

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 386 317**

51 Int. Cl.:
C12N 15/11 (2006.01) **G06F 17/30** (2006.01)
C12N 15/12 (2006.01)
C12N 15/62 (2006.01)
C12N 15/86 (2006.01)
C12N 5/10 (2006.01)
C07K 14/47 (2006.01)
C07K 16/18 (2006.01)
A61K 31/713 (2006.01)
C12Q 1/68 (2006.01)
G01N 33/577 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **08010460 .7**
96 Fecha de presentación: **09.01.2001**
97 Número de publicación de la solicitud: **1985701**
97 Fecha de publicación de la solicitud: **29.10.2008**

54 Título: **Genes expresados diferencialmente en cáncer de mama**

30 Prioridad:
10.01.2000 US 175462 P

45 Fecha de publicación de la mención BOPI:
17.08.2012

45 Fecha de la publicación del folleto de la patente:
17.08.2012

73 Titular/es:
NOVARTIS VACCINES AND DIAGNOSTICS, INC.
4560 HORTON STREET
EMERYVILLE, CA 94608, US

72 Inventor/es:
Kaufmann, Joerg;
Harrowe, Greg;
Reinhard, Cristoph y
Kang, Sanmao

74 Agente/Representante:
Carpintero López, Mario

ES 2 386 317 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

Genes expresados diferencialmente en cáncer de mama

Campo técnico

5 La presente invención se refiere a procedimientos para predecir el comportamiento de tumores. Más particularmente, la invención se refiere a procedimientos en los que se examina una muestra tumoral para expresión de una secuencia de genes específicos de este modo para indicar propensión a dispersión metastásica.

Antecedentes de la invención

10 El cáncer de mama es una de las más comunes enfermedades malignas con aproximadamente 1.000.000 de nuevos casos por año en todo el mundo. A pesar de la utilización de un número de marcadores histoquímicos, genéticos e inmunológicos, los médicos tienen aún dificultades para predecir a tiempo qué tumores metastatizarán a otros órganos. Algunos pacientes están en la necesidad de terapia coadyuvante para evitar recurrencia y metástasis y otros no. Sin embargo, distinguir entre estas subpoblaciones de pacientes no es sencillo y el curso de tratamiento no es fácil de someter a seguimiento. Hay una necesidad en la técnica de marcadores nuevos para distinguir entre los tumores que metastatizarán o han metastatizado y aquellos que es menos probable que metastatizen.

15 **Sumario de la invención**

En un primer aspecto, la invención proporciona una molécula de ácido nucleico aislada, preferentemente una molécula de ADN aislada, que comprende un polinucleótido seleccionado del grupo que consiste en:

- (a) un polinucleótido que codifica los aminoácidos del 1 al 273 de SEC ID N.º: 2;
- (b) un polinucleótido que codifica los aminoácidos del 2 al 273 de SEC ID N.º: 2;
- 20 (c) un polinucleótido que codifica los aminoácidos del 26 al 273 de SEC ID N.º: 2;
- (d) el complemento de polinucleótidos del polinucleótido de (a), (b) o (c); y
- (e) un polinucleótido al menos el 90 % idéntico al polinucleótido de (a), (b), (c), o (d) que se expresa diferencialmente entre las células de cáncer de mama con capacidad de metástasis alta y las células de cáncer no metastáticas o con capacidad metastásica baja.

25 En un aspecto adicional, la invención proporciona una molécula de ácido nucleico aislada que comprende un polinucleótido que codifica un polipéptido en el que, salvo por al menos una sustitución de aminoácido conservadora, dicho polipéptido tiene una secuencia de aminoácidos seleccionada del grupo que consiste en:

- (a) los aminoácidos 1 a 273 de SEC ID N.º: 2;
- (b) los aminoácidos 2 a 273 de SEC ID N.º: 2; y
- 30 (c) los aminoácidos 26 a 273 de SEC ID N.º: 2;

en la que dicho polipéptido se expresa diferencialmente entre las células de cáncer de mama con capacidad de metástasis alta y las células de cáncer no metastáticas o con capacidad de metástasis baja.

35 En un aspecto aún adicional, la invención se refiere a un procedimiento de preparar un vector recombinante que comprende insertar una molécula de ácido nucleico según se define anteriormente dentro de un vector en unión operable a un promotor.

La invención también proporciona un vector recombinante que comprende una molécula de ácido nucleico según se define anteriormente en unión operable a un promotor.

Un procedimiento de preparar de una célula huésped recombinante introduciendo el vector recombinante dentro de una célula huésped y las células huésped que comprenden tales vectores recombinantes se proporcionan también.

40 En otro aspecto, la invención se refiere a un polipéptido aislado que comprende aminoácidos al menos el 95% idénticos a aminoácidos seleccionados del grupo que consiste en:

- (a) los aminoácidos 1 a 273 de SEC ID N.º: 2;
- (b) los aminoácidos 2 a 273 de SEC ID N.º: 2; y
- (c) los aminoácidos 26 a 273 de SEC ID N.º: 2,

45 en la que dicho polipéptido se expresa diferencialmente entre las células de cáncer de mama con capacidad de

metastásis alta y las células de cáncer no metastásicas o con capacidad de metastásis baja.

En un aspecto aún adicional, la invención se refiere a una parte portadora de epítipo del polipéptido de la SEC ID N.º: 2 (con la condición de que dicha parte que lleva epítipo no sea los aminoácidos 85-273 de la SEC ID N.º: 2 o un fragmento de la misma) y a un anticuerpo aislado que se une específicamente a uno de los polipéptidos definidos anteriormente (con la condición de que dicho anticuerpo no se una a los aminoácidos 85-273 de SEC ID N.º: 2).

Finalmente, la invención proporciona un procedimiento para comparar el potencial metastásico de las células tumorales, preferentemente células de cáncer de mama, en una primera y una segunda muestra, comprendiendo medir en dichas muestras de tejido un producto de expresión de un gen que comprende un polinucleótido de la invención.

10 **Breve descripción de los dibujos**

La figura 1 ilustra la secuencia de polinucleótidos de Out at First de seres humanos (SEC ID N.º: 1).

La figura 2 ilustra el aminoácido de la secuencia codificada por SEC ID N.º: 1 (SEC ID N.º: 2).

La figura 3 ilustra el péptido señalizador teórico (SEC ID N.º: 3).

15 La figura 4 ilustra la traducción de la SEC ID N.º: 1 (SEC ID N.º: 1, polinucleótido; SEC. ID. N.º: 2, secuencia de aminoácidos).

La figura 5 ilustra la expresión de hsOAF en relación con β -actina en líneas celulares tumorales y tejidos tumorales de ratones SCID desarrollados a partir de las líneas celulares. "PT" se refiere a tumor primario.

20 La figura 6 ilustra el crecimiento de colonias por células MDA-MB-435 en agar blando tras el tratamiento con oligo antisentido SEC ID N.º: 4 (66-2as) o con control reverso de SEC ID N.º: 5 (66-2rc), en relación a las células no tratadas (TS).

La figura 7 es una alineación de la secuencia de aminoácidos OAF humana con la secuencia de aminoácidos OAF de *Drosophila*.

Figura 8. La figura 8A ilustra la expresión de proteína hsOAF en líneas celulares COS-7 y MCF-7. La figura 8B ilustra la expresión de proteína hsOAF en líneas celulares de carcinoma mamario.

25 La figura 9 ilustra la expresión de hsOAF en tejidos humanos normales.

Figura 10. La figura 10A ilustra los cambios morfológicos vistos en células MDA-MB-435 tras el tratamiento con oligo antisentido (SEC ID N.º: 4). AS = antisentido; RC = control reverso (SEC ID N.º: 5); M = medio acondicionado. La figura 10B ilustra la invasión celular tras el tratamiento de células MDA-MB-435 con AS, RC y RC + M.

La figura 11 ilustra la secuencia señal predicha de OAF humana (subrayado doble).

30 Figura 12. La figura 12A y 12B ilustran la secreción de hsOAF por células MDA-MB-435 tratadas con oligo antisentido (SEC ID N.º: 4) u oligo control reverso (SEC ID N.º: 5).

Descripción detallada de la invención

35 La causa más letal en pacientes con cáncer de mama es la metástasis de carcinomas de mama y su proliferación en loci distantes (pulmón y hueso, principalmente). La metástasis en un procedimiento de varias etapas por el que las células tumorales emigran a partir del tumor primario, se diseminan a través de la sangre y los vasos linfáticos y después se depositan en órganos objetivo específicos donde recolonizan. Schirmacher, V., Adv. Cancer Res. 43:1-73, 1985 y Liotta, L.A. y cols., Cell 64 (2): 327-36 (1991). Durante este procedimiento la invasividad de células tumorales es crucial, ya que deben encontrar y atravesar numerosas membranas basales. Liotta L.A., Am. J. Pathol. 117 (3): 339-48 (1984) y Fidler, I.J., Cancer Res. 38(9):2651-60 (1978). Por lo tanto, la elucidación de las causas moleculares de la invasión celular tumoral y la metástasis es esencial para el desarrollo de procedimientos de tratamiento eficaces para pacientes de cáncer de mama. Los genes expresados diferencialmente en la metástasis tumoral de mama son objetivos potenciales que juegan papeles críticos durante la metástasis. La identificación de tales genes y su función biológica contribuirá de manera significativa al desarrollo de terapia y diagnóstico de cáncer de mama.

45 Se han descubierto algunos genes importantes involucrados en la metástasis tumoral de mama. La pérdida del receptor de estrógenos y presencia de vimentina se han asociado con invasividad tumoral mamaria humana y mal pronóstico y también correlacionan con la invasividad y el potencial metastásico de líneas celulares de cáncer de mama humano. Aamdal S., y cols., Cancer 53 (11): 2525-9 (1984); Clark, G.M., y cols., Semin Oncol., 2, Supl. 1: 20-5 (1988); Raymond, W.A. y cols., J. Pathol. 157 (4): 299-306 (1989); Raymond, W.A., y cols., J. Pathol 158 (2): 107-14 (1989); y Thompson, E.W. y cols., J. Cell Physiol. 150(3):534-44 (1992). La subexpresión de e-cadherina se ha implicado en la invasividad de tumores mamarios. Vleminckx, K., y cols., Cell 66 (1): 107-19 (1991) y Oka, H., y cols.,

- 5 Cancer Res. 53(7):1696-701 (1993). Maspin, un inhibidor de proteasa expresado en células epiteliales mamarias normales pero no en la mayoría de las líneas celulares de carcinoma de mama, ha sido capaz de suprimir la capacidad de las células MDA-MB-435 para inducir tumores y metastatizar en ratones y para invadir membrana basal *in vitro*. La pérdida de expresión de maspin ocurrió lo más frecuentemente en cánceres avanzados. Zou, Z., y cols., Science 263 (5146): 526-9 (1994) y Seftor, R.E., y cols., Cancer Res. 58 (24): 5681-5 (1998).
- 10 La sobreexpresión de TTMP-4 (inhibidor de tejido de metaloproteinasas-4) o CLCA2 (Ca²⁺ canal de cloruro activado 2) en células MDA-MB-435 por transfección inhibió la tumorigénesis, invasividad y capacidad metastásica de las células. Wang, M., y cols., Oncogene 14 (23): 2767-74 (1997) y Gruber, A.D., y cols., Cancer Res. 59(21):5488-91 (1999). La sobreexpresión de los receptores de factores de crecimiento IGF-IR y p185^{ErbB-2} se ha encontrado que está implicada en metástasis de cáncer de mama. Surmacz, P., y cols., Breast Cancer Res. Treat 47 (3): 255-67 (1998); Dunn, S.E., y cols., Cancer Res. 58 (15): 3353-61 (1998); Tan, M., y cols., Cancer Res. 57 (6): 1199-205 (1997); Dhingra, K., y cols., Semin Oncol. 23 (4): 436-45 (1996); y Revillion, F., y cols., Eur. J. Cancer 34 (6): 791-808 (1998).
- 15 Se ha informado que la aspartilproteasa catepsina D es un marcador de mal pronóstico para pacientes de cáncer de mama y hay una correlación significativa entre la concentración de catepsina D alta en el citosol del cáncer de mama primario y el desarrollo de metástasis, aunque no se encontró ninguna correlación entre secreción de catepsina D y capacidad de invasión de las líneas celulares de cáncer de mama. Rochefort, H., Breast Cancer Res Treat 16 (1): 3-13 (1990); Johnson, M.D., y cols., Cancer Res. 53 (4): 873-7 (1993); y Rochefort, H., y cols., Clin Chim Acta. 291(2):157-70 (2000). Osteopontina, una glucoproteína de unión a integrinas segregada que se piensa que está implicada en resorción de hueso y formación de hueso, puede inducir migración e invasión de células de carcinoma mamario. Los niveles de osteopontina (célula tumoral o niveles plasmáticos) se han asociado con malignidad potenciada de cáncer de mama. Denhardt, D.T., y cols., FASEB J 7 (15):1475-82 (1993); Denhardt, D.T., y cols., J. Cell Biochem Suppl., 30-31: 92-102 (1998); Tuck, A.B., y cols., J. Cell Biochem. 78 (3): 465-75 (2000); Tuck, A.B. y cols., Oncogene 18 (29): 4237-46 (1999); y Singhal, H. y cols., Clin Cancer Res. 3 (4): 605-11 (1997).
- 20
- 25 La invención se refiere a la clonación de un gen novedoso sobreexpresado en líneas celulares de cáncer de mama humano altamente metastásicas. Ello codifica una proteína segregada y se ha confirmado que su secreción proteica es mucho mayor en líneas celulares de cáncer de mama humano altamente metastásicas que en líneas celulares de capacidad metastásica baja/no metastásicas. La desactivación de la secreción de esta proteína de la línea celular agresiva MDA-MB-435 por tecnología de oligos antisentido dio como resultado alteración morfológica junto con
- 30 invasión reducida y velocidad de proliferación de las células reducida. El gen se llama hsOAF basado en su homología con el gen OAF de *Drosophila* (out at first). Bergstrom, D.E., y cols., Genetics 139 (3): 1331-46 (1995) y Merli, y cols., Genes Dev. 10 (10): 1260-70 (1996)
- 35 Es un descubrimiento de la presente invención que un polinucleótido se expresa diferencialmente entre las células de cáncer de mama con alta capacidad metastásica y las células cancerosas no metastásicas o con baja capacidad metastásica. El polinucleótido por lo tanto se refiere a un gen marcador metastásico. Esta información puede utilizarse para fabricar reactivos de diagnóstico específicos para los productos de expresión del gen expresado diferencialmente. Ello se puede usar también en procedimientos de diagnóstico y de pronóstico que ayudarán a los médicos en planear regímenes de tratamiento apropiados para cánceres, especialmente de la mama.
- 40 El polinucleótido se muestra en la figura 1 (SEC ID N.º: 1) y el marco de lectura abierto predicho (ORF) codifica un polipéptido mostrado en la figura 2 (SEC ID N.º: 2). Los primeros 30 residuos aminoácidos (SEC ID N.º: 3) comprenden un péptido señal teórico, con un sitio de escisión de la proteasa predicho indicada por "***": APLLG * TGAPA (entre aminoácidos en las posiciones 25 y 26 de la SEC ID N.º: 3).
- 45 La secuencia de polinucleótidos de la invención, muestra alguna homología con un gen de *Drosophila* conocido como "Out at First" (oaf). La transcripción de oaf da como resultado tres clases de ARN poliadenilados alternativamente, la expresión de los cuales se reguló según el desarrollo. Todos los transcritos de oaf contienen dos ORF adyacentes separados por un único codón de parada UGA. La supresión del codón UGA durante la traducción podría conducir a la producción de proteínas diferentes a partir de la misma molécula de ARN. Durante la oogénesis, el ARN de oaf se expresa en células asociadas a oocito de todas las edades y se aporta maternalmente al huevo.
- 50 Durante el desarrollo embrionario, la transcripción cigótica del gen oaf tiene lugar en pequeños racimos de células en la mayoría de todos los segmentos al tiempo de la extensión de banda germinal y más tarde en un patrón repetido segmentalmente en el sistema nervioso central en desarrollo. El gen oaf también se expresa en las gónadas embrionarias, larvarias y adultas de ambos sexos. (Bergstrom, D.E. y cols., Genetics 139: 1331-1346, 1995).
- 55 El polinucleótido de la invención se expresó diferencialmente en siete pares de líneas celulares de cáncer de mama con alta capacidad metastásica o líneas celulares de cáncer de mama con baja capacidad metastásica. Las líneas celulares usadas son MDA-MB-361 (derivada de adenocarcinoma de mama humano), MDA-MB-231 (células de cáncer de mama humano); MDA-MB-468 (derivada de células de cáncer de mama negativas en receptor de estrógeno humano); MCF-7 (células de cáncer de mama humano no metastásicas); ZR-75-1 (derivada de carcinomas de mama humanos positivos para receptores de estrógenos), Engle y cols., Cancer Res. 38: 3352-64

(1978)); y MDA-MB-435 (derivada de células de carcinoma de mama humano negativas en receptor de estrógenos humanos, Rishi y cols., Cancer Res. 56: 5246-5252 (1996)).

El perfil de expresión es como sigue:

Tabla 1

Par de líneas celulares	Relación de expresión
MDA-MB-361/MDA-MB-231	0,11
MDA-MB-468/MDA-MB-231	0,44
MCF-7/MDA-MB-231	0,17
ZR-75-1/MDA-MB-231	0,12
MDA-MB-361/MDA-MB-435	0,06
MDA-MB-468/MDA-MB-435	0,36
MCF-7/MDA-MB-435	0,03

5

La regulación al alza de la expresión de ARNm humano se confirmó por análisis de bandas de northern usando el ARN total de las líneas celulares (figura 5).

10 Las líneas celulares en las que se comparó la expresión del polinucleótido de la invención representan los cánceres de mama humanos de potencial metastásico variable. Los cultivos de línea celular ZR-75-1 se derivaron de efusión ascítica maligna de un paciente de cáncer de mama. Las líneas celulares cultivadas *in vitro* se parecen mucho a la morfología vista en biopsias o preparaciones celulares a partir de los donantes de las células originales. Las células ZR-75-1 se estimulan específicamente por estrógeno y han sido usados como un sistema modelo para estudiar responsividad a estrógenos. Engel, L.W. y cols., Cancer Res. 38: 3352-3364, 1978.

15 La línea celular MDA-MB-435 es una línea celular negativa en receptor de estrógenos que se ha estudiado como un modelo de cáncer de mama humano, por ejemplo, para estudiar el mecanismo de acción de la inhibición del crecimiento en presencia de ácido retinoico. Rishi, A. K. y cols., Cancer Res. 56: 5246-5252, 1996. La inhibición de crecimiento por retinoides se ha estudiado también en células MCF-7 y en células MDA MB 468. Tin-U, C. K. y cols., Am. Soc. Clin. Onc. Proceedings, Vol. 17, 2125, 1998.

20 La línea celular MDA-MB-361 se derivó de un adenocarcinoma mamario humano, específicamente desde un sitio maligno. Número de ATCC HTB-27. La expresión diferencial de genes Wnt de ser humano se ha estudiado en esta línea celular. Huguet, E.L. y cols., Cancer Res. 54:2615-2621, 1994.

25 Una vez la metástasis tiene lugar, las células tumorales primarias mamarias invaden las membranas basales y se dispersan a otros órganos del cuerpo y la probabilidad de supervivencia de pacientes con cáncer de mama llega a ser bastante escasa. Es esencial identificar genes que participan en la invasión y la metástasis de cáncer de mama en nombre del diagnóstico clínico y la terapia. Tales genes son marcadores potenciales para diagnosis u objetivos candidatos del desarrollo de fármacos terapéuticos. Por ejemplo, la presencia de vimentina en tumor de mama humano se ha asociado con carencia de receptor de estrógenos e invasividad tumoral como un marcador de pronóstico pobre. Raymond, W.A. y cols., J. Pathol. 157 (4): 299-306 (1989); Raymond, W.A., y cols., J. Pathol. 158 (2): 107-14 (1989); y Thompson, E.W. y cols., J. Cell Physiol. 150(3):534-44 (1992). Actividades incrementadas de metaloproteinasas de matriz están relacionadas con el fenotipo metastásico de carcinomas, especialmente cáncer de mama. Basset, P., y cols., Nature 348 (6303): 699-704 (1990) y Basset, P., y cols., Cancer 74 (3 Supl.): 1045-9 (1994). Osteopontina, una glucoproteína de unión a integrina segregada, es capaz de inducir invasividad incrementada de células epiteliales mamarias humanas y se ha asociado con malignidad potenciada en cáncer de mama. Tuck, A.B. y cols., J. Cell Biochem. 78 (3): 465-75 (2000); Tuck, A.B. y cols., Oncogene 18 (29): 4237-46 (1999); y Singhal, H. y cols., Clin Cancer Res. 3 (4): 605-11 (1997).

35 La invención se refiere a la identificación de una proteína segregada novedosa (hsOAF) implicada en metástasis de cáncer de mama. Las líneas celulares de cáncer de mama humano usadas en dilucidar el papel de la proteína se dividen en tres grupos de acuerdo con su capacidad metastásica: altamente metastásicas, de baja capacidad metastásica y no metastásicas. Tomando ventaja de diferentes potenciales metastásicos entre estos grupos de línea celular y utilizando la tecnología de microensayos avanzada, se identificaron genes que se expresan diferencialmente entre líneas celulares de cáncer de mama humanas altamente metastásicas y las de baja capacidad metastásica/no metastásicas. Estos genes son buenos candidatos para estudios de validación de la expresión y de función que podrían conducir a mejor comprensión del mecanismo molecular de metástasis de cáncer de mama. El gen hsOAF es el centro de esta invención ya que codifica una proteína segregada novedosa. La

40

expresión de ARNm de hsOAF es bastante común en diversos tejidos humanos normales. Sin embargo, la secreción de proteína hsOAF es mucho más fuerte por las líneas celulares de cáncer de mama metastásicas humanas que por las de la baja capacidad metastásica/no metastásicas y MDA-MB-435 tiene la secreción de hsOAF más grande. El gen de hsOAF se nombró en base a su homología con el gen de OAF de *Drosophila* (out at first). Sin embargo, la proteína OAF de *Drosophila* puede no ser una proteína segregada dado que no posee una secuencia de péptido señal típica en el extremo N-terminal (fig. 7).

Para tratar la importancia de la proteína hsOAF segregada en metástasis de cáncer de mama, se usó tecnología de oligos antisentido para inactivar específicamente la expresión de hsOAF. La tecnología de oligo antisentido es una manera eficiente, rápida para reducir drásticamente la expresión génica para estudios funcionales de genes. Stein, C.A., y cols., *Science* 261 (5124): 1004-12 (1993); Defacque, H. y cols., *J. Cell Physiol.* 178(7):109-19 (1999). La desactivación de secreción de proteínas HsOAF de células MDA-MB-435 altamente metastásicas dio como resultado cambio en la forma celular, invasividad celular reducida y proliferación celular más lenta. El tratamiento de células con el medio acondicionado (medio de cultivo de células MDA-MB-435 normales) fue capaz de recuperar todas aquellas alteraciones fenotípicas causadas por la desactivación de secreción de proteínas hsOAF en algún grado. Aunque los autores de la presente invención no se adhieren a un mecanismo específico, se cree que la proteína hsOAF segregada es esencial para la invasividad y proliferación de células MDA-MB-435. Sin embargo, la desactivación de la secreción de las proteínas hsOAF de otra línea celular altamente metastásica, MDA-MB-231, por tecnología de oligos antisentido no causó cambio celular significativo alguno. MDA-MB-435 y MDA-MB-231 son líneas celulares metastásicas bastante diferentes y MDA-MB-435 muestra secreción de proteínas hsOAF mucho más fuerte de lo que lo hace MDA-MB-213.

El gen de hsOAF se sitúa en la región 11q23 cromosómica donde la pérdida de heterocigosis tiene lugar frecuentemente en tumores de mama humanos. Negrini, M., y cols., *Cancer Res* 55 (14): 3003-7 (1995) y Tomlinson, I.P., y cols., *J. Clin. Pathol.* 48 (5): 424-8 (1995). Se ha comunicado que la pérdida de heterocigosis en 11q23 en tumores de mama humanos primarios está asociada con supervivencia pobre después de metástasis. Winqvist, R., y cols., *Cancer Res.* 55(12):2660-4 (1995). 11q23 también contiene loci tales como ATM (ataxia-telangiectasia, mutado) y MLL (que se desbarata frecuentemente por reordenación cromosómica en leucemia aguda). Rasio, D., y cols., *Cancer Res.* 55 (24): 6053-7 (1995) y Rubnitz, J.E., y cols., *Leukemia* 10 (1): 74-82 (1996). La relación entre mutación q23 en el cromosoma 11 y expresión génica HsOAF en metástasis de cáncer de mama permanece sin aclarar.

La proteína hsOAF segregada puede ser un objetivo adecuado para desarrollo de fármacos contra cáncer de mama y un buen marcador de diagnóstico para la malignidad de tumor mamario. La SEC ID N.º: 1 y los polinucleótidos que comprenden esta secuencia son por lo tanto útiles como marcadores metastásicos. La referencia a nucleótidos marcadores metastásicos o a secuencias aminoacídicas incluye variantes que tienen patrones de expresión similares en células con alta capacidad metastásica en relación a células no metastásicas o con baja capacidad metastásica. Los polipéptidos marcadores metastásicos pueden diferir en longitud de proteínas marcadoras metastásicas de longitud completa y pueden contener al menos 6, 8, 10, 12, 15, 18, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 120, 140, 160, 180, 200, 220, 240, 260, 265, 270 o 271 o más aminoácidos contiguos de una proteína marcadora metastásica. Los polinucleótidos ejemplares incluyen aquellos que codifican aminoácidos desde aproximadamente 1 hasta aproximadamente 273; desde 1 hasta 273; desde aproximadamente 2 hasta aproximadamente 273; desde 2 hasta 273; desde aproximadamente 26 hasta aproximadamente 273; y desde 26 hasta 273 de la SEC ID N.º: 2.

Pueden aparecer también variantes de las proteínas marcadoras y los polipéptidos marcadores. Pueden aparecer de forma natural o de forma no natural variantes de proteínas o polipéptidos marcadores de metástasis. Se encuentran variantes proteicas o polipeptídicas de marcadores metastásicos que aparecen de forma natural en seres humanos u otras especies y comprenden secuencias aminoacídicas que son sustancialmente idénticas a una proteína codificada por un gen correspondiente a la secuencia nucleotídica mostrada en SEC ID N.º: 1 o su complemento. Las variantes proteicas o polipeptídicas marcadoras metastásicas que no aparecen de forma natural que retienen sustancialmente los mismos patrones de expresión diferencial en células de cáncer de mama con alta capacidad metastásica en relación a células de cáncer de mama con baja capacidad metastásica o a células de cáncer de mama no metastásicas al igual que las variantes proteicas o polipeptídicas marcadoras metastásicas se incluyen también aquí. Preferentemente, las variantes proteicas o polipeptídicas marcadoras metastásicas que aparecen de forma natural o que aparecen de forma no natural tienen secuencias aminoacídicas que son al menos el 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, o 95 % idénticas a secuencias aminoacídicas codificadas por la secuencia de nucleótidos mostrada en SEC ID N.º: 1. Más preferentemente, las moléculas son al menos el 96 %, 97 %, 98 % o 99 % idénticas. La identidad de secuencia porcentual entre una proteína o polipéptido de tipo silvestre y una variante se determina alineando la proteína o polipéptido de tipo silvestre con la variante para obtener el mayor número de coincidencias aminoacídicas, como se conoce en la técnica, contar el número de coincidencias aminoacídicas entre el tipo silvestre y la variante y dividir el número total de coincidencias por el número total de residuos aminoacídicos de la secuencia de tipo silvestre.

Preferentemente, cambios aminoacídicos en las variantes proteicas o polipeptídicas marcadoras metastásicas son cambios aminoacídicos conservadores, es decir, sustituciones de aminoácidos cargados o no cargados. Un cambio aminoacídico conservador implica la sustitución de uno de una familia de aminoácidos que están relacionados en

5 sus cadenas laterales. Los aminoácidos que se dan en la naturaleza se dividen generalmente en cuatro familias: aminoácidos ácidos (aspartato, glutamato), básicos (lisina, arginina, histidina), no polares (alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano) y polares no cargados (glicina, asparragina, glutamina, cistina, serina, treonina, tirosina). Fenilalanina, triptófano y tirosina se clasifican algunas veces de forma conjunta como aminoácidos aromáticos.

10 Es razonable esperar que la sustitución aislada de una leucina con una isoleucina o valina, un aspartato con un glutamato, una treonina con una serina, o la sustitución similar de un aminoácido con un aminoácido estructuralmente relacionado no tendrá un efecto considerable en las propiedades biológicas de la variante proteica o polipeptídica marcadora metastásica resultante. Las propiedades y funciones de las variantes proteicas o polipeptídicas marcadoras metastásicas son del mismo tipo que una proteína o polipéptido marcador metastásico resultante que comprende secuencias aminoacídicas codificadas por la secuencia de nucleótidos mostrada en SEC ID N.º: 1, aunque las propiedades y funciones de las variantes pueden diferir en grado. Si un cambio aminoacídico da como resultado una variante proteica o polipeptídica marcadora metastásica con el patrón de expresión diferencial apropiado puede determinarse fácilmente. Por ejemplo, las sondas nucleotídicas pueden seleccionarse a partir de las secuencias de genes marcadoras reveladas en el presente documento y usadas para detectar un ARNm en prueba de bandas de northern o en secciones de tejidos, como se conoce en la técnica. Alternativamente, los anticuerpos que se unen específicamente a productos proteicos de genes marcadores metastásicos se pueden usar para detectar expresión de proteínas marcadoras metastásicas o variantes de las mismas.

20 Las variantes de marcadores metastásicos incluyen formas glucosiladas, conjugados agregados con otras moléculas y conjugados covalentes con restos químicos no relacionados. Las variantes de marcadores metastásicos incluyen también variantes alélicas, variantes de especies y muteínas. Las truncaciones o deleciones de regiones que no afectan la expresión diferencial de los genes marcadores metastásicos son también variantes de marcadores metastásicos. Las variantes covalentes se pueden preparar por funcionalidades de unión a grupos que se encuentran en la cadena de aminoácidos o en el residuo N-terminal o C-terminal, como se conoce en la técnica.

25 Se reconocerá en la técnica que alguna secuencia de aminoácidos del polipéptido de la invención se puede variar sin efecto significativo sobre la estructura o función de la proteína. Si tales diferencias en la secuencia se contemplan, debería recordarse que hay áreas críticas en la proteína que determinan actividad. En general, es posible reemplazar residuos que forman la estructura terciaria, dado que los residuos que llevan a cabo una función similar se usan. En otros casos, el tipo de residuo puede ser totalmente despreciable si la alteración tiene lugar en una región no crítica de la proteína. La sustitución de aminoácidos puede también cambiar la selectividad de unión a receptores de superficie celular. Ostade y cols., Nature 361: 266-268 (1993) describe ciertas mutaciones que dan como resultado unión selectiva de TNF-alfa a sólo uno de los dos tipos conocidos de receptores de TNF. Así, los polipéptidos de la presente invención pueden incluir una o más sustituciones, deleciones o adiciones aminoacídicas, bien a partir de mutaciones naturales o bien a partir de manipulación humana.

35 La invención incluye adicionalmente variaciones del polipéptido reveladas que muestran patrones de expresión comparables o que incluyen regiones antigénicas. Tales mutantes incluyen deleciones, inserciones, inversiones, repeticiones y sustituciones de tipo. Se puede encontrar orientación referente a qué cambios aminoacídicos es probable que sean fenotípicamente silenciosos en Bowie, J.U., y cols., "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions", Science 247: 1306-1310 (1990).

40 De particular interés son sustituciones de aminoácidos cargados con otro aminoácido cargado y con aminoácidos cargados neutra o negativamente. Lo último da como resultado proteína con carga positiva reducida para mejorar las características de la proteína revelada. La prevención de agregación es altamente deseable. La agregación de proteínas no solamente da como resultado una pérdida de actividad sino que también puede ser problemática cuando se preparan formulaciones farmacéuticas debido a que pueden ser inmunógenas. (Pinckard y cols., Clin. Exp. Immunol. 2: 331-340 (1967); Robbins y cols., Diabetes 36: 838-845 (1987); Cleland y cols., Crit. Rev. Therapeutic Drug Carrier Systems 10: 307-377 (1993)).

50 Se pueden identificar aminoácidos en los polipéptidos de la presente invención que son esenciales para la función por procedimientos conocidos en la técnica, tales como mutagénesis dirigida a sitio o mutagénesis de rastreo de alanina (Cunningham y Wells, Science 244: 1081-1085 (1989)). Este último procedimiento introduce mutaciones de alanina individuales en cada residuo en la molécula. Las moléculas mutantes resultantes se probaron después para actividad biológica tal como unión a receptor, o actividad proliferativa *in vitro*. Los sitios que son críticos para unión ligando-receptor se pueden determinar también por análisis estructural tal como cristalización, resonancia magnética nuclear o marcado por fotoafinidad (Smith y cols., J. Mol. Biol. 224: 899-904 (1992) y de Vos y cols. Science 255: 306-312 (1992)).

55 Como se indica, los cambios son preferentemente de una naturaleza secundaria, tales como sustituciones aminoacídicas conservadoras que no afectan significativamente al plegamiento o a la actividad de la proteína. Por supuesto, el número de sustituciones aminoacídicas que haría un trabajador experto depende de muchos factores, incluyendo aquellos anteriormente descritos. Hablando en general, el número de sustituciones para cualquier polipéptido dado no será más de 50, 40, 30, 25, 20, 15, 10, 5 o 3.

Se pueden extraer proteínas marcadoras metastásicas de longitud total, usando procedimientos bioquímicos estándar, a partir de células humanas que producen proteínas marcadoras metastásicas, tales como células de cáncer de mama metastásicas. Una proteína o polipéptido marcador metastásico aislado y purificado se separa de otros compuestos que normalmente están asociados con una proteína o polipéptido marcador metastático en una célula, tal como ciertas proteínas, carbohidratos, lípidos u orgánulos subcelulares. Una preparación de proteínas o polipéptidos marcadores metastásicos aislados y purificados es al menos pura al 80 %; preferentemente, las preparaciones son puras al 90 %, al 95 %, o al 99 %.

Un gen humano que codifica SEC ID N.º: 2 se puede identificar y aislar usando procedimientos conocidos en la técnica. De acuerdo con un procedimiento, la SEC ID N.º: 1 se prepara en un formato legible por computador. La secuencia se comparó con secuencias de polinucleótidos de un genoma humano y se identifican una o más secuencias del genoma humano que tienen al menos el 95 % de identidad de secuencia con SEC ID N.º: 1, por ejemplo usando el algoritmo de Smith-Waterman usando una búsqueda de hueco afín con una penalización abierta del hueco de 12 y una penalización de la extensión del hueco de 1 como parámetros. Las sondas basadas en las regiones de homología entre la SEC ID N.º: 1 y las secuencias del genoma humano se preparan y usan para aislar polinucleótidos de ADN genómico humano, usando procedimientos conocidos en la técnica. A partir de la fecha de presentación un polinucleótido humano que corresponde al polinucleótido completo de SEC ID N.º: 1 no se identificó en las bases de datos públicas. Así, la invención incluye ADN genómico humano que comprende la región codificante de SEC ID N.º: 1 y cualesquiera regiones no traducidas que comparten homología con la SEC ID N.º: 1 pero que son contiguas con las regiones homólogas. Tal ADN genómico incluye pero no se limita a intrones, promotores y otras regiones reguladoras funcionalmente asociadas con un gen humano que tiene una región SEC ID N.º: 2 codificante.

Proteínas y polipéptidos marcadores metastásicos pueden también producirse por procedimientos de ADN recombinante o por procedimientos químicos de síntesis. Para la producción de proteínas o polipéptidos marcadores metastásicos, las secuencias codificantes seleccionadas a partir de las secuencias nucleotídicas mostradas en SEC ID N.º: 1, o variantes de aquellas secuencias que codifican proteínas marcadoras metastásicas, se pueden expresar en sistemas de expresión procariotas o eucariotas conocidos (véase más adelante). Se pueden usar sistemas de expresión bacterianos, de levaduras, de insectos o de mamíferos, como se conoce en la técnica.

Alternativamente, procedimientos químicos sintéticos, tales como síntesis peptídica en fase sólida, se pueden usar para sintetizar una proteína o un polipéptido marcador metastásico. Los medios generales para la producción de péptidos, análogos o derivados se resumen en CHEMISTRY AND BIOCHEMISTRY OF AMINO ACIDS, PEPTIDES, AND PROTEINS - A SURVEY OF RECENT DEVELOPMENTS, Weinstein, B. ed., Marcell Dekker, Inc., publ., Nueva York (1983). Además, la sustitución de D-aminoácidos por el estereoisómero L normal se puede llevar a cabo para incrementar la semivida de la molécula. Las variantes de marcadores metastásicos se pueden producir de forma similar.

También se pueden construir las proteínas de fusión que aparecen de forma no natural que comprenden al menos 6, 8, 10, 12, 15, 18, 20, 25, 30, 35, 40, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 120, 140, 160, 180, 200, 220, 240, 250, 260, 265, 270 o 271 o más aminoácidos marcadores metastásicos contiguos. Las proteínas de fusión marcadoras metastásicas humanas son útiles para generar anticuerpos contra las secuencias aminoacídicas marcadoras metastásicas y para usar en diversos sistemas de ensayo. Por ejemplo, las proteínas de fusión de marcadores metastásicos se pueden usar para identificar proteínas que interactúan con proteínas marcadoras metastásicas y que influyen en sus funciones. Se pueden usar también para este propósito los procedimientos físicos, tales como cromatografía de afinidad a proteínas, o los ensayos basados en bibliotecas para interacciones proteína-proteína, tales como los sistemas de expresión en doble híbrido de levaduras o los sistemas de expresión en fagos. Tales procedimientos son bien conocidos en la técnica y pueden usarse también como revisiones en busca de fármacos.

Una proteína de condensación de marcador metastásico comprende dos segmentos de proteínas condensados conjuntamente por medio de un enlace peptídico. El primer segmento proteico comprende al menos 6, 8, 10, 12, 15, 18, 20, 25, 30, 35, 40, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 120, 140, 160, 180, 200, 220, 240, 250, 260, 265, 270 o 271 o más aminoácidos contiguos de una proteína marcadora metastásica. Los aminoácidos pueden seleccionarse a partir de las secuencias de aminoácidos codificados por la secuencia de nucleótidos mostrada en la SEC ID N.º: 1 o a partir de las variantes de la secuencia, tales como aquellas descritas anteriormente. El primer segmento de proteínas puede comprender también una proteína marcadora metastásica de longitud completa.

En una realización preferida, el primer segmento proteico comprende el polipéptido mostrado en la SEC ID N.º: 2. En una variación de esta realización, el primer segmento proteico consta de los aminoácidos 31-287 de la SEC ID N.º: 2. Esta proteína de fusión carece del péptido señal de la SEC ID N.º: 2 y sería adecuada para la retención de la proteína de condensación expresada dentro de la célula.

El segundo segmento proteico puede ser una proteína de longitud completa o un fragmento de la proteína o polipéptido no encontrado adyacente al primer segmento proteico en la proteína nativa codificada por SEC ID N.º: 1. La proteína de condensación puede marcarse con un marcador detectable, como se conoce en la técnica, tal como un marcador radiactivo, fluorescente, quimioluminiscente, o biotinilado. El segundo segmento proteico puede ser una

enzima que generará un producto detectable, tal como β -galactosidasa. El primer segmento proteico puede ser N-terminal o C-terminal, según sea conveniente.

5 Las técnicas para fabricar proteínas de fusión, uniendo bien recombinantemente o bien covalentemente dos segmentos proteicos, también se conocen bien. Los procedimientos de ADN recombinante se pueden usar para preparar proteínas de fusión de marcadores metastásicos, por ejemplo, elaborando una construcción de ADN que comprende secuencias codificantes de SEC ID N.º: 1 en fase de lectura apropiada con nucleótidos que codifican el segundo segmento proteico y que expresan la construcción de ADN en una célula huésped, según se describe a continuación. La fase de lectura abierta de la SEC ID N.º: 1 se muestra en la figura 4.

10 Las proteínas, polipéptidos, variantes o proteínas de fusión, marcadores metastásicos aislados y purificados pueden usarse como inmunógenos, para obtener preparaciones de anticuerpos que se unen específicamente a una proteína marcadora metastásica. Los anticuerpos pueden usarse, *entre otras cosas*, para detectar proteínas marcadoras metastásicas de tipo silvestre en tejido humano y fracciones de las mismas. Los anticuerpos también pueden usarse para detectar la presencia de mutaciones en los genes marcadores metastásicos que dan como resultado subexpresión o sobreexpresión de una proteína marcadora metastásica o expresión de una proteína marcadora metastásica con tamaño o movilidad electroforética alterados.

15 Las preparaciones de anticuerpos policlonales o monoclonales pueden elaborarse usando procedimientos estándar. Se pueden preparar también anticuerpos de cadena individual. Un inmunógeno preferido es un polipéptido que comprende la SEC ID N.º: 2. Los anticuerpos de cadena simple que se unen específicamente a las proteínas, polipéptidos, variantes o proteínas de fusión marcadoras, pueden aislarse, por ejemplo, a partir de bibliotecas de expresión de inmunoglobulinas de cadena individual, como se muestra en la técnica. La biblioteca se "examina" frente a las secuencias aminoácidas de proteína marcadora metastásica de SEC ID N.º: 2 y se puede aislar un número de anticuerpos de cadena única que se unen con afinidad alta a epítopos diferentes de proteínas marcadoras metastásicas. Hayashi y cols., 1995, Gene 160: 129-30. Los anticuerpos de cadena individual se pueden construir también usando un procedimiento de amplificación de ADN, tal como la reacción en cadena de la polimerasa (PCR), usando ADNc de hibridoma como una plantilla. Thirion y cols., 1996, Eur. J. Cancer Prev. 5: 507-11.

20 Los anticuerpos específicos de marcador metastásico se unen específicamente a epítopos presentes en una proteína marcadora metastásica de longitud completa que tiene una secuencia aminoácida codificada por una secuencia de nucleótidos mostrada en la SEC ID N.º: 1, a polipéptidos marcadores metastásicos, o a variantes de marcadores metastásicos, bien solos o bien como parte de una proteína de fusión. Preferentemente, los epítopos de marcadores metastásicos no están presentes en otras proteínas humanas. Típicamente, se requieren al menos 6, 8, 10, o 12 aminoácidos contiguos para formar un epítipo. Sin embargo, los epítopos que implican aminoácidos no contiguos pueden requerir más, *por ejemplo*, al menos 15, 25, o 50 aminoácidos.

25 Los anticuerpos que se unen específicamente a las proteínas, polipéptidos, proteínas de fusión o variantes marcadoras metastásicas proporcionan una señal de detección al menos 5, 10, o 20 veces más alta que una señal de detección proporcionada con otras proteínas cuando se usan en pruebas de bandas de western o en otros ensayos inmunoquímicos. Preferentemente, los anticuerpos que se unen específicamente a epítopos de marcadores metastásicos no detectan otras proteínas en ensayos inmunoquímicos y pueden inmunoprecipitar una proteína, polipéptido, proteína de fusión, o variante marcadora metastásica a partir de solución.

30 Los polinucleótidos subgenómicos contienen menos de un cromosoma completo. Preferentemente, los polinucleótidos están libres de intrones. En una realización preferida, las moléculas polinucleótídicas comprenden una secuencia contigua de 10, 15, 20, 25, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300 o 2350 nucleótidos de SEC ID N.º: 1 o los complementos de las mismas. El complemento de una secuencia de nucleótidos mostrada en SEC ID N.º: 1 es una secuencia de nucleótidos contiguos que forman pares de bases de Watson y Crick con una secuencia de nucleótidos contiguos mostrada en la SEC ID N.º: 1.

35 Las secuencias de nucleótidos degeneradas que codifican secuencias de aminoácidos de proteína marcadora metastásica o variantes, así como secuencias nucleotídicas homólogas que comprenden un polinucleótido al menos idéntico al 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, o al 99 % a la región codificante de la secuencia de nucleótidos mostrada en SEC ID N.º: 1, son también polinucleótidos subgenómicos marcadores metastásicos. Típicamente, secuencias polinucleótídicas subgenómicas marcadoras metastásicas homólogas se pueden confirmar por hibridación según condiciones restrictivas, como se muestra en la técnica. La identidad de secuencia en porcentaje entre las secuencias de tipo silvestre y de variante homóloga se determina alineando el polinucleótido de tipo silvestre con la variante para obtener el mayor número de coincidencias de nucleótidos, como se conoce en la técnica, contando el número de coincidencias nucleotídicas entre el tipo silvestre y la variante y dividiendo el número total de las coincidencias por el número total de nucleótidos de la secuencia de tipo silvestre. Un algoritmo preferido para calcular la identidad porcentual es el algoritmo de búsqueda de homología Smith-Waterman según se implementa en el programa MPSRCH (Oxford Molecular) usando una búsqueda de hueco afín con los siguientes parámetros de búsqueda: penalización abierta del hueco de 12 y penalización de la extensión del hueco de 1.

Un polinucleótido subgenómico marcador metastásico que comprende secuencias codificantes de proteína marcadora metastásica puede usarse en una construcción de expresión. Preferentemente, el polinucleótido subgenómico marcador metastásico está insertado dentro de un plásmido de expresión (por ejemplo, el sistema Ecdyson, pIND, In vitro Gene). Los polinucleótidos subgenómicos marcadores metastásicos se pueden propagar en vectores y líneas celulares que usan técnicas bien conocidas en la técnica. Los polinucleótidos subgenómicos marcadores metastásicos pueden estar en moléculas lineales o circulares. Ellos pueden estar en moléculas replicantes autónomamente o en moléculas sin secuencias de replicación. Se pueden regular por sí mismas o por otras secuencias reguladoras, como se conoce en la técnica.

Una célula huésped que comprende una construcción de expresión de marcador metastásico se puede usar después para expresar toda o una parte de una proteína marcadora metastásica. Las células huésped que comprenden las construcciones de expresión de marcador metastásico pueden ser procariotas o eucariotas. Una diversidad de células huésped están disponibles para su uso en sistemas de expresión bacterianos, de levaduras, de insectos o humanos y se pueden usar para expresar o para propagar las construcciones de expresión de marcador metastásico (véase más adelante). Pueden introducirse construcciones de expresión en células huésped usando cualquier técnica conocida en la técnica. Estas técnicas incluyen transferencia de ADN mediada por transferrina-policación, transfección con ácidos nucleicos desnudos o encapsulados, fusión celular mediada por liposomas, transporte intracelular de perlas de látex revestidas de ADN, fusión de protoplastos, infección vírica, electroporación y transfección mediada por fosfato de calcio.

Una construcción de expresión de marcador metastásico comprende un promotor que es funcional en una célula huésped elegida. El trabajador experto puede seleccionar fácilmente un promotor apropiado a partir del gran número de promotores específicos de tipo celular conocidos y usados en la técnica. La construcción de expresión puede contener también un terminador de transcripción que es funcional en la célula huésped. La construcción de expresión comprende un segmento de polinucleótidos que codifica todo o una parte de la proteína, variante, proteína de fusión, anticuerpo, o ribozima marcadora metastásica. El segmento polinucleotídico está localizado más adelante en la cadena a partir del promotor. La transcripción del segmento polinucleotídico se inicia en el promotor. La construcción de expresión puede ser lineal o circular y puede contener secuencias, si se desea, para replicación autónoma.

Los sistemas bacterianos para expresar las construcciones de expresión de marcador metastásico incluyen aquellos descritos en Chang y cols., *Nature* (1978) 275: 615, Goeddel y cols., *Nature* (1979) 281: 544, Goeddel y cols., *Nucleic Acids Res.* (1980) 8: 4057, documento EP 36.776, Patente de los EE.UU. N.º: 4.551.433, deBoer y cols., *Proc. Nat'l Acad. Sci. EE.UU.* (1983) 80: 21-25 y Siebenlist y cols., *Cell* (1980) 20: 269.

Los sistemas de expresión en levaduras incluyen aquellos descritos en Hinnen y cols., *Proc. Nat'l Acad. Sci. U.S.A.* (1978) 75: 1929; Ito y cols., *J. Bacteriol.* (1983) 153: 163; Kurtz y cols., *Mol. Cell. Biol.* (1986) 6: 142; Kunze y cols., *J. Basic Microbiol.* (1985) 25: 141; Gleeson y cols., *J. Gen. Microbiol.* (1986) 132: 3459, Roggenkamp y cols., *Mol. Gen. Genet.* (1986) 202: 302; Das y cols., *J. Bacteriol.* (1984) 158: 1165; De Louvencourt y cols., *J. Bacteriol.* (1983) 154: 737, Van den Berg y cols., *Bio/Technology* (1990) 8: 135; Kunze y cols., *J. Basic Microbiol.* (1985) 25: 141; Cregg y cols., *Mol. Cell. Biol.* (1985) 5: 3376, la Patente de los EE.UU. N.º: 4.837.148, la Patente de los EE.UU. N.º: 4.929.555; Beach y Nurse, *Nature* (1981) 300: 706; Davidow y cols., *Curr. Genet.* (1985) 10: 380, Gaillardin y cols., *Curr. Genet.* (1985) 10: 49, Ballance y cols., *Biochem. Biophys. Res. Commun.* (1983) 112: 284-289; Tilbum y cols., *Gene* (1983) 26: 205-221, Yelton y cols., *Proc. Nat'l Acad. Sci. U.S.A.* (1984) 81: 1470-1474, Kelly e Hynes, *EMBO J.* (1985) 4: 475479; documentos EP 244.234 y WO 91/00357.

Las construcciones de expresión de marcador metastásico en insectos se pueden llevar a cabo como se describe en la Patente de los EE.UU. N.º: 4.745.051, Friesen y cols. (1986) "The Regulation of Baculovirus Gene Expression" en: *THE MOLECULAR BIOLOGY OF BACULOVIRUSES* (W. Doerfler, ed.), documentos EP 127.839, EP 155.476, y Vlak y cols., *J. Gen. Virol.* (1988) 69: 765-776, Miller y cols., *Ann. Rev. Microbiol.* (1988) 42: 177, Carbonell y cols., *Gene* (1988) 73: 409, Maeda y cols., *Nature* (1985) 315: 592-594, Lebacqz-Verheyden y cols., *Mol. Cell. Biol.* (1988) 8: 3129; Smith y cols., *Proc. Nat'l Acad. Sci. U.S.A.*, 1985 (82): 8404, Miyajima y cols., *Gene* (1987) 58: 273; y Martin y cols., *DNA* (1988) 7: 99. Se describen numerosas cepas y variantes baculovirales y células huésped de insecto permisivas correspondientes a partir de huéspedes descritos en Luckow y cols., *Bio/Technology* (1988) 6: 47-55, Miller y cols., en *GENETIC ENGINEERING* (Setlow, J.K. y cols. eds.), Vol. 8 (Plenum Publishing, 1986), páginas 277-279 y Maeda y cols., *Nature*, (1985) 315: 592-594.

La expresión en mamíferos de construcciones de expresión de marcadores metastásicos puede lograrse como se describe en Dijkema y cols., *EMBO J.* (1985) 4: 761, Gorman y cols., *Proc. Nat'l Acad. Sci. U.S.A.* (1982b) 79: 6777, Boshart y cols., *Cell* (1985) 41: 521 y Patente de los EE.UU. N.º: 4.399.216. Otras características de expresión en mamíferos de construcciones de expresión de marcadores metastásicos se pueden facilitar como se describe en Ham y Wallace, *Meth. Enz.* (1979) 58: 44, Bames y Sato, *Anal. Biochem.* (1980) 102: 255, Patente de los EE.UU. N.º: 4.767.704, Patente de los EE.UU. N.º: 4.657.866, Patente de los EE.UU. N.º: 4.927.762, Patente de los EE.UU. N.º: 4.560.655, documentos WO 90/103430, WO 87/00195 y U.S. RE 30.985.

Los polinucleótidos subgenómicos de la invención pueden usarse también en vehículos de administración génica, para el propósito de administrar un ARNm u oligonucleótido marcador metastásico (bien con la secuencia del ARNm

marcador metastásico nativo o bien con su complemento), proteína marcadora metastásica de longitud total, proteína de fusión marcadora metastásica, polipéptido marcador metastásico o ribozima específica de marcador metastásico en un anticuerpo de cadena simple, dentro de una célula, preferentemente una célula eucariota. Un vehículo de administración génica puede ser, por ejemplo, ADN plasmídico desnudo, un vector de expresión vírico que comprende un polinucleótido subgenómico marcador metastásico, o un polinucleótido subgenómico marcador metastásico en conjunción con un liposoma o un agente condensante.

Se describe también en el presente documento un procedimiento de detectar expresión génica de marcador metastásico. La detección de expresión génica de marcador metastásico es útil, por ejemplo, para identificar metástasis o para determinar el potencial metastásico en una muestra de tejido, preferentemente un tumor. Los regímenes de tratamiento apropiados se pueden diseñar después para pacientes que estén en riesgo de desarrollar cánceres metastásicos en otros órganos del cuerpo.

La muestra corporal puede ser, por ejemplo, un tejido sólido o una muestra de fluido. El polipéptido nativo codificado por SEC ID N.º: 1 es una proteína segregada teórica y es probable que se detecte en los fluidos corporales incluyendo sangre y fluido linfático, particularmente aquellos que se extraen desde los sitios tumorales en el cuerpo. Se puede detectar proteína o productos de expresión de ácidos nucleicos en la muestra corporal. En una realización, la muestra corporal se somete a ensayos para la presencia de una proteína marcadora metastásica. Una proteína marcadora metastásica comprende una secuencia codificada por una secuencia de nucleótidos mostrada en la SEC ID N.º: 1 o su complemento y puede detectarse usando los anticuerpos específicos de proteínas marcadoras de la presente invención. Los anticuerpos pueden marcarse, por ejemplo, con una marca radioactiva, fluorescente, biotinilada, o enzimática y detectarse directamente, o pueden detectarse usando procedimientos inmunoquímicos indirectos, usando un anticuerpo secundario marcado. La presencia de las proteínas marcadoras metastásicas se puede someter a ensayo, por ejemplo, en las secciones de tejido por inmunocitoquímica, o en lisados usando transferencia de western, como se muestra en la técnica. La presencia de la proteína marcadora es indicativa de que la muestra de tejido es metastásica.

En otra realización, la muestra corporal se somete a ensayo para detectar la presencia de ARNm de proteínas marcadoras. Una muestra se pone en contacto con una sonda de hibridación capaz de hibridar con el ARNm que corresponde al polipéptido seleccionado. Aún adicionalmente, la muestra puede someterse a una técnica de transferencia de northern para detectar ARNm, indicando la expresión del polipéptido. Para aquellas técnicas en las que se detecta el ARNm, la muestra puede someterse a un procedimiento de amplificación de ácidos nucleicos en el que la molécula de ARNm o una parte seleccionada de la misma se amplifica usando cebadores de nucleótidos apropiados. Otras técnicas de detección de ARN pueden usarse también, incluyendo, pero no limitadas a, la hibridación *in situ*.

Las sondas específicas de la proteína marcadora pueden generarse usando la secuencia de ADNc descrita en la SEC ID N.º: 1. Las sondas son preferentemente de al menos 15 a 50 nucleótidos de longitud, aunque pueden ser de al menos 8, 10, 11, 12, 20, 25, 30, 35, 40, 45, 60, 75, o 100 o más nucleótidos de longitud. Una región preferible para seleccionar sondas está dentro de las posiciones nucleotídicas 446-1173 de la SEC ID N.º: 1. Las sondas se pueden sintetizar químicamente o se pueden generar a partir de polinucleótidos más largos usando enzimas de restricción. Las sondas pueden marcarse, por ejemplo, con una marca radioactiva, biotinilada, o fluorescente.

Opcionalmente, se puede cuantificar el nivel de un producto de expresión de marcador metastásico particular en una muestra corporal. La cuantificación puede llevarse a cabo, por ejemplo, comparando el nivel del producto de expresión detectado en la muestra corporal con las cantidades de producto presentes en una curva estándar. Se puede hacer una comparación visualmente o se puede usar una técnica tal como densitometría, con o sin asistencia computerizada. Para usar como controles, se pueden aislar muestras corporales a partir de otros seres humanos, otros órganos no cancerosos del paciente que se está poniendo a prueba, o cáncer de mama no metastásico del paciente que se está poniendo a prueba. Como se indica por los resultados en el presente documento, la expresión de la SEC ID N.º: 1 en células cancerosas con baja capacidad metastásica o no metastásicas está entre el 3 % y el 44 % de los niveles de expresión en células de cáncer de mama altamente metastásicas. Si la expresión en una muestra de prueba es al menos 2 veces mayor que en una muestra control adecuada, esto es indicativo de células metastásicas.

Los polinucleótidos que codifican reactivos específicos de marcador metastásico de la invención, tales como anticuerpos y sondas nucleotídicas, se puede suministrar en un kit para detectar productos de expresión génica marcadores en una muestra biológica. El kit puede también contener tampones o componentes de marcado, así como instrucciones para usar los reactivos para detectar la expresión de marcadores en la muestra biológica.

La expresión de un gen marcador metastásico puede alterarse usando una secuencia de oligonucleótidos antisentido. La secuencia antisentido es complementaria a al menos una parte de la secuencia de codificación (nucleótidos 365-1173) de un gen marcador metastásico que tiene una secuencia de nucleótidos mostrada en SEC ID N.º: 1. Preferentemente, la secuencia de oligonucleótidos antisentido es de al menos seis nucleótidos de longitud, pero puede ser al menos aproximadamente de 8, 12, 15, 20, 25, 30, 35, 40, 45, o 50 nucleótidos de largo. Se pueden usar también secuencias más largas. Las moléculas de oligonucleótidos antisentido se pueden proporcionar en una construcción de ADN y se pueden introducir en células cuya división está disminuyéndose. Tales células

incluyen células de cáncer de mama altamente metastásicas.

Los oligonucleótidos antisentido pueden comprender desoxirribonucleótidos, ribonucleótidos, o una combinación de ambos. Se pueden sintetizar oligonucleótidos manualmente o por un sintetizador automatizado, uniendo covalentemente el extremo 5' de un nucleótido con el extremo 3' de otro nucleótido con uniones internucleotídicas no fosfodiéster tales como alquilfosfonatos, fosforotioatos, fosforoditioatos, alquilfosfonotioatos, alquilfosfonatos, fosforamidatos, ésteres de fosfato, carbamatos, acetamidato, ésteres carboximetílicos, carbonatos y triésteres de fosfato. Véase Brown, 1994, Meth. Mol. Biol. 20: 1-8; Sonveaux, 1994, Meth. Mol. Biol. 26: 1-72; Uhlmann y cols., 1990, Chem. rev. 90: 543-583.

Los anticuerpos de la invención que se unen específicamente a una proteína marcadora metastásica también pueden usarse para alterar la expresión génica marcadora metastásica. Por anticuerpos se quiere decir anticuerpos y partes o derivados de los mismos, tales como anticuerpos de cadena simple, que mantienen unión específica para la proteína. Los anticuerpos específicos se unen a proteínas marcadoras metastásicas y evitan que las proteínas funcionen en la célula. Los polinucleótidos que codifican los anticuerpos específicos de la invención se pueden introducir dentro de las células, según se describe anteriormente.

Las proteínas marcadoras de la presente invención se pueden usar para rastrear en busca de fármacos que tienen un efecto anti-metastásico terapéutico. El efecto de un compuesto de prueba de síntesis de proteína marcadora metastásica también se pueden usar para identificar compuestos de prueba que modulan metástasis. Los compuestos de prueba que pueden examinarse incluyen cualesquiera sustancias, sean productos naturales o sintéticos, que pueden administrarse al sujeto. Las bibliotecas o mezclas de compuestos se pueden probar. Los compuestos o sustancias pueden ser aquellos para los que un efecto farmacéutico se conoce o se desconoce previamente.

La síntesis de proteínas marcadoras metastásicas se puede medir por cualesquiera medios para medir la síntesis de proteínas conocida en la técnica, tales como la incorporación de aminoácidos marcados dentro de la proteína y la detección de proteínas marcadoras metastásicas en un gel de poliacrilamida. La cantidad de proteínas marcadoras metastásicas se puede detectar, por ejemplo, usando anticuerpos específicos de la proteína marcadora metastásica de la invención en bandas de western. La cantidad de las proteínas marcadoras metastásicas sintetizadas en presencia o ausencia de un compuesto de ensayo se puede determinar por cualesquiera medios conocidos en la técnica, tales como la comparación de la cantidad de proteína marcadora metastásica sintetizada con la cantidad de las proteínas marcadoras metastásicas presentes en una curva estándar.

El efecto de un compuesto de prueba de síntesis de proteína marcadora metastásica también se puede medir por análisis de bandas de northern, midiendo la cantidad de expresión de ARNm de proteína marcadora metastásica en respuesta al compuesto de ensayo usando sondas nucleotídicas específicas de proteínas marcadoras metastásicas de la invención, como se conoce en la técnica.

Típicamente, una muestra biológica se pone en contacto con un intervalo de concentraciones del compuesto de prueba, tales como 1,0 nM, 5,0 nM, 10 nM, 50 nM, 100 nM, 500 nM, 1 mM, 10 mM, 50 mM, y 100 mM. Preferentemente, el compuesto de prueba incrementa o disminuye la expresión de una proteína marcadora metastásica en un 60 %, 75 %, u 80 %. Más preferentemente, se logra un incremento o disminución del 85 %, 90 %, 95 %, o 98 %.

Las composiciones para incrementar o disminuir la expresión de proteína marcadora metastásica se describen en el presente documento. Estas composiciones comprenden los polinucleótidos que codifican todo o al menos una parte de un producto de expresión génica de proteína marcadora metastásica. Preferentemente, la composición terapéutica contiene una construcción de expresión que comprende un promotor y un segmento polinucleotídico que codifica al menos una parte de la proteína marcadora metastásica que es eficaz para disminuir el potencial metastásico. Las partes de genes marcadores metastásicos o de proteínas marcadoras metastásicas que son efectivas para disminuir el potencial metastásico de una célula se pueden determinar, por ejemplo, introduciendo partes de genes o polipéptidos marcadores metastásicos en líneas celulares metastásicas, tales como MDA-MB-231, MDA-MB-435, Km12C, o Km12L4 y sometiendo a ensayo la velocidad de división de las células o la capacidad de las células para formar metástasis cuando se implantan *in vivo*, como se muestra en la técnica. Las líneas celulares no metastásicas, tales como MCF-7, pueden usarse para analizar la capacidad de una parte de una proteína marcadora metastásica para incrementar la expresión de un gen marcador metastásico.

Típicamente, una composición marcadora de metástasis terapéutica se prepara como un inyectable, bien como una solución líquida o bien como una suspensión líquida, sin embargo, se pueden preparar también formas sólidas adecuadas para solución en, o suspensión en, vehículos líquidos antes de la inyección. Una composición marcadora metastásica se puede formular también dentro de un comprimido revestido entérico o de una cápsula de gel de acuerdo con procedimientos conocidos en la técnica, tales como aquellos descritos en la Patente de los EE.UU. N.º: 4.853.230, EP 225.189, AU 9.224.296 y AU 9.230.801.

La administración de un agente terapéutico de marcadores metastásicos puede incluir administración local o sistémica, incluyendo inyección, administración oral, pistola de partículas, o administración cateterizada y

administración tópica. Se pueden usar diversos procedimientos para administrar una composición marcadora metastásica terapéutica directamente a un sitio específico en el cuerpo.

5 Para el tratamiento de tumores, incluyendo lesiones metastásicas, por ejemplo, una composición marcadora metastásica terapéutica se puede inyectar varias veces en varias localizaciones diferentes dentro del cuerpo del tumor. Por otra parte, las arterias que suministran a un tumor pueden identificarse y se puede inyectar una composición terapéutica dentro de una arteria tal, con el fin de administrar la composición directamente dentro del tumor.

10 Un tumor que tiene un centro necrótico puede aspirarse y la composición puede inyectarse directamente dentro del centro ahora vacío del tumor. Una composición marcadora metastásica terapéutica puede administrarse directamente a la superficie de un tumor, por ejemplo, por aplicación tópica de la composición. La formación de imágenes de rayos x se puede usar para ayudar en ciertos de los procedimientos de administración anteriores. Los agentes terapéuticos de combinación, incluyendo proteínas marcadoras metastásicas o polipéptido marcador metastásico o un polinucleótido subgenómico marcador metastásico y otros agentes terapéuticos, se pueden administrar simultáneamente o secuencialmente.

15 Alternativamente, una composición terapéutica marcadora metastásica puede introducirse en células humanas ex vivo y las células pueden reemplazarse en el ser humano. Las células se pueden retirar a partir de una diversidad de localizaciones incluyendo, por ejemplo, a partir de un tumor seleccionado o a partir de un órgano afectado. Además, una composición terapéutica puede insertarse en células no afectadas, por ejemplo, fibroblastos dérmicos o leucocitos de sangre periférica. Si se desea, las fracciones particulares de células tales como un subgrupo de
20 células T o de células madre se pueden retirar específicamente de la sangre (véase, por ejemplo, documento PCT WO 91/16116). Las células retiradas se pueden poner en contacto con una composición terapéutica marcadora metastásica utilizando cualquiera de las técnicas descritas anteriormente, seguido por el retorno de las células al ser humano, preferentemente en o en las proximidades de un tumor u otro lugar a tratarse. Los procedimientos descritos anteriormente pueden comprender adicionalmente las etapas de deplecionar fibroblastos u otras células tumorales
25 no contaminantes subsiguientes a retirar células tumorales de un ser humano, y/o la etapa de inactivar las células, por ejemplo, por irradiación.

Tanto la dosis de una composición marcadora metastásica como los medios de administración se pueden determinar en base a las cualidades específicas de la composición terapéutica, la afección, edad y peso del
30 paciente, la progresión de la enfermedad y otros factores relevantes. Preferentemente, una composición terapéutica disminuye la expresión de los genes marcadores metastásicos en un 50 %, 60 %, 70 %, u 80 %. Lo más preferentemente, la expresión de los genes marcadores metastásicos está disminuida en un 90 %, 95 %, 99 %, o 100 %. La efectividad del mecanismo elegido para alterar la expresión de genes marcadores metastásicos puede valorarse usando procedimientos bien conocidos en la técnica, tales como hibridación de sondas de nucleótidos a
35 ARNm de los genes marcadores metastásicos, RT-PCR cuantitativa, o detección de las proteínas marcadoras metastásicas usando anticuerpos específicos de la invención.

Si la composición contiene las proteínas marcadoras metastásicas, el polipéptido marcador metastásico, o el anticuerpo marcador metastásico, las dosificaciones efectivas de la composición están en el intervalo de
40 aproximadamente 5 µg a aproximadamente 50 µg/kg de peso corporal del paciente, aproximadamente 50 µg a aproximadamente 5 mg/kg, aproximadamente 100 µg a aproximadamente 500 µg/kg de peso corporal del paciente y aproximadamente 200 a aproximadamente 250 µg/kg.

Las composiciones terapéuticas que contienen polinucleótidos subgenómicos marcadores metastásicos se pueden administrar en un intervalo de aproximadamente 100 ng a aproximadamente 200 mg de ADN para administración local. Los intervalos de concentración de aproximadamente 500 ng a aproximadamente 50 mg, aproximadamente 1
45 µg a aproximadamente 2 mg, aproximadamente 5 µg a aproximadamente 500 µg y aproximadamente 20 µg a aproximadamente 100 µg de ADN también se pueden usar durante un protocolo de terapia génica. Factores tales como el procedimiento de acción y la eficacia de la transformación y la expresión son consideraciones que afectarán a la dosificación requerida para la eficacia máxima de los polinucleótidos subgenómicos marcadores metastásicos. Donde se desea expresión más grande en un área de tejido más grande, se pueden requerir grandes cantidades de polinucleótidos subgenómicos marcadores metastásicos o se pueden readministrar las mismas cantidades en un
50 protocolo sucesivo de administraciones, o en varias administraciones a diferentes partes tisulares adyacentes o cercanas a, por ejemplo, un sitio tumoral, para lograr un resultado terapéutico positivo. En todos los casos, la experimentación de rutina en ensayos clínicos determinará los intervalos específicos para efecto terapéutico óptimo.

La expresión de un gen marcador metastásico endógeno en una célula puede alterarse introduciendo en fase con el gen marcador metastásico endógeno una construcción de ADN que comprende una secuencia que dirige proteínas
55 marcadoras metastásicas, una secuencia reguladora, un exón y un sitio donante de ajustamiento desapareado por recombinación homóloga, tal que se forma una célula recombinante homológamente que comprende la construcción de ADN. La unidad de transcripción nueva se puede usar para activar o desactivar el gen marcador metastásico según se desea. Este procedimiento de afectar la expresión de genes endógenos se enseña en la Patente de los EE.UU. N.º: 5.641.670.

Se puede administrar también un polinucleótido subgenómico de marcador metastásico a sujetos para el propósito de rastrear compuestos de ensayo en busca de aquellos que son útiles para mejorar la transferencia de polinucleótidos subgenómicos marcadores metastásicos a la célula o para potenciar efectos biológicos subsiguientes de polinucleótidos subgenómicos marcadores metastásicos dentro de la célula. Tales efectos biológicos incluyen hibridación a ARNm marcador metastásico complementario e inhibición de su traducción, expresión de un polinucleótido subgenómico marcador metastásico para formar ARNm marcador metastásico y/o proteína marcadora metastásica y replicación e integración de un polinucleótido subgenómico marcador metastásico. El sujeto puede ser un cultivo celular o a un animal, preferentemente un mamífero, más preferentemente un ser humano.

La divulgación anterior describe la presente invención en general. Puede obtenerse una comprensión más completa por referencia a los ejemplos específicos siguientes que se proporcionan en el presente documento solamente con fines ilustrativos y no se pretende limitar el ámbito de la invención.

Ejemplos

Materiales y procedimientos

Cultivo celular.

Se cultivaron líneas celulares de cáncer de mama humano MDA-MB-435, MDA-MB-231, ALAB, MDA-MB-468, MDA-MB-361, ZR-75-1, MCF-7, MDA-MB-453 y SK-BR-3 (obtenidas de Chiron Master Culture Collection, Chiron Corporation) a 37 °C en CO₂ al 5 % en DMEM + F-12 de HAM (1:1) (Bio*Whittaker, Walkersville, MD) que contiene L-glutamina 2 mM, piruvato de sodio 1 mM, 100 U/ml de penicilina y 100 µg/ml de estreptomina (Bio*Whittaker, Walkersville, MD), solución vitamínica 1x, aminoácidos no esenciales 1x (Irvine Scientific, Santa Ana, CA), y suero fetal bovino inactivado por calor al 10 % (Life Technologies, Rockville, MD). Las células COS-7 se obtuvieron a partir de ATCC y se cultivaron a 37 °C en CO₂ al 5 % en DMEM con suero bovino fetal inactivado por calor al 10 % (Life Technologies).

Concentración de sobrenadante de Opti-MEM1.

Se concentraron medios de cultivo Opti-MEM1 (Life Technologies) a través de columnas de Centricon YM-10 y/o de Microcon YM-10 (Millipore Corporation, Bedford, MA). Se añadió después tampón de carga de muestras SDS-PAGE y las muestras se hirvieron.

Hibridación de prueba de bandas de northern.

Se prepararon ARN totales a partir de las líneas celulares de cáncer de mama cultivadas y tejidos tumorales de ratones SCID transplantados con líneas celulares de cáncer de mama con kit RNeasy Maxi (Qiagen, Valencia, CA). Aproximadamente 20 µg de ARN total por carril se cargaron sobre un gel de formaldehído/agarosa para electroforesis, después se transfirieron a una membrana de nylon Hybond-N+ (Amersham Life Science, Little Chalfont, Inglaterra). La banda se fijó por irradiación de UV. Se precalentó tampón de Rapid-Hyb (Amersham Life Science) con 5 mg/ml de ADN de esperma de cadena simple desnaturalizado a 65 °C y la prueba de bandas se prehibridó en el tampón con agitación a 65 °C durante 30 minutos. Se añadió un fragmento de ADNc de hsOAF o un fragmento de ADNc de β-actina como sonda marcada con [³²P]dCTP (3000 Ci/mmol, Amersham Pharmacia Biotech Inc., Piscataway, NJ) (kit Prime-It RmT, Stratagene, La Jolla, CA) y se purificó con ProbeQuant™ G-50 Micro Column (Amersham Pharmacia Biotech, Inc.) y se hibridó con la prueba de bandas con agitación a 65 °C durante toda una noche. Las bandas de western se lavaron en SSC 2x, SDS al 0,1 % (p/v) a temperatura ambiente durante 20 minutos, dos veces en SSC 1x, SDS al 0,1 % (p/v) a 65 °C durante 15 minutos, después se expusieron a Hyperfilms (Amersham Life Science).

Inmunotransferencia.

Las muestras proteicas se sometieron a electroforesis en geles de SDS-PAGE al 10-20 % después se transfirieron a membranas de PVDF (0,2 µm) por electrotransferencia en Tris 25 mM, glicina 192 mM, metanol al 20 % (v/v), pH 8,3. Las membranas se bloquearon en TBST (pH 7,5) conteniendo leche desnatada en polvo al 10 %, después se sometieron a prueba de bandas en PBS (pH 7,4) conteniendo BSA al 1 % con un suero anti-hsOAF de conejo (1:1000), seguido por sondeo con un anticuerpo secundario IgG anti-conejo de cabra conjugado con fosfatasa alcalina (1:2000) (Santa Cruz Biotechnology, Inc., Santa Cruz, CA). Las bandas de proteína se visualizaron después por reactivo NBT/BCIP (Boehringer Mannheim, Alemania).

Transfección transitoria.

La región codificante (356-1174) de ADNc de hsOAF se clonó dentro de un vector de expresión de pRetro-On modificado (Clontech, Palo Alto, CA). El vector pRetro-On que alberga hsOAF o el vector pRetro-On de control con GFP se transfirió en células COS-7 en una placa de cultivo de 100 mm usando kit de reactivo de transfección Effectene™ (Qiagen) como se instruyó en el protocolo proporcionado por el fabricante. Las células se recuperaron en DMEM con FBS al 10 % durante toda una noche después se cambiaron a Opti-MEM1. Después de dos días más, el sobrenadante se recogió y se concentraron para análisis de bandas de western.

Transfección de oligos antisentido.

Las células MDA-MB-435 se sembraron en placas de cultivo de 6 pocillos un día antes de la transfección proporcionando una densidad del 90 % en transfección. Se diluyeron 100 μ M de oligo control antisentido o reverso a 2 μ M en Opti-MEM1 para transfección. 0,5 mM de lipitoide I estéril se diluyeron a una proporción de 1,5 nmoles de lipitoide I: 1 μ g de oligo en el mismo volumen de Opti-MEM1. El oligo diluido y el lipitoide I diluido se mezclaron e inmediatamente se añadieron a células en medio de cultivo a una concentración final de oligo de 100, 200, o 300 nM. Después de 6 horas, la mezcla de transfección se reemplazó con medios de cultivo normales y las células se incubaron para recuperación durante toda una noche. La secuencia del oligo antisentido es AGCTGCGGAT-GCCACACTTGTAGG (SEC ID N.º: 4) y la secuencia del oligo de control reverso es GGATGTTTACACCGTAG-GCGTCGA (SEC ID N.º: 5).

Ensayo de invasión de matrigel.

Las células se tripsinizaron, se lavaron y se resuspendieron en medios para recuento. Se lavaron 4 x 10⁴ células y se resuspendieron en 100 μ l de medios en hielo. Se añadieron 200 μ l de matrigel (Collaborative Biomedical Products, Bedford, MA) a las células en hielo. El matrigel y las células se mezclaron cuidadosamente después se dispensaron dentro de un pocillo de placa de cultivo y se solidificaron a 37 °C durante 30 minutos. La mezcla de matrigel-células se coronó con 0,5 ml de medio y se incubó a 37 °C en CO₂ al 5 % durante 6 días. El medio se reabasteció cada 2 días.

Ensayo de proliferación.

Las células se tripsinizaron, se lavaron y se resuspendieron en medios para recuento. Las células se transfirieron después en placas de 96 pocillos (5000 células/pocillo) para incubación. Los números celulares se midieron con el kit de Ensayo de Proliferación Celular de Quantos™ (Stratagene, La Jolla, CA) cada día.

Ejemplo 1**Identificación de una secuencia de ADN humano**

El ADN que codifica un homólogo humano teórico del gen Out at First (oaf) de *Drosophila* (oaf) se muestra en la SEC ID N.º: 1. Un alineamiento de hsOAF y OAF de *Drosophila* se muestra en la Figura 7. El polinucleótido comprende 2366 pares de bases y se identifica una fase de lectura abierta. Una traducción del ORF, un polipéptido de 273 aminoácidos, se muestra en la SEC ID N.º: 2. La figura 4 proporciona el ADN y las secuencias de aminoácidos, indicando la posición del ORF. Los primeros 30 aminoácidos forman un péptido señal, indicando que la proteína puede secretarse. La secuencia de aminoácidos del péptido señal es: MRLPGVPLARPALLLLPLLLAPLLG#TGAPA (SEC ID N.º: 3). "#" indica la localización del sitio de corte de proteasa predicho.

Ejemplo 2**Expresión diferencial de SEC ID N.º: 1, en líneas celulares de cáncer de mama**

Se comparó la expresión de la SEC ID N.º: 1 en las siguientes líneas celulares de cáncer de mama humano:

- MDA-MB-361, derivada de adenocarcinoma de mama humano;
- MDA-MB-231, derivada de células de cáncer de mama humano metastásicas para hueso y/o pulmón;
- MDA-MB-468, derivada de células de cáncer de mama humano negativas para receptor de estrógenos;
- MDA-MB-435, derivada de células de carcinoma de mama humano negativas en receptor de estrógenos;
- MCF-7, derivada de células de carcinoma de mama humano no metastásicas; y
- ZR-75-1, derivada de células de carcinoma de mama humano positivas en receptor de estrógenos.

La expresión de la SEC ID N.º: 1 se midió en las líneas celulares de cáncer de mama altamente metastásicas MDA-MB 231 y MDA-MB-435 y se comparó con líneas celulares de cáncer de mama con baja capacidad metastásica o no metastásicas. La expresión en MDA-MB-361 fue del 11 % del nivel en MDA-MB-231; la expresión en MDA-MD-468 fue del 44 % del nivel en MDA-MB-231; la expresión en MCF-7 fue del 17 % del nivel en MDA-MB-231; y la expresión en ZR-75-1 fue del 12 % del nivel en MDA-MB-231.

La expresión en MDA-MB-361 fue del 6 % del nivel de MDA-MB-435; la expresión en MDA-MB-468 fue del 36 % del nivel de MDA-MB-435; y la expresión en MCF-7 fue del 3 % del nivel de MDA-MB-435. Así, como se muestra en la Tabla 2, hay una clara tendencia de expresión incrementada de SEC ID N.º: 1, en líneas celulares de cáncer de mama derivadas de tumores humanos con alto potencial metastásico.

Tabla 2

Línea celular con alta capacidad metastásica	Líneas celulares con baja capacidad metastásica: % de expresión relativa a línea celular con alta capacidad metastásica			
	MDA-MB-361	MDA-MB-468	MCF-7	ZR-75-1
MDA-MB-231	11 %	44 %	17 %	12 %
MDA-MB-435	6 %	36 %	3 %	N.D.

Un mismo patrón de expresión de este gen se mantuvo en muestras de tejido tumoral de ratones SCID transplantados con líneas celulares de carcinoma mamario tumorogénicas. (Figura 6).

5 Ejemplo 3

hsOAF codifica una proteína segregada y los niveles de secreción de proteína hsOAF son consistentes con niveles de expresión de ARNm de hsOAF de líneas celulares de carcinoma mamario.

Una secuencia peptídica señal predicha se localiza en el extremo N-terminal de la secuencia de aminoácidos deducida del gen hsOAF (fig. 3). Verificando la secreción de proteína de HsOAF, se llevó a cabo transfección transitoria de células COS-7 y de células MCF-7 con vector pRetro-On que alberga ADnc de hsOAF. Mientras tanto, el vector pRetro-On que alberga GFP se usó como control. Usando un antisuero de conejo de hsOAF, se detectó proteína hsOAF segregada en medios de cultivo Opti-MEM1 de ambas líneas celulares después de transfección con hsOAF por inmunotransferencia (fig. 8A). La proteína de hsOAF se glicosiló probablemente dado que se vieron múltiples bandas con pesos moleculares patentes más altos (el peso molecular predicho de proteína hsOAF segregada es 28 Kda). El mismo antisuero de hsOAF se usó detectando la secreción de proteína hsOAF por diversas líneas celulares de carcinoma mamario. Los niveles de secreción de proteína hsOAF fueron consistentes con los niveles de expresión de ARNm de hsOAF entre estas líneas celulares en general: las líneas celulares altamente metastásicas mostraron secreción de hsOAF mucho más fuerte que las líneas celulares con baja capacidad metastásica/no metastásicas (fig. 8B). MDA-MB-435 tuvo la secreción de proteína hsOAF más fuerte.

20 Ejemplo 4

Desactivación de expresión de hsOAF en células MDA-MB-435 por oligo antisentido causó cambio morfológico, invasividad de células reducida y velocidad de proliferación más lenta.

Determinando si es esencial nivel alto de expresión de gen hsOAF para el potencial metastásico de células de carcinoma mamario humano, se usó tecnología de oligos antisentido desactivando expresión de HsOAF, después se observaron los efectos consecuentes. Se eligió MDA-MB-435 dado que está línea celular altamente metastásica mostró la secreción de proteína hsOAF más fuerte entre todas las líneas celulares de cáncer de mama examinadas. Varios pares de oligos de hsOAF antisentido (AS) y de control reverso (RC) se eligieron probando su capacidad desactivando expresión génica de hsOAF a nivel del ARNm. Se llevó a cabo análisis de RT-PCR cuantitativo en tiempo real en Lightcycler (Roche Diagnostics, Indianápolis, IN) midiendo niveles de ARNm de hsOAF en células. Kang, S. y cols., Cancer Research 60 (18): 5296-5302 (2000). El mejor par se seleccionó después para la valoración de concentración de trabajo de oligo. Se prefiere concentración de oligo baja reduciendo la toxicidad de oligo potencial para las células. Los resultados indicaron que el tratamiento con 100 nM del oligo antisentido fue suficiente reduciendo significativamente la secreción de proteína hsOAF de células MDA-MB-435. (Figura 12). Este par de oligos (SEC ID N.º: 4 (AS) y 5 (RC)) a concentración de trabajo de 100 nM se usó para todos los experimentos siguientes.

Después del tratamiento de células MDA-MB-435 con oligo antisentido hsOAF, se observó alteración morfológica drástica de células junto con secreción de proteínas de hsOAF reducida (fig. 10A). Las células llegaron a ser más esféricas y perdieron sus protrusiones de dispersión. Mientras tanto, las células tratadas con oligo de control reverso permanecieron similares a las células MDA-MB-435 cultivadas de tejido normal. Además, el medio de cultivo de células MDA-MB-435 normales que contenían nivel alto de proteína hsOAF como el medio acondicionado añadido a células tratadas con oligo antisentido fue capaz de impedir este cambio morfológico, aunque no completamente. Esta alteración de la forma celular puede ser una indicación de capacidad de invasión reducida de las células.

El ensayo de invasión de matrigel se llevó a cabo después estimando la invasividad de las células. Se ha comunicado que una morfología invasiva, estrellada de células de cáncer de mama embebidas en matrigel correlaciona con su potencial metastásico (Thompson, E.W. y cols., J Cell Physiol. 150 (3): 534-44 (1992); Sugiura, T. y cols. J. Cell Biol, 146 (6): 1375-89 (1999); Albini, A. y cols., Cancer Res. 47 (12): 3239-45 (1987); y Kramer, R.H. y cols., Cancer Res. 46 (4 Pt. 2): 1980-89 (1986)) y esto se confirmó con diversas líneas celulares de cáncer de mama cultivadas en matrigel. Las células se tripsinizaron, se contaron y se mezclaron con matrigel. Los medios se

ponen encima sobre la mezcla de células-matrigel. Después de 6 días de incubación, se examinó la invasión celular (fig. 10B). Los resultados mostraron que las células tratadas con oligo de control reverso de hsOAF formaron estructuras penetrantes, invasivas, tridimensionales semejantes a redes, como hicieron las células de MDA-MB-435 normales; por otro lado, las células tratadas con oligo antisentido de hsOAF solamente forman colonias esféricas, lisas. De nuevo, se observaron también colonias penetrantes en células tratadas con oligo antisentido de hsOAF incubadas en el medio condicionado. Estos datos demuestran que la proteína hsOAF segregada se requiere para la invasividad y el potencial metastásico de células MDA-MB-435.

Se llevaron a cabo experimentos adicionales examinando si la proteína hsOAF segregada estuvo implicada en cultivo de células MDA-MB-435. Los resultados de ensayo de proliferación celular indicaron que la desactivación de secreción de proteína hsOAF ciertamente ralentiza la velocidad de proliferación de células MDA-MB-435, aunque el cambio era moderado.

Ejemplo 5

Análisis de bandas de northern de expresión de ARN en líneas celulares de cáncer de mama humanas y en tejidos humanos

Como se muestra en la figura 5, la expresión de ARNm se reguló al alza en líneas celulares metastásicas MDA-MB-231 y MDA-MB-435. El ARN se preparó usando el kit RNeasy de Quiagen. Se llevó a cabo un análisis de bandas de northern usando aislado de ARN total de 20-30 µg por extracción de tiocianato de guanidinio/fenol-cloroformo a partir de líneas celulares, a partir de tumores primarios, o a partir de metástasis en el pulmón. Se desarrollaron tumores primarios y metástasis pulmonares a partir de las líneas celulares inyectadas en ratones scid de acuerdo con procedimientos bien conocidos en la técnica. Los plásmidos que contienen clones parciales de ADNc clonados en vector pCR2.0-TA (Invitrogen) se radiomarcaron y se hibridaron a 65 °C en Express-hyb (Clontech). Entre todos los tejidos examinados, mostraron expresión de hsOAF significativa hígado, páncreas, bazo, ovario e intestino delgado. La expresión de ARNm de hsOAF se detectó también en corazón, músculo esquelético, riñón, próstata, colon y médula ósea. (Figura 9).

La tabla 3 muestra el porcentaje de positivos en hsOAF en una diversidad de tumores y tejidos normales.

Tabla 3

Inmunohistoquímica: Porcentaje de positivos en hsOAF		
	Tumor	Normal
Páncreas	9/11	0/9
Esófago	5/8	0/1
Hígado	3/6	0/13
Estómago	6/7	6/10
Mama	1/1	
Linfoma de Hodgkin	1/8	

Ejemplo 6

Ensayo en agar blando

Ensayo en agar blando: La fase inferior constaba de 2 ml de agar al 0,6 % en medios plaqueados recién preparados unas horas después de formación de una fase sobre las células. Para la fase celular, se retiraron células MDA-MB-435 según se describe anteriormente de la placa en tripsina al 0,05 % y se lavaron dos veces en medios. Las células se contaron en un contador de Coulter y se resuspendieron a 106 por ml en medios. Se situaron alícuotas de 10 ml con medios de comunicación en placas de 96 pocillos (controlando la cuenta con WST1), o se diluyeron adicionalmente para ensayo en agar blando. Se diluyeron 2000 células en 800 ml de agar al 0,4 % en pocillos duplicados por encima de una fase de fondo de agar al 0,6 %.

Fase de medios: Después de la capa solidificada de agar, se extrajeron 2 ml de medios en la parte superior y se añadió oligo de control reverso o antisentido sin vehículos de administración. Se añadieron medios y oligos recién preparados cada 3-4 días.

Las colonias se contaron en 10 días hasta 3 semanas. Los campos de las colonias se contaron a simple vista. Los valores de metabolismo de Wst-1 se usaron compensando pequeñas diferencias en número de células de partida.

Se pueden examinar campos mayores para el registro visual de las diferencias. Los resultados se muestran en la figura 6, en la que las células MDA-MB-435 tratadas con el antisentido formaron menos colonias comparadas con las células expuestas al oligonucleótido de control.

Listado de secuencias

5 <110> Chiron Corporation

<120> Genes expresados diferencialmente en cáncer de mama

<130> PP01656.003/200130

10

<140> PCT/US01/00899

<141> 9-1-2.001

<160> 6

15

<170> FastSEQ for Windows Versión 4.0

<210> 1

<211> 2366

20

<212> ADN

<213> Homo sapiens

<400> 1

ES 2 386 317 T3

ccgcgaggtg	cgcggtctct	ttaaggcggg	tcctggtggt	ttctgtttcc	tgaaggaagt	60
gacggggggg	gggattgaat	gaaaagtgca	aaacacagge	tcgcagcgct	ggagcccggg	120
gcgcgggagc	cgggccgggg	cagcgccgtc	tccgcctcgg	ggccgcccgg	ggcgccctgc	180
tgagcgctac	ccacgtgcgt	ccgcgccacc	tcgcgggcga	ccccgcggcc	aaggcccccg	240
gcgagcgggc	tcccggggcg	cccgaactag	cccccaactt	tgggcgaagt	ttgcctgcgc	300
ctctccccgc	ccccacgcgg	cgcgccgggg	ccgcggacgg	cagcggcccc	cggggatgcg	360
ccttccccgg	gtacccttgg	cgcgccctgc	gctgctgctg	ctgctgccgc	tgctcgcgcc	420
gctgctggga	acgggtgcgc	cggccgagct	gcgggtccgc	gtgcggctgc	cggacggcca	480
ggtgaccgag	gagagcctgc	aggcggacag	cgacgcggac	agcatcagcc	tcgagctgcg	540
caagcccgac	ggcaccctcg	tctccttcac	cgccgacttc	aagaaggatg	tgaaggcttt	600
ccgggcccctg	atcctggggg	agctggagaa	ggggcagagt	cagttccagg	ccctctgctt	660
tgtcaccag	ctgcagcaca	atgagatcat	ccccagtgag	gccatggcca	agctccggca	720
gaaaaatccc	cgggcagtcg	ggcaggcgga	ggaggttcgg	ggtctggagc	atctgcacat	780
ggatgtcgct	gtcaacttca	gccagggggc	cctgctgagc	ccccatctcc	acaacgtgtg	840
tgccgagggc	gtggatgcca	tctacaccgc	ccaggaggat	gtccggttct	ggctggagca	900
agggtgtggc	agttctgtgt	tcgaggctct	gcccaaggcc	tcagagcagg	cggagctgcc	960
tcgctgcagg	caggtggggg	accgcgggaa	gccctgcgtc	tgccactatg	gcctgagcct	1020
ggcctggtac	ccctgcattg	tcaagtactg	ccacagccgc	gaccggccca	cgccctacaa	1080
gtgtggcctc	cgcagctgcc	agaagagcta	cagcttcgac	ttctacgtgc	cccagaggca	1140
gctgtgtctc	tgggatgagg	atccctaccg	aggctagggt	gggagcaacc	tgggcggttg	1200
gctgctctgg	gcccactgct	cttcaccagc	cactagaggg	ggtggcaacc	cccacctgag	1260
gccttatttc	ccctccctccc	cactccctcg	gccttagagc	ctgggcccct	ctggccccat	1320
ctcacatgac	tgtgaagggg	gtgtggcatg	gcagggggtc	tcatgaaggc	acccccattc	1380
ccaccctgtg	ccttccttgc	gggcagagag	ggagagaagg	gctccccaga	tctacacccc	1440
tcctcctctg	atctccctcg	gagtgttcac	ttgcaagctg	ccaaaacatg	atggcctctg	1500
gttgttctgt	tgaactcctt	gaacgtttag	accctaaaag	gagtctatac	ctggacaccc	1560
acctccccag	acacaactcc	cttccccatg	cacacatctg	gaaggagctg	gcccctcagt	1620
cccttcctac	tccccaaaca	ggggctcact	atccccaaag	aaggagctgt	tggggaccca	1680
cgacgcagcc	cctgtactgg	attacagcat	attctcatct	ctggccccga	ggctgctgtg	1740
ggggcgagtg	gagacctccc	atcactgaga	cagatcacag	accacgagtg	ccttccccgg	1800
acctggacgt	tgctccaaa	acaggcacca	gctctttccc	tctctagaca	gaaatatttt	1860
tgtaaaggtt	tggggcaggg	agggagcatg	aagtacgagg	aaaacttgaa	ttccagattt	1920
llaatgcaaa	gtatttatca	tttctaccag	aaataaacgt	tttaagtttt	tacttgacta	1980
atgagaccca	gagtttggag	aaaacttttg	gccaatgctg	ccacctgatg	tcagaaagtg	2040
tccccacacc	ctagcagtg	cctatcttgg	aacaagaact	tcgaaagcac	ctactgtgtg	2100
ctcagccatt	tgaggaagga	aggaggagaa	ggaagatggt	actaggggaag	gatgagataa	2160
aacttctgca	cccaagacaa	tgagacagac	ataactgcaa	ccgtagtaag	ccagtcagaa	2220
atagccagcg	cgaaggcaag	agatgggggtg	gagattggaa	ccccgcttca	gatctgggct	2280
cggctactta	cctgctgtgc	agccatgggt	caagttgctt	gacctctctg	tgctccact	2340

cccttagcta taaaatgagc ttactt 2366

<210> 2

<211> 273

5 <212> PRT

<213> Homo sapiens

<400> 2

ES 2 386 317 T3

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

<210> 3

<211> 30

5 <212> PRT

<213> Homo sapiens

<400> 3

Met Arg Leu Pro Gly Val Pro Leu Ala Arg Pro Ala Leu Leu Leu Leu
 1 5 10 15
 Leu Pro Leu Leu Ala Pro Leu Leu Gly Thr Gly Ala Pro Ala
 20 25 30

10

<210> 4

<211> 24

<212> ADN

15 <213> Secuencia artificial

ES 2 386 317 T3

<220>

<223> Oligo antisentido usado para desactivar expresión de hsOAF

5 <400> 4

agctgcggat gccacacttg tagg 24

<210> 5

<211> 24

10 <212> ADN

<213> Secuencia artificial

<220>

<223> Oligo control reverso usado para desactivar expresión de hsOAF

15

<400> 5

ggatgttcac accgtaggcg tcga 24

<210> 6

20 <211> 332

<212> PRT

<213> Drosophila melanogaster

<400> 6

25

ES 2 386 317 T3

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

305 310 315 320
 Val Val Arg Gln Lys Gln Gln Cys Leu Trp Asp Glu
 325 330

REIVINDICACIONES

1. Una molécula de ácido nucleico aislada que comprende un polinucleótido seleccionado del grupo que consiste en:
 - (a) un polinucleótido que codifica los aminoácidos del 1 al 273 de SEC ID N.º: 2;
 - 5 (b) un polinucleótido que codifica los aminoácidos del 2 al 273 de SEC ID N.º: 2;
 - (c) un polinucleótido que codifica los aminoácidos del 26 al 273 de SEC ID N.º: 2;
 - (d) el complemento polinucleotídico del polinucleótido de (a), (b) o (c); y
 - (e) un polinucleótido al menos el 90 % idéntico al polinucleótido de (a), (b), (c), o (d) que se expresa diferencialmente entre las células de cáncer de mama con capacidad de metastásis alta y las células de

10 cáncer no metastásicas o con capacidad metastásica baja.
2. La molécula de ácido nucleico aislada de la reivindicación 1, que es ADN.
3. Una molécula de ácido nucleico aislada que comprende un polinucleótido que codifica un polipéptido en el que, salvo por al menos una sustitución aminoacídica conservadora, dicho polipéptido tiene una secuencia de aminoácidos seleccionada del grupo que consiste en:
 - 15 (a) los aminoácidos 1 a 273 de SEC ID N.º: 2;
 - (b) los aminoácidos 2 a 273 de SEC ID N.º: 2; y
 - (c) los aminoácidos 26 a 273 de SEC ID N.º: 2,

en la que dicho polipéptido se expresa diferencialmente entre las células de cáncer de mama con capacidad metastásica alta y las células de cáncer no metastásicas o con capacidad metastásica baja.
- 20 4. Un procedimiento de elaborar un vector recombinante que comprende insertar una molécula de ácido nucleico de la reivindicación 1 en un vector en unión operable a un promotor.
5. Un vector recombinante que comprende una molécula de ácido nucleico de la reivindicación 1 en unión operable a un promotor.
- 25 6. Un procedimiento de elaborar una célula huésped recombinante que comprende introducir el vector recombinante de la reivindicación 5 en una célula huésped.
7. Una célula huésped recombinante que comprende el vector recombinante de la reivindicación 5.
8. Un procedimiento recombinante de producir un polipéptido, que comprende cultivar la célula huésped recombinante de la reivindicación 7 en condiciones tales que dicho polipéptido se expresa y recuperar dicho polipéptido.
- 30 9. Un polipéptido aislado que comprende aminoácidos al menos idénticos al 95 % a aminoácidos seleccionados del grupo que consiste en:
 - (a) los aminoácidos 1 a 273 de SEC ID N.º: 2;
 - (b) los aminoácidos 2 a 273 de SEC ID N.º: 2; y
 - (c) los aminoácidos 26 a 273 de SEC ID N.º: 2,

en el que dicho polipéptido se expresa diferencialmente entre las células de cáncer de mama con capacidad

35 metastásica alta y las células de cáncer no metastásicas o con capacidad metastásica baja.
10. Un polipéptido aislado de acuerdo con la reivindicación 9, en el que salvo por al menos una sustitución aminoacídica conservadora, dicho polipéptido tiene una secuencia de aminoácidos seleccionada del grupo que consiste en:
 - 40 (a) los aminoácidos 1 a 273 de SEC ID N.º: 2,
 - (b) los aminoácidos 2 a 273 de SEC ID N.º: 2 y
 - (c) los aminoácidos 26 a 273 de SEC ID N.º: 2,

en el que dicho polipéptido se expresa diferencialmente entre las células de cáncer de mama con capacidad

metastásica alta y las células de cáncer no metastásicas o con capacidad metastásica baja.

11. Un polipéptido aislado de acuerdo con la reivindicación 9 que comprende aminoácidos seleccionados del grupo que consiste en:
- (a) los aminoácidos 1 a 273 de SEC ID N.º: 2,
 - (b) los aminoácidos 2 a 273 de SEC ID N.º: 2 y
 - (c) los aminoácidos 26 a 273 de SEC ID N.º: 2.
- 5
12. Una parte portadora de epítomos del polipéptido de la SEC ID N.º: 2 a condición de que dicha parte portadora de epítomos no sea aminoácidos 85-273 de SEC ID N.º: 2 o un fragmento de la misma.
13. Un anticuerpo aislado que se une específicamente al polipéptido de la reivindicación 9, 10 u 11 con la condición de que dicho anticuerpo no se una a los aminoácidos 85-273 de SEC ID N.º: 2.
- 10
14. Un procedimiento para comparar potencial metastásico de células tumorales en una primera y segunda muestra, que comprende: medir en dichas muestras de tejido un producto de expresión de un gen que comprende un polinucleótido seleccionado del grupo que consiste en:
- (a) un polinucleótido que codifica los aminoácidos del 1 al 273 de SEC ID N.º: 2;
 - (b) un polinucleótido que codifica los aminoácidos del 2 al 273 de SEC ID N.º: 2;
 - (c) un polinucleótido que codifica los aminoácidos del 26 al 273 de SEC ID N.º: 2;
 - (d) el complemento polinucleotídico del polinucleótido de (a), (b) o (c); y
 - (e) un polinucleótido al menos el 90 % idéntico al polinucleótido de (a), (b), (c), o (d) que se expresa diferencialmente entre las células de cáncer de mama con capacidad de metastásis alta y las células de cáncer no metastásicas o con capacidad metastásica baja,
- 15
- 20 en el que al menos una expresión dos veces mayor del producto en la primera muestra de tejido indica un potencial metastásico mayor comparado con la segunda muestra de tejido.
15. El procedimiento de la reivindicación 14, en el que dicho tumor es un cáncer de mama.

ES 2 386 317 T3

730 740 750 760 770 780 790 800 810
GAAAAATCCCGGGCAGTGCAGGAGGAGGTTGGGGTCTGGAGCATCTGCACATGGATGTCGCTGTCAACTCAGCCAGGGGGC
CTTTTAGGGGCCGTACGCGCTCCGCCTCCTCAAGCCCAGACCTCGTAGACGTGTACCTACAGCGACAGTTGAAGTCGGTCCCGG
K N P R A V R Q A E E V R G L E H L H M D V A V N F S Q G A >
TRADUCCIÓN DE OAF HUMANA [A] >

820 830 840 850 860 870 880 890 900
CCTGCTGAGCCCCATCTCCACAACGTGTGTGCCGAGGCGGTGGATGCCATCTACACCCGCCAGGAGGATGTCGGTTCGGCTGGAGCA
GGACGACTCGGGGTAGAGGTGTTGCACACAGGCTCCGGCACCTACGGTAGATGTGGGCGTCTCCTACAGGCCAAGACCGACCTCGT
L L S P H L H N V C A E A V D A I Y T R Q E D V R F W L E Q >
TRADUCCIÓN DE OAF HUMANA [A] >

910 920 930 940 950 960 970 980 990
AGGTGTGGACAGTTCGTGTTCCGAGGCTCTGCCAAGGCCCTCAGAGCAGGCGGAGTGCCTCGCTGCAGGCAGGTGGGGACCGGGGAA
TCCACACCTGTCAAGACACAAGCTCCGAGACGGTTCGGAGTCTCGTCCGCCTCGACGGAGCGAGTCCGTCCACCCCTGGCGCCCTT
G V D S S V F E A L P K A S E Q A E L P R C R Q V G D R G K >
TRADUCCIÓN DE OAF HUMANA [A] >

1000 1010 1020 1030 1040 1050 1060 1070 1080
GCCCTGCGTCTGCCACTATGGCCTGAGCCTGGCTGGTACCCCTGCATGCTCAAGTACTGCCACAGCCGCGACCGGCCACGCCCTACAA
CGGGACGCAGCGGTGATACCGGACTCGGACCGGACCATGGGGACGTACGAGTTCATGACGGTGTCCGGCTGCCCAGGAGGATGTT
P C V C H Y G L S L A W Y P C M L K Y C H S R D R P T P Y K >
TRADUCCIÓN DE OAF HUMANA [A] >

1090 1100 1110 1120 1130 1140 1150 1160 1170
GTGTGGCATCCCGAGCTGCCAGAAGAGCTACAGCTTCGACTTCTACGTGCCCCAGAGGCAGCTGTGTCCTGCGATGAGGATCCCTACCC
CACACCGTAGGGGTGACGGTCTTCTCGATGTCGAAGCTGAAGATGCACGGGTCTCCGTGACACAGAGACCTACTCCTAGGGATGGG
C G I R S C Q K S Y S F D F Y V P Q R Q L C L W D E D P Y P >
TRADUCCIÓN DE OAF HUMANA [A] >

1180 1190 1200 1210 1220 1230 1240 1250 1260
AGGCTAGGGTGGGAGCAACCTGGGGGGTGGCTGCTCTGGGCCACTGCTCTTACCAGCCACTAGAGGGGGTGGCAACCCACCTGAG
TCCGATCCACCCCTCGTTGGACCCGCCACCGACGAGACCCGGGTGACGAGAAGTGGTGGTGTCTCCCCACCGTTGGGGTGGACTC
G *

1270 1280 1290 1300 1310 1320 1330 1340 1350
GCCTTATTTCCCTCCCTCCCACTCCCTGGCCCTAGAGCCTGGGCCCTCTGGCCCATCTCACATGACTGTGAAGGGGGTGGCATG
CGGAATAAAGGGAGGGGGTGGGGGACCGGGATCTCGGACCCGGGGAGACCGGGTAGAGTGTACTGACACTTCCCCACACCGTAC

Fig. 4B

ES 2 386 317 T3

1360 1370 1380 1390 1400 1410 1420 1430 1440
GCAGGGGGTCTCATGAAGGCACCCCATTTCCACCGCTGTCCTTCCCTGCGGGCAGAGAGGGAGAGAAGGGCTCCCCAGATCTACACCCC
CGTCCCCAGAGTACTTCCGTGGGGTAAGGGTGGGACACGGAAGGAACGCCCGTCTCCCTCTTCCCGAGGGTCTAGATGTGGGG

1450 1460 1470 1480 1490 1500 1510 1520 1530
TCCCTCCTGCATCTCCCTGGAGTGTTCACCTGCAAGCTGCCAAAACATGATGGCCTCTGGTGTCTGTGTAACCTTGAACGTTTAG
AGGGAGGACGTAGAGGGGACCTCACAAGTGAACGTTGACGGTTTTGTACTACCGGAGACCAACAAGACAACCTGAGGAACCTGCAAAATC

1540 1550 1560 1570 1580 1590 1600 1610 1620
ACCTAAAAGGAGTCTATACCTGGACACCCACCTCCCCAGACACAACCTCCCTTCCCATGCACACATCTGGAAGGAGCTGGCCCTCAGT
TGGGATTTTCTCAGATATGGACCTGTGGTGGAGGGGTCTGTGTTGAGGGGAAGGGGTACGTGTGTAGACCTTCTCGACCGGGGAGTCA

1630 1640 1650 1660 1670 1680 1690 1700 1710
CCCTTCTACTCCCCAACAGGGGCTCACTATCCCAAGAAGGAGCTGTTGGGGACCCAGCAGCCCTGTACTGGATTACAGCAT
GGGAAGGATGAGGGGTTGTTCCCGAGTGATAGGGTTTCTTCTCGACAACCCCTGGGTCTGCGTGGGGACATGACCTAATGTGCTA

1720 1730 1740 1750 1760 1770 1780 1790 1800
ATTCTCATCTCTGGCCCCGAGGCTGCTGTGGGGCAGTGGAGACCTCCCATCACTGAGACAGATCACAGACCAGTGCCTTTCCCGG
TAAGAGTAGAGACCGGGCTCCGACGGACCCCGCTCACCTCTGGAGGGTAGTGACTCTGTCTAGTGTCTGGTGTCTACGGAAAGGGCC

1810 1820 1830 1840 1850 1860 1870 1880 1890
ACCTGGACGTTGCCTCCAAAACAGGCACCAGCTCTTTCCCTCTCTAGACAGAAATATTTTGTAAAGTTCTGGGGCAGGGAGGGAGCATG
TGGACCTGCAACGGAGGTTTTGTCCGGTGGTGGAGAAAGGGAGAGATCTGCTTTATAAAAACATTCCAAGACCCCGTCCCTCCCTCGTAC

1900 1910 1920 1930 1940 1950 1960 1970 1980
AAGTACGAGGAAAACCTGAATCCAGATTTTTAATGCAAAGTATTTATCATTCTACCAGAAATAAACGTTTTAAGTTTTACTTGACTA
TTCATGCTCCTTTTGAACCTAAGGTCTAAAAATTACGTTTCATAAATAGTAAAGATGGTCTTTATTTGCAAAATTCAAAAATGAACCTGAT

1990 2000 2010 2020 2030 2040 2050 2060 2070
ATGAGACCCAGAGTTTGGAGAAAACTTTTGGCCAATGCTGCCACCTGATGTCAGAAAGTGTCCCCACACCCTAGCAGTGGCTATCTTGG
TACTCTGGGTCTCAAACCTCTTTTGAACCCTTTTACGACGGTGGACTACAGTCTTTCACAGGGGTGTGGGATCGTACC GGATAGAACC

2080 2090 2100 2110 2120 2130 2140 2150 2160
AACAGAACCTCGAAAGCACCTACTGTGTGCTCAGCCATTTGAGGAAGGAAGGAGGAGAAGGAAGATGTTACTAGGAAGGATGAGATAA
TTGTTCTTGAAGCTTTCTGTGGATGACACAGAGTGGTAAACTCTTCCCTCCTCCTCTCTACAATGATCCCTTCTACTCTATT

Fig. 4C

ES 2 386 317 T3

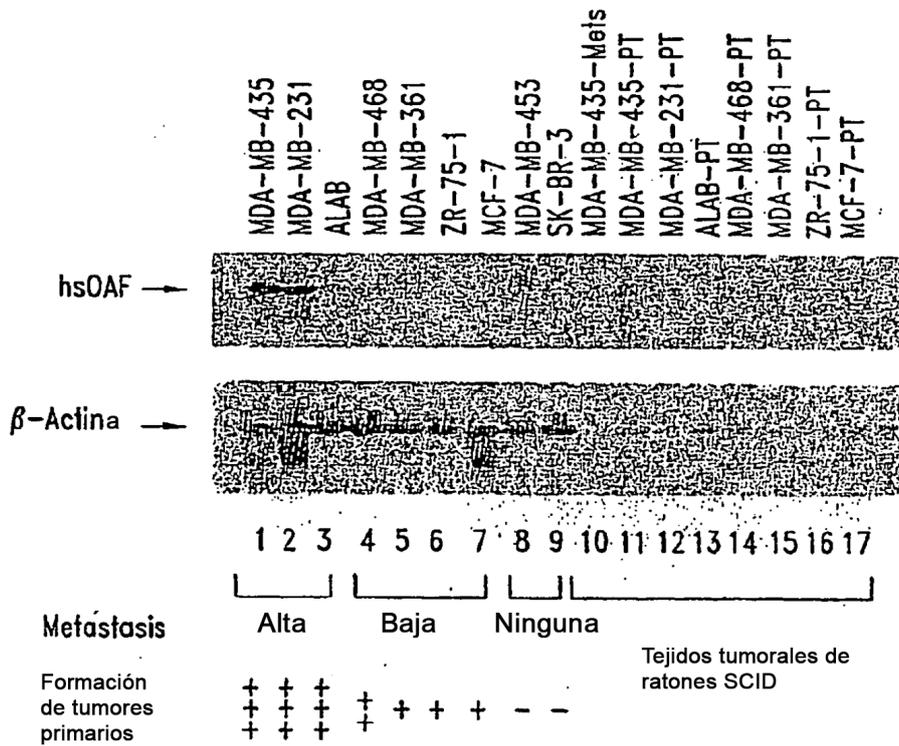
2170 2180 2190 2200 2210 2220 2230 2240 2250
AACTTCTGCACCCAAGACAATGAGACAGACATAACTGCAACCGTAGTAAGCCAGTCAGAAATAGCCAGCGCGAAGGCAAGAGATGGGGTG
TTGAAGACGTGGGTTCTGTTACTCTGTCTGTATTGAGGTTGGCATTCATTGGTCAGTCTTTATCGGTCGCGCTCCGTTCTTACCCAC

2260 2270 2280 2290 2300 2310 2320 2330 2340
GAGATTGGAACCCCGCTTCAGATCTGGGCTCGGCTACTTACCTGCTGTGCAGCCATGGGTCAAGTTGCTTGACCTCTCTGTGCCTCCACT
CTCTAACCTTGGGGCGAAGTCTAGACCCGAGCCGATGAATGGACGACACGTCGGTACCCAGTTCAACGAACTGGAGAGACACGGAGGTGA

2350 2360
CCCTTAGCTATAAAATGAGCTTACTT
GGGAATCGATATTTACTCGAATGAA

Fig. 4D

5



10

Fig. 5

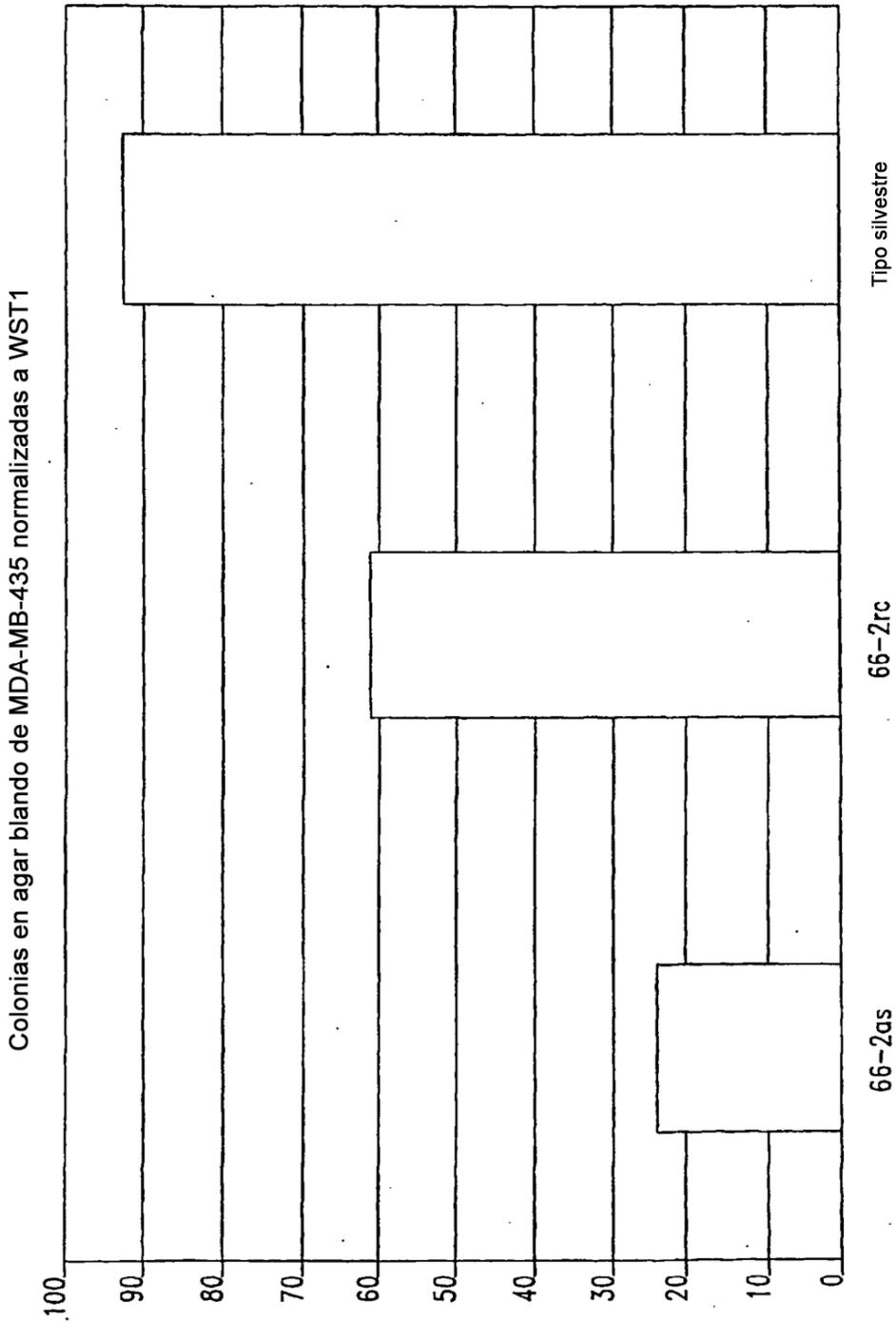


Fig. 6

F 1 MRIPG-----VPLAR 10
 1 MILKEEHPHQSIETAANAARQAQVRWRMAHLKALSRTRTAHGNCCGRVVSKNHFFKHSR 60

F 11 PALLLPLAPLLGTAFAELNVAVRLPDGQVTEESLDADSDADSISELRKPDGILVS 70
 61 AFLMRLLCNLMNADAFASQLINVCNOGGEVQESITSNIGEDLITLHCKIDGILIT 120

F 71 FTADIRKQVKVFRALITGELKKGQSQFOALCFMTOQHNEITPSEAMAKLRQKNPRAVRQ 130
 121 QVIDIRNEVQTLKALVLGEEKRGQSQYQWCFATKFNKGFHSSAAMAKLRQKNPHTIRT 180

F 131 AEEVRGLEHLMDVAVNFSQGALLSPHLHNVC AEAVDAIVTRDEDVRFAL ED- QVDSVVF 189
 181 REEIKGHEITFMSSVQLNRSLEITRHLQGLCAEAMDATVROVDLKAWELEPSSISL 240

F 190 KALPKASEQAELEPCROVGDRCNFCVCHVGLSLAVVPCMLKVCHSR-----D----- 236
 241 KAATEKFPTLSTRONEVSSLWAPCLCNLETCIQVPCGLKVCKGKVAGADSSGAQQQA 300

F 237 RPTVVKGIRSCCKSYSDFVVPDRQLCWDEDPYPG 273
 301 QPTVRCGIKTRKCTQFTYVVRQKCCCLWDE 332

Fig. 7

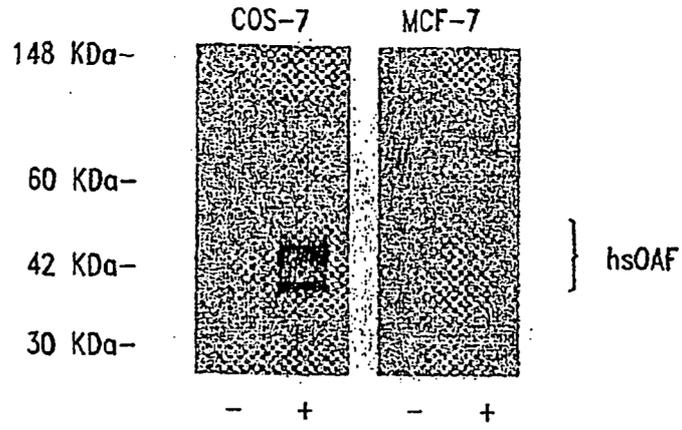


Fig. 8A

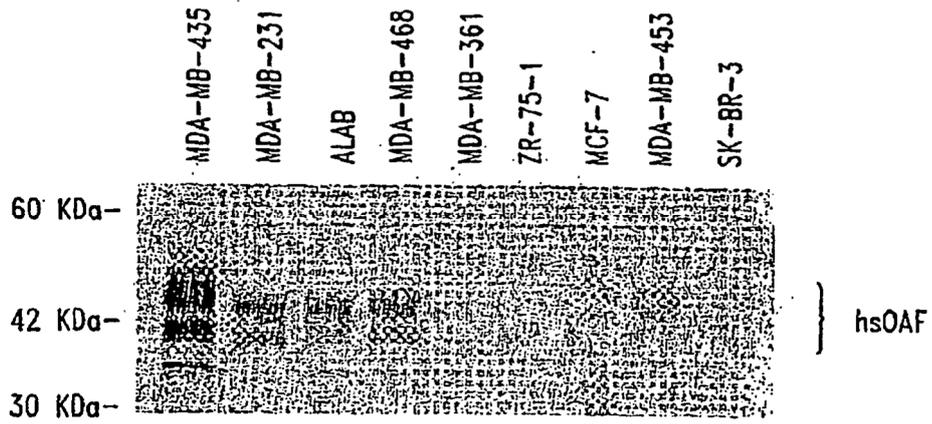
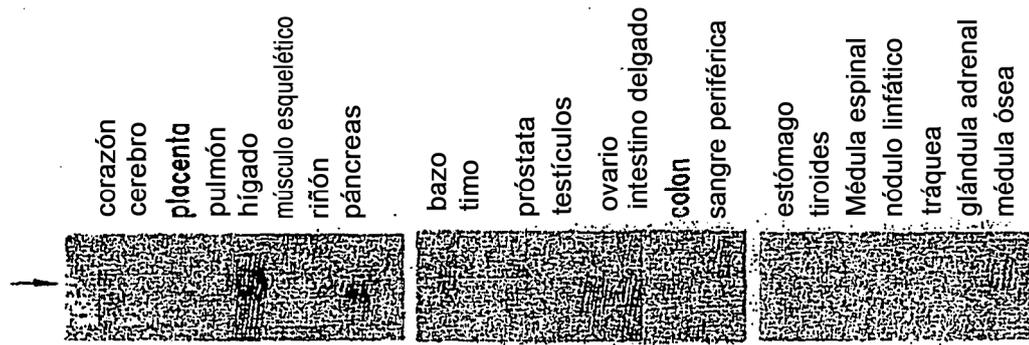


Fig. 8B

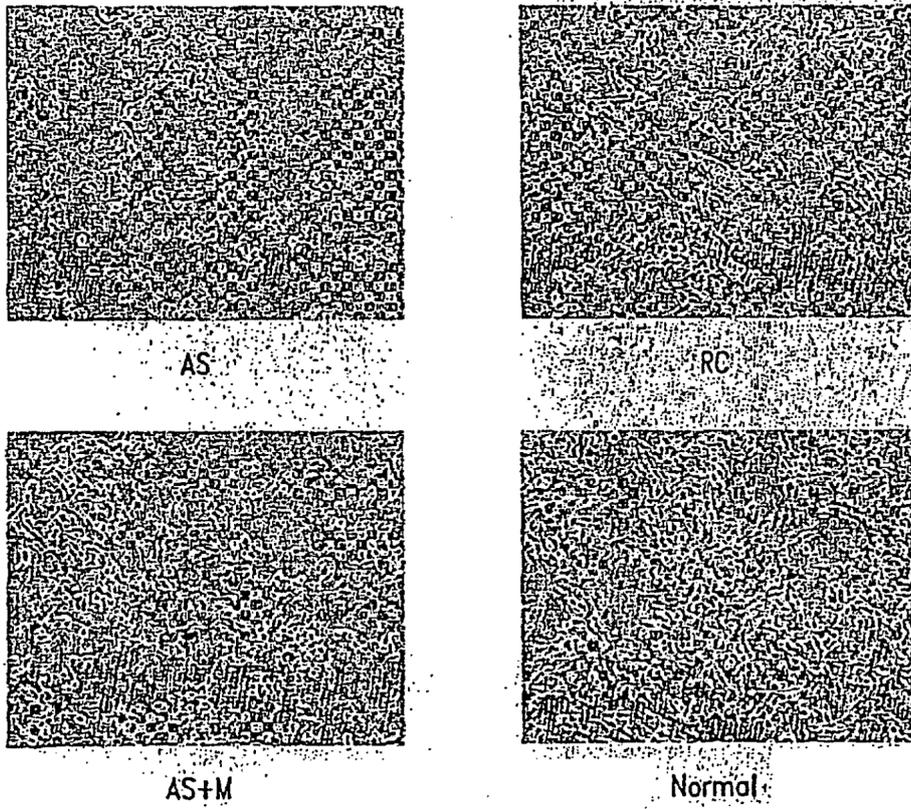
5



10

Fig. 9

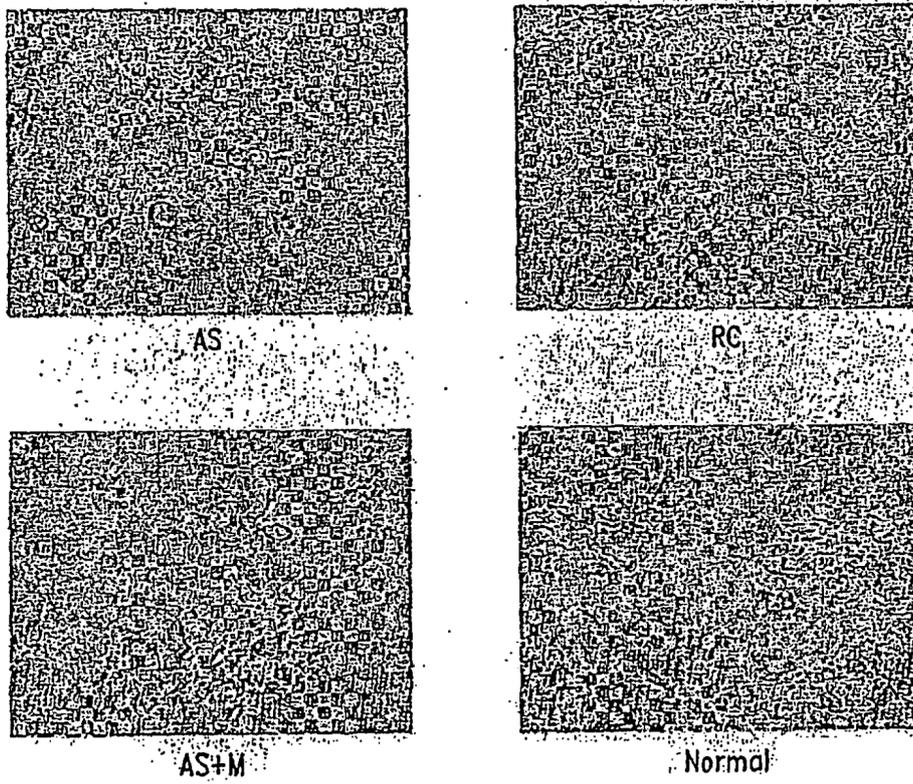
5



10

Fig. 10A

5



10

Fig. 10B

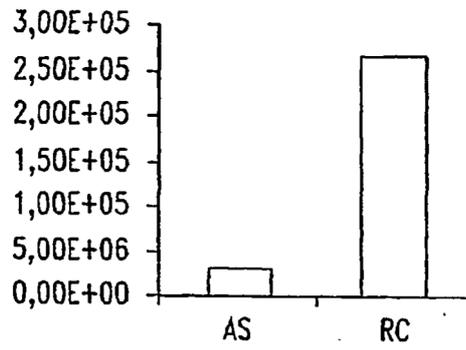


Fig. 12A

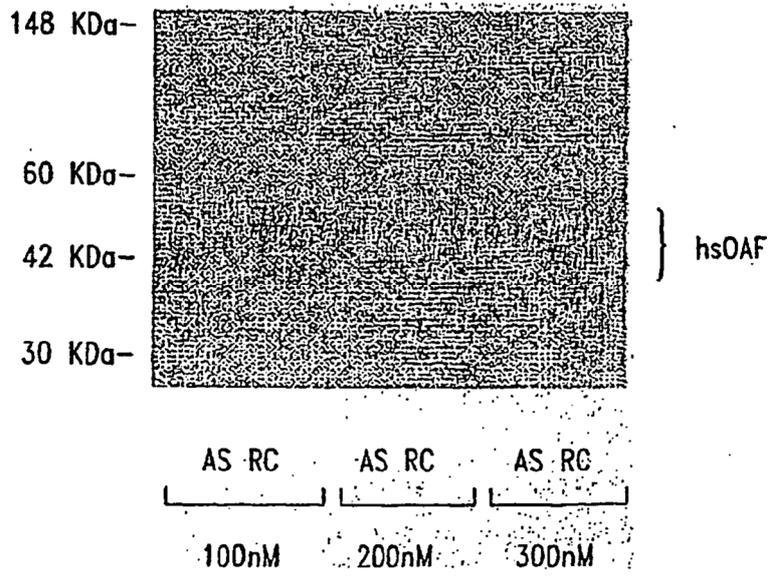


Fig. 12B