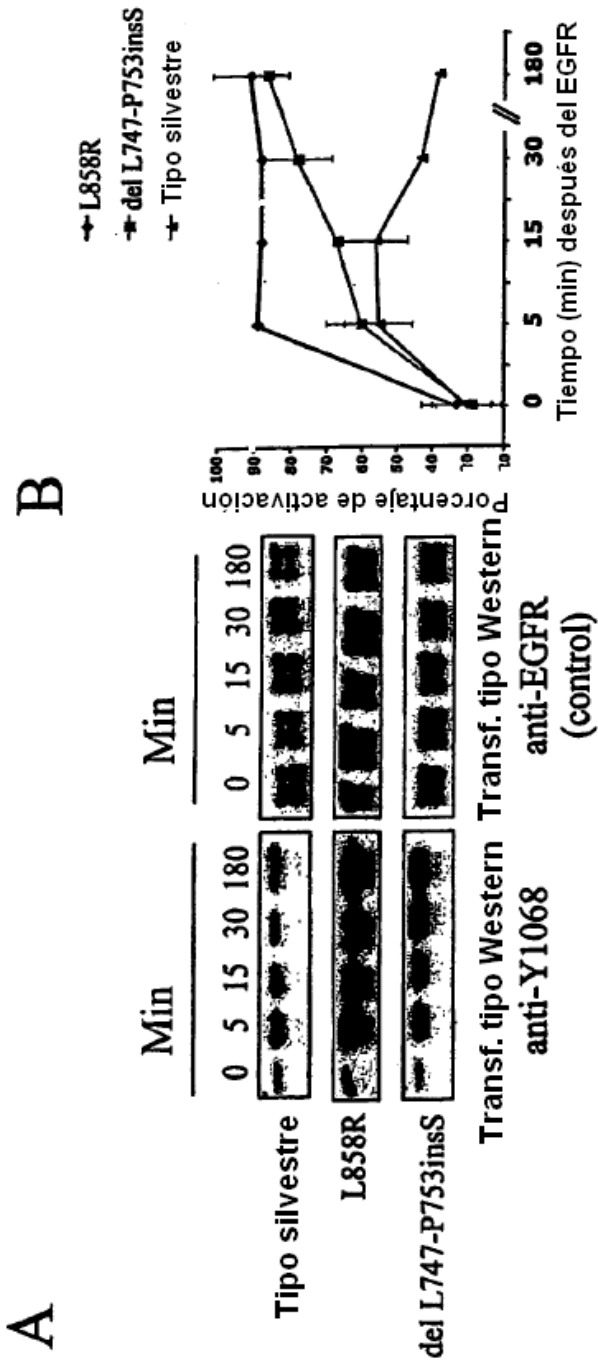


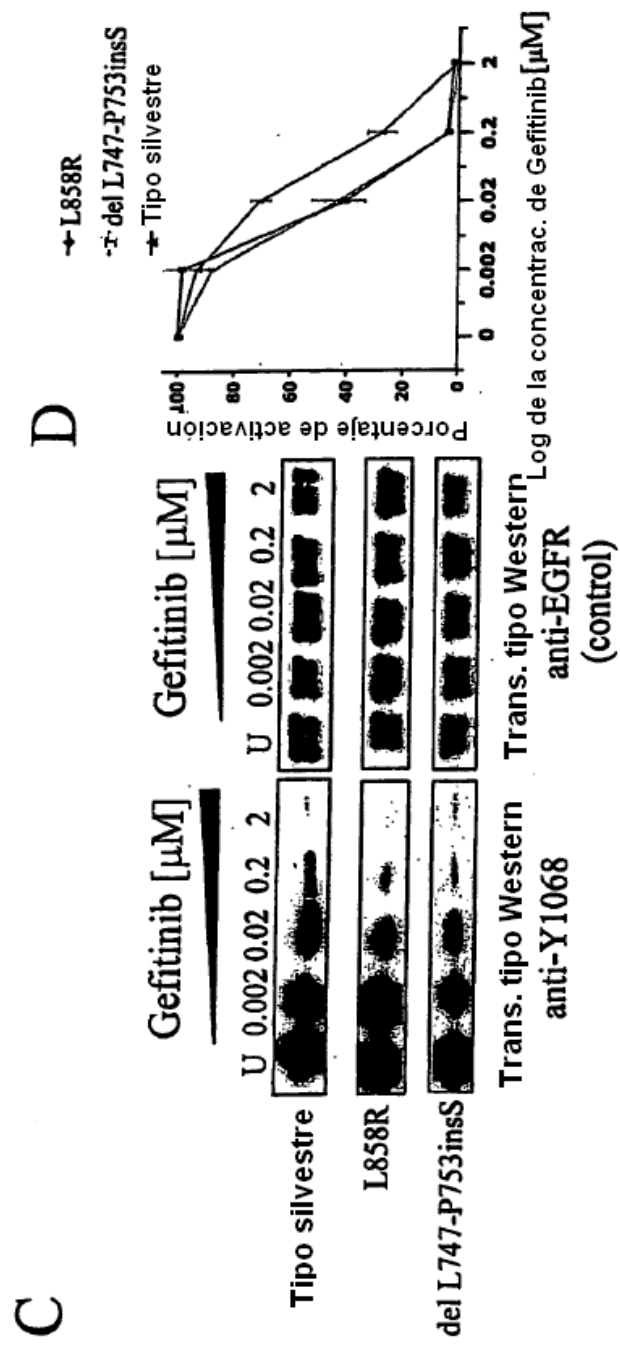
Figura 2F



Figuras 3A y 3B



Figuras 3C y 3D



	(SEQ. ID. NO. 509)	K I P V A I K E L R E A T S P K A N	756	856	(SEQ. ID. NO. 510)	F G L A K L L O	863
Gen EGFR	739	AAATTCGGCTGGCTATCAAGGAATTAGAGAGCAACATCTCCGAAGCCAC	2268	2566	TTTGGCTGGCCAAACTGCTGGT		2589
Proteína EGFR	2215	(SEQ. ID. NO. 483)			(SEQ. ID. NO. 494)		
Paciente							
Caso 1		AAATTCGGCTGGCTATCAAGGAATTAGAGAGCAACATCTCCGAAGCCAC	(SEQ. ID. NO. 495)		TTTGGCTGGCCAAACTGCTGGT	(SEQ. ID. NO. 496)	
Caso 2		AAATTCGGCTGGCTATCAAGGAATTAGAGAGCAACATCTCCGAAGCCAC	(SEQ. ID. NO. 497)		TTTGGCTGGCCAAACTGCTGGT	(SEQ. ID. NO. 498)	
Casos 3 & 4		AAATTCGGCTGGCTATCAAGGAATTAGAGAGCAACATCTCCGAAGCCAC	(SEQ. ID. NO. 499)		TTTGGCTGGCCAAACTGCTGGT	(SEQ. ID. NO. 500)	
Casos 5 & 6		AAATTCGGCTGGCTATCAAGGAATTAGAGAGCAACATCTCCGAAGCCAC	(SEQ. ID. NO. 501)		TTTGGCTGGCCAAACTGCTGGT	(SEQ. ID. NO. 502)	
Caso 7		AAATTCGGCTGGCTATCAAGGAATTAGAGAGCAACATCTCCGAAGCCAC	(SEQ. ID. NO. 503)		TTTGGCTGGCCAAACTGCTGGT	(SEQ. ID. NO. 504)	
Exón 19							
Exón 21							

Figuras 4B - C

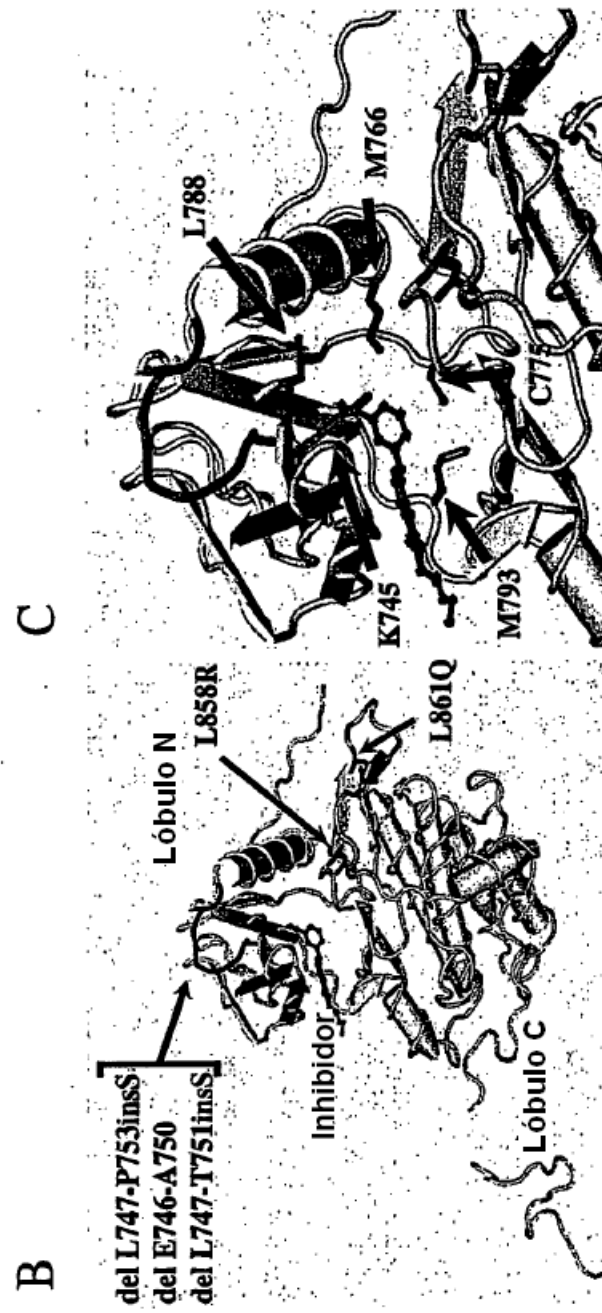


Figura 5

CCCGGCGCAGCGCGGCCGCGAGCAGCCTCCGCCCCCGCACGGTGTGAGCGCCCGACGCGG -185

 CCGAGGCGGCCGAGTCCCGAGCTAGCCCCGGCGGCCGCGCCCGCCAGACCGGACGACA -125

 GGCCACCTCGTCGGCGTCCGCCCCGAGTCCCCGCCTCGCCGCCAACGCCACAACCACCGCG -65

 CACGGCCCCCTGACTCCGTCCAGTATTGATCGGGAGAGCCGGAGCGAGCTCTTCGGGGAG -5

 CAGCGATGCGACCCTCCGGGACGGCCGGGGCAGCGCTCCTGGCGCTGCTGGCTGCGCTCT 55
-M--R--P--S--G--T--A--G--A--A--L--L--A--L--L--A--A--L-- 18
 GCCCCGCGAGTCGGGCTCTGGAGGAAAAGAAAGTTTGCCAAGGCACGAGTAACAAGCTCA 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
 TAAAGACCATCCAGGAGGTGGCTGGTTATGTCTCATTTGCCCTCAACACAGTGGAGCGAA 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
 ATTTACAGGAAATCCTGCATGGCGCCGTGCGGTTTCAGCAACAACCTGCCCTGTGCAACG 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
 ACTTCCAGAACCACCTGGGCGAGCTGCCAAAAGTGTGATCCAAGCTGTCCCAATGGGAGCT 595
 D--F--Q--N--H--L--G--S--C--Q--K--C--D--P--S--C--P--N--G--S-- 198
 GCTGGGGTGCAGGAGAGGAGAACTGCCAGAACTGACCAAAATCATCTGTGCCCGAGCAGT 655
 C--W--G--A--G--E--E--N--C--Q--K--L--T--K--I--I--C--A--Q--Q-- 218
 GCTCCGGGCGCTGCCGTGGCAAGTCCCCAGTGACTGCTGCCACAACCAAGTGTGCTGCAG 715
 C--S--G--R--C--R--G--K--S--P--S--D--C--C--H--N--Q--C--A--A-- 238
 GCTGCACAGGCCCCCGGGAGAGCGACTGCCTGGTCTGCCGCAAATTCGAGACGAAGCCA 775
 G--C--T--G--P--R--E--S--D--C--L--V--C--R--K--F--R--D--E--A-- 258
 CGTGCAAGGACACCTGCCCCCACTCATGCTCTACAACCCACACGTACCAGATGGATG 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

Figura 5 (continuación)

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
 TCAAAAACCTGCACCTCCATCAGTGGCGATCTCCACATCCTGCCGGTGGCATTATAGGGGTG 1135
 F--K--N--C--T--S--I--S--G--D--L--H--I--L--P--V--A--F--R--G-- 378
 ACTCCTTCACACATACTCCTCCTCTGGATCCACAGGAACCTGGATATTCTGAAAACCGTAA 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
 GAGATGTGATAATTTTCAGGAAACAAAAATTGTGCTATGCAATACAATAAACTGGAAAA 1435
 G--D--V--I--I--S--G--N--K--N--L--C--Y--A--N--T--I--N--W--K-- 478
 AACTGTTTGGGACCTCCGGTCAGAAAACCAAATTATAAGCAACAGAGGTGAAAACAGCT 1495
 K--L--F--G--T--S--G--Q--K--T--K--I--I--S--N--R--G--E--N--S-- 498
 GCAAGGCCACAGGCCAGGTCTGCCATGCCTTGTGCTCCCCGAGGGCTGTGGGGCCCGG 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
 ACCCAGAGTGCTGCCTCAGGCCATGAACATCACCTGCACAGGACGGGGACCAGACAAC 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
 TCATGGGAGAAAAACAACACCCTGGTCTGGAAGTACGCAGACGCCGGCCATGTGTGCCACC 1855
 V--M--G--E--N--N--T--L--V--W--K--Y--A--D--A--G--H--V--C--H-- 618
 TGTGCCATCCAACTGCACCTACGGATGCACTGGGCCAGGTCTTGAAGGCTGTCCAACGA 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

Figura 5 (continuación)

TGCGGAGGCTGCTGCAGGAGAGGGAGCTTGTGGAGCCTCTTACACCCAGTGGAGAAGCTC 2095
 L--R--R--L--L--Q--E--R--E--L--V--E--P--L--T--P--S--G--E--A-- 698
 CCAACCAAGCTCTCTTGAGGATCTTGAAGGAACTGAATTCAAAAAGATCAAAGTGCTGG 2155
 P--N--Q--A--L--L--R--I--L--K--E--T--E--F--K--K--I--K--V--L-- 718
 GCTCCGGTGCCTTCGGCACGGTGTATAAGGGACTCTGGATCCCAGAAGGTGAGAAAGTTA 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
 CCAAAGTGTGGGTGCGGAAGAGAAAGAATACCATGCAGAAGGAGGCAAAGTGCCTATCA 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
 CAAAGTTCCGTGAGTTGATCATCGAATTCTCCAAAATGGCCCCGAGACCCCCAGCGCTACC 2935
 P--K--F--R--E--L--I--I--E--F--S--K--M--A--R--D--P--Q--R--Y-- 978
 TTGTCATTTCAGGGGGATGAAAGAATGCATTTGCCAAGTCCTACAGACTCCAACCTTCTACC 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
 CACAGCAGGGCTTCTTCAGCAGCCCCCTCCACGTACGGACTCCCCTCCTGAGCTCTCTGA 3115
 P--Q--Q--G--F--F--S--S--P--S--T--S--R--T--P--L--L--S--S--L-- 1038
 GTGCAACCAGCAACAATTCACCGTGGCTTGCAATTGATAGAAATGGGCTGCAAAGCTGTC 3175
 S--A--T--S--N--N--S--T--V--A--C--I--D--R--N--G--L--Q--S--C-- 1058

Figura 5 (continuación)

CCATCAAGGAAGACAGCTTCTTGCAGCGATACAGCTCAGACCCACAGGCGCCTTGACTG 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 (SEQ ID NO 511)
 R--V--A--P--Q--S--S--E--F--I--G--A--*- 1210 (SEQ ID NO 512)

FIG. 6

A		SEQ ID NO	
	bucle de activación		
L858R	KTPQHVKITDFG E AKLLGAEEKEYH	870	477
EGFR	KTPQHVKITDFGLAKLLGAEEKEYH	870	478
BRAF	H E DLTVKIGD E GLAT V KSRWSGSHQ	608	479
	*** *****		
B			
	bucle P		
G719S	ETEFKKIKVL S SGAFGTVYKGLWIP	733	480
EGFR	ETEFKKIKVLGSGAFGTVYKGLWIP	733	481
BRAF	DGQITVGQRI S SG S FGTVYKGKWHG	477	482
	*** ***** *		
C			
	742 750 752		
Del-1	VAIK --- T-SPKANKEILDEAYV	765	483
Del-2	VAIKELREAT- --- LDEAYV	765	484
Del-3	VAIKE --- PT-SPKANKEILDEAYV	765	485
Del-4	VAIKE --- - --- SKANKEILDEAYV	765	486
Del-5	VAIKV --- - --- PKANKEILDEAYV	765	487
EGFR	VAIKELREAT-SPKANKEILDEAYV	765	488
BRAF	VAVKMLNVTAPT P QQLQAFKNEVG V	503	489
	** * * * *		

FIG. 7

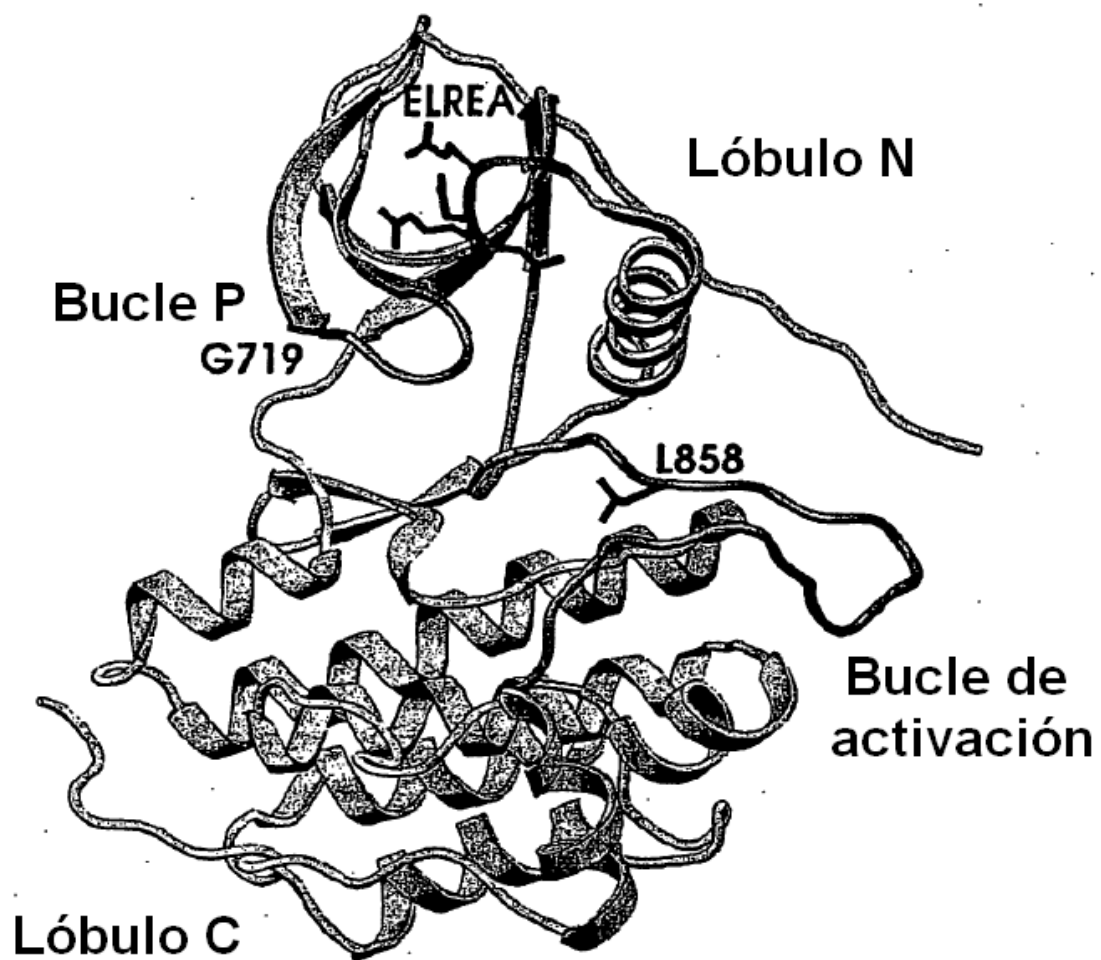


FIG. 8

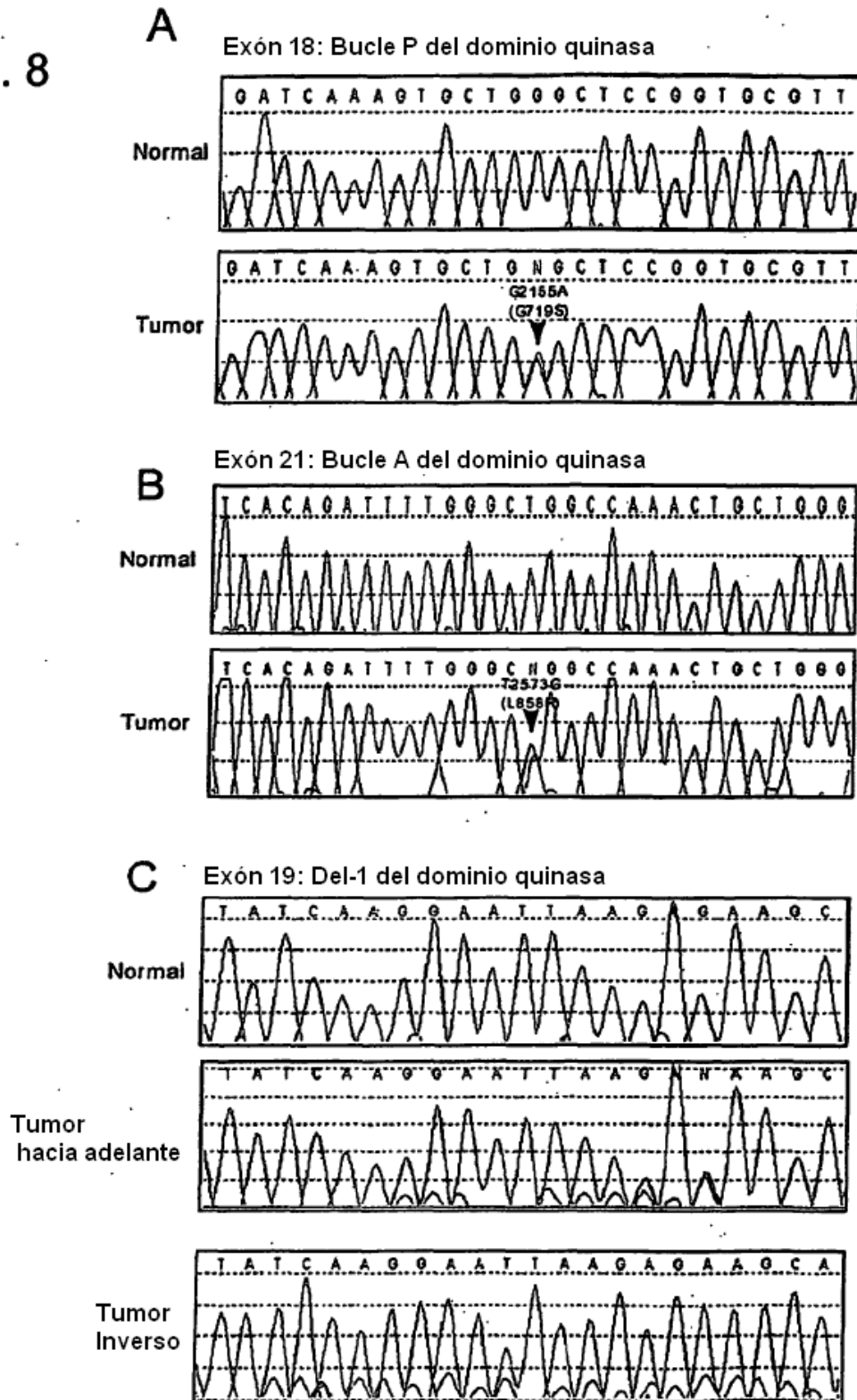


FIG. 8 (cont.)

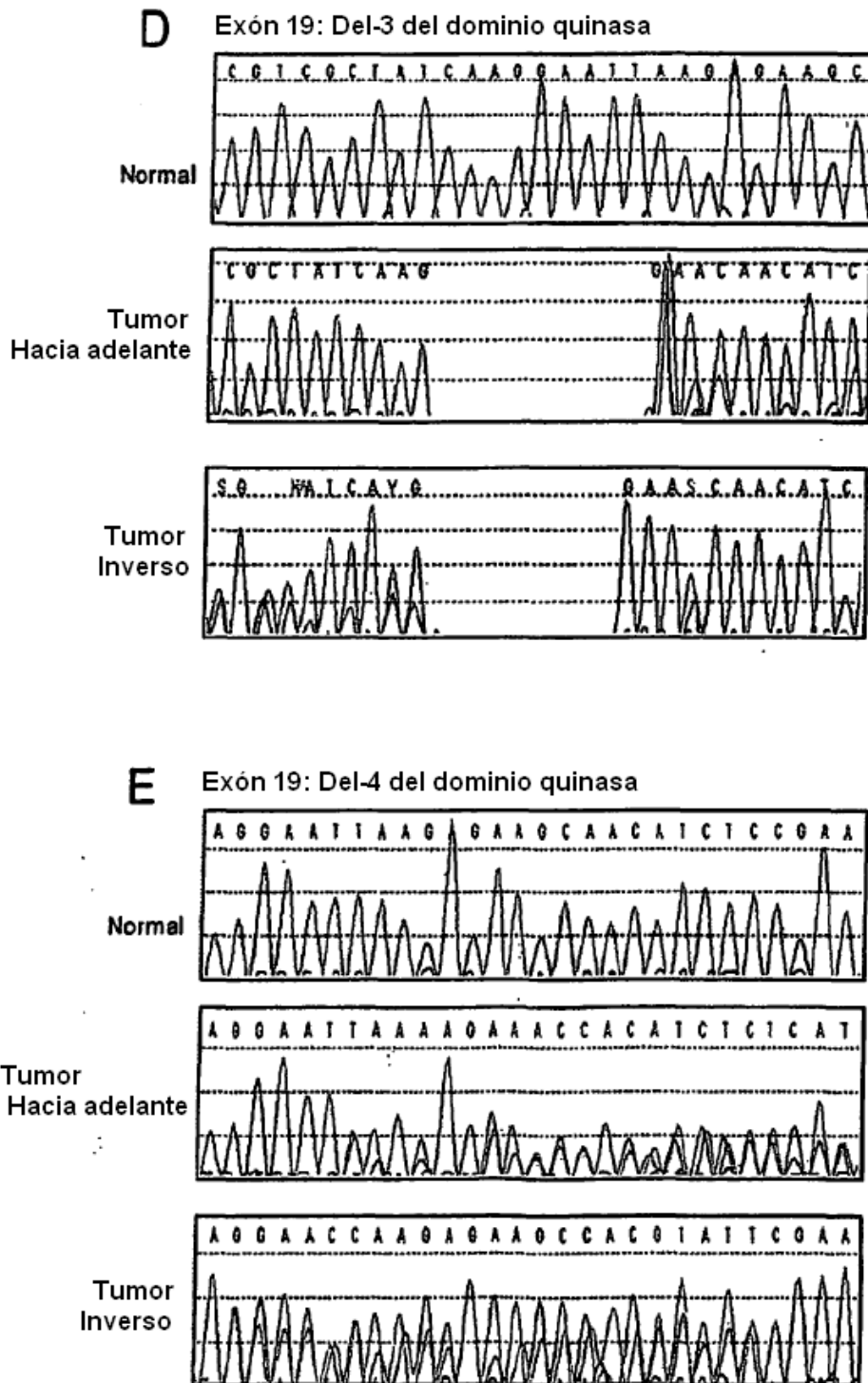
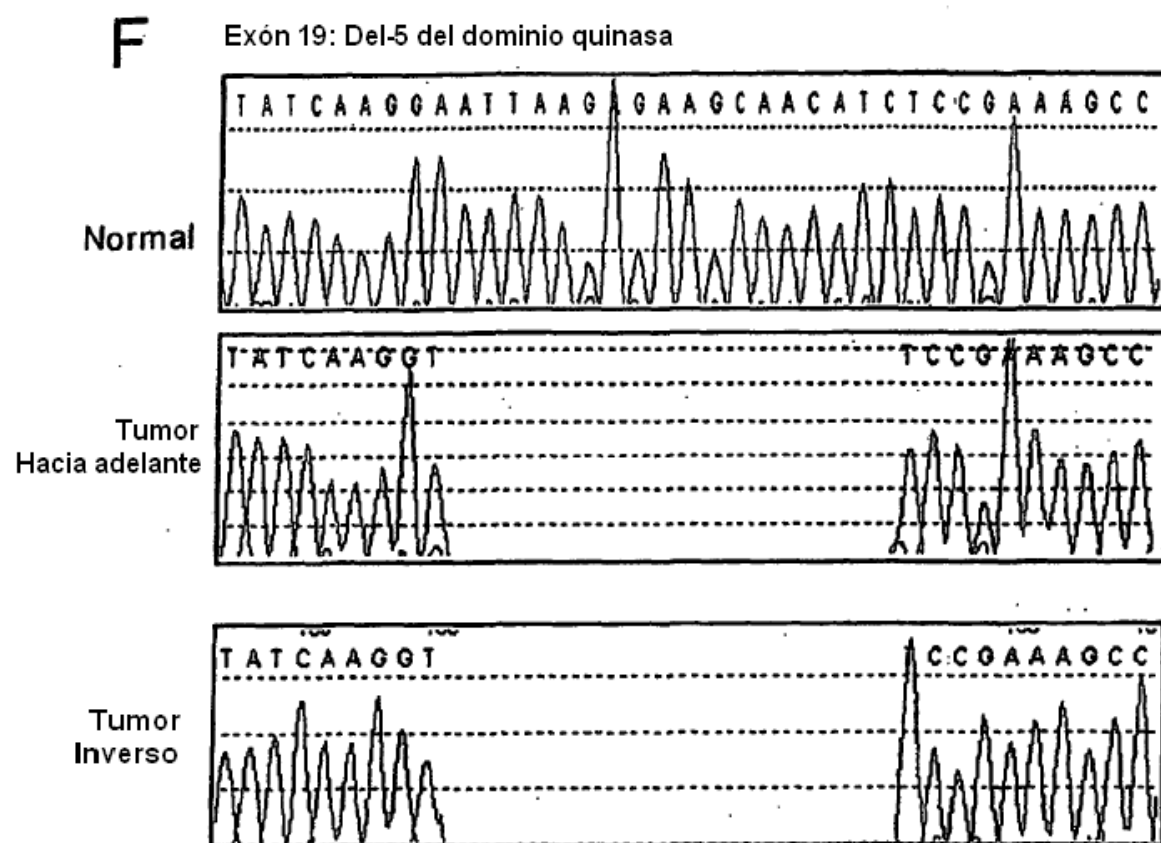


FIG. 8 (cont.)



Alineación de Abl1 con EGFR

712

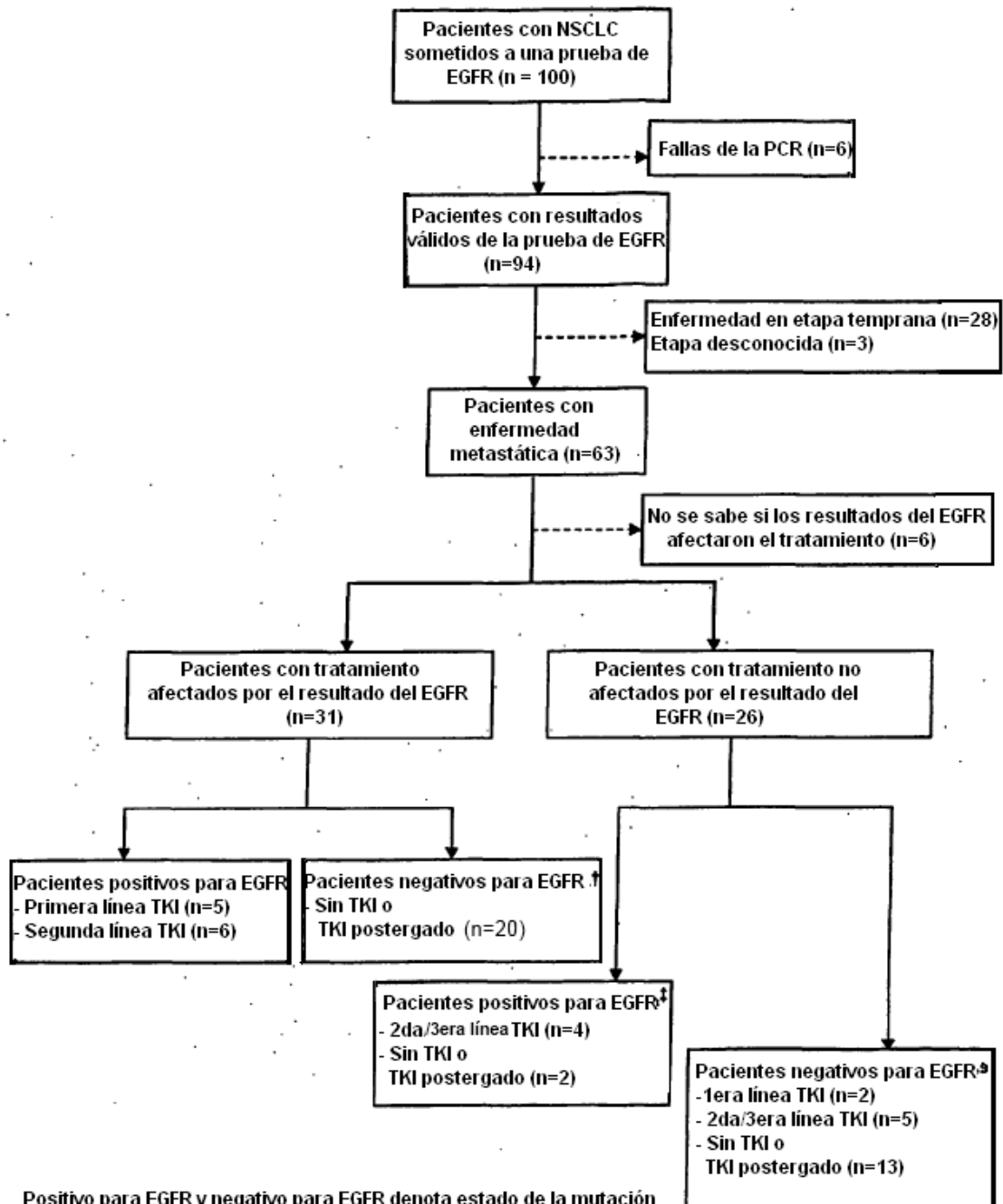
Ab11 (242) ITMKHKLGGQYGEVYGVNK-----KYSLTVAVTKLKEDT--MEVEEFKEAAVMKEIKHPNLVQLLGVCTREPPFYIIT
EGFR (712) FKIKVLGSGAFGTYYKGLWPEGEKVKIPVAIKELREATSPKANKEILD EAYVMASVDNPHVCRLLGICLTS-TVQLIT
Consenso LG G FG VY GLW K I VAIK LKE T E L EA VM I P L LGIC IIT
791 869

Ab11 (316) EPWTYGNLLDYLR ECRNQEVNAVLLYMATQISSAMEYLEKKNF IHRDLAARNCLVGENHLVKVADFGLSRLMTGDTYTA
EGFR (791) QLMPFGCLLDYVREHKDN-IGSQYLLNWCVCQIAKGMNYLEDRLVHRDLAARNVLVKTQPHVKITDFGLAKLLGAEEKEY
Consenso M FG LLDYLR N I A LL QIA AM YLE K IHRDLAARN LV VKI DFGLAKLL AD
870 949

Ab11 (396) HA-GAKFPKWTAPESLAYNKFSTKSDVWAFGVLLWEIATYGMSPYPIDLSQVYELLEKDYMERPEGCPEKVYELMRA
EGFR (870) HAEGGKVPKWMALESILHRIYTHQSDVWSYGVTVWELMTFGSKPYDGPASEISSILEKGERLPQPPICTIDVYIMVW
Consenso HA GAK PIKW A ESI H FS SDVWAFGV LWEI TFG PY GI S I ILEK RL P C VY IM
950 979

Ab11 (475) CWQWNPDRPSFAEIHQAF----- (SEQ ID NO 491)
EGFR (950) CW MIDADSRPKFRELIIEFSKWARDPQRYL (SEQ ID NO 492)
Consenso. CW RP F EI F

Figura 10: El proceso de toma de decisión para pacientes con NSCLC metastático que son sometidos a una prueba de EGFR



Positivo para EGFR y negativo para EGFR denota estado de la mutación

[^] 2 pacientes inscritos en un ensayo clínico de 1era línea de TKI

[†] 2 pacientes que reciben TKI antes del ensayo

[‡] 3 pacientes que reciben TKI antes del ensayo

[§] 1 paciente inscrito en un ensayo clínico de 1era línea de TKI y 4 pacientes que reciben TKI antes del ensayo

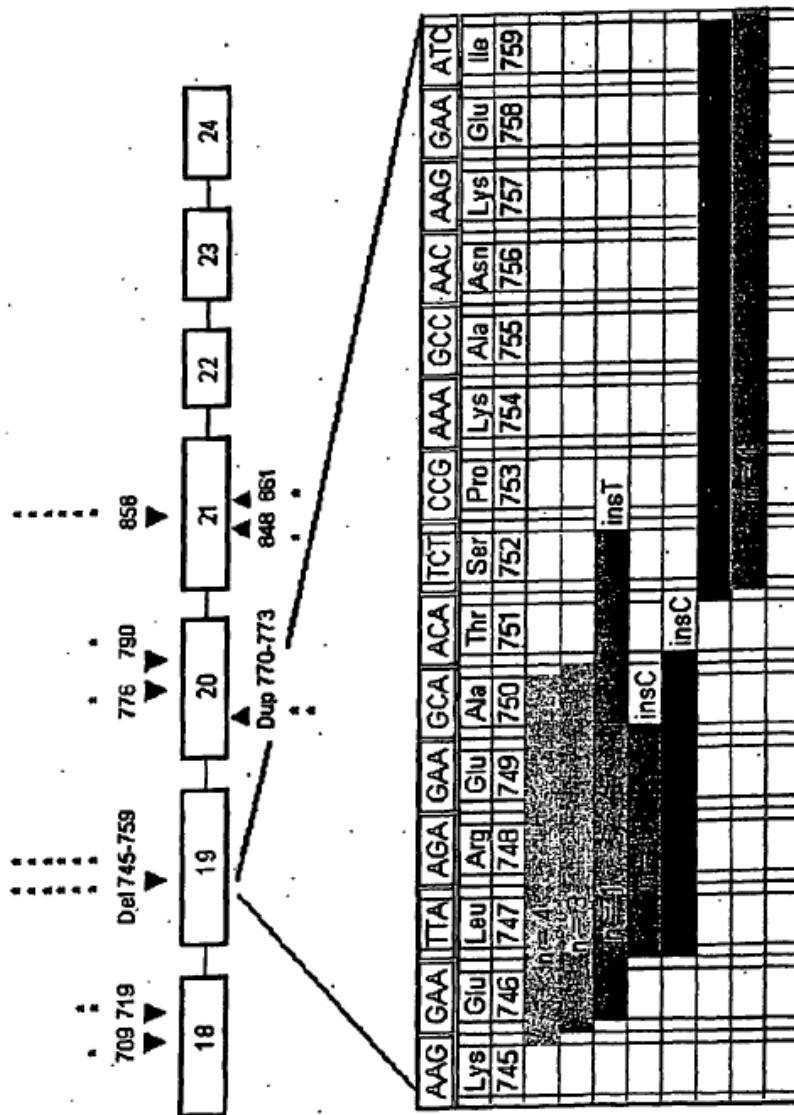


Figura 11. Diagrama de exones 18 - 24 del EGFR, que codifican la región de TK (no está a escala). Las fechas describen la ubicación de mutaciones identificadas. Los asteriscos denotan el número de pacientes con mutaciones en cada ubicación. El diagrama ampliado representa la superposición de las supresiones del exón 19, y el número de pacientes (n) con cada supresión