

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 372 522**

51 Int. Cl.:
C12N 15/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **01971167 .0**
96 Fecha de presentación: **19.09.2001**
97 Número de publicación de la solicitud: **1328624**
97 Fecha de publicación de la solicitud: **23.07.2003**

54 Título: **MOLÉCULAS DE TIPO B7 Y USOS DE LAS MISMAS.**

30 Prioridad:
20.09.2000 US 233867 P

45 Fecha de publicación de la mención BOPI:
23.01.2012

45 Fecha de la publicación del folleto de la patente:
23.01.2012

73 Titular/es:
**AMGEN INC.
ONE AMGEN CENTER DRIVE
THOUSAND OAKS, CA 91320-1799, US**

72 Inventor/es:
**FOX, Gary Michael;
SULLIVAN, John, K.;
HOLST, Paige y
YOSHINAGA, Steven, Kiyoshi**

74 Agente: **Miltenyi, Peter**

ES 2 372 522 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

Moléculas de tipo B7 y usos de las mismas.

Esta solicitud reivindica el derecho de prioridad de la solicitud de patente estadounidense provisional n.º 60/233.867, presentada el 20 de septiembre de 2000.

5 Campo de la invención

La presente invención se refiere a polipéptidos de tipo B7 (B7-L) y moléculas de ácido nucleico que codifican los mismos. La invención también se refiere a agentes de unión selectiva, vectores, células huésped, y métodos para producir polipéptidos B7-L. La invención se refiere además a composiciones farmacéuticas que comprenden polipéptidos B7-L.

10 Antecedentes de la invención

Avances técnicos en la identificación, clonación, expresión y manipulación de moléculas de ácido nucleico y el descifrado del genoma humano han acelerado enormemente el descubrimiento de productos terapéuticos novedosos. Las técnicas de secuenciación rápida de ácido nucleico pueden generar ahora información de secuencia a velocidades sin precedentes y, acoplado con análisis informáticos, permiten el ensamblaje de secuencias solapantes en genomas parciales o enteros y la identificación de regiones codificantes de polipéptidos. Una comparación de una secuencia de aminoácidos predicha con una compilación de base de datos de secuencias de aminoácidos conocidas permite determinar el grado de homología con secuencias previamente identificadas y/o puntos de referencia estructurales. La clonación y expresión de una región codificante de polipéptido de una molécula de ácido nucleico proporciona un producto de polipéptido para análisis estructurales y funcionales. La manipulación de moléculas de ácido nucleico y polipéptidos codificados puede conferir propiedades ventajosas a un producto para su uso como producto terapéutico.

A pesar de los avances técnicos significativos en la investigación del genoma en la última década, el potencial de desarrollo de productos terapéuticos novedosos todavía está ampliamente sin realizar. Todavía no se han identificado muchos genes que codifican productos terapéuticos polipeptídicos posiblemente beneficiosos o los que codifican polipéptidos que pueden actuar como "dianas" para moléculas terapéuticas. Por consiguiente, un objeto de la invención es identificar polipéptidos novedosos, y moléculas de ácido nucleico que codifican los mismos, que tienen beneficio de diagnóstico o terapéutico.

Sumario de la invención

Se ha identificado un polipéptido novedoso, que está relacionado con proteínas en la ruta coestimulante de células T. Este polipéptido, denominado polipéptido de tipo B7 (B7-L), representa un polipéptido relacionado con B7 que tiene una estructura similar a, y que comparte homología con, CD80, CD86, B7RP-1 y B7-H1. El polipéptido humano B7-L comparte una mayor identidad de aminoácidos con B7-H1 humano (35%) que con otros miembros de la familia B7 (B7rp-1, 20%; B7-1, 12%). Además, los ortólogos de ratón y humano de polipéptido B7-L y B7-H1 comparten un mayor grado de similitud de aminoácidos (aproximadamente el 70%) que los ortólogos de ratón y humano de B7-1, B7-2 o B7rp-1 (aproximadamente el 40%). Además, tanto B7-H1 como B7-L se unen al mismo receptor, PD-1.

Se dan a conocer secuencias relacionadas con los números de acceso de NCBI AK001872 y AF142780.

La presente invención se refiere a moléculas de ácido nucleico B7-L novedosas y polipéptidos codificados.

La invención proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos seleccionada del grupo que consiste en:

- (a) la secuencia de nucleótidos tal como se expone en SEQ ID NO: 1;
- (b) una secuencia de nucleótidos que codifica el polipéptido tal como se expone en SEQ ID NO: 2; y
- (c) una secuencia de nucleótidos complementaria a las secuencias de (a) - (b).

La invención también proporciona un polipéptido B7-L aislado que comprende una secuencia de aminoácidos tal como se expone en SEQ ID NO: 2.

También se proporcionan polipéptidos de fusión que comprenden secuencias de aminoácidos B7-L.

La presente invención también proporciona un vector de expresión que comprende las moléculas de ácido nucleico aisladas tal como se exponen en el presente documento, células huésped recombinantes que comprenden las moléculas de ácido nucleico recombinante tal como se exponen en el presente documento, y un método para producir un polipéptido B7-L que comprende cultivar las células huésped y opcionalmente aislar el polipéptido así producido.

También se proporcionan derivados de los polipéptidos B7-L de la presente invención.

Se proporcionan adicionalmente anticuerpos que pueden unirse específicamente a los polipéptidos B7-L de la invención. Tales anticuerpos pueden ser agonistas o antagonistas.

La invención también abarca composiciones farmacéuticas que comprenden los ácidos nucleicos, polipéptidos o anticuerpos de la invención y uno o más agentes de formulación farmacéuticamente aceptables. Las composiciones farmacéuticas se usan para proporcionar cantidades terapéuticamente eficaces de los ácidos nucleicos o polipéptidos de la presente invención.

Los polipéptidos B7-L y moléculas de ácido nucleico de la presente invención pueden usarse para tratar, prevenir, mejorar y/o detectar enfermedades y trastornos, incluyendo los mencionados en el presente documento.

Breve descripción de las figuras

Las figuras 1A-1B ilustran la secuencia de nucleótidos del gen B7-L humano (SEQ ID NO: 1) y la secuencia de aminoácidos deducida del polipéptido B7-L humano (SEQ ID NO: 2). Se indican el péptido señal predicho (subrayado) y el dominio transmembrana (doble subrayado).

Las figuras 2A-2C ilustran una alineación de secuencia de aminoácidos de polipéptido B7-L humano (GA16817596; SEQ ID NO: 2), CD80 o B7-1 humano (Cd80_Human; SEQ ID NO: 4; n.º de registro GenBank P33681), CD86 o B7-2 humano (Cd86_Human; SEQ ID NO: 5; n.º de registro GenBank U04343), B7-H1 humano (B7-H1_Human; SEQ ID NO: 6; n.º de registro GenBank AF177937), B7rp-1 humano (B7rp-1_Human; SEQ ID NO: 7; n.º de registro GenBank AF199028), PRO352 humano (Pro352_Human; SEQ ID NO: 8; n.º de registro GenBank Y41705), butirofilina BTF1 humana (Btf1_Human; SEQ ID NO: 9; n.º de registro GenBank U90543), butirofilina BTF2 humana (Btsf2a2_Hu; SEQ ID NO: 10; n.º de registro GenBank U90550), butirofilina BTF4 humana (Btf4_Human; SEQ ID NO: 11; n.º de registro GenBank U90546), butirofilina BTF3 humana (Btn3a3_Human; SEQ ID NO: 12; n.º de registro GenBank U90548) y butirofilina (Btn_Human; SEQ ID NO: 13; n.º de registro GenBank U39576).

Las figuras 3A-3E ilustran una parte de la secuencia de nucleótidos genómica para el polipéptido B7-L humano (SEQ ID NO: 14). Se indica la ubicación de la secuencia de aminoácidos deducida del exón 1 (SEQ ID NO: 19).

La figura 4 ilustra una parte de la secuencia de nucleótidos genómica para el polipéptido B7-L humano (SEQ ID NO: 15).

Las figuras 5A-5F ilustran una parte de la secuencia de nucleótidos genómica para el polipéptido B7-L humano (SEQ ID NO: 16). Se indica la ubicación de la secuencia de aminoácidos deducida del exón 2 (SEQ ID NO: 20).

Las figuras 6A-6B ilustran una parte de la secuencia de nucleótidos genómica para el polipéptido B7-L humano (SEQ ID NO: 17). Se indica la ubicación de la secuencia de aminoácidos deducida del exón 3 (SEQ ID NO: 21).

Las figuras 7A-7M ilustran una parte de la secuencia de nucleótidos genómica para el polipéptido B7-L humano (SEQ ID NO: 18). Se indican las ubicaciones de la secuencia de aminoácidos deducida de los exones 4 (SEQ ID NO: 22), 5 (SEQ ID NO: 23) y 6.

La figura 8 muestra los resultados de un análisis de transferencia de tipo Northern de la expresión de ARNm de B7-L humano.

La figura 9 muestra los resultados de un análisis de citometría de flujo activada por fluorescencia (FACS) de células CHO D- transfectadas con vector solo (A), o con vectores que codifican PD-1 (B), CRP-1/ICOS (C) o CD28 (D), e incubadas o bien con FITC solo, o bien con las siguientes proteínas de fusión: CRP-1/Fc, polipéptido B7-L/Fc, B7rp-1/Fc, B7-H1/Fc o B7-2/Fc.

La figura 10 muestra los resultados obtenidos en ensayos de proliferación de células T mediada por anticuerpos anti-CD3 usando polipéptido B7-L/Fc, B7RP-1/Fc o B7-1/Fc.

La figura 11 muestra los resultados de análisis de FACS de células de sangre periférica humanas usando FITC solo (A), control de Fc (B), B7rp-1/Fc (C), polipéptido B7-L/Fc (D), polipéptido B7-L/Fc y anticuerpo anti-CD3 (E), o polipéptido B7-L/Fc y anticuerpo anti-CD19 (F).

Descripción detallada de la invención

Los títulos de sección usados en el presente documento son únicamente para fines de organización y no deben interpretarse como limitativos del contenido descrito.

Definiciones

Las expresiones "gen B7-L" o "molécula de ácido nucleico B7-L" o "polinucleótido B7-L" se refieren a una molécula de ácido nucleico que comprende una secuencia de nucleótidos tal como se expone en SEQ ID NO: 1, una secuencia de nucleótidos que codifica el polipéptido tal como se expone en SEQ ID NO: 2 y moléculas de ácido nucleico tal como se define en el presente documento.

La expresión “variante alélica de polipéptido B7-L” se refiere a una de varias formas alternativas posibles que se producen de manera natural de un gen que ocupa un locus dado en un cromosoma de un organismo o una población organismos.

5 La expresión “variante de corte y empalme de polipéptido B7-L” se refiere a una molécula de ácido nucleico, habitualmente ARN, que se genera procesando alternativamente secuencias de intrones en un transcrito de ARN de secuencia de aminoácidos del polipéptido B7-L tal como se expone en SEQ ID NO: 2.

10 La expresión “molécula de ácido nucleico aislada” se refiere a una molécula de ácido nucleico de la invención que (1) se ha separado de al menos aproximadamente el 50 por ciento de las proteínas, lípidos, hidratos de carbono u otros materiales con los que se encuentra naturalmente cuando se aísla ácido nucleico total de las células originales, (2) no está unida a la totalidad o una parte de un polinucleótido al que la “molécula de ácido nucleico aislada” está unida en la naturaleza, (3) está operativamente unida a un polinucleótido al que no está unida en la naturaleza, o (4) no se produce en la naturaleza como parte de una secuencia de polinucleótido más grande. Preferiblemente, la molécula de ácido nucleico aislada de la presente invención está sustancialmente libre de cualquier otra molécula de ácido nucleico contaminante u otros contaminantes que se encuentran en su entorno natural que interferirían con su uso en la producción de polipéptidos o su uso terapéutico, de diagnóstico, profiláctico o de investigación.

15 La expresión “secuencia de ácido nucleico” o “molécula de ácido nucleico” se refiere a una secuencia de ADN o ARN. La expresión abarca moléculas formadas a partir de cualquiera de los análogos de bases conocidos de ADN y ARN tales como, pero sin limitarse a, 4-acetilcitosina, 8-hidroxi-N6-metiladenosina, aziridinil-citosina, pseudoisocitosina, 5-(carboxihidroximetil)uracilo, 5-fluorouracilo, 5-bromouracilo, 5-carboximetilaminometil-2-tiouracilo, 5-carboxi-metilaminometiluracilo, dihidrouracilo, inosina, N6-iso-penteniladenina, 1-metiladenina, 1-metilpseudouracilo, 1-metilguanina, 1-metilinosina, 2,2-dimetil-guanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-metiladenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiamino-metil-2-tiouracilo, beta-D-manosilqueosina, 5'-metoxycarbonil-metiluracilo, 5-metoxiuracilo, 2-metil-N6-isopenteniladenina, éster metílico del ácido uracil-5-oxiacético, ácido uracil-5-oxiacético, oxibutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, éster metílico del ácido N-uracil-5-oxiacético, ácido uracil-5-oxiacético, pseudouracilo, queosina, 2-tiocitosina, y 2,6-diaminopurina.

El término “vector” se usa para referirse a cualquier molécula (por ejemplo, ácido nucleico, plásmido o virus) usada para transferir información de codificación a una célula huésped.

30 La expresión “vector de expresión” se refiere a un vector que es adecuado para la transformación de una célula huésped y contiene secuencias de ácido nucleico que dirigen y/o controlan la expresión de secuencias de ácido nucleico heterólogas insertadas. La expresión incluye, pero no se limita a, procesos tales como transcripción, traslación, y corte y empalme de ARN, si hay intrones presentes.

35 La expresión “operativamente unido” se usa en el presente documento para referirse a una disposición de secuencias flanqueantes en la que las secuencias flanqueantes así descritas están configuradas o ensambladas para realizar su función habitual. Por tanto, una secuencia flanqueante operativamente unida a una secuencia codificante puede ser capaz de realizar la replicación, transcripción y/o traducción de la secuencia codificante. Por ejemplo, una secuencia codificante está operativamente unida a un promotor cuando el promotor puede dirigir la transcripción de la secuencia codificante. Una secuencia flanqueante no necesita ser contigua a la secuencia codificante, siempre que funcione correctamente. Por tanto, por ejemplo, pueden estar presentes secuencias transcritas intermedias todavía no traducidas entre una secuencia promotora y la secuencia codificante y la secuencia promotora todavía puede considerarse “operativamente unida” a la secuencia codificante.

40 La expresión “célula huésped” se usa para referirse a una célula que se ha transformado, o puede transformarse, con una secuencia de ácido nucleico y entonces expresar un gen seleccionado de interés. La expresión incluye la progenie de la célula original, tanto si la progenie es idéntica en cuanto a la morfología y la constitución genética a la original como si no, siempre que el gen seleccionado esté presente.

La expresión “polipéptido B7-L” se refiere a un polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 2.

50 La expresión “fragmento de polipéptido B7-L” se refiere a un polipéptido que comprende un truncamiento en el extremo amino terminal (con o sin una secuencia líder) y/o un truncamiento en el carboxilo terminal del polipéptido tal como se expone en SEQ ID NO: 2. La expresión “fragmento de polipéptido B7-L” también se refiere a truncamientos amino terminal y/o carboxilo terminal de ortólogos de polipéptido B7-L, derivados de polipéptido B7-L, o variantes de polipéptido B7-L, o a truncamientos amino terminal y/o carboxilo terminal de los polipéptidos codificados por variantes alélicas de polipéptido B7-L o variantes de corte y empalme de polipéptido B7-L. Los fragmentos de polipéptido B7-L pueden resultar de corte y empalme de ARN alternativo o de actividad proteasa *in vivo*. La presente invención también contempla formas unidas a membrana de un polipéptido B7-L. En realizaciones preferidas, truncamientos y/o deleciones comprenden aproximadamente 10 aminoácidos, o aproximadamente 20 aminoácidos, o aproximadamente 50 aminoácidos, o aproximadamente 75 aminoácidos, o aproximadamente 100 aminoácidos, o más de aproximadamente 100 aminoácidos. Los fragmentos de polipéptido así producidos comprenderán aproximadamente 25 aminoácidos contiguos, o aproximadamente 50 aminoácidos, o

aproximadamente 75 aminoácidos, o aproximadamente 100 aminoácidos, o aproximadamente 150 aminoácidos, o aproximadamente 200 aminoácidos. Tales fragmentos de polipéptido B7-L pueden comprender opcionalmente un resto de metionina amino terminal. Se apreciará que tales fragmentos pueden usarse, por ejemplo, para generar anticuerpos frente a polipéptidos B7-L.

- 5 La expresión “ortólogo de polipéptido B7-L” se refiere a un polipéptido de otra especie que corresponde a la secuencia de aminoácidos de polipéptido B7-L tal como se expone en SEQ ID NO: 2. Por ejemplo, los polipéptidos B7-L de ratón y humano se consideran ortólogos entre sí.

- 10 La expresión “variantes de polipéptido B7-L” se refiere a polipéptidos B7-L que comprenden secuencias de aminoácidos que tienen una o más sustituciones, delecciones (tales como delecciones internas y/o fragmentos de polipéptido B7-L), y/o adiciones (tales como adiciones internas y/o polipéptidos de fusión B7-L) de secuencia de aminoácidos en comparación con la secuencia de aminoácidos de polipéptido B7-L tal como se expone en SEQ ID NO: 2 (con o sin una secuencia líder). Las variantes pueden producirse de manera natural (por ejemplo, variantes alélicas de polipéptido B7-L, ortólogos de polipéptido B7-L y variantes de corte y empalme de polipéptido B7-L) o construirse artificialmente. Tales variantes de polipéptido B7-L pueden prepararse a partir de las moléculas de ácido nucleico correspondientes que tienen una secuencia de ADN que varía por consiguiente de la secuencia de ADN tal como se expone en SEQ ID NO: 1. En realizaciones preferidas, las variantes tienen desde 1 hasta 3, o desde 1 hasta 5, o desde 1 hasta 10, o desde 1 hasta 15, o desde 1 hasta 20, o desde 1 hasta 25, o desde 1 hasta 50, o desde 1 hasta 75, o desde 1 hasta 100, o más de 100 sustituciones, inserciones, adiciones y/o delecciones de aminoácido, en las que las sustituciones pueden ser conservativas, o no conservativas, o cualquier combinación de las mismas.

La expresión “derivados de polipéptido B7-L” se refiere al polipéptido tal como se expone en SEQ ID NO: 2, tal como se define en el presente documento, que se ha modificado químicamente.

- 25 La expresión “polipéptido B7-L maduro” se refiere a un polipéptido B7-L que carece de una secuencia líder. Un polipéptido B7-L maduro también puede incluir otras modificaciones tales como procesamiento proteolítico del amino terminal (con o sin una secuencia líder) y/o el carboxilo terminal, escisión de un polipéptido más pequeño a partir de un precursor más grande, glicosilación unida a N y/o unida a O, y similares. Un polipéptido B7-L maduro a modo de ejemplo se representa por la secuencia de aminoácidos de SEQ ID NO: 3.

- 30 La expresión “polipéptido de fusión B7-L” se refiere a una fusión de uno o más aminoácidos (tales como un péptido o proteína heteróloga) en el amino o carboxilo terminal del polipéptido tal como se expone en SEQ ID NO: 2, fragmentos de polipéptido B7-L, ortólogos de polipéptido B7-L, variantes de polipéptido B7-L o derivados de B7-L, tal como se define en el presente documento. La expresión “polipéptido de fusión B7-L” también se refiere a una fusión de uno o más aminoácidos en el extremo amino o carboxilo terminal del polipéptido codificado por variantes alélicas del polipéptido B7-L o variantes de corte y empalme del polipéptido B7-L, tal como se define en el presente documento.

- 35 La expresión “polipéptidos B7-L biológicamente activos” se refiere a polipéptidos B7-L que tienen al menos una actividad característica del polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 2. Además, un polipéptido B7-L puede ser activo como inmunógeno; es decir, el polipéptido B7-L contiene al menos un epítipo frente al cual pueden prepararse anticuerpos.

- 40 La expresión “polipéptido aislado” se refiere a un polipéptido de la presente invención que (1) se ha separado de al menos aproximadamente el 50 por ciento de polinucleótidos, lípidos, hidratos de carbono u otros materiales con los que se encuentra de manera natural cuando se aísla de la célula original, (2) no está unido (mediante interacción covalente o no covalente) a la totalidad o una parte de un polipéptido al que está unido el “polipéptido aislado” en la naturaleza, (3) está operativamente unido (mediante interacción covalente o no covalente) a un polipéptido al que no está unido en la naturaleza, o (4) no se produce en la naturaleza. Preferiblemente, el polipéptido aislado está sustancialmente libre de cualquier otros polipéptidos contaminantes u otros contaminantes que se encuentran en su entorno natural que interferirían con su uso terapéutico, de diagnóstico, profiláctico o de investigación.

- 45 El término “identidad”, tal como se conoce en la técnica, se refiere a una relación entre las secuencias de dos o más moléculas de polipéptido o dos o más moléculas de ácido nucleico, tal como se determina comparando las secuencias. En la técnica, “identidad” también significa el grado de relación de secuencia entre moléculas de ácido nucleico o polipéptidos, según sea el caso, tal como se determina mediante la coincidencia entre cadenas de dos o más nucleótidos o dos o más secuencias de aminoácidos. “Identidad” mide el porcentaje de coincidencias idénticas entre la menor de dos o más secuencias tratándose las alineaciones de huecos (si los hay) mediante un modelo matemático particular o programa informático (es decir, “algoritmos”).

- 50 El término “similitud” es un término relacionado, pero al contrario que “identidad”, “similitud” se refiere a una medida de la relación que incluye tanto coincidencias idénticas como coincidencias de sustitución conservativa. Si dos secuencias de polipéptido tienen, por ejemplo, 10/20 aminoácidos idénticos, y el resto son todo sustituciones no conservativas, entonces el porcentaje de identidad y similitud serán ambos del 50%. Si en el mismo ejemplo, hay cinco posiciones más en las que hay sustituciones conservativas, entonces el porcentaje de identidad sigue siendo del 50%, pero el porcentaje de similitud será del 75% (15/20). Por tanto, en casos en los que hay sustituciones

conservativas, el porcentaje de similitud entre dos polipéptidos será mayor que el porcentaje de identidad entre esos dos polipéptidos.

La expresión “que se produce de manera natural” o “nativo” cuando se usan en relación con materiales biológicos tales como moléculas de ácido nucleico, polipéptidos, células huésped y similares, se refiere a materiales que se encuentran en la naturaleza y no se manipulan por el ser humano. De manera similar, “que no se produce de manera natural” o “no nativo” tal como se usa en el presente documento se refiere a un material que no se encuentra en la naturaleza o que se ha modificado estructuralmente o sintetizado por el ser humano.

Las expresiones “cantidad eficaz” y “cantidad terapéuticamente eficaz” se refieren cada una a la cantidad de un polipéptido B7-L o molécula de ácido nucleico B7-L usada para soportar un nivel observado de una o más actividades biológicas de los polipéptidos B7-L tal como se expone en el presente documento.

La expresión “vehículo farmacéuticamente aceptable” o “vehículo fisiológicamente aceptable” tal como se usa en el presente documento se refiere a uno o más materiales de formulación adecuados para lograr o potenciar la entrega del polipéptido B7-L, molécula de ácido nucleico B7-L, o agente de unión selectiva de B7-L como composición farmacéutica.

El término “antígeno” se refiere a una molécula o una parte de una molécula a la que puede unirse un agente de unión selectiva, tal como un anticuerpo, y adicionalmente que puede usarse en un animal para producir anticuerpos que pueden unirse a un epítipo de ese antígeno. Un antígeno puede tener uno o más epítopos.

La expresión “agente de unión selectiva” se refiere a un anticuerpo o anticuerpos que tienen especificidad por un polipéptido B7-L. Tal como se usan en el presente documento, los términos “específico” y “especificidad” se refieren a la capacidad de los agentes de unión selectiva para unirse a polipéptidos B7-L humanos y para no unirse a polipéptidos humanos distintos de B7-L. Sin embargo, se apreciará que los agentes de unión selectiva también pueden unirse a ortólogos del polipéptido tal como se expone en SEQ ID NO: 2, es decir, versiones interespecies del mismo, tal como polipéptidos B7-L de ratón y de rata.

El término “transducción” se usa para hacer referencia a la transferencia de genes de una bacteria a otra, habitualmente mediante un fago. “Transducción” también se refiere a la adquisición y transferencia de secuencias celulares eucariotas mediante retrovirus.

El término “transfección” se usa para hacer referencia a la captación de ADN extraño o exógeno por una célula, y una célula se ha “transfectado” cuando el ADN exógeno se ha introducido en el interior de la membrana celular. En la técnica se conocen bien varias técnicas de transfección y se dan a conocer en el presente documento. Véase, por ejemplo, Graham *et al.*, 1973, *Virology* 52: 456; Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratories, 1989); Davis *et al.*, *Basic Methods in Molecular Biology* (Elsevier, 1986); y Chu *et al.*, 1981, *Gene* 13: 197. Tales técnicas pueden usarse para introducir uno o más restos de ADN exógenos en células huésped adecuadas.

El término “transformación” tal como se usa en el presente documento se refiere a un cambio en las características genéticas de una célula, y una célula se ha transformado cuando se ha modificado para contener un nuevo ADN. Por ejemplo, una célula se transforma cuando se modifica genéticamente a partir de su estado nativo. Tras la transfección o transducción, el ADN transformante puede recombinarse con el de la célula integrándose físicamente en un cromosoma de la célula, puede mantenerse de manera transitoria en un elemento episómico sin replicarse, o puede replicarse independientemente como un plásmido. Se considera que una célula se ha transformado de manera estable cuando el ADN se replica con la división de la célula.

Relación de moléculas de ácido nucleico y/o polipéptidos

Se entiende que las moléculas de ácido nucleico relacionadas incluyen variantes alélicas o de corte y empalme de la molécula de ácido nucleico de SEQ ID NO: 1, e incluyen secuencias que son complementarias a cualquiera de las secuencias de nucleótidos anteriores. Las moléculas de ácido nucleico relacionadas también incluyen una secuencia de nucleótidos que codifica un polipéptido que comprende o consiste esencialmente en una sustitución, modificación, adición y/o delección de uno o más restos de aminoácido en comparación con el polipéptido tal como se expone en SEQ ID NO: 2. Tales polipéptidos B7-L relacionados pueden comprender, por ejemplo, una adición y/o una delección de uno o más sitios de glicosilación unida a N o unida a O o una adición y/o una delección de uno o más restos de cisteína.

Las moléculas de ácido nucleico relacionadas también incluyen fragmentos de moléculas de ácido nucleico B7-L que codifican un polipéptido de al menos aproximadamente 25 aminoácidos contiguos, o aproximadamente 50 aminoácidos, o aproximadamente 75 aminoácidos, o aproximadamente 100 aminoácidos, o aproximadamente 150 aminoácidos, o aproximadamente 200 aminoácidos, o más de 200 restos de aminoácido del polipéptido B7-L de SEQ ID NO: 2.

Además, las moléculas de ácido nucleico B7-L relacionadas también incluyen las moléculas que comprenden secuencias de nucleótidos que se hibridan en condiciones moderada o altamente rigurosas tal como se define en el presente documento con la secuencia completamente complementaria de la molécula de ácido nucleico B7-L de

SEQ ID NO: 1, o de una molécula que codifica un polipéptido, polipéptido que comprende la secuencia de aminoácidos tal como se muestra en SEQ ID NO: 2, o de un fragmento de ácido nucleico tal como se define en el presente documento, o de un fragmento de ácido nucleico que codifica un polipéptido tal como se define en el presente documento. Pueden prepararse sondas de hibridación usando las secuencias de B7-L proporcionadas en el presente documento para examinar bibliotecas de ADNc, ADN genómico o sintético para seleccionar secuencias relacionadas. Las regiones del ADN y/o secuencia de aminoácidos de polipéptido B7-L que muestran una identidad significativa con respecto a secuencias conocidas se determinan fácilmente usando algoritmos de alineación de secuencias tal como se describe en el presente documento y esas regiones pueden usarse para diseñar sondas para la selección.

La expresión "condiciones altamente rigurosas" se refiere a aquellas condiciones que están diseñadas para permitir la hibridación de cadenas de ADN cuyas secuencias son altamente complementarias, y para excluir la hibridación de ADN con apareamiento significativamente erróneo. La rigurosidad de hibridación se determina principalmente por la temperatura, la fuerza iónica y la concentración de agentes desnaturizantes tales como formamida. Ejemplos de "condiciones altamente rigurosas" para la hibridación y el lavado son cloruro de sodio 0,015 M, citrato de sodio 0,0015 M a 65-68°C o cloruro de sodio 0,015 M, citrato de sodio 0,0015 M, y formamida al 50% a 42°C. Véase Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual* (2ª ed., Cold Spring Harbor Laboratory, 1989); Anderson *et al.*, *Nucleic Acid Hybridisation: A Practical Approach* Ch. 4 (IRL Press Limited).

También pueden usarse condiciones más rigurosas (tales como temperatura superior, fuerza iónica inferior, más formamida u otro agente desnaturizante), sin embargo la tasa de hibridación se verá afectada. Pueden incluirse otros agentes en los tampones de hibridación y lavado con fines de reducir la hibridación no específica y/o de fondo. Ejemplos son albúmina de suero bovino al 0,1%, polivinilpirrolidona al 0,1%, pirofosfato de sodio al 0,1%, dodecilsulfato de sodio al 0,1%, NaDodSO₄, (SDS), ficoll, disolución de Denhardt, ADN de esperma de salmón sonificado (u otro ADN no complementario), y sulfato de dextrano, aunque también pueden usarse otros agentes adecuados. La concentración y los tipos de estos aditivos pueden cambiarse sin afectar sustancialmente a la rigurosidad de las condiciones de hibridación. Los experimentos de hibridación se llevan a cabo habitualmente a pH 6,8-7,4; sin embargo, a condiciones de fuerza iónica típicas, la tasa de hibridación es casi independiente del pH. Véase Anderson *et al.*, *Nucleic Acid Hybridisation: A Practical Approach* Ch. 4 (IRL Press Limited).

Factores que afectan a la estabilidad del dúplex de ADN incluyen la composición de bases, longitud y grado de apareamiento erróneo de pares de bases. Un experto en la técnica puede ajustar las condiciones de hibridación con el fin de adaptarse a esas variables y permitir que ADN de diferentes relaciones de secuencia formen híbridos. La temperatura de fusión de un dúplex de ADN perfectamente apareado puede estimarse mediante la siguiente ecuación:

$$T_m (^{\circ}\text{C}) = 81,5 + 16,6 (\log [\text{Na}^+]) + 0,41 (\% \text{ G+C}) - 600/\text{N} - 0,72 (\% \text{ de formamida})$$

en la que N es la longitud del dúplex formado, [Na⁺] es la concentración molar del ión sodio en la disolución de hibridación o lavado, % G+C es el porcentaje de bases (guanina+citosina) en el híbrido. Para híbridos apareados de manera imperfecta, la temperatura de fusión se reduce en aproximadamente 1°C por cada 1% de apareamiento erróneo.

La expresión "condiciones moderadamente rigurosas" se refiere a condiciones en las que puede formarse un dúplex de ADN con un mayor grado de apareamiento erróneo de pares de bases de lo que se produciría en "condiciones altamente rigurosas". Ejemplos de "condiciones moderadamente rigurosas" típicas son cloruro de sodio 0,015 M, citrato de sodio 0,0015 M a 50-65°C o cloruro de sodio 0,015 M, citrato de sodio 0,0015 M, y formamida al 20% a 37-50°C. A modo de ejemplo, "condiciones moderadamente rigurosas" de 50°C en ión sodio 0,015 M permitirán aproximadamente un 21% de apareamiento erróneo.

Los expertos en la técnica apreciarán que no hay ninguna distinción absoluta entre "condiciones altamente rigurosas" y "condiciones moderadamente rigurosas". Por ejemplo, a ión sodio 0,015 M (sin formamida), la temperatura de fusión de ADN largo perfectamente apareado es de aproximadamente 71°C. Con un lavado a 65°C (a la misma fuerza iónica), esto permitiría aproximadamente un 6% de apareamiento erróneo. Para capturar secuencias relacionadas más distantemente, un experto en la técnica puede sencillamente reducir la temperatura o aumentar la fuerza iónica.

Una buena estimación de la temperatura de fusión en NaCl* 1 M para sondas de oligonucleótidos de hasta aproximadamente 20 nt viene dada por:

$$T_m = 2^{\circ}\text{C por par de bases A-T} + 4^{\circ}\text{C por par de bases G-C}$$

*La concentración de ión sodio en 6X sal citrato de sodio (SSC) es de 1 M. Véase Suggs *et al.*, *Developmental Biology Using Purified Genes* 683 (Brown y Fox, eds., 1981).

Las condiciones de lavado de alta rigurosidad para oligonucleótidos son habitualmente a una temperatura de 0-5°C por debajo de la T_m del oligonucleótido en 6X SSC, SDS al 0,1%.

Diferencias en la secuencia de ácido nucleico pueden dar como resultado modificaciones conservativas y/o no

conservativas de la secuencia de aminoácidos con respecto a la secuencia de aminoácidos de SEQ ID NO: 2.

Las modificaciones conservativas con respecto a la secuencia de aminoácidos de SEQ ID NO: 2 (y las correspondientes modificaciones con respecto a los nucleótidos codificantes) producirán un polipéptido que tiene características funcionales y químicas similares a las de los polipéptidos B7-L. En cambio, pueden lograrse modificaciones sustanciales en las características funcionales y/o químicas de los polipéptidos B7-L seleccionando sustituciones en la secuencia de aminoácidos de SEQ ID NO: 2 que se diferencian significativamente en su efecto sobre el mantenimiento de (a) la estructura del esqueleto molecular en la zona de la sustitución, por ejemplo, tal como una conformación en lámina o helicoidal, (b) la carga o hidrofobia de la molécula en el sitio diana, o (c) el volumen de la cadena lateral.

Por ejemplo, una “sustitución de aminoácidos conservativa” puede suponer una sustitución de un resto de aminoácido nativo por un resto no nativo de tal manera que hay poco o ningún efecto sobre la polaridad o la carga del resto de aminoácido en esa posición. Además, cualquier resto nativo en el polipéptido también puede sustituirse por alanina, tal como se describió anteriormente para “mutagénesis mediante alanina”.

Las sustituciones de aminoácidos conservativas también abarcan restos de aminoácido que no se producen de manera natural que se incorporan normalmente mediante síntesis de péptidos química en vez de mediante síntesis en sistemas biológicos. Estos incluyen compuestos peptidomiméticos, y otras formas revertidas o invertidas de restos de aminoácido.

Los restos que se producen de manera natural pueden dividirse en clases basándose en propiedades de cadena lateral comunes:

- 1) hidrófobos: norleucina, Met, Ala, Val, Leu, Ile;
- 2) hidrófilos neutros: Cys, Ser, Thr;
- 3) ácidos: Asp, Glu;
- 4) básicos: Asn, Gln, His, Lys, Arg;
- 5) restos que influyen sobre la orientación de cadena: Gly, Pro; y
- 6) aromáticos: Trp, Tyr, Phe.

Por ejemplo, sustituciones no conservativas pueden suponer el intercambio de un miembro de una de esas clases por un miembro de otra clase. Tales restos sustituidos pueden introducirse en regiones del polipéptido B7-L humano que son homólogas a polipéptidos B7-L no humanos, o en las regiones no homólogas de la molécula.

Realizando tales cambios, puede considerarse el índice hidropático de aminoácidos. A cada aminoácido se le ha asignado un índice hidropático basándose en su hidrofobia y características de carga. Los índices hidropáticos son: isoleucina (+4,5); valina (+4,2); leucina (+3,8); fenilalanina (+2,8); cisteína/cistina (+2,5); metionina (+1,9); alanina (+1,8); glicina (-0,4); treonina (-0,7); serina (-0,8); triptófano (-0,9); tirosina (-1,3); prolina (-1,6); histidina (-3,2); glutamato (-3,5); glutamina (-3,5); aspartato (-3,5); asparagina (-3,5); lisina (-3,9); y arginina (-4,5).

La importancia del índice hidropático de aminoácidos al conferir función biológica interactiva a una proteína se entiende generalmente en la técnica (Kyte *et al.*, 1982, J. Mol. Biol. 157: 105-31). Se sabe que determinados aminoácidos pueden sustituirse por otros aminoácidos que tienen una puntuación o índice hidropático similar y todavía conservar una actividad biológica similar. Al realizar cambios en el índice hidropático, se prefiere la sustitución de aminoácidos cuyos índices hidropáticos se encuentran dentro de ± 2 , se prefieren particularmente aquellos que se encuentran dentro de ± 1 , y se prefieren incluso más particularmente aquellos que se encuentran dentro de $\pm 0,5$.

También se entiende en la técnica que la sustitución de aminoácidos similares puede realizarse de manera eficaz basándose en la hidrofilia, particularmente cuando la proteína o péptido funcionalmente equivalente de manera biológica así creado está previsto para su uso en realizaciones inmunológicas, tal como en el presente caso. La mayor hidrofilia promedio local de una proteína, regido por la hidrofilia de sus aminoácidos adyacentes, se correlaciona con su inmunogenicidad y antigenicidad, es decir, con una propiedad biológica de la proteína.

Los siguientes valores de hidrofilia se han asignado a estos restos de aminoácido: arginina (+3,0); lisina (+3,0); aspartato (+3,0 $\pm$ 1); glutamato (+3,0 $\pm$ 1); serina (+0,3); asparagina (+0,2); glutamina (+0,2); glicina (0); treonina (-0,4); prolina (-0,5 $\pm$ 1); alanina (-0,5); histidina (-0,5); cisteína (-1,0); metionina (-1,3); valina (-1,5); leucina (-1,8); isoleucina (-1,8); tirosina (-2,3); fenilalanina (-2,5); y triptófano (-3,4). Al realizar cambios basándose en valores de hidrofilia similares, se prefiere la sustitución de aminoácidos cuyos valores de hidrofilia se encuentran dentro de ± 2 , se prefieren particularmente aquellos que se encuentran dentro de ± 1 , y se prefieren incluso más particularmente aquellos que se encuentran dentro de $\pm 0,5$. También pueden identificarse epítomos a partir de secuencias de aminoácidos primarias basándose en la hidrofilia. Estas regiones también se denominan “regiones principales epitópicas”.

Los expertos en la técnica pueden determinar sustituciones de aminoácidos deseadas (ya sean conservativas o no conservativas) en el momento en que se deseen tales sustituciones. Por ejemplo, pueden usarse sustituciones de aminoácidos para identificar restos importantes del polipéptido B7-L, o para aumentar o disminuir la afinidad de los polipéptidos B7-L descritos en el presente documento. En la tabla I se exponen sustituciones de aminoácidos a modo de ejemplo.

Tabla I

Sustituciones de aminoácidos

Restos originales	Sustituciones a modo de ejemplo	Sustituciones preferidas
Ala	Val, Leu, Ile	Val
Arg	Lys, Gln, Asn	Lys
Asn	Gln	Gln
Asp	Glu	Glu
Cys	Ser, Ala	Ser
Gln	Asn	Asn
Glu	Asp	Asp
Gly	Pro, Ala	Ala
His	Asn, Gln, Lys, Arg	Arg
Ile	Leu, Val, Met, Ala, Phe, Norleucina	Leu
Leu	Norleucina, Ile, Val, Met, Ala, Phe	Ile
Lys	Arg, ácido 1,4-diamino-butírico, Gln, Asn	Arg
Met	Leu, Phe, Ile	Leu
Phe	Leu, Val, Ile, Ala, Tyr	Leu
Pro	Ala	Gly
Ser	Thr, Ala, Cys	Thr
Thr	Ser	Ser
Trp	Tyr, Phe	Tyr
Tyr	Trp, Phe, Thr, Ser	Phe
Val	Ile, Met, Leu, Phe, Ala, Norleucina	Leu

Un experto en la técnica podrá determinar variantes adecuadas del polipéptido tal como se expone en SEQ ID NO: 2 usando técnicas bien conocidas. Para identificar zonas adecuadas de la molécula que pueden cambiarse sin destruir la actividad biológica, un experto en la técnica puede seleccionar como diana zonas que no se cree que sean importantes para la actividad. Por ejemplo, cuando se conocen polipéptidos similares con actividades similares de la misma especie o de otras especies, un experto en la técnica puede comparar la secuencia de aminoácidos de un polipéptido B7-L con tales polipéptidos similares. Con una comparación de este tipo, pueden identificarse restos y partes de las moléculas que se conservan entre polipéptidos similares. Se apreciará que es menos probable que cambios en zonas de la molécula de B7-L que no se conservan con respecto a tales polipéptidos similares afecten adversamente a la actividad biológica y/o estructura de un polipéptido B7-L. Un experto en la técnica también sabrá que, incluso en regiones relativamente conservadas, pueden sustituirse los restos que se producen de manera natural por aminoácidos químicamente similares al tiempo que se conserva la actividad (sustituciones de restos de aminoácido conservativas). Por tanto, incluso zonas que pueden ser importantes para la actividad biológica o para la estructura pueden someterse a sustituciones de aminoácidos conservativas sin destruir la actividad biológica o sin afectar adversamente a la estructura del polipéptido.

Adicionalmente, un experto en la técnica puede revisar estudios de estructura-función que identifican restos en polipéptidos similares que son importantes para la actividad o la estructura. A la vista de una comparación de este tipo, puede predecirse la importancia de restos de aminoácido en un polipéptido B7-L que corresponde a restos de aminoácido que son importantes para la actividad o estructura en polipéptidos similares. Un experto en la técnica puede elegir sustituciones de aminoácidos químicamente similares para tales restos de aminoácido importantes predichos de polipéptidos B7-L.

Un experto en la técnica también puede analizar la estructura tridimensional y la secuencia de aminoácidos en relación con esa estructura en polipéptidos similares. A la vista de tal información, un experto en la técnica puede predecir la alineación de restos de aminoácido de polipéptido B7-L con respecto a su estructura tridimensional. Un experto en la técnica puede elegir no realizar cambios radicales en restos de aminoácido que se predice que están en la superficie de la proteína, ya que tales restos pueden participar en interacciones importantes con otras moléculas. Además, un experto en la técnica puede generar variantes de prueba que contienen una sustitución de un único aminoácido en cada resto de aminoácido. Las variantes pueden examinarse usando ensayos de actividad conocidos por los expertos en la técnica. Tales variantes pueden usarse para recopilar información sobre variantes adecuadas. Por ejemplo, si se descubre que un cambio en un resto de aminoácido particular dio como resultado una actividad destruida, reducida de manera indeseable o no adecuada, se evitarán variantes con tal cambio. En otras palabras, basándose en la información recopilada a partir de tales experimentos de rutina, un experto en la técnica puede determinar fácilmente los aminoácidos en los que deben evitarse sustituciones adicionales o bien solas o bien en combinación con otras mutaciones.

Se han desarrollado varias publicaciones científicas dedicadas a la predicción de la estructura secundaria. Véase Moulton, 1996, *Curr. Opin. Biotechnol.* 7: 422-27; Chou *et al.*, 1974, *Biochemistry* 13: 222-45; Chou *et al.*, 1974, *Biochemistry* 113: 211-22; Chou *et al.*, 1978, *Adv. Enzymol. Relat. Areas Mol. Biol.* 47: 45-48; Chou *et al.*, 1978, *Ann. Rev. Biochem.* 47: 251-276; y Chou *et al.*, 1979, *Biophys. J.* 26: 367-84. Además, actualmente se dispone de programas informáticos para ayudar a la predicción de la estructura secundaria. Un método para predecir la estructura secundaria se basa en la modelización de la homología. Por ejemplo, dos polipéptidos o proteínas que tienen una identidad secuencia superior al 30%, o similitud superior al 40%, con frecuencia tienen topologías estructurales similares. El reciente crecimiento de la base de datos estructurales de proteínas (PDB) ha proporcionado una predictibilidad potenciada de la estructura secundaria, incluyendo el posible número de pliegues dentro de la estructura de un polipéptido o proteína. Véase Holm *et al.*, 1999, *Nucleic Acids Res.* 27: 244-47. Se ha sugerido que hay un número limitado de pliegues en un polipéptido o proteína dado y que una vez resuelto un número crítico de estructuras, la predicción estructural se volverá drásticamente más precisa (Brenner *et al.*, 1997, *Curr. Opin. Struct. Biol.* 7: 369-76).

Métodos adicionales de predicción de la estructura secundaria incluyen "hilvanado" (Jones, 1997, *Curr. Opin. Struct. Biol.* 7: 377-87; Sippl *et al.*, 1996, *Structure* 4: 15-19), "análisis de perfiles" (Bowie *et al.*, 1991, *Science*, 253: 164-70; Gribskov *et al.*, 1990, *Methods Enzymol.* 183: 146-59; Gribskov *et al.*, 1987, *Proc. Nat. Acad. Sci. U. S. A.* 84: 4355-58) y "unión evolutiva" (véase Holm *et al.*, citado anteriormente, y Brenner *et al.*, citado anteriormente).

Además, el polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 2 puede fusionarse con un polipéptido heterólogo para formar un heterodímero. Los péptidos y polipéptidos heterólogos incluyen, pero no se limitan a: un epítopo para permitir la detección y/o el aislamiento de un polipéptido de fusión B7-L; una proteína de receptor transmembrana o una parte de la misma, tal como un dominio extracelular o un dominio transmembrana e intracelular; un ligando o una parte del mismo que se une a una proteína de receptor transmembrana; una enzima o una parte de la misma que es catalíticamente activa; un polipéptido o péptido que fomenta la oligomerización, tal como un dominio de cremallera de leucina; un polipéptido o péptido que aumenta la estabilidad, tal como una región constante de inmunoglobulina; y un polipéptido que tiene una actividad terapéutica diferente del polipéptido que comprende la secuencia de aminoácidos tal como se expone en SEQ ID NO: 2.

Pueden realizarse fusiones o bien en el amino terminal o bien en el carboxilo terminal del polipéptido que comprende la secuencia de aminoácidos expuesta en SEQ ID NO: 2. Las fusiones pueden ser directas sin ligador o molécula adaptadora o pueden ser mediante un ligador o molécula adaptadora. Un ligador o molécula adaptadora puede ser uno o más restos de aminoácido, normalmente desde aproximadamente 20 hasta aproximadamente 50 restos de aminoácido. También puede diseñarse un ligador o molécula adaptadora con un sitio de escisión para una endonucleasa de restricción de ADN o para una proteasa para permitir la separación de los restos fusionados. Se apreciará que una vez construidos, los polipéptidos de fusión pueden derivatizarse según los métodos descritos en el presente documento.

En una realización adicional de la invención, el polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 2 se fusiona con uno o más dominios de una región Fc de IgG humana. Los anticuerpos comprenden dos partes funcionalmente independientes, un dominio variable conocido como "Fab", que se une a un antígeno, y un dominio constante conocido como "Fc", que participa en funciones efectoras tales como activación de complemento y ataque mediante células fagocíticas. Un Fc tiene una larga semivida en suero, mientras que un Fab tiene una vida corta. Capon *et al.*, 1989, *Nature* 337: 525-31. Cuando se construye junto con una proteína terapéutica, un dominio Fc puede proporcionar una semivida más larga o incorporar funciones tales como unión a receptor de Fc, unión a proteína A, fijación a complemento, y quizás incluso transferencia en la placenta *Id.* La tabla II resume el uso de

determinadas fusiones de Fc conocidas en la técnica.

Tabla II

Fusión de Fc con proteínas terapéuticas

Forma de Fc	Pareja de fusión	Implicaciones terapéuticas	Referencia
IgG1	Extremo N-terminal de CD30-L	Enfermedad de Hodgkin; linfoma anaplásico; leucemia de células T	Patente estadounidense n.º 5.480.981
Fcγ2a murino	IL-10	Antiinflamatorio; rechazo de trasplante	Zheng <i>et al.</i> , 1995, J. Immunol. 154: 5590-600
IgG1	Receptor de TNF	Choque séptico	Fisher <i>et al.</i> , 1996, N. Engl. J. Med. 334: 1697-1702; Van Zee <i>et al.</i> , 1996, J. Immunol. 156: 2221-30
IgG, IgA, IgM, o IgE (excluyendo el primer dominio)	Receptor de TNF	Inflamación, trastornos autoinmunitarios	Patente estadounidense n.º 5.808.029
IgG1	Receptor de CD4	SIDA	Capon <i>et al.</i> , 1989, Nature 337: 525-31
IgG1, IgG3	Extremo N-terminal de IL-2	Anticancerígeno, antiviral	Harvill <i>et al.</i> , 1995, Immunotech. 1: 95-105
IgG1	Extremo C-terminal de OPG	Osteoartritis; densidad ósea	Documento WO 97/23614
IgG1	Extremo N-terminal de leptina	Antiobesidad	Documento PCT/US 97/23183, presentado el 11 de diciembre de 1997
IgCγ1 humana	CTLA-4	Trastornos autoinmunitarios	Linsley, 1991, J. Exp. Med., 174: 561-69

- 5 En un ejemplo, puede fusionarse una región de bisagra, CH2 y CH3 de IgG humana o bien en el amino terminal o bien en el carboxilo terminal de los polipéptidos B7-L usando métodos conocidos por el experto en la técnica. En otro ejemplo, puede fusionarse una región de bisagra, CH2 y CH3 de IgG humana o bien en el amino terminal o bien en el carboxilo terminal de un fragmento de polipéptido B7-L (por ejemplo, la parte extracelular predicha de polipéptido B7-L).
- 10 El polipéptido de fusión B7-L resultante puede purificarse mediante el uso de una columna de afinidad de proteína A. Se ha encontrado que los péptidos y proteínas fusionados con una región Fc muestran una semivida *in vivo* sustancialmente superior a la del equivalente no fusionado. Además, una fusión a una región Fc permite la dimerización/multimerización del polipéptido de fusión. La región Fc puede ser una región Fc que se produce de manera natural, o puede alterarse para mejorar determinadas cualidades, tales como cualidades terapéuticas, tiempo de circulación o agregación reducida.

15 La identidad y similitud de moléculas de ácido nucleico y polipéptidos relacionados se calculan fácilmente mediante métodos conocidos. Tales métodos incluyen, pero no se limitan a, los descritos en Computational Molecular Biology (A. M. Lesk, ed., Oxford University Press 1988); Biocomputing: Informatics and Genome Projects (D. W. Smith, ed., Academic Press 1993); Computer Analysis of Sequence Data (Part 1, A. M. Griffin y H. G. Griffin, eds., Humana Press 1994); G. von Heinle, Sequence Analysis in Molecular Biology (Academic Press 1987); Sequence Analysis Primer (M. Gribskov y J. Devereux, eds., M. Stockton Press 1991); y Carillo *et al.*, 1988, SIAM J Applied Math., 48: 1073.

20 Métodos preferidos para determinar la identidad y/o similitud están diseñados para proporcionar la mayor coincidencia entre las secuencias sometidas a prueba. Se describen métodos para determinar la identidad y similitud en programas informáticos disponibles para el público. Los métodos de programa informático preferidos para determinar la identidad y similitud entre dos secuencias incluyen, pero no se limitan a, el paquete de programas GCG, que incluye GAP (Devereux *et al.*, 1984, Nucleic Acids Res. 12: 387; Genetics Computer Group, University of

Wisconsin, Madison, WI), BLASTP, BLASTN, y FASTA (Altschul *et al.*, 1990, J. Mol. Biol. 215: 403-10). El programa BLASTX está disponible al público en el Centro nacional de información en biotecnología (NCBI) y otras fuentes (Altschul *et al.*, BLAST Manual (NCB NLM NIH, Bethesda, MD); Altschul *et al.*, 1990, citado anteriormente). También puede usarse el algoritmo de Smith Waterman bien conocido para determinar la identidad.

- 5 Determinados esquemas de alineación para alinear dos secuencias de aminoácidos pueden dar como resultado la coincidencia de tan sólo una región corta de las dos secuencias, y esta pequeña región alineada puede tener una identidad de secuencia muy alta aunque no haya relación significativa entre las dos secuencias de longitud completa. Por consiguiente, en una realización preferida, el método de alineación seleccionado (programa GAP) dará como resultado una alineación que abarca al menos 50 aminoácidos contiguos del polipéptido reivindicado.
- 10 Por ejemplo, usando el algoritmo informático GAP (Genetics Computer Group, University of Wisconsin, Madison, WI), se alinean dos polipéptidos para los cuales debe determinarse el porcentaje de identidad de secuencia para obtener una coincidencia óptima de sus aminoácidos respectivos (la "extensión con coincidencia", según se determina por el algoritmo). Se usan una penalización por apertura de hueco (que se calcula como 3X la diagonal promedio; la "diagonal promedio" es el promedio de la diagonal de la matriz de comparación que se usa; la
- 15 "diagonal" es la puntuación o el número asignado a cada coincidencia de aminoácido perfecta por la matriz de comparación particular) y una penalización por extensión de hueco (que es habitualmente 0,1X la penalización por apertura de hueco), así como una matriz de comparación tal como PAM 250 o BLOSUM 62, junto con el algoritmo. También se usa una matriz de comparación convencional por el algoritmo (véase Dayhoff *et al.*, 5 Atlas of Protein Sequence and Structure (Sup. 3 1978) (matriz de comparación PAM250); Henikoff *et al.*, 1992, Proc. Natl. Acad. Sci
- 20 USA 89: 10915-19 (matriz de comparación BLOSUM 62)).

Los parámetros preferidos para la comparación de secuencias de polipéptidos incluyen los siguientes:

Algoritmo: Needleman y Wunsch, 1970, J. Mol. Biol. 48: 443-53;

Matriz de comparación: BLOSUM 62 (Henikoff *et al.*, citado anteriormente);

Penalización por hueco: 12

- 25 Penalización por longitud de hueco: 4

Umbral de similitud: 0

El programa GAP es útil con los parámetros anteriores. Los parámetros mencionados anteriormente son los parámetros por defecto para comparaciones de polipéptidos (junto con ninguna penalización para huecos terminados) usando el algoritmo GAP.

- 30 Los parámetros preferidos para la comparación de secuencias de moléculas ácido nucleico incluyen los siguientes:

Algoritmo: Needleman y Wunsch, citado anteriormente;

Matriz de comparación: coincidencias = +10, coincidencia errónea = 0

Penalización por hueco: 50

Penalización por longitud de hueco: 3

- 35 El programa GAP también es útil con los parámetros anteriores. Los parámetros mencionados anteriormente son los parámetros por defecto para comparaciones de moléculas de ácido nucleico.

- 40 Pueden usarse otros algoritmos, penalizaciones por apertura de huecos, penalizaciones por extensión de huecos, matrices de comparación y umbrales de similitud a modo de ejemplo, incluyendo los expuestos en el manual del programa, paquete Wisconsin, versión 9, septiembre de 1997. Las elecciones particulares que deben realizarse resultarán evidentes para los expertos en la técnica y dependerán de la comparación específica que debe realizarse, tal como ADN con ADN, proteína con proteína, proteína con ADN; y adicionalmente, de si la comparación es entre pares de secuencias dados (en cuyo caso se prefieren generalmente GAP o BestFit) o entre una secuencia y una gran base de datos de secuencias (en cuyo caso se prefiere FASTA o BLASTA).

Moléculas de ácido nucleico

- 45 Las moléculas de ácido nucleico que codifican un polipéptido que comprende la secuencia de aminoácidos de un polipéptido B7-L pueden obtenerse fácilmente de una variedad de maneras incluyendo, sin limitación, síntesis química, exploración de biblioteca de ADNc o genómica, exploración de biblioteca de expresión, y/o amplificación por PCR de ADNc.

- 50 Los métodos de ADN recombinante usados en el presente documento son generalmente los expuestos en Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1989) y/o Current Protocols in Molecular Biology (Ausubel *et al.*, eds., Green Publishers Inc. y Wiley and Sons 1994). La invención

proporciona moléculas de ácido nucleico tal como se describe en el presente documento y métodos para obtener tales moléculas.

5 Cuando se ha identificado un gen que codifica la secuencia de aminoácidos de un polipéptido B7-L a partir de una especie, puede usarse la totalidad o parte de ese gen como sonda para identificar ortólogos o genes relacionados de la misma especie. Las sondas o los cebadores pueden usarse para explorar bibliotecas de ADNc de diversas fuentes tisulares que se cree que expresan el polipéptido B7-L. Además, también puede usarse la totalidad o parte de una molécula de ácido nucleico que tiene la secuencia tal como se expone en SEQ ID NO: 1 para explorar una biblioteca genómica para identificar y aislar un gen que codifica la secuencia de aminoácidos de un polipéptido B7-L. Normalmente, se emplearán condiciones de rigurosidad moderada o alta para explorar para minimizar el número de falsos positivos obtenidos en la exploración.

10 También pueden identificarse moléculas de ácido nucleico que codifican la secuencia de aminoácidos de polipéptidos B7-L mediante clonación de expresión que emplea la detección de clones positivos basándose en una propiedad de la proteína expresada. Normalmente, se exploran bibliotecas de ácido nucleico mediante la unión de un anticuerpo u otra pareja de unión (por ejemplo, receptor o ligando) a proteínas clonadas que se expresan y se presentan en la superficie de una célula huésped. El anticuerpo o la pareja de unión se modifica con un marcador detectable para identificar las células que expresan el clon deseado.

15 Pueden seguirse técnicas de expresión recombinante realizadas según las descripciones expuestas a continuación para producir estos polinucleótidos y para expresar los polipéptidos codificados. Por ejemplo, insertando una secuencia de ácido nucleico que codifica la secuencia de aminoácidos de un polipéptido B7-L en un vector apropiado, un experto en la técnica puede producir fácilmente grandes cantidades de la secuencia de nucleótidos deseada. Las secuencias pueden usarse entonces para generar sondas de detección o cebadores de amplificación. Alternativamente, puede insertarse un polinucleótido que codifica la secuencia de aminoácidos de un polipéptido B7-L en un vector de expresión. Introduciendo el vector de expresión en un huésped apropiado, puede producirse el polipéptido B7-L codificado en grandes cantidades.

25 Otro método para obtener una secuencia de ácido nucleico adecuada es la reacción en cadena de la polimerasa (PCR). En este método, se prepara ADNc a partir de poli (A)+ARN o ARN total usando la enzima transcriptasa inversa. Entonces se añaden dos cebadores, normalmente complementarios a dos regiones separadas de ADNc que codifica la secuencia de aminoácidos de un polipéptido B7-L, al ADNc junto con una polimerasa tal como polimerasa de Taq, y la polimerasa amplifica la región de ADNc entre los dos cebadores.

30 Otro medio de preparación de una molécula de ácido nucleico que codifica la secuencia de aminoácidos de un polipéptido B7-L es la síntesis química usando métodos bien conocidos por el experto en la técnica tales como los descritos por Engels *et al.*, 1989, Angew. Chem. Intl. Ed. 28: 716-34. Estos métodos incluyen, entre otros, los métodos de fosfotriéster, fosforamidita, y H-fosfonato para la síntesis de ácido nucleico. Un método preferido para tal síntesis química es la síntesis soportada en polímero usando química de fosforamidita convencional. Normalmente, el ADN que codifica la secuencia de aminoácidos de un polipéptido B7-L tendrá varios cientos de nucleótidos de longitud. Pueden sintetizarse ácidos nucleicos de más de aproximadamente 100 nucleótidos como varios fragmentos usando estos métodos. Entonces pueden ligarse los fragmentos juntos para formar la secuencia de nucleótidos de longitud completa de un gen B7-L. Habitualmente, el fragmento de ADN que codifica el extremo amino terminal del polipéptido tendrá un ATG, que codifica un resto de metionina. Esta metionina puede estar presente o no en la forma madura del polipéptido B7-L, dependiendo de si el polipéptido producido en la célula huésped está diseñado para secretarse de esa célula. También pueden usarse otros métodos conocidos por el experto en la técnica.

35 En determinadas realizaciones, las variantes de ácido nucleico contienen codones que se han alterado para una expresión óptima de un polipéptido B7-L en una célula huésped dada. Las alteraciones de codones particulares dependerán del polipéptido B7-L y la célula huésped seleccionada para la expresión. Tal "optimización de codón" puede llevarse a cabo mediante una variedad de métodos, por ejemplo, seleccionando codones que se prefieren para su uso en genes altamente expresados en una célula huésped dada. Pueden usarse algoritmos informáticos que incorporan tablas de frecuencia de codones tales como "Eco_high.Cod" para la preferencia de codones de genes bacterianos altamente expresados y se proporcionan por el paquete de la Universidad de Wisconsin, versión 9.0 (Genetics Computer Group, Madison, WI). Otras tablas de frecuencia de codones útiles incluyen "Celegans_high.cod", "Celegans_low.cod", "Drosophila_high.cod", "Human_high.cod", "Maize_high.cod", y "Yeast_high.cod".

45 En algunos casos, puede ser deseable preparar moléculas de ácido nucleico que codifican variantes de polipéptido B7-L. Pueden producirse moléculas de ácido nucleico que codifican variantes usando mutagénesis dirigida al sitio, amplificación por PCR u otros métodos apropiados, en los que el/los cebador(es) tiene(n) las mutaciones puntuales deseadas (véase Sambrook *et al.*, citado anteriormente, y Ausubel *et al.*, citado anteriormente, para descripciones de técnicas de mutagénesis). También puede usarse síntesis química usando métodos descritos por Engels *et al.*, citado anteriormente, para preparar tales variantes. También pueden usarse otros métodos conocidos por el experto en la técnica.

Vectores y células huésped

Se inserta una molécula de ácido nucleico que codifica la secuencia de aminoácidos de un polipéptido B7-L en un vector de expresión apropiado usando técnicas de ligación convencionales. El vector se selecciona normalmente para que sea funcional en la célula huésped particular empleada (es decir, el vector es compatible con la maquinaria de la célula huésped de tal manera que puede producirse la amplificación del gen y/o la expresión del gen). Puede amplificarse/expresarse una molécula de ácido nucleico que codifica la secuencia de aminoácidos de un polipéptido B7-L en células huésped procariotas, de levadura, de insecto (sistemas de baculovirus) y/o eucariotas. La selección de la célula huésped dependerá en parte de si debe modificarse tras la traducción un polipéptido B7-L (por ejemplo, glicosilarse y/o fosforilarse). Si es así, son preferibles células huésped de levadura, de insecto o de mamífero. Para una revisión de vectores de expresión, véase Meth. Enz., vol. 185 (D. V. Goeddel, ed., Academic Press 1990).

Normalmente, los vectores de expresión usados en cualquiera de las células huésped contendrán secuencias para el mantenimiento de plásmidos y para la clonación y expresión de secuencias de nucleótidos exógenas. Tales secuencias, denominadas colectivamente "secuencias flanqueantes" en determinadas realizaciones, incluirán normalmente una o más de las siguientes secuencias de nucleótidos: un promotor, una o más secuencias potenciadoras, un origen de replicación, una secuencia de terminación transcripcional, una secuencia de intrón completo que contiene un sitio de corte y empalme dador y aceptor, una secuencia que codifica una secuencia líder para la secreción del polipéptido, un sitio de unión a ribosoma, una secuencia de poliadenilación, una región de poliligador para insertar el ácido nucleico que codifica el polipéptido que va a expresarse, y un elemento marcador seleccionable. Cada una de estas secuencias se comenta a continuación.

Opcionalmente, el vector puede contener una secuencia codificante de "etiqueta", es decir, una molécula de oligonucleótido ubicada en el extremo 5' o 3' de la secuencia codificante del polipéptido B7-L; la secuencia de oligonucleótido codifica poliHis (tal como hexaHis), u otra "etiqueta" tal como FLAG, HA (hemaglutinina del virus de la gripe), o *myc* para la que existen anticuerpos comercialmente disponibles. Esta etiqueta se fusiona normalmente con el polipéptido tras la expresión del polipéptido, y puede servir como medio para la purificación por afinidad del polipéptido B7-L a partir de la célula huésped. La purificación por afinidad puede lograrse, por ejemplo, mediante cromatografía en columna usando anticuerpos frente a la etiqueta como matriz de afinidad. Opcionalmente, la etiqueta puede retirarse posteriormente del polipéptido B7-L purificado mediante diversos medios tales como usando determinadas peptidasas para la escisión.

Las secuencias flanqueantes pueden ser homólogas (es decir, de la misma especie y/o cepas que la célula huésped), heterólogas (es decir, de una especie distinta de la especie o cepa de la célula huésped), híbridas (es decir, una combinación de secuencias flanqueantes de más de una fuente), o sintéticas, o las secuencias flanqueantes pueden ser secuencias nativas que funcionan normalmente para regular la expresión del polipéptido B7-L. Como tal, la fuente de una secuencia flanqueante puede ser cualquier organismo procariota o eucariota, cualquier organismo vertebrado o invertebrado, o cualquier planta, siempre que la secuencia flanqueante sea funcional en, y pueda activarse por, la maquinaria de la célula huésped.

Pueden obtenerse secuencias flanqueantes útiles en los vectores de esta invención mediante cualquiera de diversos métodos bien conocidos en la técnica. Normalmente, las secuencias flanqueantes útiles en el presente documento (distintas de las secuencias flanqueantes del gen B7-L) se habrán identificado previamente mediante mapeo y/o mediante digestión con endonucleasas de restricción y por tanto pueden aislarse de la fuente tisular apropiada usando las endonucleasas de restricción apropiadas. En algunos casos, puede conocerse la secuencia de nucleótidos completa de una secuencia flanqueante. En este caso, la secuencia flanqueante puede sintetizarse usando los métodos descritos en el presente documento para la síntesis o clonación de ácido nucleico.

Cuando se conoce la totalidad o sólo una parte de la secuencia flanqueante, puede obtenerse usando PCR y/o mediante exploración de una biblioteca genómica con un oligonucleótido y/o fragmento de secuencia flanqueante adecuado de la misma especie o de otra. Cuando no se conoce la secuencia flanqueante, puede aislarse un fragmento de ADN que contiene una secuencia flanqueante a partir de un trozo más grande de ADN que puede contener, por ejemplo, una secuencia codificante o incluso otro gen o genes. Puede lograrse el aislamiento mediante digestión con endonucleasa de restricción para producir el fragmento de ADN apropiado seguido por aislamiento usando purificación en gel de agarosa, columna de cromatografía Qiagen® (Chatsworth, CA) u otros métodos conocidos por el experto en la técnica. La selección de enzimas adecuadas para lograr este fin resultará fácilmente evidente para un experto en la técnica.

Un origen de replicación es normalmente una parte de los vectores de expresión procariotas comercialmente adquiridos, y el origen ayuda en la amplificación del vector en una célula huésped. La amplificación del vector hasta un determinado número de copias puede ser importante, en algunos casos, para la expresión óptima de un polipéptido B7-L. Si el vector de elección no contiene un sitio de origen de replicación, puede sintetizarse uno químicamente basándose en una secuencia conocida y ligarse en el vector. Por ejemplo, el origen de replicación del plásmido pBR322 (New England Biolabs, Beverly, MA) es adecuado para la mayoría de las bacterias gram-negativas y diversos orígenes (por ejemplo, SV40, polioma, adenovirus, virus de estomatitis vesicular (VEV), o virus del papiloma tales como VPH o VPB) son útiles para los vectores de clonación en células de mamíferos. Generalmente, el componente de origen de replicación no es necesario para vectores de expresión de mamíferos (por ejemplo, el origen de SV40 se usa con frecuencia sólo porque contiene el promotor temprano).

Una secuencia de terminación de la transcripción se encuentra ubicada normalmente en 3' del extremo de una región codificante de polipéptido y sirve para terminar la transcripción. Habitualmente, una secuencia de terminación de la transcripción en células procariotas es un fragmento rico en G-C seguido por una secuencia de poli-T. Aunque la secuencia se clona fácilmente a partir de una biblioteca o incluso se adquiere comercialmente como parte de un vector, también puede sintetizarse fácilmente usando métodos para la síntesis de ácido nucleico tales como los descritos en el presente documento.

Un elemento de gen de marcador seleccionable codifica una proteína necesaria para la supervivencia y el crecimiento de una célula huésped que crece en un medio de cultivo selectivo. Los genes de marcadores de selección típicos codifican proteínas que (a) confieren resistencia frente a antibióticos u otras toxinas, por ejemplo, ampicilina, tetraciclina, o kanamicina para células huésped procariotas; (b) complementan deficiencias auxotróficas de la célula; o (c) suministran nutrientes críticos no disponibles a partir de medios complejos. Marcadores seleccionables preferidos son el gen de resistencia a kanamicina, el gen de resistencia a ampicilina y el gen de resistencia a tetraciclina. También puede usarse un gen de resistencia a neomicina para la selección en células huésped procariotas y eucariotas.

Pueden usarse otros genes de selección para amplificar el gen que se expresará. La amplificación es el proceso en el que los genes que están en mayor demanda para la producción de una proteína crítica para el crecimiento se reiteran en tándem dentro de los cromosomas de generaciones sucesivas de células recombinantes. Ejemplos de marcadores seleccionables adecuados para células de mamíferos incluyen dihidrofolato reductasa (DHFR) y timidina cinasa. Los transformantes de células de mamíferos se colocan bajo presión de selección en la que sólo los transformantes están únicamente adaptados para sobrevivir gracias al gen de selección presente en el vector. La presión de selección se impone cultivando las células transformadas en condiciones en las que la concentración del agente de selección en el medio se cambia sucesivamente, conduciendo así a la amplificación tanto del gen de selección como del ADN que codifica un polipéptido B7-L. Como resultado, se sintetizan cantidades aumentadas de polipéptido B7-L a partir del ADN amplificado.

Habitualmente es necesario un sitio de unión a ribosoma para el inicio de la traducción de ARNm y se caracteriza por una secuencia de Shine-Dalgarno (procariotas) o una secuencia de Kozak (eucariotas). El elemento se encuentra ubicado normalmente en 3' con respecto al promotor y en 5' con respecto a la secuencia codificante de un polipéptido B7-L que va a expresarse. La secuencia de Shine-Dalgarno es variada pero normalmente es una polipurina (es decir, que tiene un alto contenido en A-G). Se han identificado muchas secuencias de Shine-Dalgarno, cada una de las cuales puede sintetizarse fácilmente usando métodos expuestos en el presente documento y usarse en un vector procariota.

Puede usarse una secuencia líder, o señal, para dirigir un polipéptido B7-L al exterior de la célula huésped. Normalmente, una secuencia de nucleótidos que codifica la secuencia señal se encuentra colocada en la región codificante de una molécula de ácido nucleico B7-L, o directamente en el extremo 5' de una región codificante de polipéptido B7-L. Se han identificado muchas secuencias señal, y puede usarse cualquiera de las que son funcionales en la célula huésped seleccionada junto con una molécula de ácido nucleico B7-L. Por tanto, una secuencia señal puede ser homóloga (que se produce de manera natural) o heteróloga con respecto a la molécula de ácido nucleico B7-L. Adicionalmente, una secuencia señal puede sintetizarse químicamente usando métodos descritos en el presente documento. En la mayoría de los casos, la secreción de un polipéptido B7-L a partir de la célula huésped mediante la presencia de un péptido señal dará como resultado la eliminación del péptido señal del polipéptido B7-L secretado. La secuencia señal puede ser un componente del vector, o puede ser una parte de una molécula de ácido nucleico B7-L que se inserta en el vector.

Se incluye dentro del alcance de esta invención el uso de o bien una secuencia de nucleótidos que codifica una secuencia señal de polipéptido B7-L nativa unida a una región codificante de polipéptido B7-L o bien una secuencia de nucleótidos que codifica una secuencia señal heteróloga unida a una región codificante de polipéptido B7-L. La secuencia señal heteróloga seleccionada debe ser una que se reconoce y se procesa, es decir, se escinde por una señal peptidasa, por la célula huésped. Para células huésped procariotas que no reconocen y procesan la secuencia señal de polipéptido B7-L nativa, se sustituye la secuencia señal por una secuencia señal procariota seleccionada, por ejemplo, del grupo de líderes de fosfatasa alcalina, penicilinas, o enterotoxina II termoestable. Para la secreción de levadura, la secuencia señal de polipéptido B7-L nativa puede sustituirse por líderes de invertasa, factor alfa o fosfatasa ácida de levadura. En la expresión en células de mamíferos, la secuencia señal nativa es satisfactoria, aunque otras secuencias señal de mamíferos pueden ser adecuadas.

En algunos casos, tales como cuando se desea la glicosilación en un sistema de expresión de célula huésped eucariota, pueden manipularse las diversas secuencias previas para mejorar la glicosilación o el rendimiento. Por ejemplo, puede alterarse el sitio de escisión de peptidasa de un péptido señal particular, o añadirse secuencias posteriores, que también pueden afectar a la glicosilación. El producto de proteína final puede tener, en la posición -1 (con respecto al primer aminoácido de la proteína madura) uno o más aminoácidos adicionales que inciden en la expresión, que pueden no haberse retirado totalmente. Por ejemplo, el producto de proteína final puede tener uno o dos restos de aminoácido encontrados en el sitio de escisión de la peptidasa, unidos al extremo amino terminal. Alternativamente, el uso de algunos sitios de escisión por enzimas puede dar como resultado una forma ligeramente troncada del polipéptido B7-L deseado, si la enzima corta en tal zona dentro del polipéptido maduro.

En muchos casos, la transcripción de una molécula de ácido nucleico se aumenta mediante la presencia de uno o más intrones en el vector; esto es particularmente cierto cuando se produce un polipéptido en células huésped eucariotas, especialmente células huésped de mamífero. Los intrones usados pueden producirse de manera natural dentro del gen B7-L especialmente cuando el gen usado es una secuencia genómica de longitud completa o fragmento de la misma. Cuando el intrón no se produce de manera natural dentro del gen (como para la mayoría de los ADNc), el intrón puede obtenerse de otra fuente. La posición del intrón con respecto a las secuencias flanqueantes y el gen B7-L es generalmente importante, ya que el intrón debe transcribirse para ser eficaz. Por tanto, cuando está transcribiéndose una molécula de ADNc de B7-L, la posición preferida para el intrón es 3' con respecto al sitio de inicio de la transcripción y 5' con respecto a la secuencia de terminación de la transcripción poli-A. Preferiblemente, el intrón o los intrones se encontrarán ubicados a un lado o al otro (es decir, en 5' o en 3') del ADNc de tal manera que no interrumpan la secuencia codificante. Puede usarse cualquier intrón de cualquier fuente, incluyendo organismos virales, procariotas y eucariotas (planta o animal), para la práctica de esta invención, siempre que sea compatible con la célula huésped en la que se inserta. También se incluyen en el presente documento intrones sintéticos. Opcionalmente, puede usarse más de un intrón en el vector.

Los vectores de expresión y clonación de la presente invención contendrán normalmente un promotor que se reconoce por el organismo huésped y está operativamente unido a la molécula que codifica el polipéptido B7-L. Los promotores son secuencias no transcritas ubicadas aguas arriba (es decir, en sentido 5') del codón de iniciación de un gen estructural (generalmente dentro de aproximadamente 100 a 1000 pb) que controlan la transcripción del gen estructural. Los promotores se agrupan convencionalmente en una de dos clases: promotores inducibles y promotores constitutivos. Los promotores inducibles inician niveles aumentados de transcripción a partir de ADN bajo su control en respuesta a algún cambio en las condiciones de cultivo, tales como la presencia o ausencia de un nutriente o un cambio en la temperatura. Los promotores constitutivos, por otro lado, inician la producción de producto génico continua; es decir, hay poco o ningún control sobre la expresión génica. Se conocen bien un gran número de promotores, reconocidos por una variedad de posibles células huésped. Un promotor adecuado se une operativamente al ADN que codifica el polipéptido B7-L retirando el promotor del ADN original mediante digestión con enzimas de restricción e inserción de la secuencia promotora deseada en el vector. La secuencia promotora de B7-L nativa puede usarse para dirigir la amplificación y/o la expresión de una molécula de ácido nucleico B7-L. Sin embargo, se prefiere un promotor heterólogo si permite una mayor transcripción y mayores rendimientos de la proteína expresada en comparación con el promotor nativo, y si es compatible con el sistema de célula huésped que se ha seleccionado para su uso.

Promotores adecuados para su uso con huéspedes procariotas incluyen los sistemas de promotor de beta-lactamasa y lactosa; fosfatasa alcalina; un sistema de promotor de triptófano (trp); y promotores híbridos tales como el promotor tac. Otros promotores bacterianos conocidos también son adecuados. Sus secuencias se han publicado, permitiendo así a un experto en la técnica ligarlos a la secuencia de ADN deseada, usando ligadores o adaptadores según sea necesario para proporcionar cualquier sitio de restricción útil.

En la técnica también se conocen bien promotores adecuados para su uso con levaduras. Se usan ventajosamente potenciadores de levaduras con promotores de levadura. Se conocen bien los promotores adecuados para su uso con células huésped de mamífero e incluyen, pero no se limitan a, los obtenidos a partir de los genomas de virus tales como virus del poliovirus, virus de la viruela aviar, adenovirus (tal como adenovirus 2), virus del papiloma bovino, virus del sarcoma aviar, citomegalovirus, retrovirus, virus de la hepatitis B y lo más preferiblemente virus del simio 40 (SV40). Otros promotores de mamífero adecuados incluyen promotores de mamífero heterólogos, por ejemplo, promotores de choque térmico y el promotor de actina.

Promotores adicionales que pueden ser de interés en el control de la expresión del gen B7-L incluyen, pero no se limitan a: la región de promotor temprano de SV40 (Berneis y Chambon, 1981, Nature 290: 304-10); el promotor de CMV; el promotor contenido en la repetición terminal larga en 3' del virus del sarcoma de Rous (Yamamoto, *et al.*, 1980, Cell 22: 787-97); el promotor de timidina cinasa del herpes (Wagner *et al.*, 1981, Proc. Natl. Acad. Sci. U. S. A. 78: 1444-45); las secuencias reguladoras del gen de metalotionina (Brinster *et al.*, 1982, Nature 296: 39-42); vectores de expresión procariotas tales como el promotor de beta-lactamasa (Villa-Kamaroff *et al.*, 1978, Proc. Natl. Acad. Sci. U. S. A., 75: 3727-31); o el promotor tac (DeBoer *et al.*, 1983, Proc. Natl. Acad. Sci. U. S. A., 80: 21-25). También resultan interesantes las siguientes regiones de control transcripcional de animales, que muestran especificidad de tejido y se han utilizado en animales transgénicos: la región de control del gen de la elastasa I que es activa en células acinares pancreáticas (Swift *et al.*, 1984, Célula 38: 639-46; Ornitz *et al.*, 1986, Cold Spring Harbor Symp. Quant. Biol. 50: 399-409 (1986); MacDonald, 1987, Hepatology 7: 425-515); la región de control del gen de la insulina que es activa en células beta pancreáticas (Hanahan, 1985, Nature 315: 115-22); la región de control del gen de la inmunoglobulina que es activa en células linfoides (Grosschedl *et al.*, 1984, Cell 38: 647-58; Adames *et al.*, 1985, Nature 318: 53338; Alexander *et al.*, 1987, Mol. Cell. Biol., 7: 1436-44); la región de control del virus de tumor de mama del ratón que es activa en células testiculares, de mama, linfoides y mastocitos (Leder *et al.*, 1986, Cell 45: 485-95); la región de control del gen de la albumina que es activa en el hígado (Pinkert *et al.*, 1987, Genes and Devel. 1: 268-76); la región de control del gen de la alfa-feto-proteína que es activa en el hígado (Krumlauf *et al.*, 1985, Mol. Cell. Biol., 5: 163948; Hammer *et al.*, 1987, Science 235: 53-58); la región de control del gen de la 1-antitripsina alfa que es activa en el hígado (Kelsey *et al.*, 1987, Genes and Devel. 1: 161-71); la región de control del gen de la beta-globina que es activa en células mieloides (Mogam *et al.*, 1985, Nature 315: 338-40; Kollias *et al.*, 1986, Célula 46: 89-94); la región de control del gen de la proteína básica mielina que es activa en

oligodendrocitos en el cerebro (Readhead *et al.*, 1987, Cell 48: 703-12); la región de control del gen de la cadena ligera 2 de la miosina que es activa en el músculo esquelético (Sani, 1985, Nature 314: 283-86); y la región de control del gen de la hormona de liberación gonadotrópica que es activa en el hipotálamo (Mason *et al.*, 1986, Science 234: 1372-78).

Puede insertarse una secuencia potenciadora en el vector para aumentar la transcripción de un ADN que codifica un polipéptido B7-L de la presente invención por eucariotas superiores. Los potenciadores son elementos que actúan en cis de ADN, en general de aproximadamente 10-300 pb de longitud, que actúan sobre el promotor para aumentar la transcripción. Los potenciadores son relativamente independientes de la orientación y la posición. Se han encontrado en 5' y en 3' con respecto a la unidad de transcripción. Se conocen varias secuencias potenciadoras disponibles de genes de mamíferos (por ejemplo, globina, elastasa, albúmina, alfa-feto-proteína e insulina). Sin embargo, normalmente se usará un potenciador de un virus. El potenciador de SV40, el potenciador del promotor temprano de citomegalovirus, el potenciador de poliovirus, y potenciadores de adenovirus son elementos potenciadores a modo de ejemplo para la activación de promotores eucariotas. Aunque un potenciador puede someterse a corte y empalme en el interior del vector en una posición en 5' o en 3' con respecto a una molécula de ácido nucleico B7-L, normalmente se encuentra ubicado en un sitio en 5' con respecto al promotor.

Pueden construirse vectores de expresión de la invención a partir de un vector de partida tal como un vector comercialmente disponible. Tales vectores pueden contener o no todas las secuencias flanqueantes deseadas. Cuando una o más de las secuencias flanqueantes descritas en el presente documento no están ya presentes en el vector, pueden obtenerse individualmente y ligarse en el vector. Un experto en la técnica conoce bien los métodos usados para obtener cada una de las secuencias flanqueantes.

Los vectores preferidos para poner en práctica esta invención son aquellos compatibles con células huésped bacterianas, de insectos y de mamíferos. Tales vectores incluyen, entre otros, pCRII, pCR3, y pcDNA3,1 (Invitrogen, San Diego, CA), pBSII (Stratagene, La Jolla, CA), pET15 (Novagen, Madison, WI), pGEX (Pharmacia Biotech, Piscataway, NJ), pEGFP-N2 (Clontech, Palo Alto, CA), pETL (BlueBacII, Invitrogen), pDSR-alfa (publicación internacional n.º WO 90/14363) y pFastBacDual (Gibco-BRL, Grand Island, NY).

Vectores adecuados adicionales incluyen, pero no se limitan a, cósmidos, plásmidos o virus modificados, pero se apreciará que el sistema de vector debe ser compatible con la célula huésped seleccionada. Tales vectores incluyen, pero no se limitan a, plásmidos tales como derivados del plásmido Bluescript® (un fagémido basado en ColE1 de alto número de copias; Stratagene Cloning Systems, La Jolla CA), plásmidos de clonación por PCR diseñados para clonar productos de PCR amplificados por Taq (por ejemplo, TOPO™ TA Cloning® Kit y derivados del plásmido PCR2.1®; Invitrogen), y vectores de mamífero, de levadura o de virus tales como un sistema de expresión de baculovirus (derivados del plásmido pBacPAK; Clontech).

Tras haberse construido el vector y haberse insertado una molécula de ácido nucleico que codifica un polipéptido B7-L en el sitio apropiado del vector, puede insertarse el vector completado en el interior de una célula huésped adecuada para su amplificación y/o expresión del polipéptido. La transformación de un vector de expresión para un polipéptido B7-L en una célula huésped seleccionada puede lograrse mediante métodos bien conocidos incluyendo métodos tales como transfección, infección, cloruro de calcio, electroporación, microinyección, lipofección, método de DEAE-dextrano u otras técnicas conocidas. El método seleccionado dependerá en parte del tipo de célula huésped que va a usarse. El experto en la técnica conoce bien estos métodos y otros métodos adecuados, y se exponen, por ejemplo, en Sambrook *et al.*, citado anteriormente.

Las células huésped pueden ser células huésped procariotas (tales como *E. coli*) o células huésped eucariotas (tales como una célula de levadura, de insecto o de vertebrado). La célula huésped, cuando se cultiva en condiciones apropiadas, sintetiza un polipéptido B7-L que puede recogerse posteriormente del medio de cultivo (si la célula huésped lo secreta en el medio) o directamente de la célula huésped que lo produce (si no se secreta). La selección de una célula huésped apropiada dependerá de diversos factores, tales como niveles de expresión deseados, modificaciones de polipéptido que son deseables o necesarias para la actividad (tales como glicosilación o fosforilación) y facilidad de plegamiento para dar una molécula biológicamente activa.

En la técnica se conocen varias células huésped adecuadas y muchas están disponibles de la Colección Americana de Cultivos Tipo (ATCC), Manassas, VA. Ejemplos incluyen, pero no se limitan a, células de mamífero, tales como células de ovario de hámster chino (CHO), células CHO DHFR (-) (Urlaub *et al.*, 1980, Proc. Natl. Acad. Sci. U. S. A. 97: 4216-20), células 293T o 293 de riñón embrionario humano (HEK), o células 3T3. La selección de células huésped de mamífero adecuadas y métodos para la transformación, el cultivo, la amplificación, la selección, la producción de producto y la purificación se conocen bien en la técnica. Otras líneas de células de mamífero adecuadas son las líneas de células COS-1 y COS-7 de mono, y la línea de células CV-1. Células huésped de mamífero a modo de ejemplo adicionales incluyen líneas de células de primate y líneas de células de roedores, incluyendo líneas de células transformadas. También son adecuadas células diploides normales, cepas celulares derivadas del cultivo *in vitro* de tejido primario, así como explantes primarios. Las células candidatas pueden ser genotípicamente deficientes en el gen de selección, o pueden contener un gen de selección que actúa predominantemente. Otras líneas de células de mamífero adecuadas incluyen, pero no se limitan a, células N2A de neuroblastoma de ratón, HeLa, células L-929 de ratón, líneas 3T3 derivadas de ratones Swiss, Balb-c o NIH, líneas de células de hámster BHK o HaK. Los expertos en la técnica de la expresión de proteínas conocen y disponen de

cada una de estas líneas celulares.

Las células bacterianas también son similarmente útiles como células huésped adecuadas para la presente invención. Por ejemplo, las diversas cepas de *E. coli* (por ejemplo, HB101, DH5α, DH10, y MC1061) se conocen bien como células huésped en el campo de la biotecnología. También pueden emplearse en este método diversas cepas de *B. subtilis*, *Pseudomonas spp.*, otros *Bacillus spp.*, *Streptomyces spp.*, y similares.

También están disponibles muchas cepas de células de levadura conocidas por los expertos en la técnica como células huésped para la expresión de los polipéptidos de la presente invención. Células de levadura preferidas incluyen, por ejemplo, *Saccharomyces cerevisiae* y *Pichia pastoris*.

Adicionalmente, cuando se desea, pueden utilizarse sistemas de células de insectos en los métodos de la presente invención. Tales sistemas se describen, por ejemplo, en Kitts *et al.*, 1993, *Biotechniques*, 14: 810-17; Lucklow, 1993, *Curr. Opin. Biotechnol.* 4: 564-72; y Lucklow *et al.*, 1993, *J. Virol.*, 67: 4566-79. Células de insecto preferidas son Sf-9 y Hi5 (Invitrogen).

También pueden usarse animales transgénicos para expresar polipéptidos B7-L glicosilados. Por ejemplo, puede usarse un animal transgénico productor de leche (una vaca o una cabra, por ejemplo) y obtener el presente polipéptido glicosilado en la leche del animal. También pueden usarse plantas para producir polipéptidos B7-L, sin embargo, en general, la glicosilación que se produce en plantas es diferente de la que se produce en células de mamífero, y puede dar como resultado un producto glicosilado que no es adecuado para el uso terapéutico en seres humanos.

Producción de polipéptido

Pueden cultivarse células huésped que comprenden un vector de expresión de polipéptido B7-L usando medios convencionales bien conocidos por el experto en la técnica. Los medios contendrán habitualmente todos los nutrientes necesarios para el crecimiento y la supervivencia de las células. Medios adecuados para cultivar células *E. coli* incluyen, por ejemplo, caldo Luria (LB) y/o caldo Terrific (TB). Medios adecuados para cultivar células eucariotas incluyen medio 1640 del Roswell Park Memorial Institute (RPMI 1640), medio esencial mínimo (MEM) y/o medio de Eagle modificado por Dulbecco (DMEM), todos los cuales pueden complementarse con suero y/o factores de crecimiento según sea necesario para la línea celular particular que está cultivándose. Un medio adecuado para cultivos de insecto es el medio de Grace complementado con Yeastolate, hidrolizado de lactoalbúmina y/o suero de ternero fetal según sea necesario.

Normalmente, se añade un antibiótico u otro compuesto útil para el crecimiento selectivo de células transfectadas o transformadas como complemento a los medios. El compuesto que va a usarse vendrá dictado por el elemento marcador seleccionable presente en el plásmido con el que se transformó la célula huésped. Por ejemplo, cuando el elemento marcador seleccionable es resistencia a kanamicina, el compuesto añadido al medio de cultivo será kanamicina. Otros compuestos para el crecimiento selectivo incluyen ampicilina, tetraciclina y neomicina.

La cantidad de un polipéptido B7-L producido por una célula huésped puede evaluarse usando métodos convencionales conocidos en la técnica. Tales métodos incluyen, sin limitación, análisis por inmunotransferencia de tipo Western, electroforesis en SDS-gel de poli(acrilamida), electroforesis en gel no desnaturante, separación por cromatografía de líquidos de alto rendimiento (HPLC), inmunoprecipitación, y/o ensayos de actividad tales como ensayos de retardo en gel por unión de ADN.

Si se ha diseñado un polipéptido B7-L para secretarse de las células huésped, la mayoría del polipéptido puede encontrarse en el medio de cultivo celular. Sin embargo, si el polipéptido B7-L no se secreta de las células huésped, estará presente en el citoplasma y/o el núcleo (para células huésped eucariotas) o en el citosol (para células huésped de bacterias gram-negativas).

Para un polipéptido B7-L situado en el citoplasma y/o el núcleo de la célula huésped (para células huésped eucariotas) o en el citosol (para células huésped bacterianas), el material intracelular (incluyendo cuerpos de inclusión para bacterias gram-negativas) puede extraerse de la célula huésped usando cualquier técnica convencional conocida por el experto en la técnica. Por ejemplo, pueden someterse las células huésped a lisis para liberar el contenido del periplasma/citoplasma mediante prensa francesa, homogenización y/o sonicación seguido por centrifugación.

Si un polipéptido B7-L ha formado cuerpos de inclusión en el citosol, los cuerpos de inclusión pueden unirse con frecuencia a las membranas celulares interna y/o externa y por tanto se encontrarán principalmente en el material de sedimento tras la centrifugación. El material de sedimento puede tratarse entonces a pH extremos o con un agente caotrópico tal como un detergente, guanidina, derivados de guanidina, urea, o derivados de urea en presencia de un agente reductor tal como ditiotreitól a pH alcalino o tris-carboxietil-fosfina a pH ácido para liberar, romper y solubilizar los cuerpos de inclusión. Entonces puede analizarse el polipéptido B7-L solubilizado usando electroforesis en gel, inmunoprecipitación, o similares. Si se desea aislar el polipéptido B7-L, el aislamiento puede lograrse usando métodos convencionales tales como los descritos en el presente documento y en Marston *et al.*, 1990, *Meth. Enz.*, 182: 264-75.

En algunos casos, un polipéptido B7-L puede no ser biológicamente activo tras el aislamiento. Pueden usarse diversos métodos para “replegar” o convertir el polipéptido en su estructura terciaria y generar puentes disulfuro para restaurar la actividad biológica. Tales métodos incluyen exposición del polipéptido solubilizado a un pH habitualmente superior a 7 y en presencia de una concentración particular de un caotrope. La selección de caotrope es muy similar a las elecciones usadas para la solubilización de los cuerpos de inclusión, pero habitualmente el caotrope se usa a una concentración inferior y no es necesariamente el mismo que los caotrope usados para la solubilización. En la mayoría de los casos la disolución de repliegue/oxidación también contendrá un agente reductor o el agente reductor más su forma oxidada en una razón específica para generar un potencial redox particular que permite que se produzca una reorganización de puentes disulfuro en la formación de los puentes de cisteína de la proteína. Algunos de los pares redox comúnmente usados incluyen cisteína/cistamina, glutatión (GSH)/ditiobis-GSH, cloruro cúprico, ditioneitol (DTT)/ditiano-DTT, y 2-2-mercaptoetanol (bME)/ditio-b(ME). En muchos casos, puede usarse o puede necesitarse un codisolvente para aumentar la eficacia del repliegue, y los reactivos más comunes usados para este fin incluyen glicerol, polietilenglicol de diversos pesos moleculares, arginina y similares.

Si no se forman cuerpos de inclusión en un grado significativo tras la expresión de un polipéptido B7-L, entonces el polipéptido se encontrará principalmente en el sobrenadante tras la centrifugación del homogeneizado celular. El polipéptido puede aislarse adicionalmente del sobrenadante usando métodos tales como los descritos en el presente documento.

La purificación de un polipéptido B7-L de la disolución puede lograrse usando una variedad de técnicas. Si el polipéptido se ha sintetizado de tal manera que contiene una etiqueta tal como hexahistidina (polipéptido B7-L/hexaHis) u otro péptido pequeño tal como FLAG (Eastman Kodak Co., New Haven, CT) o *myc* (Invitrogen) en su extremo o bien carboxilo o bien amino terminal, puede purificarse en un procedimiento de una etapa haciendo pasar la disolución a través de una columna de afinidad en la que la matriz de la columna tiene una alta afinidad por la etiqueta.

Por ejemplo, la polihistidina se une con gran afinidad y especificidad al níquel. Por tanto, puede usarse una columna de afinidad de níquel (tal como las columnas de níquel de Qiagen®) para la purificación del polipéptido B7-L/poliHis. Véase, por ejemplo, Current Protocols in Molecular Biology §10.11.8 (Ausubel *et al.*, eds., Green Publishers Inc. y Wiley and Sons 1993).

Adicionalmente, pueden purificarse polipéptidos B7-L mediante el uso de un anticuerpo monoclonal que puede reconocer y unirse específicamente a un polipéptido B7-L.

Otros procedimientos adecuados para la purificación incluyen, sin limitación, cromatografía de afinidad, cromatografía de inmunoafinidad, cromatografía de intercambio iónico, cromatografía de exclusión molecular, HPLC, electroforesis (incluyendo electroforesis en gel nativo) seguida por elución en gel, e isoelectroenfoque preparativo (máquina/técnica “Isoprime”, Hoefer Scientific, San Francisco, CA). En algunos casos, pueden combinarse dos o más técnicas de purificación para lograr una pureza aumentada.

También pueden prepararse polipéptidos B7-L mediante métodos de síntesis química (tales como síntesis de péptidos en fase sólida) usando técnicas conocidas en la técnica tales como las expuestas por Merrifield *et al.*, 1963, J. Am. Chem. Soc. 85: 2149; Houghten *et al.*, 1985, Proc Natl Acad. Sci. USA 82: 5132; y Stewart y Young, Solid Phase Peptide Synthesis (Pierce Chemical Co. 1984). Tales polipéptidos pueden sintetizarse con o sin una metionina en el extremo amino terminal. Los polipéptidos B7-L químicamente sintetizados pueden oxidarse usando métodos expuestos en esas referencias para formar puentes disulfuro. Se espera que los polipéptidos B7-L químicamente sintetizados tengan una actividad biológica comparable a los polipéptidos B7-L correspondientes producidos de manera recombinante o purificados a partir de fuentes naturales, y por tanto pueden usarse de manera intercambiable con un polipéptido B7-L natural o recombinante.

Otros medios de obtención del polipéptido B7-L es mediante purificación a partir de muestras biológicas tales como tejidos y/o fluidos fuente en los que se encuentra de manera natural el polipéptido B7-L. Tal purificación puede realizarse usando métodos para la purificación de proteínas tal como se describe en el presente documento. La presencia del polipéptido B7-L durante la purificación puede monitorizarse, por ejemplo, usando un anticuerpo preparado contra polipéptido B7-L producido de manera recombinante o fragmentos peptídicos del mismo.

En la técnica se conocen varios métodos adicionales para producir ácidos nucleicos y polipéptidos, y los métodos pueden usarse para producir polipéptidos que tienen especificidad por el polipéptido B7-L. Véase, por ejemplo, Roberts *et al.*, 1997, Proc. Natl. Acad. Sci. U. S. A. 94: 12297-303, que describe la producción de proteínas de fusión entre un ARNm y su péptido codificado. Véase también, Roberts, 1999, Curr. Opin. Chem. Biol. 3: 268-73. Adicionalmente, la patente estadounidense n.º 5.824.469 describe métodos para obtener oligonucleótidos que pueden llevar a cabo una función biológica específica. El procedimiento implica generar una combinación heterogénea de oligonucleótidos, que tienen cada uno una secuencia aleatorizada en 5', una secuencia preseleccionada central, y una secuencia aleatorizada en 3'. La combinación heterogénea resultante se introduce en una población de células que no muestran la función biológica deseada. Entonces se examinan subpoblaciones de las células para seleccionar las que muestran una función biológica predeterminada. A partir de esa subpoblación, se aíslan oligonucleótidos que pueden llevar a cabo la función biológica deseada.

Las patentes estadounidenses n.ºs 5.763.192; 5.814.476; 5.723.323 y 5.817.483 describen procedimientos para producir péptidos o polipéptidos. Esto se realiza produciendo genes estocásticos o fragmentos de los mismos, y después introduciendo esos genes en células huésped que producen una o más proteínas codificadas por los genes estocásticos. Entonces se examinan las células huésped para identificar los clones que producen péptidos o polipéptidos que tienen la actividad deseada.

Otro método para producir péptidos o polipéptidos se describe en la publicación internacional n.º W099/15650, presentada por Athersys, Inc. Conocido como "Random Activation of Gene Expression for Gene Discovery" (RAGE-GD), el procedimiento implica la activación de la expresión de genes endógenos o la sobreexpresión de un gen mediante métodos de recombinación *in situ*. Por ejemplo, se activa o aumenta la expresión de un gen endógeno integrando una secuencia reguladora en la célula diana que puede activar la expresión del gen mediante recombinación no homóloga o ilegítima. En primer lugar se somete el ADN diana a radiación, y se inserta un promotor genético. Eventualmente el promotor localiza una rotura en la parte frontal de un gen, iniciando la transcripción del gen. Esto da como resultado la expresión del péptido o polipéptido deseado.

Se apreciará que estos métodos también pueden usarse para crear bibliotecas de expresión de polipéptido B7-L exhaustivas, que posteriormente pueden usarse para la selección fenotípica de alto rendimiento en una variedad de ensayos, tales como ensayos bioquímicos, ensayos celulares y ensayos en organismos completos (por ejemplo, planta, ratón, etc.).

Síntesis

Los expertos en la técnica apreciarán que las moléculas de ácido nucleico y polipéptido tal como se describen en el presente documento pueden producirse mediante medios recombinantes y otros.

Agentes de unión selectiva / anticuerpos

El término "agente de unión selectiva" se refiere a una molécula que tiene especificidad para uno o más polipéptidos B7-L. Los agentes de unión selectiva adecuados incluyen anticuerpos y derivados de los mismos. Pueden prepararse anticuerpos adecuados usando métodos conocidos en la técnica. Un anticuerpo selectivo para polipéptido B7-L a modo de ejemplo de la presente invención puede unirse a una determinada parte del polipéptido B7-L inhibiendo así la unión del polipéptido a un receptor de polipéptido B7-L.

Los anticuerpos y fragmentos de anticuerpo que se unen a polipéptidos B7-L están dentro del alcance de la presente invención. Los anticuerpos pueden ser policlonales incluyendo policlonal monoespecífico; monoclonal (AcM); recombinante; quimérico; humanizado, tal como con injerto de región determinante de la complementariedad (CDR); humano; de cadena sencilla; y/o biespecífico; así como fragmentos; variantes; o derivados de los mismos. Los fragmentos de anticuerpo incluyen las partes del anticuerpo que se unen a un epítipo en el polipéptido B7-L. Ejemplos de tales fragmentos incluyen fragmentos Fab y F(ab') generados mediante escisión enzimática de anticuerpos de longitud completa. Otros fragmentos de unión incluyen los generados mediante técnicas de ADN recombinante, tales como la expresión de plásmidos recombinantes que contienen secuencias de ácido nucleico que codifican regiones variables de anticuerpos.

Los anticuerpos policlonales dirigidos contra un polipéptido B7-L se producen generalmente en animales (por ejemplo, conejos o ratones) por medio de múltiples inyecciones subcutáneas o intraperitoneales de polipéptido B7-L y un adyuvante. Puede ser útil conjugar un polipéptido B7-L con una proteína portadora que es inmunogénica en la especie que va a inmunizarse, tal como hemocianina de lapa californiana, suero, albúmina, tiroglobulina bovina, o inhibidor de tripsina de semilla de soja. Además, se usan agentes de agregación tales como alumbre para potenciar la respuesta inmunitaria. Tras la inmunización, se extrae sangre de los animales y se somete el suero a ensayo para determinar el título de anticuerpos anti-B7-L.

Se producen anticuerpos monoclonales dirigidos contra polipéptidos B7-L usando cualquier método que proporcione la producción de moléculas de anticuerpo mediante líneas celulares continuas en cultivo. Ejemplos de métodos adecuados para preparar anticuerpos monoclonales incluyen los métodos del hibridoma de Kohler *et al.*, 1975, Nature 256: 495-97 y el método del hibridoma de células B humanas (Kozbor, 1984, J. Immunol. 133: 3001; Brodeur *et al.*, Monoclonal Antibody Production Techniques and Applications 51-63 (Marcel Dekker, Inc., 1987). La invención también proporciona líneas celulares de hibridoma que producen anticuerpos monoclonales reactivos con polipéptidos B7-L.

Los anticuerpos monoclonales de la invención pueden modificarse para su uso como productos terapéuticos. Una realización es un anticuerpo "quimérico" en el que una parte de la cadena pesada (H) y/o ligera (L) es idéntica u homóloga a una secuencia correspondiente en anticuerpos derivados de una especie particular o que pertenecen a una clase o subclase de anticuerpos particular, mientras que el resto de la(s) cadena(s) es idéntico u homólogo a una secuencia correspondiente en anticuerpos derivados de otra especie o que pertenecen a otra clase o subclase de anticuerpos. También se incluyen fragmentos de tales anticuerpos, siempre que muestren la actividad biológica deseada. Véase la patente estadounidense n.º 4.816.567; Morrison *et al.*, 1985, Proc. Natl. Acad. Sci. 81: 6851-55.

En otra realización, un anticuerpo monoclonal de la invención es un anticuerpo "humanizado". En la técnica se

conocen bien métodos para humanizar anticuerpos no humanos. Véase las patentes estadounidenses n.ºs 5.585.089 y 5.693.762. Generalmente, un anticuerpo humanizado tiene uno o más restos de aminoácido introducidos en el mismo procedentes de una fuente que no es humana. La humanización puede realizarse, por ejemplo, usando métodos descritos en la técnica (Jones *et al.*, 1986, Nature 321: 522-25; Riechmann *et al.*, 1998, Nature 332: 323-27; Verhoeyen *et al.*, 1988, Science 239: 1534-36), sustituyendo al menos una parte de una región determinante de la complementariedad de roedor por las regiones correspondientes de un anticuerpo humano.

La invención también abarca anticuerpos humanos que se unen a polipéptidos B7-L. Usando animales transgénicos (por ejemplo, ratones) que pueden producir un repertorio de anticuerpos humanos en ausencia de producción de inmunoglobulina endógena, se producen tales anticuerpos mediante inmunización con un antígeno de polipéptido B7-L (es decir, que tiene al menos 6 aminoácidos contiguos), opcionalmente conjugado con un portador. Véase, por ejemplo, Jakobovits *et al.*, 1993, Proc. Natl. Acad. Sci. 90: 2551-55; Jakobovits *et al.*, 1993, Nature 362: 255-58; Bruggermann *et al.*, 1993, Year in Immuno. 7: 33. En un método, tales animales transgénicos se producen incapacitando los loci endógenos que codifican cadenas de inmunoglobulina pesada y ligera en los mismos, e insertando loci que codifican proteínas de cadena pesada y ligera humanas en el genoma de los mismos. Entonces se cruzan animales parcialmente modificados (es decir, los que tienen menos que el complemento completo de modificaciones) para obtener un animal que tenga todas las modificaciones del sistema inmunitario deseadas. Cuando se administra un inmunógeno, estos animales transgénicos producen anticuerpos con secuencias de aminoácidos humanas (en vez de, por ejemplo, murinas), incluyendo regiones variables que son inmunoespecíficas para esos antígenos. Véanse las solicitudes internacionales n.ºs PCT/US96/05928 y PCT/US93/06926. Se describen métodos adicionales en la patente estadounidense n.º 5.545.807, las solicitudes internacionales n.ºs PCT/US91/245 y PCT/GB89/01207, y en las patentes europeas n.ºs 546073B1 y 546073A1. También pueden producirse anticuerpos humanos mediante la expresión de ADN recombinante en células huésped o mediante expresión en células de hibridoma tal como se describe en el presente documento.

En una realización alternativa, también pueden producirse anticuerpos humanos a partir de bibliotecas de presentación en fagos (Hoogenboom *et al.*, 1991, J. Mol. Biol. 227: 381; Marks *et al.*, 1991, J. Mol. Biol. 222: 581). Estos procesos imitan la selección inmunitaria mediante la presentación de repertorios de anticuerpos sobre la superficie de bacteriófagos filamentosos, y posterior selección de fago mediante su unión a un antígeno de elección. Una de tales técnicas se describe en la solicitud internacional n.º PCT/US98/17364, que describe el aislamiento de anticuerpos agonistas funcionales y de alta afinidad para receptores de MPL y msk usando tal enfoque.

Los anticuerpos quiméricos, con injerto de CDR y humanizados se producen normalmente mediante métodos recombinantes. Se introducen ácidos nucleicos que codifican los anticuerpos en células huésped y se expresan usando materiales y procedimientos descritos en el presente documento. En una realización preferida, los anticuerpos se producen en células huésped de mamífero, tales como células CHO. Pueden producirse anticuerpos monoclonales (por ejemplo, humanos) mediante la expresión de ADN recombinante en células huésped o mediante la expresión en células de hibridoma tal como se describe en el presente documento.

Los anticuerpos anti-B7-L de la invención pueden emplearse en cualquier método de ensayo conocido, tal como ensayos de unión competitiva, ensayos de tipo sándwich directos e indirectos y ensayos de inmunoprecipitación (Sola, Monoclonal Antibodies: A Manual of Techniques 147-158 (CRC Press, Inc., 1987)) para la detección y cuantificación de polipéptidos B7-L. Los anticuerpos se unirán a polipéptidos B7-L con una afinidad que es apropiada para el método de ensayo que está empleándose.

Para aplicaciones de diagnóstico, en determinadas realizaciones, pueden marcarse anticuerpos anti-B7-L con un resto detectable. El resto detectable puede ser uno cualquiera que puede producir, o bien directa o bien indirectamente, una señal detectable. Por ejemplo, el resto detectable puede ser un radioisótopo, tal como ^3H , ^{14}C , ^{32}P , ^{35}S , ^{125}I , ^{99}Tc , ^{111}In , o ^{67}Ga ; un compuesto fluorescente o quimioluminiscente, tal como isotiocianato de fluoresceína, rodamina o luciferina; o una enzima, tal como fosfatasa alcalina, β -galactosidasa o peroxidasa del rábano (Bayer, *et al.*, 1990, Meth. Enz. 184: 138-63).

Los ensayos de unión competitiva se basan en la capacidad de un patrón marcado (por ejemplo, un polipéptido B7-L, o una parte inmunológicamente reactiva del mismo) para competir con el analito de muestra de prueba (un polipéptido B7-L) para unirse con una cantidad limitada de anticuerpo anti-B7-L. La cantidad de un polipéptido B7-L en la muestra de prueba es inversamente proporcional a la cantidad de patrón que se une a los anticuerpos. Para facilitar la determinación de la cantidad de patrón que se une, normalmente se insolubilizan los anticuerpos antes o después de la competición, de modo que el patrón y el analito que están unidos a los anticuerpos pueden separarse convenientemente del patrón y el analito que permanece sin unirse.

Los ensayos de tipo sándwich implican normalmente el uso de dos anticuerpos, cada uno que puede unirse a una parte inmunogénica diferente, o epítipo, de la proteína que va a detectarse y/o cuantificarse. En un ensayo de tipo sándwich, el analito de muestra de prueba se une normalmente a un primer anticuerpo que está inmovilizado en un soporte sólido, y posteriormente un segundo anticuerpo se une al analito, formando así un complejo de tres partes insoluble. Véase, por ejemplo, la patente estadounidense n.º 4.376.110. El segundo anticuerpo puede estar marcado en sí mismo con un resto detectable (ensayos de tipo sándwich directos) o puede medirse usando un anticuerpo anti-inmunoglobulina que está marcado con un resto detectable (ensayos de tipo sándwich indirectos). Por ejemplo, un tipo de ensayo de tipo sándwich es un ensayo de inmunoabsorción ligado a enzimas (ELISA), en cuyo caso el

resto detectable es una enzima.

Los anticuerpos anti-B7-L, también son útiles para la obtención de imágenes *in vivo*. Puede administrarse un anticuerpo marcado con un resto detectable a un animal, preferiblemente en el torrente sanguíneo, y someterse a ensayo la presencia y ubicación del anticuerpo marcado en el huésped. El anticuerpo puede marcarse con cualquier resto que pueda detectarse en un animal, ya sea mediante resonancia magnética nuclear, radiología u otros medios de detección conocidos en la técnica.

Los anticuerpos de la invención pueden usarse como terapéuticos. Estos agentes terapéuticos son generalmente agonistas o antagonistas, porque o bien potencian o bien reducen, respectivamente, al menos una de las actividades biológicas de un polipéptido B7-L. En una realización, los anticuerpos antagonistas de la invención son anticuerpos o fragmentos de unión de los mismos que pueden unirse específicamente a un polipéptido B7-L y que pueden inhibir o eliminar la actividad funcional de un polipéptido B7-L *in vivo* o *in vitro*. En realizaciones preferidas, el agente de unión selectiva, por ejemplo, un anticuerpo antagonista, inhibirá la actividad funcional de un polipéptido B7-L en al menos aproximadamente el 50%, y preferiblemente en al menos aproximadamente el 80%. En otra realización, el agente de unión selectiva puede ser un anticuerpo anti-polipéptido B7-L que puede interaccionar con una pareja de unión de polipéptido B7-L (un ligando o receptor) inhibiendo o eliminando así la actividad de polipéptido B7-L *in vitro* o *in vivo*. Los agentes de unión selectiva, incluyendo anticuerpos anti-polipéptido B7-L agonistas y antagonistas, se identifican mediante ensayos de selección que se conocen bien en la técnica.

Micromatrices

Se apreciará que puede utilizarse la tecnología de micromatriz de ADN según la presente invención. Las micromatrices de ADN son matrices de alta densidad en miniatura de ácidos nucleicos ubicados sobre un soporte sólido, tal como vidrio. Cada celda o elemento dentro de la matriz contiene numerosas copias de una única especie de ácido nucleico que actúa como diana para la hibridación con una secuencia de ácido nucleico complementaria (por ejemplo, ARNm). En la obtención de perfiles de expresión usando tecnología de micromatriz de ADN, en primer lugar se extrae ARNm de una muestra celular o tisular y después se convierte enzimáticamente en ADNc marcado fluorescentemente. Este material se hibrida con la micromatriz y se elimina el ADNc no unido mediante lavado. Entonces se visualiza la expresión de genes diferenciados representados en la matriz cuantificando la cantidad de ADNc marcado que se une específicamente a cada molécula de ácido nucleico diana. De esta manera, puede cuantificarse la expresión de miles de genes de una manera paralela, con alto rendimiento, a partir de una única muestra de material biológico.

Esta obtención de perfiles con alto rendimiento tiene una amplia gama de aplicaciones con respecto a las moléculas de B7-L de la invención, incluyendo, pero sin limitarse a: la identificación y validación de genes relacionados con enfermedades de B7-L como dianas para productos terapéuticos; toxicología molecular de moléculas de B7-L relacionadas e inhibidores de las mismas; estratificación de poblaciones y generación de marcadores sustitutos para ensayos clínicos; y potenciación del descubrimiento de fármacos de moléculas pequeñas de polipéptidos B7-L relacionados ayudando en la identificación de compuestos selectivos en selecciones de alto rendimiento.

Derivados químicos

Un experto en la técnica puede preparar derivados químicamente modificados de polipéptidos B7-L, dadas las divulgaciones descritas en el presente documento. Los derivados de polipéptido B7-L se modifican de una manera que es diferente, o bien en el tipo o bien en la ubicación de la moléculas unidas de manera natural al polipéptido. Los derivados pueden incluir moléculas formadas mediante la delección de uno o más grupos químicos unidos de manera natural. El polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 2 puede modificarse mediante unión covalente de uno o más polímeros. Por ejemplo, el polímero seleccionado es normalmente soluble en agua de modo que la proteína a la que se une no presenta en un entorno acuoso, tal como un entorno fisiológico. Dentro del alcance de polímeros adecuados se incluye una mezcla de polímeros. Preferiblemente, para el uso terapéutico de la preparación de producto final, el polímero será farmacéuticamente aceptable.

Los polímeros pueden ser cada uno de cualquier peso molecular y pueden ser ramificados o no ramificados. Los polímeros tienen cada uno normalmente un peso molecular promedio de entre aproximadamente 2 kDa y aproximadamente 100 kDa (indicando el término "aproximadamente" que en preparaciones de un polímero soluble en agua, algunas moléculas pesarán más, algunas menos, que el peso molecular mencionado). El peso molecular promedio de cada polímero es preferiblemente de entre aproximadamente 5 kDa y aproximadamente 50 kDa, más preferiblemente de entre aproximadamente 12 kDa y aproximadamente 40 kDa y lo más preferiblemente de entre aproximadamente 20 kDa y aproximadamente 35 kDa.

Polímeros solubles en agua adecuados o mezclas de los mismos incluyen, pero no se limitan a, hidratos de carbono unidos en N o unidos en O, azúcares, fosfatos, polietilenglicol (PEG) (incluyendo las formas de PEG que se han usado para derivatizar proteínas, incluyendo mono-(C₁-C₁₀), alcoxi, o ariloxi-polietilenglicol), monometoxipolietilenglicol, dextrano (tal como dextrano de bajo peso molecular de, por ejemplo, aproximadamente 6 kD), celulosa u otros polímeros a base de hidratos de carbono, poli-(N-vinilpirrolidona)-polietilenglicol, homopolímeros de propilenglicol, copolímeros de óxido de polipropileno/óxido de etileno, polioles polioxiethylados (por ejemplo, glicerol), y poli(alcohol vinílico). La presente invención también abarca moléculas reticuladas bifuncionales

que pueden usarse para preparar multímeros de polipéptido B7-L unidos covalentemente.

En general, puede realizarse la derivatización química en cualquier condición adecuada usada para hacer reaccionar una proteína con una molécula de polímero activada. Los métodos para preparar derivados químicos de polipéptidos comprenderán generalmente las etapas de: (a) hacer reaccionar el polipéptido con la molécula de polímero activada (tal como un derivado de aldehído o éster reactivo de la molécula de polímero) en condiciones en las que el polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 2 se une a una o más moléculas de polímero y (b) obtener los productos de reacción. Las condiciones de reacción óptimas se determinarán basándose en parámetros conocidos y el resultado deseado. Por ejemplo, cuanto mayor es la razón de molécula de polímeros con respecto a proteína, mayor es el porcentaje de molécula de polímero unida. En una realización, el derivado de polipéptido B7-L puede tener un único resto de molécula de polímero en el extremo amino terminal. Véase, por ejemplo, la patente estadounidense n.º 5.234.784.

La pegilación de un polipéptido puede llevarse a cabo específicamente usando cualquiera de las reacciones de pegilación conocidas en la técnica. Tales reacciones se describen, por ejemplo, en las siguientes referencias: Francis *et al.*, 1992, Focus on Growth Factors 3: 4-10; patentes europeas n.ºs 0154316 y 0401384; y patente estadounidense n.º 4.179.337. Por ejemplo, la pegilación puede llevarse a cabo mediante una reacción de acilación o una reacción de alquilación con una molécula de polietilenglicol reactiva (o un polímero soluble en agua reactivo análogo) tal como se describe en el presente documento. Para las reacciones de acilación, un polímero seleccionado debe tener un único grupo éster reactivo. Para la alquilación reductora, un polímero seleccionado debe tener un único grupo aldehído reactivo. Un aldehído reactivo es, por ejemplo, propionaldehído de polietilenglicol, que es soluble en agua, o derivados de monoalcoxilo C₁-C₁₀ o ariloxilo del mismo (véase la patente estadounidense n.º 5.252.714).

En otra realización, pueden acoplarse químicamente polipéptidos B7-L a biotina. Entonces se permite que las moléculas de biotina/polipéptido B7-L se unan a avidina, dando como resultado moléculas tetravalentes de avidina/biotina/polipéptido B7-L. Los polipéptidos B7-L también pueden acoplarse covalentemente a dinitrofenol (DNP) o trinitrofenol (TNP) y precipitarse los conjugados resultantes con IgM anti-DNP o anti-TNP para formar conjugados decaméricos con una valencia de 10.

Generalmente, los estados que pueden aliviarse o modularse mediante la administración de los presentes derivados de polipéptido B7-L incluyen los descritos en el presente documento para polipéptidos B7-L. Sin embargo, los derivados de polipéptido B7-L dados a conocer en el presente documento pueden tener actividades adicionales, actividad biológica potenciada o reducida, u otras características, tales como semivida aumentada o reducida, en comparación con las moléculas no derivatizadas.

Animales no humanos modificados mediante ingeniería genética

Pueden prepararse animales no humanos tales como ratones, ratas u otros roedores; conejos, cabras, ovejas u otros animales de granja, en los que se han alterado (es decir, "desactivado") los genes que codifican polipéptido B7-L nativo de tal manera que el nivel de expresión de polipéptido B7-L se disminuye significativamente o se elimina completamente. Tales animales pueden prepararse usando técnicas y métodos tales como los descritos en la patente estadounidense n.º 5.557.032.

Pueden prepararse animales no humanos tales como ratones, ratas u otros roedores; conejos, cabras, ovejas u otros animales de granja, en los que el animal sobreexpresa o bien la forma nativa de un gen B7-L para ese animal o bien un gen B7-L heterólogo, creando así un animal "transgénico". Tales animales transgénicos pueden prepararse usando métodos bien conocidos tales como los descritos en la patente estadounidense n.º 5.489.743 y la publicación internacional n.º WO 94/28122.

En tales animales no humanos, o bien se activa o bien se inactiva el promotor para uno o más de los polipéptidos B7-L de la presente invención (por ejemplo, usando métodos de recombinación homóloga) para alterar el nivel de expresión de uno o más de los polipéptidos B7-L nativos.

Estos animales no humanos pueden usarse para selección de candidatos farmacológicos. En tal selección, puede medirse el impacto de un candidato farmacológico sobre el animal. Por ejemplo, los candidatos farmacológicos pueden aumentar o disminuir la expresión del gen B7-L. En determinadas realizaciones, la cantidad de polipéptido B7-L que se produce puede medirse tras la exposición del animal al candidato farmacológico. Adicionalmente, puede detectarse el impacto real del candidato farmacológico sobre el animal. Por ejemplo, la sobreexpresión de un gen particular puede dar como resultado, o estar asociada con, una enfermedad o estado patológico. En tales casos, puede someterse a prueba la capacidad de un candidato farmacológico para reducir la expresión del gen o su capacidad para prevenir o inhibir un estado patológico. En otros ejemplos, la producción de un producto metabólico particular tal como un fragmento de un polipéptido, puede dar como resultado, o estar asociada con, una enfermedad o estado patológico. En tales casos, puede someterse a prueba la capacidad del candidato farmacológico para disminuir la producción de tal producto metabólico o su capacidad para prevenir o inhibir un estado patológico.

Ensayo de otros moduladores de la actividad del polipéptido B7-L

En algunas situaciones, puede ser deseable identificar moléculas que son moduladoras, es decir, agonistas o antagonistas, de la actividad de polipéptido B7-L. Pueden identificarse moléculas naturales o sintéticas que modulan el polipéptido B7-L usando uno o más ensayos de selección, tales como los descritos en el presente documento. Tales moléculas pueden administrarse o bien de una manera *ex vivo* o bien de una manera *in vivo* mediante inyección, o mediante administración oral, dispositivo de implantación o similares.

“Molécula de prueba” se refiere a una molécula que está en evaluación para determinar su capacidad para modular (es decir, aumentar o disminuir) la actividad de un polipéptido B7-L. Lo más comúnmente, una molécula de prueba interaccionará con un polipéptido B7-L. Sin embargo, también se contempla que una molécula de prueba también puede modular la actividad de polipéptido B7-L indirectamente, tal como afectando a la expresión del gen B7-L, o uniéndose a una pareja de unión de polipéptido B7-L (por ejemplo, receptor o ligando). En una realización, una molécula de prueba se unirá a un polipéptido B7-L con una constante de afinidad de al menos aproximadamente 10^{-6} M, preferiblemente de aproximadamente 10^{-8} M, más preferiblemente de aproximadamente 10^{-9} M, e incluso más preferiblemente de aproximadamente 10^{-10} M.

Puede incubarse un polipéptido B7-L con una molécula de prueba en condiciones que permiten la interacción de la molécula de prueba con un polipéptido B7-L, y se mide el grado de la interacción. La molécula de prueba puede seleccionarse en una forma sustancialmente purificada o en una mezcla cruda.

Un agonista o antagonista de polipéptido B7-L puede ser una proteína, péptido, hidrato de carbono, lípido o molécula de pequeño peso molecular que interacciona con el polipéptido B7-L para regular su actividad. Moléculas que regulan la expresión del polipéptido B7-L incluyen ácidos nucleicos que son complementarios a ácidos nucleicos que codifican un polipéptido B7-L, o son complementarios a secuencias de ácido nucleico que dirigen o controlan la expresión de polipéptido B7-L, y que actúan como reguladores antisentido de la expresión.

Una vez se ha identificado que una molécula de prueba interacciona con un polipéptido B7-L, la molécula puede evaluarse adicionalmente para determinar su capacidad para aumentar o disminuir la actividad del polipéptido B7-L. La medición de la interacción de una molécula de prueba con polipéptido B7-L puede llevarse a cabo en diversos formatos, incluyendo ensayos de unión basados en células, ensayos de unión a membrana, ensayos en fase de disolución e inmunoensayos. En general, se incuba una molécula de prueba con un polipéptido B7-L durante un periodo de tiempo especificado, y se determina la actividad del polipéptido B7-L mediante uno o más ensayos para medir la actividad biológica.

También puede someterse a ensayo la interacción de moléculas de prueba con polipéptidos B7-L directamente usando anticuerpos policlonales o monoclonales en un inmunoensayo. Alternativamente, pueden usarse formas modificadas de polipéptidos B7-L que contienen las etiquetas de epítipo tal como se describe en el presente documento en disolución e inmunoensayos.

En el caso de que los polipéptidos B7-L presenten actividad biológica mediante una interacción con una pareja de unión (por ejemplo, un receptor o un ligando), puede usarse una variedad de ensayos *in vitro* para medir la unión de un polipéptido B7-L a la correspondiente pareja de unión (tal como un agente de unión selectiva, receptor o ligando). Estos ensayos pueden usarse para examinar moléculas de prueba para determinar su capacidad para aumentar o disminuir la tasa y/o el grado de unión de un polipéptido B7-L a su pareja de unión. En un ensayo, se inmoviliza un polipéptido B7-L en los pocillos de una placa de microtitulación. Entonces pueden añadirse la pareja de unión de polipéptido B7-L radiomarcada (por ejemplo, pareja de unión de polipéptido B7-L yodada) y una molécula de prueba o bien por separado (en cualquier orden) o bien simultáneamente a los pocillos. Tras la incubación, pueden lavarse los pocillos y contarse para determinar la radioactividad, usando un contador de centelleo, para determinar el grado en el que se une la pareja de unión al polipéptido B7-L. Normalmente, se someterá a prueba una molécula a lo largo de un intervalo de concentraciones, y puede usarse una serie de pocillos de control que carecen de uno o más elementos de los ensayos de prueba para la precisión en la evaluación de los resultados. Una alternativa a este método implica invertir las “posiciones” de las proteínas, es decir, inmovilizar la pareja de unión de polipéptido B7-L a los pocillos de la placa de microtitulación, incubarla con la molécula de prueba y polipéptido B7-L radiomarcado, y determinar el grado de unión de polipéptido B7-L. Véase, por ejemplo, *Current Protocols in Molecular Biology*, chap. 18 (Ausubel *et al.*, eds., Green Publishers Inc. y Wiley and Sons 1995).

Como alternativa al radiomarcado, puede conjugarse un polipéptido B7-L o su pareja de unión con biotina, y entonces puede detectarse la presencia de proteína biotinilada usando estreptavidina unida a una enzima, tal como peroxidasa del rábano (HRP) o fosfatasa alcalina (AP), que puede detectarse colorimétricamente, o mediante marcado fluorescente de estreptavidina. También puede usarse un anticuerpo dirigido hacia un polipéptido B7-L o hacia una pareja de unión de polipéptido B7-L, y que está conjugado con biotina, para fines de detección tras la incubación del complejo con estreptavidina unida a enzima unida a AP o HRP.

También puede inmovilizarse un polipéptido B7-L o una pareja de unión de polipéptido B7-L mediante unión a perlas de agarosa, perlas acrílicas u otros tipos de tales sustratos de fase sólida inertes. El complejo sustrato-proteína puede colocarse en una disolución que contiene la proteína complementaria y el compuesto de prueba. Tras la incubación, pueden precipitarse las perlas mediante centrifugación, y puede evaluarse la cantidad de unión entre un

polipéptido B7-L y su pareja de unión usando los métodos descritos en el presente documento. Alternativamente, el complejo sustrato-proteína puede inmovilizarse en una columna pasando la molécula de prueba y la proteína complementaria a través de la columna. Entonces puede evaluarse la formación de un complejo entre un polipéptido B7-L y su pareja de unión usando cualquiera de las técnicas descritas en el presente documento (por ejemplo, radiomarcaje o unión a anticuerpo).

Otro ensayo *in vitro* que es útil para identificar una molécula de prueba que aumenta o disminuye la formación de un complejo entre una proteína de unión a polipéptido B7-L y una pareja de unión de polipéptido B7-L es un sistema detector por resonancia de plasmón superficial tal como el sistema de ensayo BIAcore (Pharmacia, Piscataway, NJ). El sistema BIAcore se utiliza según especifica el fabricante. Este ensayo implica esencialmente la unión covalente o bien de polipéptido B7-L o bien de una pareja de unión de polipéptido B7-L a un chip sensor recubierto con dextrano que se encuentra ubicado en un detector. Entonces pueden inyectarse el compuesto de prueba y la otra proteína complementaria, o bien simultáneamente o bien secuencialmente, en la cámara que contiene el chip sensor. Puede evaluarse la cantidad de proteína complementaria que se une basándose en el cambio de masa molecular que está físicamente asociada al lado recubierto con dextrano del chip sensor, midiéndose el cambio de masa molecular por el sistema detector.

En algunos casos, puede ser deseable evaluar dos o más compuestos de prueba juntos para determinar su capacidad para aumentar o disminuir la formación de un complejo entre un polipéptido B7-L y una pareja de unión de polipéptido B7-L. En estos casos, los ensayos expuestos en el presente documento pueden modificarse fácilmente añadiendo tal(es) compuesto(s) de prueba adicional(es) o bien simultáneamente con, o bien tras, el primer compuesto de prueba. El resto de las etapas en el ensayo son tal como se expone en el presente documento.

Pueden usarse ventajosamente ensayos *in vitro* tales como los descritos en el presente documento para examinar grandes números de compuestos para determinar un efecto sobre la formación de un complejo entre un polipéptido B7-L y una pareja de unión de polipéptido B7-L. Los ensayos pueden automatizarse para examinar compuestos generados en bibliotecas de presentación en fagos, de péptidos sintéticos y de síntesis química.

También pueden seleccionarse compuestos que aumentan o disminuyen la formación de un complejo entre un polipéptido B7-L y una pareja de unión de polipéptido B7-L en cultivo celular usando células y líneas celulares que expresan o bien polipéptido B7-L o bien pareja de unión de polipéptido B7-L. Pueden obtenerse células y líneas celulares de cualquier mamífero, pero preferiblemente serán de fuente humana o de otro primate, de canino o de roedor. La unión de un polipéptido B7-L a células que expresan pareja de unión de polipéptido B7-L en la superficie se evalúa en presencia o ausencia de moléculas de prueba, y el grado de unión puede determinarse, por ejemplo, mediante citometría de flujo usando un anticuerpo biotinilado frente a una pareja de unión de polipéptido B7-L. Pueden usarse ventajosamente ensayos en cultivo celular para evaluar adicionalmente compuestos con puntuación positiva en ensayos de unión a proteína descritos en el presente documento.

También pueden usarse cultivos celulares para examinar el impacto de un candidato farmacológico. Por ejemplo, los candidatos farmacológicos pueden disminuir o aumentar la expresión del gen B7-L. En determinadas realizaciones, la cantidad de polipéptido B7-L o un fragmento de polipéptido B7-L que se produce puede medirse tras la exposición del cultivo celular al candidato farmacológico. En determinadas realizaciones, puede detectarse el impacto real del candidato farmacológico sobre el cultivo celular. Por ejemplo, la sobreexpresión de un gen particular puede tener un impacto particular sobre el cultivo celular. En tales casos, puede someterse a prueba la capacidad de un candidato farmacológico para aumentar o disminuir la expresión del gen o su capacidad para prevenir o inhibir un impacto particular sobre el cultivo celular. En otros ejemplos, la producción de un producto metabólico particular tal como un fragmento de un polipéptido, puede dar como resultado, o estar asociada con, una enfermedad o estado patológico. En tales casos, puede someterse a prueba la capacidad de un candidato farmacológico para disminuir la producción de tal producto metabólico en un cultivo celular.

45 Internalización de proteínas

Puede usarse la secuencia de proteína *tat* (de VIH) para internalizar proteínas en una célula. Véase, por ejemplo, Falwell *et al.*, 1994, Proc. Natl. Acad. Sci. U. S. A. 91: 664-68. Por ejemplo, se ha descrito que una secuencia de 11 aminoácidos (Y-G-R-K-K-R-R-Q-R-R-R; SEQ ID NO: 24) de la proteína *tat* de VIH (denominada "dominio de transducción de proteína" o TAT PDT) media el suministro a través de la membrana citoplasmática y la membrana nuclear de una célula. Véase Schwarze *et al.*, 1999, Science 285: 1569-72; y Nagahara *et al.*, 1998, Nat. Med. 4: 1449-52. En estos procedimientos, se preparan constructos de FITC (G-G-G-G-Y-G-R-K-K-R-R-Q-R-R-R marcado con FITC; SEQ ID NO: 25), que penetran en tejidos tras la administración intraperitoneal, y se detecta la unión de tales constructos a células mediante análisis de citometría de flujo activada por fluorescencia (FACS). Las células tratadas con una proteína de fusión *tat*- β -gal mostrarán actividad de β -gal. Tras la inyección, puede detectarse la expresión de una construcción de este tipo en varios tejidos, incluyendo tejido hepático, renal, pulmonar, cardíaco y cerebral. Se cree que tales construcciones experimentan algún grado de despliegue con el fin de entrar en la célula, y como tal, pueden requerir un repliegue tras su entrada en la célula.

Por tanto, se apreciará que la secuencia de proteína *tat* puede usarse para internalizar un polipéptido deseado en una célula. Por ejemplo, usando la secuencia de proteína *tat*, puede administrarse por vía intracelular un antagonista de B7-L (tal como un agente de unión selectiva anti-B7-L, molécula pequeña, receptor soluble u oligonucleótido

antisentido) para inhibir la actividad de una molécula de B7-L. Tal como se usa en el presente documento, la expresión "molécula de B7-L" se refiere tanto a moléculas de ácido nucleico B7-L como a polipéptidos B7-L tal como se define en el presente documento. Cuando se desea, también puede administrarse la propia proteína B7-L internamente a una célula usando estos procedimientos. Véase también, Straus, 1999, Science 285: 1466-67.

5 Identificación de fuente de célula usando el polipéptido B7-L

Según determinadas realizaciones de la invención, puede ser útil poder determinar la fuente de un determinado tipo de célula asociado con un polipéptido B7-L. Por ejemplo, puede ser útil determinar el origen de una enfermedad o estado patológico como ayuda para seleccionar una terapia apropiada. En determinadas realizaciones, pueden usarse ácidos nucleicos que codifican un polipéptido B7-L como sonda para identificar células tal como se describe en el presente documento examinando los ácidos nucleicos de las células con una sonda de este tipo. En otras realizaciones, pueden usarse anticuerpos anti-polipéptido B7-L para someter a prueba para determinar a presencia de polipéptido B7-L en células, y por tanto, determinar si tales células son de los tipos descritos en el presente documento.

Composiciones de polipéptido B7-L y administración

15 Las composiciones terapéuticas están dentro del alcance de la presente invención. Tales composiciones farmacéuticas de polipéptido B7-L pueden comprender una cantidad terapéuticamente eficaz de un polipéptido B7-L o una molécula de ácido nucleico B7-L en mezcla con un agente de formulación farmacéutica o fisiológicamente aceptable seleccionado por su idoneidad con el modo de administración. Las composiciones farmacéuticas pueden comprender una cantidad terapéuticamente eficaz de uno o más agentes de unión selectiva de polipéptido B7-L en
20 mezcla con un agente de formulación farmacéutica o fisiológicamente aceptable seleccionado por su idoneidad con el modo de administración.

Los materiales de formulación aceptables son preferiblemente no tóxicos para los receptores a las dosificaciones y concentraciones empleadas.

25 La composición farmacéutica puede contener materiales de formulación para modificar, mantener o conservar, por ejemplo, el pH, osmolaridad, viscosidad, transparencia, color, isotonicidad, olor, esterilidad, estabilidad, tasa de disolución o liberación, adsorción, o penetración de la composición. Materiales de formulación adecuados incluyen, pero no se limitan a, aminoácidos (tales como glicina, glutamina, asparagina, arginina, o lisina), agentes antimicrobianos, antioxidantes (tales como ácido ascórbico, sulfito de sodio, o hidrogenosulfito de sodio), tampones (tales como borato, bicarbonato, Tris-HCl, citratos, fosfatos, u otros ácidos orgánicos), agentes de carga (tales como manitol o glicina), agentes quelantes (tales como ácido etilendiamina-tetraacético (EDTA)), agentes complejantes (tales como cafeína, polivinilpirrolidona, beta-ciclodextrina, o hidroxipropilbeta-ciclodextrina), productos de relleno, monosacáridos, disacáridos, y otros hidratos de carbono (tales como glucosa, manosa, o dextrinas), proteínas (tales como albúmina sérica, gelatina, o inmunoglobulinas), agentes colorantes, aromatizantes y diluyentes, agentes emulsionantes, polímeros hidrófilos (tales como polivinilpirrolidona), polipéptidos de bajo peso molecular, contraiones de formación de sales (tales como sodio), conservantes (tales como cloruro de benzalconio, ácido benzóico, ácido salicílico, timerosal, alcohol fenético, metilparabeno, propilparabeno, clorhexidina, ácido sórbico, o peróxido de hidrógeno), disolventes (tales como glicerina, propilenglicol, o polietilenglicol), alcoholes de azúcares (tales como manitol o sorbitol), agentes de suspensión, tensioactivos o agentes humectantes (tales como plurónicos; PEG; ésteres de sorbitano; polisorbato tales como polisorbato 20 o polisorbato 80; tritón; trometamina; lecitina; colesterol o tiloxapal), agentes potenciadores de la estabilidad (tales como sacarosa o sorbitol), agentes potenciadores de la
35 de tonicidad (tales como haluros de metal alcalino, preferiblemente cloruro de sodio o potasio, o manitol-sorbitol), vehículos de administración, excipientes y/o adyuvantes farmacéuticos. Véase Remington's Pharmaceutical Sciences (18ª Ed., A. R. Gennaro, ed., Mack Publishing Company 1990).

45 Un experto en la técnica determinará la composición farmacéutica óptima dependiendo, por ejemplo, de la vía de administración prevista, el formato de administración y la dosificación deseada. Véase, por ejemplo, Remington's Pharmaceutical Sciences, citado anteriormente. Tales composiciones pueden influir sobre el estado físico, la estabilidad, la tasa de liberación *in vivo*, y la tasa de aclaramiento *in vivo* de la molécula de B7-L.

El portador o vehículo primario en una composición farmacéutica puede ser de naturaleza o bien acuosa o bien no acuosa. Por ejemplo, un vehículo o portador adecuado para su inyección puede ser agua, solución salina fisiológica, o líquido cefalorraquídeo artificial, posiblemente complementado con otros materiales comunes en composiciones para la administración parenteral. La solución salina tamponada neutra o solución salina mezclada con albúmina sérica son vehículos adicionales a modo de ejemplo. Otras composiciones farmacéuticas a modo de ejemplo comprenden tampón Tris de aproximadamente pH 7,0-8,5, o tampón acetato de aproximadamente pH 4,0-5,5, que pueden incluir además sorbitol o un sustituto adecuado. En una realización de la presente invención, pueden prepararse composiciones de polipéptido B7-L para su almacenamiento mezclando la composición seleccionada que tiene el grado deseado de pureza con agentes de formulación opcionales (Remington's Pharmaceutical Sciences, citado anteriormente) en forma de una torta liofilizada o una disolución acuosa. Además, el producto de polipéptido B7-L puede formularse como liofilizado usando excipientes apropiados tales como sacarosa.

Las composiciones farmacéuticas de polipéptido B7-L pueden seleccionarse para su administración parenteral.

Alternativamente, las composiciones pueden seleccionarse para su inhalación o para su administración a través del tubo digestivo, tal como por vía oral. La preparación de tales composiciones farmacéuticamente aceptables está dentro de la habilidad del experto en la técnica.

5 Los componentes de formulación están presentes en concentraciones que son aceptables para el sitio de administración. Por ejemplo, los tampones se usan para mantener la composición a pH fisiológico o a un pH ligeramente inferior, normalmente dentro de un intervalo de pH de desde aproximadamente 5 hasta aproximadamente 8.

10 Cuando se contempla la administración parenteral, las composiciones terapéuticas para su uso en esta invención pueden estar en forma de una disolución acuosa libre de pirógenos, parenteralmente aceptable, que comprende la molécula de B7-L deseada en un vehículo farmacéuticamente aceptable. Un vehículo particularmente adecuado para la inyección parenteral es agua destilada estéril en la que se formula una molécula de B7-L como una disolución isotónica estéril, apropiadamente conservada. Aún otra preparación puede implicar la formulación de la molécula deseada con un agente, tal como microesferas inyectables, partículas bioerosionables, compuestos poliméricos (tales como poli(ácido láctico) o poli(ácido glicólico)), perlas, o liposomas, que proporciona la liberación controlada o sostenida del producto que entonces puede administrarse mediante una inyección de depósito. También puede usarse ácido hialurónico, y esto puede tener el efecto de fomentar la duración sostenida en la circulación. Otros medios adecuados para la introducción de la molécula deseada incluyen dispositivos de administración de fármacos implantables.

20 En una realización, puede formularse una composición farmacéutica para su inhalación. Por ejemplo, puede formularse polipéptido B7-L como un polvo seco para su inhalación. También pueden formularse disoluciones de inhalación de polipéptido B7-L o molécula de ácido nucleico con un propelente para la administración como aerosol. En aún otra realización, pueden nebulizarse disoluciones. La administración pulmonar se describe adicionalmente en la publicación internacional n.º WO 94/20069, que describe la administración pulmonar de proteínas químicamente modificadas.

25 También se contempla que determinadas formulaciones puedan administrarse por vía oral. En una realización de la presente invención, pueden formularse polipéptidos B7-L que se administran de esta manera con o sin los vehículos habitualmente usados en la preparación de formas farmacéuticas sólidas tales como comprimidos y cápsulas. Por ejemplo, puede diseñarse una cápsula para liberar la parte activa de la formulación en el punto en el tracto gastrointestinal en el que se maximiza la biodisponibilidad se minimiza la degradación presistémica. Pueden incluirse agentes adicionales para facilitar la absorción del polipéptido B7-L. También pueden emplearse diluyentes, aromatizantes, ceras de bajo punto de fusión, aceites vegetales, lubricantes, agentes de suspensión, agentes disgregantes de comprimidos, y aglutinantes.

30 Otra composición farmacéutica puede implicar una cantidad eficaz de polipéptidos B7-L en una mezcla con excipientes no tóxicos que son adecuados para la fabricación de comprimidos. Disolviendo los comprimidos en agua estéril, u otro vehículo apropiado, pueden prepararse disoluciones en forma de dosis unitaria. Los excipientes adecuados incluyen, pero no se limitan a, diluyentes inertes, tales como carbonato de calcio, carbonato o bicarbonato de sodio, lactosa, o fosfato de calcio; o agentes aglutinantes, tales como almidón, gelatina, o goma arábiga; o agentes lubricantes tales como estearato de magnesio, ácido esteárico, o talco.

40 Composiciones farmacéuticas de polipéptido B7-L adicionales resultarán evidentes para los expertos en la técnica, incluyendo formulaciones que implican polipéptidos B7-L en formulaciones de administración sostenida o controlada. Los expertos en la técnica también conocen técnicas para formular una variedad de otros medios de administración sostenida o controlada, tales como vehículos liposómicos, micropartículas bioerosionables o perlas porosas e inyecciones de depósito. Véase, por ejemplo, la solicitud internacional n.º PCT/US93/00829, que describe la liberación controlada de micropartículas porosas para la administración de composiciones farmacéuticas.

45 Ejemplos adicionales de preparaciones de liberación sostenida incluyen matrices poliméricas semipermeable en forma de artículos conformados, por ejemplo películas o microcápsulas. Las matrices de liberación sostenida pueden incluir poliésteres, hidrogeles, polilactidas (patente estadounidense n.º 3.773.919 y patente europea n.º 058481), copolímeros de ácido L-glutámico y L-glutamato de gamma-etilo (Sidman *et al.*, 1983, Biopolymers 22: 547-56), poli(metacrilato de 2-hidroxietilo) (Langer *et al.*, 1981, J. Biomed. Mater. Res. 15: 167-277 y Langer, 1982, Chem. Tech. 12: 98-105), etileno-acetato de vinilo (Langer *et al.*, citado anteriormente) o poli(ácido D(-)-3-hidroxibutírico) (patente europea n.º 133988). Las composiciones de liberación sostenida también pueden incluir liposomas, que pueden prepararse mediante cualquiera de varios métodos conocidos en la técnica. Véase, por ejemplo, Eppstein *et al.*, 1985, Proc. Natl. Acad. Sci. USA 82: 3688-92; y las patentes europeas n.ºs 036676, 088046 y 143949.

55 Normalmente, la composición farmacéutica de B7-L que va a usarse para la administración *in vivo* debe ser estéril. Esto puede lograrse mediante filtración a través de membranas de filtración estériles. Cuando se liofiliza la composición, la esterilización usando este método puede realizarse o bien antes, o bien después, de la liofilización y reconstitución. La composición para su administración parenteral puede almacenarse en forma liofilizada o en una disolución. Además, las composiciones parenterales se colocan generalmente en un recipiente que tiene un orificio de acceso estéril, por ejemplo, una bolsa de disolución intravenosa o vial que tiene un tapón que puede perforarse por una aguja de inyección hipodérmica.

60

Una vez formulada la composición farmacéutica, puede almacenarse en viales estériles como disolución, suspensión, gel, emulsión, sólido, o como polvo deshidratado o liofilizado. Tales formulaciones pueden almacenarse o bien en una forma lista para su uso o bien en una forma (por ejemplo, liofilizada) que requiere su reconstitución antes de la administración.

5 La cantidad eficaz de una composición farmacéutica de B7-L que debe emplearse terapéuticamente dependerá, por ejemplo, del contexto terapéutico y de los objetivos. Un experto en la técnica apreciará que los niveles de dosificación apropiados para el tratamiento variarán por tanto dependiendo, en parte, de la molécula administrada, la indicación para la que está usándose la molécula de B7-L, la vía de administración, y el tamaño (peso corporal, superficie corporal o tamaño de órgano) y estado (la edad y salud general) del paciente. Por consiguiente, el médico
10 puede titular la dosificación y modificar la vía de administración para obtener el efecto terapéutico óptimo. Una dosificación típica puede oscilar desde aproximadamente 0,1 µg/kg hasta aproximadamente 100 mg/kg o más, dependiendo de los factores mencionados anteriormente. En otras realizaciones, la dosificación puede oscilar desde 0,1 µg/kg hasta aproximadamente 100 mg/kg; o desde 1 µg/kg hasta aproximadamente 100 mg/kg; o desde 5 µg/kg hasta aproximadamente 100 mg/kg.

15 La frecuencia de dosificación dependerá de los parámetros farmacocinéticos de la molécula de B7-L en la formulación que está usándose. Normalmente, un médico administrará la composición hasta que se alcance una dosificación que logra el efecto deseado. Por tanto la composición puede administrarse como una única dosis, como dos o más dosis (que pueden contener o no la misma cantidad de la molécula deseada) a lo largo del tiempo, o como una infusión continua mediante un dispositivo de implantación o catéter. Los expertos en la técnica realizan de
20 manera rutinaria un refinamiento adicional de la dosificación apropiada y está dentro del ámbito de las tareas que realizan de manera rutinaria. Las dosificaciones apropiadas pueden determinarse mediante el uso de datos de respuesta a la dosis apropiados.

La vía de administración de la composición farmacéutica es de acuerdo con métodos conocidos, por ejemplo, por vía oral; mediante inyección por vías intravenosa, intraperitoneal, intracerebral (intraparenquimal),
25 intracerebroventricular, intramuscular, intraocular, intraarterial, intraporta, o intralesional; mediante sistemas de liberación sostenida; o mediante dispositivos de implantación. Cuando se desea, las composiciones pueden administrarse mediante inyección en bolo o de manera continua mediante infusión, o mediante dispositivo de implantación.

De manera alternativa o adicional, la composición puede administrarse de manera local por medio de implantación
30 de una membrana, esponja u otro material apropiado sobre el que se ha absorbido o encapsulado la molécula deseada. Cuando se usa un dispositivo de implantación, el dispositivo puede implantarse en cualquier tejido u órgano adecuado, y la administración de la molécula deseada puede realizarse mediante difusión, bolo de liberación programada, o administración continua.

En algunos casos, puede ser deseable usar composiciones farmacéuticas de polipéptido B7-L de una manera *ex vivo*. En tales casos, se exponen células, tejidos u órganos que se han extraído del paciente a composiciones
35 farmacéuticas de polipéptido B7-L tras lo cual vuelven a implantarse posteriormente las células, tejidos u órganos en el paciente.

En otros casos, un polipéptido B7-L puede administrarse implantando determinadas células que se han modificado genéticamente, usando métodos tales como los descritos en el presente documento, para expresar y secretar el
40 polipéptido B7-L. Tales células pueden ser células de animal o humanas, y pueden ser autólogas, heterólogas o exógenas. Opcionalmente, las células pueden inmortalizarse. Con el fin de reducir la posibilidad de una respuesta inmunológica, pueden encapsularse las células para evitar la infiltración de tejidos circundantes. Los materiales de encapsulación son normalmente recintos biocompatibles, semipermeables, poliméricos o membranas que permiten la liberación del/de los producto(s) de proteína pero impiden la destrucción de las células por el sistema inmunitario
45 del paciente o por otros factores perjudiciales de los tejidos circundantes.

Tal como se comenta en el presente documento, puede ser deseable tratar poblaciones celulares aisladas (tales como células madre, linfocitos, glóbulos rojos, condrocitos, neuronas, y similares) con uno o más polipéptidos B7-L. Esto puede lograrse exponiendo directamente las células aisladas al polipéptido, cuando está en una forma que es permeable a la membrana celular.

50 También pueden contemplarse células y métodos (por ejemplo, recombinación homóloga y/u otros métodos de producción recombinante) tanto para la producción *in vitro* de polipéptidos terapéuticos como para la producción y administración de polipéptidos terapéuticos mediante terapia génica o terapia celular. Pueden usarse métodos de recombinación homóloga y otros para modificar una célula que contiene un gen B7-L normalmente silencioso para la transcripción, o un gen subexpresado, y así producir una célula que expresa cantidades terapéuticamente eficaces
55 de polipéptidos B7-L.

La recombinación homóloga es una técnica desarrollada originalmente para seleccionar como diana genes para inducir o corregir mutaciones en genes transcripcionalmente activos. Kucherlapati, 1989, Prog. in Nucl. Acid Res. & Mol. Biol. 36: 301. La técnica básica se desarrolló como un método para introducir mutaciones específicas en regiones específicas del genoma de mamíferos (Thomas *et al.*, 1986, Cell 44: 419-28; Thomas y Capecchi, 1987,

Cell 51: 503-12; Doetschman *et al.*, 1988, Proc. Natl. Acad. Sci. U. S. A. 85: 8583- 87) o para corregir mutaciones específicas dentro de genes defectuosos (Doetschman *et al.*, 1987, Nature 330: 576-78). Técnicas de recombinación homóloga a modo de ejemplo se describen en la patente estadounidense n.º 5.272.071; patentes europeas n.ºs 9193051 y 505500; solicitud internacional n.º PCT/US90/07642, y publicación internacional n.º WO 91/09955.

5 Mediante recombinación homóloga, puede dirigirse la secuencia de ADN que va a insertarse en el genoma a una región específica del gen de interés uniéndola a ADN de direccionamiento. El ADN de direccionamiento es una secuencia de nucleótidos que es complementaria (homóloga) a una región del ADN genómico. Se ponen pequeñas partes de ADN de direccionamiento que son complementarias a una región específica del genoma en contacto con la cadena original durante el proceso de replicación de ADN. Una propiedad general del ADN que se ha insertado en una célula es hibridarse, y por tanto recombinarse con otras partes de ADN endógeno mediante regiones homólogas compartidas. Si esta cadena complementaria se une a un oligonucleótido que contiene una mutación o una secuencia diferente o un nucleótido adicional, éste también se incorpora en la cadena recién sintetizada como resultado de la recombinación. Como resultado de la función de corrección de lectura, la nueva secuencia de ADN puede servir como molde. Por tanto, el ADN transferido se incorpora en el genoma.

10 Unidas a estas partes de ADN de direccionamiento hay regiones de ADN que pueden interaccionar con, o controlar, la expresión de un polipéptido B7-L, por ejemplo, secuencias flanqueantes. Por ejemplo, se inserta un elemento promotor/potenciador, un supresor, o un elemento modulador de la transcripción exógeno en el genoma de la célula huésped prevista en la proximidad y en orientación suficiente para influir sobre la transcripción de ADN que codifica el polipéptido B7-L deseado. El elemento de control controla una parte del ADN presente en el genoma de la célula huésped. Por tanto, puede lograrse la expresión del polipéptido B7-L deseado no mediante transfección del ADN que codifica el propio gen B7-L, sino más bien mediante el uso de ADN de direccionamiento (que contiene regiones de homología con el gen endógeno de interés) acoplado con segmentos reguladores de ADN que proporcionan la secuencia génica endógena con señales reconocibles para la transcripción de un gen B7-L.

15 En un método a modo de ejemplo, la expresión de un gen seleccionado deseado en una célula (es decir, un gen celular endógeno deseado) se altera mediante recombinación homóloga en el genoma celular en un sitio preseleccionado, mediante la introducción de ADN que incluye al menos una secuencia reguladora, un exón, y un sitio donador de corte y empalme. Estos componentes se introducen en el ADN cromosómico (genómico) de tal manera que esto da como resultado, en efecto, la producción de una nueva unidad de transcripción (en la que la secuencia reguladora, el exón, y el sitio donador de corte y empalme presentes en el constructo de ADN están operativamente unidos al gen endógeno). Como resultado de la introducción de estos componentes en el ADN cromosómico, se altera la expresión del gen endógeno deseado.

20 La expresión génica alterada, tal como se describe en el presente documento, abarca activar (o provocar que se exprese) un gen que normalmente es silencioso (no se expresa) en la célula según se obtiene, así como aumentar la expresión de un gen que no se expresa a niveles fisiológicamente significativos en la célula según se obtiene. Las realizaciones abarcan además cambiar el patrón de regulación o inducción de tal manera que sea diferente del patrón de regulación o inducción que se produce en la célula según se obtiene, y reducir (incluyendo eliminar) la expresión de un gen que se expresa en la célula según se obtiene.

25 Un método mediante el cual puede usarse la recombinación homóloga para aumentar, o provocar, la producción de polipéptido B7-L a partir del gen B7-L endógeno de una célula implica usar en primer lugar la recombinación homóloga para colocar una secuencia de recombinación de un sistema de recombinación específico del sitio (por ejemplo, Cre/loxP, FLP/FRT) (Sauer, 1994, Curr. Opin. Biotechnol., 5: 521-27; Sauer, 1993, Methods Enzymol., 225: 890-900) aguas arriba (es decir, en sentido 5') de la región codificante de polipéptido B7-L genómica endógena de la célula. Se introduce un plásmido que contiene un sitio de recombinación homólogo al sitio que se colocó justo en aguas arriba de la región codificante de polipéptido B7-L genómica en la línea celular modificada con la enzima recombinasa apropiada. Esta recombinasa provoca que el plásmido se integre, mediante el sitio de recombinación del plásmido, en el sitio de recombinación ubicado justo aguas arriba de la región codificante de polipéptido B7-L genómica en la línea celular (Baubonis y Sauer, 1993, Nucleic Acids Res. 21: 2025-29; O'Gorman *et al.*, 1991, Science 251: 1351-55). Cualquier secuencia flanqueante que se sabe que aumenta la transcripción (por ejemplo, potenciador/promotor, intron, potenciador de traslación), si se coloca apropiadamente en este plásmido, se integrará de una manera tal como para crear una unidad de transcripción nueva o modificada que da como resultado la producción de polipéptido B7-L *de novo* o aumentada a partir del gen B7-L endógeno de la célula.

30 Un método adicional para usar la línea celular en la que se había colocado la secuencia de recombinación específica del sitio justo aguas arriba de la región codificante de polipéptido B7-L genómica endógena de la célula es usar la recombinación homóloga para introducir un segundo sitio de recombinación en otra parte en el genoma de la línea celular. Entonces se introduce la enzima recombinasa apropiada en la línea celular con dos sitios de recombinación, provocando un acontecimiento de recombinación (deleción, inversión, y traslocación) (Sauer, 1994, Curr. Opin. Biotechnol., 5: 521-27; Sauer, 1993, Methods Enzymol., 225: 890-900) que creará una unidad de transcripción nueva o modificada dando como resultado la producción de polipéptido B7-L *de novo* o aumentada a partir del gen B7-L endógeno de la célula.

35 Un enfoque adicional para aumentar, o provocar, la expresión de polipéptido B7-L a partir de un gen B7-L endógeno de una célula implica aumentar, o provocar, la expresión de un gen o genes (por ejemplo, factores de transcripción)

y/o disminuir la expresión de un gen o genes (por ejemplo, represores transcripcionales) de manera que dé como resultado la producción *de novo* o aumentada del polipéptido B7-L a partir del gen B7-L endógeno de la célula. Este método incluye la introducción de un polipéptido que se produce de manera no natural (por ejemplo, un polipéptido que comprende un dominio de unión a ADN específico del sitio fusionado con un dominio de factor de transcripción) en la célula de manera que resulta la producción *de novo* o aumentada del polipéptido B7-L a partir del gen B7-L endógeno de la célula.

Si se conoce la secuencia de un gen particular, tal como la secuencia de ácido nucleico del polipéptido B7-L presentada en el presente documento, puede sintetizarse un trozo de ADN que es complementario a una región seleccionada del gen u obtenerse de otro modo, tal como mediante la restricción apropiada del ADN nativo en sitios de reconocimiento específicos que delimitan la región de interés. Este trozo sirve como una secuencia de direccionamiento tras la inserción en la célula y se hibridará con su región homóloga dentro del genoma. Si esta hibridación se produce durante la replicación del ADN, este trozo de ADN, y cualquier secuencia adicional unida al mismo, actuarán como un fragmento de Okazaki y se incorporará en la cadena hija de ADN recién sintetizada. La presente invención, por tanto, incluye nucleótidos que codifican un polipéptido B7-L, nucleótidos que pueden usarse como secuencias de direccionamiento.

También se contempla la terapia celular con polipéptido B7-L, por ejemplo, la implantación de células que producen polipéptidos B7-L. Esta realización implica implantar células que pueden sintetizar y secretar una forma biológicamente activa del polipéptido B7-L. Tales células que producen el polipéptido B7-L pueden ser células que son productores naturales de polipéptidos B7-L o pueden ser células recombinantes cuya capacidad para producir polipéptidos B7-L se ha aumentado mediante la transformación con un gen que codifica el polipéptido B7-L deseado o con un gen que aumenta la expresión de polipéptido B7-L. Puede lograrse una modificación de este tipo por medio de un vector adecuado para suministrar el gen así como para fomentar su expresión y secreción. Para minimizar una posible reacción inmunológica en pacientes a los que se les administra un polipéptido B7-L, tal como puede producirse con la administración de un polipéptido de una especie foránea, se prefiere que las células naturales que producen el polipéptido B7-L sean de origen humano y produzcan polipéptido B7-L humano. Asimismo, se prefiere que las células recombinantes que producen el polipéptido B7-L se transformen con un vector de expresión que contiene un gen que codifica un polipéptido B7-L humano.

Pueden encapsularse las células implantadas para evitar la infiltración del tejido circundante. Pueden implantarse células animales humanas o no humanas en recintos o membranas poliméricos biocompatibles, semipermeables que permiten la liberación del polipéptido B7-L, pero que impiden la destrucción de las células por parte del sistema inmunitario del paciente o por otros factores perjudiciales del tejido circundante. Alternativamente, pueden implantarse las propias células del paciente, transformadas para producir polipéptidos B7-L *ex vivo*, directamente en el paciente sin tal encapsulación.

Se conocen en la técnica, técnicas para la encapsulación de células vivas, y puede realizarse de manera rutinaria la preparación de las células encapsuladas y su implantación en pacientes. Por ejemplo, Baetge *et al.* (publicación internacional n.º WO 95/05452 y solicitud internacional n.º PCT/US94/09299) describen cápsulas de membranas que contienen células genéticamente modificadas para el suministro eficaz de moléculas biológicamente activas. Las cápsulas son biocompatibles y pueden recuperarse fácilmente. Las cápsulas encapsulan células transfectadas con moléculas de ADN recombinante que comprenden secuencias de ADN que codifican moléculas biológicamente activas unidas operativamente a promotores que no son objeto de regulación por disminución *in vivo* tras la implantación en un huésped mamífero. El dispositivo prevé el suministro de las moléculas de células vivas a sitios específicos dentro de un receptor. Además, véanse las patentes estadounidenses n.ºs 4.892.538; 5.011.472; y 5.106.627. Se describe un sistema para encapsular células vivas en la publicación internacional n.º WO 91/10425 (Aebischer *et al.*). Véanse también, la publicación internacional n.º WO 91/10470 (Aebischer *et al.*); Winn *et al.*, 1991, Exper. Neurol. 113: 322-29; Aebischer *et al.*, 1991, Exper. Neurol. 111: 269-75; y Tresco *et al.*, 1992, ASAI0 38: 17-23.

También se prevé la administración mediante terapia génica *in vivo* e *in vitro* de polipéptidos B7-L. Un ejemplo de una técnica de terapia génica es usar un gen B7-L (o bien ADN genómico, ADNc y/o ADN sintético) que codifica un polipéptido B7-L que puede unirse operativamente a un promotor constitutivo o inducible para formar un "constructo de ADN para terapia génica". El promotor puede ser homólogo o heterólogo con respecto al gen B7-L endógeno, siempre que sea activo en la célula o tipo tisular en el que se insertará el constructo. Otros componentes del constructo de ADN para terapia génica pueden incluir opcionalmente moléculas de ADN diseñadas para la integración específica del sitio (por ejemplo, secuencias endógenas útiles para la recombinación homóloga), promotores, potenciadores o silenciadores específicos de tejidos, moléculas de ADN que pueden proporcionar una ventaja selectiva con respecto a la célula original, moléculas de ADN útiles como marcadores para identificar células transformadas, sistemas de selección negativa, agentes de unión específicos de la célula (como, por ejemplo, para el direccionamiento celular), factores de internalización específicos de la célula, factores de transcripción que potencian la expresión desde un vector, y factores que permiten la producción de vectores.

Luego pueden introducirse una construcción de ADN para terapia génica en células (o bien *ex vivo* o bien *in vivo*) usando vectores virales o no virales. Un medio para introducir el constructo de ADN para terapia génica es por medio de vectores virales tal como se describe en el presente documento. Ciertos vectores, tales como vectores

retrovirales, suministrarán la construcción de ADN al ADN cromosómico de las células, y el gen puede integrarse en el ADN cromosómico. Otros vectores funcionarán como episomas, y el constructo de ADN para terapia génica permanecerá en el citoplasma.

Pueden incluirse elementos reguladores para la expresión controlada del gen B7-L en la célula diana. Tales elementos se activan en respuesta a un efector apropiado. De esta manera, puede expresarse un polipéptido terapéutico cuando se desee. Un medio de control convencional implica el uso de dimerizadores de molécula pequeña o rapálogos para dimerizar proteínas quiméricas que contienen un dominio de unión a molécula pequeña y un dominio que puede iniciar un proceso biológico, tal como una proteína de unión a ADN o proteínas de activación transcripcional (véanse las publicaciones internacionales n.ºs WO 96/41865, WO 97/31898 y WO 97/31899). Puede usarse la dimerización de las proteínas para iniciar la transcripción del transgén.

Una tecnología de regulación alternativa usa un método de almacenamiento de proteínas expresadas desde el gen de interés dentro de la célula como un agregado o agrupamiento (*cluster*). El gen de interés se expresa como una proteína de fusión que incluye un dominio de agregación condicional que da como resultado la retención de la proteína agregada en el retículo endoplasmático. Las proteínas almacenadas son estables e inactivas dentro de la célula. Las proteínas pueden liberarse, sin embargo, administrando un fármaco (por ejemplo, ligando de molécula pequeña) que elimina el dominio de agregación condicional y separa de ese modo específicamente los agregados a agrupamientos de modo que las proteínas pueden secretarse desde la célula. Véanse Aridor et al., 2000, Science 287: 816-17 y Rivera et al., 2000, Science 287: 826-30.

Otros medios de control o interruptores génicos adecuados incluyen, pero no se limitan a, los sistemas descritos en el presente documento. Se usa mifepristona (RU486) como antagonista de la progesterona. La unión de un dominio de unión a ligando del receptor de progesterona modificado al antagonista de la progesterona activa la transcripción formando un dímero de los dos factores de transcripción que luego pasa al núcleo para unirse al ADN. El dominio de unión a ligando se modifica para eliminar la habilidad del receptor para unirse al ligando natural. El sistema de receptores de hormonas esteroideas modificado se describe adicionalmente en la patente estadounidense n.º 5.364.791 y las publicaciones internacionales n.ºs WO 96/40911 y WO 97/10337.

Aún otro sistema de control usa ecdisona (una hormona esteroidea de la mosca de la fruta), que se une a y activa un receptor de ecdisona (receptor citoplasmático). El receptor se transloca luego al núcleo para unirse a un elemento de respuesta al ADN específico (promotor del gen sensible a ecdisona). El receptor de ecdisona incluye un dominio de transactivación, dominio de unión a ADN y dominio de unión a ligando para iniciar la transcripción. El sistema de ecdisona se describe adicionalmente en la patente estadounidense n.º 5.514.578 y las publicaciones internacionales n.ºs WO 97/38117, WO 96/37609 y WO 93/03162.

Otro medio de control usa un transactivador positivo controlable mediante tetraciclina. Este sistema implica un dominio de unión a ADN de proteína represora tet mutada (cambios de aminoácidos R-4 de tet mutada que dieron como resultado una proteína transactivadora regulada por tetraciclina inversa, es decir, se une a un operador tet en presencia de tetraciclina) unida a un polipéptido que activa la transcripción. Tales sistemas se describen en las patentes estadounidenses n.ºs 5.464.758, 5.650.298 y 5.654.168.

Se describen sistemas de control de la expresión y construcciones de ácido nucleico adicionales en las patentes estadounidenses n.ºs 5.741.679 y 5.834.186, concedidas a Innovir Laboratories Inc.

Puede realizarse la terapia génica *in vivo* introduciendo el gen que codifica el polipéptido B7-L en células mediante inyección local de una molécula de ácido nucleico B7-L o mediante otros vectores de suministro virales o no virales apropiados. Hefti 1994, Neurobiology 25: 1418-35. Por ejemplo, una molécula de ácido nucleico que codifica un polipéptido B7-L puede estar contenida en un vector de virus adenoasociado (VAA) para el suministro a las células seleccionadas como diana (véanse, por ejemplo, Johnson, publicación internacional n.º WO 95/34670; solicitud internacional n.º PCT/US95/07178). El genoma de VAA recombinante normalmente contiene repeticiones terminales invertidas de VAA que flanquean una secuencia de ADN que codifica un polipéptido B7-L operativamente unido a un promotor funcional y secuencias de poliadenilación.

Los vectores virales adecuados alternativos incluyen, pero no se limitan a, vectores de retrovirus, adenovirus, virus del herpes simple, lentivirus, virus de la hepatitis, parvovirus, papovavirus, poxvirus, alfavirus, coronavirus, rabdovirus, paramixovirus y virus del papiloma. La patente estadounidense n.º 5.672.344 describe un sistema de transferencia génica mediada por virus *in vivo* que implica un vector de VHS-1 neurotrófico recombinante. La patente estadounidense n.º 5.399.346 proporciona ejemplos de un proceso para proporcionar a un paciente una proteína terapéutica mediante el suministro de células humanas que se han tratado *in vitro* para insertar un segmento de ADN que codifica una proteína terapéutica. Se describen métodos y materiales adicionales para la práctica de técnicas de terapia génica en las patentes estadounidenses n.ºs 5.631.236 (que implican vectores adenovirales), 5.672.510 (que implican vectores retrovirales), 5.635.399 (que implican vectores retrovirales que expresan citocinas).

Los métodos de suministro no viral incluyen, pero no se limitan a, transferencia mediada por liposomas, suministro de ADN desnudo (inyección directa), transferencia mediada por receptor (complejo de ADN-ligando), electroporación, precipitación con fosfato de calcio y bombardeo con micropartículas (por ejemplo, pistola génica). Los materiales y métodos para terapia génica también pueden incluir promotores inducibles, potenciadores-

promotores específicos de tejido, secuencias de ADN diseñadas para la integración específica del sitio, secuencias de ADN que pueden proporcionar una ventaja selectiva con respecto a la célula original, marcadores para identificar células transformadas, sistemas de selección negativa y sistemas de control de la expresión (medidas de seguridad), agentes de unión específicos de la célula (para el direccionamiento celular), factores de internalización específicos de la célula y factores de transcripción para potenciar la expresión mediante un vector así como métodos para la producción de vectores. Tales métodos y materiales adicionales para la práctica de técnicas de terapia génica se describen en las patentes estadounidenses n.ºs 4.970.154 (que implica técnicas de electroporación), 5.679.559 (que describe un sistema que contiene lipoproteína para inserción génica), 5.676.954 (que implica vehículos de liposomas), 5.593.875 (que describe métodos para la transfección con fosfato de calcio), y 4.945.050 (que describe un procedimiento en el que se impulsan partículas biológicamente activas en células a una velocidad mediante la cual las partículas penetran en la superficie de las células y se incorporan al interior de las células), y la publicación internacional n.º WO 96/40958 (que implica ligandos nucleares).

También se contempla que la terapia con el gen B7-L o la terapia celular puedan incluir el suministro de uno o más polipéptido(s) adicional(es) en la(s) misma(s) o diferente(s) célula(s). Tales células pueden introducirse por separado en el paciente, o las células pueden estar contenidas en un único dispositivo implantable, tal como la membrana de encapsulación descrita anteriormente, o las células pueden modificarse por separado por medio de vectores virales.

Un medio para aumentar la expresión endógena del polipéptido B7-L en una célula mediante terapia génica es insertar uno o más elementos potenciadores en el promotor del polipéptido B7-L, en el que los elementos potenciadores pueden servir para aumentar la actividad transcripcional del gen B7-L. Los elementos potenciadores usados se seleccionarán basándose en el tejido en el que se desea activar el gen, se seleccionarán los elementos potenciadores que se sabe que confieren activación del promotor en ese tejido. Por ejemplo, si un gen que codifica un polipéptido B7-L va a "activarse" en células T, puede usarse el elemento potenciador del promotor *Ick*. En este caso, la parte funcional del elemento transcripcional que va a añadirse puede insertarse en un fragmento de ADN que contiene el promotor del polipéptido B7-L (y opcionalmente, insertarse en un vector y/o secuencias flanqueantes en 5' y/o 3') usando técnicas de clonación convencionales. Este constructo, conocido como un "constructo de recombinación homóloga" puede introducirse luego en las células deseadas o bien *ex vivo* o bien *in vivo*.

También puede usarse la terapia génica para disminuir la expresión del polipéptido B7-L modificando la secuencia de nucleótidos del promotor endógeno. Tal modificación se logra normalmente mediante métodos de recombinación homóloga. Por ejemplo, puede modificarse mediante ingeniería genética una molécula de ADN que contiene la totalidad o una parte del promotor del gen B7-L seleccionado para la inactivación, para eliminar y/o sustituir trozos del promotor que regulan la transcripción. Por ejemplo, puede delecionarse la caja TATA y/o el sitio de unión de un activador transcripcional del promotor usando técnicas de biología molecular convencionales; tal delección puede inhibir la actividad del promotor reprimiendo de ese modo la transcripción del correspondiente gen B7-L. La delección de la caja TATA o el sitio de unión del activador de la transcripción en el promotor puede lograrse generando una construcción de ADN que comprende la totalidad o la parte relevante del promotor del polipéptido B7-L (de la misma especie o una relacionada que el gen B7-L que va a regularse) en el que uno o más de los nucleótidos de la caja TATA y/o el sitio de unión del activador transcripcional se mutan mediante sustitución, delección y/o inserción de uno o más nucleótidos. Como resultado, la caja TATA y/o el sitio de unión del activador tiene una actividad disminuida o se vuelve completamente inactivo. Este constructo, que también contendrá normalmente al menos aproximadamente 500 bases de ADN que corresponden a las secuencias de ADN en 5' y 3' (endógenas) adyacentes al segmento promotor que se ha modificado, pueden introducirse en las células apropiadas (o bien *ex vivo* o bien *in vivo*) o bien directamente o bien mediante un vector viral tal como se describe en el presente documento. Normalmente, la integración del constructo en el ADN genómico de las células será mediante recombinación homóloga, en la que las secuencias de ADN en 5' y 3' en el constructo de promotor pueden servir para ayudar a integrar la región promotora modificada mediante hibridación con el ADN cromosómico endógeno.

Usos terapéuticos

Pueden usarse moléculas de ácido nucleico B7-L, polipéptidos, y agonistas y antagonistas de los mismos para tratar, diagnosticar, mejorar o prevenir varias enfermedades, trastornos o estados, incluyendo los citados en el presente documento.

Los agonistas y antagonistas del polipéptido B7-L incluyen aquellas moléculas que regulan la actividad del polipéptido B7-L y o bien aumentan o bien disminuyen al menos una actividad de la forma madura del polipéptido B7-L. Los agonistas o antagonistas pueden ser cofactores, tales como una proteína, un péptido, un hidrato de carbono, un lípido o una molécula de pequeño peso molecular, que interaccionan con el polipéptido B7-L y regulan de ese modo su actividad. Los posibles agonistas o antagonistas del polipéptido incluyen anticuerpos que reaccionan con formas o bien solubles o bien unidas a membrana de polipéptidos B7-L que comprenden parte o la totalidad de los dominios extracelulares de dichas proteínas. Las moléculas que regulan la expresión del polipéptido B7-L normalmente incluyen ácidos nucleicos que codifican polipéptido B7-L que pueden actuar como reguladores antisentido de la expresión.

Puesto que los ratones transgénicos que expresan un miembro relacionado de la familia B7 mostraron hiperplasia de la vesícula seminal (solicitud de patente estadounidense en tramitación junto con la presente n.º 09/729.264, presentada el 28 de noviembre de 2000), los agonistas y antagonistas del polipéptido B7-L pueden ser útiles en el

tratamiento de trastornos reproductivos y trastornos proliferativos.

La sobreexpresión del polipéptido B7-L puede desempeñar un papel en el crecimiento y el mantenimiento de células cancerígenas provocando hiperplasia de la vesícula seminal. Por consiguiente, los agonistas o antagonistas para el polipéptido B7-L pueden ser útiles para el diagnóstico o el tratamiento del cáncer. Los ejemplos de tales cánceres incluyen, pero no se limitan a, cáncer de vesícula seminal, cáncer de pulmón, cáncer cerebral, cáncer de mama, cánceres del sistema hematopoyético, cáncer de próstata, cáncer de ovario y cáncer de testículos. Otros cánceres están abarcados dentro del alcance de la invención.

La sobreexpresión del polipéptido B7-L puede desempeñar un papel en la proliferación inapropiada de células provocando hiperplasia de la vesícula seminal. El polipéptido B7-L puede desempeñar un papel en la proliferación inapropiada de células basada en la sobreexpresión que provoca hiperplasia de la vesícula seminal. Por consiguiente, los agonistas o antagonistas para el polipéptido B7-L pueden ser útiles para el diagnóstico o tratamiento de enfermedades asociadas con la proliferación celular anómala. Los ejemplos de tales enfermedades incluyen, pero no se limitan a, arteriosclerosis y reestenosis vascular. Otras enfermedades que se ven influenciadas por la proliferación inapropiada de células están abarcadas dentro del alcance de la invención.

La sobreexpresión del polipéptido B7-L puede desempeñar un papel en el sistema reproductor provocando hiperplasia de la vesícula seminal. Por consiguiente, los agonistas o antagonistas para el polipéptido B7-L pueden ser útiles para el diagnóstico o tratamiento de enfermedades asociadas con el sistema reproductor. Los ejemplos de tales enfermedades incluyen, pero no se limitan a, esterilidad, aborto espontáneo, parto prematuro y parto, y endometriosis. Otras enfermedades del sistema reproductor están abarcadas dentro del alcance de la invención.

Preferiblemente, las moléculas de ácido nucleico B7-L, polipéptidos, y agonistas y antagonistas de la presente invención se usan para tratar, diagnosticar, mejorar o prevenir enfermedades asociadas con la función de las células T (por ejemplo, funcionar como señuelo receptor de células T). Por ejemplo, pueden usarse anticuerpos, proteínas solubles que comprenden dominios extracelulares, u otros reguladores del polipéptido B7-L que dan como resultado una activación prolongada o potenciada de células T, para aumentar la respuesta inmunitaria a tumores.

Las moléculas de ácido nucleico B7-L, polipéptidos, y agonistas y antagonistas de la presente invención pueden usarse en el tratamiento de la enfermedad autoinmunitaria, supervivencia de injertos, activación de células inmunitarias para inhibir el crecimiento de células tumorales, enfermedades mediadas por células B dependientes de células T e inmunoterapia génica contra el cáncer. En una realización, los agonistas de la función del polipéptido B7-L, polipéptidos B7-L solubles, o derivados del polipéptido B7-L pueden ser beneficiosos para aliviar los síntomas en enfermedades con disfunción crónica de las células inmunitarias. Pueden tratarse enfermedades autoinmunitarias, tales como lupus eritematoso sistémico, artritis reumatoide, osteoartritis, púrpura trombocitopénica inmune (PTI) y psoriasis, con agonistas de la función del polipéptido B7-L, polipéptidos B7-L solubles, o derivados del polipéptido B7-L. Además, también pueden tratarse enfermedades inflamatorias crónicas, tales como enfermedad inflamatoria del intestino (enfermedad de Crohn y colitis ulcerosa), enfermedad de Grave, tiroiditis de Hashimoto y diabetes mellitus, con agonistas de la función del polipéptido B7-L, polipéptidos B7-L solubles, o derivados del polipéptido B7-L.

Pueden usarse agonistas de la función del polipéptido B7-L, polipéptidos B7-L solubles, o derivados del polipéptido B7-L como agentes inmunosupresores para el trasplante de órganos y médula ósea y pueden usarse para prolongar la supervivencia de injertos. Tales agonistas de la función del polipéptido B7-L, polipéptidos B7-L solubles, o derivados del polipéptido B7-L pueden proporcionar ventajas significativas con respecto a los tratamientos existentes. La terapia de trasplante de órganos y médula ósea debe competir con el rechazo mediado por células T de las células o el tejido foráneos por parte del huésped. Los regímenes terapéuticos actuales para inhibir el rechazo mediado por células T implican el tratamiento con los fármacos ciclosporina o FK506. Aunque los fármacos son eficaces, los pacientes experimentan efectos secundarios graves, incluyendo hepatotoxicidad, nefrotoxicidad y neurotoxicidad. La diana para la clase de ciclosporina/FK506 de agentes terapéuticos es la calcineurina, una fosfatasa con expresión ubicua. Los agonistas de la función del polipéptido B7-L, polipéptidos B7-L solubles, o derivados del polipéptido B7-L pueden carecer de los efectos secundarios graves observados con el uso de los presentes agentes inmunoterapéuticos. Pueden usarse los agonistas de la función del polipéptido B7-L, polipéptidos B7-L solubles, o derivados del polipéptido B7-L como agentes inmunosupresores para trastornos autoinmunitarios, tales como artritis reumatoide, osteoartritis psoriasis, esclerosis múltiple, diabetes y lupus eritematoso sistémico. También pueden usarse los agonistas de la función del polipéptido B7-L, polipéptidos B7-L solubles, o derivados del polipéptido B7-L para aliviar el síndrome de choque tóxico, enfermedad inflamatoria del intestino, alopecia, alosensibilización debido a transfusiones sanguíneas, enfermedades mediadas por células B dependientes de células T y el tratamiento de enfermedad la enfermedad de injerto contra huésped.

Por ejemplo, muchas vacunas actúan provocando una respuesta de anticuerpos eficaz y específica. Algunas vacunas, especialmente aquéllas contra microorganismos intestinales (por ejemplo, virus de la hepatitis A y *Salmonella*), provocan una respuesta de anticuerpos de corta vida. Es deseable potenciar y prolongar esta respuesta para aumentar la eficacia de la vacuna. Por tanto, los polipéptidos B7-L solubles pueden servir como adyuvantes de vacuna.

A la inversa, puesto que B7-L puede tener funciones reguladoras inmunitarias negativas, la inhibición de la actividad

de B7-L usando anticuerpos, moléculas pequeñas, peptidocuerpos, u otros antagonistas de la función de B7-L pueden dar como resultado potenciación inmunitaria y actividad antitumoral.

También pueden potenciarse respuestas antivirales mediante activadores o agonistas de la ruta del polipéptido B7-L. La potenciación de funciones inmunitarias células por los antagonistas del polipéptido B7-L también puede ser beneficiosa en la eliminación de células infectadas por virus. De modo complementario, los antagonistas del polipéptido B7-L también pueden tener efectos sobre funciones inmunitarias humorales que pueden potenciar respuestas mediadas por anticuerpos y que pueden funcionar para ayudar a eliminar virus del organismo.

A la inversa, existen varios estados clínicos que mejorarían mediante la inhibición de la producción de anticuerpos. La hipersensibilidad es una respuesta inmunitaria normalmente beneficiosa que es exagerada o inapropiada, y conduce a reacciones inflamatorias y daño tisular. Las reacciones de hipersensibilidad que están mediadas por anticuerpos pueden ser particularmente susceptibles al antagonismo por parte de agonistas de la función del polipéptido B7-L, polipéptidos B7-L solubles, o derivados del polipéptido B7-L. Las alergias, fiebre del heno, asma y edema agudo provocan reacciones de hipersensibilidad de tipo I, y estas reacciones pueden suprimirlas los agonistas de la función del polipéptido B7-L, polipéptidos B7-L solubles, o derivados del polipéptido B7-L.

Pueden tratarse enfermedades que provocan reacciones de hipersensibilidad mediadas por anticuerpos, incluyendo lupus eritematoso sistémico, artritis (artritis reumatoide, artritis reactiva y artritis psoriásica), nefropatías (glomerulonefritis, tubulopatías membranosa, mesangiocapilar, segmentaria y focal, necrotizante focal, extracapilar y proliferativa), trastornos cutáneos (pénfigo, penfigoide y eritema nodoso), endocrinopatías (tiroiditis, enfermedad de Grave, enfermedad de Hashimoto, diabetes mellitas insulino dependiente), diversas neumopatías (especialmente alveolitis extrínseca), diversas vasculopatías, enfermedad celíaca, con producción aberrante de IgA, muchas anemias y trombocitopenias, síndrome de Guillain-Barre y miastenia grave, con agonistas de la función del polipéptido B7-L, polipéptidos B7-L solubles, o derivados del polipéptido B7-L.

Además, pueden inhibirse trastornos linfoproliferativos, tales como mieloma múltiple, macroglobulinemia de Waldenstrom y crioglobulinemias, mediante agonistas de la función del polipéptido B7-L, polipéptidos B7-L solubles, o derivados del polipéptido B7-L. Finalmente, la enfermedad de injerto contra huésped, un trastorno inmunitario "artificial", puede beneficiarse de la inhibición de la producción de anticuerpos mediante agonistas de la función del polipéptido B7-L, polipéptidos B7-L solubles, o derivados del polipéptido B7-L.

Pueden usarse agonistas o antagonistas de la función del polipéptido B7-L (simultánea o secuencialmente) en combinación con una o más citocinas, factores de crecimiento, antibióticos, antiinflamatorios y/o agentes quimioterápicos según sea apropiado para el estado que esté tratándose.

Otras enfermedades provocadas por o mediadas por niveles no deseados de polipéptidos B7-L están abarcadas dentro del alcance de la invención. Los niveles no deseados incluyen niveles excesivos de polipéptidos B7-L niveles por debajo de lo normal de polipéptidos B7-L.

El polipéptido B7-L es un ligando para un regulador negativo de respuestas inmunitarias, PD-1 (Nishimura *et al.*, 1999, Immunity 11: 141-51). Por tanto, es probable que los agonistas de esta ruta del polipéptido B7-L inhiban las respuestas inmunitarias y los antagonistas de la ruta pueden potenciar la respuestas inmunitarias. Sin embargo, los agonistas o antagonistas de la función del polipéptido B7-L pueden producir resultados inesperados debido a factores biológicos desconocidos.

Usos de ácidos nucleicos y polipéptidos B7-L

Pueden usarse moléculas de ácido nucleico de la invención para mapear las ubicaciones del gen B7-L y genes relacionados en los cromosomas. Puede realizarse el mapeo mediante técnicas conocidas en la técnica, tales como amplificación por PCR e hibridación *in situ*.

Las moléculas de ácido nucleico B7-L (incluyendo aquéllas que no codifican por sí mismas polipéptidos biológicamente activos), pueden ser útiles como sondas de hibridación en ensayos de diagnóstico, o bien de manera cualitativa o bien de manera cuantitativa, para determinar la presencia de una molécula de ácido nucleico B7-L en muestras de líquido corporal o tejido de mamífero.

También pueden emplearse otros métodos en los que es deseable inhibir la actividad de uno o más polipéptidos B7-L. Puede efectuarse tal inhibición mediante moléculas de ácido nucleico que son complementarias a y se hibridan con secuencias de control de la expresión (formación de triple hélice) o con ARNm de B7-L. Por ejemplo, pueden introducirse moléculas de ADN o ARN antisentido, que tienen una secuencia que es complementaria a al menos una parte de un gen B7-L, en la célula. Pueden diseñarse sondas antisentido mediante técnicas disponibles usando la secuencia del gen B7-L dada a conocer en el presente documento. Normalmente, cada una de tales moléculas antisentido será complementaria al sitio de inicio (extremo 5') de cada gen B7-L seleccionado. Cuando la molécula antisentido se hibrida entonces con el ARNm de B7-L correspondiente, se impide o reduce la traducción de este ARNm. Los inhibidores antisentido proporcionan información relativa a la disminución o ausencia de un polipéptido B7-L en una célula u organismo.

Alternativamente, puede emplearse la terapia génica para crear un inhibidor dominante negativo de uno o más

polipéptidos B7-L. En esta situación, puede prepararse el ADN que codifica un polipéptido mutante de cada polipéptido B7-L seleccionado e introducirse en las células de un paciente usando métodos o bien virales o bien no virales tal como se describe en el presente documento. Cada mutante de este tipo está diseñado normalmente para competir con el polipéptido endógeno en su papel biológico.

Además, un polipéptido B7-L, ya sea biológicamente activo o no, puede usarse como inmunógeno, es decir, el polipéptido contiene al menos un epítipo frente al que pueden producirse anticuerpos. Pueden usarse agentes de unión selectiva que se unen a un polipéptido B7-L (tal como se describe en el presente documento) para fines de diagnóstico *in vivo* e *in vitro*, incluyendo, pero sin limitarse a, el uso en forma marcada para detectar la presencia del polipéptido B7-L en una muestra de células o líquido corporal. También pueden usarse los anticuerpos para prevenir, tratar o diagnosticar varias enfermedades y trastornos, incluyendo los citados en el presente documento. Los anticuerpos pueden unirse a un polipéptido B7-L de modo que se disminuya o bloquee al menos una actividad característica de un polipéptido B7-L, pueden unirse a un polipéptido para aumentar al menos una actividad característica de un polipéptido B7-L (incluyendo aumentando la farmacocinética del polipéptido B7-L).

Pueden usarse polipéptidos B7-L para clonar ligandos de B7-L usando una estrategia de "clonación por expresión". Puede usarse polipéptido B7-L radiomarcado (¹²⁵yodo) o polipéptido B7-L "con etiqueta de afinidad/actividad" (tal como una fusión con Fc o una fusión con fosfatasa alcalina) en ensayos de unión para identificar un tipo celular, línea celular o tejido que expresa un ligando de B7-L. El ARN aislado a partir de tales células o tejidos puede convertirse entonces en ADNc, clonarse en un vector de expresión de mamífero, y transfectarse en células de mamífero (por ejemplo, COS o 293) para crear una biblioteca de expresión. Entonces puede usarse el polipéptido B7-L radiomarcado o con etiqueta como un reactivo de afinidad para identificar y aislar el subconjunto de células en esta biblioteca que expresa un ligando de B7-L. Entonces se aísla el ADN de estas células y se transfecta en células de mamífero para crear una biblioteca de expresión secundaria en la que la fracción de células que expresan el ligando de B7-L sería muchas veces mayor que en la biblioteca original. Este proceso de enriquecimiento puede repetirse iterativamente hasta que se aísla un único clon recombinante que contiene el ligando de B7-L. El aislamiento de ligandos de B7-L es útil para identificar o desarrollar agonistas y antagonistas novedosos de la ruta de señalización de B7-L. Tales agonistas y antagonistas incluyen ligandos de B7-L, anticuerpos anti-ligandos de B7-L, moléculas pequeñas u oligonucleótidos antisentido.

Se prepararon depósitos de ADNc que codifican el polipéptido B7-L humano, se subclonaron en pGEM-T-Easy (Promega, Madison, WI), y que tenían el número de acceso PTA 2481, en la Colección Americana de Cultivos Tipo, 10801 University Boulevard, Manassas, VA 20110-2209 el 19 de septiembre de 2000.

Los ácidos nucleicos de B7-L humano de la presente invención también son herramientas útiles para aislar los correspondientes genes cromosómicos del polipéptido B7-L. Puede usarse el ADN genómico de B7-L humano para identificar enfermedades hereditarias de degeneración tisular.

Los siguientes ejemplos ilustran la invención.

Ejemplo 1: Clonación de los genes del polipéptido B7-L humano

Generalmente, se usaron los materiales y métodos descritos en Sambrook *et al* citado anteriormente para clonar y analizar los genes que codifican polipéptidos B7-L humanos y murinos.

Se realizó una búsqueda de la base de datos Genbank-EMBL usando el programa TBLASTX (<http://blast.wustl.edu>) y B7-H1 como la secuencia de consulta. Se identificó que un clon genómico humano (Celera Genomics, Rockville, MD, GA_16817596) contenía la secuencia de ácido nucleico que codificaba para un supuesto nuevo miembro de la familia B7. La secuencia de ADNc predicha para este supuesto nuevo miembro de la familia B7 se ensambló usando un clon que contenía una secuencia parcial de ácido nucleico de B7-L (n.º de registro GenBank AK001872).

Se usó ADN de plásmido de diversas bibliotecas de ADNc y bibliotecas de ADNc de Marathon (Clontech, Palo Alto, CA) como molde en amplificaciones por PCR realizadas con los cebadores 2515-27 (5'-C-A-T-A-A-T-A-G-A-G-C-A-T-G-G-C-A-G-C-A-A-T-G-T-G-A-C-3'; SEQ ID NO: 26) y 2524-63 (5'-G-G-G-T-C-C-T-G-G-A-G-T-G-G-C-T-G-G-T-G-T-T-G-3'; SEQ ID NO: 27). Se diseñaron los cebadores de PCR para que correspondiesen a secuencias dentro de un supuesto exón en la secuencia de ADNc identificada anteriormente. Se realizaron amplificaciones por PCR usando técnicas convencionales.

Se obtuvo el fragmento de PCR de 400 pb esperado a partir de bibliotecas de ADNc generadas a partir de estómago, timo, cuero cabelludo, bóveda craneal, fémur, mesenterio, bazo, columna vertebral, tráquea y placenta fetales humanos. Además, se obtuvo el fragmento esperado en células T adultas humanas, protuberancia cerebral/médula, y NVL del mesencéfalo. Además, se obtuvo el fragmento esperado en una línea celular de linfoma y en tumores de colon, mama, ovario y pulmón. Las bibliotecas de ADNc de Marathon para glándula suprarrenal, cerebro, riñón, hígado, pulmón, bazo y timo fetales humanos, y médula ósea, corazón, riñón, hígado, pulmón, páncreas, placenta, retina, músculo esquelético, intestino delgado, bazo, testículos y timo de adulto también produjeron el fragmento esperado. También se obtuvo el fragmento de 400 pb a partir de una subcombinación (E4) de ADNc de tejido mixto humano (conteniendo la subcombinación aproximadamente 15.000 clones). La subcombinación se derivó de una biblioteca sintetizada a medida (LTI-FL, Life Technologies Inc., Rockville, MD)

optimizada para clones de ADNc de longitud completa.

El fragmento de 400 pb obtenido en amplificaciones por PCR de ADNc de tejido mixto humano se aisló y clonó en el vector pGEM-T-Easy® (Promega, Madison, WI). Se determinó la secuencia de ADN de un clon seleccionado para confirmar que la secuencia del clon era idéntica a la de la secuencia genómica identificada originariamente. Entonces, se escindió el fragmento de 400 pb del vector y se marcó mediante incorporación de ³²P-dCTP. Se usó el fragmento marcado para seleccionar 150.000 colonias bacterianas derivadas de la subcombinación de 15.000 clones de la biblioteca de ADNc de LTI-FL que dio positivo en la prueba en el experimento de amplificación por PCR anterior. Se transfirieron las colonias de placas con LB/ampicilina a filtros de nitrocelulosa, se prehibridaron en SSC 6X, SDS al 0,5%, disolución de Denhardt 1X y ADN de esperma de salmón desnaturalizado 100 µg/ml durante 3 horas a 60°C. Tras la adición de 1 x 10⁶ cpm/ml de la sonda marcada con ³²P, se realizó la hibridación durante la noche en las mismas condiciones. Se lavaron los filtros dos veces durante 30 minutos a temperatura ambiente en SSC 2X y SDS al 0,1% y luego dos veces durante 30 minutos a 65°C en SSC 0,1X y SDS al 0,1%. Luego se expusieron los filtros a película de rayos X durante 6 horas a -80°C con pantallas intensificadoras.

Se identificaron varias colonias positivas de esta manera y se preparó ADN de plásmido a partir de estos clones mediante métodos convencionales. Los insertos de ADNc de estas colonias tenían todos aproximadamente 2,2 kb de longitud. El análisis de la secuencia de ADN confirmó que los clones contenían la supuesta región codificante de un nuevo miembro de la familia B7, B7-L, y que la secuencia de ácido nucleico identificada en los clones era idéntica a la identificada en el clon genómico GA_16817596.

El análisis de secuencia del ADNc de longitud completa para el polipéptido B7-L humano indicó que el gen comprende un marco de lectura abierto de 819 pb que codifica una proteína de 273 aminoácidos (figura 1). Usando un secuenciador de proteínas HP G100, se secuenció el extremo amino-terminal del polipéptido B7-L maduro y se encontró que comprendía la secuencia de aminoácidos L-F-T-V-T-V-P-K-E-L-Y-I-I-E (SEQ ID NO: 28), estableciendo de ese modo que el polipéptido B7-L tiene un péptido señal de 19 aminoácidos de longitud en su extremo amino-terminal (véase la figura 1; péptido señal predicho indicado mediante subrayado).

La secuencia de ácido nucleico identificada en los clones, aunque contenía todo el marco de lectura abierto, no contenía la región no traducida en 5' completa del transcrito de B7-L. Se determina esta secuencia usando rondas sucesivas de 5'-RACE usando los cebadores 2515-24 (5'-G-T-G-G-C-T-C-T-T-C-A-C-G-T-G-T-G-G-G-A-T-G-3'; SEQ ID NO: 29) y 2538-68 (5'-C-C-A-G-T-G-T-C-A-A-A-G-T-T-G-C-A-T-T-C-C-A-G-G-T-3'; SEQ ID NO: 30) y los siguientes moldes: bibliotecas de ADNc de Marathon derivadas de hígado, bazo y timo fetales, y médula ósea, pulmón y bazo de adulto; bibliotecas de ADNc de una línea celular de linfoma y tumor de ovario; y bibliotecas de ADNc de bazo, placenta fetales, y células T y columna vertebral de adulto. Se emplearon protocolos para RACE convencionales. Se obtuvieron bandas claras para la amplificación por RACE de la línea celular de linfoma, bazo, placenta fetales, y células T de adulto. Se ligaron estos productos en el vector pGEM-T-Easy®. Se determina la secuencia de la región no traducida en 5' del transcrito de B7-L mediante la secuenciación de clones seleccionados de estas reacciones de transformación.

El producto de proteína predicho del gen B7-L se refiere a la familia B7 de proteínas. Estas proteínas son miembros de la superfamilia de inmunoglobulinas y funcionan como reguladores de la respuesta inmunitaria mediada por células T. Los miembros de la familia B7 de proteínas son proteínas de membrana de tipo-1 con un pequeño dominio citoplasmático y regiones extracelulares que contienen dominios de inmunoglobulina V (variable) y C (constante). Los elementos conocidos de la familia B7 incluyen CD80 (B7-1), CD86 (B7-2), B7-rpl y B7-H1. B7-1 y B7-2 interactúan con CD28 y CTLA-4 y son mediadores de la ruta coestimuladora de las células T. B7-rpl se une a un receptor distinto (ICOS; coestimulador inducible) y también es un estimulador de la proliferación de células T. B7-H1 también coestimula la proliferación de células T, pero no se une a CD28, CTLA-4 o ICOS. Las secuencias de proteína de esta familia se conservan escasamente y por consiguiente, son difíciles de distinguir de otras moléculas relacionadas usando métodos informáticos, especialmente cuando se compara sólo una parte de la secuencia de región codificante de longitud completa. Otras proteínas que muestran homología de secuencia con la familia B7 incluyen las butirofilinas y PRO352. Están relacionadas de manera más lejana todavía las proteínas de mielina de los oligodendrocitos (MOGs). Las figuras 2A-2C ilustran una alineación de secuencias de aminoácidos de las proteínas humanas polipéptido B7-L (GA16817596), CD80 (B7-1), CD86 (B7-2), B7-H1, B7rp-1, PRO352, butirofilina BTF1, butirofilina BTF2, butirofilina BTF4, butirofilina BTF3 y butirofilina.

La secuencia de longitud completa del ortólogo de ratón de B7-L se depositó en GenBank el 1 de junio de 1999 (n.º de registro AF142780). El ortólogo de ratón de B7-L se situó con precisión en la familia de B7/butirofilina en el momento del depósito. Se depositó una secuencia parcial de ADNc de B7-L humano, que codificaba para 173 aminoácidos de la parte C-terminal del polipéptido B7-L en GenBank el 23 de febrero de 2000 (n.º de registro AK001872). El polipéptido parcial codificado por este clon no se asignó a una familia de proteínas particular en el momento del depósito.

Se derivó una estructura parcial de intrón-exón para el gen B7-L a partir de los clones genómicos GA_43440610, GA_43068628, GA_43068627, y GA_43440600 (Celera Genomics). Las figuras 3A-3E ilustran una parte de la secuencia de nucleótidos genómica para el polipéptido B7-L humano (SEQ ID NO: 14). La ubicación de la secuencia de aminoácidos deducida del exón 1 (SEQ ID NO: 19) se muestra en la figura 3C. La secuencia mostrada en las figuras 3A-3E está separada por un hueco de tamaño desconocido de la parte de la secuencia de nucleótidos

genómica mostrada en la figura 4 (SEQ ID NO: 15). La secuencia mostrada en la figura 4 está separada por una secuencia enmascarada de aproximadamente 2400 bases de la secuencia de nucleótidos genómica mostrada en las figuras 5A-5F (SEQ ID NO: 16). La ubicación de la secuencia de aminoácidos deducida del exón 2 (SEQ ID NO: 20) se muestra en la figura 5D. La secuencia mostrada en las figuras 5A-5F está separada por un hueco de aproximadamente 9600 bases de la secuencia de nucleótidos genómica mostrada en las figuras 6A-6B (SEQ ID NO: 17). La ubicación de la secuencia de aminoácidos deducida del exón 3 (SEQ ID NO: 21) se muestra en 6A. La secuencia mostrada en las figuras 6A-6B está separada por una secuencia enmascarada de aproximadamente 720 bases de la secuencia de nucleótidos genómica mostrada en las figuras 7A-7M (SEQ ID NO: 18). Las ubicaciones de la secuencia de aminoácidos deducida de los exones 4 (SEQ ID NO: 22), 5 (SEQ ID NO: 23), y 6 se muestran en 7D-7E, 7H, y 7L, respectivamente.

Ejemplo 2: Expresión de ARNm de B7-L

Se hibridó mediante transferencia de tipo Northern (Clontech) un tejido humano múltiple a una sonda de 874 pb correspondiente a los nucleótidos 33-906 de la secuencia de ADNc de B7-L humano. Se marcó de manera radiactiva la sonda usando un kit de marcaje de cebadores al azar Prime-It RmT (Stratagene) según las instrucciones del fabricante. Se hibridaron las transferencias de tipo Northern y se lavaron según las instrucciones del fabricante, y luego se expusieron a autorradiografía. La figura 8 ilustra los resultados del análisis de transferencia de tipo Northern.

La expresión del ARNm de B7-L se localiza mediante hibridación *in situ*. Se fija un panel de tejidos de ratón embrionarios y de adulto normales en paraformaldehído al 4%, se incluyen en parafina, y se seccionan a 5 µm. Se permeabilizan los tejidos seccionados en HCl 0,2 M, se digieren con proteinasa K, y se acetilan con trietanolamina y anhídrido acético. Se prehibridan las secciones durante 1 hora a 60°C en disolución de hibridación (NaCl 300 mM, Tris-HCl 20 mM, pH 8,0, EDTA 5 mM, disolución de Denhardt 1X, SDS al 0,2%, DTT 10 mM, ARNt 0,25 mg/ml, poliA 25 µg/ml, poliC 25 µg/ml y formamida al 50%) y luego se hibridaron durante la noche a 60°C en la misma disolución que contenía dextrano al 10% y 2×10^4 cpm/µl de una ribsonda antisentido marcada con ^{33}P complementario al gen B7-L humano. Se obtiene la ribsonda mediante transcripción *in vitro* de un clon que contiene secuencias de ADNc de B7-L humano usando técnicas convencionales.

Tras la hibridación, se enjuagaron las secciones con disolución de hibridación, se trataron con ARNasaA para digerir la sonda no hibridada, y luego se lavaron con SSC 0,1X a 55°C durante 30 minutos. Luego se sumergieron las secciones en emulsión NTB-2 (Kodak, Rochester, NY), se expusieron durante 3 semanas a 4°C, se revelaron, y se contratiñeron con hematoxilina y eosina. Se analizaron simultáneamente la morfología tisular y la señal de hibridación mediante campo oscuro e iluminación convencional para el cerebro (una sección sagital y dos coronales), tubo digestivo (esófago, estómago, duodeno, yeyuno, íleo, colon proximal y colon distal), hipófisis, hígado, pulmón, corazón, bazo, timo, ganglios linfáticos, riñón, glándula suprarrenal, vejiga, páncreas, glándula salivar, órganos reproductores masculinos y femeninos (ovario, trompa de Falopio y útero en la mujer; y testículos, epidídimo, próstata, vesícula seminal y conducto deferente en el hombre), TAM y TAB (subcutáneo, perirrenal), hueso (fémur), piel, mama y músculo esquelético.

Ejemplo 3: Producción de polipéptidos B7-L

A. Expresión de polipéptidos B7-L en bacterias

Se usó PCR para amplificar secuencias de ADN molde que codifican un polipéptido B7-L usando cebadores correspondientes a los extremos 5' y 3' de la secuencia. Pueden modificarse los productos de ADN modificados para que contengan sitios para enzimas de restricción para permitir la inserción en vectores de expresión. Se purifican en gel los productos de PCR y se insertan en vectores de expresión usando metodología de ADN recombinante convencional. Se digiere un vector a modo de ejemplo, tal como pAMG21 (n.º ATCC 98113) que contiene el promotor lux y un gen que codifica resistencia a kanamicina, con Bam HI y Nde I para la clonación direccional del ADN insertado. Se transforma la mezcla ligada en una cepa huésped de *E. coli* mediante electroporación y se seleccionan los transformantes por la resistencia a kanamicina. Se aísla ADN de plásmido de colonias seleccionadas y se somete a secuenciación de ADN para confirmar la presencia del inserto.

Se incuban células huésped transformadas en medio 2xYT que contiene kanamicina 30 µg/mL a 30°C antes de la inducción. Se induce la expresión génica mediante la adición de N-(3-oxohexanoil)-dl-homoserina-lactona hasta una concentración final de 30 ng/mL seguido por incubación a o bien 30°C o bien 37°C durante seis horas. Se evalúa la expresión del polipéptido B7-L mediante centrifugación del cultivo, resuspensión y lisis de los sedimentos bacterianos, y análisis de las proteínas de las células huésped mediante electroforesis en gel de SDS-poliacrilamida.

Se purifican los cuerpos de inclusión que contienen el polipéptido B7-L tal como sigue. Se sedimentan células bacterianas mediante centrifugación y se resuspenden en agua. Se lisa la suspensión celular mediante sonicación y se sedimenta mediante centrifugación a 195.000 xg durante de 5 a 10 minutos. Se desecha el sobrenadante, y se lava el sedimento y se transfiere a un homogeneizador. Se homogeneiza el sedimento en 5 mL de una disolución de Percoll (Percoll líquido al 75% y NaCl 0,15 M) hasta que se suspende uniformemente y luego se diluye y se centrifuga a 21.600 xg durante 30 minutos. Se recuperan fracciones de gradiente que contienen los cuerpos de inclusión y se combinan. Se analizan los cuerpos de inclusión aislados mediante SDS-PAGE.

Se escinde del gel una única banda en un gel de SDS-poliacrilamida correspondiente al producido en *E. coli*, y se determina la secuencia de aminoácidos N-terminal esencialmente como describen Matsudaira *et al.*, 1987, J. Biol. Chem. 262: 10-35.

B. Expresión del polipéptido B7-L en células de mamífero

- 5 Se usa PCR para amplificar secuencias de ADN molde que codifican un polipéptido B7-L usando cebadores correspondientes a los extremos 5' y 3' de la secuencia. Pueden modificarse los productos de ADN amplificados para que contengan sitios para enzimas de restricción para permitir la inserción en vectores de expresión. Se purifican en gel los productos de PCR y se insertan en vectores de expresión usando metodología de ADN recombinante convencional. Puede usarse un vector a modo de ejemplo, pCEP4 (Invitrogen, Carlsbad, CA), que
10 contiene un origen de replicación del virus Epstein-Barr, para la expresión de polipéptidos B7-L en células 293-EBNA-1. Se ligan los productos de PCR purificados en gel y amplificados en vector pCEP4 y se introducen en células 293-EBNA mediante lipofección. Se seleccionan las células transfectadas en higromicina 100 µg/ml y se hacen crecer los cultivos resistentes a fármacos resultantes hasta confluencia. Luego se cultivan las células medios libres de suero durante 72 horas. Se eliminan los medios condicionados y se analiza la expresión del polipéptido B7-L mediante SDS PAGE.
15

Puede detectarse la expresión del polipéptido B7-L mediante tinción con plata. Alternativamente, se produce el polipéptido B7-L como una proteína de fusión con una etiqueta epitópica, tal como un dominio constante de IgG o un epítipo FLAG, que puede detectarse mediante análisis de inmunotransferencia de tipo Western usando anticuerpos contra la etiqueta peptídica.

- 20 Pueden escindirse polipéptidos B7-L de un gel de SDS-poliacrilamida, o se purifican proteínas de fusión de B7-L mediante cromatografía de afinidad para la etiqueta epitópica, y se someten a análisis de la secuencia de aminoácidos N-terminal tal como se describe en el presente documento.

C. Expresión de la proteína de fusión polipéptido B7-L/Fc en células de mamífero

- 25 Se amplificó un fragmento de 660 pb de ADNc de B7-L, que codificaba para 220 aminoácidos en el extremo amino-terminal del polipéptido B7-L, y que incluía el péptido señal y el dominio extracelular del polipéptido B7-L, mediante PCR usando cebadores adecuados y técnicas convencionales. Se usó el ADN de plásmido aislado de la biblioteca sintetizada a medida descrita en el ejemplo 1 como molde en las amplificaciones por PCR.

- Luego se clonó el fragmento de 660 pb en el vector de expresión, pCEP4/Fc, que contiene la secuencia de ácido nucleico que codifica los 235 aminoácidos carboxi-terminales de Fc humano (IgG-1). Se seleccionaron colonias tras
30 la transformación de células bacterianas, y se aisló ADN de plásmido de pCEP4-B7-L/Fc usando técnicas convencionales.

- Se usó el ADN de plásmido aislado para transfectar células 293-EBNA-1 usando el reactivo de transfección FuGENE 6 (Roche Molecular Biochemicals, Indianápolis, IN) según las instrucciones del fabricante. Tras la transfección, se cultivaron las células a 37°C en medio DMEM, complementado con suero bovino fetal al 10%,
35 durante 24 horas, y luego se transfirieron a medio libre de suero DMEM y se hicieron crecer a 37°C durante 5 días. Se purificó la proteína de fusión polipéptido B7-L/Fc de los medios condicionados en una columna de proteína A HiTrap (Amersham Pharmacia Biotech, Piscataway, NJ) según las instrucciones del fabricante, y luego se dializó la proteína de fusión en PBS. Se midió la concentración de la proteína purificada usando reactivo de ensayo Coomassie® Plus Protein (Pierce, Rockford, IL) y BSA como patrón, según las instrucciones del fabricante.

- 40 Ejemplo 4: Identificación del polipéptido B7-L como ligando de PD-1 novedoso

Para determinar si el polipéptido B7-L funciona en una de las rutas coestimuladores mediadas por B7 conocidas, se realizó análisis de FACS en células CHO D- que expresan CD28, CRP-1/ICOS o PD-1 tras tratamiento con una proteína de fusión polipéptido B7-L/Fc.

- Se obtuvo la secuencia de ácido nucleico de longitud completa que codifica PD-1 murino (mPD-1; Ishida *et al.*, 1992, EMBO J. 11: 3887-95), y que incluye el péptido señal de PD-1 nativo, mediante amplificación por PCR de una biblioteca de ADNc de linfocitos de bazo activado murino usando cebadores que incorporan sitios de endonucleasas de restricción Hind III y Sal I. Se digirió el producto de PCR resultante con Hind III y Sal I y luego se ligó en el vector pDSRa-19. Tras la transformación en células bacterianas, se seleccionaron clones y se secuenciaron. Se linealizó un clon que contenía la secuencia de ADNc de mPD-1 de longitud completa (clon 1,5) usando la endonucleasa de
50 restricción Pvu I, y se usó el plásmido linealizado para transfectar células CHO D- mediante el método con fosfato de calcio. Se cultivaron las células CHO D- transfectadas, y se aislaron colonias individuales mediante clonación en anillo varias semanas después. Se identificó un transfectante de alta expresión (clon 3,36) por su capacidad para unirse específicamente a la proteína de fusión B7-HI/Fc (Dong *et al.*, 1999, Nat. Med. 5: 1365-69).

- Se cultivaron células CHO D- que expresaban sólo el vector, mPD-1 (clon 3,36), CD28 humano, o mCRP1/ICOS (clon 1,41; Yoshinaga *et al.*, 1999, Nature 402: 827-32) en matraces T175 en DMEM complementados con dFBS al 5%, PSG 1X y NEAA 1X. Se liberaron células de los matraces de cultivo usando disolución de disociación celular (Sigma,
55

San Luis, MO), se diluyeron con tampón de lavado (PBS que contenía BSA al 0,5%), y luego se contaron.

Se hicieron reaccionar aproximadamente $3,0 \times 10^5$ células en 0,1 ml de medios durante 1 hora en hielo con 1 μ g de CRP1/Fc murino purificado, polipéptido B7-L humano/Fc, B7rp1 murino/Fc, o B7-2/Fc murino. También se incubaron células CHO D- con B7-H1 murino/Fc tal como se describió anteriormente, excepto porque las células se expusieron a 10 μ g de proteína de fusión en 1 ml de medios condicionados libres de suero recogidos de células CHO D- que expresaban mB7H1/Fc. Tras la incubación, se diluyeron las células con 5 ml de tampón de lavado, se centrifugaron y luego se lavaron dos veces con 5 ml de tampón de lavado. Luego se resuspendieron las células en 100 μ l de tampón de lavado que contenía 10 μ g/ml de un anticuerpo de detección conjugado con FITC específico de Fc de IgG de cabra anti-humano (Chemicon, Temecula CA). Se permitió que las células reaccionaran durante 30 minutos en hielo, y luego se lavaron tal como se describió anteriormente. Tras el lavado, se resuspendieron las células en un volumen final de 1 ml de tampón de lavado y luego se analizaron usando un aparato FACS Star (Beckman Dickinson, Franklin Lakes, NJ).

Tal como se muestra en la figura 9, polipéptido B7-L se une específicamente al receptor PD-1, pero no a los receptores CD28 o CRP-1/ICOS. Las proteína de fusión B7-2/Fc y B7rp-1/Fc se unieron a los receptores CD28 y CRP-1/ICOS, tal como se esperaba. Como se mostró también que B7-H1 se unía específicamente al receptor PD-1 (figura 9B), el análisis de FACS indica que tanto el polipéptido B7-L como B7-H1 son ligandos para el receptor PD-1. Puesto que PD-1 es un regulador negativo de la proliferación de células T (Nishimura *et al.*, 1999, Immunity 11: 141-51), la activación de esta ruta mediante el polipéptido B7-L soluble puede dar como resultado una respuesta inmunitaria reducida. A la inversa, los anticuerpos antagonistas contra el polipéptido B7-L pueden aumentar las funciones inmunitarias.

Ejemplo 5: Inhibición de la proliferación de células T por el polipéptido B7-L

Para determinar si el polipéptido B7-L desempeña un papel en la proliferación de células T, se realizaron ensayos de proliferación de células T humanas usando proteínas de fusión polipéptido B7-L/Fc. Se aislaron células T humanas altamente purificadas (> 98% de CD3+) mediante selección negativa de células mononucleares de sangre periférica (CMSP) frescas o descongeladas, de adherencia reducida usando el kit de aislamiento de células T Pan (Miltenyi Biotec, Auburn, CA). Se recubrieron previamente placas de cultivo celular de fondo redondo, de 96 pocillos durante la noche a 4°C con 0,5 o 1,5 μ g/ml de anticuerpo anti-CD3 (PharMingen, San Diego, CA), Fc de IgG anti-humana 10 μ g/ml (Sigma), y PBS, en un volumen de 0,1 ml. Se eliminó la disolución de recubrimiento previo, se lavaron las placas una vez con PBS, y se añadieron a los pocillos 0,1 ml de las proteínas de fusión polipéptido B7-L/Fc, B7-2/Fc o B7rp-1/Fc diluidas hasta 20 μ g/ml en medios RPMI 1640 complementados con FCS al 10%, o medios solos. Luego se incubaron las placas durante 4 horas a 37°C. Tras la incubación, se eliminaron los medios y se añadieron y 0,2 ml de las células T purificadas (que contenían 10×10^5 células) a cada pocillo recubierto previamente. Luego se incubaron las placas durante 48 horas a 37°C. Tras la incubación, se añadió 1 μ Ci/pocillo de [³H]-timidina a cada pocillo y se incubaron los cultivos durante 18 h adicionales a 37°C. Luego se recogieron las células y se midieron las cuentas por minuto (CPM).

Tal como se muestra en la figura 10, la proteína de fusión polipéptido B7-L/Fc pudo inhibir la proliferación de células T mediada por el anticuerpo anti-CD3 a cualquier concentración del anticuerpo anti-CD3. A la inversa, las proteínas de fusión B7rp-1/Fc y B7-1/Fc mostraron que coestimulaban la proliferación de células T en las mismas condiciones. Estos resultados indican que el polipéptido B7-L es un regulador negativo de la proliferación de células T. Además, estos resultados sugieren que la proteína de fusión polipéptido B7-L/Fc, u otros agonistas solubles del polipéptido B7-L, pueden usarse para inhibir la función inmunitaria, proporcionando de ese modo resultados terapéuticos favorables. También podrían usarse ensayos *in vitro*, tales como el ensayo descrito en el presente documento para seleccionar anticuerpos, proteínas solubles, o inhibidores de molécula pequeña de la actividad del polipéptido B7-L.

Ejemplo 6: Los receptores del polipéptido B7-L se expresan en células mononucleares de sangre periférica humanas

Se usó el análisis de FACS para identificar receptores del polipéptido B7-L en células T y B en CMSP humanas. Usando la centrifugación en gradiente Ficoll-Paque (Amersham Pharmacia Biotech), se purificaron CMSP a partir de sangre obtenida de voluntarios humanos sanos, y se estimularon las células mediante incubación en lipopolisacárido (LPS) 10 μ g/ml durante 3 días. Tras el tratamiento con LPS, se bloquearon 5×10^5 células en un volumen de 0,2 ml con Fc de IgG humana 100 μ g/ml durante 10 minutos en hielo. Luego se incubaron las células con 10 μ g/ml de diversas proteínas de fusión con Fc biotiniladas (o controles adecuados) durante 30 minutos en hielo. Tras la incubación con proteínas de fusión con Fc, se incubaron las muestras celulares indicadas en la figura 11 con anticuerpos anti-CD3 (PharMingen) o antiCD19 (Becton Dickinson) durante 30 minutos en hielo. Tras la incubación, se lavaron las células dos veces en tampón de lavado (PBS que contenía BSA al 0,5%), y luego se tiñeron con FITC-avidina (dilución 1:100) durante 30 minutos en hielo. Se resuspendieron las células en 1 ml de tampón de lavado y se analizaron en un aparato FACS Star.

Tal como se muestra en la figura 11, la proteína polipéptido B7-L/Fc se unió a poblaciones significativas de células mononucleares de sangre periférica (CMSP). La doble tinción con la proteína de fusión polipéptido B7-L/Fc y marcadores de células T (CD3) o células B (CD19) indicó que una parte de las células a las que se unió la proteína de fusión polipéptido B7-L/Fc eran células T y células B. También se mostró que números significativos de las

células que expresaban el receptor del polipéptido B7-L, sin embargo, no eran células T ni células B. Por tanto, células distintas de linfocitos, así como, linfocitos en CMSP pueden estar reguladas por el polipéptido B7-L. Este patrón de unión concuerda con el patrón de expresión de PD-1 (Ishida *et al.*, 1992, EMBOJ. 11: 3887-95).

Ejemplo 7: Producción de anticuerpos anti-polipéptido B7-L

5 Pueden obtenerse anticuerpos contra los polipéptidos B7-L mediante inmunización con proteína purificada o con péptidos B7-L producidos mediante síntesis biológica o química. Los procedimientos adecuados para generar anticuerpos incluyen los descritos en Hudson y Bay, Practical Immunology (2nd ed., Blackwell Scientific Publications).

10 En un procedimiento para la producción de anticuerpos, se les inyectó a animales (normalmente ratones o conejos) un antígeno de B7-L (tal como un polipéptido B7-L), y aquéllos con niveles de títulos séricos suficientes según se determina mediante ELISA se seleccionan para la producción del hibridoma. Se recogen los bazo de los animales inmunizados y se preparan como suspensiones de células individuales de las que se recuperan los esplenocitos. Se fusionan los esplenocitos con células de mieloma de ratón (tales como células Sp2/0-Ag14), se incuban en primer lugar en DMEM con penicilina 200 U/mL, sulfato de estreptomicina 200 µg/ml y glutamina 4 mM, y luego se incuban en medio de selección HAT (hipoxantina, aminopterina y timidina). Tras la selección, se toman los sobrenadantes de cultivo tisular de cada pocillo de fusión y se someten a prueba para determinar la producción de anticuerpos anti-B7-L mediante ELISA.

20 También pueden emplearse procedimientos alternativos para obtener los anticuerpos anti-B7-L, tal como la inmunización de ratones transgénicos que portan loci de Ig humana para la producción de anticuerpos humanos, y la selección de bibliotecas de anticuerpos sintéticos, tales como las generadas mediante mutagénesis de un dominio variable de anticuerpo.

Ejemplo 8: Expresión del polipéptido B7-L en ratones transgénicos

25 Para evaluar la actividad biológica del polipéptido B7-L, se prepara una construcción que codifica una proteína de fusión polipéptido B7-L/Fc bajo el control de un promotor de ApoE específico del hígado. Se espera que la administración de este constructo provoque cambios patológicos que son informativos en cuanto a la función del polipéptido B7-L. De manera similar, se prepara una construcción que contiene el polipéptido B7-L de longitud completa bajo el control del promotor de beta-actina. Se espera que la administración de este constructo dé como resultado la expresión ubicua.

30 Para generar estas construcciones, se usa PCR par amplificar secuencias de ADN molde que codifican un polipéptido B7-L usando cebadores que corresponden a los extremos 5' y 3' de la secuencia deseada y que incorporan sitios de enzimas de restricción para permitir la inserción del producto amplificado en un vector de expresión. Tras la amplificación, se purifican en gel los productos de PCR, se digieren los las enzimas de restricción apropiadas, y se ligan en un vector de expresión usando técnicas de ADN recombinante convencionales. Por ejemplo, pueden clonarse secuencias del polipéptido B7-L amplificadas en un vector de expresión bajo el control del promotor de β-actina humana tal como se describe por Graham *et al.*, 1997, Nature Genetics, 17: 272-74 y Ray *et al.*, 1991, Genes Dev. 5: 2265-73.

40 Tras la ligación, se usan mezclas de reacción para transformar una cepa huésped de *E. coli* mediante electroporación y se seleccionan los transformantes por la resistencia a fármacos. Se aísla ADN de plásmido de colonias seleccionadas y se somete a secuenciación de ADN para confirmar la presencia de un inserto apropiado y ausencia de mutación. El vector de expresión del polipéptido B7-L se purifica a través de dos tandas de centrifugación en gradiente de densidad de CsCl, se escinden con una enzima de restricción adecuada, y se purifica el fragmento linealizado que contiene el transgén del polipéptido B7-L mediante electroforesis en gel. Se resuspende el fragmento purificado en Tris 5 mM, pH 7,4, y EDTA 0,2 mM a una concentración de 2 mg/mL.

45 Se inyectan embriones de una sola célula de ratones cruzados BDF1 x BDF1 tal como se describe (publicación internacional n.º WO 97/23614). Se cultivan los embriones durante la noche en un incubador de CO₂ y se transfirieron 15-20 embriones de dos células a los oviductos de un ratón hembra CD1 pseudopreñada. La descendencia obtenida de la implantación de embriones microinyectados se selecciona mediante amplificación por PCR del transgén integrado en muestras de ADN genómico tal como sigue. Se digirieron trozos de oreja en 20 mL de tampón para orejas (Tris 20 mM, pH 8,0, EDTA 10 mM, SDS al 0,5% y proteinasa K 500 mg/mL) a 55°C durante la noche. Luego se diluye la muestra con 200 mL de TE, y se usan 2 mL del tampón para orejas en una reacción de PCR reacción usando cebadores apropiados.

55 A las 8 semanas de edad, se sacrifican animales fundadores transgénicos y animales control para realizar la necropsia y el análisis anatomopatológico. Se extirpan partes del bazo y se aísla el ARN celular total de los bazo usando el kit de extracción de ARN total (Qiagen) y se determinó la expresión del transgén mediante RT-PCR. Se convierte el ARN recuperado de los bazo en ADNc usando el sistema de preamplificación SuperScript™ (Gibco-BRL) tal como sigue. Se usa un cebador adecuado, ubicado en la secuencia del vector de expresión y en 3' con respecto al transgén del polipéptido B7-L, para cebar la síntesis de ADNc a partir de los transcritos de transgén. Se incuban diez mg de ARN de bazo total de fundadores transgénicos y controles con 1 mM de cebador durante 10

minutos a 70°C y se ponen en hielo. Se complementa entonces la reacción con Tris-HCl 10 mM, pH 8,3, KCl 50 mM, MgCl₂ 2,5 mM, 10 mM de cada dNTP, DTT 0,1 mM y 200 U de transcriptasa inversa SuperScript II. Tras la incubación durante 50 minutos a 42°C, se detiene la reacción calentando durante 15 minutos a 72°C y se digiere con 2 U de ARNasa H durante 20 minutos a 37°C. Luego se amplifican las muestras mediante PCR cebadores específicos para el polipéptido B7-L.

Ejemplo 9: Actividad biológica del polipéptido B7-L en ratones transgénicos

Antes del sacrificio, se pesan los animales transgénicos, se anestesian mediante isoflurano y se extrae sangre mediante punción cardiaca. Se someten las muestras análisis de hematología y bioquímica sérica. Se realiza una radiografía tras la exanguinación terminal. Tras la disección macroscópica, se someten los principales órganos viscerales a análisis de peso.

Tras la disección macroscópica, se extirpan los tejidos (es decir, hígado, bazo, páncreas, estómago, todo el tubo digestivo, riñón, órganos reproductores, piel y glándulas mamarias, hueso, cerebro, corazón, pulmón, timo, tráquea, esófago, tiroides, glándulas suprarrenales, vejiga urinaria, ganglios linfáticos y músculo esquelético) y se fijan en Zn-formalina tamponado al 10% para el examen histológico. Tras la fijación, se procesan los tejidos en bloques de parafina, y se obtienen secciones de 3 mm. Se tiñen todas las secciones con hematoxilina y eosina, y luego se someten a análisis histológico.

Se someten el bazo, ganglio linfático y placas de Peyer tanto de los ratones transgénicos como control a análisis inmunohistológico con anticuerpos específicos de células B y células T tal como sigue. Las secciones incluidas en parafinas fijadas con formalina se desparafinan y se hidratan con agua desionizada. Se enfrían bruscamente las secciones con peróxido de hidrógeno al 3%, se bloquean con bloqueo de proteínas (Lipshaw, Pittsburgh, PA), y se incuban en anticuerpo monoclonal de rata anti-B220 y CD3 de ratón (Harlan, Indianápolis, IN). Se detecta la unión a anticuerpos mediante inmunoglobulinas de conejo anti-rata biotiniladas y estreptavidina conjugada con peroxidasa (BioGenex, San Ramón, CA) con DAB como cromógeno (BioTek, Santa Bárbara, CA). Se contratiñen las secciones con hematoxilina.

Tras la necropsia, se extirpan los NLM y secciones de bazo y timo de crías de animales transgénicos y control. Se preparan suspensiones de células individuales mediante la trituración suave de los tejidos con el extremo plano de una jeringa contra el fondo de un filtro Cell Strainer de nilón de 100 mm (Becton Dickinson, Franklin Lakes, NJ). Se lavan las células dos veces, se cuentan, y luego se incuban aproximadamente 1×10^6 células de cada tejido durante 10 minutos con 0,5 µg de bloqueo de Fc CD16/32 (FcγIII/II) en un volumen de 20 µl. Luego se tiñen las muestras durante 30 minutos a 2-8°C en un volumen de 100 µl de PBS (que carece de Ca⁺ y Mg⁺), albúmina sérica bovina al 0,1%, y azida de sodio al 0,01% con 0,5 µg anticuerpo de anticuerpos monoclonales conjugados con PE o FITC contra CD90.2 (Thy-1,2), CD45R (B220), CD11b (Mac-1), Gr-1, CD4 o CD8 (PharMingen, San Diego, CA). Tras la unión a anticuerpos, se lavan las células y luego se analizan mediante citometría de flujo en un aparato FACScan (Becton Dickinson).

LISTA DE SECUENCIAS

<110> Fox, Michael

Sullivan, John K.

Holst, Paige

Yoshinaga, Steven Kiyoshi

<120> Polipéptidos de tipo B7 y usos de los mismos

<130> 00,759-B

<140>

<141>

<150> 60/233.867

<151> 2000-09-20

<160> 30

<170> PatentIn Ver. 2.0

<210> 1

<211> 1209

<212> ADN

<213> *Homo sapiens*

<220>

<221> CDS

5 <222> (33)..(854)

<220>

<221> sig_peptide

<222> (33)..(89)

<220>

10 <221> misc_feature

<222> (693)..(755)

<223> dominio de transmembrana predicho

<400> 1

```

cagaaagaga cctatatgat caaatacaga ac atg atc ttc ctc ctg cta atg 53
                                   Met Ile Phe Leu Leu Leu Met
                                   1                               5

ttg agc ctg gaa ttg cag ctt cac cag ata gca gct tta ttc aca gtg 101
Leu Ser Leu Glu Leu Gln Leu His Gln Ile Ala Ala Leu Phe Thr Val
    10                               15                               20

aca gtc cct aag gaa ctg tac ata ata gag cat ggc agc aat gtg acc 149
Thr Val Pro Lys Glu Leu Tyr Ile Ile Glu His Gly Ser Asn Val Thr
    25                               30                               35

ctg gaa tgc aac ttt gac act gga agt cat gtg aac ctt gga gca ata 197
Leu Glu Cys Asn Phe Asp Thr Gly Ser His Val Asn Leu Gly Ala Ile
    40                               45                               50                               55

aca gcc agt ttg caa aag gtg gaa aat gat aca tcc cca cac cgt gaa 245
Thr Ala Ser Leu Gln Lys Val Glu Asn Asp Thr Ser Pro His Arg Glu
    60                               65                               70

aga gcc act ttg ctg gag gag cag ctg ccc cta ggg aag gcc tcg ttc 293

```

ES 2 372 522 T3

```

Arg Ala Thr Leu Leu Glu Glu Gln Leu Pro Leu Gly Lys Ala Ser Phe
      75                      80                      85

cac ata cct caa gtc caa gtg agg gac gaa gga cag tac caa tgc ata 341
His Ile Pro Gln Val Gln Val Arg Asp Glu Gly Gln Tyr Gln Cys Ile
      90                      95                      100

atc atc tat ggg gtc gcc tgg gac tac aag tac ctg act ctg aaa gtc 389
Ile Ile Tyr Gly Val Ala Trp Asp Tyr Lys Tyr Leu Thr Leu Lys Val
      105                      110                      115

aaa gct tcc tac agg aaa ata aac act cac atc cta aag gtt cca gaa 437
Lys Ala Ser Tyr Arg Lys Ile Asn Thr His Ile Leu Lys Val Pro Glu
      120                      125                      130                      135

aca gat gag gta gag ctc acc tgc cag gct aca ggt tat cct ctg gca 485
Thr Asp Glu Val Glu Leu Thr Cys Gln Ala Thr Gly Tyr Pro Leu Ala
      140                      145                      150

gaa gta tcc tgg cca aac gtc agc gtt cct gcc aac acc agc cac tcc 533
Glu Val Ser Trp Pro Asn Val Ser Val Pro Ala Asn Thr Ser His Ser
      155                      160                      165

agg acc cct gaa ggc ctc tac cag gtc acc agt gtt ctg cgc cta aag 581
Arg Thr Pro Glu Gly Leu Tyr Gln Val Thr Ser Val Leu Arg Leu Lys
      170                      175                      180

cca ccc cct ggc aga aac ttc agc tgt gtg ttc tgg aat act cac gtg 629
Pro Pro Pro Gly Arg Asn Phe Ser Cys Val Phe Trp Asn Thr His Val
      185                      190                      195

agg gaa ctt act ttg gcc agc att gac ctt caa agt cag atg gaa ccc 677
Arg Glu Leu Thr Leu Ala Ser Ile Asp Leu Gln Ser Gln Met Glu Pro
      200                      205                      210                      215

agg acc cat cca act tgg ctg ctt cac att ttc atc ccc tcc tgc atc 725
Arg Thr His Pro Thr Trp Leu Leu His Ile Phe Ile Pro Ser Cys Ile
      220                      225                      230

att gct ttc att ttc ata gcc aca gtg ata gcc cta aga aaa caa ctc 773
Ile Ala Phe Ile Phe Ile Ala Thr Val Ile Ala Leu Arg Lys Gln Leu
      235                      240                      245

tgt caa aag ctg tat tct tca aaa gac aca aca aaa aga cct gtc acc 821
Cys Gln Lys Leu Tyr Ser Ser Lys Asp Thr Thr Lys Arg Pro Val Thr
      250                      255                      260

aca aca aag agg gaa gtg aac agt gct atc tga acctgtgggc ttgggagcca 874
Thr Thr Lys Arg Glu Val Asn Ser Ala Ile
      265                      270

gggtgacctg atatgacatc taaagaagct tctggactct gaacaagaat tcggtggcct 934

gcagagcttg ccatttgacac ttttcaaagt cctttggatg acccagcact ttaatctgaa 994

acctgcaaca agactagcca acacctggcc atgaaacttg ccccttcact gatctggact 1054

cacctotgga gcoatatggt ttaagcaagc actactgcac ttacagaat taccocactg 1114

gatcctggac ccacagaatt ccttcaggat ccttcttgct gccagactga aagcaaaagg 1174

aattatttcc cctcaagttt tctaagtgat ttcca 1209

```

<210> 2

<211> 273

ES 2 372 522 T3

<212> PRT

<213> *Homo sapiens*

<400> 2

```

Met Ile Phe Leu Leu Leu Met Leu Ser Leu Glu Leu Gln Leu His Gln
 1           5           10           15

Ile Ala Ala Leu Phe Thr Val Thr Val Pro Lys Glu Leu Tyr Ile Ile
      20           25           30

Glu His Gly Ser Asn Val Thr Leu Glu Cys Asn Phe Asp Thr Gly Ser
      35           40           45

His Val Asn Leu Gly Ala Ile Thr Ala Ser Leu Gln Lys Val Glu Asn
      50           55           60

Asp Thr Ser Pro His Arg Glu Arg Ala Thr Leu Leu Glu Glu Gln Leu
      65           70           75           80

Pro Leu Gly Lys Ala Ser Phe His Ile Pro Gln Val Gln Val Arg Asp
      85           90           95

Glu Gly Gln Tyr Gln Cys Ile Ile Ile Tyr Gly Val Ala Trp Asp Tyr
      100          105          110

Lys Tyr Leu Thr Leu Lys Val Lys Ala Ser Tyr Arg Lys Ile Asn Thr
      115          120          125

His Ile Leu Lys Val Pro Glu Thr Asp Glu Val Glu Leu Thr Cys Gln
      130          135          140

Ala Thr Gly Tyr Pro Leu Ala Glu Val Ser Trp Pro Asn Val Ser Val
      145          150          155          160

Pro Ala Asn Thr Ser His Ser Arg Thr Pro Glu Gly Leu Tyr Gln Val
      165          170          175

Thr Ser Val Leu Arg Leu Lys Pro Pro Pro Gly Arg Asn Phe Ser Cys
      180          185          190

Val Phe Trp Asn Thr His Val Arg Glu Leu Thr Leu Ala Ser Ile Asp
      195          200          205

Leu Gln Ser Gln Met Glu Pro Arg Thr His Pro Thr Trp Leu Leu His
      210          215          220

Ile Phe Ile Pro Ser Cys Ile Ile Ala Phe Ile Phe Ile Ala Thr Val
      225          230          235          240

Ile Ala Leu Arg Lys Gln Leu Cys Gln Lys Leu Tyr Ser Ser Lys Asp
      245          250          255

Thr Thr Lys Arg Pro Val Thr Thr Thr Lys Arg Glu Val Asn Ser Ala
      260          265          270

Ile

```

5 <210> 3

<211> 254

<212> PRT

<213> *Homo sapiens*

<220>

<221> TRANSMEM

<222> (202)..(222)

<400> 3

```

Leu Phe Thr Val Thr Val Pro Lys Glu Leu Tyr Ile Ile Glu His Gly
 1           5           10           15
Ser Asn Val Thr Leu Glu Cys Asn Phe Asp Thr Gly Ser His Val Asn
          20           25           30
Leu Gly Ala Ile Thr Ala Ser Leu Gln Lys Val Glu Asn Asp Thr Ser
          35           40           45
Pro His Arg Glu Arg Ala Thr Leu Leu Glu Glu Gln Leu Pro Leu Gly
          50           55           60
Lys Ala Ser Phe His Ile Pro Gln Val Gln Val Arg Asp Glu Gly Gln
          65           70           75           80
Tyr Gln Cys Ile Ile Ile Tyr Gly Val Ala Trp Asp Tyr Lys Tyr Leu
          85           90           95
Thr Leu Lys Val Lys Ala Ser Tyr Arg Lys Ile Asn Thr His Ile Leu
          100          105          110
Lys Val Pro Glu Thr Asp Glu Val Glu Leu Thr Cys Gln Ala Thr Gly
          115          120          125
Tyr Pro Leu Ala Glu Val Ser Trp Pro Asn Val Ser Val Pro Ala Asn
          130          135          140
Thr Ser His Ser Arg Thr Pro Glu Gly Leu Tyr Gln Val Thr Ser Val
          145          150          155          160
Leu Arg Leu Lys Pro Pro Pro Gly Arg Asn Phe Ser Cys Val Phe Trp
          165          170          175
Asn Thr His Val Arg Glu Leu Thr Leu Ala Ser Ile Asp Leu Gln Ser
          180          185          190
Gln Met Glu Pro Arg Thr His Pro Thr Trp Leu Leu His Ile Phe Ile
          195          200          205
Pro Ser Cys Ile Ile Ala Phe Ile Phe Ile Ala Thr Val Ile Ala Leu
          210          215          220
Arg Lys Gln Leu Cys Gln Lys Leu Tyr Ser Ser Lys Asp Thr Thr Lys
          225          230          235          240
Arg Pro Val Thr Thr Thr Lys Arg Glu Val Asn Ser Ala Ile
          245          250

```

5

<210> 4

<211> 224

<212> PRT

<213> *Homo sapiens*

10

<400> 4

Met	Gly	His	Thr	Arg	Arg	Gln	Gly	Thr	Ser	Pro	Ser	Lys	Cys	Pro	Tyr
1				5					10					15	
Leu	Asn	Phe	Phe	Gln	Leu	Leu	Val	Leu	Ala	Gly	Leu	Ser	His	Phe	Cys
			20					25					30		
Ser	Gly	Val	Ile	His	Val	Thr	Lys	Glu	Val	Lys	Glu	Val	Ala	Thr	Leu
		35					40					45			
Ser	Cys	Gly	His	Asn	Val	Ser	Val	Glu	Glu	Leu	Ala	Gln	Thr	Arg	Ile
	50					55					60				
Tyr	Trp	Gln	Lys	Glu	Lys	Lys	Met	Val	Leu	Thr	Met	Met	Ser	Gly	Asp
65					70					75					80
Met	Asn	Ile	Trp	Pro	Glu	Tyr	Lys	Asn	Arg	Thr	Ile	Phe	Asp	Ile	Thr
				85					90					95	
Asn	Asn	Leu	Ser	Ile	Val	Ile	Leu	Ala	Leu	Arg	Pro	Ser	Asp	Glu	Gly
			100					105					110		
Thr	Tyr	Glu	Cys	Val	Val	Leu	Lys	Tyr	Glu	Lys	Asp	Ala	Phe	Lys	Arg
		115					120					125			
Glu	His	Leu	Ala	Glu	Val	Thr	Leu	Ser	Val	Lys	Ala	Asp	Phe	Pro	Thr
	130					135					140				
Pro	Ser	Ile	Ser	Asp	Phe	Glu	Ile	Pro	Thr	Ser	Asn	Ile	Arg	Arg	Ile
145					150					155					160
Ile	Cys	Ser	Thr	Ser	Gly	Gly	Phe	Pro	Glu	Pro	His	Leu	Ser	Trp	Leu
				165					170					175	
Glu	Asn	Gly	Glu	Glu	Leu	Asn	Ala	Ile	Asn	Thr	Thr	Val	Ser	Gln	Asp
			180					185					190		
Pro	Glu	Thr	Glu	Leu	Tyr	Ala	Val	Ser	Ser	Lys	Leu	Asp	Phe	Asn	Met
		195					200					205			
Thr	Thr	Asn	His	Ser	Phe	Met	Cys	Leu	Ile	Lys	Tyr	Gly	His	Leu	Arg
	210					215					220				

<210> 5

<211> 323

<212> PRT

5 <213> *Homo sapiens*

<400> 5

ES 2 372 522 T3

Met	Gly	Leu	Ser	Asn	Ile	Leu	Phe	Val	Met	Ala	Phe	Leu	Leu	Ser	Gly
1				5					10					15	
Ala	Ala	Pro	Leu	Lys	Ile	Gln	Ala	Tyr	Phe	Asn	Glu	Thr	Ala	Asp	Leu
			20					25					30		
Pro	Cys	Gln	Phe	Ala	Asn	Ser	Gln	Asn	Gln	Ser	Leu	Ser	Glu	Leu	Val
		35					40					45			
Val	Phe	Trp	Gln	Asp	Gln	Glu	Asn	Leu	Val	Leu	Asn	Glu	Val	Tyr	Leu
	50					55					60				
Gly	Lys	Glu	Lys	Phe	Asp	Ser	Val	His	Ser	Lys	Tyr	Met	Gly	Arg	Thr
65					70					75					80
Ser	Phe	Asp	Ser	Asp	Ser	Trp	Thr	Leu	Arg	Leu	His	Asn	Leu	Gln	Ile
				85					90					95	
Lys	Asp	Lys	Gly	Leu	Tyr	Gln	Cys	Ile	Ile	His	His	Lys	Lys	Pro	Thr
			100					105					110		
Gly	Met	Ile	Arg	Ile	His	Gln	Met	Asn	Ser	Glu	Leu	Ser	Val	Leu	Ala
	115						120					125			
Asn	Phe	Ser	Gln	Pro	Glu	Ile	Val	Pro	Ile	Ser	Asn	Ile	Thr	Glu	Asn
	130					135					140				
Val	Tyr	Ile	Asn	Leu	Thr	Cys	Ser	Ser	Ile	His	Gly	Tyr	Pro	Glu	Pro
145					150					155				160	
Lys	Lys	Met	Ser	Val	Leu	Leu	Arg	Thr	Lys	Asn	Ser	Thr	Ile	Glu	Tyr
				165					170					175	
Asp	Gly	Ile	Met	Gln	Lys	Ser	Gln	Asp	Asn	Val	Thr	Glu	Leu	Tyr	Asp
			180					185					190		
Val	Ser	Ile	Ser	Leu	Ser	Val	Ser	Phe	Pro	Asp	Val	Thr	Ser	Asn	Met
	195						200					205			
Thr	Ile	Phe	Cys	Ile	Leu	Glu	Thr	Asp	Lys	Thr	Arg	Leu	Leu	Ser	Ser
210						215					220				
Pro	Phe	Ser	Ile	Glu	Leu	Glu	Asp	Pro	Gln	Pro	Pro	Pro	Asp	His	Ile
225					230					235				240	
Pro	Trp	Ile	Thr	Ala	Val	Leu	Pro	Thr	Val	Ile	Ile	Cys	Val	Met	Val
				245					250					255	
Phe	Cys	Leu	Ile	Leu	Trp	Lys	Trp	Lys	Lys	Lys	Lys	Arg	Pro	Arg	Asn
			260					265					270		
Ser	Tyr	Lys	Cys	Gly	Thr	Asn	Thr	Met	Glu	Arg	Glu	Glu	Ser	Glu	Gln
		275					280					285			
Thr	Lys	Lys	Arg	Glu	Lys	Ile	His	Ile	Pro	Glu	Arg	Ser	Asp	Glu	Ala
	290					295					300				
Gln	Arg	Val	Phe	Lys	Ser	Ser	Lys	Thr	Ser	Ser	Cys	Asp	Lys	Ser	Asp
305					310					315					320
Thr	Cys	Phe													

<210> 6

<211> 290

<212> PRT

<213> *Homo sapiens*

5 <400> 6

```

Met Arg Ile Phe Ala Val Phe Ile Phe Met Thr Tyr Trp His Leu Leu
 1           5           10           15

Asn Ala Phe Thr Val Thr Val Pro Asp Lys Leu Tyr Val Val Glu Tyr
      20           25           30

Gly Ser Asn Met Thr Ile Glu Cys Lys Phe Pro Val Glu Lys Gln Leu
      35           40           45

Asp Leu Ala Ala Leu Ile Val Tyr Trp Glu Met Glu Asp Lys Asn Ile
      50           55           60

Ile Gln Phe Val His Gly Glu Glu Asp Leu Lys Val Gln His Ser Ser
      65           70           75           80

Tyr Arg Gln Arg Ala Arg Leu Leu Lys Asp Gln Leu Ser Leu Gly Asn
      85           90           95

Ala Ala Leu Gln Ile Thr Asp Val Lys Leu Gln Asp Ala Gly Val Tyr
      100          105          110

Arg Cys Met Ile Ser Tyr Gly Gly Ala Asp Tyr Lys Arg Ile Thr Val
      115          120          125

Lys Val Asn Ala Pro Tyr Asn Lys Ile Asn Gln Arg Ile Leu Val Val
      130          135          140

Asp Pro Val Thr Ser Glu His Glu Leu Thr Cys Gln Ala Glu Gly Tyr
      145          150          155          160

Pro Lys Ala Glu Val Ile Trp Thr Ser Ser Asp His Gln Val Leu Ser
      165          170          175

Gly Lys Thr Thr Thr Thr Asn Ser Lys Arg Glu Glu Lys Leu Phe Asn
      180          185          190

Val Thr Ser Thr Leu Arg Ile Asn Thr Thr Thr Asn Glu Ile Phe Tyr
      195          200          205

Cys Thr Phe Arg Arg Leu Asp Pro Glu Glu Asn His Thr Ala Glu Leu
      210          215          220

Val Ile Pro Glu Leu Pro Leu Ala His Pro Pro Asn Glu Arg Thr His
      225          230          235          240

Leu Val Ile Leu Gly Ala Ile Leu Leu Cys Leu Gly Val Ala Leu Thr
      245          250          255

Phe Ile Phe Arg Leu Arg Lys Gly Arg Met Met Asp Val Lys Lys Cys
      260          265          270

Gly Ile Gln Asp Thr Asn Ser Lys Lys Gln Ser Asp Thr His Leu Glu
      275          280          285

Glu Thr

```

<210> 7

<211> 302

<212> PRT

<213> *Homo sapiens*

<400> 7

```

Met Arg Leu Gly Ser Pro Gly Leu Leu Phe Leu Leu Phe Ser Ser Leu
 1           5           10           15

Arg Ala Asp Thr Gln Glu Lys Glu Val Arg Ala Met Val Gly Ser Asp
          20           25           30

Val Glu Leu Ser Cys Ala Cys Pro Glu Gly Ser Arg Phe Asp Leu Asn
 35           40           45

Asp Val Tyr Val Tyr Trp Gln Thr Ser Glu Ser Lys Thr Val Val Thr
 50           55           60

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

Arg Asn Arg Ala Leu Met Ser Pro Ala Gly Met Leu Arg Gly Asp Phe
          85           90           95

Ser Leu Arg Leu Phe Asn Val Thr Pro Gln Asp Glu Gln Lys Phe His
 100          105          110

Cys Leu Val Leu Ser Gln Ser Leu Gly Phe Gln Glu Val Leu Ser Val
 115          120          125

Glu Val Thr Leu His Val Ala Ala Asn Phe Ser Val Pro Val Val Ser
 130          135          140

Ala Pro His Ser Pro Ser Gln Asp Glu Leu Thr Phe Thr Cys Thr Ser
145          150          155          160

Ile Asn Gly Tyr Pro Arg Pro Asn Val Tyr Trp Ile Asn Lys Thr Asp
          165          170          175

Asn Ser Leu Leu Asp Gln Ala Leu Gln Asn Asp Thr Val Phe Leu Asn
 180          185          190

Met Arg Gly Leu Tyr Asp Val Val Ser Val Leu Arg Ile Ala Arg Thr
 195          200          205

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

Asn Leu Thr Val Gly Ser Gln Thr Gly Asn Asp Ile Gly Glu Arg Asp
225          230          235          240

Lys Ile Thr Glu Asn Pro Val Ser Thr Gly Glu Lys Asn Ala Ala Thr
          245          250          255

Trp Ser Ile Leu Ala Val Leu Cys Leu Leu Val Val Val Ala Val Ala
          260          265          270

Ile Gly Trp Val Cys Arg Asp Arg Cys Leu Gln His Ser Tyr Ala Gly

```

	275		280		285								
Ala	Trp	Ala	Val	Ser	Pro	Glu	Thr	Glu	Leu	Thr	Gly	His	Val
	290					295					300		

<210> 8

<211> 316

5 <212> PRT

<213> *Homo sapiens*

<220>

<221> INCIERTO

<222> (233)

10 <223> "Xaa" puede ser cualquier aminoácido que se produce de manera natural

<400> 8

Met	Leu	Arg	Arg	Arg	Gly	Ser	Pro	Gly	Met	Gly	Val	His	Val	Gly	Ala	1	5	10	15
Ala	Leu	Gly	Ala	Leu	Trp	Phe	Cys	Leu	Thr	Gly	Ala	Leu	Glu	Val	Gln	20	25	30	
Val	Pro	Glu	Asp	Pro	Val	Val	Ala	Leu	Val	Gly	Thr	Asp	Ala	Thr	Leu	35	40	45	
Cys	Cys	Ser	Phe	Ser	Pro	Glu	Pro	Gly	Phe	Ser	Leu	Ala	Gln	Leu	Asn	50	55	60	
Leu	Ile	Trp	Gln	Leu	Thr	Asp	Thr	Lys	Gln	Leu	Val	His	Ser	Phe	Ala	65	70	75	80
Glu	Gly	Gln	Asp	Gln	Gly	Ser	Ala	Tyr	Ala	Asn	Arg	Thr	Ala	Leu	Phe	85	90	95	
Pro	Asp	Leu	Leu	Ala	Gln	Gly	Asn	Ala	Ser	Leu	Arg	Leu	Gln	Arg	Val	100	105	110	
Arg	Val	Ala	Asp	Glu	Gly	Ser	Phe	Thr	Cys	Phe	Val	Ser	Ile	Arg	Asp	115	120	125	
Phe	Gly	Ser	Ala	Ala	Val	Ser	Leu	Gln	Val	Ala	Ala	Pro	Tyr	Ser	Lys	130	135	140	
Pro	Ser	Met	Thr	Leu	Glu	Pro	Asn	Lys	Asp	Leu	Arg	Pro	Gly	Asp	Thr	145	150	155	160
Val	Thr	Ile	Thr	Cys	Ser	Ser	Tyr	Gln	Gly	Tyr	Pro	Glu	Ala	Glu	Val	165	170	175	
Phe	Trp	Gln	Asp	Gly	Gln	Gly	Val	Pro	Leu	Thr	Gly	Asn	Val	Thr	Thr	180	185	190	
Ser	Gln	Met	Ala	Asn	Glu	Gln	Gly	Leu	Phe	Asp	Val	His	Ser	Val	Leu	195	200	205	
Arg	Val	Val	Leu	Gly	Ala	Asn	Gly	Thr	Tyr	Ser	Cys	Leu	Val	Arg	Asn	210	215	220	
Pro	Val	Leu	Gln	Gln	Asp	Ala	His	Xaa	Ser	Val	Thr	Ile	Thr	Gly	Gln	225	230	235	240

ES 2 372 522 T3

Pro	Met	Thr	Phe	Pro	Pro	Glu	Ala	Leu	Trp	Val	Thr	Val	Gly	Leu	Ser
				245					250					255	
Val	Cys	Leu	Ile	Ala	Leu	Leu	Val	Ala	Leu	Ala	Phe	Val	Cys	Trp	Arg
			260					265					270		
Lys	Ile	Lys	Gln	Ser	Cys	Glu	Glu	Glu	Asn	Ala	Gly	Ala	Glu	Asp	Gln
		275						280				285			
Asp	Gly	Glu	Gly	Glu	Gly	Ser	Lys	Thr	Ala	Leu	Gln	Pro	Leu	Lys	His
	290					295					300				
Ser	Asp	Ser	Lys	Glu	Asp	Asp	Gly	Gln	Glu	Ile	Ala				
305					310					315					

<210> 9

<211> 276

<212> PRT

5 <213> *Homo sapiens*

<400> 9

Met	Glu	Ser	Ala	Ala	Ala	Leu	His	Phe	Ser	Arg	Pro	Ala	Ser	Leu	Leu			
1				5					10					15				
Leu	Leu	Leu	Leu	Ser	Leu	Cys	Ala	Leu	Val	Ser	Ala	Gln	Phe	Ile	Val			
			20					25					30					
Val	Gly	Pro	Thr	Asp	Pro	Ile	Leu	Ala	Thr	Val	Gly	Glu	Asn	Thr	Thr			
		35					40					45						
Leu	Arg	Cys	His	Leu	Ser	Pro	Glu	Lys	Asn	Ala	Glu	Asp	Met	Glu	Val			
	50					55					60							
Arg	Trp	Phe	Arg	Ser	Gln	Phe	Ser	Pro	Ala	Val	Phe	Val	Tyr	Lys	Gly			
65					70					75					80			
Gly	Arg	Glu	Arg	Thr	Glu	Glu	Gln	Met	Glu	Glu	Tyr	Arg	Gly	Arg	Thr			
				85					90					95				
Thr	Phe	Val	Ser	Lys	Asp	Ile	Ser	Arg	Gly	Ser	Val	Ala	Leu	Val	Ile			
			100					105					110					
His	Asn	Ile	Thr	Ala	Gln	Glu	Asn	Gly	Thr	Tyr	Arg	Cys	Tyr	Phe	Gln			
		115					120					125						
Glu	Gly	Arg	Ser	Tyr	Asp	Glu	Ala	Ile	Leu	His	Leu	Val	Val	Ala	Gly			
	130					135					140							
Leu	Gly	Ser	Lys	Pro	Leu	Ile	Ser	Met	Arg	Gly	His	Glu	Asp	Gly	Gly			
145					150					155					160			
Ile	Arg	Leu	Glu	Cys	Ile	Ser	Arg	Gly	Trp	Tyr	Pro	Lys	Pro	Leu	Thr			
				165					170					175				
Val	Trp	Arg	Asp	Pro	Tyr	Gly	Gly	Val	Ala	Pro	Ala	Leu	Lys	Glu	Val			
			180					185					190					
Ser	Met	Pro	Asp	Ala	Asp	Gly	Leu	Phe	Met	Val	Thr	Thr	Ala	Val	Ile			
		195					200					205						

Ile	Arg	Asp	Lys	Ser	Val	Arg	Asn	Met	Ser	Cys	Ser	Ile	Asn	Asn	Thr
210						215					220				
Leu	Leu	Gly	Gln	Lys	Lys	Glu	Ser	Val	Ile	Phe	Ile	Pro	Glu	Ser	Phe
225					230					235					240
Met	Pro	Ser	Val	Ser	Pro	Cys	Ala	Val	Ala	Leu	Pro	Ile	Ile	Val	Val
				245					250					255	
Ile	Leu	Met	Ile	Pro	Ile	Ala	Val	Cys	Ile	Tyr	Trp	Ile	Asn	Lys	Leu
			260					265					270		
Gln	Lys	Glu	Lys												
			275												

<210> 10

<211> 523

<212> PRT

5 <213> *Homo sapiens*

<400> 10

Met	Glu	Pro	Ala	Ala	Ala	Leu	His	Phe	Ser	Leu	Pro	Ala	Ser	Leu	Leu	1	5	10	15
Leu	Leu	Leu	Leu	Leu	Leu	Leu	Leu	Ser	Leu	Cys	Ala	Leu	Val	Ser	Ala	20	25	30	
Gln	Phe	Thr	Val	Val	Gly	Pro	Ala	Asn	Pro	Ile	Leu	Ala	Met	Val	Gly	35	40	45	
Glu	Asn	Thr	Thr	Leu	Arg	Cys	His	Leu	Ser	Pro	Glu	Lys	Asn	Ala	Glu	50	55	60	
Asp	Met	Glu	Val	Arg	Trp	Phe	Arg	Ser	Gln	Phe	Ser	Pro	Ala	Val	Phe	65	70	75	80
Val	Tyr	Lys	Gly	Gly	Arg	Glu	Arg	Thr	Glu	Glu	Gln	Met	Glu	Glu	Tyr	85	90	95	
Arg	Gly	Arg	Ile	Thr	Phe	Val	Ser	Lys	Asp	Ile	Asn	Arg	Gly	Ser	Val	100	105	110	
Ala	Leu	Val	Ile	His	Asn	Val	Thr	Ala	Gln	Glu	Asn	Gly	Ile	Tyr	Arg	115	120	125	
Cys	Tyr	Phe	Gln	Glu	Gly	Arg	Ser	Tyr	Asp	Glu	Ala	Ile	Leu	Arg	Leu	130	135	140	
Val	Val	Ala	Gly	Leu	Gly	Ser	Lys	Pro	Leu	Ile	Glu	Ile	Lys	Ala	Gln	145	150	155	160
Glu	Asp	Gly	Ser	Ile	Trp	Leu	Glu	Cys	Ile	Ser	Gly	Gly	Trp	Tyr	Pro	165	170	175	
Glu	Pro	Leu	Thr	Val	Trp	Arg	Asp	Pro	Tyr	Gly	Glu	Val	Val	Pro	Ala	180	185	190	
Leu	Lys	Glu	Val	Ser	Ile	Ala	Asp	Ala	Asp	Gly	Leu	Phe	Met	Val	Thr	195	200	205	

Thr Ala Val Ile Ile Arg Asp Lys Tyr Val Arg Asn Val Ser Cys Ser
 210 215 220
 Val Asn Asn Thr Leu Leu Gly Gln Glu Lys Glu Thr Val Ile Phe Ile
 225 230 235 240
 Pro Glu Ser Phe Met Pro Ser Ala Ser Pro Trp Met Val Ala Leu Ala
 245 250 255
 Val Ile Leu Thr Ala Ser Pro Trp Met Val Ser Met Thr Val Ile Leu
 260 265 270
 Ala Val Phe Ile Ile Phe Met Ala Val Ser Ile Cys Cys Ile Lys Lys
 275 280 285
 Leu Gln Arg Glu Lys Lys Ile Leu Ser Gly Glu Lys Lys Val Glu Gln
 290 295 300
 Glu Glu Lys Glu Ile Ala Gln Gln Leu Gln Glu Glu Leu Arg Trp Arg
 305 310 315 320
 Arg Thr Phe Leu His Ala Ala Asp Val Val Leu Asp Pro Asp Thr Ala
 325 330 335
 His Pro Glu Leu Phe Leu Ser Glu Asp Arg Arg Ser Val Arg Arg Gly
 340 345 350
 Pro Tyr Arg Gln Arg Val Pro Asp Asn Pro Glu Arg Phe Asp Ser Gln
 355 360 365
 Pro Cys Val Leu Gly Trp Glu Ser Phe Ala Ser Gly Lys His Tyr Trp
 370 375 380
 Glu Val Glu Val Glu Asn Val Met Val Trp Thr Val Gly Val Cys Arg
 385 390 395 400
 His Ser Val Glu Arg Lys Gly Glu Val Leu Leu Ile Pro Gln Asn Gly
 405 410 415
 Phe Trp Thr Leu Glu Met Phe Gly Asn Gln Tyr Arg Ala Leu Ser Ser
 420 425 430
 Pro Glu Arg Ile Leu Pro Leu Lys Glu Ser Leu Cys Arg Val Gly Val
 435 440 445
 Phe Leu Asp Tyr Glu Ala Gly Asp Val Ser Phe Tyr Asn Met Arg Asp
 450 455 460
 Arg Ser His Ile Tyr Thr Cys Pro Arg Ser Ala Phe Thr Val Pro Val
 465 470 475 480
 Arg Pro Phe Phe Arg Leu Gly Ser Asp Asp Ser Pro Ile Phe Ile Cys
 485 490 495
 Pro Ala Leu Thr Gly Ala Ser Gly Val Met Val Pro Glu Glu Gly Leu
 500 505 510
 Lys Leu His Arg Val Gly Thr His Gln Ser Leu
 515 520

<210> 11

<211> 263

<212> PRT

<213> *Homo sapiens*

5 <400> 11

Phe	His	Val	Ser	Leu	Leu	Leu	Val	Gln	Leu	Leu	Thr	Pro	Cys	Ser	Ala	1	5	10	15
Gln	Phe	Ser	Val	Leu	Gly	Pro	Ser	Gly	Pro	Ile	Leu	Ala	Met	Val	Gly	20	25	30	
Glu	Asp	Ala	Asp	Leu	Pro	Cys	His	Leu	Phe	Pro	Thr	Met	Ser	Ala	Glu	35	40	45	
Thr	Met	Glu	Leu	Lys	Trp	Val	Ser	Ser	Ser	Leu	Arg	Gln	Val	Val	Asn	50	55	60	
Val	Tyr	Ala	Asp	Gly	Lys	Glu	Val	Glu	Asp	Arg	Gln	Ser	Ala	Pro	Tyr	65	70	75	80
Arg	Gly	Arg	Thr	Ser	Ile	Leu	Arg	Asp	Gly	Ile	Thr	Ala	Gly	Lys	Ala	85	90	95	
Ala	Leu	Arg	Ile	His	Asn	Val	Thr	Ala	Ser	Asp	Ser	Gly	Lys	Tyr	Leu	100	105	110	
Cys	Tyr	Phe	Gln	Asp	Gly	Asp	Phe	Tyr	Glu	Lys	Ala	Leu	Val	Glu	Leu	115	120	125	
Lys	Val	Ala	Ala	Leu	Gly	Ser	Asn	Leu	His	Val	Glu	Val	Lys	Gly	Tyr	130	135	140	
Glu	Asp	Gly	Gly	Ile	His	Leu	Glu	Cys	Arg	Ser	Thr	Gly	Trp	Tyr	Pro	145	150	155	160
Gln	Pro	Gln	Ile	Gln	Trp	Ser	Asn	Ala	Lys	Gly	Glu	Asn	Ile	Pro	Ala	165	170	175	
Val	Glu	Ala	Pro	Val	Val	Ala	Asp	Gly	Val	Gly	Leu	Tyr	Glu	Val	Ala	180	185	190	
Ala	Ser	Val	Ile	Met	Arg	Gly	Gly	Ser	Gly	Glu	Gly	Val	Ser	Cys	Ile	195	200	205	
Ile	Arg	Asn	Ser	Leu	Leu	Gly	Leu	Glu	Lys	Thr	Ala	Ser	Ile	Ser	Ile	210	215	220	
Ala	Asp	Pro	Phe	Phe	Arg	Ser	Ala	Gln	Pro	Trp	Ile	Ala	Ala	Leu	Ala	225	230	235	240
Gly	Thr	Leu	Pro	Ile	Leu	Leu	Leu	Leu	Leu	Ala	Gly	Ala	Ser	Tyr	Phe	245	250	255	
Leu	Trp	Arg	Gln	Gln	Lys	Glu										260			

<210> 12

<211> 584

<212> PRT

<213> *Homo sapiens*

5 <400> 12

Met	Lys	Met	Ala	Ser	Ser	Leu	Ala	Phe	Leu	Leu	Leu	Asn	Phe	His	Val	1	5	10	15
Ser	Leu	Phe	Leu	Val	Gln	Leu	Leu	Thr	Pro	Cys	Ser	Ala	Gln	Phe	Ser	20	25	30	
Val	Leu	Gly	Pro	Ser	Gly	Pro	Ile	Leu	Ala	Met	Val	Gly	Glu	Asp	Ala	35	40	45	
Asp	Leu	Pro	Cys	His	Leu	Phe	Pro	Thr	Met	Ser	Ala	Glu	Thr	Met	Glu	50	55	60	
Leu	Arg	Trp	Val	Ser	Ser	Ser	Leu	Arg	Gln	Val	Val	Asn	Val	Tyr	Ala	65	70	75	80
Asp	Gly	Lys	Glu	Val	Glu	Asp	Arg	Gln	Ser	Ala	Pro	Tyr	Arg	Gly	Arg	85	90	95	
Thr	Ser	Ile	Leu	Arg	Asp	Gly	Ile	Thr	Ala	Gly	Lys	Ala	Ala	Leu	Arg	100	105	110	
Ile	His	Asn	Val	Thr	Ala	Ser	Asp	Ser	Gly	Lys	Tyr	Leu	Cys	Tyr	Phe	115	120	125	
Gln	Asp	Gly	Asp	Phe	Tyr	Glu	Lys	Ala	Leu	Val	Glu	Leu	Lys	Val	Ala	130	135	140	
Ala	Leu	Gly	Ser	Asp	Leu	His	Ile	Glu	Val	Lys	Gly	Tyr	Glu	Asp	Gly	145	150	155	160
Gly	Ile	His	Leu	Glu	Cys	Arg	Ser	Thr	Gly	Trp	Tyr	Pro	Gln	Pro	Gln	165	170	175	
Ile	Lys	Trp	Ser	Asp	Thr	Lys	Gly	Glu	Asn	Ile	Pro	Ala	Val	Glu	Ala	180	185	190	
Pro	Val	Val	Ala	Asp	Gly	Val	Gly	Leu	Tyr	Ala	Val	Ala	Ala	Ser	Val	195	200	205	
Ile	Met	Arg	Gly	Ser	Ser	Gly	Gly	Gly	Val	Ser	Cys	Ile	Ile	Arg	Asn	210	215	220	
Ser	Leu	Leu	Gly	Leu	Glu	Lys	Thr	Ala	Ser	Ile	Ser	Ile	Ala	Asp	Pro	225	230	235	240
Phe	Phe	Arg	Ser	Ala	Gln	Pro	Trp	Ile	Ala	Ala	Leu	Ala	Gly	Thr	Leu	245	250	255	
Pro	Ile	Ser	Leu	Leu	Leu	Leu	Ala	Gly	Ala	Ser	Tyr	Phe	Leu	Trp	Arg	260	265	270	
Gln	Gln	Lys	Glu	Lys	Ile	Ala	Leu	Ser	Arg	Glu	Thr	Glu	Arg	Glu	Arg	275	280	285	
Glu	Met	Lys	Glu	Met	Gly	Tyr	Ala	Ala	Thr	Glu	Gln	Glu	Ile	Ser	Leu	290	295	300	

Arg Glu Lys Leu Gln Glu Glu Leu Lys Trp Arg Lys Ile Gln Tyr Met
 305 310 315 320
 Ala Arg Gly Glu Lys Ser Leu Ala Tyr His Glu Trp Lys Met Ala Leu
 325 330 335
 Phe Lys Pro Ala Asp Val Ile Leu Asp Pro Asp Thr Ala Asn Ala Ile
 340 345 350
 Leu Leu Val Ser Glu Asp Gln Arg Ser Val Gln Arg Ala Glu Glu Pro
 355 360 365
 Arg Asp Leu Pro Asp Asn Pro Glu Arg Phe Glu Trp Arg Tyr Cys Val
 370 375 380
 Leu Gly Cys Glu Asn Phe Thr Ser Gly Arg His Tyr Trp Glu Val Glu
 385 390 395 400
 Val Gly Asp Arg Lys Glu Trp His Ile Gly Val Cys Ser Lys Asn Val
 405 410 415
 Glu Arg Lys Lys Gly Trp Val Lys Met Thr Pro Glu Asn Gly Tyr Trp
 420 425 430
 Thr Met Gly Leu Thr Asp Gly Asn Lys Tyr Arg Ala Leu Thr Glu Pro
 435 440 445
 Arg Thr Asn Leu Lys Leu Pro Glu Pro Pro Arg Lys Val Gly Ile Phe
 450 455 460
 Leu Asp Tyr Glu Thr Gly Glu Ile Ser Phe Tyr Asn Ala Thr Asp Gly
 465 470 475 480
 Ser His Ile Tyr Thr Phe Pro His Ala Ser Phe Ser Glu Pro Leu Tyr
 485 490 495
 Pro Val Phe Arg Ile Leu Thr Leu Glu Pro Thr Ala Leu Thr Ile Cys
 500 505 510
 Pro Ile Pro Lys Glu Val Glu Ser Ser Pro Asp Pro Asp Leu Val Pro
 515 520 525
 Asp His Ser Leu Glu Thr Pro Leu Thr Pro Gly Leu Ala Asn Glu Ser
 530 535 540
 Gly Glu Pro Gln Ala Glu Val Thr Ser Leu Leu Leu Pro Ala His Pro
 545 550 555 560
 Gly Ala Glu Val Ser Pro Ser Ala Thr Thr Asn Gln Asn His Lys Leu
 565 570 575
 Gln Ala Arg Thr Glu Ala Leu Tyr
 580

<210> 13

<211> 526

<212> PRT

<213> *Homo sapiens*

5 <400> 13

Met Ala Val Phe Pro Ser Ser Gly Leu Pro Arg Cys Leu Leu Thr Leu
 1 5 10 15
 Ile Leu Leu Gln Leu Pro Lys Leu Asp Ser Ala Pro Phe Asp Val Ile
 20 25 30
 Gly Pro Pro Glu Pro Ile Leu Ala Val Val Gly Glu Asp Ala Glu Leu
 35 40 45
 Pro Cys Arg Leu Ser Pro Asn Ala Ser Ala Glu His Leu Glu Leu Arg
 50 55 60
 Trp Phe Arg Lys Lys Val Ser Pro Ala Val Leu Val His Arg Asp Gly
 65 70 75 80
 Arg Glu Gln Glu Ala Glu Gln Met Pro Glu Tyr Arg Gly Arg Ala Thr
 85 90 95
 Leu Val Gln Asp Gly Ile Ala Lys Gly Arg Val Ala Leu Arg Ile Arg
 100 105 110
 Gly Val Arg Val Ser Asp Asp Gly Glu Tyr Thr Cys Phe Phe Arg Glu
 115 120 125
 Asp Gly Ser Tyr Glu Glu Ala Leu Val His Leu Lys Val Ala Ala Leu
 130 135 140
 Gly Ser Asp Pro His Ile Ser Met Gln Val Gln Glu Asn Gly Glu Ile
 145 150 155 160
 Cys Leu Glu Cys Thr Ser Val Gly Trp Tyr Pro Glu Pro Gln Val Gln
 165 170 175
 Trp Arg Thr Ser Lys Gly Glu Lys Phe Pro Ser Thr Ser Glu Ser Arg
 180 185 190
 Asn Pro Asp Glu Glu Gly Leu Phe Thr Val Ala Ala Ser Val Ile Ile
 195 200 205
 Arg Asp Thr Ser Thr Lys Asn Val Ser Cys Tyr Ile Gln Asn Leu Leu
 210 215 220
 Leu Gly Gln Glu Lys Lys Val Glu Ile Ser Ile Pro Ala Ser Ser Leu
 225 230 235 240
 Pro Arg Leu Thr Pro Trp Ile Val Ala Val Ala Val Ile Leu Met Val
 245 250 255
 Leu Gly Leu Leu Thr Ile Gly Ser Ile Phe Phe Thr Trp Arg Leu Tyr
 260 265 270
 Asn Glu Arg Pro Arg Glu Arg Arg Asn Glu Phe Ser Ser Lys Glu Arg
 275 280 285
 Leu Leu Glu Glu Leu Lys Trp Lys Lys Ala Thr Leu His Ala Val Asp
 290 295 300
 Val Thr Leu Asp Pro Asp Thr Ala His Pro His Leu Phe Leu Tyr Glu
 305 310 315 320
 Asp Ser Lys Ser Val Arg Leu Glu Asp Ser Arg Gln Lys Leu Pro Glu

				325						330						335			
Lys	Thr	Glu	Arg	Phe	Asp	Ser	Trp	Pro	Cys	Val	Leu	Gly	Arg	Glu	Thr				
			340					345					350						
Phe	Thr	Ser	Gly	Arg	His	Tyr	Trp	Glu	Val	Glu	Val	Gly	Asp	Arg	Thr				
		355					360					365							
Asp	Trp	Ala	Ile	Gly	Val	Cys	Arg	Glu	Asn	Val	Met	Lys	Lys	Gly	Phe				
	370					375					380								
Asp	Pro	Met	Thr	Pro	Glu	Asn	Gly	Phe	Trp	Ala	Val	Glu	Leu	Tyr	Gly				
385					390					395					400				
Asn	Gly	Tyr	Trp	Ala	Leu	Thr	Pro	Leu	Arg	Thr	Pro	Leu	Pro	Leu	Ala				
				405					410					415					
Gly	Pro	Pro	Arg	Arg	Val	Gly	Ile	Phe	Leu	Asp	Tyr	Glu	Ser	Gly	Asp				
			420					425					430						
Ile	Ser	Phe	Tyr	Asn	Met	Asn	Asp	Gly	Ser	Asp	Ile	Tyr	Thr	Phe	Ser				
		435					440					445							
Asn	Val	Thr	Phe	Ser	Gly	Pro	Leu	Arg	Pro	Phe	Phe	Cys	Leu	Trp	Ser				
	450					455					460								
Ser	Gly	Lys	Lys	Pro	Leu	Thr	Ile	Cys	Pro	Ile	Ala	Asp	Gly	Pro	Glu				
465					470				475						480				
Arg	Val	Thr	Val	Ile	Ala	Asn	Ala	Gln	Asp	Leu	Ser	Lys	Glu	Ile	Pro				
				485				490						495					
Leu	Ser	Pro	Met	Gly	Glu	Glu	Ser	Ala	Pro	Arg	Asp	Ala	Asp	Thr	Leu				
			500					505					510						
His	Ser	Lys	Leu	Ile	Pro	Thr	Gln	Pro	Ser	Gln	Gly	Ala	Pro						
		515					520					525							

<210> 14

<211> 7819

<212> ADN

5 <213> *Homo sapiens*

<220>

<221> exón

<222> (4599)..(4652)

<223> parte codificante del exón 1

10 <220>

<221> incierto

<222> (145)..(292)

<223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 <222> (1977)..(2459)
 5 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 <222> (2465)..(2576)
 <223> "N" puede ser A, C, T o G
 10 <220>
 <221> incierto
 <222> (2593)..(3410)
 <223> "N" puede ser A, C, T o G
 <220>
 15 <221> incierto
 <222> (3490)..(3533)
 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 20 <222> (6111)..(6194)
 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 <222> (6836)..(6870)
 25 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 <222> (7091)..(7112)
 <223> "N" puede ser A, C, T o G
 30 <220>
 <221> incierto
 <222> (7118)..(7420)
 <223> "N" puede ser A, C, T o G
 <220>
 35 <221> incierto
 <222> (7554)..(7575)
 <223> "N" puede ser A, C, T o G

<220>

<221> incierto

<222> (7694)..(7718)

<223> "N" puede ser A, C, T o G

5 <400> 14

```

ataagaagct gaattaaggt gatggcagtg ggggtggaaga aaggagagcc accatgcaaa 60
aagtatccag gagggagaat taacaggact aggggatggg ccatatttgc aagatgagaa 120
atgcagaggt ctaagattct agctnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 180
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 240
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nngcataaac 300
catattttcc ccagaggagg attagtagga aaggaagctg ctggttgga agtatcttta 360
tagcagtgtc tggttcctcg tttgctcaag gggacagtgt gccaggaaag tccccgtgga 420
agggcaagga agaaggggaa gttaaagcca gtggcaggtg atccaagaat cttttctggt 480
gctagagcta tggtacatgc tgtcctttca tgctctaaaa ataagagtgc tggcaagtgc 540

```

caggcctggt ggtgcagctt aagatgatac ctttcttgga tatatatgca tctgaataag 600
 gaaggctatc ttctgggtcaa gctaagggtat gccatgagca tttccctgtg gaaagcactt 660
 aattctgttc ccagttgtta cctgctgtaa gatctccctt tctaaaataa aaacaagaat 720
 acagctcact gaggacctta catttccctc tagctactga ctcatctctc ttctcctttt 780
 tatagcactc ttcttgagag agttgcctat atttgttgcc acatctttac ccattctctt 840
 ttgaacctat tcaagctttc atctgtacaa aactcactga tactgtgctt gtcaggatca 900
 tccatgacct ccatactgct aaatgcaact ctcaagagta tttggctcta ctgatcactc 960
 ctttgtagca ctgtgtttta aaatataggt tttattatta tttaggtatg gtgaggccaa 1020
 tatatcagga aatgactgtc gttgaaaaaa gtatgttgta ctacagatc ccaagagaag 1080
 gggggcacac catgccacaa agggccacat ggggaagcac cagggtcagc caggaggtgg 1140
 gtgggggggtg cgcaagatct ttattgtggt ttcaacagga agaaatgggt gaagcagggt 1200
 gagtggattt aggattagct gatataaata atttcagcag gctctggggc ataggggctg 1260
 tccctagtct tctggtaactt ggccctgggg tgattaaggc agttgcatag tgttggaat 1320
 gtgaaagccc ccaataaatg aggcagttgt gggatatggc tctgaaatgg gttggtttgc 1380
 atttgaaagg tgtgctcatg ggcaagtgtt ttactctctc ttagagggtta gaattggcta 1440
 accctgggag cggcagctcc ttccagggtca gcaaggcccc aggtgtcaaa gcatcagaat 1500
 acagaaaata aaatgcatgg ataatacaca ctgccatttg cctttgtacc cttcctttca 1560
 atcttctctg ctggtgaccg ctcttcacaa agatctataa atgttggaat acccatgtc 1620
 tcagtctctg ggcaactctt ttctatctc tctgtagggt atgtaatgca gatatccatg 1680
 actttaaatc tttaacactt ctgcattgat gactcctaaa tttacatctc taccccaact 1740
 gcctactaaa cacctccact tggctatcta ataggcattt caaaccaaat ctacaacaaa 1800
 cgtaactctt tttcccttcc cttaatttgc ttctcccca gccttctcca ttttaataaa 1860
 cagcatctcc attgccttag tgactcaagc cccaaactta ggaattttcc cagatttccc 1920
 tctttttctc aaactatata tctagcctgt cagcagttcc cttcaggtct ttttcnnnn 1980
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2040
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2100
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2160
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2220
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2280
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2340

nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 2400
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnna 2460
 ctaannnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 2520
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnngctc 2580
 catttatatt tannnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 2640
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 2700
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 2760
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 2820
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 2880
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 2940
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 3000
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 3060
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 3120
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 3180
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 3240
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 3300
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 3360
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn ttgtgtttta 3420
 aatatatata cacacttaga cacatataac cctctttogt atatcaatta tacttttaata 3480
 aagctgttgn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnctcgcca 3540
 acagaccact tctcaccctt actagcccct ctactctag ctatcagcat ttctgcaaac 3600
 acattcctaa ccgatctcac tgcttgtaat cttgccagca acctctccct ctgagcaata 3660
 gtctattgcc tacaccaag cttagtgtgc tottaatgat gtaaagtagg ttctatcatt 3720
 ctcttgacct aaaccctcca ctgcttttca tcacactcag agcagctctg ctgttgccctg 3780
 atttagatgt atggctccaa cagatttccc ctgaagaaat gattccatgg ctgataaaag 3840
 ttggaaagcc tctcagttt cagaccatta tcagattagc tgtgtgctct gtccctttcc 3900
 tcaaccataa gaagtccatg gataaagaaa gcttcagagt aaaggagaaa gcatgggagg 3960
 tacagcagga ccaagggtgg gcattcgcag cccccacct catcagagcc agttccctac 4020
 tctccctgtc taaacctctt agtaagaggt agttcaagag aggggcaaac tcaattccag 4080
 cactcaaaag cacttgacta ctttgcctcag tcaactagca agtattttatt gagaatgtag 4140
 ctctgttcta tggagtotta ttttcaagtg tcagactccc agacatccag tccaggtaaa 4200

gaagatggtg tccattatc atttgacaaa caaagttggg gttcaagggc cagctattga 4260
 aaaaagctat ggaaagcttc atgagacgtg caggtaactg ccaatatgtg tggttcacaa 4320
 ggactggttc atattcagaa acggccatta gaaaaggaag aagaacttct catttggtt 4380
 tataaagagt gtcttgttta ctcttaattt atatcttctc ttctccagga aatcaacct 4440
 taacttctcc tcccagctcc actctaccat ggtctgtcac ctccccaaa tgatttgta 4500
 ttccctgtt ttcaaaagtg aacaaagaac caaagaccca gcaaagtttc acaaggccct 4560
 gagactttca attgtctatt tcagatcaaa tacagaac atg atc ttc ctg cta 4616
 atg ttg agc ctg gaa ttg cag ctt cac cag ata gca ggtaagaaag 4662
 gacaaagga gaggttaag aaagaagagc aggtggtggt tcctagccaa agccaaaaat 4722
 gagaatgtg ccctcaggct gagggcttct tttgagagga cgtatgattt ctgggtctatt 4782
 ccaagcacca caaaaaaaaa aagagtcccc atggtggctt atacatgcca atgtccctat 4842
 ctgacagaaa cgggtgactga gaatattgct ccactatcct ccactatcca gtgagggtta 4902
 tgacaagaag acaggatcac tcagaccatg taaatctaaa ctgatacaag agggcagggg 4962
 ttgagttccc ttaaagggtga gatgccaagc agctgtcccc ttctttctg gcaggagag 5022
 taaggagaca atggccaggg aacaccgtta ctctaaagat aatgtcttga agacattctg 5082
 catattatta gttgtttctg tgagtttctt ttttgaaaag caacaatagc agccgttgg 5142
 cattcatacc ttaatgtggt ttaactgagtc ttctaaaaac ccaaataaac aatgaacctt 5202
 aaggctatcc ctttggtactt gaagaaagga cttctatttg aggatgaggg tgagcagaaa 5262
 gaaagcagt ttcacagttg gttgttctcc tggggaaggt agttcagacc attcgagggt 5322
 gtagttagaa ccatgagtg actattttg atgaacacca ggagctaaga gagtaacata 5382
 gaggtgtgga cagaggatta agtcctcaag acaatagccc cagcccatg ggaaatcatc 5442
 tttctgtca tgattgagaa ataatggctc ccttggcact tgataacctt tcgaagagct 5502
 ttctctccc tactagctgg ttccagatca ctcttcccc agtcacattc ctctactca 5562
 cttgagctgc ccagcctgg ctggcactag agacatgcac ttggggccct cctcaaagga 5622
 agaccctgag atattctgt tactctact ctgctcctgc ctgcagggcc agctaaagga 5682
 acttttcatg ttttctttgc aaggaacct gcctggctgg catttttagag acaagcaaaa 5742
 ggggcaataa cttccttgct acaaaacagc ttcaagtttc catagagtga taagggaaat 5802
 gagggccaaa agacactgtt cccatcctg tggcaggact gggggcttca ggagaaaact 5862
 tggggaatgt gtaacctctg tgggtttgta gcttaaaaac actgagatcc tgggttttct 5922
 gtctttgtt tttgccttt ctcttaggaa aggagtgagc tagggtgaca aggggcaaca 5982

ttttttatcc ctcatgggct cttttctacag aggaaggatc tttttctcta agataatcag 6042
 cacaagacaa tgaagatagg cactagctcc cagttaggta tactaatggg gcaaaaaggaa 6102
 gagcattttn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6162
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnattttgca ggtgaggga ctagaactca 6222
 gaaaaggtag ttaatttccc caagattaca tagttattag gtgacaccgg caagatttca 6282
 acaaagctaa tgtcctttct actttactgt gctaccatga tgatggtaat caaaaatggc 6342
 agacaacca taaatottcc aactttggaa taggtttttg cactgaagtc tgaatatgga 6402
 tacgtattga atgtttattc tggatattca cagaatcaaa aaatatgtgt aatgaattat 6462
 gttgctgaat taactgaaag gaaagtaaaa atgtagcgct ttctcatttt cttcacgaat 6522
 ttggaattct tttctgcttt ccactatgca gataacatca gttcagacaa atattaaata 6582
 cctacctaaa ttagaatgcc ttctcctcat gggatttttt taaaatcttg tcatttcatg 6642
 tctctttaat taaagagttt tgatttcaga ggagggtacc tgcaaaaagaa aacaacaaaa 6702
 aaactaaagg atctgagaaa taattagtgt ttacttctgg ggaggggagg aggtctggga 6762
 tgggggtaaa aaggatagtc ttatctatta tgtatattca ggtttttggt ttttacaaga 6822
 agcatgtatt aggnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnca agaaatttct 6882
 catataaaaa tatgaaagta atcagactgc aacactcagt gcctgagaca gagctacagc 6942
 tatcaggggtg tccagacaga cagaagatta cattttcttc cttgctcctt gtacagcccc 7002
 agacctgcat gcttcattga aaagaaaaga agatacctga attaaatcaa tgtgatgctt 7062
 agtaccctat cagtgcacat ttcttttcnn nnnnnnnnnn nnnnnnnnnn cacttnnnnn 7122
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7182
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7242
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7302
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7362
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnca 7422
 cttatggcgg tattctcagt cattacaaat aaataaaaac aatccatag ccctggagaa 7482
 tttgattcca ggagtaggtc tagaagaact tcaactggag aatggataga gaaatcatgg 7542
 tatatttgca gnnnnnnnnn nnnnnnnnnn nnngatagca tgtgaataaa ttaattacaa 7602
 aaacatatga ctacatctat tattatatag catgtagata aattacaaaa acatgtaact 7662
 acatctatga atcttagagc ataattattga gnnnnnnnnn nnnnnnnnnn nnnnnnggcc 7722
 agaagataac acatagcaca atgtcctttt cataaataaa tatattgctt aagcatacct 7782
 tatatataga agataaagct taaaaagtaa agaagag 7819

<210> 15
 <211> 650
 <212> ADN
 <213> *Homo sapiens*

5 <220>
 <221> incierto
 <222> (20)..(175)
 <223> "N" puede ser A, C, T o G
 <400> 15

```
attatcactt atgagggtgn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 120
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnctgcc 180
ctttatTTTT tcatgaaaga aattgctgaa gaggactaaa agaagtttta gtaagcattc 240
aataaatgta tgttctttat agtttccaaa tcagcaaata tagacatcct gcatttttta 300
ggagatttat atattttatt ggacatgctg taattttatt aaccacttcc ctgttggtag 360
acattatttc cattttcttc tgctagatta atgcttgaaa aaaatgtgtg cctcctaaag 420
actgtgatga aagttgcctc tgaataaaaac tcaaacaaat cattaatcat taactctttc 480
cttacttgta tgctcttttg atgctctact gtgttatcta taaaataaag tttgaagtga 540
aaaattaggg taaaacattt tatatcattt ttaaaggata tatacatgga tgtacttaca 600
tatgcatggt taaatttata taccataaca tttatttctt tttttaaaaa 650
```

10 <210> 16
 <211> 9179
 <212> ADN
 <213> *Homo sapiens*

15 <220>
 <221> exón
 <222> (5211)..(5516)
 <223> exón 2
 <220>

20 <221> incierto
 <222> (1054)..(1198)
 <223> "N" puede ser A, C, T o G
 <220>

<221> incierto
 <222> (1297)..(1430)
 <223> "N" puede ser A, C, T o G
 <220>
 5 <221> incierto
 <222> (1482)..(1658)
 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 10 <222> (3286)..(3529)
 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 <222> (3580)..(3620)
 15 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 <222> (4193)..(4416)
 <223> "N" puede ser A, C, T o G
 20 <220>
 <221> incierto
 <222> (4506)..(4682)
 <223> "N" puede ser A, C, T o G
 <220>
 25 <221> incierto
 <222> (6166)..(6223)
 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 30 <222> (6635)..(6803)
 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 <222> (7946)..(8351)
 35 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto

<222> (8578)..(8609)

<223> "N" puede ser A, C, T o G

<220>

<221> incierto

5 <222> (8870)..(9178)

<223> "N" puede ser A, C, T o G

<400> 16

```
tatatttggt ttttttctta catttttatt tcaaaatcta aggacatctt ataaccacaga 60
aatatttttt ataccttgct atgtcttaga ggaaagagcc accccagctt tttttcattg 120
atgtttttct tctctcttcg tactccagag gtagatgaaa accagagggc cacaatgacc 180
atggtgatgc ctgaggtcat tctggggcac agacctcagc ctaggttact ccaattcgcc 240
tatcttttaga tccaaaacta ccctgctgac tgctgagata aacaaaggag aataatcagg 300
ttggggaaaag gatttctatg cgaagacatg tctccatgca gtcctcctac actgagcaga 360
gcatgagtca ggtgcttaga gcaggatttt gtcctaaacc aggaacttca gagttttctg 420
aagaatgtgg ctatgtaaag cccccccca cccaccctt acttctcaag tacattacgt 480
```

ggcaagtctg aaaaaactta cacttctgtt gttaaatgtg ggggataaaa tataaactta 540
 gtttcaagag gaagctatct tgggaggtaa tgcaaataat togttgtgtg tttcctgaat 600
 aagtgcagag tgctgactac cattgatgct tcattgcaat aaaatgcaaa gctcccccaa 660
 gaatttttga aatgcatcaa gctaggtgtt ctaatctagc aaaaggacct gcatacatga 720
 atttttcatg cttttgcaa gtcttttgcc ctttagttta gtttaaggcc ccacatgaat 780
 ggaaagcctg tgttgtcagc ttaattttgt agttgtggaa accttccagt tttctccttt 840
 gtctaatacc ttcaggagtt caatcctagg ttgaagotta atttaataac catgtggcat 900
 gtaaagtaga aaacaaaaca tcttttctt agcatacagc aaaaaaaaaa aaaaaactca 960
 ctcatggatg tagtgtacac atgccagtgg atatatagtc ataactgcag tcattggtag 1020
 cacagaaata aatgtgcatt gaagacacag agannnnnnn nnnnnnnnnn nnnnnnnnnn 1080
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1140
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnngg 1200
 gctagcccag tgtctgactt actgtgttta agaaatatca actattacgc tacttcccag 1260
 tgacagtcca aatgcagacc agtgttataa ctctacnnnn nnnnnnnnnn nnnnnnnnnn 1320
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1380
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn aggcaaggtg 1440
 gtaaaagacg cttgcacgtg caaatgttac tttgtgtaac tnnnnnnnnn nnnnnnnnnn 1500
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1560
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1620
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnct tagtgaatct aatttgcagc 1680
 tgggcttgag aaaaaacctc ttcatagaat tgtttgcac agtgtcttga ttgcctctgt 1740
 aacttacaat aagcaagaat gtttcaggat ttcaaaaatc tattgcattg cctaaacctc 1800
 ttatttttga tggagtaatc aagctcaaag tttgcatgtc ttagaaaactt tacttggggc 1860
 aaaattagac caagtaacaa ttaatcttct aggtattctg agctattcag acatatgatt 1920
 catgtttgct aattgctctt ttctcttgta aatattagct gaaaaatgtc acctgtctga 1980
 caagtagcat attttatgcc tatcactcct ggcacgcatt cttacaaggc agacaggaaa 2040
 aataggaaga aaatggactt ttatcaaagg ccagggcagt aaagagggga gttotgctgt 2100
 aagctaaggg gagttccaga ggaagttata ggcgttcctt ttcttatgac aagaaagcat 2160
 agtgcagtaa ataaatttgc taaatagatt caacagtctc taccocaaagt catctattta 2220
 attcttgttg ttatgcagac tcagcaacta accttcttg taagccocat tttcttccct 2280

gtttcctggt tatcaaagt aattaaaca gagaagtatt atagaagagt aaaagtagta 2340
 ggtaattcct gaacttggca tatgattact acatatattga tgaatagttg aatattattc 2400
 ttcaaggaca gattggattt ggtatcaggt ggctctgcat taagttataa gggacttaat 2460
 aactcaagta ttttaaggacg gcttccatca taaagggatc tgcccttaag aggggtcccat 2520
 tatggagatt ctgaggtgag agctattcca agtgtgcagt ggattaaaat aaaagaatca 2580
 tacaggaaat ctctttttac atgccttatt ccagggtcct tgcaacctgg cacagcaagt 2640
 gcagatatga ttagcattgt tttacacatg tacactcacc ttatagccct gccctgtgc 2700
 ccctcctgca caaaagaatg ctgggcacac gtgaactcct ctctgtagaa aggcacatta 2760
 atgttctagc catgggttaa acagggatag aggcaagcca aaaatgtogg tcatttgaaa 2820
 taaatctcaa gtttgtgcat atcactatca agtgtgctgt gtggcaatta agaatgcaa 2880
 tttgtgtgat cacaggcaag ttgcagtttg atgaaaggaa agcagaggtg aatatataac 2940
 cagggtcatc ctttctttct ccctctctct ctttctgtca tttatttgcc aagctcttaa 3000
 ctagaacttg ctatgtgcta ggtactggat atatcaaagc aaactcagcc tggctcttgc 3060
 cttcaaagat ttgcaggata gtgggaagaa aaacttgaat cagaggacat ctgcagtggg 3120
 aatcattcaa gcagcagaaa acccaaaagt tacttatact gtgaaatctg atcagagaat 3180
 ggactgtcct ggtagtaaa ataccctgga ggataaagat tggccatgca ttccacatat 3240
 gaattaccac tttccaaga attaaaacat ggtacgaaag aaaggnnnnn nnnnnnnnnn 3300
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3360
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3420
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3480
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn aacaacttct 3540
 ggaatagcta atgcttagaa gcagctccca aatatttgtn nnnnnnnnnn nnnnnnnnnn 3600
 nnnnnnnnnn nnnnnnnnnn gcaagctcta ctgaacataa tttgatctaa tcttctgtga 3660
 ttattcagaa actacttcaa gattttccta tacctccatc ataataaata cccattcatt 3720
 aatgatggaa gcagcctaatt tttgtcattt ttcacacttt attgatgtaa cactaccttt 3780
 actagtttgg ccaactcctta tgcttttttt atagaactat ttagatcaat tcaactttta 3840
 aaaaataaag ccacataccc ctgtggtaga tgaaaaaaca gtatcatttg cactggtaaa 3900
 tagagaatag gaagaaaaat aaatgcagtg aaaataaagc agtggtatca aatcctaccc 3960
 agatactgtt atctaccogg aagcttcctg tttgattaaa aggaaaaata gccagtgtta 4020
 gaggtgtgga agtctagttg aaattatatg caattgaagg attaaaatag aattgaaaag 4080
 ggaataaatt cctctctgaa taatttaact ccctttaggc tttgattctg cctcatctaa 4140

aatcatctta catactttota gtggcgtgtc cctcacattt tggtaaactc tgnnnnnnnn 4200
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4260
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4320
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4380
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnccat ttttcctttc cctttattgt 4440
 cagaaaatag aaagcatcta cagtgggctt gtatgatgtg gtggttagaa atacctgatc 4500
 tgattnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4560
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4620
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4680
 nnaccatcac tggtagagag aagtgatacc tggcacaaaa atatatggat taatcaatat 4740
 ggattgaggg aaacaaacct ggagaatagg atgtgaaggt atttaagtaa catgagctca 4800
 gaccttgatg gtagggaagt cgaaaggaag cattttgttc ttatatgaca gatgacctgg 4860
 aatgactgca gggcttgggg ggtcagggac tggaggtggg agaggcctct gagagcaagc 4920
 agtgctgtcc accagaagct cttgctgggg tgcccagaga ggagcaaagg gcagtcagct 4980
 gcacaggagg gaatgtttgg aggagagagc cacctcagat cagcgggtca agaatccac 5040
 tcttgcccag atggatgggg caaaggagaa aaaggattcg ccacgggaat gtccagataa 5100
 gacaggtgcc ttttgaaaaa tgggggtgag atgggtotca ggttacactt cgtaagaact 5160
 ggaatgtaaa gtaaaggcag acaatgacaa aatatcttgt tttcttttca gct tta 5216
 ttc⁴aca gtg aca gtc cct aag gaa ctg tac ata ata gag cat ggc agc 5264
 aat gtg acc ctg gaa tgc aac ttt gac act gga agt cat gtg aac ctt 5312
 gga gca ata aca gcc agt ttg caa aag gtg gaa aat gat aca tcc cca 5360
 cac cgt gaa aga gcc act ttg ctg gag gag cag ctg ccc cta ggg aag 5408
 gcc tcg ttc cac ata cct caa gtc caa gtg agg gac gaa gga cag tac 5456
 caa tgc ata atc atc tat ggg gtc gcc tgg gac tac aag tac ctg act 5504
 ctg aaa gtc aaa ggtgagtgg gtcaaggact agaatccatg gaagctctct 5556
 ccaacagagg atctgcaagt cacagaaacc cattaaagggt agctcaagca aaaacaagca 5616
 ggctgctttt aaggagacag ctatttcaga gaaaatgaaa gcatctgctc ggaaataatt 5676
 tttgacatct gagtacaaag cagccgaagt acaagtgaaa gggggtagga cctataggaa 5736
 taaaatggga ctggaggaag ccaggaaaat tagtccctga aatgtgggag ggtatgaaaa 5796
 ataagctttg cctaattcac aattctocca tggaacatcc ctgacttgat tattaagata 5856

ctctttttca atagttttata ccttgaatcc agagttttta aaaccatggt ttgccgccca 5916
 ttcatggatt aaaatatcaa tttagttagt agcaaccaga tgcacgtttc ccgcccttta 5976
 aaaaataatg tatagaagag aatagacaga gtagatcaga cgatatcaca gagtaggact 6036
 gagtactgta aaactaatit ctgagggacg tgtgtgtgtg tgtgcgtgtt gggtcatggt 6096
 ataaattttt tttttcttac tttggatcat aaaaagttac aagtttgga aacactgctc 6156
 aaatgcaagn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6216
 nnnnnnnaga gtcccatgaa gtctatagct gtccctattc ctctggattc agggatctct 6276
 ccaactccagc acaattgaaa atctaaatat aaagagaatc ttcacactct tgtttgttct 6336
 agaaaagggtg atttgaggaa agacatataa caactataaa aaatagattt tgcttgttca 6396
 ttggcttatg gtctccaggc ttgaatgctc tgagataaat gatgccaata tttctctggc 6456
 ctcttcccat cccacgcatt ggacctcaga tggctctgtac tgtcttctag agggtttgtg 6516
 ggttttggcc ccaaaaaacc attaaccttg gcagaaagtg tgtgacttta tgatctggta 6576
 caaagaagga caaactagag ggactggaca tgaggatgaa tattgtgttc gcccttatnn 6636
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6696
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6756
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnaaa gaccctcagt 6816
 tgttacaggg gcagtgcct cctcacacct caaccatcaa tgagtcacca ggaaagccat 6876
 tagcctagat gtaactgttt tctatcttta ttgcatttcc tacatccagg cagcagctgg 6936
 gaggaactct agaactga agtttgtctg agttccctta atgtaaggct gtacattctc 6996
 aggatgcctt gatgtactcg aatatctgca accctaaatc accacctctg tttttattga 7056
 tctctatctg aatgctgtat taatgggcca ggccttctgc ccattctctc aaactgagaa 7116
 ctgtctctca ttcttgggga ggcacctgc ctactcctta cctagatcag ggatttctca 7176
 gttgtggaga gatttgttcc ttatagtgtt ggtcatcaaa ctgggatatt tggggattac 7236
 aaagactttt caagggatgt atgggcacag gcagttttag gaagtgagtt cctagatcct 7296
 catcttcccc aaatactgt tccaaaaatt gacgagcctg acaatgtgca tgccaggcaa 7356
 ggctcttggg gttccctaa aacacttctt cttttaagcc taccactcac tcatcatgaa 7416
 tatagtccat tgtcccaggg tgtaaaacct tctatagtgt taaataaaag aatgattggg 7476
 aacattgaca cctgatggaa ctgttatgac taaaaacct tttgcaaata atgtggtatc 7536
 taattttctg ctttcaacaa aattgaagga ggccttata aagttaataa ctgataatca 7596
 aaaatgagta atttttgcca tgtaaatcag gtcaaagaat gaaatggcat tgcgtgaacg 7656
 aaactgcttc cattccatt gatttactca tacgaacaag attccttagc ctttataagc 7716

tacaaaaaaa tgaaaaatag aaatagaatt gaggctgaat totattatat aaaatcattc 7776
caaccatgtc atatggttct toggattcat gaataatttg gaaaagagag ccatatccat 7836
cttattaagg gacacattcc caataaattt tcatctttca tgtttaataa ttatcaatat 7896
tcataacatt ttacattttg atcaaatatg tgtaataaat aatagaaatn nnnnnnnnnn 7956
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8016
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8076
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8136
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8196
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8256
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8316
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnattag atcttaatgc agaacaccct 8376
gaacatttaa agcttcatag tcacaagaga aaagttttca tttcaatagc tataaatatt 8436
ttgttgttgt aaagacatat aacgataatc aatacaaaat ctgtcaaaca aaaatatgtt 8496
acattaagat aaaattctgt aggggaaggtg aaattggaag tgagtttcaa tgaatgaaaa 8556
gaaacaattt agacagagaa gnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnncttagaa 8616
ggaattgaat agattaggtt ccttaccocaa aaagcctctg ttatttgtct tattttattta 8676
ttctcttttt tccacattct ccagtctcat tccccttttt taacacagga aattattcca 8736
gcatgtttca tacatattct tttgtttgta agagcttatt taaaatatgt aatattgttt 8796
tagatgcata tatttttttt cttgtggaaa ctatattgta ctatatatat atatttttaga 8856
aatggacaca ttannnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8916
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8976
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9036
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9096
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9156
nnnnnnnnnn nnnnnnnnnn nnt 9179

- <210> 17
- <211> 1814
- <212> ADN
- <213> *Homo sapiens*
- 5 <220>
- <221> exón
- <222> (738)..(1010)
- <223> exón 3
- <220>
- 10 <221> incierto
- <222> (72)..(118)
- <223> "N" puede ser A, C, T o G
- <220>
- <221> incierto
- 15 <222> (405)..(630)
- <223> "N" puede ser A, C, T o G
- <400> 17

tagatctcag ctttcttgag gcaggagacc atatctgttt aattcactca gcatatactg 60
 caaagaagca gnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnta 120
 gggagaagtg cagaataaat atacccaact ctttactatg tatagacatt atctaggtct 180
 ttattttttt tctcttttta atctcaaaga aaacagagga aaggaggaag taaaaagtaa 240
 atttttgctt gaagatgttt ggaaaaaata ccaaataaag tgagatagtg ggtaatctag 300
 tgatttttat ttttcogtcc tctttctggc ctccaattgt gaaataattt atagcactgt 360
 aagaaagaag ccacaaattg tggtagcttg gaccactgtt gaggnnnnnn nnnnnnnnnn 420
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 480
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 540
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 600
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn acttaatat gatagtgata attttattca 660
 tttcacatgg catgaagtac caagctctat aggaatcaga aaataaagtc ttattttctt 720
 ttcttctota ttgtcca gct tcc tac agg aaa ata aac act cac atc cta 770
 aag gtt cca gaa aca gat gag gta gag ctc acc tgc cag gct aca ggt 818
 tat cct ctg gca gaa gta tcc tgg cca aac gtc agc gtt cct gcc aac 866
 acc agc cac tcc agg acc cct gaa ggc ctc tac cag gtc acc agt gtt 914
 ctg cgc cta aag cca ccc cct ggc aga aac ttc agc tgt gtg ttc tgg 962
 aat act cac gtg agg gaa ctt act ttg gcc agc att gac ctt caa agt 1010
 aagagctgcc ccacttcct aggtctatca gttaggggtc agacaagaaa cagatggcat 1070
 actcgagtga tttgaggaga gtgtaataaa gggactgttt acaaagggtg gatcaccatt 1130
 tggagaaact acaaaggata gtgcagaaca ctggggcttc aatggttgga gggcaattac 1190
 cactgttgga gaagttactg gaatcagaag ggagctgtag ggaaagcccc acttcccagg 1250
 agctgtagcc acagaatagg gaagctgcca catgcagcga ctccaaaggg tggaaactgg 1310
 atgaatgaat accccaactc attctcctcc caccctccaa tctcctgcta gcacctccca 1370
 ttggctgaac ccagctagaa gtcagagaat acaagggtcc actggtgtat tccataaaag 1430

tcaacttctc agggctcaga gcaatattga catgtacaga atagatctgg agaggaaaca 1490
gaaaatatct agtacaatag ctaatcactg tgattcatgc acagtgtcat gagccagcag 1550
gatgaatatt cctttgctgt acttgctgcc agtcagctgg ttatggggtt ttccaagaaa 1610
tttgggtctc aacaaaattc ttcagagcct ttactgacta tgctggatat ttttggaagg 1670
gatcccatat ttttgaactt catacagcag aatttcaaac aatcttggga aaataacaac 1730
ttttatctgc ccagtaagga caactaacac ctagtatcat aatcatttcg taagagacag 1790
gtatattcat caccgagtgc atat 1814

<210> 18

<211> 19217

<212> ADN

5 <213> *Homo sapiens*

<220>

<221> exón

<222> (6412)..(6543)

<223> exón 4

10 <220>

<221> exón

<222> (11953)..(12003)

<223> exón 5

<220>

15 <221> exón

<222> (18746)..(18751)

<223> parte codificante del exón 6

<220>

<221> incierto

20 <222> (722)..(1277)

<223> "N" puede ser A, C, T o G

<220>

<221> incierto

<222> (1740)..(1924)

25 <223> "N" puede ser A, C, T o G

<220>

<221> incierto

<222> (3538)..(3783)

<223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 <222> (3929)..(4224)
 5 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 <222> (5233)..(5352)
 <223> "N" puede ser A, C, T o G
 10 <220>
 <221> incierto
 <222> (5974)..(6030)
 <223> "N" puede ser A, C, T o G
 <220>
 15 <221> incierto
 <222> (6048)..(6073)
 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 20 <222> (6091)..(6306)
 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 <222> (7576)..(7688)
 25 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 <222> (8259)..(8368)
 <223> "N" puede ser A, C, T o G
 30 <220>
 <221> incierto
 <222> (9787)..(9885)
 <223> "N" puede ser A, C, T o G
 <220>
 35 <221> incierto
 <222> (9926)..(10110)
 <223> "N" puede ser A, C, T o G

<220>
 <221> incierto
 <222> (10113)..(10241)
 <223> "N" puede ser A, C, T o G
 5 <220>
 <221> incierto
 <222> (11000)..(11438)
 <223> "N" puede ser A, C, T o G
 <220>
 10 <221> incierto
 <222> (12528)..(12844)
 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 15 <222> (13720)..(13963)
 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 <222> (13983)..(14275)
 20 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 <222> (14529)..(14551)
 <223> "N" puede ser A, C, T o G
 25 <220>
 <221> incierto
 <222> (14585)..(14621)
 <223> "N" puede ser A, C, T o G
 <220>
 30 <221> incierto
 <222> (14703)..(14749)
 <223> "N" puede ser A, C, T o G
 <220>
 <221> incierto
 35 <222> (16380)..(16436)
 <223> "N" puede ser A, C, T o G
 <220>

<221> incierto

<222> (16484)..(16517)

<223> "N" puede ser A, C, T o G

<220>

5 <221> incierto

<222> (17022)..(17366)

<223> "N" puede ser A, C, T o G

<400> 18

```

ctactgagaa gggatatact ctcagactaa aggacagtcc ctagtactga ttcaatctgg 60
ctttatagaa aattcactat attgtcattg tatttcacag tttgcccttt gtcttagctg 120
gtaagacaga gcctatgata aggacttggtg tggcatgcag gtatttaatt ggcaacccca 180
gagggcagaa gcaagagatt taggagttta agagagggtg atataagagt atattatcaa 240
agttgtagtg tggacaacag aaactcaaatt attcaaggac cagcatgtag acagcctcct 300
aagatgtcta ctcagacaaa gaatttcagg tgggaaggact tgttcatctg cttcacgccc 360
attggttgac aggaatatga actccattct gctgctgggc tagacatgca tgtgggctga 420
gtgagctttc cccagtatcc gtagcatcag aaaagtcgca gggcagaaag aaaagtatcc 480
aatttgaggt gaattactga ccttgaagtg agtgtaagcc taactagaat tctaccccag 540
ctggctgaag tgaaagggtg ggctgagagg aaataaggca ggactgcaca gtccccaatt 600
gtactgttca aatccactca tgcccttcat taagtcagct ctgccactga gccttccagc 660
tgggaggcag ccacaatctc tgcagaagat ttaatatata ccagtttgtg gaacaagctg 720
tnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 780
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 840
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 900
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 960

```

10

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1020
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1080
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1140
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1200
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1260
 nnnnnnnnnn nnnnnnnnacc tgcagcactg ggatagccct ggtacagacg tagaacatgc 1320
 ttgaggaggt tgtcaggagg aaatgagtta gaccttgcac agaactacca ccatcaagta 1380
 cagttgggggt gaggcagaca tggatatgagt tgaggcatca gagatatctg atgctttatg 1440
 ccaaattaaa attaatTTTT tcatggagtg aactgatcc acagaccaga ctccaagaac 1500
 tttgcagtga ctaaataccc atctcatcat aactttcctg gtattttctt ctggaaaaaa 1560
 ttcttccctg atacagtttt cagaggcagc tagatgcaact gtcattctctc cctttttccc 1620
 acttccctac ctatccacaa tttactaccc aatgccacaa ctaaagtttag cccaacttcc 1680
 ttctaactaa attattagtt tagaaggaaa gagaggagtc atgctaagga tcttaactgn 1740
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1800
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1860
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1920
 nnnngggggag aggaggggag aggaggtagg gaggggaaag aaagaaacta gaaatccatc 1980
 aatttttagga ccaacttcag gtaaaaaaat gaattaggca agttgggtctt tcaacattct 2040
 ctactctctt ttatatcatg gttgagacca cagacttctc acctcatgaa agatgaactc 2100
 taactaatc atactaaagc taaagcctct aaagaggatt aaatatgagc aatcccacga 2160
 gaactttttt cccctggaat tgtttattca actgtcgttc gttatatgga atttctctgc 2220
 tggttaagt taggccagta ctttgatga attgtagttt tctagaaaga cgcttcttat 2280
 ataagaacct ctccaggga acaggggcct gtatgagatg aattgagaa taactttaca 2340
 coactgatfa tgtcagtgtt ctattctgca tggtagagat gtgaaagggc agactgacca 2400
 ttgctctgga agcctttacg ctgtgagaag ttaacagtgg agtaaaatgg ccaactccact 2460
 ctcttcatgg aagccaacat ggcttactaa atagtcaaca accatgggag agacctgtgg 2520
 ggtcttcac agagctcagg atctcctagg gtatcactca taaatacagc catcagggag 2580
 atggagaaat ctttgtgcag ccagaaatc tcaacctggg tttaacctc ctcccaact 2640
 ttgtattcgt cctactgttt actgacatgg atctctgtc tcattaacca tcccttctc 2700
 accacatgct ctctgaactt ggctgcacct tttctacctc catgccttct ttgctcaggt 2760

ttttccacat aaatatcatt atttccctct ctactagctc caagcccacc ctctctctgg 2820
 ggcagctcag tcactccagg gcacaagggg gtctttccct catccacat tttgagacct 2880
 actacctgga ccatttggtt gccttgtaac tatgcttgcc tttttaattg ctattttatt 2940
 ttccatgtat tttcattggt cacacaagtc ttctttattc cacactaagg caaaagcaga 3000
 gtcctgtggt cataataagt gctcaacaaa tgttgggttg attgggttg agattccatc 3060
 ttagataatc gcagtcccat catgccagct accagactgt gtggacagcc aggtcagagc 3120
 agccaaatga tattctagct tgtggcacia ataccagcaa caaaataacc aaagtcacac 3180
 atctgcctct gagttcctgg cttctatttc tcaagggcat ttttaagttg tcttatgact 3240
 gttccctttc tactcattct cataaattga gctgtggact gctgtgacct acaagcttct 3300
 ccggaagtca atgtataaaa caaacacgga aacgaagagt atggtgggtg gagggctact 3360
 cactgactct agaatggatg actgaacatt ccaaatttca agcacaagtt agggagcaac 3420
 agatcatttt ccttttgaaa tagggtttct tctgctcagc cagttgttgt attttcatta 3480
 ggaaatggaa tgggactaca gcacaaaaaa taaatataaa aggacccttg tagggctggc 3540
 agaaaagaga atccttccta ggagacctgg aggtgattcc aggcnnnnnn nnnnnnnnnn 3600
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3660
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3720
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3780
 nnntcagaaa gtgtgcaaac agtaaaaaaa aatggtatat ctagcaagtt gcatgcctta 3840
 cttgtgagtt catgaagttg tggcaaggat aagacaaata ttttttgcca ttgcatcatt 3900
 atatcattgc taagagtatg ccattattnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3960
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4020
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4080
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4140
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4200
 nnnnnnnnnn nnnnnnnnnn nnnttaatt aattaaatta aagaaacgac aaaagagtat 4260
 gcaagaattt taaaacaact tagaggaata tgtatgagga tacaggctaa gctaccataa 4320
 tgaagagacc tcgaaataca gtgagaagcg agacagaagt atctttcggt ccatgtaaca 4380
 ctcagggtgt tcagagcagc taagcagcta tgttccatag agtcattcag tgatccagat 4440
 tattttcatc tgttgctctg ccattctcca ggatgttgct cctataaaat tgtcaaagct 4500
 cagtcagtgc caaacccatg tttcaacctt cagaaagtaa acgagtgggt gaaaacacat 4560
 tcaatgtttt aaggccaaga ccttgaaaac tcactctctt agcctgaact tagattacat 4620

ggctgggccc acttaactat aggggaggct tggaaacata gtctctgaga agccatgtgt 4680
 ccagctaatt ccctaatact aaagttgaaa gaaagaatgg attaaccagc agtataccac 4740
 aaggtaacaa atgactagga ggatcaggct aggtggacta gaaaagagac agtcaattca 4800
 gtgcaacaat tccatattga cacttttcat gtagctgttg cttggctcta tctagagagg 4860
 actcagaggt agtttagata aggcctttgc cctccaaata cagtctaagc agactgattt 4920
 cctactggat gttcaacttt ggagtcttca gggatgagta gggcttctgt acgtggaaga 4980
 gactatgagg gaacctgcac aggacaaggg tttgcataaa gacactgagg tagggacctc 5040
 tctgtgtgtg gggacagtga gaggcccagg tctccttgac tcacaaagtg cttactaagc 5100
 acttactaga aattaagaag cagattataa tcaatatggg ttatccaatg tttggatgag 5160
 caaggctcct tatcttttct tcgttaatgt taatcacact cttttggatg gagacaaata 5220
 tctgtggggg ctnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 5280
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 5340
 nnnnnnnnnn nnggctcaga ctaaataatg tetaatctct tctccagtaa aacaatccgt 5400
 ggttctcaga tagcactgtg ctggaggtag tggggtttga gggctgggaa gttgggagga 5460
 ctgagccctt cccgctgagc agtttctgtc agtttttctt gtaccagcct gtcagtgtta 5520
 ttccatgtga atgactccag aggcaaaatt caagcttttg aatagggcac aaattaactt 5580
 gagtaccctt tcatttccct gtaggtgaac actcctctag cctgccttt tgtcagtctg 5640
 gagcccttgt tctaattctgt acacaccaga ggactttaca aggctttccc cagcctccag 5700
 aattattctt ctgatccacc ctctactaaa ctcacccttt cctcagtgtc aggacgttga 5760
 aaaaccgaaa caaggcaaag ggccaattgt aataattcac actaaggcat gagtgactag 5820
 gtttagtata ttaacactac ctaggatatt ctatttcttc caaaaggatc ctgttaatcc 5880
 ttgaaattta acaactaatg gtatagattc taagcactgt gagtacttgt cagtggggga 5940
 aagacatttt tgggctgaga gactttgcca ctgnnnnnnn nnnnnnnnnn nnnnnnnnnn 6000
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn aagaaagatg attagatnnn nnnnnnnnnn 6060
 nnnnnnnnnn nnnaaaaaat aacatgagag nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6120
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6180
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6240
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6300
 nnnnnncaca gagctagccg tgttggtgt caccactca tgtggccagc ctgttggtct 6360
 acctcttagt tgccatgtaa caggattctg gtgttttcc tttgccagg t cag atg 6417

ES 2 372 522 T3

gaa ccc agg acc cat cca act tgg ctg ctt cac att ttc atc ccc tcc 6465
 tgc atc att gct ttc att ttc ata gcc aca gtg ata gcc cta aga aaa 6513
 caa ctc tgt caa aag ctg tat tct tca aaa ggtaagtga ttttattcat 6563
 ggtaacccaa tgcactgggt gtctgcagca tgagccactg ctttgactg caggcctatg 6623
 gcttgctgct ttcattgctaa acccactcag agcttatgaa ccactttgag cttgtcttga 6683
 tgattatattt tccccagaag aaaatggctc tcactcgtcag tgagctgaac ttcttacct 6743
 gagtttttta aagggaatgt tttgttctta tgtctgaaag agtttgtctt attctttgag 6803
 ccaagagctt tcactcagcct catgagagtg atgttatattt ggcaatgcag agagctacgt 6863
 gctccgattt tgctgggtggg aggttgccag gatcctttct gaggattcct tccattttca 6923
 cccctctttt cccagctctg gatatgacct ggggttaaacc cccccctct cccaggaatc 6983
 tcaacctcac ggttgggtaa ggaaaggaga aaggtttggt aggccatttg gggataagga 7043
 aacagctggt tgggtggtgca ttaacgtctt tcagcagctc ccttcgagtt tctccttagc 7103
 ctgttggtatt cttaccaaca cactcctggt ctgttggtacc agctgggaca gagcatgctg 7163
 aagcctttca gccctgattt cattgcttca ttgttcattg gtctgtcttt ggtttcctgg 7223
 gtggagcctg cccacaaaac cccagaatg tatgcaggcc tagctggtgc tttcctaaac 7283
 ggctcccttg tctgcactca atgaacttct ccaaagatct atacatggcc tcacttatag 7343
 aaagagaaat gacatgtgga aataattcag taggagtttg cagcagcact atctgaggac 7403
 taggggaatt ttaagtgggt gttatcttac atttatactc ataacttcta tattttcatc 7463
 tgccataaaa tattgtcatg ttctatttgt ccattgccct atgtgtgtat gtattcactt 7523
 ggggtgctgac cacaatattt ctaactgtag aatgcaagga attgttgcca aannnnnnnn 7583
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7643
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn tggtaaagga 7703
 ggaggtatgt ccaacagaac ttcgactttt aaatagaacc acttcagaga gttgtgtcag 7763
 gtgcacctca gttgtcttat cttctgccat tcttctttta cctctcacac ccataacctca 7823
 gggttcaagg cctggggcct gaggactcct taataacttc agaaatgagc agctgagtg 7883
 tccgttccag ctgtcttttg gagaatggaa tggagtcaca ctcaaagata gagtggaaat 7943
 aaatcctctc ctcatccttc accccaatct taagagtga tgaggatatc agtagctccg 8003
 agctgggagg taaagctcaa gttctaactg tgattaggag acctttctta caaataagaa 8063
 ttaagtgaat aaatgtgcaa acaatttctt ttatatattt aatgaaccag agagaaatca 8123
 tggttgccta tataaccctt gtctccaact cacttgcatc cagatctgct ttcttacctg 8183
 tgtctgccat gcacacaaac ttgtgtgcca tggaaaaggg ttgagaactg ctgggtgatgc 8243

agacagagct ttaaannnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8303
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8363
 nnnnngcagc aagagaagga acattttaca gcttattggc cgaacttcac tgccgctagt 8423
 gtggttcaac ttggactaca gagaaatcct cctaactggt ttccctgtat tcactcctgc 8483
 tacctccaac ttggtctggt ctcaactttt gctataatag gcttttaaaa atcataaatc 8543
 taccatgtgt cctctgtcca gaccttctcc atggcttctt attgctcact ggatgaagtt 8603
 ccaacgagcc caggatgggt tgactcatgt ctccagcttt aactgcatca ccatcacctt 8663
 cattgtctaa agctctaacc acacaggatt ttctagtcct cagaggcatg gcagtctttc 8723
 aattccgagt tttctcatac aatattgtct cttcttaaaa tattttttct tgttgtccac 8783
 ctgagttgga gtcacttttt aaatctcagc taagcttata cttcatcaag tctttcctaa 8843
 ttctacctcc acgcaccaca ccattacat taaatcccct tattatatgt ttccatagca 8903
 cctactttct tcttttcagt atactcagca cacaatcaca tgtctaggat ctgttttaat 8963
 agcttggact accaattaaa ttgcatccct ttttaattgt cattgattcc tcaagtacct 9023
 acatgcccat cttagcaaga agttcagtggt ctccctctta tagcatgtac ttctccacct 9083
 cccacaaact gccagaaagc ttacttagcc cacagggcca gtgctaggca gctagggttag 9143
 tctccagag ggccctgggt ttgagcagtt gctgtctact ccggccatgc agaactctctg 9203
 gtccttccag atgtctcctat ccactgtgca aaggtaacct tgctgggtcc gatccccaca 9263
 cagaccacag tgctacaaga ttacagttct tatgggtccc caacacatgc tctgtcattg 9323
 gtcöcaaagc aggacccta tgggttgatg aggtaggagg aggtccctgc cttagccaca 9383
 gctgcacaca gccagcctct tcccttctag gccctcatgt tgagcctggg acgccagtcc 9443
 taacttcctt ctcttcagtt cctcttaggg ccattgggtat cctgaatttc ttagtccatt 9503
 gcaaagttaa gtaaagaagc agcaggcttg gtöcccttcc ttccagatgg cttcttagct 9563
 cctgaacaga tttaccacc tatacctcag tgactagctc tgtgtactaa agtgtattgg 9623
 gagggcagcc attattgggt cataaaaggt cctgcttacc attttcccct aagaggaacc 9683
 attcaacagt ttggggctcg aggggtgacct gctgggctct agagaagaag ctggcaactt 9743
 ctgttgcaaa ataatgttaa attctgcttc atctgcttgt cttnnnnnnn nnnnnnnnnn 9803
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9863
 nnnnnnnnnn nnnnnnnnnn nntaaacatt gaacctacta tatgcaggtg agtatgctag 9923
 atnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9983
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10043

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10103
 nnnnnnntgn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10163
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10223
 nnnnnnnnnn nnnnnnnnga tttgaataca ggtctgtttg actccaaaac ttgtggccta 10283
 tttgttgcaa aagtgcctaa tacaaattgg ttcagtcaat attattatct ttgaacaatg 10343
 gaaggagaaa gtaagtttca atccaaaata attgagtgcac ttatacattg acttgctgag 10403
 ccaatggcaa agtcaagtta gaatccagca gaagtcacca gctacagaat ctagatcttt 10463
 agaacatgtc ttcagatctt cagaacagtg tttcttaaac tctcttgtga aggaacagtt 10523
 atcatcatag gotggtaaca gttcacctac cagcaccagc ccatgaacca gactctaagt 10583
 ggcacagccc tagaagattg agccagaatt ttacagaggt ttaaagacca aatatgctgg 10643
 tttatggtta cctgtggccc acagagaatg gcagcactaa cctcaggcat aaatgaggta 10703
 cccactgaag ccaacattca agagcaattc ctatgggtta accattgggc tcctttcaaa 10763
 tgcaaaccct catgaaagag actacagtgc tgaatagaga cctccaaatt ccaggccaag 10823
 ctcaggatag tcatgaggga attactaaaa acctggtata tagggcaaaa gcagaattag 10883
 gaatggactg atttcaggaa ccaggcaat ggcaggagtt gggcattaaa tcctaaaaga 10943
 gaatcagagt gggagggaat atgtgaaatc agaggttaag aaaaaagtga aaacctnnnn 11003
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11063
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11123
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11183
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11243
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11303
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11363
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11423
 nnnnnnnnnn nnnnnnaaaaa attaaaagaa agatgtgaaa tcaaggaaac ttactggtga 11483
 gcagcatccc attatgtgaa cttgtgcttc tgaaccagta acttgagtta ctttgagoca 11543
 gtatcagtca cttatacctt agtgcaaaat taattgatca gacattctga cctggaccag 11603
 ggaaggcagg cagaagtagc agtcaagact aaagcagaaa agggagagct aattctgcag 11663
 ccagacattt cctggattga ataccctaat tagtccctca gcctttaagt gcctgagggc 11723
 caggagtaga cagaggaatg gaaagtgtga gacttctttg ttcacactct ttgocagggg 11783
 gccagatttt gctttatgca ttaccatccg aagtcccagg ccacagtga catttgggct 11843
 tcgctatgtg gatttattta gatttacttt ttgtcctgcc atattttaat ctataagcca 11903

aacagttttc tcattaatct tattccatth ctggaatttt tccttttca gac aca aca 11961
 aaa aga cct gtc acc aca aca aag agg gaa gtg aac agt gct 12003
 gtgagtaagc atgattttta cttttctttc ttactttott ttctctctca gcttgaattt 12063
 taaagtaacc actgttctat taattcatgg aaggcaactg aatagttcca gcttatagaa 12123
 tcttctctgtt tggtagcatt tcagcgaagc ctggttctta gcccagaac aatcatgcca 12183
 tcttttgctc ggtctatatt cctaagcact cctagatgat actgcactgg acctctggtc 12243
 tcacatagtt agaaacagag ttaaaatoga acagcaaaga gaagatattc aactgcgatg 12303
 caattgacaa tggatgtttt tgcaacaaac aatgattaag aagtacattg ttgtgggctc 12363
 tgagtcaaga gtaatatggg aaaaacacaa gtctcttcat gaggttgaca ggtttgagc 12423
 tggaatctgt ggaggaggaa ggatatgac taggggtcag aagaagtggg ttactaaaat 12483
 cattaagcct ggttgatga aaagcttaga ctgaggggaa gcagnnnnnn nnnnnnnnnn 12543
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 12603
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 12663
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 12723
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 12783
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 12843
 nggcaccaag agatgagggg ggcagtcttg gccatatatt tggctgaagt aagtcaattt 12903
 gtcattcctg catgagcctt tataaacaga agtaagtaac caactactat ttggtcattg 12963
 gagttgtcca agaggccagg gttctgtcta atacctgttc atgcatgaac atgccaacct 13023
 agattgcatg cagactacca gttttgggtt tttgtttagt tcagcaggat ttttctcagc 13083
 tcaactgctc tcaaactttc agcaacaaaa ggacatctgt gatatcagaa tctaccactc 13143
 taagtatttg gatgcaatag caatgaatat ctgagtaaatt ctaggtgggg agtgggggca 13203
 ccctgtagcc aaaatgattt aacaaaatca aacaaaatt ttggaaatga tgccttggtg 13263
 caatgaagag actacttgag gtaggtttga cttatctaatt atcttatttt ctttaccat 13323
 acctaatgag gaatttaaat atttctagat agctttggaa aggtccctta aagaggcacc 13383
 agcataccac tgccagatct aatccccca aacaotgttt tcatcatcat catgtcatct 13443
 cttgtctcta tagatcatat caaatccttc ccagagtttt tcaggccttt tgacaactag 13503
 ccacatttca ctaagccaac tcatctacca ctcttcaaca aaacttttcc tcaagttgag 13563
 ctgctocacc aacaccactg ccatgagctc attcccactt ctgtggcttt gctcatgttg 13623
 gttatttttt tggagtgtcc tccctattcc ttcttaactg tccaatccc aacttttggc 13683

atggtctact ttaagataca gtaatgagta actttannnn nnnnnnnnnn nnnnnnnnnn 13743
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 13803
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 13863
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 13923
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn tttttaaaat ttaattcaan 13983
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 14043
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 14103
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 14163
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 14223
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nntacataca 14283
 caacactctc ttgcttaatc tgtaagactc tctccccac tcataccttt ttatttttcc 14343
 tctgcattgt acacacaatc tataccactc ttaagcacat gattacagcg ttattttctg 14403
 gctgcttcta tgtgtctata ttttaggtcc acctgggtcaa tataataaag tgggatatta 14463
 gtgttaatgc aactatatgg tatttgatat ttgtctttct gtccgtttat caatgtttct 14523
 tatagnnnnn nnnnnnnnnn nnnnnnnngg tgtctgattt tgaccaaatt tgactaaata 14583
 cnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnngt tgagtctatc aactcaaaaa 14643
 agaataacct acaacaataa caagtttcag aacatttttt aaattactga ttttatgagn 14703
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnngaatt ccattgcta 14763
 gaaattgcca tgggttaagat tttaatattg ctcaggccca gacagctcag ggctttgaca 14823
 ttcccacacc cattctctgc catcccagtt ctatctcatc ccaaaacoat ccattatgag 14883
 gagagtgtac agctctaggc tgcccgggag ccatcccgca ctctcatttt gtgactcggc 14943
 atcttgggag atggagtctt gggacttagc ctggacatgt cccttgcatg tacttcttac 15003
 aagactttta ttcagatgaa tattttccct tccaacttaa gaagcacagg gcttgctggt 15063
 ttgcttcac taaccagcaa ctgaagcaag acctgacttg tgaaaatgcc taatagagtt 15123
 cagtattagc gctgtttcac catttcctgg atctcttgcc tttgtgcaca tgatagaatt 15183
 gcacttctct gtgattaatt tgtagttaag tgtggtcatg taactcgtt tggatcaatta 15243
 aatgtaagca taagtgatgc gtgttatttc tgggtagaag atgtaagagt tggcatatgc 15303
 ttgcatat tttctttatc catctggcat ggtaaccagt aacattctag gtagtaattg 15363
 ctccatcagt ctcagtctct gagtgactaa aattgacaga gtcccctgct gaccctcaat 15423
 gtacatggaa catgaacaag aataagcttt tgttttttat attgagattt tggagttggt 15483
 tgttcctaca gcattaccta gtttactcta atacaacatg gaaaaaactg gaacctataa 15543

taaatagacc ctacgttgcc atttaaactt ctagttctga ggaataataa tgtggggaaa 15603
 tactttctat ataataaaaa aatagaaaat tgcaaaaataa aaatatactt atgtatcatt 15663
 catgtcctat taaaaatggt atttatagac tcaccatatt cccttcctcc agaaaaatag 15723
 aagtaaaaat atgaaaatgc ctgtaatcat gtttttggat tatggaatca agtattgctt 15783
 tttactttta tgttttctga atttttgttg tacttcacta catttttgag tgccctgatg 15843
 tattactttc aaaaagaaga agaatacttt ctgaagccat ttcaaccatc cccactcacc 15903
 tctctagatc ccagtaacca aatacattat ataggactct tcatcagtcc ttatcaagtt 15963
 taggaagggc gatgctatac cttctttaaa ggacacctac caatgtctta gttgcctttc 16023
 aaagactcct agcacagcta aatgtgatgg atatgctcta aggatataag agctgaagtg 16083
 acttgcataa ggtcatatca taacttactg ttagaaatgg agctagaact cagaccact 16143
 gagtccttgt ctgtgacaca ctgccctttc catttggtga agttgttctt gtatctaact 16203
 ttatctgtgc tactatttgg gcctagccat tctccctctt atgcagacaa gcagataaac 16263
 agtaaaactt taggagtga ttatgatacc atagatatat atcatctatc ctttacaaaa 16323
 tagttattac agtcatcaag ccttggttag agtttacaga ccatgtatcc tagctannnn 16383
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnntgaaaca 16443
 gtagcagacc aaaagaagtc atgattccca gcatagtgtc nnnnnnnnnn nnnnnnnnnn 16503
 nnnnnnnnnn nnnnaaacct gtattaagtt tctgttattt gcaagacct gtcagttagg 16563
 ccatgtggga actagaagga tgaatttatc agtcatccaa gattcttaca attaagtatt 16623
 accgataagg tactcaagaa acagttctca ttcacataat ttgggttaaa acaaaaagaa 16683
 gccagctttc tatatacttt tgggtccagtc tttacgtttt ttgttttggt ttgtttgttt 16743
 tcatgagtat cccgacttcc ttctaagaac ttccacctga gaactgacca cagcgtcagc 16803
 attccacatg ggtgtgtttc ctttcccctt tcccatttca gtgggttcca atttcttttt 16863
 cttttggcac tataaacctt tcgcaaagga aatattagac agaactccta catgtcaagc 16923
 aaattaaaat agtggtgaaa ttagagtga ggacataatc accctatcat ataggctatt 16983
 tgtccatatc atatttgtcc ctacaaaggc ctctaaggnn nnnnnnnnnn nnnnnnnnnn 17043
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 17103
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 17163
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 17223
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 17283
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 17343

nnnnnnnnnn nnnnnnnnnn nnnctctaag ataccttggtg cttcttgga catatttgga 17403
 aaatcatgta gctctcaaatt tatccctatg tcctgaagcc caccttacca tcaatcctca 17463
 gaaataccca acccatgtca ggcaacttca cactttcttt cttcaggcag cacagttgtc 17523
 tcagggaggg aggagagtgc tattagcaag aggagtcact aacagcttca ctcacctgtg 17583
 cccatgaatt tttaatggtg tgaaaagtct gtgtatTTTT tatggttctc tatggctcat 17643
 ataagagggg caaacaatac atgaaagttc agagatgggg acaaatatca ctttaagctg 17703
 gggataatca gggagaaga aagttttcat ggagaagggtg acttttgaat caggaagaaa 17763
 tggaaacagc aaactcttct agatagagga gacactaaca ggaaaggcag agaggcagga 17823
 aggtgtggga aagtgcacgt gacccogttc agagagaaaag ccagggtgtga gataaggggg 17883
 aaagactggt agggcatgta ttgtaaacca ctaattccag gcaaaagtta gattttactt 17943
 actaagcaag agtgcttcag ttagatccta gcaggaaatg gagggtatgc ttagaagagg 18003
 taactgaggc aagtttaatt tataaagggtg tgtgcagcat taagggaac cagcaaggga 18063
 tactgagcat gccaggatgc aagagcaggc agggaagggtg actattccta ggtctgaagg 18123
 agaaagggga gggagcagtt cccagaaccc tagtaaaaat ggcaatgaga aaggccatc 18183
 tggcaggacc tatggtcttt aacagaggga caaagtcaac ccacaacttg tctgggaggt 18243
 tgctgaggaa tagatacccc aacctctctc tcaaccact gcaacactct ttttcccta 18303
 gactgagccc agtcaaagac agagggagga gccagtgat gcagtctgca atgtcatcat 18363
 cctggagcat gaatagagt cagcagggtg aataatgagt ctgcaggaat taatagaaat 18423
 atctgacaca atagggaaact ataagaggtt ttgaatagga gagggccctg aaatgtgctc 18483
 caatattact gaactatgtg tggcccaaag aatggaagag gaacagctct tgcaataggt 18543
 ctgaggagag aagctgaaga cttggactag ggcaatggta aaaactgtgg aaagaagttt 18603
 taaatgaaaa gttttaaacc atgcggcttc cagctagatg aactttttta aaaaaattag 18663
 ttctcactc aaattttggg gaggttatat attttctaata cataaaaaat gatttttctt 18723
 atttgtgggc ttttctcccc ag atc tga acctgtggtc ttgggagcca ggggtgacctg 18781
 atatgacatc taaagaagct totggactct gaacaagaat tcggtggcct gcagagcttg 18841
 ccatttgac ttttcaaatt ctttggatg acccagcact ttaatctgaa acctgcaaca 18901
 agactagcca acacctggcc atgaaaactg ccccttcaact gatctggact cacctctgga 18961
 gctatgggt ttaagcaagc actactgcac ttacagaat taccctactg gatcctggac 19021
 ccacagaatt ccttcaggat cttcttctg gccagactga aagcaaaagg aattatttcc 19081
 cctcaagttt totaagtgat ttccaaaagc agaggtgtgt ggaaatttcc agtaacagaa 19141
 acagatgggt tgccaataga gttatTTTT atctatagct tcctctgggt actagaagag 19201

gctattgaga ctatga

19217

<210> 19

<211> 18

<212> PRT

5 <213> *Homo sapiens*

<400> 19

Met	Ile	Phe	Leu	Leu	Leu	Met	Leu	Ser	Leu	Glu	Leu	Gln	Leu	His	Gln
1				5					10				15		

Ile Ala

<210> 20

<211> 102

10 <212> PRT

<213> *Homo sapiens*

<400> 20

Ala	Leu	Phe	Thr	Val	Thr	Val	Pro	Lys	Glu	Leu	Tyr	Ile	Ile	Glu	His
1				5					10					15	

Gly	Ser	Asn	Val	Thr	Leu	Glu	Cys	Asn	Phe	Asp	Thr	Gly	Ser	His	Val
			20					25					30		

Asn	Leu	Gly	Ala	Ile	Thr	Ala	Ser	Leu	Gln	Lys	Val	Glu	Asn	Asp	Thr
		35					40					45			

Ser	Pro	His	Arg	Glu	Arg	Ala	Thr	Leu	Leu	Glu	Glu	Gln	Leu	Pro	Leu
	50					55					60				

Gly	Lys	Ala	Ser	Phe	His	Ile	Pro	Gln	Val	Gln	Val	Arg	Asp	Glu	Gly
65					70					75					80

Gln	Tyr	Gln	Cys	Ile	Ile	Ile	Tyr	Gly	Val	Ala	Trp	Asp	Tyr	Lys	Tyr
				85					90					95	

Leu	Thr	Leu	Lys	Val	Lys
			100		

<210> 21

15 <211> 91

<212> PRT

<213> *Homo sapiens*

<400> 21

ES 2 372 522 T3

Ala Ser Tyr Arg Lys Ile Asn Thr His Ile Leu Lys Val Pro Glu Thr
1 5 10 15
Asp Glu Val Glu Leu Thr Cys Gln Ala Thr Gly Tyr Pro Leu Ala Glu
20 25 30
Val Ser Trp Pro Asn Val Ser Val Pro Ala Asn Thr Ser His Ser Arg
35 40 45
Thr Pro Glu Gly Leu Tyr Gln Val Thr Ser Val Leu Arg Leu Lys Pro
50 55 60
Pro Pro Gly Arg Asn Phe Ser Cys Val Phe Trp Asn Thr His Val Arg
65 70 75 80
Glu Leu Thr Leu Ala Ser Ile Asp Leu Gln Ser
85 90

<210> 22

<211> 44

5 <212> PRT

<213> *Homo sapiens*

<400> 22

Gln Met Glu Pro Arg Thr His Pro Thr Trp Leu Leu His Ile Phe Ile
1 5 10 15
Pro Ser Cys Ile Ile Ala Phe Ile Phe Ile Ala Thr Val Ile Ala Leu
20 25 30
Arg Lys Gln Leu Cys Gln Lys Leu Tyr Ser Ser Lys
35 40

<210> 23

10 <211> 17

<212> PRT

<213> *Homo sapiens*

<400> 23

Asp Thr Thr Lys Arg Pro Val Thr Thr Thr Lys Arg Glu Val Asn Ser
1 5 10 15

Ala

15 <210> 24

<211> 11

<212> PRT
 <213> Virus de inmunodeficiencia humana de tipo 1
 <400> 24
 Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg
 1 5 10
 5 <210> 25
 <211> 15
 <212> PRT
 <213> Secuencia artificial
 <220>
 10 <223> Descripción de la secuencia artificial: dominio de internalización derivado de proteína tat de VIH
 <400> 25
 Gly Gly Gly Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg
 1 5 10 15
 <210> 26
 <211> 27
 15 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: cebador de oligonucleótido 2515-27
 <400> 26
 cataatagag catggcagca atgtgac 27
 20 <210> 27
 <211> 23
 <212> ADN
 <213> Secuencia artificial
 25 <220>
 <223> Descripción de la secuencia artificial: cebador de oligonucleótido 2524-63
 <400> 27
 gggtcctgga gtggctggtg ttg 23
 <210> 28
 30 <211> 14
 <212> PRT
 <213> *Homo sapiens*
 <400> 28

Leu Phe Thr Val Thr Val Pro Lys Glu Leu Tyr Ile Ile Glu
 1 5 10

<210> 29

<211> 25

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: cebador de oligonucleótido 2515-24

<400> 29

gtggctcttt cacggtgtgg ggatg

25

10 <210> 30

<211> 26

<212> ADN

<213> Secuencia artificial

<220>

15 <223> Descripción de la secuencia artificial: cebador de oligonucleótido 2538-68

<400> 30

ccagtgtcaa agttgcattc cagggt

26

REIVINDICACIONES

1. Una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos seleccionada del grupo que consiste en:
 - (a) la secuencia de nucleótidos tal como se expone en SEQ ID NO: 1;
 - 5 (b) una secuencia de nucleótidos que codifica el polipéptido tal como se expone en SEQ ID NO: 2; y
 - (c) una secuencia de nucleótidos complementaria a las secuencias de (a) – (b).
2. Un vector que comprende la molécula de ácido nucleico de la reivindicación 1.
3. Una célula huésped aislada que comprende el vector de la reivindicación 2.
4. La célula huésped de la reivindicación 3, que es una célula eucariota.
- 10 5. La célula huésped de la reivindicación 3, que es una célula procariota.
6. Un procedimiento para producir un polipéptido B7-L que comprende cultivar la célula huésped según la reivindicación 3, cuando la reivindicación 3 hace referencia a la reivindicación 1(a) o 1(b), en condiciones adecuadas para expresar el polipéptido, y aislar opcionalmente el polipéptido del cultivo.
- 15 7. El procedimiento según la reivindicación 6, en el que la molécula de ácido nucleico comprende ADN promotor distinto del ADN promotor para el polipéptido B7-L nativo operativamente unido al ADN que codifica el polipéptido B7-L.
8. Un polipéptido B7-L producido mediante el procedimiento de la reivindicación 6 ó 7.
9. Un polipéptido B7-L aislado que comprende la secuencia de aminoácidos tal como se expone en SEQ ID NO: 2.
- 20 10. Un polipéptido de fusión que comprende el polipéptido de la reivindicación 8 ó 9 y una secuencia de aminoácidos heteróloga.
11. El polipéptido de fusión de la reivindicación 10, en el que la secuencia de aminoácidos heteróloga es un dominio constante de IgG o fragmento del mismo.
- 25 12. El polipéptido de cualquiera de las reivindicaciones 8-11, que está covalentemente modificado con un polímero soluble en agua.
13. El polipéptido de la reivindicación 12, en el que el polímero soluble en agua se selecciona del grupo que consiste en polietilenglicol, monometoxipolietilenglicol, dextrano, celulosa, polietilenglicol de poli(N-vinilpirrolidona), homopolímeros de propilenglicol, copolímeros de poli(óxido de propileno/óxido de etileno), polioles polietoxilados y poli(alcohol vinílico).
- 30 14. Una composición que comprende un polipéptido B7-L de cualquiera de las reivindicaciones 8-13, y un agente de formulación farmacéuticamente aceptable.
15. Un anticuerpo que se une a un epítipo en un polipéptido B7-L humano de SEQ ID NO: 2.
16. El anticuerpo de la reivindicación 15, seleccionado del grupo que consiste en:
 - (a) un anticuerpo humanizado;
 - 35 (b) un anticuerpo humano o fragmento del mismo;
 - (c) un anticuerpo policlonal o fragmento del mismo;
 - (d) un anticuerpo monoclonal o fragmento del mismo;
 - (e) un anticuerpo quimérico o fragmento del mismo; y
 - (f) un anticuerpo con injerto de CDR o fragmento del mismo.
- 40 17. El anticuerpo de la reivindicación 15 ó 16, que es un fragmento de región variable.
18. El fragmento de región variable de la reivindicación 17, que es un fragmento Fab o Fab'.
19. Una composición que comprende una molécula de ácido nucleico de la reivindicación 1, y un agente de formulación farmacéuticamente aceptable.

20. La composición de la reivindicación 19, en la que dicha molécula de ácido nucleico está contenida en un vector viral.
21. Un vector viral que comprende una molécula de ácido nucleico de la reivindicación 1.
22. Una molécula de ácido nucleico según la reivindicación 1, unida a un soporte sólido.
- 5 23. Una matriz de moléculas de ácido nucleico que comprende al menos una molécula de ácido nucleico de la reivindicación 1.

FIG. 1A

cagaaagaga cctatatgat caaatacaga ac	atg atc ttc ctc ctg cta atg	53
	<u>Met Ile Phe Leu Leu Leu Met</u>	
	1 5	
ttg agc ctg gaa ttg cag ctt cac cag ata gca gct tta ttc aca gtg	101	
<u>Leu Ser Leu Glu Leu Gln Leu His Gln Ile Ala Ala</u>	<u>Leu Phe Thr Val</u>	
	10 15 20	
aca gtc cct aag gaa ctg tac ata ata gag cat ggc agc aat gtg acc	149	
<u>Thr Val Pro Lys Glu Leu Tyr Ile Ile Glu His Gly Ser Asn Val Thr</u>		
	25 30 35	
ctg gaa tgc aac ttt gac act gga agt cat gtg aac ctt gga gca ata	197	
<u>Leu Glu Cys Asn Phe Asp Thr Gly Ser His Val Asn Leu Gly Ala Ile</u>		
	40 45 50 55	
aca gcc agt ttg caa aag gtg gaa aat gat aca tcc cca cac cgt gaa	245	
<u>Thr Ala Ser Leu Gln Lys Val Glu Asn Asp Thr Ser Pro His Arg Glu</u>		
	60 65 70	
aga gcc act ttg ctg gag gag cag ctg ccc cta ggg aag gcc tcg ttc	293	
<u>Arg Ala Thr Leu Leu Glu Glu Gln Leu Pro Leu Gly Lys Ala Ser Phe</u>		
	75 80 85	
cac ata cct caa gtc caa gtg agg gac gaa gga cag tac caa tgc ata	341	
<u>His Ile Pro Gln Val Gln Val Arg Asp Glu Gly Gln Tyr Gln Cys Ile</u>		
	90 95 100	
atc atc tat ggg gtc gcc tgg gac tac aag tac ctg act ctg aaa gtc	389	
<u>Ile Ile Tyr Gly Val Ala Trp Asp Tyr Lys Tyr Leu Thr Leu Lys Val</u>		
	105 110 115	
aaa gct tcc tac agg aaa ata aac act cac atc cta aag gtt cca gaa	437	
<u>Lys Ala Ser Tyr Arg Lys Ile Asn Thr His Ile Leu Lys Val Pro Glu</u>		
	120 125 130 135	
aca gat gag gta gag ctc acc tgc cag gct aca ggt tat cct ctg gca	485	
<u>Thr Asp Glu Val Glu Leu Thr Cys Gln Ala Thr Gly Tyr Pro Leu Ala</u>		
	140 145 150	
gaa gta tcc tgg cca aac gtc agc gtt cct gcc aac acc agc cac tcc	533	
<u>Glu Val Ser Trp Pro Asn Val Ser Val Pro Ala Asn Thr Ser His Ser</u>		
	155 160 165	
agg acc cct gaa ggc ctc tac cag gtc acc agt gtt ctg cgc cta aag	581	
<u>Arg Thr Pro Glu Gly Leu Tyr Gln Val Thr Ser Val Leu Arg Leu Lys</u>		
	170 175 180	
cca ccc cct ggc aga aac ttc agc tgt gtg ttc tgg aat act cac gtg	629	
<u>Pro Pro Pro Gly Arg Asn Phe Ser Cys Val Phe Trp Asn Thr His Val</u>		
	185 190 195	

FIG. 1B

```

agg gaa ctt act ttg gcc agc att gac ctt caa agt cag atg gaa ccc 677
Arg Glu Leu Thr Leu Ala Ser Ile Asp Leu Gln Ser Gln Met Glu Pro
200                205                210                215

agg acc cat cca act tgg ctg ctt cac att ttc atc ccc tcc tgc atc 725
Arg Thr His Pro Thr Trp Leu Leu His Ile Phe Ile Pro Ser Cys Ile
                220                225                230

att gct ttc att ttc ata gcc aca gtg ata gcc cta aga aaa caa ctc 773
Ile Ala Phe Ile Phe Ile Ala Thr Val Ile Ala Leu Arg Lys Gln Leu
                235                240                245

tgt caa aag ctg tat tct tca aaa gac aca aca aaa aga cct gtc acc 821
Cys Gln Lys Leu Tyr Ser Ser Lys Asp Thr Thr Lys Arg Pro Val Thr
                250                255                260

aca aca aag agg gaa gtg aac agt gct atc tga acctgtgggc ttgggagcca 874
Thr Thr Lys Arg Glu Val Asn Ser Ala Ile
                265                270

gggtgacctg atatgacatc taaagaagct totggactct gaacaagaat tcggtggcct 934
gcagagcttg ccatttgcac ttttcaaattg cctttggatg acccagcact ttaatctgaa 994
acctgcaaca agactagcca acacctggcc atgaaacttg ccccttcact gatctggact 1054
cacctctgga gcctatggct ttaagcaagc actactgcac tttacagaat taccctactg 1114
gatcctggac ccacagaatt ccttcaggat ccttcttgct gccagactga aagcaaaagg 1174
aattatttcc cctcaagttt tctaagtgat ttcca 1209

```

FIG. 2A

Cd80_Human	: MGHTRQGTSPSKCPYINFFQLLVLAG. LSH. FC. SGV. ... IHWTKKE. ... VKEVATISC. ... GHNVSVEELAQTRTYWQKEKMWLT. : 75
Cd86_Human	: MGLSNILFVMAFLISCAAPKIQ. ... AYFNETADLPQOFANSONOSISELV. ... VFWQDOENIVIN. : 60
GA16817596	: MTF. ... LILMLISLEIQHQ. IAAALFTVTVPEKELYII. ... EHGSNVT. IECNFDTGSHVNLGAI. ... ASLOK. : 61
B7-H1_Huma	: MRIFAFAFIEM. ... FYWHL. INA. FTVTVEDKLYV. ... EYGSNMT. IECKEPVEKQDLAAL. ... IVYEMEDKNII. Q. : 66
B7rp-1_Hum	: MRLGSPG. ... LFLFFSLRADTQEKVR. ... AMVGSDELSCACPEGRFDLNDV. ... YVYWTSESRTV. : 62
Pro352_Hum	: MLRRRGSPG. G. H. V. GAALGALWFC. LTGALEVOFEDPVV. ... ALVGTDAITCCSFSPEPGFSLAQL. ... NLIWQLTDTK. ... Q. : 74
Btf1_Human	: MESPAALHFSRPAS. ... LLLLLLSICAL. VSAQFIIVGPTDPIL. ... AVUGENTTTRCHLSPEK. ... NAEDM. ... EVRWERSQFS. ... P. : 73
Btsf2a2_Hu	: MEPAALHFSLPASILLILLILLLSICAL. VSAQFTVVGFPANPIL. ... AMVGENTTTRCHLSPEK. ... NAEDM. ... EVRWERSQFS. ... P. : 77
Btf4_Human	: FHSILLVQLLTPC. ... SAQFSVLGPGSPIL. ... AMVGEDADLPCHLFPITM. ... SAETM. ... ELKWVSSSLR. ... Q. : 61
Btn3a3_Hum	: ... MKMASSLAFILINFHVSILEFLVQLTPCSAQFSVLGPGSPIL. ... AMVGEDADLPCHLFPITM. ... SAETM. ... ELRWVSSSLR. ... Q. : 74
Btn_Human	: ... MAVFPSSGIPROCLLTLLQLLPKIDSAFEDVIGFPEPIL. ... AMVGEDAELPCRLSPNA. ... SAEHL. ... ELRWERKKVS. ... P. : 72
Cd80_Human	: ... MMS. ... GDM. ... NIWPEYKNRT. ... IFD. ... ITNNLSIVILARSPSDEGTYECVVLKYKDAFKREHLAEVTLVSKADPPTESISDFEIPTS. : 155
Cd86_Human	: ... EVYL. ... GKEKPSVHSKYMGR. ... SFD. ... SDSWTLRLHNLQKDKGLYQCIHHKKPFCMIRIHOHNSLSVLANESQPEI. ... V. FI. : 138
GA16817596	: ... VENDT. ... SPHERATLIE. EOPPLGKASFHIEQVQVRDEGQYQCIILY. ... GVA. WLYKYLTLKVKASY. ... RKINTHILKVP. : 134
B7-H1_Huma	: ... FVH. ... GEEDLKVCHSSYRQARLIK. DQSLGNAALQITVKKQDAGVYRCMISY. ... GGADY. KRITVKVNA. ... PYNKINQRIIV. : 143
B7rp-1_Hum	: ... VTYHIPQNSSTNVDSPRYANRALMS. PAGKRGDFSLRLFNVDQDKFCHLVISL. GFQEVLSVETVLHVAANFSVEVV. ... S. APH. : 147
Pro352_Hum	: ... LVHSFAEG. ... QDGSAYANRTALFP. ILLAOGNALSIRLQWVADEGSFTCFVSIRDF. S. ... SAAVLSQVAPYISKESM. ... TLEPN. : 152
Btf1_Human	: ... AVFVYKGGRETEREQMEYRGRTEVSKDI. ISRGVALVIHNTIQAENGTYRCYF. ... QE. ... CRS. YDEAILHLVWAGLSKPLISM. : 153
Btsf2a2_Hu	: ... AVFVYKGGRETEREQMEYRGRTEVSKDI. INRGVALVIHNVTAQENGTYRCYF. ... QE. ... GRS. YDEAILRLVWAGLSKPLIEI. : 157
Btf4_Human	: ... VWNVYADGKEVEDRQSAPYRGRTSILR. DGITAGKAALRIHNVTAQENGTYRCYF. ... QD. ... GDF. YEKALVELKVAALGSNLHVEV. : 141
Btn3a3_Hum	: ... VWNVYADGKEVEDRQSAPYRGRTSILR. DGITAGKAALRIHNVTAQENGTYRCYF. ... QD. ... GDF. YEKALVELKVAALGSNLHVEV. : 154
Btn_Human	: ... AVLVHRDGRGCEAEQMEYRGRATLVQ. DGHAKGRVALRIRGRNRSDDGEYTCFER. ED. ... CS. YEEALVHLKVAALGSDEHISM. : 152
Cd80_Human	: ... NTR. ... RIICTSTSGGFFEP. HLSWLENGEELNAINTV. ... SODPETELYAVSSKUDF. ... NMTTNHSTMCILKY. ... GHLR. : 224
Cd86_Human	: ... SNITENVYINLTCSSIIHGYPEPKMSVL. ... LRTKNSTIEYDGTIMQKSDQDNVTLEYDVSI. SVSFPDVTSNMT. ECIET. ... DKT. : 219
GA16817596	: ... ET. ... DEVELTQCAT. GYPLAE. VSWPN. ... VSPANTSHSR. ... TP. EGLYQVTSVIRLKP. ... PPGRNFSV. ... WNTH. : 198
B7-H1_Huma	: ... VDPVT. SEHELTCQA. EGYKAE. VIMTSSDHQVLSGKTTTNSK. ... REEKLNFVTSVIRLKP. ... TTNEIFYCTFRRLDPEEN. : 219
B7rp-1_Hum	: ... SPSG. ... DELTFTCTSINGYERPN. VYW. ... INKTDNSLLDQALQNDTVFLNMRGLYDVSVIRIAR. ... TPSVNTGCCIEENVLLQON. : 225
Pro352_Hum	: ... KDLRPGDVTITCSSYQGYPEAE. VFWQDGGQGVPLTGNVTTSQ. ... MANEQGLFDVHSHVIRVL. ... GANGTYSCLVRNPVLQ. : 229
Btf1_Human	: ... RGHEDGGIRLECHISRGWYKPL. TVWRDP. YGGVAPAL. ... KEVSMP. DADGLFMVTTAVIIRD. ... KSVRNMSCSINNTLLGQK. : 229
Btsf2a2_Hu	: ... KAQEDGSIWLECHISRGWYKPL. TVWRDP. YGGVAPAL. ... KEVSIA. DADGLFMVTTAVIIRD. ... KYVRNVVQSVNNTLLGQE. : 233
Btf4_Human	: ... KGYEDGGIHLICRSTGWHQFQ. IQWSDA. KGENIPAV. ... EAPVVA. DQGLGYEVAASVIMRG. ... GSGEGVSCIIRNSLLGIE. : 217
Btn3a3_Hum	: ... KGYEDGGIHLICRSTGWHQFQ. IQWSDT. KGENIPAV. ... EAPVVA. DQGLGYEVAASVIMRG. ... SSGGCVSCIIRNSLLGIE. : 230
Btn_Human	: ... QVQNGEIGICLECTSVGWYPEPQ. VQWRTS. KGEKFPST. ... SESRNP. DEEGLEFVTAASVIIRD. ... TSTKNVSCYIQNILLGQE. : 228

FIG. 2B

Cd80_Human	: ..VNOTFNWNTTKQEHFPDNLIPGMAITLISVNGIFVICCTYCF.....APR.....C....RERRRNERL.RRESVR.PV	280	*	300	*	320	*	340	*	360
Cd86_Human	: RLLSPFSIELEDPQPPDHI..PW...ITAVLPVILICVMWFCILILWKKKKRPRNSYKCGTNTMERESEQTKKREKIHIEPERSDEAQR									288
GA16817596	: A.....RELTLASIDLQSQMEPRTHPTWLLHIFPSCIIFAFIATVIALRK..OLCQKLYSSKDTTKRPVTTTKREVNSAI									306
B7-H1_Huma	: HT.....AELVIPELPLAHPNERTHLAVILGAILL..CLGVALTFI...FRIRKGRMDVKKCGIQDTSNKKQSDTHIEET									273
B7rp-1_Hum	: LTVGSQTGNDIGERDKITENPVSTGEKNAFTWSILAVILCLVWVAVAI GWVC.....RDRCLQHSYAGAWAVSPETELTGHV									290
Pro352_Hum	:DAHXSVTITGQPMTFPE..ALMVTGSLSVCLIALLVALAFVC..WRKIKQSC.EEENAGAEQDGEGEKGTALQPLKHSDS									302
Btfl_Human	: RESVTFIPESFMPSV.....SPCAVALPIIV.....VIL...MIPAVCIYWINKIQKEK									307
Btsf2a2_Hu	: KETVIFIESEFMPSA.....SPWVALAVILTASPWMSVIVILAVFIIFMAVSI CCIKKLOREKKILSGEKKVEQEEKTAQQLQEELR									276
Btf4_Human	: KTASISIADEFFERSA.....QPWIAALAGTLPILLILLIAGASYFLMRQOKE									318
Btn3a3_Hum	: KTASISIADEFFERSA.....QPWIAALAGTLPISLILLIAGASYFLMRQOKEKIALS..RETEREREMKEMGYAATFQEISIREKIQEELK									263
Btn_Human	: KKVEISIPASSIPRL.....TPWIVAVAVILMVGLLITIGSHFTWR.....LY..NERPRER.....RNEFSSKERLLEELK									313
										294
Cd80_Human	: VFKSSTSSCDKSDTCF	380	*	400	*	420	*	440	*	460
GA16817596	: B7-H1_Huma									323
B7rp-1_Hum	: Pro352_Hum									316
Btfl_Human	: Btsf2a2_Hu									392
Btf4_Human	: Btn3a3_Hum									405
Btn_Human	: B7-H1_Huma									368
Cd80_Human	: VMTVGVCRRHSYE.RKGEVLLIPQNGFWTLEMF.GNQYRALSSPERILPIKESLQVGVFLDYEGDVSVFYNMRRSHIYTCPRSAFTVVRP	480	*	500	*	520	*	540	*	482
Cd86_Human	: EWHIGVCCKNERKKGWVKMTPENGYTWGLTDGNKYRALTEPRNKLPEPPRKVGIFLDYETGELISFYFNATDGSHTYTFPHASFSEPLYP									497
GA16817596	: B7-H1_Huma									458
B7rp-1_Hum	: Pro352_Hum									
Btfl_Human	: Btsf2a2_Hu									
Btf4_Human	: Btn3a3_Hum									
Btn_Human	: B7-H1_Huma									

FIG. 2C

Cd80_Human :	560	*	580	*	600	*	620	*	...	-
Cd86_Human :									...	-
GA16817596 :									...	-
B7-H1_Huma :									...	-
B7rp-1_Hum :									...	-
Pro352_Hum :									...	-
Btfl_Human :									...	-
Btsf2a2_Hu :									...	523
Btf4_Human :									...	-
Btn3a3_Hum :									...	584
Btn_Human :									...	526

FFRL.GSDDSPFFICEALTGASGVMVPEEGLKLRVGTQSL

VFRILLTLEPTALTICEIPKEVESSPDPLVPDHSLETPLTPGLANESGEPOAEVTSLLLPAHPGAEVSPSATTNQNHKLQARTEALY : 584
FFCLWSSGKKPLTICEFIADGPERVTVIANAQDLKSKEIPLSPMGEESAPRDADTLHSKLIPTQPSQGAP : 526

FIG. 3A

ataagaagct gaattaaggt gatggcagtg ggggtgaaga aaggagagcc accatgcaaa 60
 aagtatccag gagggagaaat taacaggact aggggatggg ccatatttgc aagatgagaa 120
 atgcagaggt ctaagattct agctnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 180
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 240
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nngcataaac 300
 catattttcc ccagaggagg attagtagga aaggaagctg ctggttgga agtatcttta 360
 tagcagtgtc tgttcctcgg tttgctcaag gggacagtgt gccaggaaag tccccgtgga 420
 agggcaagga agaaggggaa gttaaagcca gtggcaggtg atccaagaat cttttctgtt 480
 gctagagcta tgttacatgc tgtcctttca tgctctaaaa ataagagtgc tggcaagtgc 540
 caggcctgtt ggtgcagctt aagatgatac ctttcttgga tatatatgca tctgaataag 600
 gaaggctatc ttctgggtcaa gctaaggat gccatgagca tttccctgtg gaaagcactt 660
 aattctgttc ccagttgtta cctgctgtaa gatctccctt tctaaaataa aaacaagaat 720
 acagctcact gaggacctta catttccctc tagctactga ctcatctctc ttctcctttt 780
 tatagcactc ttcttgagag agttgcctat atttgttgcc acatctttac ccattctctt 840
 ttgaacctat tcaagctttc atctgtacaa aactcactga tactgtgctt gtcaggatca 900
 tccatgacct ccatactgct aaatgcaact ctcaagagta tttggctcta ctgatcactc 960
 cttttagca ctgtgtttta aaatataggt tttattatta tttaggtagt gtgaggccaa 1020
 tatatcagga aatgactgtc gttgaaaaaa gtatgttgta ctacagatc ocaagagaag 1080
 gggggcacac catgccacaa agggccacat ggggaagcac cagggtcagc caggaggtgg 1140
 gtgggggggtg cgcaagatct ttattgtggt ttcaacagga agaaatgggt gaagcagggt 1200
 gagtggattt aggattagct gatataaata atttcagcag gctctggggc ataggggctg 1260
 tccctagtct tctggtactt ggccctgggg tgattaaggc agttgcatag tgttggaat 1320
 gtgaaagccc ccaataaatg aggcagttgt gggtagggc tctgaaatgg gttggtttgc 1380
 atttgaaagg tgtgtcatg ggcaagtggg ttactctctc ttagaggtta gaattggcta 1440
 accctgggag cggcagtcct ttcagggtca gcaaggcccc aggtgtcaaa gcatcagaat 1500
 acagaaaata aaatgcatgg ataatacaca ctgccatttg ctttgtacc cttcctttca 1560
 atcttctctg ctggtgacgg ctcttcacaa agatctataa atgttggaat accccatgtc 1620

FIG. 3B

```

tcagtccttg ggcactctct ttcctatctc tctgtaggtg atgtaatgca gatatccatg 1680
acttttaaate tttaacactt ctgcattgat gactcctaaa ttacatctc taccccaact 1740
gcctactaaa cacctccact tggctatcta ataggcattt caaaccaaat ctacaacaaa 1800
cgtaactctt tttccccttc cttaatttgc ttctcccccga gccttctcca ttttaataaa 1860
cagcatctcc attgccttag tgactcaagc cccaaactta ggaattttcc cagatttccc 1920
tctttttctc aaactatata tctagcctgt cagcagttcc cttcaggtct tttttcnnnn 1980
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2040
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2100
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2160
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2220
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2280
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2340
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2400
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2460
ctaannnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2520
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnngctc 2580
catttatatt tannnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2640
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2700
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2760
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2820
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2880
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2940
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3000
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3060
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3120
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3180
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3240

```

FIG. 3C

```

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3300
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3360
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn ttgtgtttta 3420
aatatatata cacacttaga cacatataac cctctttcgt atatcaatta tactttaata 3480
aagctgttgn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnctcgcca 3540
acagaccact tctcacccct actagccct ctactctaag ctatcagcat ttctgcaaac 3600
acattcctaa ccatctcac tgcctgtaat cttgccagca acctctccct ctgagcaata 3660
gtctattgcc tacaccaaag cttagtgtgc tcttaatgat gtaaagagg ttctatcatt 3720
ctcctgaccc aaaccctcca ctgcttttca tcacactcag agcagctctg ctgttgccctg 3780
atttagatgt atggctcaa cagatttccc ctgaagaaat gattccatgg ctgataaaag 3840
ttggaaagcc tctcagttt cagaccatta tcagattagc tgtgtgctct gtccctttcc 3900
tcaaccataa gaagtccatg gataaagaaa gcttcagagt aaaggagaaa gcatgggagg 3960
tacagcagga ccaaggtggg gcattgcag cccccaccct catcagagcc agttccctac 4020
tctccctgtc taaacctctt agtaagaggt agttcaagag aggggcaaac tcaattccag 4080
cactcaaaag cacttgacta ctttgctcag tcaactagca agtatttatt gagaatgtag 4140
ctctgttcta tggagtctta ttttcaagtg tcagactccc agacatccag tccaggtaaa 4200
gaagatggtg tccattattc atttgacaaa caaagttggg gttcaagggc cagctattga 4260
aaaaagctat ggaaagcttc atgagacgtg caggtaactg ccaatatgtg tggttcacia 4320
ggactgggtc atattcagaa acggccatta gaaaggaag aagaacttct catttgatt 4380
tataaagagt gtcttgttta ctottaattt atatottotc ttctccagga aatcaacct 4440
taacttctcc tccagctcc actctaccat ggtctgtcac cttccccaaa tgatttgta 4500
ttcccctgtt ttcaaaagtg aacaagaac caaagacca gcaaagtttc acaaggccct 4560
gagactttca attgtctatt tcagatcaaa tacagaac atg atc ttc ctc ctg cta 4616
Met Ile Phe Leu Leu Leu 6
atg ttg agc ctg gaa ttg cag ctt cac cag ata gca ggtaagaaag 4662
Met Leu Ser Leu Glu Leu Gln Leu His Gln Ile Ala 18
gacaaagga gaggttaag aaagaagagc aggtgggtgg ttctagccaa agccaaaaat 4722
gagaatgtgg cctcaggct gagggcttcc tttagagga cgtatgattt ctgggctatt 4782

```

FIG. 3D

```

ccaagcacca caaaaaaaaa aagagtcccc atggtggcctt atacatgcc aatgtccctat 4842
ctgacagaaa cgggtgactga gaattattgct ccatctattc ccactatcca gtgagggttaa 4902
tgacaagaag acaggatcac tcagaccatg taaatctaaa ctgatacaag agggcagggg 4962
ttgagttccc ttaaagggtga gatgccaaag agctgtcccc ttcttttctg gcagggagag 5022
taaggagaca atggccaggg aacacogtta ctctaaagat aatgtcttga agacattctg 5082
catattatta gttgtttctg tgagtttctt ttttgaaaag caacaatagc agccgtttgg 5142
cattcatacc ttaatgtggt ttactgagtc ttctaaaaac ccaaatgaac aatgaacctt 5202
aaggctatcc ctttggactt gaagaaagga ctctatttgg aggatgaggg tgagcagaaa 5262
gaaaagcagt ttcacagttg gttgttctcc tggggaaggt agttcagacc attcgagggt 5322
gtagttagaa ccatgagtgc actattttgg atgaacacca ggagctaaga gagtaacata 5382
gaggtgtgga cagaggatta agtcctcaag acaatagccc cagccccatg ggaaatcatc 5442
tttctgctca tgattgagaa ataatggctc ccttggcact tgataacctt tcgaagagct 5502
ttctctccc tactagctgg ttccagatca ctcttcccc agtcacattc ctctcactca 5562
cttgagctgc ccagcctggt ctggcactag agacatgcac ttggggccct cctcaaagga 5622
agaccctgag atattctgct tacttctact ctgctcctgc ctgcagggcc agctaaagga 5682
acttttcatg ttttctttgc aaggaacctt gcctggctgg catttttagag acaagcaaaa 5742
ggggcaataa ctctcttgc acaaaacagc ttcaagtttc catagagtga taagggaat 5802
gagggccaaa agacactgtt ccccatcctg tggcaggact gggggcttca ggagaaaact 5862
tggggaatgt gtaacctctg tgggtttgta gcttaaaaac actgagatcc tgggttttct 5922
gtctttgttt tttgcctttt ctcttaggaa aggagtgagc tagggtgaca aggggcaaca 5982
ttttttatcc ctcatggct ctttctacag aggaaggatc ttttcttcta agataatcag 6042
cacaagacaa tgaagatagg cactagctcc cagttaggta tactaatggg gcaaaaggaa 6102
gagcatttnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6162
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnattttgca ggtgaggga ctagaactca 6222
gaaaaggta ttaatttccc caagattaca tagttattag gtgacaccgg caagatttca 6282
acaaagctaa tgtcctttct actttactgt gctacatga tgatggtaat caaaaatggc 6342
agacaacca taaatcttcc aactttggaa taggtttttg cactgaagtc tgaatatgga 6402

```

FIG. 3E

```

taogtattga atgtttattc tggatattca cagaatcaaa aaatatgtgt aatgaattat 6462
gttgctgaat taactgaaag gaaagtaaaa atgtagcgct ttctcatttt cttcacgaat 6522
ttggaattct tttctgcttt ccactatgca gataacatca gttcagacaa atattaaata 6582
cctacctaaa ttagaatgcc ttctcctcat gggatttttt taaaatcttg tcatttcattg 6642
tctctttaat taaagagttt tgatttcaga ggagggtacc tgcaaaagaa aacaacaaaa 6702
aaactaaagg atctgagaaa taattagtgt ttacttctgg ggaggggagg aggtctggga 6762
tgggggtaaa aaggatagtc ttatctatta tgtatattca ggtttttggt ttttacaaga 6822
agcatgtatt agnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnca agaaatttct 6882
catataaaaa tatgaaagta atcagactgc aacactcagt gcctgagaca gagctacagc 6942
tatcagggtg tccagacaga cagaagatta cattttcttc cttgctcctt gtacagcccc 7002
agacctgcat gcttcattga aaagaaaaga agatacctga attaaatcaa tgtgatgctt 7062
agtaccctat cagtgcacat ttcttttcnn nnnnnnnnnn nnnnnnnnnn cacttnnnnn 7122
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7182
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7242
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7302
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7362
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnca 7422
cttatggcgg tattctcagt cattacaaat aaataaaaac aatccatatg ccctggagaa 7482
tttgattcca ggagtaggtc tagaagaact tcaactggag aatggataga gaaatcatgg 7542
tatatttgca gnnnnnnnnn nnnnnnnnnn nnnngatagca tgtgaataaa ttaattacaa 7602
aaacatatga ctacatctat tattatatag catgtagata aattacaaaa acatgtaact 7662
acatctatga atcttagagc ataattattga gnnnnnnnnn nnnnnnnnnn nnnnnnggcc 7722
agaagataac acatagcaca atgtcctttt cataaataaa tatattgctt aagcatacct 7782
tatatataga agataaagct taaaaagtaa agaagag 7819

```

FIG. 4

```

attatcactt atgaggggtgn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 120
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnctgcc 180
ctttattttt tcatgaaaga aattgctgaa gaggactaaa agaagtttta gtaagcattc 240
aataaatgta tgttctttat agtttccaaa tcagcaaata tagacatcct gcatttttaa 300
ggagatttat atattttatt ggacatgctg taattttatt aaccacttcc ctgttggtag 360
acattatttc cattttcttc tgctagatta atgcttgaaa aaaatgtgtg cctoctaaag 420
actgtgatga aagttgcctc tgaataaaaac tcaaacaaat cattaatcat taactctttc 480
cttacttgta tgctctttgg atgctctact gtgttatcta taaaataaag tttgaagtga 540
aaaattaggg taaaacattt tatatcattt ttaaaggata tatacatgga tgtacttaca 600
tatgcatggt taaatttata taccataaca tttatttctt tttttaaaaa 650

```

FIG. 5A

tatatttggt ttttttctta catttttatt tcaaaatcta aggacatctt ataaccaga 60
 aatatttttt atacctgtc atgtcttaga ggaaagagcc acccagctt tttttcattg 120
 atgtttttct tctctcttcg tactccagag gtagatgaaa accagagggc cacaatgacc 180
 atggtgatgc ctgaggtcat tctggggcac agacctcagc ctaggttact ccaottogcc 240
 tatcttttaga tccaaaacta ccctgctgac tgctgagata aacaaaggag aataatcagg 300
 ttggggaaag gattttctatg cgaagacatg tctccatgca gtctctctac actgagcaga 360
 gcatgagtca ggtgcttaga gcaggatttt gtccataacc aggaacttca gagttttctg 420
 aagaatgtgg ctatgtaaag cccccccca cccaccctt acttctcaag tacattacgt 480
 ggcaagtctg aaaaaactta cacttctgtt gttaaagtgt ggggataaaa tataaactta 540
 gtttcaagag gaagctatct tgggaggtaa tgcaaataat tcgttggtgtg tttctgaat 600
 aagtgcagg tgctgactac cattgatgct tcattgcaat aaaatgcaa gctccccaa 660
 gaatttttga aatgcacaa gctaggtgtt ctaatctagc aaaaggacct gcatacatga 720
 atttttcatg cttttgccaa gtcttttgcc ctttagttta gttaagggcc ccacatgaat 780
 ggaaagcctg tgttgctcagc ttaattttgt agttgtggaa acctccagt tttctcttt 840
 gtctaatacc ttcaggagtt caatcctagg ttgaagctta atttaataac catgtggcat 900
 gtaaagtaga aaacaaaaca tcttttctt agcatacagc aaaaaaaaa aaaaaactca 960
 ctcatggatg tagtgtacac atgccagtgg atatatagtc ataactgcag tcattggtag 1020
 cacagaaata aatgtgcatt gaagacacag agannnnnnn nnnnnnnnnn nnnnnnnnn 1080
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnn 1140
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnng 1200
 gctagcccag tgtctgactt actgtgttta agaaatatca actattacgc tacttcccag 1260
 tgacagtcca aatgcagacc agtgttataa ctctacnnnn nnnnnnnnnn nnnnnnnnn 1320
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnn 1380
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn aggcaagggtg 1440
 gtaaaagacg cttgcacgtg caaatgttac tttgtgtaac tnnnnnnnnn nnnnnnnnn 1500
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnn 1560
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnn 1620

FIG. 5B

```

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnct tagtgaatct aatttgcagc 1680
tgggcttgag aaaaaacctc ttcatagaat tgtttgcac agtgtcttga ttgcctctgt 1740
aacttacaat aagcaagaat gtttcaggat ttcaaaaatc tattgcattg cctaaacctc 1800
ttattttgta tggagtaatc aagctcaaag tttgcatgtc ttagaaactt tacttggggc 1860
aaaattagac caagtaacaa ttaatcttct aggtattctg agctattcag acatatgatt 1920
catgtttgct aattgctctt ttctcttgta aatattagct gaaaaatgtc acctgtctga 1980
caagtagcat attttatgcc tatcactcct ggcacgcatt cttacaaggc agacaggaaa 2040
aataggaaga aaatggactt ttatcaaagg ccagggcagt aaagagggga gttctgctgt 2100
aagctaaggg gagttccaga ggaagttata ggcgttcctt ttcttatgac aagaaagcat 2160
agtgcagtaa ataaatttgc taaatagatt caacagtctc taccctaaagt catctattta 2220
attcttggtt ttatgcagac tcagcaacta acctctcttg taagcccat tttcttcct 2280
gtttctgtt tatcaaagt aattaaacaa gagaagtatt atagaagagt aaaagtagta 2340
ggtaattctt gaacttggca tatgattact acatatttga tgaatagttg aatattattc 2400
ttcaaggaca gattggattt ggtatcaggt ggctctgcat taagttataa gggacttaat 2460
aactcaagta ttaaggacg gcttccatca taaagggac tgcccttaag aggggtcccat 2520
tatggagatt ctgaggtgag agctattcca agtgtgcagt ggattaaaat aaaagaatca 2580
tacaggaaat ctctttttac atgccttatt ccagggctct tgcaacctgg cacagcaagt 2640
gcagatatga ttagcattgt ttacacatg tacttcacc ttatagccct gccctgtgc 2700
ccctctgca caaaagaatg ctgggcacac gtgaactcct ctctgtagaa aggcacatta 2760
atgttctagc catggttaaa acagggatag aggcaagcca aaaatgtcgg tcatttgaaa 2820
taaactctca gtttgtgcat atcactatca agtgtgctgt gtggcaatta agaatgccaa 2880
tttgtgtgat cacaggcaag ttgcagttt atgaaaggaa agcagaggtg aatatataac 2940
cagggtcac cttctcttct cctctctct cttctgtca tttatttggc aagctcttaa 3000
ctagaacttg ctatgtgcta ggtactggat atatcaaagc aaactcagcc tggctcttgc 3060
cttcaaagat ttgcaggata gtgggaagaa aaacttgaat cagaggacat ctgcagtggg 3120
aatcattcaa gcagcagaaa acccaaaagt tacttatact gtgaaatctg atcagagaat 3180
ggactgtcct ggtagtaaa atatcctgga ggataaagat tggccatgca ttccacatat 3240

```

FIG. 5C

```

gaattaccac tttcccaaga attaaaacat ggtacgaaag aaaggnnnnn nnnnnnnnnn 3300
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3360
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3420
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3480
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnna aacaacttct 3540
ggaatagcta atgcttagaa gcagctccca aatatttgtn nnnnnnnnnn nnnnnnnnnn 3600
nnnnnnnnnn nnnnnnnnnn gcaagctcta ctgaacataa tttgatctaa tcttctgtga 3660
ttattcagaa actacttcaa gattttccta tacctccatc ataatgaata cccattcatt 3720
aatgatggaa gcagcctaata tttgtcattt ttcacacttt attgatgtaa cactaccttt 3780
actagtttgg ccactcotta tgcttttttt atagaactat ttagatcaat tcaactttta 3840
aaaaataaag ccacataccc ctgtggtaga tgaaaaacaa gtatcatttg cactggtaaa 3900
tagagaatag gaagaaaaat aaatgcagtg aaaataaagc agtggtatca aatcctaccc 3960
agatactgtt atctaccogg aagcttcctg tttgattaaa aggaaaaata gccagtgtta 4020
gagggtgtgga agtctagttg aaattatatg caattgaagg attaaaatag aattgaaaag 4080
ggaataaatt cctctctgaa taatttaact ccctttaggg tttgattctg cctcatctaa 4140
aatcatctta catacttcta gtggcgtgtc cctcacattt tggtaaactc tgnnnnnnnn 4200
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4260
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4320
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4380
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnccat ttttcctttc cctttattgt 4440
cagaaaaatag aaagcatcta cagtgggctt gtatgatgtg gtgggtagaa atacctgatc 4500
tgattnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4560
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4620
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4680
nnaccatcac tggtagagagg aagtgaacc tggcacaaaa atatatggat taatcaatat 4740
ggattgaggg aaacaaacct ggagaatagg atgtgaaggt atttaagtaa catgagctca 4800
gaccttgatg gtagggaagt cgaaaggaag cattttgttc ttatatgaca gatgacctgg 4860

```

FIG. 5D

aatgactgca gggcttgggg ggtcagggac tggaggtggg agaggcctct gagagcaagc 4920
agtgtgtgcc accagaagct cttgtctggg tgcccagaga ggagcaaagg gcagtcagct 4980
gcacaggagg gaatgtttgg aggagagagc cacctcagat cagcgggtca agaatcccac 5040
tcttgcccag atggatgggg caaaggagaa aaaggattcg ccacgggaat gtccagataa 5100
gacaggtgcc ttttgaaaaa tgggggtgag atgggtctca ggttacactt cgtaagaact 5160
ggaatgtaaa gtaaaggcag acaatgacaa aatatcttgt tttcttttca gct tta 5216
Ala Leu 20
ttc aca gtg aca gtc cct aag gaa ctg tac ata ata gag cat ggc agc 5264
Phe Thr Val Thr Val Pro Lys Glu Leu Tyr Ile Ile Glu His Gly Ser 36
aat gtg acc ctg gaa tgc aac ttt gac act gga agt cat gtg aac ctt 5312
Asn Val Thr Leu Glu Cys Asn Phe Asp Thr Gly Ser His Val Asn Leu 52
gga gca ata aca gcc agt ttg caa aag gtg gaa aat gat aca tcc cca 5360
Gly Ala Ile Thr Ala Ser Leu Gln Lys Val Glu Asn Asp Thr Ser Pro 68
cac cgt gaa aga gcc act ttg ctg gag gag cag ctg ccc cta ggg aag 5408
His Arg Glu Arg Ala Thr Leu Leu Glu Glu Gln Leu Pro Leu Gly Lys 84
gcc tcg ttc cac ata cct caa gtc caa gtg agg gac gaa gga cag tac 5456
Ala Ser Phe His Ile Pro Gln Val Gln Val Arg Asp Glu Gly Gln Tyr 100
caa tgc ata atc atc tat ggg gtc gcc tgg gac tac aag tac ctg act 5504
Gln Cys Ile Ile Ile Tyr Gly Val Ala Trp Asp Tyr Lys Tyr Leu Thr 116
ctg aaa gtc aaa ggtgagtggg gtcaaggact agaattccatg gaagctctct 5556
Leu Lys Val Lys 120
ccaacagagg atctgcaagt cacagaaacc cattaaagggt agctcaagca aaaacaagca 5616
ggctgctttt aaggagacag ctatttcaga gaaaatgaaa gcatctgtct ggaaataatt 5676
tttgacatct gagtacaaag cagccgaagt acaagtgaaa gggggtagga cctataggaa 5736
taaaatggga ctggaggaag ccaggaaaat tagtccctga aatgtgggag ggtatgaaaa 5796
ataagctttg cctaattcac aattctccca tggaacatcc ctgaactgat tattaagata 5856
ctctttttca atagtttata ccctgaatcc agagttttta aaaccatggg ttgccgcca 5916
ttcatggatt aaaatatcaa tttagtgtgt agcaaccaga tgcacgtttc ccgccttta 5976
aaaaataatg tatagaagag aatagacaga gtagatcaga cgatatcaca gagtaggact 6036
gagtactgta aaactaattt ctgaggggacg tgtgtgtgtg tgtgcgtgtt gggatcatgg 6096
ataaattttt ttttcttac tttggatcat aaaaagttac aagtttggaa aacactgtct 6156

FIG. 5E

```

aaatgcaagn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6216
nnnnnnnnnaga gtcccatgaa gtctatagct gtccctattc ctctggattc agggatctct 6276
ccactccagc acaattgaaa atctaaatat aaagagaatc ttcacactct tgtttgttct 6336
agaaaagggtg atttgaggaa agacatataa caactataaa aaatagattt tgcttgttca 6396
ttggotttatg gtctccaggo ttgaatgctc tgagataaat gatgccaata tttctctggc 6456
ctcttcccat cccacgcatt ggacotcaga tggctgttac tgtcttctag agggtttgtg 6516
ggtttttgcc ccaaaaaacc attaaccttg gcagaaagtg tgtgacttta tgatctggta 6576
caaagaagga caaactagag ggactggaca tgaggatgaa tattgtgttc gcccttatnn 6636
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6696
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6756
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnaaa gaccctcagt 6816
tgttacaggg gcagtgcct cctcacact caaccatcaa tgagtcacca ggaaagccat 6876
tagcctagat gtaactgtt tctatcttta ttgcatttcc tacatccagg cagcagctgg 6936
gaggaactct agaactga agtttgtctg agttccotta atgtaaggct gtacattctc 6996
aggatgcctt gatgtactcg aatatctgca accctaaatc accacctctg tttttattga 7056
tctctatctg aatgctgtat taatgggcca ggccttctgc ccattctctc aaactgagaa 7116
ctgtctctca ttctgggga ggcacctgc ctactcotta cctagatcag ggatttctca 7176
gttgtggaga gatttgttcc ttatagtgtt ggtcatcaaa ctgggatatt tggggattac 7236
aaagactttt caagggatgt atgggcacag gcagttttag gaagtgagtt cctagatcct 7296
catcttcccc aaatactctg tcccaaaatt gacgagcctg acaatgtgca tgccaggcaa 7356
ggctcttggg gttcccctaa aacacttctt cttttaagcc taccactcac tcatcatgaa 7416
tatagtccat tgtcccaggg tgtaaaaccc totatagtgt taaataaaag aatgattggg 7476
aacattgaca cctgatggaa ctgttatgac taaaaaccot tttgcaaata atgtggtatc 7536
taattttctg ctttcaacaa aattgaagga ggccottata aagttaataa ctgataatca 7596
aaaatgagta atttttgcca tgtaaatcag gtcaaagaat gaaatggcat tgctgtaacg 7656
aaactgcttc cattcccatt gatttactca taogaacaag attccttagc ctttataagc 7716
tacaaaaaaa tgaaaaatag aaatagaatt gaggtgaat tctattatat aaaatcattc 7776

```

FIG. 5F

```

caaccatgtc atatggttct tcggattcat gaataatttg gaaaagagag ccatatccat 7836
ottattaagg gacacattcc caataaattt tcatctttca tgtttaataa ttatcaatat 7896
tcataacatt ttacattttg atcaaatatg tgttaataat aatagaaatn nnnnnnnnnn 7956
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8016
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8076
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8136
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8196
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8256
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8316
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnattag atcttaatgc agaacaccct 8376
gaacatttaa agcttcatag tcacaagaga aaagttttca tttcaatagc tataaatatt 8436
ttgttgttgt aaagacatat aacgataatc aatacaaaat ctgtcaaaca aaaatatgtt 8496
acattaagat aaaattctgt aggggaagtg aaattggaag tgagtttcaa tgaatgaaaa 8556
gaaacaattt agacagagaa gnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnncttagaa 8616
ggaattgaat agattagggt ccttaccocaa aaagcctctg ttatttgtct tatttattta 8676
ttctcttttt tccacattct ccagtctcat tccctttttt taacacagga aattattcca 8736
gcatgtttca tacatattct tttgtttgta agagcttatt taaaatatgt aatattgttt 8796
tagatgcata tatttttttt cttgtggaaa ctatattgta ctatatatat atattttaga 8856
aatggacaca ttannnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8916
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8976
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9036
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9096
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9156
nnnnnnnnnn nnnnnnnnnn nnt 9179

```

FIG. 6A

```

tagatctcag ctttcttgag gcagggagcc atatctgttt aattcactca gcatatactg 60
caaagaagca gnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnta 120
gggagaagtg cagaataaat atacccaact ctttactatg tatagacatt atctaggtct 180
ttatTTTTTT tctcttctta atctcaaaga aaacagagga aaggaggaag taaaaagtaa 240
atTTTTgcct gaagatgttt ggaaaaaata ccaaataaag tgagatagtg ggtaatctag 300
tgatTTTTtat ttttccgtcc tctttctggc ctccaattgt gaaataattt atagcactgt 360
aagaaagaag ccacaaattg tggtagcttg gaccactgtt gaggnnnnnn nnnnnnnnnn 420
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 480
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 540
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 600
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn acttaatat gatagtgata attttattca 660
tttcacatgg catgaagtac caagctctat aggaatcaga aaataaagtc ttatttcttt 720
ttctttctta ttgtcca gct tcc tac agg aaa ata aac act cac atc cta      770
                Ala Ser Tyr Arg Lys Ile Asn Thr His Ile Leu      131

aag gtt cca gaa aca gat gag gta gag ctc acc tgc cag gct aca ggt      818
Lys Val Pro Glu Thr Asp Glu Val Glu Leu Thr Cys Gln Ala Thr Gly      147

tat cct ctg gca gaa gta tcc tgg cca aac gtc agc gtt cct gcc aac      866
Tyr Pro Leu Ala Glu Val Ser Trp Pro Asn Val Ser Val Pro Ala Asn      163

acc agc cac tcc agg acc cct gaa ggc ctc tac cag gtc acc agt gtt      914
Thr Ser His Ser Arg Thr Pro Glu Gly Leu Tyr Gln Val Thr Ser Val      179

ctg cgc cta aag cca ccc cct ggc aga aac ttc agc tgt gtg ttc tgg      962
Leu Arg Leu Lys Pro Pro Pro Gly Arg Asn Phe Ser Cys Val Phe Trp      195

aat act cac gtg agg gaa ctt act ttg gcc agc att gac ctt caa agt      1010
Asn Thr His Val Arg Glu Leu Thr Leu Ala Ser Ile Asp Leu Gln Ser      211

aagagctgcc ccacttcct aggtctatca gttaggggtc agacaagaaa cagatggcat 1070
actcgagtga tttgaggaga gtgtaataaa gggactgttt acaaagggtg gatcaccatt 1130
tgagaaaact acaaaggata gtgcagaaca ctggggcttc aatgttggga gggcaattac 1190
cactgttggga gaagttactg gaatcagaag ggagctgtag ggaaagcccc acttcccagg 1250
agctgtagcc acagaatagg gaagctgcc catgcagcga ctcaaaggg tggaactgg 1310
atgaatgaat accccaactc atttctctcc caccctccaa tctctgcta gcacctccca 1370

```

FIG. 6B

```

ttggctgaac ccagctagaa gtcagagaat acaaggggtcc actggtgtat tccataaaag 1430
tcaacttctc agggctcaga gcaatattga catgtacaga atagatctgg agaggaaaca 1490
gaaaatatct agtacaatag ctaatcactg tgattcatgc acagtgtcat gagccagcag 1550
gatgaatatt cctttgctgt acttgctgcc agtcagctgg ttatggggtt ttccaagaaa 1610
tttggctctc aacaaaattc ttcagagcct ttactgacta tgctggatat ttttgggaagg 1670
gatcccatat ttttgaactt catacagcag aatttcaaac aatottggga aaataacaac 1730
ttttatctgc ccagtaagga caactaacac ctagtatcat aatcatttcg taagagacag 1790
gtaatttcac caccagtgat atat 1814

```

FIG. 7A

ctactgagaa gggatatact ctcagactaa aggacagtcc ctagtactga ttcaatctgg 60
 ctttatagaa aattcactat attgtcattg tatttcacag tttgcccttt gtcttagctg 120
 gtaagacaga gcctatgata aggacttgtg tggcatgcag gtattttaatt ggcaacccca 180
 gagggcagaa gcaagagatt taggagttta agagagggta atataagagt atattatcaa 240
 agttgtagtg tggacaacag aaactcaaatt attcaaggac cagcatgtag acagcctcct 300
 aagatgtcta ctcagacaaa gaatttcagg tgggaaggact tgttcactctg cttcacgccc 360
 attggttgac aggaatatga actccattct gctgctgggc tagacatgca tgtgggctga 420
 gtgagctttc ccagtatcc gtagcatcag aaaagtcgca gggcagaaaag aaaagtatcc 480
 aatttgaggt gaattactga ccttgaagtg agtghtaagcc taactagaat tctaccccag 540
 ctggctgaag tgaaaggtga ggctgagagg aaataaggca ggactgcaca gtccccaatt 600
 gtactgttca aatccactca tgccttcat taagtcagct ctgccactga gccttcagc 660
 tgggaggcag ccacaatctc tgcagaagat ttaatataca ccagtttgtg gaacaagctg 720
 tnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 780
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 840
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 900
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 960
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 1020
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 1080
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 1140
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 1200
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 1260
 nnnnnnnnnnn nnnnnnnnacc tgcagcactg ggatagccct ggtacagacg tagaacatgc 1320
 ttgaggaggt tgtcaggagg aatgagtta gaccttgac agaactacca ccatcaagta 1380
 cagttgggggt gaggcagaca tggatgagt tggagcatca gagatatctg atgctttatg 1440
 ccaaattaaa attaatTTTT tcatggagt acactgatcc acagaccaga ctccaagaac 1500
 tttgcagtga ctaaataccc atctcatcat aactttcctg gtattttctt ctggaaaaaa 1560
 ttcttccttg atacagtttt cagaggcagc tagatgcaact gtcattctct cctttttccc 1620

FIG. 7B

acttccctac ctatccacaa ttactaccc aatgccaaaca ctaaagttag cccaacttcc 1680
 ttctaactaa attattagtt tagaaggaaa gagaggagtc atgctaagga tcttaactgn 1740
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1800
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1860
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1920
 nnnnggggag aggaggggag aggaggtagg gaggggaaag aaagaaacta gaaatccatc 1980
 aatttttagga ccaacttcag gtaaaaaaat gaattaggca agttggtctt tcaacattct 2040
 ctacctctct ttatatcatg gttgagacca cagacttctc acctcatgaa agatgaactc 2100
 taactaatto atactaaagc taaagcctct aaagaggatt aaatatgagc aatcccacga 2160
 gaactttttt cccctggaat tgtttattca actgtcgttc gttatatgga atttcctgcc 2220
 tggttaagtg taggccagta ctttgatga attgtagttt tctagaaaga cgcttcttat 2280
 ataagaacct ctccaggga acaggggact gtatgagatg aattgagaaa taactttaca 2340
 cactgatta tgtcagtgtt ctattctgca tggtagagat gtgaaagggc agactgacca 2400
 ttgctctgga agcctttacg ctgtgagaag ttaacagtgg agtaaaatgg cactccact 2460
 ctcttcatgg aagccaacat ggcttaactaa atagtcaaca accatgggag agacctgtgg 2520
 ggtcttcatc agagctcagg atctcctagg gtatcactca taaatacagc catcagggag 2580
 atggagaaat ctttgtgcag ccagaaattc tcaacctggt ttaccctc cttcccaact 2640
 ttgtattogt cctactgttt actgacatgg atcctotgct tcattaacca tcccttctc 2700
 accacatgct ctctgaactt ggctgcacct tttctacctc catgccttct ttgctcaggt 2760
 tttccacat aaatatcatt atttccctct ctactagctc caagcccacc ctctctctgg 2820
 ggcagctcag tcaactccagg gcacaagggg gtctttccct catccacat ttgagacct 2880
 actacctgga ccatttggtt gccttgtaac tatgcttgcc tttttaattg ctattttatt 2940
 ttccatgtat ttccattgtt cacacaagtc ttctttattc cactaagg caaaagcaga 3000
 gtctgtgtt cataataagt gctcaacaaa tgttgggttg attgggttg agattccatc 3060
 ttagataatc gcagtcccat catgccagct accagactgt gtggacagcc aggtcagagc 3120
 agccaaatga tattctagct tgtggcacia ataccagcaa caaataacc aaagtcacac 3180
 atctgcctct gagttcctgg ctctatttct tcaagggcat ttttaagttg tcttatgact 3240

FIG. 7C

```

gttccttttc tactcattct cataaattga gctgtggact gctgtgaccc acaagcttct 3300
ccggaagtca atgtataaaa caaacacgga aacgaagagt atgggtgggtg gaggggtactc 3360
cactgactct agaatggatg actgaacatt ccaaatttca agcacaagtt agggagcaac 3420
agatcatttt ccttttgaaa tagggtttct totgtcagc cagttgttgt attttcatta 3480
ggaaatggaa tgggactaca gcacaaaaaa taaatataaa aggacccttg tagggctggc 3540
agaaaagaga atccttccta ggagacctgg aggtgattcc aggcnnnnnn nnnnnnnnnn 3600
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3660
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3720
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3780
nnntcagaaa gtgtgcaaac agtaaaaaaa aatggtatat ctagcaagtt gcatgcctta 3840
cttgtgagtt catgaagttg tggcaaggat aagacaaata ttttttgcca ttgcatcatt 3900
atatcattgc taagagtatg ccattattnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3960
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4020
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4080
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4140
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4200
nnnnnnnnnn nnnnnnnnnn nnnnttaatt aattaaatta aagaaacgac aaaagagtat 4260
gcaagaattt taaaacaact tagaggaata tgtatgagga tacaggctaa gctaccataa 4320
tgaagagacc tcgaaataca gtgagaagcg agacagaagt atctttcggt ccatgtaaca 4380
ctcaggtggt tcagagcagc taagcagcta tgttccatag agtcattcag tgatccagat 4440
tattttcatc tgttgctctg ccattctcca ggatgttggt cctataaaaat tgtcaaagct 4500
cagtcagtgc caaaccocatg tttcaacctt cagaaagtaa acgagtgggt gaaaacacat 4560
tcaatgtttt aaggccaaga ccttgaaaac tcaactctctt agcctgaact tagattacat 4620
ggctgggccc acttaactat aggggaggct tggaacata gtctctgaga agccatgtgt 4680
ccagctaatt ccctaatact aaagttgaaa gaaagaatgg attaaccagc agtataccac 4740
aaggtaacaa atgactagga ggatcaggct aggtggacta gaaaagagac agtcaattca 4800
gtgcaacaat tccatattga cacttttcat gtagctgttg cttggctcta tctagagagg 4860

```

FIG. 7D

```

actcagaggt agttagata aggcctttgc cctccaaata cagtctaagc agactgattt 4920
cctactggat gttcaacttt ggagctttca gggatgagta gggcttctgt acgtggaaga 4980
gactatgagg gaacctgcac aggacaaggg tttgcataaa gacactgagg tagggacctc 5040
tctgtttgtg gggacagtga gaggcccagg tctccttgac tcacaaagtg cttactaagc 5100
acttactaga aattaagaag cagattataa tcaatatggg ttatccaatg tttggatgag 5160
caaggctcct tatcttttct tcgttaatgt taatcacact cttttggatg gagacaaata 5220
tctgtggggg ctnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 5280
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 5340
nnnnnnnnnn nnggctcaga ctaaataatg tetaatctct tctccagtaa aacaatccgt 5400
ggttctcaga tagcactgtg ctggaggtag tggggtttga gggctgggaa gttgggagga 5460
ctgagccctt cccgctgagc agtttcgtcc agtttttctt gtaccagcct gtcattgttta 5520
ttccatgtga atgactccag aggcaaaatt caagcttttg aatagggcac aaattaactt 5580
gagtaccctt tcatttccct gtaggtgaac actcctctag ccctgccttt tgtcagtctg 5640
gagcccttgt tetaatctgt acacaccaga ggactttaca aggctttccc cagcctccag 5700
aattattctt ctgatccacc ctctactaaa ctcacccttt cctcagtgtc aggacgttga 5760
aaaaccgaaa caaggcaaag ggccaattgt aataattcac actaaggcat gagtgactag 5820
gtttagtata ttaacactac ctaggatatt ctatttcttc caaaaggatc ctgttaatcc 5880
ttgaaattta acaactaatg gtatagattc taagcactgt gagtacttgt cagtggggga 5940
aagacatttt tgggctgaga gactttgcca ctgnnnnnnn nnnnnnnnnn nnnnnnnnnn 6000
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn aagaaagatg attagatnnn nnnnnnnnnn 6060
nnnnnnnnnn nnnaaaaaat aacatgagag nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6120
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6180
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6240
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6300
nnnnnncaca gagctagccg tgttggtgt caccactca tgtggccagc ctgttggtct 6360
acctcttagt tgccatgtaa caggattctg gtgcttttcc tttgccagg t cag atg 6417
                                           Gln Met 213

```

FIG. 7E

gaa ccc agg acc cat cca act tgg ctg ctt cac att ttc atc ccc tcc 6465
 Glu Pro Arg Thr His Pro Thr Trp Leu Leu His Ile Phe Ile Pro Ser 229

 tgc atc att gct ttc att ttc ata gcc aca gtg ata gcc cta aga aaa 6513
 Cys Ile Ile Ala Phe Ile Phe Ile Ala Thr Val Ile Ala Leu Arg Lys 245

 caa ctc tgt caa aag ctg tat tct tca aaa ggtaagtga ttttattcat 6563
 Gln Leu Cys Gln Lys Leu Tyr Ser Ser Lys 255

 ggtaacccaa tgcactgggt gtctgcagca tgagccactg ctttgactg caggcctatg 6623
 gcttgctgct ttcattgctaa acccactcag agcttatgaa ccactttgag cttgtcttga 6683
 tgattatattt tccccagaag aaaatggctc tcatcgtcag tgagctgaac ttcttacact 6743
 gagtttttta aagggaatgt tttgttctta tgtctgaaag agtttgtott attctttgag 6803
 ccaagagctt tcatcagcct catgagagtg atgttatattt ggcaatgcag agagctacgt 6863
 gctccgattt tgctgggtggg aggttgccag gatcctttct gaggattcct tccattttca 6923
 cccctctttt cccagtcctg gatatgacct gggttaaacc cccccctct cccaggaatc 6983
 tcaacctcac gggtgggtaa ggaaaggaga aaggtttgtg aggccatttg gggataagga 7043
 aacagctggt tgggtggtgca ttaacgtctt tcagcagctc ccttcgagtt tctccttagc 7103
 ctgttgattt cttaccaaca cactcctggt ctgttgtagc agctgggaca gagcatgctg 7163
 aagcctttca gccctgattt cattgcttca ttgttcattg gtctgtcttt ggtttcctgg 7223
 gtggagcctg cccacaaaac cccagaatg tatgcaggcc tagctggtgc tttcctaaac 7283
 ggctcccttg totgcactca atgaacttct ccaaagatct atacatggcc tcatctatag 7343
 aaagagaaat gacatgtgga aataattcag taggagtttg cagcagcact atctgaggac 7403
 taggggaatt ttaagtgggt gttatcttac atttatactc ataacttcta tattttcatc 7463
 tgccataaaa tattgtcatg ttctatttgt ccattgccct atgtgtgtat gtattcactt 7523
 ggggtgctgac cacaatattt ctaactgtag aatgcaagga attggtgcca aannnnnnnn 7583
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7643
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnncttca tggtaaagga 7703
 ggaggtatgt ccaacagaac ttcgactttt aaatagaacc acttcagaga gttgtgtcag 7763
 gtgcacctca gttgtcttat cttctgccat tcttctttta cctctcacac ccataacctca 7823
 gggttcaagg cctggggcct gaggactcct taataacttc agaaatgagc agctgagtgt 7883

FIG. 7F

tccgttccag ctgtctttgg gagaatggaa tggagtcaca ctcaaagata gagtggaaat 7943
 aaatcctctc ctcatccttc accccaatct taagagtggag tgaggatata agtagctccg 8003
 agctgggagg taaagctcaa gttotaactg tgattaggag acctttctta caaataagaa 8063
 ttaagtgaat aaatgtgcaa acaatttctt ttatatTTTT aatgaaccag agagaaatca 8123
 tggttgccta tataaccctt gtotccaact cacttgcatt cagatctgct ttcttacatg 8183
 tgtctgccat gcacacaaac ttgtgtgcca tggaaaaggg ttgagaactg ctggtgatgc 8243
 agacagagct ttaaannnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8303
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 8363
 nnnnngcagc aagagaagga acattttaca gcttattggc cgaacttcac tgccgctagt 8423
 gtggttcaac ttggactaca gagaaatctt cctaactggg ttccctgtat tcaactcctgc 8483
 tacctccaac ttggtctggt ctcaactttt gctataatag gcttttāaaa atcataaatc 8543
 taccatgtgt cctctgtcca gaccttctcc atggcttctt attgctcact ggatgaagtt 8603
 ccaacgagcc caggatgggt tgactcatgt ctccagcttt aactgcatca ccatcacctt 8663
 cattgtctaa agctctaacc acacaggatt ttctagtcct cagaggcatg gcagtctttc 8723
 aattccgagt tttctcatac aatattgtct cttcttāaaa tattttttct tgttgtccac 8783
 ctgagttgga gtcacttttt aaatctcagc taagcttata cttcatcaag tctttcctaa 8843
 ttctacotcc acgcaccaca ccattacat taaatccctt tattatatgt ttccatagca 8903
 cctactttct tcttttcagt atactcagca cacaatcaca tgtctaggat ctgttttaat 8963
 agcttggact accaattāaa ttgcatccct tttāattgtc cattgattcc tcaagtaccc 9023
 acatgcccac cttagcaaga agttcagtgt ctccctctta tagcatgtac ttctccacct 9083
 ccacāaaact gccagāaagc ttacttagcc cacagggccā gtgctaggca gctagggttag 9143
 tcctccagag ggccctgggt ttgagcagtt gctgtctact ccggccatgc agaatctctg 9203
 gtccttccag atgtctccat ccactgtgca aaggtaacct tgctgggtcc gatcccaca 9263
 cagaccacag tgctacaaga ttacagttct tatggttccc caacacatgc tctgtcattg 9323
 gtccāāaagc aggaccctta tgggttgatg aggtaggagg aggtccctgc cttagccaca 9383
 gctgcacaca gccagcctct tccttcttag gccctcatgt tgagcctggg acgccagtcc 9443
 taacttcctt ctcttcagtt cctcttaggg ccattgggtat cctgaatttc ttagtccatt 9503

FIG. 7G

gcaaagttaa gtaaagaagc agcaggcttg gtccctttcc ttccagatgg cttcttagct 9563
 cctgaacaga tttaccacc tatacctcag tgactagctc tgtgtactaa agtgtattgg 9623
 gagggcagcc attattggtc cataaaaggt cctgcttacc atttccctcct aagaggaacc 9683
 attcaacagt ttggggctcg aggggtgacct gctgggctct agagaagaag ctggcaactt 9743
 ctgttgcaaa ataatgttaa attctgcttc atctgcttgt cttnnnnnnnn nnnnnnnnnn 9803
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9863
 nnnnnnnnnn nnnnnnnnnn nntaaacatt gaacctacta tatgcagggtg agtatgctag 9923
 atnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9983
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10043
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10103
 nnnnnnnntgn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10163
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10223
 nnnnnnnnnn nnnnnnnnga tttgaataca ggtctgtttg actccaaaac ttgtggccta 10283
 tttgttgcaa aagtgttaa tacaaattgg ttcagtcaat attattatct ttgaacaatg 10343
 gaaggagaaa gtaagtttca atccaaaata attgagtgc ttatacattg acttgctgag 10403
 ccaatggcaa agtcaagtta gaatccagca gaagtcacca gctacagaat ctagatcttt 10463
 agaacatgtc ttcagatctt cagaacagtg tttcttaaac tctcttgtga aggaacagtt 10523
 atcatcatag gctggttaaca gttcacctac cagcaccagc ccatgaacca gactctaagt 10583
 ggcacagccc tagaagattg agccagaatt ttacagaggt ttaaagacca aatagtctgg 10643
 tttatggtta cctgtggccc acagagaatg gcagcactaa cctcaggcat aaatgaggta 10703
 cccactgaag ccaacattca agagcaatc ctatgggtta accattgggc tcctttcaaa 10763
 tgcaaacct catgaaagag actacagtgc tgaatagaga cctccaaatt ccaggccaag 10823
 ctcaggatag tcatgaggga attactaaaa acctggtata tagggcaaaa gcagaattag 10883
 gaatggactg atttcaggaa ccaggcaat ggcaggagtt gggcattaaa tcctaaaaga 10943
 gaatcagagt gggagggaat atgtgaaatc agagggttaag aaaaaagtga aaacctnnnn 11003
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11063
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11123

FIG. 7H

```

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11183
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11243
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11303
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11363
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11423
nnnnnnnnnn nnnnnnaaaaa attaaaagaa agatgtgaaa tcaaggaaac ttactggtga 11483
gcagcatccc attatgtgaa cttgtgcttc tgaaccagta acttgagtta ctttgagcca 11543
gtatcagtca cttatacctt agtgcaaaat taattgatca gacattctga cctggaccag 11603
ggaaggcagg cagaagtagc agtcaagact aaagcagaaa agggagagct aattctgcag 11663
ccagacattt cctggattga atacccaaat tagtccctca gcctttaagt gcctgagggc 11723
caggagtaga cagaggaatg gaaagtgtga gacttctttg ttcacactct ttgcctaggg 11783
gccagatttt gctttatgca ttaccatccg aagtcccagg ccacagtga catttgggct 11843
tcgctatgtg gatttatatta gatttacttt ttgtcctgcc atattttaat ctataagcca 11903
aacagttttc tcattaatct tattccattt ctggaatttt tccttttca gac aca aca 11961
                                     Asp Thr Thr    258

aaa aga cct gtc acc aca aca aag agg gaa gtg aac agt gct      12003
Lys Arg Pro Val Thr Thr Thr Lys Arg Glu Val Asn Ser Ala      272

gtgagtaagc atgattttta cttttctttc ttactttctt ttctctctca gottgaattt 12063
taaagtaacc actgttctat taattcatgg aaggcaactg aatagttcca gcttatagaa 12123
tcttctgtgt tggtagcatt tcagcgaagc ctggttctta gcccagaac aatcatgcca 12183
tcttttgctc ggtctatatt cctaagcact cctagatgat actgcactgg acctctggtc 12243
tcacatagtt agaaacagag ttaaaatcga acagcaaaga gaagatattc aactgcgatg 12303
caattgacaa tggatgtttt tgcaacaaac aatgattaag aagtaacattg ttgtgggctc 12363
tgagtcaaga gtaatatggg aaaaacacaa gtctcttcat gaggttgaca ggtttgagc 12423
tggaatctgt ggaggaggaa ggatatgatc taggggtcag aagaagtggg ttactaaaat 12483
cattaagcct ggttgatga aaagcttaga ctcaggggaa gcagnnnnnn nnnnnnnnnn 12543
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 12603
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 12663

```

FIG. 7I

```

nnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 12723
nnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 12783
nnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 12843
nggcaccaag agatgagggg ggcagtcttg gccatatatt tggetgaagt aagtcaattt 12903
gtcattcctg catgagcctt tataaacaga agtaagtaac caactactat ttggtcattg 12963
gagttgtcca agaggccagg gttctgtcta atacctgttc atgcatgaac atgccaacct 13023
agattgcatg cagactacca gttttgggtt tttgtttagt tcagcaggat ttttctcagc 13083
tcactgcctc tcaaactttc agcaacaaaa ggacatctgt gatatcagaa totaccactc 13143
taagtatttg gatgcaatag caatgaatat ctgagtaaatt ctaggtgggg agtgggggca 13203
ccctgtagcc aaaatgattt aacaaaatca aacaaaatt ttggaaatga tgccttggtg 13263
caatgaagag actacttgag gtaggtttga cttatctaatt atcttatttt ctttaccaat 13323
acctaagag gaatttaaatt atttctagat agctttggaa aggtccctta aagaggcacc 13383
agcataccac tgccagatct aatcccccca aacactgttt tcatcatcat catgtcatct 13443
cttgtctcta tagatcatat caaatccttc ccagagtttt tcaggccttt tgacaactag 13503
ccacatttca ctaagccaac tcactacca ctcttcaaca aaacttttcc tcaagttgag 13563
ctgtccacc aacaccactg ccattgagctc attcccactt ctgtggcttt gctcatgttg 13623
gttatttttt tggagtgtcc tccctattcc ttcttacttg tcccaatccc aacttttggc 13683
atggtctact ttaagataca gtaatgagta actttannnn nnnnnnnnnn nnnnnnnnnn 13743
nnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 13803
nnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 13863
nnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 13923
nnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn tttttaaaat ttaattcaan 13983
nnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 14043
nnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 14103
nnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 14163
nnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 14223
nnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnn nntacataca 14283

```

FIG. 7J

caacactctc ttgcttaatc tgtaagactc tctccccac tcataccttt ttatttttcc 14343
 tctgcattgt acacacaatc tataccactc ttaagcacat gattacagcg ttattttctg 14403
 gotgcttcta tgtgtctata ttttaggtcc acctggtcaa tataataaag tgggatatta 14463
 gtgttaatgc aactatatgg tatttgatat ttgtctttct gtccgtttat caatgtttct 14523
 tatagnnnnn nnnnnnnnnn nnnnnnnngg tgtctgattt tgaccaaatt tgactaaata 14583
 cnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnngt tgagtctatc aactcaaaaa 14643
 agaataacct acaacaataa caagtttcag aacatttttt aaattactga ttttatgagn 14703
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnngaatt ccattgccta 14763
 gaaattgcc a tggttaagat tttaatattg ctcaggccca gacagctcag ggctttgaca 14823
 ttcccacacc cattctctgc catcccagtt ctatctcacc ccaaaaccat ccattatgag 14883
 gagagtgtac agctctaggc tgcccgggag ccaccccgca ctctcatttt gtgactoggc 14943
 atcttgggag atggagtctt gggacttagc ctggacatgt ccttgcatg tacttcttac 15003
 aagactttta ttcagatgaa tattttccct tccaacttaa gaagcacagg gcttgcctgg 15063
 tttgcttcac taaccagcaa ctgaagcaag acctgacttg tgaaaatgcc taatagagtt 15123
 cagtattagc gctgtttcac catttccctg atctcttgcc tttgtgcaca tgatagaatt 15183
 gcacttctct gtgattaatt tgtagttaag tgtgggtcatg taactcgctt tggtaatta 15243
 aatgtaagca taagtgatgc gtgttatttc tgggtagaag atgtaagagt tggcatatgc 15303
 tttgccatat tttctttatc catctggcat ggtaaccagt aacattctag gtagtaattg 15363
 ctccatcagt ctcatctct gagtgactaa aattgacaga gtcccctgct gaccctcaat 15423
 gtacatggaa catgaacaag aataagcttt tgttttttat attgagattt tggagttggt 15483
 tgttcttaca gcattacctt gtttactcta atacaacatg gaaaaaactg gaacctataa 15543
 taaatagacc ctacgttgcc atttaaactt ctagttctga ggaataataa tgtggggaaa 15603
 tactttctat ataataaaaa aatagaaaat tgcaaaataa aaatatactt atgtatcatt 15663
 catgtcttat taaaaatgtt atttatagac tcaccataatt ccttctctcc agaaaaatag 15723
 aagtaaaaat atgaaaatgc ctgtaatcat gtttttggat tatggaatca agtattgctt 15783
 tttactttta tgttttctga atttttgttg tacttcaacta catttttgag tgccctgatg 15843
 tattactttc aaaaagaaga agaatacttt ctgaagccat ttcaaccatc cccactcacc 15903

FIG. 7K

tctctagatc ccagtaacca aatacattat ataggactct tcatcagtc ttatcaagtt 15963
 taggaagggc gatgctatac cttcttttaa ggacacotac caatgtctta gttgcctttc 16023
 aaagactcct agcacagcta aatgtgatgg atatgctota aggatataag agctgaagtg 16083
 acttgcataa ggtcatatca taacttactg ttagaaatgg agctagaact cagaccact 16143
 gagtccttgt ctgtgacaca ctgccctttc cttttgtgga agttgttctt gtatctaact 16203
 ttatctgtgc tactatttgg gcctagccat tctccctctt atgcagacaa gcagataaac 16263
 agtaaaactt taggagtggg ttatgatacc atagatatat atcatctatc ctttacaaaa 16323
 tagttattac agtcatcaag ccttggttag agtttacaga ccatgtatcc tagctannnn 16383
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nntgaaaca 16443
 gtagcagacc aaaagaagtc atgattccca gcatagtgct nnnnnnnnnn nnnnnnnnnn 16503
 nnnnnnnnnn nnnnaaacct gtattaagtt tctgttattt gcaagaccct gtcagttagg 16563
 ccatgtggga actagaagga tgaatttato agtcatccaa gattcttaca attaagtatt 16623
 accgataagg tactcaagaa acagttctca ttcacataat ttgggttaaa acaaaaagaa 16683
 gccagctttc tatatacttt tgggtccagtc tttacgtttt ttgttttggt ttgttttggt 16743
 tcatgagtat cccgacttcc ttctaagaac ttccacctga gaactgacca cagcgtcagc 16803
 attccacatg ggtgtgtttc ctttccctt tccatttca gtgggtttca atttcttttt 16863
 cttttggcac tataaacctt tcgcaaagga aatattagac agaactocta catgtcaagc 16923
 aaattaaaat agtggtgaaa ttagagtggg ggacataatc accctatcat ataggctatt 16983
 tgtccatato atatttgtcc ctacaaaggc ctctaagggn nnnnnnnnnn nnnnnnnnnn 17043
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 17103
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 17163
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 17223
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 17283
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 17343
 nnnnnnnnnn nnnnnnnnnn nnnctctaag ataccttgtg cttcttgga catatttggg 17403
 aaatcatgta gctctcaaat tatccctatg tctgaagcc cacctacca tcaatctca 17463
 gaaataccca acccatgtca ggcaacttca cactttcttt cttcaggcag cacagttgtc 17523

FIG. 7L

tcagggaggg aggagagtgc tattagcaag aggagtcact aacagcttca ctacactgtg 17583
 cccatgaatt tttaatggtg tgaaaagtct gtgtatTTTT tatggttctc tatggctcat 17643
 ataagagggg caaacaatac atgaaagtcc agagatgggg acaaatatca ctttaagctg 17703
 gggataatca gggaagaaga aagttttcat ggagaagggtg acttttgaat caggaagaaa 17763
 tggaaacagc aaactcttct agatagagga gacactaaca ggaaaggcag agaggcagga 17823
 aggtgtggga aagtgcacgt gaccccggtc agagagaaaag ccaggtgtga gataaggggg 17883
 aaagactgtt agggcatgta ttgtaaacca ctaattccag gcaaaagtta gattttactt 17943
 actaagcaag agtgcttcag ttagatccta gcaggaaatg gagggtatgc ttagaagagg 18003
 taactgaggc aagtttaatt tataaagggtg tgtgcagcat taagggaac cagcaaggga 18063
 tactgagcat gccaggatgc aagagcaggc agggaagggtg actattccta ggtctgaagg 18123
 agaaagggga gggagcagtt ccagaaccc tagtaaaaat ggcaatgaga aaggtccatc 18183
 tggcaggacc tatggtcttt aacagagggg caaagtcaac ccacaacttg totgggaggt 18243
 tgctgaggaa tagatacccc aacctctctc tcaaccact gcaacactct ttttcccta 18303
 gactgagccc agtcaaagac agagggagga gccagtgat gcagtctgca atgtcatcat 18363
 cctggagcat gaatagagtg cagcagggtg aataatgagt ctgcaggaat taatagaaat 18423
 atctgacaca atagggaact ataagaggtt ttgaatagga gagggccctg aaatgtgctc 18483
 caatattact gaactatgtg tggcccaaag aatggaagag gaacagctct tgcaataggt 18543
 ctgaggagag aagctgaaga cttggactag ggcaatggta aaaactgtgg aaagaagttt 18603
 taaatgaaaa gttttaaac atgcggcttc cagctagatg aactttttta aaaaaattag 18663
 ttcctcactc aaattttggg gaggttatat attttctaata cataaaaaat gatttttctt 18723
 atttgtgggc ttttctcccc ag atc tga acctgtggtc ttgggagcca gggtgacctg 18781
 Ile 273
 atatgacatc taaagaagct tctggactct gaacaagaat tcggtggcct gcagagcttg 18841
 ccatttgac ttttcaaag ctttggatg accagcact ttaatctgaa acctgcaaca 18901
 agactagcca acacctggcc atgaaacttg ccccttcaact gatctggact cacctctgga 18961
 gcctatggct ttaagcaagc actactgcac ttacagaat taccactg gatcctggac 19021
 ccacagaatt ccttcaggat cttcttggct gccagactga aagcaaaagg aattatttcc 19081

FIG. 7M

```
cctcaagttt totaagtgat ttccaaaagc agaggtgtgt ggaaatttcc agtaacagaa 19141
acagatgggt tgccaataga gttatTTTTT atctatagct tcctctgggt actagaagag 19201
gctattgaga ctatga 19217
```

FIG. 8

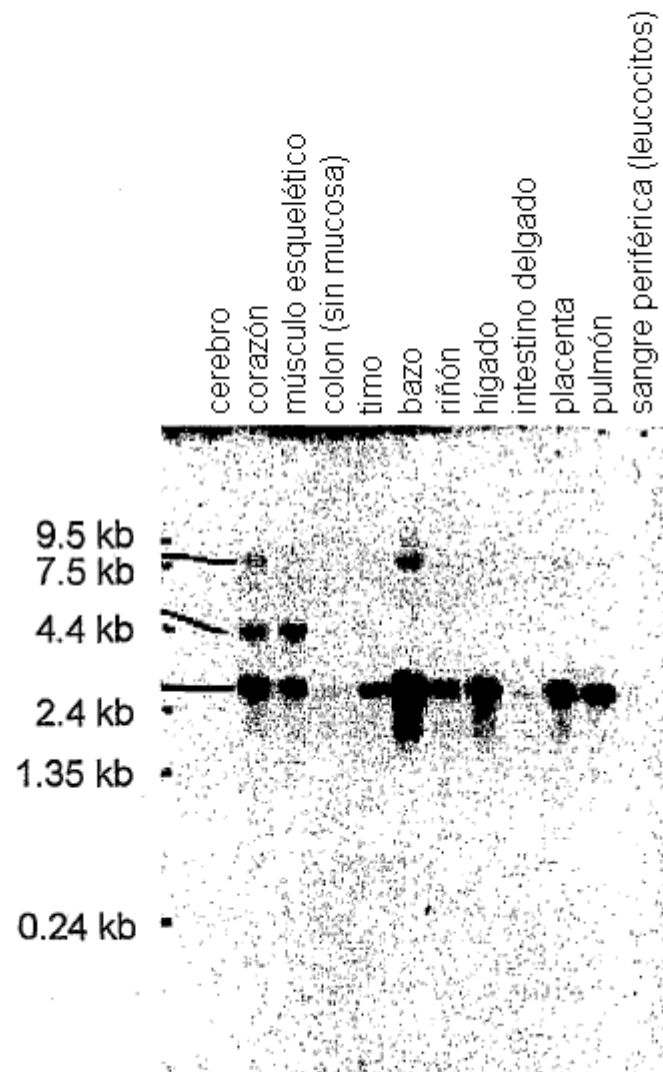


FIG. 9

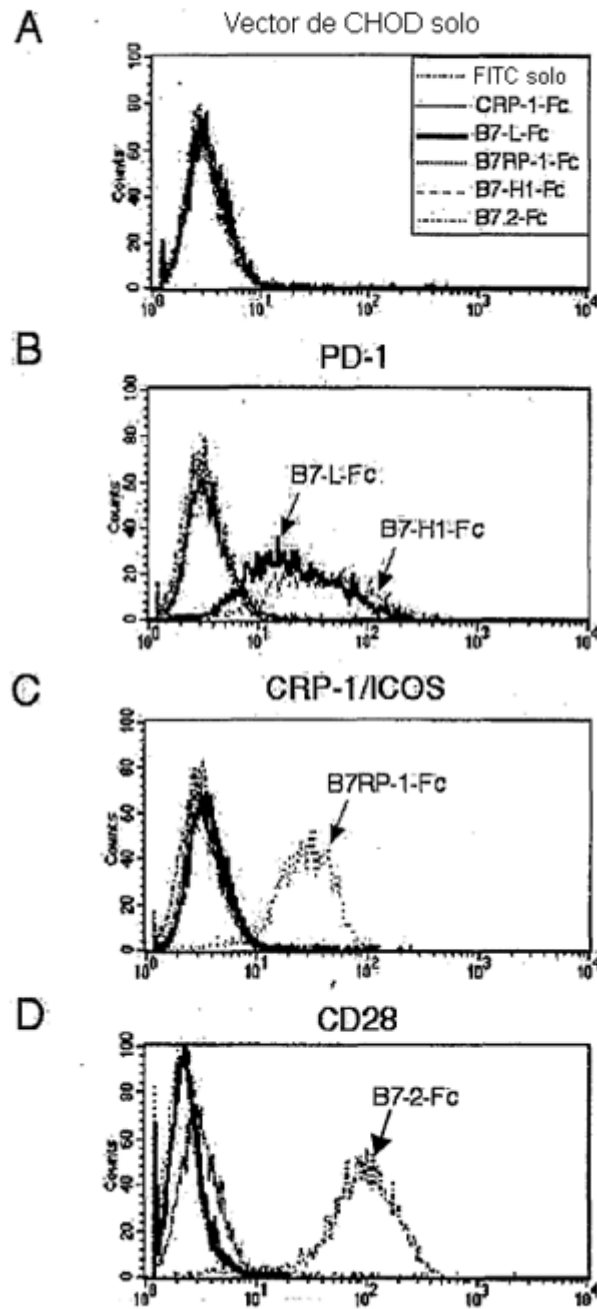


FIG. 10

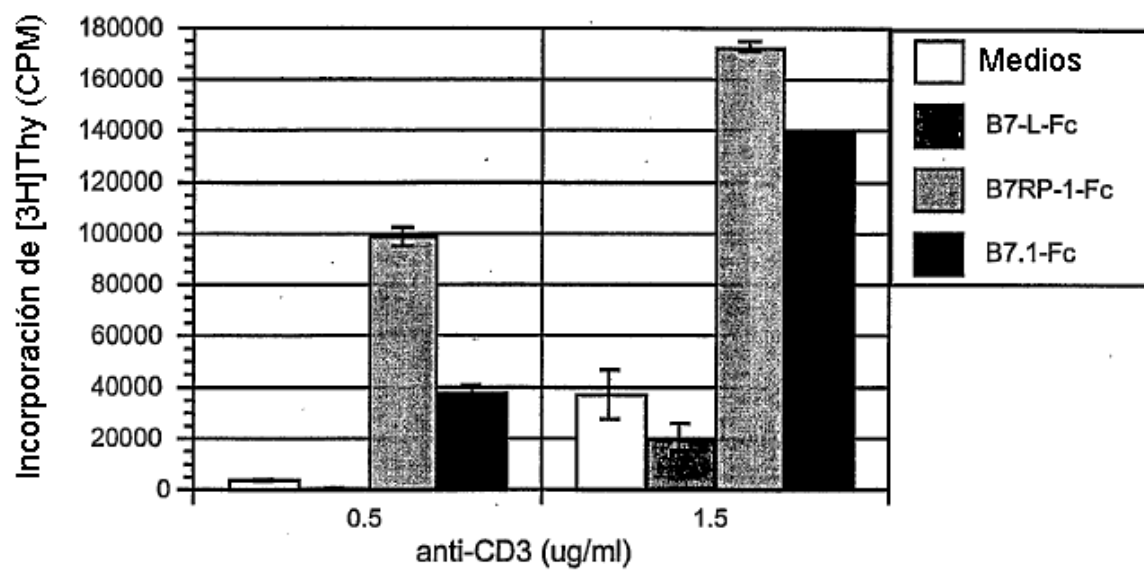


FIG. 11

