

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 368 072**

51 Int. Cl.:
C07K 14/71 (2006.01)
G01N 33/74 (2006.01)
A61K 38/17 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **05818361 .7**
96 Fecha de presentación: **08.12.2005**
97 Número de publicación de la solicitud: **1824881**
97 Fecha de publicación de la solicitud: **29.08.2007**

54 Título: **RECEPTOR DE PROTEÍNAS SIMILARES A TGR3.**

30 Prioridad:
08.12.2004 GB 0426960

45 Fecha de publicación de la mención BOPI:
14.11.2011

45 Fecha de la publicación del folleto de la patente:
14.11.2011

73 Titular/es:
ARES TRADING S.A.
ZONE INDUSTRIELLE DE L'OURIETTAZ
1170 AUBONNE, CH

72 Inventor/es:
MICHALOVICH, David;
MITTER, Richard James;
FAGAN, Richard Joseph y
POWER, Christine

74 Agente: **de Elzaburu Márquez, Alberto**

ES 2 368 072 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

Receptor de proteínas similares a TGR3

5 La presente invención se refiere a novedosas proteínas que contienen un dominio que corresponde a la mitad C terminal de un dominio ZP similar al identificado en el Receptor III de TGF-beta y al uso de estas proteínas y secuencias de ácido nucleico de los genes codificantes en el diagnóstico, la prevención y el tratamiento de enfermedades. En particular, la presente invención se refiere al uso de estas proteínas para modular la actividad de las proteínas que pertenecen a la superfamilia TGF-beta.

10 **Antecedentes**

El proceso de descubrimiento de fármacos actualmente está experimentando una revolución fundamental a medida que la genómica funcional crece. La expresión "genómica funcional" se aplica a un abordaje que utiliza herramientas bioinformáticas para atribuir funciones a secuencias de proteínas de interés. Tales herramientas se están volviendo cada vez más necesarias a medida que la velocidad de generación de datos de secuencias supera rápidamente la capacidad de los laboratorios de investigación de asignar funciones a estas secuencias de proteínas.

20 A medida que las herramientas bioinformáticas aumentan en potencia y precisión, estas herramientas reemplazan rápidamente las técnicas convencionales de caracterización bioquímica. En efecto, las herramientas bioinformáticas avanzadas usadas para la identificación en la presente invención son ahora capaces de proporcionar resultados en los que se puede depositar un alto grado de confianza.

25 Varias instituciones y organizaciones comerciales están examinando los datos de las secuencias a medida que van quedando disponibles y se están haciendo descubrimientos importantes continuamente. Sin embargo, permanece una necesidad constante de identificar y caracterizar genes adicionales y los polipéptidos que éstos codifican, como objetivos para la investigación y el descubrimiento de fármacos.

Introducción a proteínas secretadas

30 La capacidad de las células de producir y secretar proteínas extracelulares es fundamental para muchos procesos biológicos. Las enzimas, factores de crecimiento, proteínas de matriz extracelular y moléculas de señalización son todas secretadas por células. Esto se produce a través de la fusión de una vesícula de secreción con la membrana plasmática. En la mayoría de los casos, pero no en todos, las proteínas se dirigen al retículo endoplasmático y a las vesículas de secreción mediante un péptido señal. Los péptidos señal son secuencias que actúan en cis que afectan el transporte de cadenas de polipéptidos desde el citoplasma a un compartimiento unido a la membrana tal como una vesícula de secreción. Los polipéptidos que están dirigidos a las vesículas de secreción se secretan dentro de la matriz extracelular o se retienen en la membrana plasmática. Los polipéptidos que se retienen en la membrana plasmática tendrán uno o más dominios transmembrana. Ejemplos de proteínas secretadas que tienen un rol central en el funcionamiento de una célula son citoquinas, hormonas, proteínas de matriz extracelular (moléculas de adhesión), proteasas y factores de crecimiento y diferenciación.

Introducción a receptores de superficie celular

45 Los receptores de superficie celular son componentes importantes de cualquier célula procariota o eucariota y se ha demostrado que tienen roles importantes en una gran cantidad de funciones fisiológicas diversas, muchas de las cuales participan en procesos de enfermedades. La alteración de su actividad es un medio para alterar el fenotipo de la enfermedad relevante y, como tal, la identificación de nuevos receptores de superficie celular es altamente relevante.

50 **Introducción a proteínas de Zona Pelúcida**

55 El dominio Zona Pelúcida, o dominio 'ZP', se caracteriza por 8 cisteínas altamente conservadas y otros residuos polares o hidrófobos distribuidos en 260 aminoácidos y se reconoció por primera vez en los receptores ZP2 y ZP3 de esperma del ratón. Sin embargo, muchas otras proteínas que no son proteínas de recubrimiento de ovocitos en mamíferos se clasifican ahora como proteínas que contienen dominios ZP, incluidas uromodulina, glicoproteína-2, tectorinas α y β , el receptor III del factor de crecimiento transformante beta (TGR3) y endogлина (Bork P and Sander C, FEBS Lett. 1992, 300: 237-40; Jovine L *et al.*, Nat Cell Biol. 2002, 4: 457-61). Este dominio generalmente se incluye en proteínas transmembrana tipo multidominio, en donde se localiza en una sola copia en la sección C terminal de la porción extracelular y puede llevar a cabo varias funciones. Puede mediar la polimerización de polipéptidos extracelulares pero también puede participar en la unión de un ligando. De hecho, los elementos C terminales que permiten la asociación a la membrana celular (un dominio transmembrana simple o un ancla de glicosilfosfatidilinositol) pueden separarse proteolíticamente de la porción extracelular y la célula secreta una forma madura de la proteína.

65 El dominio ZP puede distinguirse estructuralmente y funcionalmente en dos mitades, al parecer generadas mediante duplicación (Yonezawa N and Nakano M, Biochem Biophys Res Commun. 2003, 307: 877-82; Jovine L *et al.*, Proc

Natl Acad Sci USA. 2004, 101: 5922-7). En particular, el Receptor III de TGF-beta contiene un sitio de unión a TGF-beta en la región C terminal del dominio ZP (WO 95/10610; Kaname S and Ruoslahti E, Biochem J. 1996, 315: 815-20).

5 Entre las proteínas que contienen el dominio ZP, el Receptor III de TGF-beta reviste particular importancia, ya que se une a TGF-beta, un ligando en sí mismo que pertenece a una superfamilia de proteínas que incluye no solo TGF-beta, sino otras muchas proteínas, tales como GDF (Factores de Crecimiento/Diferenciación), BMP (Proteínas Morfogenéticas Óseas), Activinas o Inhibinas, que tienen relevancia terapéutica (Tsuchida K *et al.*, Curr Drug Targets Immune Endocr Metabol Disord. 2004, 4: 157-66; Chang H *et al.*, Endocr Rev. 2002, 23: 787-823; de Caestecker M, Cytokine Growth Factor Rev. 2004, 15: 1-11). De hecho, se han descrito fragmentos de receptores para proteínas similares a TGF-beta en la bibliografía que tienen utilidad terapéutica (WO 95/10610; WO 97/05892; Bandyopadhyay A *et al.*, Cancer Res. 2002, 62: 4690-5; Liu M *et al.*, Am J Respir Crit Care Med. 2002, 165: 419-23; Vilchis-Landeros MM *et al.*, Biochem J. 2001, 355:215-22).

15 La variedad de proteínas conocidas por contener el dominio ZP y sus roles fisiológicos sugieren que las proteínas que contienen el dominio ZP pueden ser proteínas importantes que participan en patologías humanas tales como cánceres, retinopatías, neuropatías, trastornos reproductivos y trastornos del desarrollo.

20 Por consiguiente, el conocimiento en aumento de las proteínas que contienen el dominio ZP, y más preferiblemente aquellos homólogos del Receptor III de TGF-beta que pueden unirse a uno o más miembros de la superfamilia TGF-beta, es de extrema importancia para aumentar la comprensión de las vías subyacentes que conducen a los estados de enfermedad mencionados anteriormente, así como también a otros estados de enfermedad, en particular aquellos relacionados con módulos y fragmentos específicos, y para desarrollar terapias más efectivas para tratar trastornos, tales como cánceres.

25 WO058473 divulga la secuencia de nucleótidos GSN:AAC74735. WO058473 también divulga GSP:AAB40526.

EM_EST:BG752983 divulga la secuencia de ARNm 602732441F1 NIH_MGC_43 Homo sapiens cDNA clon IMAGE:4875922 5'.

30 LA INVENCIÓN

La invención proporciona un polipéptido que comprende la secuencia de aminoácidos como se describe en SEQ ID NO:8.

35 La invención también proporciona un polipéptido que comprende SEQ ID NO:10.

40 La invención también proporciona un polipéptido que es un fragmento o un equivalente funcional tal como se describe en cualquiera de las reivindicaciones precedentes, que tiene más de 80% de identidad de secuencia con la secuencia de aminoácidos descrita en SEQ ID NO:8, preferiblemente más de 85%, 90%, 95%, 98% o 99% de identidad de secuencia, y contiene un subdominio ZP/TGR3 como se encuentra en cualquiera de las secuencias de aminoácidos codificadas por SEQ ID NO:10, SEQ ID NO:12 o SEQ ID NO:14.

45 La presente invención proporciona novedosas secuencias que contienen un dominio que corresponde a la mitad C terminal de un dominio ZP similar al identificado en el Receptor III de TGF-beta (de aquí en adelante, un subdominio ZP/TGR3) que puede tener características estructurales y funcionales como se indica en la bibliografía (Yonezawa N and Nakano M, Biochem Biophys Res Commun. 2003, 307: 877-82; Jovine L et al, Proc Natl Acad Sci USA. 2004, 101: 5922-7). Esta secuencia, sola o fusionada con otras secuencias, puede tener un efecto terapéutico en el tratamiento de enfermedades, en particular aquellas asociadas con la expresión y secreción excesiva de una proteína que pertenece a la superfamilia de las proteínas TGF-beta. En particular, la invención se basa en el descubrimiento de que polipéptidos de la presente invención son inhibidores de factores de crecimiento. Preferiblemente, los factores de crecimiento son proteínas similares a TGF-beta. Preferiblemente, la proteína similar a TGF-beta es TGF-beta 1.

50 En un primer aspecto de la invención se proporciona un polipéptido que:

55 (i) comprende la secuencia de aminoácidos tal como se describe en SEQ ID NO:8.

También se describe en la presente un polipéptido que:

(ii) es un fragmento del mismo que contiene un subdominio ZP/TGR3 o tiene un determinante antigénico en común con los polipéptidos de (i); o

60 (iii) es un equivalente funcional de (i) o (ii).

También se describe en la presente un polipéptido de acuerdo con la parte (i) del primer aspecto de la invención que puede comprender la secuencia de aminoácidos tal como se describe en SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:19 o SEQ ID NO:21.

Un polipéptido de acuerdo con la parte (i) del primer aspecto de la invención puede consistir en la secuencia de aminoácidos tal como se describe en SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:19 o SEQ ID NO:21.

5 Un polipéptido de acuerdo con la parte (ii) de la lista anterior puede comprender la secuencia de aminoácidos tal como se describe en SEQ ID NO:10. Tales polipéptidos pueden comprender la secuencia de aminoácidos tal como se describe en SEQ ID NO:12 o SEQ ID NO:14. Tales polipéptidos pueden consistir en la secuencia de aminoácidos tal como se describe en SEQ ID NO:10, SEQ ID NO:12 o SEQ ID NO:14.

10 También se describe en la presente un polipéptido que es un equivalente funcional de acuerdo con la parte (iii) de la lista anterior que puede ser homólogo a la secuencia de aminoácidos tal como se describe en SEQ ID NO:8, y preferiblemente contiene un subdominio ZP/TGR3.

15 La invención proporciona un polipéptido que es un fragmento o un equivalente funcional tal como se describe anteriormente que preferiblemente tiene más de 80% de identidad de secuencia con la secuencia de aminoácidos descrita en SEQ ID NO:8 o con un fragmento activo de la misma, preferiblemente más de 85%, 90%, 95%, 98% o 99% de identidad de secuencia y preferiblemente contiene un subdominio ZP/TGR3.

20 También se describe en la presente un polipéptido que es un equivalente funcional tal como se describió anteriormente, que preferiblemente exhibe homología estructural importante con un polipéptido que tiene la secuencia de aminoácidos descrita en SEQ ID NO:8, y preferiblemente contiene un subdominio ZP/TGR3.

Un polipéptido que es un fragmento tal como se describió anteriormente que tiene un determinante antigénico en común con el polipéptido de la parte (i) anterior puede consistir en 7 o más residuos de aminoácidos de la secuencia de aminoácidos descrita en SEQ ED NO: 8.

25 Aunque la Solicitante no desea ceñirse a esta teoría, se postula que los primeros 16 aminoácidos del polipéptido INSP215 de longitud completa forman un péptido señal. La secuencia del polipéptido INSP215 de longitud completa sin la secuencia señal postulada se describe en SEQ ID NO:4. El polipéptido INSP215 de región extracelular sin la secuencia señal postulada se describe en SEQ ID NO:8. El polipéptido INSP215 de exones 1-4 sin la secuencia señal postulada se describe en SEQ ID NO:14.

30 El polipéptido que tiene la secuencia descrita en SEQ ID NO:2 se denomina de aquí en adelante "el polipéptido INSP215 de longitud completa". El polipéptido que tiene la secuencia descrita en SEQ ID NO:4 se denomina de aquí en adelante "el polipéptido INSP215 de longitud completa maduro". El polipéptido que tiene la secuencia descrita en SEQ ID NO:6 se denomina de aquí en adelante "el polipéptido INSP215 de región extracelular". El polipéptido que tiene la secuencia descrita en SEQ ID NO:8 se denomina de aquí en adelante "el polipéptido INSP215 de región extracelular maduro". El polipéptido que tiene la secuencia descrita en SEQ ID NO:10 se denomina de aquí en adelante "el polipéptido INSP215 de subdominio ZP/TGR3". El polipéptido que tiene la secuencia descrita en SEQ ID NO:12 se denomina de aquí en adelante "el polipéptido INSP215 de exones 1-4". El polipéptido que tiene la secuencia descrita en SEQ ID NO:14 se denomina de aquí en adelante "el polipéptido INSP215 de exones 1-4 maduro". El polipéptido que tiene la secuencia descrita en SEQ ID NO:19 se denomina de aquí en adelante "el polipéptido INSP215 de longitud completa con etiqueta 6HIS". El polipéptido que tiene la secuencia descrita en SEQ ID NO:21 se denomina de aquí en adelante "el polipéptido INSP215 de región extracelular con etiqueta 6HIS".

35 La expresión "polipéptidos INSP215", tal como se usa en la presente, incluye polipéptidos que comprenden o consisten en el polipéptido INSP215 de región extracelular maduro de la SEQ ID NO:8, tal como el polipéptido INSP215 de longitud completa, el polipéptido INSP215 de longitud completa maduro, el polipéptido INSP215 de región extracelular, el polipéptido INSP215 de región extracelular maduro, el polipéptido INSP215 de subdominio ZP/TGR3, el polipéptido INSP215 de exones 1-4, el polipéptido INSP215 de exones 1-4 maduro, el polipéptido INSP215 de longitud completa con etiqueta 6HIS y el polipéptido INSP215 de región extracelular con etiqueta 6HIS.

40 La expresión "que contiene un subdominio ZP/TGR3" se refiere a polipéptidos que contienen un dominio que corresponde a la mitad C terminal de un dominio ZP similar al identificado en el Receptor III de TGF-beta (el subdominio ZP/TGR3) que puede tener características estructurales y funcionales tales como se indican en la bibliografía (Yonezawa N and Nakano M, Biochem Biophys Res Commun. 2003, 307: 877-82; Jovine L *et al.*, Proc Natl Acad Sci USA. 2004, 101: 5922-7). Preferiblemente, los polipéptidos que contienen un subdominio ZP/TGR3 comprenderán el polipéptido INSP215 de subdominio ZP/TGR3 de la SEQ ID NO:10.

45 Se sabe que una cantidad de receptores de superficie celular que contienen el dominio ZP se secretan de la célula luego de la escisión de la cadena de polipéptidos en un sitio específico entre el dominio ZP y la región transmembrana (o sitio de unión de ancla de GPI). Por consiguiente, se prevén formas secretadas de los polipéptidos INSP215 y, por ejemplo, pueden comprender o consistir en una secuencia de aminoácidos tal como se describe en SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14 o SEQ ID NO:21.

50 Tal como se indicó anteriormente, la presente invención proporciona nuevas secuencias de polipéptidos que contienen una secuencia de dominio similar a la mitad C terminal de un dominio ZP identificado en el Receptor III de TGF-beta (el subdominio ZP/TGR3).

5 Un marco de lectura abierto de 879 nucleótidos (SEQ ID NO:1) codifica el polipéptido INSP215 de longitud completa (SEQ ID NO:2), una proteína de 292 aminoácidos (ver Figura 1). El polipéptido INSP215 de longitud completa contiene un péptido señal previsto (residuos 1-16) que conduce a una secuencia madura de 276 aminoácidos (SEQ ID NO:4), codificada por una secuencia de ADN de 831 nucleótidos (SEQ ID NO:3). El polipéptido INSP215 de longitud completa contiene un región transmembrana prevista (residuos 245-268) que separa una región intracelular (residuos 269-292) y una región extracelular (residuos 1-244; SEQ ID NO:5 y SEQ ID NO:6). Son características adicionales de la región extracelular la secuencia madura (SEQ ID NO:7 y SEQ ED NO:8) que contiene el subdominio ZP/TGR3 previsto (SEQ ID NO:9 y SEQ ID NO:10) y un sitio de N-glicosilación (localizado en el residuo 139). El subdominio ZP/TGR3 se encuentra en la secuencia de aminoácidos codificada por los primeros cuatro de los seis exones que codifican el polipéptido INSP215 de longitud completa (SEQ ID NOs 11-14).

15 Si el polipéptido INSP215 de longitud completa se usa como una consulta con BLAST en bases de datos públicas de secuencias no redundantes, se descubre que un segmento de región extracelular INSP215 es altamente homólogo a la mitad C terminal del dominio ZP/TGR3 (Bork P and Sander C, FEBS Lett. 1992, 300: 237-40; Jovine L *et al.*, Nat Cell Biol. 2002, 4: 457-61). La similitud entre INSP215 y el Receptor III de TGF-beta en el dominio ZP/TGR3 no está limitada a los residuos de cisteína separados regularmente, sino que también implica otros diversos residuos polares o hidrófobos, así como también un sitio de N-glicosilación.

20 La expresión del polipéptido INSP215 está respaldada por la existencia de numerosos EST en humanos, ratones y ratas (ver los Ejemplos en la presente).

25 La bibliografía muestra que el dominio ZP/TGR3 puede distinguirse estructuralmente y funcionalmente en dos mitades, y que un sitio de unión a TGF-beta está localizado en la región C terminal del dominio ZP (WO 95/10610; Kaname S y Ruoslahti E, Biochem J. 1996, 315: 815-20). Por consiguiente, se espera que los polipéptidos INSP215 y fragmentos funcionales y/o proteínas de fusión en base a los polipéptidos INSP215, en particular polipéptidos que comprenden la secuencia de aminoácidos descritas en SEQ ID NO:10, se unan a una o más proteínas que pertenecen a la superfamilia de proteínas TGF-beta. Cuando estas secuencias relacionadas con INSP215 son proteínas secretadas o solubles, pueden actuar como antagonistas de proteínas que pertenecen a la superfamilia de proteínas TGF-beta.

35 En base a la anotación de los polipéptidos INSP215 en la presente, ahora es posible diseñar experimentos para detectar la presencia de los transcritos de INSP215 a través de distintos tipos de tejidos humanos para determinar su expresión tisular, en particular en tejidos normales y enfermos para establecer más específicamente la relevancia de INSP215 en contextos patológicos.

40 La clonación de las secuencias de ADN relacionadas con INSP215 (como secuencias de ADNc o genómicas exónicas, y ejemplificadas en SEQ ID NO:1, 3, 5, 7, 9, 11 y 13) de tejidos de origen humano o mamífero permitirá la expresión de las correspondientes secuencias proteicas codificadas en sistemas de expresión procariotas o eucariotas y su posterior purificación y caracterización como proteínas recombinantes.

45 Pueden producirse secuencias de INSP215 recombinantes alternativas clonando en el vector de expresión apropiado el ADN que codifica la secuencia completa (SEQ ID NO:2), o sus fragmentos seleccionados, en forma de secuencias que contienen el péptido señal (SEQ ID NO:6 y 12) o maduras (SEQ ID NO:4, 8, 10 y 14) manteniendo las propiedades de unión esperadas de la proteína completa hacia proteínas que pertenecen a la superfamilia TGF-beta. Por ejemplo, las transcripciones que incluyen exones 1-4 (que abarcan el subdominio ZP/TGR3 entero y la secuencia señal) pueden identificarse y/o clonarse mediante hibridación de cebador o mediante PCR.

50 Siempre que sea apropiado, las secuencias de ADN heterólogas que codifican las secuencias proteicas que ayudan en la expresión, secreción y/o detección de estas secuencias recombinantes pueden agregarse en el mismo marco del ADN que codifica INSP215, o las secuencias proteicas derivadas identificadas anteriormente. Son ejemplos de tales secuencias heterólogas péptidos señal, regiones Fc de inmunoglobulina y etiquetas de histidina.

55 Se espera que variantes solubles de INSP215 se unan a uno o más miembros de la superfamilia de proteínas TGF-beta, capturando y reteniendo la interacción de esta proteína con los receptores de señalización, con la consiguiente inhibición de la proliferación y diferenciación de múltiples tipos de células inducidas normalmente por las proteínas similares a TGF-beta. Por lo tanto, las porciones específicas de INSP215 indicadas en la presente como proteínas secretadas o solubles, que pueden generarse in vivo o in vitro luego de la digestión proteolítica o empalme alternativo, pueden tener propiedades in vivo específicas, en particular como antagonistas de una o más proteínas que pertenecen a la superfamilia TGF-beta. Por lo tanto se espera que tales polipéptidos sean útiles para el tratamiento de enfermedades, en particular cánceres, fibrosis, cicatrización, hipertensión, aterosclerosis, trastornos reproductivos, trastornos hepáticos, trastornos pulmonares, trastornos inmunes, trastornos del desarrollo, trastornos de la piel o trastornos de tejidos conjuntivos (Tsuchida K *et al.*, Curr Drug Targets Immune Endocr Metabol Disord. 2004, 4: 157-66; Chang H *et al.*, Endocr Rev. 2002, 23: 787-823; de Caestecker M, Cytokine Growth Factor Rev. 2004, 15: 1-11; WO 95/10610; WO 97/05892; Bandyopadhyay A *et al.*, Cancer Res. 2002, 62: 4690-5; Esparza-Lopez J *et al.*, J Biol Chem. 2001, 276: 14588-96; Vilchis-Landeros MM *et al.*, Biochem J. 2001, 355: 215-22).

- Los polipéptidos divulgados en la presente pueden ser útiles para el tratamiento de enfermedades tales como cáncer metastásico, cáncer de pulmón de células no pequeñas, cáncer de colon, leucemia, cáncer testicular, mieloma, linfoma, cáncer colorrectal, cáncer de próstata, cánceres cerebrales, cánceres oculares, retinoblastoma, cáncer del útero, cáncer intestinal, cáncer del cuello de útero, carcinoma endometrial, cáncer de pulmón, cáncer de páncreas, condrosarcoma, linfomas no Hodgkin, tumores estromales de los cordones sexuales del ovario, carcinomas ováricos, tumores gonadales, cáncer de células renales, leucemia linfocítica crónica B, carcinoma gástrico, carcinomas pancreáticos, adenocarcinoma pancreático, cáncer de pulmón de células pequeñas, gliomas, gliomas de alto grado, gliomas malignos, gliosarcomas, astrocitomas anaplásicos, glioblastomas, meduloblastoma, ependimoma, afecciones malignas de células T, trasplantes de médula ósea, fibrosis, fibrosis pulmonar idiopática, esclerodermia, cicatrización post operatoria seguida de cirugía de glaucoma, enfermedad renal, diabetes, nefropatía diabética, aterosclerosis, cardiomiopatía restrictiva, angiogénesis, bronquiolitis obliterante, caquexia, enfermedad pulmonar obstructiva crónica, cicatrización de heridas, síndrome de Crouzon o glomerulonefritis.
- Preferiblemente, la leucemia es leucemia linfocítica, leucemia linfocítica aguda, leucemia de células B aguda, leucemia de células T aguda, leucemia mielógena aguda, leucemia mielógena crónica o leucemia linfocítica crónica.
- Preferiblemente, el cáncer testicular es seminoma, coriocarcinoma, carcinoma embrionario, teratoma o tumor de saco vitelino.
- Preferiblemente, el cáncer de pulmón se selecciona de tumores primarios benignos o malignos o de metástasis de cánceres primarios de muchos otros órganos y tejidos. Preferiblemente, el cáncer de pulmón se selecciona de tumores de pulmón primarios incluidos carcinoma broncogénico, carcinoide bronquial, hamartoma condromatoso (benigno), linfoma solitario, sarcoma (maligno) o linfomas multifocales. Preferiblemente, el carcinoma broncogénico se selecciona de carcinoma de células escamosas, carcinoma indiferenciado de células pequeñas, carcinoma indiferenciado de células grandes, adenocarcinoma o carcinoma bronquioloalveolar. Preferiblemente, el carcinoma broncogénico es carcinoma de células escamosas o carcinoma de células pulmonares no pequeñas.
- Preferiblemente el cáncer de pulmón se selecciona de metástasis de cánceres primarios de piel, mama, colon, próstata, riñón, tiroides, estómago, cuello uterino, recto, testículo y óseo y de melanoma.
- Preferiblemente, el linfoma es enfermedad de Hodgkin, linfoma linfocítico pequeño (SLL/CLL), linfoma de células del manto (MCL), linfoma folicular (FL), linfoma de zona marginal (MZL) incluido extranodal (linfoma MALT), nodal (linfoma de células B monocitoides) o esplénico, linfoma difuso de células grandes, linfoma de Burkitt, linfoma de alto grado similar a Burkitt o linfoma linfoblástico.
- INSP215 recombinante y los polipéptidos INSP215 identificados en la presente pueden usarse para generar anticuerpos monoclonales o policlonales específicos que podrían usarse luego en la caracterización bioquímica de INSP215, así como también para aplicaciones de diagnóstico y/o terapéuticas (en virtud de la información divulgada en la bibliografía sobre fragmentos generados a partir de estas familias de proteínas). De forma alternativa, INSP215 recombinante y los polipéptidos INSP215 identificados en la presente pueden usarse en una amplia variedad de ensayos de detección, en particular los que involucran proteínas incluidas en la superfamilia de proteínas TGF-beta, para identificar ligandos específicos y actividades biológicas asociadas.
- Preferiblemente, la actividad de un polipéptido soluble de la presente invención, solo o como parte de una proteína de fusión, o un antagonista de un polipéptido unido a membrana de la presente invención, puede confirmarse en al menos uno de los siguientes ensayos:
- en la modulación de apoptosis (por ejemplo mediante tinción de túnel de caspasa 3 activada por IHC), o
 - en la modulación de la proliferación celular (por ejemplo, uso de Ki-67 o PCNA por parte de IHC), o
 - en la modulación de angiogénesis o neovascularización (por ejemplo, CD34 por inmunohistoquímica (IHC), o
 - en la modulación de la inmunosupresión (por ejemplo, actividad de células NK, activación de célula T, marcadores de activación de macrófago por IHC), o
 - en la modulación de migración celular (por ejemplo, marcadores de EMT por IHC, vimentina, E-cadherina, SNAIL), o
 - en la modulación de la interacción tumor-huésped (por ejemplo activación de vía de factor de crecimiento; HGF, EGF, PDGF, VEGF), o
 - en la modulación del crecimiento tumoral, o
 - en la modulación de la metástasis, o
 - en la modulación de la supervivencia celular, o
 - en la modulación de la maduración de ovocito, o
 - en la modulación del crecimiento folicular, o
 - en el ensayo tal como se describe en el Ejemplo 4 en la presente, o
 - en el modelo de carcinoma mamario murino 4T1 singénico, o
 - en el modelo de ratón de alelo TGFBR2 "floxed", o

- o) en modelos de ratón de cáncer de mama metastásico (por ejemplo, los ratones MMTV/Neu), o
- p) en conejos sometidos a cirugía de glaucoma modificada, o
- q) en un modelo de rata de diabetes, o
- r) en ratas con gliomas intracraneales establecidos, o
- s) en el modelo de glioma de rata C6, o
- t) en un modelo de rata de bronquiolitis obliterante tal como se describe en Liu *et al.* (Am. J. Respir. Crit. Care Med. 2002. Vol.165, pp.419-423), o
- u) en el modelo de xenoinjerto tal como se describe en Bandyopadhyay *et al.* (Prostate 2005, Vol.63, Número 1, pp.81-90).

Adicionalmente, la actividad de un INSP215 soluble o un antagonista de INSP215 unido a membrana puede demostrarse en modelos de cáncer descritos por Kamb y Lassota (Drug Discovery Today: Disease Models; Vol. 1, No. 1, 2004, pp. 31-36) y Grunt *et al.* (Drug Discovery Today: Disease Models; Vol. 1, No. 1, 2004, pp. 17-23), en modelos de cánceres de pulmón descritos por Lahm y Fischer (Drug Discovery Today: Disease Models; Vol. 1, No. 1, 2004, pp. 25-30), en modelos de mieloma múltiple descritos por Caers *et al.* (Drug Discovery Today: Disease Models; Vol. 1, No. 4, 2004, pp. 373-380), en modelos de cáncer de colon descritos por Albanese *et al.* (Drug Discovery Today: Disease Models; en prensa), en modelos de cáncer pancreático descritos por Thomas y Lowy (Drug Discovery Today: Disease Models; en prensa) o en modelos de cáncer de mama descritos por Matulka y Wagner (Drug Discovery Today: Disease Models; en prensa) y Gupta y Kuperwasser (Drug Discovery Today: Disease Models; Vol. 1, No. 1, 2004, pp. 9-16).

Un "determinante antigénico" descrito en la presente puede ser una parte de un polipéptido de la presente invención que se une a un sitio de combinación del anticuerpo o a un receptor de células T (TCR). De forma alternativa, un "determinante antigénico" puede ser un sitio sobre la superficie de un polipéptido de la presente invención al que se une una única molécula de anticuerpo. Generalmente un antígeno tiene varios o muchos determinantes antigénicos diferentes y hace reacción con anticuerpos de muchas especificidades diferentes. Preferiblemente, el anticuerpo es inmuno-específico para un polipéptido de la invención. Preferiblemente, el anticuerpo es inmuno-específico para un polipéptido de la invención que no es parte de una proteína de fusión. Preferiblemente, el anticuerpo es inmuno-específico para INSP215 o un fragmento del mismo. Los determinantes antigénicos generalmente consisten en agrupamientos de moléculas superficiales químicamente activos, tales como aminoácidos o cadenas laterales de azúcares, y pueden tener características estructurales tridimensionales específicas, así como también características de carga específicas. Preferiblemente, el "determinante antigénico" se refiere a un grupo químico particular en un polipéptido de la presente invención que es antigénico, es decir, que provoca una respuesta inmune específica.

La secuencia de polipéptidos ORFX ORF290 humana, SEQ ID NO:580 divulgada en WO200058473 (SEQ ID NO:46 en la presente) y la PROTEÍNA IPI00645072.1 de 33 KDA (SEQ ID NO:47 en la presente) y sus secuencias de ácido nucleico codificantes se excluyen específicamente del alcance de esta invención.

En un segundo aspecto, la invención proporciona una molécula de ácido nucleico purificada que codifica un polipéptido del primer aspecto de la invención.

La expresión "molécula de ácido nucleico purificada" preferiblemente se refiere a una molécula de ácido nucleico de la invención que (1) se ha separado de al menos aproximadamente 50 por ciento de proteínas, lípidos, carbohidratos u otros materiales con los que se encuentra naturalmente cuando el ácido nucleico total se aísla de las células fuente, (2) no está ligada a la totalidad o una porción de un polinucleótido al que la "molécula de ácido nucleico purificada" está ligada naturalmente, (3) está ligada operativamente a un polinucleótido al que no está ligado naturalmente o (4) no se produce en la naturaleza como parte de una secuencia de polinucleótidos mayor. Preferiblemente, la molécula de ácido nucleico aislada de la presente invención está sustancialmente libre de cualquier otra molécula(s) de ácido nucleico contaminante u otros contaminantes que se encuentran en su ambiente natural que interferirían con su uso en la producción de polipéptidos o su uso terapéutico, diagnóstico, profiláctico o para investigación. En una realización preferida, el ADN genómico se excluye específicamente del alcance de la invención. Preferiblemente, el ADN genómico mayor a 10 kbp (kilopares de bases), 50 kbp, 100 kbp, 150 kbp, 200 kbp, 250 kbp o 300 kbp se excluye específicamente del alcance de la invención. Preferiblemente, la "molécula de ácido nucleico purificada" consiste solamente en ADNc.

Preferiblemente, la molécula de ácido nucleico purificada comprende la secuencia de ácido nucleico como se describe en SEQ ID NO:1 (que codifica el polipéptido INSP215 de longitud completa), SEQ ID NO:3 (que codifica el polipéptido INSP215 de longitud completa maduro), SEQ ID NO:5 (que codifica el polipéptido INSP215 de región extracelular), SEQ ID NO:7 (que codifica el polipéptido INSP215 de región extracelular maduro), SEQ ID NO:9 (que codifica el polipéptido INSP215 de subdominio ZP/TGR3), SEQ ID NO:11 (que codifica el polipéptido INSP215 de exones 1-4), SEQ ID NO:13 (que codifica el polipéptido INSP215 de exones 1-4 maduro), SEQ ID NO:16 (la secuencia de ADNc prevista de INSP215), SEQ ID NO:17 (el inserto de ADNc en el clon Image 4875922), SEQ ID NO:18 (que codifica el polipéptido INSP215 de longitud completa con etiqueta 6HIS) o SEQ ID NO:20 (que codifica el polipéptido INSP215 extracelular con etiqueta 6HIS), o es un equivalente o fragmento redundante de estas secuencias.

- 5 La invención además prevé que la molécula de ácido nucleico purificada consista en la secuencia de ácido nucleico como se describe en SEQ ID NO:1 (que codifica el polipéptido INSP215 de longitud completa), SEQ ID NO:3 (que codifica el polipéptido INSP215 de longitud completa maduro), SEQ ID NO:5 (que codifica el polipéptido INSP215 de región extracelular), SEQ ID NO:7 (que codifica el polipéptido INSP215 de región extracelular maduro), SEQ ID NO:9 (que codifica el polipéptido INSP215 de subdominio ZP/TGR3), SEQ ID NO:11 (que codifica el polipéptido INSP215 de exones 1-4), SEQ ID NO:13 (que codifica el polipéptido INSP215 de exones 1-4 maduro), SEQ ID NO:16 (la secuencia de ADNc prevista de INSP215), SEQ ID NO:17 (el inserto de ADNc en el clon Image 4875922), SEQ ID NO:18 (que codifica el polipéptido INSP215 de longitud completa con etiqueta 6HIS) o SEQ ID NO:20 (que codifica el polipéptido INSP215 extracelular con etiqueta 6HIS) o un equivalente o fragmento redundante de aquellas secuencias.
- 10 En un tercer aspecto, la divulgación proporciona una molécula de ácido nucleico purificada que se hibridiza en condiciones de rigurosidad alta con una molécula de ácido nucleico del segundo aspecto de la invención. Las condiciones de hibridación de rigurosidad alta se definen como incubación durante toda la noche a 42°C en una solución que comprende 50% de formamida, 5XSSC (150mM de NaCl, 15mM de citrato trisódico), 50mM de fosfato de sodio (pH 7,6), solución de Denhardt 5x, 10% de sulfato de dextrano y 20 microgramos/ml de ADN de esperma de salmón cizallado desnaturalizado, seguida por lavado de los filtros en 0,1X SSC a aproximadamente 65°C.
- 15 En un cuarto aspecto, la invención proporciona un vector, tal como un vector de expresión, que contiene una molécula de ácido nucleico del segundo o tercer aspecto de la divulgación. Los vectores preferidos incluyen aquellos descritos en los Ejemplos en la presente.
- 20 En un quinto aspecto, la invención proporciona una célula huésped transformada con un vector del cuarto aspecto de la invención.
- 25 En un sexto aspecto, la invención proporciona un ligando que se une específicamente a un polipéptido del primer aspecto de la invención.
- 30 En un séptimo aspecto, la divulgación proporciona un compuesto que es efectivo para alterar la expresión de un gen natural que codifica un polipéptido del primer aspecto de la invención o para regular la actividad de un polipéptido del primer aspecto de la invención.
- Tales compuestos pueden identificarse usando los ensayos y métodos de detección divulgados en la presente.
- 35 Un compuesto del séptimo aspecto de la divulgación puede aumentar (agonizar) o disminuir (antagonizar) el nivel de expresión del gen o la actividad del polipéptido.
- 40 En manera significativa, la identificación de la función de los polipéptidos INSP215 permite el diseño de métodos de detección capaces de identificar compuestos que son efectivos en el tratamiento y/o diagnóstico de enfermedades. Los ligandos y compuestos de acuerdo con el sexto y séptimo aspecto de la divulgación pueden identificarse usando tales métodos. Estos métodos se incluyen como aspectos de la presente divulgación.
- 45 Otro aspecto de esta invención reside en el uso de un gen o polipéptido INSP215 como objetivo para la detección sistemática de moduladores de fármacos candidatos, particularmente fármacos candidatos activos contra las enfermedades mencionadas en la presente.
- 50 También se describen en la presente métodos de detección sistemática de compuestos para terapia de las enfermedades mencionadas en la presente, que comprenden determinar la capacidad de un compuesto para unirse a un gen o polipéptido INSP215, o un fragmento de los mismos.
- 55 También se describen en la presente métodos de detección sistemática de compuestos para terapia de las enfermedades mencionadas en la presente, que comprenden evaluar la modulación de la actividad de un gen o polipéptido INSP215, o un fragmento de los mismos.
- 60 En un octavo aspecto, la invención proporciona un polipéptido del primer aspecto de la invención, o una molécula de ácido nucleico del segundo o tercer aspecto de la divulgación o un vector del cuarto aspecto de la invención, o una célula huésped del quinto aspecto de la invención, o un ligando del sexto aspecto de la invención, o un compuesto del séptimo aspecto de la divulgación para uso en la terapia o diagnóstico de enfermedades en las que están implicados polipéptidos que contienen un dominio ZP. Tales enfermedades incluyen trastornos proliferativos celulares, cáncer, trastornos cardiovasculares, trastornos neurológicos, trastornos metabólicos, infección, trastornos inmunes, trastornos autoinmunes, inflamación, trastornos genéticos y particularmente retinopatías, trastornos reproductivos y trastornos del desarrollo. Preferiblemente, la enfermedad es un enfermedad en la que están implicadas proteínas que contienen un dominio ZP/TGR3, tales como cánceres, fibrosis, cicatrización, hipertensión, aterosclerosis, trastornos reproductivos, trastornos hepáticos, trastornos de pulmón, trastornos inmunes, trastornos del desarrollo, trastornos de la piel y trastornos de tejidos conjuntivos. Estos restos del primer, segundo, tercer,
- 65

cuarto, quinto, sexto o séptimo aspecto de la divulgación también pueden usarse en la fabricación de un medicamento para el tratamiento de tales enfermedades.

5 En particular, los restos del primer, segundo, tercer, cuarto, quinto, sexto o séptimo aspecto de la divulgación pueden ser útiles para el tratamiento de enfermedades tales como cáncer metastásico, cáncer de pulmón de células no pequeñas, cáncer de colon, leucemia, cáncer testicular, mieloma, linfoma, cáncer colorrectal, cáncer de próstata, cánceres cerebrales, cánceres oculares, retinoblastoma, cáncer del útero, cáncer intestinal, cáncer del cuello de útero, carcinoma endometrial, cáncer de pulmón, cáncer de páncreas, condrosarcoma, linfomas no Hodgkin, tumores estromales de los cordones sexuales del ovario, carcinomas ováricos, tumores gonadales, cáncer de células renales, leucemia linfocítica crónica B, carcinoma gástrico, carcinomas pancreáticos, adenocarcinoma pancreático, cáncer de pulmón de células pequeñas, gliomas, gliomas de alto grado, gliomas malignos, gliosarcomas, astrocitomas anaplásicos, glioblastomas, meduloblastoma, ependimoma, afecciones malignas de células T, trasplantes de médula ósea, fibrosis, fibrosis pulmonar idiopática, esclerodermia, cicatrización post operatoria seguida de cirugía de glaucoma, enfermedad renal, diabetes, nefropatía diabética, aterosclerosis, cardiomiopatía restrictiva, angiogénesis, bronquiolitis obliterante, caquexia, enfermedad pulmonar obstructiva crónica, cicatrización de heridas, síndrome de Crouzon o glomerulonefritis.

20 Preferiblemente, la leucemia es leucemia linfocítica, leucemia linfocítica aguda, leucemia de células B aguda, leucemia de células T aguda, leucemia mielógena aguda, leucemia mielógena crónica o leucemia linfocítica crónica. Preferiblemente, el cáncer testicular es seminoma, coriocarcinoma, carcinoma embrionario, teratoma o tumor de saco vitelino.

25 Preferiblemente, el cáncer de pulmón se selecciona de tumores primarios benignos o malignos o de metástasis de cánceres primarios de muchos otros órganos y tejidos. Preferiblemente el cáncer de pulmón se selecciona de tumores de pulmón primarios incluidos carcinoma broncogénico, carcinoide bronquial, hamartoma condromatoso (benigno), linfoma solitario, sarcoma (maligno) o linfomas multifocales. Preferiblemente, el carcinoma broncogénico se selecciona de carcinoma de células escamosas, carcinoma indiferenciado de células pequeñas, carcinoma indiferenciado de células grandes, adenocarcinoma o carcinoma bronquioloalveolar. Preferiblemente, el carcinoma broncogénico es carcinoma de células escamosas o carcinoma de células pulmonares no pequeñas. Preferiblemente el cáncer de pulmón se selecciona de metástasis de cánceres primarios de piel, mama, colon, próstata, riñón, tiroides, estómago, cuello uterino, recto, testículo y óseo y de melanoma.

35 Preferiblemente, el linfoma es enfermedad de Hodgkin, linfoma linfocítico pequeño (SLL/CLL), linfoma de células del manto (MCL), linfoma folicular (FL), linfoma de zona marginal (MZL) incluido extranodal (linfoma MALT), nodal (linfoma de células B monocitoides) o esplénico, linfoma difuso de células grandes, linfoma de Burkitt, linfoma de alto grado similar a Burkitt o linfoma linfoblástico.

40 En un noveno aspecto, la divulgación proporciona un método para diagnosticar una enfermedad en un paciente, que comprende evaluar el nivel de expresión de un gen natural que codifica un polipéptido del primer aspecto de la invención o la actividad de un polipéptido del primer aspecto de la invención en tejido de dicho paciente y comparar dicho nivel de expresión o actividad con un nivel testigo, en donde un nivel que es diferente a dicho nivel testigo es indicador de enfermedad. Tal método preferiblemente se llevará a cabo in vitro. Pueden usarse métodos similares para monitorear el tratamiento terapéutico de la enfermedad en un paciente, en donde la alteración del nivel de expresión o actividad de un polipéptido o molécula de ácido nucleico durante el período de tiempo hacia un nivel testigo es indicador de una regresión de la enfermedad.

50 Un método preferido para detectar polipéptidos del primer aspecto de la invención comprende los pasos de: (a) poner en contacto un ligando, tal como un anticuerpo, del sexto aspecto de la invención con una muestra biológica en condiciones adecuadas para la formación de un complejo ligando-polipéptido; y (b) detectar dicho complejo.

55 Existe una cantidad de tales métodos diferentes de acuerdo con el noveno aspecto de la divulgación, como el lector experto apreciará, tales como métodos de hibridación de ácido nucleico con sondas cortas, análisis de mutación puntual, amplificación por reacción en cadena de la polimerasa (PCR) y métodos que usan anticuerpos para detectar niveles de proteínas aberrantes. Pueden usarse métodos similares a corto o largo plazo para permitir el tratamiento terapéutico de una enfermedad que se monitoreará en un paciente. La invención también proporciona kits que son útiles en estos métodos para diagnosticar enfermedades.

60 Preferiblemente, la enfermedad diagnosticada mediante un método del noveno aspecto de la divulgación es una enfermedad en la que están implicados polipéptidos que contienen el dominio ZP, o una enfermedad en la que están implicadas proteínas que contienen el dominio ZP/TGR3, tal como se describió anteriormente.

65 En un décimo aspecto, la divulgación proporciona el uso de un polipéptido del primer aspecto de la invención como un receptor de superficie celular o una forma soluble del mismo. Los usos adecuados de los polipéptidos de esta invención como receptores de superficie celular serán claros para el lector experto. Ejemplos particulares de usos adecuados incluyen métodos de detección para identificar ligandos y otras moléculas agonistas o antagonistas que son útiles en la terapia y diagnóstico de enfermedades y afecciones en las que está implicada esta categoría de

proteínas. En particular, la invención proporciona el uso de una forma transmembrana o soluble de un polipéptido del primer aspecto de la invención para unirse a una proteína que pertenece a la superfamilia de proteínas TGF-beta.

- 5 En un decimoprimer aspecto, la divulgación proporciona una composición farmacéutica que comprende un polipéptido del primer aspecto de la invención, o una molécula de ácido nucleico del segundo o tercer aspecto de la divulgación o un vector del cuarto aspecto de la invención, o una célula huésped del quinto aspecto de la invención, o un ligando del sexto aspecto de la invención, o un compuesto del séptimo aspecto de la divulgación, en conjunción con un portador farmacéuticamente aceptable.
- 10 En un decimosegundo aspecto, la divulgación proporciona un polipéptido del primer aspecto de la invención, o una molécula de ácido nucleico del segundo o tercer aspecto de la divulgación o un vector del cuarto aspecto de la invención, o una célula huésped del quinto aspecto de la invención, o un ligando del sexto aspecto de la invención, o un compuesto del séptimo aspecto de la divulgación, para uso en la fabricación de un medicamento para el diagnóstico o tratamiento de una enfermedad.
- 15 Preferiblemente, la enfermedad es una enfermedad en la que están implicados polipéptidos que contienen el dominio ZP, o un enfermedad en la que están implicadas proteínas que contienen el dominio ZP/TGR3, tal como se describió anteriormente.
- 20 En un decimoterter aspecto, la divulgación proporciona un método para tratar una enfermedad en un paciente que comprende administrar al paciente un polipéptido del primer aspecto de la invención, o una molécula de ácido nucleico del segundo o tercer aspecto de la divulgación o un vector del cuarto aspecto de la invención, o una célula huésped del quinto aspecto de la invención, o un ligando del sexto aspecto de la invención, o un compuesto del séptimo aspecto de la divulgación.
- 25 Preferiblemente, la enfermedad es una enfermedad en la que están implicados polipéptidos que contienen el dominio ZP, o un enfermedad en la que están implicadas proteínas que contienen el dominio ZP/TGR3, tal como se describió anteriormente.
- 30 Para enfermedades en las que la expresión de un gen natural que codifica un polipéptido del primer aspecto de la invención, o en las que la actividad de un polipéptido del primer aspecto de la invención, es más baja en un paciente enfermo cuando se compara con el nivel de expresión o actividad en un paciente sano, el polipéptido, molécula de ácido nucleico, vector, célula huésped, ligando o compuesto administrado al paciente debería ser un agonista. En cambio, para enfermedades en las que la expresión del gen natural o actividad del polipéptido es más alta en un paciente enfermo cuando se compara con el nivel de expresión o actividad en un paciente sano, el polipéptido, molécula de ácido nucleico, vector, célula huésped, ligando o compuesto administrado al paciente debería ser un antagonista. Ejemplos de tales antagonistas incluyen moléculas de ácido nucleico, ribozimas y ligandos antisentido, tales como anticuerpos.
- 35
- 40 En un decimocuarto aspecto, la divulgación proporciona animales no humanos transgénicos o genéticamente desactivados que se han transformado para expresar niveles más altos, más bajos o ausentes de un polipéptido del primer aspecto de la invención. Tales animales transgénicos son modelos muy útiles para el estudio de la enfermedad y también pueden usarse en regímenes de detección para la identificación de compuestos que son efectivos en el tratamiento o diagnóstico de tal enfermedad.
- 45 Preferiblemente, la enfermedad es una enfermedad en la que están implicados polipéptidos que contienen el dominio ZP, o un enfermedad en la que están implicadas proteínas que contienen el dominio ZP/TGR3, tal como se describió anteriormente.
- 50 Tal como se usa en la presente, "equivalente funcional" se refiere a una proteína o molécula de ácido nucleico que posee características funcionales o estructurales que son sustancialmente similares a un polipéptido o molécula de ácido nucleico de la presente invención. Un equivalente funcional de una proteína puede contener modificaciones dependiendo de la necesidad de tales modificaciones para el funcionamiento de una función específica. Se pretende que la expresión "equivalente funcional" incluya los fragmentos, mutantes, híbridos, variantes, análogos o derivados químicos de una molécula.
- 55 Preferiblemente, el "equivalente funcional" puede ser una proteína o molécula de ácido nucleico que exhibe cualquier una o más de las actividades funcionales de los polipéptidos de la presente invención.
- 60 Preferiblemente, el "equivalente funcional" puede ser una proteína o molécula de ácido nucleico que muestra actividad sustancialmente similar en comparación con INSP215 o fragmentos del mismo en un ensayo adecuado para la medición de la actividad o función biológica. Preferiblemente, el "equivalente funcional" puede ser una proteína o molécula de ácido nucleico que muestra actividad idéntica o más alta en comparación con INSP215 o fragmentos del mismo en un ensayo adecuado para la medición de la actividad o función biológica. Preferiblemente, el "equivalente funcional" puede ser una proteína o molécula de ácido nucleico que muestra 50%, 60%, 70%, 80%,
- 65

90%, 95%, 98%, 99% 100% o mayor actividad en comparación con INSP215 o fragmentos del mismo en un ensayo adecuado para la medición de la actividad o función biológica.

5 Preferiblemente, el "equivalente funcional" puede ser una proteína o polipéptido capaz de exhibir una actividad sustancialmente similar in vivo o in vitro a la de los polipéptidos de la invención. Preferiblemente, el "equivalente funcional" puede ser una proteína o polipéptido capaz de interactuar con otras moléculas celulares o extracelulares en un modo sustancialmente similar al modo en que lo haría la porción correspondiente de los polipéptidos de la invención. Por ejemplo, un "equivalente funcional" sería capaz, en un inmunoensayo, de disminuir la unión de un anticuerpo con el péptido correspondiente (es decir, el péptido la secuencia de aminoácidos de la cual fue modificada para obtener el "equivalente funcional") del polipéptido de la invención, o con el polipéptido de la invención en sí mismo, donde el anticuerpo fue generado contra el péptido correspondiente del polipéptido de la invención. Una concentración equimolar del equivalente funcional disminuirá la unión del péptido correspondiente mencionada anteriormente en al menos aproximadamente 5%, preferiblemente entre aproximadamente 5% y 10%, más preferiblemente entre aproximadamente 10% y 25%, incluso más preferiblemente entre aproximadamente 25% y 50% y más preferiblemente entre aproximadamente 40% y 50%.

20 Por ejemplo, los equivalentes funcionales pueden ser completamente funcionales o pueden carecer de función en una o más actividades. Por consiguiente, en la presente invención, las variaciones pueden afectar la función, por ejemplo, de las actividades del polipéptido que refleja su posesión de un subdominio ZP/TGR3.

25 A continuación se proporciona un compendio de técnicas y procedimientos estándares que pueden emplearse para aplicar la invención. Se comprenderá que esta invención no está limitada a la metodología, protocolos, líneas celulares, vectores y reactivos particulares descritos. También se comprenderá que la terminología usada en la presente es solamente a efectos de describir realizaciones particulares y no se pretende que esta terminología deba limitar el alcance de la presente invención. El alcance de la invención está limitado solamente por los términos de las reivindicaciones adjuntas.

En esta memoria descriptiva se usan abreviaturas estándares para nucleótidos y aminoácidos.

30 La puesta en práctica de la presente invención empleará, a menos que se indique de otra forma, técnicas convencionales de biología molecular, microbiología, tecnología de ADN recombinante e inmunología, que con conocidas por los expertos en la técnica.

35 Tales técnicas están explicadas ampliamente en la bibliografía. Ejemplos de textos particularmente adecuados para consulta incluyen los siguientes: Sambrook Molecular Cloning; A Laboratory Manual, Segunda edición (1989); DNA Cloning, Volúmenes I y II (D.N Glover ed. 1985); Oligonucleotide Synthesis (M.J. Gait ed. 1984); Nucleic Acid Hybridization (B.D. Hames & S.J. Higgins eds. 1984); Transcription and Translation (B.D. Hames & S.J. Higgins eds. 1984); Animal Cell Culture (R.I. Freshney ed. 1986); Immobilized Cells and Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide to Molecular Cloning (1984); la serie Methods in Enzymology (Academic Press, Inc.), especialmente los volúmenes 154 y 155; Gene Transfer Vectors for Mammalian Cells (J.H. Miller y M.P. Calos eds. 1987, Cold Spring Harbor Laboratory); Immunochemical Methods in Cell and Molecular Biology (Mayer and Walker, eds. 1987, Academic Press, Londres); Scopes, (1987) Protein Purification: Principles and Practice, Second Edition (Springer Verlag, N.Y.); y Handbook of Experimental Immunology, Volúmenes I-IV (D.M. Weir and C. C. Blackwell eds. 1986).

45 Tal como se usa en la presente, el término "polipéptido" incluye cualquier polipéptido o proteína que comprende dos o más aminoácidos unidos entre sí mediante enlaces peptídicos o enlaces peptídicos modificados, es decir, isómeros peptídicos. Este término se refiere a cadenas cortas (péptidos y oligopéptidos) y a cadenas más largas (proteínas).

50 El polipéptido de la presente invención puede estar en forma de una proteína madura o puede ser una pre-, pro- o prepro-proteína que puede activarse mediante escisión de la pre-, pro- o prepro-porción para producir un polipéptido maduro activo. En tales polipéptidos, la pre-, pro- o prepro-secuencia puede ser una secuencia líder o secretora o puede ser una secuencia que se emplea para la purificación de la secuencia de polipéptidos madura.

55 El polipéptido del primer aspecto de la invención puede formar parte de una proteína de fusión. Por ejemplo, con frecuencia es ventajoso incluir una o más secuencias de aminoácidos adicionales que pueden contener secuencias secretoras o líderes, pro-secuencias, secuencias que pueden ayudar en la purificación o secuencias que confieren estabilidad proteica más alta, por ejemplo durante la producción recombinante. De forma alternativa o adicional, el polipéptido maduro puede estar fusionado con otro compuesto, tal como un compuesto para aumentar la semivida del polipéptido (por ejemplo, polietilenglicol).

60 Los polipéptidos de la invención son útiles por sí solos, como componentes de proteínas de fusión tales como fusión Fc y/o en combinación con otro agente. Preferiblemente, la fusión Fc comprende un polipéptido soluble de la invención, es decir, un Fc:receptor soluble. Un Fc:receptor soluble es una proteína quimérica del fragmento constante (Fc) de inmunoglobulina G y el dominio extracelular o fragmento del mismo del receptor. Se sabe que los Fc:receptores solubles son útiles en el tratamiento de tumores (ver por ejemplo Yingling *et al.* Nature Reviews 2004,

Vol.3, pp.1011-1022). Preferiblemente, el péptido soluble consiste en el dominio extracelular o un fragmento del mismo. Preferiblemente, el polipéptido soluble consiste en la secuencia clonada (SEQ ID NO:21).

Preferiblemente el agente se selecciona entre:

- 5 1) TWIST, o
- 2) inhibidores de quinasa de molécula pequeña, o
- 3) inhibidores de señalización de TGF-beta de molécula grande, o
- 4) agente terapéutico ERBB2/HER2 tal como un inhibidor del receptor de tirosina quinasa Erbb2, o
- 10 5) inhibidor de la enzima de conversión de angiotensina (ACE) tal como lisinopril, o
- 6) inhibidor del receptor de tirosina quinasa de la familia EGFR tal como gefitinib, o
- 7) inhibidor de tirosina quinasa Abl tal como imatinib, o
- 8) Inmunomoduladores.

15 Preferiblemente, el inhibidor de quinasa de molécula pequeña se dirige al complejo receptor de TGF- β . Preferiblemente, el complejo receptor de TGF- β es el RI de TGF- β . Preferiblemente, el inhibidor de quinasa del RI de TGF- β del inhibidor de molécula pequeña se selecciona de inhibidores en base a un andamio de dihidropirrolpirazol tales como LY550410 o LY580276 desarrollados por Elli Lilly, inhibidores en base a un andamio de imidazol tales como SB-505124 desarrollado por GlaxoSmithKline, inhibidores en base a un andamio de pirazolopiridina tales como los desarrollados por Elli Lilly, inhibidores en base a un andamio de pirazol tales como los desarrollados por GlaxoSmithKline, inhibidores en base a un andamio de imidazopiridina tales como aquellos desarrollados por Biogen, inhibidores en base a un andamio de triazol tales como los desarrollados por Pfizer, inhibidores en base a un andamio de piridopirimidina tales como los desarrollados por Johnson & Johnson, inhibidores en base a un andamio de isotiazol tales como los desarrollados por Pfizer, un inhibidor derivado de quinazolina tal como SD-093 o el inhibidor de quinasa SD-208.

25 Preferiblemente, el inhibidor de señalización de TGF-beta de molécula grande es un receptor soluble, un anticuerpo neutralizante o un oligonucleótido antisentido.

30 Preferiblemente, el receptor soluble es el antagonista del receptor tipo II de Fc: soluble SR2F. Bandyopadhyay *et al.* divulga el uso de un betaglicano recombinante soluble (TGF β RIII) denominado sBG útil como un agente terapéutico para el cáncer de próstata después de la administración sistémica en un modelo de xenoinjerto (Prostate 2005, Vol.63, Número 1, pp.81-90). Liu *et al.* divulgaron una transfección génica del receptor tipo II de TGF β soluble capaz de inhibir la obliteración fibrosa de las vías respiratorias en un modelo de rata de bronquiolitis obliterante (Am. J. Respir. Crit. Care Med. 2002. Vol.165, pp.419-423).

35 Preferiblemente, el anticuerpo neutralizante es un anticuerpo monoclonal de inmunoglobulina G₄TGF- β ₂ recombinante humana tal como lerdelimumab, un anticuerpo monoclonal que se dirige a TGF- β tal como metelimumab o un anticuerpo anti-TGF- β pan-específico humano tal como GC-1008.

40 Preferiblemente, el oligonucleótido antisentido se dirige al ARNm del ligando TGF- β . Preferiblemente, el oligonucleótido antisentido es una vacuna de TGF- β ₂ antisentido (una vacuna de células tumorales alogénicas modificada por TGF- β ₂ antisentido), AP-12009, NeuGene, un inhibidor de oligonucleótido antisentido como el desarrollado por NeXstar, o un oligonucleótido antisentido de fosforotioato específico de TGF- β ₁ tal como AP-11014.

45 Preferiblemente, el inhibidor del receptor de tirosina quinasa Erbb2 es Trastuzumab.

50 Como tales, los antagonistas de TGF-beta son conocidos y usados para el tratamiento de varias enfermedades. Antisense Pharma GmbH desarrolló AP-11014 para el tratamiento de cánceres de pulmón de células no pequeñas, colorrectales y de próstata y AP-12009 para el tratamiento de gliomas, en particular gliomas de alto grado, preferiblemente astrocitomas anaplásicos o glioblastomas. AVI BioPharma desarrolló NeuGene para trasplantes de médula ósea. Biogen desarrolló inhibidores del receptor de TGF-beta para el tratamiento de fibrosis. Cambridge Antibody Technology desarrolló GC-1008 para el tratamiento de fibrosis, en particular de fibrosis pulmonar idiopática, lerdelimumab para el tratamiento o para la prevención de cicatrización postoperatoria luego de cirugía de glaucoma y metelimumab para el tratamiento de enfermedad renal, fibrosis, esclerodermia y diabetes, en particular de nefropatía diabética. Eli Lilly desarrolló inhibidores de TGF-beta para el tratamiento de aterosclerosis, cáncer e inflamación. NCI/NIH desarrolló un antagonista de TGF-beta para el tratamiento de cáncer, en particular de cáncer metastásico. NeXstar desarrolló un antagonista de TGF-beta para el tratamiento de un trastorno de angiogénesis. Scios desarrolló un antagonista de TGF-beta 1 para el tratamiento de enfermedad pulmonar obstructiva crónica, cardiomiopatía restrictiva y cáncer. Sumimoto Seika Chemicals desarrolló un antagonista de TGF-beta para el tratamiento de nefropatía diabética. Telios desarrolló anticuerpos CAT anti-TGF-beta para el tratamiento de fibrosis. NovaRx desarrolló una vacuna de TGF- β ₂ antisentido para el tratamiento de gliomas, en particular gliomas malignos, por ejemplo, gliosarcomas.

65 Preferiblemente, los polipéptidos de la presente invención y/o el agente modulan terapéuticamente la señalización de TGF- β para afectar la metástasis sin afectar adversamente la naturaleza de supresor tumoral de TGF- β . Se sabe que cánceres epiteliales tempranos permanecen sometidos al rol supresor tumoral de TGF-beta y retienen una vía

- 5 de inhibición de crecimiento mediada por TGF-beta funcional (Yingling *et al.* 2004). A medida que las lesiones premalignas se desarrollan, los mecanismos de señalización de TGF-β se alteran y el rol de TGF-β cambia progresivamente de supresor tumoral a promotor tumoral. TGF-β tiene un rol en la conversión de tumores epiteliales tempranos en tumores metastásicos invasivos mediante promoción de transición epitelio-mesenquimal (EMT). TGF-β actúa para promover el progreso y metástasis tumoral una vez que el tumor se ha convertido en un fenotipo mesenquimal.
- 10 Los polipéptidos de la presente invención, solos, como parte de una proteína de fusión y/o en combinación con un agente, preferiblemente se administran a cánceres donde TGF-β actúa como un promotor tumoral. Preferiblemente, los polipéptidos de la presente invención, solos, como parte de una proteína de fusión y/o en combinación con un agente, se administran a cánceres epiteliales no tempranos o lesiones no primarias. Preferiblemente, los cánceres no epiteliales son cánceres invasivos y/o metastásicos. Preferiblemente, los polipéptidos de la presente invención, solos, como parte de una proteína de fusión y/o en combinación con un agente, se administran a tumores que se han convertido en un fenotipo mesenquimal.
- 15 La administración puede lograrse a través de inyecciones subconjuntivales o intratumoralmente mediante administración potenciada por convección o inyección tópica.
- 20 En una realización preferida, un polipéptido de la invención que puede comprender una secuencia que tiene al menos 85% de homología con un polipéptido INSP215, es una proteína de fusión.
- 25 Estas proteínas de fusión pueden obtenerse mediante clonación de un polinucleótido que codifica un polipéptido que comprende una secuencia que tiene al menos 85% de homología con un polipéptido INSP215 en el mismo marco de la secuencia codificante para una secuencia proteica heteróloga. Las proteínas de fusión preferidas incluyen aquellas descritas en los Ejemplos en la presente, cuya secuencia de aminoácidos se proporciona en la SEQ ID NO:19 y SEQ ID NO:21.
- 30 Se pretende que el término "heterólogo", cuando se usa en la presente, designe cualquier polipéptido distinto a un polipéptido INSP215 humano. Ejemplos de secuencias heterólogas, que pueden estar comprendidas en las proteínas de fusión en los extremos N o C, incluyen: dominios extracelulares de proteína unida a membrana, regiones constantes de inmunoglobulina (regiones Fc), dominios de multimerización, dominios de proteínas extracelulares, secuencias señal, secuencia de exportación y secuencias que permiten la purificación mediante cromatografía de afinidad.
- 35 Muchas de estas secuencias heterólogas están disponibles en el mercado en plásmidos de expresión ya que estas secuencias se incluyen comúnmente en proteínas de fusión para proporcionar propiedades adicionales sin deteriorar considerablemente la actividad biológica específica de la proteína fusionada a ellas (Terpe K, 2003, Appl Microbiol Biotechnol, 60:523-33). Ejemplos de tales propiedades adicionales son una semivida más prolongada en fluidos corporales, la localización extracelular o un procedimiento de purificación más fácil como se permite mediante la prolongación de Histidinas formando la denominada "etiqueta de histidina" (Gentz *et al.* 1989, Proc Natl Acad Sci USA, 86:821-4) o mediante la etiqueta "HA", un epítipo derivado de la proteína hemaglutinina de la gripe (Wilson *et al.* 1994, Cell, 37:767-78). Si es necesario, la secuencia heteróloga puede eliminarse mediante una escisión proteolítica, por ejemplo insertando un sitio de escisión proteolítica entre la proteína y la secuencia heteróloga y exponiendo la proteína de fusión purificada a la proteasa apropiada. Estas características tienen particular importancia para las proteínas de fusión ya que facilitan su producción y uso en la preparación de composiciones farmacéuticas. Por ejemplo, el polipéptido INSP215 puede purificarse por medio de un péptido hexa-histidina fusionado en el extremo C terminal de INSP215. Cuando la proteína de fusión comprende una región de inmunoglobulina, la fusión puede ser directa, o a través de un péptido ligador corto que puede ser de tan solo 1 a 3
- 40 residuos de aminoácidos de longitud o más largos, por ejemplo, 13 residuos de aminoácidos de longitud. Dicho ligador puede ser un tripéptido de la secuencia E-F-M (Glu-Phe-Met), por ejemplo, o una secuencia ligadora de 13 aminoácidos que comprende Glu-Phe-Gly-Ala-Gly-Leu-Val-Leu-Gly-Gly-Gln-Phe-Met (SEQ ID NO: 15) introducida entre la secuencia de las sustancias de la invención y la secuencia de inmunoglobulina. La proteína de fusión resultante tiene propiedades mejoradas, tales como un tiempo de residencia extendido en fluidos corporales (es decir, una semivida aumentada), actividad específica aumentada, nivel de expresión aumentado o se facilita la purificación de la proteína de fusión.
- 45
- 50
- 55 En una realización preferida, la proteína se fusiona con la región constante de una molécula de Ig. Preferiblemente, se fusiona con regiones de cadena pesada, como por ejemplo los dominios CH2 y CH3 de IgG1 humana. Otras isoformas de moléculas de Ig también son adecuadas para la generación de proteínas de fusión de acuerdo con la presente invención, tales como las isoformas IgG2 o IgG4, u otras clases de Ig, como por ejemplo IgM o IgA. Las proteínas de fusión pueden ser monoméricas o multiméricas, hetero- u homomultiméricas.
- 60
- 65 En una realización preferida adicional, el derivado funcional comprende al menos un resto unido a uno o más grupos funcionales, que se producen como una o más cadenas laterales en los residuos de aminoácidos. Preferiblemente, es resto es un resto de polietileno (PEG). La PEGilación puede llevarse a cabo mediante métodos conocidos, tales como los descritos en W099/55377.

- 5 Los polipéptidos pueden contener aminoácidos distintos a los aminoácidos codificados de 20 genes, modificados mediante procesos naturales, tal como mediante procesamiento post-traduccional o mediante técnicas de modificación química que se conocen bien en la técnica. Entre las modificaciones conocidas que pueden estar presentes comúnmente en polipéptidos de la presente invención están glicosilación, enlace lipídico, sulfatación, gamma-carboxilación, por ejemplo de residuos de ácido glutámico, hidroxilación y ribosilación de ADP. Otras modificaciones posibles incluyen acetilación, acilación, amidación, enlace covalente de flavina, enlace covalente de un resto hemo, enlace covalente de un nucleótido o derivado de nucleótido, enlace covalente de una derivado de lípido, enlace covalente de fosfatidilinositol, reticulación, ciclización, formación de puente disulfuro, desmetilación, formación de reticulados covalentes, formación de cisteína, formación de piroglutamato, formilación, formación de ancla de GPI, yodación, metilación, miristoilación, oxidación, procesamiento proteolítico, fosforilación, prenilación, racemización, selenoilación, adición de aminoácidos mediada por ARN de transferencia a proteínas tales como arginilación y ubiquitinación.
- 10 Las modificaciones pueden ocurrir en cualquier lugar en un polipéptido, incluso en la estructura principal del péptido, las cadenas laterales de aminoácidos y los extremos terminales amino o carboxilo. De hecho, el bloqueo del extremo terminal amino o carboxilo en un polipéptido, o ambos, mediante una modificación covalente es común en polipéptidos de origen natural y sintéticos y tales modificaciones pueden estar presentes en polipéptidos de la presente invención.
- 15 Las modificaciones que ocurren en un polipéptido con frecuencia estarán en función de cómo se hace el polipéptido. Para polipéptidos que se hacen de forma recombinante, la naturaleza y alcance de las modificaciones en gran medida serán determinados por la capacidad de modificación post-translacional de la célula huésped particular y las señales de modificación que están presentes en la secuencia de aminoácidos del polipéptido en cuestión. Por ejemplo, los patrones de glicosilación varían entre tipos diferentes de células huésped.
- 20 Los polipéptidos de la presente invención pueden prepararse de cualquier modo adecuado. Tales polipéptidos incluyen, donde el polipéptido es un polipéptido de origen natural, polipéptidos de origen natural aislados (por ejemplo purificados a partir de cultivo celular), polipéptidos producidos de forma recombinante (incluidas las proteínas de fusión), polipéptidos producidos sintéticamente y polipéptidos que se producen mediante una combinación de estos métodos. El término "aislado" no denota el método mediante el cual se obtiene el polipéptido o el nivel de pureza de la preparación. Por consiguiente, tales especies aisladas pueden producirse de forma recombinante, aislarse directamente de la célula o tejido de interés o producirse sintéticamente en base a las secuencias determinadas.
- 25 Los polipéptidos funcionalmente equivalentes del primer aspecto de la invención pueden ser polipéptidos que son homólogos a los polipéptidos INSP215. Se dice que dos polipéptidos son "homólogos", tal como se usa el término en la presente, si la secuencia de uno de los polipéptidos tiene un grado suficientemente alto de identidad o similitud con la secuencia del otro polipéptido. "Identidad" indica que en cualquier posición particular en las secuencias alineadas, el residuo de aminoácido es idéntico entre las secuencias. "Similitud" indica que en cualquier posición particular en las secuencias alineadas, el residuo de aminoácido es de un tipo similar entre las secuencias. Los grados de identidad y similitud pueden calcularse fácilmente (Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, Nueva York, 1988; Biocomputing. Informatics and Genome Projects, Smith, D.W., ed., Academic Press, Nueva York, 1993; Computer Analysis of Sequence Data, Parte 1, Griffin, A.M., and Griffin, H.G., eds., Humana Press, Nueva Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; y Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, Nueva York, 1991).
- 30 Los polipéptidos homólogos, por lo tanto, incluyen variantes biológicas naturales (por ejemplo, variantes alélicas o variantes geográficas dentro de las especies de las cuales derivan los polipéptidos) y mutantes (tales como mutantes que contienen sustituciones, inserciones o supresiones de aminoácidos) de los polipéptidos INSP215. Tales mutantes pueden incluir polipéptidos en los que uno o más de los residuos de aminoácidos se sustituyen por un residuo de aminoácido conservado o no conservado (preferiblemente un residuo de aminoácido conservado) y tal residuo de aminoácido sustituido puede o no ser uno codificado por el código genético. Tales sustituciones típicas son entre Ala, Val, Leu e Ile; entre Ser y Thr; entre los residuos ácidos Asp y Glu; entre Asn y Gln; entre los residuos básicos Lys y Arg; o entre los residuos aromáticos Phe y Tyr. Se prefieren particularmente las variantes en las que varios, es decir, entre 5 y 10, 1 y 5, 1 y 3, 1 y 2 o solo 1 aminoácido se sustituye, suprime o agrega en cualquier combinación. Se prefieren especialmente las sustituciones, adiciones o supresiones imperceptibles, que no alteran las propiedades y actividades de la proteína. También se prefieren especialmente en este sentido las sustituciones conservadoras. Tales mutantes también incluyen polipéptidos en los que uno o más de los residuos de aminoácidos incluyen un grupo sustituyente.
- 35 De acuerdo con la presente invención, cualquier sustitución debe ser preferiblemente una sustitución "conservadora" o "segura", que comúnmente se define como una sustitución que introduce un aminoácido que tiene propiedades químicas suficientemente similares (por ejemplo, un aminoácido básico cargado positivamente debe reemplazarse por otro aminoácido básico cargado positivamente), para conservar la estructura y la función biológica de la molécula.
- 40
- 45
- 50
- 55
- 60
- 65

- 5 La bibliografía proporciona muchos modelos en los que la selección de sustituciones de aminoácidos conservadoras pueden llevarse a cabo en base a estudios estadísticos y físico-químicos en la secuencia y/o en la estructura de las proteínas (Rogov SI and Nekrasov AN, 2001). Los experimentos de diseño de proteínas mostraron que el uso de subconjuntos específicos de aminoácidos puede producir proteínas plegables y activas, ayudando en la clasificación de sustituciones "sinónimas" de aminoácidos que pueden alojarse más fácilmente en la estructura de la proteína, y que pueden usarse para detectar homólogos y parálogos funcionales y estructurales (Murphy LR *et al.*, 2000). Los grupos de aminoácidos sinónimos y los grupos de aminoácidos sinónimos más preferidos se muestran en la Tabla 1.
- 10 También pueden introducirse mutaciones no conservadoras específicas en los polipéptidos de la invención con diferentes fines. Las mutaciones que reducen la afinidad de la proteína pueden aumentar su capacidad para reutilizarse y reciclarse, aumentando potencialmente su potencia terapéutica (Robinson CR, 2002). Los epítomos inmunogénicos presentes finalmente en los polipéptidos de la invención pueden aprovecharse para desarrollar vacunas (Stevanovic S, 2002), o eliminarse modificando su secuencia siguiendo métodos conocidos para seleccionar mutaciones para aumentar la estabilidad de la proteína, y corregirlas (van den Burg B y Eijsink V, 2002; WO 02/05146, WO 00/34317, WO 98/52976).
- 15 Los grupos sinónimos alternativos preferidos para derivados de aminoácidos incluidos en los miméticos peptídicos son aquellos definidos en la Tabla 2. Una lista no taxativa de derivados de aminoácidos también incluye ácido aminoisobutírico (Aib), hidroxiprolina (Hyp), 1,2,3,4-tetrahidro-isoquinolina-3-COOH, ácido indolina-2-carboxílico, 4-difluoro-prolina, ácido L-tiazolidina-4-carboxílico, L-homoprolina, 3,4-dehidro-prolina, 3,4-dihidroxi-fenilalanina, ciclohexil-glicina y fenilglicina.
- 20 Por "derivado de aminoácido" se pretende significar un aminoácido o entidad química similar a un aminoácido distinta de uno de los 20 aminoácidos de origen natural codificados genéticamente. En particular, el derivado de aminoácido puede contener restos alquilo sustituidos o no sustituidos, lineales, ramificados o cíclicos, y puede incluir uno o más heteroátomos. Los derivados de aminoácidos pueden hacerse de novo u obtenerse a partir de fuentes comerciales (Calbiochem-Novabiochem AG, Suiza; Bachem, EE.UU.).
- 25 En la bibliografía se divulgan varias metodologías para incorporar derivados de aminoácidos no naturales en proteínas, usando sistemas de traducción *in vitro* e *in vivo*, para sondear y/o mejorar la estructura y función de la proteína (Dougherty DA, 2000). Las técnicas para la síntesis y el desarrollo de miméticos peptídicos, así como también miméticos no peptídicos, también se conocen en la técnica (Golebiowski A *et al.*, 2001; Hruby VJ and Balse PM, 2000; Sawyer TK, en "Structure Based Drug Design", editado por Veerapandian P, Marcel Dekker Inc., pg. 557-663, 1997).
- 30 Tales mutantes también incluyen polipéptidos en los que uno o más de los residuos de aminoácidos incluyen un grupo sustituyente.
- 35 Típicamente, más de 30% de identidad entre dos polipéptidos se considera un indicador de equivalencia funcional. Preferiblemente, los polipéptidos funcionalmente equivalentes del primer aspecto de la invención tienen un grado de identidad de secuencia con los polipéptidos INSP215, o con fragmentos activos de los mismos, mayor a 80%. Los polipéptidos más preferidos tienen un grado de identidad mayor a 85%, 90%, 95%, 98% o 99%, respectivamente.
- 40 Los polipéptidos funcionalmente equivalentes del primer aspecto de la invención también pueden ser polipéptidos que se han identificado usando una o más técnicas de alineación estructural. Por ejemplo, la tecnología Inpharmatica Genome Threader que forma un aspecto de las herramientas de búsqueda usadas para generar la base de datos de búsqueda Biopendium puede usarse (ver solicitud PCT PCT/GB01/01105, también pendiente, publicada como WO 01/69507) para identificar polipéptidos de función desconocida actualmente que, aunque tienen identidad de secuencia baja en comparación con los polipéptidos INSP215, se prevé que sean polipéptidos que contienen el subdominio ZP/TGR3, en virtud de compartir homología estructural considerable con las secuencias de polipéptidos INSP215. Por "homología estructural considerable" se pretende significar que el Inpharmatica Genome Threader prevé que dos proteínas comparten homología estructural con una certeza de 10% y superior.
- 45 Los polipéptidos del primer aspecto de la invención también incluyen fragmentos de los polipéptidos INSP215 y fragmentos de los equivalentes funcionales de los polipéptidos INSP215, siempre que dichos fragmentos contengan un subdominio ZP/TGR3 o tengan un determinante antigénico en común con los polipéptidos INSP215.
- 50 Tal como se usa en la presente, el término "fragmento" se refiere a un polipéptido que tiene una secuencia de aminoácidos que es igual a una parte, pero no a la totalidad, de la secuencia de aminoácidos de los polipéptidos INSP215 o de uno de sus equivalentes funcionales. Los fragmentos deben comprender al menos *n* aminoácidos consecutivos de la secuencia y, dependiendo de la secuencia particular, *n* preferiblemente es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20 o más). Los fragmentos pequeños pueden formar un determinante antigénico.
- 55 Los ácidos nucleicos de acuerdo con la invención tienen preferiblemente 10-2000 nucleótidos de longitud, preferiblemente 100-1750 nucleótidos, preferiblemente 500-1500, preferiblemente 600-1200, preferiblemente 750-1000 nucleótidos de longitud. Los polipéptidos de acuerdo con la invención tienen preferiblemente 10-700

aminoácidos de longitud, preferiblemente 50-600, preferiblemente 100-500, preferiblemente 200-400, preferiblemente 300-375 aminoácidos de longitud.

5 Los fragmentos de los polipéptidos de exón INSP215 y los polipéptidos INSP215 pueden consistir en combinaciones de 2, 3, 4, 5 o 6 secuencias de exones vecinas en el polipéptido INSP215, respectivamente. Preferiblemente, los exones se combinan para coincidir con los dominios en INSP215 identificados en la presente.

10 Tales fragmentos pueden ser "independientes", es decir, no ser parte o estar fusionados con otros aminoácidos o polipéptidos, o pueden estar comprendidos dentro de un polipéptido mayor del que forman una parte o región. Cuando está comprendido dentro de un polipéptido mayor, el fragmento de la invención forma más preferiblemente una única región continua. Por ejemplo, ciertas realizaciones preferidas se refieren a un fragmento que tienen una región pre y/o pro polipéptido fusionada al extremo terminal amino del fragmento y/o una región adicional fusionada al extremo terminal carboxilo del fragmento. Sin embargo, varios fragmentos pueden estar comprendidos dentro de un único polipéptido mayor.

15 Los polipéptidos de la presente invención o sus fragmentos inmunogénicos (que comprenden al menos un determinante antigénico) pueden usarse para generar ligandos, tales como anticuerpos policlonales o monoclonales, que son inmunoespecíficos para los polipéptidos. Tales anticuerpos pueden emplearse para aislar o para identificar clones que expresan los polipéptidos de la invención o para purificar los polipéptidos mediante cromatografía de afinidad. Los anticuerpos también pueden emplearse como auxiliares de diagnóstico o terapéuticos, entre otras aplicaciones, como será evidente para el lector experto.

20 El término "inmunoespecífico" significa que los anticuerpos tienen afinidad sustancialmente mayor con los polipéptidos de la invención que su afinidad con otros polipéptidos relacionados en la técnica anterior. Tal como se usa en la presente, el término "anticuerpo" se refiere a moléculas intactas así como también a fragmentos de las mismas, tales como Fab, F(ab')₂ y Fv, que son capaces de unirse al determinante antigénico en cuestión. Tales anticuerpos por consiguiente se unen a los polipéptidos del primer aspecto de la invención.

25 Por "afinidad sustancialmente mayor" pretendemos significar que hay un aumento medible en la afinidad con un polipéptido de la invención en comparación con la afinidad con polipéptidos que contienen el dominio ZP/TGR3 conocidos. Preferiblemente, la afinidad es de al menos 1,5 veces, 2 veces, 5 veces, 10 veces, 100 veces, 10³ veces, 10⁴ veces, 10⁵ veces o 10⁶ veces mayor con un polipéptido de la invención que con polipéptidos que contienen el dominio ZP/TGR3 conocidos.

30 Preferiblemente, hay un aumento medible en la afinidad con un polipéptido de la invención en comparación con proteínas que contienen el dominio ZP conocidas.

35 Si se desean anticuerpos policlonales, un mamífero seleccionado, tal como un ratón, conejo, cabra o caballo, puede inmunizarse con un polipéptido del primer aspecto de la invención. El polipéptido usado para inmunizar el animal puede derivarse mediante tecnología de ADN recombinante o puede sintetizarse químicamente. Si se desea, el polipéptido puede conjugarse con una proteína portadora. Los portadores usados comúnmente a los que los polipéptidos pueden estar químicamente acoplados incluyen albúmina de suero bovino, tiroglobulina y hemocianina de lapa californiana. El polipéptido acoplado luego se usa para inmunizar el animal. Se recoge suero del animal inmunizado y se trata de acuerdo con procedimientos conocidos, por ejemplo mediante cromatografía de inmunoafinidad.

40 Los anticuerpos monoclonales para los polipéptidos del primer aspecto de la invención también pueden ser producidos fácilmente por un experto en la técnica. La metodología general para producir anticuerpos monoclonales usando tecnología de hibridomas se conoce bien (ver, por ejemplo, Kohler, G. and Milstein, C, Nature 256: 495-497 (1975); Kozbor *et al.*, Immunology Today 4: 72 (1983); Cole *et al.*, 77-96 en Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc. (1985).

45 Los paneles de anticuerpos monoclonales producidos contra los polipéptidos del primer aspecto de la invención pueden detectarse sistemáticamente para varias propiedades, es decir, para isotipo, epítipo, afinidad, etc. Los anticuerpos monoclonales son particularmente útiles en la purificación de los polipéptidos individuales contra los que se dirigen. De forma alternativa, los genes que codifican los anticuerpos monoclonales de interés pueden aislarse a partir de hibridomas, por ejemplo, mediante técnicas de PCR conocidas en la técnica, y clonarse y expresarse en vectores apropiados.

50 Los anticuerpos quiméricos, en los que se unen o fusionan regiones variables no humanas con regiones constantes humanas (ver, por ejemplo, Liu *et al.*, Proc. Natl. Acad. Sci. EE.UU., 84, 3439 (1987)), también pueden usarse. El anticuerpo puede modificarse para hacerlo menos inmunogénico en un individuo, por ejemplo mediante humanización (ver Jones *et al.*, Nature, 321, 522 (1986); Verhoeyen *et al.*, Science, 239, 1534 (1988); Kabat *et al.*, J. Immunol., 147, 1709 (1991); Queen *et al.*, Proc. Natl Acad. Sci. EE.UU., 86, 10029 (1989); Gorman *et al.*, Proc. Natl Acad. Sci. EE.UU., 88, 34181 (1991); y Hodgson *et al.*, Bio/Technology, 9, 421 (1991)). La expresión "anticuerpo

- 5 humanizado", tal como se usa en la presente, se refiere a moléculas de anticuerpo en las que los aminoácidos CDR y otros aminoácidos seleccionados en los dominios variables de las cadenas pesadas y/o ligeras de un anticuerpo donante no humano se han sustituido en lugar de aminoácidos equivalentes en un anticuerpo humano. El anticuerpo humanizado por consiguiente es muy similar a un anticuerpo humano pero tiene la capacidad de unión del anticuerpo donante.
- En una alternativa adicional, el anticuerpo puede ser un anticuerpo "biespecífico", que es un anticuerpo que tiene dos dominios de unión a antígeno diferentes, estando cada dominio dirigido contra un epítipo diferente.
- 10 La tecnología de presentación de fagos puede utilizarse para seleccionar genes que codifican anticuerpos con actividades de unión hacia polipéptidos de la invención a partir de repertorios de PCR ampliados, genes V de linfocitos de humanos detectados por poseer los anticuerpos relevantes, o a partir de bibliotecas naive (McCafferty, J. *et al.*, (1990), Nature 348, 552-554; Marks, J. *et al.*, (1992) Biotechnology 10, 779-783). La afinidad de estos anticuerpos también puede mejorarse mediante intercambio de cadenas (Clackson, T. *et al.*, (1991) Nature 352, 624-628).
- 15 Los anticuerpos generados mediante las técnicas anteriores, ya sean policlonales o monoclonales, tienen utilidad adicional porque pueden emplearse como reactivos en inmunoensayos, radioinmunoensayos (RIA) o ensayo inmunoabsorbente ligado a enzima (ELISA). En estas aplicaciones, los anticuerpos pueden etiquetarse con un reactivo detectable analíticamente tal como un radioisótopo, una molécula fluorescente o una enzima.
- 20 Las moléculas de ácido nucleico preferidas del segundo y tercer aspecto de la divulgación son aquellas que codifican las secuencia de polipéptidos descrita en SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ED NO:8, SEQ ED NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ED NO:19 o SEQ ED NO:21 y polipéptidos funcionalmente equivalentes. Estas moléculas de ácido nucleico pueden usarse en los métodos y aplicaciones descritas en la presente. Las moléculas de ácido nucleico de la invención comprenden preferiblemente al menos n nucleótidos consecutivos de las secuencias divulgadas en la presente donde, dependiendo de la secuencia particular, n es 10 o más (por ejemplo, 12,14,15,18, 20, 25, 30, 35, 40 o más).
- 25 Las moléculas de ácido nucleico de la invención también incluyen secuencias que son complementarias a moléculas de ácido nucleico descritas anteriormente (por ejemplo, con fines antisentido o sondeo).
- 30 Las moléculas de ácido nucleico de la presente invención pueden estar en forma de ARN, tal como ARNm, o en forma de ADN, incluidos, por ejemplo ADNc, ADN sintético o ADN genómico. Tales moléculas de ácido nucleico pueden obtenerse mediante clonación, mediante técnicas sintéticas químicas o mediante una combinación de las mismas. Las moléculas de ácido nucleico pueden prepararse, por ejemplo, mediante síntesis química usando técnicas tales como síntesis química de fosoramidita en fase sólida, a partir de bibliotecas genómicas o de ADNc o mediante separación a partir de un organismo. Las moléculas de ARN pueden generarse en general mediante la transcripción in vitro o in vivo de secuencias de ADN.
- 35 Las moléculas de ácido nucleico pueden ser de cadena doble o de cadena simple. El ADN de cadena simple puede ser la hebra codificadora, también conocida como la hebra con sentido, o puede ser la hebra no codificadora, también denominada hebra anti-sentido.
- 40 La expresión "molécula de ácido nucleico" también incluye análogos de ADN y ARN, tales como aquellos que contienen estructuras principales modificadas, y ácidos nucleicos de péptido (PNA). El término "PNA", tal como se usa en la presente, se refiere a un molécula antisentido o un agente antígeno que comprende un oligonucleótido de al menos cinco nucleótidos de longitud ligado a una estructura principal de péptido de residuos de aminoácidos, que preferiblemente termina en lisina. La lisina terminal confiere solubilidad a la composición. Los PNAs pueden pegilarse para extender su tiempo de vida en una célula, donde preferentemente se unen complementariamente a ADN y ARN de cadena simple y detienen la elongación de la transcripción (Nielsen, P.E. *et al.* (1993) Anticancer Drug Des. 8:53-63).
- 45 Una molécula de ácido nucleico que codifica el polipéptido de la SEQ ID NO:2 puede ser idéntica a la secuencia codificante de la molécula de ácido nucleico que se muestra en la SEQ ID NO:1. Una molécula de ácido nucleico que codifica el polipéptido de la SEQ ID NO:4 puede ser idéntica a la secuencia codificante de la molécula de ácido nucleico que se muestra en la SEQ ID NO:3. Una molécula de ácido nucleico que codifica el polipéptido de la SEQ ID NO:6 puede ser idéntica a la secuencia codificante de la molécula de ácido nucleico que se muestra en la SEQ ID NO:5. Una molécula de ácido nucleico que codifica el polipéptido de la SEQ ID NO:8 puede ser idéntica a la secuencia codificante de la molécula de ácido nucleico que se muestra en la SEQ ID NO:7. Una molécula de ácido nucleico que codifica el polipéptido de la SEQ ID NO:10 puede ser idéntica a la secuencia codificante de la molécula de ácido nucleico que se muestra en la SEQ ID NO:9. Una molécula de ácido nucleico que codifica el polipéptido de la SEQ ID NO:12 puede ser idéntica a la secuencia codificante de la molécula de ácido nucleico que se muestra en la SEQ ID NO:11. Una molécula de ácido nucleico que codifica el polipéptido de la SEQ ID NO:14 puede ser idéntica a la secuencia codificante de la molécula de ácido nucleico que se muestra en la SEQ ID NO:13. Una molécula de ácido nucleico que codifica el polipéptido de la SEQ ID NO:19 puede ser idéntica a la secuencia codificante de la
- 50
- 55
- 60
- 65

molécula de ácido nucleico que se muestra en la SEQ ID NO:18. Una molécula de ácido nucleico que codifica el polipéptido de la SEQ ID NO:21 puede ser idéntica a la secuencia codificante de la molécula de ácido nucleico que se muestra en la SEQ ID NO:20.

5 Estas moléculas también pueden tener una secuencia diferente que, como resultado de la degeneración del código genético, codifica un polipéptido de SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:19 o SEQ ID NO:21. Tales moléculas de ácido nucleico que codifican el polipéptido de SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:19 o SEQ ID NO:21 pueden incluir, a modo no taxativo, la secuencia codificante para el polipéptido maduro en sí mismo; la secuencia codificante para el polipéptido maduro y secuencias codificantes adicionales, tales como aquellas que codifican una secuencia secretora o líder, tales como una secuencia de pro, pre o prepro polipéptido; la secuencia codificante para el polipéptido maduro, con o sin las secuencias codificantes adicionales mencionadas anteriormente, junto además con secuencias no codificantes adicionales, incluidas las secuencias no codificantes 5' y 3', tales como las secuencias transcritas no traducidas que tienen un rol en la transcripción (incluidas las señales de terminación), unión del ribosoma y estabilidad de ARNm. Las moléculas de ácido nucleico también pueden incluir secuencias adicionales que codifican aminoácidos adicionales, tales como aquellos que proporcionan funcionalidades adicionales.

20 Las moléculas de ácido nucleico del segundo y tercer aspecto de la divulgación también pueden codificar los fragmentos o los equivalentes funcionales de los polipéptidos y fragmentos del primer aspecto de la invención. Tal molécula de ácido nucleico puede ser una variante de origen natural tal como una variante alélica de origen natural, o la molécula puede ser una variante que se sabe que no tiene origen natural. Tales variantes que no tienen origen natural de la molécula de ácido nucleico pueden hacerse mediante técnicas de mutagénesis, incluidas aquellas aplicadas a moléculas de ácido nucleico, células u organismos.

25 Entre las variantes en este sentido están las variantes que difieren de las moléculas de ácido nucleico mencionadas anteriormente mediante sustituciones, supresiones o inserciones de nucleótidos. Las sustituciones, supresiones o inserciones pueden involucrar uno o más nucleótidos. Las variantes pueden alterarse en regiones codificantes o no codificantes o en ambas. Las alteraciones en las regiones codificantes pueden producir sustituciones, supresiones o inserciones de aminoácidos conservadoras o no conservadoras.

35 Las moléculas de ácido nucleico de la invención también pueden manipularse, usando métodos generalmente conocidos en la técnica, por una variedad de razones, incluso modificar la clonación, procesamiento y/o expresión del producto del gen (el polipéptido). El barajado de ADN mediante fragmentación aleatoria y reensamblaje por PCR de fragmentos de gen y oligonucleótidos sintéticos se incluyen como técnicas que pueden usarse para manipular las secuencias de nucleótidos. La mutagénesis sitio-dirigida puede usarse para insertar nuevos sitios de restricción, alterar patrones de glicosilación, cambiar la preferencia codónica, producir variantes de empalme, introducir mutaciones, etc.

40 Las moléculas de ácido nucleico que codifican un polipéptido del primer aspecto de la invención pueden ligarse a una secuencia heteróloga de modo tal que la molécula de ácido nucleico codifica una proteína de fusión. Tales moléculas de ácido nucleico combinadas se incluyen dentro del segundo o tercer aspecto de la invención. Por ejemplo, para detectar en bibliotecas de péptidos inhibidores de la actividad del polipéptido, puede ser útil expresar, usando tal molécula de ácido nucleico combinada, una proteína de fusión que puede reconocerse mediante un anticuerpo disponible en el mercado. Una proteína de fusión también puede manipularse para contener un sitio de escisión ubicado entre la secuencia del polipéptido de la invención y la secuencia de una proteína heteróloga de modo tal que el polipéptido puede escindirse y purificarse fuera de la proteína heteróloga.

50 Las moléculas de ácido nucleico de la invención también incluyen moléculas antisentido que son parcialmente complementarias a las moléculas de ácido nucleico que codifican polipéptidos de la presente invención y que por lo tanto hibridizan para las moléculas de ácido nucleico codificantes (hibridación). Tales moléculas antisentido, tales como oligonucleótidos, pueden diseñarse para reconocer, unirse específicamente y evitar la transcripción de un ácido nucleico objetivo que codifica un polipéptido de la invención, como será conocido para los expertos en la técnica (ver, por ejemplo, Cohen, J.S., Trends in Pharm. Sci., 10, 435 (1989), Okano, J. Neurochem. 56, 560 (1991); O'Connor, J. Neurochem 56, 560 (1991); Lee *et al.*, Nucleic Acids Res 6, 3073 (1979); Cooney *et al.*, Science 241, 456 (1988); Dervan *et al.*, Science 251, 1360 (1991).

60 El término "hibridación" tal como se usa en la presente se refiere a la asociación de dos moléculas de ácido nucleico entre sí mediante un enlace de hidrógeno. Típicamente, se fijará una molécula a un soporte sólido y la otra estará libre en la solución. Luego, las dos moléculas pueden colocarse en contacto entre sí en condiciones que favorecen el enlace de hidrógeno. Los factores que afectan este enlace incluyen: el tipo y volumen de disolvente; temperatura de reacción; tiempo de hibridación; agitación; agentes para bloquear el acoplamiento no específico de la molécula de fase líquida al soporte sólido (reactivo de Denhardt o BLOTTO); la concentración de las moléculas; el uso de compuestos para aumentar la tasa de asociación de moléculas (sulfato de dextrano o polietilenglicol); y la rigurosidad de las condiciones de lavado luego de la hibridación (ver Sambrook *et al.* [supra]).

65

La inhibición de la hibridación de una molécula completamente complementaria para obtener una molécula objetivo puede examinarse usando un ensayo de hibridación, como se conoce en la técnica (ver, por ejemplo, Sambrook *et al.* [supra]). Una molécula sustancialmente homóloga luego competirá e inhibirá la unión de una molécula completamente homóloga con la molécula objetivo en diversas condiciones de rigurosidad, como se describe en Wahl, G.M. and S.L. Berger (1987; *Methods Enzymol.* 152:399-407) y Kimmel, A.R. (1987; *Methods Enzymol.* 152:507-511).

"Rigurosidad" se refiere a condiciones en una reacción de hibridación que favorecen la asociación de moléculas muy similares con respecto a la asociación de moléculas que difieren.

Las condiciones de hibridación de rigurosidad alta se definen como incubación durante toda la noche a 42°C en una solución que comprende 50% de formamida, 5XSSC (150mM de NaCl, 15mM de citrato trisódico), 50mM de fosfato de sodio (pH 7,6), solución de Denhardt 5x, 10% de sulfato de dextrano y 20 microgramos/ml de ADN de esperma de salmón cizallado desnaturalizado y luego lavado de los filtros en 0,1X SSC a aproximadamente 65°C. Las condiciones de rigurosidad baja implican que la reacción de hibridación se lleve a cabo a 35°C (ver Sambrook *et al.* [supra]). Preferiblemente, las condiciones usadas para la hibridación son aquellas de rigurosidad alta.

Las realizaciones preferidas de este aspecto de la invención son moléculas de ácido nucleico que son al menos 70% idénticas en su longitud total a una molécula de ácido nucleico que codifica los polipéptidos INSP215 (SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:19 o SEQ ID NO:21) y moléculas de ácido nucleico que son sustancialmente complementarias a tales moléculas de ácido nucleico. Preferiblemente, una molécula de ácido nucleico de acuerdo con este aspecto de la invención comprende una región que es al menos 80% idéntica en su longitud total a tales secuencias codificantes, o es una molécula de ácido nucleico que es complementaria a las mismas. En este sentido, las moléculas de ácido nucleico que son al menos 90%, preferiblemente al menos 95%, más preferiblemente al menos 98%, 99% o más idénticas en su longitud total a la misma son particularmente preferidas. Las realizaciones preferidas en este sentido son moléculas de ácido nucleico que codifican polipéptidos que retienen sustancialmente la misma función o actividad biológica que los polipéptidos INSP215.

También se describe en la presente un proceso para detectar una molécula de ácido nucleico de la invención, que comprende los pasos de: (a) poner con contacto una sonda nucleica de acuerdo con la invención con una muestra biológica en condiciones de hibridación para formar duplex;es y (b) detectar cualquiera de tales duplexes que se formaron.

Tal como se describe adicionalmente a continuación en conexión con ensayos que pueden utilizarse de acuerdo con la invención, puede usarse una molécula de ácido nucleico como se describió anteriormente como una sonda de hibridación para ARN, ADNc o ADN genómico, para aislar clones de ADNcs de longitud completa y genómicos que codifican polipéptidos INSP215 y para aislar clones de ADNc y genómicos de genes homólogos u ortólogos que tienen una alta similitud de secuencia con el gen que codifica este polipéptido.

En este sentido, las siguientes técnicas, entre otras conocidas en la técnica, pueden utilizarse y se describen a continuación a efectos de ilustración. Los métodos para secuenciación y análisis de ADN se conocen y están en general disponibles en la técnica y pueden, en efecto, usarse para practicar muchas de las realizaciones de la invención descritas en la presente. Tales métodos pueden emplear tales enzimas como el fragmento Klenow de la ADN polimerasa I, Secuensa (US Biochemical Corp, Cleveland, OH), polimerasa Taq (Perkin Elmer), polimerasa T7 termoestable (Amersham, Chicago, IL) o combinaciones de polimerasas y exonucleasas correctoras tales como las encontradas en ELONGASE Amplification System comercializado por Gibco/BRL (Gaithersburg, MD). Preferiblemente, el proceso de secuenciación puede automatizarse usando máquinas tales como Hamilton Micro Lab 2200 (Hamilton, Reno, NV), Peltier Thermal Cycler (PTC200; MJ Research, Watertown, MA) y Secuenciadores de ADN ABI Catalyst y 373 y 377 (Perkin Elmer).

Un método para aislar una molécula de ácido nucleico que codifica un polipéptido con una función equivalente a la de los polipéptidos INSP215 es sondear una biblioteca genómica o de ADNc con una sonda natural o diseñada artificialmente usando procedimientos estándares que se reconocen en la técnica (ver, por ejemplo, "Current Protocols in Molecular Biology", Ausubei *et al.* (eds). Greene Publishing Association and John Wiley Interscience, Nueva York, 1989,1992). Las sondas que comprenden al menos 15, preferiblemente al menos 30 y más preferiblemente al menos 50 bases contiguas que corresponden o son complementarias a secuencias de ácido nucleico del gen codificante apropiado (SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:18 o SEQ ID NO:20) son sondas particularmente útiles. Tales sondas pueden etiquetarse con un reactivo detectable analíticamente para facilitar su identificación. Los reactivos útiles incluyen, a modo no taxativo, radioisótopos, tintes fluorescentes y enzimas que son capaces de catalizar la formación de un producto detectable. Usando estas sondas, un experto en la técnica será capaz de aislar copias complementarias de polinucleótidos de ADN, ADNc o ARN genómico que codifican proteínas de interés a partir de fuentes humanas, mamíferas o de otros animales y detectar sistemáticamente tales fuentes para secuencias relacionadas, por ejemplo, para miembros adicionales de la familia, tipo y/o subtipo.

En muchos casos, las secuencias de ADNc aisladas estarán incompletas ya que la región que codifica el polipéptido se interrumpirá, normalmente en el extremo 5'. Existen diversos métodos disponibles para obtener ADNcs de longitud completa, o para extender ADNcs cortos. Tales secuencias pueden extenderse utilizando una secuencia de nucleótidos parcial y empleando diversos métodos conocidos en la técnica para detectar secuencias corriente arriba tales como promotores y elementos reguladores. Por ejemplo, un método que puede emplearse se basa en el método de Amplificación Rápida de extremos de ADNc (RACE; ver, por ejemplo, Frohman *et al.*, PNAS USA 85, 8998-9002, 1988). Las modificaciones recientes a esta técnica, ejemplificadas mediante la tecnología Marathon™ (Clontech Laboratories Inc.), por ejemplo, han simplificado considerablemente la búsqueda de ADNcs más largos. Una técnica ligeramente diferente, denominada PCR de "sitio de restricción", usa cebadores universales para recuperar secuencias de ácido nucleico desconocidas adyacentes a un locus conocido (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). También puede usarse PCR inversa para ampliar o extender secuencias usando cebadores divergentes en base a una región conocida (Triglia, T. *et al.* (1988) Nucleic Acids Res. 16:8186). Otro método que puede usarse es PCR de captura que implica amplificación PCR de fragmentos de ADN adyacentes a una secuencia conocida en humanos y ADN de cromosoma artificial de levadura (Lagerstrom, M. *et al.* (1991) PCR Methods Applic. 1, 111-119). Otro método que puede usarse para recuperar secuencias desconocidas es el de Parker, J.D. *et al.* (1991); Nucleic Acids Res. 19:3055-3060). Adicionalmente, se puede usar PCR, cebadores anidados y bibliotecas PromoterFinder™ para desplazar el ADN genómico (Clontech, Palo Alto, CA). Este proceso evita la necesidad de detectar sistemáticamente bibliotecas y es útil para encontrar uniones intrón/exón.

Cuando se detecta sistemáticamente para ADNcs, es preferible usar bibliotecas que se han seleccionado por tamaño para incluir ADNcs más largos. También son preferibles bibliotecas cebadas aleatoriamente, ya que contendrán más secuencias que contienen las regiones 5' de los genes. El uso de una biblioteca cebada aleatoriamente puede ser especialmente preferible para situaciones en las que una biblioteca de oligo d(T) no proporciona un ADNc de longitud completa. Las bibliotecas genómicas pueden ser útiles para extensión de la secuencia en regiones 5' reguladoras no transcritas.

En la presente también se describen las moléculas de ácido nucleico de la presente invención que pueden usarse para localización de cromosomas. En esta técnica, una molécula de ácido nucleico se dirige específicamente y puede hibridarse con una ubicación particular en un cromosoma humano individual. El mapeo de secuencias relevantes para cromosomas de acuerdo con la presente invención es un paso importante en la correlación confirmatoria de aquellas secuencias con las enfermedades asociadas con el gen. Una vez que se mapeó una secuencia para una ubicación cromosómica precisa, la posición física de la secuencia en el cromosoma puede correlacionarse con datos de mapa genético. Tales datos se encuentran, por ejemplo, en V. McKusick, Mendelian Inheritance in Man (disponible en línea a través de Johns Hopkins University Welch Medical Library). Las relaciones entre genes y enfermedades que se han mapeado en la misma región cromosómica luego se identifican a través de análisis de ligamiento (coherencia de genes físicamente adyacentes). Esto proporciona información valiosa para investigadores que buscan genes de enfermedades usando clonación posicional u otras técnicas de descubrimiento de genes. Una vez que la enfermedad o síndrome se ha localizado rudimentariamente mediante ligamiento genético a una región genómica particular, cualquier mapeo de secuencias a esa área puede representar genes asociados o reguladores para investigación adicional. La molécula de ácido nucleico también puede usarse para detectar diferencias en la ubicación cromosómica debido a la translocación, inversión, etc. entre individuos normales, portadores o afectados.

Las moléculas de ácido nucleico de la presente invención también son valiosas para localización de tejido. Tales técnicas permiten la determinación de patrones de expresión del polipéptido en tejidos mediante detección de las ARNms que los codifican. Estas técnicas incluyen técnicas de hibridación in situ y técnicas de amplificación de nucleótidos, tales como PCR. Los resultados de estos estudios proporcionan un indicio de las funciones del polipéptido en el organismo. Adicionalmente, los estudios comparativos del patrón de expresión normal de ARNms con el de ARNms codificados mediante un gen mutante proporcionan abordajes valiosos sobre el rol de polipéptidos mutantes en una enfermedad. Tal expresión inapropiada puede ser de naturaleza temporal, espacial o cuantitativa.

Los abordajes de silenciamiento génico también pueden ejecutarse para regular hacia abajo la expresión endógena de un gen que codifica un polipéptido de la invención. La interferencia de ARN (ARNi) (Elbashir, SM *et al.*, Nature 2001, 411, 494-498) es un método de silenciamiento génico post-transcripcional específico de secuencia que puede emplearse. Los oligonucleótidos de ARNs se sintetizan in vitro y se introducen en una célula. La unión específica de secuencia de estos oligonucleótidos de ARNs disipa la degradación del ARNm objetivo, provocando la reducción o ablación de la expresión de la proteína objetivo.

La eficacia de los abordajes de silenciamiento génico evaluados anteriormente puede evaluarse a través de la medición de la expresión de polipéptidos (por ejemplo, mediante transferencia Western) y a nivel del ARN usando metodologías en base a TaqMan.

Los vectores de la presente invención comprenden moléculas de ácido nucleico de la invención y pueden ser vectores de clonación o expresión. Las células huésped de la invención, que pueden transformarse, transfectarse o transducirse con los vectores de la invención pueden ser procariontas o eucariotas.

Los polipéptidos de la invención pueden prepararse en forma recombinante mediante expresión de sus moléculas de ácido nucleico codificantes en vectores contenidos dentro de una célula huésped. Los expertos en la técnica conocen tales métodos de expresión y muchos están descritos en detalle por Sambrook et al (supra) y Fernandez & Hoeffler (1998, eds. "Gene expression systems. Using nature for the art of expression". Academic Press, San Diego, Londres, Boston, Nueva York, Sidney, Tokio, Toronto).

Generalmente, puede usarse cualquier sistema o vector que es adecuado para mantener, propagar o expresar moléculas de ácido nucleico para producir un polipéptido en el huésped necesario. La secuencia de nucleótidos apropiada puede insertarse en una sistema de expresión mediante cualquiera de una variedad de técnicas habituales y conocidas, tales como, por ejemplo, las descritas en Sambrook *et al.*, (supra). Generalmente, el gen codificante puede colocarse bajo el control de un elemento de control tal como un promotor, sitio de unión al ribosoma (para expresión bacteriana) y, opcionalmente, un operador, de modo tal que la secuencia de ADN que codifica el polipéptido deseado se transcriba en ARN en la célula huésped transformada.

Ejemplos de sistemas de expresión adecuados incluyen, por ejemplo, sistemas cromosómicos, episomáticos y derivados de virus, incluidos, por ejemplo, vectores derivados de: plásmidos bacterianos, bacteriófagos, transposones, episomas de levadura, elementos de inserción, elementos cromosómicos de levadura, virus tales como baculovirus, virus papova tales como SV40, virus vacuna, adenovirus, virus de la varicela, virus de la seudorrabia y retrovirus, o combinaciones de los mismos, tales como aquellos derivados de elementos genéticos de plásmidos y bacteriófagos, incluidos cósmidos y fagómidos. Los cromosomas humanos artificiales (HACs) también pueden emplearse para administrar fragmentos más grandes de ADN de los que pueden estar contenidos y expresados en un plásmido.

Los vectores pDONR221-INSP215-6HIS, pDONR221-INSP215-EC-6HIS, pDEST12.2_INSP215-EC-6HIS, pEAK12d_INSP215-EC-6HIS, pENTR_INSP215-6HIS, pEAK12d_INSP215-6HIS y pDEST12.2_INSP215-6HIS son ejemplos preferidos de vectores adecuados para uso de acuerdo con esta invención.

Sistemas de expresión particularmente adecuados incluyen microorganismos tales como bacterias transformadas con bacteriófago recombinante, vectores de expresión de ADN plásmidos o cósmidos; levadura transformada con vectores de expresión de levadura; sistemas celulares de insectos infectados con vectores de expresión de virus (por ejemplo, baculovirus); sistemas celulares vegetales transformados con sistemas de expresión de virus (por ejemplo, virus del mosaico de la coliflor, CaMV; virus del mosaico del tabaco, TMV) o con vectores de expresión bacteriana (por ejemplo, plásmidos Ti o pBR322); o sistemas celulares animales. También pueden emplearse sistemas de traducción libres de células para producir los polipéptidos de la invención.

La introducción de moléculas de ácido nucleico que codifican un polipéptido de la presente invención en células huésped puede llevarse a cabo mediante métodos descritos en muchos manuales de laboratorio estándares, tales como Davis *et al.*, Basic Methods in Molecular Biology (1986) y Sambrook *et al.*, [supra]. Los métodos particularmente adecuados incluyen transfección del fosfato de calcio, transfección mediada por DEAE-dextrano, transvección, microinyección, transfección mediada por lípido catiónico, electroporación, transducción, carga de raspaduras, introducción balística o infección (ver Sambrook *et al.*, 1989 [supra]; Ausubel *et al.*, 1991 [supra]; Spector, Goldman & Leinwald, 1998). En células eucariotas, los sistemas de expresión pueden ser transitorios (por ejemplo, episomáticos) o permanentes (integración cromosómica) de acuerdo con las necesidades del sistema.

La molécula de ácido nucleico codificante puede o no incluir una secuencia que codifica una secuencia de control, tal como un péptido señal o secuencia líder, como se desee, por ejemplo, para secreción de un polipéptido traducido en el lumen del retículo endoplasmático, en el espacio periplásmico o en el ambiente extracelular. Estas señales pueden ser endógenas al polipéptido o pueden ser señales heterólogas. Las secuencias líderes pueden eliminarse mediante el huésped bacteriano en el procesamiento post-translacional.

Además de secuencias de control, puede ser deseable agregar secuencias reguladoras que permiten la regulación de la expresión del polipéptido relativo al crecimiento de la célula huésped. Son ejemplos de secuencias reguladoras aquellas que provocan que la expresión de un gen aumente o disminuya en respuesta a un estímulo químico o físico, incluso la presencia de un compuesto regulador o a diversas condiciones de temperatura o metabólicas. Las secuencias reguladoras son aquellas regiones no traducidas del vector, tales como potenciadores, promotores y regiones sin traducir 5' y 3'. Estas interactúan con las proteínas de célula huésped para llevar a cabo la transcripción y traducción. Tales secuencias reguladoras pueden variar en su potencia y especificidad. Dependiendo del sistema de vectores y huésped utilizados, puede usarse cualquier cantidad de elementos de transcripción y traducción, incluidos promotores constitutivos e inducibles. Por ejemplo, para clonación en sistemas bacterianos, pueden usarse promotores inducibles tales como el promotor híbrido lacZ del fagómido Bluescript (Stratagene, LaJolla, CA) o plásmido pSport1™ (Gibco BRL) y similares. El promotor de polihedrina de baculovirus puede usarse en células de insectos. Los promotores o potenciadores derivados de los genomas de células vegetales (por ejemplo, genes de proteínas de choque térmico, RUBISCO y almacenamiento) o de virus vegetales (por ejemplo, promotores virales o secuencias líderes) pueden clonarse en el vector. En sistemas de células mamíferas, se prefieren promotores de genes mamíferos o de virus mamíferos. Si es necesario generar una línea celular que contiene múltiples copias de la secuencia, pueden usarse vectores en base a SV40 o EBV con un marcador seleccionable apropiado.

- 5 Un vector de expresión se construye de modo que la secuencia codificante de ácido nucleico particular se localiza en el vector con las secuencias reguladoras apropiadas, siendo la posición y la orientación de la secuencia codificante con respecto a las secuencias reguladoras tales que la secuencia codificante se transcribe bajo el "control" de las secuencias reguladoras, es decir, la polimerasa ARN que se une a la molécula de ADN en las secuencias de control transcribe la secuencia codificante. En algunos casos puede ser necesario modificar la secuencia de modo que se acople a las secuencias de control con la orientación apropiada; es decir, para mantener el marco de lectura.
- 10 Las secuencias de control y otras secuencias reguladoras pueden estar ligadas a la secuencia codificante de ácido nucleico antes de la inserción en un vector. De forma alternativa, la secuencia codificante puede clonarse directamente en un vector de expresión que ya contiene las secuencias de control y un sitio de restricción apropiado.
- 15 Para producción de alto rendimiento a largo plazo de un polipéptido recombinante, se prefiere la expresión estable. Por ejemplo, las líneas celulares que expresan establemente el polipéptido de interés pueden transformarse usando vectores de expresión que pueden contener orígenes de replicación virales y/o elementos de expresión endógenos y un gen marcador seleccionable en el mismo vector o en un vector separado. Después de la introducción del vector, se puede dejar que las células crezcan durante 1-2 días en un medio enriquecido antes de cambiarlas a un medio selectivo. El propósito del marcador selectivo es conferir resistencia a la selección y su presencia permite el crecimiento y la recuperación de células que expresan con éxito las secuencias introducidas. Los clones resistentes de células transformadas establemente pueden proliferarse usando técnicas de cultivo tisular apropiadas para el tipo celular.
- 20 Las líneas celulares mamíferas disponibles como huéspedes para expresión se conocen en la técnica e incluyen muchas líneas celulares inmortalizadas disponibles en la Colección Americana de Cultivos Tipo (ATCC) incluyendo, a modo no taxativo, ovario de hámster chino (CHO), HeLa, riñón de hámster bebé (BHK), riñón de mono (COS), C127, 3T3, BHK, HEK 293, células de melanoma de Bowes y carcinoma hepatocelular humano (por ejemplo Hep G2) y una cantidad de otras líneas celulares.
- 25 En el sistema baculovirus, los materiales para sistemas de expresión de células de baculovirus/insecto están disponibles en el mercado en forma de kit de Invitrogen, San Diego CA (el kit "MaxBac") entre otros. Estas técnicas son generalmente conocidas para los expertos en la técnica y se describen completamente en Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987). Las células huésped particularmente adecuadas para uso en este sistema incluyen células de insectos tales como las células *Drosophila* S2 y *Spodoptera* Sf9.
- 30 Existen muchos cultivos celulares vegetales y sistemas de expresión genética vegetal enteros conocidos en la técnica. Ejemplos de sistemas de expresión genética celular vegetal adecuados incluyen aquellos descritos en US 5.693.506; US 5.659.122; y US 5.608.143. Ejemplos adicionales de expresión genética en cultivos celulares vegetales se han descrito en Zenk, *Phytochemistry* 30, 3861-3863 (1991).
- 35 En particular, pueden utilizarse todas las plantas de las que se pueden aislar y cultivar protoplastos para proporcionar plantas enteras regeneradas, de modo tal que se recuperan plantas enteras que contienen el gen transferido. Prácticamente todas las plantas pueden regenerarse a partir de células o tejidos cultivados, incluidas, a modo no taxativo, todas las especies principales de caña de azúcar, remolacha azucarera, algodón, árboles frutales y otros, legumbres y hortalizas.
- 40 Ejemplos de células huésped bacterianas particularmente preferidas incluyen células de *streptococci*, *staphylococci*, *E. coli*, *Streptomyces* y *Bacillus subtilis*.
- 45 Ejemplos de células huésped particularmente adecuadas para expresión fúngica incluyen células de levadura (por ejemplo, *S. cerevisiae*) y células de *Aspergillus*.
- 50 En la técnica se conocen una cantidad de sistemas de selección que pueden usarse para recuperar líneas celulares transformadas. Los ejemplos incluyen los genes de quinasa timidina del virus del herpes simple (Wigler, M. *et al.* (1977) *Cell* 11:223-32) y de adenina fosforribosiltransferasa (Lowy, I. *et al.* (1980) *Cell* 22:817-23) que pueden emplearse en células tk- o apt±, respectivamente.
- 55 También, la resistencia herbicida o antibiótica, antimetabolito puede usarse como base para la selección; por ejemplo, dihidrofolato reductasa (DHFR) que confiere resistencia a metotrexato (Wigler, M. *et al.* (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); npt, que confiere resistencia a aminoglucósidos neomicina y G-418 (Colbere-Garapin, F. *et al.* (1981) *J. Mol. Biol.* 150:1-14) y als o pat, que confiere resistencia a clorsulfurón y fosfinotricina acetiltransferasa, respectivamente. Se han descrito genes seleccionables adicionales, ejemplos de los cuales serán claros para los expertos en la técnica.
- 60 Aunque la presencia o ausencia de expresión de gen marcador sugiere que el gen de interés también está presente, puede ser necesario confirmar su presencia y expresión. Por ejemplo, si se inserta la secuencia relevante dentro de una secuencia de gen marcador, las células transformadas que contienen las secuencias apropiadas pueden
- 65

identificarse por la ausencia de la función del gen marcador. De forma alternativa, un gen marcador puede colocarse conjuntamente con una secuencia que codifica un polipéptido de la invención bajo el control de un promotor simple. La expresión del gen marcador en respuesta a inducción o selección generalmente indica también la expresión del gen conjunto.

5 De forma alternativa, las células huésped que contienen una secuencia de ácido nucleico que codifica un polipéptido de la invención y que expresan dicho polipéptido pueden identificarse mediante una variedad de procedimientos conocidos por los expertos en la técnica. Estos procedimientos incluyen, a modo no taxativo, hibridaciones ADN-ADN o ADN-ARN y bioanálisis de proteínas, por ejemplo, separación de células activada por fluorescencia (FACS) o técnicas de inmunoensayo (tales como ensayo inmunoabsorbente ligado a enzima [ELISA] y radioinmunoensayo [RIA]), que incluyen tecnologías en base a membrana, solución o chip para la detección y/o cuantificación de ácido nucleico o proteína (ver Hampton, R. *et al.* (1990) *Serological Methods, a Laboratory Manual*, APS Press, St Paul, MN) y Maddox, D.E. *et al.* (1983) *J. Exp. Med.*, 158, 1211-1216).

10 Los expertos en la técnica conocen una amplia variedad de etiquetas y técnicas de conjugación y pueden usarse en diversos ensayos de ácido nucleico y aminoácidos. Los medios para producir hibridación etiquetada o sondas de PCR para detectar secuencias relacionadas con moléculas de ácido nucleico que codifican polipéptidos de la presente invención incluyen oligoetiquetado, translación de mellas, etiquetado del extremo o amplificación por PCR usando un polinucleótido etiquetado. De forma alternativa, las secuencias que codifican el polipéptido de la invención pueden clonarse en un vector para la producción de una sonda ARNm. Tales vectores se conocen en la técnica, están disponibles en el mercado y pueden usarse para sintetizar sondas de ARN in vitro mediante la adición de una polimerasa de ARN apropiada tal como T7, T3 o SP6 y nucleótidos etiquetados. Estos procedimientos pueden llevarse a cabo usando una variedad de kits disponibles en el mercado (Pharmacia & Upjohn, (Kalamazoo, MI); Promega (Madison WI); y U.S. Biochemical Corp., Cleveland, OH)).

15 Las moléculas reporteras o etiquetas adecuadas, que pueden usarse para facilitar la detección, incluyen radionucleidos, enzimas y agentes fluorescentes, quimioluminiscentes o cromogénicos así como también sustratos, cofactores, inhibidores, partículas magnéticas y similares.

20 Las moléculas de ácido nucleico de acuerdo con la presente invención también pueden usarse para crear animales transgénicos, particularmente animales roedores. Esto puede hacerse localmente mediante modificación de células somáticas o mediante terapia de línea germinal para incorporar modificaciones heredables. Tales animales transgénicos pueden ser particularmente útiles en la generación de modelos animales para moléculas de fármaco efectivas como moduladores de los polipéptidos de la presente invención.

25 El polipéptido puede recuperarse y purificarse de los cultivos celulares recombinantes mediante métodos conocidos que incluyen precipitación en sulfato de amonio o etanol, extracción de ácido, cromatografía de intercambio aniónico o catiónico, cromatografía sobre fosfocelulosa, cromatografía de interacción hidrófoba, cromatografía de afinidad, cromatografía de hidroxilapatita y cromatografía de lectina. La cromatografía líquida de alto rendimiento es particularmente útil para purificación. Las técnicas conocidas para replegamiento de proteínas pueden emplearse para regenerar una conformación activa cuando el polipéptido se desnaturaliza durante el aislamiento y/o purificación.

30 Las construcciones de vector especializado también pueden usarse para facilitar la purificación de proteínas, como se desee, uniendo secuencias que codifican los polipéptidos de la invención con una secuencia de nucleótido que codifica un dominio polipéptido que facilitará la purificación de proteínas solubles. Ejemplos de tales dominios que facilitan la purificación incluyen péptidos quelantes de metal tales como módulos de histidina-triptófano que permiten la purificación sobre metales inmovilizados, dominios de proteína A que permiten la purificación sobre inmunoglobulina inmovilizada y el dominio utilizado en el sistema de purificación de extensión/afinidad FLAGS (Immunex Corp., Seattle, WA). La inclusión de secuencias ligadoras escindibles tales como aquellas específicas para el Factor XA o enteroquinasa (Invitrogen, San Diego, CA) entre el dominio de purificación y el polipéptido de la invención puede usarse para facilitar la purificación. Un vector de expresión proporciona la expresión de una proteína de fusión que contiene el polipéptido de la invención fusionado con varios residuos de histidina que preceden un sitio de escisión de tiorredoxina o enteroquinasa. Los residuos de histidina facilitan la purificación mediante IMAC (cromatografía de afinidad con ión metálico inmovilizado como se describe en Porath, J. *et al.* (1992), *Prot. Exp. Purif.* 3: 263-281) mientras que el sitio de escisión de tiorredoxina o enteroquinasa proporcionan un medio para purificar el polipéptido a partir de la proteína de fusión. Una descripción de vectores que contienen proteínas de fusión se proporciona en Kroll, D.J. *et al.* (1993; *DNA Cell Biol.* 12:441-453).

35 Si el polipéptido debe expresarse para uso en ensayos de detección sistemática, generalmente se prefiere que se produzca en la superficie de la célula huésped en la que se expresa. En este caso, las células huésped pueden cosecharse antes del uso en el ensayo de detección sistemática, por ejemplo usando técnicas tales como separación de células activada por fluorescencia (FACS) o técnicas de inmunoafinidad. Si el polipéptido se secreta en el medio, el medio puede recuperarse para recuperar y purificar el polipéptido expresado. Si el polipéptido se produce intracelularmente, las células deben lisarse primero antes de recuperar el polipéptido.

- De forma alternativa, puede preferirse que los polipéptidos de la invención se expresen como proteínas de fusión de superficie celular. En este caso, las células huésped pueden cosecharse antes del uso en el ensayo de detección, por ejemplo usando técnicas tales como separación de células activada por fluorescencia (FACs) o técnicas de inmunoafinidad.
- 5 Tal como se indicó anteriormente, la presente invención también proporciona nuevos objetivos y métodos para la detección sistemática de fármacos candidatos o precursores. Estos métodos de detección sistemática incluyen ensayos de unión y/o ensayos funcionales, y pueden llevarse a cabo in vitro, en sistemas celulares o en animales.
- 10 En este sentido, un objeto particular de esta invención reside en el uso de un polipéptido INSP215 como objetivo para detectar sistemáticamente fármacos candidatos para tratar o prevenir trastornos en los que participan proteínas que contienen el dominio ZP, tales como el receptor III de TGF beta.
- 15 También se describen en la presente métodos para seleccionar compuestos biológicamente activos, comprendiendo tales métodos poner en contacto un compuesto candidato con un gen o polipéptido INSP215 y seleccionar compuestos que se unen a dicho gen o polipéptido.
- 20 También se describen en la presente métodos para seleccionar compuestos biológicamente activos, comprendiendo tales métodos poner en contacto un compuesto candidato con una célula huésped recombinante que expresa un polipéptido INSP215 con un compuesto candidato y seleccionar compuestos que se unen a dicho polipéptido INSP215 en la superficie de dichas células y/o que modulan la actividad del polipéptido INSP215.
- 25 Un compuesto "biológicamente activo" denota todo compuesto que tiene actividad biológica en un sujeto, preferiblemente actividad terapéutica, más preferiblemente un compuesto que puede usarse para tratar trastornos en los que participan proteínas que contienen el dominio ZP, tales como un receptor III de TGF beta, o como un precursor para desarrollar fármacos para tratar trastornos en los que participan proteínas que contienen el dominio ZP, tales como receptor III de TGF beta. Un compuesto "biológicamente activo" preferiblemente es un compuesto que modula la actividad de un polipéptido INSP215.
- 30 Los métodos anteriores pueden llevarse a cabo in vitro, usando diversos dispositivos y condiciones, incluso con reactivos inmovilizados, y pueden comprender además un paso adicional de ensayo de la actividad de los compuestos seleccionados en un modelo de trastorno en el que participan proteínas que contienen el dominio ZP, tales como el receptor III de TGF beta, tal como un modelo animal.
- 35 La unión a un gen o polipéptido objetivo proporciona un indicio de la capacidad del compuesto para modular la actividad de dicho objetivo, y por consiguiente, afectar una vía que conduce a un trastorno en un sujeto. La determinación de la unión puede llevarse a cabo mediante diversas técnicas, tales como mediante etiquetado del compuesto candidato, mediante competencia con un ligando de referencia etiquetado, etc.
- 40 Para ensayos de unión in vitro, los polipéptidos pueden usarse en forma esencialmente pura, en suspensión, inmovilizados sobre un soporte o expresados en una membrana (célula intacta, preparación de membrana, liposoma, etc.).
- 45 La modulación de la actividad incluye, a modo no taxativo, estimulación de la expresión superficial de un receptor y modulación de multimerización de dicho receptor (por ejemplo, la formación de complejos multiméricos con otras subunidades), etc. Las células usadas en los ensayos pueden ser cualquier célula recombinante (es decir, cualquier célula que comprende un ácido nucleico recombinante que codifica un polipéptido INSP215) o cualquier célula que expresa un polipéptido INSP215 endógeno. Ejemplos de tales células incluyen, a modo no taxativo, células procariotas (tales como bacterias) y células eucariotas (tales como células de levadura, células mamíferas, células de insectos, células vegetales, etc.). Los ejemplos específicos incluyen levaduras de *E. coli*, *Pichia pastoris*, *Hansenula polymorpha*, *Schizosaccharomyces pombe*, *Kluyveromyces* o *Saccharomyces*, líneas celulares mamíferas (por ejemplo, células Vero, células CHO, células 3T3, células COS, etc.) así como también cultivos celulares mamíferos primarios o establecidos (por ejemplo, producidos a partir de fibroblastos, células embrionarias, células epiteliales, células nerviosas, adipocitos, etc.).
- 50
- 55 Los compuestos seleccionados preferidos son agonistas de INSP215, es decir, compuestos que pueden unirse a INSP215 e imitar la actividad de un ligando endógeno del mismo.
- 60 También se describe en la presente un método para seleccionar compuestos biológicamente activos, comprendiendo dicho método poner en contacto in vitro un compuesto de prueba con un polipéptido INSP215 de acuerdo con la presente invención y determinar la capacidad de dicho compuesto de prueba para modular la actividad de dicho polipéptido INSP215.
- 65 También se describe en la presente un método para seleccionar compuestos biológicamente activos, comprendiendo dicho método poner en contacto in vitro un compuesto de prueba con un gen INSP215 de acuerdo con la presente invención y determinar la capacidad de dicho compuesto de prueba para modular la actividad de dicho gen INSP215, preferiblemente para estimular la expresión del mismo.

5 También se describe en la presente un método para detectar sistemáticamente, seleccionar o identificar compuestos activos, comprendiendo el método poner en contacto un compuesto de prueba con una célula huésped recombinante que comprende un constructo reportero, comprendiendo dicho constructo reportero un gen reportero bajo el control de un promotor de gen INSP215 y seleccionar los compuestos de prueba que modulan (por ejemplo, estimulan o reducen, preferiblemente estimulan) la expresión del gen reportero.

10 El polipéptido de la invención puede usarse para detectar sistemáticamente bibliotecas de compuestos en cualquiera de una variedad de técnicas de detección sistemática de fármacos. Tales compuestos pueden activar (agonizar) o inhibir (antagonizar) el nivel de expresión del gen o la actividad del polipéptido de la invención. Los compuestos preferidos son efectivos para alterar la expresión de un gen natural que codifica un polipéptido del primer aspecto de la invención o para regular la actividad de un polipéptido del primer aspecto de la invención.

15 Los compuestos agonistas o antagonistas puede aislarse de, por ejemplo, células, preparaciones libres de células, bibliotecas químicas o mezclas de producto natural. Estos agonistas o antagonistas pueden ser sustratos, ligandos, enzimas, receptores o miméticos estructurales o funcionales naturales o modificados. Para una evaluación adecuada de tales técnicas de detección sistemática, ver Coligan *et al.*, Current Protocols in Immunology I (2):Capítulo 5 (1991).

20 Los compuestos que más probablemente serán buenos antagonistas son moléculas que se unen al polipéptido de la invención sin inducir los efectos biológicos del polipéptido luego de unirse al él. Los antagonistas potenciales incluyen moléculas orgánicas pequeñas, péptidos, polipéptidos y anticuerpos que se unen al polipéptido de la invención y de este modo inhiben o extinguen su actividad. De este modo, la unión del polipéptido a moléculas de unión celular normales pueden inhibirse de modo que se evita la actividad biológica normal del polipéptido.

25 El polipéptido de la invención que se emplea en tal técnica de detección sistemática puede estar libre en solución, fijado a un soporte sólido, suspendido en una superficie celular o localizado intracelularmente. En general, tales procedimientos de detección sistemática pueden implicar el uso de células o membranas celulares apropiadas que expresan el polipéptido que se ponen en contacto con un compuesto de prueba para observar la unión, o estimulación o inhibición de una respuesta funcional. La respuesta funcional de las células que se pusieron en contacto con el compuesto de prueba luego se compara con células testigo que no se pusieron en contacto con el compuesto de prueba. Tal ensayo puede evaluar si el compuesto de prueba resulta en una señal generada mediante activación del polipéptido usando un sistema de detección sistemática apropiado. Los inhibidores de activación generalmente se ensayan en presencia de un agonista conocido y se observa el efecto en la activación mediante el agonista en presencia del compuesto de prueba.

40 Los métodos para generar señales detectables en los tipos de ensayos descritos en la presente serán conocidos para los expertos en la técnica. Un ejemplo particular es cotransfectar un constructo que expresa un polipéptido de acuerdo con la invención o un fragmento tal como el LBD, en fusión con el dominio de unión a ADN de GAL4, en una célula junto con un plásmido reportero, como por ejemplo pFR-Luc (Stratagene Europe, Amsterdam, Países Bajos). Este plásmido particular contiene un promotor sintético con cinco repeticiones conjuntas de sitios de unión a GAL4 que controlan la expresión del gen luciferasa. Cuando un ligando potencial se agrega a las células, unirá la fusión GAL4-polipéptido e inducirá la transcripción del gen luciferasa. El nivel de expresión de luciferasa puede monitorearse por su actividad usando un lector de luminiscencia (ver, por ejemplo, Lehman *et al.* JBC 270, 12953, 1995; Pawar *et al.* JBC, 277,39243, 2002).

Un método preferido para identificar un compuesto agonista o antagonista de un polipéptido de la presente invención comprende:

- 50 (a) poner en contacto una célula que expresa en la superficie de la misma el polipéptido de acuerdo con el primer aspecto de la invención, estando asociado el polipéptido con un segundo componente capaz de proporcionar una señal detectable en respuesta a la unión de un compuesto con el polipéptido, con un compuesto que se detectará sistemáticamente en condiciones para permitir la unión con el polipéptido; y
- 55 (b) determinar si el compuesto se une y activa o inhibe el polipéptido midiendo el nivel de una señal generada a partir de la interacción del compuesto con el polipéptido.

Un método preferido adicionalmente para identificar un agonista o antagonista de un polipéptido de la presente invención comprende:

- 60 (a) poner en contacto una célula que expresa en la superficie de la misma el polipéptido, estando asociado el polipéptido con un segundo componente capaz de proporcionar una señal detectable en respuesta a la unión de un compuesto con el polipéptido, con un compuesto que se detectará sistemáticamente en condiciones para permitir la unión con el polipéptido; y
- 65 (b) determinar si el compuesto se une y activa o inhibe el polipéptido comparando el nivel de una señal generada a partir de la interacción del compuesto con el polipéptido, con el nivel de una señal en ausencia del compuesto.

Por ejemplo, podría usarse un método tal como detección FRET de una unión de un ligando unido al polipéptido en presencia de coactivadores de péptido (Norris *et al.*, Science 285, 744,1999).

Los métodos generales que se describieron anteriormente pueden comprender además llevar a cabo la identificación del agonista o antagonista en presencia de un ligando etiquetado o no etiquetado para el polipéptido.

5 En otra realización del método para identificar un agonista o antagonista de un polipéptido de la presente invención comprende:

determinar la inhibición de la unión de un ligando con células que tienen un polipéptido en la superficie de las mismas, o con membranas celulares que contienen tal polipéptido, en presencia de un compuesto candidato en condiciones para permitir la unión con el polipéptido, y determinar la cantidad de unión de ligando con el polipéptido. Se considera que un compuesto capaz de provocar reducción de unión de un ligando es un agonista o antagonista. 10 Preferiblemente el ligando está etiquetado.

Más particularmente, un método para la detección sistemática de un compuesto antagonista o agonista de polipéptido comprende los pasos de:

15 (a) incubar un ligando etiquetado con una célula entera que expresa un polipéptido de acuerdo con la invención en la superficie celular, o una membrana celular que contiene un polipéptido de la invención,
 (b) medir la cantidad de ligando etiquetado unido a la célula entera o con la membrana celular;
 (c) agregar un compuesto candidato a la mezcla del ligando etiquetado y la célula entera o membrana celular del paso (a) y permitir que la mezcla alcance el equilibrio;
 20 (d) medir la cantidad de ligando etiquetado unido con la célula entera o con la membrana celular después del paso (c); y
 (e) comparar la diferencia del ligando etiquetado unido en el paso (b) y (d), de modo tal que el compuesto que provoca la reducción en la unión en el paso (d) se considera que es un agonista o antagonista.

25 Pueden usarse ensayos de unión simple, en los que la adherencia de un compuesto de prueba a una superficie que soporta el polipéptido se detecta mediante una etiqueta asociada directamente o indirectamente con el compuesto de prueba o en un ensayo que implica la competencia con un competidor etiquetado. En otra realización, pueden usarse ensayos de detección sistemática de fármaco competitivo, en los que compiten anticuerpos neutralizantes que son capaces de unirse al polipéptido específicamente con un compuesto de prueba para unión. De este modo, los anticuerpos pueden usarse para detectar la presencia de cualquier compuesto de prueba que posee afinidad de 30 unión específica para el polipéptido.

Los ensayos de detección sistemática conocidos en la técnica que son adecuados para la identificación de moduladores de los polipéptidos de esta invención incluyen el ensayo GAL4 (por ejemplo, Jurgen M. Lehmann *et al.*, J. Biol. Chem., Jun 1995; 270: 12953- 12956; Raivio T. *et al.*, J. Clin. Endocrinol. Metab. 86 (2001) 1539- 44), técnicas biofísicas tales como resonancia de plasmones superficiales (comercializada por Biacore AB, Uppsala, Suecia) y ensayos en base a polarización en fluorescencia o FRET para unión inducida por ligando del péptido coactivador (Mukherjee R. *et al.*, J Steroid Biochem. Mol. Biol. 2002 Jul; 81(3):217-25).

40 También pueden diseñarse ensayos para detectar el efecto de compuestos de prueba agregados en la producción de ARNm que codifica el polipéptido en las células. Por ejemplo, puede construirse un ELISA que mide los niveles secretados o asociados con la célula de polipéptido usando anticuerpos monoclonales o policlonales mediante métodos estándares conocidos en la técnica, y este puede usarse para buscar compuestos que pueden inhibir o potenciar la producción del polipéptido a partir de células o tejidos manipulados adecuadamente. La formación de complejos de unión entre el polipéptido y el compuesto que se está probando puede medirse luego. 45

Otra técnica para detección sistemática de fármaco que puede usarse proporciona detección sistemática de alto rendimiento de compuestos que tienen afinidad de unión adecuada con el polipéptido de interés (ver Solicitud de patente internacional WO84/03564). En este método, se sintetizan grandes cantidades de diferentes compuestos de prueba pequeños sobre un sustrato sólido, que luego pueden hacerse reaccionar con el polipéptido de la invención y lavarse. Un modo de inmovilizar el polipéptido es usar anticuerpos no neutralizantes. El polipéptido unido luego puede detectarse usando métodos que son conocidos en la técnica. El polipéptido purificado también puede revestirse directamente sobre placas para uso en las técnicas de detección sistemática de fármaco mencionadas anteriormente. 50

55 Los métodos de ensayo incluyen aquellos que implican el uso de los genes y polipéptidos de la invención en ensayos de sobreexpresión o ablación. Tales ensayos implican la manipulación de niveles de estos genes/polipéptidos en las células y la evaluación del impacto de este evento de manipulación sobre la fisiología de las células manipuladas. Por ejemplo, tales experimentos revelan detalles de las vías de señalización y metabólicas en las que participan los genes/polipéptidos particulares, generan información con respecto a las identidades de polipéptidos con los que los polipéptidos estudiados interactúan y proporcionan pistas sobre métodos con los que se regulan genes y proteínas relacionados. 60

Otra técnica para detección sistemática de fármaco que puede usarse proporciona detección sistemática de alto rendimiento de compuestos que tienen afinidad de unión adecuada con el polipéptido de interés (ver Solicitud de patente internacional WO84/03564). En este método, se sintetizan grandes cantidades de diferentes compuestos de prueba pequeños sobre un sustrato sólido, que luego pueden hacerse reaccionar con el polipéptido de la invención y 65

lavarse. Un modo de inmovilizar el polipéptido es usar anticuerpos no neutralizantes. El polipéptido unido luego puede detectarse usando métodos que son conocidos en la técnica. El polipéptido purificado también puede revestirse directamente sobre placas para uso en las técnicas de detección sistemática de fármaco mencionadas anteriormente.

5

El polipéptido de la invención puede usarse para identificar receptores unidos a membrana o solubles, a través de técnicas de unión a receptor estándares que se conocen en la técnica, tales como ensayos de unión a ligando y reticulación en los que el polipéptido se etiqueta con un isótopo radioactivo, se modifica químicamente, o se fusiona a una secuencia de péptidos que facilita su detección sistemática o purificación y se incubaba con una fuente del receptor putativo (por ejemplo, una composición de células, membranas celulares, sobrenadantes celulares, extractos tisulares o fluidos corporales). La eficacia de la unión puede medirse usando técnicas biofísicas tales como resonancia de plasmones superficiales (comercializada por Biacore AB, Uppsala, Suecia) y espectroscopía. Los ensayos de unión pueden usarse para la purificación y clonación del receptor, pero también pueden identificar agonistas y antagonistas del polipéptido, que compiten con la unión del polipéptido con su receptor. Los métodos estándares para llevar a cabo ensayos de detección sistemática se conocen en la técnica.

10

15

También se describe en la presente el uso de un polipéptido INSP215 o fragmento del mismo, donde el fragmento es preferiblemente un fragmento específico del gen INSP215, para aislar o generar un agonista o estimulador del polipéptido INSP215 para el tratamiento de un trastorno, en donde dicho agonista o estimulador se selecciona del grupo que consiste en:

20

1. un anticuerpo específico o fragmento del mismo que incluye: un anticuerpo a) quimérico, b) humanizado o c) completamente humano, así como también;
2. un anticuerpo biespecífico o multiespecífico,
- 25 3. una cadena simple (por ejemplo, scFv) o
4. anticuerpo de dominio simple, o
5. un mimético peptídico o no peptídico derivado de dichos anticuerpos o
6. un mimético de anticuerpo tal como a) una anticalina o b) una molécula de unión en base a fibronectina (por ejemplo, trinectina o adnectina).

25

La generación de miméticos peptídicos o no peptídicos a partir de anticuerpos se conoce en la técnica (Saragovi *et al.*, 1991 y Saragovi *et al.*, 1992).

30

Las anticalinas también se conocen en la técnica (Vogt *et al.*, 2004). Las moléculas de unión en base a fibronectina se describen en US6818418 y WO2004029224.

35

Adicionalmente, el compuesto de prueba puede ser de origen, naturaleza y composición variadas, tales como cualquier molécula pequeña, ácido nucleico, lípido, péptido, polipéptido, incluso un anticuerpo tal como un anticuerpo quimérico, humanizado o completamente humano o un fragmento de anticuerpo, mimético peptídico o no peptídico derivado de los mismos así como también un anticuerpo biespecífico o multiespecífico, un anticuerpo de cadena simple (por ejemplo, scFv) o un anticuerpo de dominio simple o un mimético de anticuerpo tal como una anticalina o molécula de unión en base a fibronectina (por ejemplo, trinectina o adnectina), etc., en forma aislada o en mezcla o combinaciones.

40

También se describe en la presente un kit de detección sistemática útil en los métodos para identificar agonistas, antagonistas, ligandos, receptores, sustratos y enzimas que se describieron anteriormente.

45

La divulgación incluye los agonistas, antagonistas, ligandos, receptores, sustratos y enzimas y otros compuestos que modulan la actividad o antigenicidad del polipéptido de la invención descubiertos mediante los métodos descritos anteriormente.

50

Tal como se menciona anteriormente, se prevé que los diversos restos de la invención (es decir, los polipéptidos del primer aspecto de la invención, una molécula de ácido nucleico del segundo o tercer aspecto de la invención, un vector del cuarto aspecto de la invención, una célula huésped del quinto aspecto de la invención, un ligando del sexto aspecto de la invención, un compuesto del séptimo aspecto de la divulgación) pueden ser útiles en la terapia o diagnóstico de enfermedades. Para evaluar la utilidad de los restos de la invención para tratar o diagnosticar una enfermedad pueden llevarse a cabo uno o más de los ensayos descritos en la presente. Debe observarse que aunque algunos de los ensayos se refieren al compuesto de prueba como una proteína/polipéptido, un experto en la técnica podrá adaptar fácilmente los siguientes ensayos de modo que otros restos de la invención también puedan usarse como el "compuesto de prueba".

55

60

También se describen en la presente composiciones farmacéuticas que comprenden un polipéptido, ácido nucleico, ligando o compuesto de la divulgación en combinación con un portador farmacéutico adecuado. Estas composiciones pueden ser adecuadas como reactivos terapéuticos o de diagnóstico, como vacunas, o como otras composiciones inmunogénicas, como se describen en detalle a continuación.

65

De acuerdo con la terminología usada en la presente, una composición que contiene un polipéptido, ácido nucleico, ligando o compuesto [X] está "sustancialmente libre de impurezas [en la presente, Y] cuando al menos 85% en peso del total X+Y en la composición es X. Preferiblemente, X comprende al menos aproximadamente 90% en peso del total de X+Y en la composición, más preferiblemente al menos aproximadamente 95%, 98% o incluso 99% en peso.

Las composiciones farmacéuticas deben comprender preferiblemente una cantidad terapéuticamente efectiva del polipéptido, molécula de ácido nucleico, ligando o compuesto de la invención. La expresión "cantidad terapéuticamente efectiva" tal como se usa en la presente se refiere a una cantidad de un agente terapéutico necesario para tratar, aliviar o prevenir una enfermedad o afección objetivo, o para exhibir un efecto terapéutico o preventivo detectable. Para cualquier compuesto, la dosis terapéuticamente efectiva puede estimarse inicialmente en ensayos de cultivo celular, por ejemplo, de células neoplásicas, o en modelos animales, generalmente ratones, conejos, perros o cerdos. El modelo animal también puede usarse para determinar el rango de concentración y vía de administración adecuados. Tal información luego puede usarse para determinar dosis y vías para administración útiles en humanos.

La cantidad efectiva precisa para un sujeto humano dependerá de la severidad del estado de enfermedad, salud general del sujeto, edad, peso y sexo del sujeto, dieta, tiempo y frecuencia de administración, combinaciones de fármacos), sensibilidades a la reacción y tolerancia/respuesta a la terapia. Esta cantidad puede determinarse mediante experimentación habitual y pertenece al criterio del médico. Generalmente, una dosis efectiva será de 0,01 mg/kg a 50 mg/kg, preferiblemente 0,05 mg/kg a 10 mg/kg. Las composiciones pueden administrarse individualmente a un paciente o pueden administrarse en combinación con otros agentes, fármacos u hormonas.

Una composición farmacéutica también puede contener un portador farmacéuticamente aceptable, para administración de un agente terapéutico. Tales portadores incluyen anticuerpos y otros polipéptidos, genes y otros agentes terapéuticos tales como liposomas, siempre que el portador no induzca por sí mismo la producción de anticuerpos dañinos para el individuo que recibe la composición, y que pueda administrarse sin toxicidad indebida. Los portadores adecuados pueden ser macromoléculas grandes metabolizadas lentamente tales como proteínas, polisacáridos, ácidos polilácticos, ácidos poliglicólicos, aminoácidos poliméricos, copolímeros de aminoácidos y partículas de virus inactivo.

Las sales farmacéuticamente aceptables pueden usarse en la misma, por ejemplo, sales de ácidos minerales tales como clorhidratos, bromhidratos, fosfatos, sulfatos y similares; y las sales de ácidos orgánicos tales como acetatos, propionatos, malonatos, benzoatos y similares. Una amplia descripción de portadores farmacéuticamente aceptables está disponible en Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991).

Los portadores farmacéuticamente aceptables en composiciones terapéuticas pueden contener adicionalmente líquidos tales como agua, solución salina, glicerol y etanol. Adicionalmente, en tales composiciones pueden estar presentes sustancias auxiliares, tales como agentes humectantes o emulsionantes, sustancias amortiguadoras de pH y similares. Tales portadores permiten que las composiciones farmacéuticas se formulen como comprimidos, pastillas, grageas, cápsulas, líquidos, geles, jarabes, soluciones, suspensiones y similares, para ingesta del paciente.

Una vez formuladas, las composiciones descritas en la presente, pueden administrarse directamente al sujeto. Los sujetos a tratar pueden ser animales; en particular, pueden tratarse sujetos humanos.

Las composiciones farmacéuticas descritas en la presente pueden administrarse mediante cualquier cantidad de vías incluso, a modo no taxativo, aplicaciones orales, intravenosas, intramusculares, intraarteriales, intramedulares, intratecales, intraventriculares, transdérmicas o transcutáneas (por ejemplo, ver WO98/20734), medios subcutáneos, intraperitoneales, intranasales, entéricos, tópicos, sublinguales, intravaginales o rectales. También pueden usarse cañones o hipopulverizadores génicos para administrar las composiciones farmacéuticas de la invención. Típicamente, las composiciones terapéuticas pueden prepararse como inyectables, como soluciones líquidas o suspensiones; también pueden prepararse formas sólidas adecuadas para solución o suspensión en vehículos líquidos antes de la inyección.

La administración directa de las composiciones generalmente se alcanzará mediante inyección, subcutáneamente, intraperitonealmente, intravenosamente o intramuscularmente, o se administrará al espacio intersticial de un tejido. Las composiciones también pueden administrarse a una lesión. El tratamiento de dosificación puede ser un esquema de dosis simple o un esquema de dosis múltiples.

Si la actividad de un polipéptido de la invención es excesiva en un estado de enfermedad particular, existen diversos abordajes disponibles. Un abordaje comprende administrar a un sujeto un compuesto inhibidor (antagonista) como se describió anteriormente, junto con un portador farmacéuticamente aceptable en una cantidad efectiva para inhibir la función del polipéptido, tal como bloquear la unión de los ligandos, sustratos, enzimas, receptores, o inhibir una segunda señal, y aliviar de este modo la afección anormal. Preferiblemente, tales antagonistas son anticuerpos. Más preferiblemente, tales anticuerpos son quiméricos y/o humanizados para minimizar su inmunogenicidad, como se describió anteriormente.

En otro abordaje, pueden administrarse formas solubles de los polipéptidos que retienen la afinidad de unión para el ligando, sustrato, enzima, receptor, en cuestión. Típicamente, el polipéptido puede administrarse en forma de fragmentos que retienen las porciones relevantes.

5 En un abordaje alternativo, la expresión del gen que codifica el polipéptido puede inhibirse usando técnicas de bloqueo de expresión, tales como el uso de moléculas de ácido nucleico antisentido (como se describieron anteriormente), generadas internamente o administradas separadamente. Las modificaciones de la expresión génica pueden obtenerse diseñando secuencias complementarias o moléculas antisentido (ADN, ARN o APN) a las regiones de control, 5' o reguladoras (secuencia señal, promotores, potenciadores e intrones) del gen que codifica el polipéptido. De forma similar, la inhibición puede lograrse usando metodología de apareamiento de bases "triple hélice". El apareamiento de triple hélice es útil porque provoca la inhibición de la capacidad de la doble hélice para abrirse suficientemente para la unión de polimerasas, factores de transcripción o moléculas reguladoras. Los avances terapéuticos recientes usando ADN triple se han descrito en la bibliografía (Gee, J.E. *et al.* (1994) En: Huber, B.E. and B.I. Carr, *Molecular and Immunologic Approaches*, Futura Publishing Co., Mt. Kisco, NY). La secuencia complementaria o la molécula antisentido también puede diseñarse para bloquear la traducción de ARNm evitando la transcripción de unión a ribosomas. Tales oligonucleótidos pueden administrarse o pueden generarse in situ a partir de expresión in vivo.

20 Adicionalmente, la expresión del polipéptido de la invención puede evitarse usando ribozimas específicas para su secuencia codificante de ARNm. Las ribozimas son ARNs activos catalíticamente que pueden ser naturales o sintéticas (ver, por ejemplo, Usman, N, *et al.*, *Curr. Opin. Struct. Biol* (1996) 6(4), 527-33). Las ribozimas sintéticas pueden diseñarse para escindir específicamente ARNms en posiciones seleccionadas evitando de este modo la traducción de los ARNms en un polipéptido funcional. Las ribozimas pueden sintetizarse con una estructura principal de ribosa-fosfato natural y bases naturales, como se encuentran normalmente en moléculas de ARN. De forma alternativa, las ribozimas pueden sintetizarse con estructuras principales no naturales, por ejemplo, 2'-O-metil ARN, para proporcionar protección de la degradación de ribonucleasa y pueden contener bases modificadas.

30 Las moléculas de ARN pueden modificarse para aumentar la estabilidad y semivida intracelular. Las posibles modificaciones incluyen, a modo no taxativo, la adición de secuencias flanqueadoras en los extremos 5' y/o 3' de la molécula o el uso de fosforotioato o 2' O-metil en lugar de enlaces de fosfodiesterasa dentro de la estructura principal de la molécula. Este concepto es inherente en la producción de APNs y puede extenderse a todas estas moléculas mediante la inclusión de bases no tradicionales tales como inosina, queosina y butosina, así como también formas modificadas con acetil, metil, tio y similares de adenina, citidina, guanina, timina y uridina que no son tan fácilmente reconocidas por las endonucleasas endógenas.

40 Para tratar afecciones anormales relacionadas con una expresión baja del polipéptido de la invención y su actividad, también existen diversos abordajes. Un abordaje comprende administrar a un sujeto una cantidad terapéuticamente efectiva de un compuesto que activa el polipéptido, es decir, un agonista como se describió anteriormente, para aliviar la afección anormal. De forma alternativa, puede administrarse una cantidad terapéutica del polipéptido en combinación con un portador farmacéutico adecuado para restaurar el equilibrio fisiológico relevante del polipéptido. La terapia génica puede emplearse para lograr la producción endógena del polipéptido mediante las células relevantes en el sujeto. La terapia génica se usa para tratar permanentemente la producción inapropiada del polipéptido mediante el reemplazo de un gen defectuoso por un gen terapéutico corregido.

45 La terapia génica descrita en la presente puede ocurrir in vivo o ex vivo. La terapia génica ex vivo requiere el aislamiento y purificación de las células del paciente, la introducción de un gen terapéutico y la introducción de las células alteradas genéticamente nuevamente al paciente. En cambio, la terapia génica in vivo no requiere el aislamiento y purificación de células de un paciente.

50 El gen terapéutico típicamente se "envasa" para administración a un paciente. Los vehículos de administración génica pueden ser no virales, tales como liposomas, o virus de replicación deficiente, tales como adenovirus como describe Berkner, K.L., in *Curr. Top. Microbiol. Immunol.*, 158, 39-66 (1992) o vectores de virus asociado a adeno (AAV) como describe Muzyczka, N., in *Curr. Top. Microbiol. Immunol.*, 158, 97-129 (1992) y la Patente de los Estados Unidos No. 5.252.479. Por ejemplo, una molécula de ácido nucleico que codifica un polipéptido de la invención puede manipularse para expresión en un vector retroviral de replicación defectuosa. Este constructo de expresión puede aislarse luego e introducirse dentro de una célula de envasado transducida con un vector plásmido retroviral que contiene ARN que codifica el polipéptido, de modo que la célula de envasado produzcan partículas virales infecciosas que contienen el gen de interés. Estas células productoras pueden administrarse a un sujeto para manipular células in vivo y para expresión del polipéptido in vivo (ver Capítulo 20, *Gene Therapy and other Molecular Genetic-based Therapeutic Approaches*, (y referencias citadas en las mismas) en *Human Molecular Genetics* (1996), T Strachan and A P Read, BIOS Scientific Publishers Ltd).

60 Otro abordaje es la administración de "ADN desnudo" en el que el gen terapéutico se inyecta directamente al flujo sanguíneo o tejido muscular.

En situaciones en las que los polipéptidos o moléculas de ácido nucleico de la invención son agentes que provocan enfermedades, la invención proporciona su uso en vacunas para aumentar los anticuerpos contra el agente que provoca la enfermedad.

5 Las vacunas descritas en la presente pueden ser profilácticas (es decir, para evitar una infección) o terapéuticas (es decir, para tratar la enfermedad después de la infección). Tales vacunas comprenden antígeno(s) inmunizante(s), inmunógeno(s), polipéptido(s), proteína(s) o ácido nucleico, generalmente en combinación con portadores farmacéuticamente aceptables como se describieron anteriormente, que incluyen cualquier portador que no induce por sí mismo la producción de anticuerpos dañinos para el individuo que recibe la composición. Adicionalmente, estos portadores pueden funcionar como agentes inmunoestimulantes ("adyuvantes"). Adicionalmente, el antígeno o inmunógeno puede conjugarse con un toxoide bacteriano, tal como un toxoide de difteria, tétano, cólera, H. pilori y otros patógenos.

10 Ya que los polipéptidos pueden desintegrarse en el estómago, las vacunas que comprenden polipéptidos preferiblemente se administran parenteralmente (por ejemplo, por inyección subcutánea, intramuscular, intravenosa o intradérmica). Las formulaciones adecuadas para administración parenteral incluyen soluciones para inyección estériles acuosas y no acuosas que pueden contener antioxidantes, soluciones amortiguadoras, bacteriostatos y solutos que hacen que la formulación sea isotónica con la sangre del recipiente, y suspensiones estériles acuosas y no acuosas que pueden incluir agentes de suspensión o agentes espesantes.

15 Las formulaciones de vacunas descritas en la presente pueden presentarse en recipientes de dosis unitaria o dosis múltiple. Por ejemplo, ampollas y viales sellados, y pueden almacenarse en condiciones de secado por congelamiento requiriendo solo la adición del portador líquido estéril inmediatamente antes del uso. La dosificación dependerá de la actividad específica de la vacuna y puede determinarse fácilmente mediante experimentación habitual.

20 La administración genética de anticuerpos que se unen a polipéptidos de acuerdo con la invención también puede llevarse a cabo, por ejemplo, como se describe en la Solicitud de patente internacional WO98/55607.

25 La tecnología denominada inyección a presión (ver, por ejemplo, www.powderject.com) también puede ser útil en la formulación de composiciones para vacuna.

30 Se describe una cantidad de métodos adecuados para vacunación y sistemas de administración de vacunas en la Solicitud de patente internacional WO00/29428.

35 Esta invención también se refiere al uso de moléculas de ácido nucleico de acuerdo con la presente invención como reactivos de diagnóstico. La detección sistemática de una forma mutada del gen caracterizado por las moléculas de ácido nucleico de la invención que se asocia con una disfunción proporcionará una herramienta de diagnóstico que puede aportar o definir un diagnóstico de una enfermedad, o la susceptibilidad a una enfermedad, que resulta de la baja expresión, sobreexpresión o expresión espacial o temporal alterada del gen. Los individuos que portan mutaciones en el gen pueden detectarse a nivel del ADN mediante una variedad de técnicas.

40 La sobre-expresión de ciertos polipéptidos de la presente invención, en particular de formas unidas a membrana, puede correlacionarse con evolución o pronóstico de enfermedad. Como tal, los polipéptidos de la presente invención, preferiblemente las formas unidas a membranas, son útiles como marcadores para evolución y/o pronóstico de enfermedad.

45 Se cree que ciertos tejidos en mamíferos con cáncer o enfermedad inflamatoria contienen una cantidad considerablemente mayor de copia génica de proteína INSP215 y expresan niveles considerablemente mejorados de la proteína INSP215 y ARNm que codifica la proteína INSP215 cuando se compara con un mamífero "estándar" correspondiente, es decir, un mamífero de la misma especie que no tiene cáncer ni enfermedad inflamatoria. Los niveles mejorados de la proteína INSP215 se detectarán en ciertos fluidos corporales (por ejemplo, suero, plasma, orina, líquido cefalorraquídeo y sinovial) de mamíferos con cáncer o enfermedad inflamatoria cuando se compara con suero de mamíferos de la misma especie que no tienen cáncer ni enfermedad inflamatoria. Por consiguiente, la invención proporciona un método útil durante el diagnóstico de un tumor o enfermedad inflamatoria, que implica someter a ensayo el nivel de expresión del gen que codifica la proteína INSP215 o la cantidad de copia génica en células o fluido corporal de mamíferos y comparar el nivel de expresión génica o la cantidad de copia génica con un nivel de expresión génica de proteína INSP215 o cantidad de copia génica estándar, en donde un aumento en el nivel de expresión génica o cantidad de copia génica con respecto al estándar es un indicador de ciertos tumores o enfermedad inflamatoria.

50 Cuando ya se ha hecho un diagnóstico de un tumor o enfermedad inflamatoria de acuerdo con métodos convencionales, la presente invención es útil como un indicador de pronóstico.

55 Por "someter a ensayo el nivel de expresión del gen que codifica la proteína INSP215" se pretende significar una medición o estimación cualitativa o cuantitativa del nivel de la proteína INSP215 o el nivel del ARNm que codifica la

- INSP215 en una primera muestra biológica directamente (por ejemplo, determinando o estimando el nivel de proteína o el nivel de ARNm absoluto) o relativamente (por ejemplo, comparando el nivel de proteína INSP215 o el nivel de ARNm en una segunda muestra biológica). Por "someter a ensayo la cantidad de copias del gen que codifica la proteína INSP215" se pretende significar una medición o estimación cualitativa o cuantitativa de la cantidad de copia génica en una primera muestra biológica directamente (por ejemplo, determinando o estimando la cantidad de copia génica absoluta) o relativamente (por ejemplo, comparando la cantidad de copia génica de proteína INSP215 en una segunda muestra biológica).
- Preferiblemente, el nivel de proteína INSP215, el nivel de ARNm o la cantidad de copia génica en la primera muestra biológica se mide o estima y compara con un nivel de proteína INSP215, nivel de ARNm o cantidad de copia génica estándar, tomando el estándar de una segunda muestra biológica obtenida de un individuo que no tiene cáncer ni enfermedad inflamatoria. De forma alternativa, cuando el método se usa como un indicador de pronóstico, la primera y segunda muestra biológica pueden tomarse de individuos que tienen el cáncer o enfermedad inflamatoria y los niveles de expresión o cantidad de copia relativos se medirán para determinar el pronóstico. Como se apreciará en la técnica, una vez que se conoce el nivel de proteína INSP215, el nivel de ARNm o cantidad de copia génica estándar, puede usarse repetidamente como un estándar para comparación.
- Por "muestra biológica" se pretende significar cualquier muestra biológica obtenida de un individuo, línea celular, cultivo tisular u otra fuente que contiene proteína INSP215 o ARNm. Los métodos para obtener biopsias de tejido y fluidos corporales de mamíferos se conocen en la técnica. Cuando la muestra biológica debe incluir ARNm, se prefiere como fuente una biopsia de tejido.
- La presente invención es útil para detectar cáncer y enfermedad inflamatoria en mamíferos. En particular la invención es útil durante el diagnóstico o pronóstico en mamíferos de los tipos de cánceres y enfermedades inflamatorias mencionados en la presente invención. Los mamíferos preferidos incluyen monos, simios, gatos, perros, vacas, cerdos, caballos, conejos y humanos. Se prefieren particularmente los humanos.
- Las moléculas de ácido nucleico para diagnóstico pueden obtenerse de células de un sujeto, tales como de sangre, orina, saliva, biopsia de tejido o material de autopsia. El ADN genómico puede usarse directamente para detección o puede amplificarse enzimáticamente usando PCR, reacción en cadena de ligasa (LCR), amplificación por desplazamiento de hebra (SDA) u otras técnicas de amplificación (ver Saiki *et al.*, Nature, 324, 163-166 (1986); Bej, *et al.*, Crit. Rev. Biochem. Molec. Biol., 26, 301-334 (1991); Birkenmeyer *et al.*, J. Virol. Meth., 35, 117-126 (1991); Van Brunt, J., Bio/Technology, 8, 291-294 (1990)) antes del análisis.
- También se describe en la presente un método para diagnosticar una enfermedad en un paciente, que comprende evaluar el nivel de expresión de un gen natural que codifica un polipéptido de acuerdo con la invención y comparar dicho nivel de expresión con un nivel testigo, en donde un nivel que es diferente a dicho nivel testigo es indicador de enfermedad. El método puede comprender los pasos de:
- poner en contacto una muestra del tejido del paciente con una sonda de ácido nucleico en condiciones rigurosas que permiten la formación de un complejo híbrido entre una molécula de ácido nucleico de la invención y la sonda;
 - poner en contacto una muestra testigo con dicha sonda en las mismas condiciones usadas en el paso a);
 - y detectar la presencia de complejos híbridos en dichas muestras;
- en donde la detección de niveles del complejo híbrido en la muestra del paciente que difieren de niveles de complejo híbrido en la muestra testigo es un indicador de enfermedad.
- También se describe en la presente un método de diagnóstico que comprende los pasos de:
- obtener una muestra de tejido de una paciente que se está evaluando para detectar una enfermedad;
 - aislar una molécula de ácido nucleico de acuerdo con la invención de dicha muestra de tejido; y
 - diagnosticar al paciente para la enfermedad detectando la presencia de una mutación en la molécula de ácido nucleico que se asocia con la enfermedad.
- Para ayudar en la detección de moléculas de ácido nucleico en los métodos mencionados anteriormente, puede incluirse un paso de amplificación, por ejemplo usando PCR.
- Las supresiones e inserciones pueden detectarse mediante un cambio en el tamaño del producto amplificado en comparación con el genotipo normal. Las mutaciones puntuales pueden identificarse mediante hibridación de ADN amplificado en ARN etiquetado de la invención o de forma alternativa, secuencias de ADN antisentido etiquetadas de la invención. Las secuencias apareadas perfectamente pueden distinguirse de los duplexes incorrectamente apareados mediante digestión de RNasa o mediante evaluación de diferencias en las temperaturas de fusión. La presencia o ausencia de la mutación en el paciente puede detectarse poniendo en contacto ADN con una sonda de ácido nucleico que se hibridiza en el ADN en condiciones rigurosas para formar una molécula de doble hebra híbrida, teniendo la molécula de doble hebra híbrida una porción no hibridada de la hebra de sonda de ácido nucleico en cualquier porción correspondiente a una mutación asociada con la enfermedad; y detectando la presencia o ausencia de una porción no hibridada de la hebra de sonda como un indicador de la presencia o ausencia de una mutación asociada con la enfermedad en la porción correspondiente de la hebra de ADN.

Tales diagnósticos son particularmente útiles para pruebas prenatales e incluso neonatales.

5 Las mutaciones puntuales y otras diferencias secuenciales entre el gen de referencia y los genes "mutantes" pueden identificarse mediante otras técnicas conocidas, tales como secuenciación de ADN directo o polimorfismo conformacional de hebra única, (ver Orita *et al.*, *Genomics*, 5, 874-879 (1989)). Por ejemplo, un cebador de secuenciación puede usarse con un producto de PCR de doble hebra o una molécula usada como plantilla de hebra única generada mediante PCR modificada. La determinación de secuencia se lleva a cabo mediante procedimientos convencionales con nucleótidos radioetiquetados o mediante procedimientos de secuenciación automáticos con etiquetas fluorescentes. Los segmentos de ADN clonados también pueden usarse como sondas para detectar
10 segmentos de ADN específicos. La sensibilidad de este método se potencia considerablemente cuando se combina con PCR. Adicionalmente, las mutaciones puntuales y otras variaciones secuenciales, tales como polimorfismos, pueden detectarse como se describió anteriormente, por ejemplo, a través del uso de oligonucleótidos específicos de alelo para amplificación por PCR de secuencias que difieren por nucleótidos simples.

15 Las diferencias secuenciales de ADN también pueden detectarse mediante alteraciones en la movilidad electroforética de fragmentos de ADN en geles, con o sin agentes desnaturizantes o mediante secuenciación de ADN directa (por ejemplo, Myers *et al.*, *Science* (1985) 230:1242). Los cambios en secuencias en ubicaciones específicas también pueden revelarse mediante ensayos de protección de nucleasa, tales como protección de RNasa y S1 o el método de escisión química (ver Cotton *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* (1985) 85: 4397-4401).
20 Adicionalmente a la electrofóresis en gel convencional y secuenciación de ADN, las mutaciones tales como microsupresiones, aneuploidias, translocaciones, inversiones, también pueden detectarse mediante análisis in situ (ver, por ejemplo, Keller *et al.*, *DNA Probes*, 2da. Ed., Stockton Press, Nueva York, N.Y., EE.UU. (1993)), es decir, pueden analizarse secuencias de ADN o ARN en células por mutaciones sin necesidad de aislarlas y/o inmovilizarlas en una membrana. Actualmente la hibridación in situ fluorescente (FISH) es el método más
25 comúnmente aplicado y han aparecido numerosas descripciones de FISH (ver, por ejemplo, Trachuck *et al.*, *Science*, 250, 559-562 (1990), y Trask *et al.*, *Trends, Genet*, 7, 149-154 (1991)).

También se describe en la presente un conjunto de sondas de oligonucleótidos que comprenden una molécula de ácido nucleico de acuerdo con la invención que pueden construirse para llevar a cabo la detección sistemática
30 eficiente de variantes, mutaciones y polimorfismos genéticos. Se conoce un conjunto de métodos de tecnología y estos tienen aplicabilidad general y pueden usarse para abordar una variedad de interrogantes en genética molecular incluso expresión génica, enlace genético y variabilidad genética (ver, por ejemplo: M.Chee *et al.*, *Science* (1996), Vol 274, pp 610-613).

35 También se describe en la presente, como se prepara y usa el conjunto de acuerdo con los métodos descritos en la solicitud PCT W095/11995 (Chee *et al.*); Lockhart, D. J. *et al.* (1996) *Nat. Biotech.* 14: 1675-1680); y Schena, M. *et al.* (1996) *Proc. Natl. Acad. Sci.* 93: 10614-10619). Los pares de oligonucleótidos pueden variar de dos a más de un millón. Los oligómeros se sintetizan en áreas designadas sobre un sustrato usando un proceso químico dirigido por luz. El sustrato puede ser papel, nylon u otro tipo de membrana, filtro, chip, portaobjetos o cualquier otro soporte
40 sólido adecuado. Un oligonucleótido puede sintetizarse en la superficie del sustrato usando un procedimiento de acoplamiento químico y un aparato de aplicación de inyección de tinta, como se describe en la aplicación PCT W095/25116 (Balteschweiler *et al.*). Un conjunto "cuadrículado" análogo a una transferencia en mancha (o en ranura) puede usarse para organizar y enlazar fragmentos de ADNc u oligonucleótidos a la superficie de un sustrato usando un sistema al vacío, térmico, UV, mecánico o procedimientos de unión química. Puede producirse un
45 conjunto, tal como los descritos anteriormente, a mano o usando dispositivos (aparatos de transferencia en mancha o en ranura), materiales (cualquier soporte sólido adecuado) y máquinas (incluso instrumentos robóticos) disponibles, y puede contener 8, 24, 96, 384, 1536 o 6144 oligonucleótidos, o cualquier otra cantidad entre dos y más de un millón que permite el uso eficiente de instrumentos disponibles en el mercado.

50 Adicionalmente a los métodos descritos anteriormente, las enfermedades pueden diagnosticarse mediante métodos que comprenden determinar, a partir de una muestra derivada de un sujeto, un nivel aumentado o disminuido anormalmente de polipéptido o ARNm. La expresión aumentada o disminuida puede medirse en el nivel de ARN usando cualquiera de los métodos conocidos en la técnica para la cuantificación de polinucleótidos, tales como, por ejemplo, amplificación de ácido nucleico, por ejemplo PCR, RT-PCR, protección de RNasa, transferencia Northern y otros métodos de hibridación.

55 Las técnicas de ensayo que pueden usarse para determinar los niveles de un polipéptido de la presente invención en una muestra derivada de un huésped son conocidas para los expertos en la técnica y se discuten en más detalle a continuación (incluidos radioinmunoensayos, ensayos de unión competitiva, análisis de transferencia Western y ensayos ELISA). Este aspecto de la invención proporciona un método de diagnóstico que comprende los pasos de:
60 (a) poner en contacto un ligando, como se describió anteriormente, con una muestra biológica en condiciones adecuadas para la formación de un complejo ligando-polipéptido; y (b) detectar dicho complejo.

65 Los protocolos tales como ELISA, RIA y FACS para medir niveles de polipéptido pueden proporcionar adicionalmente una base para diagnosticar niveles alterados o anormales de expresión de polipéptido. Los valores normales o estándares para expresión de polipéptido se establecen combinando fluidos corporales o extractos celulares tomados de sujetos mamíferos normales, preferiblemente humanos, con anticuerpo para el polipéptido en

condiciones adecuadas para formación de complejos. La cantidad de formación de complejos estándar puede cuantificarse mediante diversos métodos, tales como medios fotométricos.

5 Los anticuerpos que se unen específicamente a un polipéptido de la invención pueden usarse para diagnóstico de afecciones o enfermedades caracterizadas por expresión del polipéptido, o en ensayos para monitorear pacientes que se tratan con los polipéptidos, moléculas de ácido nucleico, ligandos y otros compuestos de la invención. Los anticuerpos útiles para propósitos de diagnóstico pueden prepararse en el mismo modo como el descrito anteriormente para propósitos terapéuticos. Los ensayos de diagnóstico para el polipéptido incluyen métodos que utilizan el anticuerpo y una etiqueta para detectar el polipéptido en fluidos corporales o extractos de células o tejidos humanos. Los anticuerpos pueden usarse con o sin modificación, y pueden etiquetarse uniéndolos, covalentemente o no covalentemente, con una molécula reportera. Una amplia variedad de moléculas reporteras conocidas en la técnica pueden usarse, varias de las cuales se describieron anteriormente.

10 Las cantidades de polipéptido expresado en las muestras del sujeto, testigo y de enfermedad de tejidos biopsiados se comparan con valores estándares. La desviación entre los valores estándares y del sujeto establece los parámetros para el diagnóstico de la enfermedad. Los ensayos de diagnóstico pueden usarse para distinguir entre la ausencia, presencia y exceso de expresión del polipéptido y para monitorear la regulación de los niveles del polipéptido durante la intervención terapéutica. Tales ensayos también pueden usarse para evaluar la eficacia de un régimen de tratamiento terapéutico particular en estudios de animales, en pruebas clínicas o en el monitoreo del tratamiento de un paciente individual.

20 Un kit de diagnóstico descrito en la presente puede comprender:

- (a) una molécula de ácido nucleico de la presente invención;
- (b) un polipéptido de la presente invención; o
- (c) un ligando de la presente invención.

25 También se describe en la presente un kit de diagnóstico que puede comprender un primer recipiente que contiene una sonda de ácido nucleico que se hibridiza en condiciones rigurosas con una molécula de ácido nucleico de acuerdo con la invención; un segundo recipiente que contiene cebadores útiles para amplificar la molécula de ácido nucleico; e instrucciones para usar la sonda y los cebadores para facilitar el diagnóstico de la enfermedad. El kit puede comprender adicionalmente un tercer recipiente que contiene un agente para digerir ARN no hibridado.

30 También se describe en la presente, un kit de diagnóstico que puede comprender un conjunto de moléculas de ácido nucleico, al menos una de las mismas puede ser una molécula de ácido nucleico de acuerdo con la invención.

35 Para detectar un polipéptido de acuerdo con la invención, un kit de diagnóstico puede comprender uno o más anticuerpos que se unen a un polipéptido de acuerdo con la invención; y un reactivo útil para la detección de una reacción de unión entre el anticuerpo y el polipéptido.

40 Tales kits se usarán en el diagnóstico de una enfermedad o susceptibilidad a una enfermedad en la que están implicados polipéptidos que contienen el dominio ZP, o una enfermedad en la que están implicadas proteínas que contienen el subdominio ZP/TGR3, tal como se describió anteriormente. Diversos aspectos y realizaciones de la presente invención se describirán en más detalle a modo de ejemplo, con referencia particular a los polipéptidos INSP215.

45 Se apreciará que pueden hacerse modificaciones a los detalles sin apartarse del alcance de la invención.

Breve descripción de las figuras

50 **Figura 1:** Secuencia de proteína INSP215 con la secuencia señal prevista subrayada (residuos 1-16), el subdominio ZP/TGR3 recuadrado (residuos 39-183), la región de transmembrana doble subrayada (residuos 245-268), y el sitio de N-glicosilación resaltado (residuo 139).

55 **Figura 2:** Alineación de la secuencia de proteína INSP215 con el ADN genómico humano que codifica esta proteína (como se determina en base al ensamblaje hg16 de julio de 2003).

60 **Figura 3:** Alineación del subdominio ZP/TGR3 del INSP215 y del receptor tipo III de TGF-beta (hTGFR3; NCBI Acc. No. Q03167). Los residuos idénticos se indican con un asterisco (*). Las cisteínas también están resaltadas. Los residuos hidrófobos conservados se indican con 'h'. Los residuos polares conservados se indican con 'p'. El sitio de N-glicosilación conservado está recuadrado. La cantidad de los primeros y de los últimos residuos indicados en la alineación se escribe al lado del nombre de la proteína para cada línea (la numeración refleja la correspondiente a una proteína completa).

Figura 4: Secuencia de nucleótidos de predicción de INSP215 con traducción.

65 **Figura 5:** Secuencia de nucleótidos con traducción del producto ensamblado INSP215 por PCR.

TABLA 1

Aminoácido	Grupos sinónimos	Grupos sinónimos más preferidos
Ser	Gly, Ala, Ser, Thr, Pro	Thr, Ser
Arg	Asn, Lys, Gln, Arg, His	Arg, Lys, His
Leu	Phe, Ile, Val, Leu, Met	Ile, Val, Leu, Met
Pro	Gly, Ala, Ser, Thr, Pro	Pro
Thr	Gly, Ala, Ser, Thr, Pro	Thr, Ser
Ala	Gly, Thr, Pro, Ala, Ser	Gly, Ala
Val	Met, Phe, Ile, Leu, Val	Met, Ile, Val, Leu
Gly	Ala, Thr, Pro, Ser, Gly	Gly, Ala
Ile	Phe, Ile, Val, Leu, Met	Ile, Val, Leu, Met
Phe	Trp, Phe, Tyr	Tyr, Phe
Tyr	Trp, Phe, Tyr	Phe, Tyr
Cys	Ser, Thr, Cys	Cys
His	Asn, Lys, Gln, Arg, His	Arg, Lys, His
Gln	Glu, Asn, Asp, Gln	Asn, Gln
Asn	Glu, Asn, Asp, Gln	Asn, Gln
Lys	Asn, Lys, Gln, Arg, His	Arg, Lys, His
Asp	Glu, Asn, Asp, Gln	Asp, Glu
Glu	Glu, Asn, Asp, Gln	Asp, Glu
Met	Phe, Ile, Val, Leu, Met	Ile, Val, Leu, Met
Trp	Trp, Phe, Tyr	Trp

5 TABLA 2

Aminoácido	Grupos sinónimos
Ser	D-Ser, Thr, D-Thr, allo-Thr, Met, D-Met, Met(O), D-Met(O), L-Cys, D-Cys
Arg	D-Arg, Lys, D-Lys, homo-Arg, D-homo-Arg, Met, Ile, D-.Met, D-Ile, Orn, D-Orn
Leu	D-Leu, Val, D-Val, AdaA, AdaG, Leu, D-Leu, Met, D-Met
Pro	D-Pro, ácido L-l-tiozolidina-4-carboxílico, ácido D- o L-l-oxazolidina-4-carboxílico
Thr	D-Thr, Ser, D-Ser, allo-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
Ala	D-Ala, Gly, Aib, B-Ala, Acp, L-Cys, D-Cys
Val	D-Val, Leu, D-Leu, Ile, D-Ile, Met, D-Met, AdaA, AdaG
Gly	Ala, D-Ala, Pro, D-Pro, Aib, .beta.-Ala, Acp
Ile	D-Ile, Val, D-Val, AdaA, AdaG, Leu, D-Leu, Met, D-Met
Phe	D-Phe, Tyr, D-Thr, L-Dopa, His, D-His, Trp, D-Trp, Trans-3,4, or 5-fenilprolina, AdaA, AdaG, cis-3,4 o 5-fenilprolina, Bpa, D-Bpa
Tyr	D-Tyr, Phe, D-Phe, L-Dopa, His, D-His
Cys	D-Cys, S--Me--Cys, Met, D-Met, Thr, D-Thr
Gln	D-Gln, Asn, D-Asn, Glu, D-Glu, Asp, D-Asp
Asn	D-Asn, Asp, D-Asp, Glu, D-Glu, Gln, D-Gln
Lys	D-Lys, Arg, D-Arg, homo-Arg, D-homo-Arg, Met, D-Met, Ile, D-Ile, Orn, D-Orn
Asp	D-Asp, D-Asn, Asn, Glu, D-Glu, Gln, D-Gln
Glu	D-Glu, D-Asp, Asp, Asn, D-Asn, Gln, D-Gln
Met	D-Met, S--Me--Cys, Ile, D-Ile, Leu, D-Leu, Val, D-Val

Ejemplos

Ejemplo 1: Identificación de INSP215

Un marco de lectura abierto de 879 nucleótidos (SEQ ID NO:1) codifica el polipéptido INSP215 de longitud completa (SEQ ID NO:2), una proteína de 292 aminoácidos (ver Figura 1). El polipéptido INSP215 de longitud completa contiene un péptido señal previsto (residuos 1-16) que conduce a una secuencia madura de 276 aminoácidos (SEQ ID NO:4), codificada por una secuencia de ADN de 831 nucleótidos (SEQ ID NO:3). El polipéptido INSP215 de longitud completa contiene una región transmembrana prevista (residuos 245-268) que separa una región intracelular (residuos 269-292) y una región extracelular (residuos 1-244; SEQ ID NO:5 y SEQ ID NO:6). Son características adicionales de la región extracelular la secuencia madura (SEQ ID NO:7 y SEQ ED NO:8) que contiene el subdominio ZP/TGR3 previsto (SEQ ID NO:9 y SEQ ID NO:10) y un sitio de N-glicosilación (localizado en el residuo 139).

El subdominio ZP/TGR3 se encuentra en la secuencia de aminoácidos codificada por los primeros cuatro de los seis exones que codifican el polipéptido INSP215 de longitud completa (SEQ ID NOs 11-14; Figura 2).

Si INSP215 se usa como una consulta BLAST en bases de datos públicas de secuencias no redundantes, un segmento de región extracelular de INSP215 es altamente homólogo a la mitad C terminal del dominio ZP/TGR3 (Bork P and Sander C, FEBS Lett. 1992, 300: 237-40; Jovine L *et al.*, Nat Cell Biol. 2002, 4: 457-61, Yonezawa N and Nakano M, Biochem Biophys Res Commun. 2003, 307: 877-82; Jovine L *et al.*, Proc Natl Acad Sci EE.UU. 2004, 101: 5922-7). Esta similitud entre INSP215 y el Receptor III de TGF-beta no está limitada a las seis cisteínas separadas regularmente (residuos 85, 106, 147, 152, 176 y 183 en INSP215; figuras 1 y 3), sino que también implica varios otros residuos polares o hidrófobos, así como también un sitio de N-glicosilación (ubicados en la posición 139 y 695 de estas secuencias, respectivamente; ver Figuras 1 y 3).

Los datos preliminares sobre la expresión de INSP215 muestran que esta proteína tiene un perfil de regulación hacia arriba/abajo específico, en particular con referencia a células tumorales aisladas a partir de tejidos de colon, próstata, cerebro y neuronales.

La bibliografía muestra que el dominio ZP/TGR3 puede distinguirse estructuralmente y funcionalmente en dos mitades y que un sitio de unión a TGF-beta está localizado en la región C terminal del dominio ZP/TGR3 (WO 95/10610; Kaname S and Ruoslahti E, Biochem J. 1996, 315: 815-20). Por consiguiente, se espera que los polipéptidos INSP215 y fragmentos funcionales y/o proteínas de fusión en base a los polipéptidos INSP215, en particular polipéptidos que comprenden la secuencia de aminoácidos descritas en SEQ ID NO:10, se unan a una o más proteínas que pertenecen a la superfamilia de proteínas TGF-beta. Cuando estas secuencias relacionadas a INSP215 son proteínas secretadas o solubles, pueden actuar como antagonistas de proteínas que pertenecen a la superfamilia de proteínas TGF-beta.

En base a la anotación de los polipéptidos INSP215 en la presente, ahora es posible diseñar experimentos para detectar la presencia de los transcritos de INSP215 en un rango de tipos de tejidos humanos para determinar su expresión tisular, en particular en tejidos normales y enfermos para establecer más específicamente la relevancia de INSP215 en contextos patológicos.

La clonación de las secuencias de ADN relacionadas con INSP215 (como secuencias ADNc o genómicas exónicas y ejemplificadas en SEQ ID NO:1, 3, 5, 7, 9, 11 y 13) de tejidos de origen humano o mamífero permitirá la expresión de las correspondientes secuencias proteicas codificadas en sistemas de expresión procariotas o eucariotas y su posterior purificación y caracterización como proteínas recombinantes.

Pueden producirse secuencias de INSP215 recombinantes alternativas clonando en el vector de expresión apropiado el ADN que codifica la secuencia completa (SEQ ID NO:2), o sus fragmentos seleccionados, en forma de contenido de péptido señal (SEQ ID NO:6 y 12) o secuencias maduras (SEQ ID NO:4, 8, 10 y 14) manteniendo las propiedades de unión esperadas de la proteína completa hacia proteínas que pertenecen a la superfamilia TGF-beta. Por ejemplo, las transcripciones que incluyen exones 1-4 (cubriendo el subdominio ZP/TGR3 entero y la secuencia señal) pueden identificarse y/o clonarse mediante hibridación de cebador o mediante PCR.

Siempre que sea apropiado, las secuencias de ADN heterólogas que codifican las secuencias proteicas que ayudan en la expresión, secreción y/o detección sistemática de estas secuencias recombinantes pueden agregarse en el mismo marco del ADN que codifica INSP215, o las secuencias proteicas derivadas identificadas anteriormente. Son ejemplos de tales secuencias heterólogas péptidos señal, regiones Fc de inmunoglobulina, dominios de homo/heterodimerización y etiquetas de histidina. Por ejemplo, las proteínas de fusión que contienen el dominio extracelular del Receptor III de TGF-beta y un dominio enrollado de heterodimerización se dimerizan eficientemente y se unen a TGF-beta con afinidad alta, proporcionando una molécula con propiedades antagonistas potenciadas contra TGF-beta (De Crescenzo G *et al.*, J Biol Chem. 2004: 279: 26013-8).

Se espera que variantes solubles de INSP215 se unan a uno o más miembros de la superfamilia de proteínas TGF-beta, capturando y reteniendo la interacción de esta proteína con los receptores de señalización, con la consiguiente inhibición de la proliferación y diferenciación de múltiples tipos de células inducidas normalmente por las proteínas similares a TGF-beta. Por lo tanto, las porciones específicas de INSP215 indicadas en la presente como proteínas secretadas o solubles, que pueden generarse in vivo o in vitro luego de la digestión proteolítica o empalme alternativo, pueden tener propiedades in vivo específicas, en particular como antagonistas de una o más proteínas que pertenecen a la superfamilia TGF-beta.

Por lo tanto se espera que tales polipéptidos sean útiles para el tratamiento de enfermedades, en particular cánceres, fibrosis, cicatrización, hipertensión, aterosclerosis, trastornos reproductivos, trastornos hepáticos, trastornos pulmonares, trastornos inmunes, trastornos del desarrollo, trastornos de la piel o trastornos de tejidos conjuntivos (Tsuchida K *et al.*, Curr Drug Targets Immune Endocr Metabol Disord. 2004, 4: 157-66; Chang H *et al.*, Endocr Rev. 2002, 23: 787-823; de Caestecker M, Cytokine Growth Factor Rev. 2004, 15: 1-11; WO 95/10610; WO 97/05892; Bandyopadhyay A *et al.*, Cancer Res. 2002, 62: 4690-5; Liu M *et al.*, Am J Respir Crit Care Med. 2002, 165: 419-23; Vilchis-Landeros MM *et al.*, Biochem J. 2001, 355: 215-22).

Los polipéptidos divulgados en la presente pueden ser útiles para el tratamiento de enfermedades tales como cáncer metastásico, cáncer de pulmón de células no pequeñas, cáncer de colon, leucemia, cáncer testicular, mieloma, linfoma, cáncer colorrectal, cáncer de próstata, cánceres cerebrales, cánceres oculares, retinoblastoma, cáncer del útero, cáncer intestinal, cáncer del cuello de útero, carcinoma endometrial, cáncer de pulmón, cáncer de páncreas, condrosarcoma, linfomas no Hodgkin, tumores estromales de los cordones sexuales del ovario, carcinomas ováricos, tumores gonadales, cáncer de células renales, leucemia linfocítica crónica B, carcinoma gástrico, carcinomas pancreáticos, adenocarcinoma pancreático, cáncer de pulmón de células pequeñas, gliomas, gliomas de alto grado, gliomas malignos, gliosarcomas, astrocitomas anaplásicos, glioblastomas, meduloblastoma, ependimoma, afecciones malignas de células T, trasplantes de médula ósea, fibrosis, fibrosis pulmonar idiopática, esclerodermia, cicatrización post operatoria seguida de cirugía de glaucoma, enfermedad renal, diabetes, nefropatía diabética, aterosclerosis, cardiomiopatía restrictiva, angiogénesis, bronquiolitis obliterante, caquexia, enfermedad pulmonar obstructiva crónica, cicatrización de heridas, síndrome de Crouzon o glomerulonefritis.

Preferiblemente, la leucemia es leucemia linfocítica, leucemia linfocítica aguda, leucemia de células B aguda, leucemia de células T aguda, leucemia mielógena aguda, leucemia mielógena crónica o leucemia linfocítica crónica. Preferiblemente, el cáncer testicular es seminoma, coriocarcinoma, carcinoma embrionario, teratoma o tumor de saco vitelino.

Preferiblemente, el cáncer de pulmón se selecciona de tumores primarios benignos o malignos o de metástasis de cánceres primarios de muchos otros órganos y tejidos. Preferiblemente el cáncer de pulmón se selecciona de tumores de pulmón primarios incluidos carcinoma broncogénico, carcinoma bronquial, hamartoma condromatoso (benigno), linfoma solitario, sarcoma (maligno) o linfomas multifocales. Preferiblemente, el carcinoma broncogénico se selecciona de carcinoma de células escamosas, carcinoma indiferenciado de células pequeñas, carcinoma indiferenciado de células grandes, adenocarcinoma o carcinoma bronquioloalveolar. Preferiblemente, el carcinoma broncogénico es carcinoma de células escamosas o carcinoma de células pulmonares no pequeñas. Preferiblemente el cáncer de pulmón se selecciona de metástasis de cánceres primarios de piel, mama, colon, próstata, riñón, tiroides, estómago, cuello uterino, recto, testículo y óseo y de melanoma.

Preferiblemente, el linfoma es enfermedad de Hodgkin, linfoma linfocítico pequeño (SLL/CLL), linfoma de células del manto (MCL), linfoma folicular (FL), linfoma de zona marginal (MZL) incluido extranodal (linfoma MALT), nodal (linfoma de células B monocitoides) o esplénico, linfoma difuso de células grandes, linfoma de Burkitt, linfoma de alto grado similar a Burkitt o linfoma linfoblástico.

La INSP215 recombinante y las secuencias proteicas derivadas identificadas anteriormente pueden usarse para generar anticuerpos monoclonales o policlonales específicos que podrían usarse luego en la caracterización bioquímica de INSP215, así como también para aplicaciones de diagnóstico y/o terapéuticas (en virtud de la información divulgada en la bibliografía sobre fragmentos generados a partir de estas familias de proteínas). De forma alternativa, la INSP215 recombinante, y las secuencias proteicas derivadas identificadas anteriormente, pueden usarse en una amplia variedad de ensayos de detección sistemática, en particular los que involucran proteínas incluidas en la superfamilia de proteínas TGF-beta, para identificar ligandos específicos y actividades biológicas asociadas.

Ejemplo 2: Clonación de INSP215

INSP215 era una predicción de longitud completa para una proteína que contiene un dominio ZP de 292 aminoácidos (876 bp) codificada en 6 exones. Se identificó un péptido señal previsto dentro de la secuencia. Un dominio de transmembrana se predijo abarcando los aminoácidos 245-268.

La interrogación de bases de datos de secuencias EST públicas usando la secuencia INSP215 identificó una cantidad de ESTs que corresponden a la misma. Estas ESTs se derivaron en su mayoría de plantillas de ADNc de ojo, cerebro, pulmón, testículos o páncreas. Una secuencia EST que se derivó de ADNc de epitelio de retina, Acceso de GenBank BG752983, pareció representar INSP215 del exón 2 hasta aproximadamente el extremo de la cds prevista. Esta secuencia corresponde al clon IMAGE ID 4875922. El clon IMAGE se adquirió en ATCC y el inserto se secuenció usando los cebadores T7 y SP6. Se descubrió que el clon IMAGE contenía la totalidad de cds de INSP215 desde el exón 2 en adelante, como se esperaba, y continuaba en el 3'-UTR hasta un tracto de poliA. La secuencia corriente arriba del comienzo del exón 2 de INSP215 en el clon IMAGE (15 bp) no correspondía a INSP215 y no contenía una metionina. Se hicieron varios intentos para clonar la secuencia INSP215, como se describe a continuación.

Se diseñaron dos pares de cebadores de PCR anidados, INSP215-F1/INSP215-R1 e INSP215-F1nest/INSP215-R1nest (Tabla 3, Figura 4), para amplificar la longitud completa de la secuencia INSP215 en productos de 913 bp y 876 bp, respectivamente. Estos pares de cebadores se probaron como cebadores anidados (es decir, el par de cebadores INSP215-F1nest/INSP215-R1nest se usó en PCR2 con los productos de PCR1 de INSP215-F1/INSP215-R1 como plantillas) usando AmpliTaq en la biblioteca de ADNc candidata y plantillas individuales de ADNc, y luego en toda la colección de bibliotecas de ADNc y reservas de plantillas de ADNc. No se obtuvieron productos relevantes.

Un tercer par de cebadores, INSP215-F2/INSP215-R2 (Tabla 3, Figura 4), se diseñó para amplificar un producto de 795 bp que contenía el dominio extracelular de INSP215 previsto (INSP215-EC). Este par se probó en PCR usando AmpliTaq en la colección de bibliotecas de ADNc y reservas de plantillas de ADNc, pero aquí tampoco se obtuvieron productos relevantes.

Los 3 pares de cebadores INSP215-F1/INSP215-R1, INSP215-F1nest/INSP215-R1nest e INSP215-F2/INSP215-R2 luego se probaron en ADNc ocular y ADNc de retina y plantillas de biblioteca (las plantillas candidatas más fuertes) usando Platinum Taq HiFi. Los productos INSP215-F1/INSP215-R1 luego se usaron como plantillas para PCR anidada usando INSP215-F1nest/INSP215-R1nest e INSP215-F2/INSP215-R2. No se obtuvieron productos relevantes a partir de cualquiera de estas amplificaciones.

Un cuarto par de cebadores, INSP215-F3/INSP215-R3 (Tabla 3, Figura 4), se diseñó para amplificar un producto de 732 bp que contenía nuevamente el dominio extracelular INSP215 previsto. El cebador INSP215-F2 se posicionó en la secuencia genómica de flujo ascendente de la cds prevista, pero el cebador INSP215-F3 se posicionó dentro de la cds prevista. Por consiguiente incluso si hubiera un intrón no previsto presente justo corriente arriba del sitio de inicio, que provocaría que los cebadores INSP215-F2/INSP215-R2 fallaran, los cebadores INSP215-F3/INSP215-R3 deberían funcionar. El par de cebadores INSP215-F3/INSP215-R3 se probó en las plantillas de ADNc de retina y ojo usando ambos AmpliTaq y Platinum Taq HiFi. No se obtuvieron productos relevantes.

Un quinto par de cebadores, INSP215-F4/INSP215-R4 (Tabla 3, Figura 4), se diseñó para amplificar la región de la predicción (exón 2 - exón 5) representada en la secuencia EST, Acceso de GenBank BG752983 (ver anteriormente). Este par de cebadores también se probó en las plantillas de ADNc de retina y ojo usando ambos AmpliTaq y Platinum Taq HiFi. No se obtuvieron productos relevantes.

La cds de INSP215 prevista por lo tanto se creó mediante adición del exón 1 (60 nucleótidos) al extremo 5' del exón 2 de INSP215 (en el clon Image 4875922) mediante PCR. El ADNc de longitud completa resultante luego se subclonó con una etiqueta 6HIS en pEAK12d para expresión en células HEK293 con el nombre INSP215 (Figura 5). La proteína resultante estaba por debajo del límite de detección mediante análisis de transferencia western, usando anticuerpos dirigidos contra la etiqueta 6HIS. La falta de secreción se debió casi con seguridad al hecho de que como el ADNc de longitud completa contiene un dominio de transmembrana no se secretaría. Por consiguiente, se generó un constructo de expresión que solamente contenía el dominio extracelular de INSP215 (aminoácidos 1-234) mediante supresión de los dominios de transmembrana e intracelular de INSP215 en el clon de entrada Gateway pDONR221-INSP215-6HIS (plásmido ID. 17435) para generar pDONR221-INSP215-EC-6HIS (plásmido ID. 18175) que posteriormente se transfirió a vectores de expresión celular de mamíferos pDEST12.2 y pEAK12d para producción de INSP215 EC-6HIS recombinante in vivo e in vitro respectivamente. Los plásmidos resultantes son pDEST12.2_INSP215-EC-6HIS y pEAK12d_INSP215-EC-6HIS. El plásmido pEAK12d_INSP215-EC-6HIS se transfirió en células HEK293/EBNA para producir INSP215EC-6HIS recombinante. Se secretó la proteína etiquetada con 6HIS.

Ejemplo 3: Construcción de vectores de expresión celular de mamíferos para INSP215

El plásmido que contiene la cds de INSP215 desde el exón 2 en adelante (clon IMAGE ID 4875922) se usó como plantilla de PCR para generar clones de expresión pEAK12d y pDEST12.2 que contenían la secuencia del ORF de INSP215 con una secuencia 3' que codifica una etiqueta 6HIS usando la metodología de clonación Gateway™ (Invitrogen). El exón 1 previsto (las primeras 60 bases de la predicción que incluyen el codón de inicio) se agregó mediante PCR.

3.1 Generación de ORF de INSP215 compatible con Gateway fusionado a una secuencia con etiqueta 6HIS en el mismo marco.

La primera etapa del proceso de clonación Gateway implica una reacción por PCR de cuatro pasos que genera el ORF de INSP215 flanqueado en el extremo 5' por un sitio de recombinación attB1 y secuencia de Kozak, y flanqueado en el extremo 3' por una secuencia que codifica una etiqueta de histidina 6 (6HIS) en el mismo marco, un codón de finalización y el sitio de recombinación attB2 (ADNc compatible con Gateway). Las 60 bases faltantes en el extremo 5' de la predicción INSP215 se agregaron mediante una reacción por PCR de dos pasos con los cebadores INSP215-EX1 y INSP215-EX2.

La primera reacción por PCR (en un volumen final de 50 µl) contiene respectivamente: 1 µl (30 ng) de clon IMAGE ID 4875922 plásmido, 1,5 µl de dNTPs (10 mM), 10 µl de solución amortiguadora polimerasa 10X Pfx, 1 µl de MgSO₄ (50 mM), solución PCR_x Enhancer 2X (Invitrogen), 0,5 µl de cada cebador específico de gen (100 µM) (INSP215-EX1 e INSP215-EX4) y 0,5 µl de polimerasa de ADN Platinum Pfx (Invitrogen). La reacción por PCR se llevó a cabo usando un paso de desnaturalización inicial de 95°C durante 2 min, seguido por 20 ciclos de 94°C durante 15 s; 55°C durante 30 s y 68°C durante 2 min; y un ciclo de mantenimiento de 4°C. El producto de amplificación de PCR1 se limpió directamente usando el Sistema de Purificación de ADN Wizard PCR Clean-up (Promega) y se recuperó en 50 µl de agua estéril de acuerdo con las instrucciones del fabricante. Se visualizó una alícuota de 10 µl sobre 0,8 % de gel de agarosa en solución amortiguadora 1 X TAE (Invitrogen) para verificar que el producto tenía el peso molecular esperado (859 bp).

Tabla 3:

Cebador	Secuencia (5' - 3')
INSP127-F1	CCCAAACACTCCGACCTCA
INSP127-R1	TTCTCTACTGGGACCTCCT
INSP127-F1 nido	ATGCTGGGCACCGTGCTGCTG
INSP127-R1 nido	CTGGGACCTCCTGGGCTGGG
INSP127-F2	CCCCAAACACTCCGACCTCA
INSP127-R2	AGCACGAAGGCTGCCAACAC
INSP127-F3	ATGCTGGGCACCGTGCTG
INSP127-R3	CGGCGCGGGTCCAGGGC
INSP127-F4	TGCGCAGACCCCTCTTCAGC
INSP127-R4	TGCACGGCCGGGACACTCTT
INSP215-EX1	CTCCAGGGATCACCACCTTACCAGCGGGCCACCTGCTCCCCCGTTC
INSP215-EX2	ATGCTGGGCACCGTGCTGCTGCTGGCCCTGCTCCAGGGATCACCACC
INSP215-EX3	GCAGGCTTCGCCACCATGCTGGGCACCGTGCTG
INSP215-EX4	TGATGGTGATGGTGCTGGGACCTCCTGGGCTG
INSP215 Mut FP	CGCCCTGAGCCTCCCGCGCACCATCACCATCACCAT
INSP215 Mut RP	ATGGTGATGGTGATGGTGCGCGGGAGGCTCAGGGCG
GCP Directo	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGCC ACC
GCP Inverso	GGGGACCACTTTGTACAAGAAAGCTGGGTTTCAATGGTGATGGTGATGGTG
pEAK12F	GCCAGCTTGGCACTTGATGT
pEAK12R	GATGGAGGTGGACGTGTCAG
21M13	TGTAAAACGACGGCCAGT
M13REV	CAGGAAACAGCTATGACC
T7	TAATACGACTCACTATAGG
T3	ATTAACCTCACTAAAGG

- Secuencia subrayada = Secuencia Kozak
- Secuencia resaltada } = Codón de finalización
- Secuencia en cursiva = Etiqueta His
- Secuencia en **negrita** = Secuencia de exón 1 agregada mediante PCR

La segunda reacción por PCR (en un volumen final de 50µl) contiene respectivamente: 5 µl de producto de PCR1 purificado, 1,5 µl de dNTPs (10 mM), 10 µl de solución amortiguadora de polimerasa 10X Pfx, 1 µl de MgSO4 (50 mM), solución PCR_x Enhancer 2X (Invitrogen), 0,5 µl de cada cebador específico de gen (100 µM) (INSP215 -EX2 e INSP215-EX4) y 0,5 µl de polimerasa de ADN Platinum Pfx (Invitrogen). La reacción PCR se llevó a cabo usando un paso de desnaturalización inicial de 95°C durante 2 min, seguido por 12 ciclos de 94°C durante 15 seg; 55°C durante 30 seg y 68°C durante 2 min; y un ciclo de mantenimiento de 4°C. El producto de amplificación se purificó directamente usando el Sistema de Limpieza de PCR y Wizard SV Gel (Promega) y se recuperó en 50 µl de agua estéril de acuerdo con las instrucciones del fabricante.

La tercera reacción por PCR (en un volumen final de 50) contiene respectivamente: 5 µl de producto de PCR1 purificado, 1,5 µl de dNTPs (10 mM), 10 µl de solución amortiguadora de polimerasa 10X Pfx, 1 µl de MgSO4 (50 mM), solución PCR_x Enhancer 2X (Invitrogen), 0,5 µl de cada cebador específico de gen (100 µM) (INSP215 -EX3 e INSP215-EX4), y 0,5 µl de polimerasa de ADN Platinum Pfx (Invitrogen). La reacción PCR se llevó a cabo usando un paso de desnaturalización inicial de 95°C durante 2 min, seguido por 12 ciclos de 94°C durante 15 seg; 55°C durante 30 seg y 68°C durante 2 min; y un ciclo de mantenimiento de 4°C. El producto de amplificación se purificó directamente usando el Sistema de Limpieza de PCR y Wizard SV Gel (Promega) y se recuperó en 50 µl de agua estéril de acuerdo con las instrucciones del fabricante.

La cuarta reacción por PCR (en un volumen final de 50µl) contenía 10 µl de producto de PCR2 purificado, 1,5 µl de dNTPs (10 mM), 10 µl de solución amortiguadora de polimerasa 10X Pfx, 1 µl de MgSO4 (50 mM), solución PCR_x Enhancer 2X (Invitrogen), 0,5 µl de cebador de conversión Gateway (100 pM) (GCP directo y GCP inverso) y 0,5 µl de polimerasa de ADN Platinum Pfx. La condiciones para la 4a reacción por PCR fueron: 95°C durante 1 min; 4

ciclos de 94°C, 15 seg; 50°C, 30 seg y 68°C durante 2 min; 25 ciclos de 94°C, 15 seg; 55°C, 30 seg y 68°C, 2 min; seguido de un ciclo de mantenimiento de 4°C. Se visualizó una alícuota de 10 µl sobre 0,8 % de gel de agarosa en solución amortiguadora 1 X TAE (Invitrogen) para verificar que el producto tuviera el peso molecular esperado (876 bp + 70 = 946 bp). Los restantes 40 µl se cargaron sobre 0,8 % de gel de agarosa en solución amortiguadora 1 X TAE y la banda se purificó usando el Wizard SV Gel y el Sistema de Limpieza de PCR (Promega) y se recuperó en 50 µl en agua estéril de acuerdo con las instrucciones del fabricante.

3.2 Subclonación de ORF de INSP215 compatible con Gateway en vector de entrada Gateway pDONR221 y vectores de expresión pEAK12d y pDEST12.2

La segunda etapa del proceso de clonación Gateway implica la subclonación del producto de PCR modificado por Gateway en el vector de entrada Gateway pDONR221 (Invitrogen) como sigue: se incubaron 5 µl de producto purificado de PCR4 con 1,5 µl de vector pDONR221 (0,1 µg/µl), 2 µl de solución amortiguadora BP y 1,5 µl de mezcla de enzima BP clonasa (Invitrogen) en un volumen final de 10 µl a TA durante 1 h. La reacción se detuvo mediante adición de 1 µl de proteinasa K (2 µg/µl) y se incubó a 37°C durante 10 min adicionales. Una alícuota de esta reacción (1 µl) se usó para transformar células de E. coli DH10B mediante electroporación como sigue: una alícuota de 25 µl de células DH10B electrocompetentes (Invitrogen) se descongeló sobre hielo y se agregó 1 µl de la mezcla de reacción BP. La mezcla se transfirió a una cubeta de electroporación de 0,1 cm helada y las células se electroporaron usando un BioRad Gene-Pulser™ de acuerdo con el protocolo recomendado por el fabricante. Se agregó un medio SOC (0,5 ml) que se entibió previamente hasta alcanzar temperatura ambiente inmediatamente después de la electroporación. La mezcla se transfirió a un tubo con tapa a presión de 15 ml y se incubó, con agitación (220 rpm) durante 1 h a 37°C. Luego se colocaron en placas alícuotas de la mezcla de transformación (10 µl y 50 µl) sobre placas de caldo de cultivo L (LB) que contenían canamicina (40 µg/ml) y se incubaron durante toda la noche a 37°C.

Se preparó miniprep de ADN de plásmido a partir de 5 ml de cultivo de 8 de las colonias resultantes usando un sistema Qiaprep BioRobot 8000 (Qiagen). El ADN de plásmido (150-200 ng) se sometió a secuenciación de ADN con cebadores 21M13 y M13Rev como se describieron anteriormente usando el sistema BigDyeTerminator (Applied Biosystems cat. no. 4336919) de acuerdo con las instrucciones del fabricante. Las secuencias de los cebadores se muestran en la Tabla 3. Las reacciones de secuenciación se purificaron usando placas de limpieza Montage SEQ 96 (Millipore cat. no. LSKS09624) luego se analizaron en un secuenciador Applied Biosystems 3700.

Luego se usó el eluido de plásmido (20 aproximadamente 150 ng) de uno de los clones que contienen la secuencia correcta (pENTR_INSP215-6HIS) en una reacción de recombinación que contenía 1 µl del vector pEAK12d o del vector pDEST12.2 (0,1 µg/µl), 2 µl de solución amortiguadora LR y 1,5 µl de clonasa LR (Invitrogen) en un volumen final de 10 µl. La mezcla se incubó a TA durante 1 h, se detuvo mediante adición de proteinasa K (2 µg) y se incubó a 37°C durante 10 min adicionales. Una alícuota de esta reacción (1 µl) se usó para transformar células de E. coli DH10B mediante electroporación como sigue: una alícuota de 25 µl de células DH10B electrocompetentes (Invitrogen) se descongeló sobre hielo y se agregó 1 µl de la mezcla de reacción LR. La mezcla se transfirió a una cubeta de electroporación de 0,1 cm helada y las células se electroporaron usando un BioRad Gene-Pulser™ de acuerdo con el protocolo recomendado por el fabricante. Se agregó un medio SOC (0,5 ml) que se entibió previamente hasta alcanzar temperatura ambiente inmediatamente después de la electroporación. La mezcla se transfirió a un tubo con tapa a presión de 15 ml y se incubó, con agitación (220 rpm) durante 1 h a 37°C. Luego se colocaron en placas alícuotas de la mezcla de transformación (10 µl y 50 µl) sobre placas de caldo de cultivo L (LB) que contenían ampicilina (100 µg/ml) y se incubaron durante toda la noche a 37°C.

Se preparó miniprep de ADN de plásmido a partir de 5 ml de cultivo de 6 de las colonias resultantes subclonadas en cada vector usando un sistema Qiaprep BioRobot 8000 (Qiagen). El ADN de plásmido (200-500 ng) en el vector pEAK12d se sometió a secuenciación de ADN con los cebadores pEAK12F y pEAK12R como se describieron anteriormente. El ADN de plásmido (200-500 ng) en el vector pDEST12.2 se sometió a secuenciación de ADN con los cebadores 21M13 y M13Rev como se describieron anteriormente. Las secuencias de los cebadores se muestran en la Tabla 3.

Se preparó maxiprep de ADN purificado con gradiente de CsCl a partir de un cultivo de 500 ml del clon de secuencia verificada (pEAK12d_INSP215-6HIS) usando el método descrito por Sambrook J. *et al.*, 1989 (en Molecular Cloning, un Laboratory Manual, 2a edición, Cold Spring Harbor Laboratory Press), el ADN de plásmido se resuspendió en una concentración de 1 µg/µl en agua estéril (o 10 mM de Tris-HCl, pH 8,5) y se almacenó a -20°C.

Se preparó maxiprep de ADN libre de endotoxina a partir de un cultivo de 500 ml del clon de secuencia verificada (pDEST12.2_INSP215-6HIS) usando el kit EndoFree Plasmid Mega (Qiagen) de acuerdo con las instrucciones del fabricante. El ADN de plásmido purificado se resuspendió en solución amortiguadora TE libre de endotoxina en una concentración final de al menos 3 µg/µl y se almacenó a -20°C.

3.3 Generación de ORF de dominio EC de INSP215 compatible con Gateway fusionando a una secuencia con etiqueta 6HIS en el mismo marco.

Se suprimió la secuencia que codifica los dominios transmembrana e intracelular (aminoácidos 235-292) del inserto de ADNc de INSP215 en el plásmido pDONR221 usando mutagénesis sitio dirigida usando pDONR221_INSP215-6HIS como plantilla.

3.31 Cebadores de clonación específica de gen para mutagénesis sitio dirigida de ORF de INSP215-6HIS

Los cebadores de PCR, INSP215 Mut FP e INSP215 Mut RP (Tabla 3), se diseñaron de modo que los cebadores se alinearan con hebras opuestas de la secuencia de plásmido pDONR221_INSP215-6HIS y cada cebador se alineó con 15-25 bases en cada lado del aminoácido que habría de mutarse. Los cebadores de PCR se diseñaron siguiendo las instrucciones proporcionadas por el manual de instrucciones para QuickChange® II XL Site-Directed Mutagenesis Kit (Stratagene).

3.32 Mutagénesis sitio dirigida de ORF de INSP215-6HIS

La mutagénesis sitio dirigida se llevó a cabo usando el QuikChange® II Site-Directed Mutagenesis Kit (Stratagene) de acuerdo con las instrucciones del fabricante. La reacción se llevó a cabo en un volumen final de 150 µl que contenía solución amortiguadora de reacción 1X, 10 ng de ADN de plantilla de plásmido (pDONR221_INSP215-6HIS), 125 ng de INSP215 Mut FP e INSP215 Mut RP, 1 µl de mezcla dNTP y 2,5 unidades de polimerasa de ADN PfuUltra HF. Se llevó a cabo un ciclo térmico usando un MJ Research DNA Engine, programado de la siguiente forma: 95°C, 1 min; 18 ciclos de 95°C, 30 seg, 55°C, 1 min y 68°C, 3 min 30 seg; seguido de un ciclo de mantenimiento a 4°C.

Se usó digestión con Dpn I para digerir la plantilla de ADN parental metilado y hemimetilado (plásmido pDONR221_INSP215-6HIS en la muestra de reacción). Se agregó 1 µl de enzima de restricción Dpn I (10 U/ml, Stratagene) a los productos de las reacciones de amplificación. Las reacciones se mezclaron suavemente y se incubaron a 37°C durante 1 hora. Luego la mezcla de reacción se transfirió a células supercompetentes XL 1-Blue (Stratagene) como sigue.

Una alícuota de 50 µl de células XL 1-Blue se descongeló sobre hielo y se agregó 1 µl de ADN tratado con Dpn I.

La mezcla se incubó durante 30 min sobre hielo y luego se sometió a choque térmico mediante incubación a 42°C durante exactamente 45 s. Las muestras se volvieron a colocar en hielo durante 2 min y se agregaron 250 µl de medio NZY entibiado previamente (42°C). Las muestras se incubaron con agitación (220 rpm) durante 1 h a 37°C. La mezcla de transformación (250 µl en cada una de 2 placas) se colocó en placas de caldo de cultivo L (LB) que contenían canamicina (50 µg/ml). Las placas se incubaron durante toda la noche a 37°C.

3.33 Preparación y secuenciación de ADN de plásmido

Se seleccionó un transformante y se preparó miniprep de ADN de plásmido a partir de cultivos de 5 ml usando el kit QIAprep Spin Miniprep (Qiagen). El ADN de plásmido (150-200 ng) se sometió a secuenciación de ADN con los cebadores 21M13 y M13Rev usando el Quick Start Kit de secuenciación cíclica con tinte terminador de ciclo CEQ (Beckman Coulter P/N 608120) de acuerdo con las instrucciones del fabricante. Las secuencias de los cebadores se muestran en la Tabla 3. Las reacciones de secuenciación se analizaron en el sistema de análisis de ADN CEQ 2000 XL (Beckman Coulter P/N 608450). El análisis de secuencia identificó un clon que contenía la secuencia de inserto INSP215 esperada (pDONR221_INSP215-EC-6HIS).

3.4 Subclonación de ORG de INSP215-EC-6HIS compatible con Gateway en vectores de expresión pEAK12d y pDEST12.2

El ADN de plásmido (1 µl o aproximadamente 150 ng) de pDONR221_INSP215 -EC-6HIS luego se usó en una reacción de recombinación que contenía 1 µl del vector pEAK12d o del vector pDEST12.2 (0,1 µg/µl), 2 µl de solución amortiguadora LR y 1,5 µl de clonasa LR (Invitrogen) en un volumen final de 10 µl.

La reacción se detuvo mediante adición de 1 µl de proteinasa K (2 µg/µl) y se incubó a 37°C durante 10 min adicionales. Una alícuota de esta reacción (2 µl) se usó para transformar la cepa DH5α (Invitrogen) como sigue: una alícuota de 50 µl de células DH5α se descongeló en hielo y se agregaron 2 µl de mezcla de reacción. La mezcla se incubó durante 30 min sobre hielo y luego se sometió a choque térmico mediante incubación a 42°C durante exactamente 45 s. Las muestras se volvieron a colocar en hielo y se agregaron 250 µl de medio SOC tibio (temperatura ambiente). Las muestras se incubaron con agitación (250 rpm) durante 1 h a 37°C. Luego la mezcla de transformación se colocó sobre placas de caldo de cultivo L (LB) que contenían ampicilina (100 µg/ml) y se incubaron durante toda la noche a 37°C.

Se recogieron cinco transformantes y se ajustaron en placas de agar LB que contenían ampicilina (100 µg/ml) y se incubaron durante toda la noche a 37°C. Se resuspendió una cuchara del cultivo de la placa ajustada en 50 µl de agua y se hirvió durante 5 minutos para lisar las células. El lisado celular se centrifugó para retirar los granulosos celulares y el sobrenadante que se obtuvo se usó como plantilla para la detección sistemática por PCR de colonias.

La mezcla de PCR (en un volumen final de 25 µl) contenía 10 µl del lisado celular centrifugado, 2,0 µl de dNTPs (10 mM), 2,5 µl de solución amortiguadora de polimerasa Taq, 0,5 µl de cebadores de detección sistemática (para proporcionar una concentración final de 100 picomoles) y 0,5 µl de ADN de polimerasa Taq. Los clones pEAK12d se detectaron sistemáticamente usando los cebadores pEAK12 FP y pEAK12 RP y los clones pDEST12.2 se detectaron sistemáticamente usando los cebadores 21M13FP y M13Rev RP.

Las condiciones para la reacción por PCR de detección sistemática fueron: 95°C durante 2 min, seguido de 30 ciclos de 94°C durante 30 s; 60°C durante 30 s y 72°C durante 1 min; y un ciclo de extensión final de 72°C durante 5 minutos y un ciclo de mantenimiento de 4°C. Los productos de PCR se cargaron en 1,0% de gel de agarosa para verificar el tamaño del fragmento.

Se seleccionó un clon positivo y se preparó miniprep de ADN de plásmido a partir de cultivos de 5 ml usando el kit QIAprep Spin Miniprep (Qiagen).

El ADN de plásmido (150 - 200 ng) en el vector pEAK12d se sometió a secuenciación de ADN con los cebadores de secuenciación pEAK12 FP y pEAK12 RP como se describieron anteriormente. El ADN de plásmido (150 - 200 ng) en el vector pDEST12.2 se sometió a secuenciación de ADN con los cebadores de secuenciación 21M13 FP y M13Rev RP como se describieron anteriormente.

Los clones confirmados por secuencia se designaron pEAK12d_INSP215-EC-6HIS y pDESTI 2.2_INSP215-EC-6HIS.

3.5 Maxipreps de ADN

Se preparó maxiprep de ADN a partir de un cultivo de 500 ml de los clones verificados por secuencia (pEAK12d_INSP215-EC-6HIS) usando un kit Qiagen Mega Plasmid Prep (cat no. 12183). El ADN de plásmido se resuspendió a una concentración mínima de 1 µg/µl en agua estéril (o 10 mM de Tris-HCl, pH 8,5) y se almacenó a -20°C.

Se preparó maxiprep de ADN libre de endotoxina a partir de un cultivo de 500 ml de los clones de secuencia verificada (pDEST12.2_INSP215-6HIS) usando el kit EndoFree Plasmid Mega (Qiagen) de acuerdo con las instrucciones del fabricante. El ADN de plásmido purificado se resuspendió en solución amortiguadora TE libre de endotoxina en una concentración final de al menos 3 µg/µl y se almacenó a -20°C.

Ejemplo 4: Ensayo de cáncer

Un INSP215 soluble se expresa solo o como una proteína de fusión con Fc en células HEK 293. La proteína se purifica usando columna de proteína A o cromatografía convencional y se evalúa en un ensayo en base a células. El ensayo en base a células se desarrolla en un modelo animal existente de enfermedad de cáncer de mama metastásico. En síntesis, las células MDA-MB-231 se tratan con TGFβ1 (5 ng/ml) en presencia o ausencia de cantidad diferente de INSP215 o INSP215-Fc soluble durante 48 horas. Los medios condicionados se recogen y se ensayan para las cantidades de IL-11 producidas mediante estas células en respuesta a TGF-β. El kit IL-11 ELISA se adquirió en R&D Systems. Se mide la producción mediada por TGFβ de IL-11 por parte de células MDA-MB-231 de adenocarcinoma de mama metastásico. Se usan anticuerpos Anti-TGFβ (R&D Systems) y OPG-Fc (Serono) como testigos positivos y negativos, respectivamente.

Ejemplo 5: Estudios de micromatrices

Las micromatrices a medida se fabricaron usando el proceso de síntesis in situ sin contacto de Agilent Technologies (Agilent Technologies Inc, Palo Alto, CA) para imprimir sondas de oligonucleótidos de 60 mer de longitud, base por base, a partir de archivos de secuencias digitales. Esto se logra con un proceso de inyección de tinta que administra volúmenes exactos, extremadamente pequeños (picolitros) de los químicos que se van a detectar. La química de fosforamidita estándar usada en las reacciones permite que se mantengan eficiencias de acoplamiento muy altas en cada paso en la síntesis del oligonucleótido de longitud completa. Las cantidades precisas se depositan de forma reproducible "sobre la marcha". Esta característica de ingeniería se logra sin parar para hacer contacto con la superficie del portaobjeto y sin introducir anomalías por contacto con la superficie, resultando en uniformidad y trazabilidad consistentes de las manchas (Hughes *et al.* (2001) Nat. Biotech. Apr; 19(4): 342-7. Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer).

Síntesis de sonda

Las metodologías se llevaron a cabo de acuerdo con las instrucciones de Agilent. Esencialmente, la síntesis de ADNc y la posterior amplificación por polimerasa T7 de sonda de ARNc etiquetada con Cianina 3(5)-CTP se llevaron a cabo usando el kit de amplificación lineal fluorescente de baja producción de ARN de Agilent a partir de una plantilla de 5 µg de ARN total de acuerdo con el protocolo del kit (versión 2 de agosto de 2003, Agilent, Palo Alto, CA). Luego el ARNc se fragmenta usando el kit de hibridación In Situ de Agilent y se hibridiza de acuerdo con el protocolo de Agilent (protocolo de procesamiento de micromatriz oligo 60-mer de Agilent versión 4.1 de abril de 2004, Agilent, Palo Alto, CA).

Diseño de chip de micromatriz

- Hay 10.536 sondas en la matriz
- 5.557 de las sondas están diseñadas específicamente para detectar secuencias secretadas de interés primario
- 1.000 sondas están diseñadas como testigos negativos
- 500 sondas están diseñadas como testigos positivos

Las sondas restantes se diseñaron para secuencias de dominio público que se conoce que son proteínas extracelulares solubles o proteínas unidas a membrana secretadas con un dominio extracelular en contacto con el medio extracelular.

Ejemplo 6: Estudios adicionales

La funcionalidad y el uso de los polipéptidos de la invención pueden evaluarse adicionalmente como se establece a continuación:

Esto puede implicar la preparación de plantillas de ADNc humano, por ejemplo, a partir de una variedad de muestras de ARN total de tejido humano normal (que pueden adquirirse en Clontech, Stratagene, Ambion, Biochain Institute y prepararse internamente) usando una enzima tal como Transcriptasa inversa Superscript II RNasa H (Invitrogen). Las bibliotecas de ADNc humano (en vectores del bacteriófago lambda (λ) pueden adquirirse en Clontech, Invitrogen o pueden hacerse internamente en vectores de λ GT10. Los pares de cebadores de PCR pueden diseñarse para amplificar la secuencia codificante completa del ADNc virtual usando un programa informático tal como el Primer Designer Software (Scientific & Educational Software, PO Box 72045, Durham, NC 27722-2045, EE.UU.). Los cebadores de PCR se optimizaron para tener una Tm cercana a 55 + 10°C y un contenido de GC de 40-60%. Deberían seleccionarse los cebadores que tienen selectividad alta para la secuencia objetivo (INSP215) con poca o ninguna actividad cebadora no específica. Luego puede usarse PCR para amplificar la secuencia del gen de interés. Los procedimientos necesarios para clonar la secuencia, tales como subclonación de productos de PCR, PCR de colonias, preparación de ADN de plásmido y secuenciación y finalmente construcción de vectores de expresión celular de mamíferos se conocen la técnica (ver, por ejemplo, WO03/055913). Luego pueden llevarse a cabo experimentos adicionales para determinar la distribución en tejido y los niveles de expresión del polipéptido INSP215 in vivo, en base a las secuencias de nucleótidos y aminoácidos divulgadas en la presente.

Por ejemplo, la presencia de los transcritos para INSP215 puede investigarse mediante PCR de ADNc de diferentes tejidos humanos. Los transcritos de INSP215 pueden estar presentes en niveles muy bajos en las muestras evaluadas. Por lo tanto, es necesario tener extremo cuidado en el diseño de los experimentos para establecer la presencia de un transcrito en diversos tejidos humanos ya que una pequeña cantidad de contaminación genómica en la preparación ARN proporcionará un resultado falso positivo. Por consiguiente, todo ARN debe tratarse con ADNasa antes del uso para transcripción inversa. Adicionalmente, para cada tejido puede montarse una reacción testigo en la que no se lleva a cabo transcripción inversa (un testigo -vo de RT).

Por ejemplo, 1µg de ARN total de cada tejido puede usarse para generar ADNc usando transcriptasa inversa Multiscript (ABI) y cebadores de hexámeros aleatorios. Para cada tejido, se monta una reacción testigo en la que se agregan todos los constituyentes excepto la transcriptasa inversa (testigo -vo de RT). Las reacciones por PCR se montan para cada tejido en las muestras de ARN transcritas inversamente y los testigos negativos de RT. Los cebadores específicos para INSP215 pueden diseñarse fácilmente en base a la información de secuencia proporcionada en la presente. La presencia de un producto de peso molecular correcto en la muestra transcrita inversamente junto con la ausencia de un producto en el testigo negativo de RT puede tomarse como evidencia de la presencia de un transcrito en ese tejido. Cualquier biblioteca de ADNc adecuada puede usarse en la detección sistemática de los transcritos de INSP215, no solo aquellas generadas como se describió anteriormente.

El patrón de distribución en tejido de los polipéptidos INSP215 proporcionará información adicional útil en relación con la función de estos polipéptidos.

Adicionalmente, la sobreexpresión o supresión de la expresión de los polipéptidos en líneas celulares puede usarse para determinar el efecto en la activación transcripcional del genoma de la célula huésped. Las parejas de dimerización, co-activadores y co-represores del polipéptido INSP215 pueden identificarse mediante inmunoprecipitación combinada con transferencia Western e inmunoprecipitación combinada con espectroscopía de masas.

Listado de secuencias

```

<110>  ARES TRADING S.A.
<120>  Receptor de proteína similar a TGR3
<130>  P039409WO
<150>  GB 0426960.1
<151>  2004-12-08
<160>  47
<170>  SeqWin99, versión 1.02
<210>  1
<211>  879
<212>  ADN
<213>  Homo sapiens

<220>
<221>  Secuencia de nucleótidos INSP215 completa

<400>  1
atgctgggca ccgtgctgct gctggccctg ctcccaggga tcaccacctt acccagcggg 60
ccacctgctc ccccgttccc cgcggcgccc ggcccctggc tgcgcagacc cctcttcagc 120
ctgaagctgt ccgacacaga ggacgtcttt cctcgccgcg cggggccgct cgaggtcccg 180
gccgacagcc gcgtgttctg gcaggcggcc ttggcccgtc cctccccgcg ctggggcctg 240
gccctgcacc gctgctcagt gacgcccgtc tcacgcccgg ccccggggcc cgccttggct 300
ctgctgctg agggctgccc cgcgcacacc tetgtgctt tcccggcacc gccgcccggc 360
agcccgggtg ccgcccggcc cgcgcgtttc agcttccgcc tgcgcccggg ctccaacgcc 420
tcggtgcaat tccctgcaat ccagctgagc cgtgcccgcc gcctccgggg agtccgcccg 480
ggcctgctg ctctgacgcc gccgcccggc ccgcccgcct cgcgggtgtt gcctcaggac 540
gagggctgct ccgacactgg cagtggcagc gccgagggcc tggtgctgta cggccccac 600
ctgcacacgc tgacgcagcc tatcgtggtc accgtgccc gcccgcccc caggcccgcc 660
aagagtgtcc ccggcctgct agtgcgccct gagcctccc cgcgggcccc cgcggcctg 720
gaacccgcgc cgggtggtgg gctggtggtg gcagccttc tgcggggcgc cgcgctggcc 780
gccgggctgg gtctcgtctg tgcgcactca gcgccccag cccctggccc gccgcgaga 840
gcctgcacca gcgggtccca gccaggagg tcccagtga 879

<210>  2
<211>  292
<212>  PRT
<213>  Homo sapiens

<220>
<221>  Secuencia de proteína INSP215 completa

<400>  2
Met Leu Gly Thr Val Leu Leu Leu Ala Leu Leu Pro Gly Ile Thr Thr
1          5          10          15

Leu Pro Ser Gly Pro Pro Ala Pro Pro Phe Pro Ala Ala Pro Gly Pro
20          25          30

Trp Leu Arg Arg Pro Leu Phe Ser Leu Lys Leu Ser Asp Thr Glu Asp
35          40          45

```

ES 2 368 072 T3

Val Phe Pro Arg Arg Ala Gly Pro Leu Glu Val Pro Ala Asp Ser Arg
 50 55 60

Val Phe Val Gln Ala Ala Leu Ala Arg Pro Ser Pro Arg Trp Gly Leu
 65 70 75 80

Ala Leu His Arg Cys Ser Val Thr Pro Ser Ser Arg Pro Ala Pro Gly
 85 90 95

Pro Ala Leu Ala Leu Leu Arg Glu Gly Cys Pro Ala Asp Thr Ser Val
 100 105 110

Ala Phe Pro Pro Pro Pro Pro Pro Ser Pro Gly Ala Ala Arg Pro Ala
 115 120 125

Arg Phe Ser Phe Arg Leu Arg Pro Val Phe Asn Ala Ser Val Gln Phe
 130 135 140

Leu His Cys Gln Leu Ser Arg Cys Arg Arg Leu Arg Gly Val Arg Arg
 145 150 155 160

Ala Pro Ala Pro Leu Thr Pro Pro Pro Pro Pro Pro Ser Arg Cys
 165 170 175

Leu Pro Gln Asp Glu Ala Cys Ala Asp Thr Gly Ser Gly Ser Ala Glu
 180 185 190

Gly Leu Ala Ala Asp Gly Pro His Leu His Thr Leu Thr Gln Pro Ile
 195 200 205

Val Val Thr Val Pro Arg Pro Pro Pro Arg Pro Pro Lys Ser Val Pro
 210 215 220

Gly Arg Ala Val Arg Pro Glu Pro Pro Ala Pro Ala Pro Ala Ala Leu
 225 230 235 240

Glu Pro Ala Pro Val Val Ala Leu Val Leu Ala Ala Phe Val Leu Gly
 245 250 255

Ala Ala Leu Ala Ala Gly Leu Gly Leu Val Cys Ala His Ser Ala Pro
 260 265 270

His Ala Pro Gly Pro Pro Ala Arg Ala Ser Pro Ser Gly Pro Gln Pro
 275 280 285

Arg Arg Ser Gln
 290

<210> 3
 <211> 831
 <212> ADN
 <213> Homo sapiens

<220>
 <221> Secuencia de nucleótidos INSP215 madura

<400> 3
 ttaccacagcg ggcacactgc tccccggttc cccgaggcgc cgggccctg gctgcccaga 60
 cccctcttca gctgaagct gtccgacaca gaggactct ttctcgccg cgcggggccg 120
 ctcgaggtcc cggccgacag ccgctgttcc gtgcaggcgg ccttggcccg tcctccccg 180

ES 2 368 072 T3

```

cgtggggcc tggcctgca ccgctgctca gtgacgctt cctcagccc ggccccggg 240
ccgcctctgg ctctgtgctg tgagggtgctg cccgcccgaca cctctgtctgc cttcccgcga 300
ccgcccgcgc cgagcccggg tgcccgcgc cccgctgctt tcagcttccg cctgcccgcg 360
gtcttcaacg cctcgggtgca gttctgcaac tgccagctga gccgctgccc cgcctccgg 420
ggagtcggcc gggcgctgca gcctctgacg cccgcccgc cgcccgcgc atcggggtgt 480
ctgcctcagg acgagggctg cgcggacact ggcagtgga gccgcccagg cctgggtgt 540
gacggcccc acctgcacac gctgacgca cctatctgtg tcaccgtgccc gggcccgcgc 600
cccaggccc ccaagagtggt ccccggcctg gcagtgccc ctgagcctcc cgcgcccgc 660
cccgcccgc tggaaaccgc gccggtggtg gcctggtgt tggcagcctt cgtgctggg 720
gccgctgtg ccgcccgggt gggctctctg tgtgctgact cagcgcacca cgcctctgg 780
ccgcccgcga gagcctgccc cagcgtccc cagcccagga ggtcccagtg a 831

```

```

<210> 4
<211> 276
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> Secuencia de proteína INSP215 madura

```

```

<400> 4
Leu Pro Ser Gly Pro Pro Ala Pro Pro Phe Pro Ala Ala Pro Gly Pro
1 5 10 15

Trp Leu Arg Arg Pro Leu Phe Ser Leu Lys Leu Ser Asp Thr Glu Asp
20 25 30

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

Val Phe Val Gln Ala Ala Leu Ala Arg Pro Ser Pro Arg Trp Gly Leu
50 55 60

Ala Leu His Arg Cys Ser Val Thr Pro Ser Ser Arg Pro Ala Pro Gly
65 70 75 80

Pro Ala Leu Ala Leu Leu Arg Glu Gly Cys Pro Ala Asp Thr Ser Val
85 90 95

Ala Phe Pro Pro Pro Pro Pro Pro Ser Pro Gly Ala Ala Arg Pro Ala
100 105 110

Arg Phe Ser Phe Arg Leu Arg Pro Val Phe Asn Ala Ser Val Gln Phe
115 120 125

Leu His Cys Gln Leu Ser Arg Cys Arg Arg Leu Arg Gly Val Arg Arg
130 135 140

Ala Pro Ala Pro Leu Thr Pro Pro Pro Pro Pro Pro Ser Arg Cys
145 150 155 160

Leu Pro Gln Asp Glu Ala Cys Ala Asp Thr Gly Ser Gly Ser Ala Glu
165 170 175

Gly Leu Ala Ala Asp Gly Pro His Leu His Thr Leu Thr Gln Pro Ile
180 185 190

Val Val Thr Val Pro Arg Pro Pro Pro Arg Pro Pro Lys Ser Val Pro
195 200 205

```

Gly Arg Ala Val Arg Pro Glu Pro Pro Ala Pro Ala Pro Ala Ala Leu
 210 215 220

Glu Pro Ala Pro Val Val Ala Leu Val Leu Ala Ala Phe Val Leu Gly
 225 230 235 240

Ala Ala Leu Ala Ala Gly Leu Gly Leu Val Cys Ala His Ser Ala Pro
 245 250 255

His Ala Pro Gly Pro Pro Ala Arg Ala Ser Pro Ser Gly Pro Gln Pro
 260 265 270

Arg Arg Ser Gln
 275

<210> 5
 <211> 732
 <212> ADN
 <213> Homo sapiens

<220>
 <221> Secuencia de nucleótidos INSP215 de región extracelular completa

<400> 5
 atgctgggca ccgtgctgct gctggccctg ctcccagggg tcaccacctt acccagcggg 60
 ccacctgctc ccccgttccc cggggcgccc ggcccctggc tgcgcagacc cctcttcagc 120
 ctgaagctgt ccgacacaga ggacgtcttt cctcgccgcg cggggccgct cgaggtcccg 180
 gccgacagcc gcgtgttcgt gcaggcggcc ttggcccgtc cctccccgcg ctggggcctg 240
 gccctgcacc gctgctcagt gacgcctcc tcacgcccgg ccccggggcc cgccttggt 300
 ctgctgcgtg agggctgccc cgcgcacacc tetgtgcct tcccgccacc gccgcgcgcg 360
 agcccgggtg ccgcccgcc cgcgcgttcc agcttccgcc tgcgcccggg cttcaacgcc 420
 tcgggtcagt tctgcaactg ccagctgagc cgctgcgcc gccctccgggg agtccgcggg 480
 gcgctgcgc ctctgacgcc gccgcgcgc cgcgcgccat cgcggtgtct gcctcaggac 540
 gaggcgtgcg ccgacactgg cagtggcagc gccgagggcc tggtgtctga cggccccac 600
 ctgcacacgc tgacgcagcc tatcgtggtc accgtgcgc ggccgcccc caggccgccc 660
 aagagtgtcc ccggccgtgc agtgcgccct gaggctccc cgcgcgcccc cgcggccctg 720
 gaaccgcgc cg 732

<210> 6
 <211> 244
 <212> PRT
 <213> Homo sapiens

<220>
 <221> Secuencia de proteína INSP215 de región extracelular completa

<400> 6
 Met Leu Gly Thr Val Leu Leu Leu Ala Leu Leu Pro Gly Ile Thr Thr
 1 5 10 15

Leu Pro Ser Gly Pro Pro Ala Pro Pro Phe Pro Ala Ala Pro Gly Pro
 20 25 30

Trp Leu Arg Arg Pro Leu Phe Ser Leu Lys Leu Ser Asp Thr Glu Asp
 35 40 45

Val Phe Pro Arg Arg Ala Gly Pro Leu Glu Val Pro Ala Asp Ser Arg
 50 55 60

Val Phe Val Gln Ala Ala Leu Ala Arg Pro Ser Pro Arg Trp Gly Leu

<213> Homo sapiens

<220>

<221> Secuencia de proteína INSP215 de región extracelular madura

<400> 8

Leu Pro Ser Gly Pro Pro Ala Pro Pro Phe Pro Ala Ala Pro Gly Pro
1 5 10 15

Trp Leu Arg Arg Pro Leu Phe Ser Leu Lys Leu Ser Asp Thr Glu Asp
20 25 30

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

Val Phe Val Gln Ala Ala Leu Ala Arg Pro Ser Pro Arg Trp Gly Leu
50 55 60

Ala Leu His Arg Cys Ser Val Thr Pro Ser Ser Arg Pro Ala Pro Gly
65 70 75 80

Pro Ala Leu Ala Leu Leu Arg Glu Gly Cys Pro Ala Asp Thr Ser Val
85 90 95

Ala Phe Pro Pro Pro Pro Pro Pro Ser Pro Gly Ala Ala Arg Pro Ala
100 105 110

Arg Phe Ser Phe Arg Leu Arg Pro Val Phe Asn Ala Ser Val Gln Phe
115 120 125

Leu His Cys Gln Leu Ser Arg Cys Arg Arg Leu Arg Gly Val Arg Arg
130 135 140

Ala Pro Ala Pro Leu Thr Pro Pro Pro Pro Pro Pro Pro Ser Arg Cys
145 150 155 160

Leu Pro Gln Asp Glu Ala Cys Ala Asp Thr Gly Ser Gly Ser Ala Glu
165 170 175

Gly Leu Ala Ala Asp Gly Pro His Leu His Thr Leu Thr Gln Pro Ile
180 185 190

Val Val Thr Val Pro Arg Pro Pro Pro Arg Pro Pro Lys Ser Val Pro
195 200 205

Gly Arg Ala Val Arg Pro Glu Pro Pro Ala Pro Ala Pro Ala Ala Leu
210 215 220

Glu Pro Ala Pro
225

<210> 9

<211> 435

<212> ADN

<213> Homo sapiens

<220>

<221> Secuencia de nucleótidos INSP215 de subdominio ZP/TGR3

<400> 9

```

ttcagcctga agctgtccga cacagaggac gtctttcctc gccgcgcggg gccgctcgag 60
gtcccggccg acagccgcgt gttcgtgcag gggccttgg cccgtccctc cccgcgctgg 120
ggcctggccc tgcaccgctg ctcaagtgcg ccgtcctcac gcccgcccc ggggcccgcc 180
ctggetctgc tgcgtgaggg ctgccccgcc gacacctctg tcgcttccc gccaccgccc 240
ccgcccagcc cgggtgccgc ccgccccgcc cgtttcagct tccgcctgcg cccggtcttc 300
aacgctcgg tgcagttcct gcaactgccag ctgagccgct gccgcgcct cccggggagtc 360
cgccgggcgc ctgcgctctt gacgccgcgc ccgcccgcgc cgccatcgcg gtgtctgct 420
caggacgagg cgtgc 435

```

```

<210> 10
<211> 145
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> Secuencia de proteína INSP215 de subdominio ZP/TGR3

```

```

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

```

Cys
145

```

<210> 11
<211> 651
<212> ADN
<213> Homo sapiens

```

```

<220>
<221> Secuencia de nucleótidos INSP215 de exones 1-4

```

```

<400> 11
atgctgggca ccgtgctgct gctggccctg ctcccaggga tcaccaoctt acccagcggg 60
ccacctgctc ccccgttccc cgcgggcgcc ggcccctggc tgcgcagacc cctcttcagc 120
ctgaagctgt ccgacacaga ggacgtcttt cctcgcgcgc cggggccgct cgaggtcccc 180

```

```

gccgacagcc gcggtgttcgt gcaggcggcc ttggcccgtc cctccccgcg ctggggcctg 24(
gccctgcacc gctgctcagt gacgcccgtc tcacgcccgg ccccggggcc cgccctggct 30(
ctgctgcgtg agggctgccc cgcgacacc tctgtgcct tcccgccacc gccgcgccc 36(
agcccgggtg ccgcccgcc cgcgcgtttc agcttccgcc tgcgcccggg cttcaacgcc 42(
tcgggtgcagt tctgcaactg ccagctgagc cgtctccgcc gcctccgggg agtccgccc 48(
gcgctgcgc ctctgacgcc gccgcccggc ccgcccgcct cgcggtgtct gcctcaggac 54(
gaggcgtgcg ccgacactgg cagtggcagc gccgagggcc tggctgctga cggccccccac 60(
ctgcacacgc tgacgcagcc tctcgtggtc accgtgcccg ggcccgcctc c 65(

```

```

<210> 12
<211> 217
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> Secuencia de proteína INSP215 de exones 1-4

```

```

<400> 12
Met Leu Gly Thr Val Leu Leu Leu Ala Leu Leu Pro Gly Ile Thr Thr
1 5 10 15

Leu Pro Ser Gly Pro Pro Ala Pro Pro Phe Pro Ala Ala Pro Gly Pro
20 25 30

Trp Leu Arg Arg Pro Leu Phe Ser Leu Lys Leu Ser Asp Thr Glu Asp
35 40 45

Val Phe Pro Arg Arg Ala Gly Pro Leu Glu Val Pro Ala Asp Ser Arg
50 55 60

Val Phe Val Gln Ala Ala Leu Ala Arg Pro Ser Pro Arg Trp Gly Leu
65 70 75 80

Ala Leu His Arg Cys Ser Val Thr Pro Ser Ser Arg Pro Ala Pro Gly
85 90 95

Pro Ala Leu Ala Leu Leu Arg Glu Gly Cys Pro Ala Asp Thr Ser Val
100 105 110

Ala Phe Pro Pro Pro Pro Pro Pro Ser Pro Gly Ala Ala Arg Pro Ala
115 120 125

Arg Phe Ser Phe Arg Leu Arg Pro Val Phe Asn Ala Ser Val Gln Phe
130 135 140

Leu His Cys Gln Leu Ser Arg Cys Arg Arg Leu Arg Gly Val Arg Arg
145 150 155 160

Ala Pro Ala Pro Leu Thr Pro Pro Pro Pro Pro Pro Ser Arg Cys
165 170 175

Leu Pro Gln Asp Glu Ala Cys Ala Asp Thr Gly Ser Gly Ser Ala Glu
180 185 190

Gly Leu Ala Ala Asp Gly Pro His Leu His Thr Leu Thr Gln Pro Ile
195 200 205

Val Val Thr Val Pro Arg Pro Pro Pro
210 215

```

<210> 13
 <211> 603
 <212> ADN
 <213> Homo sapiens

<220>
 <221> Secuencia de nucleótidos INSP215 de exones 1-4 madura

<400> 13
 ttaccacagcg ggccacctgc tcccccgttc cccgcggcgc ccggcccctg gctgcgcaga 60
 ccctcttca gctgaagct gtccgacaca gaggcgtct ttctcgcgcg cgcggggccg 120
 ctcgaggtcc cggccgacag ccgcgtgttc gtgcaggcgg ccttggcccg tcctccccg 180
 cgctggggcc tggccctgca ccgctgtca gtgacgcctg cctcagccc ggccccggg 240
 cccgccctgg ctctgctgcg tgagggtgc cccgcgcaca cctctgtcgc ctccccgca 300
 ccgcccgcgc cgagcccggg tgccgcccgc cccgcgcgtt tcagcttccg cctgcgcccg 360
 gtctcaacg cctcgggtgca gttcctgcac tgccagctga gcogctgccc ccgctccgg 420
 ggagtcgcc gggcgccctgc gcctctgac ccgcccgcgc cgcgcgcgcc atcggggtgt 480
 ctgcctcagg acgagggcgtg cgcgcgacct ggcagtgga gcgcccaggg cctggctgct 540
 gacggcccc acctgcacac gctgacgcag cctatcgtgg tcaccgtgcc gcggccgccc 600
 ccc 603

<210> 14
 <211> 201
 <212> PRT
 <213> Homo sapiens

<220>
 <221> Secuencia de proteína INSP215 de exones 1-4 madura

<400> 14
 Leu Pro Ser Gly Pro Pro Ala Pro Pro Phe Pro Ala Ala Pro Gly Pro
 1 5 10 15
 Trp Leu Arg Arg Pro Leu Phe Ser Leu Lys Leu Ser Asp Thr Glu Asp
 20 25 30
 Val Phe Pro Arg Arg Ala Gly Pro Leu Glu Val Pro Ala Asp Ser Arg
 35 40 45
 Val Phe Val Gln Ala Ala Leu Ala Arg Pro Ser Pro Arg Trp Gly Leu
 50 55 60
 Ala Leu His Arg Cys Ser Val Thr Pro Ser Ser Arg Pro Ala Pro Gly
 65 70 75 80
 Pro Ala Leu Ala Leu Leu Arg Glu Gly Cys Pro Ala Asp Thr Ser Val
 85 90 95
 Ala Phe Pro Pro Pro Pro Pro Pro Ser Pro Gly Ala Ala Arg Pro Ala
 100 105 110
 Arg Phe Ser Phe Arg Leu Arg Pro Val Phe Asn Ala Ser Val Gln Phe
 115 120 125
 Leu His Cys Gln Leu Ser Arg Cys Arg Arg Leu Arg Gly Val Arg Arg
 130 135 140
 Ala Pro Ala Pro Leu Thr Pro Pro Pro Pro Pro Pro Pro Ser Arg Cys
 145 150 155 160

Leu Pro Gln Asp Glu Ala Cys Ala Asp Thr Gly Ser Gly Ser Ala Glu
 165 170 175
 Gly Leu Ala Ala Asp Gly Pro His Leu His Thr Leu Thr Gln Pro Ile
 180 185 190
 Val Val Thr Val Pro Arg Pro Pro Pro
 195 200

<210> 15
 <211> 13
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Secuencia ligadora

<220>
 <221> Secuencia ligadora de proteína de fusión

<400> 15
 Glu Phe Gly Ala Gly Leu Val Leu Gly Gly Gln Phe Met
 1 5 10

<210> 16
 <211> 1163
 <212> ADN
 <213> Secuencia prevista

<220>
 <221> Secuencia de ANDc de INSP215 prevista

<400> 16
 ccttgccaag gacagggccg cctgcccacg tgcgggacct ggggtgcagcc cagctgggtgt 60
 ctgctctggc cccaacact ccgacctcag ctggtccagc atgctgggca ccgtgctgct 120
 gctggccctg ctcccaggga tcaccacctt acccagcggg ccacctgctc ccccgttccc 180
 cgcggcgccc ggcccctggc tgcgcagacc cctcttcagc ctgaagctgt ccgacacaga 240
 ggacgtcttt cctcgccgcg cggggccgct cgaggtcccg gccgacagcc gcgtgttcgt 300
 gcaggcggcc ttggcccgtc cctcccgcg ctggggcctg gccctgcacc gctgctcagt 360
 gacgcccgtc tcacgcccgg ccccggggcc cgcctggct ctgctgctg agggctgccc 420
 cgccgacacc tctgtcgct tcccgccacc gccgcccgg agcccgggtg ccgcccgcc 480
 cgcgcttcc agcttcggc tgcgcccgg cttcaacgcc tcgggtgcagt tctgactg 540
 ccagctgagc cgtgcccgc gcctccggg agtccgccc gcgctgctc ctctgacgcc 600
 gccgcccgg ccgcccgat cgcggtgtct gcctcaggac gaggcgtgct ccgacactgg 660
 cagtggcagc gccgagggcc tggctgctga cggccccac ctgcacacgc tgacgcagcc 720
 tatcgtggtc accgtgccc ggccgcccc caggccccc aagagtgtcc ccggccgtgc 780
 agtgcgccc gagcctccc cgcgggccc cgcggcccctg gaaccgcgc cgggtggtggc 840
 gctgggtgtg gcagccttcg tgcctgggccc cgcgctggcc gccgggctgg gtctcgtctg 900
 tgcgactca gcgcccacg cccctggccc gcccgcgaga gcctcgcca gcggtcccca 960
 gccagggagg tcccagtgag gaagggatgg tgcgccccca acatggtccg gagatacacc 1020
 cagctaccaa ttcgggacca ggaccaacag gaccggacc gcctccctgg acctcggacc 1080
 tgatgaggcc acgaccctg cgcttctctc ctcccctgt cctcccacc tgtgctcaaa 1140
 ataaacctct ggactgaccg gct 1163

<210> 17
 <211> 993
 <212> ADN
 <213> Secuencia artificial

<220>

<223> Inserto de clon

<220>

<221> Inserto de ADNc en clon Image 4875922

<400> 17

```

ggcacgaggc cgccagctcc cccgttcccc ggggcccgc gccctggct ggcgagacc 60
ctcttcagcc tgaagctgtc cgacacagag gacgtctttc ctgcccgcgc ggggcccgtc 120
gaggtcccgg ccgacagccg cgtgttcgtg caggcggcct tggcccgtcc ctccccgcgc 180
tggggcctgg ccctgcaacc ctgtcagtg acgcccgtcc cagcccggc cccggggccc 240
gccctggctc tgetgcgtga gggctgcccc gccgacacct ctgtgcctt ccgcccaccg 300
ccgcccgcga gcccggtgc cggcccgcgc ggcggtttca gcttcgcct ggcgcccgtc 360
ttcaacgcct cgggtcagtt cctgcaactg cagctgagcc gctgcccgc cctccgggga 420
gtccgcccgg cgcctgcgcc tctgacggc ccgcccgcgc cgcggccatc gcggtgtctg 480
cctcaggacg aggcgtgcgc cgacactggc agtggcagcg ccgagggcct ggctgctgac 540
ggccccacc tgacacagct gacgcagcct atcgtggtca ccgtgcccgc gccgcccccc 600
agggcccaca agagtgtccc cggccgtgca gtgcgcccgc agcctcccgc gccggcccc 660
gcggcccctg aaccgcgcgc ggtggtggcg ctggtggttg cagccttctg gctggggccc 720
gcgctggccc ccgggtggg tctcgtctgt gcgcaactag cgcgccacgc cctgggccc 780
cccgcgagag cctcgcaccg cggccccag ccagggaggt ccagtgagg aagggatggt 840
gcgcccccaa catggtccgg agatacacc agctaccaat tcgggaccag gaccaacagg 900
accggaccgc cctccctgga cctcggacct gataggcca cgaccctgc gcttctctcc 960
tcccctgtc cctcccaact gtgctcaaaa taa 993

```

<210> 18

<211> 894

<212> ADN

<213> Secuencia artificial

<220>

<223> Secuencia codificante de proteína de fusión

<220>

<221> Secuencia de nucleótidos INSP215 de longitud completa con etiqueta 6HI:

<400> 18

```

atgctgggca ccgtgctgct gctggcccctg ctcccagga tcaccacctt acccagcggg 60
ccacctgtc ccccgctccc cgcggcgcgc gggcccctggc tgcgcagacc cctcttcagc 120
ctgaagctgt ccgacacaga ggacgtcttt cctcgcgcgc cggggcccgt cgaggtccc 180
gccgacagcc gcgtgttcgt gcaggcggcc ttggcccgtc cctccccgcg ctggggcctg 240
gccctgacc gctgctcagt gacgcccgtc tcagcccgcg ccccggggccc cgcctggct 300
ctgctgcgtg agggctgccc cgcgacacc tetgtgcct tcccgcacc gccgcccgcg 360
agcccgggtg ccgcccgcgc cgcggtttc agcttcgcgc tgcgcccgtt ctcaacgcc 420
tgggtgcagt tctgcaactg ccagctgagc cgtgcgcgc gctccgggg agtcgcgcg 480
gcgctgcgc ctctgacgcc gccgcccgcg ccgcccacc cgcggtgtct gcctcaggac 540
gagggctgcg ccgacactgg cagtggcagc gccgagggcc tggctgctga cggccccacc 600
ctgcacacgc tgacgcagcc tatcgtggtc accgtgcccg ggcccgcgcc caggcccgcc 660
aagagtgtcc ccggccgtgc agtgcgccc gagcctccc cgcgggccc cgcggcccctg 720
gaaccgcgc cgggtggtggc gctggtggtg gcagccttcg tctggggcgc cgcgctggcc 780
gcggggtggt gtctcgtctg tgcgcaacta gcgcccacg cccctggccc gccgcgaga 840
gcctcgcaca gcggtcccca gccagggag tcccagcacc atcaccatca ccat 894

```

<210> 19

<211> 298

<212> PRT

<213> Proteína de fusión

<220>

<221> Secuencia de proteína INSP215 de longitud completa con etiqueta 6HIS

ES 2 368 072 T3

<400> 19
 Met Leu Gly Thr Val Leu Leu Leu Ala Leu Leu Pro Gly Ile Thr Thr
 1 5 10 15
 Leu Pro Ser Gly Pro Pro Ala Pro Pro Phe Pro Ala Ala Pro Gly Pro
 20 25 30
 Trp Leu Arg Arg Pro Leu Phe Ser Leu Lys Leu Ser Asp Thr Glu Asp
 35 40 45
 Val Phe Pro Arg Arg Ala Gly Pro Leu Glu Val Pro Ala Asp Ser Arg
 50 55 60
 Val Phe Val Gln Ala Ala Leu Ala Arg Pro Ser Pro Arg Trp Gly Leu
 65 70 75 80
 Ala Leu His Arg Cys Ser Val Thr Pro Ser Ser Arg Pro Ala Pro Gly
 85 90 95
 Pro Ala Leu Ala Leu Leu Arg Glu Gly Cys Pro Ala Asp Thr Ser Val
 100 105 110
 Ala Phe Pro Pro Pro Pro Pro Pro Ser Pro Gly Ala Ala Arg Pro Ala
 115 120 125
 Arg Phe Ser Phe Arg Leu Arg Pro Val Phe Asn Ala Ser Val Gln Phe
 130 135 140
 Leu His Cys Gln Leu Ser Arg Cys Arg Arg Leu Arg Gly Val Arg Arg
 145 150 155 160
 Ala Pro Ala Pro Leu Thr Pro Pro Pro Pro Pro Pro Ser Arg Cys
 165 170 175
 Leu Pro Gln Asp Glu Ala Cys Ala Asp Thr Gly Ser Gly Ser Ala Glu
 180 185 190
 Gly Leu Ala Ala Asp Gly Pro His Leu His Thr Leu Thr Gln Pro Ile
 195 200 205
 Val Val Thr Val Pro Arg Pro Pro Pro Arg Pro Pro Lys Ser Val Pro
 210 215 220
 Gly Arg Ala Val Arg Pro Glu Pro Pro Ala Pro Ala Pro Ala Ala Leu
 225 230 235 240
 Glu Pro Ala Pro Val Val Ala Leu Val Leu Ala Ala Phe Val Leu Gly
 245 250 255
 Ala Ala Leu Ala Ala Gly Leu Gly Leu Val Cys Ala His Ser Ala Pro
 260 265 270
 His Ala Pro Gly Pro Pro Ala Arg Ala Ser Pro Ser Gly Pro Gln Pro
 275 280 285
 Arg Arg Ser Gln His His His His His His
 290 295
 <210> 20
 <211> 720

```

<212>  ADN
<213>  Secuencia codificante de proteína de fusión

<220>
<221>  Secuencia de nucleótidos INSP215 extracelular con etiqueta 6HIS

<400>  20
atgctgggca ccgctgctgct gctggcctg ctcccagga tcaccacctt acccagcggg 60
ccacctgctc ccccgttccc cgcggcgccc ggcccctggc tgcgcagacc cctcttcagc 120
ctgaagctgt cgcacacaga ggacgtctt cctcgccgcg cggggccgct cgaggtcccg 180
gccgacagcc gcgtgttcgt gcaggcgccc ttggcccgtc cctcccgcg ctggggcctg 240
gcctgcacc gctgctcagt gacgcccgtc tcaegcccgg ccccggggcc cgcctggct 300
ctgctgcgtg agggctgccc cgcgcacacc tctgtgcct tcccgccacc gccgcccgcg 360
agcccgggtg cgcgccgccc cgcgcgtttc agcttcgcgc tgcgcccggc cttcaacgcc 420
tcgggtgagt tctgcaactg ccagctgagc cgtgcccgcc gcctccgggg agtccgccg 480
gcgctgcgc ctctgacgcc gcccccgcg ccgcccgat cgcggtgtct gcctcaggac 540
gaggcgtgcg ccgacactgg cagtggcagc gccgagggcc tggctgtgta cggccccac 600
ctgcacacgc tgacgcagcc tatcgtggtc accgtgcgc gcccgcccc caggccgccc 660
aagagtgtcc cggcgcgtgc agtgcgcctt gagcctccc cgcaccatca ccatcaccat 720

<210>  21
<211>  240
<212>  PRT
<213>  Proteína de fusión

<220>
<221>  Secuencia de proteína INSP215 extracelular con etiqueta 6HIS

<400>  21
Met Leu Gly Thr Val Leu Leu Leu Ala Leu Leu Pro Gly Ile Thr Thr
1 5 10 15
Leu Pro Ser Gly Pro Pro Ala Pro Pro Phe Pro Ala Ala Pro Gly Pro
20 25 30
Trp Leu Arg Arg Pro Leu Phe Ser Leu Lys Leu Ser Asp Thr Glu Asp
35 40 45
Val Phe Pro Arg Arg Ala Gly Pro Leu Glu Val Pro Ala Asp Ser Arg
50 55 60
Val Phe Val Gln Ala Ala Leu Ala Arg Pro Ser Pro Arg Trp Gly Leu
65 70 75 80
Ala Leu His Arg Cys Ser Val Thr Pro Ser Ser Arg Pro Ala Pro Gly
85 90 95
Pro Ala Leu Ala Leu Leu Arg Glu Gly Cys Pro Ala Asp Thr Ser Val
100 105 110
Ala Phe Pro Pro Pro Pro Pro Pro Ser Pro Gly Ala Ala Arg Pro Ala
115 120 125
Arg Phe Ser Phe Arg Leu Arg Pro Val Phe Asn Ala Ser Val Gln Phe
130 135 140
Leu His Cys Gln Leu Ser Arg Cys Arg Arg Leu Arg Gly Val Arg Arg
145 150 155 160
Ala Pro Ala Pro Leu Thr Pro Pro Pro Pro Pro Pro Ser Arg Cys

```


ES 2 368 072 T3

			165					170					175			
Leu	Pro	Gln	Asp	Glu	Ala	Cys	Ala	Asp	Thr	Gly	Ser	Gly	Ser	Ala	Glu	
			180					185					190			
Gly	Leu	Ala	Ala	Asp	Gly	Pro	His	Leu	His	Thr	Leu	Thr	Gln	Pro	Ile	
		195					200					205				
Val	Val	Thr	Val	Pro	Arg	Pro	Pro	Pro	Arg	Pro	Pro	Lys	Ser	Val	Pro	
	210					215					220					
Gly	Arg	Ala	Val	Arg	Pro	Glu	Pro	Pro	Ala	His	His	His	His	His	His	
225					230					235					240	

<210> 22
 <211> 19
 <212> **ADN**
 <213> **Secuencia artificial**

<220>
 <223> **Cebador de PCR de oligonucleótido**

<220>
 <221> INSP215-F1

<400> 22
 cccaaacact ccgacctca 19

<210> 23
 <211> 19
 <212> **ADN**
 <213> **Secuencia artificial**

<220>
 <223> **Cebador de PCR de oligonucleótido**

<220>
 <221> INSP215-R1

<400> 23
 ttctcactg ggacctcct 19

<210> 24
 <211> 21
 <212> **ADN**
 <213> **Secuencia artificial**

<220>
 <223> **Cebador de PCR de oligonucleótido**

<220>
 <221> INSP215-Flnest

<400> 24
 atgctgggca ccgtgctgct g 21

<210> 25
 <211> 20
 <212> **ADN**

<213> **Secuencia artificial**
 <220>
 <223> **Cebador de PCR de oligonucléotido**
 <220>
 <221> **INSP215-Rlnest**
 <400> 25
 ctgggacctc ctgggctggg 20
 <210> 26
 <211> 20
 <212> **ADN**
 <213> **Secuencia artificial**
 <220>
 <223> **Cebador de PCR de oligonucléotido**
 <220>
 <221> **INSP215-F2**
 <400> 26
 ccccaaacac tccgacctca 20
 <210> 27
 <211> 20
 <212> **ADN**
 <213> **Secuencia artificial**
 <220>
 <223> **Cebador de PCR de oligonucléotido**
 <220>
 <221> **INSP215-R2**
 <400> 27
 agcacgaagg ctgccaacac 20
 <210> 28
 <211> 18
 <212> **ADN**
 <213> **Secuencia artificial**
 <220>
 <223> **Cebador de PCR de oligonucléotido**
 <220>
 <221> **INSP215-F3**
 <400> 28
 atgctgggca ccgtgctg 18
 <210> 29
 <211> 18
 <212> **ADN**
 <213> **Secuencia artificial**
 <220>
 <223> **Cebador de PCR de oligonucléotido**

<220>
 <221> INSP215-R3

 <400> 29
 cggcgcgggt tccagggc 18

 <210> 30
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Cebador de PCR de oligonucléotido

 <220>
 <221> INSP215-F4

 <400> 30
 tgcgcagacc cctcttcagc 20

 <210> 31
 <211> 21
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Cebador de PCR de oligonucléotido

 <220>
 <221> INSP215-R4

 <400> 31
 tgcacggcgc gggacactct t 21

 <210> 32
 <211> 48
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Cebador de PCR de oligonucléotido

 <220>
 <221> INSP215-EX1

 <400> 32
 ctcccagga tcaccacctt acccagcggg ccacctgctc ccccgctc 48

 <210> 33
 <211> 48
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Cebador de PCR de oligonucléotido

 <220>
 <221> INSP215-EX2

<400> 33
atgctgggca ccgctgctgct gctggcctg ctcccagga tcaccacc 48

<210> 34
<211> 33
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador de PCR de oligonucléotido

<220>
<221> INSP215-EX3

<400> 34
gcaggcttcg ccaccatgct gggcaccgtg ctg 33

<210> 35
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador de PCR de oligonucléotido

<220>
<221> INSP215-EX4

<400> 35
tgatggtgat ggtgctggga cctcctgggc tg 32

<210> 36
<211> 36
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador de PCR de oligonucléotido

<220>
<221> INSP215 Mut FP

<400> 36
cgccctgagc ctcccgcgca ccatcaccat caccat 36

<210> 37
<211> 36
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador de PCR de oligonucléotido

<220>
<221> INSP215 Mut RP

<400> 37
atggtgatgg tgatggtgcg cgggaggctc agggcg 36

<210> 38

<211> 37
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Cebador de PCR de oligonucléotido

 <220>
 <221> GCP directo

 <400> 38
 ggggacaagt ttgtacaaaa aagcaggctt cgccacc 37

 <210> 39
 <211> 51
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Cebador de PCR de oligonucléotido

 <220>
 <221> GCP inverso

 <400> 39
 ggggaccact ttgtacaaga aagctgggtt tcaatggtga tggatgatg g 51

 <210> 40
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Cebador de PCR de oligonucléotido

 <220>
 <221> pEAK12F

 <400> 40
 gccagcttgg cacttgatgt 20

 <210> 41
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Cebador de PCR de oligonucléotido

 <220>
 <221> pEAK12R

 <400> 41
 gatggagggtg gacgtgtcag 20

 <210> 42
 <211> 18
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Cebador de PCR de oligonucléotido

<220>
 <221> 21M13

<400> 42
 tgtaaaacga cggccagt 18

<210> 43
 <211> 18
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Cebador de PCR de oligonucléotido

<220>
 <221> M13REV

<400> 43
 caggaaacag ctatgacc 18

<210> 44
 <211> 19
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Cebador de PCR de oligonucléotido

<220>
 <221> T7

<400> 44
 taatagcact cactatagg 19

<210> 45
 <211> 18
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Cebador de PCR

<220>
 <221> T3

<400> 45
 attaaccctc actaaagg 18

<210> 46
 <211> 95
 <212> PRT
 <213> Homo sapiens

<220>
 <221> Secuencia de polipéptidos ORFX ORF290 humana SEQ ID NO:580 en
 WO200058473

ES 2 368 072 T3

<400> 46
 Met Leu Gly Thr Val Leu Leu Leu Ala Leu Leu Pro Gly Ile Thr Thr
 1 5 10 15
 Leu Pro Ser Gly Pro Pro Ala Pro Pro Phe Pro Ala Ala Pro Gly Pro
 20 25 30
 Trp Leu Arg Arg Pro Leu Phe Ser Leu Lys Leu Ser Asp Thr Glu Asp
 35 40 45
 Val Phe Pro Arg Arg Ala Gly Pro Leu Glu Val Pro Ala Asp Ser Arg
 50 55 60
 Val Phe Val Gln Ala Ala Leu Ala Arg Pro Ser Pro Arg Trp Gly Leu
 65 70 75 80
 Ala Leu His Arg Cys Ser Val Thr Pro Ser Ser Arg Pro Ala Pro
 85 90 95

<210> 47
 <211> 316
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Homo sapiens

<220>
 <221> PROTEÍNA 33 KDA IPI00645072.1

<400> 47
 Met Gly Glu Ser Ala Ala Ala Thr Ala Ser Leu Phe Gln Arg Arg Arg
 1 5 10 15
 Arg Gly Arg Gly Gly Arg Val Thr Phe Pro Gly Gly Leu Lys Gly Ser
 20 25 30
 Ala Arg Phe Leu Ser Phe Gly Pro Pro Phe Pro Ala Pro Pro Ala Pro
 35 40 45
 Pro Phe Pro Ala Ala Pro Gly Pro Trp Leu Arg Arg Pro Leu Phe Ser
 50 55 60
 Leu Lys Leu Ser Asp Thr Glu Asp Val Phe Pro Arg Arg Ala Gly Pro
 65 70 75 80
 Leu Glu Val Pro Ala Asp Ser Arg Val Phe Val Gln Ala Ala Leu Ala
 85 90 95
 Arg Pro Ser Pro Arg Trp Gly Leu Ala Leu His Arg Cys Ser Val Thr
 100 105 110
 Pro Ser Ser Arg Pro Ala Pro Gly Pro Ala Leu Ala Leu Leu Arg Glu
 115 120 125
 Gly Cys Pro Ala Asp Thr Ser Val Ala Phe Pro Pro Pro Pro Pro Pro
 130 135 140
 Ser Pro Gly Ala Ala Arg Pro Ala Arg Phe Ser Phe Arg Leu Arg Pro
 145 150 155 160

ES 2 368 072 T3

Val Phe Asn Ala Ser Val Gln Phe Leu His Cys Gln Leu Ser Arg Cys
 165 170 175
 Arg Arg Leu Arg Gly Val Arg Arg Ala Pro Ala Pro Leu Thr Pro Pro
 180 185 190
 Pro Pro Pro Pro Pro Ser Arg Cys Leu Pro Gln Asp Glu Ala Cys Ala
 195 200 205
 Asp Thr Gly Ser Gly Ser Ala Glu Gly Leu Ala Ala Asp Gly Pro His
 210 215 220
 Leu His Thr Leu Thr Gln Pro Ile Val Val Thr Val Pro Arg Pro Pro
 225 230 235 240
 Pro Arg Pro Pro Lys Ser Val Pro Gly Arg Ala Val Arg Pro Glu Pro
 245 250 255
 Pro Ala Pro Ala Pro Ala Ala Leu Glu Pro Ala Pro Val Val Ala Leu
 260 265 270
 Val Leu Ala Ala Phe Val Leu Gly Ala Ala Leu Ala Ala Gly Leu Gly
 275 280 285
 Leu Val Cys Ala His Ser Ala Pro His Ala Pro Gly Pro Pro Ala Arg
 290 295 300
 Ala Ser Pro Ser Gly Pro Gln Pro Arg Arg Ser Gln
 305 310 315

REIVINDICACIONES

1. Un polipéptido que comprende la secuencia de aminoácidos tal como se describe en SEQ ID NO:8.
2. Un polipéptido de acuerdo con la reivindicación 1 que consiste en la secuencia de aminoácidos tal como se describe en SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:19 o SEQ ID NO:21.
3. Un polipéptido que comprende SEQ ID NO:10.
4. Un polipéptido de acuerdo con la reivindicación 3 que consiste en la secuencia de aminoácidos tal como se describe en SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:19 o SEQ ID NO:21.
5. Un polipéptido que es un fragmento o un equivalente funcional tal como se describe en cualquiera de las reivindicaciones precedentes, que tiene más de 80% de identidad de secuencia con la secuencia de aminoácidos descrita en SEQ ID NO:8, preferiblemente más de 85%, 90%, 95%, 98% o 99% de identidad de secuencia, y contiene un subdominio ZP/TGR3 como se encuentra en cualquiera de las secuencias de aminoácidos codificadas por SEQ ID NO:10, SEQ ID NO:12 o SEQ ID NO:14.
6. Una proteína de fusión que comprende el polipéptido de acuerdo con cualquiera de las reivindicaciones precedentes.
7. El polipéptido de la reivindicación 6, en donde dicho polipéptido comprende una etiqueta de histidina.
8. El polipéptido de la reivindicación 7, cuya secuencia se describe en SEQ ID NO: 19 o SEQ ID NO:21.
9. El polipéptido de cualquiera de las reivindicaciones precedentes, en donde dicho polipéptido comprende un péptido señal.
10. El polipéptido de acuerdo con la reivindicación 9, cuya secuencia se describe en SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:12, SEQ ID NO:19 o SEQ ID NO:21.
11. Un vector que comprende una molécula de ácido nucleico que codifica un polipéptido de acuerdo con cualquiera de las reivindicaciones 1 a 4 o 6 a 10.
12. Un vector de acuerdo con la reivindicación 11 que comprende la secuencia de ácido nucleico tal como se describe en SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18 o SEQ ID NO:20.
13. Un vector de acuerdo con la reivindicación 12, en donde la molécula de ácido nucleico consiste en la secuencia de ácido nucleico tal como se describe en SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18 o SEQ ID NO:20.
14. Una célula huésped transformada con un vector de acuerdo con cualquiera de las reivindicaciones 11 a 13.
15. Un anticuerpo que se une específicamente al subdominio ZP/TGR3 como se encuentra en cualquiera de las secuencias de aminoácidos codificadas por SEQ ID NO:10, SEQ ID NO:12 o SEQ ID NO:14 del polipéptido de acuerdo con cualquiera de las reivindicaciones 1 a 10.
16. Un polipéptido de acuerdo con cualquiera de las reivindicaciones 1 a 10, un vector de acuerdo con cualquiera de las reivindicaciones 11 a 13, una célula huésped de acuerdo con la reivindicación 14, un anticuerpo de acuerdo con la reivindicación 15, para uso en terapia o diagnóstico de enfermedades.
17. Un polipéptido de acuerdo con cualquiera de las reivindicaciones 1 a 10 para uso en la unión con una proteína que pertenece a la superfamilia de proteínas TGF-beta.

Figura 1

1 MLGTVLLLLALLPGITTLPSGPPAPFFPAAPGPWLRRLFSLKLSDTEDVFPRRAGLEVF 60

61 ADSRVFVQAALARPSRWGLALHRC SVTPSSRPAPGPALALLREGCPADTSVAFPPPPPP 120

121 SPGAARPARFSFRLRPVFASVQFLHCQLSRCRRLRGVRRAPAPLTPPPPPPSRCLPQD 180

181 EACADTGSGSAEGLAADGPHLHTLTQPIVVTVPRPPRPPKSVPGRAVRPEPPAPAPAAL 240

241 EPAPVVALVLA AFVLGAALAAGLGLVCAHSAPHAGPPPARASPSGPQPRRSQ 292

Figura 2

INSP215	1	MLGTVLLLALLPGITTLPSGPP	A:A[gct]	P
hg16_genómico	743	acgagcccgcggcgaatcagccGGTATGTA	Intrón 1	CAGCTc
		ttgcttttcttctgctcctcgccc	<1-----[810 : 2292]-1>	c
		ggccggggcgccagcccaccgat		c
INSP215	25	FFPAAPGFWLRRLFLSLKLSDEDVFPRRAGPLEVPADSRVAVQ		
hg16_genómico	2298	ctcggcgctccacctacactgaggggtcccggcggcgacgtgc		
		ctccccggtggcttgtatcacaattcggcgctatccaggtta		
		gccggcccggcaccggccagccttcggggcgccccggcgg		
INSP215	69		AALARPSRWGLALHRCVTPSSRPA	
hg16_genómico	2430	GTGGGGA	Intrón 2	CAGgggtgcctcctgcgcccttgacttcgg
		<0-----[2430 : 2507]-0>	cctcgcccgggtctaggtccccggc	gcgtccgcgcgcgccagggcaccg
INSP215	95	PGPALALLREGCPADTSVAFFPPPPPSGAARPARFSFRLRPVFNASVQ		
hg16_genómico	2586	cgcgcgcccgggtcggatgggtccccccacgggcccgtatccccgtagtg		
		cgctcttgaggccacctctcccccgcgcccggctgtgtgcttaccta		
		ggccgtgggtgccccctcccagggggcgtccccgtccccgcgccccggg		
INSP215	144	FLHCQLSRCRRLRGVRRAPAPLTPPPPPPPSR		
hg16_genómico	2733	tcttccactccccggcgcgcgccacccccctc		
		ttagatgggggtgggtggccctcccccccccg		
		cgccggccccccgaccgggtgtggggggggagg		
INSP215	176		CLPQDEACADTSGSASAEGLAADGPHL	
hg16_genómico	2829	GTGCGCG	Intrón 3	CAGTccccgggtggagagagggcggggccc
		<0-----[2829 : 3358]-0>	gtcaaacgcacggggcagtcagcat	tgtgcgccctctcccgcgtccccg
INSP215	202	HTLTQPIVVTVPRPPP	R:R[agg]	PKSVPG
hg16_genómico	3437	cacaccaggagccccAGTGAGCA	Intrón 4	CAGGGccaagcg
		actcactttctcgccc	<1-----[3486 : 3586]-1>	ccagtcg
		cggggtcgccggggcc		gcgtccc
INSP215	226	RAVRPEPPAPAPAALEPAPVVALVLAFLVLAAGLGLVCAHS		
hg16_genómico	3610	cgcccgccgcgcggcgcgcgggcggtgggtcgggcgggcgcggtgct		
		gctgcacccccctacccttcttctcttgcctccgtggtgca		
		tagctgtcgccgcgacgggggggaccggcggcgcggtcctgca		
INSP215	271		-:A[gcg]	PHAPGPPARASPSGPQP
hg16_genómico	3745	GGTACCGA	Intrón 5	CAGCGccgcgcggagtcagccc
		<1-----[3746 : 3942]-1>	caccgcccggcgccac	ccctcgcgacgcctcgc
INSP215	289	RRSQ		
hg16_genómico	3996	aatc		
		ggca		
		ggcg		

Figura 3

```

INSP215_39-98      FSLKLSDTEDVFPRRAGPLEVPADSRVVFVQAALARPSRWGLALHRChSVTPSSRPAPGPA
HTGFR3_591-650    FNMELYNTDLEFLVPSQGVFSVPENGHVYVEVSVTKAEQELGFAIQTChFISPYSNPDRMSH
                   * hp* p*p hh      * h ** p p*h*phhhpp    p *h*hp * h * *p*

INSP215_99-154    LALLREGhChPADTSVAFPPP-----PPSPGAARPARFSFRLRPVENASVQFLHhQLSRhCRR
HTGFR3_651-710    YTIENIChPKDESVKFYSEKRVHFPPIQADMDKKRFSEVFKPVENTSLLFLOhELThLThTK
                   hhpp ** * ** * *      *      **** hp*** *h **p*p* * p

INSP215_155-183   LRGVRRAPALTPPPPPPPSRhCLPQDEACh
HTGFR3_711-728    MEKH-----PQKLPKhCVPPDEACh
                   hp      *      p*h* ****
    
```


Figura 4 Cont.

```

901  tgcgcactca gcgccccacg ccctggccc gcccgcgaga gcctcgcca gcggtccca
    c a h s a p h a p g p p a r a s p s g p
961  gcccaggagg tcccagtgag gaagggatgg tgcgccccca acatgggccg gagatacacc
    q p r r s q
    ←
    INSP215-R1
    ←
    INSP215-R1nest
1021 cagctaccaa ttcgggacca ggaccaacag gaccggacc gcctccctgg acctcggacc
1081 tgatgaggcc acgaccctg cgtttctctc ctcccctgt cctcccacc tgtgctcaa
1141 ataaacctct ggactgaccg gct

```

Dominio transmembrana previsto

Posición y sentido de cebadores de PCR →

Figura 5

```

1  ATGCTGGGCA CCGTGCTGCT GCTGGCCCTG CTCCCAGGGA TCACCACCTT ACCCAGCGGG
   M L G T V L L L A L L P G I T T L P S G
61  CCACCTGCTC CCCCGTTCCC CGCGGCGCCC GGCCCCTGGC TGCGCAGACC CCTCTTCAGC
   P P A P P F P A A P G P W L R R P L F S
121 CTGAAGCTGT CCGACACAGA GGACGTCTTT CCTCGCCGCG CGGGGCCGCT CGAGGTCCCCG
   L K L S D T E D V F P R R A G P L E V P
181 GCCGACAGCC GCGTGTTTCGT GCAGGCGGGC TTGGCCCGTC CCTCCCCGCG CTGGGGCCTG
   A D S R V F V Q A A L A R P S P R W G L
241 GCCCTGCACC GCTGCTCAGT GACGCCGTCC TCACGCCCGG CCCCggggcc CGCCCTGGCT
   A L H R C S V T P S S R P A P G P A L A
301 CTGCTGCGTG AGGGCTGCCC CGCCGACACC TCTGTGCGCT TCCCGCCACC GCCGCCGCCG
   L L R E G C P A D T S V A F P P P P P P
361 AGCCCCGGTG CCGCCCCGCC CGCGCGTTTC AGCTTCCGCC TGCGCCCCGT CTTCAACGCC
   S P G A A R P A R F S F R L R P V F N A
421 TCGGTGCAGT TCCTGCACTG CCAGCTGAGC CGCTGCCGCC GCCTCCGGGG AGTCCGCCGG
   S V Q F L H C Q L S R C R R L R G V R R
481 GCGCCTGCGC CTCTGACGCC GCCGCGCCG CCGCCGCCAT CGCGGTGTCT GCCTCAGGAC
   A P A P L T P P P P P P P S R C L P Q D
541 GAGGCGTGCG CCGACACTGG CAGTGGCAGC GCCGAGGGCC TGGCTGCTGA CGGCCCCCAC
   E A C A D T G S G S A E G L A A D G P H
601 CTGCACACGC TGACGCAGCC TATCGTGGTC ACCGTGCCGC GGCCGCCCC CAGGCCGCC
   L H T L T Q P I V V T V P R P P P R P P
661 AAGAGTGTCC CCGGCCGTGC AGTGCGCCCT GAGCCTCCCG CGCCGGCCCC CGCGGCCCTG
   K S V P G R A V R P E P P A P A P A A L
721 GAACCCGCGC CGGTGGTGGC GCTGGTGTG GCAGCCTTCG TGCTGGGCGC CGCGCTGGCC
   E P A P V V A L V L A A F V L G A A L A
781 GCCGGGCTGG GTCTCGTCTG TGCGCACTCA GCGCCCCACG CCCCTGGCCC GCCCGCGAGA
   A G L G L V C A H S A P H A P G P P A R
841 GCCTCGCCCA GCGGTCCCCA GCCAGGAGG TCCCAGCACC ATCACCATCA CCAT
   A S P S G P Q P R R S Q H H H H H H

```