



(19)



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 311 016**

(51) Int. Cl.:

<b>C12N 15/11</b> (2006.01)	<b>C12N 15/12</b> (2006.01)
<b>C12N 15/62</b> (2006.01)	<b>C12N 15/86</b> (2006.01)
<b>C12N 5/10</b> (2006.01)	<b>C07K 14/47</b> (2006.01)
<b>C07K 16/18</b> (2006.01)	<b>A61K 31/713</b> (2006.01)
<b>C12Q 1/68</b> (2006.01)	<b>G01N 33/577</b> (2006.01)
<b>G06F 17/30</b> (2006.01)	

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Número de solicitud europea: **01942365 .6**

(96) Fecha de presentación : **09.01.2001**

(97) Número de publicación de la solicitud: **1248841**

(97) Fecha de publicación de la solicitud: **16.10.2002**

(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:  
**01.02.2009**

(45) Fecha de la publicación del folleto de la patente:  
**01.02.2009**

(73) Titular/es: **Novartis Vaccines and Diagnostics, Inc.**  
**4560 Horton Street**  
**Emeryville, California 94608, US**

(72) Inventor/es: **Kaufmann, Joerg;**  
**Harrowe, Greg;**  
**Reinhard, Christoph y**  
**Kang, Sanmao**

(74) Agente: **Elzaburu Márquez, Alberto**

ES 2 311 016 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.

**5 Campo técnico**

Esta invención se refiere al uso de moléculas de oligonucleótidos antisentido en la preparación de una composición farmacéutica para el tratamiento del cáncer de mama y/o de metástasis de cáncer de mama.

**10 Antecedentes de la invención**

El cáncer de mama es una de las enfermedades malignas más comunes, con aproximadamente 1 000 000 casos nuevos al año en el mundo. A pesar de los diversos marcadores histoquímicos, genéticos e inmunológicos, los clínicos todavía tienen dificultad para predecir qué tumores metastatizarán en otros órganos. Algunos pacientes necesitan un tratamiento adyuvante para prevenir la recidiva y metástasis y otros no. Sin embargo, distinguir entre estas dos subpoblaciones de pacientes no es sencillo y el plan de tratamiento no es fácil de hacer. Existe en la técnica la necesidad de nuevos marcadores para distinguir entre los tumores que metastatizarán, o que han metastatizado, y los que tienen menos probabilidad de metastatizar.

**20 Sumario de la invención**

Según la presente invención, se proporciona el uso de una molécula de oligonucleótido antisentido en la preparación de una composición farmacéutica para el tratamiento del cáncer de mama y/o de metástasis de cáncer de mama. El oligonucleótido antisentido comprende una secuencia antisentido de al menos 8 nucleótidos de longitud que es complementaria a una porción de la secuencia codificadora que abarca los nucleótidos del 365 al 1173 del SEQ ID NO: 1.

**Breve descripción de los dibujos**

30 La Figura 1 ilustra la secuencia polinucleotídica de la “*Out at First*” humana (SEQ ID NO: 1).

La Figura 2 ilustra la secuencia de aminoácidos codificada por el SEQ ID NO: 1 (SEQ ID NO: 2).

La Figura 3 ilustra el péptido señal putativo (SEQ ID NO: 3).

35 La Figura 4 ilustra la traducción del SEQ ID NO: 1 (SEQ ID NO: 1, polinucleótido; SEQ ID NO: 2, secuencia de aminoácidos).

La Figura 5 ilustra la expresión de hsOAF en comparación con la de  $\beta$ -Actina en líneas de células tumorales y en tejidos tumorales procedentes de ratones SCID, desarrollados a partir de las líneas celulares. “PT” hace referencia al tumor primario.

La Figura 6 ilustra el crecimiento de colonias producido por células MDA-MB-435 en agar blando después del tratamiento con un oligonucleótido antisentido, SEQ ID NO: 4 (66-2as), o con un oligonucleótido inverso de control, SEQ ID NO: 5 (66-2rc), en comparación con células no tratadas (WT).

La Figura 7 es un alineamiento de la secuencia de aminoácidos de la OAF humana con la secuencia de aminoácidos de la OAF de *Drosophila*.

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

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

55 Figura 10. La Figura 10A ilustra los cambios morfológicos observados en las células MDA-MB-435 después del tratamiento con un oligonucleótido antisentido (SEQ ID NO: 4). AS = antisentido; RC = control inverso (del inglés, “*Reverse Control*”) (SEQ ID NO: 5); M = medio acondicionado. La Figura 10B ilustra la invasión celular después del tratamiento de las células MDA-MB-435 con AS, RC y RC + M.

60 La Figura 11 ilustra la secuencia señal pronosticada de la OAF humana (subrayado doble).

Figura 12. La Figura 12A y la Figura 12B ilustran la secreción de la hsOAF por las células MDA-MB-435 tratadas con un oligonucleótido antisentido (SEQ ID NO: 4) o con un oligonucleótido inverso de control (SEQ ID NO: 5).

**65 Descripción detallada de la invención**

La causa más letal en pacientes con cáncer de mama es la metástasis de los carcinomas de mama y su proliferación en lugares alejados (principalmente, pulmón y hueso). La metástasis es un proceso multietapa mediante el cual las

células tumorales migran desde el tumor primario, se diseminan a través de los vasos sanguíneos y linfáticos y luego se depositan en órganos humanos específicos en donde recolonizan. Schirmacher, V., *Adv. Cancer Res.* 43: 1-73, 1985 y Liotta, L. A. y col., *Cell* 64(2): 327-36 (1991). Durante este proceso, la invasividad de las 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 ello el esclarecimiento de las causas moleculares de la invasión de las células tumorales y de la metástasis es esencial para el desarrollo de procedimientos terapéuticos eficientes para pacientes con cáncer de mama. Los genes que presentan una expresión diferencial en metástasis de tumores de mama son dianas potenciales que desempeñan papeles críticos durante la metástasis. La identificación de esos genes y su función biológica contribuirá significativamente al desarrollo del tratamiento y diagnóstico del cáncer de mama.

Se han descubierto algunos genes importantes implicados en la metástasis de tumores de mama. La pérdida de receptores estrogénicos y la presencia de vimentina se han asociado con la invasividad de tumores de mama humano y con un mal pronóstico de los mismos, y también se correlacionan con la invasividad y con el potencial metastásico de líneas celulares humanas de cáncer de mama. Aamdal S. y col., *Cancer* 53(11): 2525-9 (1984); Clark, G.M. y col., *Semin. Oncol.*, 2 Suplemento 1: 20-5 (1988); Raymond, W.A. y col., *J. Pathol.* 157(4): 299-306 (1989); Raymond, W.A., y col., *J. Pathol.* 158(2): 107-14 (1989); y Thompson, E.W. y col., *J. Cell Physiol.* 150(3): 534-44 (1992). Una infraexpresión de E-caderina se ha involucrado en la invasividad de tumores mamarios. Vleminckx, K. y col., *Cell* 66(1): 107-19 (1991) y Oka, H. y col., *Cancer Res.* 53(7): 1696-701 (1993). La maspina, un inhibidor de proteasas expresado en células epiteliales mamarias normales, pero no en la mayoría de las líneas celulares de carcinoma de mama, fue capaz de suprimir la capacidad de las células MDA-MB-435 para inducir tumores y metastatizar en ratones, y para invadir la membrana basal *in vitro*. La pérdida de la expresión de maspina tuvo lugar muy frecuentemente en cánceres avanzados. Zou, Z. y col., *Science* 263(5146): 526-9 (1994) y Seftor, R.E., y col., *Cancer Res.* 58(24): 5681-5 (1998).

La sobreexpresión de TIMP-4 (inhibidor tisular de las metaloproteinasas-4) o de CLCA2 (canal 2 de cloruro activado por  $\text{Ca}^{2+}$ ) en células MDA-MB-435, producida por transfección, inhibió la tumorigenicidad, la invasividad y la capacidad de metástasis de las células. Wang, M. y col., *Oncogene* 14(23): 2767-74 (1997) y Gruber, A.D. y col., *Cancer Res.* 59(21): 5488-91 (1999). Se ha descubierto que la sobreexpresión de los receptores de los factores de crecimiento IGF-IR y p185<sup>ErbB-2</sup> está implicada en la metástasis de cáncer de mama. Surmacz, E. y col., *Breast Cancer Res. Treat* 47(3): 255-67 (1998); Dunn, S.E. y col., *Cancer Res.* 58(15): 3353-61 (1998); Tan, M. y col., *Cancer Res.* 57(6): 1199-205 (1997); Dhingra, K. y col., *Semin Oncol.* 23(4):436-45 (1996); y Revillion, F. y col., *Eur. J. Cancer* 34(6): 791-808 (1998).

Se ha informado de que la aspartil-proteasa cathepsina D es un marcador de mal pronóstico para pacientes con cáncer de mama y existe una correlación significativa entre una concentración elevada de cathepsina D en el citosol de un cáncer de mama primario y el desarrollo de metástasis, aunque no se encontró correlación entre la secreción de cathepsina D y la capacidad de invasión de líneas celulares de cáncer de mama. Rochefort, H., *Breast Cancer Res Treat* 16(1): 3-13 (1990); Johnson, M. D. y col., *Cancer Res.* 53(4): 873-7 (1993); y Rochefort, H. y col., *Clin Chim Acta.* 291(2):157-70 (2000). La osteopontina, una glicoproteína secretada de unión a integrina, que se piensa que está implicada en la resorción ósea y en la formación de hueso, es capaz de inducir una migración e invasión de células de carcinoma mamario. Los niveles de osteopontina (niveles en células tumorales o en plasma) se han asociado con un aumento en la malignidad del cáncer de mama. Denhardt, D.T. y col., *FASEB J.* 7(15): 1475-82 (1993); Denhardt, D.T. y col., *J. Cell Biochem Suppl.*, 30-31: 92-102 (1998); Tuck, A.B. y col., *J. Cell Biochem.* 78(3): 465-75 (2000); Tuck, A.B. y col., *Oncogene* 18(29): 4237-46 (1999); y Singhal, H. y col., *Clin Cancer Res.* 3(4):605-11 (1997).

Se describe la clonación de un nuevo gen sobreexpresado en líneas celulares de cáncer de mama humano altamente metastásico. Este gen codifica una proteína secretada y la secreción de esta proteína se ha confirmado que es mucho mayor en líneas celulares de cáncer de mama humano altamente metastásicas que en líneas celulares poco metastásicas/no metastásicas. La supresión (del inglés “*knock out*”) de la secreción de esta proteína de la línea celular MDA-MB-435 agresiva por medio de la tecnología basada en oligonucleótidos antisentido produjo un cambio morfológico significativo junto con una disminución de la invasividad y de la velocidad de proliferación de las células. A este gen se le denomina hsOAF basándose en su homología con el gen OAF (“*out at first*”) de *Drosophila*. Bergstrom, D.E. y col., *Genetics* 139(3): 1331-46 (1995) y Merli, C. y col., *Genes Dev.* 10(10): 1260-70 (1996).

Un descubrimiento de la presente invención es que un polinucleótido se expresa de manera diferencial en células de cáncer de mama altamente metastásicas y en células cancerosas no metastásicas o poco metastásicas. Este polinucleótido por lo tanto está en relación con un gen marcador metastásico. Esta información se puede utilizar para fabricar reactivos diagnósticos específicos para los productos de expresión del gen de expresión diferencial. También se puede usar en métodos diagnósticos y pronósticos que ayudarán a los clínicos a planificar regímenes de tratamiento adecuados para los cánceres, especialmente para el cáncer de mama.

Este polinucleótido se muestra en la Figura 1 (SEQ ID NO: 1), y el marco de lectura abierto (ORF) (del inglés, “*Open Reading Frame*”) pronosticado codifica un polipéptido mostrado en la Figura 2 (SEQ ID NO: 2). Los 30 primeros residuos de aminoácidos (SEQ ID NO: 3) comprenden un péptido señal putativo, con un sitio de escisión por proteasa pronosticado, indicado con “\*”: APLLG \* TGAPA (entre los aminoácidos situados en las posiciones 25 y 26 del SEQ ID NO: 3).

La secuencia polinucleotídica descrita comparte cierta homología con un gen de *Drosophila* conocido como “*Out at First*” (oaf). La transcripción de oaf produce tres clases de RNA poliadenilados de forma alternativa, cuya expresión está regulada por el desarrollo. Todos los transcritos de oaf contienen dos ORF adyacentes, separados por un único codón UGA de terminación. La supresión del codón UGA durante la traducción podría conducir a la producción de diferentes proteínas a partir de la misma molécula de RNA. Durante la oogenesis, el RNA de oaf se expresa en células nutricias de todas las edades y su contribución al huevo es por el lado materno.

Durante el desarrollo embrionario, la transcripción zigótica del gen oaf tiene lugar en pequeñas agrupaciones de células en la mayoría o en todos los segmentos en el momento de la extensión de la banda germinal y posteriormente en un patrón repetido segmentalmente en el desarrollo del sistema nervioso central. El gen oaf se expresa también en las gónadas de embriones, larvas y adultos, de ambos sexos. (Bergstrom, D.E. y col., *Genetics* 139: 1331-1346, 1995).

El polinucleótido se expresó de manera diferencial en siete pares de líneas celulares metastásicas de cáncer de mama, frente a líneas celulares de cáncer de mama no metastásicas o poco metastásicas. Las líneas celulares usadas son MDA-MB-361 (derivada de un adenocarcinoma de mama humano), MDA-MB-231 (células de cáncer de mama humano, metastásicas en hueso y/o pulmón); MDA-MB-468 (derivada de células de cáncer de mama humano, negativas a receptores estrogénicos); MCF-7 (células de cáncer de mama humano, no metastásicas); ZR-75-1 (derivada de carcinomas de mama humanos, positivos a receptores estrogénicos, Engle y col., *Cancer Res.* 38: 3352-64 (1978)); MDA-MB-435 (derivada de células de carcinoma de mama humano, negativas a receptores estrogénicos, Rishi y col., *Cancer Res.* 56: 5246-5252 (1996)).

El perfil de expresión es el siguiente:

TABLA 1

25

30

35

40

Pareja de líneas celulares	Cociente 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

45

La regulación por aumento de la expresión del mRNA se confirmó mediante análisis de transferencia Northern usando RNA total procedente de las líneas celulares (Figura 5).

50

Las líneas celulares en las que se comparó la expresión del polinucleótido representan cánceres de mama humanos de potencial metastásico variable. Los cultivos de la línea celular ZR-75-1 derivaron de un derrame ascítico neoplásico de una paciente con cáncer de mama. Las líneas celulares cultivadas *in vitro*, se asemejaban muy estrechamente a la morfología observada en biopsias o en preparaciones celulares procedentes de los donantes de las células originales. Las células ZR-75-1 son estimuladas específicamente por estrógenos y se han usado como un sistema modelo para estudiar la capacidad de respuesta a los estrógenos. Engel, L.W. y col., *Cancer Res.* 38: 3352-3364, 1978.

55

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

60

La línea celular MDA-MB-361 derivaba de un adenocarcinoma de mama humano, específicamente de un sitio neoplásico. ATCC Número HTB-27. La expresión diferencial de genes Wnt humanos se ha estudiado en esta línea celular. Huguet, E.L. y col., *Cancer Res.* 54: 2615-2621, 1994.

65

Una vez que se ha producido la metástasis, las células del tumor primario de mama invaden las membranas basales y se diseminan por otros órganos del cuerpo, y la probabilidad de supervivencia de los pacientes con cáncer de mama se hace bastante escasa. Es crítico identificar genes que participan en la invasión y metástasis del cáncer de mama en beneficio del diagnóstico clínico y del tratamiento. Esos genes son marcadores potenciales de diagnóstico o dianas candidatas para el desarrollo de fármacos terapéuticos. Por ejemplo, la presencia de vimentina en tumores humanos

de mama se ha asociado con la falta de receptores estrogénicos y con la invasividad del tumor, considerada como un marcador de mal pronóstico. Raymond, W.A. y col., *J. Pathol.* 157(4): 299-306 (1989); Raymond, W.A. y col., *J. Pathol.* 158(2): 107-14 (1989); y Thompson, E.W. y col., *J. Cell Physiol.* 150(3): 534-44 (1992). Las actividades incrementadas de las metaloproteinasas de la matriz se relacionan con el fenotipo metastásico de los carcinomas, especialmente del cáncer de mama. Basset, P. y col., *Nature* 348(6303): 699-704 (1990) y Basset, P. y col., *Cancer* 74 (suplemento 3): 1045-9 (1994). La osteopontina, una glicoproteína secretada de unión a integrinas, es capaz de inducir un incremento en la invasividad de células humanas del epitelio mamario y se la ha asociado con un aumento de la malignidad en el cáncer de mama. Tuck, A.B. y col., *J. Cell Biochem.* 78(3): 465-75 (2000); Tuck, A.B. y col., *Oncogene* 18(29): 4237-46 (1999); y Singhal, H. y col., *Clin Cancer Res.* 3(4): 605-11 (1997).

En el presente documento se describe la identificación de una nueva proteína secretada (hsOAF) implicada en la metástasis del cáncer de mama. Las líneas celulares de cáncer de mama humano usadas para esclarecer el papel de esta proteína, se dividen en tres grupos según sus potenciales metastásicos: altamente metastásicas, poco metastásicas y no metastásicas. Aprovechando los diferentes potenciales metastásicos entre estos grupos de líneas celulares y utilizando la avanzada tecnología basada en micromatrices, se identificaron genes que se expresan de manera diferencial en las líneas celulares de cáncer de mama humano altamente metastásicas y en las líneas poco metastásicas/no metastásicas. Estos genes son buenos candidatos para la validación de la expresión y para estudios funcionales que pudieran conducir a un mejor conocimiento del mecanismo molecular de la metástasis del cáncer de mama. El gen hsOAF es el centro de esta investigación ya que codifica una nueva proteína secretada. La expresión del mRNA de hsOAF es bastante común en diferentes tejidos humanos normales. Sin embargo, la secreción de hsOAF es mucho más fuerte en líneas celulares humanas de cáncer de mama altamente metastásicas, que en las poco metastásicas/no metastásicas, y MDA-MB-435 tiene la secreción de hsOAF más alta. La denominación del gen hsOAF se basó en su homología con el gen OAF ("out at first") de *Drosophila*. Sin embargo, la proteína OAF de *Drosophila* es posible que no sea una proteína secretada ya que no posee un secuencia típica de péptido señal en el extremo N-terminal (Fig. 7).

Para abordar la importancia de la proteína hsOAF secretada en la metástasis del cáncer de mama, se usó una tecnología basada en oligonucleótidos antisentido para suprimir la expresión de hsOAF de manera específica. La tecnología basada en oligonucleótidos antisentido es una manera eficiente y rápida de disminuir espectacularmente la expresión génica en estudios funcionales de genes. Stein, C.A. y col., *Science* 261(5124): 1004-12 (1993); Defacque, H. y col., *J. Cell Physiol.* 178(1): 109-19 (1999). La supresión de la secreción de la proteína hsOAF en células MDA-MB-435 altamente metastásicas produjo un cambio en la forma de las células, una disminución de la invasividad de las células y una la proliferación celular más lenta. El tratamiento de las células con el medio acondicionado (medio de cultivo de las células MDA-MB-435 normales) fue capaz de recuperar todos estos cambios fenotípicos producidos por la supresión de la secreción de la proteína hsOAF en cierta medida. Aunque los autores de la invención no se inclinan a por un mecanismo específico, se piensa que la proteína hsOAF secretada es esencial para la invasividad y proliferación de las células MDA-MB-435. Sin embargo, la supresión de la secreción de la proteína hsOAF de otra línea celular altamente metastásica, la MDA-MB-231, utilizando la tecnología basada en oligonucleótidos antisentido no produjo cambios celulares significativos. MDA-MB-435 y MDA-MB-231 son líneas celulares metastásicas bastante diferentes y MDA-MB-435 muestra una secreción de la hsOAF más fuerte que MDA-MB-231.

El gen hsOAF se localiza en la región cromosómica 11q23 en la que frecuentemente tiene lugar una pérdida de heterozigosis en tumores humanos de mama. Negrini, M. y col., *Cancer Res.* 55(14): 3003-7 (1995) y Tomlinson, I.P. y col., *J. Clin. Pathol.* 48(5): 424-8 (1995). La pérdida de heterozigosis en 11q23 en tumores primarios humanos de mama se ha descrito que está asociada con una supervivencia escasa después de la metástasis. Winqvist, R. y col., *Cancer Res.* 55(12): 2660-4 (1995). La región 11q23 también contiene loci tales como ATM (Ataxia-telangiectasia, mutado) y MLL (que frecuentemente está alterado debido a reordenamientos cromosómicos en la leucemia aguda). Rasio, D. y col., *Cancer Res.* 55(24): 6053-6057 (1995) y Rubnitz, J.E. y col., *Leukemia* 10(1): 74-82 (1996). La relación entre una mutación en la región cromosómica 11q23 y la expresión del gen hsOAF en metástasis de cáncer de mama no está aún esclarecida.

La proteína hsOAF secretada puede ser una diana adecuada para el desarrollo de fármacos contra el cáncer de mama y un buen marcador diagnóstico de la malignidad de un tumor de mama. El SEQ ID NO: 1 y los polinucleótidos que comprenden esta secuencia son por lo tanto útiles como marcadores metastásicos. La referencia a secuencias de nucleótidos o de aminoácidos marcadores metastásicos incluye las variantes que tienen patrones de expresión similares en células altamente metastásicas en comparación con los que tienen en células no metastásicas o poco metastásicas. Los polipéptidos marcadores metastásicos pueden diferir en longitud con respecto a las proteínas marcadores metastásicos de longitud completa y contienen 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 ó 271 o más aminoácidos contiguos de una proteína marcador metastásico. Los polinucleótidos ilustrativos incluyen los que codifican los aminoácidos desde aproximadamente el 1 hasta aproximadamente el 273; desde el 1 al 273; desde aproximadamente el 2 hasta aproximadamente el 273; desde el 2 hasta el 273; desde aproximadamente el 26 hasta aproximadamente el 273; y desde el 26 al 273 del SEQ ID NO: 2.

También se pueden existir variantes de proteínas y polipéptidos marcadores. Las variantes de proteínas o polipéptidos marcadores metastásicos pueden estar o no presentes en la naturaleza. Las variantes de proteínas o polipéptidos marcadores metastásicos, presentes en la naturaleza, se encuentran en seres humanos o en otras especies y comprenden secuencias de aminoácidos que son sustancialmente idénticas a una proteína codificada por un gen que corresponde a la secuencia de nucleótidos mostrada en el SEQ ID NO: 1 o su complemento. Las variantes de proteínas o poli-

péptidos marcadores metastásicos, no presentes en la naturaleza, que conservan sustancialmente los mismos patrones de expresión diferencial en células de cáncer de mama altamente metastásicas en comparación con los de células de cáncer de mama poco metastásicas o no metastásicas, al igual que las proteínas o polipéptidos marcadores metastásicos presentes en la naturaleza, también están aquí incluidas. Preferiblemente, las variantes de proteínas o polipéptidos marcadores metastásicos, presentes o no presentes en la naturaleza, tienen secuencias de aminoácidos que son idénticas en al menos el 85%, 90%, 91%, 92%, 93%, 94% o 95% con respecto a las secuencias de aminoácidos codificadas por la secuencia de nucleótidos mostrada en el SEQ ID NO: 1. Más preferiblemente, las moléculas son idénticas en al menos el 96%, 97%, 98% o 99%. El porcentaje de identidad de las secuencias entre una proteína o polipéptido de tipo salvaje y una variante, se determina alineando la proteína o el polipéptido de tipo salvaje con la variante, para obtener el mayor número de coincidencias en aminoácidos, según se conoce en la técnica, contando el número de coincidencias en aminoácidos entre el tipo salvaje y la variante, y dividiendo el número total de coincidencias, entre el número de residuos de aminoácidos de la secuencia de tipo salvaje.

Preferiblemente, los cambios de aminoácidos en las variantes de proteínas o polipéptidos marcadores metastásicos son cambios conservativos de aminoácidos, es decir, sustituciones de aminoácidos que tienen una carga similar o de aminoácidos sin carga. Un cambio conservativo de aminoácidos implica la sustitución de un aminoácido de una familia de aminoácidos que están relacionados por sus cadenas laterales. Los aminoácidos presentes en la naturaleza, de modo general se dividen en cuatro familias: aminoácidos ácidos (aspartato, glutamato), básicos (lisina, arginina, histidina), non-polares (alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano) y aminoácidos polares sin carga (glicina, asparagina, glutamina, cistina, serina, treonina, tirosina). La fenilalanina, el triptófano y la tirosina a veces se clasifican de manera conjunta como aminoácidos aromáticos.

Es razonable esperar que una sustitución aislada de una leucina por una isoleucina o valine, de un aspartato por un glutamato, de una treonina por una serina, o una sustitución similar de un aminoácido por otro aminoácido estructuralmente relacionado, no tendrá un efecto importante sobre las propiedades biológicas de la variante de la proteína o del polipéptido marcador metastásico resultante. Las propiedades y funciones de las variantes de proteínas o polipéptidos marcadores metastásicos son del mismo tipo que las de una proteína o polipéptido marcador metastásico que comprende secuencias de aminoácidos codificadas por la secuencia de nucleótidos mostrada en el SEQ ID NO: 1, aunque las propiedades y funciones de las variantes pueden diferir en el grado. Se puede determinar fácilmente si un cambio de aminoácido da como resultado una variante de una proteína o un polipéptido marcador metastásico con el patrón de expresión diferencial adecuado. Por ejemplo, se pueden seleccionar sondas nucleotídicas entre las secuencias de genes marcadores aquí descritas, y se pueden usar para detectar el mRNA del gen marcador en transferencias Northern o en secciones de tejidos, según se conoce en la técnica. Como alternativa, los anticuerpos que se unen específicamente a los productos proteínicos de genes marcadores metastásicos, se pueden usar para detectar la expresión de proteínas marcadores metastásicos o de sus variantes.

Las variantes de marcadores metastásicos incluyen formas glicosiladas, conjugados agregativos con otras moléculas y conjugados covalentes con restos químicos no relacionados. Las variantes de marcadores metastásicos también incluyen variantes alélicas, variantes de especie y muteínas. Los truncamientos o delecciones de regiones que no afectan a la expresión diferencial de genes marcadores metastásicos, son también variantes de marcadores metastásicos. Las variantes covalentes se pueden preparar uniando funcionalidades a grupos que se encuentran en la cadena de aminoácidos o en el residuo N-terminal o C-terminal, según se conoce en la técnica.

Se reconocerá en la técnica que algunas de las secuencias de aminoácidos del polipéptido de la invención, se puede variar sin que se produzca un efecto significativo sobre la estructura o función de la proteína. Si se contemplan esas diferencias en secuencia, se debe recordar que en la proteína existen zonas críticas que determinan su actividad. En general, es posible sustituir residuos que forman la estructura terciaria, con la condición de que se usen residuos que desempeñen una función similar. En otros casos, el tipo de residuo puede no tener en absoluto importancia si el cambio tiene lugar en una región no crítica de la proteína. La sustitución de aminoácidos puede también cambiar la selectividad de la unión a receptores de superficie celular. Ostade y col., *Nature* 361: 266-268 (1993) describen determinadas mutaciones que producen la unión selectiva del TNF-alfa a sólo uno de los dos tipos conocidos de receptores para TNF. Así, los polipéptidos de la presente invención pueden incluir una o más sustituciones, delecciones o adiciones de aminoácidos, ya sean producidas por mutaciones naturales o por manipulación humana.

También se describen variaciones del polipéptido descrito que muestran patrones de expresión comparables o que incluyen regiones antigénicas. Esos mutantes incluyen delecciones, inserciones, inversiones, repeticiones y sustituciones típicas. La orientación con respecto a los cambios de aminoácidos que es probable que sean fenotípicamente silenciosos se puede encontrar en Bowie, J.U. y col., "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions" *Science* 247: 1306-1310 (1990).

Tienen particular interés las sustituciones de aminoácidos con carga por otros aminoácidos con carga y por aminoácidos neutros o con carga negativa. Estos últimos producen proteínas con una menor carga positiva para mejorar las características de la proteína descrita. También es muy aconsejable la prevención una agregación. La agregación de proteínas no sólo produce una pérdida de actividad sino que también puede ser problemática cuando se preparan formulaciones farmacéuticas, debido a que pueden ser inmunogénicas. (Pinckard y col., *Clin. Exp. Immunol.* 2: 331-340 (1967); Robbins y col., *Diabetes* 36: 838-845 (1987); Cleland y col., *Crit. Rev. Therapeutic Drug Carrier Systems* 10: 307-377 (1993)).

Los aminoácidos de los polipéptidos descritos que son esenciales para la función, se pueden identificar mediante métodos conocidos en la técnica, tales como la mutagénesis dirigida a sitios específicos o la mutagénesis por barrido de alaninas (Cunningham y Wells, *Science* 244: 1081-1085 (1989)). Este último procedimiento introduce mutaciones de una alanina en cada uno de los residuos de la molécula. Las moléculas mutantes resultantes se ensayan luego con respecto a su actividad biológica, tal como con respecto a la unión a receptores, o con respecto a su actividad proliferativa *in vitro*. Los sitios que son críticos para la unión ligando-receptor también se pueden determinar mediante análisis estructural, tal como cristalización, resonancia magnética nuclear o marcaje por fotoafinidad (Smith y col., *J. Mol. Biol.* 224: 899-904 (1992) y de Vos y col. *Science* 255: 306-312 (1992)).

Según se indica, los cambios son preferiblemente de naturaleza menor, tales como sustituciones conservativas de aminoácidos que no afectan significativamente al plegamiento o a la actividad de la proteína. Por supuesto, el número de sustituciones de aminoácidos que un experto en la técnica puede realizar depende de muchos factores, que incluyen los anteriormente descritos. Hablando en términos generales, el número de sustituciones en el caso de un polipéptido determinado no será de más de 50, 40, 30, 25, 20, 15, 10, 5 ó 3.

Las proteínas marcadores metastásicos de longitud completa se pueden extraer, usando métodos bioquímicos estándar, de células humanas productoras de proteínas marcadores metastásicos, tales como las células de cáncer de mama metastásico. Una proteína o un polipéptido marcador metastásico, aislado y purificado, está separado de otros compuestos que normalmente están asociados con una proteína o un polipéptido marcador metastásico en una célula, tales como determinadas proteínas, carbohidratos, lípidos u orgánulos subcelulares. Una preparación de proteínas o polipéptidos marcadores metastásicos aislados y purificados es pura al menos en el 80%; preferiblemente, las preparaciones son puras en el 90%, 95% o 99%.

Un gen humano que codifica el SEQ ID NO: 2 se puede identificar y aislar usando métodos conocidos en la técnica. Según uno de esos métodos, el SEQ ID NO: 1 se prepara en un formato legible por ordenador. La secuencia se compara con secuencias polinucleotídicas de un genoma humano, y se identifican una o más secuencias genómicas humanas que presentan una identidad de la secuencia de la menos el 95% con el SEQ ID NO: 1, por ejemplo, usando el algoritmo de Smith-Waterman que usa como parámetros una búsqueda por afinidad de huecos con una penalización por apertura de hueco de 12 y una penalización por extensión del hueco de 1. Se preparan sondas basadas en las regiones de homología entre el SEQ ID NO: 1 y las secuencias genómicas humanas y se usan para aislar polinucleótidos procedentes de DNA genómico humano, usando métodos conocidos en la técnica. A fecha de presentación del presente documento, no se ha identificado ningún polinucleótido humano que corresponda al polinucleótido completo del SEQ ID NO: 1 en bases de datos de acceso público. Por lo tanto, la invención incluye un DNA genómico humano que comprende la región codificadora del SEQ ID NO: 1 y las regiones que no se traducen que no comparten homología con el SEQ ID NO: 1 pero que están contiguas a las regiones homólogas. Ese DNA genómico incluye, pero no se limita a, intrones, promotores y otras regiones reguladoras funcionalmente asociadas con un gen humano que contiene una región que codifica el SEQ ID NO: 2.

Las proteínas y polipéptidos marcadores metastásicos también se pueden producir mediante métodos de DNA recombinante o mediante métodos de síntesis química. En el caso de la producción de proteínas y polipéptidos marcadores metastásicos recombinantes, las secuencias codificadoras seleccionadas entre las secuencias de nucleótidos mostradas en el SEQ ID NO: 1, o las variantes de las secuencias que codifican proteínas marcadores metastásicos, se pueden expresar en sistemas de expresión procariotas o eucariotas (véase más adelante). Se pueden usar sistemas de expresión de bacterias, levaduras, insectos o mamíferos, según se conoce en la técnica.

Como alternativa, se pueden usar métodos de síntesis química, tales como la síntesis de péptidos en fase sólida, para sintetizar una proteína o un polipéptido marcador metastásico. Medios generales para la producción de péptidos, análogos o derivados, se describen en *CHEMISTRY AND BIOCHEMISTRY OF AMINO ACIDS, PEPTIDES, AND PROTEINS-A SURVEY OF RECENT DEVELOPMENTS*, compilado por Weinstein, B., Marcell Dekker, Inc., publ., Nueva York (1983). Además, se puede llevar a cabo una sustitución del L-estereoisómero normal por D-aminoácidos para aumentar la semivida de la molécula. Las variantes de marcadores metastásicos se pueden producir de manera similar.

También se pueden construir proteínas de fusión no presentes en la naturaleza 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 ó 271 o más aminoácidos contiguos de un marcador metastásico. Las proteínas humanas de fusión con marcadores metastásicos son útiles para generar anticuerpos contra secuencias de aminoácidos de marcadores metastásicos y para usarlas en diversos sistemas analíticos. Por ejemplo, se pueden usar proteínas de fusión con marcadores metastásicos para identificar proteínas que interaccionan con proteínas de marcadores metastásicos e influir en sus funciones. Los métodos físicos tales como la cromatografía por afinidad a proteínas, o ensayos basados en bibliotecas con respecto a interacciones proteína-proteína, tales como el sistema del doble híbrido de levaduras o los sistemas de presentación en fagos, también se pueden usar para este fin. Estos métodos son muy conocidos en la técnica y se pueden usar como métodos de escrutinio de fármacos.

Una proteína de fusión con un marcador metastásico comprende dos segmentos de proteínas fusionados entre sí por medio de un enlace peptídico. El primer segmento de proteína 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 ó 271 o más aminoácidos contiguos de una proteína marcador metastásico. Los aminoácidos se pueden seleccionar entre las

## ES 2 311 016 T3

secuencias de aminoácidos codificadas por la secuencia de nucleótidos mostrada en el SEQ ID NO: 1 o entre variantes de esa secuencia, tales como las anteriormente descritas. El primer segmento de proteína también puede comprender una proteína marcador metastásico de longitud completa.

5 En una de las realizaciones preferidas, el primer segmento de proteína comprende el polipéptido mostrado en el SEQ ID NO: 2. En una variación de esta realización, el primer segmento de proteína consiste en los aminoácidos 31-287 del SEQ ID NO: 2. Esta proteína de fusión carece del péptido señal del SEQ ID NO: 2 y podría ser adecuada para retener la proteína de fusión expresada dentro de la célula.

10 El segundo segmento de proteína puede ser una proteína de longitud completa o un fragmento de proteína o un polipéptido que no se encuentra adyacente al primer segmento de proteína en la proteína natural codificada por el SEQ ID NO: 1. La proteína de fusión puede estar marcada con un marcador detectable, según se conoce en la técnica, tal como un marcador radiactivo, fluorescente, quimioluminiscente o biotinilado. El segundo segmento de proteína puede ser una enzima que generará un producto detectable, tal como la  $\beta$ -galactosidasa. El primer segmento de proteína  
15 puede ser N-terminal o C-terminal, según convenga.

También son muy conocidas técnicas para fabricar proteínas de fusión, ya sea mediante métodos recombinantes o uniendo covalentemente dos segmentos de proteínas. Se pueden usar métodos de DNA recombinante para preparar proteínas de fusión con marcadores metastásicos, por ejemplo, fabricando una construcción de DNA que comprende  
20 secuencias codificadoras del SEQ ID NO: 1, en un marco de lectura correcto, junto con nucleótidos que codifican el segundo segmento de proteína, y expresando esa construcción de DNA en una célula hospedadora, según se describe más adelante. El marco de lectura abierto del SEQ ID NO: 1 se muestra en la Figura 4.

Las proteínas, polipéptidos, variantes o proteínas de fusión de marcadores metastásicos, aislados y purificados, se  
25 pueden usar como inmunógenos para obtener preparaciones de anticuerpos que se unen específicamente a una proteína marcador metastásico. Los anticuerpos se pueden usar, *inter alia*, para detectar proteínas marcadores metastásicos de tipo salvaje en tejidos humanos y sus fracciones. Los anticuerpos también se pueden usar para detectar la presencia de mutaciones en genes de marcadores metastásicos que dan como resultado una infraexpresión o una sobreexpresión de una proteína marcador metastásico, o que dan como resultado la expresión de una proteína marcador metastásico que  
30 tiene un tamaño o con una movilidad electroforética alterados.

Se pueden fabricar preparaciones de anticuerpos policlonales o monoclonales usando métodos estándar. También se pueden preparar anticuerpos monocatenarios. Un inmunógeno preferido es un polipéptido que comprende el SEQ ID NO: 2. Anticuerpos monocatenarios que se unen específicamente a proteínas, polipéptidos, variantes o proteínas  
35 de fusión marcadores metastásicos se pueden aislar, por ejemplo, a partir de bibliotecas de presentación de inmunoglobulinas monocatenarias, según se conoce en la técnica. La biblioteca se expone usando la técnica de tapizado de una superficie (del inglés, "*panned*") frente a secuencias de aminoácidos de proteínas marcadores metastásicos del SEQ ID NO: 2, y se pueden aislar varios anticuerpos monocatenarios que se unen con una elevada afinidad a diferentes epítopos de las proteínas marcadores metastásicos. Hayashi y col., 1995, *Gene* 160: 129-130. También se pueden  
40 construir anticuerpos monocatenarios usando un método de amplificación del DNA, tal como la reacción en cadena de la polimerasa (PCR), usando como molde cDNA de hibridomas. Thirion y col., 1996, *Eur. J. Cancer Prev.* 5: 507-511.

Los anticuerpos específicos de marcadores metastásicos se unen específicamente a epítopos presentes en una pro-  
45 teína marcador metastásico de longitud completa, que tiene una secuencia de aminoácidos codificada por una secuencia de nucleótidos que se muestra en el SEQ ID NO: 1, a polipéptidos marcadores metastásicos o a variantes de marcadores metastásicos, ya sea solos o como parte de una proteína de fusión. Preferiblemente, los epítopos de marcadores metastásicos no están presentes en otras proteínas humanas. Típicamente, se requieren al menos 6, 8, 10 ó 12 aminoácidos contiguos para formar un epítipo. Sin embargo, los epítopos en los que participan aminoácidos no contiguos  
50 pueden requerir más, p. ej., al menos 15, 25 ó 50 aminoácidos.

Los anticuerpos que se unen específicamente a proteínas, polipéptidos, proteínas de fusión o variantes marcadores metastásicos, proporcionan una señal de detección al menos 5, 10 ó 20 veces más alta que una señal de detección proporcionada por otras proteínas, cuando se usan en transferencias Western o en otros ensayos inmunoquímicos.  
55 Preferiblemente, 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 marcador metastásico en el seno de una disolución.

Los polinucleótidos subgenómicos contienen menos que un cromosoma completo. Preferiblemente, esos poli-  
60 nucleótidos están libres de intrones. En una realización preferida, las moléculas polinucleotídicas 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 ó 2350 nucleótidos del SEQ ID NO: 1 o sus complementos. El complemento de una secuencia de nucleótidos mostrada en el SEQ ID NO: 1 es una secuencia de nucleótidos contiguos que forma pares de bases según Watson-Crick con una secuencia de nucleótidos contiguos  
65 mostrada en el SEQ ID NO: 1.

Las secuencias de nucleótidos degenerados que codifican secuencias de aminoácidos de proteínas o variantes marcadores metastásicos, así como las secuencias de nucleótidos homólogas que comprenden un polinucleótido idéntico



al menos en el 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% con la región codificadora de la secuencia de nucleótidos mostrada en el SEQ ID NO: 1, son también polinucleótidos subgenómicos marcadores metastásicos. Típicamente, las secuencias homólogas de polinucleótidos subgenómicos marcadores metastásicos pueden ser confirmadas mediante hibridación en condiciones rigurosas, según se conoce en la técnica. El porcentaje de identidad de las secuencias entre la secuencia de tipo salvaje y secuencias variantes homólogas se determina alineando el polinucleótido de tipo salvaje con la variante para obtener el mayor número de coincidencias entre los nucleótidos, según se conoce en la técnica, contando el número de coincidencias de nucleótidos entre el tipo salvaje y la variante, y dividiendo el número total de coincidencias entre el número total de nucleótidos de la secuencia de tipo salvaje. Un algoritmo preferido para calcular el porcentaje de identidad es el algoritmo de búsqueda de homologías de Smith-Waterman según se implementa en el programa MPSRCH (*Oxford Molecular*) que usa una búsqueda por afinidad de huecos con los siguientes parámetros de búsqueda: una penalización por apertura de hueco de 12 y una penalización por extensión del hueco de 1.

Un polinucleótido subgenómico marcador metastásico que comprende secuencias codificadoras de proteínas marcadores metastásicos se puede usar en una construcción de expresión. Preferiblemente, el polinucleótido subgenómico marcador metastásico se inserta en un plásmido de expresión (por ejemplo, el sistema Ecdyson, pIND, *In vitro Gene*). Los polinucleótidos subgenómicos marcadores metastásicos se pueden propagar en vectores y en líneas celulares usando métodos muy conocidos en la técnica. Los polinucleótidos subgenómicos marcadores metastásicos pueden estar en moléculas lineales o circulares. Pueden estar en moléculas de replicación autónoma o en moléculas que carecen de secuencias de replicación. Pueden estar regulados por sus propias secuencias reguladoras o por otras secuencias reguladoras, según se conoce en la técnica.

Una célula hospedadora que comprende una construcción de expresión de un marcador metastásico se puede usar luego la totalidad o una porción de una proteína marcador metastásico. Las células hospedadoras que comprenden construcciones de expresión de marcadores metastásicos pueden ser procariotas o eucariotas. Existen disponibles diversas células hospedadoras para usar en sistemas de expresión en bacterias, levaduras, insectos y seres humanos, y se pueden usar para expresar o para propagar construcciones de expresión de marcadores metastásicos (véase más adelante). Las construcciones de expresión se pueden introducir en las células hospedadoras usando cualquiera de los métodos conocidos en la técnica. Estos métodos incluyen la transferencia de DNA mediada por policones y transferrina, transfección con ácidos nucleicos desnudos o encapsulados, fusión celular mediada por liposomas, transporte intracelular de bolitas de látex revestidas de DNA, fusión de protoplastos, infección vírica, electroporación y transfección mediada por fosfato cálcico.

Una construcción de expresión de un marcador metastásico comprende un promotor que es funcional en una célula hospedadora elegida. El experto en la técnica puede seleccionar fácilmente un promotor adecuado de entre un gran número de promotores específicos del tipo de célula, conocidos y usados en la técnica. La construcción de expresión puede contener también un terminador de la transcripción que es funcional en la célula hospedadora. La construcción de expresión comprende un segmento polinucleotídico que codifica la totalidad o una porción de la proteína, variante, proteína de fusión, anticuerpo o ribozima marcador metastásico. El segmento polinucleotídico está situado aguas abajo 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, encargadas de su replicación autónoma.

Los sistemas bacterianos adecuados para expresar construcciones de expresión de marcadores metastásicos incluyen los descritos en Chang y col., *Nature* (1978) 275: 615, Goeddel y col., *Nature* (1979) 281: 544, Goeddel y col., *Nucleic Acids Res.* (1980) 8: 4057, documento EP 36.776, Patente de EE. UU. núm. 4.551.433, deBoer y col., *Proc. Natl. Acad. Sci. USA* (1983) 80: 21-25 y Siebenlist y col., *Cell* (1980) 20: 269.

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

La expresión de construcciones de expresión de marcadores metastásicos en insectos se puede llevar a cabo según se describe en la Patente de EE. UU. núm. 4.745.051, Friesen y col. (1986) "The Regulation of Baculovirus Gene Expression" en: *THE MOLECULAR BIOLOGY OF BACULOVIRUSES* (W. Doerfler, compilador), documento EP 127.839, documento EP 155.476 y Vlak y col., *J. Gen. Virol.* (1988) 69: 765-776, Miller y col., *Ann. Rev. Microbiol.* (1988) 42: 177, Carbonell y col., *Gene* (1988) 73: 409, Maeda y col., *Nature* (1985) 315: 592-594, Lebacqz-Verheyden y col., *Mol. Cell. Biol.* (1988) 8: 3129; Smith y col., *Proc. Natl. Acad. Sci. USA* (1985) 82: 8404, Miyajima y col., *Gene* (1987) 58: 273; y Martin y col., *DNA* (1988) 7: 99. Numerosas cepas y variantes de baculovirus y las correspondientes células hospedadoras permisivas de insectos, procedentes de hospedadores, se describen en Luckow y col., *Bio/Technology* (1988) 6: 47-55, Miller y col., en *GENETIC ENGINEERING* (Setlow, J.K. y col. compiladores), Vol. 8 (Plenum Publishing, 1986), págs. 277-279 y Maeda y col., *Nature*, (1985) 315:592-594.

La expresión en mamíferos de construcciones de expresión de marcadores metastásicos se puede llevar a cabo según se describe en Dijkema y col., *EMBO J.* (1985) 4: 761, Gorman y col., *Proc. Natl. Acad. Sci. USA* (1982b) 79: 6777, Boshart y col., *Cell* (1985) 41: 521 y Patente de EE. UU. núm. 4.399.216. Otras características de la expresión en mamíferos de construcciones de expresión de marcadores metastásicos se pueden facilitar según se describe en  
 5 Ham y Wallace, *Meth. Enz.* (1979) 58: 44, Barnes y Sato, *Anal. Biochem.* (1980) 102: 255, Patente de EE. UU. núm. 4.767.704, Patente de EE. UU. núm. 4.657.866, Patente de EE. UU. núm. 4.927.762, Patente de EE. UU. núm. 4.560.655, documento WO 90/103430, documento WO 87/00195 y documento U.S. RE 30.985.

Los polinucleótidos subgenómicos de la presente invención también se pueden usar en vehículos de suministro de  
 10 genes, con el fin de suministrar un mRNA o un oligonucleótido marcador metastásico (ya sea junto con la secuencia del mRNA marcador metastásico natural o con su complemento), una proteína marcador metastásico de longitud completa, una proteína de fusión marcador metastásico, un polipéptido marcador metastásico, o una ribozima específica de un marcador metastásico, o un anticuerpo monocatenario, a una célula, preferiblemente a una célula eucariota. Según la presente invención, un vehículo de suministro de genes puede ser, por ejemplo, un DNA plasmídico desnudo, un  
 15 vector vírico de expresión que comprende un polinucleótido subgenómico marcador metastásico, o un polinucleótido subgenómico marcador metastásico juntamente con un liposoma o un agente condensante.

También se describe un método para detectar la expresión de genes marcadores metastásicos en una muestra biológica. La detección de la expresión de genes marcadores metastásicos es útil, por ejemplo, para identificar metástasis  
 20 o para determinar el potencial metastásico de una muestra de tejido, preferiblemente un tumor. Se pueden diseñar, por lo tanto, regímenes de tratamiento adecuados para pacientes en riesgo de presentar cánceres metastásicos en otros órganos del cuerpo.

La muestra corporal puede ser, por ejemplo, un tejido sólido o una muestra de fluido. El polipéptido natural codificado por el SEQ ID NO: 1 es una proteína secretada putativa, y es probable que se detecte en fluidos corporales que incluyen sangre y fluido linfático, en particular los que vacían desde sitios tumorales del organismo. Los productos de expresión de proteínas o ácidos nucleicos pueden ser detectados en la muestra corporal. En una de las realizaciones, la muestra corporal se ensaya con respecto a la presencia de una proteína marcador metastásico. Una proteína marcador metastásico comprende una secuencia codificada por una secuencia de nucleótidos mostrada en el SEQ ID NO: 1, o  
 30 su complemento, y se puede detectar usando los anticuerpos específicos de proteína marcador de la presente invención. Los anticuerpos pueden estar marcados, por ejemplo, con una secuencia identificadora radioactiva, fluorescente, biotinilada o enzimática, y se pueden detectar directamente, o se pueden detectar usando métodos inmunoquímicos indirectos, usando un anticuerpo secundario marcado. La presencia de las proteínas marcadores metastásicos se puede ensayar, por ejemplo, en secciones de tejidos mediante inmunohistoquímica, o en lisados, usando transferencias Western, según se conoce en la técnica. La presencia de la proteína marcador es indicativa de que la muestra de tejido es metastásica.  
 35

En otra realización, la muestra corporal se ensaya con respecto a la presencia del mRNA de una proteína marcador. Una muestra se pone en contacto con una sonda de hibridación a ácidos nucleicos, capaz de hibridarse con el mRNA correspondiente al polipéptido seleccionado. Y aún más, la muestra se puede someter a una técnica de transferencia Northern para detectar mRNA, que indica la expresión del polipéptido. En el caso de las técnicas en las que se detecta mRNA, la muestra se puede someter a un procedimiento de amplificación de ácidos nucleicos en el que la molécula de mRNA, o una de sus partes seleccionadas, se amplifica usando cebadores nucleotídicos adecuados. También se pueden usar otras técnicas de detección de mRNA, que incluyen, pero no se limitan a, la hibridación *in situ*.  
 40

Se pueden generar sondas específicas de proteínas marcadores usando la secuencia de DNA descrita en el SEQ ID NO: 1. Las sondas tienen preferiblemente una longitud de al menos 15 a 50 nucleótidos, aunque pueden tener una longitud de al menos 8, 10, 11, 12, 20, 25, 30, 35, 40, 45, 60, 75 ó 100 o más nucleótidos. Una región preferible para seleccionar sondas es la incluida en las posiciones nucleotídicas 446-1173 del SEQ ID NO: 1. Las sondas se  
 50 pueden sintetizar por métodos químicos o se pueden generar a partir de polinucleótidos más largos usando enzimas de restricción. Las sondas se pueden marcar, por ejemplo, con una secuencia identificadora radiactiva, biotinilada o fluorescente.

Opcionalmente, se puede cuantificar el nivel de un producto particular de expresión de un marcador metastásico en una muestra corporal. La cuantificación se puede realizar, 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 usando una técnica tal como la densitometría, con o sin asistencia computerizada. Para usar como controles, se pueden aislar muestras corporales procedentes de otros seres humanos, de otros órganos no cancerosos del paciente sometido al ensayo, o de cáncer de mama no metastásico procedente del paciente sometido  
 60 al ensayo. Según se indica en los presentes resultados, la expresión del SEQ ID NO: 1 en células de cáncer de mama poco metastásicas 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 a ensayar es al menos 2 veces mayor que en una muestra control adecuada, esto es indicativo de células metastásicas.

Los polinucleótidos que codifican reactivos de la invención, específicos para marcadores metastásicos, tales como anticuerpos y sondas nucleotídicas, se pueden suministrar en un kit de detección de productos de expresión de genes marcadores en una muestra biológica. El kit también puede contener tampones y componentes de marcaje, así como instrucciones para el uso de los reactivos para detectar los productos de expresión marcadores en la muestra biológica.  
 65

La expresión de un gen marcador metastásico se puede alterar usando una secuencia de oligonucleótido antisentido. La secuencia antisentido es complementaria a al menos una porción de la secuencia codificadora (nucleótidos 365-1173) de un gen marcador metastásico que tiene una secuencia de nucleótidos mostrada en el SEQ ID NO: 1. Preferiblemente, la secuencia de oligonucleótido antisentido tiene una longitud de al menos seis nucleótidos, pero puede tener una longitud de al menos aproximadamente 8, 12, 15, 20, 25, 30, 35, 40, 45 ó 50 nucleótidos. También se pueden usar secuencias más largas. Las moléculas de oligonucleótidos antisentido se pueden proporcionar en una construcción de DNA y se puede introducir en células cuya división se ha de disminuir. Esas células incluyen células de cáncer de mama altamente metastásicas.

Los oligonucleótidos antisentido pueden comprender desoxirribonucleótidos, ribonucleótidos o una combinación de ambos. Los oligonucleótidos se pueden sintetizar manualmente o mediante un sintetizador automático, uniendo covalentemente el extremo 5' de uno de los nucleótidos con el extremo 3' de otro nucleótido, con uniones internucleotídicas no de tipo fosfodiéster tales como alquilfosfonatos, fosforotioatos, fosforoditioatos, alquilfosfonotioatos, alquilfosfonatos, fosforamidatos, ésteres fosfato, carbamatos, acetamidato, ésteres de carboximetilo, carbonatos y triésteres fosfato. Véase, Brown, 1994, *Meth. Mol. Biol.* 20: 1-8; Sonveaux, 1994, *Meth. Mol. Biol.* 26: 1-72; Uhlmann y col., 1990, *Chem. Rev.* 90: 543-583.

Los anticuerpos aquí descritos, que se unen específicamente a una proteína marcador metastásico, también se pueden usar para alterar la expresión de un gen marcador metastásico. Por anticuerpos se entiende anticuerpos y sus partes o derivados, tales como anticuerpos monocatenarios, que conservan la especificidad de unión a la proteína. Los anticuerpos específicos se unen a proteínas marcadores metastásicos y evitan que las proteínas actúen en la célula. Los polinucleótidos que codifican los anticuerpos específicos de la invención se pueden introducir en las células, según se ha descrito en lo que antecede.

Las proteínas marcadores aquí descritas se pueden usar en el escrutinio de fármacos que presentan un efecto terapéutico antimetastásico. El efecto de un compuesto a ensayar sobre la síntesis de proteínas marcadores metastásicos se puede usar también para identificar compuestos a ensayar que modulan la metástasis. Los compuestos a ensayar que se pueden someter a escrutinio incluyen cualquiera de las sustancias, ya sean productos naturales o sustancias sintéticas, que puedan ser administradas al sujeto. Se pueden realizar ensayos en bibliotecas o en mezclas de compuestos. Los compuestos o las sustancias pueden ser aquellos para los que se conoce o desconoce un efecto farmacéutico con anterioridad.

La síntesis de proteínas marcadores metastásicos se puede medir mediante cualquiera de los métodos de medida de síntesis de proteínas conocidos en la técnica, tales como la incorporación de aminoácidos marcados a proteínas y la detección de proteínas marcadores metastásicos marcadas en un gel de poliacrilamida. La cantidad de proteínas marcadores metastásicos se puede detectar, por ejemplo, usando los anticuerpos específicos para proteínas marcadores metastásicos de la invención en transferencias Western. La cantidad de proteínas marcadores metastásicos, sintetizadas en presencia o ausencia de un compuesto a ensayar se puede determinar mediante cualquiera de los medios conocidos en la técnica, tales como la comparación de la cantidad de proteína marcador metastásico sintetizada con la cantidad de las proteínas marcador metastásico presentes en una curva estándar.

El efecto de un compuesto a ensayar sobre la síntesis de proteínas marcadores metastásicos también se puede medir mediante análisis de transferencia Northern, midiendo la cantidad de expresión del mRNA de las proteínas marcadores metastásicos en respuesta al compuesto a ensayar, usando sondas nucleotídicas específicas de proteínas marcadores metastásicos de la invención, según se conoce en la técnica.

Típicamente, una muestra biológica se pone en contacto con un intervalo de concentraciones del compuesto a ensayar, 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. Preferiblemente, el compuesto a ensayar aumenta o disminuye la expresión de una proteína marcador metastásico en el 60%, 75% u 80%. Más preferiblemente, se alcanza un aumento o una disminución del 85%, 90%, 95% o 98%.

Se describen además composiciones para aumentar o disminuir la expresión de proteínas marcadores metastásicos. Estas composiciones comprenden polinucleótidos que codifican la totalidad o al menos una porción de un producto de expresión de genes de proteínas marcadores metastásicos. Preferiblemente, la composición terapéutica contiene una construcción de expresión que comprende un promotor y un segmento polinucleotídico que codifica al menos una porción de la proteína marcador metastásico que es efectiva para disminuir el potencial metastásico. Las porciones de genes o proteínas marcadores metastásicos que son efectivas para disminuir el potencial metastásico de una célula, se pueden determinar, por ejemplo, introduciendo porciones de genes o polipéptidos marcadores metastásicos en líneas de células metastásicas, tales como MDA-MB-231, MDA-MB-435, Km12C o Km12L4, y realizando ensayos para determinar la velocidad de división de las células o el potencial de las células para formar metástasis cuando se implantan *in vivo*, según se conoce en la técnica. Se pueden usar líneas de células no metastásicas, tales como la MCF-7, para realizar ensayos con respecto a la capacidad de una porción de una proteína marcador metastásico para aumentar la expresión de un gen marcador metastásico.

Típicamente, una composición terapéutica de un marcador metastásico se prepara en forma de un inyectable, ya sea en forma de una disolución líquida o de una suspensión; sin embargo, también se pueden preparar formas sólidas adecuadas para su disolución, o para su suspensión, en vehículos líquidos antes de su inyección. Una composición de marcador metastásico también se puede formular en forma de un comprimido con recubrimiento entérico o de una

cápsula de gel, según los métodos conocidos en la técnica, tales como los descritos en la Patente de EE. UU. núm. 4.853.230, documento EP 225.189, documento AU 9.224.296 y documento AU 9.230.801.

La administración de los agentes terapéuticos marcadores metastásicos puede incluir la administración local o sistémica, incluyendo inyecciones, administración oral, pistola de partículas, o administración cateterizada, y administración tópica. Se pueden usar diversos métodos para administrar una composición terapéutica de marcador metastásico directamente en un sitio específico del cuerpo.

Para el tratamiento de tumores, incluyendo las lesiones metastásicas, por ejemplo, una composición terapéutica de marcador metastásico se puede inyectar varias veces en varios lugares diferentes del interior de la parte principal del tumor. Como alternativa, se pueden identificar las arterias que dan servicio a un tumor y una composición terapéutica se puede inyectar en una arteria de esas con el fin de suministrar directamente la composición al interior del tumor.

Un tumor que tiene un centro necrótico puede ser aspirado y la composición se puede inyectar directamente en el interior del centro, ahora vacío, del tumor. Una composición terapéutica de marcador metastásico se puede administrar directamente en la superficie de un tumor, por ejemplo, mediante aplicación tópica de la composición. La formación de imágenes por rayos X se puede usar para ayudar en algunos de los métodos de suministro anteriormente mencionados. Una combinación de agentes terapéuticos, que incluyen proteínas o polipéptidos marcadores metastásicos o un polinucleótido subgenómico marcador metastásico y otros agentes terapéuticos, se pueden administrar de manera simultánea o secuencial.

Como alternativa, una composición terapéutica de marcador metastásico se puede introducir en células humanas *ex vivo*, y las células se devuelven luego al ser humano. Las células se pueden extraer de diversos lugares que incluyen, por ejemplo, un tumor seleccionado o un órgano afectado. Además, una composición terapéutica se puede introducir en células no afectadas, por ejemplo, en fibroblastos dérmicos o en leucocitos de sangre periférica. Si se desea, de la sangre también se pueden extraer específicamente fracciones particulares de células, tales como un subgrupo de células T o células madre (véase, por ejemplo, el documento PCT WO 91/16116). Las células extraídas se pueden luego poner en contacto con una composición terapéutica de marcador metastásico utilizando cualquiera de las técnicas anteriormente descritas, seguido del retorno de las células al ser humano, preferiblemente a, o en, las inmediaciones de un tumor o en otro sitio a tratar. Los métodos anteriormente descritos pueden comprender adicionalmente las etapas de producir una merma de fibroblastos o de otras células tumorales no contaminantes, después de extraer a un ser humano las células tumorales, y/o de la etapa de inactivar las células, por ejemplo, mediante irradiación.

Tanto la dosis de una composición de marcador metastásico como los medios de administración se pueden determinar sobre la base de las cualidades específicas de la composición terapéutica, el estado, la edad y el peso del paciente, la progresión de la enfermedad y otros factores relevantes. Preferiblemente, una composición terapéutica de la invención disminuye la expresión de los genes marcadores metastásicos en el 50%, 60%, 70% u 80%. Más preferiblemente, la expresión de los genes marcadores metastásicos se hace disminuir en el 90%, 95%, 99% o 100%. La efectividad del mecanismo elegido para alterar la expresión de los genes marcadores metastásicos se puede determinar usando métodos muy conocidos en la técnica, tales como la hibridación de sondas nucleotídicas al mRNA de los genes marcadores metastásicos, RT-PCR cuantitativa, o detección de las proteínas marcadores metastásicos usando anticuerpos específicos de la invención.

Si la composición contiene las proteínas, polipéptidos o anticuerpos marcadores metastásicos, las dosificaciones efectivas de la composición están en el intervalo desde aproximadamente 5  $\mu\text{g/kg}$  hasta aproximadamente 50  $\mu\text{g/kg}$  de peso corporal del paciente, desde aproximadamente 50  $\mu\text{g/kg}$  hasta aproximadamente 5  $\text{mg/kg}$ , desde aproximadamente 100  $\mu\text{g/kg}$  hasta aproximadamente 500  $\mu\text{g/kg}$  de peso corporal del paciente, y desde aproximadamente 200  $\mu\text{g/kg}$  hasta aproximadamente 250  $\mu\text{g/kg}$ .

Las composiciones terapéuticas que contienen polinucleótidos subgenómicos marcadores metastásicos se pueden administrar en un intervalo desde aproximadamente 100 ng hasta aproximadamente 200 mg of DNA en el caso de una administración local. También se pueden usar intervalos de concentraciones desde aproximadamente 500 ng hasta aproximadamente 50 mg, desde aproximadamente 1  $\mu\text{g}$  hasta aproximadamente 2 mg, desde aproximadamente 5  $\mu\text{g}$  hasta aproximadamente 500  $\mu\text{g}$  y desde aproximadamente 20  $\mu\text{g}$  hasta aproximadamente 100  $\mu\text{g}$  de DNA, durante un protocolo de terapia génica. Factores tales como el método de actuación y la eficacia de la transformación y expresión son consideraciones que afectarán a la dosificación requerida para la eficacia final de los polinucleótidos subgenómicos marcadores metastásicos. Cuando se desea obtener una mayor expresión en una mayor superficie de tejido más grande, se pueden requerir mayores cantidades de polinucleótidos subgenómicos marcadores metastásicos o las mismas cantidades readministradas en protocolos de administración sucesivos, o varias administraciones en porciones de tejido diferentes, adyacentes o próximas, por ejemplo, de un sitio tumoral, para lograr un resultado terapéutico positivo. En todos los casos, una experimentación de rutina en ensayos clínicos determinará los intervalos para un efecto terapéutico óptimo.

La expresión de un gen marcador metastásico endógeno en una células también se puede alterar introduciendo en el marco de lectura junto con el gen marcador metastásico endógeno, una construcción de DNA que comprende una secuencia de direccionamiento, una secuencia reguladora, un exón y un sitio dador de corte y empalme desparejado de una proteína marcador metastásico, mediante recombinación homóloga de manera que se forma una célula recombinante por recombinación homóloga que comprende la construcción de DNA. Esta nueva unidad de transcripción se

puede usar para conectar o desconectar el gen marcador metastásico según se desee. Este método para afectar a la expresión de un gen endógeno se ilustra en la Patente de EE. UU. núm. 5.641.670.

Un polinucleótido subgenómico marcador metastásico también puede ser administrado a sujetos con el fin de escrutar compuestos de ensayo con respecto a los compuestos que son útiles para aumentar la transferencia de polinucleótidos subgenómicos marcadores metastásicos a la célula, o para aumentar los subsiguientes efectos biológicos de los polinucleótidos subgenómicos marcadores metastásicos dentro de la célula. Esos efectos biológicos incluyen la hibridación a un mRNA complementario de marcadores metastásicos e inhibición de su traducción, la expresión de un polinucleótido subgenómico marcador metastásico para formar mRNA de un marcador metastásico y/o una proteína de un marcador metastásico, y la replicación e integración de un polinucleótido subgenómico marcador metastásico. El sujeto puede ser un cultivo celular o un animal, preferiblemente un mamífero, más preferiblemente, un ser humano.

La anterior descripción describe de manera general la presente invención. Se puede obtener un conocimiento más completo con referencia a los siguientes ejemplos específicos que aquí se proporcionan con fines de ilustración solamente y que no pretenden limitar el alcance de la invención.

## Ejemplos

### Materiales y métodos

**Cultivo de células.** Las 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 la *Chiron Master Culture Collection*, Chiron Corporation) se cultivaron a 37°C en CO<sub>2</sub> al 5% in DMEM + HAM F-12 (1:1) (Bio\*Whittaker, Walkersville, MD) que contenía L-glutamina 2 mM, piruvato sódico 1 mM, penicilina 100 U/ml y estreptomicina 100 µg/ml (Bio\*Whittaker, Walkersville, MD), Disolución vitamínica 1X, Aminoácidos no esenciales 1X (Irvine Scientific, Santa Ana, CA) y suero bovino fetal inactivado por calor al 10% (Life Technologies, Rockville, MD). Se obtuvieron células COS-7 procedentes de la ATCC y se cultivaron a 37°C en presencia de CO<sub>2</sub> al 5% en DMEM con suero bovino fetal inactivado por calor al 10% (Life Technologies).

**Concentración del sobrenadante en Opti-MEM1.** Los medios de cultivo Opti-MEM1 (Life Technologies) se concentraron a través de columnas Centricon YM-10 y/o Microcon YM-10 (Millipore Corporation, Bedford, MA). Después se les añadió un tampón de carga de muestras para SDS-PAGE y las muestras se sometieron a ebullición.

**Hibridación en transferencias Northern.** Se prepararon los RNA totales procedentes de las líneas celulares de cáncer de mama cultivadas y de tejidos tumorales de ratones SCID transplantados con líneas celulares de cáncer de mama, usando el *RNeasy Maxi Kit* (Qiagen, Valencia, CA). Se cargaron aproximadamente 20 µg de RNA total por pista sobre un gel de formaldehído/agarosa para su electroforesis y luego se transfirieron a una membrana de nilón Hybond-N+ (Amersham Life Science, Little Chalfont, Inglaterra). La transferencia se fijó con irradiación UV. El tampón *Rapid-Hyb* (Amersham Life Science), con una concentración de 5 mg/ml de DNA de esperma, monocatenario y desnaturizado, se precalentó hasta 65°C y la transferencia se pre-hibridó en este tampón, con agitación, a 65°C durante 30 minutos. Se añadió un fragmento del cDNA de hsOAF o un fragmento del cDNA de β-actina como sonda marcada con [α-<sup>32</sup>P]dCTP (3000 Ci/mmol, Amersham Pharmacia Biotech, Inc., Piscataway, NJ) (*Prime-It RmT Kit*, Stratagene, La Jolla, CA) y purificada con una Microcolumna *ProbeQuant™ G-50* (Amersham Pharmacia Biotech, Inc.) y se hibridó con la transferencia, con agitación a 65°C durante la noche. La transferencia se lavó 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 y después se expuso a *Hyperfilms* (Amersham Life Science).

**Inmunotransferencias.** Las muestras de las proteínas se sometieron a electroforesis sobre geles de SDS-PAGE al 10%-20% y después se transfirieron a membranas de PVDF (0,2 µm) mediante 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) que contenía leche desnatada al 10%, después se transfirieron en PBS (pH 7,4) que contenía BSA al 1%, con un suero anti-hsOAF procedente de conejo (1:1000), seguido de una hibridación con una IgG anti-conejo, procedente de cabra, conjugada con fosfatasa alcalina, como anticuerpo secundario, (1:2000) (Santa Cruz Biotechnology, Inc., Santa Cruz, CA). Las bandas de las proteínas se visualizaron luego con el reactivo NBT/BCIP (Boehringer Mannheim, Alemania).

**Transfección transitoria.** La región codificadora (356-1174) del cDNA de hsOAF se clonó en un vector de expresión pRetro-On modificado (Clontech, Palo Alto, CA). El vector pRetro-On que alberga hsOAF o el vector pRetro-On de control con GFP, se usó para transfectar células COS-7 en una placa de cultivo de 100 mm usando *Effectene™ Transfection Reagent Kit* (Qiagen) según las instrucciones del protocolo proporcionado por el fabricante. Las células se recuperaron en DMEM que contenía FBS al 10% durante la noche y seguidamente se cambiaron a Opti-MEM1. Después de dos días más, el sobrenadante se recogió y se concentró para el análisis de transferencia Western.

**Transfección con un oligonucleótido antisentido.** Las células MDA-MB-435 se sembraron en placas de cultivo de 6 pocillos un día antes de la transfección para obtener una densidad del 90% en el momento de la transfección. Un oligonucleótido antisentido 100 µM, o un oligonucleótido inverso 100 µM de control, se diluyó hasta alcanzar una concentración 2 µM en Opti-MEM1 para la transfección. Lipitoid1 esterilizado 0,5 mM se diluyó en una proporción de 1,5 nmoles de lipitoid1 : 1 µg de oligonucleótido, en un volumen igual de Opti-MEM1. El oligonucleótido diluido y el lipitoid1 diluido se mezclaron e inmediatamente se añadieron a las células del medio de cultivo hasta alcanzar

## ES 2 311 016 T3

una concentración final del oligonucleótido 100 nM, 200 nM o 300 nM. Después de 6 h, la mezcla de transfección se reemplazó por medio de cultivo normal y las células se incubaron durante la noche para su recuperación. La secuencia del oligonucleótido antisentido es: AGCTGCGGATGCCACACTTGTAGG (SEQ ID NO: 4) y la secuencia del oligonucleótido inverso de control es: GGATGTTACACCGTAGGCGTCGA (SEQ ID NO: 5).

*Ensayo de invasión en Matrigel.* Las células se trataron con tripsina, se lavaron y se resuspendieron en el medio para su recuento. Se lavaron  $4 \times 10^4$  células y se resuspendieron en 100  $\mu$ l de medio sobre hielo. A las células enfriadas en hielo se añadieron 200  $\mu$ l de Matrigel (*Collaborative Biomedical Products*, Bedford, MA). El Matrigel y las células se mezclaron cuidadosamente y luego se dispensaron en un pocillo de una placa de cultivo de 24 pocillos y se solidificaron a 37°C durante 30 min. La mezcla de Matrigel-células se cubrió 0,5 ml de medio y se incubó a 37°C en presencia de CO<sub>2</sub> al 5% durante 6 días. El medio se reemplazó cada 2 días.

*Ensayo de proliferación.* Las células se trataron con tripsina, se lavaron y se resuspendieron en el medio para su recuento. Las células se transfirieron luego a placas de 96 pocillos (5000 células/pocillo) para su incubación. El número de células se midió con un *Quantos™ Cell Proliferation Assay Kit* (*Stratagene*, La Jolla, CA) todos los días.

### Ejemplo 1

#### Identificación de una secuencia de cDNA humana

Un DNA que codifica un homólogo humano putativo del gen *Out at First (oaf)* de *Drosophila* se muestra en el SEQ ID NO: 1. Un alineamiento de hsOAF y de OAF de *Drosophila* se muestra en la Figura 7. El polinucleótido comprende 2366 pares de bases y se identifica un marco de lectura abierto. Una traducción del ORF, un polipéptido de 273 aminoácidos, se muestra en el SEQ ID NO: 2. La Figura 4 proporciona las secuencias de DNA y de aminoácidos e indica la posición del ORF. Los primeros 30 aminoácidos constituyen un péptido señal, lo que indica que la proteína puede ser secretada. La secuencia de aminoácidos del péptido señal es: MRLPGVPLARPALLLLLLPLLAPLLG#TGAPA (SEQ ID NO: 3). “#” indica el lugar del sitio de corte por proteasa pronosticado.

### Ejemplo 2

#### Expresión diferencial del SEQ ID NO: 1 en líneas celulares de cáncer de mama

Se comparó la expresión del SEQ ID NO: 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 en hueso y/o pulmón;

MDA-MB-468, derivada de células de cáncer de mama humano, negativas a receptores estrogénicos;

MDA-MB-435, derivada de células de carcinoma de mama humano, negativas a receptores estrogénicos;

MCF-7, derivada de células de cáncer de mama humano no metastásicas, y

ZR-75-1, derivada de células de carcinoma de mama humano, positivas a receptores estrogénicos.

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

La expresión en MDA-MB-361 fue el 6% del nivel en MDA-MB-435; la expresión en MDA-MB-468 fue el 36% del nivel en MDA-MB-435; y la expresión en MCF-7 fue el 3% del nivel en MDA-MB-435. Por lo tanto, como se muestra en la Tabla 2, existe una clara tendencia a un aumento de expresión del SEQ ID NO: 1 en las líneas celulares de cáncer de mama derivadas de tumores humanos con un alto potencial metastásico.

TABLA 2

Línea celular altamente metastásica	Líneas celulares poco metastásicas: % de expresión en comparación con una línea celular altamente 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%	ND

Un patrón similar de expresión de este gen continuó apareciendo en muestras de tejido tumoral procedentes de ratones SCID trasplantados con líneas celulares de carcinoma mamario tumorigénicas (Figura 6).

#### Ejemplo 3

*hsOAF codifica una proteína secretada y los niveles de secreción de la proteína hsOAF concuerdan con los niveles de expresión del mRNA de hsOAF de líneas celulares de carcinoma mamario*

Una secuencia de péptido señal pronosticada se localiza en el extremo N-terminal de la secuencia de aminoácidos deducida del gen hsOAF (Figura 3). Para verificar la secreción de la proteína hsOAF, se llevó a cabo una transfección transitoria de células COS-7 y de células MCF-7 con el vector pRetro-On que alberga el cDNA de hsOAF. Mientras tanto, el vector pRetro-On que alberga GFP se usó como control. Usando un antisuero contra hsOAF, procedente de conejo, se detectó la proteína hsOAF secretada en los medios de cultivo Opti-MEM1 de ambas líneas celulares después de la transfección con hsOAF, mediante inmunotransferencia (Fig. 8A). La proteína hsOAF secretada probablemente se encontraba glicosilada ya que se observaron múltiples bandas con pesos moleculares aparentes más altos (el MW pronosticado de la proteína hsOAF secretada es 28 kDa). El mismo antisuero contra hsOAF se usó para detectar la secreción de la proteína hsOAF por diferentes líneas celulares de carcinoma mamario. Los niveles de secreción de la proteína hsOAF concordaron con los niveles de expresión del mRNA de hsOAF entre estas líneas celulares en su totalidad: las líneas celulares altamente metastásicas mostraron una secreción de hsOAF mucho más fuerte que las líneas celulares poco metastásicas/no metastásicas (Fig. 8B). MDA-MB-435 presentó la secreción más fuerte de proteína hsOAF.

#### Ejemplo 4

*La supresión de la expresión de hsOAF en células MDA-MB-435 producida por oligonucleótidos antisentido produjo un cambio morfológico, una disminución de la invasividad de las células y una ralentización de la velocidad de proliferación*

Para determinar si un alto nivel de expresión génica de hsOAF es esencial para el potencial metastásico de células de carcinoma mamario humano, se usó la tecnología basada en oligonucleótidos antisentido para suprimir la expresión de hsOAF, y después se observaron los consiguientes efectos. Se eligió la línea celular MDA-MB-435 ya que esta línea celular altamente metastásica mostró la secreción más fuerte de la proteína hsOAF entre todas las líneas celulares de cáncer de mama estudiadas. Se eligieron varios pares de oligonucleótidos hsOAF antisentido (AS) y oligonucleótidos hsOAF inversos de control (RC) (del inglés, "*Reverse Control*") para realizar ensayos con respecto a su capacidad para bloquear la expresión génica de hsOAF a nivel del mRNA. Se realizó un análisis cuantitativo en tiempo real de RT-PCR en un *Lightcycler* (Roche Diagnostics, Indianapolis, IN) para medir los niveles del mRNA de hsOAF en las células. Kang, S. y col., *Cancer Research* 60(18): 5296-5302 (2000). El par mejor se seleccionó entonces para la titulación de la concentración de los oligonucleótidos de trabajo. Se prefiere una concentración de oligonucleótidos baja para disminuir la potencial toxicidad de los oligonucleótidos para con las células. Los resultados indicaron que el tratamiento con una concentración 100 nM del oligonucleótido antisentido fue suficiente para disminuir significativamente la secreción de la proteína hsOAF de las células MDA-MB-435 (Figura 12). Este par de oligonucleótidos (SEQ ID NO: 4 (AS) y SEQ ID NO: 5 (RC)), a una concentración de trabajo 100 nM, se usó en todos los experimentos que se detallan a continuación.

Después del tratamiento de las células MDA-MB-435 con un oligonucleótido hsOAF antisentido, se observó un espectacular cambio morfológico de las células junto con una disminución de la secreción de la proteína hsOAF (Fig. 10A). Las células se hicieron más esféricas y perdieron sus profusas protuberancias. Al mismo tiempo, las células tratadas con un oligonucleótido inverso de control permanecieron igual que las células MDA-MB-435 normales en cultivo de tejidos. Además, el medio de cultivo de las células MDA-MB-435 normales que contenía un nivel alto de proteína hsOAF, al igual que el medio acondicionado añadido a las células tratadas con un oligonucleótido antisentido, fue capaz de prevenir este cambio morfológico, aunque no completamente. Este cambio de la forma de la célula puede ser una indicación de la disminución de la capacidad de invasión de las células.

El ensayo de invasión en Matrigel se llevó luego a cabo para estimar la capacidad de invasión de las células. Se ha informado de que una morfología estrellada invasiva de las células de cáncer de mama embebidas en el Matrigel, se correlaciona con el potencial metastásico de estas células (Thompson, E.W. y col., *J. Cell Physiol.* 150(3): 534-44 (1992); Sugiura, T. y col., *J. Cell Biol.*, 146(6): 1375-89 (1999); Albini, A. y col., *Cancer Res.* 47(12): 3239-45 (1987); y Kramer, R.H. y col., *Cancer Res.* 46(4 Pt 2): 1980-89 (1986)) y esto se confirmó con diferentes líneas celulares de cáncer de mama cultivadas en Matrigel. Las células se trataron con tripsina, se contaron y se mezclaron con el Matrigel. Los medios se dispusieron luego sobre la mezcla de células-Matrigel. Después de 6 días de incubación, se estudió la invasión celular (Fig. 10B). Los resultados mostraron que las células tratadas con un oligonucleótido hsOAF inverso de control formaron estructuras tridimensionales, penetrantes e invasivas, a modo de red, al igual que hicieron las células MDA-MB-435 normales; por otro lado, las células tratadas con un oligonucleótido hsOAF antisentido únicamente formaron colonias lisas y esféricas. De nuevo, se observaron también colonias penetrantes en las células tratadas con un oligonucleótido hsOAF antisentido e incubadas en el medio acondicionado. Estos datos demuestran que la proteína hsOAF secretada se requiere para la invasividad y el potencial metastásico de las células MDA-MB-435.

Se realizaron experimentos adicionales para examinar si la proteína hsOAF secretada estaba implicada en el crecimiento de las células MDA-MB-435. Los resultados del ensayo de proliferación celular indicaron que la supresión de la secreción de la proteína hsOAF sin duda ralentizaba la velocidad de proliferación de las células MDA-MB-435, aunque el cambio era moderado.

#### Ejemplo 5

*Análisis de transferencia Northern de la expresión de RNA en líneas celulares de cáncer de mama humano y en tejidos humanos*

Como se muestra en la Figura 5, la expresión del mRNA estaba regulada por aumento en las líneas celulares metastásicas MDA-MB-231 y MDA-MB-435. El RNA total se preparó usando el *RNeasy Kit* procedente de Quiagen. El análisis de transferencia Northern se realizó usando 20 µg - 30 µg de RNA total aislado mediante extracción en tiocianato de guanidinio/fenol, cloroformo, a partir de las líneas celulares, de los tumores primarios o de metástasis pulmonares. Los tumores primarios y las metástasis pulmonares se desarrollaron a partir de líneas celulares inyectadas en ratones SCID según métodos muy conocidos en la técnica. Los plásmidos que contenían clones de cDNA parcial del hOAF clonado en el Vector pCR2.0-TA (*In vitro* gen), se marcaron radiactivamente y se hibridaron a 65°C en *Express-hyb* (Clontech). Entre todos los tejidos estudiados, el hígado, páncreas, bazo, ovario e intestino delgado mostraron una expresión de hsOAF significativa. La expresión del mRNA 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 a hsOAF en diversos tumores y tejidos normales.

TABLA 3

Inmunohistoquímica: Porcentaje de positivos a 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	
Hodgkin	1/8	

#### Ejemplo 6

##### *Ensayo en agar blando*

*Ensayo en agar blando:* La capa inferior consistió en 2 ml de agar al 0,6% en medios recién dispensados en las placas a las pocas horas de haber extendido las células. En lo que se refiere a la capa de células, las células MDA-MB-435, según se ha descrito en lo que antecede, se separaron de la placa en tripsina al 0,05% y se lavaron dos veces en el medio. El recuento de células se efectuó en un contador Coulter y las células se resuspendieron hasta alcanzar 106 por ml de medio. Partes alícuotas de 10 ml se pusieron con el medio en placas de 96 pocillos (para comprobar el recuento usando WST1), o se diluyeron más para el ensayo en agar blando. Se diluyeron 2000 células en 800 µl de agar al 0,4%, en pocillos duplicados, sobre la capa inferior de agar al 0,6%.

*Capa de medio:* Después de que el agar de la capa de células hubo solidificado, 2 ml de medio se difundió sobre la superficie y se añadió un oligonucleótido antisentido o un oligonucleótido inverso de control sin vehículos de suministro. Cada 3-4 días se añadieron medios recién preparados y oligonucleótidos.

El recuento de las colonias se realizó en de 10 días a 3 semanas. Los campos de colonias se contaron de manera visual. Los valores de metabolismo de Wst-1 se usaron para compensar las pequeñas diferencias en el número de células de partida. Los campos de mayor tamaño se pueden recorrer para realizar un registro visual de las diferencias. Los resultados se muestran en la Figura 6, en la que las células MDA-MB-435 tratadas con oligonucleótido antisentido formaron un número menor de colonias en comparación con las células expuestas al oligonucleótido de control.

Los expertos en la técnica reconocerán, o serán capaces de encontrar, usando sólo una experimentación de rutina, muchas realizaciones equivalentes a las realizaciones específicas de la invención aquí descritas. Esas realizaciones específicas y equivalentes se tiene intención de que estén abarcadas en las siguientes reivindicaciones.



REIVINDICACIONES

1. Uso de una molécula de oligonucleótido antisentido que comprende una secuencia antisentido de al menos 8 nucleótidos de longitud que es complementaria a una porción de la secuencia codificadora que abarca los nucleótidos del 365 al 1173 del SEQ ID NO: 1, para la preparación de una composición farmacéutica para el tratamiento del cáncer de mama y/o de metástasis de cáncer de mama.

2. Uso según la reivindicación 1, en el que la secuencia antisentido tiene una longitud de al menos 12, 15, 18, 20, 25, 30, 35, 40, 45 ó 50 nucleótidos.

3. Uso según la reivindicación 1 ó 2, en el que la molécula de oligonucleótido antisentido comprende ribonucleótidos y/o desoxirribonucleótidos.

4. Uso según cualquiera de las reivindicaciones de 1 a 3, en el que la molécula de oligonucleótido antisentido tiene la secuencia del SEQ ID NO: 4.

5. Uso según cualquiera de las reivindicaciones de 1 a 4, en el que la molécula de oligonucleótido antisentido se proporciona en una construcción de DNA.

6. Uso según cualquiera de las reivindicaciones de 1 a 5, en el que la composición es para administración local o sistémica.

7. Uso según cualquiera de las reivindicaciones de 1 a 6, en el que la composición disminuye la expresión de los genes marcadores metastásicos en el 50%.

8. Uso de la reivindicación 1, en el que la molécula de oligonucleótido antisentido comprende al menos 20 nucleótidos contiguos, complementarios a la secuencia codificadora que abarca los nucleótidos del 365 al 1173 del SEQ ID NO: 1.

SEQ ID NO:1

CCGCGAGGTGCGCGGTCTCTTTAAGCGGGTCTGCTGGTTTCTGTTTCTGAAGGAAGTGACGGGGGTGGGATTGAATGAAAAGTCAAAA  
CACAGGCTCGCABCGCTGGAGCCCGGGGCGCGGAGCCGGGCGGGGACGCGCGTCTCGGCTCGGGGCGCGGGGGGCGCCCTGCTGAGCG  
CTACCCAGGTGCGTCCGCGCCACCTCGCGGGCGACCCCGGGCCAAAGCCCCCGCGGAGCGGCTCCGGGCGCCCGAACTAGCCCCAACT  
TTGGGCGAAGTTTGCTGCGCTCTCCCGCCCCACGCGGCGCGCGGGGCGCGGACGCGAGCGGCCCCGGGGATGCGCTTCCCGGGT  
ACCCCTGGCGCGCCCTGCGTGTCTGTCTGCTGCGCTGCTCGCGCGCTGCTGGGAACGGGTGCGCGGCGGAGCTGCGGGTCCGCGTGCG  
GCTGCGGACGGCGAGGTGACCGAGGAGAGCTGCAGGCGGACAGCGGCGGACAGCATCAGCTCGAGCTGCGCAAGCCGAGCGCACCT  
CGTCTCCTTCAACCGCGACTTCAAGAAGGATGTGAAGGTCTTCCGGCCCTGATCCTGGGGAGCTGGAGAAGGGCAGAGTCAGTTCAGGC  
CCTCTGCTTTGTACCCAGCTGCAGCACAATGAGATCATCCCCAGTGAGCCATGGCCAAAGCTCGGCGAGAAAATCCCGGGCAGTGCGGCA  
GGCGGAGGAGTTCCGGGTCTGGAGCATCTGCACATGATGTGCGTGTCAACTTCAGCCAGGGGGCCCTGCTGAGCCCCATCTCCACAACGT  
GTGTGCGGAGGCGGTGGATGCCATCTACCCCGCCAGGAGGATGTCCGGTTCTGGCTGGAGCAAGGTGTGGACAGTTCTGTGTTCGAGGCTCT  
GCCAAGGCCCTCAGAGCAGGCGGAGTGCCTCGCTGCAGGCGAGTGGGGACCGGGGAAGCCCTGCGTCTGCCACTATGCCCTGAGCCTGGC  
CTGGTACCCCTGCATGCTCAAGTACTGCCACAGCCGCGACCGGCCACGCCCTACAAGTGTGGCATCCGAGCTGCCAGAAGAGCTACAGCTT  
CGACTTCTACGTGCCCCAGAGGAGCTGTGTCTCTGGATGAGGATCCCTACCAAGCTAGGGTGGGAGCAACCTGGGCGGGTGGTGTCTG  
GGCCACTGCTCTTACCAGCCACTAGAGGGGTGGCAACCCCACTGAGGGCTTATTTCCCTCCCTCCCACTCCCTGGCCCTAGAGCCT  
GGGCCCCCTCTGGCCCCATCTCAGTACTGTGAAGGGGTGTGGCATGGCAGGGGTCTCATGAAGGCACCCCATTCACCCCTGTGCTTCT  
CTTGCGGGCAGAGAGGGAGAGAAGGGTCCCGAGATCTACACCCCTCCCTCCTGCATCTCCCTGGAGTGTTCAGTTGCAAGCTGCCAAAACA  
TGATGGCTCTGGTGTCTGTGAAGTCTTGAAGCTTGAAGCTTAAAGAGTCTATACCTGGACACCCACCTCCCGAGACACAACCTCC  
TTCCCATGCACATCTGGAAGGAGTGGCCCTCAGTCCCTTCTACTCCCCAAGGGGCTCACTATCCCCAAGAGGAGCTGTGGG  
GACCCAGCAGCGAGCCCTGTACTGGATTACAGCATATCTCATCTCTGGCCCCGAGGCTGCCTGTGGGGCAGTGGAGACCTCCCATCACTG  
AGACAGATCACAGACCAGAGTGCCTTCCCGACCTGGACGTTGCTCCAAAACAGGCACCGAGCTCTTCCCTCTCTAGACAGAAATATTT  
TGTAAGGTTCTGGGCGAGGAGGAGCATGAAGTACGAGGAAAACCTGAATTCAGATTTTAAATGCAAAGTATTTATCTTCTACAGAAA  
TAAACGTTTTAAGTTTTTACTTGACTAATGAGACCCAGAGTTTGGAGAAAACCTTTGGCCAATGCTGCCACCTGATGTCAGAAAGTGTCCCA  
CACCTAGCAGTGGCTATCTTGAACAAGAACTTGAAGACCACTACTGTGTGCTCAGCCATTTGAGGAAGGAAGGAGGAGAAGGAAGTGT  
TACTAGGAAGGATGAGATAAACTTCTGCACCAAGACAATGAGACAGACATACTGCAACCGTAGTAAGCCAGTCAGAAATAGCCAGCGCG  
AAGGAAGAGATGGGTGGAGATTGAACCCCGCTTCAGATCTGGCTCGGCTACTTACCTGCTGTGAGCCATGGGTCAAGTGTGTTGACCT  
CTCTGTGCTCCACTCCCTTAGCTATAAAATGAGCTTACTT

*Fig. 1*

SEQ ID NO:2

MRLPGVPLARPALLLLPLLAPLLGTGAPAE LRVRRLPDQVTEESLQADSDADSLSLELRKPDGTLVSFTADFKDKVKVFRALILGELEKG  
QSOFQALCFVTQLQHNEIIPSEAMAKLRQKNPRAVRQAEEVRGLEHLHMDVAVNFSGALLSPHLHNVCAEAVDAIYTRQEDYRFLWLEQVDS  
SVFEALPKASEQAE LPRCRQVGRGKPCVHYGLSLAWYPCMLKYCHSRDRPTPYKCGIRSCQKSYSFDFYVPRQLCLWDEDPYPG\*

*Fig. 2*

SEQ ID NO:3

MRLPGVPLARPALLLLPLLAPLLG TGAPA

*Fig. 3*

## Intervalo de la secuencia: 1 a 2366

[illegible]

*Fig. 4A*

730 740 750 760 770 780 790 800 810  
 GAAAAATCCCGGGCAGTGCGGCAGGCGGAGGAGTTCTGGGGTCTGGAGCATCTGCACATGGATGTCGCTGTCAACTTCAGCCAGGGGGC  
 CTTTTTAGGGGCCCGTCACGCCGTCGCCCTCCTCAAGCCCCAGACCTCGTAGACGTGTACCTACAGCGACAGTTGAAGTCGGTCCCCCG  
 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 LA OAF HUMANA [A] \_\_\_\_\_>

820 830 840 850 860 870 880 890 900  
 CCTGCTGAGCCCCATCTCCACAAGTGTGTGCCGAGGCCGTGGATGCCATCTACACCCGCCAGGAGGATGTCCGGTCTGGCTGGAGCA  
 GGACGACTCGGGGTAGAGGTGTTGCACACACGGCTCCGGCACCTACGGTAGATGTGGCGGTCTCTACAGGCCAAGACCGACCTCGT  
 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 LA OAF HUMANA [A] \_\_\_\_\_>

910 920 930 940 950 960 970 980 990  
 AGGTGTGGACAGTTCTGTGTTCCGAGGCTCTGCCAAGGCCTCAGAGCAGCGGAGCTGCCTCGCTGCAGGCAGGTGGGGGACCGCGGAA  
 TCCACACCTGTCAAGACACAAGCTCCGAGACGGGTTCCGGAGTCTCGTCCGCTCGACGGAGCGACGTCCGTCCACCCCTGGCGCCCTT  
 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 LA OAF HUMANA [A] \_\_\_\_\_>

1000 1010 1020 1030 1040 1050 1060 1070 1080  
 GCCCTGCGTCTGCCACTATGGCCTGAGCCTGGCCTGGTACCCCTGCATGCTCAAGTACTGCCACAGCCGCGACCGGCCACGCCCTACAA  
 CGGACGCGAGACGGTGATACCGGACTCGGACCGGACCATGGGACGTACGAGTTCATGACGGTGTGGCGCTGGCCGGGTGCGGGATGTT  
 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 LA OAF HUMANA [A] \_\_\_\_\_>

1090 1100 1110 1120 1130 1140 1150 1160 1170  
 GTGTGGCATCCGAGCTGCCAGAAGAGCTACAGCTTCGACTTCTACGTGCCCCAGAGGCAGCTGTGTCTCTGGGATGAGGATCCCTACCC  
 CACACCGTAGGCGTCGACGGTCTTCTCGATGTGGAAGCTGAAGATGCACGGGCTCTCGTGCACACAGAGACCTACTCTAGGGATGGG  
 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 LA OAF HUMANA [A] \_\_\_\_\_>

1180 1190 1200 1210 1220 1230 1240 1250 1260  
 AGGCTAGGGTGGGAGCAACCTGGGCGGGTGGCTGCTCTGGGCCACTGCTCTTACCAGCCACTAGAGGGGTGGCAACCCCCACCTGAG  
 TCCGATCCACCCCTCGTGTGACCCGCCACCGACGAGACCCGGGTGACGAGAAGTGGTGGTGATCTCCCCACCGTTGGGGTGGGACTC  
 G \*  
 \_\_\_\_\_ TRADUCCIÓN DE LA OAF HUMANA [A] \_\_\_\_\_>

1270 1280 1290 1300 1310 1320 1330 1340 1350  
 GCCTTATTTCCCTCCCTCCCACTCCCTGGCCCTAGAGCCTGGGCCCTCTGGCCCATCTCACATGACTGTGAAGGGGTGTGGCATG  
 CGGAATAAGGGAGGGAGGGGTGAGGGGACCGGGATCTCGACCCGGGAGACCGGGTAGAGTGTACTGACACTTCCCCACACCGTAC

Fig. 4B

1360 1370 1380 1390 1400 1410 1420 1430 1440  
GCAGGGGGTCTCATGAAGGCACCCCAATCCCAACCTGTGCCTTCTTGCGGGCAGAGAGGGAGACAAGGGCTCCCCAGATCTACACCC  
CGTCCCCCAGAGTACTTCGGTGGGGGTAAAGGTGGGACCGGAAGGAACGCCCGTCTCTCCCTCTCTTCCCGAGGGGTCTAGATGTGGGG

1450 1460 1470 1480 1490 1500 1510 1520 1530  
TCCCTCCTGCATCTCCCCTGGAGTGTTCACTTGCAAGCTGCCAAAACATGATGGCCTCTGTTGTTCTGTTGAACCTCTGAACGTTTAG  
AGGGAGGACGTAGAGGGGACCTCACAAGTGAACTGCGACGTTTTGTACTACCGGAGACCAACAAGACAACCTGAGGAACCTGCAATC

1540 1550 1560 1570 1580 1590 1600 1610 1620  
ACCCATAAAGGAGTCTATACCTGGACACCCACCTCCCCAGACACAATCCCTTCCCCATGCACACATCTGGAAGGAGCTGGCCCTCAGT  
TGGGATTTTCTCAGATATGGACCTGTGGGTGGAGGGTCTGTGTGTAGGGAAGGGGTACGTGTGTAGACCTTCTCGACCGGGGAGTCA

1630 1640 1650 1660 1670 1680 1690 1700 1710  
CCCTTCCTACTCCCCAACAGGGGCTCACTATCCCCAAGAAGGAGCTGTTGGGGACCCACGACGCAGCCCGTGTACTGSAATACAGCAT  
GGGAAGGATGAGGGGTTGTTCCCGAGTGATAGGGGTTCTTCCTCGACAACCCCTGGGTGCTGCGTCGGGGACATGACCTAATGTCGTA

1720 1730 1740 1750 1760 1770 1780 1790 1800  
ATTCTCATCTCTGGCCCGAGGCTGCCTGTGGGGCGAGTGAGACCTCCCATCACTGAGACAGATCACAGACCAGAGTGCCTTTCGGG  
TAAGAGTAGAGACCGGGGCTCCGACGGACACCCCGCTCACCTCTGGAGGGTAGTACTCTGTCTAGTGTCTGTGCTCACGGAAGGGCC

1810 1820 1830 1840 1850 1860 1870 1880 1890  
ACCTGACGTTGCCTCCAAAACAGGCACCAAGCTCTTTCCCTCTCTAGACAGAAATATTTTGTAGGTTCTGGGCAGGGAGGGAGCATG  
TGGACCTGCAACCGAGGTTTGTCCGTGGTCGAGAAAGGAGAGATCTGTCTTTATAAAAACATTCGAAGACCCGTCCTCCCTCGTAC

1900 1910 1920 1930 1940 1950 1960 1970 1980  
AAGTACGAGGAAACCTGAATCCAGATTTTAAATGCAAAGTATTATCATTCTACCAGAAATAAACGTTTTAAGTTTTACTTGACTA  
TTCATGCTCCTTTTGAACCTTAAGGTCTAAAAATTACGTTTCATAAATAGTAAGATGGTCTTTATTGCAAAATCAAAAATGAAGTAT

1990 2000 2010 2020 2030 2040 2050 2060 2070  
ATGAGACCCAGAGTTTGGAGAAAACCTTTTGCCCAATGCTGCCACCTGATGTCAGAAAGTGTCCCAACCCCTAGCAGTGGCCTATCTTGG  
TACTCTGGGTCTCAAACCTCTTTTGAAAACCGGTACGACGGTGGACTACAGTCTTTCACAGGGGTGTGGGATCGTCACCGGATAGAAC

2080 2090 2100 2110 2120 2130 2140 2150 2160  
AACAGAACTTCGAAAGCACCTACTGTGTGCTCAGCCATTGAGGAAGGAAGGAGGAGAAGGAAGATGTTACTAGGGAAGGATGAGATAA  
TTGTTCTTGAAGCTTTCTGGATGACACACGAGTCGGTAAACTCCTTCCTTCCTCCTTCCTTCCTCTACAATGATCCCTTCCTACTCTATT

*Fig. 4C*

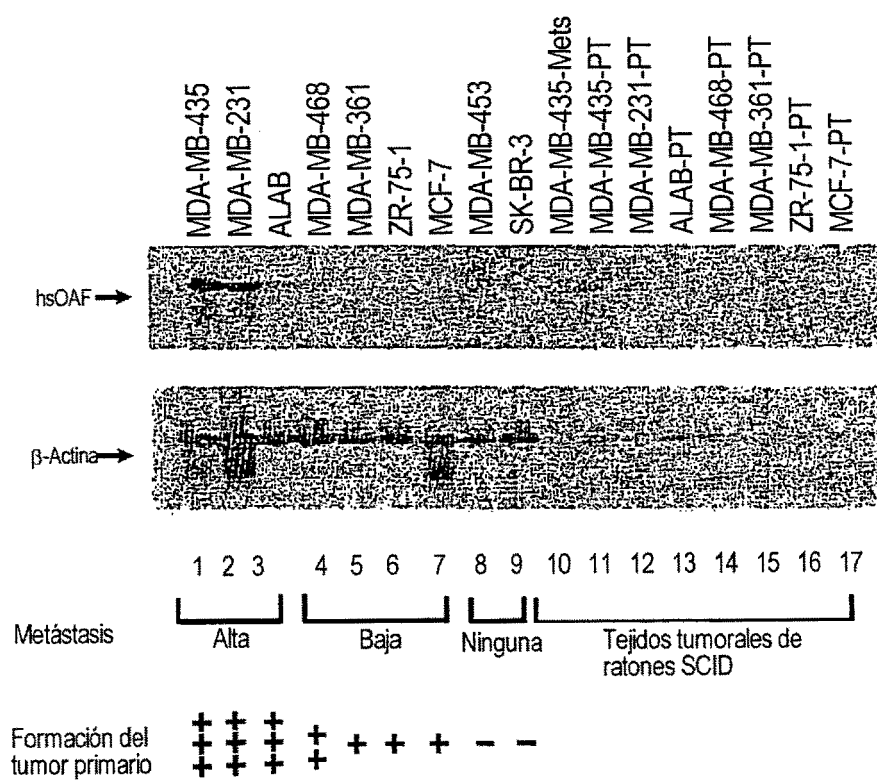
ES 2 311 016 T3

2170 2180 2190 2200 2210 2220 2230 2240 2250  
AACTTCTGCACCCAAGACAATGAGACAGACATAACTGCAACCGTAGTAAGCCAGTCAGAAATAGCCAGCGCGAAGGCAAGAGATGGGGTG  
TTGAAGACGTGGGTTCTGTTACTCTGTCTGTATTGACGTTGGCATCATTCGGTCAGTCTTTATCGGTCGCGCTTCGGTTCCTACCCAC

2260 2270 2280 2290 2300 2310 2320 2330 2340  
GAGATTGGAACCCCGCTTCAGATCTGGGCTCGGCTACTTACCTGCTGTGCAGCCATGGGTCAAGTTGCTTGACCTCTCTGTGCCTCCACT  
CTCTAACCTTGGGGCGAAGTCTAGACCCGAGCCGATGAATGGACGACACGTCGGTACCCAGTTCAACGAACTGGAGAGACACGGAGGTGA

2350 2360  
CCCTTAGCTATAAAATGAGCTTACTT  
GGGAATCGATATTTACTCGAATGAA

*Fig. 4D*



*Fig. 5*

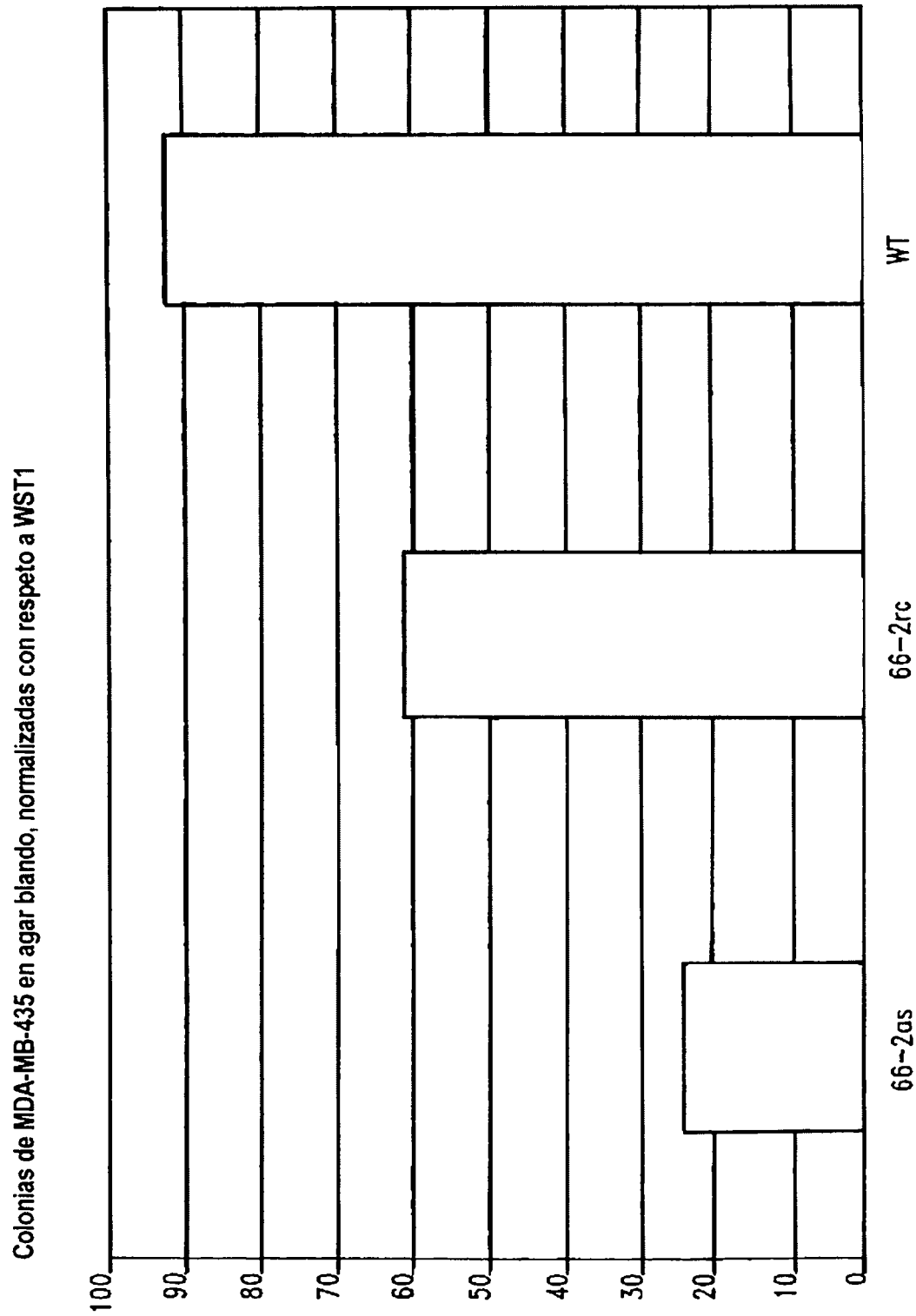


Fig. 6



F 1 MILKEEHPHOSIETAANAARQAQVRWRMAHLKALSRITTAHGNCCGRVYSKNHFFKSR 10  
 1 MILKEEHPHOSIETAANAARQAQVRWRMAHLKALSRITTAHGNCCGRVYSKNHFFKSR 60

F 11 PALLILLPLAPLLGTGAPAEKVRRLPGQVTEESLQADSDASTSLEKPDGILVS 70  
 61 AFLWFLLCNLVMNADARAISQLINVONOGGEVQESTISNIGEDLITLFFKDTGILIT 120

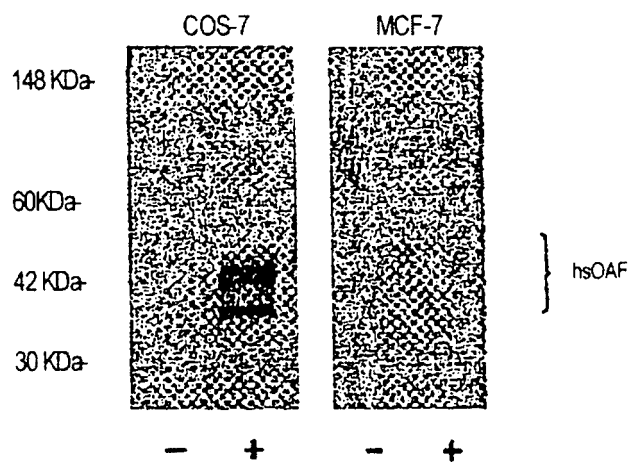
F 71 FTADIKKDVKVFRAITLGEIKKGQSOFQALCFMTQDHNEITSEAMAKLRQKNPRAVRQ 130  
 121 QVIDIRNEVOILKALVLGEIKRGQSQVQVNCFAIKFNKGFISSEAMAKLRQKNPHTIRT 180

F 131 AEEVRGLEHLNDVAVNFSGALLSPILHNVC AEAVDAIMTQEDVRFMEDGVDSVIF 189  
 181 REEDKGRFTFTSSWQLNPSLITRHLQGLCAEANDATVNRVDLKAAELRGSSISL 240

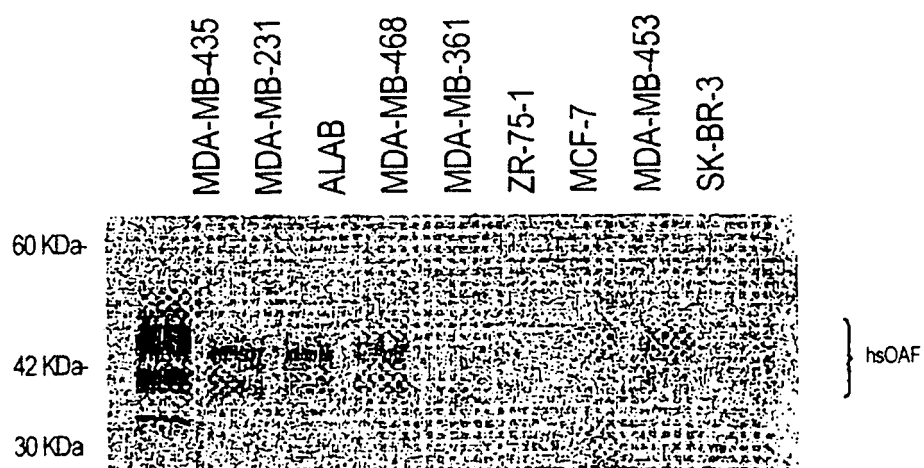
F 190 KALPKASEQALPRCQVGGORGKPCVCHYGLSLWVPCMLKVCHER-----D----- 236  
 241 KAATEKFPTILSTRQNEVSSLWAPCLCNLETCIGWVPCGLKVCKKGVAGADSSGAQQA 300

F 237 RPTVVKCGIRSCOKSYSEFVVPDRQLWDEDPYPG 273  
 301 OPTNVRCGIKTCRKCTOPTVVRKQCLWDE 332

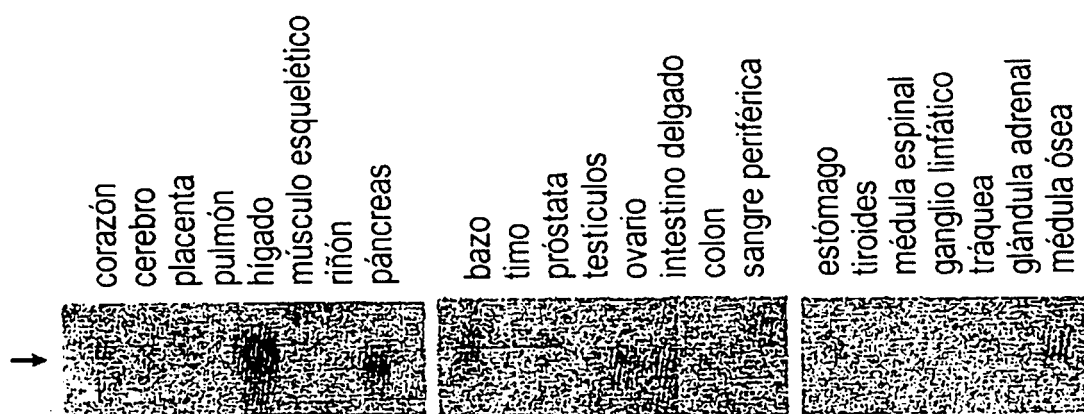
Fig. 7



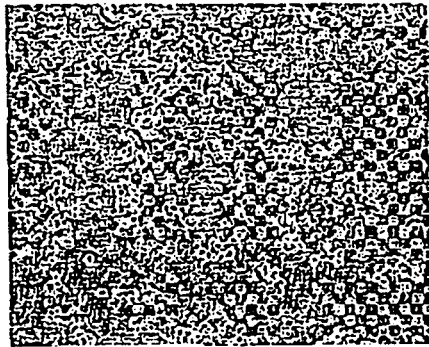
*Fig. 8A*



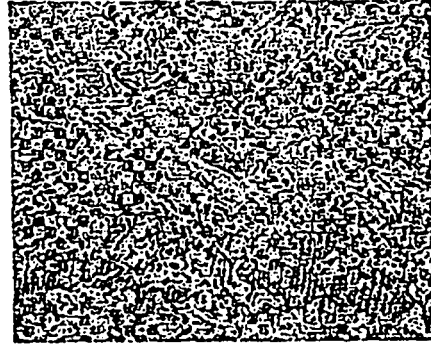
*Fig. 8B*



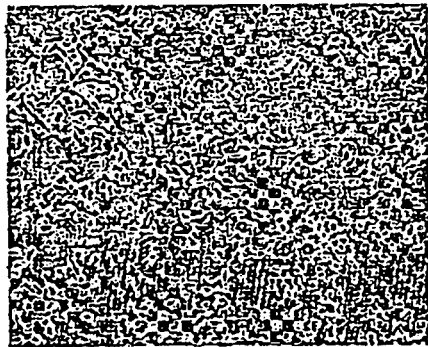
*Fig. 9*



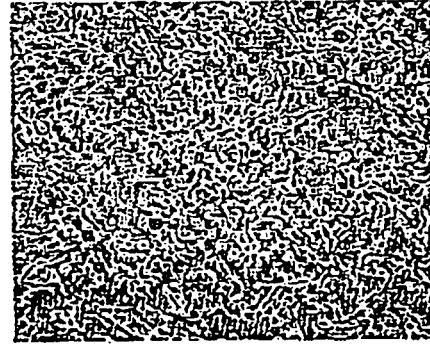
AS



RC

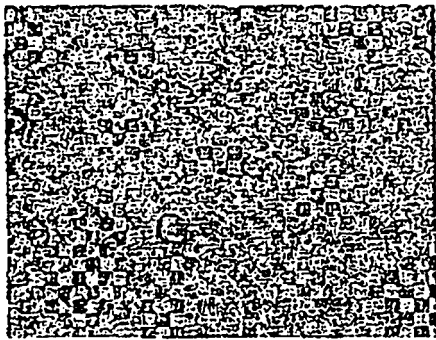


AS+M

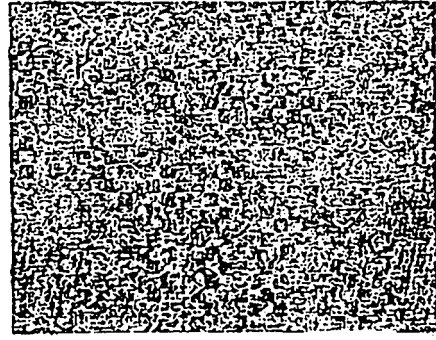


Normal

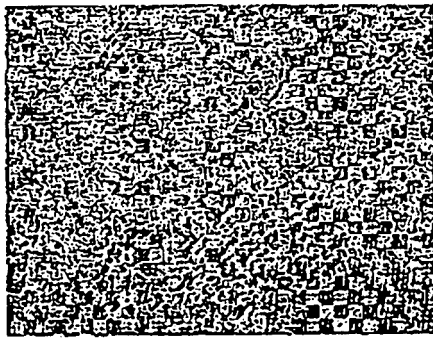
*Fig. 10A*



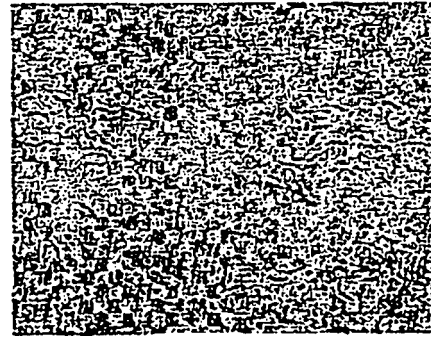
AS



RC



AS+M

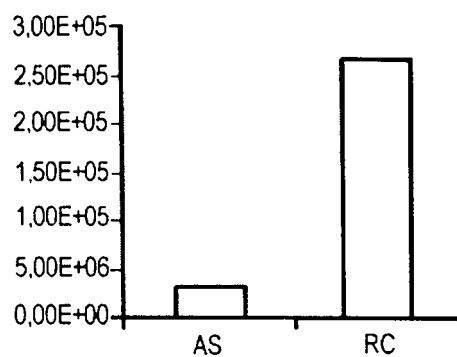


Normal

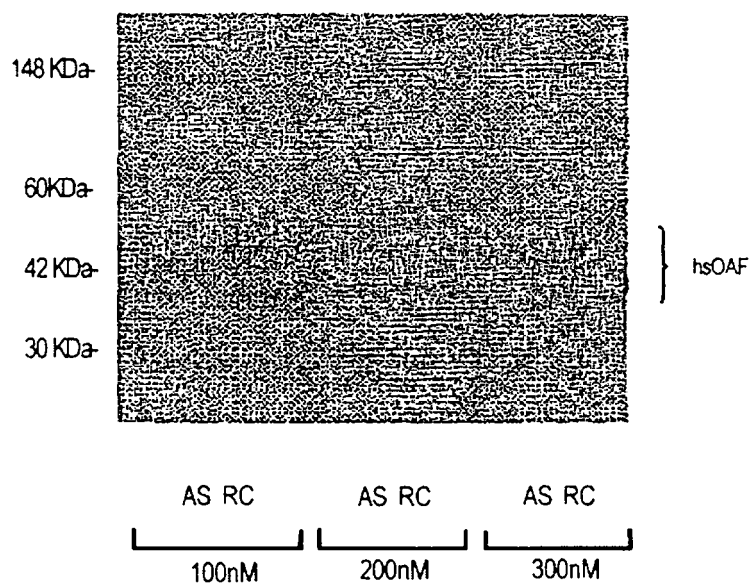
*Fig. 10B*

1 ccgcgaggtgcgcggtctctttaaggcgggtcctggtggtttctgtttcctgaaggaagtacggggggtgggattgaatgaaaagtg  
 89 caaaacacaggtcgcagcgtggagccccgggcccggagccggggcgggagcgcggtctccgcctcggggccgccccggccct  
 179 GCTGAGCGCTACCCACGTGCGTCCGCGCCACCTCGCGGGCGACCCCGCGGCCAAGGCCCCCGCGGAGCGGCTCCGCGGCGCCCCGAACT  
 269 AGCCCCCAACTTTGGGCGAAGTTTGCTGCGCTCTCCCCGCCCCACGCGGCGCGCGGGGCCGCGGACGGCAGCGGCCCCCGGGATG  
 1  
 359 CGCCTTCCCGGGGTACCCCTGGCGCGCCCTGCGTGTGCTGTGCTGCGCGCTGCTCGCGCCGCTGCTGGGAACGGGTGCGCCGCCGAG  
 2 R L P G V P L A R P A L L L L L P L L A P L L G T G A P A E  
 449 CTGCGGGTCCCGTGC GGCTGCCGACGGCCAGGTGACCGAGGAGAGCTGCAGGCGGACGACGCGGACAGCATCAGCCTCGAGCTG  
 32 L R V R V R L P D G Q V T E E S L Q A D S D A D S I S L E L  
 539 CGCAAGCCCCGACGGCACCCCTCGTCTCCTTACCGCCGACTTCAAGAAGGATGTGAAGGTCTTCCGGGCCCTGATCCTGGGGAGCTGGAG  
 62 R K P D G T L V S F T A D F K K D V K V F R A L I L G E L E  
 629 AAGGGGAGAGTCAGTTCAGGCCCTCTGCTTTGTACCCAGCTGCAGCACATGAGATCATCCCAGTGAGGCCATGGCCAGCTCCGG  
 92 K G Q S Q F Q A L C F V T Q L Q H N E I I P S E A M A K L R  
 719 CAGAAAAATCCCGGCGAGTGCGGCGAGGCGAGGAGGTTGCGGGTCTGGAGCATCTGCACATGGATGTGCTGTCAACTTCAGCCAGGGG  
 122 Q 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  
 809 GCCCTGCTGAGCCCCCATCTCCACAAGCTGTGTGCCGAGGCGGTGGATGCCATCTACACCGCCAGGAGGATGTCCGTTCTGGCTGGAG  
 152 A 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  
 899 CAAGGTGTGGACAGTTCGTGTTGAGGCTCTGCCAAGGCCCTCAGAGCAGGCGGAGCTGCCTCGCTGCAGGCAGGTGGGGACCGCGGG  
 182 Q 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  
 989 AAGCCCTGCGTCTGCCACTATGGCCTGAGCCTGGCCTGTACCCCTGCATGCTCAAGTACTGCCACAGCCGCGACCGGCCACGCCCTAC  
 212 K 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  
 1079 AAGTGTGGCATCCGAGCTGCCAGAAGAGCTACAGCTTCGACTTCTACGTGCCCCAGAGCAGCTGTGTCTCTGGGATGAGGATCCCTAC  
 242 K 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  
 1169 ccaggctagggtgggagcaacctgggagggtggtgctctgggcccactgctcttcaccagccactagaggggtggcaacccccacctg  
 272 p g \*  
 1259 aggccttatttccctccctccccactccctggccctagagcctgggcccctctggccccatctcacatgactgtgaaggggtgtggca  
 1349 tggcaggggtctcatgaaggcacccttccacccctgtgccttcccttgcgggcagagaggagagaagggtccccagatctacacc  
 1439 cctccctcctgcatctcccttgaggtgttacttgcaagctgccaaaacatgatggcctctggttgttctgttgaactccttgaacgttt  
 1529 agaccctaaaaggagtctatacctggacaccacctccccagacacaactcccttccccatgcacacatctggaaggagctggccctca  
 1619 gtccttccctactccccaaagggtcactatccccaaaggagctgttgggacccacgacgagccctgtactggattacagc  
 1709 atattctcatctctggcccgaggtgctgtggggcgagtgagacctcccatcactgagacagatcacagaccagagtgcttccc  
 1799 ggacctggagcttgctccaaaacaggcaccagctcttccctctctagacagaaatattttgtaagggtctggggcagggagggagca  
 1889 tgaagtacgaggaacttgaattccagatttttaagtcaaagtatttatcttaccagaaataaacgttttaagttttacttgac  
 1979 taatgagaccagagtttggagaaaactttggccaatgctgccacctgatgtcagaaagtgtccccacaccctagcagtgccctatctt  
 2069 ggaacaagaacttcgaaagcactactgtgtgctcagccatttgaggaaggaaggaggagaaggaagatgttactaggaaggatgagat  
 2159 aaaactctgcaccaagacaatgagacagacataactgcaacctgtagtaagccagtcagaaatagccagcgcaaggcaagagatgggg  
 2249 tggagattggaacccgcttcagatctgggctcggtacttacctgctgtgcagccatgggtcaagttgcttgacctctctgtgcctcca  
 2339 ctcccttagctataaatgagcttactt-polyA

Fig. 11



*Fig. 12A*



*Fig. 12B*