

(11) Número de Publicação: **PT 1248841 E**

(51) Classificação Internacional:

**C12N 15/11** (2007.10) **C12N 15/12** (2007.10)  
**C12N 15/62** (2007.10) **C12N 15/86** (2007.10)  
**C12N 5/10** (2007.10) **C07K 14/47** (2007.10)  
**C07K 16/18** (2007.10) **A61K 31/713** (2007.10)  
**C12Q 1/68** (2007.10) **G01N 33/577** (2007.10)  
**G06F 17/30** (2007.10)

(12) **FASCÍCULO DE PATENTE DE INVENÇÃO**

(22) Data de pedido: **2001.01.09**

(30) Prioridade(s): **2000.01.10 US 175462 P**

(43) Data de publicação do pedido: **2002.10.16**

(45) Data e BPI da concessão: **2008.07.23**  
**180/2008**

(73) Titular(es):

**NOVARTIS VACCINES AND DIAGNOSTICS, INC.**  
**4560 HORTON STREET EMERYVILLE, CA 94608**  
**US**

(72) Inventor(es):

**JOERG KAUFMANN** **US**  
**GREG HARROWE** **US**  
**CHRISTOPH REINHARD** **US**  
**SANMAO KANG** **US**

(74) Mandatário:

**PEDRO DA SILVA ALVES MOREIRA**  
**RUA DO PATROCÍNIO, N.º 94 1399-019 LISBOA** **PT**

(54) Epígrafe: **GENES EXPRESSOS DIFERENCIALMENTE NO CANCRO DA MAMA**

(57) Resumo:

## **RESUMO**

### **"GENES EXPRESSOS DIFERENCIALMENTE NO CANCRO DA MAMA"**

É associada uma sequência polinucleotídica, como apresentada na SEQ ID N°: 1, com o potencial metastático de células de cancro, especialmente, de cancro da mama. São proporcionados métodos para determinar se uma amostra de tecido de um tumor expressa um polipéptido um ARNm codificado por um polinucleótido como apresentado na SEQ ID N°: 1. São, também, proporcionados métodos e composições terapêuticos.

## **DESCRIÇÃO**

### **"GENES EXPRESSOS DIFERENCIALMENTE NO CANCRO DA MAMA"**

#### CAMPO DA TÉCNICO

Esta invenção refere-se à utilização de moléculas oligonucleotídicas anti-sentido para a preparação de uma composição farmacêutica, para o tratamento do cancro da mama e/ou metástases do cancro da mama.

#### ANTECEDENTES DA INVENÇÃO

O cancro da mama é uma das doenças malignas mais comuns, com, cerca de, 1000000 novos casos por ano, em todo o mundo. Apesar da utilização de vários de marcadores histoquímicos, genéticos e imunológicos, os médicos têm ainda dificuldades para prever quais os tumores que irão metastizar para outros órgãos. Alguns doentes necessitam de terapia adjuvante para prevenir a reincidência e as metástases e outros não. No entanto, a distinção entre estas sub-populações de doentes não é fácil e a via de tratamento não é facilmente identificada. Existe uma necessidade na técnica de novos marcadores para distinguir entre tumores que irão metastizar ou tenham metastizado e os que tenham uma menor probabilidade metastizar.

## SUMÁRIO da INVENÇÃO

De acordo com a presente invenção, é proporcionada a utilização de uma molécula oligonucleotídica anti-sentido para a preparação de uma composição farmacêutica para o tratamento do cancro da mama e/ou metástases do cancro da mama. O oligonucleótido anti-sentido compreende uma sequência anti-sentido com, pelo menos, 8 nucleótidos de comprimento a qual é complementar a uma porção da sequência codificante que abrange os nucleótidos 365 a 1173 da SEQ ID N°: 1.

## BREVE DESCRIÇÃO DOS DESENHOS

A Figura 1 ilustra a sequência polinucleotídica do *Out at First* humano (SEQ ID N°: 1).

A Figura 2 ilustra a sequência de aminoácidos codificada pela SEQ ID N°: 1 (SEQ ID N°: 2).

A Figura 3 ilustra o péptido sinal putativo (SEQ ID N°: 3).

A Figura 4 ilustra a tradução da SEQ ID N°: 1 (SEQ ID N°: 1, polinucleótido; SEQ ID N°: 2, sequência de aminoácidos).

A Figura 5 ilustra a expressão de hsOAF em relação à  $\beta$ -Actina, em linhas celulares tumorais e tecidos tumorais de murganhos SCID, desenvolvidos a partir das linhas celulares. "PT" refere-se a tumor primário.

A Figura 6 ilustra o desenvolvimento de colónias por células MDA-MB-435 em agar mole, após tratamento com o

oligonucleótido anti-sentido da SEQ ID N°: 4 (66-2as) ou o controlo reverso da SEQ ID N°: 5 (66-2rc), em relação a células não tratadas (WT).

A Figura 7 é um alinhamento da sequência de aminoácidos da OAF humana com a sequência de aminoácidos da OAF de *Drosophila*.

Figura 8. A Figura 8A ilustra a expressão da proteína hsOAF nas linhas celulares COS-7 e MCF-7. A Figura 8B ilustra a expressão da proteína hsOAF em linhas celulares de carcinoma mamário.

A Figura 9 ilustra a expressão de hsOAF em tecidos humanos normais.

Figura 10. A Figura 10A ilustra as modificações morfológicas observadas em células MDA-MB-435, após tratamento com o oligonucleótido anti-sentido (SEQ ID N°: 4). AS = anti-sentido; RC = controlo reverso (SEQ ID N°: 5); M = meio condicionado. A Figura 10 B ilustra a invasão celular após tratamento de células MDA-MB-435 com AS, RC e RC + M.

A Figura 11 ilustra a sequência sinal prevista da OAF humana (sublinhado duplo).

Figura 12. As Figuras 12A e 12B ilustram a secreção de hsOAF por células MDA-MB-435 tratadas com oligonucleótido anti-sentido (SEQ ID N°: 4) ou oligonucleótido de controlo reverso (SEQ ID N°: 5).

## DESCRIÇÃO DETALHADA DA INVENÇÃO

A causa mais letal em doentes com cancro da mama é a metástase dos carcinomas da mama e a sua proliferação em locais distantes (principalmente, pulmões e ossos). A metástase é um processo de múltiplos passos, pelos quais as células tumorais migram a partir do tumor primário, disseminam-se através do sangue e dos vasos linfáticos e são, seguidamente, depositadas em órgãos alvo específicos, onde estas recolonizam, Schirrmacher, V., *Adv. Cancer Res.* **43**:1-73, 1985 e Liotta, L. A. *et al.*, *Cell* **64**(2):327-36 (1991). A invasividade das células tumorais é crucial durante este processo, uma vez que estas devem contactar e atravessar numerosas membranas basais. Liotta, L. A., *Am. J. Pathol.* **117**(3):339-48 (1984) e Fidler, I. J., *Cancer Res.* **38**(9):2651-60 (1978). Consequentemente, a elucidação das causas moleculares da invasão das células tumorais e da metástase é essencial para o desenvolvimento de processos de tratamento eficientes de doentes de cancro da mama. Os genes expressos diferencialmente na metástase de tumores da mama são alvos potenciais que desempenham papéis decisivos durante a metástase. A identificação desses genes e da sua função biológica irá contribuir, significativamente, para o desenvolvimento de terapia e diagnóstico para o cancro da mama.

Foram descobertos alguns genes importantes envolvidos nas metástases do tumor da mama. A perda de receptor do estrogénio e presença de vimentina foram associadas à invasividade do tumor da mama humano e a um prognóstico desfavorável e, também, se correlaciona com a invasividade e com o potencial metastático de linhas celulares do cancro da mama humano. Aamdal S., *et al.*, *Cancer* **53** (11):2525-9 (1984); Clark, G. M., *et al.*, *Semin*

*Oncol.*, **2** Supl:20-5 (1988); Raymond, W. A., et al., *J. Pathol.* **157**(4):299-306 (1989); Raymond, W. A., et al., *J. Pathol.* **158**(2):107-14 (1989); e Thompson, E. W. et al., *J. Cell Physiol.* **150**(3):534-44 (1992). A subexpressão da E-caderina tem sido associada à invasividade do tumor mamário. Vleminckx, K., et al., *Cell* **66**(1):107-19 (1991) e Oka, H., et al., *Cancer Res.* **53**(7):1696-701 (1993). A maspina, um inibidor de proteases expresso em células epiteliais mamárias normais, mas não na maioria das linhas celulares do carcinoma da mama, foi capaz de suprimir a capacidade das células MDA-MB-435 para induzirem tumores e metastizarem em murganhos e invadirem a membrana basal *in vitro*. A perda da expressão da maspina ocorreu, mais frequentemente, em cancros avançados. Zou, Z., et al., *Science* **263**(5146):526-9 (1994) e Seftor, R. E., et al., *Cancer Res.* **58**(24):5681-5 (1998).

A sobreexpressão de TIMP-4 (inibidor tecidual de metaloproteinases-4) ou CLCA2 (canal-2 de cloreto activado por  $\text{Ca}^{2+}$ -2) em células MDA-MB-435 por transfecção, inibiu a tumorigenicidade, a invasividade e a capacidade de metastizar das células. Wang, M., et al., *Oncogene* **14**(23):2767-74 (1997) e Gruber, A. D., et al., *Cancer Res.* **59**(21):5488-91 (1999). Foi verificado que a sobreexpressão dos receptores do factor de crescimento IGF-IR e p185<sup>ErbB-2</sup> está envolvida na metástase do cancro da mama. Surmacz, E., et al., *Breast Cancer Res. Treat* **47**(3):255-67 (1998); Dunn, S. E., et al., *Cancer Res.* **58**(15):3353-61 (1998); Tan, M., et al., *Cancer Res.* **57**(6):1199-205 (1997); Dhingra, K., et al., *Semin Oncol.* **23**(4):436-45 (1996); e Revillion, F., et al., *Eur. J. Cancer* **34**(6):791-808 (1998).

Foi descrito que a protease de aspartilo catepsina D é um marcador de prognóstico desfavorável de doentes de cancro da mama e existe uma correlação significativa entre a concentração elevada de catepsina D no citossol do cancro da mama primário e o desenvolvimento de metáteses, embora não tenha sido encontrada qualquer correlação entre a secreção da catepsina D e a capacidade de invasão das linhas celulares do cancro da mama. Rochefort, H., *Breast Cancer Res Treat* **16**(1):3-13 (1990); Johnson, M. D., et al., *Cancer Res.* **53**(4):873-7 (1993); e Rochefort, H., et al., *Clin Chim Acta.* **291**(2):157-70 (2000). A osteopontina, uma glicoproteína segregada de ligação à integrina que se pensa estar envolvida na reabsorção óssea e na formação óssea, pode induzir a migração e a invasão de células de carcinomas mamários. Os níveis de osteopontina (níveis em células tumorais ou plasmáticos) têm sido associados com a malignidade intensificada do cancro da mama. Denhardt, D. T., et al. *FASEB J.* **7**(15):1475-82 (1993); Denhardt, D. T., et al., *J. Cell Biochem Suppl.*, **30-31**:92-102 (1998); Tuck, A. B., et al., *J. Cell Biochem.* **78**(3):465-75 (2000); Tuck, A. B., et al., *Oncogene* **18**(29):4237-46 (1999); e Singhal, H., et al., *Clin Cancer Res.* **3**(4):605-11 (1997).

É divulgada a clonagem de um novo gene sobreexpresso em linhas celulares do cancro da mama humano altamente metastáticas. Este codifica uma proteína segregada e a sua secreção proteica foi confirmada como sendo muito superior em linhas celulares do cancro da mama humano altamente metastáticas do que em linhas celulares pouco metastáticas/não-metastáticas. A inativação da secreção desta proteína da linha celular MDA-MB-435 agressiva, pela tecnologia de oligonucleótidos anti-sentido, resultou em alteração morfológica significativa, conjuntamente com invasividade e taxa de proliferação reduzidas



das células. O gene é denominado hsOAF, com base na sua homologia com o gene OAF de *Drosophila* (*out at first*). Bergstrom, D. E., et al. *Genetics* **139**(3):1331-46 (1995) e Merli, C., et al., *Genes Dev.* 10(10):1260-70 (1996).

É uma verificação da presente invenção, o facto de um polinucleótido ser diferencialmente expresso entre células do cancro da mama altamente metastáticas e células de cancro não-metastáticas ou pouco metastáticas. Este polinucleótido refere-se, consequentemente, a um gene marcador metastático. Esta informação pode ser utilizada para preparar reagentes de diagnóstico, específicos para os produtos de expressão do gene diferencialmente expresso. Este pode ser, também, utilizado em métodos de diagnóstico e de prognóstico, que irão auxiliar os médicos no planeamento de regimes de tratamento adequados para cancros, especialmente, da mama.

O polinucleótido é apresentado na Figura 1 (SEQ ID N°: 1) e a grelha de leitura aberta (ORF) prevista codifica um polipéptido apresentado na Figura 2 (SEQ ID N°: 2). Os 30 primeiros resíduos de aminoácido (SEQ ID N°: 3) compreendem um péptido sinal putativo, com um local de clivagem por proteases previsto indicado por "\*" : APLLG \* TGAPA (entre as posições de aminoácidos 25 e 26 da SEQ ID N°: 3).

A sequência polinucleotídica divulgada partilha alguma homologia com um gene de *Drosophila* conhecido como *Out at First* (oaf). A transcrição do oaf resulta em três classes de ARN alternativamente poliadenilados, cuja expressão é regulada do ponto de vista do desenvolvimento. Todas os transcritos do oaf contêm duas ORF adjacentes, separadas por um único codão de terminação UGA. A supressão do codão UGA durante a tradução pode

originar a produção de diferentes proteínas, a partir da mesma molécula de ARN. Durante a oogénese, o ARN do oaf é expresso em células foliculares de todas as idades e é contribuído para o ovo por via materna.

Durante o desenvolvimento embrionário, a transcrição zigótica do gene oaf ocorre em pequenos grupos de células na maioria ou na totalidade dos segmentos, no momento de extensão da banda germinal e, mais tarde, num padrão repetido segmentadamente no sistema nervoso central em desenvolvimento. O gene oaf é, também, expresso nas gónadas embrionárias, larvais e adultas de ambos os sexos. (Bergstrom, D. E., *et al.*, *Genetics* **139**:1331-1346, 1995.)

O polinucleótido foi expresso diferencialmente em sete pares de linhas celulares do cancro da mama altamente metastáticas em comparação com não-metastáticas ou pouco metastáticas. As linhas celulares utilizadas são MDA-MB-361 (obtida de adenocarcinoma da mama humano), MDA-MB-231 (células do cancro da mama humano, metastáticas para osso e/ou pulmão); MDA-MB-468 (obtida de células do cancro da mama humano, negativas para o receptor estrogénio); MCF-7 (células não-metastáticas do cancro da mama humano); ZR-75-1 (obtida de carcinomas da mama humanos positivos para o receptor do estrogénio, Engle *et al.*, *Cancer Res.* **38**:3352-64 (1978)); e MDA-MB-435 (obtida de células de carcinoma da mama humano, negativas para o receptor do estrogénio, Rishi *et al.*, *Cancer Res.* **56**; 5246-5252 (1996)).

O perfil de expressão é como se segue:

Tabela 1

<b>Par de Linhas Celulares</b>	<b>Proporção de Expressão</b>
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

A regulação positiva da expressão do ARNm foi confirmada por análise de transferência de Northern, utilizando ARN total das linhas celulares (Figura 5).

As linhas celulares nas quais foi comparada a expressão do polinucleótido, representam cancros da mama humanos com potencial metastático variável. As culturas da linha celular ZR-75-1 derivaram de efusão ascítica maligna de um doente de cancro da mama. As linhas celulares cultivadas *in vitro* assemelharam-se, intimamente, com a morfologia observada em biopsias ou em preparações celulares dos dadores das células originais. As células ZR-75-1 são especificamente estimuladas pelo estrogénio e têm sido utilizadas como um sistema modelar para estudar a sensibilidade ao estrogénio. Engel, L. W., et al., *Cancer Res.* **38**:3352-3364, 1978.

A linha celular MDA-MB-435 é uma linha celular negativa para o receptor do estrogénio que foi estudada como um modelo do cancro da mama humano, por exemplo, para estudar o mecanismo de acção da inibição do crescimento na presença de ácido retinóico.

Rishi, A. K., et al., *Cancer Res.* **56**:5246-5252, 1996. A inibição do crescimento por retinóides foi, também, estudada em células MCF-7 e células MDA MB 468. Tin-U, C. K., et al., *Am. Soc. Clin. Onc. Proceedings*, Vol. **17**, 2125, 1998.

A linha celular MDA-MB-361 derivou de um adenocarcinoma da mama humano, especificamente, de um local maligno. Número ATCC HTB-27. Foi estudada a expressão diferencial de genes humanos Wnt nesta linha celular. Huguet, E. L., et al., *Cancer Res.* **54**:2615-2621, 1994.

Quando ocorre uma metástase, as células tumorais primárias mamárias invadem as membranas basais e disseminam-se para outros órgãos do corpo e a possibilidade de sobrevivência dos doentes com cancro da mama torna-se bastante reduzida. É fundamental identificar genes que participam na invasão do cancro da mama e na metástase, no se refere ao diagnóstico clínico e à terapia. Esses genes são marcadores potenciais para diagnóstico ou são alvos candidatos para o desenvolvimento de fármacos terapêuticos. Por exemplo, a presença de vimentina no tumor da mama humano tem sido associada com a ausência de receptor de estrogénio e invasividade tumoral, como um marcador de prognóstico desfavorável. Raymond, W. A., et al., *J. Pathol.* **157**(4):299-306 (1989); Raymond, W. A., et al., *J. Pathol.* **158**(2): 107-14 (1989); e Thompson, E. W., et al., *J. Cell Physiol.* **150**(3):534-44 (1992). As actividades aumentadas das metaloproteinases da matriz estão relacionadas com o fenótipo metastático de carcinomas, especialmente, do cancro da mama. Basset, P., et al., *Nature* **348**(6303):6 699-704 (1990) e Basset, P., et al., *Cancer* **74**(Supl 3); 1045-9 (1994). A osteopontina, uma glicoproteína segregada de ligação à integrina, é capaz de induzir invasividade aumentada de células epiteliais mamárias

humanas e tem sido associada com malignidade intensificada no cancro da mama. Tuck, A. B., et al., *J. Cell Biochem.* **78**(3):465-75 (2000); Tuck, A. B., et al., *Oncogene* **18**(29):4237-46 (1999); e Singhal, H., et al., *Clin Cancer Res.* **3**(4):605-11 (1997).

É aqui divulgada a identificação de uma nova proteína segregada (hsOAF) envolvida nas metástases do cancro da mama. As linhas celulares do cancro da mama humano, utilizadas na elucidação do papel da proteína, são divididas em três grupos, de acordo com as suas capacidades metastáticas: altamente metastática, pouco metastática e não-metastática. Aproveitando os diferentes potenciais metastáticos dentro destes grupos de linhas celulares e utilizando tecnologia de *microarrays* avançada, foram identificados genes que são expressos diferencialmente entre linhas celulares do cancro da mama humano altamente metastáticas e pouco metastáticas e não-metastáticas. Estes genes são bons candidatos para a validação da expressão e estudos funcionais que possam originar uma melhor compreensão do mecanismo molecular da metástase do cancro da mama. O gene hsOAF é o foco desta invenção, dado que este codifica uma nova proteína segregada. A expressão do ARNm do hsOAF é bastante comum em vários tecidos humanos normais. No entanto, a secreção de proteína hsOAF é muito mais significativa em linhas celulares altamente metastáticas do cancro da mama humano do que nas pouco metastáticas/não-metastáticas e a MDA-MB-435 apresenta a maior secreção de hsOAF. O gene hsOAF foi denominado com base na sua homologia com o gene OAF (*out at first*) de *Drosophila*. No entanto, a proteína OAF de *Drosophila* pode não ser uma proteína segregada, uma vez que esta não possui uma sequência peptídica sinal típica, no terminal N (Fig. 7).

Para abordar a importância da proteína hsOAF segregada na metástase do cancro da mama, foi utilizada tecnologia de oligonucleótidos anti-sentido, para inativar, especificamente, a expressão do hsOAF. A tecnologia de oligonucleótidos anti-sentido é uma forma rápida, eficiente, para reduzir, drasticamente, a expressão génica, para estudos funcionais génicos. Stein, C. A., et al., *Science* **261**(5124):1004-12 (1993); Defacque, H., et al., *J. Cell Physiol.* **178**(1):109-19 (1999). A inativação da secreção da proteína hsOAF de células MDA-MB-435 altamente metastáticas resultou na alteração da forma celular, invasividade celular reduzida e proliferação celular mais lenta. O tratamento de células com meio condicionado (meio de cultura de células MDA-MB-435 normais) foi capaz de recuperar todas as alterações fenotípicas causadas pela inativação da secreção da proteína hsOAF, em algum grau. Embora os requerentes não estejam limitados a um mecanismo específico, pensa-se que a proteína hsOAF segregada é essencial para a invasividade e a proliferação de células MDA-MB-435. No entanto, a inativação da secreção da proteína hsOAF de outra linha celular altamente metastática, MDA-MB-231, pela tecnologia de oligonucleótidos anti-sentido, não causou qualquer modificação celular significativa. MDA-MB-435 e MDA-MB-231 são linhas celulares metastáticas bastante diferentes e a MDA-MB-435 apresenta uma secreção de proteína hsOAF muito mais significativa do que a MDA-MB-213.

O gene hsOAF está localizado na região q23 do cromossoma 11, onde a perda de heterozigotocidade ocorre frequentemente em tumores da mama humanos. Negrini, M., et al., *Cancer Res* **55**(14):3003-7 (1995) e Tomlinson, I. P., et al., *J. Clin. Pathol.* **48**(5):424-8 (1995). A perda de heterozigotocidade na 11q23 em tumores da mama humanos primários tem sido descrita como associada com a baixa sobrevivência após metástase.

Winqvist, R., et al., *Cancer Res.* **55**(12):2660-4 (1995). A 11q23 contém, também, loci, tais como ATM (Ataxia-telangiectasia, mutado) e MLL (o qual é frequentemente interrompido pelo rearranjo cromossômico na leucemia aguda). Rasio, D., et al., *Cancer Res.* **55**(24):6053-7 (1995) e Rubnitz, J. E., et al., *Leukemia* **10**(1):74-82 (1996). A relação entre a mutação no cromossoma 11q23 e expressão do gene hsOAF na metástase do cancro da mama permanece pouco evidente.

A proteína hsOAF segregada pode ser um alvo adequado para o desenvolvimento de fármacos contra o cancro da mama e um bom marcador para diagnóstico da malignidade do tumor da mama. A SEQ ID N°: 1 e os polinucleótidos compreendendo esta sequência são, consequentemente, úteis como marcadores metastáticos. A referência a sequências nucleotídicas aminoacídicas marcadoras metastáticoas inclui variantes que têm padrões de expressão semelhantes em células metastáticas elevadas, em relação a não-metastáticas ou pouco metastáticas. Os polipéptidos marcadores metastáticos podem diferir, em termos de comprimento, das proteínas marcadoras metastáticas completas e conter, pelo 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 ou 271 ou mais aminoácidos contíguos de uma proteína marcadora metastática. Os exemplos de polinucleótidos incluem os que codificam os aminoácidos de cerca de 1 a cerca de 273; de 1 a 273; de cerca de 2 a cerca de 273; de 2 a 273; de cerca de 26 a cerca de 273; e de 26 a 273 da SEQ ID N°: 2.

Podem, também, ocorrer variantes de proteínas e polipéptidos marcadores. As variantes de proteínas ou polipéptidos marcadores metastáticos podem ser de ocorrência natural ou não natural. As variantes de proteínas ou de

polipéptidos marcadores metastáticos de ocorrência natural são encontrados em humanos ou em outras espécies e compreendem sequências de aminoácidos que são substancialmente idênticas a uma proteína codificada por um gene correspondente à sequência nucleotídica apresentada na SEQ ID N°: 1 ou o seu complemento. São, também, aqui incluídas variantes de proteínas ou de polipéptidos marcadores metastáticos de ocorrência não-natural que conservam, substancialmente, os mesmos padrões de expressão diferenciais, em células do cancro da mama metastáticas elevadas em relação às pouco metastáticas ou às não-metastáticas, do que as proteínas marcadoras metastáticas de ocorrência natural ou as variantes do polipéptido. De um modo preferido, as variantes de proteínas ou polipéptidos marcadores metastáticos de ocorrência natural ou não-natural têm sequências de aminoácidos que são, pelo menos, 85%, 90%, 91%, 92%, 93%, 94% ou 95% idênticos a sequências de aminoácidos codificadas pela sequência nucleotídica apresentada na SEQ ID N°: 1. De um modo mais preferido, as moléculas são, pelo menos, 96%, 97%, 98% ou 99% idênticas. A identidade de sequência percentual entre uma proteína ou polipéptido do tipo selvagem e uma variante é determinada pelo alinhamento da proteína ou polipéptido do tipo selvagem com a variante, para obter o maior número de correspondências de aminoácidos, como é conhecido na técnica, contando o número de correspondências de aminoácido entre o tipo selvagem e a variante e dividindo o número total de correspondências pelo número total de resíduos de aminoácidos da sequência do tipo selvagem.

De um modo preferido, as alterações de aminoácidos nas variantes proteicas ou polipetídicas marcadoras metastáticas são modificações conservadoras de aminoácidos, *i. e.*, substituições de aminoácidos carregados ou não carregados de um modo



semelhante. Uma modificação conservadora de aminoácido envolve a substituição de um de uma família de aminoácidos, os quais estão relacionados em termos das suas cadeias laterais. Os aminoácidos de ocorrência natural são, geralmente, divididos em quatro famílias: aminoácidos acídicos (aspartato, glutamato), básicos (lisina, arginina, histidina), não-polares (alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptofano) e polares não carregados (glicina, asparagina, glutamina, cistina, serina, treonina, tirosina). A fenilalanina, triptofano e tirosina são, por vezes, classificados conjuntamente como aminoácidos aromáticos.

É razoável presumir que uma substituição isolada de uma leucina por uma isoleucina ou valina, um aspartato por um glutamato, uma treonina por uma serina ou uma substituição semelhante de um aminoácido por um aminoácido estruturalmente relacionado, não terá um efeito significativo sobre as propriedades biológicas da variante da proteína ou do polipéptido marcador metastático resultante. As propriedades e funções das variantes da proteína ou do polipéptido marcador metastático são do mesmo tipo de uma proteína ou polipéptido marcador metastático compreendendo sequências de aminoácidos codificadas pela sequência nucleotídica apresentada na SEQ ID N°: 1, embora as propriedades e as funções das variantes possam diferir no grau. Pode ser facilmente determinado se uma modificação de aminoácido resulta numa variante de proteína ou polipéptido marcador metastático com o padrão de expressão diferencial adequado. Por exemplo, podem ser seleccionadas sondas nucleotídicas a partir das sequências génicas marcadoras aqui divulgadas e serem utilizadas para detectar o ARNm do gene marcador, em transferências de Northern ou em secções de tecido, como é conhecido na técnica. Alternativamente, podem ser

utilizados anticorpos que se ligam, especificamente, a produtos proteicos de genes marcadores metastáticos, para detectar a expressão de proteínas marcadoras metastáticas ou suas variantes.

As variantes marcadoras metastáticas incluem formas glicosiladas, conjugados agregativos com outras moléculas e conjugados covalentes com partes químicas não relacionadas. As variantes marcadoras metastáticas incluem, também, variantes alélicas, variantes de espécie e muteínas. As truncagens ou as eliminações de regiões que não afectam a expressão diferencial de genes marcadores metastáticos são, também, variantes de marcadores metastáticos. Podem ser preparadas variantes covalentes, por ligação de funcionalidades a grupos que se encontram na cadeia de aminoácidos ou no resíduo do terminal N ou C, como é conhecido na técnica.

Será reconhecido na técnica que algumas sequências de aminoácidos do polipéptido da invenção podem ser variadas sem efeito significativo na estrutura ou função da proteína. Se forem contempladas essas diferenças na sequência, deve ser tido em consideração que existem áreas críticas na proteína as quais determinam a actividade. Em geral, é possível substituir resíduos que formam a estrutura terciária, contanto que sejam utilizados resíduos que realizem uma função semelhante. Em outros casos, o tipo do resíduo pode ser totalmente indiferente se a alteração ocorrer numa região não crítica da proteína. A substituição de aminoácidos pode, também, modificar a selectividade de ligação a receptores da superfície celular. Ostade *et al.*, *Nature* **361**:266-268 (1993) descrevem determinadas mutações que resultam em ligação selectiva do TNF-alfa a, apenas, um de dois tipos conhecidos de receptores do TNF. Deste

modo, os polipéptidos de presente invenção podem incluir uma ou mais substituições, eliminações ou adições de aminoácidos, por mutações naturais ou por manipulação humana.

São, também, divulgadas variações do polipéptido divulgado, as quais apresentam padrões de expressão comparáveis ou as quais incluem regiões antigénicas. Esses mutantes incluem eliminações, inserções, inversões, repetições e substituições de tipo. Em Bowie, J. U., et al., "Deciphering the Message in Protein Sequence: Tolerance to Amino Acid Substitutions," *Science* **247**:1306-1310 (1990) pode ser encontrada orientação sobre quais modificações de aminoácidos são prováveis de ser fenotipicamente silenciosas.

São de particular interesse as substituições de aminoácidos carregados por outros aminoácidos carregados e por aminoácidos neutros ou carregados negativamente. O último resulta em proteínas com carga positiva reduzida, para melhorar as características da proteína divulgada. É altamente desejável a prevenção da agregação. A agregação de proteínas não resulta, apenas, numa perda de actividade mas pode ser, também, problemática quando são preparadas formulações farmacêuticas, dado que estas podem ser imunogénicas. (Pinckard et al., *Clin. Exp. Immunol.* **2**:331-340 (1967); Robbins et al., *Diabetes* **36**:838-845 (1987); Cleland et al., *Crit. Rev. Therapeutic Drug Carrier Systems* **10**:307-377 (1993)).

Os aminoácidos nos polipéptidos divulgados que são essenciais para a função podem ser identificados por métodos conhecidos na técnica, tais como mutagénese dirigida para um local ou mutagénese de varrimento da alanina (Cunningham e Wells, *Science* **244**:1081-1085 (1989)). O último processo introduz

mutações de alanina única em cada resíduo na molécula. As moléculas mutantes resultantes são, então, testadas para a actividade biológica, tal como ligação ao receptor ou actividade proliferativa *in vitro*. Os locais que são críticos para ligação ligando-receptor podem ser, também, determinados por análise estrutural, tais como, cristalização, ressonância magnética nuclear ou marcação por fotoafinidade (Smith *et al.*, *J. Mol. Biol.* **224**:899-904 (1992) e de Vos *et al.* *Science* **255**:306-312 (1992)).

Como indicado, as modificações são, de um modo preferido, de uma natureza menor, tais como substituições de aminoácido conservadoras que não afectam, significativamente, o enrolamento ou a actividade da proteína. Obviamente, o número de substituições de aminoácido que um especialista na técnica irá realizar depende de muitos factores, incluindo os descritos acima. Em geral, o número de substituições de qualquer polipéptido determinado não será superior a 50, 40, 30, 25, 20, 15, 10, 5 ou 3.

As proteínas marcadoras metastáticas de tamanho total podem ser extraídas, utilizando métodos bioquímicos padrão, de células humanas produtoras de proteínas marcadoras metastáticas, tais como células do cancro da mama metastáticas. Uma proteína ou polipéptido marcador metastático isolado e purificado está separado de outros compostos que estão, normalmente, associados com uma proteína ou polipéptido marcador metastático numa célula, tais como, algumas proteínas, hidratos de carbono, lípidos ou organelos subcelulares. Uma preparação de proteínas ou polipéptidos marcadores metastáticos isolados e purificados está, pelo menos, pura a 80%; de um modo preferido, as preparações estão puras a 90%, 95% ou 99%.

Um gene humano codificando a SEQ ID N°: 2 pode ser identificado e isolado utilizando métodos conhecidos na técnica. De acordo com um método, a SEQ ID N°: 1 é preparada num formato legível em computador. A sequência é comparada com sequências polinucleotídicas de um genoma humano e são identificadas uma ou mais sequências do genoma humano, tendo identidade de sequência de, pelo menos, 95% com a SEQ ID N°: 1, utilizando, por exemplo, o algoritmo de Smith-Waterman, utilizando uma pesquisa de espaços relacionada, com uma penalização para a abertura de espaços de 12 e uma penalização para a extensão de espaço de 1, como parâmetros. São preparadas sondas baseadas nas regiões de homologia entre a SEQ ID N°: 1 e sequências do genoma humano e são utilizadas para isolar polinucleótidos do ADN genómico humano, utilizando métodos conhecidos na técnica. Até à data do pedido não foi identificado qualquer polinucleótido humano correspondente ao polinucleótido completo da SEQ ID N°: 1 nas bases de dados públicas. Deste modo, a invenção inclui ADN genómico humano compreendendo a região codificante da SEQ ID N°: 1 e qualquer região não traduzida, a qual não partilha homologia com a SEQ ID N°: 1 mas é contígua a regiões homólogas. Esse ADN de genómico inclui mas não está limitado a intrões, promotores e outras regiões reguladoras, funcionalmente associadas com um gene humano, tendo uma região codificando a SEQ ID N°: 2.

Podem ser, também, produzidas proteínas e polipéptidos marcadores metastáticos por métodos de ADN recombinante ou por métodos químicos sintéticos. Para a produção de proteínas ou polipéptidos marcadores metastáticos recombinantes, podem ser expressas sequências codificantes, seleccionadas das sequências nucleotídicas apresentadas na SEQ ID N°: 1 ou variantes dessas

sequências, as quais codificam proteínas marcadoras metastáticas, em sistemas de expressão procariotas ou eucariotas conhecidos (ver abaixo). Podem ser utilizados sistemas de expressão bacterianos, de leveduras, de insectos ou mamíferos, como é conhecido na técnica.

Alternativamente, podem ser utilizados métodos de síntese química, tal como síntese de péptidos em fase sólida, para sintetizar uma proteína ou polipéptido marcador metastático. As formas gerais de produção de péptidos, análogos ou derivados são descritos em *CHEMISTRY AND BIOCHEMISTRY OF AMINO ACIDS, PEPTIDES AND PROTEINS - A SURVEY OF RECENT DEVELOPMENTS*, Weinstein, B. ed., Marcell Dekker, Inc., publ., Nova Iorque (1983). Além disso, pode ser realizada a substituição de D-aminoácidos pelo L-estereoisómero normal, para aumentar a meia-vida da molécula. Podem ser produzidas do mesmo modo variantes marcadores metastáticos.

Podem ser, também, construídas proteínas de fusão de ocorrência não natural, compreendendo, pelo 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 ou 271 ou mais aminoácidos marcadores metastáticos contíguos. As proteínas de fusão marcadoras metastáticas humanas são úteis para produzir anticorpos contra sequências de aminoácidos marcadores metastáticos e para utilização em vários sistemas de ensaio. Por exemplo, as proteínas de fusão marcadoras metastáticas podem ser utilizadas para identificar as proteínas que interagem com proteínas marcadoras metastáticas e influenciam as suas funções. Podem ser, também, utilizados métodos físicos, tais como cromatografia de proteínas de afinidade ou ensaios baseados em bibliotecas, para interacções

proteína-proteína, tais como os sistemas de dois híbridos de levedura ou de apresentação de fagos, com esta finalidade. Esses métodos são bem conhecidos na técnica e podem ser, também, utilizados em pesquisa de fármacos.

Uma proteína de fusão marcadora metastática compreende dois segmentos de proteína fundidos em conjunto através de uma ligação peptídica. O primeiro segmento de proteína compreende, pelo 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 ou 271 ou mais aminoácidos contíguos de uma proteína marcadora metastática. Os aminoácidos podem ser seleccionados das sequências de aminoácidos codificadas pela sequência nucleotídica apresentada na SEQ ID N°: 1 ou de variantes da sequência, tais como as descritos acima. O primeiro segmento de proteína pode, também, compreender uma proteína marcadora metastática de tamanho total.

Numa forma de realização preferida, o primeiro segmento de proteína compreende o polipéptido apresentado na SEQ ID N°: 2. Numa variação desta forma de realização, o primeiro segmento de proteína consiste nos aminoácidos 31-287 da SEQ ID N°: 2. Esta proteína de fusão não possui o péptido sinal da SEQ ID N°: 2 e será adequada para a retenção da proteína de fusão expressa dentro da célula.

O segundo segmento de proteína pode ser uma proteína completa ou um fragmento de proteína ou polipéptido não encontrado adjacente ao primeiro segmento de proteína na proteína nativa codificada pela SEQ ID N°: 1. A proteína de fusão pode ser marcada com um marcador detectável, como é conhecido na técnica, tal como um marcador radioactivo,

fluorescente, quimioluminescente ou biotinilado. O segundo segmento de proteína pode ser uma enzima, a qual irá produzir um produto detectável, tal como  $\beta$ -galactosidase. O primeiro segmento de proteína pode ser N-terminal ou C-terminal, como for conveniente.

As técnicas para preparar proteínas de fusão, recombinantemente ou por ligação covalente de dois segmentos de proteína, são, também, bem conhecidas. Podem ser utilizados métodos de ADN recombinante para preparar proteínas de fusão marcador metastático, por exemplo, preparando uma construção de ADN que compreende sequências codificantes da SEQ ID N°: 1 na grelha de leitura correcta, com nucleótidos codificando o segundo segmento da proteína e expressando a construção de ADN numa célula hospedeira, como descrito abaixo. Na Figura 4 é apresentada a grelha de leitura aberta da SEQ ID N°: 1.

As proteínas, polipéptidos, variantes ou proteínas de fusão marcadoras metastáticas isoladas e purificadas, podem ser utilizadas como imunogénios, para obter preparações de anticorpos que se ligam, especificamente, a uma proteína marcadora metastática. Os anticorpos podem ser utilizados, *inter alia*, para detectar proteínas marcadoras metastáticas do tipo selvagem em tecidos humanos e suas fracções. Os anticorpos podem ser, também, utilizados para detectar a presença de mutações em genes marcadores metastáticos, as quais resultem em sub ou sobreexpressão de uma proteína marcadora metastática ou na expressão de uma proteína marcadora metastática com tamanho ou mobilidade electroforética alterada.

Podem ser realizadas preparações de anticorpos policlonais ou monoclonais, utilizando métodos padrão. Podem ser, também,



preparados anticorpos de cadeia única. Um imunogénio preferido é um polipéptido compreendendo a SEQ ID N°: 2. Podem ser isolados anticorpos de cadeia única que se ligam, especificamente, a proteínas, polipéptidos, variantes ou proteínas de fusão marcador metastático, por exemplo, a partir de bibliotecas de apresentação de imunoglobulina de cadeia única, como é conhecido na técnica. A biblioteca é “rastreada” contra sequências de aminoácidos de proteínas marcadoras metastáticas da SEQ ID N°: 2 e podem ser isolados vários anticorpos de cadeia única que se ligam com elevada afinidade a diferentes epitopos de proteínas marcadoras metastáticas. Hayashi *et al.*, 1995, *Gene* **160**:129-30. Podem ser, também, construídos anticorpos de cadeia única, utilizando um método de amplificação de ADN, tal como a reacção em cadeia da polimerase (PCR), utilizando ADNc de hibridoma como molde. Thirion *et al.*, 1996, *Eur. J. Cancer Prev.* **5**:507-11.

Os anticorpos específicos para os marcadores metastáticos ligam-se, especificamente, a epitopos presentes numa proteína marcadora metastática de tamanho total, tendo uma sequência de aminoácidos codificada por uma sequência nucleotídica apresentada na SEQ ID N°: 1, a polipéptidos marcadores metastático ou a variantes marcadoras metastáticas, quer isoladamente, quer como parte de uma proteína de fusão. De um modo preferido, os epitopos marcadores metastáticos não estão presentes em outras proteínas humanas. Tipicamente, são necessários, pelo menos, 6, 8, 10 ou 12 aminoácidos contíguos para formar um epitopo. No entanto, os epitopos que envolvem aminoácidos não-contíguos podem requerer mais, e. g., pelo menos, 15, 25 ou 50 aminoácidos.

Os anticorpos que se ligam especificamente a proteínas, polipéptidos, proteínas de fusão ou variantes marcadoras

metastáticas, proporcionam um sinal de detecção, pelo menos, 5, 10 ou 20 vezes mais elevado do que um sinal de detecção proporcionado com outras proteínas, quando utilizados em transferências de Western ou outros ensaios imunoquímicos. De um modo preferido, os anticorpos que se ligam especificamente a epitopos de marcadores metastáticos não detectam outras proteínas em ensaios de imunoquímicos e podem imunoprecipitar uma proteína, polipéptido, proteína de fusão ou variante marcadora metastática a partir de uma solução.

Os polinucleótidos subgenómicos contêm menos de um cromossoma completo. De um modo preferido, os polinucleótidos são isentos de intrões. Numa forma de realização preferida, as moléculas polinucleotídicas compreendem uma sequência contígua 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 ou 2350 nucleótidos da SEQ ID N°: 1 ou os seus complementos. O complemento de uma sequência nucleotídica apresentada na SEQ ID N°: 1 é uma sequência nucleotídica contígua, a qual forma pares de bases de Watson-Crick com uma sequência nucleotídica contígua apresentada na SEQ ID N°: 1.

As sequências nucleotídicas degeneradas codificando as sequências de aminoácidos da proteína ou variantes marcadoras metastáticas, assim como sequências nucleotídicas homólogas, que compreendem um polinucleótido idêntico a, pelo menos, 85% 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ou 99%, à região codificante da sequência nucleotídica apresentada na SEQ ID N°: 1, são, também, polinucleótidos marcadores metastáticos subgenómicos. Tipicamente, as sequências polinucleotídicas marcadoras metastáticas homólogas subgenómicas

podem ser confirmadas por hibridação sob condições restritivas, como é conhecido na técnica. A identidade de sequência percentual entre sequências do tipo selvagem e variantes homólogas é determinada pelo alinhamento do polinucleótido do tipo selvagem com a variante, para obter o maior número de correspondências de nucleótidos, como é conhecido na técnica, contando o número de correspondências de nucleótidos entre o tipo selvagem e a variante e dividindo o número total de correspondências pelo número total de nucleótidos da sequência do tipo selvagem. Um algoritmo preferido para calcular a identidade percentual é o algoritmo de pesquisa de homologia de Smith-Waterman, como implementado no programa MPSRCH (Oxford Molecular), utilizando uma pesquisa de espaços relacionada com os seguintes parâmetros de pesquisa: penalização de abertura de espaços de 12 e penalização de extensão de espaços de 1.

Pode ser utilizado um polinucleótido marcador metastático subgenómico, compreendendo sequências codificando uma proteína marcadora metastática, numa construção de expressão. De um modo preferido, o polinucleótido marcador metastático subgenómico é inserido num plasmídeo de expressão (por exemplo, no sistema Ecdyson, pIND, In vitro Gene). Os polinucleótidos marcadores metastáticos subgenómicos podem ser propagados em vectores e linhas celulares, utilizando técnicas bem conhecidas na técnica. Os polinucleótidos marcadores metastáticos subgenómicos podem situar-se em moléculas lineares ou circulares. Estes podem situar-se em moléculas de replicação autónoma ou em moléculas sem sequências de replicação. Estes podem ser autoregulados ou por outras sequências reguladoras, como é conhecido na técnica.

Uma célula hospedeira compreendendo uma construção de expressão de um marcador metastático pode ser, então, utilizada

para expressar a totalidade ou uma porção de uma proteína marcadora metastática. As células hospedeiras compreendendo construções de expressão de marcadores metastáticos podem ser procariotas ou eucariotas. Estão disponíveis várias células hospedeiras para utilização em sistemas de expressão bacterianos, de levedura, de insecto e humanos e podem ser utilizadas para expressar ou propagar construções de expressão de marcadores metastáticos (ver abaixo). As construções de expressão podem ser introduzidas em células hospedeiras, utilizando qualquer técnica conhecida na técnica. Estas técnicas incluem transferência de ADN mediada pelo polícatião de transferrina, transfecção com ácidos nucleicos nus ou encapsulados, fusão celular mediada por lipossomas, transporte intra-celular de esferas de látex revestidas por ADN, fusão de protoplastos, infecção viral, electroporação e transfecção mediada por fosfato de cálcio.

Uma construção de expressão de marcadores metastáticos compreende um promotor que é funcional numa célula hospedeira seleccionada. Um especialista na técnica pode seleccionar, facilmente, um promotor adequado, de um grande número de promotores específicos para células-tipo, conhecidos e utilizados na técnica. A construção de expressão pode, também, conter um terminador da transcrição que é funcional na célula hospedeira. A construção de expressão compreende um segmento polinucleotídico que codifica a totalidade ou uma porção da proteína, variante, proteína de fusão, anticorpo ou ribozima marcadora metastática. O segmento polinucleotídico está localizado a jusante do promotor. A transcrição do segmento polinucleotídico inicia no promotor. A construção de expressão pode ser linear ou circular e conter sequências para a replicação autónoma, se pretendido.

Os sistemas bacterianos para expressar construções de expressão de marcadores metastáticos incluem os descritos em Chang *et al.*, *Nature* (1978) **275**:615, Goeddel *et al.*, *Nature* (1979) **281**:544, Goeddel *et al.*, *Nucleic Acids Res.* (1980) **8**:4057, documento EP 36776, Patente U. S. N°. 4551433, de Boer *et al.*, *Proc. Nat'l Acad Sci. USA* (1983) **80**:21-25 e Siebenlist *et al.* *Cell* (1980) **20**:269.

Os sistemas de expressão em leveduras incluem os descritos em Hinnen *et al.*, *Proc. Nat'l Acad. Sci. USA* (1978) **75**:1929; Ito *et al.*, *J Bacteriol.* (1983) **153**:163; Kurtz *et al.*, *Mol. Cell. Biol.* (1986) **6**:142; Kunze *et al.*, *J. Basic Microbiol.* (1985) **25**:141; Gleeson *et al.*, *J. Gen. Microbiol.* (1986) **132**:3459, Roggenkamp *et al.*, *Mol. Gen. Genet.* (1986) **202**:302) Das *et al.*, *J. Bacteriol.* (1984) **158**:1165; De Louvencourt *et al.*, *J. Bacteriol.* (1983) **154**:737, Van den Berg *et al.*, *Bio/Technology* (1990) **8**:135; Kunze *et al.*, *J. Basic Microbiol.* (1985) **25**:141; Cregg *et al.*, *Mol. Cell. Biol.* (1985) **5**:3376, Patente U.S. N°: 4837148, Patente U.S. N°: 4929555; Beach e Nurse, *Nature* (1981) **300**:706; Davidow *et al.*, *Curr. Genet.* (1985) **10**:380, Gaillardin *et al.*, *Curr. Genet.* (1985) **10**:49, Ballance *et al.*, *Biochem. Biophys. Res. Commun.* (1983) **112**:284-289; Tilburn *et al.*, *Gene* (1983) **26**:205-221, Yelton *et al.*, *Proc. Nat'l Acad. Sci. USA* (1984) **81**:1470-1474, Kelly e Hynes, *EMBO J.* (1985) **4**:475479; documento EP 244234 e documento WO 91/00357.

A expressão da construções de expressão de marcadores metastáticos em insectos pode ser realizada como descrito na Patente U.S. N°: 4745051, Friesen *et al.* (1986) "The Regulation of the Baculovirus Gene Expression" em: *THE MOLECULAR BIOLOGY of BACULOVIRUSES* (W. Doerfler, ed.), documento EP 127839, documento

EP 155476 e Vlak et al., *J. Gen. Virol.* (1988) **69**:765-776, Miller et al., *Ann. Rev. Microbiol.* (1988) **42**:177, Carbonell et al., *Gene* (1988) **73**:409, Maeda et al., *Nature* (1985) **315**:592-594, Lebacqz-Verheyden et al., *Mol. Cell. Biol.* (1988) **8**:3129; Smith et al., *Proc. Nat'l Acad. Sci. USA* (1985) **82**:8404, Miyajima et al., *Gene* (1987) **58**:273; e Martin et al., *DNA* (1988) **7**:99. Em Luckow et al., *Bio/Technology* (1988) **6**:47-55, Miller et al., em *GENETIC ENGINEERING* (Setlow, J. K. et al. ed.), volume 8 (Plenum Publishing, 1986), pp, 277-279 e Maeda et al., *Nature*, (1985) **315**:592-594 são descritas numerosas estirpes e variantes de baculovírus e células hospedeiras de insecto permissivas correspondentes, de hospedeiros.

A expressão de construções de expressão de marcadores metastáticos em mamíferos pode ser realizada como descrito em Dijkema et al., *EMBO J.* (1985) **4**:761, Gorman et al., *Proc. Nat'l Acad. Sci. USA* (1982b) **79**:6777, Boshart et al., *Cell* (1985) **41**:521 e Patente U. S. N°: 4399216. Podem ser facilitadaa outras características da expressão de construções de expressão de marcadores metastáticos em mamíferos, como descrito em Ham e Wallace, *Meth. Enz.* (1979) **58**:44, Barnes e Sato, *Anal. Biochem.* (1980) **102**:255, Patente U. S. N°: 4767704, Patente U.S. N°: 4657866, Patente U.S. N°: 4927762, Patente U.S. N°: 4560655, documento WO 90/103430, documento WO 87/00195 e documento U.S. RE 30985.

Os polinucleótidos subgenómicos da invenção podem ser, também, utilizados em veículos de transferência de genes, com o objectivo de distribuir um ARNm ou oligonucleótido marcador metastático (com a sequência do ARNm marcador metastático nativo ou do seu complemento), proteína marcadora metastática completa, proteína de fusão marcadora metastática, polipéptido marcador

metastático ou ribozima ou anticorpo de cadeia única, específicos para um marcador metastático, numa célula, de um modo preferido, numa célula eucariota. De acordo com a presente invenção, um veículo de transferência de genes pode ser, por exemplo, ADN plasmídico nu, um vector de expressão viral compreendendo um polinucleótido marcador metastático subgenómico ou um polinucleótido marcador metastático subgenómico, em conjunto com um lipossoma ou um agente condensador.

É, também, divulgado um método de detecção da expressão de um gene marcador metastático numa amostra biológica. A detecção da expressão do gene marcador metastático é útil, por exemplo, para identificar metástases ou para determinar o potencial metastático numa amostra de tecido, de um modo preferido, num tumor. Podem ser, então, concebidos regimes de tratamento adequados para doentes que estão em risco de desenvolver cancros metastáticos em outros órgãos do corpo.

A amostra corporal pode ser, por exemplo, um tecido sólido ou uma amostra fluida. O polipéptido nativo codificado pela SEQ ID N°: 1 é uma proteína putativa segregada e provável de ser detectada em fluidos corporais, incluindo sangue e fluido linfático, particularmente, os que drenam a partir de localizações tumorais no corpo. A proteína ou os produtos de expressão de ácidos nucleicos podem ser detectados na amostra corporal. Numa forma de realização, a amostra corporal é ensaiada para a presença de uma proteína marcadora metastática. Uma proteína marcadora metastática compreende uma sequência codificada por uma sequência nucleotídica apresentada na SEQ ID N°: 1 ou o seu complemento e pode ser detectada utilizando anticorpos específicos para a proteína marcador da presente invenção. Os anticorpos podem ser marcados, por exemplo, com uma

marcação radioactiva, fluorescente, biotinilada ou enzimática e podem ser detectados directamente ou podem ser detectados utilizando métodos imunoquímicos indirectos, utilizando um anticorpo secundário marcado. A presença das proteínas marcadoras metastáticas pode ser ensaiada, por exemplo, em secções de tecido, por imunocitoquímica ou em lisados, utilizando, transferência de Western, como é conhecido na técnica. A presença da proteína marcadora é indicadora de que a amostra de tecido é metastática.

Noutra forma de realização, a amostra corporal é ensaiada para a presença de ARNm da proteína marcadora. Uma amostra pode ser feita contactar com uma sonda de hibridação de ácido nucleico capaz de hibridar com o ARNm correspondente ao polipéptido seleccionado. Ainda além disso, a amostra pode ser submetida a uma técnica de transferência de Northern para detectar ARNm, indicando a expressão do polipéptido. Para as técnicas em que é detectado ARNm, a amostra pode ser submetida a um processo de amplificação de ácidos nucleicos pelo qual a molécula de ARNm ou uma sua parte seleccionada é amplificada, utilizando iniciadores nucleotídicos adequados. Podem ser, também, utilizadas outras técnicas de detecção de ARN, incluindo, mas não limitadas a, hibridação *in situ*.

Podem ser produzidas sondas específicas para proteínas marcadoras, utilizando a sequência de ADNc divulgada na SEQ ID N°: 1. As sondas têm, de um modo preferido, pelo menos, 15 a 50 nucleótidos de comprimento, embora estas possam ter, pelo menos, 8, 10, 11, 12, 20, 25, 30, 35, 40, 45, 60, 75 ou 100 ou mais, nucleótidos de comprimento. Uma região preferida para seleccionar sondas situa-se dentro das posições nucleotídicas 446-1173 da SEQ ID N°: 1. As sondas podem ser sintetizadas



quimicamente ou serem produzidas a partir de polinucleótidos mais longos, utilizando enzimas de restrição. As sondas podem ser marcadas, por exemplo, com uma marcação radioactiva, biotinilada ou fluorescente.

Opcionalmente, pode ser quantificado o nível de um produto de expressão marcador metastático particular numa amostra corporal. A quantificação pode ser realizada, por exemplo, comparando o nível do produto de expressão detectado na amostra corporal com as quantidades de produto presentes numa curva padrão. A comparação pode ser realizada visualmente ou utilizando uma técnica, tal como densitometria, com ou sem assistência computadorizada. Para utilização como controlos, as amostras corporais podem ser isoladas de outros humanos, outros órgãos não-cancerosos do doente a ser testado ou cancro da mama não-metastático do doente a ser testado. Como indicado pelos presentes resultados, a expressão da SEQ ID N°: 1 em células do cancro da mama pouco metastático ou não-metastático situa-se entre 3% e 44% dos níveis de expressão em células do cancro da mama altamente-metastático. Se a expressão numa amostra de teste for, pelo menos, 2 vezes superior à de numa amostra de controlo adequada, isto é indicador de células metastáticas.

Os polinucleótidos codificando reagentes específicos para marcadores metastáticos da invenção, tais como anticorpos e sondas nucleotídicas, podem ser proporcionados num kit para detectar produtos de expressão do gene marcador numa amostra biológica. O kit pode, também, conter tampões ou componentes de marcação, assim como instruções de utilização dos reagentes para detectar os produtos de expressão marcadores na amostra biológica.

A expressão de um gene marcador metastático pode ser alterada, utilizando uma sequência oligonucleotídica anti-sentido. A sequência anti-sentido é complementar a, pelo menos, uma porção da sequência codificante (nucleótidos 365-1173) de um gene marcador metastático tendo uma sequência nucleotídica apresentada na SEQ ID N°: 1. De um modo preferido, a sequência oligonucleotídica anti-sentido tem, pelo menos, seis nucleótidos de comprimento, mas pode ter, pelo menos, cerca de 8, 12, 15, 20, 25, 30, 35, 40, 45 ou 50 nucleótidos de comprimento. Podem ser, também, utilizadas sequências mais longas. As moléculas oligonucleotídicas anti-sentido podem ser proporcionadas numa construção de ADN e serem introduzidas em células cuja divisão deverá ser reduzida. Essas células incluem células do cancro da mama altamente metastáticas.

Os oligonucleótidos anti-sentido podem compreender desoxirribonucleótidos, ribonucleótidos ou uma combinação de ambos. Os oligonucleótidos podem ser sintetizados manualmente ou por um sintetizador automatizado, por ligação covalente da extremidade 5' de um nucleótido com a extremidade 3' de outro nucleótido com ligações internucleótido não-fosfodiéster, tais como, alquilfosfonatos, fosforotioatos, fosforoditioatos, alquilfosfonotioatos, alquilfosfonatos, fosforamidatos, ésteres de fosfato, carbamatos, acetamidato, ésteres carboximéticos, carbonatos e triésteres de fosfato. Ver Brown, 1994, *Meth. Mol. Biol.* **20**:1-8; Sonveaux, 1994, *Meth. Mol. Biol.* **26**:1-72; Uhlmann *et al.*, 1990 *Chem. Rev.* **90**:543-583.

Os anticorpos aqui divulgados que se ligam, especificamente, a uma proteína marcadora metastática podem ser, também, utilizados para alterar a expressão de um gene marcador metastático. Por anticorpos pretende-se significar anticorpos e

suas partes ou derivados, tais como, anticorpos de cadeia única, que conservam a ligação específica para a proteína. Os anticorpos específicos ligam-se a proteínas marcadoras metastáticas e impedem as proteínas de funcionar na célula. Os polinucleótidos codificando os anticorpos específicos da invenção podem ser introduzidos em células, como descrito acima.

As proteínas marcadoras aqui divulgadas podem ser utilizadas no rastreamento de fármacos que têm um efeito anti-metastático terapêutico. O efeito de um composto de teste sobre a síntese de proteínas marcadoras metastáticas pode ser, também, utilizado para identificar compostos de teste que modulam a metástase. Os compostos de teste que podem ser submetidos a rastreamento incluem qualquer substância, seja produto natural ou sintético, que possa ser administrada ao indivíduo. Podem ser testadas bibliotecas ou misturas de compostos. Os compostos ou as substâncias podem ser aquelas para as quais é anteriormente conhecido ou desconhecido um efeito farmacêuticos.

A síntese de proteínas marcadoras metastáticas pode ser medida por qualquer forma de medição da síntese proteica conhecida na técnica, tais como incorporação de aminoácidos marcados em proteínas e detecção de proteínas marcadoras metastáticas marcadas, num gel de poliacrilamida. A quantidade de proteínas marcadoras metastáticas pode ser detectada, por exemplo, utilizando anticorpos específicos para a proteína marcadora metastática da invenção em transferências de Western. A quantidade de proteínas marcadoras metastáticas sintetizadas na presença ou na ausência de um composto de teste pode ser determinada por qualquer meio conhecido na técnica, tal como a comparação da quantidade da proteína marcadora metastática

sintetizada, com a quantidade de proteínas marcadoras metastáticas presentes numa curva padrão.

O efeito de um composto de teste na síntese da uma proteína marcadora metastática pode ser, também, medido por análise de transferência de Northern, medindo a quantidade de expressão de ARNm da proteína marcadora metastática, em resposta ao composto de teste, utilizando as sondas nucleotídicas específicas para a proteína marcadora metastática da invenção, como é conhecido na técnica.

Tipicamente, uma amostra biológica é feita contactar com uma gama de concentrações do composto de teste, tais como 1,0 nM, 5,0 nM, 10 nM, 50 nM, 100 nM, 500 nM, 1 mM, 10 mM, 50 mM e 100 mM. De um modo preferido, o composto de teste aumenta ou reduz a expressão de uma proteína marcadora metastática em 60%, 75% ou 80%. De um modo mais preferido, é atingido um aumento ou uma redução de 85%, 90%, 95% ou 98%.

São, adicionalmente, divulgadas composições para aumentar ou reduzir a expressão da proteína marcadora metastática. Estas composições compreendem polinucleótidos codificando a totalidade ou, pelo menos, uma porção de um produto de expressão de um gene de uma proteína marcadora metastática. De um modo preferido, a composição terapêutica contém uma construção de expressão, compreendendo um promotor e um segmento polinucleotídico codificando, pelo menos, uma porção da proteína marcadora metastática, a qual é eficaz na redução do potencial metastático. As porções de genes ou proteínas marcadoras metastáticas que são eficazes na redução do potencial metastático de uma célula podem ser determinadas, por exemplo, introduzindo as porções dos genes ou polipéptidos marcadores

metastáticos em linhas celulares metastáticas, tais como MDA-MB-231, MDA-MB-435, Kml2C ou Kml2L4 e ensaiando a taxa de divisão das células ou a capacidade das células para formarem metástases, quando implantadas *in vivo*, como é conhecido na técnica. Podem ser utilizadas linhas celulares não-metastáticas, tais como MCF-7, para ensaiar a capacidade de uma porção de uma proteína marcadora metastática para aumentar a expressão de um gene marcador metastático.

Tipicamente, uma composição terapêutica de marcador metastático é preparada sob a forma de um injectável, como uma solução líquida ou em suspensão; no entanto, podem ser, também, preparadas formas sólidas adequadas para solução ou suspensão em veículos líquidos, antes da injeção. Uma composição de marcador metastático pode ser, também, formulada num comprimido ou cápsula de gel com revestimento entérico, de acordo com os métodos conhecidos na técnica, tais como os descritos na Patente U.S. Nº: 4853230, no documento EP 225189, no documento AU 9224296 e documento AU 9230801.

A administração dos agentes terapêuticos marcadores metastáticos pode incluir a administração local ou sistémica, incluindo injeção, administração oral, acelerador de partículas ou administração cateterizada e administração tópica. Podem ser utilizados vários métodos para administrar uma composição terapêutica de marcador metastático, directamente, a um local específico no corpo.

Para o tratamento de tumores, incluindo lesões metastáticas, pode ser, por exemplo, injectada uma composição terapêutica de marcador metastático, várias vezes, em várias localizações diferentes, dentro do corpo do tumor.

Alternativamente, podem ser identificadas as artérias que abastecem um tumor e ser injectada uma composição terapêutica nessa artéria, para distribuir a composição, directamente, no tumor.

Um tumor que tem um centro necrótico pode ser aspirado e a composição injectada, directamente, no centro do tumor então vazio. Uma composição terapêutica de marcador metastático pode ser administrada, directamente, à superfície de um tumor, por exemplo, por aplicação tópica da composição. Pode ser utilizada visualização de raios-X, para auxiliar em alguns dos métodos de distribuição acima. Pode ser administrada, simultaneamente ou sequencialmente, uma combinação de agentes terapêuticos, incluindo proteínas ou polipéptidos marcadores metastáticos ou um polinucleótido marcador metastático subgenómico e outros agentes terapêuticos.

Alternativamente, pode ser introduzida uma composição terapêutica de marcador metastático em células humanas *ex vivo* e as células serem, então, substituídas no humano. As células podem ser removidas a partir de várias localizações, incluindo, por exemplo, um tumor seleccionado ou de um órgão afectado. Além disso, pode ser inserida uma composição terapêutica em células não-afectadas, por exemplo, fibroblastos dérmicos ou leucócitos do sangue periférico. Se pretendido, podem ser, também, especificamente removidas do sangue, fracções particulares de células, tais como um subconjunto de células T ou células estaminais (ver, por exemplo, documento PCT WO 91/16116). As células removidas podem ser, então, feitas contactar com uma composição terapêutica de marcador metastático, utilizando qualquer das técnicas descritas acima, seguidas pelo regresso das células ao humano, de um modo preferido, à vizinhança de um

tumor ou dentro desta ou a outro local a ser tratado. Os métodos descritos acima podem compreender, adicionalmente, os passos de empobrecimento de fibroblastos ou de outras células tumorais não-contaminantes subsequentes à remoção de células tumorais de um humano e/ou o passo de inactivação das células, por exemplo, por irradiação.

Tanto a dose de uma composição de marcador metastático, como os meios da administração podem ser determinados com base nas qualidades específicas da composição terapêutica, estado, idade e peso do doente, progressão da doença e outros factores relevantes. De um modo preferido, uma composição terapêutica da invenção reduz a expressão dos genes marcadores metastáticos em 50%, 60%, 70% ou 80%. De um modo mais preferido, a expressão dos genes marcadores metastáticos é reduzida em 90%, 95%, 99% ou 100%. A eficácia do mecanismo seleccionado para alterar a expressão dos genes marcadores metastáticos pode ser avaliada utilizando métodos bem conhecidos na técnica, tais como a hibridação de sondas nucleotídicas com ARNm dos genes marcadores metastáticos, RT-PCR quantitativo ou detecção das proteínas marcadoras metastáticas, utilizando os anticorpos específicos da invenção.

Se a composição contiver as proteínas o polipéptido ou o anticorpo marcadoras metastáticas as dosagens eficazes da composição estão na gama de cerca de 5 µg a cerca de 50 µg/kg de peso corporal do doente cerca de 50 µg a cerca de 5 mg/kg cerca de 100 µg a cerca de 500 µg/kg do peso corporal do doente e cerca de 200 a cerca de 250 µg/kg.

As composições terapêuticas contendo polinucleótidos marcadores metastáticos subgenómicos podem ser administradas

numa gama de cerca de 100 ng a cerca de 200 mg de ADN para administração local. Podem ser, também, utilizadas gamas de concentração de cerca de 500 ng a cerca de 50 mg cerca de 1 µg a cerca de 2 mg cerca de 5 µg a cerca de 500 µg e cerca de 20 µg a cerca de 100 µg de ADN, durante um protocolo de terapia génica. Factores, tais como, o método de acção e a eficácia de transformação e da expressão são considerações que irão afectar a dosagem necessária, para uma maior eficácia dos polinucleótidos marcadores metastáticos subgenómicos. Quando é pretendida uma maior expressão ao longo de uma maior área de tecido, podem ser necessárias maiores quantidades de polinucleótidos marcadores metastáticos subgenómicos ou as mesmas quantidades, readministradas num protocolo de administrações sucessivas ou várias administrações em diferentes porções de tecido adjacentes ou próximas de, por exemplo, um local tumoral, para produzir um resultado terapêutico positivo. Em todos os casos, a experimentação de rotina em ensaios clínicos irá determinar gamas específicas para um efeito terapêutico óptimo.

A expressão de um gene marcador metastático endógeno numa célula pode ser, também, alterada, pela introdução de uma construção de ADN compreendendo uma sequência que visa uma proteína marcadora metastática, uma sequência reguladora, um exão e um local dador de excisão por recombinação homóloga não emparelhado, na grelha de leitura correcta em relação ao gene marcador metastático endógeno, de forma a que seja formada uma célula recombinante homóloga, compreendendo a construção de ADN. A nova unidade de transcrição pode ser utilizada para activar ou desactivar o gene marcador metastático, como pretendido. Este método de afectar a expressão génica endógena é mostrado na Patente U.S. Nº 5641670.



Um polinucleótido marcador metastático subgenómico pode ser, também, distribuído a indivíduos alvo com o objectivo de rastrear compostos de teste, para os que sejam úteis para intensificar a transferência dos polinucleótidos marcadores metastáticos subgenómicos para célula ou para intensificar efeitos biológicos subsequentes de polinucleótidos marcadores metastáticos subgenómicos dentro da célula. Esses efeitos biológicos incluem hibridação com ARNm marcador metastático complementar e inibição da sua tradução, expressão de um polinucleótido marcador metastático subgenómico para formar ARNm marcador metastático e/ou proteína marcadora metastática e replicação e integração de um polinucleótido marcador metastático subgenómico. O indivíduo pode ser uma cultura celular ou um animal, de um modo preferido, um mamífero, de um modo mais preferido, um humano.

A divulgação acima descreve, de um modo geral, a presente invenção. Pode ser obtida uma compreensão mais completa, por referência aos seguintes exemplos, específicos, os quais são aqui proporcionados com objectivos de, apenas, ilustração e não se destinam a limitar o âmbito da invenção.

## EXEMPLOS

### Materiais e Métodos

**Cultura celular.** As linhas celulares MDA-MB-435, MDA-MB-231, ALAB, MDA-MB-466, MDA-MB-361, ZR-75-1, MCF-7, MDA-MB453 e SK-BR-3 do cancro da mama humano (obtidas da Chiron Master Culture Collection, Chiron Corporation) foram cultivadas

a 37 °C, em CO<sub>2</sub> a 5% em DMEM + HAM F-12 (1:1) (Bio\*Whittaker, Walkersville, MD) contendo L-Glutamina 2 mM, Piruvato de Sódio 1 mM, Penicilina 100 U/mL e Estreptomicina 100 µg/mL (Bio\*Whittaker, Walkersville, MD), Solução de Vitaminas 1 x, Aminoácidos Não-Essenciais 1 x (Irvine Scientific, Santa Ana, CA) e soro fetal bovino a 10% inativado por calor (Life Technologies, Rockville, MD). As células COS-7 foram obtidas de ATCC e foram cultivadas a 37 °C, em CO<sub>2</sub> a 5%, em DMEM com soro fetal bovino a 10% inativado por calor (Life Technologies).

**Concentração de Sobrenadante Opti-MEMI.** Os meios de cultura Opti-MEM1 (Life Technologies) foram concentrados por Centricon YM-10 e/ou colunas Microcon YM-10 (Millipore Corporation, Bedford, MA). Foi, então, adicionado tampão de aplicação de amostra em SDS-PAGE e as amostras foram fervidas.

**Hibridação de transferência de Northern.** Os ARN totais foram preparados a partir de linhas celulares de cancro da mama cultivadas e tecidos tumorais de murganhos SCID transplantados com linhas celulares de cancro da mama com o Kit RNeasy Maxi (Qiagen, Valência, CA). Foram aplicadas, aproximadamente, 20 µg de ARN total por pista, num gel de formaldeído/agarose para electroforese, seguidamente, foi transferido para uma membrana de nylon Hybond-N+ (Amersham Life Sciences, Little Chalfont, Inglaterra). A transferência foi fixa por irradiação com UV. Foi pré-aquecido tampão Rapid-Hyb (Amersham Life Sciences) com ADN de esperma de cadeia simples, desnaturado, 5 mg/mL, a 65 °C e a transferência foi pre-hibridada no tampão, com agitação a 65 °C durante 30 minutos. Foi adicionado um fragmento de ADNc de hSOAF ou um fragmento de ADNc de β-actina como sonda marcada com [ $\alpha$ -<sup>32</sup>P]dCTP (3000 Ci/mmol, Amersham Pharmacia Biotech Inc., Piscataway, NJ) (Kit Prime-It RmT, Stratagene, La Jolla, CA) e

foi purificado com uma ProbeQuant™ G-50 Micro Column (Amersham Pharmacia Biotech Inc.) e foi hibridado com a transferência, com agitação, a 65 °C durante a noite. A transferência foi lavada com SSC 2 x, SDS 0,1% (w/v), à temperatura ambiente, durante 20 minutos, duas vezes em SSC 1 x, SDS 0,1% (w/v), a 65 °C, durante 15 minutos, seguidamente foi exposta a Hiperfilms (Amersham Life Sciences).

**Imunotransferência.** As amostras de proteína foram submetidas a electroforese em géis de SDS-PAGE a 10-20% e, seguidamente, foram transferidas para membranas de PVDF (0,2 µm) por electrotransferência em Tris 25 mM, glicina 192 mM, metanol a 20% (v/v), pH 8,3. As membranas foram bloqueadas em TBST (pH 7,5) contendo leite magro a 10%, seguidamente, foram transferidas em PBS (pH 7,4) contendo BSA a 1% com soro anti-hsOAF de coelho (1:1000), seguindo-se sondagem com um anticorpo secundário IgG de cabra anti-coelho, conjugado com fosfatase alcalina (1:2000) (Santa Cruz Biotechnology, Inc, Santa Cruz, CA). As bandas de proteína foram, seguidamente, visualizadas através do reagente NBTBCIP (Boehringer Mannheim, Alemanha).

**Transfecção Transiente.** A região codificante (356-1174) do ADNc do hsOAF foi clonada num vector de expressão pRetro-On modificado (Clontech, Palo Alto, CA). O vector pRetro-On contendo hsOAF ou o vector pRetro-On de controlo com GFP foi transfectado em células COS-7, numa placa de cultura de 100 mm, utilizando Effectene™ Transfection Reagent Kit (Qiagen), como descrito no protocolo fornecido pelo fabricante. As células foram recuperadas em DMEM com FBS a 10%, durante a noite, seguidamente, foram transferidas para Opti-MEM1. Ao fim de mais

dois dias, o sobrenadante foi recolhido e foi concentrado, para análise de transferência de Western.

**Transfecção com oligonucleótidos anti-sentido.** As células MDA-MB-435 foram inoculadas em placas de cultura de 6 poços, um dia antes da transfecção, para produzir uma densidade de 90% na transfecção. Foram diluídos 100  $\mu$ M de oligonucleótido anti-sentido ou controlo reverso para 2  $\mu$ M, em Opti-MEM1, para a transfecção. Foi diluído lipitoid1 estéril 0,5 mM para uma proporção de 1,5 nm de lipitoid1:1  $\mu$ g de oligonucleótido, no mesmo volume de Opti-MEM1. O oligonucleótido diluído e o lipitoid1 diluído foram misturados e foram adicionados, imediatamente, a células em meios de cultura, para uma concentração final de oligonucleótido de 100, 200 ou 300 nM. Ao fim de 6 horas, a mistura de transfecção foi substituída por meios de cultura normais e as células foram incubadas, para recuperação durante a noite. A sequência do oligonucleótido anti-sentido é AGCTGCGGATGCCACACTTGTAGG (SEQ ID N°: 4) e a sequência do oligonucleótido controlo reverso é GGATGTTACACCGTAGGCGTCGA (SEQ ID N°: 5).

**Ensaio de invasão de Matrigel.** As células foram tratadas com tripsina, lavadas e ressuspensas em meio para contagem. Foram lavadas e ressuspensas  $4 \times 10^4$  células em 100  $\mu$ L de meio em gelo. Foram adicionados 200  $\mu$ L de Matrigel (Collaborative Biomedical Products, Bedford, MA) às células no gelo. O Matrigel e as células foram cuidadosamente misturados, seguidamente, foram dispensados numa placa de cultura de 24 poços e foram solidificados a 37 °C, durante 30 min. A mistura Matrigel-células foi coberta com 0,5 mL de meio e foi incubada a 37 °C, em CO<sub>2</sub> a 5%, durante 6 dias. O meio foi substituído cada 2 dias.

**Ensaio de proliferação.** As células foram tratadas com tripsina, lavadas e ressuspensas em meios para a contagem. As células foram, seguidamente, transferidas para placas de 96 poços (5000 células/poço) para incubação. Os números de células foram medidos com o Quantos<sup>TM</sup> Cell Proliferation Assay Kit (Stratagene, La Jolla, CA), diariamente.

#### EXEMPLO 1

##### IDENTIFICAÇÃO DE UMA SEQUÊNCIA DE ADNC HUMANO

Na SEQ ID N°: 1 é apresentado o ADN codificando um homólogo humano putativo do gene *Out at First* (oaf) de *Drosophila*. Na Figura 7 é apresentado um alinhamento de hsOAF e OAF de *Drosophila*. O polinucleótido compreende 2366 pares de bases e é identificada uma grelha de leitura aberta. Na SEQ ID N°: 2 é apresentada uma tradução da ORF, um polipéptido com 273 aminoácidos. A Figura 4 apresenta as sequências de ADN e de aminoácidos, indicando a posição da ORF. Os 30 primeiros aminoácidos formam um péptido sinal, indicando que a proteína pode ser segregada. A sequência de aminoácidos do péptido sinal é: MRLPGVPLARPALLLLPLLAPLLG#TGAPA (SEQ ID N°: 3). “#” indica que a posição do local previsto de corte com protease.

## EXEMPLO 2

### EXPRESSÃO DIFERENCIAL DA SEQ ID N°: 1 EM LINHAS CÉLULARES DO CANCRO DA MAMA

Foi comparada a expressão da SEQ ID N°: 1 nas seguintes linhas celulares do cancro da mama humano:

MDA-MB-361, derivada de adenocarcinoma da mama humano;

MDA-MB-231, derivada de células do cancro da mama humano, metastático no osso e/ou pulmão;

MDA-MB-468, derivada de células do cancro da mama humano, negativas para o receptor do estrogénio;

MDA-MB-435, derivada de células de carcinoma da mama humano, negativas para o receptor do estrogénio;

MCF-7, derivada de células do cancro da mama humano, não-metastático; e

ZR-75-1, derivada de células de carcinoma da mama humano, positivas para o receptor do estrogénio.

A expressão da SEQ ID N°: 1 foi medida nas linhas celulares do cancro da mama MDA-MB 231 e MDA-MB-435, altamente metastáticas e foi comparada com linhas celulares do cancro da mama pouco metastáticas ou não-metastáticas. A expressão na MDA-MB-361 foi 11% do nível na MDA-MB-231; a expressão na MDA-MD-468 foi 44% do nível na MDA-MB-231; a expressão na MCF-7

foi 17% do nível na MDA-MB-231; e a expressão na ZR-75-1 foi 12% do nível na MDA-MB-231.

A expressão na MDA-MB-361 foi 6% do nível na MDA-MB-435; a expressão na MDA-MB-468 foi 36% do nível na MDA-MB-435; e a expressão na MCF-7 foi 3% do nível na MDA-MB-435. Deste modo, como apresentado na Tabela 2, verifica-se uma tendência evidente de expressão aumentada da SEQ ID N°: 1 em linhas celulares do cancro da mama derivadas de tumores humanos com elevado potencial metastático.

Tabela 2

Linha Celular Altamente Metastática	Linhas Celulares Pouco Metastáticas: % de Expressão em Relação às Linhas Celulares Altamente Metastáticas			
	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

Permaneceu um padrão de expressão semelhante deste gene em amostras de tecido tumorais de murganhos SCID transplantados com linhas celulares de carcinoma mamário tumorigénico. (Figura 6.)

### EXEMPLO 3

HSOAF CODIFICA UMA PROTEÍNA SEGREGADA E OS NÍVEIS DE SECREÇÃO DA PROTEÍNA HSOAF SÃO COMPATÍVEIS COM OS NÍVEIS DE EXPRESSÃO DE ARNM DE HSOAF DE LINHAS CELULARES DE CARCINOMA MAMÁRIO.

É localizada uma sequência de péptido sinal prevista, no terminal N da sequência de aminoácidos deduzida a partir do gene hsOAF (Fig. 3). Foi realizada a transfecção transiente das células COS-7 e das células MCF-7 com o vector pRetro-On contendo ADNc de hsOAF, para verificar a secreção da proteína hsOAF. Entretanto, foi utilizado o vector pRetro-On contendo GFP, como controlo. Após transfecção com hsOAF, a proteína hsOAF segregada foi detectada por imunotransferência nos meios de cultura Opti-MEM1 de ambas as linhas celulares, utilizando um anti-soro hsOAF de coelho (Fig. 8A). A proteína hsOAF segregada estava, provavelmente, glicosilada, uma vez que foram observadas bandas múltiplas com pesos moleculares aparentes mais elevados (o PM previsto para a proteína hsOAF segregada é de 28 Kda). O mesmo anti-soro hsOAF foi utilizado para detectar a secreção da proteína hsOAF em várias linhas celulares de carcinoma mamário. Os níveis de secreção da proteína hsOAF foram globalmente consistentes com os níveis de expressão de ARNm de hsOAF de entre estas linhas celulares: as linhas celulares altamente metastáticas apresentaram uma secreção de hsOAF muito mais significativa do que as linhas celulares pouco metastáticas/não-metastáticas (Fig. 8B). A MDA-MB-435 apresentou a secreção mais significativa da proteína hsOAF.



#### EXEMPLO 4

A INACTIVAÇÃO DA EXPRESSÃO DE HSOAF EM CÉLULAS MDA-MB-435 PELO OLIGONUCLEÓTIDO ANTI-SENTIDO, CAUSOU MODIFICAÇÃO MORFOLÓGICA, INVASIVIDADE CELULAR REDUZIDA E TAXA DE PROLIFERAÇÃO MAIS LENTA

Foi utilizada tecnologia de oligonucleótidos anti-sentido, para inactivar a expressão do hsOAF, para determinar se o nível elevado de expressão do gene hsOAF é essencial para o potencial metastático das células do carcinoma mamário humano, seguidamente, foram observados os efeitos consequentes. Foi seleccionada a MDA-MB-435, uma vez que esta linha celular altamente metastática apresentava a secreção de proteína hsOAF mais significativa, de entre todas as linhas celulares de cancro da mama examinadas. Foram seleccionados vários pares de oligonucleótidos hsOAF anti-sentido (AS) e controlo reverso (RC), para testar para a sua capacidade para inactivar a expressão do gene hsOAF ao nível do ARNm. Foi realizada análise em tempo real quantitativa RT-PCR em Lightcycler (Roche Diagnostics, Indianapolis, IN), para medir os níveis de ARNm de hsOAF nas células. Kang, S., et al., *Cancer Research* **60**(18):5296-5302 (2000). O melhor par foi, então, seleccionado para a titulação da concentração de trabalho do oligonucleótido. é preferida uma concentração baixa de oligonucleótido, para reduzir a toxicidade potencial dos oligonucleótidos para as células. Os resultados indicaram que o tratamento com 100 nM do oligonucleótido anti-sentido foi suficiente para reduzir, significativamente, a secreção da proteína hsOAF pelas células MDA-MB-435. (Figura 12). Foi utilizado este par de oligonucleótidos (SEQ ID N°: 4 (AS) e 5 (RC)), nas concentrações de trabalho de 100 nM, em todas as experiências seguintes.

Após o tratamento das células MDA-MB-435 com oligonucleótido hsOAF anti-sentido, foi observada uma alteração morfológica significativa das células, juntamente com uma secreção reduzida da proteína hsOAF (Fig. 10A). As células tornaram-se mais esféricas e perderam as suas protrusões proliferantes. Entretanto, as células tratadas com oligonucleótido controlo reverso permaneceram semelhantes às células MDA-MB-435 normais de cultura de tecidos. Além disso, o meio de cultura das células MDA-MB-435 normais, contendo um nível elevado da proteína hsOAF, como o meio condicionado adicionado a células tratadas com oligonucleótido anti-sentido, foi capaz de prevenir esta alteração morfológica, embora não completamente. Esta alteração da forma celular pode ser uma indicação de capacidade de invasão reduzida das células.

Foi, então, realizado um ensaio de invasão de Matrigel, para estimar a invasividade das células. Tem sido descrito que uma morfologia estrelada, invasiva, das células do cancro da mama embutidas em matrigel se correlaciona com o seu potencial metastático (Thompson, E. W., *et al.*, *J. Cell Physiol.* **150**(3):534-44 (1992); Sugiura, T., *et al.* *J. Cell Biol.* **146**(6):1375-89 (1999); Albini, A., *et al.*, *Cancer Res.* **47**(12):3239-45 (1987); e Kramer, R. H., *et al.*, *Cancer Res.* **46**(4 Pt 2):1980-89 (1986)) e isto foi confirmado em várias linhas celulares do cancro da mama cultivadas em matrigel. As células foram tratadas com tripsina, foram contadas e foram misturadas com matrigel. A mistura célula-matrigel foi, então, coberta com meio. A invasão celular foi examinada, 6 dias após a incubação, (Fig. 10B). Os resultados demonstraram que as células tratadas com oligonucleótido hsOAF controlo reverso formaram estruturas tridimensionais invasivas penetrantes, semelhantes a uma rede,

tal como ocorreu com as células MDA-MB-435 normais; por outro lado, as células tratadas com oligonucleótido hsOAF anti-sentido formaram, apenas, colónias lisas, esféricas. Novamente, foram, também, observadas colónias penetrantes nas células tratadas com oligonucleótido hsOAF anti-sentido, incubadas em meio condicionado. Estes dados demonstram que é necessária a proteína hsOAF segregada para o potencial de invasividade e metastático das células MDA-MB-435.

Foram realizadas experiências adicionais para examinar se a proteína hsOAF segregada estava envolvida no crescimento das células MDA-MB-435. Os resultados do ensaio de proliferação celular indicaram que a inactivação da secreção de proteína hsOAF diminuiu, de facto, a taxa de proliferação das células MDA-MB-435, embora a modificação tivesse sido moderada.

#### EXEMPLO 5

#### ANÁLISE DE TRANSFERÊNCIA DE NORTHERN DA EXPRESSÃO DE ARN EM LINHAS CELULARES DO CANCRO DA MAMA HUMANO E EM TECIDOS HUMANOS

Como apresentado na Figura 5, a expressão de ARNm foi regulada positivamente em linhas celulares MDA-MB-231 e MDA-MB-435 metastáticas. O ARN total foi preparado utilizando o Kit RNeasy da Quiagen. Foi realizada análise de transferência de Northern, utilizando 20-30 µg de ARN total, isolado por extracção com tiocianato de guanidina/fenol clorofórmio, a partir de linhas celulares, tumores primários ou metástases pulmonares. Os tumores primários e as metástases pulmonares foram desenvolvidos a partir de linhas celulares injectadas em murganhos scid, de acordo com os métodos bem conhecidos na

técnica. Foram radiomarcados plasmídeos contendo clones de ADNc parcial de hOAF clonado no Vector pCR2.0-TA (In vitrogen) e foram hibridados a 65 °C em Express-hyb (Clontech). De todos os tecidos examinados, o fígado, o pâncreas, o baço, o ovário e o intestino delgado apresentaram expressão significativa do hsOAF. Foi, também, detectada expressão de ARNm do HsOAF no coração, músculo esquelético, rim, próstata, cólon e medula óssea. (Figura 9).

A Tabela 3 apresenta a percentagem de positivos para hsOAF em vários tumores e tecidos normais.

**Tabela 3**

<b>Imuno-histoquímica: Percentagem de positivos para hsOAF</b>		
	<b>Tumor</b>	<b>Normal</b>
Pâncreas	9/11	0/9
Esófago	5/8	0/1
Fígado	3/6	0/13
Estômago	6/7	6/10
Mama	1/1	
De Hodgkin	1/8	

#### EXEMPLO 6

#### ENSAIO COM AGAR MOLE

**Ensaio com Agar Mole:** A camada inferior consistiu em 2 mL de agar a 0,6% em meios plaqueados de novo, algumas horas após a deposição das células. Para a camada das células, as células

MDA-MB-435, como descritas acima, foram removidas da placa com tripsina a 0,05% e foram lavadas, duas vezes, nos meios. As células foram contadas num contador coulter e foram ressuspensas a  $10^6$  por mL nos meios. Foram colocadas alíquotas de 10 mL com os meios, em placas de 96 poços (para verificar a contagem com WST1) ou foram diluídas, adicionalmente, para o ensaio com agar mole. Foram diluídas 2000 células em 800 mL de agar a 0,4% em poços duplicados, sobre a camada inferior de agar a 0,6%.

Camada dos meios: Após o agar de camada de célula ter solidificado, foram vertidos 2 mL de meio do topo e foi adicionado oligonucleótido anti-sentido ou controlo reverso, sem veículos de transferência. Foram adicionados novos meios e oligonucleótidos, cada 3-4 dias.

As colónias foram contadas durante 10 dias a 3 semanas. Os campos de colónias foram contados visualmente. Foram utilizados os valores do metabolismo de Wst-1 para compensar pequenas diferenças no número inicial de células. Podem ser examinados campos maiores para o registo visual das diferenças. Os resultados são apresentados na Figura 6, na qual as células MDA-MB-435 tratadas com anti-sentido formaram menos colónias, em comparação com células expostas ao oligonucleótido de controlo.

Os especialistas na técnica irão reconhecer ou serão capazes de apurar muitos equivalentes às formas de realização específicas da invenção aqui descritas, utilizando apenas experimentação de rotina. Pretende-se que essas formas de realização específicas e equivalentes sejam abrangidas pelas reivindicações seguintes.

Lisboa, 2 de Setembro 2008

## **REIVINDICAÇÕES**

1. Utilização de uma molécula oligonucleotídica anti-sentido, compreendendo uma sequência anti-sentido com, pelo menos, 8 nucleótidos de comprimento, a qual é complementar a uma porção da sequência codificante que abrange os nucleótidos 365 a 1173 da SEQ ID N°: 1, para a preparação de uma composição farmacêutica para o tratamento do cancro da mama e/ou metástases do cancro da mama.
2. Utilização de acordo com a reivindicação 1, em que a sequência anti-sentido tem, pelo menos, 12, 15, 20, 25, 30, 35, 40, 45 ou 50 nucleótidos de comprimento.
3. Utilização de acordo com a reivindicação 1 ou 2, em que a molécula oligonucleotídica anti-sentido compreende ribonucleótidos e/ou desoxirribonucleótidos.
4. Utilização de acordo com qualquer das reivindicações 1 a 3, em que a molécula oligonucleotídica anti-sentido tem a sequência da SEQ ID N°: 4.
5. Utilização de acordo com qualquer das reivindicações 1 a 4, em que a molécula oligonucleotídica anti-sentido é proporcionada numa construção de ADN.
6. Utilização de acordo com qualquer das reivindicações 1 a 5, em que a composição se destina a administração local ou sistémica.

7. Utilização de acordo com qualquer das reivindicações 1 a 6, em que a composição reduz em 50% a expressão dos genes marcadores metastáticos.
8. Utilização de acordo com a reivindicação 1, em que a molécula oligonucleotídica anti-sentido compreende, pelo menos, 20 nucleótidos contíguos, complementares à sequência codificante que abrange os nucleótidos 365 a 1173 da SEQ ID N°: 1.

Lisboa, 2 de Setembro de 2008

SEQ ID N°: 1

CCGCGAGGTGCGGGTCTCTTTAAGGCGGGTCTGGTGGTTCTGTTTCTGAAGGAAGTGACGGGGGGTGGGATTGAATGAAAAGTGCAAAA  
CACAGGCTCGCAGCGCTGGAGCCCGGGCCGCGGAGCCGGGCCGAGCGCGTCTCCGCCTCGGGGCCCGGGGGCGCCCTGCTGAGCG  
CTACCCACGTGCGTCCGCGCCACCTCGCGGGCGACCCCGCGGCCAAGGCCCCCGCGGAGCGGGTCCCGGGCGCCCGAACTAGCCCCAACT  
TTGGGCGAAGTTTGCTGCGCTCTCCCCGCCCCACGCGGCGCGCGGGGCCGCGGACGGCAGCGGCCCCCGGGGATGCGCTTCCCGGGT  
ACCCCTGGCGGCCCTGCGCTGCTGCTGCTGCTGCCGTGCTCGCGCGGTGCTGGGAACGGGTGCGCGGCCGAGCTGCGGGTCCGCGTGCG  
GCTGCGGACGGCCAGGTGACCGAGGAGAGCTGACGGCGSACAGCGACGCGGACAGCATCAGCCTCGAGCTGCGCAAGCCCGACGGCACCT  
CGTCTCCTTACCGCCGACTTCAAGAAGGATGTGAAGGTCTTCCGGGCCCTGATCCTGGGGGAGCTGGAGAAGGGGAGAGTCAGTCCAGGC  
CCTCTGCTTGTACCCAGCTGCAGCACAATGAGATCATCCCAAGTGAGGCCATGGCCAAAGTCCGGCAGAAAAATCCCGGGCAGTGCGGCA  
GGCGGAGGAGTTGCGGTCTGGAGCATGTCACATGGATGTCGCTGTCAACTCAGCCAGGGGGCCCTGCTGAGCCCCATCTCCACAACGT  
GTGTGCCAGGCGGTGGATGCCATCTACACCGCCAGGAGGATGTCGGTTCTGGCTGGAGCAAGGTGTGGACAGTTCTGTGTCGAGGCTCT  
GCCCAAGGCTCAGAGCAGCGGAGCTGCCCTCGCTGCAGGCAGGTGGGGACCGCGGGAAGCCCTGCGTCTGCCACTATGGCTGAGCCTGGC  
CTGGTACCCCTGCATGCTCAAGTACTGCCACAGCGCGACCGGCCACGCCCTACAAGTGTGGCATCCGAGCTGCCAGAAGAGCTACAGCTT  
CGACTTCTAGTGCCCCAGAGGCAGCTGTGTCTCTGGGATGAGGATCCCTACCCAGGCTAGGGTGGGAGCAACCTGGGCGGGTGGCTGCTG  
GGCCCACTGCTCTTACCAGCCACTAGAGGGGTGGCAACCCCACTGAGGCCATTATTCCTCCCTCCCACTCCCTGGCCCTAGAGCCT  
GGGCCCCCTCTGGCCCCATCTCACATGACTGTGAAGGGGTGTGGCATGGCAGGGGTCTCATGAAGGCACCCCCATTCCACCTGTGCCTTC  
CTTGCGGGCAGAGAGGAGAGAAGGGCTCCCCAGATCTACACCCCTCCCTCCTGCATCTCCCTGGAGTGTTCACCTGCAAGTGCCAAACA  
TGATGGCTCTGGTTGTTCTGTTGAACCTCTGAACGTTTAGACCCATAAAGGAGTCTATACCTGGACACCCACCTCCCAAGACACAATCCC  
TTCCCATGCACATCTGGAAGGAGCTGGCCCTCAGTCCCTTCTACTCCCCAACAGGGGCTCACTATCCCCAAGAGGAGCTGTGGG  
GACCCAGCAGCAGCCCTGTACTGGATTACAGCATATTCTCATCTCTGCCCCGAGGCTGCCTGTGGGGCAGTGGAGACCTCCCATCACTG  
AGACAGATCACAGACCAGAGTGCCCTTCCCGACCTGGACGTTGCCCTCCAAACAGGCACCAGCTCTTCCCTCTCTAGACAGAAATATTTT  
TGTAAGGTTCTGGGGCAGGAGGAGCATGAAGTACGAGGAAACTTGAATTCCAGATTTTTAATGCAAAGTATTTATCATTTCTACCAGAAA  
TAAAGTTTTAAGTTTTACTTGACTAATGAGACCCAGAGTTTGGAGAAACTTTTGGCCAATGCTGCCACCTGATGTGAGAAAGTGTCCCA  
CACCCTAGCAGTGGCTATCTTGAACAAGAACTTCGAAAGCACCTACTGTGTGCTCAGCCATTTGAGGAAGGAAGGAGAGGAAGATGT  
TACTAGGAAGGATGAGATAAACTTCTGCACCAAGACAATGAGACAGACATAACTGCAACCGTAGTAAGCCAGTCAGAAATAGCCAGCGG  
AAGGAAGAGATGGGTGGAGATTGGAACCCGCTTCAAGTCTGGGCTCGGCTACTTACCTGCTGTGCAGCCATGGGTCAAGTTGCTTGACCT  
CTCTGTGCTCCACTCCCTTAGCTATAAAATGAGCTTACTT

*Fig. 1*

SEQ ID N°: 2

MRLPGVPLARPALLLLPLAPLLGTGAPAE LRVRVRLPDGQVTEESLQADSDADSISLELRKPDGLVSFTADFKDKVKVFRALILGELEKG  
QSQFQALCFVTQLQHNEIIPSEAMAKLRQKNPRAVRQAEVRLGHLHMDVAVNFSQGALLSPHLHNVCAEAVDAIYTRQEDVRFWLEQGVDS  
SVFEALPKASEQAE LPRCQVQDRGKPCVCHYGLSLAWYPCMLKYCHSRDPTPYKCGIRSCQKSYSFDFYVQRQLCLWDEDPYPG\*

*Fig. 2*

SEQ ID N°: 3

MRLPGVPLARPALLLLPLAPLLG TGAPA

*Fig. 3*



*Fig. 4A*

730	740	750	760	770	780	790	800	810
GAAAAATCCCGGGCAGTGCCGGCAGGCGGAGGAGGTTCTGGGGTCTGGAGCATCTGCACATGGATGTCGCTGTCAACTTCAGCCAGGGGGC								
CTTTTAGGGGCCGTACGCCGTCCGGCTCCTCCAAGCCCCAGACCTCGTAGACGTGTACCTACAGCGACAGTTGAAGTCGGTCCCCCG								
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>								
TRADUÇÃO DO OAF HUMANO [A]								
820	830	840	850	860	870	880	890	900
CCTGCTGAGCCCCATCTCCACAAGTGTGTGCCGAGGCGGTGGATGCCATCTACACCGCCAGGAGGATGTCCGGTCTGGCTGGAGCA								
GGACGACTCGGGGGTAGAGGTGTGCACACACGGCTCCGGCACCTACGGTAGATGTGGGCGGTCTCTCAGAGCCAAGACCGACCTCGT								
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>								
TRADUÇÃO DO OAF HUMANO [A]								
910	920	930	940	950	960	970	980	990
AGGTGTGGACAGTTCTGTGTTTCGAGGCTCTGCCAAGGCCTCAGAGCAGGCGGAGCTGCCTCGCTGCAGGAGGTGGGGGACCGCGGGAA								
TCCACACCTGTCAAGACACAAGCTCCGAGACGGGTTCCGGAGTCTCGTCCGCCTCGACGGAGCGACGTCCGTCCACCCCTGGCGCCCTT								
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>								
TRADUÇÃO DO OAF HUMANO [A]								
1000	1010	1020	1030	1040	1050	1060	1070	1080
GCCCTGCGTCTGCCACTATGGCCTGAGCCTGGCCTGGTACCCCTGCATGCTCAAGTACTGCCACAGCCGCGACCGGCCACGCCCTACAA								
CGGGACGACAGCGGTGATACCGGACTCGGACCGGACCATGGGGACGTACGAGTTCATGACGGTGTCCGGCTGGCCGGGTGCGGGATGTT								
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>								
TRADUÇÃO DO OAF HUMANO [A]								
1090	1100	1110	1120	1130	1140	1150	1160	1170
GTGTGGCATCCGACGTGCCAGAAGAGCTACAGCTTCGACTTCTACGTGCCCCAGAGGCAGCTGTGTCTCTGGGATGAGGATCCCTACCC								
CACACCGTAGGCGTCGACGGTCTTCTCGATGTGCAAGCTGAAGATGCACGGGGTCTCCGTGACACAGAGACCTACTCTAGGGATGGG								
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>								
TRADUÇÃO DO OAF HUMANO [A]								
1180	1190	1200	1210	1220	1230	1240	1250	1260
AGGCTAGGGTGGGAGCAACCTGGGCGGGTGGCTGCTCTGGGCCACTGCTCTTACCAGCCACTAGAGGGGGTGGCAACCCCACTGAG								
TCCGATCCCACCTCGTTGGACCCGCCACCGACGAGACCCGGGTGACGAGAAGTGGTGGTGATCTCCCCACCGTTGGGGGTGGACTC								
G *								
1270	1280	1290	1300	1310	1320	1330	1340	1350
GCCTTATTTCCCTCCCTCCCCACTCCCCTGGCCCTAGAGCCTGGGCCCTCTGGCCCCATCTCATGACTGTGAAGGGGTGTGGCATG								
CGGAATAAAGGGAGGGAGGGGTGAGGGGACCGGGATCTCGGACCCGGGGAGACCGGGGTAGAGTGTACTGACACTTCCCCACACCGTAC								

Fig. 4B

1360 1370 1380 1390 1400 1410 1420 1430 1440  
 GCAGGGGGTCTCATGAAGGCACCCCATTCACCCCTGTGCCTTCCTTGCAGGAGAGAGGAGAGAAGGGCTCCCCAGATCTACACCC  
 CGTCCCCAGAGTACTTCGTGGGGTAAGGGTGGGACACGGAAGGAACGCCCGTCTCTCCCTCTCTCCGAGGGGTCTAGATGTGGGG

1450 1460 1470 1480 1490 1500 1510 1520 1530  
 TCCCTCTGCATCTCCCTGGAGTGTCACTTGCAAGCTGCCAAAACATGATGGCTCTGGTGTCTGTGAACCTTTGAACGTTTAG  
 AGGGAGGACGTAGAGGGGACCTACAAGTGAACGTTGACGGTTTTGTACTACCGAGACCAACAAGACAATTGAGGAACCTGCAATC

1540 1550 1560 1570 1580 1590 1600 1610 1620  
 ACCCTAAAAGGAGTCTATACCTGGACACCCACCTCCCCAGACACAACCTCCCTTCCCATGCACACATCTGGAAGGAGTGGCCCTCAGT  
 TGGGATTTTCTCAGATATGGACCTGTGGGTGGAGGGGTCTGTGTTGAGGGAAGGGGTACGTGTGTAGACCTTCTCGACCGGGGAGTCA

1630 1640 1650 1660 1670 1680 1690 1700 1710  
 CCCTTCTACTCCCCAACAAGGGGCTCACTATCCCCAAGAAGGAGCTGTTGGGGACCCACGACGACGCCCTGTACTGGATTACAGCAT  
 GGGGAAGGATGAGGGGTGTTCCCGAGTGATAGGGGTTTCTCTCGACAACCCCTGGGTGCTGCGTCGGGGACATGACCTAATGTCTGA

1720 1730 1740 1750 1760 1770 1780 1790 1800  
 ATTCTCATCTCTGGCCCCGAGGCTGCCTGTGGGGCGAGTGGAGACCTCCCATCACTGAGACAGATCAGAGACCAGAGTGCCTTTCCCGG  
 TAAGAGTAGAGACCGGGGCTCCGACGGACACCCCGCTCACCTCTGGAGGGTAGTGACTCTGTCTAGTGTCTGGTGCTCACGGAAGGGCC

1810 1820 1830 1840 1850 1860 1870 1880 1890  
 ACCTGGACGTTGCCTCCAAAACAGGCACAGCTCTTTCCCTCTCTAGACAGAAATATTTTGTAAAGTTCTGGGGCAGGGAGGGAGCATG  
 TGGACCTGCAACGGAGGTTTGTCCGTGGTCGAGAAAGGGAGAGATCTGTCTTTATAAAAACATTCCAAGACCCCGTCCCTCCCTCGTAC

1900 1910 1920 1930 1940 1950 1960 1970 1980  
 AAGTACGAGGAAAACCTGAATCCAGATTTTAAATGCAAAGTATTATCATTTCTACCAGAAATAAACGTTTTAAGTTTTACTTGACTA  
 TTCATGCTCCTTTTGAACCTAAGGTCTAAAAATTACGTTTCATAAATAGTAAAGATGGTCTTTATTGCAAAATTCAAAAATGAAGTGAT

1990 2000 2010 2020 2030 2040 2050 2060 2070  
 ATGAGACCCAGAGTTTGGAGAAAACTTTGGCCAATGCTGCCACCTGATGTCAGAAAGTGTCCCACACCCTAGCAGTGGCCTATCTTGG  
 TACTCTGGGTCTCAACCTCTTTGAAAACCGTTACGACGGTGGACTACAGTCTTTCACAGGGGTGTGGGATCGTCACCGGATAGAACC

2080 2090 2100 2110 2120 2130 2140 2150 2160  
 AACAGAAGCTTCGAAAGCACCTACTGTGTGCTCAGCCATTTGAGGAAGGAAGGAGGAGAAGGAAGATGTTACTAGGGAAGGATGAGATAA  
 TTGTTCTTGAAGCTTTCGTGGATGACACAGAGTCGGTAAACTCCTTCCTCCTCCTCCTCTCTACAATGATCCCTTCTACTCTATT

*Fig. 4C*

---

2170	2180	2190	2200	2210	2220	2230	2240	2250
AACTTCTGCACCCAAGACAATGAGACAGACATAACTGCAACCGTAGTAAGCCAGTCAGAAATAGCCAGCGCGAAGGCAAGAGATGGGGTG								
TTGAAGACGTGGGTTCTGTACTCTGTCTGTATTGACGTTGGCATCATTCGGTCAGTCTTTATCGGTCGCGCTTCGGTTCTCTACCCAC								
2260	2270	2280	2290	2300	2310	2320	2330	2340
GAGATTGGAACCCCGCTTCAGATCTGGGCTCGGCTACTTACCTGCTGTGCAGCCATGGGTCAAGTTGCTTGACCTCTCTGTGCCTCCACT								
CTCTAACCTTGGGGCGAAGTCTAGACCCGAGCCGATGAATGGACGACACGTCGGTACCCAGTTCAACGAAGTGGAGAGACCGGAGGTGA								
2350	2360							
CCCTTAGCTATAAAATGAGCTTACTT								
GGGAATCGATATTTTACTCGAATGAA								

*Fig. 4D*

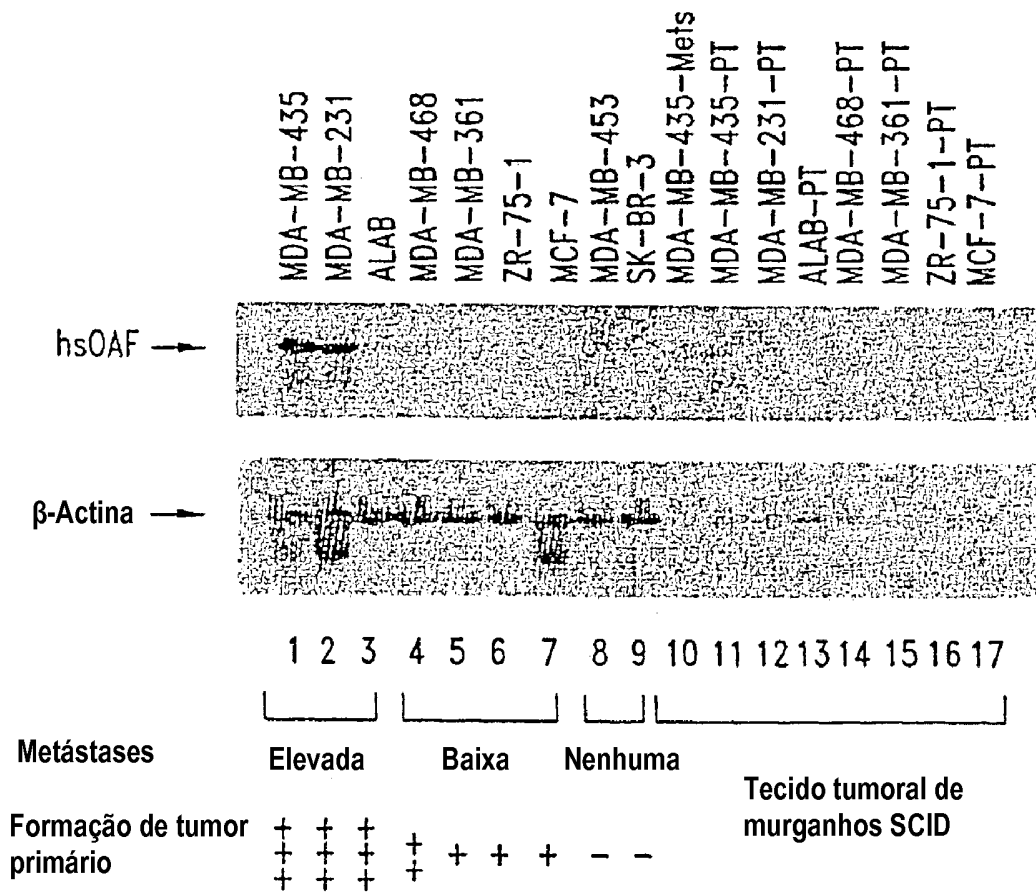
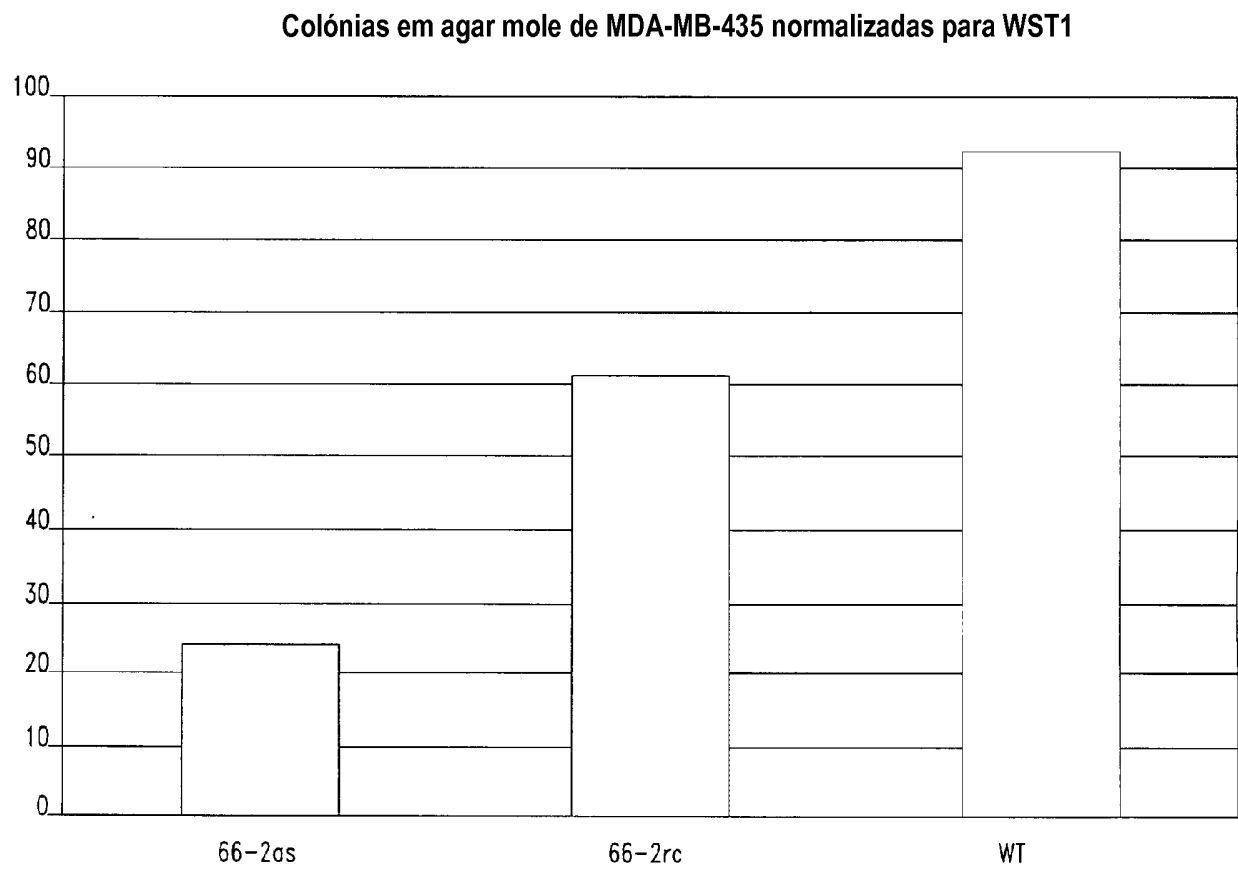


Fig. 5



*Fig. 6*

F 1 MR L PG-----VPLAR 10  
 1 MILKEEPHQSIETAANAARQAQVRWRMAHLKALSRTIRTPAHGNCCGRVVSKNHFFKHSR 60

F 11 PALLLLPL LAPLLGTGAPAE LRVRVRLPDGOVTEESLQADSDADSTISLE RKPDGILVS 70  
 61 AFLWFLLCNLVMNADAFHSQLIN VQNOGGEVTEESITSNIGEDLITL EFOKTDGILIT 120

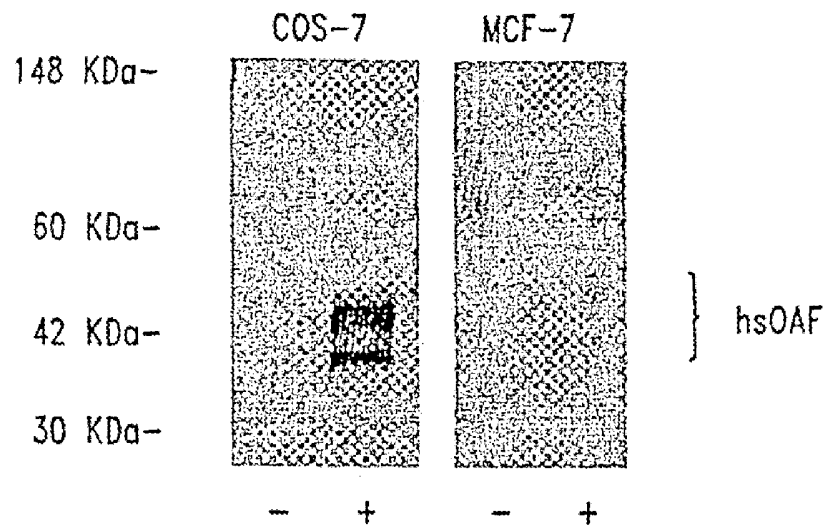
F 71 FTADIKKDVKF RALILGEL KKGQSQFQALCFVTQLQHNETIPSEAMAKLRQKNPRAVRQ 130  
 121 QVIDIRNEVQIL KALVLGEEKRGQSQYQVMCFATKFNKGDFISSAAMAKLRQKNPHTIRT 180

F 131 AEEVRGLEHLHMDVAVNFSQALLSPHLHNVC AEAVDAITVRDEDVRFMLED-GVDSSVF 189  
 181 REEDKGRETFMSSWQLNRSLPITRHLQGLCAEAMDATVVRDVLKAWAELPGSSISL 240

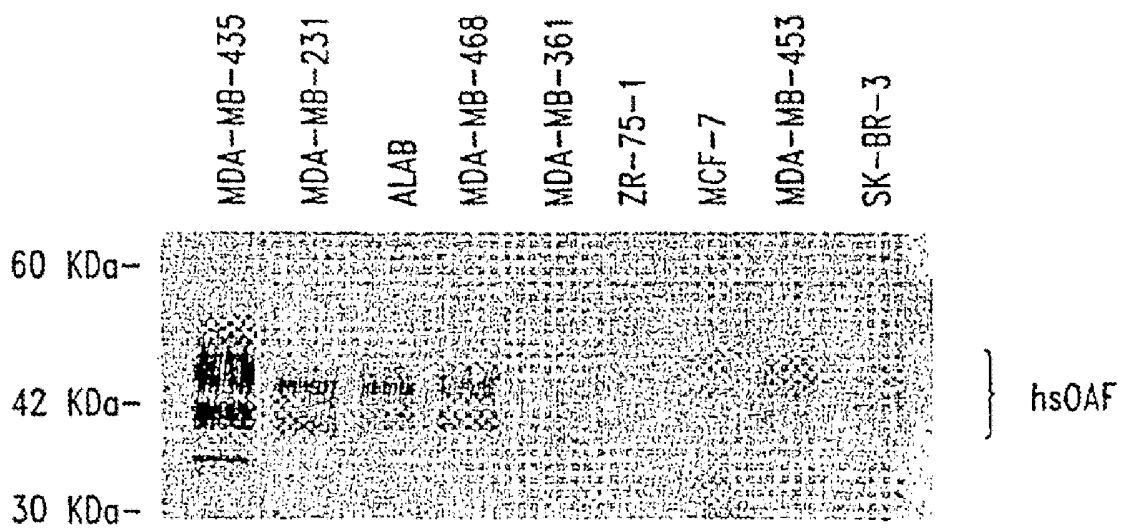
F 190 KALPKASEQAELPRCROVGDRGKPCVCHYGLSLAWVPCMLKVCHSR-----D----- 236  
 241 KAATEKFPTLSTRONEVSSLWAPCLCNLETICGWVPCGLKVCKGKGVAGADSSGAQQQA 300

F 237 RPTPVKCGIRSCQKSYSFDFVVPQRQLCLWDEDPYPG 273  
 301 OPTNVRCGIKTCRKCTOFTYVVRQKQCLWDE 332

*Fig. 7*

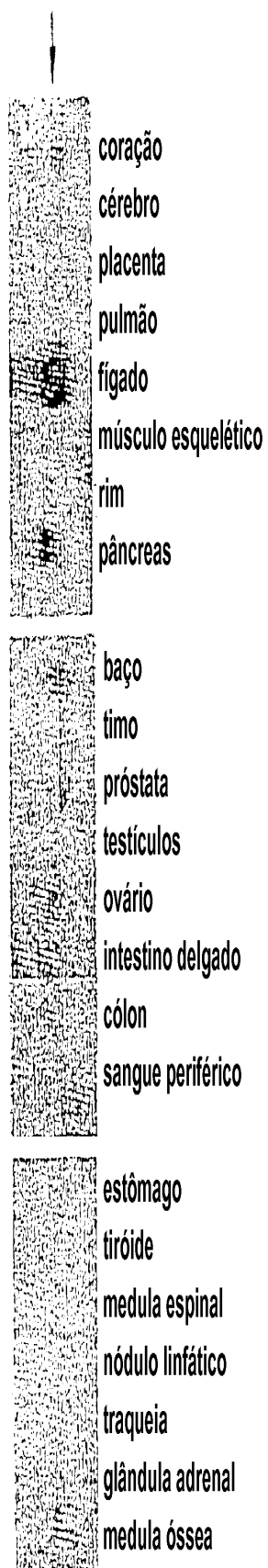


*Fig. 8A*

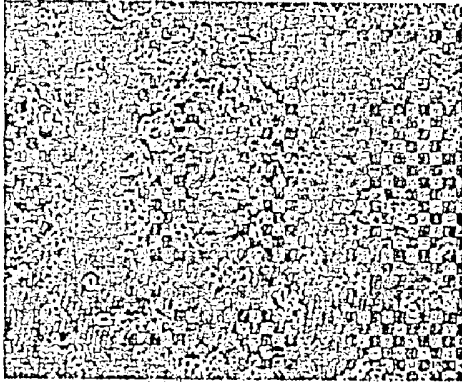


*Fig. 8B*

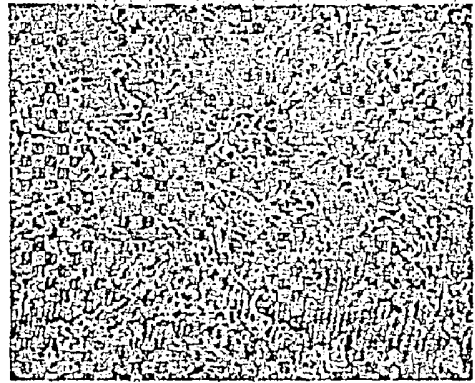




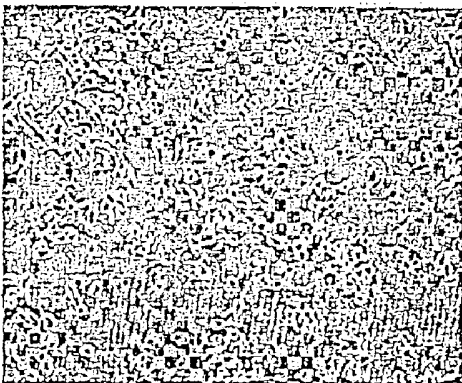
*Fig. 9*



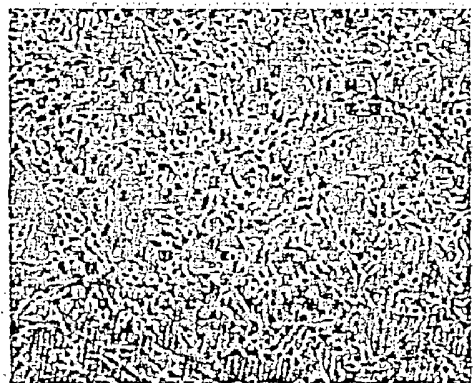
AS



RC

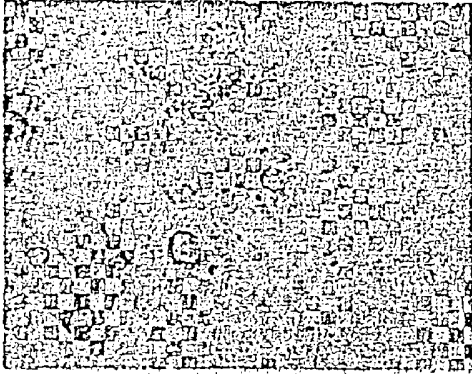


AS + M

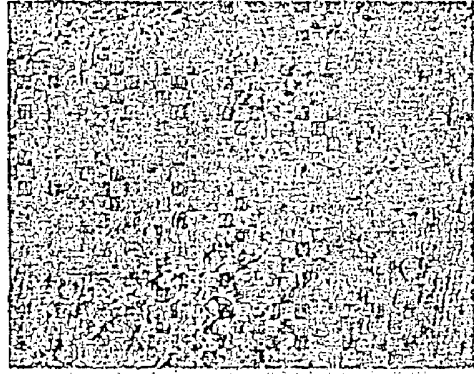


Normal

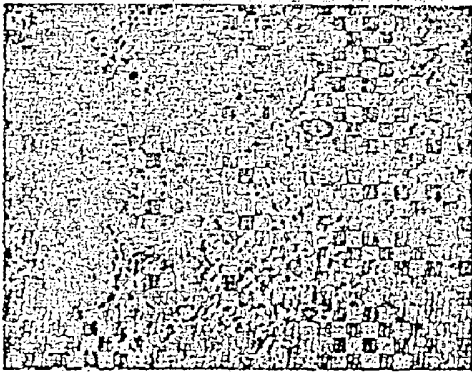
*Fig. 10A*



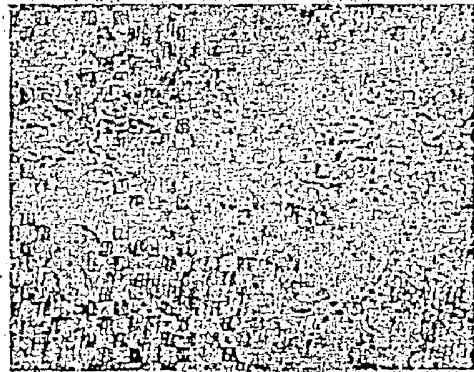
AS



RC



AS + M

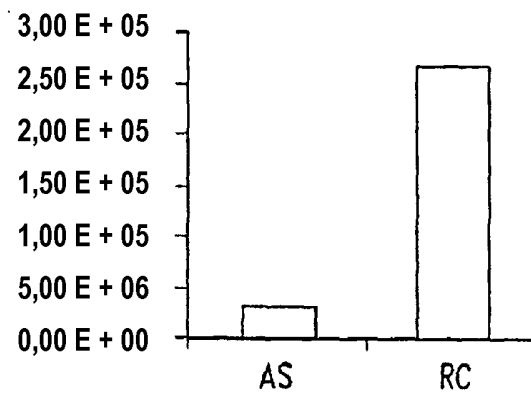
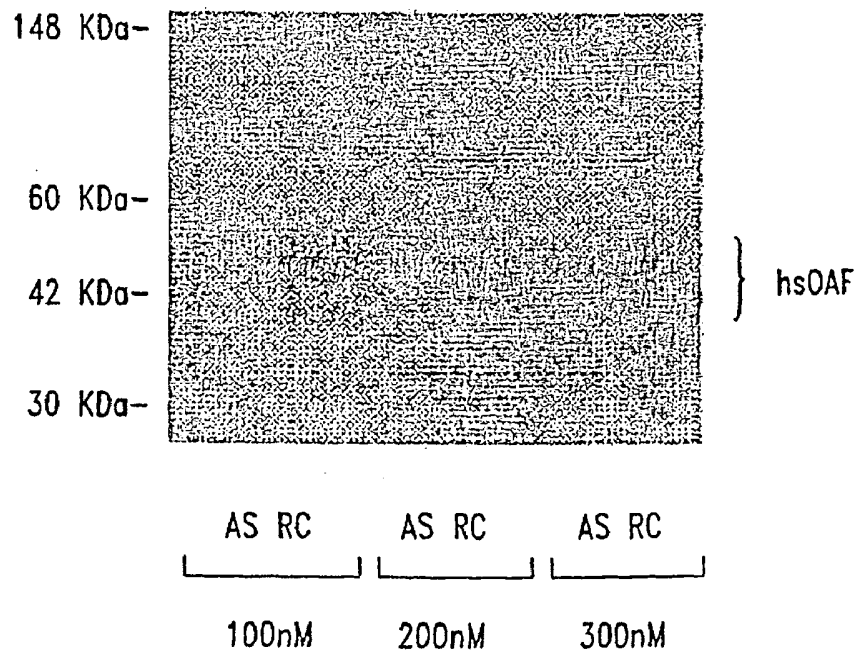


Normal

*Fig. 10B*

1 ccgcgaggtgcgcggtctctttaaggcgggtcctggtggtttctgtttcctgaaggaagtacggggggtgggattgaatgaaaagtg  
 89 caaaacacaggtctgcagcgtggagcccgggccgcggagccgggcccgggcagcgcgtctccgctcggggccgcgggggcccct  
 179 GCTGAGCGCTACCCACGTGCGTCCGCGCCACCTCGCGGGCGACCCCGCGGCCAAGGCCCCCGCGGAGCGGCTCCCGGGCGCCCCGAAC  
 269 AGCCCCCAACTTTGGGCGAAGTTTGCCTGCGCCTCTCCCGCCCCACGCGGCGCGCGGGGCGCGGACGGCAGCGGCCCCCGGGGATG  
 1  
 359 CGCCTTCCCGGGGTACCCCTGGCGCGCCTCGCGTGTGCTGTGCTGCCGTGCTCGCGCCGCTGCTGGGAACGGGTGCGCCGGCCGAG  
 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 CTGCGGGTCCCGTGCCTGCGGACGGCCAGGTGACCGAGGAGAGCCTGCAGGCGGACAGCGACCGGACAGCATCAGCCTCGAGCTG  
 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 CGCAAGCCCGACGGCACCCCTCGTCTCCTTACCGCGACTTCAAGAAGGATGTGAAGGTCTTCCGGGCCCTGATCTGGGGGAGCTGGAG  
 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 AAGGGGAGAGTCAGTTCAGGCCCTTGCTTTGTACCCAGCTGCAGCACAATGAGATCATCCCCAGTAGGGCCATGCCAAGCTCCGG  
 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 CAGAAAAATCCCGGGCAGTGGCGCAGGCGGAGGAGGTTCCGGGTCTGGAGCATCTGCACATGGATGCTGTCACTTACGCCAGGGG  
 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 GCCCTGCTGAGCCCCATCTCCACAAGTGTGTGCGAGGCGGTGGATGCCATCTACACCGCCAGGAGGATGTCCGGTTCTGGCTGGAG  
 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 CAAGGTGTGGACAGTTCTGTGTTGAGGCTCTGCCAAGGCCCTCAGAGCAGGCGGAGCTGCCTCGCTGCAGGCAGGTGGGGGACCGCGG  
 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 AAGCCCTGCGTCTGCCACTATGGCCTGAGCCTGGCCTGGTACCCCTGCATGCTCAAGTACTGCCACAGCCGCGACCGGCCACGCCCTAC  
 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 AAGTGTGGCATCCGAGCTGCCAGAAGAGCTACAGCTTCGACTTCTACGTGCCCCAGAGGCAGCTGTGTCTCTGGGATGAGGATCCCTAC  
 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 ccagggtagggtgggagcaacctgggcgggtggctgctctgggccactgctcttcaccagccactagaggggtggcaacccccacctg  
 272 p g \*  
 1259 aggccttatttccctccctccccactccccctggccctagagcctgggcccCtctggcccatctcacatgactgtgaaggggtgtggca  
 1349 tggcaggggtctcatgaaggcacccttccaccctgtgccttcccttgcgggcagagagggagagaagggtccccagatctacacc  
 1439 cctccctcctgcatctccctggagtggtcacttgcaagctgcaaaacatgatggcctctggtgttctgttgtaactcctgaacgttt  
 1529 agaccctaaaaggagtctatacctggacaccacccctccccagacacaactcccttccccatgcacacatctggaaggagctggccctca  
 1619 gtcccttccctactcccaacaagggtcactatccccaaagaaggagctgttggggacccacgacgcagccctgtactggattacagc  
 1709 atattctcatctctgccccgaggtgctgtggggcagtgagacctccatcactgagacagatcacagaccagagtgcttccc  
 1799 ggaactggaggttgctccaaaacaggcaccagctcttccctctctagacagaaatattttgtaaggttctggggcaggaggagca  
 1889 tgaagtacgaggaacttgaattccagatttttaagtcaaaagtatttatcatttctaccagaaataaacgttttaagttttacttgac  
 1979 taatgagaccagagtttgagaaaaactttggccaatgctgccacctgatgtcagaaagtgtccccacacctagcagtgccctatctt  
 2069 ggaacaagaacttcgaaagcacctactgtgtgctcagccatttgaggaaggaaggaggagaaggaagatgttactaggaaggatgagat  
 2159 aaaacttctgcaccaagacaatgagacagacataactgcaaccgtagtaagccagtcagaaatagccagcgcgaaggcaagagatgggg  
 2249 tggagattggaacccgcttcagatctgggctcggtacttacctgctgtgcagccatgggtcaagttgcttgacctctctgtgcctca  
 2339 ctcccttagctataaaatgagcttactt-polyA

*Fig. 11*

*Fig. 12A**Fig. 12B*