

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 381 766**

51 Int. Cl.:
C07K 14/82 (2006.01)
G01N 33/48 (2006.01)
G01N 33/50 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **07746891 .6**
96 Fecha de presentación: **13.06.2007**
97 Número de publicación de la solicitud: **2024500**
97 Fecha de publicación de la solicitud: **18.02.2009**

54 Título: **Familia génica (LBFL313) asociada con cáncer pancreático**

30 Prioridad:
14.06.2006 KR 20060053629

45 Fecha de publicación de la mención BOPI:
31.05.2012

45 Fecha de la publicación del folleto de la patente:
31.05.2012

73 Titular/es:
LG LIFE SCIENCES, LTD.
20, YOIDO-DONG, YOUNGDUNGPO-GU
SEOUL 150-721, KR

72 Inventor/es:
KOH, Sang-seok;
SONG, Si-young;
KIM, Sun-a;
LEE, Yang-soon;
JEON, Sun-bok;
PARK, Eui-chul y
KIM, Young-gun

74 Agente/Representante:
Zuazo Araluze, Alexander

ES 2 381 766 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

Familia génica (Ibfl313) asociada con cáncer pancreático.

5 **Campo de la invención**

La presente invención se refiere a los cambios en la expresión génica en tejidos de cáncer pancreático de pacientes con cáncer pancreático. La invención se refiere específicamente a un gen humano que se expresa de manera diferente en tejidos de cáncer pancreático en comparación con tejidos pancreáticos normales correspondientes, y en otros neoplasmas malignos.

10 **Antecedentes de la técnica**

El cáncer pancreático, la cuarta causa principal de mortalidad por cáncer tanto en hombres como en mujeres, es un problema de salud importante en el mundo desarrollado y está asociado con un pronóstico extremadamente malo (Faint *et al.* (2004) *Datamonitor DMHC2045*; Garcea *et al.* (2005) *Pancreatology* 5:514-529; Kern *et al.* (2002) *Cancer Biol Therapy* 1:607-613; Laheru y Jaffee (2005) *Nature Rev Cancer* 5: 59-467; Li *et al.* (2004) *Lancet* 363:1049-1057). Se ha estimado que aproximadamente 30.000 estadounidenses desarrollan y mueren de esta enfermedad al año. A pesar del manejo médico y quirúrgico agresivo, la esperanza de vida media es de aproximadamente 15~18 meses para pacientes con enfermedad local y regional, y de 3~6 meses para pacientes con enfermedad metastásica. Cerca del 100% de pacientes con cáncer pancreático desarrollan metástasis y mueren debido a los efectos metabólicos debilitantes de su crecimiento sin restricción y la supervivencia a los cinco años global para los grupos de pacientes que no se someten a procedimientos de resección es claramente inferior al cinco por ciento. Es particularmente agresivo con síntomas iniciales no específicos y pronóstico temprano difícil. Están en desarrollo métodos de detección temprana de cáncer pancreático pero aún no existen en práctica y las terapias contra el cáncer convencionales tienen poco impacto sobre el pronóstico o desenlace de la enfermedad. El mal pronóstico del cáncer pancreático puede atribuirse a su tendencia a una presentación tardía, invasión local agresiva, metástasis temprana y escasa respuesta a la quimioterapia.

Como muchas otras enfermedades malignas, el cáncer pancreático resulta de la acumulación de mutaciones adquiridas. Cambios epigenéticos y genéticos múltiples, incluyendo activación de protooncogenes, inactivación de genes supresores de tumores y anomalías de genes de mantenimiento, están implicados en el desarrollo, crecimiento continuado y metástasis del cáncer pancreático. Se cree que las mutaciones acumuladas en tales genes se producen en un transcurso de tiempo predecible durante los estadios de "PanIN" (neoplasia intraepitelial pancreática ("Pancreatic Intraepithelial Neoplasia")) (Hruban *et al.* (2000) *Clin Cancer Res* 6:2969-2972; Kern *et al.* (2002) *Cancer Biol Therapy* 1:607-613; Li *et al.* (2004) *Lancet* 363:1049-1057). La mutación del K-ras se produce a la mitad del PanIN-1. El estadio PanIN-2 está marcado por cambios adicionales y el aumento en la tasa de mutaciones de K-ras y por la aparición de numerosas anomalías de p16, y la sobreexpresión de la proteína p53, que puede indicar la presencia de mutaciones de p53, aparece ocasionalmente en los PanIN más avanzados. La pérdida de genes supresores de tumores, TP53, DPC4 y BRCA2, parece producirse tarde en el desarrollo de neoplasia pancreática, PanIN-3.

Más del 85% de los cánceres ductales pancreáticos tienen una mutación puntual activante del gen K-ras en el desarrollo del cáncer pancreático (Li *et al.* (2004) *Lancet* 363:1049-1057; Xiong (2004) *Cancer Chem Pharm* 54:S69-77). La mutación de K-ras da como resultado activación constitutiva de una ruta de señalización intracelular, Ras-Raf-MEK-ERK, que conduce a proliferación celular y confiere por tanto propiedades transformantes a las células que contienen mutaciones puntuales en este gen. La mutación de Ras no está asociada con el pronóstico o estadio tumoral, indicando que el oncogen K-ras puede estar relacionado con el inicio de la carcinogénesis, pero no está vinculado a potencial maligno o promoción de cáncer pancreático humano. Una de las dianas en el sentido de 3' clave de la familia Ras es fosfoinositol 3 cinasa (PI3K). La activación de PI3K está implicada en la resistencia del cáncer pancreático a la apoptosis inducida por agentes de direccionamiento molecular o quimioterápicos.

La inactivación del gen supresor de tumores p16 parece producirse poco después. El gen p16 está inactivado en prácticamente todos los adenocarcinomas ductales por mutación, delección homocigota o silenciamiento transcripcional asociado con metilación del promotor (Kern *et al.* (2002) *Cancer Biol Therapy* 1:607-613; Maitra *et al.* (2006) *Best Pract Res Clin Gastroenterol* 20:211-226). La proteína p16 regula el ciclo celular a través de la ruta p16/Rb, por tanto la inactivación genética del gen p16 significa que se pierde un regulador crítico del ciclo celular en el cáncer pancreático. De interés, las mutaciones heredadas en el gen p16 son una causa del síndrome de melanoma familiar atípico con moluscos múltiples (FAMMM), y los pacientes con FAMMM tienen un mayor riesgo de desarrollar melanoma y cáncer pancreático.

La inactivación del gen TP53 se produce casi siempre por una mutación intragénica en un alelo acoplada con la pérdida del segundo alelo (Maitra *et al.* (2006) *Best Pract Res Clin Gastroenterol* 20:211-226). El mal funcionamiento de p53 significa que dos controles críticos del número celular, el punto de control del ciclo celular G1/S y el mantenimiento de la detención en G2/M, están desregulados en la mayoría de cánceres pancreáticos.

El gen DPC4, también conocido como SMAD4, está genéticamente inactivado en más de la mitad de los cánceres pancreáticos, en el 35% por delección homocigota, y en el 20% por una mutación intragénica acoplada con la pérdida del alelo restante (Maitra *et al.* (2006) *Best Pract Res Clin Gastroenterol* 20:211-226; Wilentz *et al.* (2000) *Am J Pathol* 156:37-43). Sin embargo, la inactivación genética de DPC4 sólo aparece rara vez en otros tipos de tumor. La proteína dpc4 desempeña un papel crítico en la señalización y el control del crecimiento a través de la ruta de TGF- β .

El gen BRCA2, relacionado con la reparación del ADN, sólo se selecciona como diana en un porcentaje pequeño (~10%) de cáncer pancreático, pero provoca la agregación familiar del cáncer pancreático (Maitra *et al.* (2006) *Best Pract Res Clin Gastroenterol* 20:211-226; Murphy *et al.* (2002) *Cancer Res* 62:3789-3793). Portadores de un único par de bases del gen BRCA2 (denominado la mutación del gen BRCA26174delT) tienen un riesgo 10 veces mayor de desarrollar cáncer pancreático.

El factor nuclear κ B (NF- κ B), un factor de transcripción que existe predominantemente como heterodímero p65 (RelA)/p50, también se considera como uno de los genes relacionados con cáncer pancreático (Garcea *et al.* (2005) *Pancreatology* 5:514-529; Xiong (2004) *Cancer Chem Pharm* 54:S69-77). RelA, la subunidad p65 de NF- κ B, se activa constitutivamente en aproximadamente el 67% de los adenocarcinomas pancreáticos, pero no en tejidos pancreáticos sanos, y I κ B α se sobreexpresa en líneas celulares y tejidos de tumores pancreáticos humanos. La activación constitutiva de RelA parece estar correlacionada con la ruta de señalización en el sentido de 5', tal como Ras, en células tumorales pancreáticas. Se ha sugerido que NF- κ B desempeña un papel importante en la resistencia del tumor a la apoptosis inducida por agentes citotóxicos en el cáncer pancreático. Otros resultados también mencionan que el mecanismo principal por el que NF- κ B parece promover el crecimiento de células pancreáticas es mediante la inhibición de la apoptosis.

La entrada (n.º de registro BC009722) de la base de datos EMBL se refiere a un polinucleótido y una secuencia de polipéptido traducida correspondiente idéntica a SEQ ID NO. 1 y 2. No da a conocer el uso de dichas secuencias para el diagnóstico o tratamiento de cánceres pancreáticos.

El documento EP 1 526 381 A da a conocer un método de diagnóstico para detectar adenocarcinoma pancreático usando la proteína DDEF2, un factor de potenciamiento de diferenciación y desarrollo. No da a conocer el uso de las presentes secuencias para el diagnóstico o tratamiento de cánceres pancreáticos.

Sigue habiendo una necesidad en la técnica de materiales y métodos que permitan un diagnóstico más preciso del adenocarcinoma pancreático. Además, sigue habiendo una necesidad en la técnica de métodos para tratar y métodos para identificar agentes que puedan tratar eficazmente esta enfermedad. La presente invención cumple estas y otras necesidades.

Descripción

Problema técnico

La presente invención proporciona un material y método para un diagnóstico de manera exacta de adenocarcinoma pancreático.

La presente invención también proporciona material para su tratamiento que trata de manera eficaz el adenocarcinoma pancreático.

Solución técnica

La presente invención se basa en un gen nuevo (a continuación en el documento denominado "LBFL313") que se expresa de manera diferente en tejidos de adenocarcinoma pancreático en comparación con tejidos pancreáticos normales. La invención proporciona una composición para diagnosticar cáncer pancreático, que comprende (a) una molécula de ácido nucleico aislada que comprende SEQ ID NO: 1, (b) una molécula de ácido nucleico aislada que codifica para SEQ ID NO: 2, (c) una molécula de ácido nucleico aislada que presenta una identidad de secuencia de nucleótidos de al menos el 95% a lo largo de la secuencia contigua completa de SEQ ID NO: 1 y (d) una molécula de ácido nucleico aislada que comprende el complemento de una molécula de ácido nucleico de (a), (b) o (c).

La presente invención proporciona además una composición para diagnosticar cáncer pancreático, que comprende las moléculas de ácido nucleico operativamente unidas a uno o más elementos de control de la expresión, incluyendo vectores que comprenden las moléculas de ácido nucleico aisladas. La invención proporciona además una composición para diagnosticar cáncer pancreático, que comprende células huésped transformadas para contener las moléculas de ácido nucleico de la invención y métodos para producir una proteína que comprende la etapa de cultivar una célula huésped transformada con una molécula de ácido nucleico de la invención en condiciones en las que la proteína se expresa.

La invención proporciona además una composición para diagnosticar cáncer pancreático, que comprende una proteína o polipéptido aislado que comprende la secuencia de aminoácidos de SEQ ID NO: 2 o que presenta una identidad de secuencia de aminoácidos de al menos el 95% con SEQ ID NO: 2.

5 A. Las proteínas asociadas al cáncer pancreático

La presente invención proporciona proteínas aisladas, variantes alélicas de las proteínas y sustituciones de aminoácido conservativas de las proteínas. Tal como se usa en el presente documento, la "proteína" o el "polipéptido" se refiere, en parte, a una proteína que tiene la secuencia de aminoácidos humana representada en SEQ ID NO: 2. Los términos también se refieren a variantes alélicas y proteínas que se producen de manera natural que tienen una secuencia de aminoácidos ligeramente diferente de la mencionada de manera específica anteriormente. Las variantes alélicas, aunque poseen una secuencia de aminoácidos ligeramente diferente de las mencionadas anteriormente, tendrán todavía las mismas o similares funciones biológicas asociadas con estas proteínas.

Tal como se usa en el presente documento, la familia de proteínas relacionada con la secuencia de aminoácidos humana de SEQ ID NO: 2 se refiere a proteínas que se han aislado de organismos además de seres humanos. Los métodos usados para identificar y aislar otros miembros de la familia de proteínas relacionados con estas proteínas se describen a continuación.

Las proteínas de la presente invención están preferiblemente en forma aislada. Tal como se usa en el presente documento, se dice que una proteína está aislada cuando se emplean métodos físicos, mecánicos o químicos para separar la proteína de los constituyentes celulares que están normalmente asociados con la proteína. Un experto puede emplear fácilmente métodos de purificación convencionales para obtener una proteína aislada.

Las proteínas de la presente invención incluyen además variantes de sustitución de aminoácidos conservativa, por delección o inserción de SEQ ID NO: 2. Tal como se usa en el presente documento, una variante conservativa se refiere a alteraciones en la secuencia de aminoácidos que no afectan adversamente a las funciones biológicas de la proteína. Una sustitución, inserción o delección se dice que afecta adversamente a la proteína cuando la secuencia alterada impide o altera una función biológica asociada con la proteína. Por ejemplo, las propiedades hidrófobas/hidrófilas, estructura o carga globales de la proteína, en determinados casos, pueden alterarse sin afectar adversamente a una actividad biológica. Por consiguiente, la secuencia de aminoácidos puede alterarse, por ejemplo para hacer que el péptido sea más hidrófobo o hidrófilo, sin afectar adversamente a las actividades biológicas de la proteína.

Generalmente, las variantes alélicas, las variantes de sustitución conservativa y los miembros de la familia de proteínas, tendrán una secuencia de aminoácidos que tiene una identidad de aminoácidos de al menos aproximadamente el 50%, el 60%, el 70% o el 75% con la secuencia expuesta en SEQ ID NO: 2, más preferiblemente al menos aproximadamente el 80-90%, incluso más preferiblemente al menos aproximadamente el 92-94%, y lo más preferiblemente al menos aproximadamente el 95%, el 98% o el 99% de identidad de secuencia. La identidad u homología con respecto a tales secuencias se define en el presente documento como el porcentaje de residuos de aminoácido en la secuencia candidata que son idénticos a SEQ ID NO: 2, tras alinear las secuencias e introducir huecos, si es necesario, para lograr el porcentaje de homología máximo, y sin considerar ninguna sustitución conservativa como parte de la identidad de secuencia (véase la sección B para los parámetros relevantes). No debe interpretarse que proteínas de fusión, o extensiones, delecciones o inserciones internas, C-terminales o N-terminales en la secuencia del péptido afectan a la homología.

Por tanto, las proteínas dadas a conocer en el presente documento incluyen moléculas que tienen la secuencia de aminoácidos dada a conocer en SEQ ID NO: 2; fragmentos de la misma que tienen una secuencia consecutiva de al menos aproximadamente 3, 4, 5, 6, 10, 15, 20, 25, 30, 35 o más residuos de aminoácido de estas proteínas; variantes de secuencia de aminoácidos en las que uno o más residuos de aminoácido se han insertado de manera N- o C-terminal en, o dentro de, la secuencia codificante dada a conocer; y variantes de secuencia de aminoácidos de la secuencia dada a conocer, o sus fragmentos tal como se definieron anteriormente, que se han sustituido con al menos un residuo. Tales fragmentos, también denominados péptidos o polipéptidos, pueden contener regiones antigénicas, regiones funcionales de la proteína identificadas como regiones de la secuencia de aminoácidos que corresponden a dominios de proteína conocidos, así como regiones de hidrofobicidad pronunciada. Las regiones pueden identificarse fácilmente usando software de análisis de la secuencia de proteínas comúnmente disponible tal como MacVector (Oxford Molecular).

Las variantes contempladas incluyen además las que contienen mutaciones predeterminadas mediante, por ejemplo, recombinación homóloga, mutagénesis por PCR o dirigida al sitio, y las correspondientes proteínas de otra especie animal, incluyendo pero sin limitarse a de conejo, de ratón, de rata, porcina, bovina, ovina, equina y especies de primate no humano, y los alelos u otras variantes que se producen de manera natural de la familia de proteínas; y derivados en los que la proteína se ha modificado covalentemente por medios de sustitución, químicos, enzimáticos u otros medios apropiados con un resto distinto de un aminoácido que se produce de manera natural (por ejemplo un resto detectable tal como una enzima o un radioisótopo).

La presente invención proporciona además composiciones que comprenden una proteína o polipéptido de la invención y un diluyente. Diluyentes adecuados pueden ser disolventes acuosos o no acuosos o una combinación de los mismos, y pueden comprender componentes adicionales, por ejemplo sales solubles en agua o glicerol, que contribuyen a la estabilidad, solubilidad, actividad y/o almacenamiento de la proteína o polipéptido.

Tal como se describe a continuación, pueden usarse miembros de la familia de proteínas: (1) para identificar agentes que modulan el nivel de o al menos una actividad de la proteína, (2) para identificar parejas de unión para la proteína, (3) como antígeno para generar anticuerpos monoclonales o policlonales, (4) como agente terapéutico o diana y (5) como agente de diagnóstico o marcador de cáncer pancreático y otras enfermedades hiperplásicas.

B. Moléculas de ácido nucleico

La presente invención proporciona además moléculas de ácido nucleico que codifican para la proteína que tiene SEQ ID NO: 2 y las proteínas relacionadas descritas en el presente documento, preferiblemente en forma aislada. Tal como se usa en el presente documento, "ácido nucleico" se define como ARN o ADN que codifica para una proteína o péptido tal como se definió anteriormente, es complementario a una secuencia de ácido nucleico que codifica para tales péptidos, se hibrida con el ácido nucleico de SEQ ID NO: 1 y permanece unido de manera estable a la misma en condiciones de rigurosidad apropiadas, codifica para un polipéptido que comparte al menos aproximadamente el 50%, el 60%, el 70% o el 75%, preferiblemente al menos aproximadamente el 80-90%, más preferiblemente al menos aproximadamente el 92-94%, y lo más preferiblemente al menos aproximadamente el 95%, el 98%, el 99% o más identidad con la secuencia de péptido de SEQ ID NO: 2 o presenta al menos el 50%, el 60%, el 70% o el 75%, preferiblemente al menos aproximadamente el 80-90%, más preferiblemente al menos aproximadamente el 92-94%, e incluso más preferiblemente al menos aproximadamente el 95%, el 98%, el 99% o más identidad de secuencia de nucleótidos a lo largo de los marcos de lectura abiertos de SEQ ID NO: 1.

La presente invención da a conocer moléculas de ácido nucleico aisladas que se hibridan específicamente con el complemento de SEQ ID NO: 1, particularmente las moléculas que se hibridan específicamente a lo largo de los marcos de lectura abiertos. Tales moléculas que se hibridan específicamente con el complemento de SEQ ID NO: 1 normalmente lo hacen en condiciones de hibridación rigurosas.

Se contemplan específicamente ADN genómico, ADNc, ARNm y moléculas antisentido, así como ácidos nucleicos basados en estructuras principales alternativas o que incluyen bases alternativas, ya se deriven de fuentes naturales o se sintetizen. Tales ácidos nucleicos complementarios o de hibridación, sin embargo, se definen adicionalmente como novedosos y no obvios con respecto a cualquier ácido nucleico de la técnica anterior incluyendo los que codifican para, se hibridan en condiciones de rigurosidad apropiadas o son complementarios al ácido nucleico que codifica para una proteína según la presente invención.

La homología o identidad al nivel de secuencia de aminoácidos o nucleótidos se determina mediante análisis BLAST (Basic Local Alignment Search Tool) usando el algoritmo empleado mediante los programas blastp, blastn, blastx, tblastn y tblastx (Altschul *et al.* (1997), *Nucleic Acids Res.* 25: 3389-3402, y Karlin *et al.* (1990), *Proc. Natl. Acad. Sci. USA* 87: 2264-2268, ambos completamente incorporados como referencia) que están diseñados para búsquedas de similitud de secuencias. El enfoque usado por el programa BLAST es en primer lugar considerar segmentos similares, con y sin huecos, entre una secuencia de consulta y una secuencia de la base de datos, luego evaluar la significación estadística de todas las coincidencias que se identifican y finalmente resumir sólo las coincidencias que satisfacen un umbral de significación preseleccionado. Para una discusión de temas básicos en la búsqueda de similitudes de bases de datos de secuencias, véase Altschul *et al.* (1994), *Nat. Genet.* 6: 119-129 que se incorpora completamente como referencia. Los parámetros de búsqueda para histogramas, descripciones, alineaciones, expectativas (es decir, el umbral de significación estadística para notificar coincidencias frente a secuencias de bases de datos), punto de corte, matriz y filtro (baja complejidad) están en los parámetros por defecto. La matriz de puntuación por defecto usada por blastp, blastx, tblastn, y tblastx es la matriz BLOSUM62 (Henikoff *et al.* (1992), *Proc. Natl. Acad. Sci. USA* 89: 10915-10919, completamente incorporada como referencia), recomendada para secuencias de consulta de más de 85 nucleótidos o aminoácidos de longitud.

Para blastn, la matriz de puntuación se fija mediante las razones de M (es decir, la puntuación de recompensa por un par de residuos coincidentes) con respecto a N (es decir, la puntuación de penalización para residuos no coincidentes), en la que los valores por defecto para M y N are 5 y -4, respectivamente. Se ajustaron cuatro parámetros de blastn tal como sigue: Q=10 (penalización por creación de hueco); R=10 (penalización por extensión de hueco); wink=1 (genera aciertos de palabra en cada posición de wink^o a lo largo de la consulta); y gapw=16 (fija el ancho de la ventana dentro de la que se generan alineaciones con huecos). Los parámetros de Blastp equivalentes fueron Q=9; R=2; wink=1; y gapw=32. Una comparación Bestfit entre secuencias, disponible en el paquete GCG versión 10.0, usa los parámetros de ADN GAP=50 (penalización por creación de hueco) y LEN=3 (penalización por extensión de hueco) y los parámetros equivalentes en comparaciones de proteínas son GAP=8 y LEN=2.

"Condiciones rigurosas" son aquellas que (1) emplean baja fuerza iónica y alta temperatura para el lavado, por ejemplo, NaCl 0,015 M/citrato de sodio 0,0015 M/SDS al 0,1% a 50°C, o (2) emplean durante la hibridación un

agente desnaturalizante tal como formamida, por ejemplo, formamida al 50% (vol/vol) con albúmina sérica bovina al 0,1%/Ficoll al 0,1%/polivinilpirrolidona al 0,1%/tampón fosfato de sodio 50 mM a pH 6,5 con NaCl 750 mM, citrato de sodio 75 mM a 42°C. Otro ejemplo es hibridación en formamida al 50%, 5x SSC (NaCl 0,75 M, citrato de sodio 0,075 M), fosfato de sodio 50 mM (pH 6,8), pirofosfato de sodio al 0,1%, disolución de Denhardt 5x, ADN de esperma de salmón sonicado (50 µg/ml), SDS al 0,1% y sulfato de dextrano al 10% a 42°C, con lavados a 42°C en 0,2x SSC y SDS al 0,1%. Un experto puede determinar fácilmente y variar las condiciones de rigurosidad apropiadamente para obtener una señal de hibridación detectable y clara. Moléculas preferidas son aquellas que se hibridan en las condiciones anteriores con el complemento de SEQ ID NO: 1 y que codifican para una proteína de longitud completa o funcional. Moléculas incluso más preferidas son aquellas que se hibridan en las condiciones anteriores con la hebra complemento del marco de lectura abierto de SEQ ID NO: 1.

Tal como se usa en el presente documento, una molécula de ácido nucleico se dice que está "aislada" cuando la molécula de ácido nucleico está sustancialmente separada de moléculas de ácido nucleico contaminantes que codifican para otros polipéptidos.

La presente invención proporciona además fragmentos de las moléculas de ácido nucleico dadas a conocer. Tal como se usa en el presente documento, un fragmento de una molécula de ácido nucleico se refiere a una pequeña parte de la secuencia codificante o no codificante. El tamaño del fragmento se determinará mediante el uso previsto. Por ejemplo, si el fragmento se escoge de manera que codifica para una parte activa de la proteína, el fragmento necesitará ser lo suficientemente largo como para codificar para la región/regiones funcional(es) de la proteína. Por ejemplo, pueden prepararse fragmentos que codifican para péptidos correspondientes a regiones antigénicas pronosticadas. Si el fragmento va a usarse como sonda de ácido nucleico o cebador de PCR, entonces la longitud del fragmento se elige de manera que se obtiene un número relativamente pequeño de falsos positivos durante la exploración con sondas/cebado (véase la discusión en la sección G).

Pueden sintetizarse fácilmente fragmentos de las moléculas de ácido nucleico de la presente invención (es decir, oligonucleótidos sintéticos) que se usan como sondas o cebadores específicos para la reacción en cadena de la polimerasa (PCR), o para sintetizar secuencias de genes que codifican para proteínas de la invención, mediante técnicas químicas, por ejemplo, el método de fosforamidita de Matteucci *et al.*, ((1981) J. Am. Chem. Soc. 103: 3185-3191) o usando métodos de síntesis automatizados. Además, pueden prepararse fácilmente segmentos de ADN más grandes mediante métodos bien conocidos, tales como síntesis de un grupo de oligonucleótidos que definen diversos segmentos modulares del gen, seguido por ligamiento de oligonucleótidos para construir el gen modificado completo.

Las moléculas de ácido nucleico de la presente invención pueden modificarse adicionalmente para que contengan una etiqueta detectable para fines de exploración con sondas y diagnóstico. Se conoce en la técnica una variedad de tales etiquetas y pueden emplearse fácilmente con las moléculas codificantes descritas en el presente documento. Las etiquetas adecuadas incluyen, pero no se limitan a, biotina, nucleótidos marcados fluorescentemente o radiomarcados y similares. Un experto puede emplear fácilmente cualquier etiqueta de este tipo para obtener variantes marcadas de las moléculas de ácido nucleico de la invención.

C. Aislamiento de otras moléculas de ácido nucleico relacionadas

Tal como se describió anteriormente, la identificación y caracterización de la molécula de ácido nucleico que tiene SEQ ID NO: 1 permite a un experto aislar moléculas de ácido nucleico que codifican para otros miembros de la familia de proteínas además de las secuencias descritas en el presente documento. Además, las moléculas de ácido nucleico dadas a conocer en el presente documento permiten a un experto habitual aislar moléculas de ácido nucleico que codifican para otros miembros de la familia de proteínas además de las proteínas que tienen SEQ ID NO: 2.

Por ejemplo, un experto puede usar fácilmente la secuencia de aminoácidos de SEQ ID NO: 2 para generar sondas de anticuerpo para examinar bibliotecas de expresión preparadas a partir de células apropiadas. Normalmente, puede usarse antisuero policlonal de mamíferos tales como conejos inmunizados con la proteína purificada (tal como se describe a continuación) o anticuerpos monoclonales para explorar con sondas una biblioteca de expresión genómica o de ADNc de mamífero, tal como la biblioteca lambda gtlI, para obtener la secuencia codificante apropiada para otros miembros de la familia de proteínas. La secuencia de ADNc clonada puede expresarse como una proteína de fusión, expresarse directamente usando sus propias secuencias control, o expresarse mediante construcciones usando secuencias control apropiadas para el huésped particular usado para la expresión de la enzima.

Alternativamente, una parte de la secuencia codificante descrita en el presente documento puede sintetizarse y usarse como sonda para recuperar ADN que codifica para un miembro de la familia de proteínas de cualquier organismo mamífero. Se preparan oligómeros que contienen aproximadamente 18-20 nucleótidos (que codifican para un tramo de aproximadamente 6-7 aminoácidos) y se usan para examinar bibliotecas de ADN genómico o ADNc para obtener hibridación en condiciones rigurosas o condiciones de suficiente rigurosidad para eliminar un nivel excesivo de falsos positivos.

Adicionalmente, pueden prepararse pares de cebadores de oligonucleótidos para su uso en PCR para clonar selectivamente una molécula de ácido nucleico codificante. Un ciclo de desnaturalización/apareamiento/extensión de PCR para el uso de tales cebadores de PCR se conoce bien en la técnica y puede adaptarse fácilmente para su uso en el aislamiento de otras moléculas de ácido nucleico codificantes.

5 También pueden identificarse moléculas de ácido nucleico que codifican para otros miembros de la familia de proteínas en información genómica existente u otra información de secuencia usando cualquier método informático disponible, incluyendo pero sin limitarse a: PSI-BLAST (Altschul *et al.* (1997), Nucl. Acids Res. 25: 3389-3402); PHI-BLAST (Zhang *et al.* (1998), Nucl. Acids Res. 26: 3986-3990), 3DPSSM (Kelly *et al.* (2000), J. Mol. Biol. 299: 499-520); y otros métodos de análisis informáticos (Shi *et al.* (1999), Biochem. Biophys. Res. Commun. 262: 132-138 y Matsunami *et al.* (2000), Nature 404: 601-604).

D. Moléculas de ADNr que contienen una molécula de ácido nucleico

15 La presente invención proporciona además moléculas de ADN recombinante (ADNr) que contienen una secuencia codificante. Tal como se usa en el presente documento, una molécula de ADNr es una molécula de ADN que se ha sometido a manipulación molecular *in situ*. Se conocen bien en la técnica métodos para generar moléculas de ADNr, por ejemplo, véase Sambrook *et al.*, Molecular Cloning- A Laboratory Manual, tercera ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001. En las moléculas de ADNr preferidas, una secuencia de ADN codificante está operativamente unida a secuencias de control de la expresión y/o secuencias de vector.

20 La elección de secuencias de control de la expresión y/o de vector a las que una de las secuencias codificantes de la familia de proteínas de la presente invención está operativamente unida depende directamente, tal como se conoce bien en la técnica, de las propiedades funcionales deseadas, por ejemplo, expresión de proteínas, y la célula huésped que va a transformarse. Un vector contemplado por la presente invención tiene la capacidad de al menos dirigir la replicación o inserción en el cromosoma del huésped, y preferiblemente también la expresión, del gen estructural incluido en la molécula de ADNr.

30 Se conocen en la técnica elementos de control de la expresión que se usan para regular la expresión de una secuencia que codifica para la proteína operativamente unida e incluyen, pero sin limitarse a, promotores inducibles, promotores constitutivos, señales de secreción y otros elementos reguladores. Preferiblemente, el promotor inducible se controla fácilmente, tal como al ser sensible a un nutriente en el medio de la célula huésped.

35 En una realización, el vector que contiene una molécula de ácido nucleico codificante incluirá un replicón procariota, es decir, una secuencia de ADN que tiene la capacidad de dirigir el mantenimiento y la replicación autónoma de la molécula de ADN recombinante de manera extracromosómica en una célula huésped procariota, tal como una célula huésped bacteriana, transformada con la misma. Tales replicones se conocen en la técnica. Además, los vectores que incluyen un replicón procariota también pueden incluir un gen cuya expresión confiera un marcador detectable tal como resistencia a un fármaco. Genes de resistencia a fármacos bacterianos típicos son aquellos que confieren resistencia a ampicilina, kanamicina, cloranfenicol o tetraciclina.

45 Los vectores que incluyen un replicón procariota pueden incluir además un promotor procariota o de bacteriófago que tiene la capacidad de dirigir la expresión (transcripción y traducción) de las secuencias de genes codificantes en una célula huésped bacteriana, tal como *E. coli*. Un promotor es un elemento de control de la expresión formado por una secuencia de ADN que permite la unión de ARN polimerasa y que se produzca la transcripción. Se proporcionan normalmente secuencias promotoras compatibles con huéspedes bacterianos en vectores de plásmido que contienen sitios de restricción convenientes para la inserción de un segmento de ADN de la presente invención. Vectores de plásmido de este tipo típicos son pUC8, pUC9, pBR322 y pBR329 disponibles de BioRad Laboratories, (Richmond, CA), pPL y pKK223 disponibles de Pharmacia (Piscataway, NJ).

50 También pueden usarse vectores de expresión compatibles con células eucariotas, preferiblemente aquéllos compatibles con células de vertebrados, para formar moléculas de ADNr que contienen una secuencia codificante. Se conocen en la técnica vectores de expresión de células eucariotas, incluyendo vectores virales, y están disponibles de varias fuentes comerciales. Normalmente, se proporcionan tales vectores que contienen sitios de restricción convenientes para la inserción del segmento de ADN deseado. Vectores de este tipo típicos son pSVL y pKSV-10 (Pharmacia), pBPV-1/pML2d (International Biotechnologies, Inc.), pTDT1 (ATCC, n.º 31255), el vector pCDM8 descrito en el presente documento, y vectores de expresión eucariota similares. Pueden modificarse los vectores para que incluyan promotores específicos de tejido si es necesario.

60 Los vectores de expresión de células eucariotas usados para construir las moléculas de ADNr de la presente invención pueden incluir además un marcador seleccionable que es eficaz en una célula eucariota, preferiblemente un marcador de selección de resistencia a fármacos. Un marcador de resistencia a fármacos preferido es el gen cuya expresión da como resultado resistencia a neomicina, es decir, el gen de neomicina fosfotransferasa (neo) (Southern *et al.* (1982), J. Mol. Anal. Genet. 1:327-341). Alternativamente, el marcador seleccionable puede estar presente en un plásmido separado, y los dos vectores se introducen mediante cotransfección de la célula huésped, y se selecciona cultivando en el fármaco apropiado para el marcador seleccionable.

E. Células huésped que contienen una molécula de ácido nucleico codificante suministrada de manera exógena

La presente invención proporciona además células huésped transformadas con una molécula de ácido nucleico que codifica para una proteína de la presente invención. La célula huésped puede ser o bien procarionta o bien eucariota. Las células eucariotas útiles para la expresión de una proteína de la invención no están limitadas, siempre que la línea celular sea compatible con métodos de cultivo celular y compatible con la propagación del vector de expresión y la expresión del producto génico. Las células huésped eucariotas preferidas incluyen, pero sin limitarse a, células de levadura, de insecto y de mamífero, preferiblemente células de vertebrados tales como las de una línea celular de ratón, rata, mono o ser humano. Las células huésped eucariotas preferidas incluyen células de ovario de hámster chino (CHO) disponibles de la ATCC como CCL61, células de embrión de ratón NIH Swiss (NIH/3T3) disponibles de la ATCC como CRL 1658, células de riñón de cría de hámster (BHK) y líneas celulares de cultivo tisular eucariotas similares.

Puede usarse cualquier huésped procarionta para expresar una molécula de ADN_r que codifica para una proteína de la invención. El huésped procarionta preferido es *E. coli*.

La transformación de huéspedes celulares apropiados con una molécula de ADN_r de la presente invención se logra mediante métodos bien conocidos que normalmente dependen del tipo de vector usado y el sistema huésped empleado. Con respecto a la transformación de células huésped procariontas, se emplean normalmente métodos de tratamiento con sal y electroporación (véase, por ejemplo, Cohen *et al.* (1972), Proc. Natl. Acad. Sci. USA 69: 2110; y Sambrook *et al.*, citado anteriormente). Con respecto a la transformación de células de vertebrados con vectores que contienen ADN_r, se emplean normalmente métodos de tratamiento con sal, lípidos catiónicos o electroporación, véase, por ejemplo, Graham *et al.* (1973), Virol. 52: 456; Wigler *et al.* (1979), Proc. Natl. Acad. Sci. USA 76: 1373-1376.

Pueden identificarse células transformadas satisfactoriamente, es decir, células que contienen una molécula de ADN_r de la presente invención, mediante técnicas bien conocidas incluyendo la selección de un marcador seleccionable. Por ejemplo, pueden clonarse células que resultan de la introducción de un ADN_r de la presente invención para producir colonias individuales. Pueden recogerse células de esas colonias, lisarse y examinarse su contenido en ADN para detectar la presencia del ADN_r usando un método tal como el descrito por Southern, (1975) J. Mol. Biol. 98: 503 o Berent *et al.*, (1985) Biotech. 3: 208, o las proteínas producidas a partir de la célula sometida a ensayo mediante un método inmunológico. Los presentes inventores prepararon una *Escherichia coli* DH5@/p313-JF3, y la depositaron en colección coreana de cultivos tipo del Korea Research Institute of Bioscience and Biotechnology el 5 de junio de 2006 (n.º de depósito KCTC 10954BP).

F. Producción de proteínas recombinantes usando una molécula de ADN_r

La presente invención da a conocer además métodos para producir una proteína de la invención usando moléculas de ácido nucleico descritas en el presente documento. En términos generales, la producción de una forma recombinante de una proteína normalmente implica las siguientes etapas:

En primer lugar, se obtiene una molécula de ácido nucleico que codifica para una proteína de la invención, tal como una molécula de ácido nucleico que comprende, que consiste esencialmente en o que consiste en SEQ ID NO: 1, o los nucleótidos 53-643 o 53-640 de SEQ ID NO: 1. Si la secuencia codificante está interrumpida por intrones, tal como en estos marcos de lectura abiertos, es directamente adecuada para su expresión en cualquier huésped.

Entonces, la molécula de ácido nucleico se coloca preferiblemente en unión operativa con secuencias control adecuadas, tal como se describió anteriormente, para formar una unidad de expresión que contiene el marco de lectura abierto de la proteína. La unidad de expresión se usa para transformar un huésped adecuado y el huésped transformado se cultiva en condiciones que permiten la producción de la proteína recombinante. Opcionalmente, la proteína recombinante se aísla del medio o de las células; puede que no sea necesario la recuperación y purificación de la proteína en algunos casos en los que algunas impurezas pueden tolerarse.

Cada una de las etapas anteriores puede realizarse de una variedad de formas. Por ejemplo, pueden obtenerse las secuencias codificantes deseadas a partir de fragmentos genómicos y usarse directamente en huéspedes apropiados. La construcción de vectores de expresión que son operativos en una variedad de huéspedes se logra usando replicones y secuencias control apropiadas, tal como se expuso anteriormente. Las secuencias control, vectores de expresión y métodos de transformación dependen del tipo de célula huésped usada para expresar el gen y se comentaron en detalle anteriormente. Pueden añadirse sitios de restricción adecuados, si no están disponibles normalmente, a los extremos de la secuencia codificante de manera que proporcionan un gen que puede escindirse para insertar en estos vectores. Un experto habitual puede adaptar fácilmente cualquier sistema de huésped/expresión conocido en la técnica para su uso con las moléculas de ácido nucleico de la invención para producir proteína recombinante.

G. Química combinatoria y diseño de fármacos racional

La presente invención abarca además química combinatoria y diseño de fármacos racional. Los expertos reconocerán métodos apropiados para utilizar y explotar aspectos de la presente invención en la identificación de compuestos que pueden desarrollarse para el tratamiento de cáncer pancreático. El diseño de fármacos racional que implica polipéptidos requiere identificar y definir un primer péptido con el que el fármaco diseñado va a interactuar, y usar el primer péptido diana para definir los requisitos de un segundo péptido. Con tales requisitos definidos, puede encontrarse o prepararse un péptido o no péptido apropiado que cumple todos o sustancialmente todos los requisitos definidos. Por tanto, un objetivo del diseño de fármacos racional es producir análogos funcionales o estructurales de polipéptidos biológicamente activos de interés o de moléculas pequeñas con las que interactúan (por ejemplo, agonistas, antagonistas, compuestos nulos) con el fin de crear fármacos que son, por ejemplo, formas más o menos potentes del ligando. (Véase, por ejemplo, Hodgson (1991), *Bio. Technology* 9:19-21). La química combinatoria es la ciencia de sintetizar y someter a prueba compuestos para determinar su bioactividad en masa, en lugar de uno por uno, siendo el objetivo descubrir fármacos y materiales más rápida y económicamente de lo que era anteriormente posible. La química combinatoria y el diseño de fármacos racional se han relacionado más íntimamente en los últimos años debido al desarrollo de enfoques en el modelado de proteínas asistido por ordenador y el descubrimiento de fármacos. (Véanse por ejemplo, las patentes estadounidenses n.ºs 4.908.773; 5.884.230; 5.873.052; 5.331.573 y 5.888.738).

El uso del modelado molecular como herramienta para la química combinatoria y el diseño de fármacos racional ha aumentado drásticamente debido a la llegada de los gráficos por ordenador. No sólo es posible ver las moléculas en las pantallas de los ordenadores en tres dimensiones sino que también es posible examinar las interacciones de macromoléculas tales como enzimas y receptores y diseñar racionalmente moléculas derivadas para someterlas a prueba. (Véase Boorman (1992), *Chem. Eng. News* 70:18-26). Una gran cantidad de hardware y software fácil de usar está disponible ahora y prácticamente todas las compañías farmacéuticas tienen grupos de modelado por ordenador dedicados al diseño de fármacos racional. Molecular Simulations Inc. (www.msi.com), por ejemplo, comercializa varios programas sofisticados que permiten a un usuario empezar desde una secuencia de aminoácidos, construir un modelo bi o tridimensional de la proteína o polipéptido, compararlo con otros modelos bi o tridimensionales, y analizar las interacciones de compuestos, fármacos, y péptidos con un modelo tridimensional en tiempo real. Por consiguiente, en algunas realizaciones de la invención, se usa el software para comparar regiones de la proteína de la invención y las moléculas que interactúan con la misma (denominadas conjuntamente "parejas de unión" --por ejemplo, anticuerpos anti-proteína), y fragmentos o derivados de estas moléculas con otras moléculas, tales como péptidos, peptidomiméticos y productos químicos, de modo que puedan predecirse y diseñarse interacciones terapéuticas. (Véase Schneider (1998), *Genetic Engineering News* December: página 20; Tempczyk *et al.* (1997), *Molecular Simulations Inc. Solutions* April; y Butenhof (1998), *Molecular Simulations Inc. Case Notes* (agosto de 1998) para una discusión de modelado molecular).

H. Terapia génica

En otra realización, puede usarse terapia génica como medio para modular procesos patológicos y biológicos asociados con la función y actividad de la proteína. Esto comprende insertar en una célula cancerosa un constructo génico que codifica para una proteína que comprende todas o al menos una parte de las secuencias de SEQ ID NO: 2, o alternativamente un constructo génico que comprende toda o una parte de la región no codificante de SEQ ID NO: 1, operativamente unida a un elemento potenciador o promotor de manera que la expresión de dicha proteína provoca la supresión de dicho cáncer y en la que dicho elemento potenciador o promotor es un elemento potenciador o promotor que modula dicho constructo génico.

En los constructos descritos, la expresión de dicha proteína puede dirigirse a partir de cualquier promotor adecuado (por ejemplo, los promotores de citomegalovirus humano (CMV), virus de simio 40 (SV40) o metalotioneína), y regularse mediante cualquier elemento regulador de mamífero apropiado. Por ejemplo, si se desea, pueden usarse potenciadores que se sabe que dirigen preferentemente la expresión génica en células neurales, células T o células B para dirigir la expresión. Los potenciadores usados podrían incluir, sin limitación, los que se caracterizan como específicos de célula o tejido en su expresión. Alternativamente, si se usa un clon genómico de LBFL313 como constructo terapéutico (por ejemplo, tras su aislamiento mediante hibridación con la molécula de ácido nucleico de la invención descrita anteriormente), la regulación puede estar mediada por secuencias reguladoras relacionadas o, si se desea, por secuencias reguladoras derivadas de una fuente heteróloga, incluyendo cualquiera de los promotores o elementos reguladores descritos anteriormente.

La inserción del constructo dentro una célula cancerosa se logra *in vivo*, por ejemplo usando un vector de plásmido o viral. Tales métodos también pueden aplicarse a usos *in vitro*. Por tanto, los métodos de la presente invención pueden aplicarse fácilmente a diferentes formas de terapia génica, o bien en las que las células se modifican genéticamente *ex vivo* y luego se administran a un huésped o bien en las que la modificación génica se realiza *in vivo* usando cualquiera de varios métodos adecuados que implican vectores especialmente adecuados para tales terapias.

Pueden usarse vectores retrovirales, vectores adenovirales, vectores virales adenoasociados u otros vectores virales

con el tropismo apropiado para células que probablemente están implicadas en cáncer (por ejemplo, células epiteliales) como sistema de suministro de transferencia génica para un constructo génico terapéutico. Se conocen generalmente numerosos vectores útiles para este fin (Cozzi PJ, *et al.*, (2002) *Prostate*, 53(2):95-100; Bitzer M, Lauer U., (2002) *Dtsch Med Wochenschr.* 127(31-32):1623-1624; Mezzina y Danos (2002), *Trends Genet.* 8:241-256; Loser *et al.* (2002) *Curr. Gene Ther.* 2:161-171; Pfeifer y Verma (2001), *Annu. Rev. Genomics Hum. Genet.* 2:177-211). Los vectores retrovirales están particularmente bien desarrollados y se han usado en la práctica clínica (Anderson *et al.* (1995), patente estadounidense n.º 5.399.346). También pueden emplearse enfoques no virales para la introducción de ADN terapéutico en células que de lo contrario están predestinadas a experimentar cáncer (Jeschke *et al.* (20002) *Curr. Gene Ther.* 1:267-278; Wu *et al.* (1988), *J. Biol. Chem.* 263:14621-14624; Wu *et al.* (1989), *J. Biol. Chem.* 264:16985-16987). Por ejemplo, puede introducirse un gen en una neurona o una célula T mediante lipofección, conjugación asialorosonucleoide-polilisina o, menos preferiblemente, microinyección en condiciones quirúrgicas.

Para cualquiera de los métodos de aplicación descritos anteriormente, el constructo de ácido nucleico terapéutico se aplica preferiblemente al sitio del acontecimiento canceroso (por ejemplo, mediante inyección). Sin embargo, también puede aplicarse al tejido en las proximidades del acontecimiento canceroso o a un vaso sanguíneo que suministra a las células predestinadas a experimentar cáncer.

I. Métodos de diagnóstico

Puesto que los genes y las proteínas de la invención se expresan de manera diferente en tejidos de cáncer pancreático en comparación con tejidos pancreáticos no cancerosos, los genes y las proteínas de la invención también pueden usarse para diagnosticar o monitorizar el cáncer pancreático, rastrear la evolución de la enfermedad o diferenciar muestras de tejido pancreático de muestras tejido pancreático no canceroso. Uno de los medios de diagnóstico del cáncer usando las moléculas de ácido nucleico o las proteínas de la invención implica obtener tejido de sujetos vivos.

Los ensayos para detectar moléculas de ácido nucleico o de proteína de la invención pueden estar en cualquier formato disponible. Los ensayos típicos para moléculas de ácido nucleico incluyen formatos basados en PCR o hibridación. Los ensayos típicos para la detección de proteínas, polipéptidos o péptidos de la invención incluyen el uso de sondas de anticuerpo en cualquier formato disponible tal como ensayos de unión *in situ*, etc. (véase Harlow & Lane, *Antibodies - A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1988). En realizaciones preferidas, los ensayos se llevan a cabo con controles apropiados.

Generalmente, los diagnósticos de la invención pueden clasificarse según si la realización es un ensayo a base de ácido nucleico o proteína. Algunos ensayos de diagnóstico detectan mutaciones o polimorfismos en los ácidos nucleicos o las proteínas de la invención, que contribuyen a aberraciones cancerosas. Otros ensayos de diagnóstico identifican y distinguen defectos en la actividad de la proteína detectando un nivel de ARN o proteína de la invención en un organismo sometido a prueba que se asemeja al nivel de ARN o proteína de la invención en un organismo que padece una enfermedad, tal como cáncer, o detectando un nivel de ARN o proteína en un organismo sometido a prueba que es diferente de un organismo que no padece una enfermedad.

Adicionalmente, se contempla la fabricación de kits que incorporan los reactivos y métodos descritos en las siguientes realizaciones de manera que permiten la rápida detección e identificación de aberraciones en el nivel o la actividad de la proteína. Los kits de diagnóstico pueden incluir una sonda de ácido nucleico o un anticuerpo o combinaciones de los mismos, que detectan específicamente una forma mutante de una proteína de la invención o una sonda de ácido nucleico o un anticuerpo o combinaciones de los mismos, que pueden usarse para determinar el nivel de expresión de ARN o proteína de una o más proteínas de la invención. El componente de detección de estos kits se suministrará normalmente en combinación con uno o más de los siguientes reactivos. Se suministrará a menudo un soporte que puede absorber o unirse de otra forma al ADN, ARN o proteína. Los soportes disponibles incluyen membranas de nitrocelulosa, nailon o nailon derivatizado que pueden caracterizarse por llevar una serie de sustituyentes cargados positivamente. Pueden suministrarse en estos kits una o más enzimas de restricción, reactivos de control, tampones, enzimas de amplificación y polinucleótidos no humanos como ADN de esperma de salmón o timo de ternero.

Las técnicas de diagnóstico a base de ácido nucleico útiles incluyen, pero sin limitarse a, secuenciación de ADN directa, electroforesis en gel en gradiente, análisis de transferencia de tipo Southern, análisis de conformación monocatenaria (SSCA), ensayo de protección de ARNasa, análisis de transferencia puntual, amplificación de ácido nucleico, PCR específica de alelo y combinaciones de estos enfoques. El punto de partida para estos análisis es ácido nucleico aislado o purificado a partir de una muestra biológica. Se contempla que biopsias de tejido proporcionen una buena fuente de muestras. El ácido nucleico se extrae de la muestra y puede amplificarse mediante una técnica de amplificación de ADN tal como la reacción en cadena de la polimerasa (PCR) usando cebadores. Los expertos en la técnica reconocerán fácilmente métodos disponibles para confirmar la presencia de polimorfismos. Además, puede emplearse cualquier tecnología de alineamientos de polinucleótidos se conoce como Genechips™, y se ha descrito de manera general en la patente estadounidense n.º 5.143.854; las publicaciones

PCT WO 90/15070 y 92/10092.

Los expertos en la técnica conocen una amplia variedad de marcadores y técnicas de conjugación y pueden usarse en diversos ensayos de ácido nucleico. Existen varias maneras de producir ácidos nucleicos marcados para hibridación o PCR incluyendo, pero sin limitarse a, oligomarcaje, traducción de mellas, marcaje de extremos, o amplificación por PCR usando un nucleótido marcado. Alternativamente, un ácido nucleico que codifica para una proteína de la invención puede clonarse en un vector para la producción de una sonda de ARNm. Tales vectores se conocen en la técnica, están disponibles comercialmente y pueden usarse para sintetizar sondas de ARN *in vitro* mediante la adición de una ARN polimerasa apropiada tal como T7, T3 o Sup6 y nucleótidos marcados. Varias compañías tales como Pharmacia Biotech (Piscataway, NJ), Promega (Madison, WI) y U.S. Biochemical Corp (Cleveland, OH) suministran kits y protocolos comerciales para estos procedimientos. Los marcadores o moléculas indicadoras adecuadas incluyen los radionúclidos, enzimas, agentes fluorescentes, quimioluminiscentes o cromógenos, así como, sustratos, cofactores, inhibidores, partículas magnéticas y similares.

En el diagnóstico a base de proteína preferido, los anticuerpos de la invención se unen a un soporte en un alineamiento ordenado en el que una pluralidad de anticuerpos se unen a distintas regiones del soporte que no se solapan entre sí. Los expertos en la técnica reconocerán rápidamente ensayos disponibles que son diagnósticos a base de proteínas. Las proteínas se obtienen de muestras biológicas y se marcan mediante enfoques convencionales (por ejemplo, radiactividad, colorimétrica o fluorescentemente). Empleando patrones marcados de una concentración conocida de proteína de la invención de tipo natural y/o mutante, un investigador puede determinar con exactitud la concentración de la proteína de la invención en una muestra y a partir de esta información puede evaluar el nivel de expresión de la forma particular de la proteína. También pueden usarse métodos convencionales en densitometría para determinar con mayor exactitud la concentración o el nivel de expresión de tal proteína. Estos enfoques también se automatizan fácilmente usando tecnología conocida por los expertos en la técnica de análisis de diagnóstico de alto rendimiento. Tal como se detalló anteriormente, puede emplearse cualquier tecnología de alineamiento abordable conocida en la técnica con este aspecto de la invención y presentar los alineamientos de proteína sobre los chips en un intento por maximizar los patrones de unión de anticuerpos y la información de diagnóstico.

Tal como se comentó anteriormente, la presencia o detección de un polimorfismo en un gen o una proteína de la invención puede proporcionar un diagnóstico de un cáncer o enfermedad similar en un organismo. Las realizaciones adicionales incluyen la preparación de kits de diagnóstico que comprenden componentes de detección, tales como anticuerpos, específicos para una variante polimórfica particular del gen o la proteína de la invención. El componente de detección se suministrará normalmente en combinación con uno o más de los siguientes reactivos. Se suministrará a menudo un soporte que puede absorber o unirse de otra forma al ARN o proteína. Los soportes disponibles para este fin incluyen, pero sin limitarse a, membranas de nitrocelulosa, nailon o nailon derivatizado que puede caracterizarse por llevar una serie de sustituyentes cargados positivamente, y Genechips™ o sus equivalentes. Una o más enzimas, tales como transcriptasa inversa y/o Taq polimerasa, pueden proporcionarse en el kit, así como también dNTP, tampones o polinucleótidos no humanos como ADN de esperma de salmón o timo de ternero. Los resultados de los ensayos del kit pueden interpretarse por un profesional sanitario o un laboratorio de diagnóstico. Alternativamente, los kits de diagnóstico se fabrican y comercializan a individuos privados para autodiagnóstico.

Además de diagnosticar la enfermedad según la presencia o ausencia de un polimorfismo, algunas enfermedades que implican cáncer resultan de niveles sesgados del gen o la proteína de la invención en tejidos particulares o patrones aberrantes de expresión de la proteína de la invención. Monitorizando el nivel de expresión en diversos tejidos, por ejemplo, puede realizarse un diagnóstico o identificarse un estado patológico. De manera similar, determinando razones del nivel de expresión de diversas proteínas de la invención en tejido específicos (por ejemplo, patrones de expresión) puede realizarse un pronóstico de la salud o enfermedad. Se determinan los niveles de expresión de la proteína de la invención en diversos tejidos de individuos sanos, así como individuos que padecen cánceres. Estos valores pueden registrarse en una base de datos y pueden compararse con valores obtenidos de individuos sometidos a prueba. Adicionalmente, las razones o patrones de expresión en diversos tejidos de individuos tanto sanos como enfermos se registran en una base de datos. Estos análisis se denominan "perfiles de estado patológico" y comparando un perfil de estado patológico (por ejemplo de un individuo sano o enfermo) con respecto a un perfil de estado patológico de un individuo sometido a prueba, un médico puede diagnosticar rápidamente la presencia o ausencia de la enfermedad. Las técnicas de diagnóstico a base de proteínas y ácidos nucleicos descritas anteriormente pueden usarse para detectar el nivel o la cantidad o razón de expresión de genes o proteínas de la invención en un tejido. A través de hibridaciones de tipo Northern cuantitativas, análisis *in situ*, inmunohistoquímica, ELISA, tecnología de alineamiento genechip, PCR e inmunotransferencias de tipo Western, por ejemplo, puede determinarse rápidamente la cantidad o el nivel de expresión de ARN o proteína para una proteína particular de la invención (tipo natural o mutante) y a partir de esta información pueden establecerse razones de expresión. Alternativamente, las proteínas de la invención que van a analizarse pueden ser miembros de familias que se desconocen actualmente pero que se identifican basándose en su posesión de una o más de las regiones de homología descritas anteriormente.

Descripción de los dibujos

Figura 1. La figura 1 muestra la secuencia señal pronosticada para la secreción de LBFL313 (SEQ ID NO: 2). El análisis se ha realizado usando SignallP 3.0 Server (www.cbs.dtu.dk/services/SignallP/).

5 Figura 2. La figura 2 es el resultado del análisis de tipo Western que muestra que LBFL313 se detecta en el sobrenadante de cultivo de células.

Figura 3. La figura 3 muestra los efectos de la sobreexpresión de LBFL313 en células CHO sobre la proliferación celular (panel A), motilidad (panel B) e invasividad (panel C).

10 Figura 4. La figura 4 muestra los efectos de la sobreexpresión de LBFL313 en ratones desnudos sobre la tumorigénesis (panel A) y formación de microvasos (panel B).

15 Figura 5. La figura 5 muestra resultados representativos del análisis inmunohistoquímico de la expresión de LBFL313 usando muestras de biopsias pancreáticas y anticuerpo anti-LBFL313.

Figura 6. La figura 6 representa gráficos que muestran los efectos del anticuerpo policlonal anti-LBFL313 sobre la invasividad de líneas celulares de cáncer pancreático.

20 Mejor modo

Sin descripción adicional, se cree que un experto habitual en la técnica puede, usando la descripción anterior y los siguientes ejemplos ilustrativos, preparar y utilizar los compuestos de la presente invención y poner en práctica los métodos reivindicados. Los siguientes ejemplos de trabajo por tanto, señalan específicamente realizaciones preferidas de la presente invención, y no deben interpretarse como limitativos de ninguna manera del resto de la descripción.

25

Ejemplos**30 Ejemplo 1****Identificación de ARNm expresado de manera diferente en adenocarcinoma pancreático**

35 Se derivaron muestras de tejido de paciente de pacientes coreanos y se clasificaron en dos grupos. Un grupo consistió en pacientes a los que se les había diagnosticado adenocarcinoma pancreático. La edad de los pacientes en este grupo, seis hombres y tres mujeres, oscilaba entre 51-70. Al segundo grupo de pacientes se les había diagnosticado páncreas normal. En este grupo de tres hombres, la edad de los pacientes oscilaba entre 63-66. Se realizaron los análisis histológicos de cada una de las muestras de tejidos y se separaron las muestras en categorías de o bien no cancerosa o bien cancerosa.

40 Con modificaciones menores, el protocolo de preparación de muestras siguió el manual de análisis de la expresión Affymetrix GeneChip. En primer lugar, se molió hasta obtener un polvo tejido congelado usando el molino Spex Certiprep 6800 Freezer. Entonces se extrajo el ARN total usando Trizol (Life Technologies). A continuación, se aisló el ARNm usando el kit Oligotex mRNA Midi (Qiagen). Usando 1-5 mg de ARN, se creó ADNc bicatenario usando el sistema SuperScript Choice (Gibco-BRL). Se cebó la síntesis de ADNc de primera hebra con un oligonucleótido de T7-(dT24). Entonces, se extrajo con fenol-cloroformo el ADNc y se precipitó con etanol hasta una concentración final de 1 mg/ml.

50 A partir de 2 mg de ADNc, se sintetizó ARNc según procedimientos convencionales. Para marcar con biotina el ARNc, se añadieron los nucleótidos Bio-11-CTP y Bio-16-UTP (Enzo Diagnostics) a la reacción. Tras una incubación a 37°C durante seis horas, se limpió el ARNc marcado según el protocolo del kit Rneasy Mini (Qiagen). Entonces, se fragmentó el ARNc (tampón de fragmentación 5': Tris-Acetato 200 mM (pH 8,1), KOAc 500 mM, MgOAc 150 mM) durante treinta y cinco minutos a 94°C.

55 Se hibridaron cincuenta y cinco mg de ARNc fragmentado en el conjunto de alineamientos de Affymetrix Human Genome U133 durante veinticuatro horas a 60 rpm en una estufa de hibridación a 45°C. Se lavaron los chips y se tiñeron con estreptavidina-ficoeritrina (SAPE) (Molecular Probes) en estaciones fluidicas Affymetrix. Para amplificar la tinción, se añadió disolución SAPE dos veces con una etapa de tinción con anticuerpo biotinilado anti-estreptavidina (Vector Laboratories) entre las etapas. Se detectó la hibridación con los alineamientos de las sondas mediante exploración fluorimétrica (Hewlett Packard Gene Array Scanner). Tras la hibridación y exploración, se analizaron las imágenes de microalineamientos para control de calidad, buscando defectos en el chip importantes o anomalías en la señal de hibridación. Después de que todos los chips pasaron el CC, se usaron Affymetrix Microarray Suite (v5.0) y LIMS (v3.0).

65 Se determinó la expresión de genes diferencial entre las muestras pancreáticas cancerosas y no cancerosas usando

5 conjuntos U133 de Affymetrix human GeneChip, con los siguientes métodos estadísticos. (1) para cada gen, se determinaron los valores de señal para U133 mediante Affymetrix Microarray Suite (v5.0), que también realizó llamadas “Ausentes” (=no detectado), “Presentes” (=detectado) o “Marginales” (=no está claramente ausente o presente) para cada elemento del GeneChip. (2) Usando el criterio de al menos el 40% de llamadas presentes en grupos de muestras pancreáticas cancerosas, se seleccionó un conjunto de gen para análisis adicional. (3) Se transformaron todos los valores de señal a una escala logarítmica. (4) Se usó el método de análisis de la varianza (ANOVA) para el análisis de datos (Steel *et al.*, Principles and Procedures of Statistics: A Biometrical Approach, tercera ed., McGraw-Hill, 1997).

10 El análisis de los datos del chip mostró que la expresión del marcador LBFL313 estaba regulada por incremento significativamente (11,13 veces, $p = 0$) en muestras de adenocarcinoma pancreático en comparación con muestras de tejido pancreático normal. Estos datos indican que la regulación por incremento de LBFL313 puede ser un diagnóstico para el cáncer pancreático.

15 Puede medirse el nivel de expresión de LBFL313 (SEQ ID NO: 1) mediante el fragmento de secuencia del chip n.º 228058_at en un Affymetrix GeneChip U133. A través de la obtención de datos anteriores combinados con el GeneExpress Oncology Datasuite™ de Gene Logic, Inc. (Gaithersburg, MD), se muestran en la tabla 1 los niveles de expresión de 228058_at en diversos neoplasmas malignos, en comparación con tejidos control normales, en la que el cambio en veces y la dirección del cambio (regulación por incremento o por disminución) también se indican.
20 Un cambio en veces mayor de 1,5 se consideró significativo.

25 Se validaron los resultados de la expresión de GeneChip, determinados mediante la unión de la muestra al fragmento de secuencia del n.º 228058_at, mediante RT-PCR cuantitativa (Q-RT-PCR) usando el ensayo Taqman® (Perkin-Elmer). Se usaron cebadores de PCR diseñados a partir del archivo de información de secuencia del fragmento de Affymetrix específico (228058_at) en el ensayo. Se sometió a ensayo el gen diana en cada muestra de ARN (diez ng de ARN total) en relación con un gen de referencia añadido de manera conocida exógenamente. Para este fin, se usó el gen de resistencia a tetraciclina como adición conocida añadida exógenamente. Este enfoque proporciona la expresión relativa tal como se midió mediante valor de ciclo umbral (Ct) del ARNm diana en relación con una cantidad constante de valores de Ct de la adición conocida de Tet. El panel de muestras incluía ARN de tejido que se analizó en U133 GeneChips. Además, se usaron varias muestras nuevas que no se analizaron en el GeneChip para las validaciones de la expresión mediante Q-RT-PCR. Los datos de QRT-PCR confirman la regulación por incremento de LBFL313 observada en adenocarcinoma pancreático en comparación con muestras de biopsias pancreáticas normales.

35 Tabla 1.

Expresión de LBFL313 en neoplasias malignas

Tejido	Patología/morfología	Cambio en veces	Dirección	valor de p
Mama	Carcinoma lobular infiltrante	1,64	Aumento	0,19
	Carcinoma lobular y ductal infiltrante	1,59	Aumento	0,06
Cuello uterino	Carcinoma de células escamosas	1,61	Aumento	0,38
Colon	Adenocarcinoma mucinoso	1,88	Aumento	0,05
Duodeno	Adenocarcinoma	1,69	Aumento	0,19
Esófago	Adenocarcinoma	1,83	Aumento	0,15
Hígado	Carcinoma hepatocelular	1,52	Aumento	0,02
Pulmón	Adenocarcinoma	1,52	Aumento	0,21
Ovario	Adenocarcinoma quístico mucinoso	9,65	Aumento	0
	Adenocarcinoma quístico seroso	2,15	Aumento	0,13
Páncreas	Adenocarcinoma	8,62	Aumento	0
Piel	Carcinoma de células basales	-3,01	Disminución	0,01
	Melanoma maligno	-6,61	Disminución	0
	Carcinoma de células escamosas	-3,82	Disminución	0,03
Estómago	Carcinoma de células en anillo de sello	2,40	Aumento	0,03

40 Ejemplo 2

Clonación de ADNc humano de longitud completa (LBFL313) correspondiente a especies de ARNm expresado de manera diferente

45 Se obtuvo el ADNc de longitud completa que tiene SEQ ID NO: 1 mediante el método de oligo-precipitación. En resumen, se diseñó un oligo específico para el gen basándose en la secuencia de LBFL313. Se marcó el oligo con biotina y se usó para hibridar con 2 µg de ADN de plásmido monocatenario (recombinantes de ADNc) a partir de una biblioteca de placenta humana siguiendo los procedimientos de Sambrook *et al.* Se separaron los ADNc hibridados

mediante perlas conjugadas con estreptavidina y se eluyeron mediante calentamiento. Se convirtió el ADNc eluido en ADN de plásmido bicatenario y se usó para transformar células de *E. coli* (DH10B) y se seleccionó el ADNc más largo. Tras confirmarse la selección positiva mediante PCR usando cebadores específicos para el gen, se sometió el clon de ADNc a secuenciación de ADN.

La secuencia de nucleótidos de los ADNc humanos de longitud completa correspondiente al ARNm regulado de manera diferente detectado anteriormente se expone en SEQ ID NO: 1. El ADNc comprende 777 pares de bases.

Un marco de lectura abierto dentro de la secuencia de nucleótidos de ADNc de SEQ ID NO: 1, en los nucleótidos 53-640 (53-643 incluyendo el codón de terminación), codifica para una proteína de 196 aminoácidos. La secuencia de aminoácidos correspondiente a una proteína pronosticada codificada por SEQ ID NO: 1 se expone en SEQ ID NO: 2.

SEQ ID NO: 2 contiene un dominio de lectina de tipo jacalina: las proteínas que contienen este dominio son lectinas. Se encuentra en de 1 a 6 copias en estas proteínas. El dominio también se encuentra en la proteína de unión a espermina prostática animal (Raval *et al.* (2004) *Glycobiology* 14:1247-1263). La figura 1 muestra el resultado del análisis de la secuencia señal usando SignalIP 3.0 Server (www.cbs.dtu.dk/services/SignalIP/). Se pronostica el posible sitio de escisión de la secuencia señal para la secreción entre la posición de aminoácido 40 (Ala) y 41 (Gly).

Se realizó el análisis mediante transferencia de tipo Northern para determinar el tamaño de los transcritos de ARNm que corresponden a LBFL313. Se usó una transferencia de tipo Northern que contenía ARN totales de diversos tejidos humanos (Human 12-Lane MTN Blot, Clontech, Palo Alto, CA), y se marcó radioactivamente un EST que contenía la secuencia de 228058_at mediante el método del cebador aleatorio y se usó para explorar con sondas la transferencia. Se hibridó la transferencia en formamida al 50%, 5X SSPE, SDS al 0,1%, disolución de Denhart 5X y ADN de esperma de arenque 0,2 mg/ml a 42°C y se lavó con 0,2X SSC que contenía SDS al 0,1% a temperatura ambiente. La transferencia de tipo Northern mostró un único transcrito para este gen, que tiene aproximadamente 0,8 kb de tamaño. Esto corresponde al tamaño del clon de LBFL313 (SEQ ID NO: 1).

Ejemplo 3

Producción de líneas celulares transfectadas con LBFL313

Se amplificó la región codificante de LBFL313 mediante PCR usando cebador directo (5'-TTG GGATCCGTATAAAGGCGATGTGGAGG-3') que incorpora el sitio BamHI y cebador inverso (5'-ACC ATC TAG AGC GAC CCA CGG GTG AGT-3') que incorpora el sitio XbaI. Se realizó la PCR usando la TaqPlus precision ADN polimerasa (Stratagene, CA) según las instrucciones del fabricante. Los ciclos de amplificación por PCR implicaban desnaturalización inicial a 94°C durante 2 min., y 27 ciclos; 94°C durante 30 s, 50°C durante 30 s, 72°C durante 1 min., seguido por una extensión final a 72°C durante 10 min. Se clonó el producto de PCR en los sitios *Bam*HI y *Xba*I del vector de expresión de mamífero pADNc3.1-mycHis (Invitrogen). Se secuenció el plásmido clonado (pLFG250) a través de la región del sitio de clonación para confirmar su estructura primaria.

Se transfectaron de manera estable células CHO subconfluentes con el vector pLFG250 o con el vector pADNc3.1-mycHis solo usando reactivo LipofectaminePLUS (Invitrogen) según las instrucciones del fabricante. Tras 24 h, se cultivaron las células transfectadas en F12 de Ham (Invitrogen) que contenía FBS al 10% y G418 400 µg/ml (Sigma) para la selección. Se cambió este medio de selección cada 2 días, y tras 10-12 días, se usaron anillos de clonación para aislar clones positivos. Se expandieron los cultivos adicionalmente y se examinaron para detectar la expresión de LBFL313. Se lisaron las células en tampón de lisis que contenía HEPES 50 mM (pH 7,2), NaCl 150 mM, beta-glicerofosfato 25 mM, NaF 25 mM, EGTA 5 mM, EDTA 1 mM, NP-40 al 1%, ortovanadato de sodio 1 mM, PMSF 0,1 mM y cóctel de inhibidores de proteasas (leupeptina, pepstatina, aprotinina y antipaina cada una 5 µg/ml). Se concentraron los medios de cultivo mediante un aparato Microcon de punto de corte de 10k (Amicon) constituyendo hasta 15 µl de volumen. Se redujeron las proteínas mediante incubación durante 5 min. a 95°C en tampón de carga 4x SDS. Se resolvieron los polipéptidos a 10 mA / gel sobre geles de SDS-PAGE al 12% y se transfectaron electroforéticamente a una membrana de P-transferencia Immobilon (Millipore) de 0,45 µm durante 2 h a 200 mA / gel a 4°C. Se bloquearon las membranas durante 1 h en tampón TBST (Tris 25 mM, pH 7,5; NaCl 125 mM y Tween-20 al 0,1% (v/v)) que contenía leche desnatada al 5% (p/v). Entonces se exploraron con sondas las transferencias durante la noche con anticuerpo anti-His HRP (Santa Cruz). Entonces se visualizó el material inmunorreactivo mediante quimioluminiscencia potenciada (Elpis Biotech.) según las instrucciones del fabricante. Tal como se muestra en la figura 2, se detectó la proteína LBFL313 sólo en el sobrenadante del cultivo pero no en el extracto celular. Esto confirma la predicción de que LBFL313 codifica para una proteína secretada.

Ejemplo 4

Análisis de la sobreexpresión de LBFL313 sobre la proliferación celular, migración e invasión

Para determinar el efecto de la sobreexpresión de LBFL313 sobre la proliferación celular, se midió la tasa de crecimiento mediante recuento celular. En placas de 12 pocillos, se sembraron en placa 4 x 10⁴ células. Se

incubaron las placas a 37°C y CO₂ al 5% durante 12 días y se contó el número de células con hematocitómetro cada día. El resultado representa los valores medios + la desviación estándar de muestras por triplicado. Tal como se muestra en la figura 3A, en células CHO, la sobreexpresión de LBFL313 indujo una proliferación más rápida. Se observó un efecto similar con la línea celular de cáncer pancreático PANC-1 que sobreexpresa LBFL313.

Se estudió la migración e invasión mediante el ensayo de cámara de Boyden. Se realizaron ambos ensayos en una cámara de Boyden de 48 pocillos, una microcámara de 48 pocillos Neuro Probe (Neuroprobe). Se tripsinizaron las células y se resuspendieron en inhibidor de tripsina, disolución de inhibidor de tripsina de semilla de soja (Sigma). Se sedimentaron las células y se suspendieron hasta una concentración final de 2×10^6 células/ml en medio libre de suero. Se llenaron los pocillos inferiores de la cámara con 30 μ l de medio convencional. Se ensambló la cámara usando filtros de policarbonato (policarbonato, 8 μ m de diámetro de tamaño de poro, Neuroprobe). Para los ensayos de invasión celular, se colocó en capas 1 mg/ml de matriz de membrana basal Matrigel™ (BD biosciences) sobre cada filtro (500 μ g/filtro) y se secó al aire. Se añadieron muestras de cincuenta μ l de suspensión celular (1×10^5 células/pocillo) al compartimiento superior. Se incubó la cámara a 37°C y CO₂ al 5% y se varió el tiempo de incubación dependiendo de los tipos de células que se van a analizar: durante 24 h en el caso del ensayo de migración de CHO, durante 48 h en el caso del ensayo de invasión de CHO y durante 18 h en el caso del ensayo de migración e invasión de PANC-1. Al final de la incubación, se retiraron mecánicamente las células en la superficie superior de los filtros. Se fijaron los filtros en metanol y se tiñeron en disolución modificada de tinción de Giemsa (Sigma). Se contó el número de células migradas por campo (100x) en el microscopio óptico (Olympus). Se sometió a ensayo cada muestra por triplicado. Tal como se muestra en la figura 3B y la figura 3C, se potenciaron la invasividad y motilidad de las células CHO por la sobreexpresión de LBFL313. Se observaron efectos similares con la línea celular de cáncer pancreático PANC-1 que sobreexpresa LBFL313.

Ejemplo 5

Efectos de la sobreexpresión de LBFL313 sobre la tumorigénesis en ratones desnudos

Se determinaron los efectos biológicos de la sobreexpresión de LBFL313 sobre el crecimiento tumoral *in vivo*. Se inyectaron células CHO por vía subcutánea en los costados de ratones desnudos inmunodeficientes. Se tripsinizaron células CHO confluentes, no transfectadas o transfectadas de manera estable con el vector LBFL313 (pLFG250) o con vector solo, y se resuspendieron en PBS a una densidad de $3,3 \times 10^7$ células/ml. Se inyectaron cinco millones de células CHO de cada tipo por vía subcutánea en el costado de cinco ratones Balb/C (nu/nu) hembra de 8 semanas de edad. En el proceso de crecimiento de los ratones, se midieron las longitudes y anchuras de los tumores. Se calculó el volumen tumoral con la fórmula volumen tumoral = (longitud) x (anchura)² / 2. Tras 37 días, se sacrificaron los ratones y se analizaron los tumores. Tal como se muestra en la figura 4A, el tamaño de los tumores generados por células transfectadas con LBFL313 era más de 5 veces más grande que los formados por células control simuladas.

Para determinar el grado de angiogénesis inducida por el tumor, se tiñeron secciones en parafina de xenoinjertos de tumor con anticuerpo monoclonal anti-factor VIII (Dako). Se desparafinizaron secciones de tejido incrustadas en parafina (3-5 μ m de grosor) obtenidas de tumores de cada grupo de ratones en xileno y se rehidrataron en una serie de etanol graduado (100-90-80-70-50-30%) y se lavaron con PCS. Se bloqueó la peroxidasa endógena sumergiendo el portaobjetos en peróxido de hidrógeno al 0,3% (v/v) en metanol durante 15 min. a TA. Tras lavar tres veces con PBS durante 4 min. cada una, se bloquearon las secciones empapando en suero de burro normal al 10% (v/v) en PBS durante 1 h a TA. Tras lavar tres veces con PBS durante 4 min. cada una, se incubaron las secciones bloqueadas en anticuerpo monoclonal anti-factor VIII (factor de von Willebrand, dilución 1:50) (Dako) durante 2 h a TA. Tras 2 h, y lavar tres veces con PBS durante 4 min. cada una, se incubó el portaobjetos durante 30 min. con Link biotilado universal a TA, se lavó tres veces en PBS durante 4 min. cada una, y se incubó con conjugado de estreptoavidina-HRP (Dako) durante 15 min. a TA. Tras lavar tres veces con PBS, se incubaron las secciones con cromógeno y se lavaron en agua destilada. No se contratiñeron las secciones de factor VIII.

Se determinó el número de microvasos tal como describe Padro (Padro *et al.* (2000) Blood 95:2637-2644). Se evaluó simultáneamente el recuento de microvasos por dos investigadores experimentados independientes usando microscopía óptica. Los investigadores no eran conscientes del hallazgo clínico-patológico. Se exploró sistemáticamente la sección completa, es decir, campo por campo, a 100 aumentos para encontrar las zonas que mostraban la vascularización más intensa. Entonces se cambiaron los aumentos a 200 aumentos o a 400 aumentos, y se dejó que los investigadores recolocaran el portaobjetos hasta que el mayor número de microvasos estaba dentro del campo de 400 aumentos. Esta zona se definió como punto caliente tras lograr un consenso entre ambos investigadores, reduciendo así el error entre observadores del recuento de microvasos. En cada punto caliente, ambos investigadores realizaron el recuento de microvasos individuales en un campo de 400 aumentos. Tal como se muestra en la figura 4B, los resultados revelaron recuentos de microvasos superiores en los tumores de ratones a los que se les inyectaron células transfectadas con LBFL313 que células control simuladas. Los experimentos *in vivo* implican formación de tumores agresivos y angiogénesis potenciada por LBFL313.

Ejemplo 6Producción de anticuerpo policlonal anti-LBFL313

5 Se amplificó el ADNc de longitud completa de LBFL313 mediante PCR usando el cebador directo (5'-CTAAGGCCAGCCGGCCGGGAAGATGTATGGCCCTGGA-3') y el cebador inverso (5'-CATAGGCCACCGGCCGAGCGACCCACGGGTGAGTT-3'). Se realizó la PCR usando la TaqPlus precision ADN polimerasa (Stratagene, CA) según las instrucciones del fabricante. Los ciclos de amplificación por PCR implicaban desnaturalización inicial a 94°C durante 5 min., y 30 ciclos; 94°C durante 30 s, 58°C durante 30 s, 72°C durante 30 s, seguido por una extensión final a 72°C durante 10 min. Se visualizó el producto de PCR en un gel TAE al 1,5%, y se purificó el gel, usando el kit de recuperación Zymoclean Gel DNA (Zymo Research, CA) según las instrucciones del fabricante. Se insertó este ADN purificado en el sitio *SfiI* del vector pLFG106 para producir la proteína de fusión LBFL313-Fc recombinante. pLFG106 es un vector de expresión que contiene una secuencia señal, la región Fc de IgG humana y genes DHFR en la estructura principal de pDNAc3.1 (Invitrogen). PLFG106 también contiene una secuencia de reconocimiento de trombina entre el sitio clonación y la región Fc. Se transfectó de manera estable el plásmido de expresión de LBFL313-Fc en líneas celulares mutantes CHO deficientes en DHFR usando reactivos ExGen 500 (Fermentas, Lituania) según las instrucciones del fabricante. Se seleccionaron células transfectadas de manera estable en MEM- α libre de nucleótidos (Invitrogen, CA) que contenía 800 μ g/ml de G418 (Invitrogen, CA) y FBS dializado al 10% durante 2 semanas. Se adaptaron adicionalmente transfectantes estables en una concentración de metotrexato 100 nM (Sigma, MO) durante 2 semanas. Se identificó la proteína de fusión LBFL313-Fc recombinante, secretada en el medio de cultivo, mediante análisis de inmunotransferencia de tipo Western y ELISA usando anticuerpo anti-Fc humano (Sigma, MO).

25 Se realizó la expresión a gran escala de la proteína de fusión LBFL313-Fc recombinante usando frascos rotatorios (1750 cm²). Se hicieron crecer transfectantes estables que expresaban proteínas de fusión LBFL313-Fc recombinantes en IMDM más FBS al 5%. Se inoculó el cultivo en frasco rotatorio con 3×10^8 células y se hizo rotar el dispositivo a 5 rpm en una incubadora a 37°C. Se cultivaron las células durante 5 días, y se intercambio el medio por medio libre de suero cada tres días. Se eliminó el residuo celular en el medio libre de suero recogido mediante la centrifugación a 6.000 rpm durante 10 min. Para purificar la proteína de fusión LBFL313-Fc recombinante, se procesó adicionalmente medio libre de células mediante concentración usando Concentrator, ProFlux M12 (Amicon) a través de una membrana 10K MWCO, Pellicon 2 (Millipore). Entonces se hizo pasar el medio concentrado a través de una columna de proteína A agarosa de 10 ml (Pierce) preequilibrada en fosfato de sodio 20 mM, pH 8,0. Se retiró mediante lavado la proteína no unida de la columna con tampón fosfato de sodio 20 mM. Entonces se eluyó la columna con citrato 0,1 M, pH 3,0, recogiendo fracciones de 4,5 ml en tubos que contenían 500 μ l de Tris-HCl 1 M, pH 9,0. Se reunieron las fracciones que contenían proteína de fusión LBFL313-Fc recombinante y se purificaron adicionalmente con cromatografía de exclusión molecular usando resina sepharose 200HR (Amersham, IL).

40 Se escindió el Fc C-terminal de la proteína de fusión LBFL313-Fc recombinante mediante el kit Thrombin Cleavage Capture (Novagen), y luego se eliminó mediante columna de proteína A inmovilizada ImmunoPureR (Pierce). En resumen, se incubó la proteína de fusión LBFL313-Fc recombinante con trombina biotinizada a 20°C durante la noche. Tras la proteólisis, se eliminó la trombina biotinizada mediante perlas de estreptavidina-agarosa. La disolución resultante, una mezcla de LBFL313 recombinante y Fc, se separó haciéndola pasar dos veces a través de la columna de proteína A inmovilizada ImmunoPureR (Pierce). Se comprobó la pureza de LBFL313 recombinante sobre geles de SDS-PAGE.

45 Se usó LBFL313 recombinante purificada para inmunizar dos conejos usando técnicas convencionales. Se recogió el suero inmunizado y luego se purificó usando el kit de purificación de IgG ImmunoPureR (A Plus) (Pierce) según las instrucciones del fabricante. Se realizó la purificación adicional usando el kit de inmovilización AminoLink^R (A Plus) (Pierce) unido con Fc humano recombinante. El anticuerpo policlonal purificado detecta la proteína de aproximadamente 21 kDa en inmunotransferencias de tipo Western de medios condicionados obtenidos a partir de células de cáncer pancreático que expresan LBFL313. Se demostró adicionalmente la especificidad del anticuerpo mediante detección potenciada de la proteína LBFL313 obtenida a partir de los medios condicionados de células PANC-1 transfectadas con LBFL313, mientras que no hubo potenciación a partir de los de las células PANC-1 transfectadas sólo con vector.

Ejemplo 7Análisis inmunohistoquímico de la expresión de LBFL313

60 Se desparafinizaron portaobjetos de microalineamientos de tejidos en xileno y se rehidrataron en alcohol graduado. Se bloqueó la actividad peroxidasa endógena con metanol que contenía peróxido de hidrógeno al 0,3% a temperatura ambiente durante 20 min. Se realizó la recuperación del antígeno mediante microondas en tampón citrato (0,01 M, pH 6,0) durante 4 min. Entonces se incubaron los portaobjetos con disolución de suero de burro normal al 10% durante 1 h para reducir la tinción no específica de fondo.

65 El anticuerpo primario era anticuerpo policlonal anti-LBFL313, a una dilución de 1:500. Se incubaron las secciones

bloqueadas en anticuerpo primario durante la noche a 4°C. Se realizó la reacción posterior usando un kit LSAB+ (DakoCytomation, Carpinteria, CA, EE.UU.) y el procedimiento recomendado. Finalmente, se incubaron los portaobjetos con 3-amino-9-etilcarbazol (DakoCytomation, Carpinteria, CA, EE.UU.) y se contratiñeron con disolución de hematoxilina de Harris, modificada (Sigma-Aldrich, Inc., St. Louis, MO, EE.UU.).

En todos los casos, se observó inmunorreactividad en el citoplasma de células tumorales. Se evaluó la inmunorreactividad mediante el método de puntuación H. Se puntuó la intensidad de la tinción como 0, 1, 2 y 3 correspondiente a la presencia de tinción negativa, débil, intermedia y marrón fuerte, respectivamente. Se determinó el porcentaje de células que se teñían a diferentes intensidades, y se aplicó la siguiente fórmula: puntuación H = (% de células teñidas a puntuación de intensidad 1) + 2 X (% de células teñidas a puntuación de intensidad 2) + 3 X (% de células teñidas a puntuación de intensidad 3). Se analizaron las puntuaciones H de tejido tumoral y tejido no tumoral adyacente usando la prueba de los rangos con signo de Wilcoxon. Tal como se muestra en la figura 5, los tejidos tumorales tenían una expresión de LBFL313 significativamente superior que los tejidos no tumorales adyacentes (14 de los 16 casos) ($p < 0,05$).

Ejemplo 8

Efecto del anticuerpo anti-LBFL313 sobre la invasividad de células de cáncer pancreático

Para determinar el efecto del anticuerpo anti-LBFL313 sobre la invasión de líneas celulares de cáncer pancreático humano, se realizó un ensayo de cámara de Boyden en presencia de anticuerpo anti-LBFL313 o IgG de conejo normal. En resumen, se tripsinizaron cinco clases de líneas celulares de cáncer pancreático, CFPAC-1, MiaPaCa-2, PANC-1, AsPC-1 y BxPC-3, y se resuspendieron en disolución de inhibidor de tripsina (Sigma). Se sedimentaron las células y se suspendieron en medio libre de suero en presencia de PBS, anticuerpo anti-LBFL313 o IgG de conejo normal hasta una concentración final de $1-5 \times 10^6$ células/ml. Se llenaron los pocillos inferiores de la cámara con 30 μ l de medio convencional. Se ensambló la cámara usando filtros de policarbonato de 8 μ m de diámetro de tamaño de poro (Neuroprobe) recubiertos en el lado superior con 1 mg/ml de matriz de membrana basal Matrigel™ (BD biosciences). Se añadieron cincuenta μ l de suspensión celular al compartimiento superior. Se incubó la cámara a 37°C y CO₂ al 5% durante 24 h. Al final de la incubación, se retiraron mecánicamente las células en la superficie superior de los filtros. Se fijaron los filtros en metanol y se tiñeron en disolución modificada de tinción de Giemsa (Sigma). Se contó el número de células migradas por campo (100x) en el microscopio óptico (Olympus). Se sometió a ensayo cada muestra por triplicado. Tal como se muestra en la figura 6, el anticuerpo anti-LBFL313 inhibió eficazmente la invasividad de las líneas celulares de cáncer pancreático humano de una manera dependiente de la dosis. Se observaron efectos similares en líneas celulares de cáncer gástrico, AGS y N87.

Aplicabilidad industrial

Aunque la presente invención se ha descrito en detalle con referencia a los ejemplos anteriores, se entiende que pueden hacerse diversas modificaciones. Por consiguiente, la invención está limitada sólo por las siguientes reivindicaciones.

Lista de secuencias

<110> KOH, Sang Seok
 5 SONG, Si Young
 KIM, Sun-A
 LEE Yangsoon
 10 JEON, Sun Bok
 PARK, Euichul
 15 KIM, Young-Gun
 <120> FAMILIA GÉNICA (LBFL313) ASOCIADA CON CÁNCER PANCREÁTICO
 <130>
 20 <160> 2
 <170> FastSEQ para Windows Versión 4.0
 25 <210> 1
 <211> 777
 <212> ADN
 <213> *Homo sapiens*
 30 <220>
 <221> CDS
 <222> (53)...(643)
 <223> Clon LBFL313
 <400> 1
 35 cgcttcttcc ttctggatgg gggcccaggg ggcccaggag agtataaagg cg atg tgg
 agg gtg ccc ggc aca acc aga cgc cca gtc aca ggc gag agc cct ggg
 atg cac cgg cca gag gcc atg ctg ctg ctg ctc acg ctt gcc ctc ctg
 ggg ggc ccc acc tgg gca ggg aag atg tat ggc cct gga gga ggc aag
 tat ttc agc acc act gaa gac tac gac cat gaa atc aca ggg ctg cgg
 gtg tct gta ggt ctt ctc ctg gtg aaa agt gtc cag gtg aaa ctt gga
 gac tcc tgg gac gtg aaa ctg gga gcc tta ggt ggg aat acc cag gaa
 gtc acc ctg cag cca ggc gaa tac atc aca aaa gtc ttt gtc gcc ttc
 caa gct ttc ctc cgg ggt atg gtc atg tac acc agc aag gac cgc tat
 ttc tat ttt ggg aag ctt gat ggc cag atc tcc tct gcc tac ccc agc
 caa gag ggg cag gtg ctg gtg ggc atc tat ggc cag tat caa ctc ctt
 ggc atc aag agc att ggc ttt gaa tgg aat tat cca cta gag gag ccg
 acc act gag cca cca gtt aat ctc aca tac tca gca aac tca ccc gtg
 ggt cgc tag ggtggggtat ggggccatcc gagctgagge catctgtgtg
 gtgggtggctg atggtactgg agtaactgag tcgggacgct gaatctgaat ccaccaataa
 ataaagcttc tgcagaatca aaaaaaaaaa aaaa

<210> 2
 <211> 196
 <212> PRT
 5 <213> *Homo sapiens*
 <400> 2

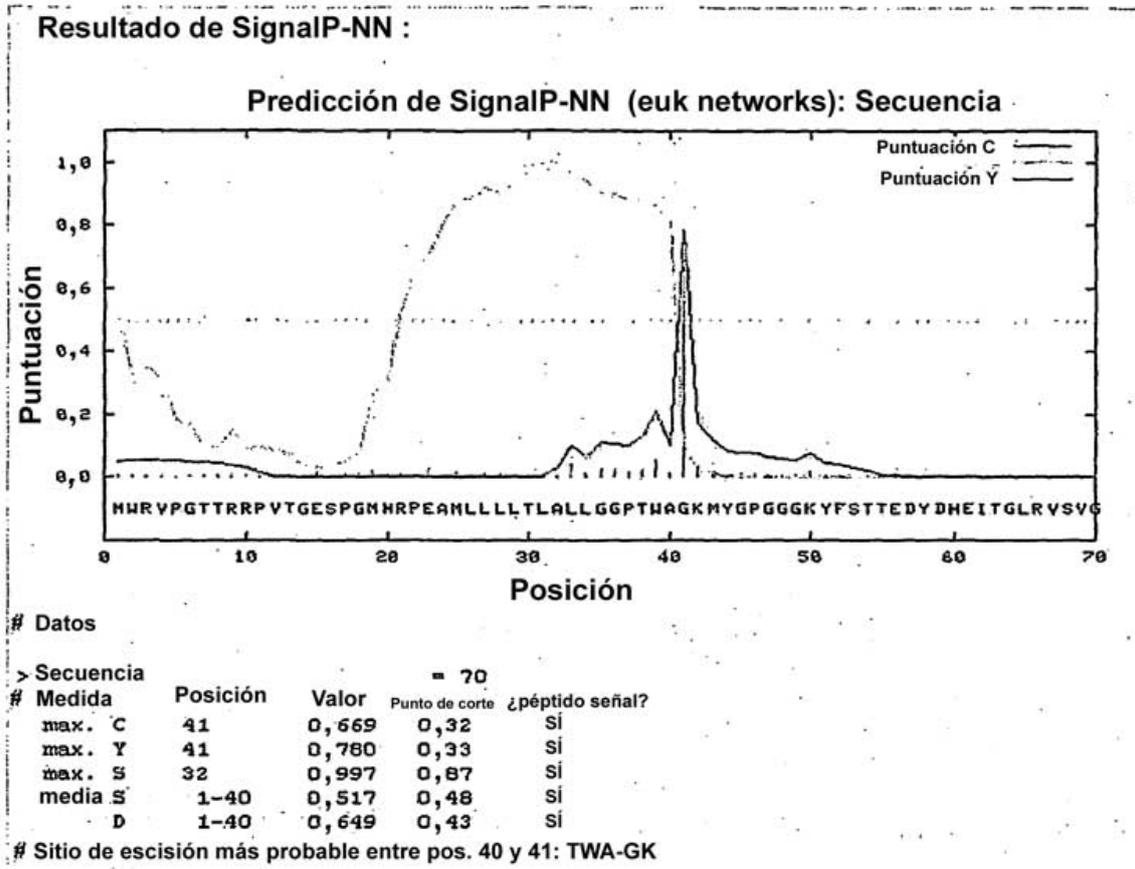
Met Trp Arg Val Pro Gly Thr Thr Arg Arg Pro Val Thr Gly Glu Ser
 Pro Gly Met His Arg Pro Glu Ala Met Leu Leu Leu Leu Thr Leu Ala
 Leu Leu Gly Gly Pro Thr Trp Ala Gly Lys Met Tyr Gly Pro Gly Gly
 Gly Lys Tyr Phe Ser Thr Thr Glu Asp Tyr Asp His Glu Ile Thr Gly
 Leu Arg Val Ser Val Gly Leu Leu Leu Val Lys Ser Val Gln Val Lys
 Leu Gly Asp Ser Trp Asp Val Lys Leu Gly Ala Leu Gly Gly Asn Thr
 Gln Glu Val Thr Leu Gln Pro Gly Glu Tyr Ile Thr Lys Val Phe Val
 Ala Phe Gln Ala Phe Leu Arg Gly Met Val Met Tyr Thr Ser Lys Asp
 Arg Tyr Phe Tyr Phe Gly Lys Leu Asp Gly Gln Ile Ser Ser Ala Tyr
 Pro Ser Gln Glu Gly Gln Val Leu Val Gly Ile Tyr Gly Gln Tyr Gln
 Leu Leu Gly Ile Lys Ser Ile Gly Phe Glu Trp Asn Tyr Pro Leu Glu
 Glu Pro Thr Thr Glu Pro Pro Val Asn Leu Thr Tyr Ser Ala Asn Ser
 Pro Val Gly Arg

10

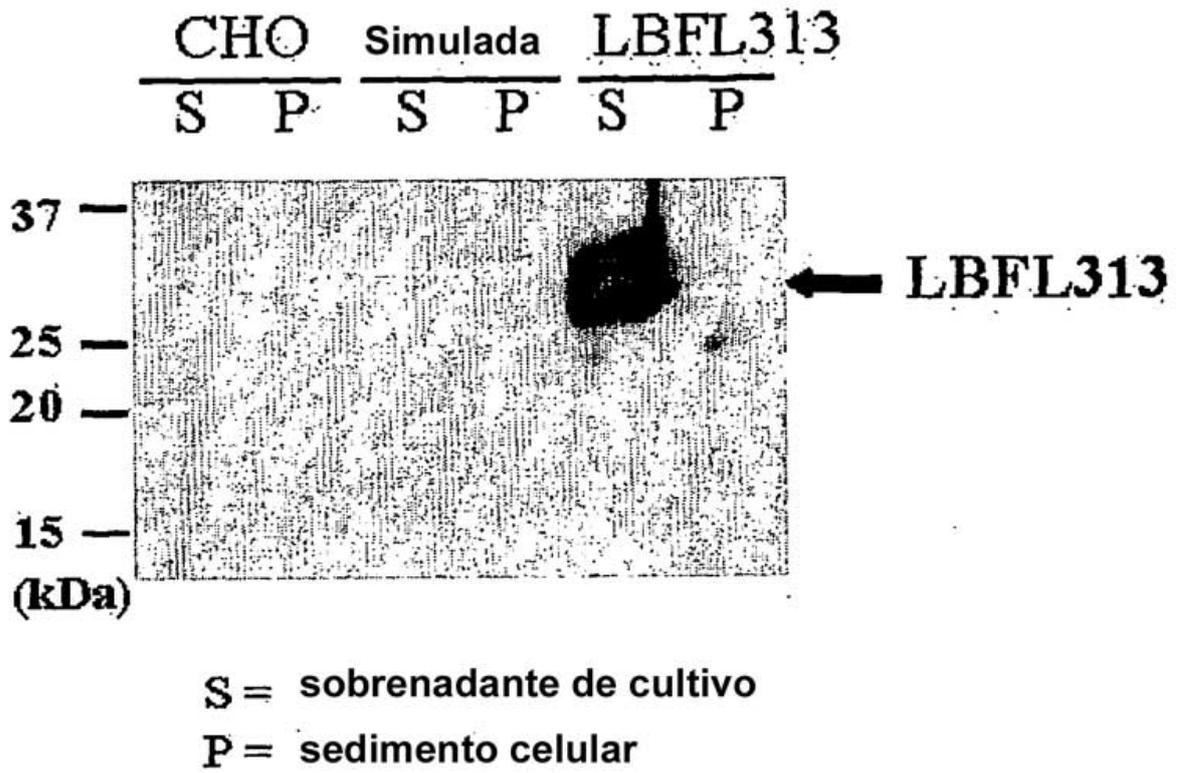
REIVINDICACIONES

1. Composición para diagnosticar cánceres pancreáticos, que comprende una molécula de ácido nucleico aislada seleccionada del grupo que consiste en: (a) una molécula de ácido nucleico aislada que comprende SEQ ID NO: 1, (b) una molécula de ácido nucleico aislada que codifica para SEQ ID NO: 2, (c) una molécula de ácido nucleico aislada que codifica para una proteína que se expresa en cáncer y que presenta una identidad de secuencia de nucleótidos de al menos aproximadamente el 95% a lo largo de la secuencia contigua completa de SEQ ID NO: 1, y (d) una molécula de ácido nucleico aislada que comprende el complemento de una molécula de ácido nucleico de (a), (b) o (c).
2. Composición según la reivindicación 1, en la que la molécula de ácido nucleico comprende los nucleótidos 53-640 de SEQ ID NO: 1.
3. Composición según la reivindicación 1, en la que la molécula de ácido nucleico comprende los nucleótidos 53-643 de SEQ ID NO: 1.
4. Composición según una cualquiera de las reivindicaciones 1 a 3, en la que dicha molécula de ácido nucleico está operativamente unida a uno o más elementos de control de la expresión.
5. Composición para diagnosticar cánceres pancreáticos, que comprende un vector que comprende una molécula de ácido nucleico aislada seleccionada del grupo que consiste en: (a) una molécula de ácido nucleico aislada que comprende SEQ ID NO: 1, (b) una molécula de ácido nucleico aislada que codifica para SEQ ID NO: 2, (c) una molécula de ácido nucleico aislada que codifica para una proteína que se expresa en cáncer y que presenta una identidad de secuencia de nucleótidos de al menos aproximadamente el 95% a lo largo de la secuencia contigua completa de SEQ ID NO: 1, y (d) una molécula de ácido nucleico aislada que comprende el complemento de una molécula de ácido nucleico de (a), (b) o (c).
6. Composición para diagnosticar cánceres pancreáticos, que comprende una célula huésped transformada para contener una molécula de ácido nucleico aislada seleccionada del grupo que consiste en: (a) una molécula de ácido nucleico aislada que comprende SEQ ID NO: 1, (b) una molécula de ácido nucleico aislada que codifica para SEQ ID NO: 2, (c) una molécula de ácido nucleico aislada que codifica para una proteína que se expresa en cáncer y que presenta una identidad de secuencia de nucleótidos de al menos aproximadamente el 95% a lo largo de la secuencia contigua completa de SEQ ID NO: 1, y (d) una molécula de ácido nucleico aislada que comprende el complemento de una molécula de ácido nucleico de (a), (b) o (c).
7. Composición para diagnosticar cánceres pancreáticos, que comprende una proteína o polipéptido aislado seleccionado del grupo que consiste en proteína que comprende la secuencia de aminoácidos de SEQ ID NO: 2 y una proteína que tiene una identidad de secuencia de aminoácidos de al menos aproximadamente el 95% con SEQ ID NO: 2.
8. Composición para diagnosticar o tratar cánceres pancreáticos, que comprende un anticuerpo aislado o fragmento de anticuerpo de unión a antígeno que se une a un polipéptido aislado seleccionado del grupo que consiste en una proteína que comprende la secuencia de aminoácidos de SEQ ID NO: 2 y una proteína que tiene una identidad de secuencia de aminoácidos de al menos aproximadamente el 95% con SEQ ID NO: 2.
9. Composición según una cualquiera según las reivindicaciones 1 a 8, en la que el cáncer pancreático es adenocarcinoma pancreático.

[Figura 1]

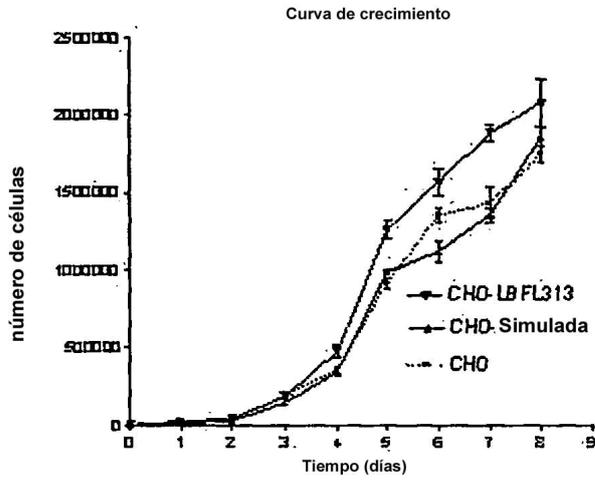


[Figura 2]

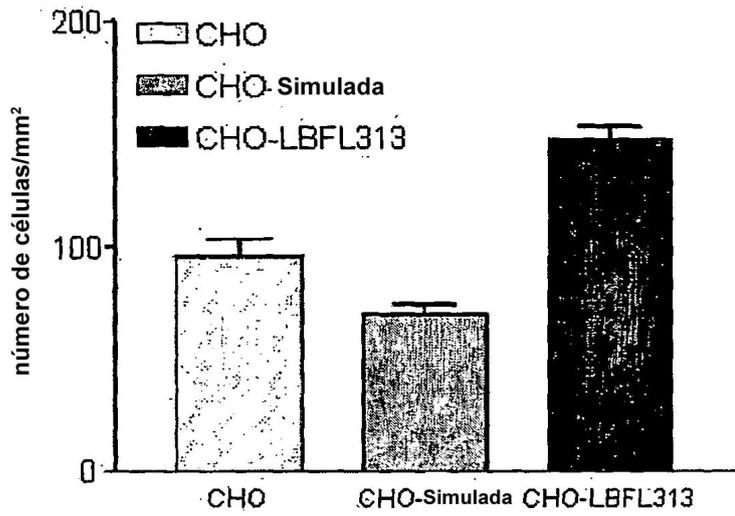


[Figura 3]

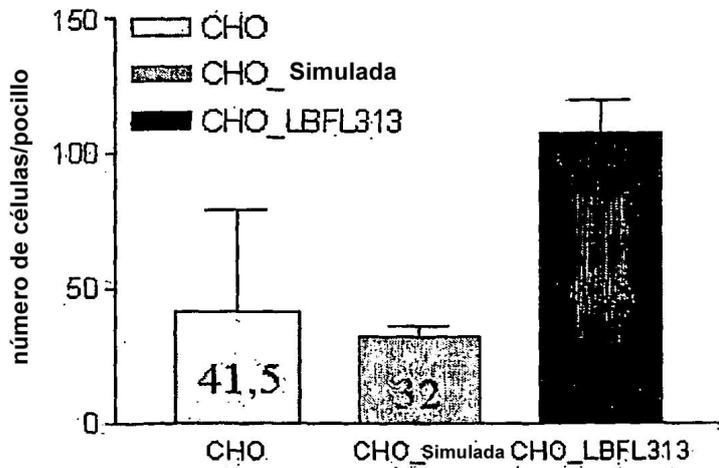
A. Proliferación



B. Migración



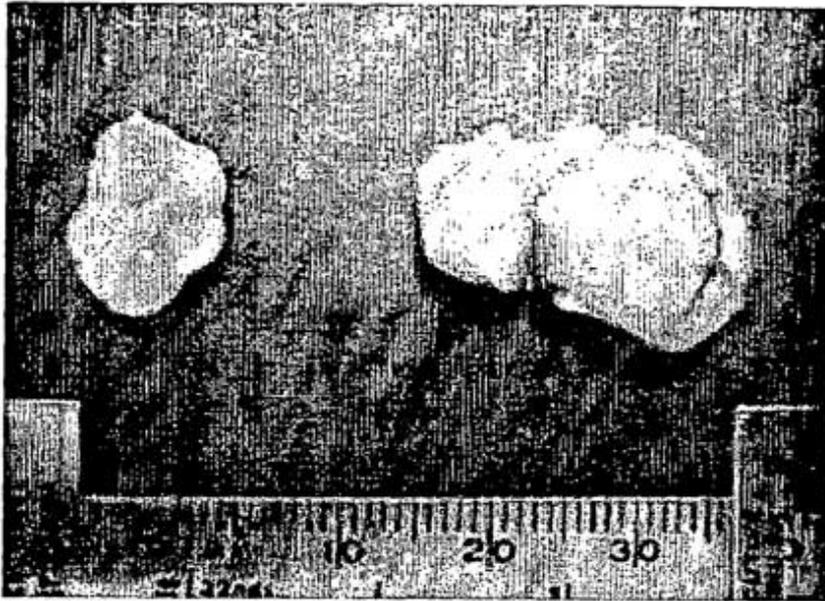
C. Invasión



[Figura 4]

A. Tamaño del tumor

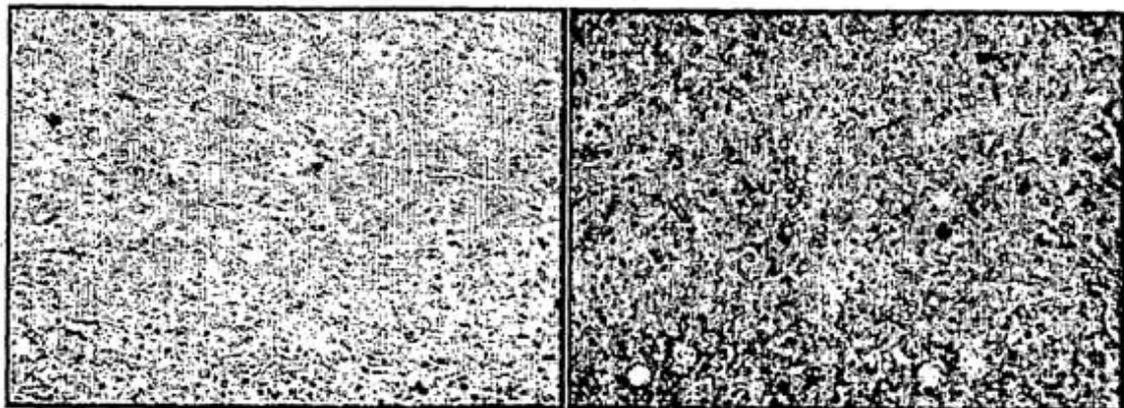
CHO_Simulada **CHO_LBFL313**



B. Tinción del factor VIII

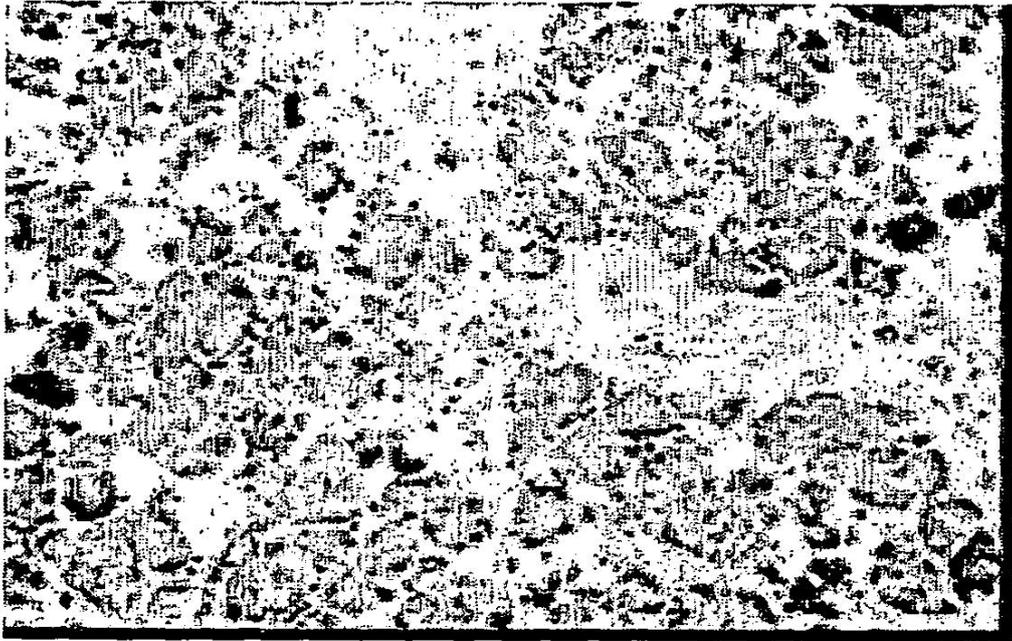
CHO_Simulada

CHO_LBFL313



[Figura 5]

Páncreas normal



Adenocarcinoma pancreático



[Figura 6]

