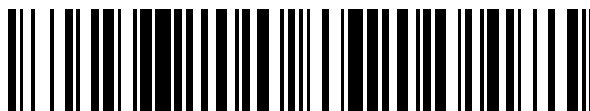


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 556 332**

51 Int. Cl.:

C07H 21/04 (2006.01)
C07K 14/00 (2006.01)
C07K 16/00 (2006.01)
G01N 33/53 (2006.01)
A61K 39/395 (2006.01)
C07K 14/705 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.11.2000** **E 10183164 (2)**

97 Fecha y número de publicación de la concesión europea: **14.10.2015** **EP 2316843**

54 Título: **Aislamiento de cinco genes novedosos que codifican nuevos melanomas de tipo receptor de Fc implicados en la patogénesis del linfoma/melanoma**

30 Prioridad:

29.11.1999 US 168151 P

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

15.01.2016

73 Titular/es:

**THE TRUSTEES OF COLUMBIA UNIVERSITY IN
THE CITY OF NEW YORK (100.0%)
West 116th Street and Broadway
New York, New York 10027, US**

72 Inventor/es:

DALLA-FAVERA, RICCARDO

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 556 332 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

Aislamiento de cinco genes novedosos que codifican nuevos melanomas de tipo receptor de Fc implicados en la patogénesis del linfoma/melanoma

- 5 La invención descrita en la presente memoria se realizó en el curso del trabajo bajo la Concesión NCI Núm. CA 44029 del Instituto Nacional del Cáncer. En consecuencia, el Gobierno de Estados Unidos tiene ciertos derechos sobre esta invención.

A lo largo de toda esta solicitud, las diversas referencias se mencionan entre paréntesis.

Las citas bibliográficas completas para estas referencias se pueden encontrar al final de esta solicitud, antes de las reivindicaciones.

10 **Antecedentes de la invención**

- Las anomalías del cromosoma 1q21 son comunes en las neoplasias malignas de células B, incluyendo el linfoma de células B y el mieloma, pero los genes elegidos como diana por estas aberraciones son en gran parte desconocidos. Mediante la clonación de los puntos de rotura de una translocación cromosómica t(1;14)(q21;q32) en una línea celular de mieloma, los autores de la presente invención han identificado dos nuevos genes, *IRTA1* e *IRTA2*, que codifican receptores de la superficie celular con homologías con las familias de Receptores de Fc e Inhibidores. Ambos genes se expresan normalmente en las células B maduras, pero con diferentes distribuciones en los órganos linfoides periféricos: *IRTA1* se expresa en células B de la zona marginal, mientras *IRTA2* también se expresa en centrocitos del centro germinal y en inmunoblastos. Como resultado de la translocación t(1;14), el péptido señal *IRTA1* se fusiona con el dominio C α de la Inmunoglobulina para producir una proteína de fusión *IRTA1*/Ca quimérica. En las líneas celulares de mieloma múltiple y de linfoma de Burkitt con anomalías en 1q21, la expresión de *IRTA2* está desregulada. Por lo tanto, *IRTA1* e *IRTA2* son inmunorreceptores novedosos con un papel potencialmente importante en el desarrollo de células B y la linfomagénesis.

- El linfoma de células B no Hodgkin (LNH-B) y el mieloma múltiple (MM) representan un grupo heterogéneo de neoplasias malignas derivadas de células B maduras con fenotipos correspondientes a células del pre-Centro Germinal (CG) (células del manto), GC (células grandes difusas, foliculares, de Burkitt), o post-CG (MM) (para la revisión, Gaidano y Dalla-Favera, 1997; Kuppers et al., 1999). Se ha obtenido una información valiosa en torno a la patogénesis de estas neoplasias malignas mediante la identificación de anomalías cromosómicas clonales recurrentes características para los subtipos específicos de las enfermedades. La consecuencia común de estas translocaciones es la desregulación transcripcional de los protooncogenes por su yuxtaposición a los elementos reguladores de la transcripción heterólogos localizados en el cromosoma compañero (Gaidano y Dalla-Favera, 1997). Estos elementos reguladores de la transcripción heterólogos pueden derivar del locus de inmunoglobulina (IG) o de otros loci cromosómicos del compañero. Los ejemplos incluyen *MYC* en t(8;14) (q24;q32) en el linfoma de Burkitt (LB) (Dalla-Favera et al., 1982; Taub et al., 1982), el gen *CCND1* desregulado por la t(11;14) (q13;q32) en el linfoma de células del manto (LCM) (Rosenberg et al., 1991) y mieloma múltiple (MM) (Ronchetti et al., 1999), *BCL2* implicado en la t(14;18) (q32;q21) en el linfoma folicular (LF) (Bakhshi et al., 1985), *BCL6* en t(3;14) (q27;q32) en linfoma difuso de células B grandes (LDCG) (Ye et al., 1993), así como *FGFR3* en t(4;14) (p16;q32). (Chesi et al., 1997), *MAF* en t(14;16) [q32;q23] (Chesi et al., 1998) y *MUM1/IRF4* en t(6;14) (p25;q32) (Iida et al., 1997) en el mieloma múltiple (MM). La identificación de estos oncogenes ha ofrecido información valiosa sobre la patogénesis y el diagnóstico de sus correspondientes neoplasias malignas.

- Las anomalías cromosómicas que están implicadas en la banda 1q21-q23 se encuentran entre las lesiones genéticas más frecuentes tanto en LNH-B como en MM. Entre los subtipos de LNH, se ha informado sobre puntos de rotura en 1q21-q23, incluyendo translocaciones y duplicaciones, a menudo como la única anomalía cromosómica, en 17-20% de los linfomas de células B grandes difusos y foliculares (LDCG), en 39% de los linfomas de células B de la zona marginal (Offit et al., 1991; Whang Peng et al., 1995; Cigudosa et al., 1999) y en de 27-38% de los linfomas de Burkitt, donde representan la segunda anomalía citogenética más común después de translocaciones que implican al proto-oncogén *MYC* (Berger y Bernheim, 1985; Kornblau et al., 1991). La hibridación comparativa del genoma (HCG) también ha identificado 1q21-q23 como un sitio recurrente para la amplificación de alto nivel en 10% de los casos de LDCG (Rao et al., 1998). En el MM, la trisomía de la región 1q21-q32 ha sido referida en 20-31% de los casos (Sawyer et al., 1995), la amplificación de la región 1q12-qter en 80% de las líneas celulares y 40% de los tumores primarios (Avet-Loiseau et al., 1997), y las translocaciones de todo el brazo desequilibradas no aleatorias de 1q, asociadas con la multiduplicación de la región adyacente 1q21-22, se encontraron en 23% de los pacientes con cariotipos anormales (Sawyer et al., 1998).

- La alta frecuencia de implicación de los reordenamientos estructurales 1q21 en las neoplasias malignas de células B sugiere que este locus puede albergar genes críticos para la patogénesis de estas enfermedades. La clonación de una t(1;14)(q21;q32) en una línea celular de leucemia linfoblástica aguda pre-B identificó previamente un nuevo gen, *BCL9* desregulado solamente en este caso (Willis et al., 1998), pero no implicado en otros casos. Un informe reciente caracterizó la t(1;22) (q22;q11) en una línea celular de linfoma folicular (LF) y se encontró que el locus *FCGR2B*, que codifica el receptor Fc de IgG de baja afinidad FCGR1b, era elegido como diana en esta línea celular

y en dos casos adicionales de FL (Callanan et al., 2000). Por último, el locus *MUC1* ha sido identificado en la proximidad del punto de rotura de una t(1;14)(q21;q32) en NHL (Dyomin et al., 2000; Gilles et al., 2000), y se han encontrado reordenamientos del locus *MUC1* en 6% de los NHL con anomalías 1q21 (Dyomin et al., 2000). Estos resultados ponen de manifiesto la heterogeneidad de los puntos de rotura de 1q21 y la necesidad de identificar oncogenes candidatos adicionales situados en este locus, ya que la gran mayoría de estas alteraciones permanecen sin explicación.

Adicionalmente, el documento EP-A2 0 330 191 describe un método para la clonación de ADNc a partir de bibliotecas de expresión de mamíferos basadas en la expresión transitoria en células anfitrionas de mamífero.

La Base de Datos EMBL, clon IMAGE: 1333716, Núm. de Acceso de la Base de datos AA811806 describe una secuencia de ácido ribonucleico mensajero (ARNm) humano de 475 pares de bases (AA811806.1) que codifica un clon de ácido desoxirribonucleico complementario humano. El documento WO-A1 00/17222, citado en virtud del Artículo 54 (3) EPC, comenta proteínas humanas secretadas y ácidos nucleicos aislados que contienen las regiones de los genes que codifican tales proteínas. La secuencia de aminoácidos codificada por el gen 11 muestra poca o ninguna homología con la proteína IRTA1 de la presente invención.

El objetivo de este estudio fue explorar más a fondo la arquitectura de los reordenamientos cromosómicos 1q21 en neoplasias malignas de las células B. Para ello, los autores de la presente invención han empleado un enfoque de clonación molecular de la t(1;14)(q21;q32) presente en la línea celular de mieloma FR4. Los autores de la presente invención han identificado dos nuevos genes que son elegidos como diana diferencialmente por anomalías de 1q21. Estos genes codifican cinco nuevos miembros de la familia de receptores de inmunoglobulina, *IRTA1*, *IRTA2*, *IRTA3*, *IRTA4* y *IRTA5* (Immunoglobulin superfamily Receptor Translocation Associated genes (Genes asociados a la translocación de receptores de la superfamilia de inmunoglobulina) 1, 2, 3, 4, y 5), que pueden ser importantes para la función de los linfocitos normales y para las neoplasias malignas de células B.

Compendio de la invención

Esta invención proporciona una molécula de ácido nucleico aislada que codifica la proteína receptora de inmunoglobulina, Asociada a la Translocación de Receptores de la superfamilia de Inmunoglobulinas, IRTA1.

Se describe un método para producir un polipéptido IRTA (proteína) que comprende: (a) introducir un vector que comprende un ácido nucleico aislado que codifica una proteína receptora de inmunoglobulina, Asociada a la Translocación de Receptores de la superfamilia de Inmunoglobulinas, IRTA, en una célula anfitriona adecuada; y (b) cultivar la célula resultante para producir el polipéptido.

Esta invención proporciona una molécula de ácido nucleico aislada que comprende al menos 15 nucleótidos contiguos capaz de hibridar específicamente con una secuencia única incluida dentro de la secuencia de la molécula de ácido nucleico aislada que codifica la proteína IRTA1, o uno o varios fragmentos de la misma, que tiene la secuencia de aminoácidos expuesta en la Figura 18A.

Esta invención proporciona un método para detectar una neoplasia maligna de células B o un tipo de neoplasia maligna de células B en una muestra de un sujeto en donde la neoplasia maligna de células B comprende una reordenación cromosómica 1q21, que comprende:

a) poner en contacto la muestra de ARN obtenida del sujeto con una molécula de ácido nucleico de al menos 15 nucleótidos contiguos capaz de hibridar específicamente con una secuencia única incluida dentro de la secuencia de un ARN aislado que codifica la proteína IRTA1 humana en condiciones que permiten la hibridación del ARN de la etapa (a) con la molécula de ácido nucleico capaz de hibridar específicamente con una secuencia única incluida dentro de la secuencia de un ARN aislado que codifica una proteína IRTA1 humana, en donde la molécula de ácido nucleico está marcada con un marcador detectable; y b) detectar cualquier hibridación en la etapa (a), en donde la detección de hibridación indica la presencia de neoplasia maligna de células B o un tipo de neoplasia maligna de células B en la muestra.

Esta invención proporciona un oligonucleótido antisentido que tiene una secuencia capaz de hibridar específicamente con una molécula de ARNm que codifica una proteína IRTA1 humana con el fin de evitar la expresión en exceso de la molécula de ARNm.

Esta invención proporciona una proteína IRTA1 purificada que comprende la secuencia de aminoácidos expuesta en la Figura 18A (SEQ ID NO: 1).

Esta invención proporciona uno o varios anticuerpos dirigidos a un epítipo de una proteína IRTA1 purificada, o uno o varios fragmentos de los mismos, que tienen la secuencia de aminoácidos expuesta en la Figura 18A.

Esta invención proporciona un anticuerpo dirigido a una proteína IRTA1 purificada.

Esta invención proporciona una composición farmacéutica que comprende una cantidad del anticuerpo dirigido a una proteína IRTA1 eficaz para unirse a las células cancerosas que expresan una proteína IRTA1 humana con el fin de

evitar el crecimiento de las células cancerosas y un portador farmacéuticamente aceptable.

Se describe una composición farmacéutica que comprende una cantidad de cualquiera de los oligonucleótidos de moléculas de ácido nucleico que codifican las proteínas IRTA descritas en la presente memoria eficaz para prevenir la expresión en exceso de una proteína IRTA humana y un portador farmacéuticamente aceptable idóneo. Esta invención proporciona un método de diagnóstico de neoplasias malignas de células B que comprende una reordenación cromosómica 1q21 en una muestra de un sujeto que comprende:

a) poner en contacto la muestra obtenida del sujeto con un anticuerpo dirigido a una proteína IRTA1 purificada capaz de unirse específicamente con una proteína IRTA1 humana sobre la superficie celular de una célula cancerosa en condiciones que permiten la unión del anticuerpo con la proteína IRTA1 humana sobre la superficie celular de la célula cancerosa, en donde el anticuerpo está marcado con un marcador detectable; y b) detectar cualquier unión en la etapa (a), en donde la detección de la unión indica un diagnóstico de neoplasia maligna de células B en la muestra.

Se describe un método para detectar la proteína IRTA humana en una muestra que comprende: a) poner en contacto la muestra con cualquiera de los anticuerpos anti-IRTA descritos anteriormente bajo condiciones que permitan la formación de un complejo entre el anticuerpo y la IRTA en la muestra; y b) detectar el complejo formado en la etapa (a), detectando de ese modo la presencia de IRTA humana en la muestra. Esta invención proporciona el uso para el tratamiento de un sujeto que tiene un cáncer de células B, que comprende administrar al sujeto una cantidad de anticuerpo anti-IRTA1 eficaz para unirse a las células cancerosas que expresan una proteína IRTA1 con el fin de evitar el crecimiento de las células cancerosas y un portador farmacéuticamente aceptable, tratando de este modo al sujeto.

Se describe un uso para el tratamiento de un sujeto que tiene un cáncer de células B, que comprende administrar al sujeto una cantidad de un oligonucleótido antisentido que tiene una secuencia capaz de hibridar específicamente con una molécula de ARNm que codifica una proteína IRTA humana con el fin de evitar la expresión en exceso de la proteína IRTA humana, con el fin de detener el crecimiento celular o inducir la muerte celular de las células cancerosas que expresan la proteína o las proteínas IRTA y un portador farmacéuticamente aceptable, tratando de este modo al sujeto.

Asimismo se describe una composición farmacéutica que comprende una cantidad eficaz de cualquiera de los oligonucleótidos descritos en la presente memoria y un portador farmacéuticamente aceptable.

La invención también proporciona una composición farmacéutica que comprende una cantidad eficaz de un anticuerpo dirigido contra un epítipo de la proteína IRTA1 descrita en la presente memoria y un portador farmacéuticamente aceptable.

Breve descripción de las figuras

Figuras 1A-1B.

Clonación molecular de la translocación t(1;14)(q21;q32) en la línea celular de mieloma múltiple FR4. Fig. 1A) Representación esquemática de los clones λ FR4B-5 y λ FR4S-a, que representan los puntos de rotura der(14) y der(1), y de la línea germinal IgH y loci 1q21. Fig. 1B) Secuencia de nucleótidos de la unión del punto de rotura y su alineamiento con las regiones correspondientes de la línea germinal del cromosoma 14. S α , región de cambio de IgA; LCR: región de control del locus 3' IgH; B, *BamHI*; H, *HindIII*; X, *XhoI*.

Figuras 2A-2B.

Mapa genómico del locus 1q21 en las proximidades del punto de rotura FR4. Fig. 2A) Mapa de endonucleasas de restricción y representación esquemática de clones genómicos, es decir, bacteriófagos (1), cromosomas artificiales P1 (PAC) (2), y cromosoma artificial de levadura (YAC) (3), que abarca el locus de 1q21 de la línea germinal en la región del punto de rotura de FR4 (punta de flecha). El nombre de cada clon se coloca directamente sobre la parte superior de su representación. Los fragmentos finales derivados de los insertos de PAC y YAC se representan como círculos, ya sea con una orientación del vector SP6/T7 (PAC), ya sea con una orientación del vector del brazo izquierdo/derecho (YAC). El panel superior de la Fig. 1A representa la organización genómica de dos genes que rodean el punto de rotura de FR4. Los dos genes fueron identificados por la captura del exón de PAC 49A16. Están estrechamente espaciados en el genoma, dentro de ≤ 30 Kb uno de otro y se denominan MUM2 y MUM3 (mieloma múltiple 2 y 3). En el esquema de sus loci genómicos, los recuadros de color negro indican los exones codificantes, mientras que los recuadros de color blanco y gris claro o medio indican exones no codificantes. Los intrones conectores son líneas. MUM3 (izquierda) da lugar a tres ARNm empalmados alternativamente, que comparten todos una región no traducida 5' común (UTR), pero diversas UTR 3' (marcadas por diferentes tonos). Los números debajo de los recuadros identifican el orden de los exones del ADNc. Los exones menores de 100 pb se representan como líneas verticales delgadas. La posición y el tamaño de cada exón se determinaron por medio de secuenciación de clones genómicos de PAC y de fagos y por hibridación de sondas de ADNc con ADN del clon digerido con endonucleasas. El mapeo de PAC y YAC se realizó por digestión parcial con enzimas de corte raro seguido de Electroforesis en Gel de Campo

Pulsado e hibridación con sondas internas y derivadas de los extremos. Las líneas discontinuas alinean las regiones de solapamiento. S, SacI; H, HindIII; S, SmaI; P, PstI; P, PstI; Fig. 2B) Mapa de conexiones genéticas Genethon de 1q21 en la región del locus MUM2/MUM3. Los sitios de secuencia expresada (STS) se ordenan a una distancia aproximada determinada previamente por Dib, C., et al., (1996) Nature, 380: 162-164. El STS WI-5435 (en negrita) está contenido dentro de YAC 23GC4 y PAC 49A16. Las líneas verticales paralelas representan segmentos interrumpidos, cuyo tamaño aproximado se representa anteriormente en megabases (MB). El dimensionamiento se estimó por el tamaño de los cóntigos de YAC no quiméricos entre dos marcadores. El gen BCL9 en el centrómero fue clonado a partir de un punto de rotura de t(1;14)(q21;q32) diferente por Willis T.G. et al., (1998) Blood 91, 6: 1873-1.881. El gen FcGRIIA está en el límite de la banda cromosómica Iq21-q22.

Figuras 3A-3C.

Estructura del ARNm y patrón de expresión de MUM2. Fig. 3A) Representación esquemática del ARNm de MUM2. Los recuadros grandes rellenos representan los dominios codificantes y los recuadros vacíos estrechos representan regiones no traducidas. SP, péptido señal; EC, dominio extracelular; TM, dominio transmembrana; CIT, dominio citoplásmico; A(n), cola poliA. La región extracelular se compone de cuatro dominios similares a inmunoglobulina tal como se representa. Las señales de poliadenilación alternativas (flechas) generan tres especies de ARNm de MUM2 (a, b, c) cuya longitud (en Kb) oscila de 2,6 a 3,5. Fig. 3B) Análisis de transferencia Northern de la expresión del ARNm de MUM2 en tejidos humanos del sistema inmunitario. La sonda de ADNc utilizada para el análisis se muestra como una barra continua de color negro debajo del esquema del ARNm en la Fig. 3A). Cada calle contiene 2 µg de ARNm del tejido correspondiente. En el lado derecho de la transferencia, se representa la posición de los marcadores de peso molecular de ARN. La posición de los transcritos de ARNm de MUM2 y GAPDH se muestra por medio de flechas. (Se incluyó una sonda de GAPDH en la hibridación como un control interno - sonda marcada 0,15 ng + no marcada 50 ng). Los resultados de este análisis muestran expresión débil de MUM2 en los ganglios linfáticos y el bazo. No se detectó expresión de MUM2 en una variedad de tejidos humanos distintos (datos no mostrados). Fig. 3C) Análisis de transferencia Northern de la expresión de MUM2 en ARN total de EREB, una línea de células linfoblastoides B transformadas con EBV condicional. EREB lleva, el genoma de EBV con una proteína de fusión de EBNA2-receptor de estrógeno, activa sólo en presencia de estrógeno. Para este experimento, las células fueron cultivadas en presencia de estrógeno (1 µg/ml), seguido de la retirada de estrógeno durante los momentos indicados. Tras la retirada de estrógeno, las células EREB se someten a la detención G0/G1, determinada por la pérdida de la expresión de c-myc. En la Fig. 3C, una transferencia Northern del ARN total de EREB [10 µg por calle] se hibridó con la sonda de ADNc de MUM2 mostrada en la Fig. 3A y la sonda de control interno de GAPDH, como en la Fig. 3B. Las flechas indican la posición de los ARNm correspondientes en la transferencia de EREB. a, la banda C corresponde a la especie MUM2 en el panel de la Fig. 3A. A continuación, la misma transferencia se eliminó y se volvió a sondear con una sonda de ADNc de c-myc (exón 2) para verificar detención en la G0/G1 celular. La cuantificación del ARNm de MUM2 por medio del uso de un análisis densitométrico Phosphorimager demuestra un aumento de 10 veces en sus niveles en el plazo de 48 horas de la retirada de estrógeno, lo que sugiere que la expresión de MUM2 es elevada ya que las células entran en una fase de reposo.

Figuras 4A-4B.

Estructura de ARNm y patrón de expresión de MUM3. Fig. 4A) Representación esquemática del ARNm de MUM3. Los recuadros grandes rellenos representan dominios codificantes y los recuadros vacíos estrechos o de color gris representan regiones no traducidas. SP, péptido señal; EC, dominio extracelular; TM, dominio transmembrana; CIT, dominio citoplásmico; A(n), cola poliA. La región extracelular se compone de dominios de tipo inmunoglobulina, tal como se representa. El empalme alternativo genera cuatro especies de ARNm con diversa localización subcelular. Se secretan las proteínas MUM3-a y d, mientras que MUM3-b contiene un tramo hidrófobo de aminoácidos en su extremo C-terminal que puede servir como señal para la adición de un anclaje de glicofosfatidil-inositol (ancla de GPI), como se muestra. MUM3-c se extiende por la membrana plasmática. La identidad de secuencia entre las especies se indica mediante un relleno idéntico. Fig. 4B) Análisis de transferencia Northern de la expresión del ARNm de MUM3 en múltiples tejidos humanos (izquierda) y en diversas líneas celulares linfoides y no linfoides (derecha). La sonda de ADNc utilizada se muestra como una barra continua debajo del esquema de ADNc en la Fig. 4A. Cada calle contiene 2 µg de ARNm del tejido o la línea celular correspondiente. La posición de los transcritos de ARNm de MUM3 y GAPDH se muestra por medio de flechas. (Se incluyó una sonda de GAPDH en la hibridación como control interno como se describe en la Fig. 3). a, b, c y d corresponden a la especie de ARNm de MUM3 mostrada en la Fig. 4A. RD, NC42 y CB33, líneas de células B linfoblastoides transformadas con virus de Epstein-Barr; EREB, línea celular linfoblastoide B transformada con EBV condicional; FR4, línea de células de plasma; MOLT4 y HUT78, líneas de células T; HL60 y U937, líneas de células mielomonocíticas; K562, línea de células eritroides. Los resultados sugieren que MUM3 se expresa exclusivamente en los tejidos del sistema inmunitario de la médula ósea, la linfa y el bazo y en particular en las células B con un fenotipo linfoblastoide.

Figura 5.

Secuencias de nucleótidos y aminoácidos de MUM2 humano. La secuencia de aminoácidos deducida se muestra

por encima de la secuencia de nucleótidos con el código de una letra y se numera a la derecha, con la posición 1 ajustada al primer codón del péptido señal. El sitio de la peptidasa señal pronosticado se obtuvo por medio de un algoritmo informático descrito por Nielsen et al., Protein Engineering 10, 1-6 (1997) y se marca por medio de una punta de flecha. La señal de poliadenilación AATAAA está subrayada. Los sitios potenciales para la N-glicosilación están subrayados también en la secuencia de aminoácidos. Un tramo hidrófobo de 16 aminoácidos que se pronostica que abarca la membrana plasmática está doblemente subrayado. Los sitios de unión a SH2 consenso se destacan por medio de un subrayado ondulado.

Figura 6A.

Secuencia de nucleótidos y de aminoácidos de MUM3-a humana. La secuencia de aminoácidos deducida se muestra por encima de la secuencia de nucleótidos con el código de una letra y se numera a la derecha, con la posición 1 ajustada al primer codón del péptido señal. El sitio pronosticado para la escisión de la peptidasa señal se obtuvo como se ha indicado anteriormente y se marca por medio de una punta de flecha. La señal de poliadenilación AATAAA está subrayada. Los sitios potenciales para la N-glicosilación también están subrayados en la secuencia de aminoácidos. La proteína carece de un dominio transmembrana y se prevé que sea secretada.

Figura 6B.

Secuencia de nucleótidos y de aminoácidos MUM3-b humana. La secuencia de aminoácidos deducida se muestra por encima de la secuencia de nucleótidos con el código de una letra y se numera a la derecha, con la posición 1 ajustada al primer codón del péptido señal. El sitio pronosticado para la escisión de la peptidasa señal se obtuvo como se ha indicado anteriormente y se marca por medio de una punta de flecha. La señal de poliadenilación AATAAA está subrayada. Los sitios potenciales para la N-glicosilación están subrayados en la secuencia de aminoácidos.

Figura 6C-1-6C-2.

Secuencia de nucleótidos y de aminoácidos de MUM3-c humana. La secuencia de aminoácidos deducida se muestra por encima de la secuencia de nucleótidos con el código de una letra y se numera a la derecha, con la posición 1 ajustada al primer codón del péptido señal. El sitio pronosticado para la escisión de la peptidasa señal se obtuvo como se ha indicado anteriormente y se marca por medio de una punta de flecha. La señal de poliadenilación AATAAA está subrayada. Los sitios potenciales para la N-glicosilación están subrayados en la secuencia de aminoácidos. Un tramo hidrófobo de 23 aminoácidos que se pronostica que abarca la membrana plasmática está doblemente subrayado. Los sitios de unión a SH2 consenso se destacan por medio de un subrayado ondulado.

Figuras 7A-7C.

t(1;14)(q21;q32) en FR4 genera un transcrito de fusión MUM2/Ca. Fig. 7A) Representación esquemática del clon genómico der(14) λFR4B-5 y del locus IgHA1 de la línea germinal. El punto de rotura de FR4 está marcado por medio de una flecha. Los recuadros rellenos y huecos representan los exones codificantes de MUM2 y Calfa y los no codificantes, respectivamente. La posición de la sonda 1 para el exón de MUM2 utilizada para el análisis de transferencia Northern se muestra por medio de una barra. Fig. 7B) El análisis de transferencia Northern con una sonda para el exón 1 de MUM2 sobre FR4 y líneas celulares adicionales detecta un mensaje anormal de 0,8 Kb, selectivamente en FR4. Las puntas de flecha apuntan a la ubicación de mensaje normal de MUM2 en el ARNm de EREB. JJN3 y U266, líneas celulares de mieloma; EREB, línea celular linfoblástico B transformada con EVB condicional. Se cargaron dos µg de ARN poliA+ por calle. Fig. 7C) Secuencia de nucleótidos y aminoácidos del ADNc de la fusión MUM2-Ca en FR4. El ADNc se amplificó por medio de RT-PCR a partir del ARN total de FR4 usando los cebadores mostrados en la Fig. 7A, y posteriormente se subclonó y se secuenció. La secuencia de aminoácidos deducida se muestra por encima de la secuencia de nucleótidos con el código de una letra y se numera a la derecha con la posición 1 ajustada al primer codón del péptido señal. El sitio pronosticado para la escisión de la peptidasa señal se obtuvo como se ha indicado anteriormente y se marca por medio de una punta de flecha. La señal de poliadenilación AATAAA está subrayada. El dominio transmembrana Calfa está subrayado. La porción MUM2 del ADNc se muestra en cursiva. H, HindIII; B, BamHI; X, XhoI; Sa, región de cambio de IgA; EC, región extracelular; TM, transmembrana; CIT, dominio citoplásmico.

Figuras 8A-8C.

Clonación molecular de la translocación t(1;14)(q21;q32) en la línea celular de mieloma múltiple FR4. Fig. 8A) Representación esquemática de los clones de fagos que representan los puntos de rotura der(14) y der(1) y los loci IGH y 1q21 de la línea germinal. Se indican 14 secuencias de cromosomas mediante una línea continua de color negro representando los recuadros de color negro los exones Ca1. Las secuencias del cromosoma 1 se muestran como una línea de color gris. Las sondas utilizadas para el mapeo cromosómico se indican debajo del mapa. Los códigos para las enzimas de restricción son: B, BamHI; H, HindIII; X, XhoI; S, SacI; E, EcoRI. Para las enzimas marcadas con un (*) sólo se muestran los sitios que delinean las sondas. Sa: región de cambio de IgA; LCR: región de control del locus 3'IgH. Fig. 8B) Secuencia de nucleótidos de las uniones de los puntos de

rotura y su alineamiento con las correspondientes regiones de la línea germinal de los cromosomas 14 y 1. Fig. 8C) A la izquierda, análisis de hibridación fluorescente in situ (FISH) sobre dispersiones de metafase humana normales abarcando el clon PAC 49A16 (Fig. 13) la región q21 de la línea germinal en el punto de rotura de FR4. Derecha, imagen teñida con DAPI de la misma dispersión de metafase.

5 Figuras 9A-9B.

Estructura de los ADNc de *IRTA1* e *IRTA2*. Figs.9A, 9B) Representación esquemática de los ADNc de *IRTA1* (Fig. 9A) e *IRTA2* (Fig. 9B) completos. Los recuadros rellenos, grandes representan dominios codificantes y los recuadros estrechos representan las regiones no traducidas (UTR). El sitio pronosticado para la escisión de la peptidasa señal está marcado por medio de una punta de flecha y se obtuvo de acuerdo con el servidor Red Informática Mundial SignallP en <http://www.cbs.dtu.dk/services/SignalP>. El algoritmo de predicción de dominios transmembrana es descrito por Tusnady et al., 1998. SP, péptido señal; EC, dominio extracelular; Ig, de tipo inmunoglobulina; TM, dominio transmembrana; CIT, dominio citoplásmico; A(n), cola poli A; GPI, glicosofatidil inositol. En la (Fig. 9A), las flechas en la UTR 3' indican diferentes sitios de adición de poliadenilación utilizados en el ADNc de *IRTA1*. En la (Fig. 9B), se somborean de manera diferente las regiones UTR 3' de las isoformas de *IRTA2*. Las barras debajo de las regiones UTR en la (Fig. 9A) y la (Fig. 9B) identifican las sondas utilizadas para el análisis de transferencia Northern en la Figura 12.

Figuras 10A-10B.

Comparación de las secuencias de aminoácidos de *IRTA1* e *IRTA2* con miembros de la familia de receptores de Fc Fig. 10A) Alineamiento de secuencias múltiples de los dos primeros (arriba) y el tercero (abajo) dominios extracelulares para Ig de *IRTA1* e *IRTA2* para los miembros de la familia de receptores de Fc. Las secuencias se compararon utilizando el programa ClustalW (Thompson et al., 1994). Los recuadros sombreados de color negro indican aminoácidos conservados entre todas las secuencias; los recuadros sombreados de color gris oscuro indican aminoácidos conservados entre al menos la mitad de las secuencias; los recuadros sombreados en color claro indican las sustituciones conservativas. Fig. 10B) Alineamiento de los dominios de unión a SH2 de *IRTA1* e *IRTA2* con los motivos de consenso ITAM e ITIM. Las posiciones de aminoácidos conservados están en negrita. El símbolo X representa cualquier aminoácido.

Figuras 11A-11B-4.

Patrón de expresión de *IRTA1*. Fig. 11A) Panel izquierdo. Análisis de transferencia Northern de la expresión del ARNm de *IRTA1* en tejidos del sistema inmunitario humano. Cada calle contiene 2 mg de ARNm. La posición de los marcadores de peso molecular de ARN se representa en el lado derecho de la transferencia. Las posiciones de los transcritos de ARNm de *IRTA1* y *GAPDH* se muestran por medio de flechas. (Se incluyó una sonda de *GAPDH* en la hibridación como un control interno - sonda marcada 0,15 ng + no marcada 50 ng). Panel derecho. Análisis de transferencia Northern de la expresión de *IRTA1* en el ARN total de la línea celular ER/EB (10 mg por calle). Para este experimento, las células fueron cultivadas en presencia de estrógeno (1 mg/ml), seguido de la retirada de estrógeno durante los momentos indicados. Las flechas indican las posiciones de los ARNm correspondientes. a, b y c corresponden a especies poliadeniladas diferencialmente de *IRTA1*. La misma transferencia se eliminó y se volvió a sondear con una sonda de ADNc MYC (exón 2) para verificar la detención de G₀/G₁ celular. El análisis densitométrico de los niveles de ARNm de *IRTA1* se representa en el gráfico de columnas adyacente. La sonda de ADNc utilizada se muestra como una barra continua por debajo del esquema de ARNm de *IRTA1* en la Figura 9A. Fig.11B-1-11B-4) Análisis de hibridación *in situ* de la expresión de *IRTA1* en secciones seriadas de amígdala humana. 1. Sonda *IRTA1* efectora 2. Sonda *IRTA1* antisentido 3. Tinción H & E 4. Señal *IRTA1* antisentido superpuesta sobre una sección teñida con H & E. GC, centro germinal, MargZ, zona marginal

Figura 12A-12B-4.

Patrón de expresión de *IRTA2*. Fig. 12A) Análisis de transferencia Northern de la expresión de ARNm de *IRTA2* en múltiples tejidos humanos (panel izquierdo) y en diversas líneas celulares linfoides y no linfoides (panel derecho). Cada calle contiene 2 mg de ARNm. Las posiciones de los transcritos de *IRTA2* y *GAPDH* se muestran por medio de flechas. a, b, c y d corresponden a isoformas de ARNm de *IRTA2* de corte y empalme alternativo. RD, NC42 y CB33, líneas de células B linfoblastoides transformadas con virus de Epstein-Barr; EREB, línea celular linfoblastoide B transformada con EBV condicional; FR4, línea de células de plasma; MOLT4 y HUT78, líneas de células T; HL60 y U937, líneas de células mielomonocíticas; K562, línea celular eritroide. La sonda de ADNc utilizada se muestra como una barra continua debajo del esquema del ARNm de *IRTA2* en la Figura 9B. Figs.12B-1-12B-4) Análisis de hibridación *in situ* de la expresión de *IRTA2* en la amígdala humana. Fig. 12B-1. Sonda de ADNc de *IRTA2* efectora, Fig. 12B-2. Sonda de ADNc de *IRTA2* antisentido, Fig. 12B-3. Tinción H & E, Fig. 12B-4. Señal de la sonda de ADNc de *IRTA2* antisentido superpuesta sobre la sección teñida con H & E. GC, centro germinal, MargZ, zona marginal

Figura 13.

Mapa de la región 1q21 de la línea germinal que abarca el punto de rotura FR4 y la organización genómica de

IRTA1 e IRTA2. Los cebadores utilizados para amplificar los exones de IRTA1 a partir de ADNc de bazo están marcados por medio de puntas de flecha en el panel superior. Los recuadros de color negro y claro indican exones codificantes y no codificantes respectivamente. Las flechas indican la posición de *BCL9*, *MUC1*, familia *IRTA* y *FCGR1B*. loci. S, SacI; H, HindIII; S, SwaI; Pc, PaeI; P, Pml; Mb, Megabases

Figuras 14A-14D.

t(1;14)(q21;q32) en FR4 genera un transcrito de fusión *IRTA1/Cα*. Fig. 14A) Representación esquemática del clon genómico der(14) 1FR4B-5 y del locus *IgCα1* de la línea germinal. El punto de rotura FR4 está marcado por una flecha. Los recuadros rellenos y vacíos representan los exones codificantes y no codificantes de *IRTA1* y *Cα1*, respectivamente. Fig. 14B) El análisis de transferencia Northern con una sonda del exón 1 de *IRTA1* (que se muestra por medio de una barra en la Fig. 14A) sobre FR4 y otras líneas celulares detecta un mensaje anormal en FR4. Las puntas de flecha apuntan a la ubicación del mensaje de IRTA1 normal en el ARNm de ER/EB. JJN3 y U266, líneas celulares de mieloma. Se cargaron dos mg de ARN poliA+ por calle. Fig. 14C) Representación esquemática del ADN de fusión *IRTA1/Cα* en FR4. El ADNc se amplificó por RT-PCR a partir de ARN total de FR4 utilizando los cebadores mostrados en la (Fig.14A), y se secuenció después de la subclonación. Fig. 14D) Análisis SDS/PAGE de los productos inmunoprecipitados obtenidos a partir de células 293-T' transfectadas con control de vector y transfectadas con constructos de expresión transitoria de *IRTA1/cα* (calles 1 y 2), o las siguientes líneas celulares: línea de células linfoblastoides positiva para mIgA Dakiki (calle 3), FR4 (calle 4), línea de células NHL positivas para mIgM Ramos (calle 5). H, HindIII; B, BamHI; X, XhoI; Sa, región de cambio de IgA; EC, región extracelular; TM, transmembrana; CIT, citoplasmática.

Figuras 15A-15B.

La expresión de *IRTA2* está desregulada en líneas celulares que portan anomalías 1q21. Figs. 15A, 15B) análisis de transferencia Northern de la expresión del ARNm de *IRTA2* en líneas celulares de linfoma de Burkitt (Fig. 15A) y mieloma múltiple (Fig.15B). La sonda de ADNc utilizada es la misma que en la Fig. 12. Cada calle contiene 2 mg de ARNm. Las posiciones de los transcritos de ARNm de *IRTA2* y *GAPDH* se muestran por medio de guiones y flechas, respectivamente. Los niveles relativos de expresión de ARNm de *IRTA2* en el panel de la izquierda (Fig. 15A) se representaron en el panel de la derecha (Fig. 15A) después de análisis densitométrico y la normalización frente a los niveles de *GAPDH*. El panel de la derecha (Fig. 15B) es un resumen de los resultados del análisis de transferencia Northern.

Figuras 16-1-16-4

Expresión de IRTA1 en tejido linfóide normal. Se tiñeron secciones incluidas en parafina de amígdala humana normal con los anticuerpos siguientes: Fig. 16-1) Control negativo; Fig. 16-2) Anticuerpo monoclonal de ratón anti-CD3 para detectar las células T; Fig. 16-3) anticuerpo monoclonal de ratón anti-IRTA1 (mIRTA); Fig. 16-4) anticuerpo policlonal de conejo anti-IRTA1 (J92884K). Las células positivas para IRTA1 se encuentran en la región perifolicular e intraepitelial de la amígdala, equivalente a la zona marginal en el bazo.

Figura 17

Expresión de IRTA1 en un linfoma de células B en Tejido Linfóide Asociado a Mucosa del estómago (MALT). Se tiñó una sección incluida en parafina de un linfoma de células B MALT de estómago con el anticuerpo monoclonal de ratón anti-IRTA1 (mIRTA) y se contratiñó con H & E. La mayoría de los linfomas MALT analizados eran positivos para IRTA1. Por tanto, este anticuerpo puede ser una herramienta eficaz en el diagnóstico diferencial del linfoma MALT. También se ha demostrado que el anticuerpo contra mIRTA1 es útil en la terapia de este tumor de células B, de manera similar al uso del anticuerpo anti-CD20 (rituximab) en la terapia de los linfomas positivos para CD20 con recaída (Foon K., *Cancer J.* 6: pág. 273).

Figura 18A.

ADNc de IRTA1 y secuencia de aminoácidos de la proteína IRTA1codificada.

Figuras 18B-1-18B-3.

ADNc de IRTA2 y secuencia de aminoácidos de la proteína IRTA2 codificada.

Figuras 18C-1-18C-2.

ADNc de IRTA3 y secuencia de aminoácidos de la proteína IRTA3 codificada.

Figuras 18D-1-18D-2.

ADNc de IRTA4 y secuencia de aminoácidos de la proteína IRTA4 codificada.

Figuras 18E-1-18E-2.

ADNc de IRTA5 y secuencia de aminoácidos de la proteína IRTA5 codificada.

Descripción detallada de la invención

Las siguientes abreviaturas convencionales se utilizan en toda la memoria para indicar nucleótidos específicos: C = citosina; A = adenosina; T = timidina y G = guanosina.

- 5 Esta invención proporciona una molécula de ácido nucleico aislada que codifica la proteína receptora de inmunoglobulina, Asociada a la Translocación de Receptores de la superfamilia de Inmunoglobulinas, IRTA1.

Según se utiliza en la presente memoria los genes "Asociados a la Translocación de Receptores de Inmunoglobulinas", "IRTA" son moléculas de ácido nucleico que codifican los receptores de la superficie celular novedosos de la superfamilia de inmunoglobulinas en las células B que son importantes en el desarrollo de células B, y cuya expresión anormal, p. ej., expresión desregulada, perturba las respuestas inmunológicas de las células B en la superficie celular y por lo tanto están implicadas en las neoplasias malignas de las células B, incluyendo la linfomagénesis.

10 Las moléculas de ácido nucleico que codifican proteínas denominadas proteínas "MUM-2" y "MUM-3" en la Primera Serie de Experimentos que ahora se denominan genes "IRTA-1" e "IRTA-2", es decir, moléculas de ácido nucleico que codifican las proteínas IRTA-1 e IRTA-2, respectivamente. Las proteínas IRTA-3, 4 y 5 son miembros de la misma la superfamilia de genes de inmunoglobulina como son las proteínas IRTA-1 e IRTA-2.

15 En una realización de la molécula de ácido nucleico aislada descrita anteriormente que codifica la proteína receptora de inmunoglobulina, Asociada a la Translocación de Receptores de la superfamilia de inmunoglobulinas, IRTA, la proteína IRTA codificada es la proteína IRTA1 que comprende la secuencia de aminoácidos expuesta en la Figura 18A (SEQ ID NO: 1).

20 En otra realización de cualquiera de las moléculas de ácido nucleico aisladas descritas anteriormente, la molécula de ácido nucleico es ADN. En realizaciones adicionales, el ADN es ADNc. En realizaciones adicionales, el ADN es ADN genómico. En otra realización, la molécula de ácido nucleico es una molécula de ARN. En otra realización adicional, la molécula de ADN es ADNc que tiene la secuencia de nucleótidos expuesta en la Figura 18A (SEQ ID NO: 2). En otra realización, la molécula de ADN es ADNc que tiene la secuencia de nucleótidos expuesta en la Figura 18A (SEQ ID NO: 4). En una realización adicional, la molécula de ADN es ADNc que tiene la secuencia de nucleótidos expuesta en la Figura 18A (SEQ ID NO: 6). En otra realización, la molécula de ADN es ADNc que tiene la secuencia de nucleótidos expuesta en la Figura 18A (SEQ ID NO: 8). En una realización, la molécula de ADN es ADNc que tiene la secuencia de nucleótidos expuesta en la Figura 18A (SEQ ID NO: 10). En realizaciones preferidas de la molécula de ácido nucleico aislada, las moléculas de ácido nucleico codifican la proteína IRTA1 humana. En realizaciones adicionales, las moléculas de ácido nucleico codifican la proteína IRTA1 de mamífero. La proteína IRTA1 de mamífero puede ser la proteína IRTA1 murina. En otra realización preferida, las moléculas aisladas de ácido nucleico están unidas operativamente a un promotor de la transcripción de ADN. En otra realización preferida más de la molécula de ácido nucleico aislada, el promotor comprende un promotor bacteriano, de levadura, de insectos, de planta o de mamífero.

35 Esta invención proporciona un vector que comprende cualquiera de las moléculas de ácido nucleico aisladas descritas anteriormente que codifican la proteína IRTA1, incluyendo pero no limitada a la proteína IRTA1 de mamífero, de las cuales se prefieren la humana y la murina. En una realización, el vector es un plásmido.

40 Esta invención proporciona una célula anfitriona que comprende el vector descrito anteriormente que comprende cualquiera de las moléculas de ácido nucleico aisladas descritas anteriormente que codifican la proteína IRTA1. Preferiblemente, las moléculas de ácido nucleico aisladas en tales vectores están unidas operativamente a un promotor de la transcripción de ADN. En otra realización de la célula anfitriona, se selecciona la célula de un grupo que consiste en una célula bacteriana, una célula vegetal, y una célula de insecto y una célula de mamífero.

45 Se proporciona un método para producir un polipéptido (proteína) IRTA que comprende: (a) introducir un vector que comprende un ácido nucleico aislado que codifica una proteína receptora de inmunoglobulina, Asociada a la Translocación de Receptores de la superfamilia de Inmunoglobulinas, IRTA, en una célula anfitriona adecuada; y (b) cultivar la célula resultante con el fin de producir el polipéptido. En realizaciones adicionales, la proteína IRTA producida por el método descrito anteriormente puede ser recuperada y en otra realización más, puede ser purificada total o parcialmente. En una realización, la proteína IRTA es IRTA1.

50 En una realización adicional, la proteína IRTA; puede ser una proteína de mamífero. En otra realización adicional, la proteína de mamífero puede ser una proteína IRTA humana o de ratón.

55 Los genes IRTA (moléculas de ácido nucleico que codifican proteínas IRTA IRTA1, IRTA2, IRTA3, IRTA4 e IRTA5) son útiles para la producción de las proteínas IRTA codificadas de ese modo. Las proteínas IRTA son útiles para la producción de anticuerpos; tales anticuerpos se utilizan como reactivos para el diagnóstico diferencial de los subtipos de linfoma en hematopatología. Los anticuerpos dirigidos contra las proteínas IRTA y que se unen específicamente a las proteínas IRTA también tienen usos terapéuticos, es decir, para dirigirse específicamente a las

células tumorales, que pueden ser utilizados y administrados de manera similar a "Rituximab" (un anticuerpo anti-CD20), que es un anticuerpo aprobado por la FDA para el tratamiento de los linfomas positivos para CD20 con recaída (Foon K., Cancer J. 6 (5):273). Los anticuerpos anti-IRTA1, anti-IRTA2, anti-IRTA3, anti-IRTA4 y anti-IRTA5 también son marcadores útiles para el aislamiento de subconjuntos específicos de células B en estudios de investigación de la biología de células B normales y tumorales. Por otra parte, los anticuerpos anti-IRTA1, anti-IRTA2, anti-IRTA3, anti-IRTA4 y anti-IRTA5 son reactivos de investigación útiles para estudiar experimentalmente la biología de señalización en las células B normales y tumorales.

Los métodos para introducir moléculas de ácido nucleico en células son bien conocidos para los expertos en la técnica. Tales métodos incluyen, por ejemplo, el uso de vectores virales y la co-precipitación con fosfato de calcio. Por consiguiente, las moléculas de ácido nucleico que codifican las proteínas IRTA IRTA1, IRTA2, IRTA3, IRTA4 e IRTA5 se pueden introducir en las células para la producción de estas proteínas IRTA.

Se pueden emplear numerosos vectores para la expresión de las proteínas de la invención IRTA1, IRTA2, IRTA3, IRTA4, e IRTA5. Tales vectores, incluyendo los vectores plasmídicos, los vectores cósmidos, los vectores de bacteriófagos y otros virus, son bien conocidos en la técnica. Por ejemplo, una clase de vectores utiliza elementos de ADN que derivan de virus animales tales como virus del papiloma bovino, virus de poliovirus, adenovirus, virus vaccinia, baculovirus, retrovirus (RSV, MMTV o MoMLV), virus del bosque Semliki o virus SV40. Adicionalmente, las células que han integrado de forma estable el ADN en sus cromosomas se pueden seleccionar introduciendo uno o más marcadores que permiten la selección de células anfitrionas transfectadas. Los marcadores pueden proporcionar, por ejemplo, prototrofia a un anfitrión auxótrofo, resistencia a biocidas o resistencia a metales pesados tales como el cobre. El gen marcador seleccionable puede ser directamente ligado a las secuencias de ADN que se van a expresar, o puede ser introducido en la misma célula por medio de transformación simultánea.

Los elementos reguladores requeridos para la expresión incluyen secuencias promotoras que se van a unir a ARN polimerasa y secuencias de inicio de la transcripción para la unión al ribosoma. También se pueden necesitar elementos adicionales para la síntesis óptima del ARNm. Estos elementos adicionales pueden incluir señales de empalme, así como potenciadores y señales de terminación. Por ejemplo, un vector de expresión bacteriano incluye un promotor tal como el promotor lac y para el inicio de la transcripción la secuencia de Shine-Dalgarno y el codón de inicio AUG. De un modo similar, un vector de expresión eucariótico incluye un promotor heterólogo u homólogo para la ARN polimerasa II, una señal de poliadenilación aguas abajo, el codón de inicio AUG, y un codón de terminación para la separación del ribosoma. Tales vectores se pueden obtener comercialmente o se pueden ensamblar a partir de las secuencias descritas por métodos bien conocidos en la técnica, por ejemplo los métodos descritos anteriormente para construir vectores en general.

Estos vectores se pueden introducir en una célula anfitriona adecuada para formar un sistema vector anfitrión para producir las proteínas de la invención. Los métodos de elaboración de sistemas de vectores anfitriones son bien conocidos por los expertos en la técnica.

Las células anfitrionas adecuadas incluyen, pero no se limitan a, células bacterianas (incluyendo células Gram positivas), células de levadura, células fúngicas, células de insecto y células animales. Las células animales adecuadas incluyen, pero no se limitan a células HeLa, células Cos, células CV1 y diversas células de mamíferos primarias. Se pueden utilizar numerosas células de mamíferos como anfitriones, incluyendo, pero no limitadas a, fibroblastos de ratón, células NIH-3T3, células CHO, células HeLa, células Ltk y células COS. Las células de mamíferos pueden ser transfectadas por métodos bien conocidos en la técnica tales como precipitación con fosfato cálcico, electroporación y microinyección.

Esta invención proporciona una molécula de ácido nucleico aislada que comprende al menos 15 nucleótidos contiguos capaces de hibridar específicamente con una secuencia única incluida dentro de la secuencia de la molécula de ácido nucleico aislada que codifica la proteína IRTA1, o uno o varios fragmentos de la misma, que tiene la secuencia de aminoácidos expuesta en la Figura 18A. En otras realizaciones, las moléculas de ácido nucleico aisladas se marcan con un marcador detectable. En otras realizaciones adicionales de las moléculas de ácido nucleico aisladas, el marcador detectable se selecciona del grupo que consiste en un isótopo radiactivo, enzima, colorante, biotina, una marca fluorescente o una marca quimioluminiscente.

Esta invención proporciona un método para detectar una neoplasia maligna de las células B o un tipo de neoplasia maligna de las células B en una muestra de un sujeto en donde la neoplasia maligna de las células B comprende una reordenación cromosómica 1q21, que comprende:

- a) poner en contacto la muestra de ARN obtenida del sujeto con la molécula de ácido nucleico de al menos 15 nucleótidos contiguos capaces de hibridar específicamente con una secuencia única incluida dentro de la secuencia de un ARN aislado que codifica la proteína IRTA1 humana en condiciones que permiten la hibridación del ARN de la etapa (a) con la molécula de ácido nucleico capaz de hibridar específicamente con una secuencia única incluida dentro de la secuencia de un ARN aislado que codifica la proteína IRTA1 humana, en donde la molécula de ácido nucleico está marcada con un marcador detectable; y b) detectar cualquier hibridación en la etapa (a) en donde la detección de hibridación indica la presencia de neoplasia maligna de las células B o un tipo de neoplasia maligna de las células B en la muestra.

La detección de la hibridación de ARN que codifica las proteínas IRTA indicará que una neoplasia maligna es una neoplasia maligna de las células B. Más específicamente, la detección de hibridación del ARN que codifica la proteína IRTA1 indica que la neoplasia maligna de células B es un linfoma de células B de Tejido Linfoide Asociado a la Mucosa (MALT). La detección de hibridación de ARN que codifica las proteínas IRTA4 e IRTA5 indica que la neoplasia maligna de las células B es un linfoma de células del manto. En una realización del método descrito anteriormente, la neoplasia maligna de células B comprende una reordenación cromosómica 1q21. Un experto utilizará el método descrito anteriormente como ayuda para el diagnóstico junto con otros métodos convencionales de detección/diagnóstico de neoplasias malignas, p. ej., patología de una muestra de tumor, que puede indicar linfoma y a continuación el método descrito anteriormente, restringirá la neoplasia maligna a un linfoma de células B o más específicamente a un linfoma de células B MALT) o un linfoma de células del manto como se ha discutido anteriormente.

Un experto en la técnica está familiarizado con los métodos conocidos de detección de hibridación de moléculas de ácido nucleico con oligonucleótidos de ácido nucleico, es decir, sondas de ácido nucleico que codifican una proteína de interés para métodos de diagnóstico. Las moléculas de ácido nucleico que codifican las proteínas IRTA de la presente invención son útiles para detectar neoplasias malignas de las células B. Un experto en la técnica reconocerá que las variaciones del método descrito anteriormente para la detección de una neoplasia maligna de las células B en una muestra incluyen, pero no se limitan a, digestión del ácido nucleico de la muestra con enzimas de restricción y separación de los fragmentos de la molécula de ácido nucleico así obtenidos por medio de fraccionamiento por tamaños antes de la hibridación.

En una realización del método descrito anteriormente para la detección de una neoplasia maligna de células B en una muestra de un sujeto, el marcador detectable es un isótopo radiactivo, una enzima, un colorante, biotina, una marca fluorescente o una marca quimioluminiscente. En una realización preferida, la neoplasia maligna de células B se selecciona del grupo que consiste en linfoma de células B, mieloma múltiple, linfoma de Burkitt, linfoma de la zona marginal, linfoma difuso de células grandes y linfoma de las células foliculares. En una realización adicional, el linfoma de células B es linfoma de células B del Tejido Linfoide Asociado a la Mucosa (MALT). En otra realización preferida, el linfoma de células B es linfoma no Hodgkin.

Esta invención proporciona un oligonucleótido antisentido que tiene una secuencia capaz de hibridar específicamente con una molécula de ARNm que codifica una proteína IRTA1 humana con el fin de evitar la expresión en exceso de la molécula de ARNm.

En una realización preferida del oligonucleótido antisentido, la proteína IRTA es IRTA1 humana. En una realización adicional de cualquiera de los oligonucleótidos de moléculas de ácido nucleico descritos anteriormente que codifican la proteína IRTA1, el ácido nucleico puede ser ADN genómico o ADNc.

Un experto en la técnica está familiarizado con los mecanismos convencionales para la hibridación de ácido nucleico de oligonucleótidos, por ejemplo, Ausubel, F.M. et al. *Current Protocols in Molecular Biology*, (John Wiley & Sons, Nueva York, 1998), por ejemplo condiciones restrictivas de 65°C en presencia de una concentración de sal elevada. Tales condiciones se utilizan para la hibridación de ácido nucleico completamente complementario, mientras que las condiciones que no son restrictivas se utilizan para la hibridación de ácidos nucleicos que no son totalmente complementarios.

Según se utiliza en la presente memoria, la frase "hibridar específicamente" representa la capacidad de una molécula de ácido nucleico para reconocer una secuencia de ácido nucleico complementaria a ella misma y para formar segmentos de doble hélice mediante enlaces de hidrógeno entre pares de bases complementarias. Según se utiliza en la presente memoria, una "secuencia única" es una secuencia específica sólo para las moléculas de ácido nucleico que codifican la proteína IRTA1. La tecnología de sondas de ácido nucleico es bien conocida para los expertos en la técnica que apreciarán fácilmente que tales sondas pueden variar mucho de longitud y pueden estar marcadas con una marca detectable, tal como un radioisótopo o colorante fluorescente, para facilitar la detección de la sonda. La detección de moléculas de ácido nucleico que codifican la proteína IRTA1 es útil como prueba de diagnóstico para cualquier proceso de enfermedad en el que los niveles de expresión de la proteína IRTA1 correspondiente están alterados. Las moléculas de sonda de ADN se producen mediante la inserción de una molécula de ADN que codifica la proteína IRTA1 de mamífero o sus fragmentos en vectores adecuados, tales como plásmidos o bacteriófagos, seguido de la inserción en las células anfitrionas bacterianas adecuadas y la replicación y la recolección de las sondas de ADN, todo ello utilizando bien los métodos conocidos en la técnica. Por ejemplo, el ADN puede ser extraído de un producto lisado celular utilizando fenol y etanol, digerido con las enzimas de restricción correspondientes a los sitios de inserción del ADN en el vector (comentado en la presente memoria), sometido a electroforesis, y separado por corte del gel resultante. Las sondas de oligonucleótidos son útiles para la hibridación 'in situ' o con el fin de localizar tejidos que expresan esta familia de genes IRTA, y para otros análisis de hibridación para determinar la presencia de estos genes (moléculas de ácido nucleico que codifican proteínas IRTA1) o su ARNm en diversos tejidos biológicos. Adicionalmente, los oligonucleótidos sintetizados (producidos por un sintetizador de ADN) complementarios a la secuencia de una molécula de ADN que codifica una proteína IRTA1 son útiles como sondas para estos genes, para su ARNm asociado, o para el aislamiento de genes relacionados por escrutinio de homología de bibliotecas genómicas o de ADNc, o mediante el uso de técnicas de amplificación tales como la Reacción en Cadena de la Polimerasa.

Esta invención proporciona una proteína IRTA1 purificada que comprende la secuencia de aminoácidos expuesta en la Figura 18A (SEQ ID NO: 1). En una realización de la proteína IRTA1 purificada, la proteína IRTA1 es IRTA1 humana.

5 Con el fin de facilitar la comprensión de la sección de Detalles Experimentales siguiente, ciertos métodos y/o términos que aparecen con frecuencia se describen mejor en Sambrook, et al. (1989) y Harlow y Lane Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratories, Cold Spring Harbor, Nueva York: 1988.

Esta invención proporciona uno o varios anticuerpos dirigidos a un epítipo de proteína IRTA1 purificada, o uno o varios fragmentos de los mismos, que tienen la secuencia de aminoácidos expuesta en la Figura 18A.

10 Según se utiliza en la presente memoria, el término "anticuerpo" incluye, pero no se limita a, anticuerpos tanto de origen natural como de origen no natural. Específicamente, el término "anticuerpo" incluye anticuerpos policlonales y monoclonales, y fragmentos de unión de los mismos. Además, el término "anticuerpo" incluye anticuerpos quiméricos y anticuerpos completamente sintéticos, y fragmentos de los mismos. Los anticuerpos policlonales y monoclonales pueden ser "purificados", lo que significa que los anticuerpos policlonales y monoclonales están libres de cualquier otro anticuerpo. Según se utiliza en la presente memoria, anticuerpo parcialmente purificado significa
15 una composición de anticuerpos que comprende anticuerpos que se unen específicamente a la proteína IRTA de la invención sujeto, y consta de menos impurezas de proteína que el suero del cual derivan los anticuerpos. Una impureza de proteína es una proteína distinta de los anticuerpos específicos para la proteína IRTA de la presente invención. Por ejemplo, los anticuerpos parcialmente purificados pueden ser una preparación de IgG.

20 Los anticuerpos policlonales (anticuerpos anti-IRTA) se pueden producir mediante la inyección a un animal anfitrión tal como un conejo, rata, cabra, ratón u otro animal del inmunógeno o los inmunógenos de esta invención, por ejemplo, una IRTA1 humana purificada descrita más abajo. Los sueros se extraen del animal anfitrión y se escrutan para obtener anticuerpos policlonales que son específicos para el inmunógeno. Los métodos de escrutinio de anticuerpos policlonales son bien conocidos para los expertos normales en la técnica tales como los descritos en Harlow y Lane Antibodies: A Laboratory Manual, (Cold Spring Harbor Laboratories, Cold Spring Harbor, Nueva York:
25 1988)

Los anticuerpos monoclonales anti-IRTA de la presente invención pueden ser producidos mediante la inmunización, por ejemplo, de ratones con un inmunógeno (los polipéptidos IRTA o fragmentos de los mismos descritos en la presente memoria). A los ratones se les inocula por vía intraperitoneal una cantidad inmunogénica del inmunógeno descrito anteriormente y a continuación se refuerza con cantidades similares de inmunógeno. Se recogen los bazo
30 de los ratones inmunizados unos pocos días después del refuerzo final y se prepara una suspensión de células a partir de los bazo para su uso en la fusión.

Los hibridomas se pueden preparar a partir de los esplenocitos y un compañero tumoral murino utilizando la técnica de hibridación de células somáticas general de Kohler, B. y Milstein, C., Nature (1975) 256: 495-497. Se pueden utilizar en la hibridación las líneas de mieloma murino disponibles, tales como las de la Colección Americana de Cultivos Tipo (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, USA. Básicamente, la técnica implica la fusión de las células tumorales y los esplenocitos utilizando un fusógeno tal como polietilenglicol. Después de la fusión las células se separan del medio de fusión y se cultivan en un medio de crecimiento selectivo, tal como medio HAT, para eliminar células parentales no hibridadas. Los hibridomas se pueden expandir, si se desea, y los sobrenadantes se pueden analizar por procedimientos de inmunoanálisis convencionales, por ejemplo
40 radioinmunoanálisis, utilizando el agente inmunizante como antígeno. Los clones positivos se pueden caracterizar adicionalmente para determinar si satisfacen los criterios de los anticuerpos de la invención.

Los hibridomas que producen tales anticuerpos se pueden hacer crecer in vitro o in vivo utilizando procedimientos conocidos. Los anticuerpos monoclonales se pueden aislar de los medios de cultivo o de los fluidos corporales, según corresponda, por medio de procedimientos de purificación de inmunoglobulinas convencionales tales como
45 precipitación con sulfato de amonio, electroforesis en gel, diálisis, cromatografía y ultrafiltración, si se desea.

En la práctica de la presente invención cualquiera de los anticuerpos descritos anteriormente se puede marcar con un marcador detectable. En una realización, el anticuerpo marcado es un anticuerpo marcado purificado. El término "anticuerpo" incluye, a modo de ejemplo, anticuerpos tanto naturales como no naturales. Específicamente, el término "anticuerpo" incluye anticuerpos policlonales y monoclonales, y fragmentos de los mismos. Además, el término
50 "anticuerpo" incluye anticuerpos quiméricos y anticuerpos completamente sintéticos, y fragmentos de los mismos. Los "radicales detectables" que funcionan como marcas detectables son bien conocidos para los expertos en la técnica e incluyen, pero no se limitan a, una marca fluorescente, un átomo radiactivo, un ion paramagnético, biotina, una marca quimioluminiscente o una marca que se puede detectar a través de una etapa enzimática secundaria o de unión. La etapa enzimática secundaria o de unión puede comprender el uso de digoxigenina, fosfatasa alcalina,
55 peroxidasa de rábano picante, β -galactosidasa, fluoresceína o estreptavidina/biotina. Los métodos de marcaje de anticuerpos son bien conocidos en la técnica.

Los métodos de recuperación de suero de un sujeto son bien conocidos por los expertos en la técnica. Los métodos de purificación parcial de anticuerpos también son bien conocidos por los expertos en la técnica, e incluyen, a modo

de ejemplo, filtración, cromatografía de intercambio iónico, y precipitación.

Los anticuerpos policlonales y monoclonales de la invención se pueden marcar con un marcador detectable. En una realización, el anticuerpo marcado es un anticuerpo marcado purificado. El marcador detectable puede ser, por ejemplo, un marcador radiactivo o fluorescente. Los métodos de marcaje de anticuerpos son bien conocidos en la técnica.

La determinación de si los anticuerpos policlonales y monoclonales de la presente invención se unen a las células, por ejemplo, células cancerosas, que expresan una proteína IRTA y forman un complejo con una o más de las proteínas IRTA descritas en la presente memoria, o fragmentos de las mismas, en la superficie de dichas células, se puede llevar a cabo de acuerdo con métodos bien conocidos por los expertos en la técnica. En la realización preferida, la determinación se lleva a cabo de acuerdo con métodos de citometría de flujo.

Los anticuerpos de la presente invención pueden estar unidos a una matriz insoluble tal como la utilizada en la cromatografía de afinidad. Las células que forman un complejo, es decir, se unen, con el anticuerpo policlonal o monoclonal inmovilizado se pueden aislar por medio de métodos convencionales bien conocidos por los expertos en la técnica. Por ejemplo, el aislamiento puede comprender cromatografía de afinidad utilizando el anticuerpo inmovilizado.

Alternativamente, el anticuerpo puede ser un anticuerpo libre. En este caso, el aislamiento puede comprender la clasificación de células utilizando anticuerpos primarios o secundarios marcados. Tales métodos de clasificación de células son convencionales y son bien conocidos para los expertos en la técnica.

Esta invención proporciona un anticuerpo dirigido a una proteína IRTA1 purificada.

En una realización preferida del anticuerpo anti-IRTA, la proteína IRTA es una proteína IRTA humana. La proteína IRTA puede ser cualquier proteína IRTA de mamífero, incluyendo una proteína IRTA murina. En una realización adicional de cualquiera de los anticuerpos descritos anteriormente, el anticuerpo es un anticuerpo monoclonal. En otra realización, el anticuerpo monoclonal es un anticuerpo monoclonal murino o un anticuerpo monoclonal humanizado. Según se utiliza en la presente memoria, "humanizado" representa un anticuerpo que tiene características de un anticuerpo humano, siendo dicho anticuerpo de origen no natural, pero creado utilizando técnicas de hibridoma en las que el anticuerpo es de origen humano, excepto para la porción determinante antigénica, que es murina. En otra realización más, el anticuerpo es un anticuerpo policlonal.

En realizaciones preferidas, cualquiera de los anticuerpos de la presente invención se puede conjugar con un agente terapéutico. En otras realizaciones preferidas, el agente terapéutico es un radioisótopo, una toxina, un toxoide o un agente quimioterapéutico. Los anticuerpos conjugados de la presente invención se pueden administrar a un sujeto que tiene un cáncer de células B mediante cualquiera de los métodos proporcionados a continuación.

Esta invención proporciona una composición farmacéutica que comprende una cantidad del anticuerpo dirigido a una proteína IRTA1 eficaz para unirse a las células cancerosas que expresan una proteína IRTA1 con el fin de prevenir el crecimiento de las células cancerosas y un vehículo farmacéuticamente aceptable. El anticuerpo anti-IRTA puede estar dirigido a un epítipo de una proteína IRTA1. La proteína puede ser una proteína IRTA humana o IRTA de ratón.

En realizaciones preferidas de la composición farmacéutica descrita anteriormente, las células cancerosas se seleccionan del grupo que consiste en células de linfoma de células B, de mieloma múltiple, de linfoma de células del manto, de linfoma de Burkitt, de linfoma de la zona marginal, de linfoma difuso de células grandes y de linfoma folicular. En otra realización preferida de la composición farmacéutica, las células de linfoma de células B son células de linfoma de células B de Tejido Linfoide Asociado a la Mucosa (MALT). En otra realización preferida de la composición farmacéutica, las células de linfoma de células B son células de linfoma no Hodgkin.

Se describe una composición farmacéutica que comprende una cantidad de cualquiera de los oligonucleótidos descritos anteriormente eficaz para prevenir la expresión en exceso de una proteína IRTA humana y un portador farmacéuticamente aceptable idóneo. En realizaciones preferidas de la composición farmacéutica el oligonucleótido es una molécula de ácido nucleico que codifica una proteína IRTA1.

La proteína IRTA puede ser una proteína IRTA humana o de ratón.

Según se utiliza en la presente memoria, "neoplasia maligna" significa susceptible de metastatizar. Según se utiliza en la presente memoria, "células tumorales" son las células que se originan a partir de un tumor, es decir, a partir de un nuevo crecimiento de tejido diferente o anormal. Las células tumorales y las células cancerosas pueden existir como parte de la masa tumoral, o pueden existir como células que flotan libremente, desprendidas de la masa tumoral a partir de la cual se originan.

Según se utiliza en la presente memoria, las células malignas incluyen, pero no se limitan en modo alguno a, linfoma de células B, mieloma múltiple, linfoma de Burkitt, linfoma de células del manto, linfoma de la zona marginal, linfoma difuso de células grandes y linfoma folicular. El linfoma de células B es linfoma de células B de Tejido Linfoide

Asociado a la Mucosa (MALT) o es linfoma no Hodgkin.

Según se utiliza en la presente memoria, "sujeto" es cualquier animal o animal modificado artificialmente. Los animales modificados artificialmente incluyen, pero no se limitan a, ratones SCID con sistemas inmunitarios humanos. En una realización preferida, el sujeto es un ser humano.

5 Esta invención proporciona un método de diagnóstico de neoplasia maligna de células B que comprende una reordenación cromosómica 1q21 en una muestra de un sujeto que comprende:

10 a) poner en contacto la muestra obtenida del sujeto con un anticuerpo dirigido a una proteína IRTA1 purificada capaz de unirse específicamente con una proteína IRTA1 humana sobre una superficie celular de una célula cancerosa en condiciones que permiten la unión del anticuerpo con la proteína IRTA1 humana sobre la superficie celular de la célula cancerosa, en donde el anticuerpo está marcado con un marcador detectable; y b) detectar cualquier unión en la etapa (a), en donde la detección de unión indica un diagnóstico de neoplasia maligna de células B en la muestra.

15 En una realización del método de diagnóstico descrito anteriormente de neoplasia maligna de células B, la proteína IRTA es IRTA1. En otra realización del método, la proteína IRTA es la proteína IRTA humana o de ratón. En una realización adicional la proteína IRTA es purificada. En una realización preferida de este método, la neoplasia maligna de células B se selecciona entre el grupo que consiste en linfoma de células B, mieloma múltiple, linfoma de Burkitt, linfoma de la zona marginal, linfoma difuso de células grandes y linfoma folicular. En otra realización más de este método, el linfoma de células B es linfoma de células B de Tejido Linfoide Asociado a la Mucosa (MALT). En otra realización preferida de este método, el linfoma de células B es linfoma no Hodgkin.

20 Se describe un método para detectar la proteína IRTA humana en una muestra que comprende: a) poner en contacto la muestra con cualquiera de los anticuerpos anti-IRTA descritos anteriormente bajo condiciones que permitan la formación de un complejo entre el anticuerpo y la IRTA en la muestra; y b) detectar el complejo formado en la etapa (a), detectando de ese modo la presencia de IRTA humana en la muestra. En una realización la proteína IRTA detectada puede ser una proteína IRTA1, que tiene una secuencia de aminoácidos expuesta en la Figura 18A.
25 Como se ha descrito anteriormente en este documento, la detección del complejo formado se puede conseguir mediante el uso de anticuerpos marcados con un marcador detectable y la determinación de la presencia del complejo marcado. La detección de la proteína IRTA humana en una muestra de un sujeto es otro método de diagnóstico de neoplasia maligna de células B en un sujeto. En una realización de este método de diagnóstico, la neoplasia maligna de células B se selecciona del grupo que consiste en linfoma de células B, mieloma múltiple,
30 linfoma de Burkitt, linfoma de la zona marginal, linfoma difuso de células grandes y linfoma folicular. En otra realización adicional de este método, el linfoma de células B es linfoma de células B de Tejido Linfoide Asociado a la Mucosa (MALT). En otra realización preferida de este método, el linfoma de células B es linfoma no Hodgkin.

35 Esta invención proporciona la utilización de un tratamiento de un sujeto que tiene un cáncer de células B, que comprende administrar al sujeto una cantidad de anticuerpo anti-IRTA eficaz para unirse a las células cancerosas que expresan una proteína IRTA con el fin de evitar el crecimiento de las células cancerosas y un portador farmacéuticamente aceptable, tratando de este modo al sujeto. El crecimiento y la proliferación de las células cancerosas es inhibido de ese modo y las células cancerosas mueren. En una realización de la utilización anteriormente descrita, la proteína IRTA es IRTA1 humana.

40 En una realización preferida de la utilización descrita anteriormente de tratamiento de un sujeto que tiene un cáncer de células B, el anticuerpo anti-IRTA es un anticuerpo monoclonal. En otra realización del método, el anticuerpo monoclonal es un anticuerpo monoclonal murino o un anticuerpo monoclonal humanizado. El anticuerpo puede ser un anticuerpo quimérico. En una realización adicional, el anticuerpo anti-IRTA es un anticuerpo policlonal. En una realización, el anticuerpo policlonal puede ser un anticuerpo policlonal murino o humano. En una realización preferida, el cáncer de células B se selecciona del grupo que consiste en linfoma de células B, mieloma múltiple,
45 linfoma de Burkitt, linfoma de células del manto, linfoma de la zona marginal, linfoma difuso de células grandes y linfoma folicular. En otra realización preferida, el linfoma de células B es linfoma de células B de Tejido Linfoide Asociado a la Mucosa (MALT). En una realización preferida adicional, el linfoma de células B es linfoma no Hodgkin. En una realización preferida del método de tratamiento de un sujeto que tiene un cáncer de células B descrito anteriormente, la administración de la cantidad de anticuerpo anti-IRTA eficaz para unirse a las células cancerosas que expresan una proteína IRTA es mediante liberación intravenosa, intraperitoneal, intratecal, intralinfática,
50 intramuscular, intralesional, parenteral, epidural, subcutánea; por infusión, por medio de liberación mediada por liposomas, liberación en aerosol; tópica, oral, nasal, anal, ocular u ótica. En otra realización preferida de los métodos descritos anteriormente, el anticuerpo anti-IRTA se puede conjugar con un agente terapéutico. En otras realizaciones preferidas, el agente terapéutico es un radioisótopo, una toxina, un toxoide o un agente quimioterapéutico.

55 Se describe una utilización para tratar a un sujeto que tiene un cáncer de células B, que comprende administrar al sujeto una cantidad de un oligonucleótido antisentido que tiene una secuencia capaz de hibridar específicamente con una molécula de ARNm que codifica una proteína IRTA humana con el fin de evitar la expresión en exceso de la proteína IRTA humana, con el fin de detener el crecimiento celular o de inducir la muerte celular de las células cancerosas que expresan la proteína o las proteínas IRTA y un portador farmacéuticamente aceptable, tratando de

este modo al sujeto.

En una realización de la utilización descrita anteriormente para el tratamiento de un sujeto que tiene un cáncer de células B, la proteína IRTA es la proteína IRTA1 humana. En una realización preferida, el cáncer de células B se selecciona del grupo que consiste en linfoma de células B, mieloma múltiple, linfoma de Burkitt, linfoma de la zona marginal, linfoma difuso de células grandes y linfoma folicular. En otra realización preferida, el linfoma de células B es linfoma de células B de Tejido Linfoide Asociado a la Mucosa (MALT). En otra realización preferida más, el linfoma de células B es linfoma no Hodgkin. En realizaciones de cualquiera de los oligonucleótidos descritos anteriormente de moléculas de ácido nucleico que codifican la proteína IRTA1, el ácido nucleico puede ser ADN genómico o ADNc. En una realización adicional preferida del método de tratamiento descrito anteriormente de un sujeto que tiene un cáncer de células B, la administración de la cantidad de oligonucleótido eficaz para prevenir la expresión en exceso de una proteína IRTA humana mediante liberación es intravenosa, intraperitoneal, intratecal, intralinfática, intramuscular, intralesional, parenteral, epidural, subcutánea; por infusión, mediante liberación mediada por liposomas, liberación en aerosol; tópica, oral, nasal, anal, ocular u ótica. En una realización de los métodos descritos anteriormente, el oligonucleótido se puede conjugar con un agente terapéutico. En otras realizaciones preferidas, el agente terapéutico es un radioisótopo, una toxina, un toxoide o un agente quimioterapéutico.

La invención también proporciona una composición farmacéutica que comprende una cantidad eficaz de los oligonucleótidos o de los anticuerpos descritos anteriormente y un portador farmacéuticamente aceptable. En la invención sujeto una "cantidad eficaz" es cualquier cantidad de un oligonucleótido o un anticuerpo que, cuando se administra a un sujeto que padece una enfermedad o anomalía contra las que el oligonucleótido o anticuerpo son eficaces, causa una reducción, remisión o regresión de la enfermedad o anomalía. En la práctica de esta invención, el "portador farmacéuticamente aceptable" es cualquier portador fisiológico conocido para los expertos normales en la técnica útil en la formulación de composiciones farmacéuticas.

Los portadores farmacéuticamente aceptables son bien conocidos para los expertos en la técnica e incluyen, pero no se limitan a tampón fosfato 0,01-0,1 M y preferiblemente 0,05 M o solución salina al 0,8%. Adicionalmente, tales portadores farmacéuticamente aceptables pueden ser soluciones, suspensiones, y emulsiones acuosas o no acuosas. Los ejemplos de disolventes no acuosos son propilenglicol, polietilenglicol, aceites vegetales tales como aceite de oliva, y ésteres orgánicos inyectables tales como oleato de etilo. Los portadores acuosos incluyen agua, soluciones, emulsiones o suspensiones alcohólicas/acuosas, incluyendo solución salina y medios tamponados. Los portadores parenterales incluyen solución de cloruro sódico, dextrosa de Ringer, dextrosa y cloruro sódico, solución de Ringer con lactato añadido o aceites fijados. Los vehículos intravenosos incluyen fluidos y reponedores de nutrientes, reponedores de electrolitos tales como los basados en dextrosa de Ringer, y similares. También pueden estar presentes otros aditivos, tales como, por ejemplo, antimicrobianos, antioxidantes, agentes quelantes, gases inertes y similares.

En una realización preferida, el portador farmacéutico puede ser un líquido y la composición farmacéutica estaría en forma de una solución. En otra realización igualmente preferida, el portador farmacéuticamente aceptable es un sólido y la composición está en la forma de un polvo o un comprimido. En una realización adicional, el portador farmacéutico es un gel y la composición está en forma de un supositorio o una crema. En una realización adicional, el compuesto se puede formular como parte de un parche transdérmico farmacéuticamente aceptable.

Un portador sólido puede incluir una o más sustancias que también pueden actuar como agentes aromatizantes, lubricantes, solubilizantes, agentes de suspensión, cargas, antiapelmazantes, coadyuvantes de compresión, aglutinantes o agentes disgregantes de comprimidos; también puede ser un material encapsulante. En los polvos, el portador es un sólido finamente dividido que está mezclado con el ingrediente activo finamente dividido. En los comprimidos, el ingrediente activo se mezcla con un portador que tiene las propiedades de compresión necesarias en proporciones adecuadas y se compacta en la forma y el tamaño deseados. Los polvos y los comprimidos contienen preferiblemente hasta 99% del ingrediente activo. Los portadores sólidos adecuados incluyen, por ejemplo, fosfato de calcio, estearato de magnesio, talco, azúcares, lactosa, dextrina, almidón, gelatina, celulosa, polivinilpirrolidina, ceras de bajo punto de fusión y resinas de intercambio iónico.

Los portadores líquidos se utilizan en la preparación de soluciones, suspensiones, emulsiones, jarabes, elixires y composiciones presurizadas. El ingrediente activo se puede disolver o suspender en un portador líquido farmacéuticamente aceptable tal como agua, un disolvente orgánico, una mezcla de ambos o aceites o grasas farmacéuticamente aceptables. El portador líquido puede contener otros aditivos farmacéuticos adecuados tales como solubilizantes, emulsionantes, tampones, conservantes, edulcorantes, agentes aromatizantes, agentes de suspensión, agentes espesantes, colorantes, reguladores de la viscosidad, estabilizantes u osmorreguladores. Los ejemplos adecuados de portadores líquidos para administración oral y parenteral incluyen agua (que contiene parcialmente aditivos como los anteriores, p. ej., derivados de celulosa, preferentemente solución de carboximetilcelulosa sódica), alcoholes (incluyendo alcoholes monohidroxilados y alcoholes polihidroxilados, p. ej. glicoles) y sus derivados, y aceites (p. ej., aceite de coco fraccionado y aceite de cacahuete). Para la administración parenteral, el portador también puede ser un éster oleoso tal como oleato de etilo y miristato de isopropilo. Los portadores líquidos estériles son útiles en composiciones estériles en forma líquida para la administración parenteral. El portador líquido para composiciones presurizadas puede ser un hidrocarburo halogenado u otro propelente farmacéuticamente aceptable.

Las composiciones farmacéuticas líquidas que son soluciones o suspensiones estériles se pueden utilizar, por ejemplo, mediante inyección intramuscular, intratecal, epidural, intraperitoneal o subcutánea. Las soluciones estériles también se pueden administrar por vía intravenosa. Los compuestos se pueden preparar como una composición sólida estéril que se puede disolver o suspender en el momento de la administración utilizando agua estéril, solución salina, u otro medio inyectable estéril apropiado. Se pretende que los portadores incluyan los aglutinantes, agentes de suspensión, lubricantes, aromatizantes, edulcorantes, conservantes, colorantes, y recubrimientos necesarios e inertes.

La composición farmacéutica que comprende el oligonucleótido o el anticuerpo se puede administrar oralmente en forma de una solución o suspensión estéril que contiene otros solutos o agentes de suspensión, por ejemplo, suficiente solución salina o glucosa para hacer la solución isotónica, sales biliares, acacia, gelatina, monooleato de sorbitán, polisorbato 80 (ésteres oleato de sorbitol y sus anhídridos copolimerizados con óxido de etileno) y similares.

La composición farmacéutica que comprende el oligonucleótido o el anticuerpo también se puede administrar por vía oral en forma de una composición líquida o sólida. Las composiciones adecuadas para administración oral incluyen formas sólidas, tales como píldoras, cápsulas, gránulos, comprimidos y polvos, y formas líquidas, tales como soluciones, jarabes, elixires, y suspensiones. Las formas útiles para la administración parenteral incluyen soluciones, emulsiones, y suspensiones estériles.

Las dosificaciones óptimas que se van a administrar pueden ser determinadas por los expertos en la técnica, y variarán con el inhibidor concreto en uso, la fuerza de la preparación, el modo de administración, y el progreso del estado de enfermedad o anomalía. Otros factores que dependen del sujeto concreto a tratar darán como resultado la necesidad de ajustar las dosificaciones, incluyendo la edad del sujeto, el peso, el sexo, la dieta y el tiempo de administración.

Esta invención se entenderá mejor a partir de los Detalles Experimentales siguientes. Sin embargo, un experto en la técnica apreciará fácilmente que los métodos y resultados específicos comentados son meramente ilustrativos de la invención como se describe más completamente en las reivindicaciones siguientes.

Detalles experimentales

Primera serie de experimentos

El análisis molecular de las translocaciones cromosómicas asociadas con el mieloma múltiple (MM) ha indicado que la patogénesis de esta neoplasia maligna puede ser heterogénea, estando asociada con varios oncogenes distintos incluyendo BCL-1, MUM-1 y FGFR3. Las anomalías estructurales del cromosoma 1q21, incluyendo las translocaciones con el cromosoma 14q32, representan aberraciones citogenéticas frecuentes asociadas con el mieloma múltiple. Con el fin de identificar los genes implicados en estas translocaciones, se clonaron las regiones del punto de rotura correspondientes a ambos derivados de una t(1;14)(q21;q32) detectable en la línea celular de plasmacitoma humano FR4. El análisis de las secuencias del punto de rotura mostró que implicaban una recombinación recíproca entre el locus de la cadena pesada de inmunoglobulina (IgH) en 14q32 y secuencias desconocidas en 1q21. El locus normal correspondiente a la región 1q21 implicada en la translocación se clonó y se identificaron los genes adyacentes a la región del punto de rotura por medio de una estrategia de captura del exón. Se encontraron dos genes, localizados a una distancia de 20 Kb entre sí, en la región que abarcaba el punto de rotura en 1q21. El primer gen, denominado MUM-2 (mieloma múltiple 2) se expresa como un transcrito de ARNm de 2,5 Kb detectable en el bazo y los ganglios linfáticos. La clonación y secuenciación del ADNc de MUM-2 completo pronostica una glicoproteína de la superficie celular de 515 aminoácidos que contiene cuatro dominios de tipo Ig extracelulares, uno transmembrana y un dominio citoplasmático y que comparten 37% de identidad (51% de homología) con el receptor I gamma de Fc sobre sus tres primeros dominios extracelulares. En las células FR4, los puntos de rotura de la translocación interrumpen el dominio codificante de MUM-2 y lo yuxtaponen al locus IgH en la misma orientación transcripcional. Como consecuencia, los transcritos MUM-2 específicos de FR4 estructuralmente anormales (3,0, 5,2 y 6,0 Kb) en los ganglios linfáticos y el bazo codifican una proteína con un dominio extracelular que contiene seis dominios de tipo Ig homólogos a los miembros de las familias de receptores Fc gamma y de adherencia de tipo Ig. La estructura de los genes MUM-2 y MUM-3 y su participación directa en una translocación asociada a MM sugieren que estos genes codifican los receptores de la superficie celular novedosos importantes para la función de los linfocitos normales y las neoplasias malignas de las células B.

Segunda serie de experimentos

Procedimientos experimentales

Líneas celulares

Las líneas celulares de MM utilizadas en este estudio (FR4, U266, JJN3, EJM, SKMM1, RPMI-8226, XG1, XG2, XG4, XG6, XG7) han sido referidas previamente (Tagawa et al., 1990), (Jernberg et al., 1987), (Hamilton et al., 1990; Jackson et al., 1989), (Eton et al., 1989), (Zhang et al., 1994). La línea celular FR4 se estableció en el laboratorio de uno de los autores (S.T.). Las líneas celulares U266, JJN3 y EJM fueron donaciones del Dr. K. Nilsson (Universidad

de Uppsala, Uppsala, Suecia) y la línea celular SKMM-1 fue una donación A.N. Houghton (Memorial Sloan Kettering Cancer Center, Nueva York, NY). Las cinco líneas de células XG se obtuvieron del Dr. Bernard Klein y se cultivaron en presencia de 1 ng/ml de IL-6 humana recombinante como se ha descrito previamente (Zhang et al., 1994). Las líneas celulares BL con anomalías 1q21 se han descrito previamente (Polito et al., 1995), (Magrath et al., 1980) y se cultivaron en RPMI, FCS al 10%.

Selección de la biblioteca genómica y de ADNc y análisis de la secuencia de ADN

Se construyeron dos bibliotecas genómicas a partir de ADN genómico de FR4 por medio de digestión completa con BamHI o por medio de digestión parcial con Sau3AI y posterior ligación de las fracciones purificadas en gel en el vector de fago 1DASH-II (Stratagene). La biblioteca de BamHI se escrutó con una sonda XhoI-BamHI de 4,2 kb derivada del locus de Ca y la biblioteca Sau3AI se escrutó con una sonda 5'Sa descrita previamente (Bergsagel et al., 1996). Se escrutó una biblioteca de ADN de placenta humana (Stratagene) con una sonda 1.0EH (Figuras 8A-8C) para obtener el locus 1q21 de la línea germinal. El escrutinio de la biblioteca y el aislamiento de la placa se llevaron a cabo de acuerdo con los procedimientos establecidos (Sambrook et al., 1989). Los clones de ADNc de *IRTA1* e *IRTA2* se aislaron a partir de una biblioteca de ADNc oligo-dT/cebada al azar construida a partir de ARN de bazo humano normal (Clontech). La sonda de ADNc de *IRTA1* utilizada para el escrutinio de la biblioteca se obtuvo a partir de la RT-PCR de ADNc de bazo humano utilizando los cebadores que flanquean los exones 1 y 3. La secuenciación del ADN se llevó a cabo en un secuenciador automatizado ABI 373 (Applied Biosystems). Las búsquedas de homología de secuencia se llevaron a cabo a través del servidor de correo electrónico BLAST en el Centro Nacional de Información Biotecnológica, Bethesda, MD.

Aislamiento de PAC y YAC y captura de exones

Los clones de PAC humanos se obtuvieron mediante escrutinio de una biblioteca PAC humana aplicada sobre membranas de nailon (Research Genetics), con la sonda de 1,0 EH (Figuras 8A-8C). Se escrutó la biblioteca de YAC humanos Zeneca (antes ICI) (Anand et al., 1990) obtenida del United Kingdom Human Genome Mapping Resource center (UK-HGMP) utilizando una estrategia de agrupamiento basada en la PCR. La captura de exones se realizó utilizando el sistema de captura de exones (Gibco BRL), de acuerdo con las instrucciones del fabricante.

Aislamiento de clones finales de PAC/YAC, electroforesis en gel de campo pulsado (PFGE) y análisis de hibridación con fluorescencia *in situ* (FISH)

La extracción de ADN de PAC se realizó de acuerdo con los métodos de lisis alcalina convencionales (Drakopoli N et al., 1996). Se utilizó un método de PCR vectorette para aislar sondas terminales de PAC y YAC (Riley et al., 1990), como se ha descrito previamente (Iida et al., 1996). El análisis de PFGE se realizó de acuerdo con protocolos convencionales (Drakopoli N et al., 1996) utilizando el sistema CHEF Mapper (BioRad, Hercules, CA). El marcaje con biotina del ADN de PAC, la preparación de cromosomas y FISH se realizaron como se ha descrito previamente (Rao et al., 1993).

Análisis de transferencia Southern y Northern, RACE y RT-PCR

Los análisis de transferencia Southern y Northern se realizaron como se ha descrito previamente (Neri et al., 1991). Para los análisis de transferencia Northern se preparó el ARN total por el método del tiocianato de guanidinio y se seleccionó el ARN poli(A) utilizando cuentas recubiertas con poli(T) (Kit Oligotex de Quigen). Para las transferencias Northern, se cargaron 2 mg de ARN poli(A) por calle. Se obtuvieron filtros Northern de múltiples tejidos de Clontech. La RACE se realizó utilizando el kit de Amplificación de ADNc Marathon (Clontech) y ADNc de bazo Marathon-Ready. La síntesis de la primera hebra de ADNc se realizó utilizando el sistema SuperScript RT-PCR (Gibco BRL).

Hibridación *in situ*

Se transcribieron sondas de ARNc antisentido y efectoras que contenían digoxigenina con ARN polimerasa T3 y T7, respectivamente, a partir de plásmidos pBluescript KS+ linealizados que contenían la región codificante de los ADNc (nucleótidos 62-1681 de *IRTA1* y 18 a 2996 de *IRTA2*). Se congeló instantáneamente tejido tonsilar humano hiperplásico extirpado quirúrgicamente de niños del Babies Hospital, Columbia Presbyterian Medical Center en hielo seco en polvo. Las secciones del criostato se almacenaron durante varios días a -80 grados C antes del procesamiento. La hibridación no radiactiva *in situ* se realizó esencialmente como se ha descrito (Frank et al., 1999), excepto que el tiempo de fijación en paraformaldehído al 4% se aumentó a 20 minutos, y se omitió el tratamiento con proteinasa K. La rigurosidad de la hibridación fue de 68 grados C, en 5X SSC, formamida al 50%. La tinción de los anticuerpos anti-digoxigenina conjugados con fosfatasa alcalina se desarrolló con el sustrato BCIP/NBT.

Transfección, inmunoprecipitación y transferencia Western. Se transfectaron de manera transitoria células 293 (ATCC), cultivadas en DMEM, FCS al 10%, de acuerdo con el método de fosfato de calcio convencional, con constructos de expresión transitoria pMT2T y pMT2T-*IRTA1*/Ca. Este último fue generado utilizando el producto de la RT-PCR de *IRTA1*/Ca de FR4. Las células (2×10^6 de transfectantes y 2×10^7 de líneas celulares restantes) se solubilizaron en tampón de lisis Triton X-100 (NaCl 150 mM, Tris-HCl 10 mM [pH 7,4], TX-100 al 1%, BSA al 0,1%) en presencia de un cóctel de inhibidores de proteasa (Roche Biochemicals). Los productos lisados se incubaron a 4°C durante 2 horas con 4 mg/ml del anticuerpo monoclonal Núm. 117-332-1 (Yu et al., 1990) (Tanox Biosystems,

Inc, Houston, Texas) que se originaron contra la porción extracelular del péptido de membrana IgA. Los complejos inmunitarios se aislaron con proteína G-Sepharose (Pharmacia) antes de la electroforesis sobre geles en gradiente al 10-20% de Tris-HCl (BioRad) y de la inmunotransferencia, utilizando 15 mg/ml del anticuerpo Núm. 117-332-1. Los resultados se visualizaron por medio de ECL (Amersham).

5 Resultados

Clonación Molecular de la t(1;14)(q21;q32)

Las translocaciones cromosómicas que involucran al locus de la cadena pesada de Ig (IGH) se producen a menudo dentro o cerca de regiones de cambio de IgH como resultado de eventos de recombinación de cambio "ilegítimos" (Dalla-Favera et al., 1983; Chesi et al., 1996; Chesi et al., 1998). Los puntos de rotura se pueden detectar mediante análisis de hibridación por transferencia Southern como alelos reordenados en los que las secuencias de la región constante (C_H) de IGH han perdido su asociación sinténica con las secuencias de la región de unión (J_H) de IGH y de cambio 5' (S) (Dalla-Favera et al., 1983; Neri et al., 1988; Neri et al., 1991; Bergsagel et al., 1996). Este análisis ha llevado a la identificación de varios compañeros cromosómicos para el locus IgH en B-NHL y MM (Taub et al., 1982; Dalla-Favera et al., 1983; Neri et al., 1988; Neri et al., 1991; Ye et al., 1993; Chesi et al., 1996; Richelda et al., 1997; lida et al., 1997; Dyomin et al., 1997; Dyomin et al., 2000). Los autores de la presente invención emplearon la misma estrategia con el fin de clonar la región del punto de rotura de 1q21 en FR4, una línea celular de mieloma que porta una t(1;14)(q21;q32), según lo determinado por análisis citogenético (Tagawa et al., 1990; Taniwaki M, resultados no publicados). Dos "fragmentos ilegítimamente reordenados se identificaron en el locus de la cadena pesada $C\alpha$ en FR4 mediante análisis de transferencia Southern (datos no mostrados), y se clonaron a partir de bibliotecas de fagos construidas a partir de ADN genómico de FR4. El mapeo de restricción, la hibridación por transferencia Southern y la secuenciación de nucleótidos parcial de dos fagos genómicos (clones lambda FR4B-5 y λ FR4S-a, Figura 8A) demostraron que contenían los puntos de rotura cromosómicos de una translocación equilibrada recíproca entre el locus $C\alpha_1$ en 14q32 y secuencias distintas de IGH. A continuación se utilizó una sonda (1.0EH) que representaba estas secuencias distintas de IgH (Figura 8A) para clonar el correspondiente locus genómico normal a partir bibliotecas genómicas humanas del cromosoma artificial del fago P1 (PAC), y el cromosoma artificial de levadura (YAC). El análisis de hibridación fluorescente in situ (FISH) de las propagaciones de la metafase humana normal utilizando el clon PAC no quimérico de 100 kb 49A16 que abarca la región del punto de rotura (véase más abajo, Figura 13), identificó el locus cromosómico compañero como derivado de la banda 1q21 (Figura 8C). El mapeo de un único locus dentro del cromosoma 1 se confirmó por hibridación de dos sondas no repetitivas a ADN de un panel de híbridos de células somáticas representativo de cromosomas humanos individuales (datos no mostrados). Estos resultados fueron compatibles con la clonación de secuencias que abarcan la t(1;14)(q21;q32) en FR4.

El análisis de secuencia de las regiones de punto de rotura en los cromosomas derivados y el alineamiento con los loci 14q32 y 1q21 de la línea germinal revelaron que el punto de rotura se había producido en el intrón entre $CH3$ y el exón transmembrana de $C\alpha_1$ en el cromosoma 14. Aunque la región del punto de rotura estaba desprovista de secuencias señal de recombinación (RSS) o secuencias de señal de cambio (Kuppers et al., 1999), la secuencia CTTAAC (subrayada en la Figura 8B) estaba presente en ambos cromosomas 14 y 1 de la línea germinal en la unión del punto de rotura. Una copia de esta secuencia estaba presente en cada uno de los cromosomas derivados, con una ligera modificación en la copia der(1) (mutación puntual en el último nucleótido: C a G). Los nucleótidos AT que preceden a CTTAAC en el cromosoma 1 también estaban presentes en ambos cromosomas derivados (Figura 8B). La translocación no dio lugar a ninguna pérdida de secuencias en el cromosoma 1. Por otro lado, en la porción del cromosoma 14 de der(1) los autores de la presente invención observaron dos deleciones aguas arriba de la unión del punto de rotura: una delección de 16 nucleótidos (GGCACCTCCCCTTAAC) y una delección de 4 nucleótidos (TGCA) 6 nucleótidos aguas arriba (Figura 8B). Estas observaciones indican que la t(1;14)(q21;q32) en las células FR4 representa una translocación recíproca equilibrada posiblemente facilitada por la presencia de secuencias homólogas. (CTTAAC) en ambos cromosomas.

La región del punto de rotura de 1q21 contiene genes que codifican nuevos miembros de la superfamilia de receptores de inmunoglobulina

Los autores de la presente invención investigaron a continuación si la región del cromosoma 1q21 que abarca el punto de rotura de la translocación en FR4 contiene una unidad transcripcional. El ADN de los clones de PAC parcialmente solapantes 49A16 y 210K22 (Figura 13) fueron clonados por medio de "pistola génica" en plásmidos, secuenciados y analizados para determinar la homología con genes conocidos en bases de datos del genoma humano. Paralelamente, se buscaron los genes candidato en el PAC 49A16 por medio de una estrategia de captura de exones (Church et al., 1994).

El mapeo de los exones candidato sobre los clones genómicos de 1q21 reveló que el punto de rotura de FR4 se había producido entre dos exones atrapados (véase más abajo, Figura 13), que pertenecían al mismo transcrito, ya que podrían estar conectados por RT-PCR utilizando ARN de bazo. A continuación, se utilizó este producto de RT-PCR como sonda para escrutar una biblioteca de ADNc de bazo con el fin de aislar clones completos correspondientes a este transcrito. Se identificaron dos conjuntos de clones de ADNc, que pertenecían a dos transcritos distintos y que compartían una identidad de secuencia de ARNm de 76% dentro de la región de la sonda de 443 pb. Los clones de ADNc completos para ambos transcritos se obtuvieron mediante amplificación rápida de

extremos de ADNc (RACE) en ADNc de bazo humano que generó productos de extensión 5' y 3'.

La estructura esquemática del ADNc que representa el primer transcrito se representa en la Figura 9A. El uso alternativo de los tres posibles sitios de poliadenilación en su región no traducida 3' da lugar a tres especies de ARNm de 2,6, 2,7 y 3,5 kb, que codifican la misma supuesta proteína de 515-aminoácidos (Figura 9A). Las características pronosticadas de esta proteína incluyen un péptido señal, de acuerdo con la regla [-3, -1] (von Heijne, 1986), cuatro dominios extracelulares de tipo Ig que llevan tres sitios de glicosilación potenciales unidos a (N) de asparragina (Figura 9A), uno transmembrana de 16 aminoácidos y un dominio citoplasmático de 106 aminoácidos, con tres supuestos dominios de unión de homología 2 a Src (SH2) consenso (Unkeless y Jin, 1997) (Figura 10B). Estos dominios de unión a (SH2) muestran características de motivos tanto ITAM (Motivo Inmunorreceptor basado en la Activación de Tirosina D/EX₇D/EX₂YXXL/IX₆₋₈YXXL/I; donde X indica los residuos no conservados) (Reth, 1989) como motivos ITIM (Motivo de Inhibición de Inmunorreceptor basado en Tirosina - S/V/L/IYXXL/V, donde X indica residuos no conservados) (Unkeless y Jin, 1997). Como se muestra en la Figura 10B, los primeros dos dominios de unión a SH2 están espaciados por 8 aminoácidos, en consonancia con el motivo ITAM consenso. A diferencia del consenso, el residuo glutamato (E) se posiciona cuatro en lugar de dos aminoácidos antes de la primera tirosina (Y) (Figura 10B), y la posición +3 con respecto a la tirosina (Y) está ocupada por valina (V) en lugar de leucina (L) o isoleucina (I) (Cambier, 1995). Los tres dominios se ajustan al consenso ITIM y cada uno está codificado por un exón separado, como es el caso de ITIM. Así, su disposición puede dar lugar a tres ITIM o posiblemente a un ITAM y un ITIM. La estructura general de esta proteína sugiere que representa un nuevo receptor transmembrana de la superfamilia de Ig y por lo tanto se denominó IRTA1 (gen 1 Asociado a la Translocación del Receptor Inmunitario 1).

El segundo ADNc comparte homología con *IRTA1* (68% de identidad de nucleótidos para la longitud del mensaje de IRTA1 que codifica su dominio extracelular) y fue denominado *IRTA2*. El locus *IRTA2* es más complejo que *IRTA1* y se transcribe en tres isoformas de ARNm principales (*IRTA2a*, *IRTA2b*, *IRTA2c*) de diferente peso molecular (2,8, 4,7 y 5,4 kb respectivamente), cada uno con una propia única región no traducida 3' (Figura 9B). Adicionalmente, un transcrito de 0,6 kb (Figura 12A) surge del uso de una señal de poliadenilación temprana en el nucleótido 536 de *IRTA2*. Las tres isoformas de la proteína *IRTA2* pronosticadas codificadas por estos transcritos comparten una secuencia de aminoácidos común hasta el residuo 560, que presenta un péptido señal común y seis dominios de tipo Ig extracelulares (Figura 9B). *IRTA2a* codifica una glicoproteína secretada de 759 aa con ocho dominios de tipo Ig seguido de 13 aminoácidos predominantemente polares únicos en su extremo C. *IRTA2b* diverge de *IRTA2a* en el residuo de aminoácido 560, y se extiende a lo largo de un corto tramo de 32 residuos adicionales, cuyo carácter hidrófobo es compatible con su acoplamiento a la membrana plasmática a través de un ancla de GPI (Ferguson y Williams, 1988). *IRTA2c* es la isoforma más larga cuya secuencia se desvía de *IRTA2a* en el aminoácido 746. Codifica una glicoproteína transmembrana de tipo I de 977 aa con nueve dominios de tipo Ig extracelulares, que alberga ocho sitios de glicosilación unidos a N potenciales, uno transmembrana de 23 aminoácidos y un dominio citoplasmático de 104 aminoácidos con tres motivos de unión a SH2 consenso (Figura 10B). Cada uno de los sitios de unión a SH2 en *IRTA2c* coincide con el consenso ITIM (Figura 10B) y es codificado por un exón separado. Estos rasgos sugieren que *IRTA2c* es un receptor transmembrana novedoso de la superfamilia de Ig con isoformas secretadas y unidas a GPI.

Homología entre las proteínas IRTA y los receptores de la superfamilia de inmunoglobulinas

El alineamiento de aminoácidos de la totalidad de los dominios extracelulares de las proteínas *IRTA1* e *IRTA2* entre sí y con otros miembros de la superfamilia de Ig reveló una homología notable entre ellos (identidad de 47% y similitud de 51%) y una homología inferior, pero llamativa con la familia de proteínas receptoras de Fc gamma. Esta homología era más fuerte en las posiciones de aminoácidos conservados entre las diferentes clases de receptores de Fc. Entre los receptores de Fc, el receptor de IgG de la alta afinidad FCGRI (CD64) compartió los más altos niveles de homología con los tres primeros dominios Ig de *IRTA1* e *IRTA2* (identidad de 37% y similitud de 50%) a lo largo de toda su porción extracelular (Figura 10A). Se observaron niveles inferiores de homología entre las proteínas *IRTA* y los dominios extracelulares de otras moléculas de la superficie celular, incluyendo la molécula de adherencia endotelial de plaquetas humanas (PECAM1), la molécula de adherencia celular de linfocitos B (CD22) y la glicoproteína biliar 1 (BGP1) (22-25% de identidad, 38-41% de homología).

No existe homología aparente entre las *IRTA* y los miembros de la familia de receptores de Fc en sus dominios citoplásmicos. En contraste, existe una homología de aminoácidos significativa entre *IRTA1* y PECAM1 (identidad de aminoácidos de 31% y homología de 45%), *IRTA2c* y BGP1 (identidad de 30%, homología de 35%) e *IRTA2c* y PECAM1 (28% de identidad, 50% de homología) (Figura 10B). Estas homologías sugieren el empleo de rutas de señalización aguas abajo similares por parte de estas proteínas diferentes.

IRTA1 e *IRTA2* normalmente se expresan en subpoblaciones específicas de células B

El patrón de expresión normal de los ARNm de *IRTA1* e *IRTA2* se analizó primero por hibridación de transferencia de Northern del ARN derivado de diferentes tejidos humanos normales y de las líneas celulares humanas que representan diferentes linajes hematopoyéticos y etapas de desarrollo de células B.

Se detectó la expresión de *IRTA1* a un nivel muy bajo en ARN de bazo y ganglios linfáticos humanos (Figura 11A, panel izquierdo) y fue indetectable en todos los otros tejidos humanos analizados, incluyendo el hígado fetal, médula ósea, pulmón, placenta, intestino delgado, riñón, hígado, colon, músculo esquelético, corazón y cerebro (datos no mostrados). Entre las líneas de células B, la expresión de *IRTA1* estaba ausente en líneas de células que representan células pre-B y células B del centro germinal, células plasmáticas y células de origen eritroide, células T y mieloide (datos no mostrados, véase Materiales y Métodos). La expresión fue detectable a niveles muy bajos solamente en líneas celulares linfoblastoides (LCL) inmortalizadas con EBV, que representan una subpoblación (inmunoblastos) situada aguas abajo de las células B del centro germinal en la diferenciación de células B. Sin embargo, la expresión se indujo en células ER/EB privadas de estrógeno que, al estar inmortalizadas por un genoma de EBV recombinante en el que el gen EBNA2 se fusiona con el receptor de estrógeno, proliferan en presencia de estrógeno, mientras se detienen en la fase G₀/G₁ tras la privación de estrógeno (Kempkes et al., 1995). La expresión de *IRTA1* era apenas detectable en estas células en presencia de estrógeno, pero fue inducida (10 veces) después de la detención G₀/G₁ tras la retirada de estrógenos (Figura 11A, panel derecho). Tomados en conjunto, estos resultados sugieren que *IRTA1* se expresa en una subpoblación linfóide presente en el bazo y los ganglios linfáticos y, presumiblemente, representada por células B en reposo.

Para investigar más el fenotipo y la distribución en los tejidos de las células que expresan *IRTA1*, los autores de la presente invención realizaron la hibridación *in situ* en tejido tonsilar humano utilizando una sonda de ADNc antisentido de *IRTA1* (Figura 11B). Se procesaron las secciones seriadas para la hibridación *in situ* con una sonda de ADNc efectora de control (Panel Núm. 1 en la figura 11B), una sonda de ADNc antisentido (Panel Núm. 2) y tinción con hematoxilina y eosina (H & E) (Panel Núm. 3) para esbozar la arquitectura del tejido linfóide. La señal de hibridación de *IRTA1* fue excluida del centro germinal y la zona de manto de los folículos y se concentró característicamente en la zona perifolicular con infiltraciones en la región intraepitelial (Figuras 11B-2, 11B-4). En esta región, sólo las células B fueron positivas tal como se documenta mediante la tinción con marcadores específicos de células B (IGD, no mostrado), y mediante análisis inmunohistoquímico con anticuerpos anti-*IRTA1* y anti-B (CD20, PAX5), anti-T (CD3), y anti-monocitos (CD68) (no mostrados; G. Cattoretti et al., manuscrito en preparación.). Esta área perifolicular es la "zona marginal" equivalente a la amígdala, lo que representa un compartimiento de células B funcionalmente distinta que contiene la mayoría de las células B de memoria y las células B monocitoides (de Wolf-Peters et al., 1997). Junto con el análisis de transferencia Northern de tejidos normales y líneas celulares, estos resultados indican que *IRTA1* se expresa en una subpoblación de células B maduras en reposo topográficamente ubicada en la región perifolicular e intraepitelial, sitios ricos en células B de memoria.

En el caso de *IRTA2*, el análisis de transferencia Northern detectó todas las especies empalmadas alternativamente en ARNm de ganglio linfático, bazo, médula ósea e intestino delgado humanos, con una preponderancia relativa de la isoforma *IRTA2a* (Figura 12A, panel izquierdo). Entre las líneas de células hematopoyéticas de origen linfóide y no linfóide sometidas a ensayo, la expresión de *IRTA2* se restringió a las líneas de células B con un fenotipo post-centro germinal inmunoblástico (Figura 12A, panel derecho). De manera similar a *IRTA1*, estaba ausente de las líneas celulares derivadas de células pre-B, centroblastos del centro germinal, células plasmáticas, células T, células eritroides y células mieloides (Figura 12A, panel derecho).

El análisis de hibridación *in situ* de tejido tonsilar humano, utilizando el ADNc de *IRTA2* como sonda, fue compatible con los resultados del análisis de transferencia Northern. El ARNm de *IRTA2* fue excluido en gran parte de la zona del manto del centro germinal, con la excepción de unas pocas células positivas (Figuras 12B-2, 12B4). Dentro del centro germinal, la zona oscura, representada por centroblastos, apareció negativa para *IRTA2*, mientras que la zona clara, rica en centrocitos, fue fuertemente positiva (Figuras 12B-2, 12B-4). Finalmente, el ARNm de *IRTA2* se detectó en la región equivalente de la "zona marginal" fuera de los folículos de los centros germinales y en las regiones intraepiteliales e interfoliculares de la amígdala. Este patrón es compatible con la especificidad de *IRTA2* para los centrocitos y las células post-B del centro germinal. Comparando sus patrones de expresión, los autores de la presente invención concluyen que ambos son específicos para las células B maduras, pero *IRTA2* tiene un patrón más amplio de expresión que incluye centrocitos y células B interfoliculares, mientras *IRTA1* está restringida a las células B de la zona marginal, muy probablemente células de memoria.

Organización genómica de los genes *IRTA1* e *IRTA2*

Para entender las consecuencias de las anomalías 1q21 sobre la estructura y la expresión de los genes *IRTA1* e *IRTA2*, los autores de la presente invención determinaron primero la organización de sus loci genómicos. El gen *IRTA1* contiene 11 exones con un tamaño genómico total de 24,5 kb (Figura 13). Se encontró que el locus *IRTA2* abarcaba una región genómica de aproximadamente 40 kb (Figura 13). Los tres productos empalmados alternativamente de *IRTA2* comparten sus primeros 8 exones, en cuyo punto *IRTA2b* no utiliza el siguiente sitio de empalme, y termina introduciendo su región UTR 3'. Las isoformas *IRTA2a* y 2c se empalman en el exón 9, entrando *IRTA2a* en su UTR 3' después del exón 11 y empalmando *IRTA2c* en el exón 12 y extendiéndose hasta el exón 18 (Figura 13).

Basándose en los datos de secuenciación, los autores de la presente invención determinaron que los genes *IRTA1* e *IRTA2* se encuentran a 21 kb el uno del otro, yuxtapuestos en la misma orientación transcripcional (Figura 13) que

se extiende desde el telómero (5') hacia el centrómero (3'). En el locus 1q21, están estrechamente conectados el uno al otro, así como a tres genes adicionales que habían clonado recientemente los autores de la presente invención a través de su homología con las IRTA (I.M., manuscrito en preparación). Los cinco genes son contiguos, cubriendo una región de ~ 300 kb en 1q21. Esta región se encuentra en el intervalo entre los puntos de rotura de 1q21 referidos anteriormente. Basándose en la distancia entre clones genómicos que albergan los genes respectivos en el mapa Whitehead Institute Radiation Hybrid, se estima que el locus de *IRTA1-2* se encuentra aproximadamente a 0,8 Mb de distancia del locus *MUC1* hacia el telómero (N.P., datos no publicados; Dyomin et al., 2000; Gilles et al., 2000) y a menos de o igual a 7 Mb del locus *FCGR1B* hacia el centrómero (N.P., datos no publicados).

La translocación t(1;14)(q21;q32) genera una proteína de fusión IRTA1/Ca₁ en la línea celular de mieloma FR4

El análisis comparativo de restricción y de la secuencia de nucleótidos de la línea germinal frente a las secuencias reordenadas de los loci Ca₁ e *IRTA1* mostró que la translocación había fusionado las secuencias en el intrón 2 del gen *IRTA1* a las secuencias intrónicas entre *CH3* y el exón transmembrana de Ca₁ en la misma orientación transcripcional (Figura 14A). Esto sugiere que, si las secuencias de *IRTA1* se expresaban en el locus translocado, el sitio donante intacto en el límite 3' del exón IRTA1 y el sitio aceptor intacto en el 5' de Ca₁ podrían ser utilizados para generar un ARNm de fusión *IRTA1/Ca₁*, y posiblemente una proteína de fusión IRTA1/Ca₁.

Con el fin de someter a ensayo esta predicción, los autores de la presente invención analizaron la expresión del ARNm de *IRTA1* en FR4 por medio de análisis de transferencia Northern utilizando una sonda de ADNc de *IRTA1* derivada del exón 1 (Figura 14A). Esta sonda detectaba un mensaje de 0,8 kb en FR4 que estaba ausente de otras líneas de células B, y era más corto que el mensaje normal de 2,5 kb detectable en células ER/EB (Figura 14B). Los autores de la presente invención clonaron este transcrito por medio de RT-PCR del ARNm de FR4 utilizando cebadores derivados de secuencias en el límite 5' del exón 1 de *IRTA1* y el límite 3' del exón citoplásmico Ca (Figura 14A). Se obtuvo un producto de RT-PCR de FR4, pero no de la línea celular DAKIKI que expresa la IgA de la superficie de tipo salvaje, u otras líneas celulares que carecen de una translocación t(1;14) (datos no mostrados). El análisis de secuenciación directa del producto de PCR indicó que el empalme había ligado de manera precisa IRTA1 y Ca₁ en los sitios de empalme canónicos y determinó que el transcrito de fusión tenía 820 pb de longitud.

El análisis del producto de proteína pronosticado indicó que el empalme *IRTA1/Ca₁* había dado lugar a una fusión entre el péptido señal IRTA1 y los dos primeros aminoácidos extracelulares, con el espaciador largo extracelular de 32 aminoácidos, el dominio transmembrana y la cola citoplásmica del receptor de IgA₁ de membrana (mIgA₁) (Figura 14C). Para el análisis de la expresión de esta proteína de fusión en extractos de proteína FR4, los autores de la presente invención utilizaron un anticuerpo dirigido contra residuos de aminoácidos extracelulares específicos para la isoforma transmembrana de Ca₁ (Yu et al., 1990) para la inmunoprecipitación, seguido de transferencia Western. Los resultados de los autores de la presente invención demuestran que las células FR4, pero no una línea celular de control (DAKIKI) que expresa IgA de la superficie de tipo salvaje, expresan una proteína de 9,8 kDa acorde con el tamaño pronosticado de la proteína de fusión IRTA1/Ca₁ (Figura 14D). Estos resultados muestran que el alelo translocado codifica una proteína de fusión, compuesta por el péptido señal y los dos primeros residuos extracelulares de IRTA1 (17 aminoácidos) fusionado a los dominios transmembrana y citoplásmicos codificados por Ca₁ (71 aminoácidos). En contraste con la expresión en exceso de *IRTA1/Ca₁* en der(14), no se detectó expresión en FR4 para el transcrito recíproco *Ca₁/IRTA1* o para el gen *IRTA2* intacto en der(1).

Con la excepción de FR4, no se detectó expresión de ARNm de IRTA1 en ninguna otra línea celular de mieloma o linfoma, independientemente del estatus de su banda cromosómica 1q21 (datos no mostrados). Por lo tanto, la fusión IRTA1/Ca representa un evento raro en las aberraciones de 1q21.

Desregulación frecuente de la expresión de IRTA2 en líneas celulares que portan anomalías de 1q21

Con el fin de establecer la relación física entre otros puntos de rotura de 1q21 y el locus de *IRTA1/2*, los autores de la presente invención realizaron un análisis FISH con el PAC 49A16 en el panel de líneas celulares BL y MM de los autores de la presente invención. De diez líneas celulares BL analizadas, siete con dup(1) (q21q32) y tres con translocaciones 1q21 (AS283A, BL104, BL136), los autores de la presente invención detectaron tres señales correspondientes al locus de *IRTA1/IRTA2* en siete de las primeras y dos de las últimas, en consonancia con dup(1) (q21q32) en el primer caso y dup(1) (q21q32) seguido de un punto de rotura de translocación en 1q21 en el segundo (Tabla 1). El análisis FISH de AS283A y BL136, utilizando sondas que abarcan el locus de *IRTA* y con clones genómicos vecinos, situó el punto de rotura de los cromosomas derivados fuera del locus de IRTA en ambas líneas celulares, a una distancia de >800 kb hacia el centrómero en AS283A y >800 kb hacia el telómero en BL136 (N.P., resultados no publicados). En consonancia con este hallazgo, el análisis de 30 casos de tumores primarios MM por medio de FISH de interfase con el YAC de 300 kb 23GC4 (Figura 13), mostró que 15 casos (50% del total analizado) tenían más de dos señales de FISH en interfase (datos no mostrados), mientras que el FISH de dos colores con dos clones de PAC que flanqueaban las fronteras centroméricas y teloméricas de YAC no detectó ninguna escisión de estas dos sondas en ninguno de los casos. Estos resultados indican que, con la excepción de FR4, los puntos de rotura de las aberraciones de 1q21 en BL o MM no están dentro de o en estrecha proximidad con la región genómica definida por IRTA1 e *IRTA2*. Sin embargo, el resultado coherente de cualquiera dup(1) (q21q32) (véase la Tabla 1) o

dup(1) (q21q32) seguido de translocaciones desequilibradas (AS283A, BL136, XG2, XG7 en la Tabla 1) es la trisomía o tetrasomía parcial de la región de 1q21 que contiene los genes de *IRTA*.

Tabla 1. Resumen de los datos del cariotipo y FISH en el locus *IRTA1/IRTA2*

Tipo de tumor	Citogenética	PAC 49A16	Número de copias del locus <i>IRTA</i> por FISH	Expresión de ARNm de <i>IRTA2</i>
Linfoma de Burkitt				
AS283A	der(4) t(1;4) (q21; q35)	der(4), normal 1	3	+++++
MC116	dup1q21	dup1q21	3	+++
CA46	dup1q21	dup1q21	3	+++
PA682	dup1q21	dup1q21	3	++
BrglgA	dup1q21	dup1q21	3	++
BL32	dup1q21	dup1q21	3	-
BL92	dup1q21	dup1q21	3	++
BL103	invdup1q21	dup1q21	3	+
BL104	t(1; 3) (q21;p25)	der(1)	2	+
BL136	der(1) (qpter1q21 :: q21)	der(1)	3	++
Mieloma múltiple				
XG2	der(1) t(1;?) (q21 ;?)	der(1), normal 1	3	++++
	der 19 t(1;19) (q12;?)	der(19)		
XG7	der(9) t(1; 9) (q12 ;?)	der(9)	4	-
	der(19) t(1;19) (q12;?)	der(19)		
	der(1) t(1 ;?) (q21 ;?) x2	der(1) x2		

- 5 A continuación, los autores de la presente invención investigaron si estas aberraciones tenían un efecto sobre la expresión del ARNm de *IRTA2*. Con este fin, los autores de la presente invención utilizaron una sonda de ADNc correspondiente a la región no traducida 5' de *IRTA2* para escrutar una transferencia Northern con un panel de líneas celulares de B-NHL y MM carentes de o que mostraban anomalías cromosómicas en 1q21. Los resultados muestran que la mayoría (diez de doce) líneas BL con cromosomas 1q21 normales carecen esencialmente de expresión de *IRTA2*, coincidiendo con el hecho de que los BL derivan de centroblastos GC que normalmente carecen de expresión de *IRTA2* (Figura 15A, panel izquierdo). Por el contrario, la mayoría de las líneas BL que portan anomalías en 1q21 (diez de doce) muestran claramente una regulación al alza del ARNm de *IRTA2* (Figura 15A, panel derecho), que oscilan de 2 a 50 veces sobre los niveles basales detectados en BL con 1q21 normal.
- 10

Entre las líneas celulares de mieloma, *IRTA2* era expresada en exceso en una de cada tres líneas que muestran anomalías en 1q21 (XG2), mientras que no se expresaba en ninguna de las siete con 1q21 normal (Figura 15B).

Estos resultados muestran una fuerte correlación entre la presencia de aberraciones cromosómicas 1q21 y la desregulación de la expresión de ARNm de *IRTA2* en BL y sugieren que las trisomías del locus *IRTA2* pueden desregular su expresión en este subtipo de linfoma (véase la Discusión).

Discusión

Los esfuerzos descritos en la presente memoria para identificar genes implicados en las aberraciones cromosómicas que afectan a la banda 1q21 en el Mieloma Múltiple y el linfoma de células B, condujeron al descubrimiento de que *IRTA1* e *IRTA2*, dos miembros fundadores de una nueva subfamilia de receptores relacionados dentro de la familia de inmunorreceptores; las secuencias de ácido nucleico completas que codifican las proteínas *IRTA1* e *IRTA2* son proporcionadas en la presente memoria, así como las secuencias de aminoácidos de las proteínas *IRTA1* e *IRTA2* codificadas. Con posterioridad tres genes adicionales de los miembros de esta subfamilia de receptores relacionados fueron aislados, *IRTA3*, *IRTA4*, e *IRTA5*, cuyas secuencias de ácido nucleico completas se proporcionan en la presente memoria, así como las secuencias de aminoácidos de las proteínas *IRTA3*, *IRTA4*, e *IRTA5* codificadas. Estos resultados tienen implicaciones para la biología normal de las células B, así como para el papel de las aberraciones de 1q21 en la linfomagénesis.

IRTA1 e *IRTA2* son miembros fundadores de una nueva subfamilia dentro de la superfamilia de Ig

Varias características compartidas entre los dos genes de *IRTA* y sus proteínas codificadas sugieren que forman una nueva subfamilia dentro de la superfamilia de inmunorreceptores. En primer lugar, comparten un mayor grado de homología entre sí en sus dominios extracelulares que con otros miembros de la superfamilia tanto en su secuencia de ARNm (identidad de 68%) como de proteína (identidad de 47%). En segundo lugar, comparten homología en sus dominios citoplásmicos, marcados por la presencia de motivos de señalización de tipo ITAM e ITIM en el contexto de las secuencias de aminoácidos homólogas. En tercer lugar, *IRTA1* e *IRTA2* pertenecen a una subfamilia más grande de cinco genes que muestra mayor homología intrafamiliar y una estrecha agrupación dentro de una región de ~ 300 kb en 1q21 (I.M. et al., Manuscrito en preparación). Su organización genómica sugiere que un gen ancestral común puede haber dado lugar a esta subfamilia, por medio de un proceso de duplicación y divergencia de secuencia, similar al mecanismo propuesto para la familia de receptores de Fc (Qiu et al., 1990).

En su dominio extracelular, las proteínas *IRTA* están estrechamente relacionadas con la subfamilia de receptores de Fc basándose en el alto grado de homología de aminoácidos compartida especialmente con el receptor *FCGR1* de alta afinidad (37-45% de identidad de aminoácidos). También sugiere un origen evolutivo común con los receptores de Fc por la posición del locus de la familia *IRTA* en el intervalo entre el locus *FCGR1* en 1q21 y los loci *FCER1* y *FCGR2-III* en 1q21-q23. Por último, los genes *IRTA* y *FCR* comparten una organización exón/intrón similar de la porción del gen que codifica su péptido señal, en particular, los dos exones líder 5' con las secuencias que codifican el sitio de la peptidasa señal ubicado dentro del segundo exón de 21 pb.

Basándose en sus motivos de tipo ITIM citoplasmáticos, las proteínas *IRTA* se pueden considerar miembros de la Superfamilia de Receptores Inhibidores (IRS), un grupo de receptores que bloquean la activación de muchos tipos de células en el sistema inmunitario (Lanier, 1998). Tales miembros incluyen *FCGR1B* y *CD22* en el ser humano (DeLisser et al., 1994) y *PIR-B* en el ratón (Kubagawa et al., 1997). De un modo análogo a los miembros de IRS, los ITIM de *IRTA1* e *IRTA2* están codificados por exones individuales. Una característica que comparten muchos miembros de IRS es la existencia de las correspondientes isoformas activadoras del receptor cuyos dominios citoplásmicos están desprovistos de ITIM (revisado en Ravetch y Lanier, 1998). Es posible que la isoforma secretada de *IRTA2*, que carece de motivos de tipo ITIM, cumpla un papel análogo al contrarrestar el efecto de la isoforma transmembrana.

Las proteínas *IRTA1* e *IRTA2* y los miembros de la subfamilia las Moléculas de Adherencia Celular (CAM) *PECAM1*, *CD22* y *BGP1* comparten una homología significativa en la secuencia y la organización general de su porción extracelular. Además, la capacidad de *IRTA2* para generar tres isoformas de la proteína con una localización subcelular distinta (una proteína transmembrana, una unida a GPI o una secretada) por corte y empalme diferencial es compartida por *NCAM*, otro miembro de la subfamilia de CAM (Dickson et al., 1987; Gower et al., 1988). Por lo tanto, la familia de *IRTA* también se relaciona con la familia de CAM, como se ha sugerido previamente para un miembro de la familia de receptores de Fc (*FCGR2* murino) debido a su homología con *PECAM1* (CAM, familia IRS) (Daeron, 1991; Newman et al., 1990; Stockinger et al., 1990).

En conclusión, la familia de *IRTA* puede representar una intersección entre las familias de Fc, IRS y CAM, que combina las características de las tres. En consecuencia, las proteínas *IRTA* pueden tener un papel en la regulación de la transducción de señales durante una respuesta inmunitaria (como receptores de Fc), en la comunicación intercelular (como miembros de las familias IRS y CAM) y en la migración celular (como miembros de la familia CAM) (DeLisser et al., 1994; Ravetch y Lanier, 2000). Los experimentos iniciales indican que *IRTA1* se puede unir débilmente a la IgA agregada por calor, mientras *IRTA2c* se puede unir específicamente a la IgG de suero humano

agregada por calor (con mayor afinidad para IgG₁ e IgG₂), pero no a IgG, IgA, IgM e IgE humanas monoméricas (datos no mostrado). Estos datos iniciales apoyan una relación funcional entre IRTA y las familias de receptores de Fc, pero no excluyen las funciones que dependen de otros ligandos para las proteínas IRTA.

Patrón de expresión diferencial de los genes de IRTA en células B maduras

- 5 Los genes *IRTA* muestran un patrón de expresión específico en diversos compartimentos de células B normales. *IRTA1* está topográficamente restringida a las células B dentro de la región perifolicular, que se denominó originalmente zona marginal en el bazo, pero también es detectable en la mayor parte de los órganos linfoides (de Wolf-Peelers et al., 1997). Los datos de hibridación *in situ* que aquí se presentan han sido confirmados por análisis
- 10 inmunohistoquímico utilizando anticuerpos anti-*IRTA1* que muestran que la proteína *IRTA1* se expresa selectivamente en células B de la zona marginal, y, entre los NHL, en el linfoma de la zona marginal, los tumores derivados de estas células (G. Cattoretti et al., manuscrito en preparación). Por otra parte, *IRTA2* tiene un patrón de expresión más amplio que incluye centrocitos GC, así como un amplio espectro de células perifoliculares, que pueden incluir inmunoblastos y células de memoria. Los datos iniciales sugieren que el patrón de expresión de
- 15 *IRTA3* es análogo al de *IRTA2*, mientras que *IRTA4* e *IRTA5* se expresan selectivamente en las células B de la zona del manto (I. Miller et al., manuscrito en preparación), el compartimiento pre-GC de las células B maduras (MacLennan, I.C., 1994). Esta restricción topográfica de la expresión génica de *IRTA* en los órganos linfoides sugiere que las moléculas de IRTA pueden desempeñar un papel en la migración o actividad de diversas subpoblaciones de células B en compartimentos específicos de células B funcionales. Además, la expresión de IRTA debería ser útil para el diagnóstico diferencial de los subtipos de NHL que derivan de diversos compartimentos de células B, en particular *IRTA1* en el diagnóstico del linfoma de la zona marginal.
- 20

Locus de *IRTA1* y anomalías de 1q21 en MM

- En la línea celular FR4, la consecuencia de la translocación t(1;14) es la formación de un gen de fusión *IRTA1/Cα₁*. A pesar del hecho de que este gen está dirigido por la región promotora de *IRTA1*, que normalmente es silencioso en las células plasmáticas, su expresión es alta en FR4, presumiblemente debido a la influencia de LCR 3' de *Cα₁*, que es retenido aguas abajo del locus *Cα₁*. El gen de fusión codifica una proteína de fusión *IRTA1/Cα₁* que contiene sólo el péptido señal y los primeros dos aminoácidos de *IRTA1* unidos al receptor de IgA de la superficie. Este último ha sido desprovisto casi completamente de su dominio extracelular, pero conserva todos sus dominios transmembrana e intracelulares. Esta estructura indica que la proteína de fusión *IRTA1/Cα₁*, aunque probablemente incapaz de unirse a cualquier ligando, puede conservar el potencial para la dimerización y la señalización. En particular, la porción extracelular derivada de IgA de membrana (m) contiene un residuo de cisteína, que puede estar implicado en enlaces disulfuro entre dos cadenas α o entre cadenas α y proteínas asociadas, tales como el receptor de la superficie coadyuvante CD19 (Leduc et al., 1997). La proteína de fusión también porta el dominio citoplásmico mIgA de 14 aminoácidos intacto, que está altamente conservado en rotación (Reth, 1992) y puede desempeñar un papel esencial en la proliferación, la supervivencia y la diferenciación de células B maduras, de forma análoga al papel de mIgG y mIgE (Kaisho et al., 1997). Por lo tanto, la aparición de la proteína *IRTA1/Cα₁* en FR4 puede haber proporcionado a las células una ventaja proliferativa y de supervivencia durante el desarrollo del tumor a través la activación independiente del ligando (antígeno) de la ruta de BCR. Este evento de fusión sin embargo, parece ser poco frecuente en las neoplasias malignas de células B, ya que hasta ahora los autores de la presente invención han sido capaces de detectarlo sólo en las células FR4.
- 25
- 30
- 35

40 Locus de *IRTA2* y anomalías de 1q21 en MM y BL

- La expresión anormal de *IRTA2* es una consecuencia frecuente de las anomalías en 1q21. Aunque este gen no se expresa normalmente en centroblastos, las presuntas contrapartes normales de BL (Kuppers et al., 1999), ni en BL con 1q21 normal, sus niveles están regulados al alza un promedio de 10 veces en líneas celulares de BL con anomalías en 1q21. Esta desregulación parece ser específica para *IRTA2* ya que todos los otros 4 genes de *IRTA* presentes dentro de 300 kb en 1q21, o bien no se expresan en BL (*IRTA1*), o bien su patrón de expresión no se corresponde con la presencia de anomalías en 1q21 (*IRTA3*, 4, 5, no mostrado). El mecanismo por el cual se produce este desregulación es difícil de determinar en ausencia de lesiones estructurales dentro o adyacentes al gen de *IRTA2*. Dado que las aberraciones heterogéneas que afectan a 1q21 ocasionan todas un número de copias en exceso del locus de *IRTA*, es posible que esto pueda conducir a trastornos de la regulación, como en el caso de un bajo nivel de amplificación de *BCL2* en FL que carece de translocaciones (14;18) (Monni et al., 1997), *REL* en linfoma difuso de células grandes (Houldsworth et al., 1996; Rao et al., 1998) y desregulación de Ciclina D1 en algunos casos de MM con trisomía 11 (Pruneri et al., 2000). Por otra parte, las anomalías en 1q21, incluyendo translocaciones y duplicaciones, cambian el contexto genómico del locus de *IRTA* y pueden conducir a la desregulación de *IRTA2* por elementos de organización de la cromatina potenciadores que actúan en cis distantes que actúan sobre su promotor como es el caso *MYC* en BL endémico (Pelicci et al., 1986) y MM (Shou et al., 2000) y para *CCND1* en el linfoma de células del manto (Bosch et al., 1994; Swerdlow et al., 1995) y MM (Pruneri et al., 2000).
- 45
- 50
- 55

Las consecuencias biológicas de una expresión desregulada de *IRTA2* son difíciles de predecir en esta etapa. La observación de que *IRTA2* tiene homología con los receptores de adherencia CAM, junto con su distribución específica en la zona clara del GC sugieren que su expresión ectópica en centroblastos puede causar una interrupción en el desarrollo y la arquitectura del GC. Por otro lado, las observaciones iniciales de los autores de la presente invención de que *IRTA2* se puede unir a complejos inmunitarios con IgG de una manera comparable a los receptores de Fc bona fide sugieren que su expresión inapropiada puede perturbar la dinámica de la regulación de la superficie celular de las respuestas inmunológicas de células B, conduciendo posiblemente a la expansión clonal. Se ha propuesto que la expresión desregulada de *FCGR2B* como consecuencia de la t(1;14)(q21;q32) en el linfoma folicular contribuye a la linfomagénesis en este tipo de tumor (Callanan et al., 2000), por medio de un mecanismo que implica la huida por parte de las células tumorales de la vigilancia inmunitaria anti-tumoral a través de su unión a Fc y de la inactivación de la IgG específica del tumor. Se han observado mecanismos de evasión similares en las células infectadas por herpesvirus que codifican Fc (Dubin et al., 1991). Es necesario probar el papel de la desregulación de *IRTA2* en la "ganancia de función" de ratones transgénicos que expresan constitutivamente *IRTA2* en el GC.

Referencias para la segunda serie de experimentos

- Anand, R., Riley, J. H., Butler, R., Smith, J. C., y Markham, A. F. (1990). A 3.5 genome equivalent multi access YAC library: construction, characterisation, screening and storage. *Nucleic Acids Res* 18, 1951-6.
- Avet-Loiseau, H., Andree-Ashley, L. E., Moore, D., 2º, Mellerin, M. P., Feusner, J., Bataille, R., and Pallavicini, M. G. (1997). Molecular cytogenetic abnormalities in multiple myeloma and plasma cell leukemia measured using comparative genomic hybridization. *Genes Chromosomes Cancer* 19, 124-33.
- Bakhshi, A., Jensen, J. P., Goldman, P., Wright, J. J., McBride, O. W., Epstein, A. L., y Korsmeyer, S. J. (1985). Cloning the chromosomal breakpoint of t(14;18) human lymphomas: clustering around JH on chromosome 14 and near a transcriptional unit on 18. *Cell* 41, 899-906.
- Berger, R., Bernheim, A. (1985). Cytogenetics of Burkitt's lymphoma-leukaemia: una revisión. *IARC Sci Publ* 60, 65-80.
- Bergsagel, P. L., Chesi, M., Nardini, E., Brents, L. A., Kirby, S. L., y Kuehl, W. M. (1996). Promiscuous translocations into immunoglobulin heavy chain switch regions in multiple myeloma. *Proc Natl Acad Sci U S A* 93, 13931-6.
- Bosch, F., Jares, P., Campo, E., Lopez-Guillermo, A., Piris, M. A., Villamor, N., Tassies, D., Jaffe, E. S., Montserrat, E., Rozman, C. et al. (1994). PRAD-1/cyclin D1 gene overexpression in chronic lymphoproliferative disorders: a highly specific marker of mantle cell lymphoma. *Blood* 84, 2726-32.
- Callanan, M. B., Le Baccon, P., Mossuz, P., Duley, S., Bastard, C., Hamoudi, R., Dyer, M. J., Klobeck, G., Rimokh, R., Sotto, J. J., y Leroux, D. (2000). The IgG Fc receptor, FcγRIIB, is a target for deregulation by chromosomal translocation in malignant lymphoma. *Proc Natl Acad Sci U S A* 97, 309-14.
- Cambier, J. C. (1995). Antigen and Fc receptor signaling. The awesome power of the immunoreceptor tyrosine-based activation motif (ITAM). *J Immunol* 155, 3281-5.
- Chesi, M., Bergsagel, P. L., Brents, L. A., Smith, C. M., Gerhard, D. S., y Kuehl, W. M. (1996). Dysregulation of cyclin D1 by translocation into an IgH gamma switch region in two multiple myeloma cell lines [see comments]. *Blood* 88, 674-81.
- Chesi, M., Bergsagel, P. L., Shonukan, O. O., Martelli, M. L., Brents, L. A., Chen, T., Schrock, E., Ried, T., and Kuehl, W. M. (1998). Frequent dysregulation of the c-maf proto-oncogene at 16q23 by translocation to an Ig locus in multiple myeloma. *Blood* 91, 4457-63.
- Chesi, M., Nardini, E., Brents, L. A., Schrock, E., Ried, T., Kuehl, W. M., y Bergsagel, P. L. (1997). Frequent translocation t(4;14)(p16.3;q32.3) in multiple myeloma is associated with increased expression and activating mutations of fibroblast growth factor receptor 3. *Nat Genet* 16, 260-4.
- Church, D. M., Stotler, C. J., Rutter, J. L., Murrell, J. R., Trofatter, J. A., y Buckler, A. J. (1994). Isolation of genes from complex sources of mammalian genomic ADN using exon amplification. *Nat Genet* 6, 98-105.
- Cigudosa, J. C., Parsa, N. Z., Louie, D. C., Filippa, D. A., Jhanwar, S. C., Johansson, B., Mitelman, F., y Chaganti, R. S. (1999). Cytogenetic analysis of 363 consecutively ascertained diffuse large B- cell lymphomas. *Genes Chromosomes Cancer* 25, 123-33.
- Daeron, M. (1991). Fc receptors, or the elective affinities of adhesion molecules. *Immunol Lett* 27, 175-81.
- Dalla-Favera, R., Bregni, M., Erikson, J., Patterson, D., Gallo, R.C., y Croce, C.M.: The human c-myc onc-gene is located on the region of chromosome 8 which is translocated in Burkitt lymphoma cells. *Proc. Nat. Acad. Sci. USA* 79:7824-7827, 1982.

- Dalla-Favera, R., Martinotti, S., Gallo, R. C., Erikson, J., y Croce, C. M. (1983). Translocation and rearrangements of the c-myc oncogene locus in human undifferentiated B-cell lymphomas. *Science* 219, 963-7.
- de Wolf-Peeters, C., Pittaluga, S., Dierlamm, J., Wlodarska, I., y Van Den Berghe, H. (1997). Marginal zone B-cell lymphomas including mucosa-associated lymphoid tissue type lymphoma (MALT), monocytoid B-cell lymphoma and splenic marginal zone cell lymphoma and their relation to the reactive marginal zone. *Leuk Lymphoma* 26, 467-78.
- DeLisser, H. M., Newman, P. J., y Albelda, S. M. (1994). Molecular and functional aspects of PECAM-1/CD31. *Immunol Today* 15, 490-5.
- Dickson, G., Gower, H. J., Barton, C. H., Prentice, H. M., Elsom, V. L., Moore, S. E., Cox, R. D., Quinn, C., Putt, W., y Walsh, F. S. (1987). Human muscle neural cell adhesion molecule (N-CAM): identification of a muscle-specific sequence in the extracellular domain. *Cell* 50, 1119-30.
- Dierlamm, J., Pittaluga, S., Wlodarska, I., Stul, M., Thomas, J., Boogaerts, M., Michaux, L., Driessen, A., Mecucci, C., Cassiman, J. J., y et al. (1996). Marginal zone B-cell lymphomas of different sites share similar cytogenetic and morphologic features [see comments]. *Blood* 87, 299-307.
- Dracopoli, C. N., Haines, J. L., Korf, B. R., Morton, C. C., Seidman, C. E., Seidman, J.G., Smith, D. R. (1997). *Current Protocols in Human Genetics* (New York: Wiley & Sons)
- Dubin, G., Socolof, E., Frank, I., Friedman, H. M. (1991). Herpes simplex virus type 1 Fc receptor protects infected cells from antibody-dependent cellular cytotoxicity. *Journal of Virology* 65, 7046-50.
- Dyomin, V. G., Palanisamy, N., Lloyd, K. O., Dyomina, K., Jhanwar, S. C., Houldsworth, J., y Chaganti, R. S. (2000). MUC1 is activated in a B-cell lymphoma by the t(1;14)(q21;q32) translocation and is rearranged and amplified in B-cell lymphoma subsets. *Blood* 95, 2666-71.
- Dyomin, V.G., Rao, P.H., Dalla-Favera, R., Chaganti, R.S.K. (1997). BCL8, a novel gene involved in translocations affecting band 15q11-13 in diffuse large-cell lymphoma. *Proc Natl Acad Sci USA* 94, 5728-32.
- Eton, O., Scheinberg, D. A., and Houghton, A. N. (1989). Establishment and characterization of two human myeloma cell lines secreting kappa light chains. *Leukemia* 3, 729-35.
- Ferguson, M. A., y Williams, A. F. (1988). Cell-surface anchoring of proteins via glycosyl-phosphatidylinositol structures. *Annu Rev Biochem* 57, 285-320.
- Frank, D., Mendelsohn, C. L., Ciccone, E., Svensson, K., Ohlsson, R., y Tycko, B. (1999). A novel pleckstrin homology-related gene family defined by Ipl/Tssc3, TDAG51, and Tih1: tissue-specific expression, chromosomal location, and parental imprinting. *Mamm Genome* 10, 1150-1159.
- Gaidano, G., y Dalla-Favera, R. (1997). Molecular Biology of Lymphomas. In: *Principles and Practice of Oncology*, Fifth Ed, DeVita, VT, Hellman, S., Rosenberg SA (eds) JB Lippincott Co (publ.), 2131-2145.
- Gilles, F., Goy, A., Remache, Y., Shue, P., y Zelenetz, A. D. (2000). MUC1 dysregulation as the consequence of a tt(1;14)(q21;q32) translocation in an extranodal lymphoma. *Blood* 95, 2930-2936.
- Gower, H. J., Barton, C. H., Elsom, V. L., Thompson, J., Moore, S. E., Dickson, G., y Walsh, F. S. (1988). Alternative splicing generates a secreted form of N-CAM in muscle and brain. *Cell* 55, 955-64.
- Hamilton, M. S., Ball, J., Bromidge, E., Lowe, J., y Franklin, I. M. (1990). Characterization of new IgG lambda myeloma plasma cell line (EJM): a further tool in the investigation of the biology of multiple myeloma. *Br J Haematol* 75, 378-84.
- Houldsworth, J., Mathew, S., Rao, P. H., Dyomina, K., Louie, D. C., Parsa, N., Offit, K., Chaganti, R. S. (1996). REL proto-oncogene is frequently amplified in extranodal diffuse large cell lymphoma. *Blood* 87, 25-9.
- Iida, S., Rao, P. H., Butler, M., Corradini, P., Boccadoro, M., Klein, B., Chaganti, R. S., y Dalla-Favera, R. (1997). Deregulation of MUM1/IRF4 by chromosomal translocation in multiple myeloma. *Nat Genet* 17, 226-30.
- Jackson, N., Lowe, J., Ball, J., Bromidge, E., Ling, N. R., Larkins, S., Griffith, M. J., y Franklin, I. M. (1989). Two new IgA1-kappa plasma cell leukaemia cell lines (JJN-1 & JJN-2) which proliferate in response to B cell stimulatory factor 2. *Clin Exp Immunol* 75, 93-9.
- Jernberg, H., Zech, L., y Nilsson, K. (1987). Cytogenetic studies on human myeloma cell lines. *Int J Cancer* 40, 811-7.
- Juliussen, G., Oscier, D. G., Fitchett, M., Ross, F. M., Stockdill, G., Mackie, M. J., Parker, A. C., Castoldi, G. L., Gunee, A., Knuutila, S., y et al. (1990). Prognostic subgroups in B-cell chronic lymphocytic leukemia defined by specific chromosomal abnormalities. *N Engl J Med* 323, 720-4.

- Kaisho, T., Schwenk, F., y Rajewsky, K. (1997). The roles of gamma 1 heavy chain membrane expression and cytoplasmic tail in IgG1 responses. *Science* 276, 412-5.
- 5 Kempkes, B., Spitkovsky, D., Jansen-Durr, P., Ellwart, J. W., Kremmer, E., Delecluss, H. J., Rottenberger, C., Bornkamm, G. W., y Hammerschmidt, W. (1995). B-cell proliferation and induction of early G1-regulating proteins by Epstein-Barr virus mutants conditional for EBNA2. *Embo J* 14, 88-96.
- Kornblau, S. M., Goodacre, A., Cabanillas, F. (1991). Chromosomal abnormalities in adult non-endemic Burkitt's lymphoma and leukemia: 22 new reports and a review of 148 cases from the literature. *Hematol Oncol* 9, 63-78.
- Kubagawa, H., Burrows, P. D., y Cooper, M. D. (1997). A novel pair of immunoglobulin-like receptors expressed by B cells and myeloid cells [see comments]. *Proc Natl Acad Sci U S A* 94, 5261-6.
- 10 Kuppers, R., Klein, U., Hansmann, M. L., y Rajewsky, K. (1999). Cellular origin of human B-cell lymphomas. *N Engl J Med* 341, 1520-9.
- Lanier, L. L. (1998). NK cell receptors. *Annu Rev Immunol* 16, 359-93.
- Leduc, I., Drouet, M., Bodinier, M. C., Helal, A., y Cogne, M. (1997). Membrane isoforms of human immunoglobulins of the A1 and A2 isotypes: structural and functional study. *Immunology* 90, 330-6.
- 15 MacLennan, I. C. (1994). Germinal Centers. *Annu Rev Immunol* 12, 117-39.
- Magrath, I. T., Pizzo, P. A., Whang-Peng, J., Douglass, E. C., Alabaster, O., Gerber, P., Freeman, C. B., y Novikovs, L. (1980). Characterization of lymphoma-derived cell lines: comparison of cell lines positive and negative for Epstein-Barr virus nuclear antigen. I. Physical, cytogenetic, and growth characteristics. *J Natl Cancer Inst* 64, 465-76.
- 20 Monni, O., Joensuu, H., Franssila, K., Klefstrom, J., Alitalo, K., y Knuutila, S. (1997). BCL2 overexpression associated with chromosomal amplification in diffuse large B-cell lymphoma. *Blood* 90, 1168-74.
- Neri, A., Barriga, F., Knowles, D. M., Magrath, I. T., y Dalla-Favera, R. (1988). Different regions of the immunoglobulin heavy-chain locus are involved in chromosomal translocations in distinct pathogenetic forms of Burkitt lymphoma. *Proc Natl Acad Sci U S A* 85, 2748-52.
- 25 Neri, A., Chang, C. C., Lombardi, L., Salina, M., Corradini, P., Maiolo, A. T., Chaganti, R. S., y Dalla-Favera, R. (1991). B cell lymphoma-associated chromosomal translocation involves candidate oncogene *lyt-10*, homologous to NF-kappa B p50. *Cell* 67, 1075-87.
- Newman, P. J., Berndt, M. C., Gorski, J., White, G. C. d., Lyman, S., Paddock, C., y Muller, W. A. (1990). PECAM-1 (CD31) cloning and relation to adhesion molecules of the immunoglobulin gene superfamily. *Science* 247, 1219-22.
- 30 Offit, K., Louie, D. C., Parsa, N. Z., Roy, P., Leung, D., Lo Coco, F., Zelenetz, A., Dalla-Favera, R., Chaganti, R. S. (1995). BCL6 gene rearrangement and other cytogenetic abnormalities in diffuse large cell lymphoma. *Leuk Lymphoma* 20, 85-9.
- Pelicci, P. G., Knowles, D. M. d., Magrath, I., y Dalla-Favera, R. (1986). Chromosomal breakpoints and structural alterations of the *c-myc* locus differ in endemic and sporadic forms of Burkitt lymphoma. *Proc Natl Acad Sci U S A* 83, 2984-8.
- 35 Polito, P., Cilia, A. M., Gloghini, A., Cozzi, M., Perin, T., De Paoli, P., Gaidano, G., y Carbone, A. (1995). High frequency of EBV association with non-random abnormalities of the chromosome region 1q21-25 in AIDS-related Burkitt's lymphoma-derived cell lines. *Int J Cancer* 61, 370-4.
- Pruneri, G., Fabris, S., Baldini, L., Carboni, N., Zagano, S., Colombi, M. A., Ciceri, G., Lombardi, L., Rocchi, M., Buffa, R., Maiolo, A. T., Neri, A. (2000). Immunohistochemical analysis of cyclin D1 shows deregulated expression in multiple myeloma with the t(11;14). *Am J Pathol* 156, 1505-13.
- 40 Qiu, W. Q., de Bruin, D., Brownstein, B. H., Pearce, R., Ravetch, J. V. (1990). Organization of the human and mouse low-affinity Fc gamma R genes: duplication and recombination. *Science* 248, 732-5.
- Rao, P. H., Houldsworth, J., Dyomina, K., Parsa, N. Z., Cigudosa, J. C., Louie, D. C., Popplewell, L., Offit, K., Jhanwar, S. C., y Chaganti, R. S. (1998). Chromosomal and gene amplification in diffuse large B-cell lymphoma. *Blood* 92, 234-40.
- 45 Rao, P. H., Murty, V. V., Gaidano, G., Hauptschein, R., Dalla-Favera, R., y Chaganti, R. S. (1993). Subregional localization of 20 single-copy loci to chromosome 6 by fluorescence in situ hybridization. *Genomics* 16, 426-30.
- Ravetch, J. V., y Lanier, L. L. (2000). Immune inhibitory receptors [In Process Citation]. *Science* 290, 84-9.

- Reth, M. (1989). Antigen receptor tail clue [letter]. *Nature* 338, 383-4.
- Reth, M. (1992). Antigen receptors on B lymphocytes. *Annu Rev Immunol* 10, 97-121.
- Richelda, R., Ronchetti, D., Baldini, L., Cro, L., Viggiano, L., Marzella, R., Rocchi, M., Otsuki, T., Lombardi, L., Maiolo, A. T., Neri, A. (1997). A novel chromosomal translocation t(4; 14)(p16.3; q32) in multiple myeloma involves the fibroblast growth-factor receptor 3 gene [see comments]. *Blood* 90, 4062-70.
- Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J. C., y Markham, A. F. (1990). A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res* 18, 2887-90.
- Ronchetti, D., Finelli, P., Richelda, R., Baldini, L., Rocchi, M., Viggiano, L., Cuneo, A., Bogni, S., Fabris, S., Lombardi, L., Maiolo, A. T., y Neri, A. (1999). Molecular analysis of 11q13 breakpoints in multiple myeloma. *Blood* 93, 1330-7.
- Rosenberg, C. L., Wong, E., Petty, E. M., Bale, A. E., Tsujimoto, Y., Harris, N. L., y Arnold, A. (1991). PRAD1, a candidate BCL1 oncogene: mapping and expression in centrocytic lymphoma. *Proc Natl Acad Sci U S A* 88, 9638-42.
- Sawyer, J. R., Tricot, G., Mattox, S., Jagannath, S., y Barlogie, B. (1998). Jumping translocations of chromosome 1q in multiple myeloma: evidence for a mechanism involving decondensation of pericentromeric heterochromatin. *Blood* 91, 1732-41.
- Sawyer, J. R., Waldron, J. A., Jagannath, S., Barlogie, B. (1995). Cytogenetic findings in 200 patients with multiple myeloma. *Cancer Genet Cytogenet* 82, 41-9.
- Shou, Y., Martelli, M. L., Gabrea, A., Qi, Y., Brents, L. A., Roschke, A., Dewald, G., Kirsch, I. R., Bergsagel, P. L., y Kuehl, W. M. (2000). Diverse karyotypic abnormalities of the c-myc locus associated with c-myc dysregulation and tumor progression in multiple myeloma. *Proc Natl Acad Sci U S A* 97, 228-33.
- Stockinger, H., Gadd, S. J., Eher, R., Majdic, O., Schreiber, W., Kasinrerker, W., Strass, B., Schnabl, E., y Knapp, W. (1990). Molecular characterization and functional analysis of the leukocyte surface protein CD31. *J Immunol* 145, 3889-97.
- Swerdlow, S. H., Yang, W. I., Zukerberg, L. R., Harris, N. L., Arnold, A., Williams, M. E. (1995). Expression of cyclin D1 protein in centrocytic/mantle cell lymphomas with and without rearrangement of the BCL1/cyclin D1 gene. *Hum Pathol* 26, 999-1004.
- Tagawa, S., Doi, S., Taniwaki, M., Abe, T., Kanayama, Y., Nojima, J., Matsubara, K., y Kitani, T. (1990). Amylase-producing plasmacytoma cell lines, AD3 and FR4, with der(14)t(8;14) and dic(8)t(1;8) established from ascites. *Leukemia* 4, 600-5.
- Taub, R., Kirsch, I., Morton, C., Lenoir, G., Swan, D., Tronick, S., Aaronson, S., Leder, P. (1982). Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. *Proc Natl Acad Sci U S A* 79, 7837-41.
- Thompson, J. D., Higgins, D. G., Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22, 4673-80.
- Tusnady, G. E., Simon, I. (1998). Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J Mol Bio* 283, 489-506.
- Unkeless, J. C., y Jin, J. (1997). Inhibitory receptors, ITIM sequences and phosphatases. *Curr Opin Immunol* 9, 338-43.
- von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res* 14, 4683-90.
- Whang-Peng, J., Knutsen, T., Jaffe, E. S., Steinberg, S. M., Raffeld, M., Zhao, W. P., Duffey, P., Condrón, K., Yano, T., Longo, D. L. (1995). Sequential analysis of 43 patients with non-Hodgkin's lymphoma: clinical correlations with cytogenetic, histologic, immunophenotyping, and molecular studies. *Blood* 85, 203-16.
- Willis, T. G., Zalcborg, I. R., Coignet, L. J., Wlodarska, I., Stul, M., Jadayel, D. M., Bastard, C., Treleaven, J. G., Catovsky, D., Silva, M. L., y Dyer, M. J. (1998). Molecular cloning of translocation tt(1;14)(q21;q32) defines a novel gene (BCL9) at chromosome 1q21. *Blood* 91, 1873-81.
- Ye, B. H., Lista, F., Lo Coco, F., Knowles, D. M., Offit, K., Chaganti, R. S., y Dalla-Favera, R. (1993). Alterations of a zinc finger-encoding gene, BCL-6, in diffuse large-cell lymphoma. *Science* 262, 747-50.
- Yu, L. M., Peng, C., Starnes, S. M., Liou, R. S., y Chang, T. W. (1990). Two isoforms of human membrane-bound alpha Ig resulting from alternative mARN splicing in the membrane segment. *J Immunol* 145, 3932-6.

Zhang, X. G., Gaillard, J. P., Robillard, N., Lu, Z. Y., Gu, Z. J., Jourdan, M., Boiron, J. M., Bataille, R., y Klein, B. (1994). Reproducible obtaining of human myeloma cell lines as a model for tumor stem cell study in human multiple myeloma. *Blood* 83, 3654-63.

Tercera Serie de Experimentos

- 5 El cromosoma 1q21 con frecuencia se ve alterado por translocaciones y duplicaciones en varios tipos de neoplasias malignas de células B, incluyendo mieloma múltiple, linfoma de Burkitt, linfomas de la zona marginal, y linfoma folicular. Para identificar los genes implicados en estas aberraciones, se clonó el punto de rotura cromosómica de una t(1;14)(q21;q32) en la línea celular de mieloma FR4. Una región de 300 kb que abarca el punto de rotura
- 10 contiene al menos cinco genes adyacentes altamente relacionados que codifican moléculas receptoras de la superficie que son miembros de la superfamilia de genes de inmunoglobulina, y por lo tanto denominados IRTA (Asociados a la Translocación Receptores de Inmunoglobulina). Las diversas moléculas IRTA tienen de tres a nueve dominios extracelulares de la superfamilia de la inmunoglobulina y están relacionados con los receptores de Fc gamma. Tienen dominios transmembrana y citoplásmicos que contienen motivos de señalización de tipo ITIM y de tipo ITAM (IRTA-1, IRTA-3, IRTA-4). Experimentos de hibridación in situ muestran que todos los genes IRTA se
- 15 expresan en el linaje de células B con distintos patrones específicos de la etapa de desarrollo: IRTA-1 se expresa en un patrón de células B marginales. IRTA-2 se expresa en centrocitos y células B más maduras. Como resultado de la translocación en FR4, IRTA-1 se rompe y produce un transcrito de fusión con el locus de inmunoglobulina. El gen IRTA-2, normalmente silencioso en centroblastos, se expresa en exceso en líneas celulares de mieloma múltiple y linfoma de Burkitt que portan anomalías en 1q21. Aquí, los datos sugieren que los genes de IRTA son moléculas
- 20 reguladoras de las células B novedosas que también pueden tener un papel en la linfomagénesis.

Lista de secuencias

<110> Dalla-Favera, Riccardo

<120> Aislamiento de cinco genes novedosos que codifican nuevos melanomas de tipo receptor de Fc implicados en la patogénesis del linfoma/melanoma

5 <130> 0575/58044-a-PCT-EPO

<140> 00 98 3778.2 (PCT/US00/32403)

<141> 28-11-2000

<160> 44

<170> PatentIn versión 3.1

10 <210> 1

<211> 515

<212> PRT

<213> Homo Sapiens

<400> 1

Met Leu Leu Trp Ala Ser Leu Leu Ala Phe Ala Pro Val Cys Gly Gln
1 5 10 15

Ser Ala Ala Ala His Lys Pro Val Ile Ser Val His Pro Pro Trp Thr
20 25 30

Thr Phe Phe Lys Gly Glu Arg Val Thr Leu Thr Cys Asn Gly Phe Gln
35 40 45

Phe Tyr Ala Thr Glu Lys Thr Thr Trp Tyr His Arg His Tyr Trp Gly
50 55 60

Glu Lys Leu Thr Leu Thr Pro Gly Asn Thr Leu Glu Val Arg Glu Ser
65 70 75 80

Gly Leu Tyr Arg Cys Gln Ala Arg Gly Ser Pro Arg Ser Asn Pro Val
85 90 95

Arg Leu Leu Phe Ser Ser Asp Ser Leu Ile Leu Gln Ala Pro Tyr Ser
100 105 110

15

ES 2 556 332 T3

Val Phe Glu Gly Asp Thr Leu Val Leu Arg Cys His Arg Arg Arg Lys
 115 120 125
 Glu Lys Leu Thr Ala Val Lys Tyr Thr Trp Asn Gly Asn Ile Leu Ser
 130 135 140
 Ile Ser Asn Lys Ser Trp Asp Leu Leu Ile Pro Gln Ala Ser Ser Asn
 145 150 155 160
 Asn Asn Gly Asn Tyr Arg Cys Ile Gly Tyr Gly Asp Glu Asn Asp Val
 165 170 175
 Phe Arg Ser Asn Phe Lys Ile Ile Lys Ile Gln Glu Leu Phe Pro His
 180 185 190
 Pro Glu Leu Lys Ala Thr Asp Ser Gln Pro Thr Glu Gly Asn Ser Val
 195 200 205
 Asn Leu Ser Cys Glu Thr Gln Leu Pro Pro Glu Arg Ser Asp Thr Pro
 210 215 220
 Leu His Phe Asn Phe Phe Arg Asp Gly Glu Val Ile Leu Ser Asp Trp
 225 230 235 240
 Ser Thr Tyr Pro Glu Leu Gln Leu Pro Thr Val Trp Arg Glu Asn Ser
 245 250 255
 Gly Ser Tyr Trp Cys Gly Ala Glu Thr Val Arg Gly Asn Ile His Lys
 260 265 270
 His Ser Pro Ser Leu Gln Ile His Val Gln Arg Ile Pro Val Ser Gly
 275 280 285
 Val Leu Leu Glu Thr Gln Pro Ser Gly Gly Gln Ala Val Glu Gly Glu
 290 295 300
 Met Leu Val Leu Val Cys Ser Val Ala Glu Gly Thr Gly Asp Thr Thr
 305 310 315 320
 Phe Ser Trp His Arg Glu Asp Met Gln Glu Ser Leu Gly Arg Lys Thr
 325 330 335
 Gln Arg Ser Leu Arg Ala Glu Leu Glu Leu Pro Ala Ile Arg Gln Ser
 340 345 350

ES 2 556 332 T3

His Ala Gly Gly Tyr Tyr Cys Thr Ala Asp Asn Ser Tyr Gly Pro Val
355 360 365

Gln Ser Met Val Leu Asn Val Thr Val Arg Glu Thr Pro Gly Asn Arg
370 375 380

Asp Gly Leu Val Ala Ala Gly Ala Thr Gly Gly Leu Leu Ser Ala Leu
385 390 395 400

Ile Leu Ala Val Ala Leu Leu Phe His Cys Trp Arg Arg Arg Lys Ser
405 410 415

Gly Val Gly Phe Leu Gly Asp Glu Thr Arg Leu Pro Pro Ala Pro Gly
420 425 430

Pro Gly Glu Ser Ser His Ser Ile Cys Pro Ala Gln Val Glu Leu Gln
435 440 445

Ser Leu Tyr Val Asp Val His Pro Lys Lys Gly Asp Leu Val Tyr Ser
450 455 460

Glu Ile Gln Thr Thr Gln Leu Gly Glu Glu Glu Glu Ala Asn Thr Ser
465 470 475 480

Arg Thr Leu Leu Glu Asp Lys Asp Val Ser Val Val Tyr Ser Glu Val
485 490 495

Lys Thr Gln His Pro Asp Asn Ser Ala Gly Lys Ile Ser Ser Lys Asp
500 505 510

Glu Glu Ser
515

<210> 2

<211> 2499

<212> DNA

5 <213> Homo Sapiens

<400> 2

ctcaatcagc tttatgcaga gaagaagctt actgagctca ctgctggtgc tgggtgtaggc 60

aagtgtgtgt ttggcaatct gggctgacct ggcttgtctc ctcagaactc cttctccaac 120

cctggagcag gcttccatgc tgctgtgggc gtccttgcctg gcctttgtct cagtctgtgg 180

ES 2 556 332 T3

acaatctgca gctgcacaca aacctgtgat ttccgtccat cctccatgga ccacattctt 240
 caaaggagag agagtgaactc tgacttgcaa tggatttcag ttctatgcaa cagagaaaac 300
 aacatggtat catcggcact actggggaga aaagttgacc ctgaacccag gaaacacccct 360
 cgaggttcgg gaatctggac tgtacagatg ccaggcccg ggctcccccac gaagtaaccc 420
 tgtgogcttg ctcttttctt cagactcctt aatcctgcag gcaccatatt ctgtgtttga 480
 aggtgacaca ttggtttctga gatgccacag aagaaggaaa gagaaattga ctgtgttgaa 540
 atatacttgg aatggaaaca ttctttccat ttctaataaa agctgggac ttcttatccc 600
 acaagcaagt tcaaataaca atggcaatta tcgatgcatt ggatatggag atgagaatga 660
 tgtattttaga tcaaatttca aaataattaa aattcaagaa ctatttccac atccagagct 720
 gaaagctaca gaactctcagc ctacagaggg gaattctgta aacctgagct gtgaaacaca 780
 gcttctcca gagcggctcag acaccccaact tcacttcaac ttcttcagag atggcgaggt 840
 catcctgtca gactggagca cgtacccgga actccagctc ccaaccgtct ggagagaaaa 900
 ctcaggatcc tattggtgtg gtgctgaaac agtgaggggt aacatccaca agcacagtc 960
 ctcgctacag atccatgtgc agcgatccc tgtgtctggg gtgctcctgg agacccagcc 1020
 ctcagggggc caggctgttg aaggggagat gctggctcct gtctgctccg tggctgaagg 1080
 cacaggggat accacattct cctggcaccg agaggacatg caggagagtc tggggaggaa 1140
 aactcagcgt tccctgagag cagagctgga gctccctgcc atcagacaga gccatgcagg 1200
 gggatactac tgtacagcag acaacagcta cggccctgtc cagagcatgg tgctgaatgt 1260
 cactgtgaga gagacccag gcaacagaga tggccttgtc gccgoggag ccaactggagg 1320
 gctgctcagt gctcttctcc tggctgtggc cctgctgttt cactgctggc gtggaggaa 1380
 gtcaggagtt ggtttcttgg gagacgaaac caggctccct cccgctccag gccaggaga 1440
 gtctcccat tccatctgcc ctgccagggt ggagcttcag tcgttgtatg ttgatgtaca 1500
 cccaaaaag ggagatttgg tatactctga gatccagact actcagctgg gagaagaaga 1560
 ggaagcta atccctcagga cacttctaga ggataaggat gtctcagttg tctactctga 1620
 ggtaaagaca caacacccag ataactcagc tggaaagatc agctctaagg atgaagaaag 1680
 ttaagagaat gaaaagtac gggaacgtcc tactcatgtg atttctccct tgtccaaagt 1740
 cccaggccca gtgcagtcct tgcggcacct ggaatgatca actcattcca gctttctaat 1800
 tcttctcatg catatgcatt cactcccagg aatactcatt cgtctactct gatgttggga 1860
 tggaatggcc tctgaaagac ttcactaaaa tgaccaggat ccacagttaa gagaagaccc 1920
 tgtagtattt gctgtgggcc tgacctaatg cattccctag ggtctgcttt agagaagggg 1980
 gataaagaga gagaaggact gttatgaaaa acagaagcac aaattttggt gaattgggat 2040

ES 2 556 332 T3

```

ttgcagagat gaaaaagact gggtagacctg gatctctgct taatacatct acaaccattg 2100
tctcaactgga gactcaacttg catcagtttg ttttaactgtg agtggctgca caggcactgt 2160
gcaaacaatg aaaagcccoct tcacttctgc ctgcacagct tacactgtca ggattcagtt 2220
gcagattaaa gaaccocatct ggaatgggtt acagagagag gaatttataa gaggacatca 2280
gaagagctgg agatgcaagc tctaggctgc gcttccaaaa gcaaatgata attatgttaa 2340
tgtcattagt gacaaagatt tgcaacatta gagaaaagag acacaaatat aaaattataa 2400
acttaagtac caactctcca aaactaaatt tgaacttaaa atattagtat aaactcataa 2460
taaactctgc ctttaaataa aaaaaaaaaa aaaaaaaaaa 2499

```

<210> 3

<211> 592

<212> PRT

5 <213> Homo Sapiens

<400> 3

```

Met Leu Leu Trp Val Ile Leu Leu Val Leu Ala Pro Val Ser Gly Gln
1           5           10           15

```

```

Phe Ala Arg Thr Pro Arg Pro Ile Ile Phe Leu Gln Pro Pro Trp Thr
20           25           30

```

```

Thr Val Phe Gln Gly Glu Arg Val Thr Leu Thr Cys Lys Gly Phe Arg
35           40           45

```

```

Phe Tyr Ser Pro Gln Lys Thr Lys Trp Tyr His Arg Tyr Leu Gly Lys
50           55           60

```

```

Glu Ile Leu Arg Glu Thr Pro Asp Asn Ile Leu Glu Val Gln Glu Ser
65           70           75           80

```

```

Gly Glu Tyr Arg Cys Gln Ala Gln Gly Ser Pro Leu Ser Ser Pro Val
85           90           95

```

```

His Leu Asp Phe Ser Ser Ala Ser Leu Ile Leu Gln Ala Pro Leu Ser
100          105          110

```

```

Val Phe Glu Gly Asp Ser Val Val Leu Arg Cys Arg Ala Lys Ala Glu
115          120          125

```

ES 2 556 332 T3

Val Thr Leu Asn Asn Thr Ile Tyr Lys Asn Asp Asn Val Leu Ala Phe
130 135 140

Leu Asn Lys Arg Thr Asp Phe His Ile Pro His Ala Cys Leu Lys Asp
145 150 155 160

Asn Gly Ala Tyr Arg Cys Thr Gly Tyr Lys Glu Ser Cys Cys Pro Val
165 170 175

Ser Ser Asn Thr Val Lys Ile Gln Val Gln Glu Pro Phe Thr Arg Pro
180 185 190

Val Leu Arg Ala Ser Ser Phe Gln Pro Ile Ser Gly Asn Pro Val Thr
195 200 205

Leu Thr Cys Glu Thr Gln Leu Ser Leu Glu Arg Ser Asp Val Pro Leu
210 215 220

Arg Phe Arg Phe Phe Arg Asp Asp Gln Thr Leu Gly Leu Gly Trp Ser
225 230 235 240

Leu Ser Pro Asn Phe Gln Ile Thr Ala Met Trp Ser Lys Asp Ser Gly
245 250 255

Phe Tyr Trp Cys Lys Ala Ala Thr Met Pro His Ser Val Ile Ser Asp
260 265 270

Ser Pro Arg Ser Trp Ile Gln Val Gln Ile Pro Ala Ser His Pro Val
275 280 285

Leu Thr Leu Ser Pro Glu Lys Ala Leu Asn Phe Glu Gly Thr Lys Val
290 295 300

Thr Leu His Cys Glu Thr Gln Glu Asp Ser Leu Arg Thr Leu Tyr Arg
305 310 315 320

Phe Tyr His Glu Gly Val Pro Leu Arg His Lys Ser Val Arg Cys Glu
325 330 335

Arg Gly Ala Ser Ile Ser Phe Ser Leu Thr Thr Glu Asn Ser Gly Asn
340 345 350

Tyr Tyr Cys Thr Ala Asp Asn Gly Leu Gly Ala Lys Pro Ser Lys Ala
355 360 365

ES 2 556 332 T3

Val Ser Leu Ser Val Thr Val Pro Val Ser His Pro Val Leu Asn Leu
370 375 380

Ser Ser Pro Glu Asp Leu Ile Phe Glu Gly Ala Lys Val Thr Leu His
385 390 395 400

Cys Glu Ala Gln Arg Gly Ser Leu Pro Ile Leu Tyr Gln Phe His His
405 410 415

Glu Asp Ala Ala Leu Glu Arg Arg Ser Ala Asn Ser Ala Gly Gly Val
420 425 430

Ala Ile Ser Phe Ser Leu Thr Ala Glu His Ser Gly Asn Tyr Tyr Cys
435 440 445

Thr Ala Asp Asn Gly Phe Gly Pro Gln Arg Ser Lys Ala Val Ser Leu
450 455 460

Ser Ile Thr Val Pro Val Ser His Pro Val Leu Thr Leu Ser Ser Ala
465 470 475 480

Glu Ala Leu Thr Phe Glu Gly Ala Thr Val Thr Leu His Cys Glu Val
485 490 495

Gln Arg Gly Ser Pro Gln Ile Leu Tyr Gln Phe Tyr His Glu Asp Met
500 505 510

Pro Leu Val Ser Ser Ser Thr Pro Ser Val Gly Arg Val Ser Phe Ser
515 520 525

Phe Ser Leu Thr Glu Gly His Ser Gly Asn Tyr Tyr Cys Thr Ala Asp
530 535 540

Asn Gly Phe Gly Pro Gln Arg Ser Glu Val Val Ser Leu Phe Val Thr
545 550 555 560

Gly Lys Cys Trp Val Leu Ala Ser Lys Pro Pro Leu Ala Glu Phe Ser
565 570 575

Leu Thr His Ser Phe Lys Asn Leu Phe Ala Leu Ser Ser Phe Leu Pro
580 585 590

<210> 4

<211> 5308

<212> DNA

<213> Homo Sapiens

ES 2 556 332 T3

<400> 4

cgggtgcagtg tectgactgt aagatcaagt ccaaacctgt tttggaattg aggaaacttc	60
tcttttgatc tcagcccttg gtggtccagg tcttcatgct gctgtgggtg atattactgg	120
tcttggtctc tgtcagtga cagtttgcaa ggacaccag gccattatt ttcctccagc	180
tctccatgga ccacagtctt ccaaggagag agagtgacct tcacttgcaa gggatttcgc	240
tctactcacc acagaaaaca aaatggtacc atcggtacct tgggaaagaa atactaagag	300
aaacccagag caatatcctt gagttcagga atctggagag tacagatgcc aggccaggg	360
ctccctctc agtagccctg tgcacttgga tttttcttca gcttcgctga tctgcaagc	420
tcactttct gtgtttgaag gagactctgt ggttctgagg tgccgggcaa aggcggaagt	480
aacactgaat aatactatctt acaagaatga taatgtcctg gcattcctta ataaaagaac	540
tgacttccat attctcatg catgtctcaa ggacaatggt gcatacgcct gtactggata	600
taaggaaagt tgttgccctg tttcttccaa tacagtcaaa atccaagtcc aagagccatt	660
tacacgtcca gtgctgagag ccagctcctt ccagccatc agcgggaacc cagtgacct	720
gacctgtgag acccagctct ctctagagag gtcagatgtc ccgtccggt tccgttctt	780
cagagatgac cagacctgg gattaggctg gagtctctcc ccgaatttcc agattactgc	840
catgtggagt aaagattcag ggttctactg gtgtaaggca gcaacaatgc ctacacgcgt	900
catatctgac agcccgaga tcttgatac aggtgcagat cctgcctct catctgtcc	960
tcactctcag cctgaaaag gctctgaatt ttgagggaac caaggtgaca cttactgtg	1020
aaaccagga agattctctg cgcactttgt acaggtttta tcatgagggg gtccctctga	1080
ggcacaagtc agtcgctgt gaaaggggag catccatcag cttctcactg actacagaga	1140
attcaggga ctactactgc acagctgaca atggccttg gcgcaagccc agtaaggctg	1200
tgagcctctc agtcactgtt cccgtgtctc atctgtcct caacctcagc tctctgagg	1260
acctgatttt tgagggagcc aaggtgacac ttactgtga agccagaga ggttactcc	1320
ccatctgta ccagtttcat catgaggatg ctgccctgga gcgtaggctg gccactctg	1380
caggaggagt ggccatcagc ttctctctga ctgcagagca ttcagggaac tactactgca	1440
cagctgacaa tggctttggc cccagcgca gtaaggcggg gagcctctcc atcactgtcc	1500
ctgtgtctca tctgtctct accctcagct ctgctgaggc cctgactttt gaaggagcca	1560
ctgtgacact tcactgtgaa gtccagagag gttcccccaca aatcctatac cagttttatc	1620

ES 2 556 332 T3

atgaggacat gccctgtgg agcagctcaa caccctctgt gggaagagtg tccttcagct	1680
tctctctgac tgaaggacat tcagggaatt actactgcac agctgacaat ggctttggtc	1740
cccagcgcag tgaagtgggtg ggtaagtgtt gggttcttgc cagtcacca cccctggctg	1800
agttctctct caccattcc tttaaaaatc tgtttgcact gtccagtttc ctccccta	1860
caacttaatc cccttcttgg ctctctctc aactaactag ctgggggttt cogtactcat	1920
aagtctggc tcagccagac cctaaaaaca gctcagtaga ttccccagct tttaccaa	1980
gaatttattt attgtatttt ctctctatc ctgttatgtt ccaacagtac gccattttt	2040
cttgatgcac ggagcgtgtc ctactctct actgacattt acatattaac ttagctacaa	2100
gcacagtctt atagataaat attggtcaag accttaaatt ctccaaagga tttccaatct	2160
tatggtagat ttggagaaag ctgctgggtg acaaaagggg aaatggctcc ctaggaacca	2220
actctcaaaa ctctggagt ttttatgac ccttggtttc taacctgcta aaatcagtat	2280
cattttattg tattatttta aaaaaactat tgttgaagta tgacatacat tcaagaaacg	2340
tgtgcaaatt gtatgtgtac gatttgggtt ctcttttaga gctaagttgc ttctgtttt	2400
acttgaatct ttgtttatag aaactggggg aaagttaact ttcttttcag agaagccaaa	2460
tggtatgata gaaaaatctt gagcctgatg tgtcagacat gccctagca taacttgttg	2520
agtaaagagg ttatttttaa aatgtgaatg ttctgagact actccaaagt cagagccaaa	2580
tctactagga agctcttaga ctctactcat tctgcaccc attactatct ttttatccat	2640
gttttacttt ctctctatat tcagcagcat cttaagcctc tttattttct gttctctgac	2700
tgtcaccctt aatgccagta gaatgtaagc ttcattgagaa cagaactgca tccatcttgg	2760
tcttcacaa atccctgtgc ctactcagtg ttggcacac agtaggtct cagtcaacat	2820
ttgtaattta gtggacagat gatatgacaa gatgataaga ggggatttaa aaaaatcatc	2880
tagcaaaagg caagaggaaa aaaaacaaag ctatttttaga aatgaaatc caatttgaag	2940
cagtaagaat agattggata tctttgaaaa ccatttaattg aatgaagaac caatttgaga	3000
aaacaataca gaatgcaaag tagaaagata cagaaataaa ggcaaaagtt ataatatgga	3060
aatcagacaa tggatttgtc tgtatccagt tatgtggata attaaaatgg agaccctcag	3120
aaaattgaac cgaagagtaa aatgaaactc aaaaatgtag tagaaattgt tgggaagtaa	3180
agaaaacttg aatatgtaga tcagaacata tatgttgatg acgttattga ctttgagggt	3240
aaaaatatat atatgtgcct atgattatgg ggaaaaaagc agtcgtctca gaaagaaaaa	3300
catcaagtta gtcttagaact ttgcagtgca ctcagtaacca aagagttacc acacaaaggg	3360
agagtggggc ttcaggagat gccgggctgg cctaacagct caggtgctcc taaactccga	3420

ES 2 556 332 T3

cacagagttc ctgctttggg tggatgcatt totcaattgt catcagcctg gtggggctac 3480
 tgcagtgtgc tgccaaatgg gacagcacac agcctgtgca catgggacat gtgatgggtc 3540
 tccccacggg ggtgcattt cacactctc cacctgtctc aaactctaag gtgggcaatt 3600
 gacaccaagg taactctctc cctgctcatg tgtcagtgtc tacctgccca agtaagtggc 3660
 tttcatacac caagtccga agttcttccc atcctaacag aagtaacca gcaagtcaag 3720
 ggcaggagga ccaggggtgc agacagaaca catactggaa cacaggaggt gctcaattac 3780
 tatttgactg actgactgaa tgaatgaatg aatgaggaag aaaactgtgg gtaatcaaac 3840
 tggcataaaa tccagtgcac tccctaggaa atccgggagg tattctggct tccctaagaaa 3900
 caacggaaga gaaggagctt ggatgaagaa actgttcagc aagaagaagg gcttcttcac 3960
 acttttatgt gcttgtggat cacctgagga totgtgaaaa tacagatact gattcagtgg 4020
 gtctgtgtag agcctgagac tgccattcta acatgttccc aggggatgct gatgctgctg 4080
 gccctgggac tgcactgcat gcatgtgaag ccttataggt ctccagagag gcccatggag 4140
 agggaatgtg tggctctggc tgcccagggc ccaactcggg tcacacggat cgtgctgctc 4200
 cctggccagc ctttggccac agcaccacca gctgctgttg ctgagagagc ttcttctctg 4260
 tgacatgttg gctttcatca gccaccctgg gaagcggaaa gtagctgcca ctatctttgt 4320
 ttccccacct caggcctcac actttcccat gaaaagggtg aatgtatata acctgagccc 4380
 tctccattca gaggtttct cccatctctg agcaatggga tgttctgttc cgtttttatg 4440
 atatccatca catcttatct tgatctttgc tcccagtggg ttgtacagtg atgactttta 4500
 agccccacgg cctgaaata aaatccttcc aagggcattg gaagctcact ccacctgaac 4560
 catggccttt catgcttcca agtgtcaggg ccttgcccag atagacaggg ctgactctgc 4620
 tgccccaacc tttcaaggag gaaaccagac acctgagaca ggagcctgta tgcagcccag 4680
 tgcagccttg cagaggacaa ggctggaggc atttgtcatc actacagata tgcaactaaa 4740
 atagacgtgg agcaagagaa atgcattccc accgaggcgg ctttttttagg cctagttgaa 4800
 agtcaagaag gacagcagca agcataggct caggattaaa gaaaaaaatc tgctcacagt 4860
 ctgttctgga ggtcacatca ccaacaaagc tcacgcccta tgcagttctg agaagggtga 4920
 ggcaccaggc tcaaaagagg aaatttagaa tttctcattg ggagagtaag gtacccccat 4980
 ccagaatga taactgcaca gtggcagaac aaactccaco ctaatgtggg tggacccccat 5040
 ccagtctgtt gaaggcctga atgtaacaaa agggcttatt ctctctcaag taagggggaa 5100
 ctctctgctt gggctgggac ataagttttt ctgctttcag acgcaaaactg aaaaatggct 5160
 cttcttgggt cttgagcttg ctggcatatg gactgaaaga aactatgcta ttggatctcc 5220
 tggatctcca gcttctgac tgcagatctt gagatatgtc agcctctaca gtcacaagag 5280
 ctaattcatt ctaataaacc aatctttc 5308

<210> 5

<211> 734

<212> PRT

<213> Homo Sapiens

5

ES 2 556 332 T3

<400> 5

Met Leu Leu Trp Leu Leu Leu Leu Ile Leu Thr Pro Gly Arg Glu Gln
1 5 10 15

Ser Gly Val Ala Pro Lys Ala Val Leu Leu Leu Asn Pro Pro Trp Ser
20 25 30

Thr Ala Phe Lys Gly Glu Lys Val Ala Leu Ile Cys Ser Ser Ile Ser
35 40 45

His Ser Leu Ala Gln Gly Asp Thr Tyr Trp Tyr His Asp Glu Lys Leu
50 55 60

Leu Lys Ile Lys His Asp Lys Ile Gln Ile Thr Glu Pro Gly Asn Tyr
65 70 75 80

Gln Cys Lys Thr Arg Gly Ser Ser Leu Ser Asp Ala Val His Val Glu
85 90 95

Phe Ser Pro Asp Trp Leu Ile Leu Gln Ala Leu His Pro Val Phe Glu
100 105 110

Gly Asp Asn Val Ile Leu Arg Cys Gln Gly Lys Asp Asn Lys Asn Thr
115 120 125

His Gln Lys Val Tyr Tyr Lys Asp Gly Lys Gln Leu Pro Asn Ser Tyr
130 135 140

Asn Leu Glu Lys Ile Thr Val Asn Ser Val Ser Arg Asp Asn Ser Lys
145 150 155 160

Tyr His Cys Thr Ala Tyr Arg Lys Phe Tyr Ile Leu Asp Ile Glu Val
165 170 175

ES 2 556 332 T3

Thr Ser Lys Pro Leu Asn Ile Gln Val Gln Glu Leu Phe Leu His Pro
180 185 190

Val Leu Arg Ala Ser Ser Ser Thr Pro Ile Glu Gly Ser Pro Met Thr
195 200 205

Leu Thr Cys Glu Thr Gln Leu Ser Pro Gln Arg Pro Asp Val Gln Leu
210 215 220

Gln Phe Ser Leu Phe Arg Asp Ser Gln Thr Leu Gly Leu Gly Trp Ser
225 230 235 240

Arg Ser Pro Arg Leu Gln Ile Pro Ala Met Trp Thr Glu Asp Ser Gly
245 250 255

Ser Tyr Trp Cys Glu Val Glu Thr Val Thr His Ser Ile Lys Lys Arg
260 265 270

Ser Leu Arg Ser Gln Ile Arg Val Gln Arg Val Pro Val Ser Asn Val
275 280 285

Asn Leu Glu Ile Arg Pro Thr Gly Gly Gln Leu Ile Glu Gly Glu Asn
290 295 300

Met Val Leu Ile Cys Ser Val Ala Gln Gly Ser Gly Thr Val Thr Phe
305 310 315 320

Ser Trp His Lys Glu Gly Arg Val Arg Ser Leu Gly Arg Lys Thr Gln
325 330 335

Arg Ser Leu Leu Ala Glu Leu His Val Leu Thr Val Lys Glu Ser Asp
340 345 350

Ala Gly Arg Tyr Tyr Cys Ala Ala Asp Asn Val His Ser Pro Ile Leu
355 360 365

Ser Thr Trp Ile Arg Val Thr Val Arg Ile Pro Val Ser His Pro Val
370 375 380

Leu Thr Phe Arg Ala Pro Arg Ala His Thr Val Val Gly Asp Leu Leu
385 390 395 400

Glu Leu His Cys Glu Ser Leu Arg Gly Ser Pro Pro Ile Leu Tyr Arg
405 410 415

Phe Tyr His Glu Asp Val Thr Leu Gly Asn Ser Ser Ala Pro Ser Gly

ES 2 556 332 T3

420	425	430
Gly Gly Ala Ser Phe Asn Leu Ser Leu Thr Ala Glu His Ser Gly Asn		
435	440	445
Tyr Ser Cys Asp Ala Asp Asn Gly Leu Gly Ala Gln His Ser His Gly		
450	455	460
Val Ser Leu Arg Val Thr Val Pro Val Ser Arg Pro Val Leu Thr Leu		
465	470	480
Arg Ala Pro Gly Ala Gln Ala Val Val Gly Asp Leu Leu Glu Leu His		
	485	490 495
Cys Glu Ser Leu Arg Gly Ser Phe Pro Ile Leu Tyr Trp Phe Tyr His		
	500	505 510
Glu Asp Asp Thr Leu Gly Asn Ile Ser Ala His Ser Gly Gly Gly Ala		
	515	520 525
Ser Phe Asn Leu Ser Leu Thr Thr Glu His Ser Gly Asn Tyr Ser Cys		
	530	535 540
Glu Ala Asp Asn Gly Leu Gly Ala Gln His Ser Lys Val Val Thr Leu		
545	550	555 560
Asn Val Thr Gly Thr Ser Arg Asn Arg Thr Gly Leu Thr Ala Ala Gly		
	565	570 575
Ile Thr Gly Leu Val Leu Ser Ile Leu Val Leu Ala Ala Ala Ala Ala		
	580	585 590
Leu Leu His Tyr Ala Arg Ala Arg Arg Lys Pro Gly Gly Leu Ser Ala		
	595	600 605
Thr Gly Thr Ser Ser His Ser Pro Ser Glu Cys Gln Glu Pro Ser Ser		
	610	615 620
Ser Arg Pro Ser Arg Ile Asp Pro Gln Glu Pro Thr His Ser Lys Pro		
625	630	635 640
Leu Ala Pro Met Glu Leu Glu Pro Met Tyr Ser Asn Val Asn Pro Gly		
	645	650 655
Asp Ser Asn Pro Ile Tyr Ser Gln Ile Trp Ser Ile Gln His Thr Lys		
	660	665 670

ES 2 556 332 T3

Glu Asn Ser Ala Asn Cys Pro Met Met His Gln Glu His Glu Glu Leu
675 680 685

Thr Val Leu Tyr Ser Glu Leu Lys Lys Thr His Pro Asp Asp Ser Ala
690 695 700

Gly Glu Ala Ser Ser Arg Gly Arg Ala His Glu Glu Asp Asp Glu Glu
705 710 715 720

Asn Tyr Glu Asn Val Pro Arg Val Leu Leu Ala Ser Asp His
725 730

<210> 6

<211> 2970

<212> DNA

5 <213> Homo Sapiens

<400> 6

```

agtgaagggg tttcccatat gaaaaataca gaaagaatta tttgaatact agcaaataca      60
caacttgata tttctagaga acccaggcac agtcttgag acattactcc tgagagactg      120
cagctgatgg aagatgagcc ccaacttcta aaaatgtatc actacoggga ttgagataca      180
aacagcattt aggaaggtct catctgagta gcagcttcct gccctccttc ttggagataa      240
gtggggcttt tggtagagaca gaatttccca accctctgcc eggcgggtgc ccatgcttct      300
gtggctgctg ctgctgatcc tgactcctgg aagagaacaa tcaggggtgg ccccaaaagc      360
tgtacttctc ctcaatcctc catggtccac agccttcaaa ggagaaaaag tggctctcat      420
atgcagcagc atatcacatt cctagccca gggagacaca tattggtatc acgatgagaa      480
gttggtgaaa ataaaacatg acaagatcca aattacagag cctggaaatt accaatgtaa      540
gacccgagga tcctccctca gaattacaga gcctggaaat taccaatgta agacccgagg      600
atctccctc agacatcctg tctttgaagg agacaatgtc attctgagat gtcaggggaa      660
agacaacaaa aacactcatc aaaagggtta ctacaaggat ggaaaacagc ttctaataag      720
ttataattta gagaagatca cagtgaattc agtctccagg gataatagca aatatcattg      780
tactgcttat aggaagtitt acatacttga cattgaagta acttcaaaac cctaaaatat      840
ccaagttcaa gagctgtttc tacatcctgt gctgagagcc agctcttcca cgcccataga      900

```

ES 2 556 332 T3

ggggagtcac atgacctga cctgtgagac ccagctctct ccacagaggc cagatgtcca 960
 gctgcaatcc tccctcttca gagatagcca gacctctgga ttgggctgga gcaggctccc 1020
 cagactccag atccctgcca tgtggactga agaactcaggg tcttactggt gtgagggtga 1080
 gacagtgact cacagcatca aaaaaaggag cctgagatct cagatacgtg tacagagagt 1140
 ccctgtgtct aatgtgaatc tagagatccg gccacccgga gggcagctga ttgaaggaga 1200
 aaatatgggc cttatttgcct cagtagccca ggggttcaggg actgtcacat tctcctggca 1260
 caaagaagga agagtaagaa gcctgggtag aaagaccag cgttccctgt tggcagagct 1320
 gcatgttctc accgtgaagg agagtgatgc agggagatac tactgtgcag ctgataacgt 1380
 tcacagcccc atcctcagca cgtggattcg agtcaccgtg agaattccgg tatctcacc 1440
 tgtctcacc ttcagggtct ccagggccca cactgtggtg ggggacctgc tggagcttca 1500
 ctgtgagtc ctgagaggct ccccccgat cctgtaccga ttttatcatg aggatgtcac 1560
 cctggggaac agctcagccc cctctggagg aggagcctcc tcaacctct ctctgactgc 1620
 agaacattct ggaaactact cctgtgatgc agacaatggc ctggggggccc agcacagtca 1680
 tggagtgagt ctcagggtca cagttccggt gtctcgcccc gtccacccc tcagggtcc 1740
 cggggcccag gctgtggtg gggacctgct ggagcttcac tgtgagtcac tgagaggctc 1800
 cttcccgatc ctgtactggt tttatcacga ggatgacacc ttggggaaca tctcgcccca 1860
 ctctggagga ggggcattct tcaacctctc tctgactaca gaacattctg gaaactactc 1920
 atgtgaggct gacaatggcc tgggggccca gcacagtaaa gtggtgacac tcaatgttac 1980
 aggaacttcc aggaacagaa caggccctac cgtcgccgga atcacggggc tgggtgctcag 2040
 cactctctgc cttgctgctg ctgctgctct gctgcattac gccagggccc gaaggaaacc 2100
 aggaggactt tctgccactg gaacatctag tcacagtcct agtgagtgtc aggagccttc 2160
 ctctccagg ccttccagga tagacctca agagccact cactctaaac cactagcccc 2220
 aatggagctg gagccaatgt acagcaatgt aaatcctgga gatagcaacc cgatttattc 2280
 ccagatctgg agcatccagc atacaaaaga aaactcagct aattgtccaa tgatgcatca 2340
 agagcatgag gaacttacag tctctattc agaactgaag aagacacacc cagacgactc 2400
 tgcaggggag gctagcagca gaggcagggc ccattgaagaa gatgatgaag aaaactatga 2460
 gaatgtacca cgtgtattac tggcctcaga ccactagccc cttaccaga gtggcccaca 2520
 ggaaacagcc tgcaccattt tttttctgt tctctccaa cacacatcat ccatctctcc 2580
 agactctgcc tctacgagg ctgggctgca gggatatgta ggctgagcaa aaggctctgca 2640
 aatctccct gtgcctgato tgtgtgttcc ccaggaagag agcaggcagc ctctgagcaa 2700
 gcaactgtgtt attttcacag tggagacag tggcaaggca ggagggccct cagctcctag 2760

ES 2 556 332 T3

ggctgtcgaa tagaggagga gagagaaatg gtctagccag ggttacaagg gcacaatcat 2820
gaccatttga tccaagtgtg atcgaaagct gttaatgtgc tctctgtata aacaatttgc 2880
tccaaatatt ttgtttccct tttttgtgtg gctggtagtgc gcattgctga tgttttggtg 2940
tatatgctgt atccttgcta ccatattggg 2970

<210> 7

<211> 508

<212> PRT

5 <213> Homo Sapiens

<400> 7

Met Leu Leu Trp Ser Leu Leu Val Ile Phe Asp Ala Val Thr Glu Gln
1 5 10 15

Ala Asp Ser Leu Thr Leu Val Ala Pro Ser Ser Val Phe Glu Gly Asp
20 25 30

Ser Ile Val Leu Lys Cys Gln Gly Glu Gln Asn Trp Lys Ile Gln Lys
35 40 45

Met Ala Tyr His Lys Asp Asn Lys Glu Leu Ser Val Phe Lys Lys Phe
50 55 60

Ser Asp Phe Leu Ile Gln Ser Ala Val Leu Ser Asp Ser Gly Asn Tyr
65 70 75 80

Phe Cys Ser Thr Lys Gly Gln Leu Phe Leu Trp Asp Lys Thr Ser Asn
85 90 95

Ile Val Lys Ile Lys Val Gln Glu Leu Phe Gln Arg Pro Val Leu Thr
100 105 110

Ala Ser Ser Phe Gln Pro Ile Glu Gly Gly Pro Val Ser Leu Lys Cys
115 120 125

Glu Thr Arg Leu Ser Pro Gln Arg Leu Asp Val Gln Leu Gln Phe Cys
130 135 140

Phe Phe Arg Glu Asn Gln Val Leu Gly Ser Gly Trp Ser Ser Ser Pro
145 150 155 160

ES 2 556 332 T3

Glu Leu Gln Ile Ser Ala Val Trp Ser Glu Asp Thr Gly Ser Tyr Trp
 165 170 175

Cys Lys Ala Glu Thr Val Thr His Arg Ile Arg Lys Gln Ser Leu Gln
 180 185 190

Ser Gln Ile His Val Gln Arg Ile Pro Ile Ser Asn Val Ser Leu Glu
 195 200 205

Ile Arg Ala Pro Gly Gly Gln Val Thr Glu Gly Gln Lys Leu Ile Leu
 210 215 220

Leu Cys Ser Val Ala Gly Gly Thr Gly Asn Val Thr Phe Ser Trp Tyr
 225 230 235 240

Arg Glu Ala Thr Gly Thr Ser Met Gly Lys Lys Thr Gln Arg Ser Leu
 245 250 255

Ser Ala Glu Leu Glu Ile Pro Ala Val Lys Glu Ser Asp Ala Gly Lys
 260 265 270

Tyr Tyr Cys Arg Ala Asp Asn Gly His Val Pro Ile Gln Ser Lys Val
 275 280 285

Val Asn Ile Pro Val Arg Ile Pro Val Ser Arg Pro Val Leu Thr Leu
 290 295 300

Arg Ser Pro Gly Ala Gln Ala Ala Val Gly Asp Leu Leu Glu Leu His
 305 310 315 320

Cys Glu Ala Leu Arg Gly Ser Pro Pro Ile Leu Tyr Gln Phe Tyr His
 325 330 335

Glu Asp Val Thr Leu Gly Asn Ser Ser Ala Pro Ser Gly Gly Gly Ala
 340 345 350

Ser Phe Asn Leu Ser Leu Thr Ala Glu His Ser Gly Asn Tyr Ser Cys
 355 360 365

Glu Ala Asn Asn Gly Leu Gly Ala Gln Cys Ser Glu Ala Val Pro Val
 370 375 380

Ser Ile Ser Gly Pro Asp Gly Tyr Arg Arg Asp Leu Met Thr Ala Gly
 385 390 395 400

ES 2 556 332 T3

Val Leu Trp Gly Leu Phe Gly Val Leu Gly Phe Thr Gly Val Ala Leu
405 410 415

Leu Leu Tyr Ala Leu Phe His Lys Ile Ser Gly Glu Ser Ser Ala Thr
420 425 430

Asn Glu Pro Arg Gly Ala Ser Arg Pro Asn Pro Gln Glu Phe Thr Tyr
435 440 445

Ser Ser Pro Thr Pro Asp Met Glu Glu Leu Gln Pro Val Tyr Val Asn
450 455 460

Val Gly Ser Val Asp Val Asp Val Val Tyr Ser Gln Val Trp Ser Met
465 470 475 480

Gln Gln Pro Glu Ser Ser Ala Asn Ile Arg Thr Leu Leu Glu Asn Lys
485 490 495

Asp Ser Gln Val Ile Tyr Ser Ser Val Lys Lys Ser
500 505

<210> 8

<211> 2580

<212> DNA

5 <213> Homo Sapiens

<400> 8

tggtgaccaa gaggacatct cttttcaaatt agctggatta ggtcctcatg ctgctgtggt	60
cattgctggt catctttgat gcagtcactg aacaggcaga ttgctgacc cttgtggcgc	120
cctctttctgt cttcgaagga gacagcatcg ttctgaaatg ccaggagaga cagaactgga	180
aaattcagaa gatggcttac cataaggata acaaagagtt atctgttttc aaaaaattct	240
cagatttcct tatccaaagt gcagttttta gtgacagtgg taactatttc tgtagtaoca	300
aaggacaact ctttctctgg gataaaaactt caaatatagt aaagataaaa gtccaagagc	360
tctttcaacg tctgtgctg actgccagct ccttcagacc catcgaaggg ggtccagtga	420
gctgaaatg tgagaccgg ctctctccac agaggttgga tgttcaactc cagttctgct	480
tcttcagaga aaaccaggtc ctggggtcag gctggagcag ctctccggag ctccagattt	540
ctgccgtgtg gagtgaagac acagggtctt actgggtgcaa ggcagaaaac gtgactcaca	600

ES 2 556 332 T3

ggatcagaaa acagagcctc caatcccaga ttcacgtgca gagaatcccc atctctaatg 660
 taagcttggg gatccggggc cccggggggac aggtgactga aggacaaaaa ctgatcctgc 720
 tctgctcagt ggctgggggt acaggaaatg tcacattctc ctggtacaga gaggccacag 780
 gaaccagtat gggaaagaaa acccagcggt cctgtgcagc agagctggag atcccagctg 840
 tgaaagagag tgatgccggc aaatattact gtagagctga caacggccat gtgcctatcc 900
 agagcaagggt ggtgaatata cctgtgagaa ttccagtgtc tcgccctgtc ctcacctca 960
 ggtctcctgg gggccagget gcagtggggg acctgctgga gcttccactgt gaggccctga 1020
 gaggtctctc cccaatcttg taccaatttt atcatgagga tgtcaccctt gggaacagct 1080
 cggcccccctc tggaggaggg gcctccttca acctctcttt gactgcagaa cattctggaa 1140
 actactcctg tgaggccaac aacggcctgg gggcccagtg cagtgagga gtgccagctc 1200
 ccatctcagg acctgatggc tatagaagag acctcatgac agctggagtt ctctggggac 1260
 tgtttggtgt ccttggtttc actggtgttg ctttgcctgt gtatgccttg ttccacaaga 1320
 tatcaggaga aagttctgcc actaatgaac ccagaggggc ttccaggcca aatctcaag 1380
 agttcaccta ttcaagccca accccagaca tggaggagct gcagccagtg tatgtcaatg 1440
 tgggtctctgt agatgtggat gtggtttatt ctccaggtctg gagcatgcag cagccagaaa 1500
 gctcagcaaa catcaggaca cttctggaga acaaggactc ccaagtcac tactcttctg 1560
 tgaagaaatc ataacacttg gaggaatcag aagggaagat caacagcaag gatggggcat 1620
 cattaagact tgctataaaa ctttatgaaa atgcttgagg cttatcacct gccacagcca 1680
 gaacgtgcct caggaggcac ctctgtcat tttgtcctg atgatgtttc ttctccaata 1740
 tcttctttta cctatcaata ttcatgaac tgctgtaca tccagacact gtgcaaataa 1800
 attatttctg ctacctctc ttaagcaatc agtgtgtaaa gatttgaggg aagaatgaat 1860
 aagagataca aggtctcacc ttcatctact gtgaagtgat gagaacagga cttgatagtg 1920
 gtgtattaac ttatttatgt gctgctggat acagtttgct aatattttgt tgagaatttt 1980
 tgcaaatatg ttcatggga atattggcct gaaattttct ttccactgt gtctctgcca 2040
 gaatgtttgt atcaggetga tgctggcttc atagaatgag ttaggcagga gcccttctc 2100
 cttgattttt tggcatagtt tcagcaggat tggtaaccagt tattctttct gcattctgta 2160
 gaattcagct atgaatccat ctggtctagg gcttttgtgt tggttggtaa gttttttatt 2220
 actaattcaa cttcagcgct tgatattggt ctaggagggg tttctgtctc ttcttggttc 2280
 aatcttggga gattgtgtgt ttccaggaat ttagccgttt cctccagatt ttcttttta 2340
 tgtgcacoga cttgagtgt aacataactt atatgcactg ggaaaccaa aaatctgtgt 2400
 gacttgcttt attgcagcat ttgttttatt ttggtagtct ggaactgaac ctgcaatata 2460
 accaaagtat gcataatagtt gcaaaaatgt gatttttgac atagtaaata tgagtatttg 2520
 caataaacta tgatattact tttgtaagta tatagaataa aatgtaaata atctataaaa 2580

ES 2 556 332 T3

<210> 9
 <211> 429
 <212> PRT
 <213> Homo Sapiens

5 <400> 9
 Met Leu Pro Arg Leu Leu Leu Leu Ile Cys Ala Pro Leu Cys Glu Pro
 1 5 10 15
 Ala Glu Leu Phe Leu Ile Ala Ser Pro Ser His Pro Thr Glu Gly Ser
 20 25 30
 Pro Val Thr Leu Thr Cys Lys Met Pro Phe Leu Gln Ser Ser Asp Ala
 35 40 45
 Gln Phe Gln Phe Cys Phe Phe Arg Asp Thr Arg Ala Leu Gly Pro Gly
 50 55 60
 Trp Ser Ser Ser Pro Lys Leu Gln Ile Ala Ala Met Trp Lys Glu Asp
 65 70 75 80
 Thr Gly Ser Tyr Trp Cys Glu Ala Gln Thr Met Ala Ser Lys Val Leu
 85 90 95
 Arg Ser Arg Arg Ser Gln Ile Asn Val His Arg Val Pro Val Ala Asp
 100 105 110
 Val Ser Leu Glu Thr Gln Pro Pro Gly Gly Gln Val Met Glu Gly Asp
 115 120 125
 Arg Leu Val Leu Ile Cys Ser Val Ala Met Gly Thr Gly Asp Ile Thr
 130 135 140
 Phe Leu Trp Tyr Lys Gly Ala Val Gly Leu Asn Leu Gln Ser Lys Thr
 145 150 155 160
 Gln Arg Ser Leu Thr Ala Glu Tyr Glu Ile Pro Ser Val Arg Glu Ser
 165 170 175

ES 2 556 332 T3

Asp Ala Glu Gln Tyr Tyr Cys Val Ala Glu Asn Gly Tyr Gly Pro Ser
 180 185 190
 Pro Ser Gly Leu Val Ser Ile Thr Val Arg Ile Pro Val Ser Arg Pro
 195 200 205
 Ile Leu Met Leu Arg Ala Pro Arg Ala Gln Ala Ala Val Glu Asp Val
 210 215 220
 Leu Glu Leu His Cys Glu Ala Leu Arg Gly Ser Pro Pro Ile Leu Tyr
 225 230 235 240
 Trp Phe Tyr His Glu Asp Ile Thr Leu Gly Ser Arg Ser Ala Pro Ser
 245 250 255
 Gly Gly Gly Ala Ser Phe Asn Leu Ser Leu Thr Glu Glu His Ser Gly
 260 265 270
 Asn Tyr Ser Cys Glu Ala Asn Asn Gly Leu Gly Ala Gln Arg Ser Glu
 275 280 285
 Ala Val Thr Leu Asn Phe Thr Val Pro Thr Gly Ala Arg Ser Asn His
 290 295 300
 Leu Thr Ser Gly Val Ile Glu Gly Leu Leu Ser Thr Leu Gly Pro Ala
 305 310 315 320
 Thr Val Ala Leu Leu Phe Cys Tyr Gly Leu Lys Arg Lys Ile Gly Arg
 325 330 335
 Arg Ser Ala Arg Asp Pro Leu Arg Ser Leu Pro Ser Pro Leu Pro Gln
 340 345 350
 Glu Phe Thr Tyr Leu Asn Ser Pro Thr Pro Gly Gln Leu Gln Pro Ile
 355 360 365
 Tyr Glu Asn Val Asn Val Val Ser Gly Asp Glu Val Tyr Ser Leu Ala
 370 375 380
 Tyr Tyr Asn Gln Pro Glu Gln Glu Ser Val Ala Ala Glu Thr Leu Gly
 385 390 395 400
 Thr His Met Glu Asp Lys Val Ser Leu Asp Ile Tyr Ser Arg Leu Arg
 405 410 415
 Lys Ala Asn Ile Thr Asp Val Asp Tyr Glu Asp Ala Met
 420 425

<210> 10

<211> 2303

<212> DNA

<213> Homo Sapiens

ES 2 556 332 T3

<400> 10

gaggcatctc taggtaccat cctgacctg gtctcatgc tgcgaggct gttgctgttg	60
atctgtgctc cactctgtga acctgccgag ctgtttttga tagccagccc ctcccatccc	120
acagagggga gccagtgac cctgacgtgt aagatgccct ttctacagag ttcagatgcc	180
cagttccagt tctgcttttt cagagacacc cgggoccttg gcccaggctg gagcagctcc	240
cccaagctcc agatcgctgc catgtggaaa gaagacacag ggtcatactg gtgcgaggca	300
cagacaatgg cgtccaaagt cttgaggagc aggagatccc agataaatgt gcacagggtc	360
cctgtcgctg atgtgagctt ggagactcag ccccaggag gacaggtgat ggaggagac	420
aggctggtcc tcactgtctc agttgctatg ggcacaggag acatcacctt cctttggtac	480
aaaggggctg taggtttaaa ccttcagtca aagaccagc gttcactgac agcagagtat	540
gagattcctt cagtgagga gagtgatgt gagcaatatt actgtgtago tgaatatggc	600
tatggtccca gccccagtgg gctggtgagc atcactgtca gaatccgggt gtctcgccca	660
atctcatgc tcagggctcc cagggccag gctgcagtgg aggatgtgct ggagcttcac	720
tgtgaggccc tgagaggctc tctccaatc ctgtactggt tttatcagc ggatatcacc	780
ctggggagca ggtcgccccc ctctggagga ggagcctct tcaaccttcc cctgactgaa	840
gaacattctg gaaactactc ctgtgaggcc aacaatggcc tgggggccc ggcagtgag	900
gggtgacac tcaacttcac agtgctact ggggccagaa gcaatcatct tacctcagga	960
gtcattgagg ggctgctcag cacccttggc ccagccaccg tggccttatt attttgctac	1020
ggcctcaaaa gaaaaatagg aagacgttca gccagggatc cactcaggag ccttcccagc	1080
cctctacccc aagagttcac ctacctaac tcacctacc cagggcagct acagcctata	1140
tatgaaaatg tgaatgttgt aagtgggat gaggtttatt cactggcgta ctataaccag	1200
cggagcagg aatcagtagc agcagaaacc ctggggacac atatggagga caaggtttcc	1260
ttagacatct attccaggct gaggaagca aacattacag atgtggacta tgaagatgct	1320

ES 2 556 332 T3

atgtaagggtt atggaagatt ctgctcttttg aaaaccatcc atgaccccaa gcctcaggcc 1380
 tgatatgttc ttcagagatc ctggggcatt agctttccag tatacctctt ctggatgcca 1440
 ttctccatgg cactattcct tcctctactg tgaagtgaag ttggcgccgc cctgaagaaa 1500
 ctacctagga gaactaatag acacaggagt gacagggact ttgttatcag aaccagattc 1560
 ctgccggctc ctttgaaaaac aggtcatatt gtgctcttct gtttacaaga ggaaacaaga 1620
 tgggaataaaa gaaattggga tottggggtg gagggacagt gaagcttaga gcacatgaac 1680
 tcaaggtttag tgactctgca ggacttcaca gagagagctg tgcccatcat tcagtccaag 1740
 tgctttctct gccagacag cacagaactc cagccccgt acttacatgg atcatcgagt 1800
 ttccacctaa aatatgattc tttttatctt ggtcactgt taccaaatta gaactaaaac 1860
 aaagttacat aaaaagttat tgtgactcca cttaatttta gtgacgtatt tttgtatata 1920
 taggccaacc tataccacat ccaaaattat gtatctatta cagccctag aagctttata 1980
 aatacagtgt gtctctcttt attcacaaaa tttttgaaat cgtggttaata tggtttgaaa 2040
 cctgtatctt aattattttt tttttaaatt gagacagggc ctactctgt cactcaatct 2100
 ggaatgcagt ggcacaatct tgctcactg caacgcctgc ctctcaggct caagcaaacc 2160
 tctcacctca gcctgctgag tagctgggac tacaggcaca tgccacccaa cttggccatt 2220
 ttttgtctta cgtagagaca agatttcacc gttttgccca ggctggctct aaactcctgg 2280
 gctcaagcaa tgtattgaat ttt 2303

<210> 11
 <211> 90
 5 <212> DNA
 <213> Homo Sapiens

<400> 11
 gggcctgaca gcaacttttc ttctactagt tcctcttaac tttatcctgg taactggcga 60
 gacaacctgt cttaagtaac tgaagggaac 90

<210> 12
 <211> 90
 10 <212> DNA
 <213> Homo Sapiens

<400> 12
 gggcctgaca gcaacttttc ttctactagt tcctcttaac aactgctct gtacggggca 60
 cgtgggcaca ggtgcacact cacactcaca 90

<210> 13
 <211> 90
 15 <212> DNA
 <213> Homo Sapiens

<400> 13
 tccactgac gcatgcagga aggggcacct ccccttaacc aactgctct gtacggggca 60
 cgtgggcaca ggtgcacact cacactcaca 90

<210> 14
 <211> 2499

ES 2 556 332 T3

<212> DNA

<213> HOMO SAPIENS

<400> 14

ctcaatcagc tttatgcaga gaagaagctt actgagctca ctgctggtgc tgggtgtaggc	60
aagtgctgct ttggcaatct gggctgaact ggcttgtctc ctcagaactc cttctccaac	120
cctggagcag gcttccatgc tgctgtgggc gtcccttgctg gcctttgctc cagtctgtgg	180
acaatctgca gctgcacaca aacctgtgat ttccgtccat cctccatgga ccacattctt	240
caaaggagag agagtgaact tgacttgcaa tggatttcag ttctatgcaa cagagaaaac	300
aacatggtat catcggaact actggggaga aaagttgacc ctgaccccag gaaacacctt	360
cgaggttcgg gaatctggac tgtacagatg ccaggcccgg ggctccccac gaagtaaccc	420
tgtgogcttg ctcttttctt cagactcctt aatcctgcag gcaccatatt ctgtgtttga	480
aggtgacaca ttggttctga gatgccacag aagaaggaaa gagaaattga ctgctgtgaa	540
atatacttgg aatggaaaca ttctttccat ttctaataaa agctgggacg ttctttatccc	600
acaagcaagt tcaaataaca atggcaatta tcgatgcatt ggatatggag atgagaatga	660
tgtattttaga tcaaatttca aaataattaa aattcaagaa ctatttccac atccagagct	720
gaaagctaca gactctcagc ctacagaggg gaattctgta aacctgagct gtgaaacaca	780

ES 2 556 332 T3

```

gcttctctcca gagcggtcag acaccccact tcacttcaac ttcttcagag atggcgaggt      840
catcctgtca gactggagca cgtaccogga actccagctc ccaaccgtct ggagagaaaa      900
ctcaggatcc tattggtgtg gtgtgaaac agtgaggggt aacatccaca agcacagtcc      960
ctcgctacag atccatgtgc agcggatccc tgtgtctggg gtgctcctgg agacccagcc     1020
ctcagggggc caggctgttg aaggggagat gctggtcctt gtctgctccg tggtgaagg      1080
cacaggggat accacattct cctggcaccc agaggacatg caggagagtc tggggaggaa      1140
aactcagcgt tccttgagag cagagctgga gctccctgcc atcagacaga gccatgcagg      1200
gggatactac tgtacagcag acaacagcta cggccctgtc cagagcatgg tgctgaatgt      1260
cactgtgaga gagaccccag gcaacagaga tggccttgtc gccgcgggag ccactggagg      1320
gctgctcagt gctctctctc tggtgtgtgc cctgctgttt cactgctggc gtcggaggaa      1380
gtcaggagtt ggtttcttgg gagacgaaac caggctccct ccgctccag gcccaggaga      1440
gtcctcccat tccatctgcc ctgccagggt ggagcttcag tcgttgtatg ttgatgtaca      1500
ccccaaaaag ggagatttgg tatactctga gatccagact actcagctgg gagaagaaga      1560
ggaagcta atccccagga cacttctaga ggataaggat gtctcagttg tctactctga      1620
ggtaagaca caacaccocag ataactcage tggaaagatc agctctaagg atgaagaaag      1680
ttaagagaat gaaaagttac gggaaagtc tactcatgtg atttctccct tgtccaaagt      1740
cccaggccca gtgcagtcct tgcggcacct ggaatgatca actcattcca gctttctaat      1800
tcttctcatg catatgcatt cactccaggt aatactcatt cgtctactct gatgttgagg      1860
tggaatggcc tctgaaagac ttcactaaaa tgaccaggat ccacagttaa gagaagacc      1920
tgtagtatct gctgtgggcc tgacctaatg cattccctag ggtctgcttt agagaagggg      1980
gataaagaga gagaaggact gttatgaaaa acagaagcac aaattttggt gaattgggat      2040
ttgcagagat gaaaaagact ggggtgacct gatctctgct taatacatct acaaccattg      2100
tctcactgga gactcacttg catcagtttg ttttaactgt agtggctgca caggcactgt      2160
gcaaaacatg aaaagccctt tcacttctgc ctgcacagct tacactgtca ggattcagtt      2220
gcagattaaa gaaccatctt ggaatggttt acagagagag gaatttataa gaggacatca      2280
gaagagctgg agatgcaagc tctaggtgtc gcttccaaaa gcaaatgata attatgttaa      2340
tgtcattagt gacaaagatt tgcaacatta gagaaaagag acacaaatat aaaattaaaa      2400
acttaagtac caactctcca aaactaaatt tgaacttaaa atattagtat aaactcataa      2460
taaaactctg ctttaaataa aaaaaaaaaa aaaaaaaaaa      2499

```

<210> 15

<211> 515

<212> PRT

<213> Homo Sapiens

5

ES 2 556 332 T3

<400> 15

Met Leu Leu Trp Ala Ser Leu Leu Ala Phe Ala Pro Val Cys Gly Gln
1 5 10 15

Ser Ala Ala Ala His Lys Pro Val Ile Ser Val His Pro Pro Trp Thr
20 25 30

Thr Phe Phe Lys Gly Glu Arg Val Thr Leu Thr Cys Asn Gly Phe Gln
35 40 45

Phe Tyr Ala Thr Glu Lys Thr Thr Trp Tyr His Arg His Tyr Trp Gly
50 55 60

Glu Lys Leu Thr Leu Thr Pro Gly Asn Thr Leu Glu Val Arg Glu Ser
65 70 75 80

Gly Leu Tyr Arg Cys Gln Ala Arg Gly Ser Pro Arg Ser Asn Pro Val
85 90 95

Arg Leu Leu Phe Ser Ser Asp Ser Leu Ile Leu Gln Ala Pro Tyr Ser
100 105 110

Val Phe Glu Gly Asp Thr Leu Val Leu Arg Cys His Arg Arg Arg Lys
115 120 125

Glu Lys Leu Thr Ala Val Lys Tyr Thr Trp Asn Gly Asn Ile Leu Ser
130 135 140

Ile Ser Asn Lys Ser Trp Asp Leu Leu Ile Pro Gln Ala Ser Ser Asn
145 150 155 160

Asn Asn Gly Asn Tyr Arg Cys Ile Gly Tyr Gly Asp Glu Asn Asp Val
165 170 175

Phe Arg Ser Asn Phe Lys Ile Ile Lys Ile Gln Glu Leu Phe Pro His
180 185 190

Pro Glu Leu Lys Ala Thr Asp Ser Gln Pro Thr Glu Gly Asn Ser Val
195 200 205

ES 2 556 332 T3

Asn Leu Ser Cys Glu Thr Gln Leu Pro Pro Glu Arg Ser Asp Thr Pro
 210 215 220
 Leu His Phe Asn Phe Phe Arg Asp Gly Glu Val Ile Leu Ser Asp Trp
 225 230 235 240
 Ser Thr Tyr Pro Glu Leu Gln Leu Pro Thr Val Trp Arg Glu Asn Ser
 245 250 255
 Gly Ser Tyr Trp Cys Gly Ala Glu Thr Val Arg Gly Asn Ile His Lys
 260 265 270
 His Ser Pro Ser Leu Gln Ile His Val Gln Arg Ile Pro Val Ser Gly
 275 280 285
 Val Leu Leu Glu Thr Gln Pro Ser Gly Gly Gln Ala Val Glu Gly Glu
 290 295 300
 Met Leu Val Leu Val Cys Ser Val Ala Glu Gly Thr Gly Asp Thr Thr
 305 310 315 320
 Phe Ser Trp His Arg Glu Asp Met Gln Glu Ser Leu Gly Arg Lys Thr
 325 330 335
 Gln Arg Ser Leu Arg Ala Glu Leu Glu Leu Pro Ala Ile Arg Gln Ser
 340 345 350
 His Ala Gly Gly Tyr Tyr Cys Thr Ala Asp Asn Ser Tyr Gly Pro Val
 355 360 365
 Gln Ser Met Val Leu Asn Val Thr Val Arg Glu Thr Pro Gly Asn Arg
 370 375 380
 Asp Gly Leu Val Ala Ala Gly Ala Thr Gly Gly Leu Leu Ser Ala Leu
 385 390 395 400
 Ile Leu Ala Val Ala Leu Leu Phe His Cys Trp Arg Arg Arg Lys Ser
 405 410 415
 Gly Val Gly Phe Leu Gly Asp Glu Thr Arg Leu Pro Pro Ala Pro Gly
 420 425 430
 Pro Gly Glu Ser Ser His Ser Ile Cys Pro Ala Gln Val Glu Leu Gln
 435 440 445

ES 2 556 332 T3

Ser Leu Tyr Val Asp Val His Pro Lys Lys Gly Asp Leu Val Tyr Ser
450 455 460

Glu Ile Gln Thr Thr Gln Leu Gly Glu Glu Glu Glu Ala Asn Thr Ser
465 470 475 480

Arg Thr Leu Leu Glu Asp Lys Asp Val Ser Val Val Tyr Ser Glu Val
485 490 495

Lys Thr Gln His Pro Asp Asn Ser Ala Gly Lys Ile Ser Ser Lys Asp
500 505 510

Glu Glu Ser
515

<210> 16

<211> 2805

<212> DNA

5 <213> Homo Sapiens

<400> 16

cggtgcagtg tccctgactgt aagatcaagt ccaaacctgt tttggaattg aggaaaacttc	60
tcttttggatc tcagcccttg gtggtccagg tottcatgct gctgtgggtg atattactgg	120
tccctggctcc tgtcagtggc cagtttgcaa ggacacccag gccattatt ttccctccagc	180
ctccatggac cacagtcttc caaggagaga gaggaccct cacttgcaag ggatttcgct	240
tctactcacc acagaaaaca aaatgggtacc atcggtacct tgggaaagaa atactaagag	300
aaaccccaga caatatcttc gaggttcagg aatctggaga gtacagatgc caggcccagg	360
gctccccctc cagtgccttc gtgcacttgg atttttcttc agcttgcctg atcctgcaag	420
ctccactttc tgtgtttgaa ggagactctg tgggtctgag gtgcgggca aaggcggaag	480
taacactgaa taatactatt tacaagaatg ataatgtcct ggcattcctt aataaaagaa	540
ctgacttcca tattctctcat gcatgtctca aggacaatgg tgcataatgc tgtactggat	600
ataaggaag ttgttgcctt gtttcttcca atacagtcaa aatccaagtc caagagccat	660
ttacacgtcc agtgctgaga gccagctctc tccagcccat cagcggaac ccagtgaacc	720
tgacctgtga gaccagctc tctctagaga ggtcagatgt cccgctccgg ttccgcttct	780
tcagagatga ccagacctg ggattaggct ggagtctctc ccgaatttc cagattactg	840

ES 2 556 332 T3

```

ccatgtggag taaagattca gggttctact ggtgtaaggc agcaacaatg cctcacagcg 900
tcatatctga cagcccgaga tcttgatac aggtgcagat cctgcacat catcctgtcc 960
tcaactctcag cctgaaaag gctctgaatt ttgagggaac caagggtgaca cttcactgtg 1020
aaaccacagga agattctctg cgcactttgt acagggtttta tcatgagggg gtccccctga 1080
ggcacaagtc agtcogctgt gaaaggggag catccatcag cttctcactg actacagaga 1140
attcagggaa ctactactgc acagctgaca atggccttgg cgccaagccc agtaaggctg 1200
tgagcctctc agtcaactgt cccgtgtctc atcctgtctc caacctcago tctcctgagg 1260
aactgatttt tgaggggagcc aagggtgacac ttcactgtga agcccagaga ggttcaactcc 1320
ccatcctgta ccagtttcat catgaggatg ctgccttggg gcttaggtcg gccaaactctg 1380
caggaggagt ggccatcagc ttctctctga ctgcagagca ttcagggaac tactactgca 1440
cagctgacaa tggctttggc ccccagcgca gtaaggcggt gagcctctcc atcactgtcc 1500
ctgtgtctca tctgtctcc accctcagct ctgctgaggc cctgactttt gaaggagcca 1560
ctgtgacact tcaactgtga gtccagagag gttccccaca aatcctatac cagttttatc 1620
atgaggacat gccctgtgg agcagctcaa caccctctgt gggaagagtg tccttcagct 1680
tctctctgac tgaaggacat tcagggaatt actactgcac agctgacaat ggctttggtc 1740
cccagcgagc tgaagtgggt agcctttttg tcaactgttc agtgtctcg cccatctca 1800
ccctcagggc tcccagggcc caggctgtgg tgggggaact gctggagctt cactgtgagg 1860
ccccgagagg ctctccccc atcctgtact ggttttatca tgaggatgtc accctgggga 1920
gcagctcagc cccctctgga ggagaagctt ctttcaacct ctctctgact gcagaacatt 1980
ctggaaaact ctcatgtgag gccacaatg gcctagtggc ccagcacagt gacacaatat 2040
cactcagtgt tatagttcca gtatctcgtc ccatctcac cttcagggct cccagggccc 2100
aggctgtggt gggggacctg ctggagcttc actgtgaggc cctgagaggc tctccccc 2160
tctgttactg gttttatcat gaagatgtca ccctgggtaa gatctcagcc cctctggag 2220
gaggggcctc cttcaacctc tctctgacta cagaacatto tggaatctac tctgtgagg 2280
cagacaatgg tctggaggcc cagcgagtg agatgggtgac actgaaagt ttgcaggtgagt 2340
gggccctgcc caccagcagc acatctgaga actgactgtg cctgtttctc ctgcagctga 2400
aatggagcc acagagctcc tcagggtgt ttgcttgtgt ggcattccag cacacttct 2460
gcctgcagaa cctccctgtg aaagtctcg atcctttgtg gtatggttcc aggaatctga 2520
tgtttccag cagtcttctt gaagatgac aaagcacctc actaaaaatg caaataagac 2580
ttttttagaa cataaactat attctgaact gaaattatta catgaaaatg aaaccaaaga 2640
attctgagca tatgtttctc tgccgtagaa aggattaagc tgtttcttgt ccgattctt 2700
ctctcattga cttctaagaa gcctctactc ttgagtctct ttcattactg gggatgtaaa 2760
tgttcttac atttccacat taaaaatcct atgttaacga aaaaa 2805

```

<210> 17

<211> 759

<212> PRT

<213> Homo Sapiens

ES 2 556 332 T3

<400> 17

Met Leu Leu Trp Val Ile Leu Leu Val Leu Ala Pro Val Ser Gly Gln
1 5 10 15

Phe Ala Arg Thr Pro Arg Pro Ile Ile Phe Leu Gln Pro Pro Trp Thr
20 25 30

Thr Val Phe Gln Gly Glu Arg Val Thr Leu Thr Cys Lys Gly Phe Arg
35 40 45

Phe Tyr Ser Pro Gln Lys Thr Lys Trp Tyr His Arg Tyr Leu Gly Lys
50 55 60

Glu Ile Leu Arg Glu Thr Pro Asp Asn Ile Leu Glu Val Gln Glu Ser
65 70 75 80

Gly Glu Tyr Arg Cys Gln Ala Gln Gly Ser Pro Leu Ser Ser Pro Val
85 90 95

His Leu Asp Phe Ser Ser Ala Ser Leu Ile Leu Gln Ala Pro Leu Ser
100 105 110

Val Phe Glu Gly Asp Ser Val Val Leu Arg Cys Arg Ala Lys Ala Glu
115 120 125

Val Thr Leu Asn Asn Thr Ile Tyr Lys Asn Asp Asn Val Leu Ala Phe
130 135 140

Leu Asn Lys Arg Thr Asp Phe His Ile Pro His Ala Cys Leu Lys Asp
145 150 155 160

Asn Gly Ala Tyr Arg Cys Thr Gly Tyr Lys Glu Ser Cys Cys Pro Val
165 170 175

ES 2 556 332 T3

Ser Ser Asn Thr Val Lys Ile Gln Val Gln Glu Pro Phe Thr Arg Pro
 180 185 190
 Val Leu Arg Ala Ser Ser Phe Gln Pro Ile Ser Gly Asn Pro Val Thr
 195 200 205
 Leu Thr Cys Glu Thr Gln Leu Ser Leu Glu Arg Ser Asp Val Pro Leu
 210 215 220
 Arg Phe Arg Phe Phe Arg Asp Asp Gln Thr Leu Gly Leu Gly Trp Ser
 225 230 235 240
 Leu Ser Pro Asn Phe Gln Ile Thr Ala Met Trp Ser Lys Asp Ser Gly
 245 250 255
 Phe Tyr Trp Cys Lys Ala Ala Thr Met Pro His Ser Val Ile Ser Asp
 260 265 270
 Ser Pro Arg Ser Trp Ile Gln Val Gln Ile Pro Ala Ser His Pro Val
 275 280 285
 Leu Thr Leu Ser Pro Glu Lys Ala Leu Asn Phe Glu Gly Thr Lys Val
 290 295 300
 Thr Leu His Cys Glu Thr Gln Glu Asp Ser Leu Arg Thr Leu Tyr Arg
 305 310 315 320
 Phe Tyr His Glu Gly Val Pro Leu Arg His Lys Ser Val Arg Cys Glu
 325 330 335
 Arg Gly Ala Ser Ile Ser Phe Ser Leu Thr Thr Glu Asn Ser Gly Asn
 340 345 350
 Tyr Tyr Cys Thr Ala Asp Asn Gly Leu Gly Ala Lys Pro Ser Lys Ala
 355 360 365
 Val Ser Leu Ser Val Thr Val Pro Val Ser His Pro Val Leu Asn Leu
 370 375 380
 Ser Ser Pro Glu Asp Leu Ile Phe Glu Gly Ala Lys Val Thr Leu His
 385 390 395 400
 Cys Glu Ala Gln Arg Gly Ser Leu Pro Ile Leu Tyr Gln Phe His His
 405 410 415

ES 2 556 332 T3

Glu Asp Ala Ala Leu Glu Arg Arg Ser Ala Asn Ser Ala Gly Gly Val
420 425 430

Ala Ile Ser Phe Ser Leu Thr Ala Glu His Ser Gly Asn Tyr Tyr Cys
435 440 445

Ala Thr Asp Asn Gly Phe Gly Pro Gln Arg Ser Lys Ala Val Ser Leu
450 455 460

Ser Ile Thr Val Pro Val Ser His Pro Val Leu Thr Leu Ser Ser Ala
465 470 475 480

Glu Ala Leu Thr Phe Glu Gly Ala Thr Val Thr Leu His Cys Glu Val
485 490 495

Gln Arg Gly Ser Pro Gln Ile Leu Tyr Gln Phe Tyr His Glu Asp Met
500 505 510

Pro Leu Trp Ser Ser Ser Thr Pro Ser Val Gly Arg Val Ser Phe Ser
515 520 525

Phe Ser Leu Thr Glu Gly His Ser Gly Asn Tyr Tyr Cys Thr Ala Asp
530 535 540

Asn Gly Phe Gly Pro Gln Arg Ser Glu Val Val Ser Leu Phe Val Thr
545 550 555 560

Val Pro Val Ser Arg Pro Ile Leu Thr Leu Arg Val Pro Arg Ala Gln
565 570 575

Ala Val Val Gly Asp Leu Leu Glu Leu His Cys Glu Ala Pro Arg Gly
580 585 590

Ser Pro Pro Ile Leu Tyr Trp Phe Tyr His Glu Asp Val Thr Leu Gly
595 600 605

Ser Ser Ser Ala Pro Ser Gly Gly Glu Ala Ser Phe Asn Leu Ser Leu
610 615 620

Thr Ala Glu His Ser Gly Asn Tyr Ser Cys Glu Ala Asn Asn Gly Leu
625 630 635 640

Val Ala Gln His Ser Asp Thr Ile Ser Leu Ser Val Ile Val Pro Val
645 650 655

ES 2 556 332 T3

Ser Arg Pro Ile Leu Thr Phe Arg Ala Pro Arg Ala Gln Ala Val Val
660 665 670

Gly Asp Leu Leu Glu Leu His Cys Glu Ala Leu Arg Gly Ser Ser Pro
675 680 685

Ile Leu Tyr Trp Phe Tyr His Glu Asp Val Thr Leu Gly Lys Ile Ser
690 695 700

Ala Pro Ser Gly Gly Gly Ala Ser Phe Asn Leu Ser Leu Thr Thr Glu
705 710 715 720

His Ser Gly Ile Tyr Ser Cys Glu Ala Asp Asn Gly Leu Glu Ala Gln
725 730 735

Arg Ser Glu Met Val Thr Leu Lys Val Ala Gly Glu Trp Ala Leu Pro
740 745 750

Thr Ser Ser Thr Ser Glu Asn
755

<210> 18

<211> 4448

<212> DNA

5 <213> Homo Sapiens

<400> 18

```

cgggtgcagtg tcttgactgt aagatcaagt ccaaacctgt ttggaattg aggaaacttc      60
tcttttgatc tcagcccttg gtggtccagg tcttcatgct gctgtgggtg atattactgg      120
tcttggctcc tgtcagtgga cagtttgcaa ggacaccag gccattatt ttctccagc      180
ctccatggac cacagtcttc caaggagaga gagtgaacct cacttgcaag ggatttcgct      240
tctactcacc acagaaaaca aaatggtaac atcggtacct tgggaaagaa atactaagag      300
aaaccccaga caatatcctt gaggttcagg aatctggaga gtacagatgc caggcccagg      360
gtcctccctct cagtagccct gtgcacttgg atttttcttc agcttcgctg atcctgcaag      420
ctccactttc tgtgtttgaa ggagactctg tggttctgag gtgcggggca aaggcggaag      480
taacactgaa taatactatt tacaagaatg ataatgtcct gccattcctt aataaaagaa      540
ctgacttcca ttttcctcat gcatgtctca aggacaatgg tgcatacgc tgtactggat      600

```

ES 2 556 332 T3

ataaggaaag ttgttgcctt gtttcttcca atacagtcaa aatccaagtc caagagccat	660
ttacacgtcc agtgctgaga gccagctcct tccagcccat cagcgggaac ccagtgaccc	720
tgacctgtga gaccagctc tctctagaga ggtcagatgt cccgctccgg ttccgcttct	780
tcagagatga ccagaccctg ggattaggtt ggagtctctc ccgaatttc cagattactg	840
ccatgtggag taaagattca gggttctact ggtgtaaggc agcaacaatg cctcacagcg	900
tcatacttga cagcccagaga tcttgatag aggtgcagat ccttgcatct cactctgtcc	960
tcaactctcag cctgaaaag gctctgaatt ttgagggaac caaggtgaca cttcactgtg	1020
aaaccagga agattctctg cgcactttgt acaggtttta tcatgagggt gtccccctga	1080
ggcacaagtc agtccgtgt gaaaggggag catccatcag cttctcactg actacagaga	1140
attcagggaa ctactactgc acagctgaca atggccttgg cgcacaagccc agtaaggctg	1200
tgagcctctc agtcaactgtt cccgtgtctc atcctgtcct caacctcagc tctcctgagg	1260
acctgatttt tgaggagacc aaggtgacac ttcactgtga agcccagaga ggttcactcc	1320
ccatcctgta ccagtttcat catgaggatg ctgccttga gcgtaggctg gccactctg	1380
caggaggagt ggccatcagc ttctctctga ctgcagagca ttcagggaac tactactgca	1440
cagctgacaa tggctttggc cccagcgca gtaaggcgtt gagcctctcc atcactgtcc	1500
ctgtgtctca tctgtctctc acctcagct ctgctgaggg cctgactttt gaaggagcca	1560
ctgtgacact tcaacttgaa gtccagagag gttccccaca aatcctatac cagttttatc	1620
atgaggacat gcccctgtgg agcagctcaa caccctctgt gggaagagtg tccctcagct	1680
tctctctgac tgaaggacat tcagggaatt actactgcac agctgacaat ggctttggtc	1740
cccagcgag tgaagtgggt agcctttttg tcaactgtaa gtgtgggtt cttgccagtc	1800
accacccctt ggctgagttc tctctcacc attcctttta aaatctgttt gcaactgtcc	1860
gtttcctccc ctaatcaact taatccctt cttggttcc tctcaacta actagctggg	1920
gttttccgta ctcataagtc ctggtcagc cagacccta aaacagctca gtagattccc	1980
cagcttttac caaatgaatt tttttattgt atttctctc cattccttgt atgttccaa	2040
agtaagccaa tttttcttga tgcaaggagc gtgtcctact tctctactga catttacata	2100
ttaacttagc tacaagcaca gtcttataga taaatattgg tcaagacctt aaattctcca	2160
aaggatttcc aatcttatgg tagatttgga gaaagctgct ggtgaacaaa gggggaaatg	2220
gtcctcagg aaccaactcc tcaaacctct ggagttttta tgatcccttg ttttctaacc	2280
tgctaaaatc agtatcattt tattgtatta ttttaaaaa actattgttg aagtatgaca	2340
tacattcaag aaacgtgtgc aaattgtatg tgaacgattt ggtgtctttt taggagctaa	2400

ES 2 556 332 T3

gttgcttctg tttttacttg aatctttgtt tatagaaact gggggaaagt ttactttctt 2460
 ttcagagaag ccaaattgta tgatagaaaa atcttgagcc tgatgtgtca gacatgcccc 2520
 tagcataact tgttgagtaa agaggttatt tttaaaatgt gaatgtttctg agactactcc 2580
 aaagtcagag ccaaattctac taggaagctt ctagacttca ctcattctgc atcccattac 2640
 tatcttttta tccatgtttt actttcttct catattcagc agcatcttaa gctcttttat 2700
 tttctgtttc ttgactgtca ccttaattgc cagtagaatg taagcttcat gagaacagaa 2760
 ctgcctccat cttgggtctc acaacatccc tgtgctact cagtgtttgg cacacagtag 2820
 gtcctcagtc aacattttgta atttagtgga cagatgatat gacaagatga taagagggga 2880
 tttaaaaaaa tcatctagca aagcccaaga ggaaaaaaa caaagctatt ttagaaatga 2940
 aataccaatt tgaagcagta agaatagatt ggatatcttt gaaaaccatt aattgaatga 3000
 agaaccaatt tgagaaaaca atacagaatg caaagtagaa agatacagaa ataaaggcaa 3060
 aagttataat atggaaaatca gacaatggat ttgtctgtat ccagttatgt ggataattaa 3120
 aatggagacc ctcagaaaat tgaaccgaag agtaaaatga aactcaaaaa tgtagtagaa 3180
 attgttggga agtaaaagaaa acttgaatat gtatgcaga acatatatgt tgatgacgtt 3240
 attgactttg aggttaaaaa tatatatatg tgctatgat tatggggaaa aaagcagtcg 3300
 tctcagaaaag aaaaacatca agttagtctt agacttttga gtgcactcag taccaaagag 3360
 agaggaggcc agacttgga ctcgagggga agaataataa ccgaaaattt tatatcaatt 3420
 caaaaagaca ttgtcaaaaa tacagggatt caggaaaactg agaatgcact aagccttctg 3480
 gaaaaaacac ctaatgacaa aatctagccc aacaagatgt aaatgaatat aaaggactca 3540
 taatgaggaa accgcattat gactggctct caaccctggc cgcattatag actcgtcaaa 3600
 gaccttttga aaaggtcaca cattgactcg tcaaagcccc tctcagact aattcaattc 3660
 agaattctac agatggggcc acagaatcag ttttttttga cacaacctca agtgagaata 3720
 ttgtgtagac aagattggaa accactgatt tagatataga aacaaaggct aatcaactgt 3780
 gagaattatg gtcacagaat agaaagtaac tattatgaac actgaaaatg taaaaaaat 3840
 gtaacaaaga aaaatagtta gaggaaggag aggaagtaaa ggaacaatca ttttctcatg 3900
 attattatta tttcagagta aattgtgagt tatttcaca ttcaaaaaga atggactgtt 3960
 ttaaaaaatt agtaatagat ttcaaaatgt ccatttttga aatcgtttct gaatactttg 4020
 tcaacagtta ctcacatta atggcttata cttcactaaa attccatgga aaaccaacta 4080
 gtgacctgta gagtcacata ggagagaaca agtgaattct ttgggtggcg caagcataga 4140
 tgttaggact gacaaaaaaa aataataaaa ataaadctgt gcattgatat gatcacaat 4200
 gatcagggaa agaggaaaca gaaactctca taogccatta ttacaagtgt aaattggttc 4260
 aaccttttct tottaattga cacattgtaa ttgtatatat ttatggaagc acagtttgat 4320
 attttgatat acatacatgg tatataacga tcaaatagg atatttaatg taccatcat 4380
 ctcatgcatt tatcatttct ttggaataaa aacattcaaa agccaaaaaa aaaaaaaaaa 4440
 aaaaaaaaaa 4448

5 <210> 19
 <211> 592
 <212> PRT
 <213> Homo Sapiens

ES 2 556 332 T3

<400> 19

Met	Leu	Leu	Trp	Val	Ile	Leu	Leu	Val	Leu	Ala	Pro	Val	Ser	Gly	Gln
1				5					10					15	

Phe Ala Arg Thr Pro Arg Pro Ile Ile Phe Leu Gln Pro Pro Trp Thr
20 25 30

Thr Val Phe Gln Gly Glu Arg Val Thr Leu Thr Cys Lys Gly Phe Arg
35 40 45

Phe Tyr Ser Pro Gln Lys Thr Lys Trp Tyr His Arg Tyr Leu Gly Lys
50 55 60

Glu Ile Leu Arg Glu Thr Pro Asp Asn Ile Leu Glu Val Gln Glu Ser
65 70 75 80

Gly Glu Tyr Arg Cys Gln Ala Gln Gly Ser Pro Leu Ser Ser Pro Val
85 90 95

His Leu Asp Phe Ser Ser Ala Ser Leu Ile Leu Gln Ala Pro Leu Ser
100 105 110

Val Phe Glu Gly Asp Ser Val Val Leu Arg Cys Arg Ala Lys Ala Glu
115 120 125

Val	Thr	Leu	Asn	Asn	Thr	Ile	Tyr	Lys	Asn	Asp	Asn	Val	Leu	Ala	Phe
130						135					140				

Leu Asn Lys Arg Thr Asp Phe His Ile Pro His Ala Cys Leu Lys Asp
145 150 155 160

ES 2 556 332 T3

Asn Gly Ala Tyr Arg Cys Thr Gly Tyr Lys Glu Ser Cys Cys Pro Val
 165 170 175

Ser Ser Asn Thr Val Lys Ile Gln Val Gln Glu Pro Phe Thr Arg Pro
 180 185 190

Val Leu Arg Ala Ser Ser Phe Gln Pro Ile Ser Gly Asn Pro Val Thr
 195 200 205

Leu Thr Cys Glu Thr Gln Leu Ser Leu Glu Arg Ser Asp Val Pro Leu
 210 215 220

Arg Phe Arg Phe Phe Arg Asp Asp Gln Thr Leu Gly Leu Gly Trp Ser
 225 230 235 240

Leu Ser Pro Asn Phe Gln Ile Thr Ala Met Trp Ser Lys Asp Ser Gly
 245 250 255

Phe Tyr Trp Cys Lys Ala Ala Thr Met Pro His Ser Val Ile Ser Asp
 260 265 270

Ser Pro Arg Ser Trp Ile Gln Val Gln Ile Pro Ala Ser His Pro Val
 275 280 285

Leu Thr Leu Ser Pro Glu Lys Ala Leu Asn Phe Glu Gly Thr Lys Val
 290 295 300

Thr Leu His Cys Glu Thr Gln Glu Asp Ser Leu Arg Thr Leu Tyr Arg
 305 310 315 320

Phe Tyr His Glu Gly Val Pro Leu Arg His Lys Ser Val Arg Cys Glu
 325 330 335

Arg Gly Ala Ser Ile Ser Phe Ser Leu Thr Thr Glu Asn Ser Gly Asn
 340 345 350

Tyr Tyr Cys Thr Ala Asp Asn Gly Leu Gly Ala Lys Pro Ser Lys Ala
 355 360 365

Val Ser Leu Ser Val Thr Val Pro Val Ser His Pro Val Leu Asn Leu
 370 375 380

Ser Ser Pro Glu Asp Leu Ile Phe Glu Gly Ala Lys Val Thr Leu His
 385 390 395 400

ES 2 556 332 T3

Cys Glu Ala Gln Arg Gly Ser Leu Pro Ile Leu Tyr Gln Phe His His
405 410 415

Glu Asp Ala Ala Leu Glu Arg Arg Ser Ala Asn Ser Ala Gly Gly Val
420 425 430

Ala Ile Ser Phe Ser Leu Thr Ala Glu His Ser Gly Asn Tyr Tyr Cys
435 440 445

Thr Ala Asp Asn Gly Phe Gly Pro Gln Arg Ser Lys Ala Val Ser Leu
450 455 460

Ser Ile Thr Val Pro Val Ser His Pro Val Leu Thr Leu Ser Ser Ala
465 470 475 480

Glu Ala Leu Thr Phe Glu Gly Ala Thr Val Thr Leu His Cys Glu Val
485 490 495

Gln Arg Gly Ser Pro Gln Ile Leu Tyr Gln Phe Tyr His Glu Asp Met
500 505 510

Pro Leu Val Ser Ser Ser Thr Pro Ser Val Gly Arg Val Ser Phe Ser
515 520 525

Phe Ser Leu Thr Glu Gly His Ser Gly Asn Tyr Tyr Cys Thr Ala Asp
530 535 540

Asn Gly Phe Gly Pro Gln Arg Ser Glu Val Val Ser Leu Phe Val Thr
545 550 555 560

Gly Lys Cys Trp Val Leu Ala Ser Lys Pro Pro Leu Ala Glu Phe Ser
565 570 575

Leu Thr His Ser Phe Lys Asn Leu Phe Ala Leu Ser Ser Phe Leu Pro
580 585 590

<210> 20

<211> 5323

<212> DNA

5 <213> Homo Sapiens

<400> 20

cgggtgcagtg tcttgactgt aagatcaagt ccaaacctgt tttggaattg aggaaacttc 60

ES 2 556 332 T3

tcttttgatc tcagccottg gtggtccagg tcttcatgct gctgtgggtg atattactgg	120
tectggetcc tgtcagtgga cagtttgcaa ggacaccag gccattatt ttctccagc	180
ctccatggac cacagtcttc caaggagaga gagtgacct cacttgcaag ggatttgcgt	240
tctactcacc acagaaaaca aaatggtacc atcggtacct tgggaaagaa atactaagag	300
aaaccccaga caatatcctt gaggttcagg aatctggaga gtacagatgc caggcccagg	360
gctccctct cagttagcct gtgcacttgg atttttcttc agcttgcgtg atctgcaag	420
ctccactttc tgtgtttgaa ggagactctg tggttctgag gtgcgggca aaggcggag	480
taacactgaa taatactatt tacaagaatg ataatgtcct ggcattcctt aataaaagaa	540
ctgacttcca tattctcat goatgtctca aggacaatgg tgcataatgc tgtactggat	600
ataaggaaag ttgttgccct gttttctcca atacagtcaa aatccaagtc caagagccat	660
ttacacgtcc agtgctgaga gccagctcct tcagcccat cagcgggaac ccagtgacc	720
tgaactgtga gaccagctc tctctagaga ggtcagatgt cccgctcgg ttccgtttct	780
tcagagatga ccagacctg ggattaggct ggagtctct cccgaatttc cagattactg	840
ccatgtggag taaagattca gggttctact ggtgtaaggc agcaacaatg cctcacagcg	900
tcatatctga cagcccgaga tcttgatac aggtgcagat cctgcatct catctgtcc	960
tcactctcag cctgaaaag gctctgaatt ttgagggaac caagtgaca cttactgtg	1020
aaaccagga agattctctg cgcactttgt acaggtttta tcatgaggt gtccccctga	1080
ggcacaagtc agtccgtgt gaaaggggag catccatcag cttctcactg actacagaga	1140
attcagggaa ctactactgc acagctgaca atggccttgg cccaagccc agtaaggtg	1200
tgagcctctc agtcactgtt cccgtgtctc atcctgtcct caacctcagc tctctgagg	1260
acctgatttt tgagggagcc aaggtgacac ttcactgtga agcccagaga ggttactcc	1320
ccatctgtga ccagtttcat catgaggatg ctgccctgga gcgtaggctg gccactctg	1380
caggaggagt ggccatcagc ttctctctga ctgcagagca ttcagggaac tactactgca	1440
cagctgacaa tggctttggc ccccagcga gtaaggcgt gagcctctcc atcactgtcc	1500
ctgtgtctca tctgtctct accctcagct ctgtgaggc cctgactttt gaaggagcca	1560
ctgtgacact tcaactgtgaa gtccagagag gttcccaca aatcctatac cagttttatc	1620
atgaggacat gccctgtgg agcagctcaa caccctctgt gggaagagtg tcttccagct	1680
tctctctgac tgaaggacat tcaggaatt actactgcac agctgacaat ggctttggtc	1740
cccagcgcag tgaagtgggt agcctttttg tcaactgtcc agtgtctgc cccatcctca	1800
ccctcagggg tcccagggcc caggctgtgg tgggggacct gctggagctt cactgtgagg	1860

ES 2 556 332 T3

ccccgagagg ctotccccca atcctgtact ggttttatca tgaggatgtc accctgggga	1920
gcagctcagc cccctctgga ggagaagctt ctttcaacct ctctctgact gcagaacatt	1980
ctggaaacta ctcatgtgag gccacaatg gcctagtggc ccagcacagt gacacaatat	2040
cactcagtgt tatagttcca gtatctcgtc ccatactcac cttcagggct cccagggccc	2100
aggctgtggt gggggacctg ctggagcttc actgtgagge cctgagagge tcttcccaa	2160
tctgtactg gttttatcat gaagatgtca ccttgggtaa gatctcagcc cctctggag	2220
gaggggctc cttcaacctc tctctgacta cagaacatto tggaatctac tctgtgagg	2280
cagacaatgg tctggaggcc cagcgcagtg agatggtgac actgaaagt gcagttcagg	2340
tgtctcgccc ggtctcacc ctcagggctc cggggaccca tgctgcggtg ggggacctgc	2400
tggagcttca ctgtgaggcc ctgagaggct ctccctgat cctgtaccgg tttttcatg	2460
aggatgtcac cctaggaat aggtcgtccc cctctggagg agcgtcctta aacctctctc	2520
tgactgcaga gcactctgga aactactcct gtgaggccga caatggcctc ggggccagc	2580
gcagtgagac agtgacaatt tatatcacag ggctgaccgc gaacagaagt ggcccttttg	2640
ccacaggagt cgcggggggc ctgctcagca tagcaggcct tgctgcgggg gcactgctgc	2700
tctactgctg gctctcgaga aaagcaggga gaaagcctgc ctctgacccc gccaggagcc	2760
cttcagactc ggactcccaa gagcccact atcacaaatgt accagcctgg gaagagctgc	2820
aaccagtgtc cactaatgca aatcctagag gaaaaaatgt ggtttactca gaagtaagg	2880
tcattccaaga gaaaaagaaa catgcagtgg cctctgaccc caggcatctc aggaacaagg	2940
gttcccctat catctactct gaagttaagg tggcgtcaac ccgggtttcc ggatccctgt	3000
tcttggttc ctcagctcct cacagatgag tccacagtc tctccaaactg ctgtttcagc	3060
ctctgcaccc caaagtccc cttgggggag aagcagcatt gaagtgggaa gatttaggct	3120
gccccagacc atatctactg gcctttgttt cacatgtcct cattctcagt ctgaccagaa	3180
tgcagggcc tgctggactg tcacctgttt ccagttaaa gcctgactg gcaggttttt	3240
taatccagtg gcaagggtgt cccactccag ggcccagcac atctcctgga ttcttagtg	3300
ggcttcagct gtggttgctg ttctgagtae tgctctcate acacccccac agagggggtc	3360
ttaccacaca aaggagagt gggccttcag gagatgcgg gctggcctaa cagctcaggt	3420
gctcctaaac tccgacacag agttcctgct ttgggtggat gcatttctca attgtcatca	3480
gcctggtggg gctactgcag tgtgctgcca aatgggacag cacacagcct gtgcacatgg	3540
gacatgtgat gggctcctcc acgggggctg catttcacac tctccacct gtctcaaact	3600
ctaaggtcgg cacttgacac caagtaact tctctcctgc tcatgtgtca gtgtctacct	3660

ES 2 556 332 T3

gcccaagtaa gtggctttca tacaccaagt cccgaagtgc ttcccatcct aacagaagta 3720
 acccagcaag tcaaggccag gaggaccag ggtgcagaca gaacacatac tggaaacacag 3780
 gaggtgctca attactatct gactgactga ctgaatgaat gaatgaatga ggaagaaaac 3840
 tgtgggtaat caaactggca taaaatccag tgcactccct aggaaatccg ggaggtattc 3900
 tggcttccta agaaacaacg gaagagaagg agcttggatg aagaaactgt tcagcaagaa 3960
 gaagggtctc ttcacacttt tatgtgcttg tggatcacct gaggatctgt gaaaatacag 4020
 atactgattc agtgggtctg tgtagagcct gagactgcca ttctaactg ttcccagggg 4080
 atgctgatgc tgcctggcct gggactgcac tgcctgcatg tgaagcccta taggtctcag 4140
 cagaggccca tggagaggga atgtgtggct ctggctgccc agggcccaac tcgggttcaca 4200
 cggatcgtgc tgcctccctg ccagccttg gccacagcac caccagctgc tgttgcctgag 4260
 agaggttctt ctctgtgaca tgttggcttt catcagccac cctgggaagc gaaagtagc 4320
 tggcactatc tttgtttccc cactcaggc ctacacttt cccatgaaa ggtgaatgt 4380
 atataacctg agccctctcc attcagagtt gttctcccat ctctgagca tgggatgttc 4440
 tgttcgctt ttatgatatc catcacatct tatcttgatc tttgctccca gtggattgta 4500
 cagtgatgac ttttaagccc caggccctg aaataaaaac ctccaaggc cattggaagc 4560
 tcaactccac tgaacctggt cttttcatgc ttccaagtgt cagggccttg ccagataga 4620
 cagggtgac tctgctgccc caacctttca aggaggaaac cagacacctg agacaggagc 4680
 ctgtatgcag cccagtgcag ccttgacag gacaaggctg gaggcatttg tcatcactac 4740
 agatatgcaa ctaaaataga cgtggagcaa gagaaatgca ttcccaccga ggcgctttt 4800
 ttaggcctag ttgaaagtca agaaggacag cagcaagcat aggotcagga ttaaagaaaa 4860
 aaatctgctc acagtctgtt ctggaggta catcaccaac aaagctcacg ccctatgcag 4920
 ttctgagaag gtggaggcac caggctcaaa agaggaaatt tagaatttct cattgggaga 4980
 gtaaggtaac cccatcccag aatgataact gcacagtggc agaacaact ccaacctaat 5040
 gtgggtggac cccatccagt ctgttgaagg cctgaatgta acaaaagggc ttattcttcc 5100
 tcaagtaagg gggaactcct gctttgggt gggacataag tttttctgct ttcagacgca 5160
 aactgaaaaa tggctcttct tgggtcttga gcttgctggc atatggactg aaagaaacta 5220
 tgcatttga tctctggat ctccagcttg ctgactgcag atcttgagat atgtcagcct 5280
 ctacagtcac aagagctaatt tcattctaatt aaaccaatct ttc 5323

<210> 21

<211> 977

<212> PRT

<213> Homo sapiens

5

ES 2 556 332 T3

<400> 21

Met Leu Leu Trp Val Ile Leu Leu Val Leu Ala Pro Tyr Ser Gly Gln
1 5 10 15

Phe Ala Arg Thr Pro Arg Pro Ile Ile Phe Leu Gln Pro Pro Trp Thr
20 25 30

Thr Val Phe Gln Gly Glu Arg Val Thr Leu Thr Cys Lys Gly Phe Arg
35 40 45

Phe Tyr Ser Pro Gln Lys Thr Lys Trp Tyr His Arg Tyr Leu Gly Lys
50 55 60

Glu Ile Leu Arg Glu Thr Pro Asp Asn Ile Leu Glu Val Gln Glu Ser
65 70 75 80

Gly Ser Tyr Arg Cys Gln Ala Gln Gly Ser Pro Leu Ser Ser Pro Val
85 90 95

His Leu Asp Phe Ser Ser Ala Ser Leu Ile Leu Gln Ala Pro Leu Ser
100 105 110

Val Phe Glu Gly Asp Ser Val Val Leu Arg Cys Arg Ala Lys Ala Glu
115 120 125

Val Thr Leu Asn Asn Thr Ile Tyr Lys Asn Asp Asn Val Leu Ala Phe
130 135 140

Leu Asn Lys Arg Thr Asp Phe His Ile Pro His Ala Cys Leu Lys Asp
145 150 155 160

Asn Gly Ala Tyr Arg Cys Thr Gly Tyr Lys Glu Ser Cys Cys Pro Val
165 170 175

Ser Ser Asn Thr Val Lys Ile Gln Val Gln Glu Pro Phe Thr Arg Pro
180 185 190

Val Leu Arg Ala Ser Ser Phe Gln Pro Ile Ser Gly Asn Pro Val Thr
195 200 205

ES 2 556 332 T3

Leu Thr Cys Glu Thr Gln Leu Ser Leu Glu Arg Ser Asp Val Pro Leu
 210 215 220
 Arg Phe Arg Phe Phe Arg Asp Asp Gln Thr Leu Gly Leu Gly Trp Ser
 225 230 235 240
 Leu Ser Pro Asn Phe Gln Ile Thr Ala Met Trp Ser Lys Asp Ser Gly
 245 250 255
 Phe Tyr Trp Cys Lys Ala Ala Thr Met Pro His Ser Val Ile Ser Asp
 260 265 270
 Ser Pro Arg Ser Trp Ile Gln Val Gln Ile Pro Ala Ser His Pro Val
 275 280 285
 Leu Thr Leu Ser Pro Glu Lys Ala Leu Asn Phe Glu Gly Thr Lys Val
 290 295 300
 Thr Leu His Cys Glu Thr Gln Glu Asp Ser Leu Arg Thr Leu Tyr Arg
 305 310 315 320
 Phe Tyr His Glu Gly Val Pro Leu Arg His Lys Ser Val Arg Cys Glu
 325 330 335
 Arg Gly Ala Ser Ile Ser Phe Ser Leu Thr Thr Glu Asn Ser Gly Asn
 340 345 350
 Tyr Tyr Cys Thr Ala Asp Asn Gly Leu Gly Ala Lys Pro Ser Lys Ala
 355 360 365
 Val Ser Leu Ser Val Thr Val Pro Val Ser His Pro Val Leu Asn Leu
 370 375 380
 Ser Ser Pro Glu Asp Leu Ile Phe Glu Gly Ala Lys Val Thr Leu His
 385 390 395 400
 Cys Glu Ala Gln Arg Gly Ser Leu Pro Ile Leu Tyr Gln Phe His His
 405 410 415
 Glu Asp Ala Ala Leu Glu Arg Arg Ser Ala Asn Ser Ala Gly Gly Val
 420 425 430
 Ala Ile Ser Phe Ser Leu Thr Ala Glu His Ser Gly Asn Tyr Tyr Cys
 435 440 445

ES 2 556 332 T3

Thr Ala Asp Asn Gly Phe Gly Pro Gln Arg Ser Lys Ala Val Ser Leu
 450 455 460
 Ser Ile Thr Val Pro Val Ser His Pro Val Leu Thr Leu Ser Ser Ala
 465 470 475 480
 Glu Ala Leu Thr Phe Glu Gly Ala Thr Val Thr Leu His Cys Glu Val
 485 490 495
 Gln Arg Gly Ser Pro Gln Ile Leu Tyr Gln Phe Tyr His Glu Asp Met
 500 505 510
 Pro Leu Trp Ser Ser Ser Thr Pro Ser Val Gly Arg Val Ser Phe Ser
 515 520 525
 Phe Ser Leu Thr Glu Gly His Ser Gly Asn Tyr Tyr Cys Thr Ala Asp
 530 535 540
 Asn Gly Phe Gly Pro Gln Arg Ser Glu Val Val Ser Leu Phe Val Thr
 545 550 555 560
 Val Pro Val Ser Arg Pro Ile Leu Thr Leu Arg Val Pro Arg Ala Gln
 565 570 575
 Ala Val Val Gly Asp Leu Leu Glu Leu His Cys Glu Ala Pro Arg Gly
 580 585 590
 Ser Pro Pro Ile Leu Tyr Trp Phe Tyr His Glu Asp Val Thr Leu Gly
 595 600 605
 Ser Ser Ser Ala Pro Ser Gly Gly Glu Ala Ser Phe Asn Leu Ser Leu
 610 615 620
 Thr Ala Glu His Ser Gly Asn Tyr Ser Cys Glu Ala Asn Asn Gly Leu
 625 630 635 640
 Val Ala Gln His Ser Asp Thr Ile Ser Leu Ser Val Ile Val Pro Val
 645 650 655
 Ser Arg Pro Ile Leu Thr Phe Arg Ala Pro Arg Ala Gln Ala Val Val
 660 665 670
 Gly Asp Leu Leu Glu Leu His Cys Glu Ala Leu Arg Gly Ser Ser Pro
 675 680 685

ES 2 556 332 T3

Ile Leu Tyr Trp Phe Tyr His Glu Asp Val Thr Leu Gly Lys Ile Ser
690 695 700

Ala Pro Ser Gly Gly Gly Ala Ser Phe Asn Leu Ser Leu Thr Thr Glu
705 710 715 720

His Ser Gly Ile Tyr Ser Cys Glu Ala Asp Asn Gly Leu Glu Ala Gln
725 730 735

Arg Ser Glu Met Val Thr Leu Lys Val Ala Val Pro Val Ser Arg Pro
740 745 750

Val Leu Thr Leu Arg Ala Pro Gly Thr His Ala Ala Val Gly Asp Leu
755 760 765

Leu Thr Glu Leu His Cys Glu Ala Leu Arg Gly Ser Pro Leu Ile Leu
770 775 780

Tyr Arg Phe Phe His Glu Asp Val Thr Leu Gly Asn Glu Leu His Cys
785 790 795 800

Glu Ala Leu Arg Gly Ser Pro Leu Ile Leu Tyr Arg Phe Phe His Glu
805 810 815

Asp Val Thr Leu Gly Asn Asn Gly Leu Gly Ala Gln Arg Ser Glu Thr
820 825 830

Val Thr Leu Tyr Ile Thr Gly Leu Thr Ala Asn Arg Ser Gly Pro Phe
835 840 845

Ala Thr Gly Val Ala Gly Gly Leu Leu Ser Ile Ala Gly Leu Ala Ala
850 855 860

Gly Ala Leu Leu Leu Tyr Cys Trp Leu Ser Arg Lys Ala Gly Arg Lys
865 870 875 880

Pro Ala Ser Asp Pro Ala Arg Ser Pro Ser Asp Ser Asp Ser Gln Glu
885 890 895

Pro Thr Tyr His Met Val Pro Ala Trp Glu Glu Leu Gln Pro Val Tyr
900 905 910

Thr Asn Ala Asn Pro Arg Gly Glu Asn Val Val Tyr Ser Glu Val Arg
915 920 925

Ile Ile Gln Glu Lys Lys Lys His Ala Val Ala Ser Asp Pro Arg His
930 935 940

Leu Arg Asn Lys Gly Ser Pro Ile Ile Tyr Ser Glu Val Lys Val Ala
945 950 955 960

Ser Thr Pro Val Ser Gly Ser Leu Phe Leu Ala Ser Ser Ala Pro His
965 970 975

Arg

ES 2 556 332 T3

<210> 22
 <211> 88
 <212> PRT
 <213> Homo Sapiens

5 <400> 22
 Met Leu Leu Trp Ala Ser Leu Leu Ala Phe Ala Pro Val Cys Gly Gln
 1 5 10 15
 Ser Gly Ser Cys Ser Val Ala Asp Trp Gln Met Pro Pro Pro Tyr Val
 20 25 30
 Val Leu Asp Leu Pro Gln Glu Thr Leu Glu Glu Glu Thr Pro Gly Ala
 35 40 45
 Asn Leu Trp Pro Thr Thr Ile Thr Phe Leu Thr Leu Phe Leu Leu Ser
 50 55 60
 Leu Phe Tyr Ser Thr Ala Leu Thr Val Thr Ser Val Arg Gly Pro Ser
 65 70 75 80
 Gly Asn Arg Glu Gly Pro Gln Tyr
 85

<210> 23
 <211> 837
 <212> DNA
 <213> Homo Sapiens

10 <400> 23
 ctcaatcagc tttatgcaga gaagaagcct actgagctca ctgctggtgc tgggtgtaggc 60
 aagtgcctgc ttggcaatct gggctgacct ggcttgcctc ctcagaactc cttctccaac 120
 cctggagcag gcttccatgc tgcctgtggc gtccttgcctg gccttgcctc cagtctgtgg 180
 acaatctggc tcttgcctct ttgcagattg gcagatgccg cctccctatg tgggtgctgga 240
 cttgcccagc gagaccctgg agggaggagc ccccgccgcc aacctgtggc ccaccaccat 300
 caccttctc accctcttcc tgcctgagcct gttctatagc acagcactga ccgtgaccag 360
 cgtccggggc ccatctggca acaggagggg ccccagtagc tgagcgggag ccggcaaggc 420
 acagggtggga gccacggagg gggatgagcc cacagtggat gaggtgggct gcagtgcctg 480
 gctaagagga gagcaccacc tgctccact gtggggggac gtgctctctc ggggggccct 540
 tcacagacac tgaggacacg cgcaggccca gggtcagggc tgagcttccc tcagtgcag 600
 taacgaggat tccgtccagg ctcccatgag caggccaggg ctgagacaga gggcgttggc 660
 aaggatgctg ctcttcaggc tgtgacctc ctgtctttgc agggaggaag tgtggaggaa 720
 cctcttgagg aagccagcta tgcctgccag aactcagccc ttccagacgt caccgacctg 780
 cccttactca catgccttcc aggtgcaata aagtggcccc aaggaaaaaa aaaaaaa 837

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

15

ES 2 556 332 T3

```

<400> 24
tcccactgac gcattgcagga aggggcacct ccccttaacc aactgctct gtacggggca      60

cgtgggcaca ggtgcacact cacactcaca      90

<210> 25
<211> 89
5 <212> DNA
   <213> Homo Sapiens

<400> 25
ggcctgacag caacttttct tctactagtt catcttaaca cactgctctg tacggggcac      60

10 gtgggcacag gtgcacactc aactcaca      89

<210> 26
<211> 89
<212> DNA
<213> Homo Sapiens

15 <400> 26
ggcctgacag caacttttct tctactagtt catcttaact ttatcctggt aactggcgag      60

acaacctgtc ttaagtaact gaagggaaa      89

<210> 27
<211> 77
<212> DNA
20 <213> Homo Sapiens

<400> 27
tcccactgac gcaggaagga tcttaagttt atcctggtta ctggcgagac aacctgtctt      60

aagtaactga agggaaa      77

<210> 28
<211> 200
25 <212> PRT
   <213> Homo Sapiens

<400> 28
Met Ala Met Glu Thr Gln Met Ser Gln Asn Val Cys Pro Arg Asn Leu
1          5          10          15

Trp Leu Leu Gln Pro Leu Thr Val Leu Leu Leu Leu Ala Ser Ala Asp
          20          25          30

Ser Gln Ala Ala Ala Pro Pro Lys Ala Val Leu Lys Leu Glu Pro Pro
          35          40          45

Trp Ile Asn Val Leu Gln Glu Asp Ser Val Thr Leu Thr Cys Cys Gly
50          55          60

```

ES 2 556 332 T3

Ala Arg Ser Pro Glu Ser Pro Ser Ile Gln Trp Phe His His Asn Gly
65 70 75 80

Asn Leu Ile Pro Ile His Thr Gln Ser Ser Tyr Arg Phe Lys Ala Asn
85 90 95

Asn Asn Asp Ser Gly Glu Tyr Thr Cys Gln Thr Gly Gln Thr Ser Leu
100 105 110

Ser Asp Pro Val His Leu Thr Val Leu Ser Glu Trp Leu Leu Leu Gln
115 120 125

Thr Pro His Leu Glu Phe Gln Glu Gly Glu Thr Ile Asn Leu Arg Cys
130 135 140

His Ser Trp Lys Asp Lys Pro Leu Val Lys Val Thr Glu Glu Gln Asn
145 150 155 160

Gly Lys Ser Gln Lys Phe Ser Arg Leu Asp Pro Thr Phe Ser Ile Pro
165 170 175

Gln Ala Asn His Ser His Ser Gly Asp Tyr His Cys Thr Gly Asn Cys
180 185 190

Gly Tyr Thr Leu Phe Ser Ser Lys
195 200

<210> 29

<211> 184

5 <212> PRT

<213> Homo Sapiens

<400> 29

Met Trp Gln Leu Leu Leu Pro Thr Ala Leu Leu Leu Leu Val Ser Ala
1 5 10 15

Gly Met Arg Thr Glu Asp Leu Pro Lys Ala Val Val Phe Leu Glu Pro
20 25 30

Gln Trp Tyr Arg Val Leu Glu Lys Asp Ser Val Thr Leu Lys Cys Cys
35 40 45

ES 2 556 332 T3

Gly Ala Tyr Ser Pro Glu Leu Asn Ser Thr Gln Trp Phe His Asn Glu
50 55 60

Ser Leu Ile Ser Glu Gln Ala Ser Ser Tyr Phe Ile Asp Ala Ala Thr
65 70 75 80

Val Asp Asp Ser Gly Glu Tyr Arg Cys Gln Thr Asn Leu Ser Thr Leu
85 90 95

Ser Asp Pro Val Gln Leu Glu Val His Ile Gly Trp Leu Leu Leu Gln
100 105 110

Ala Pro Arg Trp Val Phe Lys Glu Glu Asp Pro Ile His Leu Arg Cys
115 120 125

His Ser Trp Lys Asn Thr Ala Leu His Lys Val Thr Tyr Leu Gln Asn
130 135 140

Gly Lys Gly Arg Lys Tyr Phe His His Asn Ser Asp Phe Tyr Ile Pro
145 150 155 160

Gln Ala Thr Leu Lys Asp Ser Gly Ser Tyr Phe Cys Arg Gly Leu Phe
165 170 175

Gly Ser Lys Asn Val Ser Ser Glu
180

<210> 30

<211> 188

5 <212> PRT

<213> Homo Sapiens

<400> 30

Met Ala Pro Ala Met Glu Ser Pro Thr Leu Leu Cys Val Ala Leu Leu
1 5 10 15

Phe Phe Ala Asp Asp Gly Val Leu Ala Val Pro Gln Lys Pro Lys Val
20 25 30

Ser Leu Asn Pro Pro Trp Asn Arg Ile Phe Lys Gly Glu Asn Val Thr
35 40 45

ES 2 556 332 T3

Leu Thr Cys Asn Gly Asn Asn Phe Phe Glu Val Ser Ser Thr Lys Trp
50 55 60

Phe His Asn Gly Ser Leu Ser Glu Ser Thr Asn Ser Ser Leu Asn Ile
65 70 75 80

Val Asn Ala Lys Phe Glu Asp Ser Gly Glu Tyr Lys Cys Gln His Gln
85 90 95

Gln Val Asn Glu Ser Glu Pro Val Tyr Leu Glu Val Phe Ser Asp Trp
100 105 110

Leu Leu Leu Gln Ala Ser Ala Glu Val Val Met Glu Gly Gln Pro Leu
115 120 125

Phe Leu Arg Cys His Gly Trp Arg Asn Trp Pro Val Tyr Lys Val Ile
130 135 140

Tyr Tyr Lys Asp Gly Glu Ala Leu Lys Tyr Trp Tyr Glu Asn His Asn
145 150 155 160

Ile Ser Ile Thr Asn Ala Thr Val Glu Asp Ser Gly Thr Tyr Tyr Cys
165 170 175

Thr Gly Lys Val Trp Gln Leu Asp Tyr Glu Ser Glu
180 185

<210> 31

<211> 378

5 <212> PRT

<213> Homo Sapiens

<400> 31

Met Trp Phe Leu Thr Thr Leu Leu Leu Trp Val Pro Val Asp Gly Gln
1 5 10 15

Val Asp Thr Thr Lys Ala Val Ile Ser Leu Gln Pro Pro Trp Val Ser
20 25 30

Phe Val Gln Glu Glu Thr Val Thr Leu His Cys Glu Val Leu His Leu
35 40 45

ES 2 556 332 T3

Pro Gly Ser Ser Ser Thr Gln Trp Phe Leu Asn Gly Thr Ala Thr Gln
 50 55 60

Thr Ser Thr Pro Ser Tyr Arg Ile Thr Ser Ala Ser Val Asn Asp Ser
 65 70 75 80

Gly Glu Tyr Arg Cys Gln Arg Gly Leu Ser Gly Arg Ser Asp Pro Thr
 85 90 95

Trp Leu Glu Thr His Arg Gly Trp Leu Leu Leu Gln Tyr Ser Ser Arg
 100 105 110

Val Phe Thr Glu Gly Glu Pro Leu Ala Leu Arg Cys His Ala Trp Lys
 115 120 125

Asp Lys Leu Val Tyr Asn Val Leu Tyr Tyr Arg Asn Gly Lys Ala Phe
 130 135 140

Lys Phe Phe His Trp Asn Ser Asn Leu Ile Ile Leu Lys Ile Asn Ile
 145 150 155 160

Ser Ser His Asn Gly Thr Tyr His Cys Ser Gly Asn Gly Lys His Arg
 165 170 175

Tyr Thr Ser Ala Gly Lys His Arg Tyr Thr Ser Ala Gly Ile Ser Val
 180 185 190

Thr Val Lys Glu Leu Phe Pro Ala Pro Val Leu Asn Ala Ser Val Thr
 195 200 205

Ser Pro Leu Leu Glu Gly Asn Leu Val Thr Leu Ser Cys Glu Thr Lys
 210 215 220

Leu Leu Leu Gln Arg Pro Gly Leu Gln Leu Tyr Phe Ser Phe Tyr Met
 225 230 235 240

Gly Ser Leu Thr Leu Arg Gly Arg Asn Thr Ser Ser Glu Tyr Gln Ile
 245 250 255

Leu Thr Ala Arg Arg Glu Asp Ser Gly Leu Tyr Trp Cys Glu Ala Ala
 260 265 270

Thr Glu Asp Gly Asn Val Leu Lys Arg Ser Pro Glu Leu Glu Leu Gln
 275 280 285

ES 2 556 332 T3

Val Leu Gly Leu Gln Leu Pro Thr Pro Val Val Trp Phe His Val Leu
290 295 300

Gly Tyr Leu Ala Val Gly Ile Met Phe Leu Val Asn Thr Val Leu Trp
305 310 315 320

Val Val Thr Ile Arg Lys Glu Leu Lys Arg Lys Lys Lys Trp Asp Leu
325 330 335

Glu Ile Ser Leu Asp Ser Gly His Glu Lys Lys Val Thr Ser Ser Leu
340 345 350

Gln Glu Asp Arg His Glu Glu Glu Glu Leu Lys Cys Gln Glu Gln Lys
355 360 365

Gly Glu Gln Leu Gln Glu Gly Val His Arg
370 375

<210> 32

<211> 376

<212> PRT

5 <213> Homo Sapiens

<400> 32

Met Leu Leu Trp Ala Ser Leu Leu Ala Phe Ala Pro Val Cys Gly Gln
1 5 10 15

Ser Ala Ala Ala His Lys Pro Val Ile Ser Val His Pro Pro Trp Thr
20 25 30

Thr Phe Phe Lys Gly Glu Arg Val Thr Leu Thr Cys Asn Gly Phe Gln
35 40 45

Phe Tyr Ala Thr Glu Lys Thr Thr Trp Tyr His Arg His Tyr Trp Gly
50 55 60

Glu Lys Leu Thr Leu Thr Pro Gly Asn Thr Leu Glu Val Arg Ala Ser
65 70 75 80

Gly Leu Tyr Arg Cys Gln Ala Arg Gly Ser Pro Arg Ser Asn Pro Val
85 90 95

ES 2 556 332 T3

Arg Leu Leu Phe Ser Ser Asp Ser Leu Ile Leu Gln Ala Pro Tyr Ser
 100 105 110
 Val Phe Glu Gly Asp Thr Leu Val Leu Arg Cys His Arg Arg Arg Lys
 115 120 125
 Glu Lys Leu Thr Ala Val Lys Tyr Thr Trp Asn Gly Asn Ile Leu Ser
 130 135 140
 Ile Ser Asn Lys Ser Trp Asp Leu Leu Ile Pro Gln Ala Ser Ser Asn
 145 150 155 160
 Asn Asn Gly Asn Tyr Arg Cys Ile Gly Tyr Gly Val Glu Asn Asp Val
 165 170 175
 Phe Arg Ser Asn Gly Asp Glu Asn Asp Val Phe Arg Ser Asn Phe Lys
 180 185 190
 Ile Ile Lys Ile Gln Glu Leu Phe Pro His Pro Glu Leu Lys Ala Thr
 195 200 205
 Asp Ser Gln Pro Thr Glu Gly Asn Ser Val Asn Leu Ser Cys Glu Thr
 210 215 220
 Gln Leu Pro Pro Glu Arg Ser Asp Thr Pro Leu His Phe Asn Phe Phe
 225 230 235 240
 Arg Asp Gly Glu Val Ile Leu Ser Asp Trp Ser Thr Tyr Pro Glu Leu
 245 250 255
 Gln Leu Pro Thr Val Trp Arg Glu Asn Ser Gly Ser Tyr Trp Cys Gly
 260 265 270
 Ala Glu Thr Val Arg Gly Asn Ile His Lys His Ser Pro Ser Leu Gln
 275 280 285
 Ile His Val Gln Arg Ile Pro Val Ser Gly Val Leu Leu Glu Thr Gln
 290 295 300
 Pro Ser Gly Gly Gln Ala Val Glu Gln Glu Met Leu Val Leu Val Cys
 305 310 315 320
 Ser Val Ala Glu Gly Thr Gly Asp Thr Thr Phe Ser Trp His Arg Glu
 325 330 335

ES 2 556 332 T3

Asp Met Gln Glu Ser Leu Gly Arg Lys Thr Gln Arg Ser Leu Arg Ala
340 345 350

Glu Leu Glu Leu Pro Ala Ile Arg Gln Ser His Ala Gly Gly Tyr Tyr
355 360 365

Cys Thr Ala Asp Asn Ser Tyr Gly
370 375

<210> 33

<211> 373

<212> PRT

5 <213> Homo Sapiens

<400> 33

Met Leu Leu Trp Val Ile Leu Leu Val Leu Ala Pro Val Ser Gly Gln
1 5 10 15

Phe Ala Arg Thr Pro Arg Pro Ile Ile Phe Leu Gln Pro Pro Trp Thr
20 25 30

Thr Val Phe Gln Gly Glu Arg Val Thr Leu Thr Cys Lys Gly Phe Arg
35 40 45

Phe Tyr Ser Pro Gln Arg Thr Arg Trp Tyr His Arg Tyr Leu Gly Lys
50 55 60

Glu Ile Leu Arg Glu Thr Pro Asp Asn Ile Leu Glu Val Gln Glu Ser
65 70 75 80

Gly Glu Tyr Arg Cys Gln Ala Gln Gly Ser Pro Leu Ser Ser Pro Val
85 90 95

His Leu Asp Phe Ser Ser Ala Ser Leu Ile Leu Gln Ala Pro Leu Ser
100 105 110

Val Phe Glu Gly Asp Ser Val Val Leu Arg Cys Arg Ala Lys Ala Glu
115 120 125

Val Thr Leu Asn Asn Thr Ile Tyr Lys Asn Asp Asn Val Leu Ala Phe
130 135 140

ES 2 556 332 T3

Leu Asn Lys Arg Thr Asp Phe His Ile Pro His Ala Cys Leu Lys Asp
145 150 155 160

Asn Gly Ala Tyr Arg Cys Thr Gly Tyr Lys Glu Ser Cys Cys Pro Val
165 170 175

Ser Ser Asn Lys Glu Ser Cys Cys Pro Val Ser Ser Asn Thr Val Lys
180 185 190

Ile Gln Val Gln Glu Pro Phe Thr Arg Pro Val Leu Arg Ala Ser Ser
195 200 205

Phe Gln Pro Thr Ser Gly Asn Pro Val Thr Leu Thr Cys Glu Thr Gln
210 215 220

Leu Ser Leu Glu Arg Ser Asp Val Pro Leu Arg Phe Arg Phe Phe Arg
225 230 235 240

Asp Asp Gln Thr Leu Gly Leu Gly Trp Ser Leu Ser Pro Asn Phe Gln
245 250 255

Ile Thr Ala Met Trp Ser Lys Asp Ser Gly Phe Tyr Trp Cys Lys Ala
260 265 270

Ala Thr Met Pro His Ser Val Ile Ser Asp Ser Pro Arg Ser Trp Ile
275 280 285

Gln Val Gln Ile Pro Ala Ser His Pro Val Leu Thr Leu Ser Pro Glu
290 295 300

Lys Ala Leu Asn Phe Glu Gly Thr Lys Val Thr Leu His Cys Glu Thr
305 310 315 320

Gln Glu Asp Ser Leu Arg Thr Leu Tyr Arg Phe Tyr His Glu Gly Val
325 330 335

Pro Leu Arg His Lys Ser Val Arg Cys Glu Arg Gly Ala Ser Ile Ser
340 345 350

Phe Ser Leu Thr Thr Glu Asn Ser Gly Asn Tyr Tyr Cys Thr Ala Asp
355 360 365

Asn Gly Leu Gly Ala
370

<210> 34

<211> 26

5 <212> PRT

<213> Homo Sapiens

<220>

<221> CARACTERÍSTICA_MISC

<222> (2)..(8)

10 <223> Xaa = Cualquier aminoácido

<220>

<221> CARACTERÍSTICA_MISC

<222> (26)..(26)
 <223> Xaa = I o L

 <220>
 <221> CARACTERÍSTICA_MISC
 5 <222> (15)..(15)
 <223> Xaa = I o L

 <220>
 <221> CARACTERÍSTICA_MISC
 <222> (9)..(9)
 10 <223> Xaa = D o E

 <220>
 <221> CARACTERÍSTICA_MISC
 <222> (1)..(1)
 <223> Xaa = D o E

 15 <220>
 <221> CARACTERÍSTICA_MISC
 <222> (10)..(11)
 <223> Xaa = Cualquier aminoácido

 <220>
 20 <221> CARACTERÍSTICA_MISC
 <222> (13)..(14)
 <223> Xaa = Cualquier aminoácido

 <220>
 <221> CARACTERÍSTICA_MISC
 25 <222> (16)..(22)
 <223> Xaa = Cualquier aminoácido

 <220>
 <221> CARACTERÍSTICA_MISC
 <222> (24)..(25)
 30 <223> Xaa = Cualquier aminoácido

 <400> 34
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Tyr Xaa Xaa Xaa Xaa
 1 5 10 15

 Xaa Xaa Xaa Xaa Xaa Xaa Tyr Xaa Xaa Xaa
 20 25

 <210> 35
 <211> 63
 35 <212> PRT
 <213> Homo Sapiens

 <400> 35
 Glu Ser Ser His Ser Ile Cys Pro Ala Gln Val Glu Leu Gln Ser Leu
 1 5 10 15

 Tyr Val Asp Val His Pro Lys Lys Gly Asp Leu Val Tyr Ser Glu Ile
 20 25 30

 Gln Thr Thr Thr Leu Gly Glu Glu Glu Glu Ala Asn Thr Ser Arg
 35 40 45

 Thr Leu Leu Glu Asp Lys Asp Val Ser Val Val Tyr Ser Glu Val
 50 55 60

 <210> 36
 40 <211> 39

<212> PRT
<213> Homo Sapiens

<400> 36
Asp Asn Lys Glu Pro Leu Asn Ser Asp Val Gln Tyr Thr Glu Val Gln
1 5 10 15

Val Ser Ser Ala Glu Trp Ser His Lys Asp Leu Gly Lys Lys Asp Thr
20 25 30

Glu Thr Val Tyr Ser Glu Val
35

5 <210> 37
<211> 68
<212> PRT
<213> Homo Sapiens

10 <220>
<221> CARACTERÍSTICA_MISC
<222> (38)..(61)
<223> Xaa = Cualquier aminoácido

<400> 37
Asp Ser Asp Ser Gln Glu Pro Thr Tyr His Asn Val Pro Ala Trp Glu
1 5 10 15

Glu Leu Gln Pro Val Tyr Thr Asn Ala Asn Pro Arg Gly Glu Asn Val
20 25 30

Val Tyr Ser Glu Val Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ile Ile Tyr
50 55 60

Ser Glu Val Lys
65

15 <210> 38
<211> 65
<212> PRT
<213> Homo Sapiens

20 <220>
<221> CARACTERÍSTICA_MISC
<222> (38)..(58)
<223> Xaa = Cualquier aminoácido

<400> 38
Ala Ser Asp Gln Arg Asp Leu Thr Glu His Lys Pro Ser Val Ser Asn
1 5 10 15

His Thr Gln Asp His Ser Asn Asp Pro Pro Asn Lys Met Asn Glu Val
20 25 30

Thr Tyr Ser Thr Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
50 55 60

Lys
65

<210> 39
<211> 6
5 <212> PRT
<213> Secuencia artificial

<220>
<223> Motivo de inhibición basado en tirosina de receptor inmune

<220>
10 <221> CARACTERÍSTICA_MISC
<222> (1)..(1)
<223> Xaa = S, V, L o I

<220>
15 <221> CARACTERÍSTICA_MISC
<222> (2)..(2)
<223> Xaa = Cualquier aminoácido

<220>
20 <221> CARACTERÍSTICA_MISC
<222> (4)..(5)
<223> Xaa = Cualquier aminoácido

<220>
<221> CARACTERÍSTICA_MISC
<222> (6)..(6)
<223> Xaa = L o V

25 <400> 39
Xaa Xaa Tyr Xaa Xaa Xaa

1 5

<210> 40
30 <211> 5321
<212> DNA
<213> HOMO SAPIENS

ES 2 556 332 T3

<400> 40

gtgcagtgtc ctgactgtaa gatcaagtcc aaacctgttt tggaattgag gaaacttctc	60
ttttgatctc agcccttggg ggtccaggtc ttcattgctgc tgtgggtgat attactgggc	120
ctggctcctg tcagtggaca gtttgcaagg acacccaggc coattatttt cctccagcct	180
ccatggacca cagtcttcca aggagagaga gtgaccctca cttgcaaggg atttgccttc	240
tactcaccac agaaaacaaa atggtaccat cggtaacctg ggaaagaaat actaagagaa	300
acccagaca atatccttga ggttcaggaa totggagagt acagatgcca ggcccagggc	360
tcccctctca gtagccctgt gcacttggat ttttcttcag cttecgctgat cctgcaagct	420
ccactttctg tgtttgaagg agactctgtg gttctgaggt gccgggcaaa ggcggaagta	480
acactgaata atactattta caagaatgat aatgtcctgg cattccttaa taaaagaact	540
gacttcata ttctcatgc atgtctcaag gacaatgggt catatcgctg tactggatat	600
aaggaaagtt gttgcctgtt ttcttccaat acagtcaaaa tccaagtcca agagccattt	660
acacgtccag tgctgagagc cagctccttc cagcccatca gcgggaaccc agtgaccctg	720
acctgtgaga ccagctctc tctagagagg tcagatgtcc cgctccgggt ccgcttcttc	780
agagatgacc agaccctggg attaggetgg agtctctccc cgaatttcca gattactgcc	840
atgtggagta aagattcagg gttctactgg tgtaaggcag caacaatgcc tcacagcgtc	900
atatctgaca gcccgagatc ctggatacag gtgcagatcc ctgcactcca tctgtctctc	960
actctcagcc ctgaaaaggc totgaatttt gagggaacca aggtgacact tcaactgtgaa	1020
acccaggaag attctctgag cactttgtac aggtttttatc atgaggggtg cccctgagg	1080
cacaagtcag tccgctgtga aaggggagca tccatcagct tctcactgac tacagagaat	1140
tcagggaact actactgcac agctgacaat ggccttggcg ccaagcccag taaggctgtg	1200
agcctctcag tcaactgttc cgtgtctcat cctgtcctca acctcagctc tctgaggac	1260
ctgatTTTTG agggagccaa ggtgacactt cactgtgaag ccagagagg ttactcccc	1320
atcctgtacc agtttcatca tgaggatgct gccctggagc gtaggtcggc caactctgca	1380
ggaggagtgg ccacagctt ctctctgact gcagagcatt cagggaacta ctactgcaca	1440

ES 2 556 332 T3

gctgacaatg gctttggccc ccagcgagcgt aaggcgggtga gcctctccat cactgtccct 1500
 gtgtctcctc ctgtccctcac cctcagctct gctgaggccc tgacttttga aggagccact 1560
 gtgacacttc actgtgaagt ccagagaggt tccccacaaa tctataacca gttttatcat 1620
 gaggacatgc cctgtggag cagctcaaca cctctgtgg gaagagtgtc cttcagcttc 1680
 tctctgactg aaggacattc aggaattac tactgcacag ctgacaatgg ctttgggtccc 1740
 cagcgagcgtg aagtgggtgag cctttttgtc actgttccag tgtctcgccc catcctcacc 1800
 ctgagggttc ccaggggccc ggctgtgggtg ggggacctgc tggagcttca ctgtgaggcc 1860
 ccgagaggct ctcccccaat cctgtactgg ttttatcatg aggatgtcac cctgggggagc 1920
 agctcagccc cctctggagg agaagcttct ttcaacctct ctctgactgc agaacattct 1980
 ggaaactact catgtgaggc caacaatggc ctagtggccc agcacagtga cacaatatca 2040
 ctcaagtgtt tagttccagt atctctccc atcctcacct tcagggtccc cagggccagc 2100
 gctgtgggtg gggacctgct ggagcttcc tgtgaggccc tgagaggctc ctccccaatc 2160
 ctgtactggt tttatcatga agatgtcacc ctgggtaaga tctcagcccc ctctggagga 2220
 ggggcctcct tcaacctctc tctgactaca gaacattctg gaatctactc ctgtgaggca 2280
 gacaatggtc tggaggccca gcgcagtgag atggtgacac tgaaagtgc agttccggtg 2340
 tctcgcccg tctcaccct cagggtccc gggacctatg ctgcgggtggg ggacctgctg 2400
 gagcttcaet gtgaggccct gagaggctct cccctgatcc tglaccggtt ttttcatgag 2460
 gatgtcacc taggaaatag gtgtccccc totggaggag cgtccttaaa cctctctctg 2520
 actgcagagc actctggaaa ctactcctgt gaggcggaca atggcctcgg ggcccagcgc 2580
 agtgagacag tgacacttta tatcacagg ctgaccgcga acagaagtgg cctttttgcc 2640
 acaggagtgc ccgggggct gctcagcata gcaggccttg ctgcgggggc actgctgctc 2700
 tactgctggc tctcgagaaa agcagggaga aagcctgcct ctgacccgc caggagccct 2760
 tcagactcgg actcccaaga gccacctat cacaatgtac cagcctggga agagctgcaa 2820
 ccagtgtaca ctaatgcaaa tcttagagga gaaaatgtgg ttactcaga agtaaggatc 2880
 atccaagaga aaaagaaaca tgcagtggc totgacccca ggcactcag gaacaagggt 2940
 tcccctatca tctactctga agttaagggt gogtcaaccc cggtttcgg atccctgttc 3000
 ttggcttcc cagctcctca cagatgagtc cacacgtctc tccaactgct gtttcagcct 3060
 ctgcacccca aagtccctc tgggggagaa gcagcattga agtgggaaga tttaggctgc 3120
 ccagaccat atctactggc ctttgtttca catgtcctca ttctcagtct gaccagaatg 3180
 cagggcctg ctggactgtc aactgttcc cagttaaagc cctgactggc aggtttttta 3240

ES 2 556 332 T3

atccagtggc aaggtgctcc cactccaggg cccagcacat ctctggatt ccttagtggg 3300
 ctccagctgt ggttgctgtt ctgagtaactg ctctcatcac accccacag aggggtctt 3360
 accacacaaa gggagagtgg gccttcagga gatgcgggc tggcctaaca gtcaggtgc 3420
 tctaaactc cgacacagag ttctgcttt gggtgatgc atttctaat tgcatacagc 3480
 ctggtggggc tactgcagtg tgctgcaaaa tgggacagca cacagcctgt gcacatggga 3540
 catgtgatgg gtctccccc gggggctgca ttccacactc ctccacctgt ctcaaaactct 3600
 aaggtgggca ctgacacca aggttaactc tctctgctc atgtgtcagt gtctacctgc 3660
 ccaagtaagt ggctttcata cccaagtcc cgaagttctt cccatctaa cagaagtaac 3720
 ccagcaagtc aaggccagga ggaccagggg tgacagacaga acacatactg gaacacagga 3780
 ggtgctcaat tactatttga ctgactgact gaatgaatga atgaatgagg aagaaaactg 3840
 tgggtaatca aactggcata aaatccagtg cactccctag gaaatccggg aggtattctg 3900
 gcttctcaag aaacaacgga agagaaggag cttggatgaa gaaactgttc agcaagaaga 3960
 agggcttctt cacactttta tgtgcttgtg gatcacctga ggatctgtga aaatacagat 4020
 actgattcag tgggtctgtg tagagcctga gactgccatt ctaacatgtt cccaggggat 4080
 gctgatgctg ctggccctgg gactgcactg catgcatgtg aagccctata ggtctcagca 4140
 gaggcccatg gagagggaat gtgtggctct ggctgccag ggcccaactc ggttcacacg 4200
 gatcgtgctg ctccctggcc agcctttggc cacagcacca ccagctgctg ttgctgagag 4260
 agcttctctt ctgtgacatg ttggctttca tcagccaccc tgggaagcgg aaagtagctg 4320
 ccactatctt tgtttcccca cctcaggcct cacactttcc catgaaaagg gtgaatgtat 4380
 ataacctgag cctctccat tcagagttgt tctcccatct ctgagcaatg ggaatgtctg 4440
 ttccgttttt atgatatcca tcacatctta tcttgatctt tgcctccagt ggattgtaca 4500
 gtgatgactt ttaagcccca cggccctgaa ataaaatcct tccaagggca ttggaagctc 4560
 actccacctg aacctggtt ttctatgctt ccaagtgtca gggccttgcc cagatagaca 4620
 gggctgactc tgctgcccc acccttcaag gaggaacca gacacctgag acaggagcct 4680
 gtatgcagcc cagtgcagcc ttgcagagga caaggctgga ggcatttgc atcactacag 4740
 atatgcaact aaaatagacg tggagcaaga gaaatgcatt cccaccgagg ccgctttttt 4800
 aggcctagtt gaaagtcaag aaggacagca gcaagcatag gctcaggatt aaagaaaaaa 4860
 atctgctcac agtctgttct ggaggtcaca tcaccaacaa agctcacgcc ctatgcagtt 4920
 ctgagaaggt ggaggcacca ggctcaaaa aggaaattta gaatttctca ttgggagagt 4980
 aaggtacccc catcccagaa tgataactgc acagtggcag aacaaactcc accctaattg 5040
 gggtggaacc catccagtct gttgaaggcc tgaatgtaac aaaaggcctt attcttctc 5100
 aagtaagggg gaactcctgc ttgggctgg gacataagtt tttctgctt cagacgcaaa 5160
 ctgaaaaatg gctcttcttg ggtcttgagc ttgctggcat atggactgaa agaaaactatg 5220
 ctattggatc tcttgatct ccagcttgc gactgcagat cttgagatat gtcagcctct 5280
 acagtccaaa gagctaattc attctaataa accaatctt c 5321

<210> 41

<211> 977

<212> PRT

<213> HOMO SAPIENS

ES 2 556 332 T3

<400> 41

Met	Leu	Leu	Trp	Val	Ile	Leu	Leu	Val	Leu	Ala	Pro	Val	Ser	Gly	Gln	1	5	10	15
Phe	Ala	Arg	Thr	Pro	Arg	Pro	Ile	Ile	Phe	Leu	Gln	Pro	Pro	Trp	Thr	20	25	30	
Thr	Val	Phe	Gln	Gly	Glu	Arg	Val	Thr	Leu	Thr	Cys	Lys	Gly	Phe	Arg	35	40	45	
Phe	Tyr	Ser	Pro	Gln	Lys	Thr	Lys	Trp	Tyr	His	Arg	Tyr	Leu	Gly	Lys	50	55	60	
Glu	Ile	Leu	Arg	Glu	Thr	Pro	Asp	Asn	Ile	Leu	Glu	Val	Gln	Glu	Ser	65	70	75	80
Gly	Glu	Tyr	Arg	Cys	Gln	Ala	Gln	Gly	Ser	Pro	Leu	Ser	Ser	Pro	Val	85	90	95	
His	Leu	Asp	Phe	Ser	Ser	Ala	Ser	Leu	Ile	Leu	Gln	Ala	Pro	Leu	Ser	100	105	110	
Val	Phe	Glu	Gly	Asp	Ser	Val	Val	Leu	Arg	Cys	Arg	Ala	Lys	Ala	Glu	115	120	125	
Val	Thr	Leu	Asn	Asn	Thr	Ile	Tyr	Lys	Asn	Asp	Asn	Val	Leu	Ala	Phe	130	135	140	
Leu	Asn	Lys	Arg	Thr	Asp	Phe	His	Ile	Pro	His	Ala	Cys	Leu	Lys	Asp	145	150	155	160

ES 2 556 332 T3

Asn Gly Ala Tyr Arg Cys Thr Gly Tyr Lys Glu Ser Cys Cys Pro Val
 165 170 175

Ser Ser Asn Thr Val Lys Ile Gln Val Gln Glu Pro Phe Thr Arg Pro
 180 185 190

Val Leu Arg Ala Ser Ser Phe Gln Pro Ile Ser Gly Asn Pro Val Thr
 195 200 205

Leu Thr Cys Glu Thr Gln Leu Ser Leu Glu Arg Ser Asp Val Pro Leu
 210 215 220

Arg Phe Arg Phe Phe Arg Asp Asp Gln Thr Leu Gly Leu Gly Trp Ser
 225 230 235 240

Leu Ser Pro Asn Phe Gln Ile Thr Ala Met Trp Ser Lys Asp Ser Gly
 245 250 255

Phe Tyr Trp Cys Lys Ala Ala Thr Met Pro His Ser Val Ile Ser Asp
 260 265 270

Ser Pro Arg Ser Trp Ile Gln Val Gln Ile Pro Ala Ser His Pro Val
 275 280 285

Leu Thr Leu Ser Pro Glu Lys Ala Leu Asn Phe Glu Gly Thr Lys Val
 290 295 300

Thr Leu His Cys Glu Thr Gln Glu Asp Ser Leu Arg Thr Leu Tyr Arg
 305 310 315 320

Phe Tyr His Glu Gly Val Pro Leu Arg His Lys Ser Val Arg Cys Glu
 325 330 335

Arg Gly Ala Ser Ile Ser Phe Ser Leu Thr Thr Glu Asn Ser Gly Asn
 340 345 350

Tyr Tyr Cys Thr Ala Asp Asn Gly Leu Gly Ala Lys Pro Ser Lys Ala
 355 360 365

Val Ser Leu Ser Val Thr Val Pro Val Ser His Pro Val Leu Asn Leu
 370 375 380

Ser Ser Pro Glu Asp Leu Ile Phe Glu Gly Ala Lys Val Thr Leu His
 385 390 395 400

ES 2 556 332 T3

Cys Glu Ala Gln Arg Gly Ser Leu Pro Ile Leu Tyr Gln Phe His His
405 410 415

Glu Asp Ala Ala Leu Glu Arg Arg Ser Ala Asn Ser Ala Gly Gly Val
420 425 430

Ala Ile Ser Phe Ser Leu Thr Ala Glu His Ser Gly Asn Tyr Tyr Cys
435 440 445

Thr Ala Asp Asn Gly Phe Gly Pro Gln Arg Ser Lys Ala Val Ser Leu
450 455 460

Ser Ile Thr Val Pro Val Ser His Pro Val Leu Thr Leu Ser Ser Ala
465 470 475 480

Glu Ala Leu Thr Phe Glu Gly Ala Thr Val Thr Leu His Cys Glu Val
485 490 495

Gln Arg Gly Ser Pro Gln Ile Leu Tyr Gln Phe Tyr His Glu Asp Met
500 505 510

Pro Leu Trp Ser Ser Ser Thr Pro Ser Val Gly Arg Val Ser Phe Ser
515 520 525

Phe Ser Leu Thr Glu Gly His Ser Gly Asn Tyr Tyr Cys Thr Ala Asp
530 535 540

Asn Gly Phe Gly Pro Gln Arg Ser Glu Val Val Ser Leu Phe Val Thr
545 550 555 560

Val Pro Val Ser Arg Pro Ile Leu Thr Leu Arg Val Pro Arg Ala Gln
565 570 575

Ala Val Val Gly Asp Leu Leu Glu Leu His Cys Glu Ala Pro Arg Gly
580 585 590

Ser Pro Pro Ile Leu Tyr Trp Phe Tyr His Glu Asp Val Thr Leu Gly
595 600 605

Ser Ser Ser Ala Pro Ser Gly Gly Glu Ala Ser Phe Asn Leu Ser Leu
610 615 620

Thr Ala Glu His Ser Gly Asn Tyr Ser Cys Glu Ala Asn Asn Gly Leu
625 630 635 640

ES 2 556 332 T3

Val Ala Gln His Ser Asp Thr Ile Ser Leu Ser Val Ile Val Pro Val
 645 650 655

Ser Arg Pro Ile Leu Thr Phe Arg Ala Pro Arg Ala Gln Ala Val Val
 660 665 670

Gly Asp Leu Leu Glu Leu His Cys Glu Ala Leu Arg Gly Ser Ser Pro
 675 680 685

Ile Leu Tyr Trp Phe Tyr His Glu Asp Val Thr Leu Gly Lys Ile Ser
 690 695 700

Ala Pro Ser Gly Gly Gly Ala Ser Phe Asn Leu Ser Leu Thr Thr Glu
 705 710 715 720

His Ser Gly Ile Tyr Ser Cys Glu Ala Asp Asn Gly Leu Glu Ala Gln
 725 730 735

Arg Ser Glu Met Val Thr Leu Lys Val Ala Val Pro Val Ser Arg Pro
 740 745 750

Val Leu Thr Leu Arg Ala Pro Gly Thr His Ala Ala Val Gly Asp Leu
 755 760 765

Leu Glu Leu His Cys Glu Ala Leu Arg Gly Ser Pro Leu Ile Leu Tyr
 770 775 780

Arg Phe Phe His Glu Asp Val Thr Leu Gly Asn Arg Ser Ser Pro Ser
 785 790 795 800

Gly Gly Ala Ser Leu Asn Leu Ser Leu Thr Ala Glu His Ser Gly Asn
 805 810 815

Tyr Ser Cys Glu Ala Asp Asn Gly Leu Gly Ala Gln Arg Ser Glu Thr
 820 825 830

Val Thr Leu Tyr Ile Thr Gly Leu Thr Ala Asn Arg Ser Gly Pro Phe
 835 840 845

Ala Thr Gly Val Ala Gly Gly Leu Leu Ser Ile Ala Gly Leu Ala Ala
 850 855 860

Gly Ala Leu Leu Leu Tyr Cys Trp Leu Ser Arg Lys Ala Gly Arg Lys
 865 870 875 880

ES 2 556 332 T3

Pro Ala Ser Asp Pro Ala Arg Ser Pro Ser Asp Ser Asp Ser Gln Glu
885 890 895

Pro Thr Tyr His Asn Val Pro Ala Trp Glu Glu Leu Gln Pro Val Tyr
900 905 910

Thr Asn Ala Asn Pro Arg Gly Glu Asn Val Val Tyr Ser Glu Val Arg
915 920 925

Ile Ile Gln Glu Lys Lys Lys His Ala Val Ala Ser Asp Pro Arg His
930 935 940

Leu Arg Asn Lys Gly Ser Pro Ile Ile Tyr Ser Glu Val Lys Val Ala
945 950 955 960

Ser Thr Pro Val Ser Gly Ser Leu Phe Leu Ala Ser Ser Ala Pro His
965 970 975

Arg

<210> 42

<211> 16

<212> DNA

5 <213> HOMO SAPIENS

<400> 42

ggcacctccc ctaac 16

<210> 43

<211> 2797

10 <212> DNA

<213> HOMO SAPIENS

<400> 43

gtgcagtgtc ctgactgtaa gatcaagtc aaacctgttt tggaattgag gaaacttctc 60

ttttgatctc agcccttggt ggtccaggtc ttcattgctgc tgtgggtgat attactggtc 120

ES 2 556 332 T3

ctggtctctg tcagtggaca gtttgcaagg acacccaggc ccattatctt cctccagcct	180
ccatggacca cagtcttcca aggagagaga gtgacctca cttgcaaggg atttcgcttc	240
tactcaccac agaaaacaaa atggtaccat cggtaacctg ggaaagaaat actaagagaa	300
accccagaca atataccttga ggttcaggaa tctggagagt acagatgcca ggcccagggc	360
tcccctctca gtagccctgt gcaattggat ttttcttcag cttecgctgat cctgcaagct	420
ccactttctg tgtttgaagg agactctgtg gttctgaggt gccgggcaaa ggcggaagta	480
acactgaata atactattta caagaatgat aatgtccttg catctcttaa taaaagaact	540
gaattccata ttctcatgc atgtctcaag gacaatggg catatcgctg tactggatat	600
aaggaaagtt gttgcctgt ttcttccaat acagtcaaaa tccaagtcca agagccattt	660
acacgtccag tgctgagagc cagctccttc cagcccatca gccggaacc agtgacctg	720
acctgtgaga cccagctctc tctagagagg tcagatgtcc cgtccgggtt ccgcttcttc	780
agagatgacc agacctggg attaggctgg agtctctccc cgaatttcca gattactgcc	840
atgtggagta aagattcagg gttctactgg tgtaaggcag caacaatgcc tcacagcgtc	900
atatctgaca gcccgagatc ctggatacag gtgcagatcc ctgcattcca tctgtctc	960
actctcagcc ctgaaaaggc tctgaatttt gagggaaacca aggtgacact tcactgtgaa	1020
acccaggaag attctctgag cactttgtac aggttttata atgaggggtg cccctgagg	1080
cacaagttag tccgtgtga aaggggagca tccatcagct tctcactgac tacagagaat	1140
tcagggaact actactgcac agctgacaat gcccttggcg ccaagcccag taaggctgtg	1200
agcctctcag tcactgttcc cgtgtctcat cctgtctcca acctcagctc tctgaggac	1260
ctgatttttg agggagccaa ggtgacactt cactgtgaag cccagagagg ttcactcccc	1320
atcctgtacc agtttcatca tgaggatgct gccctggagc gtaggtcggc caactctgca	1380
ggaggagtgg ccatcagctt ctctctgact gcagagcatt cagggaacta ctactgcaca	1440
gctgacaatg gctttggccc ccagcgcagt aaggcgggta gcctctccat cactgtccct	1500
gtgtctcacc ctgtctcacc cctcagctct gctgaggccc tgacttttga aggagccact	1560
gtgacacttc actgtgaagt ccagagaggt tccccacaaa tctatacca gttttatcat	1620
gaggacatgc cctgtggag cagctcaaca cctctgtgg gaagagtgtc cttcagcttc	1680
tctctgactg aaggacattc aggaattac tactgcacag ctgacaatgg ctttggctcc	1740
cagcgcagtg aagtggtag cctttttgtc actgttcag tgtctcgccc catctcacc	1800
ctcaggggtc ccagggccca ggtgtgggtg ggggacctgc tggagcttca ctgtgaggcc	1860
ccgagaggct ctccccaat cctgtactgg ttttatcatg aggatgtcac cctggggagc	1920

ES 2 556 332 T3

```

agctcagccc cctctggagg agaagcttct ttcaacctct ctctgactgc agaacattct 1980
ggaaactact catgtgaggc caacaatggc ctagtggccc agcacagtga cacaatatca 2040
ctcagtgtta tagttccagt atctcgtccc atcctcacct tcagggtccc cagggtccag 2100
gctgtgggtg gggacctgct ggagcttcac tgtgaggccc tgagaggctc ctccccaatc 2160
ctgtactggt tttatcatga agatgtcacc ctgggtaaga tctcagcccc ctctggagga 2220
ggggcctcct tcaacctctc tctgactaca gaacattctg gaatctactc ctgtgaggca 2280
gacaatggtc tggaggccca ggcagtgag atggtgacac tgaaagttgc aggtgagtgg 2340
gccctgcccc ccagcagcac atctgagaac tgactgtgcc tgttctccct gcagctgaaa 2400
atggagccac agagctctc agggctgttt gcttgtgtgg catcccagca cacttctctc 2460
ctgcagaacc tccctgtgaa agtctcggat cctttgtggt atggttccag gaatctgatg 2520
tttcccagca gtcttcttga agatgatcaa agcacctcac taaaaatgca aataagactt 2580
ttttagaaca taaactatat tctgaactga aattattaca tgaaaatgaa accaaagaat 2640
tctgagcata tgtttctctg ccgtagaaag gattaagctg tttcttctcc ggattctctt 2700
ctcattgact tctaagaagc ctctactctt gagtctcttt cattactggg gatgtaaatg 2760
ttccttacat ttccacatta aaaatcctat gttaacg 2797

```

<210> 44

<211> 759

<212> PRT

5 <213> HOMO SAPIENS

<400> 44

```

Met Leu Leu Trp Val Ile Leu Leu Val Leu Ala Pro Val Ser Gly Gln
1           5           10           15

```

```

Phe Ala Arg Thr Pro Arg Pro Ile Ile Phe Leu Gln Pro Pro Trp Thr
                20           25           30

```

```

Thr Val Phe Gln Gly Glu Arg Val Thr Leu Thr Cys Lys Gly Phe Arg
          35           40           45

```

```

Phe Tyr Ser Pro Gln Lys Thr Lys Trp Tyr His Arg Tyr Leu Gly Lys
          50           55           60

```

```

Glu Ile Leu Arg Glu Thr Pro Asp Asn Ile Leu Glu Val Gln Gly Ser
65           70           75           80

```

ES 2 556 332 T3

Gly Glu Tyr Arg Cys Gln Ala Gln Gly Ser Pro Leu Ser Ser Pro Val
 85 90 95
 His Leu Asp Phe Ser Ser Ala Ser Leu Ile Leu Gln Ala Pro Leu Ser
 100 105 110
 Val Phe Glu Gly Asp Ser Val Val Leu Arg Cys Arg Ala Lys Ala Glu
 115 120 125
 Val Thr Leu Asn Asn Thr Ile Tyr Lys Asn Asp Asn Val Leu Ala Phe
 130 135 140
 Leu Asn Lys Arg Thr Asp Phe His Ile Pro His Ala Cys Leu Lys Asp
 145 150 155 160
 Asn Gly Ala Tyr Arg Cys Thr Gly Tyr Lys Glu Ser Cys Cys Pro Val
 165 170 175
 Ser Ser Asn Thr Val Lys Ile Gln Val Gln Glu Pro Phe Thr Arg Pro
 180 185 190
 Val Leu Arg Ala Ser Ser Phe Gln Pro Ile Ser Gly Asn Pro Val Thr
 195 200 205
 Leu Thr Cys Glu Thr Gln Leu Ser Leu Glu Arg Ser Asp Val Pro Leu
 210 215 220
 Arg Phe Arg Phe Phe Arg Asp Asp Gln Thr Leu Gly Leu Gly Trp Ser
 225 230 235 240
 Leu Ser Pro Asn Phe Gln Ile Thr Ala Met Trp Ser Lys Asp Ser Gly
 245 250 255
 Phe Tyr Trp Cys Lys Ala Ala Thr Met Pro His Ser Val Ile Ser Asp
 260 265 270
 Ser Pro Arg Ser Trp Ile Gln Val Gln Ile Pro Ala Ser His Pro Val
 275 280 285
 Leu Thr Leu Ser Pro Glu Lys Ala Leu Asn Phe Glu Gly Thr Lys Val
 290 295 300
 Thr Leu His Cys Glu Thr Gln Glu Asp Ser Leu Arg Thr Leu Tyr Arg
 305 310 315 320

ES 2 556 332 T3

Phe Tyr His Glu Gly Val Pro Leu Arg His Lys Ser Val Arg Cys Glu
325 330 335

Arg Gly Ala Ser Ile Ser Phe Ser Leu Thr Thr Glu Asn Ser Gly Asn
340 345 350

Tyr Tyr Cys Thr Ala Asp Asn Gly Leu Gly Ala Lys Pro Ser Lys Ala
355 360 365

Val Ser Leu Ser Val Thr Val Pro Val Ser His Pro Val Leu Asn Leu
370 375 380

Ser Ser Pro Glu Asp Leu Ile Phe Glu Gly Ala Lys Val Thr Leu His
385 390 395 400

Cys Glu Ala Gln Arg Gly Ser Leu Pro Ile Leu Tyr Gln Phe His His
405 410 415

Glu Asp Ala Ala Leu Glu Arg Arg Ser Ala Asn Ser Ala Gly Gly Val
420 425 430

Ala Ile Ser Phe Ser Leu Thr Ala Glu His Ser Gly Asn Tyr Tyr Cys
435 440 445

Thr Ala Asp Asn Gly Phe Gly Pro Gln Arg Ser Lys Ala Val Ser Leu
450 455 460

Ser Ile Thr Val Pro Val Ser His Pro Val Leu Thr Leu Ser Ser Ala
465 470 475 480

Glu Ala Leu Thr Phe Glu Gly Ala Thr Val Thr Leu His Cys Glu Val
485 490 495

Gln Arg Gly Ser Pro Gln Ile Leu Tyr Gln Phe Tyr His Glu Asp Met
500 505 510

Pro Leu Trp Ser Ser Ser Thr Pro Ser Val Gly Arg Val Ser Phe Ser
515 520 525

Phe Ser Leu Thr Glu Gly His Ser Gly Asn Tyr Tyr Cys Thr Ala Asp
530 535 540

Asn Gly Phe Gly Pro Gln Arg Ser Glu Val Val Ser Leu Phe Val Thr
545 550 555 560

ES 2 556 332 T3

Val Pro Val Ser Arg Pro Ile Leu Thr Leu Arg Val Pro Arg Ala Gln
565 570 575

Ala Val Val Gly Asp Leu Leu Glu Leu His Cys Glu Ala Pro Arg Gly
580 585 590

Ser Pro Pro Ile Leu Tyr Trp Phe Tyr His Glu Asp Val Thr Leu Gly
595 600 605

Ser Ser Ser Ala Pro Ser Gly Gly Glu Ala Ser Phe Asn Leu Ser Leu
610 615 620

Thr Ala Glu His Ser Gly Asn Tyr Ser Cys Glu Ala Asn Asn Gly Leu
625 630 635 640

Val Ala Gln His Ser Asp Thr Ile Ser Leu Ser Val Ile Val Pro Val
645 650 655

Ser Arg Pro Ile Leu Thr Phe Arg Ala Pro Arg Ala Gln Ala Val Val
660 665 670

Gly Asp Leu Leu Glu Leu His Cys Glu Ala Leu Arg Gly Ser Ser Pro
675 680 685

Ile Leu Tyr Trp Phe Tyr His Glu Asp Val Thr Leu Gly Lys Ile Ser
690 695 700

Ala Pro Ser Gly Gly Gly Ala Ser Phe Asn Leu Ser Leu Thr Thr Glu
705 710 715 720

His Ser Gly Ile Tyr Ser Cys Glu Ala Asp Asn Gly Leu Glu Ala Gln
725 730 735

Arg Ser Glu Met Val Thr Leu Lys Val Ala Gly Glu Trp Ala Leu Pro
740 745 750

Thr Ser Ser Thr Ser Glu Asn
755

REIVINDICACIONES

1. Un anticuerpo dirigido a una proteína IRTA1 humana, en donde la secuencia de aminoácidos de la proteína IRTA1 humana se expone en la Figura 18A.
- 5 2. Un anticuerpo dirigido a una proteína IRTA1 humana, en donde el anticuerpo está conjugado con un agente terapéutico, en donde el agente terapéutico se selecciona del grupo que consiste en un radioisótopo, una toxina, un toxoide, o un agente quimioterapéutico, y en donde la secuencia de aminoácidos de la proteína IRTA1 humana se expone en la Figura 18A.
3. El anticuerpo de la reivindicación 1 o 2, en donde el anticuerpo se une específicamente al dominio extracelular de una proteína IRTA1 humana.
- 10 4. El anticuerpo de una cualquiera de las reivindicaciones 1-3, en donde el anticuerpo es un anticuerpo monoclonal.
5. El anticuerpo de la reivindicación 4, en donde el anticuerpo monoclonal es un anticuerpo monoclonal humanizado.
6. Un anticuerpo dirigido a una proteína IRTA1 humana eficaz para unirse a las células cancerosas que expresan una proteína IRTA1 humana con el fin de evitar el crecimiento de las células cancerosas para su uso en el tratamiento de un sujeto que tiene un cáncer de células B y en donde la secuencia de aminoácidos de la proteína IRTA1 humana se expone en la Figura 18A.
- 15 7. El anticuerpo de la reivindicación 6, en donde el anticuerpo es un anticuerpo monoclonal.
8. El anticuerpo de la reivindicación 7, en donde el anticuerpo monoclonal es un anticuerpo monoclonal humanizado.
9. El anticuerpo de cualquiera de las reivindicaciones 1-8 para su uso en el tratamiento de un sujeto que tiene un cáncer de células B.
- 20 10. Una composición que comprende una cantidad del anticuerpo de una cualquiera de las reivindicaciones 1-9 eficaz para prevenir el crecimiento de células cancerosas que expresan IRTA1 humana y un portador farmacéuticamente aceptable.
11. La composición de la reivindicación 10 para su uso en el tratamiento de un cáncer de células B.
- 25 12. La composición de la reivindicación 11, en donde las células cancerosas se seleccionan del grupo que consiste en células de linfoma de células B, linfoma de células del manto, mieloma múltiple, linfoma de Burkitt, linfoma de la zona marginal, linfoma difuso de células grandes y linfoma folicular.
13. La composición de la reivindicación 10, en donde las células cancerosas son células de linfoma de Tejido Linfoide Asociado a la Mucosa (MALT) o células de Linfoma no Hodgkin.
14. La composición de la reivindicación 10, en donde el anticuerpo es un anticuerpo monoclonal.
- 30 15. La composición de la reivindicación 14, en donde el anticuerpo monoclonal es un anticuerpo monoclonal humanizado.
16. Un ácido nucleico que codifica una proteína Asociada a la Translocación de Receptores de la superfamilia de la Inmunoglobulina humana (IRTA), cuya secuencia de aminoácidos se expone en la Figura 18A.
17. Un vector que comprende el ácido nucleico de la reivindicación 16.
- 35 18. Una célula anfitriona que comprende el vector de la reivindicación 17.
19. Un ácido nucleico que comprende al menos 15 nucleótidos contiguos capaces de hibridar específicamente, bajo condiciones restrictivas, con una secuencia única incluida dentro de la secuencia de la molécula de ácido nucleico de la reivindicación 16.
- 40 20. Un método para detectar una neoplasia maligna de células B o un tipo de neoplasia maligna de células B en una muestra de un sujeto en donde la neoplasia maligna de células B comprende una reordenación cromosómica 1q21, que comprende:
 - a) poner en contacto la muestra de ácido nucleico obtenida del sujeto con el ácido nucleico de la reivindicación 17, en donde el ácido nucleico está marcado con un marcador detectable; y
 - 45 b) detectar cualquier hibridación en la etapa (a), en donde la detección de hibridación indica la presencia de neoplasia maligna de células B o un tipo de neoplasia maligna de células B en el sujeto.
21. Una proteína IRTA1 humana que comprende los aminoácidos, cuya secuencia se expone en la Figura 18A.

22. Un método de diagnóstico de una neoplasia maligna de células B que comprende una reordenación cromosómica 1q21 en una muestra que contiene células cancerosas de un sujeto que comprende:

a) poner en contacto la muestra de ácido nucleico obtenida del sujeto con el anticuerpo de la reivindicación 1, en donde el anticuerpo está marcado con un marcador detectable; y

5 b) detectar cualquier unión en la etapa (a), en donde la detección de la unión indica un diagnóstico de neoplasia maligna de células B en la muestra.

23. Un oligonucleótido antisentido que tiene una secuencia capaz de hibridar específicamente, bajo condiciones restrictivas, con al menos 15 nucleótidos contiguos de una molécula de ARNm que codifica una proteína IRTA1 humana de la reivindicación 21.

10

FIGURA 1

FIG. 1A

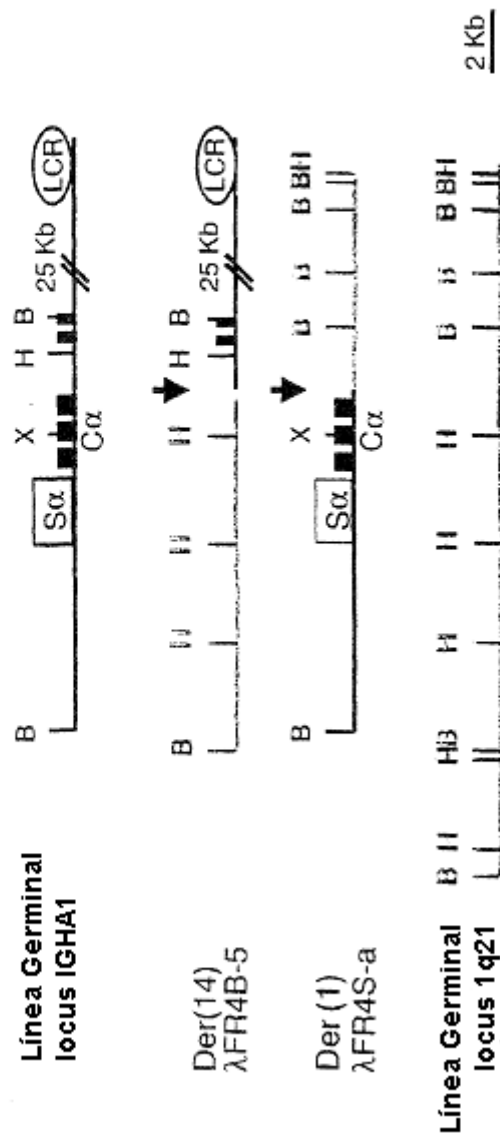


FIG. 1B

Chr 1: GGGCTGACAGCAACTTTCTCTACTAGTTTCATCTTAACTTTATCTGTGTAACCTGGGAGACAACCTGTCTTTAAGTAACCTGAAGGAAA
 |||||
 der14: GGGCTGACAGCAACTTTCTCTACTAGTTTCATCTTAACTTTATCTGTGTAACCTGGGAGACAACCTGTCTTTAAGTAACCTGAAGGAAA
 |||||
 Chr 14: TCCCACTGACGCATGACAGGAGGGGACCTCCCTTAACTGACACACTGCTCTGTACGGGGACCTGGGACAGGTGACACTCACA

FIGURA 2A

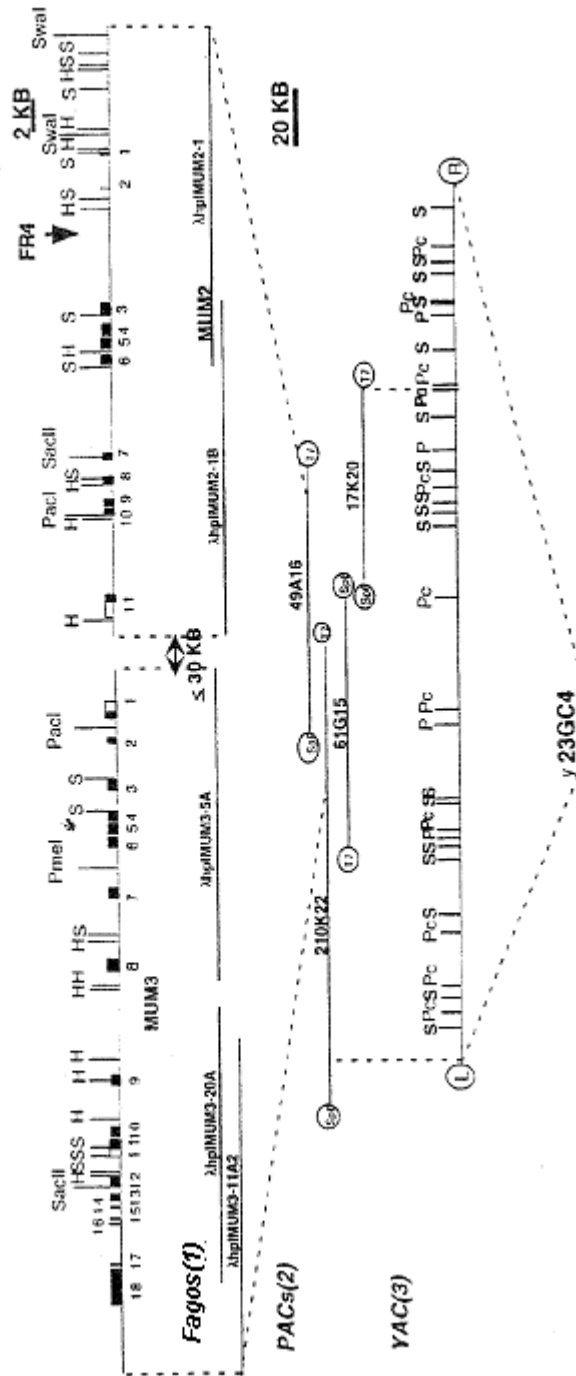


FIGURA 2B

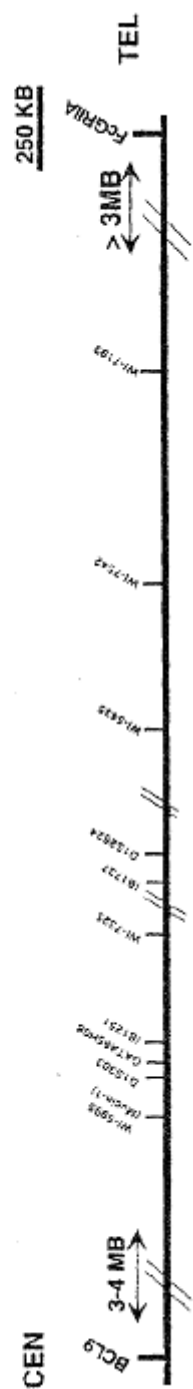


FIGURA 3

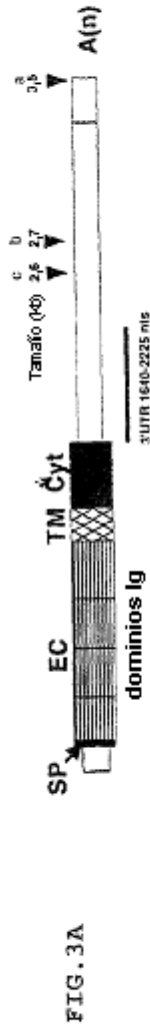


FIG. 3B

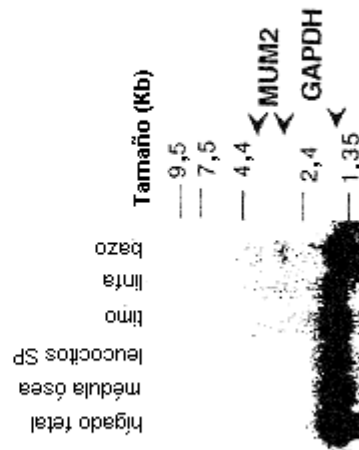


FIG. 3C

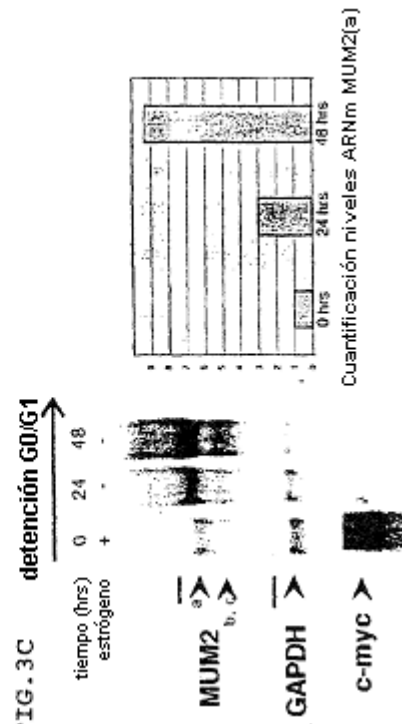


FIGURA 4

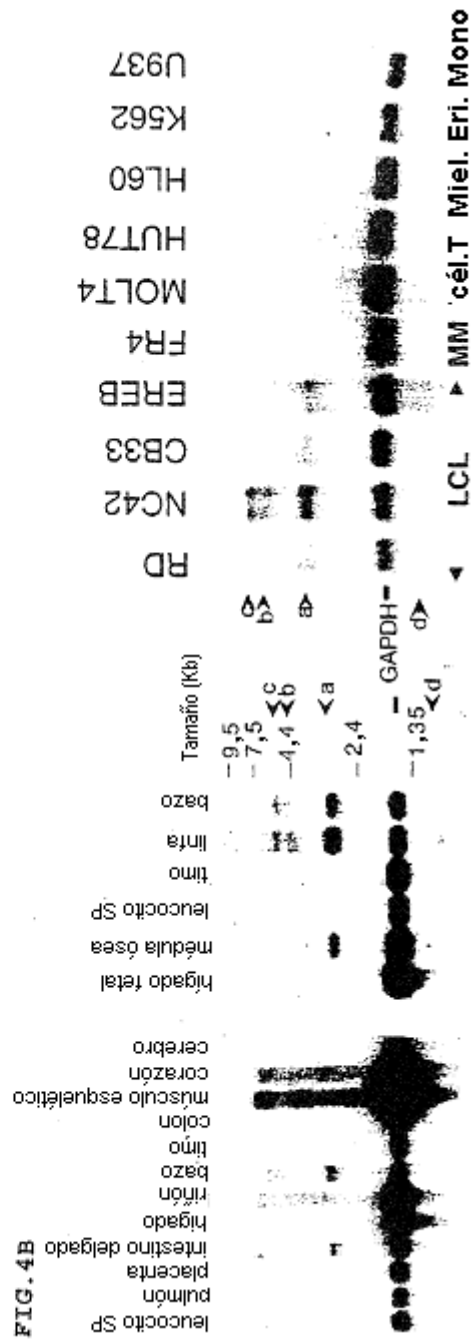
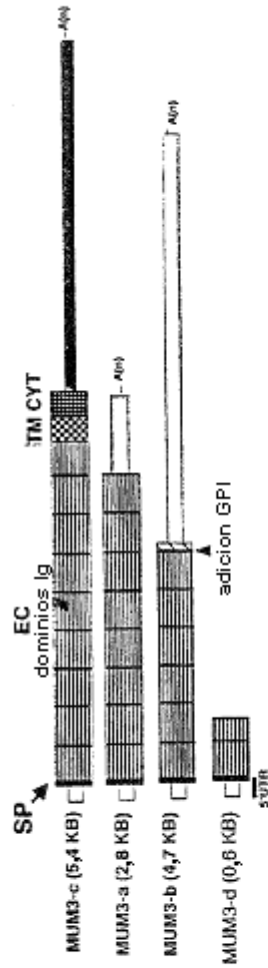


FIGURA 5

1 CTCATTCAGCTTTATGACAGAGAAGAAGCTTACTGAGCTCACTGCTGGTCTGGTGTAGGCAAGTCTGCTTGGCCAA
MLLWAS
78 TCTGSGCTGACCTGGCTTGTCTCCTCAGAACTCCTTCTCCAACCTCGAGCAGGCTTCCATGCTGCTGGGGCGTCC
LLAFAPVVCQQAASAAAHKPVISVHPPWT 12
155 TTGCTGGCTTGTGCTCAGTCTGTGGACAATCTGCAGCTGCACACAACCTGTGATTTCCGTCCCTCCTCCATGGAC
TGTGFKGGRVRLTCTCNGFQFYATTEKTTWY 58
232 CACATTTCTCAAAGGAGAGAGAGTACTCTGACTTGCATATGGATTTGAGTTCTATGACACAGAGAAAACAACATGGT
HRHWYWGEEKLLTPTGNTLEVRRESGLY 93
309 ATCATCGGCACTGGGGAAGAAAGTTGACCTGACCCGAGAAACACCTCGAGGTTCCGGGAATCTGGACTGTAC
RCQACARGSPRNSPVRLLFTSSDSLILQA 10
386 AGATGCCAGGCCCGGGGCTCCCAAGAGTAAGCTGTGGCGCTTGTCTTCTTCAGACTCCTTAATCTGCAGGCG
PYSVFEGDGLVLRCHRRRKEKELTAVK 135
463 ACCATATTCTGTTGTAAGGTCACACATTGGTCTGAGATGCCACAGAGAAGGAGAATTGACTGCTGTGA
YATWNGNINLSNKSNDLLIPQASNS 160
540 AATATCTTGAATGGAAACATCTTTCCATTTCTAATAAAGAGCTGGGATCTCTTATCCCAAGCAAGATTTCAAAT
NNGNYRCIGYGDENDVFRSNFKIKI 186
617 AACAAATGGCAATTCGATGCAATTTGGATATGGAGATGAGAATGATGATTTAGATCAAATTTCAAATAAATTAATAAT
QZELFFPHPELKATDSQPTEGNSVNLSC 212
694 TCAAGAACTATTTCACATCCAGAGCTGAAAGCTACAGACTTCAGCTACAGAGGGGAATTCGTAAACCTGAGCT
ETQQLPFFERGGSDTPLHFNFFPRDGEVIL 237
771 GTGAACACAGCTTCTCCAGAGCGGTGACAGACCCCACTTCACTTCAACTTCTTCAGAGATGGCGAGGTCACTCTG
SDNSTYPELQLPTVWRENSGSYWGCA 263
848 TCAGACTGGAGACGCTACCGGGAATCCAGCTCCCAACCGTCTGGAGAGAAATCAGGATCCTATTTGGTGTGGTGGC
ETVRGNINHKHSPSLQIHWQRIPIVSGV 289
925 TGAACACAGTGGGGTAACATCCACAAGCACAGTCCCTCGCTACAGATCCATGTGCGAGCGGATCCCTGTGTCTGGG
LLELTQPSGGGQAVEGGEMLVLCVSGAE 314
1002 TGCTCCTGGAGACCCAGCCTCAGSGGGCGAGGCTGTGTAAGGGGAGATGCTGGTCTGTCTGCTCCGTGGCTGAA
GTGACGDTTCTFSSWHREDMQESLGTGRKTRQSL 340
1079 GGCACAGGGATACCACTTCTCTGACCCAGGAGGACATCAGGAGAGTCTGGGAGGAAATCAGGCTTCCCT
RAELELPAIRQSHAGGYCYCTADNSY 366
1156 GAGAGCAGAGCTGGAGCTCCCTGCCATCAGACAGGACCATCAGGGGGGACTACTGTACAGAGCAGACAACAGCTACG
PVQGSMLVNLVTVRETPGNRNDGLVAAG 391
1233 GCGCTGTCCAGAGCATGGTGTGAATGTCACTGTGAGAGAGACCCAGGCCAACAGAGATCGCCTTGTGCGCGCGGA
ATGGLGLLAVALLFLFHGWRRRKS 417
1310 GCCATCTGAGGGCTGTCACTGCTCTTCTCTGGCTGTGGCGCTGCTGTTCACTGCTGGGCTCGAGAGGAAGTCAGG
VGFGLTGDEETRLPPAPPGPGGESSHSICPA 443
1387 AGTTGGTTCTTGGGAGAGAAACAGGCTCCCTCCGCTCCAGGCCAGGAGAGTCTCCCATTCCTCTGCGCTG
QVELQSLVVDVHMPKKGDLYXERRIQT 468
1464 CCCAGTGGAGCTTCAGTGTGTATGTGTATGATACACCCAAAAGGGAGATTGGTATCTCTGAGATTCAGAGT
TQLGEEEAANTSRTSLLEDKDVSVVYS 494
1541 ACTCAGCTGGGAGAAGAAGAGGAAGCTAATACCTCCAGGACACTTCTAGAGGATAAGGATGTCTCAGTTGTCTACTC
E V K T Q H P D N S A G K I S S K D E E S * 515
1618 TGAGGTAAGACACAAACACAGGATAACTCAGCTGGAAGAATCAGCTCTAAGAGTGAAGAAGTTAAGAGAAATGAAA
1695 AGTTACGGGAAGCTCTACTCATGTGATTTCTCCCTTGTCCAAAGTCCCAGGCCAGTGCAGTCTTGGCGGCACCTG
1772 GAATGATCACTCATTCAGCTTCTTAATTTCTTCATGCAATATGCATTCCTCCTCCGGAATCACTCATTCGTCTACT
1849 CTGATTTTGGGATGGAATGGCCCTGCAAGAGACTTCACTAAAGTACAGGATCCACAGTTAAGAGAAGACCTGTGAT
1926 TATTTGTGTGGGCTGACCTAATGATCTCCCTAGGGTTCGCTTTAGAGAAGGGGATTAAGAGAGAGAGGAGTCTG
2003 TATGAAAACAGAGGACAAATTTGGTGAATGGGATTCGAGAGATGAAAAGACTGGGTGACCTGGATCTGCTGC
2080 TTAAATACATCTACAACCATGTGCTCACTGAGAGACTCACTTGCATCAGTTTGTAACTGTGAGTGGCTGCACAGGCA
2157 CTGTGCAAAAGTGAAGAAAGCCCTTCACTTGTGCTGCACAGCTTACACTGTGAGAGTTCACTGTGAGATTAAGAA
2234 CCCATCTGGAATGGTTTACAGAGAGAGGAAATTAAGAGAGGACATCAGAGAGCTGGAGATGCTCAGCTCTTACGCTGC
2311 GCTTCCAAAAGCAAAATGATAAATATGTTAATGTTCATTAGTGACAAAGATTTGCAACATTAAGAGAAAAGAGACAAA
2388 TATATAAATTTAAAACCTTAAGTACCACTCTCCAAAGCTAAATTTGAACTTAAAATATTAAGTAACTCATATAAA
2465 CTCTGCGCTTTAATATAAAAAAAGAAAAA

FIGURA 6a

1 CGGTGACAGTGTCTGACGTGTAAGATCAAGTCCAAACCTGTTTGGAAATGAGGAAACTCTCTTTTGATCTCAGCCCTTG
H L L W V I L L V L A P U S G V Q F A R T P R 22
81 GTGGTCCAGGTCTTCAAGCTGCTGTGGGTGATATTTACTGGTCTGGCTCTGTGACGTGGACAGTTTGCAGGACACCCAG
P I I F L Q P P W T T V F Q G E R V T L T C K G F R F 49
161 GCCCATTATTTTCTCCAGCCTCCATGGACACAGTCTTCCAGGAGAGAGAGTGACCCCTCACTTGCAGGGGATTTCGCT
Y S P Q K T K W Y H R Y L G K E I L R E T P D N I L 75
241 TCTACTCACCAGAAACAAATGGTACCATCGGTACCTTGGGAAAGAAATACTAAGAGAAACCCAGACATATCCTT
E V Q E S Q L S L E R S A Q Q S P L S S P V H L D F S S 102
321 GAGGTTGAGGAATCTGAGAGTACAGATGCCAGGCCAGGGCTCCCTCTCAGTAGCCCTGTGCACTTGGATTCTTCTTC
A S L I L Q A P L S V F E G D S V V L R C R A K A E V 129
401 AGCTTCGCTGATCTGCAAGCTCCACTTTCTGTGTTTGAAGGAGACTCTGTGGTCTGAGGTGCCGGGCAAGGCGGAAG
T L N N T I Y K N D N V L A F L N K R T D F H I P H 155
481 TAACACTGAATAATCTATTTACAGAAATGATAATGTCTGGCATTCTTAAATAAAGAACTGACTTCCATATTCCTCAT
A C L K D N G A Y R C T G Y K E S C C P V S S N T V K 182
561 GCATGTCTCAAGGACAAATGGTGCATATCGCTGACTGAGATATAAGGAAAGTCTTGGCCCTGTTCTTCCAAATACAGTCAA
I Q V Q E P F T R P V L R A S S F Q P I S G N P V T 209
641 AATCCAGTCCAGAGGCCATTTACAGCTCCAGTGTCTGAGAGGCCAGCTCCTCCAGCCCATCAGCGGGAACCCAGTGACCC
T C E T Q L S L E R S D V P L R F R F P R D D Q T L 235
721 TGACCTGTGAGAACCCAGCTCTCTCTAGAGAGGTCAGATGTCCCGCTCCGGTTCGGCTTCTTCAAGATGACCCAGACCCCTG
G L G W S L S P N F O I T A M W S K D S G F Y W C K A 262
801 GGATTAGGCTGGAGTCTCTCCCGAAATTTCCAGATTACTGCCATGTGGAGTAAAGATTGAGGGTCTACTGGTGTAAAGG
A T M P H S V I S D S P R S W I Q V Q I P A S H P V 289
881 AGCAACAATGCTTCACTGAGGCTCATATCTGACAGCCGAGAGTCTGAGTACAGGTGAGATCCCTGCATCTCATCTCTCTC
T L S P E K A L N F E G T K V T L H C E T Q E D S L 315
961 TCACCTCTCAGCCCTGAAAGCTCTGAAATTTGAGGGAACCAAGGTGACACTTCACTGTGAAACCCAGGAAGATTCTCTG
R T L Y R F Y H E G V P L R H K S V R C E R G A S I S 342
1041 CGCAGCTTGTACAGGTTTATCATGAGGGTGTCCCGCTGAGGCACAAGTCACTCCGCTGTGAAAGGGGAGCATCCATCAG
F S L T T E N S G N Y V C T A D N G L G A K P S K A V 369
1121 CTTCTCACTGACTACAGAGAATTGAGGGAACACTACTGACAGCTGACAAATGGCTTGGCGGCAAGGCCAGTAAGGCTG
S L S V T V P V S H P V L N L S L E D L I F E G A 395
1201 TGAGCCTCTCAGTCACTGTTTCCCGTGTCTCATCTCTCTCAACCTCAGCTCTCTCTGAGGACCTGATTTTGGAGGGAGCC
K V T L H C E A Q R G S L P I L Y Q P H H E D A A L E 422
1281 AAGGTGACACTTCACTGTGAAAGCCAGAGAGGTTCACTCCCATCTGTACCACTTTCATCATGAGGATCTCCCGCTGGA
R R S A N S A G G V A I S F S L T A E H S G N Y V C 449
1361 GCGTAGGTCGGCAACTGTGAGGAGGAGTGGCCATCAGCTTCTCTCTCACTGACAGGACTTCAAGGAACTACTACTGCA
A D N G F G P Q R S K A V S L S I T V P V S H P V L 475
1441 CAGCTGACAAATGCTTTTGGCCCCAGAGGAGTAAAGGCGGTGAGCCCTCTCCATCACTGTCCCTGTGTCTCATCTCTCTC
T L S S A E A L T F E G A T V T L H C E V Q R G S P Q 502
1521 ACCTCTCAGCTCTGTCTGAGGCGCTGACTTTTGAAGGAGCCACTGTGACACTTCACTGTGAGGTGACAGAGGTTCCCCACA
I L Y Q F Y H E D M P L W S S S T F S V G R V S F S 529
1601 AATCCTATACAGGTTTATCATGAGGAGATGCCCTGTGGAGGAGCTCAACACCTCTGTGGGAAGAGTGTCTCTCAGCT
S L T E G H S G N Y V C T A D N G F G P Q R S E V V 555
1681 TCTCTCTGACTGAGGAGACATTGAGGAATTACTACTGACAGCTGACAAATGGCTTGGTCCCGAGCCAGTGAAGTGGTG
S L F V T V P V S R P I L T L E V P R A Q A V V G D L 582
1761 AGCCTTTTGTCACTGTTCAGCTGTCTCCCGCATCTCAACCTCAGGGTTCAGGAGCCAGGCTGTGGTGGGGAGCT
L E L H C E A P R G S P P I L Y N F Y H E D V T L G 609
1841 GCTGGAGCTTCACTGTGAGGCGCCAGAGAGGCTCTCCCGCAATCTGTACTGGTTTATCATGAGGATGTCAACCTGGGGA
S S A P S G G E A S F N L S L T A E H S G N Y S C E 635
1921 GCAGCTCAGCCCTCTGAGGAGGAGAGCTTCTTTCACTCTCTCTGACTGACAGAACTTCTGAAACTACTCATGTGAG
A N N G L V A Q H S D T I S L S V I V P V S R P I L T 662
2001 GCGCAACAATGGCTAGTGGCCAGGACAGTGCACAAATATCACTCAGTGTATAGTTCAGTATCTCTCTCCATCTCTCAC
F R A P R A Q A V V G D L L E L H C E A L R G S S P 1589
2081 CTTGAGGCTCCAGGCGCCAGGCTGTGGTGGGGAGCTGTGGAGCTTCACTGTGAGGCGCTGAGAGGCTCTCTCCCGA
L Y W F Y H E D V T L G K I S A P S G G G A S F N L 715
2161 TCTGTACTGCTTTATCATGAAGATGTCAACCTGGGTAAAGATCTCAGCCCGCTGTGGAGGAGGGGCTCTCTCAACCTC
S L T T E H S G I Y S C E A D N G L E A Q R S E V T 742
2241 TCTCTGACTACAGAACTTCTGGAATCTACTCTGTGAGGACAGCAATGGTCTGGAGGCCAGGCGAGTGAGATGGTGAC
L K V A G E W A L P T S T S E N 759
2321 ACTGAAATTTGAGGTGAGTGGCCCTGCCACAGGACACATCTGAGAACTGACTGTGGCTGTCTCTCTGAGCTGA
2401 AAATGGAGCCACAGAGCTCTCAGGGCTGTTTGGCTGTGTGGCATCCAGACACTTCTCTCTGAGAACTCTCTCTGTG
2481 AAAGTCTCGGATCTTTCTGTGATGGTTCAGGGAATCTGATGTTTCCAGCAGTCTTCTTGAAGATGATCAAAGCACCTC
2561 ACTAAAAATGCAAAATAGACTTTTGTAGAACATAAATATATCTGAACTGAAATATTACATGAAATGAAACCAAGAA
2641 ATTCTGAGCATATGTTTCTGTGGCGTAGAAAGGATTAAGCTGTTCTGTGGGATTTCTCTCATTTGACTTCTAAGAA
2721 GCTCTACTCTTGAATCTCTTTCATTAAGTGGATGTAATGTTCTTACATTTCCAGATTAAATCCCTATGTTAAGGA
AAAAA

1 CGGTGCAGTGTCTGCACTGTAGATCAAGTCCAAAGCTGTTTGGAAATGAGGAAAGCTCTCTTTTGATCTCAGCCCTTG
N L L M U V L L V L A P V S G Q A F A T R P R
81 GTGTGCCAGGTCTTCAAGCTGCTGTGGTGATATTAAGTCTGTGGCTCTGTGAGTGGACAGTTTGCAGGAGACCTCCAG
P I I F I L Q P P M T C V G G Q G E R V T L T C K G F R F
161 GGCATATATTTTCTCCAGCCTCCATGAGCCACAGCTCTTCAAGGAGAGAGTGAACCTCACTTCCAAAGGGATTTGTGGCT
Y S P Q K T X W Y H R V L G K E I L R E T P D N L
241 TCTCATCACCACAGAAAAAATGATACCTGCTGCTGATCTGGGAAAGAACTAAAGAGAAACCCGACCAATATGCTT
E V T Q E S G E Y R C Q A Q G S P L S S P V H L D F S
321 SAGGTTCCAGGAATCTGGAGAGTACAGATCCAGCCAGGGCTCCCTCTCAGTAGCCCTGTGCACCTTGGATTTTCTCTTC
A S L I L Q A P L S V F E G D S V L R C R A K A E V
401 AGCTTCTGCTGATCTGCAAGCTCCACTCTCTGTGTTTGAAGGAGATCTGTGGTCTTGAGGTGCCGCGAAGGCGGAG
T L N N T I Y K N D N V L A P L N K R T C D F H I P H
481 TAACACTGAATAACTATTTAGCAAGAAATGATATCTCTGGCATCTCTTAATAAAGAACTGACTTCCATATCTCTCAT
A C L K D N G A Y R C T G Y K E S C C P V S S N T V K
561 GGTATGCTCAAGGACAATGTGCATATCTGCTGTACTGGATATAGGAAAGTTGTGTGCGCTGTTTCTTCCAATACAGTCAA
I Q V Q E P F T R P V L R A S S F Q P I S G N P V T L
641 AATCCCAAGTCCAGAGGCAATTTACAGCTCCAGTGTGAGAGCCAGCTCTCTTCCAGCCCATCAGCGGGAACCCAGTGACCC
T C E T Q L S L E R S D V P L R F R F R D D O T L
721 TGACCTGTGAGAGCCAGCTCTCTCTAGAGAGGTGAGATGTGCGCTCGGGTCTGCGCTCTTTCAGAGATGACCAAGCCGTG
G L G W S L S P N F O I T A N W S K D S G F Y W G A
801 GGAATAGCGTGGAGTCTCTCCCGAATTTCCAGATTAAGTCTGAGGATGAGGATTAAGATTCAGGTTCTACTGCTGTAGAGC
A T M P H S V I S D S F R S W I Q V Q I P H S P U L
881 AGCAACAATGCTCCACAGGCTCATATCTGACAGCCGAGATCTGTGATGAGGTGCAAGATCCCTGCATCTCATCTCTCTGTC
T L S P E K A L N F E G T K Y T L H C E T Q E D S L
961 TCACTCTCAGCCTGAAAGGCTCTCAATTTTGAAGGAACCAAGTGCACACTCTCACTGTGAAACCGGGAAGATTTCTCTG
R T L Y R F Y H E G U P L R H K S V R C E R G A S I S
1041 GGCATCTTCTACAGGTTTATCATGAGGCTGTGCGCTCGAGGACAAGTCAAGTCCGCTGTGAAAGGGGAGCATCTCATCAG
F S L T T T E N S G N Y V C T A D N G L O A K P S K A V
1121 CTCTCTACTGACTACAGAGAATTGAGGAACTACTACTGCACAGCTGACAATGGCTTGGCGCCAGCCAGTAAGGCTG
S L S V T V P V S H P V L N L S S P G D L I F E C A
1201 TGAGGCTCTCAGTCACTGTCTCCGTGTCTCATCTGTCTCAACTCAGCTCTCTCTGAGGACCTGATTTTGAAGGGAGCC
K V T L H C E A Q R G S L C P I L Y Q F H N E D A A L E
1281 AAGGTGACACTCTCACTGTGAAGCCGAGAGGCTCACTCCCATCTGTACAGGTTTATCATAGAGGATGTGCGCTTGGG
R R S A N S A G G V A I S F S L T A E H S G N N Y C T
1361 GGTAGGTCGGCAGCTCTGAGAGGAGGTGGCATCAGCTTCTCTGACTGCAGAGCATCTAGGGAAGTCACTACTGCA
A D N G F G F Q R S H A V S L S I T V P V S H P V L
1441 CAGCTGACAATGCTTTGGCGCCGAGCAGTAAGGCGGTGAGCTTCTCCATCACTGTGCTGTGCTCATCTCTCTCTCT
T L S E A E A L T F E G A T V T L H C E V Q R G S P Q
1521 ACCCTCAGCTCTCTGAGGCTGACTTTTGAAGGAGGCACTGTGACACTTCTCACTGTGAAGTCCAGAGAGGTTGCCACAA
I L Y Q P Y H E D H P L W S S T P S V G R V S F S F
1601 AATCTCATACCAAGTTTATCATAGGACAGTCCCTCTGTGGAGGAGCTCAACACCCCTCTGTGGGAAGAGTGTCTCTCAGCT
S L T E G H S G N Y V C T A D N G G P G F P S E V
1681 TCTCTCTGACTGAAGGACATTGAGGAAATTAATCTGACAGCTGACAATGGCTTTGGTCCCGAGCGAGTGAATGGTG
S L P V T V P V S R P L T L R V F R A Q A V V G L T
1761 AGCCTTTTGTGACTGTGTCCAGTGTCTGCCCACTCTGCACTCTCAGGTTTCCAGGCGCCAGGCTGTGGTGGGGAGCT
L E L H C E A P R G G S P P I L Y N V Y H E D V T L G S
1841 GCTGAGGCTTCACTGTGAGGCGCCGAGAGGCTTCTCCCAATCTGTACTGGTTTATCATGAGGATGTGCACTCTGGGA
S E A P S G G E A S F U L S L T A E H S G N Y S C E
1921 GCAGCTCAGCCCCCTCTGGAGGAGAGGCTTCTTCAAGCTCTCTGACTGCAGAGCATCTGGAAACTACTCATGTGAG
A N N G L V A Q H S D T I S L S V I V P V S R P L T
2001 GGCACAATGCTCTAGTGGCCAGGACAGTGACACAATATCACTCAGTGTATTAGTTCAGCATCTCTGTCCTCATCTCAC
F R A P R A Q A V V G D L L E L H C E A L R G S S P I
2081 CTTCAAGGCTCCGAGGCTCAGGCTGTGGTGGGGAGCTCTGAGACTTCACTGTGAGGCTCTGACAGGCTCTCTCCCCAA
L Y M F Y H E D V T L G K I S A P E G G G A S F N L
2161 TCTCTACTGTGTTTATATAGAGATGTACCTCGGTGAAGATCTCAGCCCCCTCTGGAGGAGGGGCTCTCTCAACCTC
S L T T H E G I Y S C E A D N G L E A Q R S E M V T
2241 TCTCTGACTACAGAACTATGTGAATCTACTCTGTGAGGAGAGCAATGGTCTGGAGGCCAGCGAGTGAGTGGTGAC
L K V A V P V S R P V L T L R A F T C T A A V G D L T
2321 ACTGAAGAAGTGTGAGTGTGCGGTGTCTGCGCGGTCTCAGCTCAGGCTCCCGGAGACCATCTGTGCGGTGGGGAGCTG
S L H C E A L R G S P L I L Y R F F H E D V T L G N
2401 TGGAGCTCTGTGAGGCGCTGAGAGGCTCTGCGCTGATCTGTATACCGTTTTCATGAGGATGTCACTCTAGGAAT
R S S F S G G A S L N L S L T A E H S G N Y S C E A D
2481 AGGTGCTCCCGCTCTGGAGGAGGCTCTTAAGCTCTCTGACTGCAGAGCATCTGGAAACTACTCTCTGAGGCGGCA
N G L G A O R S E T T I L I T G L T A N K S G P F A
2561 CAATGCGCTCCGCGCCAGGCGAGTGAGACAGTGACACTTTATATACAGGCTGACCCCGAACAGAAAGTGGCCCTTTTG
T G V A G G C C C A G A A A L L I L Y V N S R
2641 CCAAGGAGTGTGCGGGGCTCTCTGACATAGCAGGCTTCTGCTGGGGGAGCTGCTGCTCTCTGCTGCTGCTGCTGCA
K A G R K P A S D P A S P S D S D S Q E P T Y H V V
2721 AAAGCAGGGAGAAAGCTCTGCTGACCCCGGAGGAGCGCTTCAAGCTGAGGATCCGAGAGCCCACTATCACAATGT
P A W E L Q F V Y T T A N P R G E N V V E E N R I
2801 ACCAGCTCTGGGAAGAGCTCAACACAGTGTACACTTAATCAAACTCTAGAGGAGAAAGATGTGGTTTACTCAGAAGTACGGA
T O E K K H A V A S D G F R H L N K G S P I I Y S
2881 TCACTCAAGAGAAAAAAGCAATGAGTGGCTGTGAGCCGAGGCACTTCAGGAAGAAAGGTTGCCCTTATCATCTACTCT

FIGURA 6c-2

E V K V A S T P V S G S L F L A S S A P H R * stop 977
 2951 GAAGTTAAGGTGGCGTCAACCCCGGTTTCCCGATCCCTGTTCTTGGCTTCCCTCAGCTCCTCAGAGATGAGTCCACACGTC
 3041 TCTCCAACTGCTGTTTCAGCCTCTGCACCCCAAGTTCCTCTTGGGGGAGAGCAGCATTTGAAGTGGGAAGATTAGGCT
 3121 GCGCCAGACCATATCTACTGGGCTTTGTTTACATGTCTCTCAATCTCAGTCTGACAGAAATGCAGGGCCCTGCTGGACTG
 3201 TCACCTGTTTCCCAAGTTAAAGCCCTGACTGGCAGGTTTAAATCCAGTGGCAAGGTGCTCCCACTCCAGGGCCAGCAC
 3281 ATCTCTGGATTCTTAGTGGGCTTCAGCTGTGGTTGCTGTTCTGAGTACTGCTCTCATCACACCCCCACAGAGGGGTC
 3361 TTACCACACAAAAGGAGAGTGGGCTTCAGGAGATGCCGGGCTGGCTAACAGCTCAGGCTCTCTTAAACTCCGACACAG
 3441 AGTTCTGCTTTGGGTGGATGCAATTTCTCAATTTGTCATCAGCTGGTGGGGCTACTGAGTGTGCTGCCAAATGGGACAG
 3521 CACACAGCCTGTGCACATGGGACATGTGATGGGTCTCCCAAGGGGCTGGCTTTCACACTCTCTCACTCTCTCAAACT
 3601 CTAGGTCGGCACTTGACACCAAGGTAACTTCTCTCTGCTCATGTGTGAGTGTCTACTGCCCCAAGTAAGTGGCTTTCA
 3681 TACACCAAGTCCCGAAGTTCTTCCCATCTAACAGAGTAACCCAGCAAGTCAGGCCAGGAGGACCCAGGGTGCAGACA
 3761 GAACACATACTGGAAACACAGGAGGTGCTCAATTACTATTTGACTGACTGACTGAATGAATGAATGAATGAGGAAGAAAAC
 3841 TGTGGTAATCAAACTGGCATAAATCCAGTGCCTCTAGGAAATCCGGGAGGTATTCTGGCTTCTAAGAAACAAAG
 3921 GAAGAGAGAGGAGCTTGGATGAAGAACTGTCAGCAAGAGAGGGCTTCTTCACTCTTTATGTGCTTTGGATCACCT
 4001 GAGGATCTGTGAAGATACAGATACTGATTGAGTGGTCTGTGTAGAGCTGAGACTGCCATTCTAACATGTTCCAGGGG
 4081 ATGCTGATCTGCTGGGCTTGGGACTGCACTGCATGTCATGTGAAGCCCTATAGGTCTCAGCAGAGGCCCATGGAGAGGA
 4161 ATGTGTGGCTCTGGCTGCCAGGGCCCACTCGGTTACACAGGATCTGTGCTCTCTCCCTGGCCAGGCTTTGGCCACAGCAC
 4241 CAGCAGCTGCTGTGCTGAGAGAGGTTCTTCTCTGTGACATGTTGGCTTTCATCAGCCACCTTGGGAAGCGGAAAGTAGC
 4321 TGCCACTATCTTTGTTTCCCACTCAGGGCTCACACTTTGCCATGAAAAGGGTGAATGTATATTAACCTGAGCCCTCTCC
 4401 ATTCAAGATTGTTCTCCCATCTCTGAGCAATGGGATGTTCTGTTCCGCTTTTATGATATCCATCACATCTTATCTTGATC
 4481 TTTGCTCCCAAGTGGATTGTACAGTGTACTTTTAAGCCCCACGGCCCTGAAATAAATCTTCCAAAGGCAATGGGAAGC
 4561 TCACTCCAGCTGAACCATGGCTTTTCATGCTTCCAAGTGTGAGGGCTTGCCCAAGATAGACAGGGCTGACTCTGCTGCC
 4641 CAACCTTTCAAGGAGGAAACACAGACCTGAGACAGGAGCTGTATGCAAGCCAGTGCAGCTTTGCAGAGGACAAAGCTG
 4721 GAGGCATTTTTCATCACTACAGATATGCAACTAAATAGAGCTGGAGCAAGGAAATGCAATTTCCACCGAGGCGGCTTTT
 4801 TTAGGCTAGTTGAAAGTCAAGAGGACAGCAGCAAGCATAGGCTCAGGATTAAAGAAAAAATCTGCTCAGCTCTGTT
 4881 CTGGAGGTCACATCACCAACAAAGCTCACGGCCATATSCAGTTCTGAGAAAGTGGAGGACACAGGCTCAAAAGAGGAAAT
 4961 TAGAAATTTCTCATTTGGGAGATTAAGGTACCCCATCCCAAGATGATAACTGCACAGTGGCAGAACAACTCCACCTAAT
 5041 GTGGGTGGACCCCATCCAGTCTGTGTAAGGCTGAATGTAAACAAAGGGCTTATTTCTTCTCAAGTAAGGGGAACTCCT
 5121 GCTTTGGGCTGGGACATAAGTTTTTCTGCTTTCCAGCGCAAACTGAAAAATGGCTCTTCTTGGGTCTTGAGCTTCTGCGC
 5201 ATATGGAGTGAAGAACTATGCTATTGGATCTCTGGATCTCCAGCTTGGTACTGCAGATCTTGAGATATGTAGCCT
 5281 CTACAGTCACAGAGCTAATTCATTTCAATAAACCAATCTTTC

FIGURA 8A

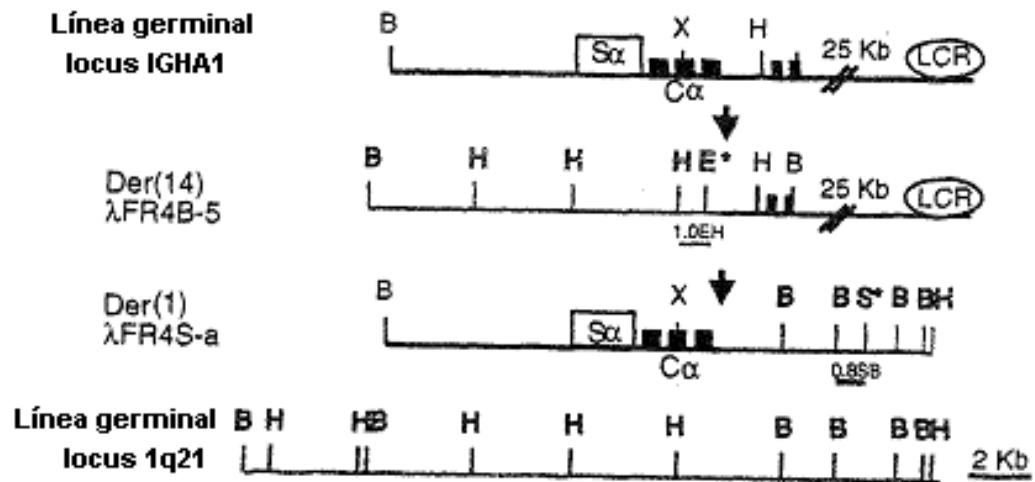


FIG. 8B

FIGURA 8

Chr	14	TCCCACTGACGCATGCAGGAAGGGGCACCTCCCTTAAC	CACACTGCTCTGTACGGGGCACAGGTGCACACTCACACTCACA
Der	(14)	GGCCTGACAGCAACTTTCTTCTACTAGTTCATCTTAA	CACACTGCTCTGTACGGGGCACAGGTGCACACTCACACTCACA
Chr	1	GGCCTGACAGCAACTTTCTTCTACTAGTTCATCTTAA	CTTTATCCTGCTAACTGGGGGACACCTGCTCTTAAGTAACTGAAGGGAAA
Chr	1	GGCCTGACAGCAACTTTCTTCTACTAGTTCATCTTAA	CTTTATCCTGCTAACTGGGGGACACCTGCTCTTAAGTAACTGAAGGGAAA
Der	(1)	TCCCACTGACGCA	-----GGAAGGATCTTAA
Chr	14	TCCCACTGACGCA	TGCAGGAGGGGGGACCTCCCTTAAC

8c

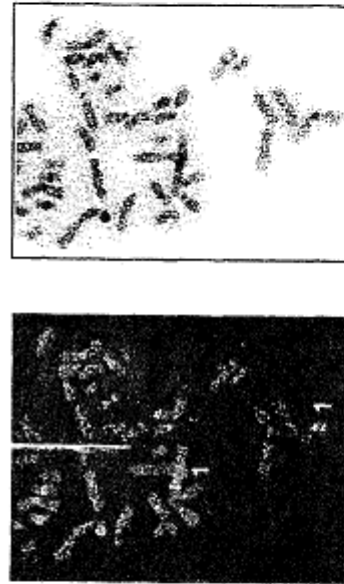


FIGURA 9

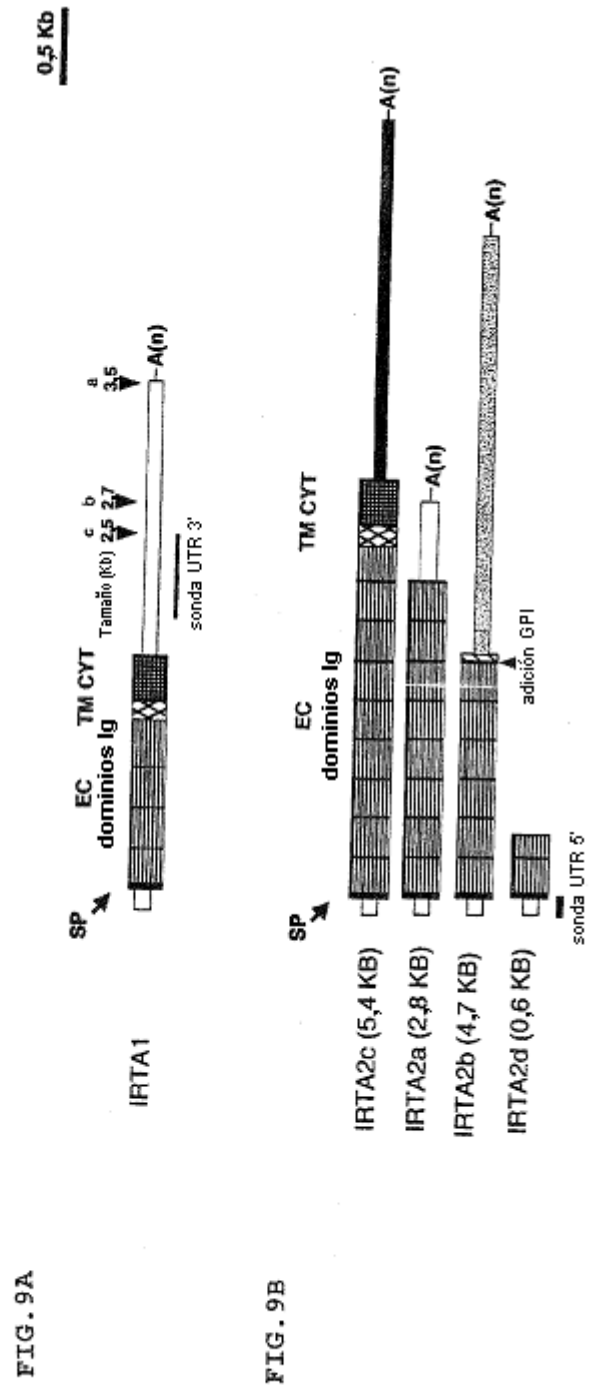


FIGURA 11A

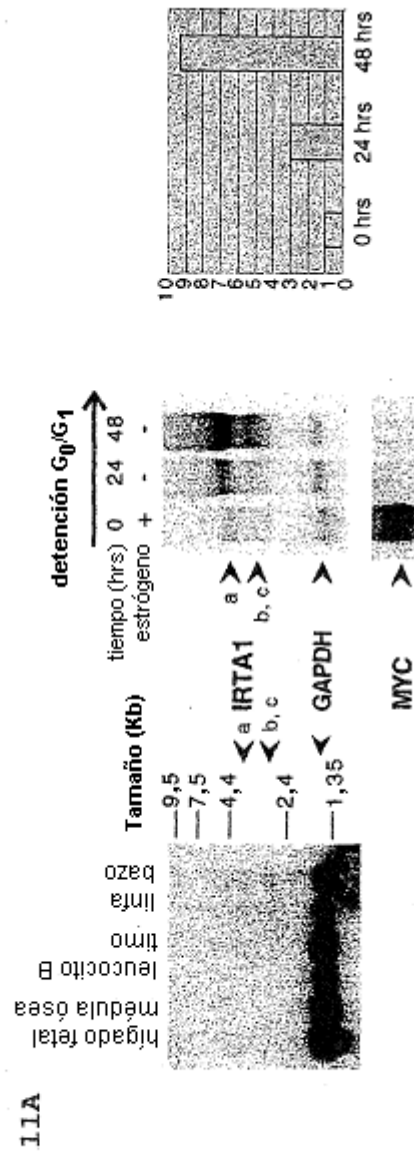


FIGURA 11B1-B4

11B1-4

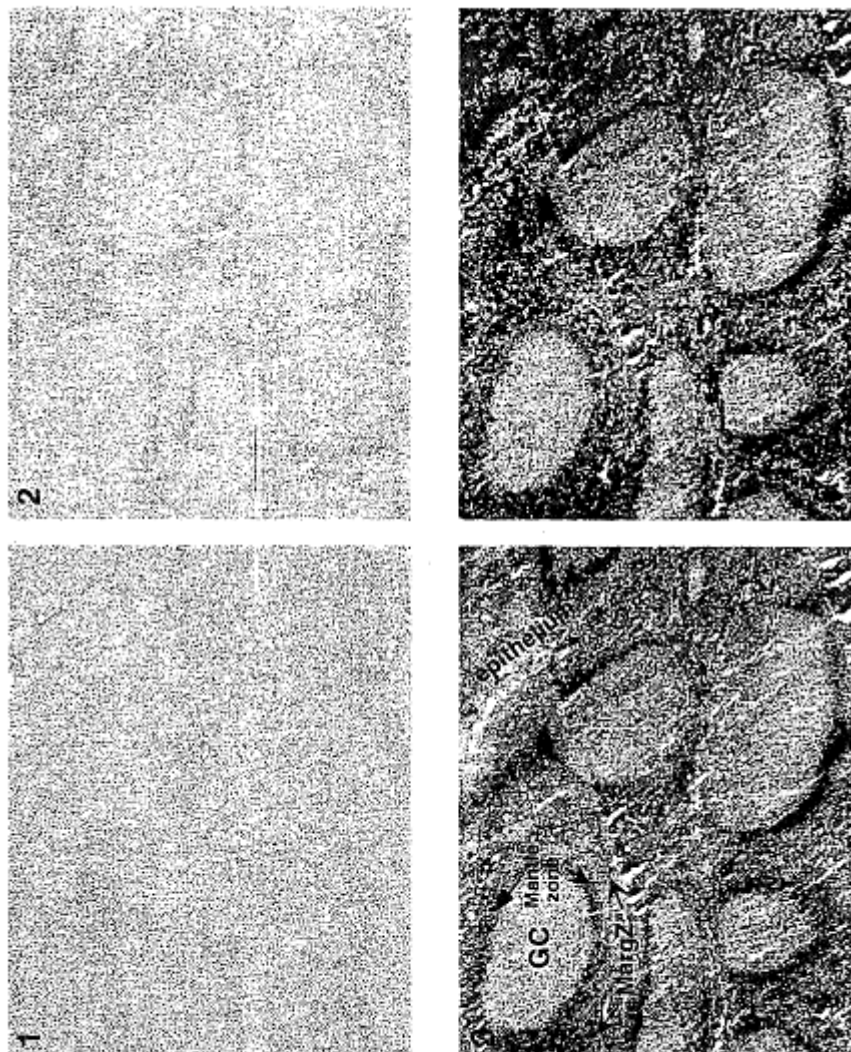


FIGURA 12A

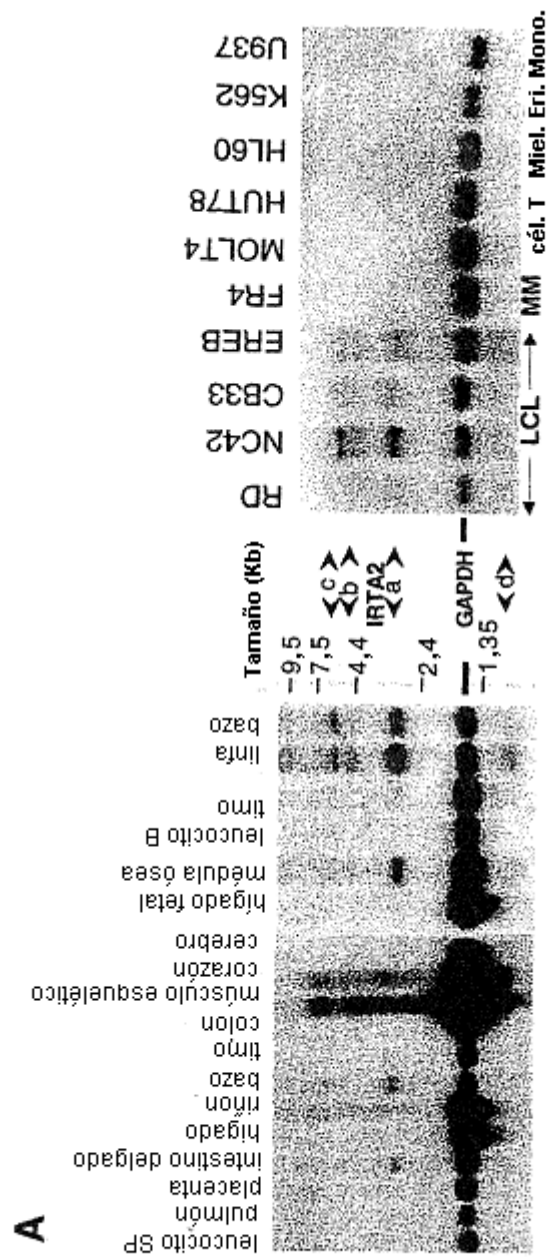


FIGURA 12B1-B4

B

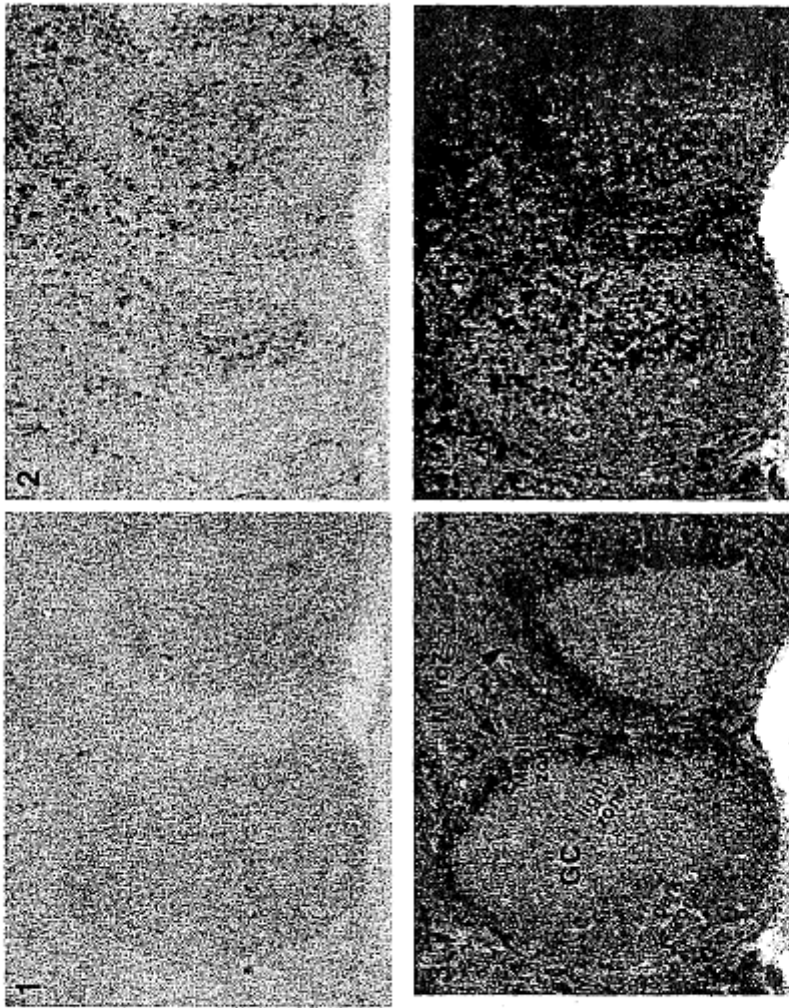


FIGURA 14

FIG. 14A

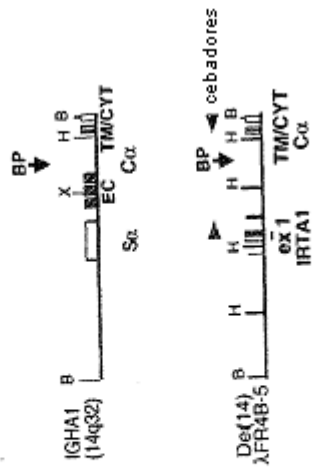


FIG. 14C



FIG. 14B

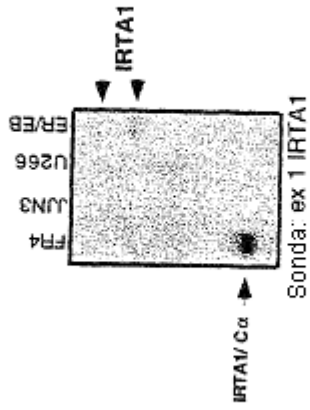
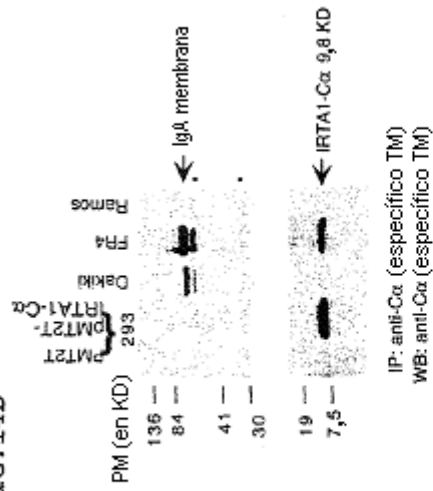


FIG. 14D



* Cadenas Pesada (arriba) y Ligera (abajo) de IgG1

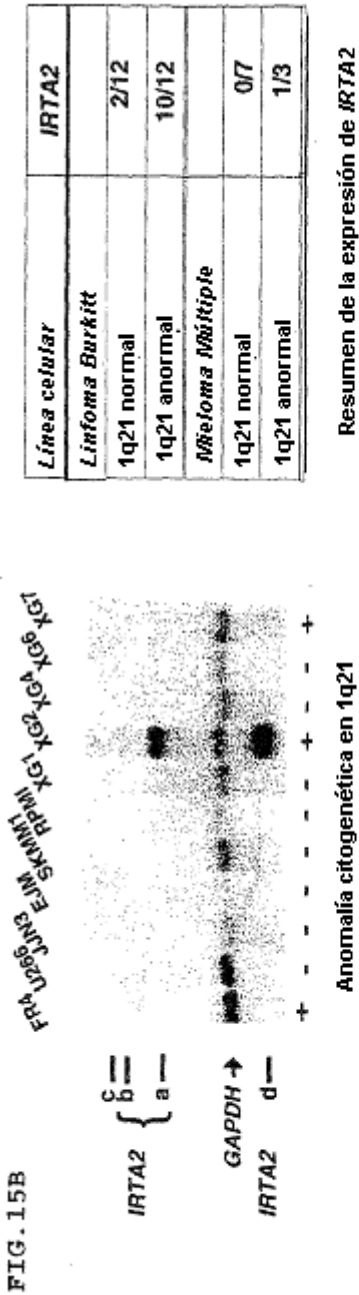
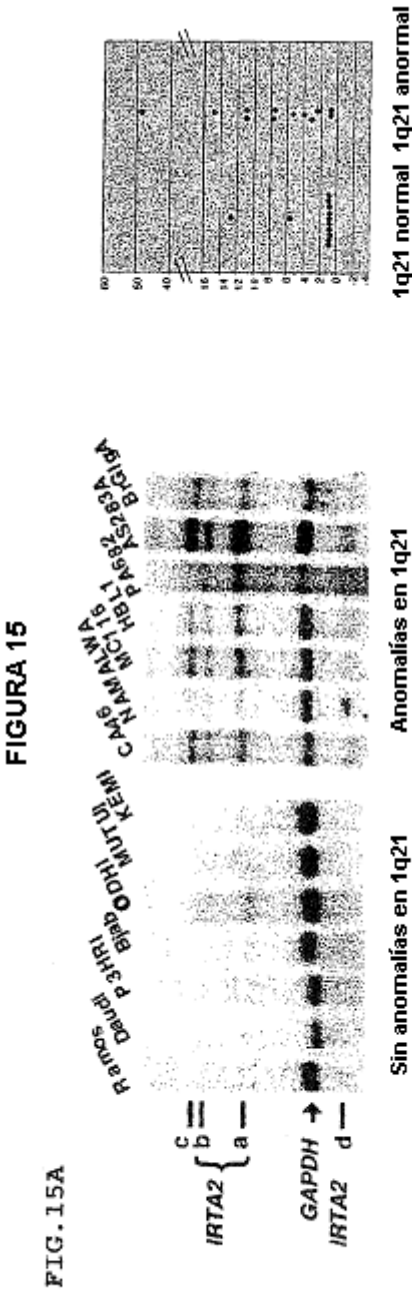


FIGURA 16 - 1 -16 - 4
Expresión de IRTA1 en tejido linfóide normal

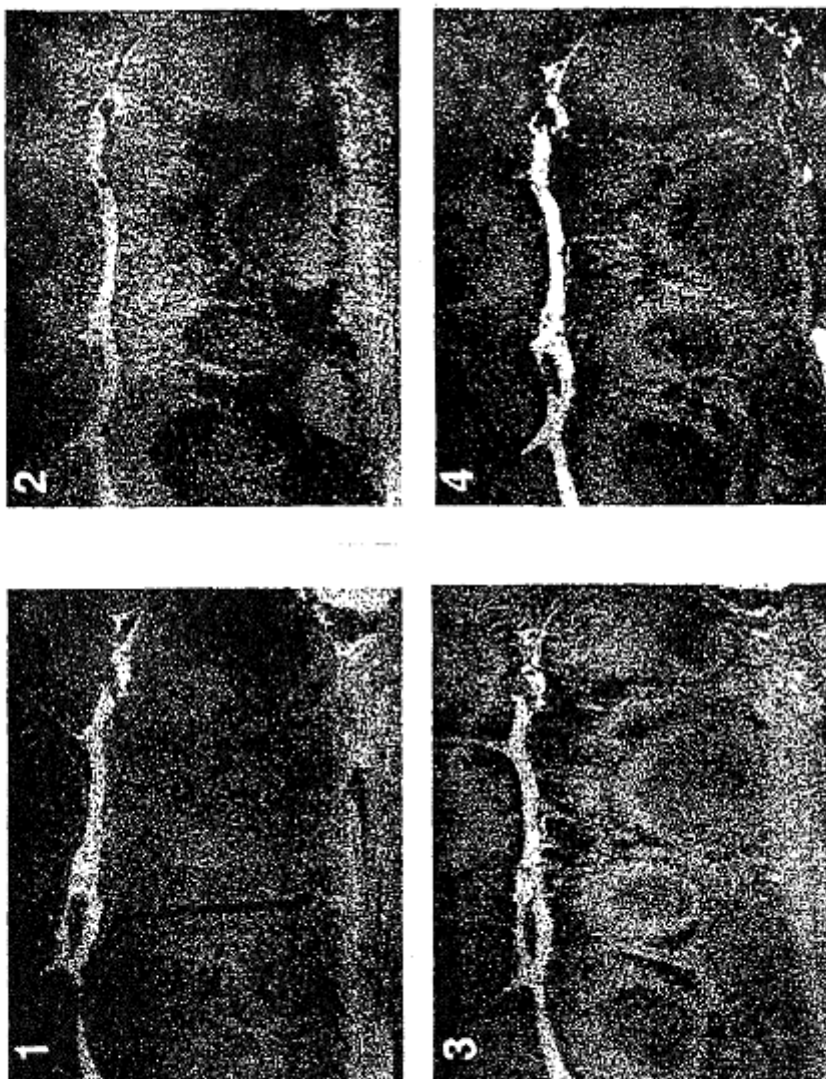


FIGURA 17

Expresión de IRTA1 en linfoma de células T de Tejido Linfoide Asociado a la Mucosa del Estomago

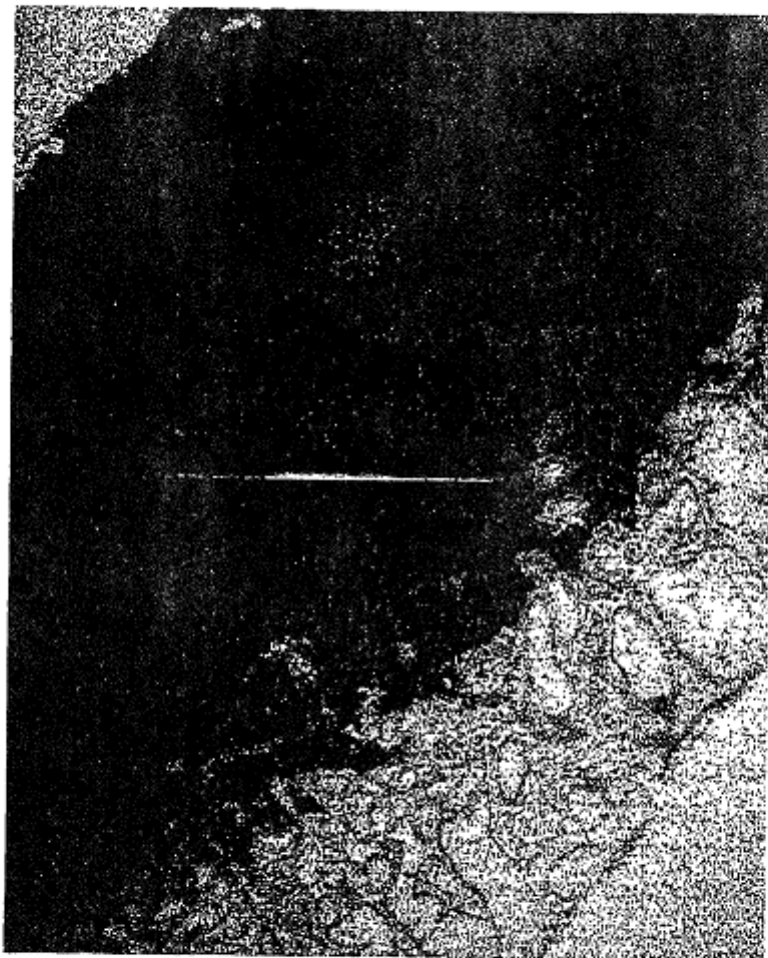


FIGURA 18B-1

IRTA2A	-----	
IRTA2C	-----	
IRTA2B	1	CGGTGCAGTGTCTGACTGTAGATCAAGTCCAAAGCTGTTTGGAAATGAGGAAACTTCTCTTTGATCTCAGCCCTTG
		M L L N V I L L V L A P V S G Q F A R T P R 22
81		GTGGTCCAGGTCTTCAGTGTCTGCTGGGTGATATTACTGGTCTCTGGCTCCCTGTCAGTGGACAGTTTTCAGGACACCCAG
		P I I F L Q P P W T T V F Q G G E R V T L T C K G F R F 49
161		GCCCATTTATTTCTCCAGCTCCATGGACACAGTCTTCCAGGAGAGAGAGTGACCTTCAGTTGCAAGGGATTTGGCT
		Y S P Q X T K W Y M R Y L G K E I L R E T P D N I L 75
241		TCTACTCACCACAGAAACAAATGGTACCATCGGTACCTTGGGAAAGAAATCTAAGAGAAACCCAGACAATATCCTT
		E V Q E S G E Y R C Q A Q G S P L S S P V H L D F S S 102
321		GAGGTTCCAGGAATCTGGAGAGTACAGATGCCAGGCCAGGGCTCCCTCTCAGTAGCCCTGTGCACTTGGATTTTCTTC
		A S L L L Q A P L S V F E G D S V V L R C R A K A E V 129
401		AGCTTCGCTGATCTTCAAGCTCCACTTCTGTGTGTGAAGGAGACTCTGTGGTCTGAGGTGCCGGGCAAGGGCGGAG
		T L I Y R N D N V L A F L N K R T D P H I P H 155
481		TAACACTGATTAATCTATTACAGAAATGATATGCTCTGGCATTCCTTAATAAJAAGTCACTTCCATATTCCTCAT
		A C L K D N G A Y R C T G Y K E S C C P V S S N T V K 182
561		GCTGTCTCAAGGACAAATGGTGCATATCGCTGTACTGGATATAAGGAAAGTTGTTCGCTGTTCCTCCATACAGTCAA
		I Q V Q E P F T R P V L R A S S F Q P I S G N P V T L 209
641		AATCCAAATCCCAAGAGCCATTACAGCPCAGCTGCTGAGAGCCAGCTCTTCCAGCCCATCAGCCGGAACCCAGTGACCC
		T C E T Q L S L E R S D V P L K F R F F R D D Q T L 235
721		TGACCTGTGAGAGCCAGCTCTCTCTAGAGAGGTGAGATGTCCGCTCCGGTTCCGCTTCCTTCAGAGATGACCCAGACCTG
		G L G T L S P N P O I T A N W S K D S G F Y W C K A 262
801		GGATTAGGCTGGATCTCTCCCGAATTTCCAGATTACTGCCATGTGGAGTAAAGATTCCAGGGTCTCTACTGGTGTAAAGC
		A T M P H S V I E D S F R S M I Q V D I P A S H P V L 289
881		AGCAATGCTCCACAGGGTCATATCTGACAGCCGAGATCTGGATACAGGTGACAGATCCCTCATCTCATCTCTGTC
		T L E P E K A L N F E G T K V T L H C E T Q E D S L 315
961		TCACTGTGAGCCCTGAAAGGCTCTGAATTTTGGGGAAACCAAGGTGACACTTCACTGTGAAAGCCAGGAAGATTCTCTG
		R T L Y R F Y N E G V P L R H K S V R C E R G A S I S 342
1041		CGCACTTTGTACAGGTTTATCATGAGGGTGTCCCGCTGAGGACACAGTCACTGCTGTGAAAGGGGAGCATCCATCAG
		I S L T T E N S G N Y Y C T A D N G L G A K P S K A V 369
1121		CTTCTCACTGACTACAGAGAATTCAGGGAACTACTACTGACAGCTGACAAATGGCTTGGCGCAAGGCCAATAGGCTG
		S L S V T V P V S H P V L S P E D L I F E G A 395
1201		TGAGCTGTGAGAGCCAGCTCTCTCTAGAGAGGTGAGATGTCCGCTCCGGTTCCGCTTCCTTCAGAGATGACCCAGACCTG
		K V T L K C E A Q R G S L P I L Y Q P H N E D A A L E 422
1281		AAGGTGAGCTTCACTGTGAGGCGGAGAGAGGTCTCTCTCCATCTGTTACCACTTTCATCATGAGGATCTGCGCTGGA
		R R S A N S A C G V A I S F S L T A E H S G N Y Y C T 449
1361		GCCTAGTCTGCGCAACTCTGACAGAGAGTGGCCATCAGCTTCTCTCTGACTGACAGCACTCAGGGAAGTCTACTGCA
		A D N G F G F Q R S K A V S L S I T V P V S H P V L 475
1441		CAGCTGACAAATGCTTTGGGCCCCAGGCGAGTAAAGGCGGTGAGCTCTCCATCAGTGTCCCTGTGTCTCATCTCTCTCT
		T L S S A E A L T F E G A T V T L H C E V Q R G S P Q 502
1521		ACCTTCAGCTCTGCTGAGGCGGCTGACTTTTGAAGGAGGCACTGTGAGACTTCACTGTGAAAGTCCAGAGAGGTTCCCGACA
		I L Y Q F Y H E D M P L M S S S T P S V G R V S P S P 529
1601		AATTCCTATACCACTTTTATCATGAGGACATGCGGCTGTGAGGAGCTCAACACCTCTCTGTTGGGAGAGTGTCTCTCAGCT
		S L T E G H S G N Y Y C T A D N G F G P Q R S E V V 555
1681		TCTCTCTGACTGAAGGACATTGAGGGAATTACTACTGACAGCTGACAATGGCTTTGGTCCCGAGGCACTGAAGTGGTG
		S L F V T V P V S R P I L T L R V P R A Q A V V G D L 582
2A, 2C1761		AGCCCTTTTGTGCTAGTGTCCAGTGTCTGCGCCCATCTCTCAGCTTCAGGGTTCAGGAGGCCAGGCTGTGGTGGGGAGCT
		G K C M V L A S H P P L A E F S L T H S F K 582
2B 1761		-----GGTAAGTGCTGGGTCTTGGCTAGTCAAGCCAGCCGCTGGCTGAGTCTCTCTCAGCCATTCCTTTAA
		L E L H C E A P R G S P P I L Y W F Y H E D V T L G S 609
2A, 2C1841		GCTGGAGCTTCACTGTGAGGCGGCGAGAGGCTCTCCCGCAATCTGTACTGGTTTATCATGAGGATGTACAGCTGGGGA
		N L F A L S S F L P * stop 592
2B 1841		AAATCTGTTTGCAGTGTCCAGTTTCTCCCGCTAATCAACTTAATCCCTCTCTGCTCTCTCACTAATAGCTGGG
		S S A P S G G E A S F L T A E H S G C E 635
2A, 2C1921		GCAGCTCAGGCGCTCTGGAGGAGAAGCTTCTTCACTCTCTGACTCCAGAACATTCGGATACCTCTCTCTCTCTCTCT
2B 1921		GTITTCCTACTCATAGTCTGCTGCTCAGCCAGACCCCTAAAACAGCTCAGTATATCCCGAGCTTTTACCAATGAATT
		A N N G L V A Q H S D T I S L S V I V F V S R P I L T 662
2A, 2C2001		GCCAACANTGGCTAGTGGCCAGCAGACAGTACACAAATATCACTCAGTGTATAGTTCAGTATCTCTGCTCCATCTCAC
2B 2001		TATTTATTTGATTTCTCTCTCATCTTGTATGTTTCAACAGTACGCCAATTTTCTTGTATGACAGGAGCTGTCTCTACT
		F R A P R A Q A V V G D L L E L H C E A L R G S S P I 689
2A, 2C2081		CTTCAGGGCTCCAGGGCCAGGCTGTGGTGGGGAGCTGTGGAGCTTCACTGTGAGGCGCTGAGAGGCTCTCTCCCAA
2B 2081		TCTCTACTGACATTTACATATTAACTTAGCTACAGCACACTCTATAGATAAATATGGTCAAGACCTTAAATCTCTCA

129

FIGURA 18B - 3

```

4081 ATGCTGATGCTGCTGGCCCTGGGACTGCACTGCATGCATGTGAAGCCCTATAGGTCTCAGCAGAGGCCCATGGAGAGGGA
4161 ATGTGTGGCTCTGGCTGCCAGGGGCCAACTCGGTTTACACGGATCGTGTCTGCTCCCTGGCCAGCCTTTGGCCACAGCAC
4241 CACCAGCTGCTGTGTGTGAGAGAGCTTCTTCTCTGTGACATGTTGGCTTTCATCAGCCACCCCTGGGAAGCGGAAAGTAGC
4321 TGCCACTATCTTTGTTTCCCCACCTCAGGCCTCAGACTTTCCCATGAAAAGGGTGAATGTATATAACCTGAGCCCTCTCC
4401 ATTCAGAGTTGTTCTCCCATCTCTGAGCAATGGGATGTTCTGTTCCGCTTTTATGATATCCATCACAATCTTATCTTGATC
4481 TTTGCTCCAGTGGATTGTACAGTGAAGTCTTTAAGCCCCACGGCCCTGAAATAAAAATCCTTCCAAGGGCATTTGGAAGC
4561 TCACTCCACCTGAACCATGGCTTTTCATGCTTCCAAGTGTGAGGGCCTTGCCAGATAGACAGGGCTGACTCTGCTGCC
4641 CAACCTTTCAAGGAGGAAACAGACACCTGAGACAGGAGCTGTATGCAAGCCAGTGACCCCTTGACAGGACAAAGGCTG
4721 GAGGCATTTGTCTACTACAGATATGCAACTAAATAGACGTGGAGCAAGAGAAATGCAATCCACCCGAGGCCCTTTT
4801 TTAGGCCTAGTTGAAAGTCAAGAAAGGACAGCAGCAAGCATAGGCTCAGGATTAAGAAAGAAAAATCTGCTCAGACTCTGTT
4881 CTGGAGGTCACATCAACCAACAAAGCTCAAGCCCTATGCAATTTCTGAGAGGTGGAGGCCACAGGCTCAAAAGAGGAAAT
4961 TAGAATTTCTCATTTGGAGAGTAAGGTACCCCATCCAGAAATGATAACTGCACAGTGGCAGAACAACTCCACCTTAAT
5041 GTGGGTGGACCCCATCCAGTCTGTTGAAGGCCGTGAATGTAAACAAAGGGCTTATTTCTTCTCAAGTAAGGGGAACTCCT
5121 GCTTTGGGCTGGGACATAAGTTTTCTGCTTTTCAAGCGCAAACTGAAAAATGGCTCTTCTTGGGTCTTGAGCTTGCTGGC
5201 ATATGGAAGTGAAGAACTATGCTATTGGATCTCCTGGATCTCCAGCTTCTGACTGCAGATCTTGAGATATGTCAGCCT
5281 CTACAGTCACAAGAGCTAATTCATTTCTAATAAACCAATCTTTC

```

FIGURA 18C - 1

1 AGTGAAGGGGTTTCCCATATGAAAAATACAGAAAGAAATTTTGAATACTA
 52 GCAGATACAGAACTTGATATTTCTAGAGAACCCAGGCACAGTCTTGGAGAC
 103 ATTACTCTCTGAGAGACTGCAGCTGATGGAGATGAGCCCCAACTTCTAAAA
 154 ATGTATCACTACCGGGATTGAGATACAAACAGCATTTAGGAAGGTCTCATC
 205 TGAGTAGCAGCTTCTGCTCTCTCTCTTGGAGATAAGTGGGCTTTTGGTG
 256 AGACAGACTTTCCCAACCTCTGCCCCGGGTGCCATGCTTCTGTGGCT
 1 M L L W L
 307 GCTGCTCTGATCCTGACTCCTGGAAGAGAACATCAGGGGTGGCCCCAA
 6 L L L I L T P G R E Q S G V A F K
 358 AGCTGTACTTCTCTCAATCTCCATGGTCCACAGCCTTCAAAGGAGAAAA
 23 A V L L L N P P W S T A F K G E K
 409 AGTGGCTCTCATATGCAGCAGCATATCATTCTCCAGCCAGGGAGACAC
 40 V A L I C S S I S H S L A Q G D T
 460 ATATTGGTATCAAGATGAGAAAGTTGTGAAAAATAAACATGACAAAGATCCA
 57 Y W Y H D E K L L K I K H D H I Q
 511 AATTACAGAGCCTGGAAATTACCAATGTAGAGCCCGAGGATCCTCCCTCAG
 74 I T E P G N Y Q C K T R G S S L S
 562 TGATGCCGTGCATGTGGAAITTTACCTGACTGGCTGATCCTGAGGCTTT
 91 D A V H V E F S E D W L I L Q A L
 613 ACATCTCTCTTTGAAGGAGACAAATGTCTTCTGAGATGTGAGGGGAAAGA
 108 H P V F E G D N V I L R C Q C K D
 664 CAACAAACACTCATCAAAGGTTTACTACAAGGATCGAAACAGCTTCC
 125 N K N T H Q K V Y Y K D G K Q L P
 715 TANTAGTTATAATTTAGAGAGATCAGAGTGAATTCAGTCTCCAGGGATAA
 142 N S Y N L E K I T V N S V S R D N
 766 TAGCAATATCATCTGTACTGCTTATAGGAAGTTTACATACTTGACATGGA
 159 S K Y H C T A Y R K F Y I L D I E
 817 AGTAACCTTCAAAACCCCTAAATATCCAAAGTTCAAGAGCTGTCTTCTACATCC
 176 V T S K P L N I Q V Q E L F L H P
 868 TGTCTGAGAGCCAGCTCTTCCACGCCCATAGAGGGAGTCCCATGACCT
 193 V L R A S S S T P I E G S P M T L
 919 GACCTGTGAGAGCCAGCTCTCTCCACAGAGGCCAGATGTCCAGCTGCAAT
 210 T C E T Q L S P Q R P D V Q L O P
 970 CTCCCTCTTCAGAGATAGCCAGACCTCGGATTTGGGCTGGAGAGGTCGCC
 227 S L F R D S Q T L G L G W S R S P
 1021 CAGACTCCAGATCCCTGCCATGTGGACTGAAGACTCAGGGTCTTACTGGTG
 244 R L Q I P A M W T E D S G S Y W C
 1072 TGAGGTGGAGACAGTCACTCACAGCATCAAAAAAGGAGCCTGAGATCTCA
 261 E V E T V T H S I E K R S L R S Q
 1123 GATAGCTGTACAGAGATCCCTGTGTCTAATGTGAATCTAGAGATCCGGCC
 278 I R V Q K V P V S N V N L E I R P
 1174 CACCGAGGGCAGCTGATTTGAAGGAGAAAAATAGGTCTTATTTGCTCAGT
 295 T G G Q L I E G E N M V L I C S V
 1225 AGCCCAAGGTTGAGGACTGTACATTTCTCCTGGCAGAAAGAGAGAGT
 312 A Q G S G T V T F S W H K E G R V
 1276 AAGAGCCTCCGTAGAAAGACCCAGCTTCCCTGTGTGGCAGAGCTGCATGT
 329 R S L G R K T Q R S L L A E L H V
 1327 TCTCACCGTGAAGGAGAGTGATCCAGGGAGATACTACTGTGCAGCTGATAA
 346 L T V K E S D A G R Y Y C A A D N
 1378 CGTTACAGCCCCATCTCAGCACCTGGATTTCGAGTCAACCTGAGAAATTC
 363 V H S P I L S T W I R V T V R I P
 1429 GGTATCTCACCTGTCTCAGCTTCCAGGGCTCCAGGGGCCACACTGTGCT
 380 V S H F V L T P R A P R A H T V V
 1480 GGGGACCTGCTGGAGCTTCACTGTGAGTCCCTGAGAGGCTCTCCCCGAT
 397 G D L L E L H C E S L R C S P P I
 1531 CCTGTACCGATTTTATCATGAGGATGTACCTTGGGGAAACAGCTCAGCCCC
 414 L Y R F Y H E D V T L G N S S A F
 1562 CTCTGGAGGAGGAGCTCCTTCAACCTCTCTCTGACTSCAGAACATTTGG
 431 S G G G A S P N L S L T A E H S G
 1633 AAATACTCTCTGTATGCAGACAAATGGCTTGGGGGCCAGCACAGTCATGG
 448 N Y S C D A D N G L G A Q H S H G
 1684 AGTGAGTCTCAGGCTCACAGTTCCGGTGTCTCGCCCCCTCTCAGCTCAG
 465 V S L R V T V F V S R P V L T L R
 1735 CGCTCCCGGGGCCAGGCTGTGTGGGGGACCTGCTGGAGCTTCACTGTGA
 483 A P G A Q A V V G D L L E L H C E
 1786 GTCCCTGAGAGGCTCCTTCCCGATCTGTACTGGTTTATACAGAGGATGA
 499 S L R G S P P I L Y W F Y H E D D
 1837 CACCTTGGGGAACATCTCGGCCCACTCTGGAGGAGGGCATCTTCAACCT
 516 T L G N I S A H S G G G A S P N L
 1888 CTCTCTGACTACAGAACTTCTGGAACTACTTATGTGAGGCTGACAAATGG
 533 S L T T E H S G N Y S C E A D N G

FIGURA 18C-2

1939 CTTGGGCGCCAGCAGTAAAGTGCTGAACACTCAATGTTCACAGSAACTTC
550 L G A Q H S K V N F T L C N V V T G T C T S
1990 CHGAGAACAGAAGGCCCTTACCCTGC GGGAATACGGVGTCTGCTCTCACG
567 R N R T T G T G L T A A G I T T G L V L S
2041 CATCTCTCGTGCTTGCTGCTGCTGCTCTGCTGCATTACGCCAGVGCCC
584 I L V L A A A A A L L H Y A R A R
2092 AAOGAACCAGGAGGLTCTTCECATGSAACACTTAGTACAGCTCTAG
601 R X P F CG GL S A T G T T S H S F S
2143 TGAGTGTCACAGGAGCTTCCTCTCTCGACGGCTTCAGGATAGACCTCAAGA
618 E C Q EPSSSRSPRSRIDPQE
2194 GCCCACTCACTCAAACACTAGCCCCAATGGAGCTGGAGCCAATTGACAG
635 P T H S K P L A P M E L E P K Y S
2245 CAATGTAAACTCTGGAGATAGCAACCGATTATTATCCAGATCTGGAGCAT
652 N V N P G D A E N T I Y S O I W S J
2296 CCAGATAACAAGGAAGTAACAGCTPAITGTCCAAATGATGCATCAAGAGICA
669 Q H T K E N S A N C P M K H Q E H
2347 TAGGAGAACTTCAGTCTCTATTCAAGACTGAAGAKACHACCCAGACGA
686 Z E B L T V L Y S E L K K T H F D D
2398 CTCTGCAGGGAGGCTAGCAGCAGAGGGCAGGGCCCATGAAGAAGATGTGA
703 S A G E A S S R G R A H E E D G E
2449 AGAAAACTATGAGATGTACCACTGTATTACTGGCTCAGACCACTAGGCC
720 E N Y E N V P R V L L A S D H
2500 CCTTACCAGAGTGCGGCCACAGGAAACAGGCTGCACCAATTTTTTTTCTGT
2551 TCCTTCCAAACACATCATCCACTCTCCAGACTCTGCCCTCTACAGAGGC
2602 TGGGCTGCAGSGTATGTAGGGCTGAGCAAAAGGCTGCACAAATCTCCCCGT
2653 GCCTGTGTCTGTGTTCCCGACAGAGAGCAGCAGGCCTCTGTAGCAAGCA
2704 GTGTGTTATTTCACAGTGGAGACAGCTGGCAAGAGCAGGAGGGCCCTCAGC
2755 TCTTAGGGCTGTGCAATGACAGAGGACAGAGAAATGTGTTCTACCGAGGGTTA
2805 CAAGGSCAAATCAGACACTTGTATCAGATGGATGTCTGAAAGCTGTAAAT
2857 GTGCTCTCTGTATRAACAATPTGCTCCAAATATTTTGTGTTCCCTTTTGT
2908 GTGGCTGGTATGGCATGTGCTGATGTTTGGTGTFATATGCTGTATCCTGTG
2959 TACCATAATTTGG

FIGURA 18D - 1

```

1  TGGTGACCAAGAGTACATCTCTTTTCAAATAGCTGGATTAGGTCTCATGC
1  M L
52  TGGTGTGGTCATTGCTGGTCATCTTTGATGCAGTCACTGAACAGGCAGATT
19  L W S L L V I F D A V T E Q A D S
103  CGCTGACCCCTTGTGGCGCCCTCTTCTGTCTTCGAAGGAGACAGCATCGTTC
36  L T L V A P S S V F E G D S I V L
154  TGAAATGCCAGGGAGAACAGAACTGGAATTCAGAAGATGGCTTACCATA
53  K C Q G E Q N W K I Q K M A Y H K
205  AGGATAACAAGAGTATCTGTGTTTCAAAAAATTCAGATTTCCTTATCC
70  D N K E L S V F K K F S D F L I Q
256  AAAGTGACAGTTTAAAGTGACAGTGGTAATTTCTGTAGTACCAAGGAC
87  S A V L S D S G N Y F C S T K G Q
307  AACTCTTTCTGCGGATAAACTTCAAATATAGTAAAGATAAAGTCCAAAG
104  L F L W D R T S N I V K I K V Q E
358  AGCTCTTTCAACGTCCTGTGCTGACTGCCAGCTCTTCCAGCCCATCGAAG
121  L F Q R P V L T A S S F Q P I E G
409  GGGGTCCAGTGAGCTGAAATGTGAGACCCGGCTCTCTCCACAGAGGTGG
138  G F V S L K C E T R L S P Q R L D
460  ATGTTCAACTCCAGTTCGCTTCTTCAGAGAAAACAGGTCCTGGGGTCAG
155  V Q L Q F C F F R E N Q V L G S G
511  GCTGGAGCAGCTCTCCGAGCTCCAGATTCTGCGCTGTGGAGTGAAGACA
172  W S S S P E L Q I S A V W S E D T
562  CAGGCTCTTACTGGTGAAGCAGAAACGGTGAATCAGAGATCAGAAAC
189  G S Y N C K A E T V T H R I R K Q
613  AGAGCCTCCAAATCCAGATTACGTCAGAGAAATCCCATCTCTAATGTAA
206  S L Q S Q I H V Q R I P I S N V S
664  GCTTGGAGATCCGGGCCCCGGGGGACAGGTGACTGAAGGACAAAATGTA
223  L E I R A P G G Q V T E G Q K L I
715  TCCTGCTCTGCTCAGTGGCTGGGGGTACAGGAAATGTACATCTCTCGGT
240  L L C S V A G G T G N V T F S W Y
766  ACAGAGAGGCCACAGGAACAGTATGGGAAAGAAAACCCAGCTTCCCTGT
257  R E A T G T S H G K K T Q R S L S
817  CAGCAGAGCTGGAGATCCAGCTGTGAAGAGAGTGAATGCCGCAAAATAT
274  A E L E I P A V K E S D A G K Y Y
868  ACTGTAGAGCTGACAACGGCCATGTGCTATCCAGAGCAAGGTGGTGAATA
291  C R A D N G H V F I Q S K V V N I
919  TCCCTGTGAGAAATCCAGTGTCTGCGCCCTGTCTCACCTCAGGTCTCTGT
308  P V R I P V S R P V L T L R S P G
970  GGGCCCAAGCTGCTGAGTGGGGGACCTGCTGGAGCTTCACTGTGAGGCCCTGA
325  A Q A A V G D L L E L H C E A L R
1021  GAGGCTCTCCCCAATCTTGTACCAATTTATCATGAGGATGTCAACCTTG
342  G S P P I L Y Q F Y H E D V T L G
1072  GGAACAGCTCGGCCCCCTCTGGAGGAGGGGCTCTCTCAACCTCTCTTTGA
359  N S S A P S G G G A S F N L S L T
1123  CTGCAGAACATCTTGGAACTACTCTGTGAGGCCAACACGGCTGGGGG
376  A E H S G N Y S C E A N N G L G A
1174  CCCAGTGCAGTGGGAGTGCAGTCTCCATCTCAGGACCTGATGGCTATA
393  Q C S E A V P V S I S G P D G Y R
1225  GAAGAGACCTCATGACAGCTGGAGTCTCTGGGGACTGTTTGGTGTCTTGT
410  R D L M T A G V L W G L F G V L G
1276  GTTTCAGTGGTGTGCTTTGCTGTGTATGCTTGTTCACAGATATCAG
427  P T G V A L L L Y A L F H K I S G
1327  GAGAAAGTCTGCCACTAATGAACCCAGAGGGGCTTCCAGGCCAAATCCTC
444  E S S A T N E P R G A S R P N P Q
1378  AAGAGTTACCTATTCAAGCCCAACCCAGACATGGAGGAGCTGCAGCCAG
461  E F T Y S S P T P D M E E L Q P V
1429  TGTATGTCAATGTGGGCTCTGTAGATGGGATGTGGTTTATCTCAGGCTT
478  Y V K V G S V D V D V V Y S Q V H
1480  GGAGCATGCAGCAGCCAGAAAGCTCAGCAACATCAGGACACTTCTGGAGA
495  S H Q Q P E S S A N I R T L L E N
1531  ACAAGGACTCCCAAGTCTACTCTCTGTGAAGAAATCATAACACTTGG
512  K D S Q V I Y S S V K K S
1582  AGGAATCAGAAAGAGATCAACAGCAAGGATGGGACATCATTAAGACTTG
1633  CTATAAAACCTTATGAAATGCTTGAGGCTTATCACTGCCACAGCCAGAA
1684  CGTGCTCAGGAGGCACTCTGTCTATTTTGTCTGATGATGTTTCTTCT
1735  CCAATATCTTCTTTACCTATCAATATTCATTGAAGTGTGCTACATCCAG
1786  ACACTGTGCAATAAATTATTTCTGTACCTTCTTTAAGCAATCAGTGTG
1837  TAAAGATTTGAGGGAAGATGAATAGAGATACAAGGCTCAGCTTCTATCT
1888  ACTGTGAAGTATGAGAACAGGACTTGTAGTGGTGTATTAACCTTATTTAT
1939  GTGCTGTGGATACAGTTTGTCTAATATTTTGTGAGAAATTTTGCATAAT

```

FIGURA 18D - 2

```

1990 GTTCATTGGGAATATTGGCCTGAAATTTCTTTTCCACTGTGTCTCTGCCA
2041 GAATGTTTGTATCAGGCTGATGCTGGCTTCATAGAATGAGTTAGGCAGGAG
2092 CCCTTCCTCCTTGATTTTGTGGCATAGTTTCAGCAGGATGGTACCAGTTA
2143 TTCTTTCTGCATCTGTAGAATTCAGCTATGAATCCATCTGGTCTAGGGCT
2194 TTTGTGTTGGTTGGTAAGTTTTTATTACTAATTCAACTTCAGCGCTTGAT
2245 ATTGGTCTAGGAGGGGTTTCTGTCTCTTCTGTTCAATCTGGGAGATTG
2296 TGTGTTTCCAGGAATTAGCCGTTTCTCTCAGATTTTCTTCTTTATGTGCA
2347 TCGACTTGAGTGTAACATAACTTATATGCACTGGGAAACCAAAAAATCTG
2398 TGTGACTTGCTTTATTGCAGCATTGTGTTTTATTTGGTAGTCTGGAAGTGA
2449 ACCTGCAATATCACCAAGTATGCATATAGTTGCAAAATGTGATTTTGA
2500 CATAGTAAATATGAGTATTTGCAATAAACTATGATATTACTTTTGTAAGTA
2551 TATAGAATAAAATGTAAATAATCTATAAAA
    
```

FIGURA 18E - 1

```

1      GAGGCATCTCTAGGTACCATCCCTGAACCTGGTCTC
37 ATGCTGCCGAGGCTGTGCTGTGATCTGTGCTCCACTCTGTGAA
   M L P R L L L L I C A P L C E
82 CCGCCGAGCTGTGTTTTGATAGCCAGCCCTCCCATCCACAGAG
   F A E L F L I A S P S H P T E
127 GGGAGCCCASTGACCTTGACGTGTAAGATGECCTTTCTACAGAGT
   G S P V 7 L T C K M P F L Q S
172 TCAGATGCCAGTTCCAGTTCTGCTTTTTCAGAGACACCCGGGOC
   S D A Q F Q F C F F R D T R A
217 TTGGGCCAGGCTGGAGCAGCTCCCAAGCTCCAGATCGCTGCC
   L G P G W S S S P K L Q I A A
262 ATGTGGAAAGACACAGGCTCATCTGTTGGGAGGCACAGACA
   M W K E D T G S Y W C E A Q T
307 ATGGCCTCCAAAGTCTTGAGGAGCAGGAGATCCAGATAAATGTG
   M A S X V L R S R R S Q I N V
352 CACAGGGTCCCTCTGCTGATGTAGCTTGGAGACTCAGCCCCCA
   H R V F V A D V S L E T Q P P
397 GGAGGACAGGTGATGGAGGAGACAGGCTGGTCTCATCTGCTCA
   G G Q V M E G D R L V L I C S
442 GTTGCTATGGGACAGGAGACATCACCTTCTTTGGTACAAAGGG
   V A M G T G D I T F L W Y K G
487 GCTCTAGGTTTAAACCTTCAGTCAAGAGCCAGGTTCACTGACA
   A V G L N L Q S K T Q R S L T
532 GCAGAGTATGAGATTCTTCAGTGAGGAGAGTATGCTGAGCAA
   A E Y E I P S V R E S D A E Q
577 TATTACTGTGTAGCTGAJAATGGCTATGGTCCAGCCCACTGGG
   Y Y C V A E N G Y G P S P S G
622 CTGGTGAGCATCACTGTGAGAATCCCGGTGTCTCGCCCAATCCTC
   L V S I T V R I P V S R P I L
667 ATGCTCAGGGTCCCGAGGCGCCAGGCTGAGTGGAGGATGTGCTG
   M L R A P R A Q A A V E D V L
712 GAGCTTCACTGTGAGGCGCTGAGAGGCTCTCTCCATCTGTAC
   E L H C E A L R G S P P I L Y
757 TGGTTTTATCAGGAGATATCACCTTGGGAGCAGGTGGGGCCCC
   W F Y H E D I T L G S R S A P
802 TCTGGAGGAGGAGGCTCTTCAACCTTTCCTGACTGAAGAACAT
   S G G G A S F N L S L T E E H
847 TCTGGAACACTCTCTGTGAGGCCAACAATGGCTGGGGGCCAG
   S G N Y S C E A N G L G A Q
892 CCACTGAGGCGGTGACACTCAACTTCACTGCTCTACTGGGGCC
   R S E A V T L N F T V P T G A
937 AGAAGCAATCATCTTACCTCAGGAGTCAATGAGGGGCTGCTCAGC
   R S N H L T S G V I E G L L S
982 ACCCTTGGTCCAGCCACCGTGGCCTTATTATTTTGTACGGGCTC
   T L G P A T V A L L F C Y G L
1027 AAAAGAAJATAGAAAGACCTTCAGCCAGGATCCACTCAGGAGC
   K R K I G R R S A R D P L R S
1072 CTTCCCGAGCCTCTACCCCAAGAGTTCACTTACCTCAACTCACCT
   L P S F L P Q E F T Y L N S P
1117 ACCCCAGGGCAGCTACAGCCTATATATGAAJATGTGAATGTTGTA
   T P G Q L Q P I Y E N V N V V
1162 AGTGGGATGAGGTTTATTCACTGGGCTACTATAACCGCCGAG
   S G D E V Y S L A Y Y N Q P E
1207 CAGGAATCAGTAGCAGCAGAAACCTGGGGACACATATGGAGGAC
   Q E S V A A E T L G T H M E D
1252 AAGSTTTCTTAGACATCTATTCCAGGCTGAGGAAGCAACATT
   K V S L D I Y S R L R K A N I
1297 ACAGATGTGACTATGAAGATGCTATGTAA 1326
   T D V D Y E D A M
      GGT ATGGAAGATT CTGCTCTTTG
1351 AAAACCATCC ATGACCCCAA GCCTCAGGCC TGATATGTC TTCAGAGATC
1401 CTGGGGCATT AGCTTCCAG TATACCTCT CTGGATGCCA TTCCTCATGG
1451 CACTATTCTCT TCATCTACTG TGAAGTGAAG TTGGCGCAGC CCGAAGAAA
1501 CTACCTAGCA GAACATAAG ACACAGGAGT GACAGGGACT TTGTTATCAG
1551 AACCAAGTTC CTGCGGCTC CTTTGAAAC AGGTATATT GTGCTCTCT
1601 CTTTACAAGA GCAACCAAGA TGAATAAAA CAAATGGGA TCTTGGGTTG
1651 GAGGACAGT GAAGCTTAGA GCACATGAAC TCAAGGTTAG TGACTCTGCA
1701 GGACTTCACA GAGAGAGCTG TGCCCATCAT TCAATCCAAG TGCTTCTCT
1751 GCGCAGACAG CACAGAACTC CAGCCCGGCT ACTTACATGG ATCATCGAGT
1801 TTCCACCTAA AATATGATTC TATTATTTT GAGTCACTGT TACCAAA7TA

```

FIGURA 18E - 2

```

1851 GAACTAAAC AAAGTTACAT AAAAAGTTAT TGTGACTCCA CTTAATTTTA
1901 GTGACGTATT TTTGTATATA TAGGCCAACC TATACCACAT CCAAAATTAT
1951 GTATCTATTA CAGCCCCTAG AAGCTTTATA AATACAGTGT GTCTTCTTTT
2001 ATTCACAAA TTTTGTAAAT CGTGGTAATA TGGTTTGAAA CCTGTATCTT
2051 AATTATTTTT TTTTAAATT GAGACAGGGT CTCACCTCTGT CACTCAATCT
2101 GGAATGCAGT GGCACAACTT TGCCTCACTG CAACGCGCTGC CTCTCAGGCT
2151 CAAGCAAACC TCTCACCTCA GCCTGCTGAG TAGCTGGGAC TACAGGCACA
2201 TGCCACCAA CTTGGCCATT TTTTGTCTTA CGTAGAGACA AGATTTCAAC
2251 GTTTTGCCCA GGCTGCTCTC AAACCTCTGG GCTCAAGCAA TGTATTGAAT
2301 TTT

```