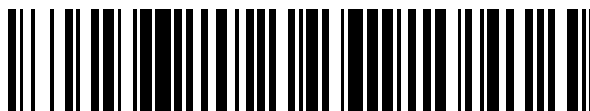


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 631 458**

51 Int. Cl.:

C12N 15/113 (2010.01)

A61K 31/713 (2006.01)

A61K 31/7105 (2006.01)

A61P 35/04 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **04.03.2011 PCT/NL2011/050152**

87 Fecha y número de publicación internacional: **09.09.2011 WO11108930**

96 Fecha de presentación y número de la solicitud europea: **04.03.2011 E 11708936 (7)**

97 Fecha y número de publicación de la concesión europea: **12.04.2017 EP 2542678**

54 Título: **Molécula de ARNmi definida por su fuente y sus usos terapéuticos en el cáncer asociado a la EMT**

30 Prioridad:

04.03.2010 US 310452 P
04.03.2010 EP 10155512

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
31.08.2017

73 Titular/es:

INTERNA TECHNOLOGIES B.V. (100.0%)
Jonkerbosplein 52
6534 AB Nijmegen, NL

72 Inventor/es:

SCHAAPVELD, ROELAND QUIRINUS JOZEF;
VERHAEGH, GERARDUS WILHELMUS
CHRISTIAAN THEODOOR;
SCHALKEN, JACOBUS ANTONIUS;
VAN PUIJENBROEK, ANDREAS ALPHONS
FRANCISCUSLUDOVICUS;
GOMMANS, WILLEMIJN MARIA y
WEIJZEN, SANNE

74 Agente/Representante:

TOMAS GIL, Tesifonte Enrique

ES 2 631 458 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Molécula de ARNmi definida por su fuente y sus usos terapéuticos en el cáncer asociado a la EMT

5 Campo de la invención

[0001] La invención se refiere al uso terapéutico de una molécula de ARNmi definida más tarde por su fuente en enfermedades y afecciones de un cáncer asociado con la EMT (transición epitelial a mesenquimal).

10 [0002] La divulgación se refiere al uso diagnóstico de una molécula de ARNmi definida por su fuente más tarde en enfermedades y afecciones asociadas con la EMT.

Antecedentes de la invención

15 [0003] Muchos tumores sólidos son epiteliales en origen (es decir carcinomas). Una pérdida de marcadores celulares epiteliales (por ejemplo, E-cadherina) y una ganancia de marcadores celulares mesenquimales (por ejemplo, N-cadherina y vimentina) se ha observado en muestras de tumores de pacientes, incluyendo cáncer de próstata (1). Las células cancerosas se pueden desdiferenciar a través de esta denominada transición epitelial a mesenquimal (EMT). Durante la EMT, las uniones celulares intercelulares se descomponen, aportando así a las
20 células tumorales la capacidad para invadir y migrar al tejido circundante o a través de las paredes de los vasos sanguíneos. Tales cambios fenotípicos se cree que desempeñan un gran papel en la diseminación de la enfermedad y en última instancia llevan a la progresión de la enfermedad, que se asocia frecuentemente a una prognosis pobre para los pacientes (2, 3).

25 [0004] La pérdida de expresión de la E-cadherina se considera como una marca del contraste molecular de EMT. La EMT en las células tumorales se produce a partir de una reprogramación transcripcional de la célula. En particular la represión transcripcional del promotor génico de la *E-cadherina* (*CDH1*) ha mostrado que desencadena el fenotipo de EMT. La proteína E-cadherina es una de las moléculas de cadherina más importantes que median los contactos célula-célula en las células/tejidos epiteliales. *CDH1* es reprimida por la
30 unión de los represores transcripcionales, SNAI1, SNAI2, TCF3, TWIST, ZEB1, ZEB2 o KLF8 (4-7), para tres denominadas E-cajas en la región del promotor proximal de *CDH1* (8-10). La inhibición de la unión de estos represores para el promotor de *CDH1* puede revertir la EMT, también denominada transición mesenquimal a epitelial (MET), e inhibe la invasión de células tumorales y la progresión tumoral (11). La WO 2009/044899 y la EP 2 208 499 divulgan que la expresión de un precursor de ARNmi-518 podría reducir la viabilidad e inducir la
35 apoptosis en unas líneas celulares de cáncer de colon y ovárico. La WO 2007/148235 divulga un método de diagnóstico de un sujeto con un cáncer específico basado en la expresión de un ARNmi dado. Cervigne *et al.* Mol Gen 2009, vol 18, no 24, p4818-4829 divulgan una sobreexpresión significativa de un subconjunto de ARNmi incluyendo ARNmi-518b tanto en leucoplasias progresivas como en carcinomas de células escamosas orales (OSCC) en comparación con la expresión en una célula normal.

40 [0005] Recientemente, la expresión de diferentes microARN ha mostrado que está enlazada con la EMT (12). Al comparar perfiles de expresión de microARN de células con un fenotipo epitelial y mesenquimal (inducido), miembros de la familia de miR-200 (miR-141, miR-200a/b/c y miR-429) y miR-205 se identificaron como miRs asociados a la EMT (13-15). Los genes diana de los microARN asociados a la EMT de la familia miR-200
45 mostraron ser *ZEB1* y *ZEB2*. MicroARN dirigidos a los otros represores transcripcionales conocidos de *CDH1* (es decir, *SNAI1*, *SNAI2*, *TCF3* y *TWIST1*) aún no se han encontrado. La identificación de estos microARN en un perfil de expresión de células que han sufrido EMT no significa necesariamente que estos microARN están implicados durante la EMT.

50 [0006] Huang *et al.* Nature Cell Biol. 2008, vol 10, no 2, p202-210 demuestran que otros miembros del agrupamiento miR-515 promueven la invasión tumoral y la metástasis. La US 2009/186353 divulga varios miembros del agrupamiento miR-515 como regulados de forma diferencial en el cáncer, incluyendo cáncer de vejiga de alto grado y el cáncer de hígado metastásico. Actualmente no existe ningún medicamento conocido eficaz que se pueda utilizar para prevenir, tratar, revertir y/o retrasar específicamente una enfermedad o
55 condición asociada a la EMT en un sujeto. Los tratamientos estándar solo comprenden quimioterapia, radioterapia, cirugía. Particularmente, la identificación de pacientes que desarrollarán o que ya han desarrollado metástasis y/o el tratamiento temprano de pacientes con marcadores de EMT que expresan tumores, tal como la expresión de cadherina mesenquimal y/o la expresión inferior de E-cadherina podría contribuir a una mejor supervivencia libre y global de la enfermedad. Por lo tanto, sigue habiendo una necesidad de marcadores de diagnóstico para EMT y para tratamientos nuevos de la enfermedad o condiciones asociadas con la EMT.
60

Descripción de la invención

65 [0007] En un primer aspecto de la invención, se proporciona una molécula de ARNmi o un equivalente de la misma donde una fuente de dicha molécula de ARNmi o equivalente de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID

NO: 1 o una composición que comprende dicha molécula de ARNmi o equivalente de la misma o dicha fuente de la misma para su uso como un medicamento para prevenir, tratar, revertir, curar y/o retrasar un cáncer asociado a la EMT por inducción de un proceso de MET, donde dicha molécula de ARNmi es un ARNmi-518b, ARNmi-520f y/o ARNmi-524. En un primer aspecto de la divulgación, se proporciona una molécula de ARNmi o un equivalente de la misma donde una fuente de dicha molécula de ARNmi o equivalente de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1 o una composición que comprende dicha molécula de ARNmi o equivalente de la misma o dicha fuente de la misma para su uso como un medicamento para prevenir, tratar, revertir, curar y/o retrasar una enfermedad o una condición asociada a EMT, donde dicha molécula de ARNmi es un ARNmi-518b, ARNmi-520f y/o ARNmi-524 o un equivalente o un mimético o un isomiR o una fuente del mismo. Los microARN (ARNmi) son ARN pequeños de 17-25 nucleótidos, que funcionan como reguladores de la expresión génica en eucariotas. Los ARNmi son inicialmente expresados en el núcleo como parte de transcritos primarios largos denominados ARNmi primarios (ARNmi-pri). Dentro del núcleo, los ARNmi-pri son digeridos parcialmente por la enzima Drosha, para formar ARNmi precursor de horquilla de 65-120 nucleótidos de longitud (ARNmi-pre) que se exportan al citoplasma para más tratamiento por Dicer en ARNmi maduros más cortos, que son las moléculas activas. En animales, estos ARN cortos comprenden una región "semilla" proximal 5' (nucleótidos 2 a 8) que parece ser el determinante primario de la especificidad de emparejamiento del ARNmi en la región no traducida 3' (3'-UTR) de un ARNm objetivo. Una explicación más detallada se da en la parte dedicada a las definiciones generales.

[0008] Cada una de las definiciones dadas a continuación sobre una molécula de ARNmi, un equivalente de ARNmi o una fuente de ARNmi se debe usar para cada uno de los ARNmi identificados o equivalentes de ARNmi o fuentes de ARNmi de esta solicitud: ARNmi-124-1, ARNmi-206, ARNmi-1, ARNmi-141, ARNmi-200a, ARNmi-200b, ARNmi-200c, ARNmi-429, ARNmi-205, ARNmi518b, ARNmi520f, ARNmi524 y fuentes de los mismos, además incluyen una fuente que comprende al menos 80 nucleótidos y que comprenden una unidad que tiene al menos 98% de identidad con la unidad representada por SEQ ID NO: 1. Secuencias maduras preferidas (como se identifica en la tabla 3), semilla (como se identifica en la tabla 5) isomiRs (como se identifica en la tabla 6) o fuente (como se identifica en las tablas 2 (precursor de ARN) o 4 (ADN que codifica un precursor de ARN)) de dicha molécula de ARNmi o equivalente de la misma respectivamente se identifican en las tablas correspondientes. En el texto entero de la solicitud a menos que se indique lo contrario, un ARNmi también se puede denominar una molécula de ARNmi, un miR, o un equivalente del mismo o una fuente o un precursor del mismo. Un equivalente preferido es un maduro, un isomiR o un mimético. Cada secuencia identificada aquí se puede identificar como SEQ ID NO como se usa en el texto de la solicitud o como SEQ ID NO correspondiente en el listado de secuencias.

[0009] En el contexto de la invención una molécula de ARNmi o un equivalente o un mimético o un isomiR de la misma puede ser un sintético o natural o recombinante o maduro o parte de un ARNmi maduro o un ARNmi humano o derivado a partir de un ARNmi humano como se define más adelante en la parte dedicada a las definiciones generales. Una molécula de ARNmi humana es una molécula de ARNmi que se encuentra en una célula, tejido, órgano o líquidos biológicos humanos (es decir molécula de ARNmi humana endógena). Una molécula de ARNmi humana también puede ser una molécula de ARNmi humana derivada a partir de una molécula de ARNmi humana endógeno por sustitución, eliminación y/o adición de un nucleótido. Una molécula de ARNmi o un equivalente o un mimético de la misma puede ser una molécula de ARN monocatenaria o bicatenaria.

[0010] En el contexto de la invención, una molécula de ARNmi preferida o un equivalente o un mimético o un isomiR de la misma es de manera que una fuente de dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma comprende o consiste en al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1 se define posteriormente de forma preferible de la siguiente manera.

SEQ ID NO: 1 es como sigue:

UCAnGCUGUGnCCCUnnAnAGGGAAGCnCUUUCUnUnGUCnnAAAnGAAAAnnA
nGnGCUnCCnUUUnGAGnnUUACnGUUUG

En la unidad representada por SEQ ID NO: 1, n puede ser cualquier base A, U, C o G. En otra forma de realización preferida, se proporciona una molécula de ARNmi o un equivalente o un mimético o un isomiR de la misma de manera que una fuente de dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO:1. Más preferiblemente, la identidad es de al menos 99% o 100%. En una forma de realización preferida, dicha fuente es un precursor de una molécula de ARNmi-518b, ARNmi-520f o ARNmi-524 o un equivalente o un mimético o un isomiR de la misma. Fuentes preferidas y precursores se definirán en la presente más adelante. La invención por lo tanto se refiere a una molécula de ARNmi o un equivalente o un

5 mimético o un isomiR de la misma o una fuente de la misma o una composición que comprende dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma o fuente de la misma, tal y como se define en los siguientes párrafos, donde la molécula de ARNmi es un ARNmi-518b, ARNmi-520f y/o ARNmi-524 o un equivalente o un mimético o un isomiR de la misma o una fuente de la misma. Las moléculas de ARNmi, equivalentes, miméticos e isomiR preferidas se definirán más tarde en la presente.

10 [0011] En una forma de realización preferida, un equivalente de un ARNmi-518b, un ARNmi-520f o de un ARNmi-524 es una molécula de ARNmi humana. Una molécula de ARNmi humana es una molécula de ARNmi que se encuentra en una célula, tejido u órgano humano. Una molécula de ARNmi humana también puede ser una molécula de ARNmi humana derivada a partir de una molécula de ARNmi humana endógena por sustitución, eliminación y/o adición de un nucleótido. En este contexto, un "nucleótido" puede referirse a 1, 2, 3, 4, 5 o más nucleótidos. Un equivalente preferido de una molécula de ARNmi-518b, ARNmi-520f o ARNmi-524 no es una molécula de ARNmi mml-mir-519a o mml-mir-520c como se identifica de aquí en adelante. Una fuente preferida o precursor de una molécula de ARNmi-518b, ARNmi-520f o ARNmi-524 no es una fuente o un precursor de una molécula de ARNmi mml-mir-519a o mml-mir-520c según se identifica de aquí en adelante. Las secuencias maduras rechazadas preferidas y precursoras de mml-mir-519a se identifican como SEQ ID NO: 108 y 109. Las secuencias maduras preferidas y precursoras de mml-mir-520c se identifican como SEQ ID NO: 110 y 111.

20 [0012] En una forma de realización preferida, una molécula de ARNmi o equivalente o mimético o isomiR de la misma comprende al menos 6 de los 7 nucleótidos presentes en la secuencia semilla de dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma (la tabla 5 muestra la secuencia semilla preferida de cada una de las moléculas de ARNmi identificadas aquí). Preferiblemente en esta forma de realización, una molécula de ARNmi o un equivalente o un mimético o isomiR de la misma puede ser de 6, 7, 8, 9, 10, 11, 12 a 30 nucleótidos de longitud y más preferiblemente comprende al menos 6 de los 7 nucleótidos presentes en la secuencia semilla de dicha molécula de ARNmi o equivalente de la misma. Aún más preferiblemente una molécula de ARNmi o un equivalente o un mimético o isomiR de la misma es de 15 a 28 nucleótidos de longitud y más preferiblemente comprende al menos 6 de los 7 nucleótidos presentes en la secuencia semilla. Aún más preferiblemente una molécula de ARNmi tiene una longitud de al menos 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 nucleótidos o más y preferiblemente comprende al menos 6 de los 7 nucleótidos presentes en la secuencia semilla. En cada una de estas formas de realización, una molécula de mirARN o un equivalente o un mimético o isomiR de la misma puede comprender los 7 nucleótidos de la secuencia semilla como se identifican en la tabla 5. Aún más preferiblemente una molécula de ARNmi tiene una longitud de al menos 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 nucleótidos o más y comprende los 7 nucleótidos presentes en la secuencia semilla como se identifican en la tabla 5.

40 [0013] Por consiguiente, una molécula preferida de ARNmi-520f o equivalente o mimético o isomiR de la misma comprende al menos 6 de los 7 nucleótidos presentes en la secuencia semilla identificada como SEQ ID NO: 104 o 105 y más preferiblemente tiene una longitud de al menos 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 nucleótidos o más. Por consiguiente, una molécula de ARNmi-518b preferida o equivalente o mimético o isomiR de la misma comprende al menos 6 de los 7 nucleótidos presentes en la secuencia semilla identificada como SEQ ID NO: 103 y más preferiblemente tiene una longitud de al menos 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 nucleótidos o más.

45 [0014] Por consiguiente, una molécula de ARNmi-524 preferida o equivalente o mimético o isomiR de la misma comprende al menos 6 de los 7 nucleótidos presentes en la secuencia semilla identificada como SEQ ID NO: 106 o 107 y más preferiblemente tiene una longitud de al menos 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 nucleótidos o más.

50 [0015] En otra forma de realización preferida, una molécula de ARNmi o un equivalente o mimético o isomiR de la misma comprende al menos 6 de los 7 nucleótidos presentes en una secuencia semilla dada como se identifica en la tabla 5 como SEQ ID NO: 87-107 y tiene al menos 80% de identidad sobre la secuencia madura entera (la tabla 3 muestra las secuencias maduras preferidas de cada uno de los ARNmi identificados aquí). Preferiblemente, la identidad es al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, o más alta, tal como 96%, 97%, 98%, 99% o 100%. Alternativamente, preferiblemente en esta forma de realización, una molécula de ARNmi o un equivalente o un mimético o un isomiR de la misma tiene una longitud no mayor de 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40 nucleótidos, comprende al menos 6 de los 7 nucleótidos presentes en una secuencia semilla dada como se identifica en la tabla 5 como SEQ ID NO: 87-107 y tiene al menos 80% de identidad sobre la secuencia madura entera como se identifica en la tabla 3 como SEQ ID NO: 2-21. Preferiblemente, la identidad es al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100%. En otra forma de realización preferida, un isomiR de una molécula de ARNmi tiene al menos 80% de identidad con la secuencia de isomiR entera (la tabla 6 muestra el isomiR preferido de cada uno de los ARNmi maduros identificados como SEQ ID NO: 118-162). Preferiblemente, la identidad es al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95% o más alto tal como 96%, 97%, 98%, 99% o 100%. Preferiblemente en esta forma de realización, un isomiR

de una molécula de ARNmi o un equivalente o un mimético del mismo tiene una longitud de al menos 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 nucleótidos o más.

[0016] La identidad se puede evaluar utilizando diferentes vías tal y como se define más adelante aquí. Sin embargo, en una forma de realización preferida, identidad se refiere al porcentaje de identidad y se calcula por el número de nucleótidos iguales entre base de datos y consulta, dividido por la longitud total de la consulta, y multiplicado por 100. Por consiguiente, una molécula de ARNmi-520f preferida o equivalente o mimético o isomiR de la misma comprende al menos 6 de los 7 nucleótidos presentes en la secuencia semilla identificada como SEQ ID NO: 104 o 105 y/o tiene al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94,95%, 96%, 97%, 98%, 99% o 100% de identidad con la SEQ ID NO: 19, 118, 119 y/o 120 y/o tiene una longitud de al menos 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40 nucleótidos o más. Por consiguiente, una molécula de ARNmi-518b preferida o equivalente o mimético o isomiR de la misma comprende al menos 6 de los 7 nucleótidos presentes en la secuencia semilla identificada como SEQ ID NO: 103 y/o tiene al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94,95%, 96%, 97%, 98%, 99% o 100% de identidad con la SEQ ID NO: 18, 121,122 y/o 123 y/o tiene una longitud de al menos 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40 nucleótidos o más. Por consiguiente, una molécula de ARNmi-524 preferida o equivalente o mimético o isomiR de la misma comprende al menos 6 de los 7 nucleótidos presentes en la secuencia semilla identificada como SEQ ID NO: 106 o 107 y/o tiene al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94,95%, 96%, 97%, 98%, 99% o 100% de identidad con la SEQ ID NO: 20, 21, 124, 125, 126,127 y/o 128 y/o tiene una longitud de al menos 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40 nucleótidos o más. Otra molécula de ARNmi preferida o equivalente o mimético o un isomiR de la misma tiene al menos 80% de identidad con una secuencia semilla (como se identifica en la tabla 5 como SEQ ID NO: 87-107) o con una secuencia madura (como se identifica en la tabla 3 como SEQ ID NO: 2-21) o con una secuencia precursora (como se identifica en la tabla 2 como SEQ ID NO: 22-35) o con un ADN que codifica un precursor de ARN (como se identifica en la tabla 4 como SEQ ID NO: 36-47) o con una secuencia de isomiR (como se identifica en la tabla 6 como SEQ ID NO: 118-162). La identidad puede ser al menos 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% o 100%. La identidad es preferiblemente evaluada en toda la SEQ ID NO según se identifica en una tabla. Sin embargo, la identidad también se puede evaluar en la parte de una SEQ ID NO dada. Parte puede referirse a al menos 50% de la longitud de la SEQ ID NO, al menos 60%, al menos 70%, al menos 80%, al menos 90% o 100%.

[0017] Un equivalente puede ser un isomiR o un mimético. Una secuencia precursora puede dar como resultado más de una secuencias de isomiR dependiendo del proceso de maduración (véase, por ejemplo, ARNmi-520f o ARNmi-518b o ARNmi-524 donde en ciertos tejidos, múltiples isomiRs se han identificado (Tabla 6: SEQ ID NO: 118-128). Un mimético es una molécula que tiene una actividad similar o idéntica a la de una molécula de ARNmi. En este contexto una actividad similar tiene el mismo sentido que un nivel aceptable de una actividad.

[0018] Cada una de las moléculas ARNmi o equivalentes o miméticos o isomiRs de la misma como se identifica aquí tiene un nivel aceptable de una actividad de un ARNmi dado del que derivan. Un nivel aceptable de una actividad es preferiblemente que dicho ARNmi o equivalente o miméticos o isomiRs del mismo sigue siendo capaz de mostrar un nivel aceptable de dicha actividad de dicho ARNmi. Una actividad de un ARNmi dado o un equivalente del mismo es por ejemplo la capacidad para inducir una MET detectable como se definirá más adelante aquí. Un nivel aceptable de una actividad es preferiblemente al menos 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% o 100% de la actividad del ARNmi del que derivan. Tal actividad puede ser según se mide en una célula de vejiga de un individuo o *in vitro* en una célula por comparación con la actividad del ARNmi del que derivan. La evaluación de la actividad se puede realizar en el nivel de ARNm, preferiblemente utilizando RT-qPCR. La evaluación de la actividad se puede realizar en el nivel de proteína, preferiblemente utilizando análisis de transferencia de Western o, análisis de inmunohistoquímica o inmunofluorescencia de secciones transversales. La evaluación de la actividad se puede realizar utilizando células que expresan un constructo de luciferasa de luciérnaga dirigido por el promotor de *CDH1* y midiendo la actividad de luciferasa. Una actividad preferida de cualquiera de las moléculas de ARNmi o equivalente de la misma como se identifica aquí (es decir, ARNmi-124-1, ARNmi-206, ARNmi181a-1, ARNmi-141, ARNmi-200a, ARNmi-200b, ARNmi-200c, ARNmi-429, ARNmi-205, ARNmi-518b, ARNmi-520f, ARNmi-524) o una actividad preferida de una molécula de ARNmi o equivalente de la misma identificada por una fuente preferida (es decir, una molécula de ARNmi o un equivalente de la misma donde una fuente de dicha molécula de ARNmi o equivalente de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1) es para inducir una MET detectable en un sujeto como se define aquí más adelante.

[0019] Una fuente de una molécula de ARNmi o una fuente de un equivalente de una molécula de ARNmi, mimético, isomiR puede ser cualquier molécula que sea capaz de inducir la producción de una molécula de ARNmi o de un equivalente de la misma tal como un mimético o isomiR según se identifica aquí y que comprende una estructura tipo horquilla y/o una molécula de ácido nucleico bicatenaria. La presencia de una estructura tipo horquilla se puede evaluar utilizando el programa RNAshapes (Steffen P., *et al.*, (2006), Bioinformatics, 22: 500-503) usando ventanas deslizantes de 80, 100 y 120 nt o más. La presencia de una

estructura tipo horquilla está normalmente presente en una fuente natural o endógena de una molécula de ARNmi mientras que una molécula de ácido nucleico bicatenaria está normalmente presente en una fuente recombinante o sintética de una molécula de ARNmi o de un equivalente de la misma. Una fuente de una molécula de ARNmi o de un equivalente o un mimético o un isomiR de la misma puede ser un ARN monocatenario, un bicatenario o un ARN parcialmente bicatenario o comprender tres cadenas, un ejemplo esto se describe en la WO2008/10558. Como se utiliza en este caso, parcialmente bicatenario se refiere a estructuras bicatenarias que también comprenden estructuras monocatenarias en los extremos 5' y/o 3'. Puede ocurrir cuando cada cadena de una molécula de ARNmi no tiene la misma longitud. En general, tal molécula de ARNmi bicatenaria parcial puede tener menos del 75% de estructura bicatenaria y más del 25% de estructura monocatenaria, o menos del 50% de estructura bicatenaria y más del 50% de estructura monocatenaria, o más preferiblemente menos del 25%, 20 % o 15% de estructura bicatenaria y más del 75%, 80%, 85% de estructura monocatenaria. Alternativamente, una fuente de una molécula de ARNmi o de un equivalente o un mimético o un isomiR de la misma es una molécula de ADN que codifica un precursor de una molécula de ARNmi o de un equivalente o un mimético o un isomiR de la misma. Las moléculas de ADN preferidas en este contexto se identifican en la tabla 4 como SEQ ID NO: 36-47. La invención abarca el uso de una molécula de ADN que codifica un precursor de una molécula de ARNmi que tiene al menos 80% de identidad con dicha secuencia como se identifica en la tabla 4. Preferiblemente, la identidad es al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100%. Preferiblemente en esta forma de realización, una molécula de ADN tiene una longitud de al menos 50, 55, 60, 70, 75, 80, 85, 90, 95, 100, 130, 150, 200, 250, 300, 350, 400 nucleótidos o más y tiene al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con una secuencia de ADN como se identifica en la tabla 4 como SEQ ID NO: 36-47. Por consiguiente, una fuente preferida de una molécula de ARNmi-520f tiene al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con SEQ ID NO: 45 y/o tiene una longitud de al menos 50, 55, 60, 70, 75, 80, 85, 90, 95, 100, 130, 150, 200, 250, 300, 350, 400 nucleótidos o más. Por consiguiente, una fuente preferida de una molécula de ARNmi-518b tiene al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con SEQ ID NO: 44 y/o tiene una longitud de al menos 50, 55, 60, 70, 75, 80, 85, 90, 95, 100, 130, 150, 200, 250, 300, 350, 400 nucleótidos o más. Por consiguiente, una fuente preferida de una molécula de ARNmi-524 tiene al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con SEQ ID NO: 46 y/o tiene una longitud de al menos 50, 55, 60, 70, 75, 80, 85, 90, 95, 100, 130, 150, 200, 250, 300, 350, 400 nucleótidos o más.

[0020] La inducción de la producción de una molécula de ARNmi dada o de un equivalente de la misma o de un mimético o un isomiR de la misma se obtiene preferiblemente cuando dicha fuente se introduce en una célula usando un ensayo tal y como se define a continuación. Las células abarcadas por la presente invención se definen más adelante. Una fuente preferida de una molécula de ARNmi o de un equivalente de la misma o de un mimético o un isomiR de la misma es un precursor de la misma, más preferiblemente un ácido nucleico que codifica dicha molécula de ARNmi o un equivalente de la misma o un mimético o un isomiR de la misma. Un precursor preferido es un precursor que se produce de forma natural. Un precursor puede ser un precursor sintético o recombinante. Un precursor preferido de una molécula de ARNmi dada se identifica en la tabla 2 como SEQ ID NO: 22-35. La invención abarca el uso de un precursor de una molécula de ARNmi o de un equivalente de la misma que tiene al menos 80% de identidad con dicha secuencia. Preferiblemente, la identidad es al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100%. Preferiblemente en esta forma de realización, una molécula de ARN tiene una longitud de al menos 50, 55, 60, 70, 75, 80, 85, 90, 95, 100, 130, 150, 200, 250, 300, 350, 400 nucleótidos o más y tiene al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con una secuencia según se identifica en la tabla 2 como SEQ ID NO: 22-35.

[0021] Por consiguiente, una fuente preferida de una molécula de ARNmi-520f tiene al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con SEQ ID NO: 33 y/o tiene una longitud de al menos 50, 55, 60, 70, 75, 80, 85, 90, 95, 100, 130, 150, 200, 250, 300, 350, 400 nucleótidos o más.

Por consiguiente, una fuente preferida de una molécula de ARNmi-518b tiene al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con SEQ ID NO: 32 y/o tiene una longitud de al menos 50, 55, 60, 70, 75, 80, 85, 90, 95, 100, 130, 150, 200, 250, 300, 350, 400 nucleótidos o más. Por consiguiente, una fuente preferida de una miRNA-524 molécula tiene al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con SEQ ID NO: 34 y/o tiene una longitud de al menos 50, 55, 60, 70, 75, 80, 85, 90, 95, 100, 130, 150, 200, 250, 300, 350, 400 nucleótidos o más. Fuentes preferidas o precursores se han definido más adelante aquí. Una fuente preferida incluye o comprende un constructo de expresión que comprende un ácido nucleico, es decir, ADN que codifica dicho precursor de dicho ARNmi, más preferiblemente dicho constructo de expresión es un vector de terapia génica vírica seleccionado de vectores de terapia génica basados en un adenovirus, un virus adeno-asociado (AAV), un virus del herpes, un virus de la sífilis y un retrovirus. Un vector de terapia génica vírico preferido es un vector AAV o Lentiviral. Otros vectores preferidos

son vectores virales oncolíticos. Tales vectores se describen más adelante aquí. Alternativamente, una fuente puede ser una molécula de ARNmi sintética o un mimético química como se define más adelante en la parte dedicada a definiciones generales. [0022] La detección de la presencia de una molécula de ARNmi o de un equivalente de la misma tal como un mimético o un isomiR se puede realizar utilizando cualquier técnica conocida por la persona experta. La evaluación del nivel de expresión o de la presencia de una molécula de ARNmi o de un equivalente de la misma se realiza preferiblemente utilizando técnicas de biología molecular tradicionales tales como qPCR (tiempo real), micromatrices, matrices de esferas, análisis de protección de ribonucleasa o análisis Northern o clonación y secuenciación. La persona experta entenderá que alternativamente o en combinación con la cuantificación de una molécula de ARNmi o de un equivalente de la misma, la cuantificación de un sustrato de una molécula de ARNmi correspondiente o de un equivalente de la misma o cualquier compuesto conocido por estar asociado con una función de dicha molécula de ARNmi o de dicho equivalente de la misma o la cuantificación de una función o actividad de dicha molécula de ARNmi o de dicho equivalente de la misma utilizando un ensayo específico está abarcado dentro del campo de la invención. Las composiciones y formulaciones preferidas se definen todas más adelante aquí. Una molécula de ARNmi o un equivalente de la misma o un mimético o un isomiR de la misma se puede utilizar como tal como una molécula desnuda, con o sin modificaciones químicas, o encapsulada en una partícula o conjugada con una fracción. Una composición preferida comprende una molécula de ARNmi o un equivalente de la misma o un mimético o un isomiR de la misma encapsulada en una nanopartícula o una estructura liposómica. Una molécula de ARNmi o equivalente de la misma o un mimético o un isomiR de la misma puede ser un híbrido de aptámero-ARNmi. Una molécula de ARNmi o equivalente de la misma puede ser un híbrido de aptámero-ARNmi. Un aptámero-ARNmi se define como un ARNmi enlazado a un nucleótido de ARN (o ADN), el último adopta una conformación que dirige la molécula del híbrido de aptámero-ARNmi a una proteína de superficie celular (por ejemplo, el antígeno de membrana específico de la próstata (PSMA)). El ARNmi marcado con el aptámero se puede enlazar a por ejemplo polietilenglicol, lo que aumenta la vida media circulante de la quimera (Dassie, J.P., *et al.* Nat. Biotechnol. 27: 839-849 (2009)). Cualquier cáncer en el que la EMT esté implicada o asociada, se puede evitar, retardar, curar y/o tratar con una molécula tal y como se define aquí. En el contexto de la invención, la Transición Epitelial-Mesenquimal (EMT) es una serie orquestada de eventos en los que las interacciones de la matriz célula-célula y célula-extracelular (ECM) se alteran para permitir liberar células epiteliales del tejido circundante. El citoesqueleto celular epitelial se reorganiza para conferir la capacidad de la célula para moverse a través de una ECM tridimensional vía la reprogramación molecular de la célula. La reprogramación molecular de una célula epitelial es necesaria para conseguir un fenotipo mesenquimal e implica la infrarregulación o la disminución de la expresión de las proteínas epiteliales, tal como E-cadherina y proteínas de unión tales como desmoplaquina, claudina y ocludina. Además, la expresión de proteínas mesenquimales es sobreexpresada o aumentada, incluyendo por ejemplo, la expresión de proteínas de ECM tales como MMPs y fibronectina y proteínas de superficie celular tales como N-Cadherina e integrina $\alpha\text{v}\beta\text{3}$. Factores de transcripción también se pueden sobreexpresar o aumentar en las células que muestran un fenotipo mesenquimal tal como por ejemplo, SNAI1 (conocido también como SNAIL), TWIST, ZEB1 (conocido también como δEF1) y ZEB2 (conocido también como SIP1). La referencia a la inducción de la "transición" de una célula epitelial a una célula que muestra un fenotipo mesenquimal debería ser entendida como una referencia para inducir los cambios genéticos, morfológicos y/o funcionales que son necesarios para cambiar una célula epitelial a una célula que muestra un fenotipo mesenquimal del tipo definido aquí. La referencia a la inducción de la transición mesenquimal a epitelial debería ser entendida como que tiene el significado inverso.

[0023] En un cáncer de la invención, la EMT puede ser detectable antes de la aparición del cáncer, es decir, antes de la aparición de un síntoma de dicho cáncer. Además también está abarcado por la presente invención que la EMT sea detectable durante el desarrollo de dicho cáncer, es decir después de la aparición de un síntoma de dicho cáncer. Además también está abarcado que la EMT sea detectable antes de la aparición del cáncer y durante el desarrollo de dicho cáncer. La EMT se puede detectar utilizando cualquier técnica conocida por la persona experta. Preferiblemente, la EMT es evaluada detectando una reducción de la expresión de E-cadherina epitelial y/o un aumento de la expresión de vimentina mesenquimal y/o cadherina mesenquimal usando inmunohistoquímica usando anticuerpos específicos dirigido contra E-cadherina y/o vimentina mesenquimal y/o cadherina mesenquimal respectivamente (2, 3). La N-Cadherina (CDH2) y la OB-cadherina (CDH11) son dos ejemplos de cadherinas mesenquimales. La evaluación de la expresión se realiza preferiblemente en una biopsia o sección tumoral en diferentes puntos temporales para un sujeto dado o en uno o más puntos temporales para un sujeto dado y un control sano. La evaluación se puede realizar cada semana, cada mes. El aumento/reducción se puede evaluar por lo tanto cada semana, mes. Preferiblemente, una reducción de la expresión de la E-cadherina epitelial se refiere a una reducción significativa, preferiblemente una reducción de al menos 5% de la expresión usando inmunohistoquímica. Más preferiblemente, una reducción se refiere a una reducción de al menos 10%, aún más preferiblemente al menos 20%, al menos 30%, al menos 40%, al menos 50%, al menos 70%, al menos 90% o 100%. En este caso, ninguna expresión es detectable. Un aumento de 25 veces de E-cadherina se obtuvo utilizando una molécula de ARNmi-520f como se identifica aquí (un ejemplo representativo se muestra en el ejemplo 2, figura 4B). El efecto de esta molécula de ARNmi en este marcador es de forma cuantitativa (al menos 1,5 veces) más pronunciado que el efecto correspondiente de la molécula de ARNmi de la familia 200 que ya se sabía que tiene un efecto en la E-cadherina como se muestra en el ejemplo 2, figura 4B. Por lo tanto se puede anticipar que una molécula de ARNmi-520f según se identifica aquí se puede considerar como una molécula atractiva para un uso según se identifica aquí. Preferiblemente, un aumento de la

expresión de vimentina mesenquimal y/o cadherina mesenquimal significa un aumento significativo, preferiblemente un aumento de al menos 5% de la expresión usando inmunohistoquímica. Más preferiblemente, un aumento significa un aumento de al menos 10%, aún más preferiblemente al menos 20%, al menos 30%, al menos 40%, al menos 50%, al menos 70%, al menos 90%, al menos 100%, o al menos 150% o más. Un cáncer en el que la EMT está implicada o asociada es preferiblemente un cáncer en el que se produce un proceso de dediferenciación. En este proceso de dediferenciación, una reducción de la expresión de E-cadherina epitelial y/o un aumento de la expresión de vimentina mesenquimal y/o cadherina mesenquimal se produce preferiblemente y se puede evaluar como se ha explicado aquí. Como ejemplo, la EMT puede ser mostrada por células cancerosas que sufren este proceso y se convierten por lo tanto en metastásicas debido a su capacidad para separarse de las células vecinas y penetrar en y a través de los tejidos circundantes. Por lo tanto, una enfermedad preferida de la invención es un cáncer (por ejemplo, maligno, metastásico) y otra enfermedad más de la divulgación es una fibrosis. Los cánceres de una forma de realización preferida de la invención incluyen un cáncer de origen epitelial o un carcinoma o un tumor sólido. Las células cancerosas pueden ser de vejiga, cerebro, pecho, colon, esófago, gastrointestinales, cabeza, riñón, hígado, nasofaringe, cuello, ovario, próstata, piel, estómago, testículo, lengua o útero. Además, el cáncer puede específicamente ser del siguiente tipo histológico, aunque no está limitado a estos: neoplasma, maligno; carcinoma; carcinoma, no diferenciado; carcinoma de célula gigante y fusiforme; carcinoma de célula pequeña; carcinoma papilar; carcinoma de célula escamosa; carcinoma de célula basal; carcinoma pilomatricial; carcinoma de célula transicional; carcinoma de célula transicional papilar; adenocarcinoma; gastrinoma, maligno; colangiocarcinoma; carcinoma hepatocelular; carcinoma hepatocelular y colangiocarcinoma combinado; adenocarcinoma trabecular; carcinoma cístico adenoideo; adenocarcinoma de pólipos adenomatosos; adenocarcinoma, poliposis coli familiar; carcinoma sólido; tumor carcinoide, maligno; adenocarcinoma branquiol-alveolar; adenocarcinoma papilar; carcinoma cromóforo; carcinoma acidófilo; adenocarcinoma oxifílico; carcinoma de basófilos; adenocarcinoma de célula clara; carcinoma de célula granulosa; adenocarcinoma folicular; adenocarcinoma papilar y folicular; carcinoma esclerosante no encapsulante; carcinoma cortical suprarrenal; carcinoma endometriode; carcinoma de apéndice de piel; adenocarcinoma apocrino; adenocarcinoma sebáceo; adenocarcinoma ceruminoso; carcinoma mucoepidermoide; cistadenocarcinoma; cistadenocarcinoma papilar; cistadenocarcinoma seroso papilar; cistadenocarcinoma mucinoso; adenocarcinoma mucinoso; carcinoma de célula en anillo de sello; carcinoma de conducto de infiltración; carcinoma medular; carcinoma lobular; carcinoma inflamatorio; enfermedad de Paget de mama; carcinoma de célula acinar; carcinoma adenoescamoso; adenocarcinoma con metaplasia escamosa; tumor estromal ovárico, maligno; androblastoma, maligno; carcinoma de células de Sertoli. En una forma de realización preferida, el cáncer asociado a la EMT es un cáncer, preferiblemente un cáncer de vejiga o de próstata. La EMT no se produce en todos los cánceres. En los sarcomas de tejido blando y las leucemias el proceso de dediferenciación de la EMT no se produce. Por lo tanto en una forma de realización preferida, un cáncer como se identifica aquí es un carcinoma y/o no es una leucemia y/o no es un sarcoma de tejido.

[0024] La EMT según la divulgación también puede producirse durante la inflamación crónica o condiciones que promueven la interrupción de tejido prolongada que pueden estimular la fibrosis, comprometiendo así la integridad del tejido y función del órgano. Una fibrosis se conoce también como fibrosis de órgano o degeneración de órgano (mencionada en Thierry J. P. *et al.*, 2009, Cell, 139: 871-890). En tejidos fibróticos, los miofibroblastos se acumulan y segregan una cantidad excesiva de colágeno que se deposita como fibras, comprometiendo así la función del órgano y llevando a su fallo. La fibrosis se origina de la conversión de una parte significativa de células epiteliales en miofibroblastos a través de un proceso de EMT (Iwano *et al.*, 2002 J. Clin. Invest. 110:341-50). Inicialmente demostrado en células diferenciadas de túbulos y conductos renales, ahora está claro que el epitelio, endotelio, hepatocitos y cardiomiocitos de la lente pueden todos ser sometidos a EMT y contribuir significativamente a fibrosis tisular. Cada fibrosis en la que la EMT se supone o se sospecha que se produce está abarcada dentro del campo de la divulgación. Una condición o enfermedad asociada a la EMT según la divulgación también puede ser la cicatrización pobre de una herida, nefropatía renal diabética, disfunción de aloinjerto, cataratas o defectos en la formación de válvula cardíaca.

[0025] Actualmente no existe ningún medicamento conocido eficaz que se pueda utilizar para prevenir, tratar, revertir, curar y/o retrasar específicamente un cáncer asociado a la EMT en un sujeto. La invención abarca el uso de una molécula de ARNmi o un equivalente de la misma donde una fuente de dicha molécula de ARNmi o equivalente de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1 o una composición que comprende dicha molécula de ARNmi o equivalente de la misma o una fuente de la misma para este fin. Las moléculas de ARNmi preferidas o equivalentes o miméticos o isomiR o fuentes de las mismas ya se han definido aquí.

[0026] Este uso incluye el aumento farmacológicamente de una actividad o el nivel de estado estable de dicha molécula de ARNmi o equivalente de la misma o de dicha fuente de la misma en un sujeto, en una célula de dicho sujeto, en un tejido de dicho sujeto o en el fluido corporal de dicho sujeto. En este uso, una actividad o nivel de estado estable de dicho ARNmi, o equivalente del mismo o fuente del mismo se aumenta para inducir una MET detectable en un sujeto. Una MET, transición mesenquimal a epitelial es lo contrario de una EMT. Por lo tanto, la inducción de la MET es idéntica a la reversión de la EMT. Preferiblemente, una MET se evalúa detectando un aumento de la expresión de E-cadherina epitelial y/o una reducción de la expresión de vimentina mesenquimal usando inmunohistoquímica utilizando un anticuerpo específico dirigido contra E-cadherina,

vimentina mesenquimal y/o cadherina mesenquimal (2, 3) respectivamente. La evaluación de la expresión se realiza preferiblemente en una biopsia o sección tumoral en diferentes puntos temporales para un sujeto dado o en uno o más puntos temporales para un sujeto dado y un control sano. La evaluación se puede realizar en intervalos de tiempo regulares, por ejemplo cada semana, cada mes. El aumento/reducción se puede evaluar por lo tanto regularmente, por ejemplo cada semana, cada mes. Una MET se ha detectado preferiblemente cuando durante al menos un punto temporal, se ha detectado un aumento de la expresión de E-cadherina epitelial y/o una reducción de la expresión de vimentina mesenquimal y/o cadherina mesenquimal. Preferiblemente, una MET se ha detectado cuando durante al menos dos, tres, cuatro, cinco puntos temporales tal aumento de la expresión de E-cadherina epitelial y/o reducción de la expresión de vimentina mesenquimal y/o cadherina mesenquimal se han detectado. Preferiblemente, un aumento de la expresión de E-cadherina epitelial significa un aumento significativo, preferiblemente un aumento de al menos 5% de la expresión usando inmunohistoquímica. Más preferiblemente, un aumento significa un aumento de al menos 10%, aún más preferiblemente al menos 20%, al menos 30%, al menos 40%, al menos 50%, al menos 70%, al menos 90%, al menos 150% o más. Preferiblemente, una reducción de la expresión de la vimentina mesenquimal y/o de la cadherina mesenquimal significa una reducción significativa, preferiblemente una reducción de al menos 5% de la expresión usando inmunohistoquímica. Más preferiblemente, una reducción significa una reducción de al menos 10%, aún más preferiblemente al menos 20%, al menos 30%, al menos 40%, al menos 50%, al menos 70%, al menos 90% o 100%. En este caso, ninguna expresión es detectable.

[0027] Una actividad o nivel de estado estable de dicha molécula de ARNmi o un equivalente o un mimético o un isomiR de la misma donde una fuente de dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1 se puede aumentar en el nivel de la molécula de ARNmi (o equivalente o mimético o isomiR de la misma) en sí, por ejemplo proporcionando dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma a un sujeto, preferiblemente a una célula de un sujeto, o a un tejido de dicho sujeto, o a un órgano de dicho sujeto o a dicho sujeto dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma siendo de una fuente exógena. Para proporcionar una molécula de ARNmi o equivalente o mimético o isomiR de la misma a partir de una fuente exógena, dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma se puede producir convenientemente por expresión de un ácido nucleico que codifica dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma o que codifica una fuente de dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma en una célula huésped adecuada como se describe más adelante o como moléculas completamente sintéticas por síntesis química.

[0028] Preferiblemente, sin embargo, una actividad o nivel de estado estable de una molécula de ARNmi o equivalente o un mimético o un isomiR de la misma se aumenta por el reglaje del nivel de expresión de una secuencia de nucleótidos que codifica dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma o que codifica una fuente de dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma. Preferiblemente, el nivel de expresión de una secuencia de nucleótidos se regula en una célula de dicho sujeto o en un tejido de dicho sujeto o en el sujeto. El nivel de expresión de una molécula de ARNmi o equivalente o mimético o isomiR de la misma o una fuente de dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma se puede aumentar por introducción de una molécula de ARNmi, o equivalente o mimético o isomiR de la misma, o una fuente de la misma, o un constructo de expresión (o vector) en una célula, tejido, órgano o fluido corporal de dicho sujeto, o en el sujeto por lo que un vector de expresión comprende una secuencia de nucleótidos que comprende una molécula de ARNmi o equivalente o un mimético o un isomiR de la misma o comprende una fuente de dicha molécula de ARNmi o equivalente o un mimético o un isomiR de la misma, y por lo cual una secuencia de nucleótidos está bajo el control de un promotor capaz de dirigir la expresión de una secuencia de nucleótidos en dicha célula, tejido, órgano, sujeto. El nivel de expresión de una molécula de ARNmi o equivalente o un mimético o un isomiR de la misma o fuente de la misma también se puede aumentar por introducción de un constructo de expresión en una célula, tejido, órgano, sujeto, por lo cual un constructo comprende una secuencia de nucleótidos que codifica un factor capaz de trans-activación de una secuencia de nucleótidos endógena que codifica una molécula de ARNmi o equivalente o mimético o isomiR de la misma.

[0029] Un uso de la invención preferiblemente comprende la etapa de administrar a un sujeto una cantidad terapéuticamente eficaz de una composición farmacéutica que comprende un constructo de ácidos nucleicos para aumentar la actividad o el nivel estable de una molécula de ARNmi o un equivalente o un mimético o un isomiR de la misma donde una fuente de dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1 o una fuente de la misma como se identifica aquí. Un constructo de ácidos nucleicos puede ser un constructo de expresión como se especifica más aquí. Preferiblemente, un constructo de expresión es un vector de terapia génica viral seleccionado de vectores de terapia genética basados en un adenovirus, un virus adeno-asociado (AAV), un virus del herpes, un virus de la sífilis, un vector viral oncolítico y un retrovirus. Un vector de terapia génica viral preferido es un vector AAV o Lentiviral. Alternativamente, un uso de la invención preferiblemente comprende la etapa de administración a un sujeto de una cantidad terapéuticamente eficaz de una composición farmacéutica que comprende una molécula de ARNmi o un equivalente o un mimético o un isomiR de la misma donde una fuente de dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma comprende al menos 80 nucleótidos y comprende una unidad con

al menos 98% de identidad con la unidad representada por SEQ ID NO: 1 o una fuente de la misma tal y como se define aquí.

[0030] En un uso de la invención, una célula, un tejido, un órgano o fluido corporal es preferiblemente de un sujeto que se sospecha que tiene un alto riesgo de tener un cáncer asociado a la EMT debido por ejemplo a su edad o a su contexto genético o a su dieta. Alternativamente, en otra forma de realización preferida, el uso de la invención se aplica a una célula, tejido, órgano o fluido corporal de un sujeto diagnosticado con un riesgo predictivo para desarrollar más tarde un cáncer asociado a la EMT. Un método de diagnóstico usado es preferiblemente una de las divulgaciones como se describe en este caso. Alternativamente, una célula, un tejido u órgano a tratar se puede seleccionar basado en el riesgo de progresión del cáncer asociado a la EMT. Tal riesgo de progresión se puede evaluar utilizando criterios patológicos clínicos tradicionales o prognosis basada en biomarcador conocidos por la persona experta. Está abarcado también por la invención la administración de una molécula de ARNmi o un equivalente o un mimético o un isomiR de la misma donde una fuente de dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1 o un precursor de la misma o una composición que comprende dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma o la fuente de la misma en un tejido o órgano de dicho sujeto. Las moléculas de ARNmi preferidas o equivalentes o miméticos o isomiR o fuentes de la misma ya se han definido aquí. En la invención, una célula preferida, tejido u órgano es una célula, tejido u órgano que es o comprende una célula o tejido de vejiga o próstata o es o comprende la vejiga o próstata como órgano.

[0031] Un tratamiento de un cáncer asociado a la EMT puede incluir un tratamiento que previene la EMT en una célula tumoral que aún no se ha metastatizado o revierte la EMT (definida como transición mesenquimal a epitelial) en una célula tumoral que ya ha formado metástasis y/o está migrando desde el tumor primario a sitios distantes en el cuerpo.

[0032] En otro uso, la invención mencionada aquí se puede combinar con tratamientos estándar de cáncer asociado a EMT tal como quimioterapia, radioterapia o cirugía. Ejemplos de agentes quimioterapéuticos se ejemplifican más adelante aquí. Aunque la terapia genética es una posibilidad para prevenir, tratar, revertir y/o retrasar un cáncer asociado a EMT, otros tratamientos posibles también se pueden prever. Por ejemplo, el tratamiento por fármacos de "molécula pequeña" para dirigir determinadas vías moleculares en la dirección deseada, también se prefieren. Estas pequeñas moléculas se identifican preferiblemente por el método de selección de la invención tal y como se define más adelante aquí.

[0033] En el contexto de la invención, prevenir, tratar, revertir, curar y/o retrasar un cáncer asociado a EMT puede significar que:

- al menos un síntoma de este cáncer ha sido mejorado, y/o
- al menos un parámetro asociado a este cáncer ha sido mejorado.

Un síntoma puede ser la presencia de metástasis como se explica más adelante. Un parámetro puede ser la evaluación de MET como se ha explicado anteriormente aquí. En el contexto de la invención, prevenir, tratar, revertir, curar y/o retrasar un cáncer asociado a la EMT se puede sustituir alcanzando un efecto antitumoral. A menos que se indique lo contrario, un efecto antitumoral se evalúa o detecta preferiblemente después de al menos una semana, dos semanas, tres semanas, cuatro semanas, un mes, dos meses, tres meses, cuatro meses, cinco meses, seis meses o más en un sujeto tratado. Un efecto antitumoral se identifica preferiblemente en un sujeto como:

- una inhibición de la proliferación de células tumorales y/o
- un aumento en la capacidad de diferenciación de células tumorales y/o
- una inducción o inducción aumentada de muerte de células tumorales y/o
- un retraso en la aparición de metástasis y/o de migración de células tumorales y/o
- una inhibición o prevención o retraso del aumento de un peso o crecimiento tumoral y/o
- una prolongación de la supervivencia del paciente de al menos un mes, varios meses o más (en comparación con aquellos no tratados o tratados con un control o en comparación con el sujeto a inicio del tratamiento).

En el contexto de la invención, un paciente puede sobrevivir y/o se puede considerar como libre de enfermedad. Alternativamente, el cáncer se puede detener o retrasar. Una inhibición de la proliferación de células tumorales puede ser de al menos 20%, 30%, 40%, 50%, 55%, 60%, 65%, 70% o 75% o más. La proliferación de células se puede evaluar utilizando técnicas conocidas. Una inducción de la muerte de células tumorales puede ser de al menos 1%, 5%, 10%, 15%, 20%, 25% o más. El crecimiento tumoral se puede inhibir al menos 5%, 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 70% o 75% o más. La muerte de células tumorales se puede evaluar utilizando técnicas conocidas por la persona experta. La muerte de células tumorales se puede evaluar utilizando MRI (formación de imágenes por resonancia magnética) o CT (tomografía computarizada). En formas de

realización determinadas, el aumento del peso del tumos o el crecimiento del tumor se puede inhibir al menos 20%, 30%, 40%, 50%, 55%, 60%, 65%, 70% o 75% o más. El peso del tumor o el crecimiento del tumor se pueden evaluar utilizando técnicas conocidas por la persona experta. La detección de crecimiento del tumor o la detección de la proliferación de células tumorales se puede evaluar *in vivo* por la medición de cambios en la utilización de glucosa por tomografía de emisión de positrón con el análogo de glucosa 2-[18F]-fluor-2-deoxi-D-glucosa (FDG-PET) o [18F]-3-fluoro-3-deoxi-L-timidina PET. Una alternativa *ex vivo* puede ser la coloración de una biopsia tumoral con Ki67.

[0034] Un retraso en la aparición de la metástasis y/o de la migración de células tumorales puede ser un retraso de al menos una semana, un mes, varios meses, un año o más tiempo. La presencia de metástasis se puede evaluar utilizando MRI, CT o ecografía o técnicas que permitan la detección de células tumorales circulantes (CTC). Ejemplos de las últimas pruebas son la prueba CellSearch CTC (Veridex), una selección magnética basada en EpCam de CTCs de sangre periférica. En ciertas formas de realización, el crecimiento tumoral se puede retardar al menos una semana, un mes, dos meses o más. En una forma de realización determinada, la aparición de metástasis se retarda al menos una semana, dos semanas, tres semanas, cuatro semanas, un mes, dos meses, tres meses, cuatro meses, cinco meses, seis meses o más. Una molécula de ARNmi o un equivalente o un mimético o un isomiR de la misma o una fuente de la misma donde una fuente de dicha molécula de ARNmi o equivalente o mimético o isomiR de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1 se ha descubierto sorprendentemente que retrasa la aparición de metástasis y/o la migración de células tumorales de una forma más pronunciada que el efecto correspondiente de la molécula de ARNmi de la familia 200 ya se sabía que tenía tal efecto como se muestra en el ejemplo 2, figura 6D. Por lo tanto, podemos anticipar que una molécula de ARNmi o una fuente de la misma como se identifica aquí se puede considerar como una molécula atractiva para su uso según se identifica aquí.

[0035] Un aumento en la capacidad de diferenciación de las células tumorales se pueden evaluar utilizando un marcador de diferenciación específico y tras la presencia de tal marcador en las células tratadas. Los marcadores o parámetros preferidos ya han sido identificados aquí, es decir, los marcadores asociados a la MET. Esto se puede conseguir usando RT-PCR, transferencia de Western o inmunohistoquímica. Un aumento de la capacidad de diferenciación puede ser al menos un aumento detectable después al menos una semana de tratamiento utilizando cualquiera de las técnicas identificadas. Preferiblemente, el aumento es de 1%, 5%, 10%, 15%, 20%, 25% o más, lo que significa que el número de células diferenciadas dentro de una muestra dada aumentará en consecuencia. En formas de realización determinadas, el crecimiento del tumor se puede retardar al menos una semana, un mes, dos meses o más. En una forma de realización determinada, la aparición de metástasis se retarda al menos una semana, dos semanas, tres semanas, cuatro semanas, un mes, dos meses, tres meses, cuatro meses, cinco meses, seis meses o más.

[0036] En otra forma de realización preferida, se proporciona una composición que comprende además otra molécula de ARNmi seleccionada de:

a) al menos uno de ARNmi-124-1, ARNmi-206, ARNmi-181a-1, ARNmi-141, ARNmi-200a, ARNmi-200b, ARNmi-200c, ARNmi-429 y ARNmi-205 y/o un equivalente o un mimético o un isomiR o una fuente del mismo.

[0037] Dado que no cada una de las moléculas de ARNmi identificadas o equivalentes o isomiRs o miméticos de la misma está previsto que tenga los mismos genes diana, se supone que el uso de una molécula de ARNmi o un equivalente o un isomiR o un mimético de la misma donde una fuente de dicha molécula de ARNmi o equivalente o isomiR o mimético de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1 opcionalmente combinada con al menos una de la molécula de ARNmi, o equivalente o isomiR o mimético de la misma o fuente de la misma identificada anteriormente bajo a) permite un tratamiento más eficaz de un cáncer asociado a la EMT. Las moléculas de ARNmi preferidas o equivalentes o miméticos o isomiR o fuentes de las mismas ya se han definido aquí. Un tumor tratado por una composición o un cóctel de al menos una molécula de ARNmi o un equivalente o un isomiR o un mimético de la misma donde una fuente de dicha molécula de ARNmi o equivalente de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1 está previsto que tenga menos posibilidades de escapar o resistir a dicho tratamiento. En otra forma de realización preferida, se abarca el diagnóstico de la expresión de cada una de las moléculas ARNmi o de sus genes diana según se identifica aquí y dependiendo del resultado la adaptación de la identidad de las moléculas ARNmi usadas al tratamiento.

[0038] Cuando la invención se refiere a una composición que comprende más de una molécula de ARNmi o equivalente o isomiR o mimético de la misma o fuente de la misma, se abarca que cada molécula de ARNmi o equivalente o isomiR o mimético de la misma o fuente de la misma pueda estar presente en cada composición separada, cada composición se administra consecutivamente o simultáneamente a un sujeto. Alternativamente, también se abarca que más de una molécula de ARNmi o equivalente o isomiRs o miméticos de las mismas o fuentes de las mismas esté presente en una composición tal y como se define aquí.

[0039] En otro aspecto, se proporciona el uso de una molécula de ARNmi o un equivalente o un isomiR o un mimético de la misma donde una fuente de dicha molécula de ARNmi o equivalente o isomiR o mimético de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1 o una fuente de la misma o una composición que comprende dicha molécula de ARNmi, un equivalente o isomiR o mimético o una fuente de la misma para la producción de un medicamento para prevenir, tratar, revertir, curar y/o retrasar un cáncer asociado a la EMT. Cada característica de este otro aspecto se ha descrito ya aquí.

[0040] En otro aspecto de la divulgación, se proporciona un método para prevenir, tratar, revertir, curar y/o retrasar un asociado a la EMT por la administración de una molécula de ARNmi o equivalente o isomiR o mimético de la misma o fuente de la misma o composición como se ha definido en la presente anteriormente a un sujeto que lo necesita. Cada característica de este otro aspecto ya se ha descrito aquí. En otro aspecto de la divulgación, se proporciona un método para diagnosticar la EMT o una enfermedad o condición asociada a la EMT en un sujeto, el método incluye las etapas de:

- (a) determinar el nivel de expresión de una molécula de ARNmi o un equivalente o isomiR o mimético de la misma donde una fuente de dicha molécula de ARNmi o equivalente o isomiR o mimético de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1 o una fuente de la misma en un sujeto, y opcionalmente
- (b) comparar el nivel de expresión de dicha molécula o equivalente de la misma o fuente de la misma tal y como se define en (a) con un valor de referencia para el nivel de expresión de dicha molécula, equivalente isomiR, mimético o fuente de la misma, el valor de referencia preferiblemente es el valor medio para el nivel de expresión de dicha molécula, equivalente, isomiR, mimético o fuente de la misma en un sujeto sano.

En el contexto de la divulgación, diagnóstico significa bien una estimación de riesgos predictiva de un sujeto para el desarrollo de una enfermedad o condición asociada a la EMT o para el desarrollo de la propia EMT. En el contexto de la divulgación, un sujeto puede ser un animal. Preferiblemente un sujeto es un mamífero. El mamífero preferido es un ser humano. Preferiblemente, un sujeto es un ser humano. Ya que los niveles de expresión de estas secuencias de nucleótidos y/o cantidades de molécula de ARNmi correspondientes o equivalentes o isomiR o miméticos de las mismas o fuente de las mismas pueden ser difíciles de medir en un sujeto, se usa preferiblemente una muestra de un sujeto. Según otra forma de realización preferida, el nivel de expresión (de una secuencia de nucleótidos o molécula de ARNmi o equivalente o isomiR o mimético o fuente de la misma) se determina *ex vivo* en una muestra obtenida de un sujeto. La muestra preferiblemente comprende un fluido corporal de un sujeto. Un fluido corporal puede comprender o derivar de la sangre, suero, plasma, deposición, orina o una biopsia de tejido o una biopsia tumoral o un tejido canceroso de origen epitelial de un sujeto. El tejido preferido es vejiga o próstata. Concretamente se contempla que la invención puede utilizarse para evaluar o diagnosticar diferencias entre etapas de la enfermedad o condición asociada a la EMT, tal como entre pre-cáncer y cáncer, o entre un tumor primario y un tumor metastatizado.

[0041] Un aumento o reducción en el nivel de expresión de una secuencia de nucleótidos (o nivel estable de la molécula de ARNmi codificada o equivalente o isomiR o mimético o fuente de la misma) se define preferiblemente como un cambio detectable del nivel de expresión de un nucleótido (o nivel estable de una molécula de ARNmi codificada o equivalente o isomiR o mimético o fuente de la misma o cualquier cambio detectable en una actividad biológica de una molécula de ARNmi o equivalente o isomiR o mimético o fuente de la misma) utilizando un método tal y como se ha definido anteriormente en comparación con el nivel de expresión de una secuencia de nucleótidos correspondiente (o nivel estable de una molécula de ARNmi codificada correspondiente o equivalente o isomiR o mimético o fuente de la misma) en un sujeto sano. Una secuencia de nucleótidos preferida es una secuencia que codifica un precursor de una molécula de ARNmi o equivalente o isomiR o mimético de la misma o una secuencia precursora de una molécula de ARNmi o equivalente o isomiR o mimético de la misma. Según una forma de realización preferida, un aumento o reducción de una actividad ARNmi se cuantifica utilizando un ensayo específico para una actividad de ARNmi. Un ensayo preferido es la evaluación de MET como se ha definido en la presente anteriormente.

[0042] Preferiblemente, una reducción del nivel de expresión de una secuencia de nucleótidos significa una reducción de al menos 5% del nivel de expresión de la secuencia de nucleótidos usando matrices. Más preferiblemente, una reducción del nivel de expresión de una secuencia de nucleótidos significa una reducción de al menos 10%, aún más preferiblemente al menos 20%, al menos 30%, al menos 40%, al menos 50%, al menos 70%, al menos 90% o 100%. En este caso, no hay expresión detectable.

[0043] Preferiblemente, una reducción del nivel de expresión de una molécula de ARNmi o equivalente o isomiR o mimético o fuente de la misma significa una reducción de al menos 5% del nivel de expresión del ARNmi usando qPCR, micromatrices o análisis de Northernblot. Preferiblemente qPCR es RT qPCR de tallo-bucle. Más preferiblemente, una reducción del nivel de expresión de una molécula de ARNmi o equivalente o isomiR o mimético o fuente de la misma significa una reducción de al menos 10%, aún más preferiblemente al menos

20%, al menos 30%, al menos 40%, al menos 50%, al menos 70%, al menos 90% o 100 %. En este caso, no hay expresión detectable.

5 [0044] Preferiblemente, una reducción de una actividad de ARNmi se refiere a una reducción de al menos 5% de una actividad de ARNmi utilizando un ensayo adecuado. Más preferiblemente, una reducción de una actividad de ARNmi se refiere a una reducción de al menos 10%, aún más preferiblemente al menos 20%, al menos 30%, al menos 40%, al menos 50%, al menos 70%, al menos 90% o 100%. En este caso, no hay expresión detectable.

10 [0045] Preferiblemente, un aumento del nivel de expresión de una secuencia de nucleótidos significa un aumento de al menos 5% del nivel de expresión de la secuencia de nucleótidos utilizando cualquiera de las técnicas mencionadas aquí. Más preferiblemente, un aumento del nivel de expresión de una secuencia de nucleótidos significa un aumento de al menos 10%, aún más preferiblemente al menos 20%, al menos 30%, al menos 40%, al menos 50%, al menos 70%, al menos 90%, al menos 150% o más.

15 [0046] Preferiblemente, un aumento del nivel de expresión de una molécula de ARNmi o equivalente o isomiR o mimético o fuente de la misma significa un aumento de al menos 5% del nivel de expresión de la molécula de ARNmi o equivalente o isomiR o mimético o fuente de la misma usando RT-qPCR, preferiblemente RT qPCR tallo-bucle. Más preferiblemente, un aumento del nivel de expresión de una molécula de ARNmi o equivalente o isomiR o mimético o fuente de la misma significa un aumento de al menos 10%, aún más preferiblemente al menos 20%, al menos 30%, al menos 40%, al menos 50%, al menos 70%, al menos 90%, al menos 150% o más.

20 [0047] Preferiblemente, un aumento de una actividad de ARNmi significa un aumento de al menos 5% de una actividad de ARNmi utilizando un ensayo adecuado. Más preferiblemente, un aumento de una actividad de ARNmi significa un aumento de al menos 10%, aún más preferiblemente al menos 20%, al menos 30%, al menos 40%, al menos 50%, al menos 70%, al menos 90%, al menos 150% o más.

25 [0048] Preferiblemente, un nivel de expresión se determina *ex vivo* en una muestra obtenida de un sujeto. Más preferiblemente, la muestra es tal y como se ha definido en la presente anteriormente y donde posteriormente, una secuencia de nucleótidos dada y/o molécula de ARNmi o equivalente o isomiR o mimético o fuente de la misma es extraída y purificada usando métodos conocidos a la persona experta. Más preferiblemente, la muestra es o comprende o se deriva de una biopsia tumoral, sangre u orina.

30 [0049] En un método de diagnóstico de la divulgación preferiblemente el nivel de expresión de más de uno, más preferiblemente de al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 moléculas de ARNmi o equivalentes o isomiRs o miméticos o fuentes de las mismas y/o los niveles estables de la molécula de ARNmi correspondientes o equivalentes o isomiRs o miméticos o fuentes de las mismas son determinados.

35 [0050] Por consiguiente en un método preferido, en la etapa (a) se determina el nivel de expresión de otra molécula de ARNmi o equivalente o fuente de la misma seleccionada de:

(a) al menos uno de ARNmi-124-1, ARNmi-206, ARNmi-181a-1, ARNmi-141, ARNmi-200a, ARNmi-200b, ARNmi-200c, ARNmi-429 and ARNmi-205 y/o un equivalente o una fuente de la misma.

40 [0051] En otro método preferido, la EMT o una enfermedad o condición asociada a la EMT se diagnostica cuando la comparación conduce al hallazgo de una reducción del nivel de expresión de dicha molécula de ARNmi o un equivalente o un mimético o un isomiR de la misma donde una fuente de dicha molécula de ARNmi o equivalente de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1. Las moléculas de ARNmi más preferidas o unos equivalentes o miméticos o isomiRs o fuentes de las mismas se han definido aquí anteriormente.

45 [0051] En otro método preferido, la EMT o una enfermedad o condición asociada a la EMT se diagnostica cuando la comparación conduce al hallazgo de una reducción del nivel de expresión de la molécula de ARNmi o un equivalente de la misma donde una fuente de dicha molécula de ARNmi o equivalente de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1 o una fuente de la misma y una reducción del nivel de expresión de al menos uno de otro ARNmi seleccionado de:

50 (a) al menos uno de ARNmi-124-1, ARNmi-206, ARNmi-181a-1, ARNmi-141, ARNmi-200a, ARNmi-200b, ARNmi-200c, ARNmi-429 and ARNmi-205 y/o un equivalente o un mimético o un isomiR o una fuente del mismo.

55 [0052] En otro aspecto de la invención, se proporciona un método para la identificación de una sustancia o una molécula capaz de prevenir, tratar, revertir, curar y/o retrasar un cáncer asociado a la EMT induciendo un proceso de MET en un sujeto, el método incluye las etapas de:

(a) proporcionar una población de célula de prueba capaz de expresar una molécula de ARNmi o un equivalente o un isomiR de la misma, donde una fuente de dicha molécula de ARNmi o equivalente o isomiR de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1 o fuente de la misma, preferiblemente la población de prueba comprende células de vejiga o próstata, y/o la población de célula de prueba comprende células cancerosas y/o la población de célula de prueba comprende células mamíferas, y/o la población de célula de prueba comprende células humanas;

(b) poner en contacto la población de células de prueba con la sustancia;

(c) determinar el nivel de expresión de dicha molécula de ARNmi o equivalente o isomiR de la misma o fuente de la misma o la actividad o nivel estable de dicha molécula de ARNmi o equivalente o isomiR de la misma o fuente de la misma en la población de células de prueba en contacto con la sustancia;

(d) comparar la expresión, actividad o nivel estable determinado en (c) con la expresión, actividad o nivel estable de dicha molécula de ARNmi o equivalente o isomiR de la misma o fuente de la misma en una población de células de prueba que no está en contacto con la sustancia; y,

(e) identificar una sustancia que produce una diferencia en el nivel de expresión, actividad o nivel estable de dicha molécula de ARNmi o equivalente o isomiR de la misma o fuente de la misma, entre la población de células de prueba que está en contacto con la sustancia y la población de células de prueba que no está en contacto con la sustancia.

[0053] Preferiblemente, en el paso a), una célula de prueba comprende un constructo de ácidos nucleicos que comprende una fuente o un precursor de una molécula de ARNmi o un equivalente o un isomiR de la misma donde una fuente o un precursor de dicha molécula de ARNmi o equivalente o isomiR de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1. Las moléculas de ARNmi más preferidas o unos equivalentes o isomiRs o fuentes de las mismas se han definido todos aquí anteriormente. Más preferiblemente, una célula de prueba comprende un constructo de luciferasa de luciérnaga impulsado por un promotor de CDH1. Preferiblemente, en un método los niveles de expresión, una actividad o niveles de estado estable superiores a los de una secuencia de nucleótidos o mayores que una molécula de ARNmi, equivalente o isomiR o fuente del mismo se comparan. Preferiblemente, en un método, una población de células de prueba comprende células mamíferas, más preferiblemente células humanas. Aún más preferiblemente, una población de células de prueba comprende células de vejiga o próstata. Una población de células de prueba preferida no expresa una molécula de ARNmi o un equivalente o un isomiR de la misma donde una fuente de dicha molécula de ARNmi o equivalente de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con la unidad representada por SEQ ID NO: 1 o fuente de la misma o tiene una expresión reducida en comparación con un homólogo epitelial normal que expresa CDH1. Más preferiblemente, una población de células de prueba comprende una célula mesenquimal con baja expresión de CDH1, pero es capaz de expresar CDH1. Alternativamente o además de las células mencionadas anteriormente, en un aspecto la invención también se refiere a una sustancia que se identifica en los métodos anteriormente mencionados. En un método preferido, los niveles de expresión, actividades o niveles estables de al menos otra molécula de ARNmi o isomiR o fuente de la misma se compara, preferiblemente donde la otra molécula de ARNmi o isomiR o fuente de la misma se selecciona de:

(a) al menos uno de ARNmi-124-1, ARNmi-206, ARNmi-181a-1, ARNmi-141, ARNmi-200a, ARNmi-200b, ARNmi-200c, ARNmi-429 y ARNmi-205 y/o un isomiR o una fuente del mismo.

Definiciones generales y tecnologías generales a las que se hace referencia aquí

[0054] Las moléculas de microARN ("ARNmi") son generalmente de 21 a 22 nucleótidos de longitud, aunque longitudes de 17 y hasta 25 nucleótidos se han descrito. Cualquier longitud de 17, 18, 19, 20, 21, 22, 23, 24, 25 está por lo tanto abarcada en la presente invención. Los ARNmi se procesan cada uno a partir de una molécula de ARN precursora más larga ("ARNmi precursor"). Los ARNmi precursores se transcriben a partir de genes no codificantes de proteína. Un precursor puede tener una longitud de al menos 50, 70, 75, 80, 85, 100, 150, 200 nucleótidos o más. Los ARNmi precursores tienen dos regiones de complementariedad que les habilita para formar una estructura de tipo tallo-bucle o plegada hacia atrás, que es escindida por enzimas llamadas Dicer y Drosha en animales. Dicer y Drosha son nucleasas de tipo ribonucleasas III. El ARNmi procesado es típicamente una porción del tallo. El ARNmi procesado (también denominado "ARNmi maduro") se vuelve parte de un complejo grande, conocido como el complejo de silenciamiento inducido por ARN (RISC), complejo para (infra)-regular un gen objetivo particular. Ejemplos de ARNmi animal incluyen aquellos que perfecta o imperfectamente hacen par de base con el ARNm objetivo, dando como resultado bien la degradación de ARNm o la inhibición de la traducción respectivamente (Olsen *et al.*, 1999; Seggerson *et al.*, 2002). Las moléculas ARNsi también son procesadas por Dicer, pero a partir de una molécula de ARN larga bicatenaria. Los ARNsi no se encuentran naturalmente en células animales, pero pueden funcionar en tales células en un complejo de silenciamiento inducido por ARN (RISC) para dirigir la escisión específica de secuencia de un ARNm objetivo (Denli *et al.*, 2003).

[0055] El estudio de moléculas de ARNmi endógenas se describe en la solicitud de patente de EE.UU. 2008/0171667. Un ARNmi está aparentemente activo en la célula cuando el ARN maduro monocatenario es

ligado por un complejo de proteína que regula la traducción de ARNm que hibridan al ARNm. Introducir moléculas ARN exógenas que afectan a células de la misma manera que los ARNm expresados endógenamente requiere que una molécula de ARN monocatenaria de la misma secuencia que el ARNm maduro endógeno sea absorbida por el complejo de proteína que facilita el control traduccional. Una variedad de diseños de molécula de ARN se han evaluado. Tres diseños generales que maximizan la absorción del ARNm monocatenario deseado por la ruta de ARNm han sido identificados. Una molécula de ARN con una secuencia de ARNm que tiene al menos uno de los tres diseños se puede denominar como un ARNm sintético. Las moléculas de ARNm de la invención pueden reemplazar o suplementar la actividad de silenciamiento génico de un ARNm endógeno. Un ejemplo de tales moléculas, características preferidas y modificaciones de tales moléculas y composiciones que comprenden tales moléculas se describe en la WO2009/091982.

[0056] Moléculas de ARNm de la invención o equivalentes o fuente de las mismas comprenden, en algunas formas de realización, dos moléculas de ARN donde un ARN es idéntico a un ARNm de origen natural maduro. La molécula de ARN que es idéntica a un ARNm maduro se denomina cadena activa. La segunda molécula de ARN, denominada cadena complementaria, es al menos complementaria parcialmente a la cadena activa. Las cadenas activas y complementarias se hibridizan para crear un ARN bicatenario, que es similar al precursor de ARNm de origen natural que es ligado por el complejo de proteína inmediatamente antes de la activación del ARNm en la célula. Maximizar la actividad de dicho ARNm requiere maximizar la absorción de la cadena activa y minimizar la absorción de la cadena complementaria por el complejo de proteína de ARNm que regula la expresión génica en el nivel de traducción. Los diseños moleculares que proporcionan actividad de ARNm óptima implican modificaciones de la cadena complementaria. Dos diseños incorporan modificaciones químicas de la cadena complementaria. La primera modificación implica la creación de un ARN complementario con un grupo diferente de un fosfato o hidroxilo en su 5' término. La presencia de la modificación 5' elimina aparentemente la absorción de la cadena complementaria y favorece posteriormente la absorción de la cadena activa por el complejo de proteína de ARNm. La modificación 5' puede ser cualquiera de una variedad de moléculas que incluyen NH₂, NHCOCH₃, biotina y otros. La segunda estrategia de modificación química que reduce significativamente la absorción de la cadena complementaria por la ruta de ARNm es la incorporación de nucleótidos con modificaciones de azúcar en los primeros 2-6 nucleótidos de la cadena complementaria. Debe observarse que las modificaciones de azúcar de acuerdo con la segunda estrategia de diseño se pueden acoplar con modificaciones en el terminal 5' de acuerdo con la primera estrategia de diseño para mejorar adicionalmente las actividades de ARNm. El tercer diseño de ARNm implica la incorporación de nucleótidos en el extremo 3' de la cadena complementaria que no son complementario de la cadena activa. Híbridos de los ARN activos resultantes y complementarios son muy estables en el extremo 3' de la cadena activa pero relativamente inestables en el extremo 5' de la cadena activa. Estudios con ARNs indican que la estabilidad de híbridos 5' es un indicador clave de absorción de ARN por el complejo de proteína que sostiene la interferencia de ARN, que está al menos relacionada con la ruta de ARNm en las células. Los inventores descubrieron que el uso juicioso de desapareamientos en la cadena de ARN complementaria mejora significativamente la actividad de dicho ARNm.

Bibliotecas de ARNm

[0057] En la divulgación, una aplicación clave para los ARNm según se identifica aquí es la evaluación o diagnóstico de la presencia de un individual o grupos de ARNm en una muestra. Poblaciones de células con cada uno de los ARNm diferentes se pueden después evaluar para identificar ARNm cuya presencia afecta a un fenotipo celular (es decir, EMT). El número de diferentes ARNm en las bibliotecas es variable. Se contempla que puede haber, al menos, o como mucho 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más, o cualquier gama derivable en esta, moléculas específicas de ARNm diferentes en la biblioteca. En formas de realización específicas, las bibliotecas tienen de 1 a 20 moléculas específicas de ARNm diferentes, o de 5 a 20 moléculas específicas de ARNm diferentes. Las moléculas específicas de ARNm "diferentes" se refieren a ácidos nucleicos que específicamente codifican ARNm con secuencias diferentes.

[0058] Los ARNm se contemplan por estar hechos principalmente de ARN, aunque en algunas formas de realización, pueden ser ARN, análogos de nucleótido, tales como ácidos nucleicos bloqueados (LNA) o ácidos nucleicos desbloqueados (UNA), ADN, o cualquier combinación de ADN, ARN, análogos de nucleótidos, y PNAS (ácidos nucleicos de péptido). Por consiguiente, se entiende que la biblioteca contiene uno o más ácidos nucleicos para estos ARNm diferentes. En formas de realización específicas, la biblioteca es específica para ARNm humanos, aunque bibliotecas para organismos múltiples se contemplan.

[0059] Una molécula de ARN de la invención tiene o comprende o consiste en una región de ARNm. En formas de realización específicas, una molécula de ARNm o equivalente de la misma tiene una secuencia que deriva de cualquiera de SEQ ID NO: 2-21 inclusive (tabla 3). Se contempla particularmente que las moléculas de ácido nucleico de la invención se pueden derivar de cualquiera de las secuencias de ARNm maduro en las SEQ ID NOs: 2-21.

[0060] Una molécula de ARNm o equivalente de la misma incluirá una secuencia que se extiende al menos de 1 a 5 nucleótidos de secuencia codificante arriba y/o abajo de la secuencia de ARNm predicha. En algunas formas

de realización, las moléculas tienen hasta 1, 2, 3, 4, 5, 6, 7 o más nucleótidos contiguos, o cualquier gama derivable de la misma, que flanquean la secuencia que codifica del ARNm procesado predominante en uno o ambos lados (extremo 5' y/o 3').

- 5 [0061] Bibliotecas de la divulgación pueden contener secuencias de ARNm de cualquier organismo que tiene ARNm, específicamente incluyendo pero no limitado a, mamíferos tales como seres humanos, primates no humanos, ratas y ratones. Específicamente se contemplan las bibliotecas que tienen, tienen al menos, o tienen como mucho 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más ARNm diferentes (esto es, moléculas específicas de ARNm que tienen secuencias diferentes derivadas de distintos genes de ARNm).
- 10 Específicamente se contemplan tales bibliotecas descritas en la frase precedente con respecto a cualquiera de las SEQ ID NO: 2-21 particularmente aquellas correspondientes a las secuencias de ARNm (secuencia madura).

Ácidos nucleicos

- 15 [0062] La presente invención se refiere a moléculas de ácidos nucleicos también llamadas fuentes o precursores de ARNm que pueden introducir ARNm en células cultivadas o en un sujeto. Los ácidos nucleicos se pueden producir en células o *in vitro* por enzimas purificadas aunque se producen preferentemente por síntesis química. Pueden ser crudos o purificados. El término "ARNm", a menos que se indique de otro modo, se refiere al ARNm procesado, después de que se haya escindido de su precursor. La tabla 2 indica qué SEQ ID NO corresponde a una secuencia precursora particular de un ARNm (SEQ ID NO: 22-35 y la tabla 3 qué SEQ ID NO corresponde a la secuencia madura de un ARNm (SEQ ID NO: 2-21). La tabla 4 identifica las secuencias de ADN clonadas en el vector lentiviral (SEQ ID NO: 36-47, que fueron usadas en la pantalla funcional como se describe en los ejemplos. La tabla 5 identifica la secuencia semilla preferida (como SEQ ID NO: 87-107) de cada uno de los ARNm maduros de la tabla 3. El nombre del ARNm se abrevia frecuentemente y se menciona sin el prefijo y se entenderá como tal, dependiendo del contexto. A menos que se indique lo contrario, los ARNm mencionados en la solicitud son secuencias humanas identificadas como mir-X o let-X, donde X es un número y/o letra.

- [0063] Se entiende que un ARNm se deriva de secuencias genómicas o un gen no codificante. En este aspecto, el término "gen" se usa por simplicidad para referirse a la secuencia genómica que codifica el ARNm precursor para un ARNm dado. Sin embargo, formas de realización de la invención pueden implicar secuencias genómicas de un ARNm que está implicado en su expresión, tal como un promotor u otras secuencias reguladoras.
- 30

- [0064] El término "recombinante" se puede utilizar y generalmente se refiere a una molécula que ha sido manipulada *in vitro* o que es el producto replicado o expresado de tal molécula.
- 35

- [0065] El término "ácido nucleico" es bien conocido en la técnica. Un "ácido nucleico" como se utiliza en este caso generalmente se refiere a una molécula (una o más cadenas) de ADN, ARN o un derivado o análogo de la misma, que comprende una nucleobase. Una nucleobase incluye, por ejemplo, una purina de origen natural o base de pirimidina encontrada en el ADN (por ejemplo, una adenina "A", a guanina "G", a timina "T" o una citosina "C") o ARN (por ejemplo, un A, un G, un uracilo "U" o un C). El término "ácido nucleico" abarca los términos "oligonucleótido" y "polinucleótido", cada uno como un subgénero del término "ácido nucleico".
- 40

- [0066] El término "ARNm" generalmente se refiere a una molécula monocatenaria, pero en formas de realización específicas, moléculas implementadas en la invención también abarcarán una región o una cadena adicional que es parcialmente (entre 10 y 50% complementaria a través de la longitud de la cadena), sustancialmente (más de 50% pero menos del 100% complementaria a través de la longitud de la cadena) o complementaria completamente de otra región de la misma molécula monocatenaria o de otro ácido nucleico. Así, los ácidos nucleicos pueden abarcar una molécula que comprende una o más cadena(s) complementarias o autocomplementarias o "complemento(s)" de una secuencia particular que comprende una molécula. Por ejemplo, el ARNm precursor puede tener una región autocomplementaria, que es hasta 100% complementaria.
- 45
- 50

- [0067] Como se utiliza en este caso, "hibridación", "hibridiza" o "capaz de hibridar" se entiende que se refiere a la formación de un molécula de cadena doble o triple o una molécula con naturaleza de cadena parcial doble o triple usando técnicas conocidas por la persona experta tales como procedimientos transferencia de Southern. El término "recocimiento" como se utiliza en este caso es sinónimo de "hibridar". El término "hibridación", "hibridiza" o "capaz de hibridación" puede significar condiciones de hibridación "baja", "media" o "alta" tal y como se define por debajo.
- 55

- [0068] Las condiciones de astringencia baja a media a alta significa pre-hibridación e hibridación a 42°C en 5X SSPE, 0,3% SDS, 200pg/ml de ADN de esperma de salmón cortado y desnaturalizado, y bien 25% 35% o 50% de formamida para astringencias bajas a medias a altas respectivamente. Posteriormente, la reacción de hibridación es lavada tres veces durante 30 minutos cada una utilizando 2XSSC, 0,2% SDS y bien 55 °C, 65 °C, o 75 °C para astringencias de bajas a medias a altas.
- 60

- [0069] Ácidos nucleicos o derivados de los mismos de la invención comprenderán, en algunas formas de realización la secuencia ARNm de cualquier ARNm descrito en SEQ ID NOs: 2-21 o son descritos en SEQ ID
- 65

NO: 22-35 o en SEQ ID NO: 36-47. Se contempla que secuencias de ácidos nucleicos de la invención derivadas de SEQ ID NO: 2-21 pueden tener, tener al menos, o tener como mucho 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, nucleótidos contiguos de SEQ ID NOs: 2-21 (o cualquier gama derivable de ello). En otras formas de realización, los ácidos nucleicos son, son al menos, o son como mucho 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100% idénticos a la secuencia de ARNm de SEQ ID NOs: 2-21 o a la secuencia precursora de cualquiera de las SEQ ID NOs: 22-35 o cualquier combinación o gama derivable de ellas.

Nucleobases

[0070] Como se utiliza en este caso una "nucleobase" se refiere a una base heterocíclica, tal como por ejemplo una nucleobase de origen natural (es decir, una A, T, G, C o U) encontrada en al menos un ácido nucleico de origen natural (es decir, ADN y ARN), y derivado(s) y análogos de origen natural o no de tal nucleobase. Una nucleobase generalmente puede formar uno o más enlaces de hidrógeno ("recocer" o "hibridar") con al menos una nucleobase de origen natural de manera que pueda sustituirse por emparejamientos de nucleobases de origen natural (por ejemplo, la unión de hidrógeno entre A y T, G y C, y A y U).

[0071] Las nucleobases de "purina" y/o "pirimidina" abarcan nucleobases de purina y/o pirimidina de origen natural y también derivado(s) y análogo(s) de las mismas, incluyendo pero no limitado a, las que tienen una purina o pirimidina sustituida por una o más fracciones de un alquilo, caboxialquilo, amino, hidroxilo, halógeno (es decir, fluoro, cloro, bromo o yodo), tiol o alquiltiol. Las fracciones preferidas de alquilo (por ejemplo, alquilo, caboxialquilo, etc.) comprenden de aproximadamente 1, aproximadamente 2, aproximadamente 3, aproximadamente 4, aproximadamente 5, a aproximadamente 6 átomos de carbono. Otros ejemplos no limitativos de una purina o pirimidina incluyen una deazapurina, una 2,6-diaminopurina, un 5-fluorouracil, una xantina, una hipoxantina, una 8- bromoguanina, un 8-cloroguanina, una bromotimina, una 8-aminoguanina, una 8- hidroxiguanina, una 8-metilguanina, una 8-tioguanina, una azaguanina, una 2- aminopurina, una 5-etilcitosina, una 5-metilcitosina, un 5-bromouracil, un 5-etiluracil, un 5-iodouracil, un 5-clorouracil, un 5-propiluracil, un tiouracil, una 2-metiladenina, una metiltioadenina, una N,N-dimetiladenina, unas azaadeninas, una 8-bromoadenina, una 8-hidroxiadenina, una 6-hidroxiaminopurina, una 6-tiopurina, una 4-(6- aminohehil/citosina) y similares.

Otros ejemplos son conocidos por los de habilidad en la técnica.

[0072] Una nucleobase puede estar comprendida en un nucleósido o nucleótido, utilizando cualquier método de síntesis química o natural descrito aquí o conocido por uno de habilidad ordinaria en la técnica. Tal nucleobase se puede marcar o puede ser parte de una molécula que está marcada y contiene la nucleobase.

Nucleósidos

[0073] Como se utiliza en este caso, un "nucleósido" se refiere a una unidad química individual que comprende una nucleobase de manera covalente fijada a una fracción enlazadora de nucleobase. Un ejemplo no limitativo de una "fracción enlazadora de nucleobase" es un azúcar que comprende átomos de 5 carbonos (es decir, un "azúcar de 5 carbonos"), incluyendo pero no limitado a una deoxiribosa, una ribosa, una arabinosa, o un derivado o un análogo de un azúcar de 5 carbonos. Ejemplos no limitativos de un derivado o un análogo de un azúcar de 5 carbonos incluyen un 2'-fluoro-2'-deoxiribosa o un azúcar carbocíclico donde un carbono se sustituye para un átomo de oxígeno en el anillo de azúcar.

[0074] Diferentes tipos de fijación(es) covalente(s) de una nucleobase a una fracción enlazadora de nucleobase se conocen en la técnica. A modo de ejemplo no limitativo, un nucleósido que comprende una nucleobase de una purina (es decir, A o G) o una 7-deazapurina de manera covalente típicamente agrega la posición 9 de una purina o una 7-deazapurina a la posición 1' de un azúcar de 5 carbonos. En otro ejemplo no limitativo, un nucleósido que comprende una nucleobase de pirimidina (es decir, C, T o U) de manera covalente típicamente agrega una posición 1 de una pirimidina a una posición 1' de un azúcar de 5 carbonos (Kornberg and Baker, 1992).

Nucleótidos

[0075] Como se utiliza en este caso, un "nucleótido" se refiere a un nucleósido que comprende además un "fracción de estructura". Una fracción de estructura de manera covalente generalmente agrega un nucleótido a otra molécula que comprende un nucleótido, o a otro nucleótido para formar un ácido nucleico. La "fracción de estructura" en los nucleótidos de origen natural típicamente comprende una fracción de fósforo, que está de manera covalente unida a un azúcar de 5 carbonos. La unión de la fracción de estructura ocurre típicamente en la posición 3' o 5' del azúcar de 5 carbonos. Sin embargo, otros tipos de uniones se conocen en la técnica, particularmente cuando un nucleótido comprende derivados o análogos de un azúcar de 5 carbonos de origen natural o fracción de fósforo.

Análogos de ácido nucleico

[0076] Un ácido nucleico puede comprender, o estar compuesto totalmente por, un derivado o análogo de una nucleobase, una fracción enlazadora de nucleobase y/o fracción de estructura que puede estar presente en un ácido nucleico de origen natural. ARN con análogos de ácido nucleico también se puede marcar según métodos de la invención. Como se utiliza en este caso un "derivado" se refiere a una forma modificada químicamente o alterada de una molécula de origen natural, mientras que los términos "mimético" o "análogo" se refieren a una molécula que puede parecerse o no estructuralmente a una molécula de origen natural o fracción, pero posee funciones similares. Como se utiliza en este caso, una "fracción" generalmente se refiere a un componente químico o molecular menor de una estructura química o molecular mayor. Los análogos o derivados de nucleobases, nucleósidos y nucleótidos se conocen bien en la técnica, y se han descrito (véase, por ejemplo, Scheit, 1980).

[0077] Ejemplos no limitativos adicionales de nucleósidos, nucleótidos o ácidos nucleicos que comprenden azúcar de 5 carbonos y/o derivados o análogos de fracción de estructura, incluyen los de: la patente de EE.UU. nº 5.681.947, que describe oligonucleótidos que comprenden derivados de purina con los que forman hélices triples y/o previenen la expresión de dsADN; las patentes estadounidenses 5.652.099 y 5.763.167, que describen ácidos nucleicos que incorporan análogos fluorescentes de nucleósidos que se encuentran en el ADN o el ARN, particularmente para su uso como sondas de ácidos nucleicos fluorescentes; la patente estadounidense 5.614.617, que describe análogos de oligonucleótidos con sustituciones en los anillos de pirimidina que poseen estabilidad de nucleasa mejorada; las patentes estadounidenses 5.670.663, 5.872.232 y 5.859.221, que describen análogos de oligonucleótidos con azúcares de 5 carbonos modificados (es decir, fracciones de T-deoxifuranosilo modificadas) usadas en detección de ácidos nucleicos; la patente estadounidense 5.446.137, que describe oligonucleótidos que comprenden al menos una fracción de azúcar de 5 carbonos sustituida en la posición 4' con un sustituyente diferente del hidrógeno que se puede usar en ensayos de hibridación; la patente estadounidense 5.886.165, que describe oligonucleótidos con ambos deoxiribonucleótidos con enlaces de internucleótidos 3'-5' y ribonucleótidos con enlaces de internucleótidos 2'-5'; la patente estadounidense 5.714.606, que describe una unión de internucleótido modificado donde un oxígeno en la posición 3' de la unión de internucleótido se sustituye por un carbono para mejorar la resistencia a la nucleasa de los ácidos nucleicos; la patente estadounidense 5.672.697, que describe oligonucleótidos que contienen uno o más enlaces de internucleótidos de fosfonato de metileno que mejoran la resistencia a la nucleasa; las patentes estadounidenses 5.466.786 y 5.792.847, que describen la unión de una fracción sustituyente que puede comprender un medicamento o marcador para el carbono 2' de un oligonucleótido para proporcionar estabilidad de nucleasa mejorada y capacidad para entregar fármacos o fracciones de detección; la patente estadounidense 5.223.618, que describe análogos de oligonucleótidos con una unión de estructura de carbono 2' o 3' que une la posición 4' y la posición 3' de la fracción de azúcar de 5 carbonos adyacente para captación celular mejorada, resistencia a nucleasas e hibridación a ARN diana; la patente estadounidense 5.470.967, que describe oligonucleótidos que comprenden al menos una unión de internucleótido de sulfamato o sulfamida que son útiles como sonda de hibridación de ácido nucleico; las patentes estadounidenses 5.378.825, 5.777.092, 5.623.070, 5.610.289 y 5.602.240, que describen oligonucleótidos con fracción enlazadora de tres o cuatro átomos que reemplaza la fracción de estructura de fosfodiéster usada para resistencia a la nucleasa mejorada, absorción celular y regular la expresión de ARN; la patente estadounidense 5.858.988, que describe el agente portador hidrofóbico fijado a la posición 2'-O de los oligonucleótidos para mejorar su permeabilidad de membrana y estabilidad; la patente estadounidense 5.214.136, que describe oligonucleótidos conjugados a antraquinona en el término 5' que poseen hibridación mejorada a ADN o ARN; estabilidad mejorada a nucleasas; la patente estadounidense 5.700.922, que describe quimeras PNA-DNA-PNA donde el ADN comprende nucleótidos de 2'-deoxi-eritropentofuranosil para resistencia a la nucleasa mejorada, afinidad de enlace, y capacidad para activar ribonucleasa H; y WO98WO98/39352, WO99/14226, WO2003/95467 y WO2007/085485, que describen nucleótidos de ARN modificados de los cuales la fracción de ribosa se modifica con un puente extra que conecta el oxígeno 2' y el carbono 4'. La ribosa bloqueada aumenta significativamente la afinidad de enlace y la especificidad; y la WO2008/147824, que describe nucleótidos de ARN modificado denominados UNA (ácido nucleico desbloqueado). Los UNA son análogos acíclicos de ARN en los que el enlace entre los átomos C2' y C3' han sido escindidos, mermando la afinidad de enlace hacia una cadena complementaria. Los UNA son compatibles con el reconocimiento de ribonucleasa H y la escisión de ARN y mejoran el silenciamiento génico mediado por ARNs; la WO2008/036127 que describe análogos de ácido nucleico de morfolino, que contienen tanto enlaces de intersubunidad sin carga y catiónicos; la WO/2007/069092 y la EP2075342 que describen ácidos nucleicos de cremallera (ZNA), que contienen derivados de espermina conjugantes como fracciones catiónicas (unidades Z) para un oligonucleótido; la patente estadounidense 5.708.154, que describe ARN enlazado a un ADN para formar un híbrido de ADN-ARN; la patente estadounidense 5.728.525, que describe el marcaje de análogos de nucleósidos con un marcador fluorescente universal.

[0078] Instrucciones adicionales para análogos de nucleósidos y análogos de ácido nucleico son la patente estadounidense 5.728.525, que describe análogos de nucleósidos que están marcado en el extremo; las patentes estadounidenses 5.637.683, 6.251.666 (sustituciones de L-nucleótidos), y 5.480.980 (nucleótidos de 7-deaza-2'-deoxiguanosina y análogos de ácido nucleico de los mismos).

[0079] El uso de otros análogos se contempla específicamente para su uso en el contexto de la presente invención. Tales análogos se pueden utilizar en las moléculas de ácido nucleico sintéticas de la invención, tanto en toda la molécula como en nucleótidos seleccionados. Incluyen, pero de forma no limitativa,

- 5 1) modificaciones de ribosa (tales como 2'F, 2' NH₂, 2'N₃, 4'thio o 2' O-CH₃) y
- 2) modificaciones de fosfato (tales como las que se encuentran en fosforotioatos, metil fosfonatos y fosforoboratos).

10 Tales análogos han sido creados para conferir estabilidad en los ARN reduciendo o eliminando su capacidad para ser escindidos por las ribonucleasas. Cuando estos análogos de nucleótido están presentes en los ARN, pueden tener efectos profundamente positivos en la estabilidad de los ARN en animales. Se contempla que el uso de análogos de nucleótidos se puede usar solo o conjuntamente con cualquiera de las modificaciones de diseño de un ARNm sintético para cualquier ácido nucleico de la invención.

15 Nucleótidos modificados

[0080] Los ARNm de la invención específicamente contemplan el uso de nucleótidos que se modifican por mejorar sus actividades. Tales nucleótidos incluyen aquellos que son en el término 5' o 3' del ARN al igual que aquellos que están dentro de la molécula. Los nucleótidos modificados usados en las cadenas complementarias de dichos ARNm bien bloquean el OH 5' o fosfato del ARN o introducen modificaciones de azúcar internas que mejoran la absorción de la cadena activa del ARNm. Modificaciones para los ARNm incluyen modificaciones de azúcar internas que mejoran la hibridación al igual que estabilizan las moléculas en las células y modificaciones terminales que además estabilizan los ácidos nucleicos en las células. Además se contemplan modificaciones que se pueden detectar por microscopía u otros métodos para identificar las células que contienen los ARNm sintéticos.

Preparación de ácidos nucleicos

30 [0081] Un ácido nucleico se puede hacer por cualquier técnica conocida para uno de habilidad en la materia, tal como por ejemplo, síntesis química, producción enzimática o producción biológica. Aunque los ARNm según la invención se podrían producir utilizando métodos recombinantes, se prefiere producir ARNm por síntesis química o producción enzimática. Los ARNm se pueden producir por un número de métodos, incluyendo métodos que implican tecnología de ADN recombinante.

35 [0082] La síntesis de ácidos nucleicos se realiza según métodos estándar. Véase, por ejemplo, Itakura and Riggs (1980). Adicionalmente, la patente estadounidense 4.704.362, la patente estadounidense 5.221.619, y la patente estadounidense 5.583.013 describen cada una varios métodos para preparar ácidos nucleicos. Ejemplos no limitativos de un ácido nucleico (por ejemplo, un oligonucleótido), incluyen un ácido nucleico hecho por síntesis *in vitro* químicamente usando química de fosfotriéster, fosfito o fosforamidita y técnicas de fase sólida tales como las descritas en la EP 266.032, o vía productos intermedios de H-fosfonato de desoxinucleósido como se describe por Froehler *et al.*, 1986 y la patente estadounidense número de serie: 5.705.629. En los métodos de la presente invención, uno o más oligonucleótidos pueden ser utilizados. Varios mecanismos diferentes de síntesis de oligonucleótidos se han descrito en, por ejemplo, las patentes estadounidenses: 4.659.774, 4.816.571, 5.141.813, 5.264.566, 4.959.463, 5.428.148, 5.554.744, 5.574.146, 5.602.244.

45 [0083] Un ejemplo no limitativo de un ácido nucleico producido enzimáticamente incluye uno producido por enzimas en reacciones de amplificación tales como PCR(TM) (véase, por ejemplo, la patente estadounidense 4.683.202 y la patente estadounidense 4.682.195), o la síntesis de un oligonucleótido descrita en la patente estadounidense n° 5.645.897.

50 [0084] La síntesis de oligonucleótidos es bien conocida por los de habilidad en la técnica. Se han descrito varios mecanismos diferentes de síntesis de oligonucleótidos en, por ejemplo, las patentes estadounidenses 4.659.774, 4.816.571, 5.141.813, 5.264.566, 4.959.463, 5.428.148, 5.554.744, 5.574.146, 5.602.244. Básicamente, la síntesis química se puede conseguir por el método de diéster, el método triéster, el método de polinucleótidos fosforilasa y por química de fase sólida. Estos métodos se discuten con más detalle más adelante.

Método de diéster

60 [0085] El método de diéster fue el primero en ser desarrollado hasta un estado utilizable, principalmente por Khorana y sus colaboradores (Khorana, 1979). La etapa básica es la unión de dos deoxinucleótidos protegidos adecuadamente para formar un dideoxinucleótido que contiene un enlace de fosfodiéster. El método de diéster está bien establecido y se ha usado para sintetizar moléculas de ADN (Khorana, 1979).

Método de triéster

65

[0086] La diferencia principal entre los métodos de diéster y triéster es la presencia en el último de un grupo protector extra en los átomos de fosfato de los reactivos y productos (Itakura *et al.*, 1975). El grupo protector de fosfato es normalmente un grupo de clorofenilo, que hace que los productos intermedios de los nucleótidos y los polinucleótidos sean solubles en solventes orgánicos. Por lo tanto, las purificaciones se hacen en soluciones de cloroformo. Otras mejoras en el método incluyen (i) el acoplamiento en bloque de trímeros y oligómeros mayores, (ii) el uso extensivo de cromatografía en fase líquida de alta eficacia para la purificación tanto de productos finales como de intermedios, y (iii) la síntesis en fase sólida.

Método de polinucleótido fosforilasa

[0087] Este es un método enzimático de síntesis de ADN que puede utilizarse para sintetizar muchos oligonucleótidos útiles (Gillam *et al.*, 1978; Gillam *et al.*, 1979). Bajo condiciones controladas, la polinucleótido fosforilasa añade predominantemente un nucleótido único a un oligonucleótido corto.

[0088] La purificación cromatográfica permite obtener el aducto único deseado. Al menos un trímero se requiere para iniciar el procedimiento, y este cebador debe ser obtenido por algún otro método. El método de polinucleótido fosforilasa funciona y tiene la ventaja de que los procedimientos implicados son conocidos por la mayoría de los bioquímicos.

Métodos de fase sólida

[0089] A partir la tecnología desarrollada para la síntesis en fase sólida de polipéptidos, ha sido posible unir el nucleótido inicial a un material de soporte sólido y proceder con la adición gradual de nucleótidos. Cualquier etapa de mezcla y de lavado se simplifica, y el procedimiento es susceptible de automatización. Estas síntesis se realizan hoy en día de forma rutinaria utilizando sintetizadores de ácido nucleico automáticos.

[0090] La química de fosforamidita (Beaucage and Lyer, 1992) se ha convertido con mucho en la química de acoplamiento más ampliamente usada para la síntesis de oligonucleótidos. Como ya saben los expertos en la técnica, la síntesis de fosforamidita de oligonucleótidos implica la activación de precursores de monómeros de nucleósido fosforamidita por reacción con un agente de activación para formar productos intermedios activados, seguido de adición secuencial de los productos intermedios activados para la cadena de oligonucleótidos creciente (anclada generalmente en un extremo para un soporte sólido adecuado) para formar el producto de oligonucleótidos.

Métodos recombinantes.

[0091] Métodos recombinantes para producir ácidos nucleicos en una célula son bien conocidos por los expertos en la técnica. Incluyen el uso de vectores, plásmidos, cósmidos y otros vehículos para la entrega de un ácido nucleico a una célula, que puede ser la célula objetivo o sencillamente una célula huésped (para producir grandes cantidades de la molécula de ARN deseada). Alternativamente, tales vehículos se pueden usar en el contexto de un sistema libre de células mientras los reactivos para generar la molécula de ARN están presentes. Tales métodos incluyen los descritos en Sambrook, 2003, Sambrook, 2001 y Sambrook, 1989. En formas de realización determinadas, la presente invención se refiere a moléculas de ácido nucleico que no son sintéticas. En algunas formas de realización, la molécula de ácido nucleico tiene una estructura química de un ácido nucleico de origen natural y una secuencia de un ácido nucleico de origen natural, tal como la secuencia exacta y entera de un ARNm primario monocatenario (véase, Lee 2002), un ARNm precursor monocatenario o un ARNm maduro monocatenario. Además del uso de tecnología recombinante, tales ácidos nucleicos no sintéticos se pueden generar químicamente, tal como utilizando tecnología usada para la creación de oligonucleótidos.

Diseño de ARNm

[0092] Los ARNm típicamente comprenden dos cadenas, una cadena activa que es idéntica en la secuencia al ARNm maduro que se está estudiando y una cadena complementaria que es al menos complementaria parcialmente de la cadena activa. La cadena activa es la molécula pertinente biológicamente y debería ser preferentemente absorbida por el complejo en las células que modulan la traducción bien a través de la degradación de ARNm o el control traduccional. La absorción preferencial de la cadena activa tiene dos resultados profundos: (1) la actividad observada de dicho ARNm aumenta espectacularmente y (2) los efectos no previstos inducidos por la absorción y la activación de la cadena complementaria son esencialmente eliminados. Según la invención, varios diseños de ARNm se pueden utilizar para asegurar la absorción preferencial de la cadena activa.

Agente de bloqueo 5'

[0093] La introducción de una fracción estable diferente del fosfato o hidroxilo en el extremo 5' de la cadena complementaria afecta a su actividad en la ruta de ARNm. Esto asegura que solo la cadena activa del ARNm se usará para regular la traducción en la célula. Las modificaciones 5' incluyen, pero de forma no limitativa, NH₂,

biotina, un grupo de amina, un grupo de alquilamina inferior, un grupo de acetil, 2' O-Me, DMTO, fluoresceína, un tiol, o acridina o cualquier otro grupo con este tipo de funcionalidad.

[0094] Otras modificaciones de cadena sentido. La introducción de modificaciones de nucleótidos tipo 2'-O Me, 2'-deoxi, T-deoxi-2'-fluoro, 2'-O-metil, 2'-O-metoxietil (2'-O-MOE), 2'-O-aminopropil (2'-O-AP), 2'-O-dimetilaminoetil (2'-O-DMAOE), 2'-O-dimetilaminopropil (2'-O-DMAP), 2'-O-dimetilaminoetiloxietil (2'-O-DMAEOE), o 2'-O-N-metilacetamido (2'-O-NMA), NH₂, biotina, un grupo de amina, un grupo de alquilamina inferior, un grupo de acetil, DMTO, fluoresceína, un tiol, o acridina o cualquier otro grupo con este tipo de funcionalidad en la cadena complementaria del ARNm puede eliminar la actividad de la cadena complementaria y mejorar la absorción de la cadena activa del ARNm.

[0095] Desapareamientos de bases en la cadena sentido. Como con los ARNs (Schwarz 2003), la estabilidad relativa de los extremos 5' y 3' de la cadena activa del ARNm determina aparentemente la absorción y la activación del activo por la ruta de ARNm. La desestabilización del extremo 5' de la cadena activa del ARNm por la colocación estratégica de desapareamientos de bases en el extremo 3' de la cadena complementaria del ARNm sintético mejora la actividad de la cadena activa y esencialmente elimina el actividad de la cadena complementaria.

Células huésped y células diana

[0096] Las células en las que un ARNm o fuente del mismo se introduce o donde la presencia de un ARNm se evalúa se pueden derivar de o estar contenidas en cualquier organismo. Preferiblemente, la célula es una célula de vertebrado. Más preferiblemente, la célula es una célula de mamífero. Aún más preferiblemente, la célula es una célula de humano. Una célula de mamífero puede ser de la línea germinal o somática, totipotente o pluripotente, divisora o no divisora, de epitelio, inmortalizada o transformada, o similar. La célula puede ser una célula no diferenciada, tal como una célula madre, o una célula diferenciada, tal como de una célula de un órgano o tejido. Alternativamente, las células se pueden clasificar como células epiteliales, de cerebro, pecho, cervix, colon, tracto gastrointestinal, corazón, riñón, intestino grueso, hígado, pulmón, ovario, páncreas, corazón, próstata, vejiga, intestino delgado, estómago, testículos o útero.

[0097] Como se utiliza en este caso, los términos "célula", "línea celular" y "cultivo celular" se pueden utilizar de forma intercambiable. Todos estos términos también incluyen su progenie, que es cualquiera y todas las generaciones posteriores formadas por división celular. Se entiende que cualquier descendiente puede no ser idéntico debido a mutaciones deliberadas o involuntarias. Una célula huésped puede ser "transfectada" o "transformada", lo que se refiere a un proceso por el que el ácido nucleico exógeno es transferido o introducido en la célula huésped. Una célula transformada incluye la célula sujeto primaria y su progenie. Como se utiliza en este caso, los términos células "diseñadas" y "recombinantes" o células huésped se refieren a una célula en la que una secuencia de ácidos nucleicos exógena, tal y como, por ejemplo, un ARN pequeño de interferencia o un constructo modelo que codifica un gen reportero ha sido introducido. Por lo tanto, las células recombinantes se pueden distinguir de las células de origen natural que no contienen un ácido nucleico recombinantemente introducido.

[0098] Un tejido puede comprender una célula huésped o células para transformar o poner en contacto con una composición de entrega de ácido nucleico y/o un agente adicional. El tejido puede ser parte o estar separado de un organismo. En formas de realización determinadas, un tejido y sus células constituyentes pueden comprender, pero no están limitadas a cerebro, células madre, hígado, pulmón, hueso, pecho, cervix, colon, endometrio, epitelio, esófago, células calciformes, riñón, ovarios, páncreas, próstata, vejiga, piel, intestino delgado, estómago, testículos, corazón, vasos sanguíneos.

[0099] En formas de realización determinadas, la célula huésped o tejido puede estar comprendido en al menos un organismo. En formas de realización determinadas, el organismo puede ser un mamífero, un humano, un primate o murino. Uno de habilidad en la técnica además entendería las condiciones bajo las que hay que incubar todas las células huésped descritas anteriormente para mantenerlas y para permitir su división para formar progenie.

Métodos de entrega

[0100] La presente invención implica en algunas formas de realización la entrega de un ácido nucleico en una célula. Esto puede realizarse como parte de un método de selección, o puede estar relacionado con una aplicación terapéutica o de diagnóstico. Las moléculas de ARN pueden ser codificadas por una molécula de ácido nucleico comprendida en un vector. El término "vector" se utiliza para referirse a una molécula de ácido nucleico portador en la que una secuencia de ácidos nucleicos se puede insertar para su introducción en una célula en la que éste puede ser replicado. Una secuencia de ácidos nucleicos puede ser "exógena", que significa que es foránea de la célula en la que el vector se ha introducido o que la secuencia es homóloga de una secuencia en la célula pero en una posición en el ácido nucleico de la célula huésped donde la secuencia no se encuentra habitualmente. Los vectores incluyen plásmidos, cósmidos, virus (bacteriófagos, virus de animales,

lentivirus y virus de plantas) y cromosomas artificiales (por ejemplo, YACs). Uno de habilidad en la técnica estaría bien equipado para construir un vector a través de técnicas estándar recombinantes, que se describen en Sambrook *et al.*, 1989 y Ausubel *et al.*, 1996. Además de codificar un polipéptido modificado tal como gelonina modificada, un vector puede codificar secuencias de polipéptidos no modificadas tales como una etiqueta o molécula de direccionamiento. Una molécula de direccionamiento es una que dirige el ácido nucleico deseado a un órgano particular, tejido, célula, u otra ubicación en el cuerpo de un sujeto.

[0101] El término "vector de expresión" se refiere a un vector que contiene una secuencia de ácidos nucleicos que codifica para al menos parte de un producto génico capaz de ser transcrito. Los vectores de expresión pueden contener una variedad de "secuencias de control", que se refieren a secuencias de ácidos nucleicos necesarias para la transcripción y posiblemente la traducción de una secuencia codificante operativamente enlazada en un organismo huésped particular. Además de las secuencias de control que rigen la transcripción y la traducción, los vectores y vectores de expresión puede contener secuencias de ácidos nucleicos que sirven para otras funciones también y son descritos

[0102] Hay muchas formas en las que los vectores de expresión se pueden introducir en las células. En ciertas formas de realización de la invención, el vector de expresión comprende un virus o vector diseñado derivado de un genoma viral. La capacidad de ciertos virus para introducir células vía endocitosis mediada por receptor, para integrar en el genoma de la célula huésped y expresar genes de forma estable y eficaz ha hecho que estos candidatos sean atractivos para la transferencia de genes foráneos en células mamíferas (Ridgeway, 1988; Nicolas and Rubenstein, 1988; Baichwal and Sugden, 1986; Temin, 1986). Los primeros virus usados como vectores génicos fueron virus de ADN incluyendo los papovaviruses (virus símico 40, papiloma virus bovino y polioma) (Ridgeway, 1988; Baichwal and Sugden, 1986) y los adenovirus (Ridgeway, 1988; Baichwal and Sugden, 1986). Tienen una capacidad relativamente baja para las secuencias de ADN foráneas y tienen un espectro de huésped restringido. Además, su potencial oncogénico y efectos citopáticos en células permisivas generan problemas de seguridad. Pueden alojar solo hasta 8 kb de material genético foráneo pero se pueden introducir fácilmente en una variedad de líneas celulares y animales de laboratorio (Nicolas and Rubenstein, 1988; Temin, 1986). Los vectores de expresión pueden contener un casete de expresión de ARNi que comprende un promotor y una o más estructuras de tallo-bucle separadas por una o más regiones separadoras (WO2006/084209). Otra forma de introducir los vectores de expresión en células, usando proteínas de fusión de avidina se describe en la US6.287.792.

[0103] Los retrovirus son un grupo de virus de ARN monocatenario caracterizados por una capacidad para convertir su ARN en ADN bicatenario en células infectadas; también pueden usarse como vectores. Otros vectores virales se pueden emplear como constructos de expresión en la presente invención. Vectores derivados de virus tales como virus vaccinia (Ridgeway, 1988; Baichwal and Sugden, 1986; Coupar *et al.*, 1988) virus adeno-asociado (AAV) (Ridgeway, 1988; Baichwal and Sugden, 1986; Hermonat and Muzycska, 1984, lentivirus (WO2008/071959, WO2004/054512), virus hemaglutinante de Japón (WO2004/035779), baculovirus (WO2006/048662) y virus de herpes se pueden emplear. Ofrecen diferentes características atractivas para varias células mamíferas (Friedmann, 1989; Ridgeway, 1988; Baichwal and Sugden, 1986; Coupar *et al.*, 1988; Horwich *et al.*, 1990).

[0104] Se cree que otros métodos adecuados de entrega de ácido nucleico para afectar la expresión de las composiciones de la presente invención incluyen prácticamente cualquier método por el que un ácido nucleico (por ejemplo, ADN, incluyendo vectores virales y no virales) se puedan introducir en un orgánulo, una célula, un tejido o un organismo, como se describe en este caso o como conocería uno de habilidad ordinaria en la técnica. Tales métodos incluyen, pero de forma no limitativa, entrega directa de ADN tal como por inyección (las patentes de EEUU números: 5.994.624, 5.981.274, 5.945.100, 5.780.448, 5.736.524, 5.702.932, 5.656.610, 5.589.466 y 5.580.859), incluyendo microinyección (Harlan and Weintraub, 1985; patente de EEUU n° 5.789.215); por electroporación (patente de EEUU n° 5.384.253); por precipitación de fosfato cálcico (Graham and Van Der Eb, 1973; Chen and Okayama, 1987; Rippe *et al.*, 1990); usando DEAE-dextrano seguido de polietilenglicol (Gopal, 1985); por carga sónica directa (Fechheimer *et al.*, 1987); por transfección mediada por liposoma (Nicolau and Sene, 1982; Fraley *et al.*, 1979; Nicolau *et al.*, 1987; Wong *et al.*, 1980; Kaneda *et al.*, 1989; Kato *et al.*, 1991); por internalización fotoquímica (WO2008/007073); por bombardeo con microproyectiles (solicitud de PCT números: WO 94/09699 y 95/06128; patentes de EEUU números: 5.610.042, 5.322.783, 5.563.055, 5.550.318, 5.538.877 y 5.538.880); por agitación con fibras de carburo de silicona (Kaepler *et al.*, 1990; patentes de EEUU números: 5.302.523 y 5.464.765); por transformación mediada por Agrobacterium (patentes de EEUU números: 5.591.616 y 5.563.055); o por transformación mediada por PEG de protoplastos (Omirulleh *et al.*, 1993; patentes de EEUU números: 4.684.611 y 4.952.500 por absorción mediada por disecación/inhibición de ADN (Potrykus *et al.*, 1985). Mediante la aplicación de técnicas tales como estas, orgánulo(s), célula(s), tejido(s) u organismo(s) se pueden transformar de forma estable o transitoria.

[0105] Una revisión proporciona diferentes maneras de formular una molécula de ARN para optimizar su internalización en una célula (Kim SS., *et al.*, Trends Mol. Med., 2009, 15: 491-500). Las siguientes publicaciones divulgan otras formas alternativas de formular una molécula de ARN para mejorar su internalización en una célula: la WO 2007/095152, describe el uso de PTD-DRBD (dominios de transducción de péptido enlazados a

dominio de unión bicatenarios) para la entrega de oligonucleótidos, la WO 2009/086558, describe el uso de SNALP (partículas lipídicas de ácidos nucleicos estables) partículas, que comprenden una mezcla de lípidos catiónicos y fusogénicos que permiten la absorción celular y la liberación endosomal de la carga de ácidos nucleicos de la partícula, la WO 2009/149418, describe emulsiones fosfolípido-aceite-ARNi, la WO 2007/121947, describe el uso de un vehículo de entrega basado en lipoplex, la WO 2009/132131, describe el uso de lípidos nuevos y partículas de ácido nucleico-lípido que proporcionan encapsulación eficiente y entrega eficaz del ácido nucleico encapsulado a células, la WO2004/091578 y WO2004/064805 describe tecnología de coqueatos de capas de lípidos alternantes que giran en espiral alrededor de una molécula de ácido nucleico, la WO2003/047494 y WO2003/047493 describen micelas inversas que incorporan ácidos nucleicos para la entrega oral y mucosa, la WO 2008/156702, describe bacterias y partícula terapéutica bacteriana (BTP), incluyendo oligonucleótidos como vehículo de entrega a las células. Cada una de las formulaciones mencionadas o descritas en estas publicaciones están abarcadas por la presente invención.

[0106] Una variedad de compuestos se han adjuntado a los extremos de los oligonucleótidos para facilitar su transporte a través de las membranas celulares. Los péptidos señal cortos encontrados en HIV TAT, HSV VP22, *Drosophila antennapedia*, y otras proteínas se ha descubierto que permiten la transferencia rápida de biomoléculas a través de las membranas (revisado por Schwarze 2000). Estos péptidos señal, denominados dominios de transducción de proteína (PTDs), se han unido a los oligonucleótidos para facilitar su entrega en las células cultivadas (Eguchi A, Dowdy SF, *Trends Pharmacol Sci.*, 2009, 7:341-5). Los colesterolos han sido conjugados a oligonucleótidos para mejorar su absorción en células en animales (MacKellar 1992). Los grupos de colesterol terminales aparentemente interactúan con receptores o lípidos en las superficies de las células y facilitan la internalización de los oligonucleótidos modificados. Asimismo, la poli-L-lisina ha sido conjugada a oligonucleótidos para reducir la carga negativa neta y mejorar la absorción en células (Leonetti 1990).

[0107] Se han desarrollado una variedad de compuestos que forman complejos con ácidos nucleicos, los entregan a las superficies de las células, y facilitan su absorción en y liberación desde los endosomas. Entre estos están: (1) una variedad de lípidos tales como DOTAP (u otro lípido catiónico), DDAB, DHDEAB y DOPE y (2) polímeros no basados en lípidos como polietilenimina, poliamidoamínica y dendrímeros de estos y otros polímeros. En ciertas de estas formas de realización una combinación de lípidos se emplea tal como DOTAP y colesterol o un derivado de colesterol (patente de EEUU 6.770.291). Varios de estos reactivos se ha demostrado que facilitan la absorción de ácido nucleico en animales.

[0108] Los componentes celulares implicados en la ruta de ARNmi se están haciendo conocidos. Las proteínas que estabilizan y/o transportan ARNmi dentro de las células pueden mejorar la estabilidad y la actividad de los ARNmi porque deberían proteger y guiar los ARNmi ligados una vez que están en las células. Mezclas de proteínas transportadoras de ARNmi y ARNmi podrían mejorar la eficacia de tratamientos basados en ARNmi. Los ARN son moléculas hidrofílicas por su estructura de fosfato aniónico y de azúcar. Aunque las nucleobases son hidrofóbicas, la hidrofiliidad domina debido a la unión de hidrógeno extensiva resultante de los residuos de fosfato y de azúcar. El carácter hidrofílico y la estructura aniónica reducen la permeación celular. La conjugación de grupos lipofílicos como el colesterol (Manoharan, 2002) y derivados de ácido láurico y litocólico con funcionalidad C32 (Lorenz *et al.*, 2004), se ha demostrado que mejoran la absorción celular. Además, la unión de oligonucleótidos conjugados de esteroide a lipoproteínas diferentes en el flujo sanguíneo, tales como LDL, protege su integridad y gobierna su biodistribución (Rump *et al.*, 2000). El colesterol fijado a las moléculas antisentido (Bijsterbosch *et al.*, 2001) y los aptámeros (Rusconi *et al.*, 2004) también se ha demostrado que estabilizan los oligonucleótidos permitiendo la unión a lipoproteínas. El colesterol se ha demostrado que mejora la absorción y la estabilidad del suero de los ARNsi *in vitro* (Lorenz *et al.*, 2004) e *in vivo* (Soutschek *et al.*, 2004). Adicionalmente, un número de moléculas pequeñas como SB-435495 (Blackie *et al.*, (2002), isradipina (Oravcova *et al.*, 1994), amlodipina (Oravcova *et al.*, 1994) y 2,2',4,4',5,5'-hexaclorobifenil (Borlakoglu *et al.*, 1990) podría mejorar la absorción celular, y mejorar la resistencia a la nucleasa promoviendo la asociación de lipoproteínas.

Selección con bibliotecas de ARNmi

[0109] Como se usa en la solicitud de patente, la selección es un proceso donde varios reactivos específicos de ARNmi se entregan separadamente en poblaciones de células individuales o animales. En uno o más momentos designados después de la entrega, las poblaciones celulares o animales se evalúan para uno o más fenotipos. Esas células o animales que tienen un fenotipo significativamente diferente de las células o animales en el grupo de control negativo se clasifican como positivos. El ARNmi que estaba siendo manipulado en la muestra se define como un hit. Los hits representan objetivos para investigación adicional y desarrollo terapéutico potencial.

En algunas formas de realización, hay un proceso multifase para la selección, en formas de realización determinadas, hay cuatro etapas generales:

(1) desarrollo de ensayo cuantitativo para controlar el proceso celular que se está estudiando.

[0110] Ensayos que miden la intensidad de un rango de fenotipo celular a partir de ensayos microscópicos que controlan el tamaño celular, el estado del ciclo celular, o el anticuerpo que tiñe para ensayos enzimáticos que

evalúan la producción de un sustrato específico en un lisado celular para dirigir las mediciones biomoléculas o moléculas pequeñas en lisados, en células, o en el medio.

[0111] Importante para el éxito de una selección es la creación de un ensayo que realmente mida el fenotipo celular y maximice la proporción de señal a ruido del ensayo. Maximizar la señal a ruido implica variables de prueba tales como el tiempo de ensayo, los componentes de ensayo, el tipo de célula y la cantidad de tiempo entre la transfección y el ensayo. Cuanto mayor es la diferencia en los resultados del ensayo entre un fenotipo positivo y un fenotipo de control negativo, mayor es la extensión en los resultados de la selección y mejor es la oportunidad de identificar genes interesantes.

(2) Optimizar las condiciones de transfección para las células deseadas.

[0112] La primera fase de este proceso es identificar un reactivo de transfección y las condiciones de cultivo que maximicen la absorción de ARNm sintéticos mientras que se mantiene una viabilidad celular alta. Encontramos útil evaluar 2-5 reactivos de transfección diferentes cuando se usan líneas celulares o 5-10 condiciones de electroporación cuando se usan células primarias o de suspensión. La transfección se puede optimizar para el reactivo o condición de electroporación que funcionó mejor entre las condiciones evaluadas. La selección de bibliotecas específicas de ARNm requiere condiciones para la transfección de alto rendimiento. En este tipo de selección, se usó la introducción lentiviral antes que la transfección. Esto puede requerir técnicas de optimización alternativas.

(3) Selección

[0113] Una vez que se han desarrollado el ensayo y el proceso de transfección, una biblioteca de ARNm sintéticos o ARNm expresados por virus se puede introducir secuencialmente en las células en una placa de 24 o 96 pocillos. Duplicar o triplicar las transfecciones para cada reactivo proporciona datos suficientes para el análisis estadístico razonable.

(4) Validación de hits

[0114] La validación de un hit implica mostrar que el fenotipo observado se debe al ARNm al que se dirige. Los hits se confirman típicamente por la entrega de una serie de diluciones del inhibidor de ARNm o ARNm sintético que se registró como hit en la célula que se evaluó originalmente. La confirmación es ligeramente diferente de la validación. La confirmación es una repetición del fenotipo inducido por el ARNm, mientras que la validación puede también incluir inversión del fenotipo antagonizando el fenotipo mediado por ARNm.

Marcado y técnicas de marcado

[0115] En algunas formas de realización, la presente invención se refiere a los ARNm que están marcados, tales como para ensayos de selección para evaluar la relevancia terapéutica o de diagnóstico de una especie de ARNm particular. Se contempla que el ARNm puede primero ser aislado (bien a partir de una célula en la que el ARNm es endógeno de la célula o a partir de una célula en la que el ARNm es exógeno de la célula) y/o purificado antes del marcado. Esto puede conseguir una reacción que marca de forma más eficaz el ARNm, a diferencia de otro ARN en una muestra en la que el ARNm no está aislado o purificado antes del marcado. En muchas formas de realización de la invención, el marcador es no radioactivo. Generalmente, los ácidos nucleicos se pueden marcar añadiendo nucleótidos marcados (proceso de una única fase) o añadiendo nucleótidos y marcando los nucleótidos añadidos (proceso de dos fases).

[0116] Además, los ARNm se pueden marcar como se describe en la patente de EEUU 7.880.010. Tales nucleótidos incluyen los que se pueden marcar con un tinte, incluyendo un tinte fluorescente, o con una molécula tal como biotina. Los nucleótidos marcados están fácilmente disponibles; se pueden adquirir comercialmente o se pueden sintetizar por reacciones conocidas por los expertos en la técnica.

Nucleótidos para marcado

[0117] Los nucleótidos para el marcado no son nucleótidos de origen natural, sino que se refieren a nucleótidos preparados que tienen una fracción reactiva en ellos. Funcionalidades reactivas específicas de interés incluyen: amino, sulfhidrilo, sulfoxilo, aminosulfhidrilo, azido, epóxido, isotiocianato, isocianato, anhídrido, monoclorotriazina, diclorotriazina, piridina mono o dihalógeno sustituida, diazina mono- o disustituida, maleimida, epóxido, aziridina, sulfonil haluro, haluro ácido, haluro de alquilo, haluro de arilo, alquilsulfonato, éster de N-hidroxisuccinimida, éster de imido, hidracina, azidonitrofenilo, azida, 3-(2-piridil ditio)-propionamida, glioxal, aldehído, iodoacetilo, éster de cianometilo, éster de p-nitrofenilo, éster de O-nitrofenilo, éster de hidroxipiridina, imidazol de carbonilo, y los otros grupos químicos de este tipo. En algunas formas de realización, la funcionalidad reactiva se puede unir directamente a un nucleótido, o se puede unir al nucleótido a través de un grupo de conexión. La fracción funcional y cualquier enlazador no puede sustancialmente perjudicar la capacidad del nucleótido que se va a añadir al ARNm o que se va a marcar. Grupos de unión representativos incluyen

carbono que contiene grupos de unión, típicamente que varían de aproximadamente 2 a 18, normalmente de aproximadamente 2 a 8 átomos de carbono, donde el carbono que contiene grupos de unión pueden o no incluir uno o más heteroátomos, por ejemplo S, O, N, etc., y pueden o no incluir uno o más sitios de insaturación. De interés particular en muchas formas de realización son los grupos de unión a alquilo, típicamente grupos de unión a alquilo inferiores de 1 a 16, normalmente de 1 a 4 átomos de carbono, donde los grupos de unión pueden incluir uno o más sitios de insaturación. Los nucleótidos funcionalizados (o cebadores) usados en los métodos anteriores de generación de objetivos funcionalizados se pueden fabricar utilizando protocolos conocidos o comprarse de vendedores comerciales, por ejemplo, Sigma, Roche, Ambion e IDT. Los grupos funcionales se pueden preparar según formas conocidas por los expertos en la técnica, con la información representativa de las patentes de EEUU números: 4.404.289, 4.405.711, 4.337.063 y 5.268.486 y Br. Pat. N° 1.529.202.

[0118] Nucleótidos modificados con amina se usan en diferentes formas de realización de la invención. El nucleótido modificado con amina es un nucleótido que tiene un grupo de amina reactivo para la fijación del marcador. Se contempla que cualquier ribonucleótido (G, A, U o C) o desoxirribonucleótido (G, A, T o C) se puede modificar para el marcado. Ejemplos incluyen, pero de forma no limitativa, los siguientes ribo- y desoxirribonucleótidos modificados: 5-(3-aminoalil)-UTP; 8-[(4-amino)butil]-amino-ATP y 8-[(6-amino)butil]-amino-ATP; N⁶-(4-amino)butil-ATP, N⁶-(6-amino)butil-ATP, N⁴-[2,2-oxi-bis(etilamina)]-CTP; N⁶-(6-Amino)hexil-ATP; 8-[(6-Amino)hexil]-amino-ATP; 5-propargilamino-CTP, 5-propargilamino-UTP; 5-(3-aminoalil)-dUTP; 8-[(4-amino)butil]-amino-dATP y 8-[(6-amino)butil]-amino-dATP; N-(4-amino)butil-dATP, N⁶-(6-amino)butil-dATP, N⁴-[2,2-oxi-to(etilamina)]-dCTP; N⁶-(6-Amino)hexil-dATP; 8-[(6-Amino)hexil]-amino-dATP; 5-propargilamino-dCTP, y 5-propargilamino-dUTP. Tales nucleótidos se pueden preparar según métodos conocidos por aquellos de habilidad en la técnica. Además, un experto en la materia podría preparar otras entidades de nucleótidos con la misma modificación con amina, tal como un 5-(3-aminoalil)-CTP, GTP, ATP, dCTP, dGTP, dTTP o dUTP en lugar de un 5-(3-aminoalil)-UTP.

Técnicas de marcado

[0119] En algunas formas de realización, los ácidos nucleicos se marcan añadiendo catalíticamente al ácido nucleico un nucleótido o nucleótidos ya marcados. Uno o más nucleótidos marcados se pueden añadir a las moléculas ARNmi. Véase la patente U.S. 6.723.509.

[0120] En otras formas de realización, un nucleótido no marcado o nucleótidos se añade catalíticamente a un ARNmi, y el nucleótido no marcado se modifica con una fracción química que permite que sea marcado posteriormente, en formas de realización de la invención, la fracción química es una amina reactiva de manera que el nucleótido es un nucleótido modificado con amina. Ejemplos de nucleótidos modificados con amina son conocidos por los de habilidad en la técnica, muchos están disponibles comercialmente tales como de Ambion, Sigma, Jena Bioscience y TriLink.

[0121] A diferencia del marcado de los ADNc durante sus síntesis, la cuestión del marcado de los ARNmi es cómo marcar la molécula ya existente. Con este fin, podemos usar una enzima capaz de utilizar un ribonucleótido de di- o trifosfato o desoxirribonucleótido como un sustrato para su adición a un ARNmi, una molécula de ARN pequeña. Además, en formas de realización específicas, implica el uso de un ribonucleótido de di- o trifosfato modificado, que se añade al extremo 3' de un ARNmi. La fuente de la enzima no es limitante. Ejemplos de fuentes para las enzimas incluyen levadura, bacterias gram-negativas tales como E. coli, lactococcus lactis y virus de la sífilis de oveja.

[0122] Enzimas capaces de añadir tales nucleótidos incluyen, pero de forma no limitativa, poli(A) polimerasa, transferasa terminal y polinucleótido fosforilasa. En formas de realización específicas de la invención, la ligasa se contempla como no siendo la enzima usada para añadir el marcador, y en cambio, una enzima sin ligasa es emplea.

[0123] Poli(A) polimerasa se ha clonado a partir de un número de organismos desde plantas a seres humanos. Se ha demostrado para catalizar la adición de tractos de homopolímero a ARN (Martin *et al.*, RNA, 4(2):226-30, 1998).

[0124] La transferasa terminal cataliza la adición de nucleótidos en el término 3' de un ácido nucleico.

[0125] La polinucleótido fosforilasa puede polimerizar difosfatos de nucleótido sin necesidad de un cebador.

Marcadores y etiquetas

[0126] Los ARNmi o las sondas de ARNmi se pueden marcar con una emisión de positrón (incluyendo radiactivo), enzimático, colorimétrico (incluye espectro visible y UV, incluyendo fluorescente), luminiscente u otro marcador o etiqueta para fines de detección o aislamiento. El marcador se puede detectar directa o

indirectamente. Las etiquetas radiactivas incluyen ¹²⁵I, ³²P, ³³P y ³⁵S. Ejemplos de etiquetas enzimáticas incluyen fosfatasa alcalina, luciferasa, peroxidasa de rábano y β-galactosidasa. Las etiquetas también pueden ser proteínas con propiedades luminiscentes, por ejemplo, proteína verde fluorescente y ficoeritrina.

5 [0127] Las etiquetas colorimétricas y fluorescentes contempladas para su uso como conjugados incluyen, pero de forma no limitativa, AMCA, tintes Alexa Fluor, tintes BODIPY, tal como BODIPY FL, BODIPY 630/650, BODIPY 650/665, BODIPY Y-R6G, BODIPY-TRX; Cascade Blue; Cascade Yellow; cumarina y sus derivados, tal como 7-amino-4-metilcumarina, aminocumarina e hidroxycumarina; tintes de cianina, tal como Cy3 y Cy5; eosinas y eritrosinas; fluoresceína y sus derivados, tal como isotiocianato de fluoresceína; quelatos macrocíclicos de iones de lantánido, tal como Quantum DyeTM; Marina Blue; Oregon Green; tintes de rodamina, tal como rojo de rodamina, tetrametilrodamina y rodamina 6G; Texas Red;

10 [0128] Ejemplos específicos de tintes incluyen, pero de forma no limitativa, aquellos identificados anteriormente y los siguientes: Alexa Fluor 350, Alexa Fluor 405, Alexa Fluor 430, Alexa Fluor 488, Alexa Fluor 500. Alexa Fluor 514, Alexa Fluor 532, Alexa Fluor 546, Alexa Fluor 555, Alexa Fluor 568, Alexa Fluor 594, Alexa Fluor 610, Alexa Fluor 633, Alexa Fluor 647, Alexa Fluor 660, Alexa Fluor 680, Alexa Fluor 700, and, Alexa Fluor 750; tintes amino-reactivos, tales como BODIPY 493/503, BODEPY 530/550, BODEPY 558/568, BODIPY 564/570, BODDPY 576/589, BODIPY 581/591, BODEPY 630/650, BODIPY 650/655, BODIPY FL, BODIPY R6G, BODEPY TMR, y, BODIPY-TR; Cy3, Cy5, 6-FAM, isotiocianato de fluoresceína, HEX, 6-JOE, Oregon Green 488, Oregon Green 500, Oregon Green 514, Pacific Blue, REG, Rhodamine Green, Rhodamine Red, Renographin, ROX, SYPRO, TAMRA, 2',4',5',7'-Tetrabromosulfonofluoresceína y TET.

15 [0129] Ejemplos específicos de ribonucleótidos fluorescentemente marcados están disponibles de Molecular Probes, e incluyen Alexa Fluor 488-5-UTP, Fluoresceína-12-UTP, BODEPY FL-14-UTP, BODIPY TMR-14-UTP, Tetrametilrodamina-6-UTP, Alexa Fluor 546-14-UTP, Texas Red-5-UTP y BODIPY TR-14-UTP. Otros ribonucleótidos fluorescentes están disponibles de Amersham Biosciences, tales como Cy3-UTP y Cy5-UTP. Ejemplos de deoxiribonucleótidos fluorescentemente marcados incluyen dinitrofenilo (DNP)-11-dUTP, Cascade Blue-7-dUTP, Alexa Fluor 488-5-dUTP, Fluoresceína-12-dUTP, Oregon Green 488-5-dUTP, BODEPY FL-14-dUTP, Rhodamine Green-5-dUTP, Alexa Fluor 532-5-dUTP, BODEPY TMR-14-dUTP, Tetramethylrhodamine-6-dUTP, Alexa Fluor 546-14-dUTP, Alexa Fluor 568-5-dUTP, Texas Red-12-dUTP, Texas Red-5-dUTP, BODEPY TR-14-dUTP, Alexa Fluor 594-5-dUTP, BODEPY 630/650-14-dUTP, BODIPY 650/665-14-dUTP; Alexa Fluor 488-7-OBEA-dCTP, Alexa Fluor 546-16-OBEA-dCTP, Alexa Fluor 594-7-OBEA-dCTP, Alexa Fluor 647-12-OBEA-dCTP. Se contempla que los ácidos nucleicos se pueden marcar con dos etiquetas diferentes.

20 [0130] Se contempla que los ARNm sintéticos se puedan marcar con más de un marcador, o con dos marcadores diferentes. Además, la transferencia de energía por resonancia de fluorescencia (FRET) se puede emplear en métodos de la invención (por ejemplo, Klostermeier *et al.*, 2002; Emptage, 2001; Didenko, 2001). Tintes de transferencia por energía fluorescente, tales como naranja de tiazol-heterodímero de etidio; y, TOTAB se pueden utilizar. Alternativamente, el marcador puede no ser detectable de por sí, sino detectable indirectamente o permitiendo el aislamiento o la separación del ácido nucleico previsto. Por ejemplo, el marcador podría ser biotina, digoxigenina, cationes polivalentes, grupos quelantes y los otros ligandos, incluyen ligandos para un anticuerpo.

Técnicas de visualización

25 [0131] Un número de técnicas para visualizar o detectar ácidos nucleicos marcados están disponibles fácilmente. La referencia de Stanley T. Crooke, 2000 tiene una discusión de tales técnicas (capítulo 6). Tales técnicas incluyen microscopía, matrices, fluorimetría, cicladores de luz u otras máquinas de PCR en tiempo real, análisis FACS, contadores de centelleo, generadores de imágenes por fósforo, contadores Geiger, MRI, CAT, métodos de detección basados en anticuerpos (Westerns, inmunofluorescencia, inmunohistoquímica), técnicas histoquímicas, HPLC (Griffey *et al.*, 1997, espectroscopia, electroforesis en gel capilar (Cummins *et al.*, 1996), espectroscopia; espectroscopia de masas; técnicas radiológicas; y técnicas de equilibrio de masa. Alternativamente, los ácidos nucleicos se pueden marcar o etiquetar para permitir para su aislamiento eficaz. En otras formas de realización de la invención, los ácidos nucleicos son biotinilados.

30 [0132] Cuando dos o más marcadores de color se emplean diferencialmente, las técnicas de transferencia de energía por resonancia fluorescente (FRET) se puede emplear para caracterizar el ARNbc. Además, una persona experta en la materia es bien consciente de las formas de visualizar, identificar y caracterizar los ácidos nucleicos marcados, y por consiguiente, tales protocolos se pueden utilizar como parte de la invención. Ejemplos de herramientas que se pueden utilizar también incluyen microscopía fluorescente, un bioanalizador, un lector de placa, Storm (Molecular Dynamics), escáner de matriz, FACS (clasificadora de células activadas fluorescentes), o cualquier instrumento que tenga la capacidad de excitar y detectar una molécula fluorescente.

Preparación de matriz

65

[0133] La presente invención se puede emplear con matrices de ARNmi, que son macromatrices o micromatrices ordenadas de moléculas (sondas) de ácidos nucleicos que son total o casi complementarias o idénticas a una pluralidad de moléculas de ARNmi o moléculas de ARNmi precursoras y que están posicionadas en un material de soporte en una organización separada espacialmente. Las macromatrices son típicamente hojas de nitrocelulosa o nylon sobre las que las sondas se han salpicado. Las micromatrices sitúan las sondas de ácidos nucleicos más densamente de manera que hasta 10.000 moléculas de ácido nucleico se pueden encontrar en una región típicamente de 1 a 4 centímetros cuadrados. Las micromatrices se pueden fabricar manchando moléculas de ácido nucleico, por ejemplo, genes, oligonucleótidos, etc., sobre sustratos o fabricando secuencias de oligonucleótidos *in situ* en un sustrato. Las moléculas de ácido nucleico manchadas o fabricadas se pueden aplicar a un modelo de matriz de alta densidad de hasta aproximadamente 30 moléculas de ácido nucleico no idénticos por centímetro cuadrado o superior, por ejemplo, hasta aproximadamente 100 o incluso 1.000 por centímetro cuadrado. Las micromatrices típicamente usan vidrio recubierto como soporte sólido, a diferencia del material basado en nitrocelulosa de las matrices de filtro. Al tener una matriz ordenada de muestras de ácido nucleico que complementan ARNmi, la posición de cada muestra se puede rastrear y enlazar a la muestra original. Una variedad de diferentes dispositivos de matriz donde una pluralidad de sondas de ácido nucleico distintas están asociadas de forma estable a la superficie de un soporte sólido se conocen por aquellos de habilidad en la técnica. Sustratos útiles para matrices incluyen nylon, vidrio y silicona. Tales matrices pueden variar en un varias formas diferentes, incluyendo longitud de sonda media, secuencia o tipos de sondas, naturaleza de enlace entre la sonda y la superficie de matriz, por ejemplo covalente o no covalente, y similares.

[0134] Métodos representativos y equipos para la preparación de una micromatriz se han descrito, por ejemplo, en las patentes de EEUU números 5.143.854; 5.202.231; 5.242.974; 5.288.644; 5.324.633; 5.384.261; 5.405.783; 5.412.087; 5.424.186; 5.429.807; 5.432.049; 5.436.327; 5.445.934; 5.468.613; 5.470.710; 5.472.672; 806; 5.525.464; 5.503.980; 5.510.270; 5.525.464; 5.527.681; 5.529.756; 5.532.128; 5.545.531; 5.547.839; 5.554.501; 5.556.752; 5.561.071; 5.571.639; 5.580.726; 5.580.732; 5.593.839; 5.599.695; 5.599.672; 5.610.287; 5.624.711; 5.631.134; 5.639.603; 5.654.413; 5.658.734; 5.661.028; 5.665.547; 5.667.972; 5.695.940; 5.700.637; 5.744.305; 5.800.992; 5.807.522; 5.830.645; 5.837.196; 5.871.928; 5.847.219; 5.876.932; 5.919.626; 6.004.755; 6.087.102; 6.368.799; 6.383.749; 6.617.112; 6.638.717; 6.720.138, al igual que WO 93/17126; WO 95/11995; WO 95/21265; WO 95/21944; WO 95/35505; WO 96/31622; WO 97/10365; WO 97/27317; WO 99/35505; WO 09923256; WO 09936760; WO0138580; WO 0168255; WO 03020898; WO 03040410; WO 03053586; WO 03087297; WO 03091426; WO03100012; WO 04020085; WO 04027093; EP 373 203; EP 785 280; EP 799 897 y UK 8 803 000. Se contempla que las matrices pueden ser matrices de alta densidad, de manera que contienen 100 o más sondas diferentes. Se contempla que pueden contener 1.000, 16.000, 65.000, 250.000 o 1.000.000 o más sondas diferentes. Las sondas se pueden dirigir a objetivos en uno o más organismos diferentes. Las sondas de oligonucleótidos varían de 5 a 50, de 5 a 45, de 10 a 40, o de 15 a 40 nucleótidos de longitud en algunas formas de realización, en algunas formas de realización, las sondas de oligonucleótidos son de 20 a 25 nucleótidos de longitud.

[0135] La ubicación y la secuencia de cada secuencia de sonda diferente en la matriz se conocen generalmente. Además, el gran número de diferentes sondas puede ocupar una área pequeña relativamente proporcionando una matriz de alta densidad que tiene una densidad de sonda generalmente mayor de aproximadamente 60, 100, 600, 1.000, 5.000, 10.000, 40.000, 100.000 o 400.000 sondas de oligonucleótidos diferentes por cm². El área de superficie de la matriz puede ser de aproximadamente o menos de aproximadamente 1, 1,6, 2, 3, 4, 5, 6, 7, 8, 9 o 10 cm². Además, una persona experta en la técnica en la materia podría fácilmente analizar datos generados utilizando una matriz. Tales protocolos se han descrito anteriormente, e incluyen información que se encuentra en WO 9743450; WO 03023058; WO 03022421; WO 03029485; WO03067217 WO 03066906; WO 03076928; WO 03093810; WO 03100448A1. Recientemente, métodos de perfilación alternativos se han hecho disponibles, basados en hibridación de solución e inmovilización e identificación posterior por ejemplo, la plataforma Illumina.

Preparación de la muestra

[0136] Se contempla que el ARNmi de una amplia variedad de muestras se puede analizar utilizando ensayos descritos aquí. Mientras que el ARNmi endógeno se contempla para su uso con algunas formas de realización, el ARNmi recombinante o sintético - incluyendo ácidos nucleicos que son idénticos a ARNmi endógeno o ARNmi precursor también se pueden manejar y analizar como se describe en este caso. Las muestras pueden ser muestras biológicas, en cuyo caso, pueden ser de sangre, fluido cerebroespinal, tejido, órganos, tumor, semen, esputo, deposición, orina, saliva, lágrimas, otro fluido corporal, folículos pilosos, piel, o cualquier muestra que contenga o esté formada por células biológicas. Alternativamente, la muestra puede no ser una muestra biológica, sino ser una mezcla química, tal como una mezcla reactiva libre de células (que puede contener una o más enzimas biológicas).

Ensayos celulares para identificar ARNmi con lazos con la enfermedad

[0137] Específicamente se contemplan dentro de la descripción las aplicaciones que incluyen la identificación de los ARNmi que contribuyen a la EMT que son partes ellos mismos de una enfermedad o condición o pueden de otro modo estar asociados a un estado de enfermedad particular. Adicionalmente, una aplicación contemplada incluye la identificación de ARNmi que sean capaces de revertir la EMT e inducir la MET. También, las funciones de ARNmi se pueden comparar entre una que se cree que es susceptible de una enfermedad particular o condición asociada a la EMT y una que se cree que no es susceptible o resistente a esa enfermedad o condición. Concretamente se contempla que las moléculas de ARN de la presente invención pueden utilizarse para tratar cualquiera de las enfermedades o condiciones discutidas en la sección precedente o para modular cualquiera de las vías celulares discutidas en la sección precedente. Aplicaciones que se contemplan específicamente incluyen la identificación de los ARNmi que contribuyen a los procesos celulares de EMT que son partes ellos mismos de una enfermedad o pueden de otro modo estar asociados a un estado de enfermedad particular. También, las funciones de ARNmi se pueden comparar entre una muestra que se cree que es susceptible de una enfermedad particular o condición asociada a la EMT y una que se cree que no es susceptible o resistente a esa enfermedad o condición.

[0138] La eficacia de diferentes fármacos terapéuticos se puede alterar mediante los ARNmi tal y como se definen y usan según la presente invención. Además, se ha descrito que las células tumorales que han sido sometidas a EMT pueden hacerse resistentes a la quimioterapia y la inmunoterapia (Thiery *et al.* 2009 Cell 139:871-90). Por lo tanto, los fármacos basados en ARNmi que inducen la inversión de la EMT pueden mejorar la susceptibilidad a, por ejemplo, a la quimioterapia y la inmunoterapia. Tales fármacos terapéuticos incluyen, pero de forma no limitativa, fármacos quimioterapéuticos. Un "agente quimioterapéutico" se utiliza para connotar un compuesto o composición que se administra en el tratamiento contra el cáncer. Estos agentes o fármacos se categorizan por su modo de actividad dentro de una célula, por ejemplo, si afectan y hasta qué grado afectan al ciclo celular. Alternativamente, un agente se puede caracterizar basado en su capacidad para reticular directamente ADN, para intercalar en ADN, o para inducir aberraciones cromosómicas y mitóticas afectando la síntesis de ácido nucleico. Muchos agentes quimioterapéuticos entran dentro de las siguientes categorías: agentes alquilantes, antimetabolitos, antibióticos antitumorales, inhibidores mitóticos y nitrosureas.

[0139] Ejemplos de agentes quimioterapéuticos incluyen agentes alquilantes tales como tiotepa y ciclosfosfamida; sulfonatos de alquilo tales como busulfano, improsulfano y piposulfano; aziridinas tales como benzodopa, carbocuaona, meturedopa y uredopa; etileniminas y metilamelaminas incluyendo altretamina, trietilenemelamina, trietilenofosforamida, trietilenotiofosforamida y trimetilolomelamina; acetogeninas (especialmente bulatacina y bulatacinona); una camptotecina (incluyendo el análogo sintético topotecán); briostatina; calistatina; CC-1065 (incluyendo sus análogos sintéticos de adocolesina, carcelesina y bicelesina); criptoficinas (particularmente criptoficina 1 y criptoficina 8); dolastatina; duocarmicina (incluyendo los análogos sintéticos, KW-2189 y CB1-TM1); eleuterobina; pancratistatina; una sarcodictina; espongistatina; mostazas de nitrógeno tales como clorambucil, clornafacina, colofosfamida, estramustina, ifosfamida, mecloretamina, hidrocloreuro de óxido de mecloretamina, melfalán, novembichina, fenesterina, prednimustina, trofosfamida, mostaza de uracilo; nitrosureas tales como carmustina, clorozotocina, fotemustina, lomustina, nimustina y ranimustina; antibióticos tales como los antibióticos de enedina (por ejemplo, caliqueamicina, especialmente caliqueamicina gama y caliqueamicina omega); dinemicina, incluyendo dinemicina A; bisfosfonatos, tales como clodronato; una esperamicina; al igual que cromóforo de neocarcinostatina y cromóforos antibióticos de cromoproteína enediina relacionados, aclacinomisinas, actinomisina, autramnicin, azaserina, bleomicinas, cactinomicina, carabicina, carminomicina, carzinoflina, cromomicinias, dactinomicina, daunorubicina, detorubicina, 6-diazo-5-oxo-L-norleucina, doxorubicina (incluyendo morfolmo-doxorubicina, cianomorfolino-doxorubicina, 2-pirrolino-doxorubicina y deoxidoxorubicina), epirubicina, esorubicina, idarubicina, marcelomicina, mitomicinas tales como mitomicina C, ácido micofenólico, nogalarnicina, olivomicinas, peplomicina, potfrimicina, puromicina, quelamicina, rodorubicina, estreptonigrina, estreptozocina, tubercidina, ubenimex, cinostatina, zorubicina; anti-metabolitos tales como metotrexato y 5-fluorouracil (5-FU); análogos de ácido fólico tal como denopterina, metotrexato, pteropterina, trimetrexato; análogos de purina tales como fludarabina, 6-mercaptopurina, tiamiprina, tioguanina; análogos de pirimidina tales como ancitabina, azacitidina, 6-azauridina, carmofur, citarabina, didesoxiuridina, doxifluridina, enocitabina, floxuridina; andrógenos tales como calusterona, dromostanolona propionato, epitioestano, mepitioestano, testolactona; anti-suprarrenales tales como aminoglutetimida, mitotano, trilostano; reforzador de ácido fólico tal como ácido frolínico; aceglatona; glicósido de aldofosfamida; ácido aminolevulínico; eniluracil; amsacrina; bestrabucil; bisantreno; edatraxato; defofamina; demecolcina; diazicuaona; elformitina; acetato de eliptinio; un epotilono; etoglúcido; nitrato de galio; hidroxiurea; lentinano; lonidainina; maitansinoides tales como maitansina y ansamitocinas; mitoguaona; mitoxantrona; mopidanmol; nitraerina; pentostatina; fenamet; pirarubicina; losoxantrona; ácido podofilínico; 2-etilhidrácido; procarbazona; complejo polisacárido PSK; razoxano; rhizoxina; sizofirana; espirogermanio; ácido tenuazónico; triazicuaona; 2,2',2"-triclorotrietilamina; tricotecenas (especialmente T-2 toxina, verracurina A, roridina A y anguidina); uretano; vindesina; dacarbazina; manomustina; mitobronitol; mitolactol; pipobromano; gacitosina; arabinósido ("Ara-C"); ciclofosfamida; tiotepa; taxoides, por ejemplo, paclitaxel y doxetaxel; clorambucil; gemcitabina; 6-tioguanina; mercaptopurina; metotrexato; complejos de coordinación de platino tales como cisplatino, oxaliplatino y carboplatino; vinblastina; platino; etopósido (VP-16); ifosfamida; mitoxantrona; vincristina; vinorelbina; novantrona; tenipósido; edatraxato; daunomicina; aminopterina; xeloda; ibandronato; irinotecán (por ejemplo, CPT-II); inhibidor de topoisomerasa RFS 2000; difluoromethylornitina (DMFO); retinoides

tales como ácido retinoico; capecitabina; y sales farmacéuticamente aceptables, ácidos o derivados de cualquiera de los anteriores.

[0140] También se incluyen en esta definición los agentes anti-hormonales que actúan para regular o inhibir la acción de la hormona en tumores tales como los anti-estrógenos y moduladores receptores de estrógenos selectivos (SERMs), incluyendo, por ejemplo, tamoxifeno, raloxifeno, droloxifeno, 4-hidroxitamoxifeno, trioxifeno, keoxifeno, LYI 17018, onapristona y toremifeno; inhibidores de aromatasa que inhiben la aromatasa enzimática, que regula la producción de estrógenos en las glándulas suprarrenales, tales como, por ejemplo, 4(5)-imidazolas, aminoglucetimidina, acetato de megestrol, exemestano, formestania, fadrozol, vorozol, letrozol y anastrozol; y anti-andrógenos tales como flutamida, nilutamida, bicalutamida, leuprolida y goserelina; al igual que troxacitabina (un análogo de 1,3-dioxolano nucleósido citosina); oligonucleótidos antisentido, particularmente los que inhiben la expresión de genes en la señalización de vías implicadas en la proliferación celular aberrante, tales como, por ejemplo, PKC- α , Raf y H-Ras; ribozimas tales como un inhibidor de expresión VEGF y un inhibidor de expresión HER2; vacunas tales como vacunas de terapia génica y sales farmacéuticamente aceptables, ácidos o derivados de cualquiera de los anteriores. Una lista de fármacos oncológicos aprobada por la FDA de EEUU con sus indicaciones aprobadas se puede encontrar en www.accessdata.fda.gov/scripts/cder/onctools/druglist.cfm. Además, se contempla que muestras que tienen diferencias en la actividad de ciertas vías también se puedan comparar. Tales vías celulares incluyen pero de forma no limitativa lo siguiente: cualquier vía de adhesión o de motilidad incluyendo, pero no limitado a, aquellas que implican AMP cíclico, proteína quinasa A, receptores de pares de proteína de G, adenilciclase, L-selectina, E-selectina, PECAM, VCAM-I, α -actinina, paxilina, cadherinas, AKT, integrina- α , integrina- β , RAF-I, ERK, PI-3 quinasa, vinculina, metaloproteinasas matriciales, Rho GTPasas, p85, factores trébol, profilina, FAK, MAP quinasa, Ras, caveolina, calpaína-1, calpaína-2, receptor de factor de crecimiento epidérmico, ICAM-1, ICAM-2, cofilina, actina, gelsolina, Rho A, Rac, quinasa de cadena ligera de la miosina, receptor de factor de crecimiento derivado de las plaquetas o eprina; cualquier vía de apoptosis incluyendo, pero no limitada a, aquellas que implican AKT, ligando Fas, NF κ B, caspasa-9, quinasa PB, caspasa-3, caspasa-7, ICAD, CAD, EndoG, granzima B, Bad, Bax, Bid, Bak, APAF-I, citocromo C, p53, ATM, Bcl-2, PARP, Chk1, Chk2, Rho-21, c-Jun, Rho73, Rad51, Mdm2, Rad50, c-Abl, BRCA-I, perforina, caspasa-4, caspasa-8, caspasa-6, caspasa-1, caspasa-2, caspasa-10, Rho, Jun quinasa, Jun quinasa quinasa, Rip2, lamina-A,, lamina-BI, lamina-B2, receptor Fas, H₂O₂, granzima A, NADPH oxidasa, HMG2, CD4; CD28, CD3, TRADD, IKK, FADD, GADD45, receptor de muerte DR3, receptor de muerte DR4/5, FLIPs, APO-3, GRB2, SHC, ERK, MEK, RAF-1, AMP cíclico, proteína quinasa A, E2F, proteína de retinoblastoma, Smac/Diablo, receptor ACH, 14- 3-3, FAK, SODD, receptor TNF, RTP, ciclina DI, PCNA, Bcl-XL, PIP2, PIP3, PTEN, ATM, Cdc2, proteína quinasa C, calcineurina, IKK α , IKK β , IKK γ , SOS-I, c-FOS, Traf-1, Traf-2, I κ B β o el proteasoma; cualquier vía de activación celular incluyendo, pero no limitada a, aquellas que implican proteína quinasa A, óxido nítrico, caveolina-1, actina, calcio, proteína quinasa C, Cdc2, ciclina B, Cdc25, GRB2, SRC proteína quinasa, factores de ribosilación de ADP (ARFs), fosfolipasa D, AKAP95; p68, Aurora B, CDK1, Eg7, histona H3, PKAc, CD80, PI3 quinasa, WASP, Arp2, Arp3, p34, p20, PP2A, angiotensina, enzima de conversión de angiotensina, receptor activado de proteasa-1, receptor activado de proteasa-4, Ras, RAF-I, PLC β , PLC γ , COX-I, receptores acoplados a proteína G, fosfolipasa A2, IP3, SUMO1, SUMO 2/3, ubiquitina, Ran, Ran-GAP, Ran-GEF, p53, glucocorticoides, receptor de glucocorticoides, componentes del complejo SWI/SNF, RanBP1, RanBP2, importinas, exportinas, RCC1, CD40, ligando CD40, p38, DCK α , IRK β , NF κ B, TRAF2, TRAF3, TRAF5, TRAF6, IL-4, receptor de IL-4, CDK5, factor de transcripción de AP-I, CD45, CD4, receptores de células T, quinasa MAP, factor de crecimiento de nervio, receptor de factor de crecimiento de nervio, c-Jun, c-Fos, Jun kinase, GRB2, SOS-I, ERK-I, ERK, JAK2, STAT4, IL-12, receptor de IL-12, sintasa de óxido nítrico, TYK2, IFN γ , elastasa, IL-8, epitelinas, IL-2, receptor de IL-2, CD28, SMAD3, SMAD4, TGF β o receptor de TGF β ; cualquier ruta de regulación de ciclo celular, señalización o diferenciación incluyendo pero, no limitado a, aquellas que implican TNFs, SRC proteína quinasa, Cdc2, ciclina B, Grb2, Sos- 1, SHC, p68, Aurora quinasas, proteína quinasa A, proteína quinasa C, Eg7, p53, ciclinas, quinasas dependientes de ciclina, factor de crecimiento neuronal, factor de crecimiento epidérmico, proteína de retinoblastoma, ATF-2, ATM, ATR, AKT, CHK1, CHK2, 14-3-3, WEE1, CDC25 CDC6, proteínas de complejo de reconocimiento de origen, p15, p16, p27, p21, ABL, c- ABL, SMADs, ubiquitina, SUMO, proteínas de choque térmico, Wnt, GSK-3, angiotensina, p73 cualquier PPAR, TGF α , TGF β , p300, MDM2, GADD45, Notch, cdc34, BRCA-I, BRCA- 2, SKP1, el proteasoma, CUL1, E2F, pi 07, hormonas esteroideas, receptores de hormonas esteroideas, I κ B α , I κ B β , Sin3A, proteínas de choque térmico, Ras, Rho, ERKs, IKKs, PI3 quinasa, Bcl-2, Bax, PCNA, quinasas MAP, dineína, RhoA, PKAc, ciclina AMP, FAK, PIP2, PIP3, integrinas, trombopoyetina, FAS, ligando Fas, PLK3, MEKs, JAKs, STATs, acetilcolina, calcineurina de paxilina, p38, importinas, exportinas, Ran, Rad50, Rad51, ADN polimerasa, ARN-polimerasa, Ran-GAP, Ran-GEF, NuMA, Tpx2, RCC1, Sonic Hedgehog, Crml, Patched (Ptc-1), MPF, CaM quinasas, tubulina, actina, proteínas asociadas a cinetocoro, proteínas de unión a centrómero, telomerasa, TERT, PP2A, c-MYC insulina, receptores de células T, receptores de células B, CBP, 1KB, NF κ B, RAC1, RAF1, EPO, diacilglicerol, c-Jun, c-Fos, Jun quinasa, factores inducibles por hipoxia, GATA4, β -catenina, α -catenina, calcio, arrestina, survivina, caspasas, procaspases, CREB, CREM, cadherinas, PECAMs, corticoesteroides, factores que estimulan colonias, calpaínas, adenilciclase, factores de crecimiento, óxido nítrico, receptores de transmembrana, retinoides, proteínas G, canales iónicos, activadores transcripcionales, coactivadores transcripcionales, represores transcripcionales, interleuquinas, vitaminas, interferones, correpresores transcripcionales, el poro nuclear, nitrógeno, toxinas, proteólisis, o fosforilación; o cualquier vía metabólica incluyendo pero, no limitada a, las que implican la biosíntesis de aminoácidos, oxidación de ácidos grasos, biosíntesis de neurotransmisores y otras moléculas de señalización celular, biosíntesis de

poliaminas, biosíntesis de lípidos y esfingolípidos, catabolismo de aminoácidos y nutrientes, síntesis de nucleótidos, eicosanoides, reacciones de transporte de electrones, degradación asociada a ER, glucólisis, fibrinólisis, formación de cuerpos cetónicos, formación de fagosomas, metabolismo de colesterol, regulación de ingesta de alimentos, homeostasis de energía, activación de protrombina, síntesis de lactosa y otros azúcares, resistencia a multifármacos, biosíntesis de fosfatidilcolina, el proteasoma, proteína precursora amiloide, Rab

5 GTPasas, síntesis de almidón, glicosilación, síntesis de fosfoglicéridos, vitaminas, el ciclo del ácido cítrico, receptor IGF-I, el ciclo de la urea, transporte vesicular o vías de salvamento. Además se contempla que las moléculas de ácidos nucleicos de la invención se puedan emplear en métodos de diagnóstico y terapéuticos con respecto a cualquiera de las vías o factores anteriores. Así, en algunas formas de realización de la invención, un

10 ARNmi inhibe, elimina, activa, induce, aumenta o de otro modo modula una o más de las vías anteriores o factores se contempla como parte de métodos de la invención. El ácido nucleico se puede usar en la divulgación para diagnosticar una enfermedad o condición basada en la relación de ese ARNmi con cualquiera de las vías anteriormente descritas.

15 Otros ensayos

[0141] Además del uso de matrices y micromatrices, se contempla que un número de ensayos de diferencia se podrían emplear para analizar los ARNmi, sus actividades y sus efectos. Tales ensayos incluyen, pero de forma no limitativa, RT-PCR, hibridación *in situ*, ensayo de protección de hibridación (HPA)(GenProbe), ensayo de ADN ramificado (bDNA) (Collins, M. L. *et al.* (1997). *Nucleic Acids Research* 25: 2979-2984), amplificación de círculo rodante (RCA), detección de hibridación de molécula única (US Genomics), ensayo invasor (ThirdWave Technologies) y ensayo de unión puente (Qiagen). Se contempla que tales métodos se pueden utilizar en el contexto de matrices, al igual que en el contexto de ensayos de diagnóstico.

25 Aplicaciones terapéuticas y de diagnóstico

[0142] Los ARNmi que afectan rasgos fenotípicos proporcionan puntos de intervención para aplicaciones terapéuticas al igual que para aplicaciones de diagnóstico (mediante la selección para la presencia o ausencia de un ARNmi particular). Concretamente se contempla que las moléculas de ARN de la presente invención pueden utilizarse para tratar cualquiera de las enfermedades o condiciones discutidas en la sección precedente. Además, cualquiera de los métodos anteriormente descritos también se puede emplear con respecto a aspectos terapéuticos de la invención. Además, cualquiera de los métodos anteriormente descritos también se puede emplear con respecto a aspectos de diagnóstico de la divulgación. Por ejemplo, métodos con respecto a la detección de ARNmi o selección de éstos también se pueden emplear en un contexto de diagnóstico. En usos terapéuticos, una cantidad eficaz de los ARNmi de la presente invención se administra a una célula, que puede o no estar en un animal. En algunas formas de realización, una cantidad terapéuticamente eficaz de los ARNmi de la presente invención se administra a un individuo para el tratamiento contra el cáncer asociado a la EMT. El término "cantidad eficaz", como se utiliza en este caso, se define como la cantidad de las moléculas de la presente invención que es necesaria para dar como resultado el cambio fisiológico deseado en la célula o tejido al que se administra. El término "cantidad terapéuticamente eficaz", como se utiliza en este caso, se define como la cantidad de moléculas de la presente invención que consigue un efecto deseado con respecto a un cáncer asociado a la EMT como se ha definido en la presente anteriormente. Un experto en la materia reconoce fácilmente que en muchos casos las moléculas pueden no proporcionar una cura pero pueden proporcionar un beneficio parcial, tal como un alivio o mejora de al menos un síntoma. En algunas formas de realización, un cambio fisiológico que tiene algún beneficio se considera también beneficioso terapéuticamente. Así, en algunas formas de realización, una cantidad de moléculas que proporciona un cambio fisiológico se considera una "cantidad efectiva" o una "cantidad terapéuticamente efectiva".

[0143] En algunas formas de realización una molécula tiene una secuencia que corresponde a la secuencia de ARNmi a partir de ese animal particular, a diferencia de otro animal. Así, en algunas formas de realización, una secuencia humana se utiliza como una molécula de ARN de la presente invención. En experimentos *in vivo*, una secuencia ARNmi usada en un animal de prueba puede diferir de una secuencia humana correspondiente. En este caso, un ARNmi que difiere de la secuencia humana se puede usar para demostrar efecto terapéutico en el animal. Los resultados obtenidos con esta secuencia evaluada en un animal se pueden extrapolar a resultados previstos en humanos con una molécula de ARNmi correspondiente.

55 Modos de administración y formulaciones

[0144] Las moléculas de ácido nucleico de la invención se pueden administrar a un sujeto solo o en forma de una composición farmacéutica para el tratamiento de una condición o enfermedad. Las composiciones farmacéuticas se pueden formular de manera convencional usando uno o más portadores fisiológicamente aceptables, diluyentes, excipientes o productos auxiliares que facilitan el procesamiento del ARNmi en preparaciones que se pueden usar farmacéuticamente. La formulación apropiada depende de la ruta de administración elegida. Para la administración tópica los ARNmi de la invención se pueden formular como soluciones, geles, pomadas, cremas, suspensiones, etc. como bien se conocen en la técnica. Las formulaciones sistémicas incluyen aquellas diseñadas para la administración por inyección, por ejemplo, inyección subcutánea, intravenosa, intramuscular,

intratecal o intraperitoneal, al igual que las diseñadas para administración transdérmica, transmucosa, inhalación, oral o pulmonar. Para la inyección, los ácidos nucleicos de la invención se pueden formular en soluciones acuosas, preferiblemente en tampones compatibles fisiológicamente tales como la solución de Hanks, la solución de Ringer, o tampón de solución salina fisiológica. La solución puede contener agentes formulatorios tales como agentes suspensión, estabilizantes y/o de dispersión.

[0145] Alternativamente, las moléculas de ácido nucleico pueden ser en polvo para la constitución con un vehículo adecuado, por ejemplo, agua sin pirógeno estéril, antes de su uso. Para la administración transmucosa, penetrantes apropiados para que la barrera sea permeada se usan en la formulación. Tales penetrantes se conocen generalmente en la técnica. Para la administración oral, los ácidos nucleicos se pueden formular fácilmente combinando las moléculas con portadores farmacéuticamente aceptables bien conocidos en la técnica. Tales portadores permiten que los ácidos nucleicos de la invención se formulen como comprimidos, píldoras, grageas, cápsulas, líquidos, geles, jarabes, lodos, suspensiones y similares, para ingestión oral por un paciente a tratar. Para formulaciones sólidas orales tales como, por ejemplo, polvos, cápsulas y comprimidos, excipientes adecuados incluyen productos de relleno tales como azúcares, por ejemplo lactosa, sacarosa, manitol y sorbitol; preparaciones de celulosa tales como almidón de maíz, almidón de trigo, almidón de arroz, almidón de patata, gelatina, goma tragacanto, metilcelulosa, hidroxipropilmetilcelulosa, carboximetilcelulosa sódica y/o polivinilpirrolidona (PVP); agentes de granulación; y agentes aglutinantes. Si se desea, agentes de desintegración pueden ser adicionados, como la polivinilpirrolidona reticulada, agar, o ácido alginico o una sal derivada tal como alginato de sodio. Si se desea, formas de dosificación sólidas puede estar recubiertas con azúcar o recubiertas entéricas utilizando técnicas estándar. Para preparaciones líquidas oral tales como, por ejemplo, suspensiones, elixires y soluciones, portadores, excipientes o diluyentes adecuados incluyen agua, glicoles, aceites, alcoholes, etc. Adicionalmente, agentes aromatizantes, conservantes, agentes colorantes y similares pueden ser adicionados. Para la administración bucal, las moléculas puede tomar la forma de comprimidos, pastillas, etc. formulados de manera convencional. Para la administración por inhalación, las moléculas para su uso según la presente invención se entregan convenientemente en forma de un aerosol desde embalajes presurizados o un nebulizador, con el uso de un propulsor adecuado, por ejemplo, diclorodifluorometano, triclorofluorometano, diclorotetrafluoroetano, dióxido de carbono u otro gas adecuado. En el caso de un aerosol presurizado la unidad de dosificación se puede determinar mediante una válvula para entregar una cantidad dosificada. Las cápsulas y cartuchos de gelatina para su uso en un inhalador o insuflador se pueden formular con una mezcla de polvo de los ácidos nucleicos y una base de polvo adecuada tal como lactosa o almidón. Las moléculas de ARN también se pueden formular en composiciones rectales o vaginales tales como supositorios o enemas de retención, por ejemplo, que contienen bases de supositorio convencionales tal como manteca de cacao u otros glicéridos.

[0146] Además de las formulaciones descritas previamente, las moléculas también se pueden formular como una preparación de depósito. Tales formulaciones de larga actuación se pueden administrar por implantación (por ejemplo subcutáneamente o intramuscularmente) o por inyección intramuscular. Así, por ejemplo, las moléculas se pueden formular con materiales poliméricos adecuados o hidrofóbicos (por ejemplo como una emulsión en un aceite aceptable) o resinas de intercambio de iones, o como derivados difícilmente solubles, por ejemplo, como una sal difícilmente soluble. Alternativamente, otros sistemas de entrega farmacéutica se pueden emplear.

[0147] Liposomas y emulsiones son ejemplos bien conocidos por vehículos de entrega que se pueden utilizar para entregar ácidos nucleicos de la invención.

[0148] Un ácido nucleico de la invención se puede administrar en combinación con un portador o lípido para aumentar la absorción celular. Por ejemplo, el oligonucleótido se puede administrar en combinación con un lípido catiónico. Ejemplos de lípidos catiónicos incluyen, pero de forma no limitativa, lipofectina, DOTMA, DOPE y DOTAP. La publicación de la WO0071096 describe formulaciones diferentes, tales como un DOTAP, colesterol o formulación derivada de colesterol que se puede usar eficazmente para terapia génica. Otras descripciones también discuten diferentes formulaciones de lípidos o liposómicas incluyendo nanopartículas y métodos de administración; estas incluyen, pero de forma no limitativa, la publicación de patente de EEUU 20030203865, 20020150626, 20030032615 y 20040048787. Los métodos usados para la formación de partículas también se describen en las patentes de EEUU números: 5.844.107, 5.877.302, 6.008.336, 6.077.835, 5.972.901, 6.200.801 y 5.972.900. Los ácidos nucleicos también se pueden administrar en combinación con una amina catiónica tal como poli-L-lisina. Los ácidos nucleicos también se pueden conjugar a una fracción química, tal como transferrina y colesterol. Además, los oligonucleótidos se pueden dirigir a ciertos órganos o tejidos mediante el enlace de grupos químicos específicos al oligonucleótido. Por ejemplo, la unión del oligonucleótido a una matriz adecuada de residuos de manosa dirigirá el oligonucleótido al hígado. Otros ligandos objetivo se describen en Liu B., Brief Funct. Genomic Proteomic 6:112-119, 2007. Ejemplos adicionales son azúcares de carbohidrato tales como galactosa, N-acetilgalactosamina, manosa; vitaminas tales como folatos; moléculas pequeñas incluyendo naproxen, ibuprofeno u otras moléculas de unión a proteína conocidas, ciclodextrina, que se dirige al receptor de transferrina, también llamada ciclodextrina modificada de transferencia (Hu-Lieskovan *et al.*, 2005), PEI (RGD-targeted PEG-PEI, Schiffelers *et al.* 2004), anisamida, RGD-péptido o miméticos de RGD, poliarginina, fragmento de anticuerpo monocatenario anti-TfR/TfRscFv, anexina A5 (que se dirige a las membranas que exponen fosfatidilserina, Garnier B. *et al.*, Bioconjug Chem., 2009, 11:2114-22), la WO 2009/126933 que describe

composiciones y métodos para la entrega específica de sitio de ácidos nucleicos mediante la combinación de éstos con los ligandos objetivo y componentes endosomolíticos. Los ligandos de objetivo que son preferentemente adecuados son proteínas de superficie celular asociadas a tumores, más preferiblemente proteínas de superficie celular asociadas a tumores de próstata. La dirección de ácidos nucleicos también se puede realizar usando tecnología de aptámero como se describe en la WO2005/111238. Además, fracciones de lípidos adicionales, tales como PEG-lípidos, colesterol, lípidos auxiliares endosomolíticos o péptidos (WO2009/046220) o la morfología general de las nanopartículas generadas (caracterizadas por la carga y el tamaño de partícula) para los vehículos de entrega anteriormente mencionados pueden conferir especificidad de objetivo para células cancerosas y/o vasculatura tumoral.

[0149] Adicionalmente, las moléculas se pueden entregar utilizando un sistema de liberación prolongada, tal como matrices semipermeables de polímeros sólidos que contienen el agente terapéutico. Varios de los materiales de liberación prolongada se han establecido y son conocidos por aquellas personas expertas en la técnica. Las cápsulas de liberación prolongada pueden, dependiendo de su naturaleza química, liberar las moléculas durante unas pocas semanas hasta más de 100 días. Dependiendo de la naturaleza química y la estabilidad biológica de las moléculas químicas, estrategias adicionales para la estabilización de moléculas se pueden emplear.

[0150] Alternativamente, las moléculas se pueden entregar utilizando un sistema de entrega a base de química de coordinación como se describe en la WO2007011217.

[0151] Los ácidos nucleicos se pueden incluir en cualquiera de las formulaciones descritas anteriormente como los ácidos libres o bases o como sales farmacéuticamente aceptables. Sales farmacéuticamente aceptables son aquellas sales que sustancialmente retienen la actividad biológica de las bases libres y que se preparan por reacción con ácidos inorgánicos. Las sales farmacéuticas tienden a ser más soluble en solventes acuosos y otros solventes próticos que las formas de base libre correspondientes.

[0152] Las composiciones farmacéuticas de la presente invención comprenden una cantidad eficaz de una o más moléculas de ARNmi disueltas o dispersadas en un portador farmacéuticamente aceptable. Las frases "farmacéutico o farmacéuticamente aceptable" se refieren a entidades moleculares y composiciones que no producen o producen reacciones adversas, alérgicas u otras reacciones adversas aceptables cuando se administran a un animal, tal como, por ejemplo, un humano, según sea apropiado. Si ciertos efectos adversos son aceptables se determina según la gravedad de la enfermedad.

La preparación de una composición farmacéutica que contiene al menos un polipéptido químico o ingrediente activo adicional será conocido por los expertos en la técnica a la luz de la presente descripción, como se ejemplifica en Remington's Pharmaceutical Sciences, 18th Ed. Mack Printing Company, 1990. Además, para la administración a un animal (por ejemplo, humano) se entiende que las preparaciones debería cumplir con los estándares de esterilidad, pirogenicidad, seguridad general y pureza como establece la FDA Office of Biological Standards.

[0153] Como se utiliza en este caso, "portador farmacéuticamente aceptable" incluye cualquiera y todos los solventes, medios de dispersión, recubrimientos, surfactantes, antioxidantes, conservantes (por ejemplo, agentes antibacterianos, agentes antifúngicos), agentes isotónicos, agentes de retraso de absorción, sales, conservantes, fármacos, estabilizadores de medicamento, geles, ligantes, excipientes, agentes de desintegración, lubricantes, agentes edulcorantes, agentes aromatizantes, tintes, materiales similares y combinaciones de los mismos, como conocería un técnico en la materia (véase, por ejemplo, Remington's Pharmaceutical Sciences, 18th Ed. Mack Printing Company, 1990, pp. 1289-1329). Excepto en la medida en que cualquier portador convencional es incompatible con la sustancia activa, su uso en las composiciones farmacéuticas o terapéuticas se contempla.

[0154] Los ARNmi pueden comprender diferentes tipos de portadores dependiendo de si son para administrar en forma de sólido, líquido o aerosol, y de si necesitan ser estériles para este tipo de formas de administración como inyección. La presente invención se puede administrar por vía intravenosa, intradérmica, intraarterial, intraperitoneal, intralesional, intracraneal, intraarticular, intraprostática, intrapleurales, intratraqueal, intranasal, intravítrea, intravaginal, intrarrectal, tópica, intratumoral, intramuscular, intraperitoneal, subcutánea, subconjuntival, intravesicular, mucosa, intrapericardial, intraumbilical, intraocular, oral, tópica, local, inhalación (por ejemplo, inhalación de aerosol), inyección, infusión, infusión continua, perfusión localizada bañando células diana directamente, vía un catéter, vía un lavado, en cremas, en composiciones lipídicas (por ejemplo, liposomas), o por otro método o cualquier combinación de lo anterior como conocería un técnico en la materia (véase, por ejemplo, Remington's Pharmaceutical Sciences, 18th Ed. Mack Printing Company, 1990).

[0155] La cantidad de dosificación real de una composición de la presente invención administrada a un animal o un paciente se puede determinar por factores físicos y fisiológicos tales como peso corporal, gravedad de condición, el tipo de enfermedad que se está tratando, intervenciones terapéuticas precedentes o concurrentes, idiopatía del paciente y en la forma de administración. El profesional responsable de la administración

determinará, de cualquier manera, la concentración de ingrediente(s) activo(s) de una composición y la dosis apropiada para el sujeto particular.

5 [0156] En ciertas formas de realización, las composiciones farmacéuticas pueden comprender, por ejemplo, al menos aproximadamente 0,1% de un compuesto activo. En otras formas de realización, un compuesto activo puede comprender entre aproximadamente 2% a aproximadamente 75% del peso de la unidad, o entre aproximadamente 25% a aproximadamente 60%, por ejemplo, o de 2% a 75% del peso de la unidad o de 25% a 60% por ejemplo y cualquier gama derivable de esta. En otros ejemplos no limitativos, una dosis también puede comprender menos de 1 microgramo/kg/peso corporal, o 1 microgramo/kg/peso corporal, desde 5
10 microgramo/kg/peso corporal, 10 microgramo/kg/peso corporal, 50 microgramo/kg/peso corporal, 100 microgramo/kg/peso corporal, 200 microgramo/kg/peso corporal, 350 microgramo/kg/peso corporal, 500 microgramo/kg/peso corporal, 1 miligramo/kg/peso corporal, 5 miligramo/kg/peso corporal, 10 miligramo/kg/peso corporal, 50 miligramo/kg/peso corporal, 100 miligramo/kg/peso corporal, 200 miligramo/kg/peso corporal, 350 miligramo/kg/peso corporal, o 500 miligramo/kg/peso corporal, a 1000 miligramo/kg/peso corporal o más por
15 administración, y cualquier gama derivable de esta. En ejemplos no limitativos de una gama derivable de las cifras enumeradas aquí, una gama de 5 miligramo/kg/peso corporal a 100 miligramo/kg/peso corporal, 5 microgramo/kg/peso corporal a 500 miligramo/kg/peso corporal, etc., se puede administrar, basada en las cifras anteriormente descritas.

20 [0157] En cualquier caso, la composición puede comprender diferentes antioxidantes para retardar la oxidación de uno o más componentes. Adicionalmente, la prevención de la acción de microorganismos puede ser provocada por conservantes tales como varios agentes antibacterianos y antifúngicos, incluyendo pero no limitados a parabenos (por ejemplo, metilparabenos, propilparabenos), clorobutanol, fenol, ácido sórbico, timerosal o combinaciones de los mismos.

25 [0158] Las moléculas se pueden formular en una composición en forma de base libre, neutral o sal. Sales farmacéuticamente aceptables incluyen las sales de adición de ácido, por ejemplo, las formadas con los grupos amino libres de una composición proteínica, o que se forman con ácidos inorgánicos tales como por ejemplo, ácidos clorhídricos o fosfóricos, o tales ácidos orgánicos como ácido acético, oxálico, tartárico o mandélico.
30 Sales formadas con los grupos carboxilo libres también pueden derivar de bases inorgánicas tales como por ejemplo, sodio, potasio, amonio, calcio o hidróxidos férricos; o tales bases orgánicas como isopropilamina, trimetilamina, histidina o procaina.

35 [0159] En formas de realización donde la composición está en una forma líquida, un portador puede ser un solvente o medio de dispersión que comprende pero no está limitado a, agua, etanol, poliol (por ejemplo, glicerol, propilenglicol, polietilenglicol líquido, etc.), lípidos (por ejemplo, triglicéridos, aceites vegetales, liposomas) y combinaciones de los mismos. La fluidez apropiada se puede mantener, por ejemplo, usando un recubrimiento, tal como lecitina; manteniendo el tamaño de partícula requerido mediante la dispersión en portadores tales como, por ejemplo poliol líquido o lípidos; por el uso de surfactantes tales como, por ejemplo hidroxipropilcelulosa; o
40 combinaciones de los mismos de tales métodos. En muchos casos, será preferible incluir agentes isotónicos, tales como, por ejemplo, azúcares, cloruro sódico o combinaciones de los mismos.

[0160] En otras formas de realización, se pueden usar colirios, soluciones o esprays nasales, aerosoles o inhalantes en la presente invención. Tales composiciones se diseñan generalmente para ser compatibles con el
45 tipo de tejido objetivo. En un ejemplo no limitativo, las soluciones nasales son normalmente soluciones acuosas diseñadas para ser administradas a los orificios nasales en gotas o esprays. Las soluciones nasales se preparan de modo que son similares en muchos aspectos a las secreciones nasales, de modo que la acción ciliar normal se mantiene. Así, en formas de realización preferidas las soluciones nasales acuosas normalmente son isotónicas o ligeramente tamponadas para mantener un pH de aproximadamente 5,5 a aproximadamente 6,5.
50 Además, conservantes antimicrobianos, similar a los usados en preparaciones oftálmicas, fármacos o estabilizadores de medicamentos apropiados, si es necesario, se pueden incluir en la formulación. Por ejemplo, varias preparaciones nasales comerciales se conocen e incluyen fármacos tales como antibióticos o antihistamínicos. En formas de realización determinadas, las moléculas se preparan para la administración por vías tales como la ingestión oral. En estas formas de realización, la composición sólida puede comprender, por
55 ejemplo, soluciones, suspensiones, emulsiones, comprimidos, píldoras, cápsulas (por ejemplo, cápsulas duras o blandas con carcasa de gelatina), formulaciones de liberación prolongada, composiciones bucales, pastillas, elixires, suspensiones, jarabes, obleas o combinaciones de los mismos. Las composiciones orales se pueden incorporar directamente con el alimento de la dieta. Portadores preferidos para administración oral comprenden diluyentes inertes, portadores comestibles asimilables o combinaciones de los mismos. En otros aspectos de la invención, la composición oral se puede preparar como un jarabe o elixir. Un jarabe o elixir, y puede comprender,
60 por ejemplo, al menos un agente activo, un agente edulcorante, un conservante, un agente aromatizante, un tinte, un conservante o combinaciones de los mismos.

65 [0161] En determinadas formas de realización preferidas una composición oral puede comprender uno o más ligantes, excipientes, agentes de desintegración, lubricantes, agentes aromatizantes y combinaciones de los mismos. En formas de realización determinadas, una composición puede comprender uno o más de los

siguientes: un ligante, tal como, por ejemplo, goma tragacanto, acacia, almidón de maíz, gelatina o combinaciones de los mismos; un excipiente, tal como, por ejemplo, fosfato dicálcico, manitol, lactosa, almidón, estearato de magnesio, sacarina de sodio, celulosa, carbonato de magnesio o combinaciones del mismo; un agente de desintegración, tal como, por ejemplo, almidón de maíz, almidón de patata, ácido algínico o combinaciones del mismo; un lubricante, tal como, por ejemplo, estearato de magnesio; un agente edulcorante, tal como, por ejemplo, sacarosa, lactosa, sacarina o combinaciones de las mismas; un agente aromatizante, tal como, por ejemplo menta, aceite de gaulteria, aromatizante de cereza, aromatizante de naranja, etc. o combinaciones de los anteriormente mencionados. Cuando la forma unitaria de dosificación es una cápsula, puede contener, además de materiales del tipo anterior, portadores tales como un portador líquido. Varios materiales más pueden estar presentes como recubrimientos o par modificar de otro modo la forma física de la unidad de dosificación. Por ejemplo, comprimidos, píldoras o cápsulas pueden estar recubiertos con goma laca, azúcar o ambas.

[0162] La composición debe ser estable bajo las condiciones de producción y almacenamiento, y conservada contra la acción contaminante de microorganismos, tales como bacterias y hongos. Se apreciará que la contaminación de endotoxina se mantenga mínimamente a un nivel seguro, por ejemplo, menos de 0,5 ng/mg proteína.

[0163] En formas de realización particulares, la absorción prolongada de una composición inyectable se puede provocar por el uso en las composiciones de agentes que retardan la absorción, tal como, por ejemplo, monoestearato de aluminio, gelatina o combinaciones de los mismos.

[0164] Cualquier forma de realización mencionada anteriormente con respecto a la entrega o transporte a las células se puede emplear también con respecto a la implementación de la entrega de compuestos medicinales discutidos en esta sección.

Dosificaciones eficaces

[0165] Las moléculas de la invención se usarán generalmente en una cantidad eficaz para conseguir el fin deseado. Para el uso para tratar o prevenir una condición de enfermedad, las moléculas de la invención, o composiciones farmacéuticas de las mismas, se administran o aplican en una cantidad terapéuticamente eficaz. Una cantidad terapéuticamente eficaz es una cantidad eficaz para mejorar o prevenir los síntomas, o prolongar la supervivencia del paciente que se va a tratar. La determinación de una cantidad terapéuticamente eficaz está bien en las capacidades de los expertos en la técnica, especialmente en vista de la divulgación detallada proporcionada aquí.

[0166] Para la administración sistémica, una dosis terapéuticamente efectiva se puede estimar inicialmente a partir de ensayos *in vitro*. Por ejemplo, una dosis se puede formular en modelos animales para conseguir una gama de concentración circulante que incluye el EC50 como se determina en el cultivo celular. Tal información puede utilizarse para determinar con más con precisión las dosis útiles en seres humanos.

[0167] Las dosificaciones iniciales también se pueden estimar a partir de datos *in vivo*, por ejemplo, modelos animales, usando técnicas que se conocen en la técnica. Un experto en la materia podrían fácilmente optimizar la administración a seres humanos basado en datos de animales.

[0168] La cantidad de dosificación y el intervalo se puede ajustar individualmente para proporcionar niveles de plasma de las moléculas que son suficientes para mantener el efecto terapéutico. Las dosificaciones de pacientes usuales para la administración por inyección varía de 0,01 a 0,1 mg/kg/día, o de 0,1 a 5 mg/kg/día, preferiblemente de 0,5 a 1 mg/kg/día o más. Niveles de suero eficaces terapéuticamente se pueden conseguir por la administración de dosis múltiples cada día.

[0169] En casos de administración local o absorción selectiva, la concentración local eficaz de las proteínas pueden no estar relacionada con la concentración de plasma. Alguien de habilidad en la técnica será capaz de optimizar las dosificaciones locales eficaces terapéuticamente sin experimentación excesiva.

[0170] La cantidad de moléculas administradas dependerá, por supuesto, del sujeto que se va a tratar, del peso del sujeto, de la gravedad de la afección, de la forma de administración y de la decisión del médico que prescribe.

[0171] La terapia se puede repetir intermitentemente cuando los síntomas son detectables o incluso cuando no son detectables. La terapia se puede proporcionar sola o en combinación con otros fármacos o tratamientos (incluyendo cirugía).

Toxicidad

[0172] Preferiblemente, una dosis terapéuticamente efectiva de las moléculas descritas aquí proporcionará beneficio terapéutico sin causar toxicidad sustancial. La toxicidad de las moléculas descritas aquí se pueden determinar por procedimientos farmacéuticos estándar en cultivos de células o animales experimentales, por ejemplo, determinando la LD50 (la dosis letal para el 50% de la población) o la LD100 (la dosis letal para el 100% de la población). La proporción de dosis entre el efecto tóxico y terapéutico es el índice terapéutico. Se prefieren las proteínas que muestran altos índices terapéuticos. Los datos obtenidos de estos ensayos de cultivo celular y estudios de animales se pueden usar en la formulación de una gama de dosificaciones que no es tóxica para su uso en seres humanos. La dosificación de las proteínas descritas aquí se extiende preferiblemente dentro de una gama de concentraciones circulantes que incluye la dosis efectiva con poca o ninguna toxicidad. La dosificación puede variar dentro de esta gama dependiendo de la forma de dosificación empleada y la forma de administración utilizada. La formulación exacta, la forma de administración y la dosificación pueden ser elegidas por el médico particular en vista de la condición del paciente. (Véase, por ejemplo, Fingl *et al.*, 1975, In: The Pharmacological Basis of Therapeutics, Ch.1, p.1).

15 Grupos colgantes

[0173] Un "grupo colgante" se puede unir o conjugar al ácido nucleico. Los grupos colgantes pueden aumentar la absorción celular del ácido nucleico. Los grupos colgantes se pueden enlazar a cualquier porción del ácido nucleico pero están comúnmente enlazados al extremo(s) de la cadena de oligonucleótidos. Ejemplos de grupos colgantes incluyen, pero de forma no limitativa: derivados de acridina (es decir, 2-metoxi-6-cloro-9-amoacridina); reticuladores tales como los derivados de psoraleno, azidofenacilo, proflavina y azidoproflavina; endonucleasas artificiales; complejos metálicos tales como EDTA-Fe(II), o-fenantrolina-Cu(I), y porfirina-Fe(II); fracciones alquilantes; nucleasas tales como amino-1-hexanolstafilococal nucleasa y fosfatasa alcalina; transferasas terminales; abzymas; fracciones de colesterol; portadores lipofílicos; conjugados de péptido; alcoholes de cadena larga; ésteres de fosfato; amino; grupos de mercapto; marcadores radiactivos; marcadores no radiactivos tales como tintes; y polilisina u otras poliaminas. En un ejemplo, el ácido nucleico se conjuga con un carbohidrato, carbohidrato sulfatado o glicano.

30 Kits

[0174] Cualquiera de las composiciones descritas aquí puede estar comprendida en un kit. En un ejemplo no limitativo, los ARNmi individuales están incluidos en un kit. El kit puede incluir además uno o más ARNmi sintéticos de control negativo que pueden utilizarse para controlar los efectos de la entrega de ARNmi sintéticos. El kit puede incluir además agua y tampón de hibridación para facilitar la hibridación de las dos cadenas de los ARNmi sintéticos. El kit también puede incluir uno o más reactivo(s) de transfección para facilitar la entrega del ARNmi a las células.

[0175] En otro ejemplo no limitativo, los ARNmi sintéticos múltiples están incluidos en un kit. El kit puede incluir además uno o más ARNmi sintéticos de control negativo que pueden utilizarse para controlar los efectos de la entrega de los ARNmi sintéticos. El kit también puede incluir uno o más reactivos de transfección para facilitar la entrega en células.

[0176] Los componentes de los kits se pueden empaquetar bien en medios acuosos o en forma liofilizada. Los medios contenedores de los kits incluirán generalmente por lo menos un vial, probeta, matraz, botella, jeringa u otros medios contenedores, en los que un componente se puede colocar, y preferiblemente, dividir en partes alícuotas adecuadamente. Donde hay más de un componente en el kit (el reactivo de etiquetado y el marcador se pueden empaquetar juntos), el kit también generalmente contendrá un segundo, tercero u otro contenedor adicional en el que los componentes adicionales se pueden colocar separadamente. Sin embargo, varias combinaciones de componentes pueden estar comprendidas en un frasco. Los kits de la presente invención también típicamente incluirán unos medios para contener los ácidos nucleicos, y cualquiera de los otros contenedores reactivos confinamiento estrecho para la venta comercial. Tales contenedores pueden incluir contenedores de plástico moldeados por inyección o soplado en los que los viales deseados se retienen.

[0177] Cuando los componentes del kit están provistos en una y/o más soluciones líquidas, la solución líquida es una solución acuosa, con una solución acuosa estéril preferida particularmente.

[0178] Sin embargo, los componentes del kit pueden estar provistos como polvo seco. Cuando los reactivos y/o componentes están provistos como un polvo seco, el polvo se puede reconstituir añadiendo un solvente adecuado. Se prevé que el solvente también se pueda proporcionar en otros medios contenedores.

[0179] Los medios contenedores incluirán generalmente por lo menos un frasco, probeta, matraz, botella, jeringa y/o otros medios contenedores, en los que las formulaciones de ácido nucleico se colocan, preferiblemente, situadas adecuadamente. Los kits también pueden comprender unos segundos medios contenedores para contener un tampón estéril farmacéuticamente aceptable y/o otro diluyente. Los kits de la presente invención también incluirán típicamente medios para contener los viales en confinamiento estrecho para la venta comercial,

tal como, por ejemplo, contenedores plásticos moldeados por inyección y/o soplado en los que los viales deseados se retienen.

5 [0180] Tales kits también pueden incluir componentes que preservan o mantienen el ARNm o que protegen contra su degradación. Tales componentes pueden ser libres de ARNasa o proteger contra las ARNasas. Tales kits generalmente comprenderán, en medios adecuados, contenedores diferentes para cada reactivo o solución individual.

10 [0181] Un kit también incluirá instrucciones para utilizar los componentes del equipo, al igual que el uso de cualquier otro reactivo no incluido en el kit. Las instrucciones pueden incluir variaciones que se pueden implementar.

15 [0182] Los kits de la invención también pueden incluir uno o más de los siguientes: ARNm, biblioteca de ARNm, biblioteca de combinación de ARNm, ARNm de control negativo, agua libre de nucleasa; contenedores libres de ribonucleasa, tales como tubos de 1,5 ml; tampón de hibridación; y reactivo(s) de transfección.

20 [0183] Se contempla que tales reactivos sean las formas de realización de los kits de la invención. Tales kits, sin embargo, no se limitan a los artículos particulares identificados por encima y pueden incluir cualquier reactivo usado para la manipulación o caracterización de ARNm.

Identidad de secuencia

25 [0184] La "identidad de secuencia" se defiende aquí como una relación entre dos o más secuencias de ácido nucleico (nucleótido, polinucleótido, ARN, ADN), como se determina por comparación de las secuencias. En la técnica, "identidad" también significa el grado de relación de secuencia entre secuencias de ácidos nucleicos, según sea el caso, como se determina por la correspondencia entre cadenas de tales secuencias.

30 [0185] En una forma de realización preferida, la identidad también significa porcentaje de identidad y se puede calcular por el número de nucleótidos iguales entre base de datos y consulta, dividido por la longitud total de la consulta, y multiplicado por 100.

35 [0186] La "identidad" y la "similitud" se pueden calcular fácilmente por métodos conocidos, incluyendo pero no limitado a, los descritos en Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heine, G., Academic Press, 1987; y Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; y Carillo, H., y Lipman, D., SIAM J. Applied Math., 48:1073 (1988).

40 [0187] Métodos preferidos para determinar la identidad se diseñan para dar la mayor correspondencia entre las secuencias evaluadas. Métodos para determinar la identidad y la similitud se codifican en programas informáticos disponibles públicamente. Métodos de programas informáticos preferidos para determinar la identidad y la similitud entre dos secuencias incluyen por ejemplo el paquete del programa GCG (Devereux, J., *et al.*, Nucleic Acids Research 12 (1): 387 (1984)), BestFit, BLASTP, BLASTN, y FASTA (Altschul, S. F. *et al.*, J. Mol. Biol. 215:403-410 (1990)). La programa BLAST X está disponible públicamente de NCBI y otras fuentes (BLAST Manual, Altschul, S., *et al.*, NCBI NLM NIH Bethesda, MD 20894; Altschul, S., *et al.*, J. Mol. Biol. 215:403-410 (1990)). El algoritmo de Smith Waterman bien conocido también se puede usar para determinar la identidad.

50 [0188] Parámetros preferidos para comparación de ácidos nucleicos incluyen el siguiente: algoritmo: Needleman y Wunsch, J. Mol. Biol. 48:443-453 (1970); Matriz de comparación: coincidencias=+10, falta de coincidencia=0; Penalización por espacio: 50; Penalización por longitud de espacio: 3. Disponible como el programa Gap de Genetics Computer Group, situado en Madison, Wis. Anteriormente se han dado los parámetros por defecto para comparaciones de ácidos nucleicos.

55 [0189] En este documento y en sus reivindicaciones, el verbo "comprender" y sus conjugaciones se usan en su sentido no limitativo para referirse a que los artículos que van después del verbo están incluidos, pero artículos que no se mencionan específicamente no están excluidos. Además el verbo "consistir" se puede sustituir por "consistir esencialmente en", que significa que una molécula de ARNm, un equivalente o una fuente de la misma o una composición tal y como se define aquí puede comprender componente(s) adicional(es) a los identificados específicamente, dicho(s) componente(s) adicional(es) no alteran la característica única de la invención. Además, la referencia a un elemento por el artículo indefinido "un" o "una" no excluye la posibilidad que más de uno del elemento esté presente, a menos que el contexto requiera claramente que haya uno y solo uno de los elementos. El artículo indefinido "un" o "una" normalmente significa "al menos uno". Los ejemplos siguientes se ofrecen para uso ilustrativo solo, y no están destinados a limitar el alcance de la presente invención de ninguna manera.

Descripción de las figuras

[0190]

5

Figura 1. Diagrama de flujo del protocolo de selección para la inversión de EMT, y visión de conjunto de los resultados obtenidos de una de las placas de la biblioteca de expresión de microARN basados en Lentivirus (ITM0081-0160). A, se añadió caldo de Lentivirus (1,0 ul; MOI=3 a 300) a células TSUp1-pEcad-Luc durante 24 horas. Dos días post-infección (día 4), se aplicó selección de puromicina, y seis días post-infección (día 8) se midió la actividad de luciferasa. B, las proporciones de luciferasa de *luciérnaga/Renilla* corregida en la base se pusieron en placas contra MOI lentiviral. Los cuatro miRs con señales significativamente por encima de la base (proporción FLuc/RLuc > media + 2 x desviación típica) están rodeados. Los cuatro valores de control representan proporciones FLuc/RLuc inducidas por miR-141 y miR-200c, que fueron suministrados en tubos separados (ambos por duplicado).

10

15

Figura 2: visión de conjunto del vector de expresión pEcad-Luc/RLuc. El promotor del gen de *CDH1* (GenBank L34545, posición -294 a +44), que contiene las tres E-cajas, se obtuvo por amplificación por PCR usando ADN genómico humano como un modelo. El fragmento promotor de *CDH1* fue clonado arriba del gen de luciferasa de *luciérnaga* en el vector básico pGL2 (Promega). El casete de luciferasa de *Renilla* dirigido por el promotor HSV-Tk se obtuvo a partir del vector pRL-TK (Promega), y el casete de resistencia a los antibióticos de zeocina dirigido por el promotor SV40 del vector pVgRXR (Invitrogen). La luciferasa de *Renilla* y los cassetes de ZeocinR fueron clonados en el vector básico pGL2 en las ubicaciones indicadas.

20

25

Figura 3: caracterización de modelos de línea celular tumoral. El ARN total fue aislado de cultivos de línea celular confluyente al 70-80%, y analizado para *CDH1* y expresión de vimentina por qPCR. Los valores de expresión relativos se representaron gráficamente, y las células con alto *CDH1*/bajo *VIM* y bajo *CDH1*/alto *VIM* fueron designadas epitelial y mesenquimal, respectivamente.

30

35

Figura 4: inducción de MET usando miméticos de ARNmi sintéticos. A, células TSUp1-pEcad-Luc se transfectaron transitoriamente con 20 a 60 nM de miméticos de miR sintéticos (Ambion o Dharmacon; véase tabla 7 para detalles). Cuatro días post-transfección, la actividad de Luc y RLuc se midió, y las proporciones Luc/RLuc se normalizaron para células transfectadas de control negativo (NC). B, las células fueron transfectadas y tratadas como se describe en la figura 4A. El ARN total fue aislado y usado para análisis de ARNmi y qPCR génica. Los niveles de expresión de *CDH1* y *CDH2*, normalizados para el gen de mantenimiento *HPRT*, como porcentaje de células transfectadas NC se muestran. C, diferentes líneas celulares tumorales parentales se transfectaron con miméticos de miR-200c y miR-520f y otros miméticos de miR y se analizaron para expresión de *CDH1* (como se describe en la Fig. 4B).

40

45

50

Figura 5: creación de sistemas de expresión de ARNmi inducibles por doxiciclina. A, los precursores de miR-200c y miR-520f fueron aislados del provirus integrado en células TSUp1 infectadas con lentivirus pCDH (en los experimentos de selección de miR) por digestión enzimática de restricción. Los precursores de miR fueron clonados en los sitios *NheI*/*EagI* del vector de expresión de miR inducible por pmRi-ZsGreen1 (Clontech PT5049-1). El mapa del vector de expresión miR-520f se muestra. B, las células PC3 (ATCC# CRL-1435) fueron transfectadas con el vector avanzado pTet-on (Clontech PT3899-5), y seleccionadas en el medio que contiene G418 (300ug/ml). Las células fueron transitoriamente transfectadas con el vector indicador pTRE-Luc, tratadas con 0, 0,1 y 1,0 ug/ml doxiciclina (DOX, Sigma D9891) durante 2 días, y luego analizadas para activación de Luc. C, las células PC3-Tet-on (clon 8) fueron luego transfectadas con el vector pmRi-ZsGreen1-miR-200c o miR-520f, y seleccionadas en el medio que contiene puromicina (5ug/ml). Los clones estables de PC3-mRi-ZsGreen-miR-X fueron evaluados para la inducción de ARNmi y expresión génica. Las células fueron tratadas con 0 y 1,0 ug/ml DOX durante 4 días, ARN total fue aislado, y ARNmi y expresión génica (tabla 8) fueron analizados por qPCR.

55

60

65

Figura 6: influencia de ARNmi en la invasión de células tumorales. A, cincuenta mil (PC3) o cuarenta mil (TSUp1) células fueron sembradas en cámaras de invasión Biocoat Matrigel, 8 micras (BD 354480), en medio sin suero. La cámara de invasión fue colocada en 24-pocillos con medio con 10% de suero fetal de ternera como quimioatrayente. Como un control, la misma cantidad de células fue sembrada sobre la superficie de una placa de cultivo de 24 pocillos. Después de 48 horas de incubación, las células en de cámara de invasión fueron retiradas por aspiración y limpieza del compartimento interno con un hisopo. La cámara de invasión se puso luego en reactivo de viabilidad celular CellTiter-GLO (CTG, Promega-G7571) e incubado durante 15 minutos. La actividad de CTG se midió en un luminómetro Victor3. B, ensayos de invasión celular fueron realizadas con células PC3-mRi-ZsGreen1-miR-X que fueron pretratadas durante 2 días con 1µg/ml DOX. Además, las células PC3 infectadas con lentivirus miR-200c y miR-520f (MOI=30), y seleccionadas en la puromicina durante 4 días, fueron usadas. El porcentaje de invasión celular fue calculado como la actividad de CTG en la parte inferior de la membrana dividida por el total de actividad de CTG (de las células crecidas en la superficie de una placa de cultivo de 24 pocillos). La inhibición de invasión celular por un ARNmi específico fue calculada por la división del porcentaje de invasión celular contra células no tratadas (-DOX) o de control (virus de vector vacío). C, los

niveles de expresión de microARN relativos en cultivos paralelos se determinaron por RT-qPCR de tallo-bucle. D, los ensayos de invasión celular fueron realizadas con células TSUpr-pEcad que fueron infectadas con lentivirus que contiene precursor miR-124-1, miR-181a-1, miR-200c, miR-206, miR-518b, miR-520f o miR-524 de vector vacío (MOI=30), y seleccionados en puomicina durante 4 días. La invasión celular fue calculada como se describe en la Fig. 6B.

Ejemplos

Ejemplo 1

Material y métodos

Generación de la biblioteca lentiviral que codifica ARNmi

constituyente	concentración	volumen	proveedor / cat. #
tampón	10X	1 µL	Stratagene / 600159
dNTPs	10 mM cada	0,2 µL	GE Healthcare / 27-18(5-8)0-04
cebador sentido	10 µM	0,2 µL	IDT (Integrated DNA Technologies)
cebador antisentido	10 µM	0,2 µL	" "
gADN	100 ng/µL	0,1 µL	fuentes privadas
Pfu ADN pol	2.5 U/µL	0,1 µL	Stratagene / 600159
H ₂ O	N/A	8,2 µL	N/A

temp. (°C)	tiempo	ciclos	
95	2 min		
95	15 s	40	
59*	15 s	40	* -0,1 °C / ciclo
72	90 s	40	
72	15 min		
4			∞

[0191] Los ARNmi humanos fueron seleccionados tanto del repositorio de ARNmi público (www.mirbase.org) como de los datos de secuenciación profunda de ARN pequeños propietarios (véase WO 2007/081204). Las secuencias de ARNmi fueron amplificadas desde su ubicación genómica con amplicones que contenían la horquilla de pre-ARNmi en toda su longitud y una secuencia flanqueante en ambos lados de 50-150 pares de bases. Los cebadores para los amplicones fueron diseñados utilizando el software Primer3 (www.geneious.com). Si el programa de diseño de cebador no pudo encontrar cebadores apropiados en las secuencias designadas, los requisitos para las secuencias flanqueantes se ajustaron a 0-200 pares de bases. Los cebadores diseñados fueron complementados con un saliente 5' GCGC y un sitio de restricción para la clonación direccional. Como predeterminado el cebador arriba del ARNmi fue complementado con un sitio de restricción BamHI (GGATCC) y el cebador abajo del ARNmi fue complementado con un sitio de restricción EcoRI (GAATTC). Los cebadores de los amplicones con sitios de restricción BamHI o EcoRI internos (es decir, que se producen en la secuencia genómica) fueron complementados con bien un sitio BglII (AGATCT) o un sitio XbaI (TCTAGA) respectivamente. Los ARNmi fueron amplificados utilizando los cebadores anteriormente mencionados a partir de ADN genómico humano de un individuo individual en la siguiente reacción por PCR:

[0192] Todos los loci de ARNmi fueron amplificados en reacciones separadas de PCR de 10 µL. Los productos fueron purificados utilizando el equipo de tampón Qiagen PCR Clean-Up y las placas de filtro Whatman Unifilter GF/C (Cat. # 7700-1101). El ADN fue eluido con 17 µL H₂O por pocillo. Los eluatos separados fueron usados en la siguiente reacción de restricción:

ES 2 631 458 T3

constituyente	concentración	volumen	proveedor / cat. #
tampón E	10X	2 μ L	Promega / R005A
EcoRI*	12 U/ μ L	0,1 μ L	Promega / R6017
BamHI*	10 U/ μ L	0,1 μ L	Promega / R6025
eluato	N/A	16 μ L	N/A
H ₂ O	N/A	1,8 μ L	N/A

*Amplicones con sitios de restricción internos para EcoRI o BamHI fueron cortados con XbaI o BglII respectivamente en su lugar. La reacción EcoRI+BglII se hizo con tampón Promega D. La reacción BamHI+XbaI se hizo con tampón Promega E.

5 [0193] Restricción durante 2 horas a 37 °C. Las reacciones de restricción separadas de 20 μ L fueron purificadas utilizando el conjunto de tampón PCR Qiagen Clean-Up y las placas de filtro Whatman Unifilter GF/C (Cat. # 7700-1101). El ADN fue eluido con 20 μ L H₂O por pocillo. Los eluatos separados fueron usados en la siguiente reacción de ligamiento:

constituyente	concentración	volumen	proveedor / cat. #
tampón	10X	2 μ L	Promega / C1263
T4 ADN ligasa	1-3 U/ μ L	0,2 μ L	Promega / M1804
pCDH* restringido	1 ng/ μ L	7,8 μ L	System Biosciences / CD510B-1
eluato	N/A	10 μ L	N/A

Ligamiento durante toda la noche a 4 °C.
*Para clonación direccional, pCDH se cortó tanto con EcoRI como con BamHI. Un constructo alterno llamado pCDH- se hizo con sitios de restricción de EcoRI y BamHI invertidos de modo que los amplicones con 5' BamHI y 3' EcoRI fueron clonados en la dirección apropiada. Los amplicones con un sitio de EcoRI interno se cortaron con XbaI y se ligaron en un vector pCDH que fue restringido con XbaI y BamHI.

10 [0194] Los ligados resultantes fueron transformados separadamente en bacterias (células competentes de etapa única Promega (KRX), cat. # L3002). 50 μ L de células competentes se diluyeron con 950 μ L de tampón de transformación II (10 mM MOPS, 75 mM CaCl₂, 10 mM RbCl, 15% glicerol, esterilizado por filtro). Por 20 μ L de ligado 20 μ L de células competentes diluidas se añadieron. La mezcla fue incubada durante 15 minutos en hielo, se sometió a choque térmico a 37 °C durante 30 segundos, y se volvió a poner en el hielo. Después de 2 minutos las bacterias transformadas fueron reconstituidas en 150 μ L de caldo de lisogenia (LB). Las bacterias se dejaron recuperar durante 20 minutos a 37 °C después de lo cual se colocaron en placas separadamente en placas de LB-agar conteniendo ampicilina (50 μ g/mL) y se dejaron crecer durante toda la noche a 37 °C.

20 [0195] Colonias individuales de cada placa se escogieron y subcultivaron durante toda la noche en 400 μ L de LB conteniendo ampicilina (50 μ g/mL). 1 μ L del subcultivo se lisó en 100 μ L de agua para fines de secuenciación. El lisado bacteriano se usó en la siguiente reacción por PCR:

constituyente	concentración	volumen	proveedor / cat. #
tampón	5X	1 μ L	fuelle privada
dNTPs	10 mM cada	0,1 μ L	GE Healthcare / 27-18(5-8)0-04
pCDH-fwd	10 μ M	0,1 μ L	IDT (Integrated DNA Technologies)
pCDH-rev	10 μ M	0,1 μ L	" "
lisato	1:100	1 μ L	N/A
Taq ADN pol	desconocido	0,02 μ L	fuelle privada

ES 2 631 458 T3

constituyente	concentración	volumen	proveedor / cat. #
H ₂ O	N/A	2,68 µL	N/A

temp. (°C)	tiempo		ciclos
95	2 min		
95	15 s		40
59*	15 s	40	* -0,1 °C / ciclo
72	90 s	40	
72	15 min		
4	∞		

pCDH-sentido CACGCTGTTTTGACCTCCATAGA
pCDH-antisentido CACTGACGGGCACCGGAG
(SEQ ID NO: 84-85)

5

[0196] Los productos de PCR fueron diluidos 25X. 1 µL del producto de PCR diluido se usó en la siguiente reacción de secuenciación de Sanger:

constituyente	concentración	volumen	proveedor / cat. #
tampón	N/A	1,9 µL	fuelle privada
BigDye v3.1	N/A	0,1 µL	ABI / 4336921
pCDH-seq	10 µM	0,1 µL	IDT (Integrated DNA Technologies)
producto de PCR	1:25	1 µL	N/A
H ₂ O	N/A	1,9 µL	N/A

10

temp. (°C)	tiempo	ciclos
94	10 seg.	
50	5 s	40
60	2 min	40
10	∞	

15 pCDH-seq GACCTCCATAGAAGATTCTAGAGCTAGC
(SEQ ID NO: 86)

[0197] 30 u de mezcla de precipitación (80% de etanol, 50 mM de acetato sódico pH 5,5) se añadió a cada uno de los productos de reacción de secuenciación. Las mezclas fueron agitadas en vórtex durante 10 segundos y centrifugadas a 5000 rcf durante 45 minutos a 4 °C. El sobrenadante fue aspirado y los gránulos de ADN fueron lavados con 30 µL de etanol al 80% helado y centrifugados a 5000 rcf durante 5 minutos a 4 °C. El sobrenadante fue aspirado y el granulado de ADN fue secado en un bloque de calor durante 10 minutos. El granulado de ADN seco fue disuelto en 10 µL H₂O. La solución de ADN resultante fue secuenciada en un analizador de ADN ABI 3730XL. Las secuencias fueron comparadas con las secuencias genómicas previstas. Los clones correctos fueron añadidos a la biblioteca. Para los clones incorrectos 4 colonias bacterianas adicionales fueron escogidas, y analizadas para secuencia de inserto.

25

[0198] Los constructos de la biblioteca fueron subcultivados durante toda la noche en 50 mL de LB conteniendo ampicilina (100 µg/mL) y aislados con el kit Qiagen QIAfilter Plasmid Midi (Cat. # 12245) suplementado con el set Qiagen EndoFree Plasmid Buffer (Cat. # 19048) según las instrucciones del fabricante. El ADN fue disuelto en el tampón de TE suministrado y llevado a una concentración final de 500 ng/µL.

5

[0199] Ordenamos los constructos que no fuimos capaces de clonar nosotros mismos como minigenes de Integrated DNA Technologies. En estos casos, la horquilla en toda su longitud más 20 pares de bases flanqueando cada sitio se clonaron en nuestro vector como un servicio por el IDT.

10

[0200] El embalaje y la producción viral fue realizado por System Biosciences como se describe en el manual del usuario de CD-500B1-CD523-A1.

Cultivo celular

15

[0201] La línea celular TSUpr1/pEcad-luc/Rluc fue generada por transfección estable de la línea celular de carcinoma de vejiga de célula transicional humana TSUpr1 con el vector de expresión pEcad-Luc/Rluc (figura 2). Un clon único, resistente a la zeocina, (clon 1.c.4) se usó para todos los experimentos.

20

[0202] Las células TSUpr1/pEcad-luc/Rluc se mantuvieron en el medio RPMI-1640 (Invitrogen, 31870), suplementadas con 10% de suero bovino fetal (Sigma, F7524), L-Glutamina (Invitrogen 25030-024) y 50µg/ml de zeocina (Invitrogen, R250-01). Las células se mantuvieron en una atmósfera humedecida a 37°C y 5% CO₂. Las células se dividieron una vez a la semana en una proporción de 1:20.

Productos químicos

25

[0203] Polibreno (2µg/ml; Sigma, H9268) se usó para aumentar la eficiencia de infección con las partículas lentivirales codificantes de ARNm. Puomicina (5µg/ml; Sigma, P8833) se usó para seleccionar células que expresan el ARNm de interés. Zeocina (50µg/ml; Invitrogen, R250-01) se usó para mantener la integración y la expresión del transgén Ecad-luc/Rluc.

30

Protocolo de selección de MET:

Día 1: siembra de células

35

[0204] Células TSUpr1/pEcad-luc/Rluc fueron sembradas en una densidad de 2.500 células por pocillo en placas de 96 pocillos (100 µl volumen total por pocillo), por duplicado.

Día 2: infección lentiviral

40

[0205] Los constructos lentivirales empaquetados se proporcionaron como partículas virales pseudotipadas congeladas VSV-G, y se almacenaron a -80°C. Antes del uso, los contenidos fueron descongelados a temperatura ambiente y puestos en hielo inmediatamente después. Para abrir los tubos se utilizó una herramienta de apertura de tapa SepraSeal (Fisher Scientific Cat. #: 50823908) y los precipitados fueron resuspendidos por pipeteo en varias veces. El medio se retiró utilizando una pipeta multicanal. Suavemente, 200µl de medio fresco (+FBS/glutamina/zeocina) conteniendo 2 µg/ml de polibreno (Sigma, H9268) se añadió a las células. Después, 1,0 µl de partículas lentivirales no diluidas se añadió a cada pocillo (por duplicado). En cada placa de 96 pocillos, se añadieron partículas lentivirales codificantes de miR-141 y miR-200c como controles positivos.

50

Día 3: medio fresco

[0206] Veinticuatro horas después de la adición de lentivirus, el medio se retiró utilizando una pipeta multicanal. Las células fueron lavadas una vez con 0,9% NaCl, 50µl por pocillo. Posteriormente, 100µl (+FBS/glutamina/zeocina) de medio fresco se añadió a cada pocillo.

55

Día 4: selección de puomicina

[0207] El medio se retiró utilizando una pipeta multicanal. Para seleccionar células transducidas, 200µl de medio fresco conteniendo 5 µg/ml de puomicina (Sigma, P8833) se añadió a las células. Nota, las células no se lavaron entremedias.

60

Día 8: ensayo de indicador de luciferasa dual (Promega)

65

[0208] El medio fue retirado utilizando una pipeta multicanal. Después, las células fueron lavadas suavemente con 0,9% de NaCl (50 µl/pocillo). Para preparar lisatos celulares, tampón de lisis pasiva 1x (Promega, E1980) se añadió a las células (20µl por pocillo), y se incubó a temperatura ambiente durante 20 minutos en una coctelera

de placa. Los lisatos celulares fueron transferidos a un placas de microtitulación de 96 pocillos blancas (Thermo Scientific, 9502887). La actividad de luciferasa de *Renilla* y *luciférnaga* se midió en un contador Victor3 Multilabel (PerkinElmer), según las instrucciones del fabricante.

5 Aislamiento de ARN total

[0209] Células TSUpr1/pEcad-luc/Rluc se sembraron en una densidad de 15.000 células por pocillo en placas de 24 pocillos. La infección lentiviral fue realizada como se ha descrito anteriormente. Debido a la cantidad limitada de partículas virales, se añadió virus a una multiplicidad de infección (MOI) de 30, y si fue posible una MOI de 100. En el día 8, el ARN total fue aislado utilizando el reactivo de Trizol (200µl por pocillo), según las instrucciones del fabricante (Invitrogen, 15596-018). La concentración y la pureza del ARN fue determinada en un espectrofotómetro Nanodrop-1000 (Thermo Scientific).

15 RT-PCR en tiempo real

[0210] Dos microgramos de ARN total se trató con ADNasa I (Invitrogen; 18068-015) y ADNc fue sintetizado utilizando cebadores de hexámero aleatorios y transcriptasa inversa de SuperScript II-MMLV (Invitrogen, 18064-014). La reacción RT (30µl) fue diluida 4 veces en H₂O.

20 [0211] La expresión génica fue determinada por qPCR SYBR Green, usando mezcla de PCR SYBR Green (Roche, 04707516001) y 2µl de ADNc como un modelo. ARN no sometido a transcriptasa inversa se usó como un control negativo para amplificación por PCR. Los cebadores específicos de gen son de la siguiente manera:

E-Cadherina	sentido1	5'-GAAAAGAGAGTGGAAGTG-3'
	antisentido1	5'-GTGAAGGGAGATGTATTG-3'
E-cadherina	sentido 2	5'-CAGGTCTCCTCTTGCTCTG-3'
	antisentido2	5'-ACTTTGAATCGGGTGTGAG-3'
N-Cadherina	sentido	5'-GAGGATTAGCCGGAACAACA-3'
	antisentido	5'-AACAAATTTCCCCATCTCC-3'
SNAIL	sentido	5'-AGGATCTCCAGGCTCGAAAG-3'
	antisentido	5'-GACATCTGAGTGGGTCTGGA-3'
SLUG	sentido	5'-TTCGGACCCACACATTACCT-3'
	antisentido	5'-TTGGAGCAGTTTTTGCCTG-3'
ZEB1	sentido	5'-ATGCGGAAGACAGAAAATGG-3'
	antisentido	5'-GTCACGTTCTTCCGCTTCTC-3'
ZEB2	sentido	5'-CGCTTGACATCACTGAAGGA-3'
	antisentido	5'-CTTGCCCACTCTGTGCATT-3'
β2M	sentido	5'-AGCAGAGAATGGAAAGTCAAA-3'
	antisentido	5'-TGCTGCTTACATGTCTCG-3'
(SEQ ID NO: 48-63).		

25 [0212] Q-PCR se realizó en un instrumento LightCycler LC480 (Roche), utilizando las siguientes condiciones de amplificación: 5 min. 95°C, seguido de 45 ciclos de 10 seg. 95°C, 20 seg. 60°C, 20 seg. 72°C. Para conjunto de cebador 1 de E-cadherina, se usó una temperatura de recocimiento de 49°C (en vez de 60°C). Los valores Cp fueron determinados utilizando el software LightCycler 480 SW 1.5 (Roche). La expresión de Beta-2-microglobulina se usó para normalización. Los niveles de expresión génica relativos se calcularon según el modelo descrito por Pfaffl (18).

30 RT-PCR tallo-bucle

35 [0213] La expresión de microARN se determinó por RT-PCR tallo-bucle como se describe (19). Para esto, 100 ng de ARN total se transcribió inversamente utilizando 0,375 pmol de cebadores tallo-bucle (SL) específicos de miR:

miR-141:

5'-GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC
CCATCT-3'

5

miR-200c:

5'-GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC
TCCATC-3'

10

miR-181a-1:

5'-GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACA
CTCAC-3'

miR-124*:

5'-GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC
GGCATT-3'

15

miR-518b:

5'- GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACA
CCTCT-3'

20

miR-520f:

5'- GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC
AACCCCT-3'

25

miR-524-5p:

5'-GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC
GAGAAA-3'

30

miR-524-3p:

5'- GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC
ACTCCA-3'

35 (SEQ ID NO: 64-71

en 1xRT tampón, conteniendo 0,25 mM dNTPs, 3,33 u/μl transcriptasa inversa SuperScript II-MMLV (Invitrogen) y 0,25 u/μl de inhibidor de ribonucleasa HRP-I (Amersham), durante 30 min. a 16°C, 30 min a 42°C y 5 min a 85°C. Los productos de RT tallo-bucle fueron diluidos 2 veces en H₂O.

40

[0214] La expresión de microARN fue determinada por qPCR SYBR Green, usando mezcla de RCP (0,5 μl cebador sentido (25 pmol/μl), 0,5 μl cebador antisentido (25 pmol/μl), 8 μl H₂O, 10 μl 2x mezcla RCP SYBR Green, Roche) y 2μl de producto RT de tallo-bucle como un modelo. El ARN no sometido a SL-RT se usó como un control negativo para amplificación de PCR. Los cebadores específicos de miR son de la siguiente manera:

45

Cebadores sentido:

5 miR-141: 5'-GCCCCGCTAACACTGTCTGGTAAAG -3'
 miR-200c: 5'-GCCCCGCTAATACTGCCGGGTAATG -3'
 miR-181a-1: 5'-TGCCAGAACATTCAACGCTGTCCG-3'
 miR-124*: 5'-TGCCAGTAAGGCACGCGGTGA-3'
 miR-518b: 5'-TGCCAGCAAAGCGCTCCCTTTAG-3'
 miR-520f: 5'-GCCCCGCAAGTGCTTCCTTTAGAG-3'
 10 miR-524-5p: 5'-GCCCCGCTACAAAGGGAAGCACT-3'
 miR-524-3p: 5'-TGCCAGGAAGGCGCTTCCTTTG-3'

Se usó un cebador universal antisentido: 5'-GTGCAGGGTCCGAGGT-3'
 (SEQ ID NO: 72-80)

15 [0215] Q-PCR se realizó en un instrumento LightCycler LC480 (Roche), utilizando las siguientes condiciones de
 amplificación: 5 min. 95°C, seguido de 45 ciclos de 10 seg. 95°C, 20 seg. 60°C, 10 seg. 72°C. Los valores Cp
 fueron determinados utilizando el software LightCycler 480 SW 1.5 (Roche). La expresión de U6 snRNA (RNA6-
 1) se usó para normalización (U6 (RNU6-1) cebadores: RT 5'-GTCATCCTTGCGCAGG-3' U6 sentido 5'-
 20 CGCTTCGGCAGCACATATAC-3' y U6 antisentido 5'-AGGGGCCATGCTAATCTTCT-3' (SEQ ID NO: 81-83). Los
 niveles de expresión de miR relativo se calcularon según el modelo descrito por Pfaffl (18).

Análisis de secuencia de ADN

25 [0216] La secuencia de los ARNm clonados en los vectores lentivirales para los hits como se describe en la tabla
 1 fue verificada de la siguiente manera. Los ácidos nucleicos totales de células transducidas lentivirales fueron
 aislados utilizando reactivo de Trizol, según las instrucciones del fabricante (Invitrogen, 15596-018). La
 concentración y la pureza de los ácidos nucleicos fue determinada en un espectrofotómetro Nanodrop-1000
 (Thermo Scientific). ADN proviral fue amplificado por PCR, usando 500 ng de ácidos nucleicos como entrada, y
 cebadores específicos de vector lentiviral de pCDH (sentido: 5'-CACGCTGTTTTGACCTCCATAGA-3',
 30 antisentido: 5'-CACTGACGGGCACCGGAG-3', (SEQ ID NO: 84-85).) durante 35 ciclos a una temperatura de
 recocimiento de 65°C. Los productos amplificados fueron purificados utilizando el sistema de purificación de las
 preparaciones de Wizard PCR (Promega). El análisis de secuencia de ADN se realizó utilizando 2 µl de amplicón
 purificado, 5 pmoles de cebador específico de pCDH (5'-GACCTCCATAGAAGATTCTAGAGCTAGC-3', (SEQ ID
 NO: 86).), y el kit Big Dye Terminator v 1.1 (Applied Biosystems). Los productos fueron analizados en un
 35 analizadores de ADN 3730 (Applied Biosystems). Los datos fueron recogidos utilizando el Collection Software
 v3.0 y analizados utilizando el programa Sequencing Analysis v5.3.1 (Applied Biosystems). La secuencia para
 todos los ARNm clonados fue correcta y se da en la tabla 4.

Resultados

40 Selección de microARN por transición mesenquimal a epitelial (MET)

[0217] EMT en células tumorales se produce a partir de una reprogramación transcripcional de la célula. En
 particular, la represión transcripcional del promotor del gen de E-cadherina (*CDH1*) se ha demostrado que
 45 desencadena el fenotipo de EMT. La proteína E-cadherina es una de las moléculas de cadherina más
 importantes que media los contactos célula-célula en células/tejidos epiteliales. La *CDH1* es reprimida por la
 unión de los represores transcripcionales, SNAI1, SNAI2, TCF3, TWIST, ZEB1 o ZEB2 (20-23), a tres de las
 denominadas E-cajas en la región del promotor proximal de *CDH1* (24-26). La inhibición de la unión de estos
 represores al promotor de *CDH1* puede revertir la EMT, llamada también transición mesenquimal a epitelial
 50 (MET), e inhibe la invasión de células tumorales y la progresión tumoral en modelos animales (27). Se plantea la
 hipótesis de que un conjunto extenso de microARN es capaz de inducir MET, mediante la orientación de uno de
 los represores transcripcionales asociados a la EMT. La supresión de los represores transcripcionales asociados
 a la EMT dará lugar a la reactivación de la expresión génica de *CDH1*. Esta reactivación de *CDH1* se puede
 monitorizar fácilmente por activación de luciferasa de luciérnaga dirigida por el promotor de *CDH1*. Por lo tanto,
 55 se clonó el elemento de núcleo del promotor del gen de *CDH1*, que contenía las tres E-cajas (24,25), en un
 vector de expresión para dirigir la expresión de luciferasa de *luciérnaga*. Como control interno, se insertó un
 casete de luciferasa de *Renilla* impulsado por el promotor HSV-Tk en el mismo vector (pEcad-Luc). Se transfirió
 de forma estable la línea celular de cáncer de vejiga TSUp1 con este constructo indicador. La expresión de
 60 ARNm de *CDH1* endógena en esta línea celular no es detectable, mientras que diferentes marcadores
 mesenquimales son expresados (24, 28).

Selección de MET inducida por microARN: configuración

65 [0218] Las células TSUp1-pEcad-Luc son susceptibles de selección de puromicina. Como primera fase en la
 configuración del modelo de selección, se determinó la eficiencia de transducción de las células TSU. Las células
 TSU fueron infectadas con un lentivirus que expresa eGFP en diferentes MOIs, y las células transducidas fueron

seleccionadas en el medio que contenía puromicina. El número de células infectadas, y por lo tanto resistentes a la puromicina, fue determinado por un ensayo de supervivencia celular MTT. En dos experimentos independientes, demostramos que a un MOI de 8, una eficiencia de transducción (MTT de célula infectada+puro/MTT de célula infectada-puro x 100%) de al menos 60% se pudo obtener. Con en estos resultados, se decidió realizar los experimentos de infección pilotos en un MOI de 3 y 30.

[0219] Recientemente, la familia miR-200 y miR-205 demostraron regular EMT dirigiéndose a ZEB1 y ZEB2 (13-15). Dos ARNmi de la familia miR-200, miR-141 y miR-200c, fueron seleccionados para configurar y optimizar nuestro ensayo de selección de MET. Las células TSU fueron infectadas con ambos vectores de ARNmi (MOI=30). Dos (día 4) y seis (días 8) días postinfección, se midió la actividad de luciferasa de luciérnaga dirigida por el promotor de E-cadherina y se normalizó contra la actividad de luciferasa de Renilla controlada por HSV-Tk. El día 8, la proporción de FLuc/RLuc fue inducida más de 2 veces por miR-200c, y más de 1,5 veces por miR-141 (figura 1; tabla 1). Para reducir el "ruido" de células no-infectados (FLuc- y RLuc+), se aplicó la selección de puromicina. La proporciones de FLuc/RLuc inducidos por miR-141 y miR-200c fueron comparables con aquellos sin selección.

[0220] Las células TSU fueron infectadas con 0,2 microlitros de lentivirus de miR no diluido. Dos días después de la infección, la selección de puromicina fue aplicada durante 4 días. Seis días post-infección las actividades de luciferasa se midieron. El resultado de la selección de los primeros 80 microARN se muestra en la figura 1. Los cuatro "hits" positivos (FLuc/RLuc > media + 2 x SD) fueron microARN conocidos por regular EMT, es decir, miR-141, miR-200c, miR-205 y miR-429. Este experimento de piloto indicó la validez del sistema de selección.

Selección de MET inducida por microARN

[0221] Diferentes constructos lentivirales en la biblioteca de miR tienen títulos muy altos. Por lo tanto, sin dilución de caldos de virus antes de la infección, en algunos casos MOIs altos de virus fueron aplicados. La adición de lentivirus de miR-141 y miR-200c en MOI altos (100 a 600) no tuvo ninguna toxicidad significativa, mientras que la inducción de la proporción FLuc/RLuc mejoró ligeramente. Por lo tanto, la biblioteca de expresión de microARN basado en lentivirus entera fue seleccionada utilizando un microlitro de lentivirus no diluido (todo por duplicado). Después de la selección todos los 1120 miRs (14 placas, cada una conteniendo 80 vectores miR), 65 "hits" positivos (FLuc/RLuc > media + 2 x SD) se descubrieron. Además de estos 65 miRs, 59 miRs adicionales fueron seleccionados basados en, por ejemplo, señal de Luc aumentada con señal de RLuc aumentada, dejando la proporción por debajo del umbral. Estos 124 miRs fueron reselectionados en el modelo TSUpr1-pEcad-luc, después de lo cual 30 miRs con una proporción FLuc/RLuc positiva reproducible permanecieron (26 de 30 del grupo de 65 "hits").

Validación de microARN de MET preliminar

[0222] Para validar adicionalmente y seleccionar miRs importantes para regulación de EMT/MET, se estudiaron sus efectos en la expresión de gen endógeno, es decir, *CDH1* (*E-cadherina*), *CDH2* (*N-Cadherina*), *SNAI1* (*SNAIL*), *SNAI2* (*SLUG*), *ZEB1* (*deltaEF1*) y *ZEB2* (*SIP1*). Células TSUpr1-pEcad-luc fueron infectadas con los 30 miRs identificados por selección de MET (MOI= 30 y 100). Seis días después de la infección, el ARN total fue aislado y usado para análisis de qPCR de los genes anteriormente mencionados (tabla 1). La regulación prevista de la expresión de *CDH1* (*E-cadherina*) fue solo observada con algunos miRs (n=10), como fue la infrarregulación de la expresión de *CDH2* (*N-Cadherina*) (n=8). De esos miRs, miR-181a-1, miR-200a, miR-429 y miR-524 dio lugar tanto a sobreexpresión de *CDH1* como infrarregulación de *CDH2*. La cinética de la expresión de cadherina por miRs asociados a EMT es dependiente de cadherina. Esto puede explicar la sobre e infra regulación no consistente de ambas cadherinas, en momento temporal elegido (6 días post introducción de miR). La marca de EMT o de inverso, MET, es conmutación de cadherina. Por lo tanto, la proporción (relativa) de expresión de *CDH1/CDH2* se usó como una medida para estudiar el papel de los miRs identificados en la regulación de EMT. Como se muestra en la tabla 1, diez miRs (de los 12) tuvieron una proporción de *CDH1/CDH2* aumentada (gama: 1,85-17,65). Los otros dos miRs indujeron la expresión de *CDH1* significativamente (en línea con la inducción de FLuc), pero estos miRs también sustancialmente (> 2 veces) indujeron la expresión de *CDH2*. De los 12 miRs que indujeron la expresión de *CDH1* endógena o indujeron un cambio de cadherina, todos también infrarregularon al menos un represor transcripcional de *CDH1* en el nivel de ARN.

Conclusiones

[0223] En el conjunto de 12 miRs seleccionados, los miRs asociados a la EMT conocidos (familia miR-200 y miR-205) están presentes, lo que confirma la especificidad de nuestro sistema. De los restantes 9 miRs, tres miRs (miR-518b, miR-520f y miR-524) pertenecen a la familia miR-515 humana. En este estudio, los miembros de la familia miR-200 en general infrarregularon la expresión de *ZEB2*, lo que está asociado a la infrarregulación de la expresión de *CDH2* (en el día 8). Por otro lado, los miembros de la familia miR-515 en general infrarregularon la expresión de *SNAI2*, que estaba asociada con la sobreexpresión de la expresión de *CDH1*. Estas diferencias destacables pueden subyacer dos mecanismos diferentes de regulación de EMT por los miembros de la familia

miR-200 y miR-515. La expresión y la función de los 3 miembros de la familia miR-515 no ha sido estudiada ni reportada en las bases de datos disponibles públicamente, y por lo tanto proporciona una primera visión sobre el papel de estos miR en la EMT.

5 Ejemplo 2

Materiales y métodos

Cultivo celular

10 [0224] Las células TSUp1/pEcad-luc/Rluc (a.k.a. TSUp1-pEcad) y la línea celular de cáncer de próstata PC-3 (ATCC# CRL-1435) se mantuvieron en medio RPMI-1640 (Invitrogen, 31870), suplementadas con 10% de suero bovino fetal (Sigma, F7524), L-glutamina (Invitrogen 25030-024) y para TSUp1/pEcad-luc/Rluc 50µg/ml zeocina (Invitrogen, R250-01). Las células se mantuvieron en una atmósfera humedecida a 37°C y 5% CO₂.

15 Generación de un sistema de expresión de ARNmi inducible

[0225] Para facilitar la expresión de ARNmi a largo plazo y controlada, se hizo uso del sistema de expresión de ARNmi de miR-X Tet-inducible (Clontech Inc.). El precursor de ARNmi fue clonado en la 3'UTR del unidad de transcripción de proteína fluorescente ZsGreen. La expresión de ARNmi se determina en el vector por el fuertemente regulado, promotor inducible Ptight, y la actividad del transactivador coexpresado (Tet-on). En presencia de doxiciclina el promotor Ptight será activado, dando como resultado la coexpresión de altos niveles de ARNmi y proteína ZsGreen1, con niveles bajos de expresión en células no inducidas.

25 [0226] Los precursores miR-200c y miR-520f fueron amplificados por PCR, usando ADN proviral de células TSUp1 infectadas con lentivirus como modelo, y los cebadores específicos de vector sentido y antisentido pCDH universales (SEQ ID NO: 84-85). Los productos de amplificación fueron digeridos con *EcoRI/NheI*; estos sitios son del sitio de clonación múltiple del vector de pCDH lentiviral. Los fragmentos digeridos fueron clonados en los sitios *EcoRI/NheI* de un vector modificado pEGFP-N3 (Clontech #6080-1), que fue obtenido por *BamHI / NotI* excitando el EGFP ORF desde el vector, y cerrando el vector después del relleno mediado por ADN polimerasa de Klenow de los salientes *BamHI* y *NotI*. El precursor de miR fue extirpado de este vector por digestión de *EagI/NheI* y clonado en los sitios *EagI/NheI* del vector pmRi-ZsGreen1 (Clontech PT5049-1, Fig. 5A). La clonación apropiada fue confirmada por análisis de secuencia de ADN. El vector de transactivador pTet-on (Clontech PT3899-5) fue transfectado en la línea celular de cáncer de próstata PC3 (ATCC# CRL-1435). Los clones resistentes a G418 fueron evaluados para actividad de pTet-on apropiada, en un ensayo de indicador de lucifereasa de pTRE transitorio.

40 [0227] Después, los vectores pmRi-ZsGreen1-miR-X fueron transfectados en las células PC3-Tet-on (clon 8), junto con una etiqueta de selección de puromicina. Las colonias resistentes a la puromicina fueron seleccionadas, y se seleccionaron rápidamente para expresión de ZsGreen1 inducible con DOX (por medición de fluorescencia en placas de 96 pocillos en el multímetro Victor3). Las células ZsGreen1 positivas fueron además analizadas para expresión ARNmi inducible con DOX.

Ensayos de invasión celular

45 [0228] Para ensayos de invasión celular, células PC3-mRi-ZsGreen1-miR-X inducibles con DOX fueron incubadas en presencia de DOX (1 µg/ml) durante 2 días, antes del ensayo de invasión. Células PC3 y TSUp1-pEcad fueron transducidas con las partículas lentivirales de expresión de miR (MOI=30), como se describe en el ejemplo 1. Las células transducidas lentivirales fueron seleccionadas por puromicina y transferidas 1 vez antes del uso. Cincuenta mil (PC3) o 40.000 (TSUp1-pEcad) células fueron sembradas en cámaras de invasión Biocoat Matrigel (8 micras; BD 354480) en medio sin suero (Fig. 6A). La cámara de invasión fue colocada en 24 pocillos que contenían medio con 10% de suero fetal de ternera como quimioatrayente. Como control, la misma cantidad de células fue sembrada en placas de cultivo de 24 pocillos. Después de 48 horas de incubación, las células de la cámara de invasión fueron retiradas por aspiración y limpieza del compartimento interno con un hisopo. La cámara de invasión se puso luego en reactivo de viabilidad celular CellTiter-GLO (CTG, Promega-G7571), se incubó durante 15 minutos, y luego se analizó en un luminómetro Victor3. La invasión celular fue calculada como la actividad de CTG en la parte inferior de la membrana dividida por la actividad de CTG de las células que crecieron en una placa de 24 pocillos. La inhibición de la invasión celular por un ARNmi específico fue calculada por la división del porcentaje de invasión celular de células tratadas con DOX versus células no tratadas, o la invasión de células transducidas contra células transducidas por el vector vacío.

RT-PCR en tiempo real

65 [0229] El mismo protocolo que se describe bajo el ejemplo 1 se usó para RT-PCR de los genes siguientes: vimentina y CDH11 y HPRT. Los cebadores siguientes fueron usados para los genes respectivos:

HPRT	sentido	5'- CTCAACTTTAACTGGAAAGAATGTC -3'
	antisentido	5'- TCCTTTTCACCAGCAAGCT -3'
		5'- GGCTCAGATTCAGGAACAGC -3'
	antisentido	5'- GCTTCAACGGCAAAGTTCTC -3'
CDH11	sentido	5'- GGTCTGGAACCAGTTCTTCG -3'
	antisentido	5'- GGCATGAATGTTCCCTGATT -3'
(SEQ ID NO: 112-117)		

Resultados

5

Validación de los microARN que inducen MET

Validación *in vitro* usando miméticos de microARN

10 [0230] Para convalidar los microARN identificados mediante selección de una biblioteca de expresión de ARNm lentiviral, se usaron miméticos de ARNm sintético. Las miméticos de miR sintético usadas fueron suministradas por Ambion o Dharmacon (Thermo Scientific); véase tabla 7 para detalles. Las miméticos de miR-200c de ambas compañías fueron comparados y mostraron igual eficacia (es decir, inducción de *CDH1*; datos no mostrados). Las células TSUp1-pEcad-luc/RLuc fueron transitoriamente transfectadas con 20 a 60 nM de miméticos de miR sintético. Cuatro días post-transfección, se midió la actividad de Luc y RLuc, y las proporciones Luc/RLuc fueron normalizadas para células transfectadas mimético de control negativo (NC). De todas las miméticos evaluadas, miR-124, 181a, 200c, 206 y miR-520f mostraron una inducción de la proporción Luc/RLuc (Fig. 4A). Salvo miR-200c y miR-520f, las otras miméticos de miR infrarregularon la actividad de RLuc, y mostraron toxicidad celular (no mostrada). De todos estos miRs, miR-200c y miR-520f fuertemente indujeron (~20-veces) expresión de *CDH1* endógena, mientras que miR-124* y miR-181a indujeron *CDH1* aproximadamente 2,5-veces (Fig. 4B). MiR-200c, miR-518 y miR-524 infrarregularon la expresión de *CDH2* más del 30% (Fig. 4B). La eficacia de las miméticos de miR fue además evaluada en otras diferentes líneas celulares tumorales con un fenotipo tipo mesenquimal (basadas en las proporciones de expresión de ARNm de vimentina y *CDH1*; figura 3). Como en TSUp1-pEcad-luc/RLuc, miR-200c y miR-520f también indujeron la expresión de *CDH1* en células TSUp1, PC3, PANC-1, MIA-PaCa2 y MDA-MB-231 de tipo salvaje (Fig. 4C).

25

Generación de un sistema de expresión de ARNm inducible

30 [0231] Clones resistentes a G418 fueron evaluados para actividad de pTet-on apropiada, en un ensayo indicador de luciferasa pTRE transitorio. Tras el tratamiento con doxiciclina (DOX) (0,1 y 1,0 ug/ml), el clon 8 mostró una inducción fuerte de expresión de luciferasa, mientras que la actividad de Luc en células no inducidas (sin DOX) no excedía los niveles básicos (Fig. 5B). Como se muestra en la figura 5C, diferentes clones transfectados con pmRi-ZsGreen1-miR-X clones mostraron expresión de miR-200c o miR-520f inducible con DOX, que fue dependiente de la dosis de DOX, alcanzando niveles de máximo en alrededor de 1 ug/ml DOX (no mostrado).

35

40 [0232] La expresión génica endógena en las células que expresan miR-200c y miR-520f inducible fue analizada por qPCR. Se observó sobreexpresión de *CDH1*, mientras que se observó infrarregulación de *CDH11* y vimentina reiteradamente para ambos miRs (tabla 8). Las células PC3 también mostraron la inducción más débil de *CDH1* por miméticos de miR-200c y miR-520f de todas las líneas celulares evaluadas. Además de la infrarregulación de los marcadores mesenquimales, los represores transcripcionales, tales como ZEB1, ZEB2 y SNAI2 fueron también infrarregulados después de la inducción de miR-200c y miR-520f (tabla 8). Colectivamente, estos datos indican que en las líneas celulares inducibles miR-200c y miR-520f, los miRs correspondientes al igual que sus genes diana de EMT directos e indirectos se pueden regular a nivel molecular.

45 Inhibición de invasión de células tumorales

50 [0233] Para estudiar si los miRs identificados también regulan la EMT en el nivel celular, se realizaron ensayos de invasión celular utilizando los clones estables inducibles miR-520f y miR-200c como generados por encima, y miR-200c y miR-520f que expresan lentivirus como se ha descrito anteriormente. Todos los clones de miR-200c y miR-520f evaluados, inhibieron la invasión celular en el rango de 48-68% (Fig. 6B). La inducción de la expresión de ARNm fue determinada en cultivos en paralelo y se encontró positiva (Fig. 6C). Los niveles de miR no mostraron una correlación significativa con el porcentaje de invasión. Estos datos indican que miR-520f, identificado en el ensayo de selección de Luc como un ARNm de inducción de MET, podría también revertir uno de los aspectos biológicos celulares clave de EMT, es decir invasión celular.

[0234] Para confirmar el efecto de los otros miembros de la familia miR-515, miR-518b y miR-524, en la invasión celular fue evaluada en el modelo de selección TSUpr1/pEcad-luc/RLuc. Las células TSUpr1-pEcad fueron infectadas con lentivirus que contiene precursor de miR. Los tres miRs de la misma familia de miR-515, miR-518b, miR-520f y miR-524, inhibieron la invasión celular en 25-53% (Fig. 6D). La expresión de los ARNm maduros correspondientes fue confirmada por qPCR de tallo-bucle (como se describe en el ejemplo 1; no mostrado).

Conclusiones

[0235] Los tres miembros de la familia miR-515, miR-518b, miR-520f y miR-524, fueron identificados en la selección de MET. Los miméticos de miR-518b y miR-524 infrarregularon la expresión del marcador mesenquimal CDH2, mientras que, los miméticos de miR-520f sobrerregularon fuertemente la expresión del marcador epitelial *CDH1* en diferentes modelos de línea celular. La invasión celular de las células PC3 a través de Matrigel fue inhibida después de la inducción de la expresión de miR-520f. El efecto anti-invasión se observó también cuando los tres ARNm fueron expresadas en el modelo de selección TSUpr1-pEcad. De hecho, miR-200c no redujo la invasión en este modelo. Esto indica la gran fuerza de estos tres miembros de miR-515 como teniendo actividad de anti-invasión en diferentes células y tipos de tumores a diferencia de miR-200c.

Inhibición de la formación de metástasis en un modelo animal

[0236] El hecho de que estos tres miembros de la familia miR-515 pudieran activar el promotor del gen E-cadherina *in vitro*, basado en la activación del indicador de luciferasa, y la inducción de *CDH1* endógena (miR-520f) o la reducción de la expresión de CDH2 (miR-518b y miR-524), y pudieran todos inhibir la invasión celular tumoral *in vitro*, indica que estos microARN son potentes herramientas para inhibir la invasión y la metástasis de células tumorales *in vivo*. Para confirmar la actividad anti-metastásica de miR-520f, miR-518b y miR-524 en una preparación *in vivo*, se configuró un experimento de metástasis tumoral experimental. Ratones desnudos, machos, inmunodeficientes BALB/c (6-8 semanas de edad) fueron inyectados con 0,5 x 10⁶ (200µl de volumen) de células PC3-mRi-ZsGreen1 que expresan bien miR-520f, miR-518b y miR-524, vía la vena de la cola (lateral) (29). A grupos de 6 ratones se les dio agua potable que contenía DOX (0,2 mg/ml) durante todo el experimento, y un grupo de control de ratones inyectados con células PC3-mRi-ZsGreen1 no tratadas con DOX que expresan bien miR-520f, miR-518b y miR-524 reciben agua libre de DOX. Los ratones se monitorizan diariamente, y si no sufren graves inconvenientes, se sacrifican después de 2 meses. Los pulmones son el sitio primario de metástasis, ya que contienen el primer lecho capilar que las células inyectadas encontrarán después de la inyección (30, 31). Se sugiere que las células que pasan el lecho capilar del pulmón, también pueden formar metástasis en otros órganos (30). El número y tamaño de metástasis se determina de forma macroscópica (es decir, por examen visual a través de un microscopio de disección) y microscópicamente (es decir, estudiando la expresión de ZsGreen en el pulmón y otras secciones de tejido). Se prevé, basado en todos los datos disponibles presentados anteriormente, que las células que expresan bien miR-520f, miR-518b o miR-524 forman menos (y menor) metástasis en este modelo animal que las células que no expresan dichos ARNm. Esta prueba de concepto preparará el terreno para otra prueba preclínica de miR-520f, miR-518b y miR-524 como un fármaco para tratar enfermedades asociadas a la EMT.

[0237] Para confirmar adicionalmente el efecto inhibitorio de miR-520f, miR-518b y miR-524 en la metástasis *in vivo*, diferentes líneas celulares tumorales, diseñadas con pmRi-ZsGreen1-miR-520f, 518b o 524, se inyectan ortotópicamente en ratones desnudos (NOD-SCID o BALB/c nu/nu). En momentos diferentes después de la inyección de células tumorales, los ratones reciben agua que contiene DOX o libre de DOX. La invasión local y la metástasis distante a los pulmones y otros órganos se monitoriza por la formación de imágenes *in vivo* de ZsGreen1 o LUC. La sobre-expresión de miR-520f, miR-518b o miR-524 (inducida por DOX en el agua potable) está prevista que reduzca el número de metástasis. Además, los ratones ortotópicamente inyectados con células tumorales, donde miR-520f, miR-518b o miR-524 se activaron, está previsto que tengan un índice de supervivencia media que se prolongará significativamente en comparación con el de los ratones que no han recibido DOX (es decir sin activación de miR-520f, miR-518b o miR-524 en células tumorales). La supervivencia de los animales se definirá por la supervivencia sin complicaciones serias.

Tabla 1: lista de microARN identificada por selección de MET (luciferasa) y regulación de genes endógenos asociados a la EMT. Células TSUpr1-pEcad-Luc fueron infectadas con 30 microARN positivos de luciferasa (véase texto). El ARN total fue transcrito inverso, y usado para análisis de PCR en tiempo real con SYBR Green de los genes de *CDH1* (*Ecad*), *CDH2* (*Ncad*), *SNAI1*, *SNAI2*, *ZEB1* y *ZEB2*. LUC, inducción de proporción FLuc/RLuc, promedio de 3 experimentos; *CDH1*, (inducción de) expresión de *CDH1* endógena; CDH2, (inhibición de) expresión de CDH2 endógena; represor, represores transcripcionales de *CDH1* que fueron infrarregularados al menos 2 veces, o entre 1,3 y 2 veces (subrayado). Solo los datos para 12 microARN que son de interés para posteriores estudios se muestran; véase texto para un razonamiento detallado de la selección estos 12 miRs.

ARNmi	LUC	<i>CDH1</i>	<i>CDH2</i>	<i>CDH1/CDH2</i>	Represor
-------	-----	-------------	-------------	------------------	----------

ARNmi	LUC	CDH1	CDH2	CDH1/CDH2	Represor
miR-124-1	1,76	1,30	2,46	0,53	SNAI2
miR-181a-1	1,59	4,22	0,89	4,74	SNAI2
miR-141	1,95	0,58	0,04	14,45	ZEB2, <u>SNAI1</u>
miR-200a	1,66	2,18	0,67	3,25	ZEB2
miR-200c	2,53	0,71	0,04	17,65	ZEB2, <u>ZEB1</u>
miR-205	1,79	0,39	0,08	4,86	ZEB2
miR-429	1,50	3,39	0,41	8,28	<u>ZEB2</u>
miR-206	1,73	2,24	2,07	1,08	SNAI2
miR-518b	1,62	3,09	1,67	1,85	SNAI2
miR-520f	1,43	12,92	1,16	11,14	SNAI2, <u>ZEB2</u>
miR-524	1,57	1,60	0,45	3,56	<u>ZEB2</u>
miR-200b	2,08	5,17	1,18	4,38	<u>ZEB2</u>

Tabla 2. Secuencias precursoras de ARNmi identificados en la selección de MET (véase tabla 1). Lista de secuencias precursoras de ARNmi (dirección 5' a 3'). Todas las secuencias fueron obtenidas de la miRBase (versión 14: Sept 2009; www.mirbase.org).

5

SEQ ID	ARNmi	Secuencia precursora
22	miR-124-1	AGGCCUCUCUCUCCGUGUUCACAGCGGACCUUGAUUUAAAUGUC CAUACAAUUAAGGCACGCGGUGAAUGCCAAGAAUGGGGCUG
23	miR-124-2	AUCAAGAUUAGAGGCUCUCUCUCCGUGUUCACAGCGGACCUUGAU UUAUGUCAUACAAUUAAGGCACGCGGUGAAUGCCAAGAGCGGAGC CUACGGCUGCACUUGAA
24	miR-124-3	UGAGGGCCCCUCUGCGUGUUCACAGCGGACCUUGAUUUAAAUGUCUA UACAAUUAAGGCACGCGGUGAAUGCCAAGAGAGGGCGCCUCC
25	miR-181a-1	UGAGUUUUGAGGUUGCUCAGUGAACAUUCAACGCUGUCGGUGAGU UUGGAAUUAUUUUCAAAACCAUCGACCGUUGAUUGUACCCUAUGGC UAACCAUCAUCUACUCCA
26	miR-141	CGGCCGGCCCUGGGUCCAUCUCCAGUACAGUGUUGGAUGGUCU AAUUGUGAAGCUCCUAAACACUGUCUGGUAAGAUGGCUCCCGGGUG GGUUC
27	miR-200a	

SEQ ID	ARNmi	Secuencia precursora
		CCGGGCCCCUGUGAGCAUCUUACCGGACAGUGCUGGAUUUCC CAGCUUGACUCUAACACUGUCUGGUAACGAUGUUCAAAGGUGA CCCGC
28	miR-200c	CCCUCGUCUUACCCAGCAGUGUUUGGGUGCGGUUGGGAGUCUC UAAUACUGCCGGGUAUGAUGGAGG
29	miR-205	AAAGAUCUCAGACAAUCCAUGUGCUUCUCUUGUCCUUCAUUC ACCGGAGUCUGUCUCAUACCCAACCAGAUUCAGUGGAGUGA AGUUCAGGAGGCAUGGAGCUGACA
30	miR-429	GCGUCUUACCAGACAUGGUUAGACCUGGCCUCUGUCUAAUAC UGUCUGGUA AAAACCGUCCAUCGCGUC
31	miR-206	UGCUUCCCGAGGCCACAUGCUUCUUUAUAUCCCCAU AUGGAUU ACUUUGCUAUGGAAUGUAAGGAAGUGUGUGGUUCCGGCAAGUG
32	miR-518b	UCAUGCUGUGGCCCUCCAGAGGGAAGCGCUUUCUGUUGUCUGAAAG AAAACAAAGCGCUCCCCUUUAGAGGUUUACGGUUUGA
33	miR-520f	UCUCAGGCUGUGACCCUCUAAAGGGAAGCGCUUUCUGUGGU CAGAAAGAAAAGCAAGUGCUUCCUUUAGAGGGUUACCGUUU GGGA
34	miR-524	UCUCAUGCUGUGACCCUACAAAGGGAAGCACUUUCUCUUGUCCAAA GGAAAAGAAGGCGCUUCCUUUGGAGUGUUACGGUUUGAGA
35	miR-200b	CCAGCUCGGGCAGCCGUGGCCAUCUUACUGGGCAGCAUUGGAUGGA GUCAGGUCUCUAAUACUGCCUGGUAUGAUGACGGCGGAGCCCUGC ACG

Tabla 3. Secuencias maduras de ARNmi identificados en la selección de MET (véase tabla 1). Lista de secuencias de ARNmi maduros (dirección 5' a 3').

microARN	ARNmi maduro	SEQ ID	ARNmi SEQ madura
hsa-miR-124-1	miR-124	2	uaaggcacgcggugaaugcc
hsa-miR-124-2	miR-124*	3	cguguucacagcggaccuugau
hsa-miR-124-3			
hsa-miR-181a-1	miR-181a	4	aacauucaacgcugucggugagu
	miR-181a*	5	accaucgaccguugauuguacc
hsa-miR-141	miR-141	6	uaacacugucugguaaagaugg
	miR-141*	7	caucuuccagucaguguugga
hsa-miR-200a	miR-200a	8	uaacacugucugguaacgaugu
	miR-200a*	9	caucuuaccggacagucugga
hsa-miR-200b	miR-200b	10	uaauacugccugguaaugauga
	miR-200b*	11	caucuucugggcagcauugga
hsa-miR-200c	miR-200c	12	uaauacugccgguaaagaugga
	miR-200c*	13	cgucuuaccagcaguguuugg
hsa-miR-205	miR-205	14	uccuucuuaccaggagucug
	miR-205*	15	gauuucaguggagugaaguuc
hsa-miR-429	miR-429	16	uaauacugucugguaaaaccgu
hsa-miR-206	miR-206	17	uggaauquaaggaagugugugg
hsa-miR-518b	miR-518b	18	caaagcgcuccuuuagaggu
hsa-miR-520f	miR-520f	19	aagugcuuccuuuagaggguu
hsa-miR-524	miR-524-5p	20	cuacaaaggaagcacuuucuc
	miR-524-3p	21	gaaggcgcuuuccuuuggagu

Tabla 4. Secuencias de ARNmi identificados en la selección de MET como clonadas en vectores lentivirales (véase tabla 1)

Seq ID	ARNmi	Secuencia clonada en vector lentiviral
36	miR-124-1	TTTCTTTCACCTTTCCTTCCTTCCTTCCTTCCTTTCTTCCTCAGGAGAA AGGCCTCTCTCTCCGTGTTACAGCGGACCTTGATTTAAATGTCCATA CAATTAAGGCACGCGGTGAATGCCAAGAATGGGGCTGGCTGAGCA CCGTGGGTTCGGCGAGGGCCCCGCCAAGGAAGGAGCGACC
37	miR-181a-1	GTTGTTTCTGTCTCCCATCCCCTTCAGATACTTACAGATACTGTAAAG TGAGTAGAATTCTGAGTTTTGAGGTTGCTTCAGTGAACATTCAACGCT GTCGGTGAGTTTGAATTTAAATCAAACCATCGACCGTTGATTGTAC CCTATGGCTAACCATCATCTACTCCATGGTGCTCAGAATTGCTGAAG ACAGGAAACCAAA
38	miR-141	

Seq ID	ARNmi	Secuencia clonada en vector lentiviral
		CCTGTAGCAACTGGTGAGCGCGCACCGTAGTTCTCTGTCGGCCGGCC CTGGGTCCATCTTCCAGTACAGTGTTGGATGGTCTAATTGTGAAGCTC CTAACACTGTCTGGTAAAGATGGCTCCCGGGTGGGTTCTCTCGGCAGT AACCTTCAGGGAGCCCTGAAGACCATGGAGGACTACTGACCAACAA CCTCTGACCTT
39	miR-200a	GTTCTTCCCTGGGCTTCCACAGCAGCCCCTGCCTGCCTGGCGGGACC CCACGTCCCTCCCGGGCCCCTGTGAGCATCTTACCGGACAGTGCTGG ATTTCCAGCTTGACTCTAACACTGTCTGGTAACGATGTTCAAAGGTG ACCCGCCGCTCGCCGGGGACACCACCGAGGCACATCCGGAGCTCCTACT
40	miR-200c	AAGCTGCCTGACCCAAGGTGGGCGGGCTGGGCGGGGGCCCTCGTCT TACCCAGCAGTGTTTGGGTGCGGTTGGGAGTCTCTAATACTGCCGGG TAATGATGGAGGCCCTGTCCCTGTGTCAGCAACATCCATCGCCTCA
41	miR-205	AGTGTCTACAGGCTGAGGTTGACATGCATCCCCACCCTCTGAGAAAAA GATCCTCAGACAATCCATGTGCTTCTCTTGTCTTCATTCCACCGGAGT CTGTCTCATACCCAACCAGATTTCAAGTGGAGTGAAGTTCAGGAGGCAT GGAGCTGACAACCATGAGGCCTCGGCAGCCACCGCCACCACCGCCGC CGCCACCACCGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAAGAGTAACT
42	miR-429	AGACACCAGCCCAGGACCCGGAGGCCACCCACACCACCGCCGGCCGA TGGGCGTCTTACCAGACATGGTTAGACCTGGCCCTCTGTCTAATACTGT CTGGTAAAACCGTCCATCCGCTGCCTGATCACCGTTAGAGGAGAGAGC TGCCTGCCCTGCAGCTCATCAGTGCAAAGCC
43	miR-206	GCAAGGAGGAAAGATGCTACAAGTGGCCCACTTCTGAGATGCGGGCT GCTTCTGGATGACACTGCTTCCCGAGGCCACATGCTTCTTTATATCCC CATATGGATTACTTTGCTATGGAATGTAAGGAAGTGTGTGGTTTCGGC AAGTGCCTCCTCGCTGGCCCCAGGGTACCACCCGGAGCACAGGTTTG GTGACCTTCTTC
44	miR-518b	

Seq ID	ARNmi	Secuencia clonada en vector lentiviral
		GCAAACAGGGCAAATAAATGCATCTTTATTTTGTGTCCATTTTAAACCT GGTCAAGGAAAATTCCAACAGCAACATCAAAAAACCAGTGTTGGAG
		CAAGAATATGTCATGCTGTGGCCCTCCAGAGGGAAGCGCTTTCTGTT GTCTGAAAGAAAACAAAGCGCTCCCCTTTAGAGGTTTACGGTTTGAG TAAAGCAGCGTTGAAGTTGATGCTGATCTTGGTAATACATTTGCAGA GCGTGCTTATCATCAG
45	miR-520f	TGTGTCCATTTAAACCTGGTCAAGGAAGATTCCCACAAAAAATCCAC GGTGCTGGAGCAAGAGGATCTCAGGCTGTGACCCTCTAAAGGGAA GCGCTTTCTGTGGTCAGAAAGAAAAGCAAGTGCTTCCTTTTAGAGG GTTACCGTTTGGGAAAAGCAATGTTGAAGTTGATGCTGATCTTGGT AAAATATTTGCAGAGCGTGCTTATCATCAG
46	miR-524	CAAACAGGGCCAATAAATGCATCCTCATTTTTGTGTCCATTTTAAACCT GGGCAAGGAAAATTCCAACAAAAAACCAGAGTTCTGGAGCAAGAA GATCTCATGCTGTGACCCTACAAAGGGAAGCACTTTCTCTTGCCAA AGGAAAAGAAGGCGCTTCCCTTTGGAGTGTTACGGTTTGAGAAAAG CAGCGTTGAAGTTGATGCTTATCTCGGTAATACATTTGTAGAGCATG CTTATCATGAGGCTTGGAC
47	miR-200b	CAGCCGGGCGGCCCCCGGACCCAGCTCGGGCAGCCGTGGCCATCTT ACTGGGCAGCATTGGATGGAGTCAGGTCTCTAATACTGCCTGGTAAT GATGACGGCGGAGCCCTGCACGCAGCGACCGGCCGACCCCGT

Tabla 5. Secuencias semilla de ARNmi identificados en la selección de MET (véase tabla 1). Lista de secuencias semilla de ARNmi (dirección 5' a 3'). Todas las secuencias fueron obtenidas de la miRBase (versión 14: Sept 2009; www.mirbase.org). La secuencia semilla se define como los nucleótidos 2-8 (dirección 5' a 3') de la secuencia de ARNmi maduro.

5

microARN	ARNmi maduro	SEQ ID	Secuencia semilla
hsa-miR-124-1	miR-124	87	aaggcac
hsa-miR-124-2	miR-124*	88	guguuca
hsa-miR-124-3			
hsa-miR-181a-1	miR-181a	89	acauuca
	miR-181a [~]	90	ccaucga
hsa-miR-141	miR-141	91	aacacug

microARN	ARNmi maduro	SEQ ID	Secuencia semilla
	miR-141 [*]	92	aucuucc
hsa-miR-200a	miR-200a	93	aacacug
	miR-200a [*]	94	aucuucac
hsa-miR-200b	miR-200b	95	aaucacug
	miR-200b [*]	96	aucuucac
hsa-miR-200c	miR-200c	97	aaucacug
	miR-200c [*]	98	gucuucac
hsa-miR-205	miR-205	99	ccuucac
	miR-205 [*]	100	auuucag
hsa-miR-429	miR-429	101	aaucacug
hsa-miR-206	miR-206	102	ggaaugu
hsa-miR-518b	miR-518b	103	aaagcgc
hsa-miR-520f	miR-520f	104	agugcuu
		105	aagugcu
hsa-miR-524	miR-524-5p	106	uacaaag
	miR-524-3p	107	aaggcgc

5 Tabla 6: Secuencias de isomiR de ARNmi identificados en la selección (véase tabla 1). Estos isomiRs han sido detectados después del análisis de 66 muestras de tejido humano usando secuenciación profunda de alto rendimiento y solo los isomiRs que representan > 5% del número total de secuencias clonadas están enumerados aquí, a menos que se indique lo contrario.

ARNmi	ARNmi maduro	Semilla (SEQ ID)	Secuencia de IsomiR (SEQ ID)
hsa-miR-520f	mir-520f	AGUGCUU(104)	AAGUGCUUCCUUUAGAGGGU (118)
		AAGUGCU(105)	AAGUGCUUCCUUUAGAGGGUU ^{*1} (119)
			CAAGUGCUUCCUUUAGAGGGU (120)
hsa-miR-518b	mir-518b	AAAGCGC (103)	CAAAGCGCUCCCCUUAGAGGU (121)
			CAAAGCGCUCCCCUUAGAGG (122)
			CAAAGCGCUCCCCUUAGAG (123)
hsa-miR-524	mir-524-3p	AAGGCGC (107)	GAAGGCGCUUCCUUUGGAGUG (124)
			GAAGGCGCUUCCUUUGGAGU (125)
	mir-524-5p	UACAAAG (106)	CUACAAAGGAAGCACUUUCU (126)
			CUACAAAGGAAGCACUUUCUC (127)
			CUACAAAGGAAGCACUUUC (128)
hsa-miR-124-1	mir-124	AAGGCAC (87)	UAAGGCACGCGGUGAAUGCCA (129)
hsa-miR-124-2			UAAGGCACGCGGUGAAUGCCA (130)
hsa-miR-124-3			UAAGGCACGCGGUGAAUGC (131)
			UAAGGCACGCGGUGAAUGCC ^{*3} (132)
hsa-miR-206	mir-206	GGAAUGU (102)	UGGAAUGUAAGGAAGUGUGUGG (133)
			UGGAAUGUAAGGAAGUGUGUGU (134)

ARNmi	ARNmi maduro	Semilla (SEQ ID)	Secuencia de IsomiR (SEQ ID)
			UGGAAUGUAAGGAAGUGUGUG (135)
hsa-miR-181a-1	mir-181a	ACAUUCA(89)	AACAUUCAACGCUGUCGGUGAGUUU (136)
			AACAUUCAACGCUGUCGGUGAGU (137)
			AACAUUCAACGCUGUCGGUGAG (138)
			AACAUUCAACGCUGUCGGUGAGUU (139)
			AACAUUCAACGCUGUCGG (140)
			AACAUUCAACGCUGUCGGU (141)
			AACAUUCAACGCUGUCGGUG (142)
			AACAUUCAACGCUGUCGGUGA (143)
hsa-miR-141	mir-141	AACACUG (91)	UAACACUGUCUGGUAAGAUGG (144)
			UAACACUGUCUGGUAAGAUG (145)
			UAACACUGUCUGGUAAGAUGGC (146)
			UAACACUGUCUGGUAAGAU (147)
			UAACACUGUCUGGUAAGA (148)
			UAACACUGUCUGGUAAGAUGGCU (149)
hsa-miR-200a	mir-200a	AACACUG (93)	UAACACUGUCUGGUAACGAUGUU (150)
			UAACACUGUCUGGUAACGAUGU (151)
hsa-miR-200b	mir-200b	AAUACUG (95)	UAAUACUGCCUGGUAUGAUGAC (152)
			UAAUACUGCCUGGUAUGAUGA (153)
	mir-200b [*]	AUCUUAC (96)	CAUCUUACUGGGCAGCAUUGGA (154)
			CAUCUUACUGGGCAGCAUUGG (155)
hsa-miR-200c	mir-200c	AAUACUG (97)	UAAUACUGCCGGGUAUGAUGGA (156)
			UAAUACUGCCGGGUAUGAUGG (157)
hsa-miR-205	mir-205	CCUUCAU (99)	UCCUUCAUCCACCGGAGUCU (158)
			UCCUUCAUCCACCGGAGUCUG (159)
			UCCUUCAUCCACCGGAGUCUGU (160)
hsa-miR-429	mir-429	AAUACUG (101)	UAAUACUGUCUGGUAACCGU (161)
			UAAUACUGUCUGGUAACCG (162)
^{*1} Secuencia anotada en la miRBase. No detectada en el análisis de secuenciación profundo ^{*2} Solo 0,9% del número total de secuencias clonadas ^{*3} Solo 0,7% del número total de secuencias clonadas			

5

Tabla 7: Lista de miméticos de microARN usados para validación de resultados de selección de MET. Se muestran los proveedores y códigos de producto para todos los precursores de miR sintéticos. Todos los miméticos enumerados son bicatenarios, con cada cadena con un 100% de alineamiento complementario entre sí. Una cadena será la secuencia exacta de las secuencias de ARNmi maduro como está depositada en la miRBase y en enumerada en la tabla 3. La otra cadena contiene ciertas modificaciones.

ARNmi mimético	Proveedor	Código de producto
Mimic Negative Control #1 **	Dharmacon*	CN-001000-01-05
hsa-miR-200c miRIDIAN mimético	Dharmacon	C-300646-05-0005

ARNmi mimético	Proveedor	Código de producto
hsa-miR-520f miRIDIAN mimético	Dharmacon	C-300779-03-0005
hsa-miR-524-5p miRIDIAN mimético	Dharmacon	C-300806-03-0005
hsa-miR-518b miRIDIAN mimético	Dharmacon	C-300798-03-0005
hsa-miR-124 Pre-miR precursor	Ambion	PM10691
hsa-miR-124* Pre-miR precursor	Ambion	PM11154
hsa-miR-181 a Pre-miR precursor	Ambion	PM10421
hsa-miR-181a* Pre-miR precursor	Ambion	PM10381
hsa-miR-200c Pre-miR precursor	Ambion	PM11714
hsa-miR-206 Pre-miR precursor	Ambion	PM10409

*, Dharmacon es ahora parte de Thermo Fisher Scientific.
 ** Este control negativo se basa en C. *Elegans* Cel-miR-67 (Seq ID 163).

5 Tabla 8. Expresión génica en las células PC3 transfectadas con microARN inducible. Células PC3-Tet-on (clon 8) fueron transfectadas de forma estable con los plásmidos de expresión pmRi-ZsGreen1-miR-X. Clones seleccionados (véase figura 5) fueron tratados con 1 µg/ml de DOX durante 4 días. ARN total fue aislado y la expresión génica fue estudiada por qPCR. *, más de 1,5 veces; =, sin cambios; N.d., no determinado.

Gen	PC3-mRi-miR-200c	PC3-mRi-miR-520f
<i>CDH1</i>	=	Sobre*
<i>CDH2</i>	<i>Weakly expressed</i>	<i>Débilmente expresado</i>
<i>CDH1</i>	Infra *	Infra
<i>VIM</i>	Infra	n.d.
<i>ZEB1</i>	Infra	Infra
<i>ZEB2</i>	Infra	=
<i>SNAI1</i>	=	=
<i>SNAI2</i>	Infra	Infra

10 Lista de referencia

[0238]

15 1. Tomita, K., van, B.A., van Leenders, G.J., Ruijter, E.T., Jansen, C.F., Bussemakers, M.J., and Schalken, J.A. Cadherin switching in human prostate cancer progression. *Cancer Res.* 60: 3650-3654 (2000).

20 2. Umbas, R., Isaacs, W.B., Bringuier, P.P., Schaafsma, H.E., Karthaus, H.F., Oosterhof, G.O., Debruyne, F.M., and Schalken, J.A. Decreased E-cadherin expression is associated with poor prognosis in patients with prostate cancer. *Cancer Res.* 54: 3929-3933 (1994).

25 3. Oort van, I., Tomita, K., van, B.A., Bussemakers, M.J., Kiemeneij, L.A., Karthaus, H.F., Witjes, J.A., and Schalken, J.A. The prognostic value of E-cadherin and the cadherin-associated molecules alpha-, beta-, gamma-catenin and p120ctn in prostate cancer specific survival: a long-term follow-up study. *Prostate.* 67: 1432-1438 (2007).

30 4. Batlle, E., Sancho, E., Franci, C., Dominguez, D., Monfar, M., Baulida, J., and Garcia, D.H. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat. Cell Biol.* 2: 84-89 (2000).

5. Cano, A., Perez-Moreno, M.A., Rodrigo, I., Locascio, A., Blanco, M.J., del Barrio, M.G., Portillo, F., and Nieto, M.A. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat. Cell Biol.* 2: 76-83 (2000).

6. Hajra, K.M., Chen, D.Y., and Fearon, E.R. The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer Res.* 62: 1613-1618 (2002).
- 5 7. Yang, J., Mani, S.A., Donaher, J.L., Ramaswamy, S., Itzykson, R.A., Come, C., Savagner, P., Gitelman, I., Richardson, A., and Weinberg, R.A. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* 117: 927-939 (2004).
- 10 8. Bussemakers, M.J., Giroldi, L.A., van Bokhoven, A., and Schalken, J.A. Transcriptional regulation of the human E-cadherin gene in human prostate cancer cell lines: characterization of the human E-cadherin gene promoter. *Biochem. Biophys. Res. Commun.* 203: 1284-1290 (1994).
- 15 9. Giroldi, L.A., Bringuier, P.P., de Weijert, M., Jansen, C., van Bokhoven, A., and Schalken, J.A. Role of E boxes in the repression of E-cadherin expression. *Biochem. Biophys. Res. Commun.* 241: 453-458 (1997).
- 20 10. Rodrigo, I., Cato, A.C.B., and Cano, A. Regulation of E-cadherin gene expression during tumor progression: the role of a new Ets-binding site and the E-pal element. *Exp. Cell Res.* 248: 358-371 (1999).
- 25 11. Conacci-Sorrell, M., Simcha, I., Ben Yedidia, T., Blechman, J., Savagner, P., and Ben Ze'ev, A. Autoregulation of E-cadherin expression by cadherin-cadherin interactions: the roles of beta-catenin signaling, Slug, and MAPK. *J. Cell Biol.* 163: 847-857 (2003).
- 30 12. Gregory, P.A., Bracken, C.P., Bert, A.G., and Goodall, G.J. MicroRNAs as regulators of epithelial-mesenchymal transition. *Cell Cycle.* 7: 3112-3118 (2008).
- 35 13. Burk, U., Schubert, J., Wellner, U., Schmalhofer, O., Vincan, E., Spaderna, S., and Brabletz, T. A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO Rep.* 9: 582-589 (2008).
- 40 14. Gregory, P.A., Bert, A.G., Paterson, E.L., Barry, S.C., Tsykin, A., Farshid, G., Vadas, M.A., Khew-Goodall, Y., and Goodall, G.J. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat. Cell Biol.* 10: 593-601 (2008).
- 45 15. Park, S.M., Gaur, A.B., Lengyel, E., and Peter, M.E. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev.* 22: 894-907 (2008).
- 50 16. Berezikov, E., van, T.G., Verheul, M., van de, B.J., van, L.L., Vos, J., Verloop, R., van de, W.M., Guryev, V., Takada, S., van Zonneveld, A.J., Mano, H., Plasterk, R., and Cuppen, E. Many novel mammalian microRNA candidates identified by extensive cloning and RAKE analysis. *Genome Res.* 16: 1289-1298 (2006).
- 55 17. Lowery, A.J., Miller, N., Devaney, A., McNeill, R.E., Davoren, P.A., Lemetre, C., Benes, V., Schmidt, S., Blake, J., Ball, G., and Kerin, M.J. MicroRNA signatures predict oestrogen receptor, progesterone receptor and HER2/neu receptor status in breast cancer. *Breast Cancer Res.* 11: R27 (2009).
- 60 18. Pfaffl, M.W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29: e45 (2001).
- 65 19. Chen, C., Ridzon, D.A., Broomer, A.J., Zhou, Z., Lee, D.H., Nguyen, J.T., Barbisin, M., Xu, N.L., Mahuvakar, V.R., Andersen, M.R., Lao, K.Q., Livak, K.J., and Guegler, K.J. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.* 33: e179 (2005).
20. Battle, E., Sancho, E., Franci, C., Dominguez, D., Monfar, M., Baulida, J., and Garcia, D.H. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat. Cell Biol.* 2: 84-89 (2000).
21. Cano, A., Perez-Moreno, M.A., Rodrigo, I., Locascio, A., Blanco, M.J., del Barrio, M.G., Portillo, F., and Nieto, M.A. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat. Cell Biol.* 2: 76-83 (2000).
22. Hajra, K.M., Chen, D.Y., and Fearon, E.R. The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer Res.* 62: 1613-1618 (2002).
23. Yang, J., Mani, S.A., Donaher, J.L., Ramaswamy, S., Itzykson, R.A., Come, C., Savagner, P., Gitelman, I., Richardson, A., and Weinberg, R.A. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* 117: 927-939 (2004).

24. Bussemakers, M.J., Giroldi, L.A., van Bokhoven, A., and Schalken, J.A. Transcriptional regulation of the human E-cadherin gene in human prostate cancer cell lines: characterization of the human E-cadherin gene promoter. *Biochem. Biophys. Res. Commun.* 203: 1284-1290 (1994).
- 5 25. Giroldi, L.A., Bringuier, P.P., de Weijert, M., Jansen, C., van Bokhoven, A., and Schalken, J.A. Role of E-boxes in the repression of E-cadherin expression. *Biochem. Biophys. Res. Commun.* 241: 453-458 (1997).
26. Rodrigo, I., Cato, A.C.B., and Cano, A. Regulation of E-cadherin gene expression during tumor progression: the role of a new Ets-binding site and the E-pal element. *Exp. Cell Res.* 248: 358-371 (1999).
- 10 27. Conacci-Sorrell, M., Simcha, I., Ben Yedidia, T., Blechman, J., Savagner, P., and Ben Ze'ev, A. Autoregulation of E-cadherin expression by cadherin-cadherin interactions: the roles of beta-catenin signaling, Slug, and MAPK. *J. Cell Biol.* 163: 847-857 (2003).
- 15 28. Bussemakers, M.J., van, B.A., Tomita, K., Jansen, C.F., and Schalken, J.A. Complex cadherin expression in human prostate cancer cells. *Int. J. Cancer.* 85: 446-450 (2000).
29. Mohanty, S., and Xu, L. Experimental metastasis assay. *J. Vis. Exp.* 1942 (2010).
- 20 30. Rubio, N., Villacampa, M.M., El, H.N., and Blanco, J. Metastatic burden in nude mice organs measured using prostate tumor PC-3 cells expressing the luciferase gene as a quantifiable tumor cell marker. *Prostate.* 44: 133-143 (2000).
- 25 31. Fukuchi, K., Steiniger, S.C., Deryugina, E., LIU, Y., Lowery, C.A., Gloeckner, C., Zhou, B., Kaufmann, G.F., Quigley, J.P., and Janda, K.D. Inhibition of tumor metastasis: functional immune modulation of the CUB domain containing protein 1. *Mol. Pharm.* 7: 245-253 (2010).

Abreviaturas

30 [0239]

- CDH1 - E-cadherina
 CDH2 - N-Cadherina
 CDH11 - OB-cadherina
- 35 CTG - CEITiter-GLO (ensayo de viabilidad celular)
 DOX - Doxiciolina
 EMT - transición epitelial a mesenquimal
 miR - microARN
 MET - transición mesenquimal a epitelial
- 40 MOI - multiplicidad de infección
 MTT - (3-(4,5-Dimetiltiazol-2-il)-2,5-difeniltetrazolio bromuro
 qPCR - PCR cuantitativo
 PCa - cáncer de próstata
- 45 SNAI1 - homólogo 1 SNAIL, alias SNAIL
 SNAI2 - homólogo 2 SNAIL, alias SLUG
 TSU - células TSUpr1-pEcad-Luc
 UBC - cáncer de vejiga urinaria
 VIM - vimentina
- 50 ZEB1 - homeocaja 1 de unión a E-caja de dedo de zinc, alias deltaEF1
 ZEB2 - homeocaja 2 de unión a E-caja de dedo de zinc, alias SIP 1

LISTADO DE SECUENCIAS

[0240]

5 <110> InteRNA Technologies BV
 <120> A miRNA molecule defined by its source and its diagnostic and therapeutic uses in diseases or conditions associated with EMT
 <130> P6029697PCT
 <150> EP 10155512.6

10 <151> 2010-03-04
 <150> US 61/310,452
 <151> 2010-03-04
 <160> 163
 <170> PatentIn version 3.3

15 <210> 1
 <211> 82
 <212> RNA
 <213> Artificial
 <220>

20 <223> motif
 <220>
 <221> misc_feature
 <222> (4)..(4)
 <223> n may be A, U, C, G

25 <220>
 <221> misc_feature
 <222> (11)..(11)
 <223> n may be A, U, C, G
 <220>

30 <221> misc_feature
 <222> (16)..(16)
 <223> n may be A, U, C, G
 <220>

35 <221> misc_feature
 <222> (17)..(17)
 <223> n may be A, U, C, G
 <220>

40 <221> misc_feature
 <222> (19)..(19)
 <223> n is a, c, g, or u
 <220>

45 <221> misc_feature
 <222> (28)..(28)
 <223> X may be A, U, C, G
 <220>

50 <221> misc_feature
 <222> (35)..(35)
 <223> n may be A, U, C, G
 <220>

55 <221> misc_feature
 <222> (37)..(37)
 <223> n may be A, U, C, G
 <220>

60 <221> misc_feature
 <222> (41)..(41)
 <223> n may be A, U, C, G
 <220>

65 <221> misc_feature
 <222> (42)..(42)
 <223> n may be A, U, C, G
 <220>

70 <221> misc_feature
 <222> (45)..(45)
 <223> n may be A, U, C, G
 <220>

75 <221> misc_feature
 <222> (45)..(45)
 <223> n may be A, U, C, G
 <220>

80 <221> misc_feature
 <222> (45)..(45)
 <223> n may be A, U, C, G
 <220>

85 <221> misc_feature
 <222> (45)..(45)
 <223> n may be A, U, C, G
 <220>

90 <221> misc_feature
 <222> (45)..(45)
 <223> n may be A, U, C, G
 <220>

95 <221> misc_feature
 <222> (45)..(45)
 <223> n may be A, U, C, G
 <220>

<222> (51)..(51)
 <223> n may be A, U, C, G
 <220>
 <221> misc_feature
 5 <222> (52)..(52)
 <223> n may be A, U, C, G
 <220>
 <221> misc_feature
 <222> (54)..(54)
 10 <223> n may be A, U, C, G
 <220>
 <221> misc_feature
 <222> (56)..(56)
 <223> n may be A, U, C, G
 15 <220>
 <221> misc_feature
 <222> (60)..(60)
 <223> n may be A, U,C, G
 <220>
 20 <221> misc_feature
 <222> (63)..(63)
 <223> n may be A, U, C, G
 <220>
 <221> misc_feature
 25 <222> (67)..(67)
 <223> n may be A, U, C, G
 <220>
 <221> misc_feature
 <222> (71)..(71)
 30 <223> n may be A, U, C, G
 <220>
 <221> misc_feature
 <222> (72)..(72)
 <223> n may be A, U, C, G
 35 <220>
 <221> misc_feature
 <222> (77)..(77)
 <223> n may be A, U, C, G
 <400> 1
 ucangcugug ncccunnana gggaagcncu uucununguc nnaangaaaa nnangngcun 60
 40 ccnuuungag nnuuacnguu ug 82
 <210> 2
 <211> 20
 <212> RNA
 <213> Homo sapiens
 45 <400> 2
 uaaggcacgc ggugaaugcc 20
 <210> 3
 <211> 22
 <212> RNA
 50 <213> Homo Sapiens
 <400> 3
 cguguucaca gcgaccuug au 22
 <210> 4
 <211> 23
 55 <212> RNA
 <213> Homo sapiens
 <400> 4
 aacauucaac gcugucggug agu 23
 <210> 5
 60 <211> 22
 <212> RNA
 <213> Homo sapiens

	<400> 5	
	accaucgacc guugauugua cc	22
	<210> 6	
	<211> 22	
5	<212> RNA	
	<213> Homo sapiens	
	<400> 6	
	uaacacuguc ugguaaagau gg	22
	<210> 7	
10	<211> 22	
	<212> RNA	
	<213> Homo sapiens	
	<400> 7	
	caucuuccag uacaguguug ga	22
15	<210> 8	
	<211> 22	
	<212> RNA	
	<213> Hom Sapiens	
	<400> 8	
20	uaacacuguc ugguaacgau gu	22
	<210> 9	
	<211> 22	
	<212> RNA	
	<213> Homo sapiens	
25	<400> 9	
	caucuuaccg gacagugcug ga	22
	<210> 10	
	<211> 22	
	<212> RNA	
30	<213> Homo sapiens	
	<400> 10	
	uaauacugcc ugguaaugau ga	22
	<210> 11	
	<211> 22	
35	<212> RNA	
	<213> Homo sapiens	
	<400> 11	
	caucuuacug ggcagcauug ga	22
	<210> 12	
40	<211> 23	
	<212> RNA	
	<213> Homo sapiens	
	<400> 12	
	uaauacugcc ggguaaugau gga	23
45	<210> 13	
	<211> 22	
	<212> RNA	
	<213> Homo sapiens	
	<400> 13	
50	cgucuuacc agcaguguuu gg	22
	<210> 14	
	<211> 22	
	<212> RNA	
	<213> Homo sapiens	
55	<400> 14	
	uccuucauuc caccggaguc ug	22
	<210> 15	
	<211> 21	
	<212> RNA	
60	<213> Homo sapiens	
	<400> 15	
	gauuucagug gagugaaguu c	21
	<210> 16	
	<211> 22	
65	<212> RNA	
	<213> Homo sapiens	

ES 2 631 458 T3

	<400> 16						
	uauuacuguc	ugguaaaacc	gu	22			
	<210> 17						
	<211> 22						
5	<212> RNA						
	<213> Homo sapiens						
	<400> 17						
	uggaauguaa	ggaagugugu	gg	22			
	<210> 18						
10	<211> 22						
	<212> RNA						
	<213> Homo sapiens						
	<400> 18						
	caaagcgcuc	ccuuuagag	gu	22			
15	<210> 19						
	<211> 22						
	<212> RNA						
	<213> Homo sapiens						
	<400> 19						
20	aagugcuucc	uuuuagaggg	uu	22			
	<210> 20						
	<211> 22						
	<212> RNA						
	<213> Homo sapiens						
25	<400> 20						
	cuacaaaggg	aagcacuuuc	uc	22			
	<210> 21						
	<211> 21						
	<212> RNA						
30	<213> Homo sapiens						
	<400> 21						
	gaaggcgcuu	ccuuuggag	u	21			
	<210> 22						
	<211> 85						
35	<212> RNA						
	<213> homo sapiens						
	<400> 22						
	agggccucucu	cuccguguuc	acagcggacc	uugauuuaaa	uguccauaca	auuaaggcac	60
	gctggugaaug	ccaagaaugg	ggcug				85
	<210> 23						
40	<211> 109						
	<212> RNA						
	<213> homo sapiens						
	<400> 23						
	aucaagauua	gaggcucugc	ucuccguguu	cacagcggac	cuugauuuua	ugucauacaa	60
	uuuaggcacg	cggugaaugc	caagagcggg	gccuacggcu	gcacuugaa		109
45	<210> 24						
	<211> 87						
	<212> RNA						
	<213> homo sapiens						
	<400> 24						
	ugagggcccc	ucugcguguu	cacagcggac	cuugauuuua	ugucuauaca	auuaaggcac	60
	gctggugaaug	ccaagagagg	cgccucc				87
50	<210> 25						
	<211> 110						
	<212> RNA						
	<213> homo sapiens						
55	<400> 25						
	ugaguuuuga	gguugcuuca	gugaacauuc	aacgcugucg	gugaguuuug	aauuuuuuuc	60
	aaaaccaucg	accguugauu	guaccuauug	gcuaaccauc	aucuacucca		110

ES 2 631 458 T3

	<210> 26		
	<211> 95		
	<212> RNA		
	<213> homo sapiens		
5	<400> 26	cggccggccc ugguccauc uuccaguaca guguuggaug gucuaauugu gaagcuccua	60
		acacugucug guaaagaugg cucccgggug gguuc	95
	<210> 27		
	<211> 90		
	<212> RNA		
10	<213> homo sapiens		
	<400> 27	ccgggccccu gugagcaucu uaccggacag ugcuggauuu cccagcuuga cucuaacacu	60
		gucugguaac gauguucaa ggugacccgc	90
	<210> 28		
	<211> 68		
15	<212> RNA		
	<213> homo sapiens		
	<400> 28	cccucgucuu acccagcagu guuugggugc gguugggagu cucuaauacu gccggguaau	60
		gauggagg	68
	<210> 29		
20	<211> 110		
	<212> RNA		
	<213> homo sapiens		
	<400> 29	aaagauccuc agacaaucca ugugcuucuc uuguccuua uuccaccgga gucugucuca	60
		uacccaacca gauuucagug gagugaagu caggaggcau ggagcugaca	110
25	<210> 30		
	<211> 70		
	<212> RNA		
	<213> homo sapiens		
	<400> 30	gcgucuuacc agacaugguu agaccuggcc cucugucuaa uacugucugg uaaaaccguc	60
30		cauccgcugc	70
	<210> 31		
	<211> 86		
	<212> RNA		
	<213> homo sapiens		
35	<400> 31	ugcuucccga ggccacaugc uucuuuauau ccccauagg auuacuuugc uauggaangu	60
		aaggaagugu gugguuucgg caagug	86
	<210> 32		
	<211> 83		
	<212> RNA		
40	<213> homo sapiens		
	<400> 32	ucaugcugug gccuccaga ggaagcgc uucuguuguc ugaaagaaa caagcgcuc	60
		ccuuuagag guuuacgguu uga	83
	<210> 33		
	<211> 87		
45	<212> RNA		
	<213> homo sapiens		

ES 2 631 458 T3

	<400> 33		
	ucucagggcug ugacccucua aaggggaagcg cuuucugugg ucagaaagaa aagcaagugc	60	
	uuccuuuuag aggguuaccg uuuggga	87	
	<210> 34		
	<211> 87		
5	<212> RNA		
	<213> homo sapiens		
	<400> 34		
	ucucaugcug ugacccuaca aaggggaagca cuuucucuug uccaaaggaa aagaaggcgc	60	
	uucccuuugg aguguuacgg uuugaga	87	
	<210> 35		
10	<211> 95		
	<212> RNA		
	<213> homo sapiens		
	<400> 35		
	ccagcucggg cagccguggc caucuacug ggcagcauug gauggaguca ggucucuaau	60	
	acugccuggu aaugaugacg gcggagcccu gcacg	95	
15	<210> 36		
	<211> 180		
	<212> DNA		
	<213> artificial DNA		
	<220>		
20	<223> cloned sequence		
	<400> 36		
	tttctttcac ctttccttcc ttccttctc ctttccttcc tcaggagaaa ggcctctctc	60	
	tccgtgttca cagcggacct tgatttaaat gtccatacaa ttaaggcacg cggatgaatgc	120	
	caagaatggg gctggctgag caccgtgggt cggcgagggc ccgccaagga aggagcgacc	180	
	<210> 37		
	<211> 205		
25	<212> DNA		
	<213> artificial DNA		
	<220>		
	<223> cloned sequence		
	<400> 37		
	gttgtttctg tctcccatcc ccttcagata cttacagata ctgtaaagtg agtagaattc	60	
	tgagttttga ggttgcttca gtgaacattc aacgctgtcg gtgagtttg aattaaaatc	120	
	aaaaccatcg accgttgatt gtaccctatg gctaaccatc atctactcca tggatgctcag	180	
30	aattcgctga agacaggaaa ccaaa	205	
	<210> 38		
	<211> 200		
	<212> DNA		
	<213> artificial DNA		
35	<220>		
	<223> cloned sequence		
	<400> 38		
	cctgtagcaa ctggtgagcg cgcaccgtag ttctctgtcg gccggccctg ggtccatctt	60	
	ccagtacagt gttggatggt ctaattgtga agctcctaac actgtctggt aaagatggct	120	
40	cccgggtggg ttctctcggc agtaaccttc agggagccct gaagaccatg gaggactact	180	
	gaccaacaac ctctgacctt	200	

ES 2 631 458 T3

<210> 39
 <211> 191
 <212> DNA
 <213> artificial DNA
 5 <220>
 <223> cloned sequence
 <400> 39
 gttcttccct gggcttccac agcagccoct gcctgcctgg cgggacccca cgtccctccc 60
 gggcccctgt gagcatctta ccggacagtg ctggatttcc cagcttgact ctaacactgt 120
 ctggtaacga tgttcaaagg tgacccgccg ctcgccgggg acaccaccga ggcacatccg 180
 gagtcctac t 191
 10 <210> 40
 <211> 140
 <212> DNA
 <213> artificial DNA
 <220>
 <223> cloned sequence
 15 <400> 40
 aagctgcctg acccaaggtg ggcgggctgg gcgggggcc tegtcttacc cagcagtgtt 60
 tgggtgcggt tgggagtctc taatactgcc gggtaatgat ggaggcccct gtccctgtgt 120
 cagcaacatc catcgcctca 140
 <210> 41
 <211> 253
 <212> DNA
 20 <213> artificial DNA
 <220>
 <223> cloned sequence
 <400> 41
 agtgtctaca ggctgaggtt gacatgcatc cccaccctct gagaaaaaga tcctcagaca 60
 atccatgtgc ttctcttgtc cttcattcca ccggagtctg tctcataccc aaccagattt 120
 cagtggagtg aagttcagga ggcattggagc tgacaaccat gaggcctcgg cagccaccgc 180
 caccaccgcc gccgccacca ccgtagcagc agcagcagca gcagcagcag cagcagcagc 240
 agcaagagta act 253
 25 <210> 42
 <211> 175
 <212> DNA
 <213> artificial DNA
 <220>
 30 <223> cloned sequence
 <400> 42
 agacaccagc ccaggaccgg gaggccaccc acaccaccgc cggccgatgg gcgtcttacc 60
 agacatgggt agacctggcc ctctgtctaa tactgtctgg taaaaccgtc catccgctgc 120
 ctgatcaccg ttagaggaga gagctgcctg ccctgcagct catcagtgca aagcc 175
 35 <210> 43
 <211> 202
 <212> DNA
 <213> artificial DNA
 <220>
 <223> cloned sequence

ES 2 631 458 T3

<400> 43
 gcaaggagga aagatgctac aagtggccca cttctgagat gcgggctgct tctggatgac 60
 actgcttccc gaggccacat gcttctttat atccccatat ggattacttt gctatggaat 120
 gtaaggaagt gtgtggtttc ggcaagtgcc tcctcgctgg ccccagggta ccaccggag 180
 cacaggttg gtgaccttct tc 202
 <210> 44
 <211> 251
 5 <212> DNA
 <213> artificial DNA
 <220>
 <223> cloned sequence
 <400> 44
 gcaaacaggg caataaatg catctttatt ttgtgtccat tttaacctgg tcaaggaaaa 60
 ttccaacagc aacatcaaaa aaccagtgtt ggagcaagaa tatgtcatgc tgtggccctc 120
 cagaggggaag cgctttctgt tgtctgaaag aaaacaaagc gctccccttt agaggtttac 180
 ggtttgagta aagcagcgtt gaagttgatg ctgatcttgg taatacattt gcagagcgtg 240
 10 cttatcatca g 251
 <210> 45
 <211> 214
 <212> DNA
 <213> artificial DNA
 15 <220>
 <223> cloned sequence
 <400> 45
 tgtgtccatt taaacctggt caaggaagat tcccacaaaa aatccacggt gctggagcaa 60
 gaggatctca ggctgtgacc ctctaaaggg aagcgctttc tgtggtcaga aagaaaagca 120
 agtgcttctt tttagagggt taccgtttgg gaaaagcaat gttgaagttg atgctgatct 180
 tggtaaaata tttgcagagc gtgcttatca tcag 214
 <210> 46
 20 <211> 253
 <212> DNA
 <213> artificial DNA
 <220>
 <223> cloned sequence
 25 <400> 46
 caaacagggc caataaatgc atcctcattt ttgtgtccat tttaacctgg gcaaggaaaa 60
 ttccaacaaa aaaccagag ttctggagca agaagatctc atgctgtgac cctacaaagg 120
 gaagcacttt ctcttgtcca aaggaaaaga aggcgcttcc ctttgagtg ttacggtttg 180
 agaaaagcag cgttgaagtt gatgcttadc tcggtaatac atttgtagag catgcttadc 240
 atgaggcttg gac 253
 <210> 47
 <211> 135
 <212> DNA
 30 <213> artificial
 <220>
 <223> DNA cloned into lentiviral vector

ES 2 631 458 T3

<400> 47
 cagccgggcg gccccggac ccagctcggg cagccgtggc catcttactg ggcagcattg 60
 gatggagtca ggtctctaata actgcctcggg aatgatgacg gcggagccct gcacgcagcg 120
 accggccgac cccgt 135
 <210> 48
 <211> 18
 5 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 <400> 48
 10 gaaaagagag tggaagtg 18
 <210> 49
 <211> 18
 <212> DNA
 <213> artificial DNA
 15 <220>
 <223> primer
 <400> 49
 gtgaaggag atgtattg 18
 <210> 50
 20 <211> 20
 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 25 <400> 50
 caggtctcct ctggctctg 20
 <210> 51
 <211> 20
 <212> DNA
 30 <213> artificial DNA
 <220>
 <223> primer
 <400> 51
 acttgaatc ggggtcggag 20
 35 <210> 52
 <211> 20
 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 <400> 52
 gaggattagc cggaacaaca 20
 <210> 53
 <211> 20
 45 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 <400> 53
 50 aacaaatttc cccatctcc 20
 <210> 54
 <211> 20
 <212> DNA
 <213> artificial DNA
 55 <220>
 <223> primer
 <400> 54
 aggatctcca ggctcgaag 20
 <210> 55
 60 <211> 20

<212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 5 <400> 55
 gacatctgag tgggtctgga 20
 <210> 56
 <211> 20
 <212> DNA
 10 <213> artificial DNA
 <220>
 <223> primer
 <400> 56
 ttcggacca cacattacct 20
 15 <210> 57
 <211> 20
 <212> DNA
 <213> artificial DNA
 <220>
 20 <223> primer
 <400> 57
 ttggagcagt tttgcactg 20
 <210> 58
 <211> 20
 25 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 <400> 58
 30 atgcggaaga cagaaaatgg 20
 <210> 59
 <211> 20
 <212> DNA
 <213> artificial DNA
 35 <220>
 <223> primer
 <400> 59
 gtcacgttct tccgcttctc 20
 <210> 60
 40 <211> 20
 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 45 <400> 60
 cgcttgacat cactgaagga 20
 <210> 61
 <211> 20
 <212> DNA
 50 <213> artificial DNA
 <220>
 <223> primer
 <400> 61
 55 ctgcccacac tctgtgcatt 20
 <210> 62
 <211> 21
 <212> DNA
 <213> artificial DNA
 <220>
 60 <223> primer
 <400> 62
 agcagagaat ggaaagtcaa a 21
 <210> 63
 <211> 18
 65 <212> DNA
 <213> artificial DNA

<220>
 <223> primer
 <400> 63
 5 tgctgcttac atgtctcg 18
 <210> 64
 <211> 50
 <212> DNA
 <213> artificial DNA
 <220>
 10 <223> primer
 <400> 64
 gtcgtatcca gtcagggtc cgaggatttc gcaactggata cgacccatct 50
 <210> 65
 <211> 50
 15 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 <400> 65
 20 gtcgtatcca gtcagggtc cgaggatttc gcaactggata cgactccatc 50
 <210> 66
 <211> 50
 <212> DNA
 <213> artificial DNA
 25 <220>
 <223> primer
 <400> 66
 gtcgtatcca gtcagggtc cgaggatttc gcaactggata cgacactcac 50
 <210> 67
 30 <211> 50
 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 <400> 67
 35 gtcgtatcca gtcagggtc cgaggatttc gcaactggata cgacggcatt 50
 <210> 68
 <211> 50
 <212> DNA
 40 <213> artificial DNA
 <220>
 <223> primer
 <400> 68
 45 gtcgtatcca gtcagggtc cgaggatttc gcaactggata cgacacctct 50
 <210> 69
 <211> 50
 <212> DNA
 <213> artificial DNA
 <220>
 50 <223> primer
 <400> 69
 gtcgtatcca gtcagggtc cgaggatttc gcaactggata cgacaaccct 50
 <210> 70
 <211> 50
 55 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 <400> 70
 60 gtcgtatcca gtcagggtc cgaggatttc gcaactggata cgacgagaaa 50
 <210> 71
 <211> 50
 <212> DNA
 <213> artificial DNA
 65 <220>
 <223> primer

<400> 71
 gtcgtatcca gtcagggtc cgaggatttc gcaactggata cgacactcca 50
 <210> 72
 <211> 24
 5 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 <400> 72
 10 gcccgctaac actgtctggt aaag 24
 <210> 73
 <211> 24
 <212> DNA
 <213> artificial DNA
 15 <220>
 <223> primer
 <400> 73
 gcccgctaact actgccgggt aatg 24
 <210> 74
 20 <211> 23
 <212> DNA
 <213> artificial
 <220>
 <223> primer
 25 <400> 74
 tgccagaaca ttcaacgctg tcg 23
 <210> 75
 <211> 21
 <212> DNA
 30 <213> artificial DNA
 <220>
 <223> primer
 <400> 75
 35 tgccagtaag gcacgcggtg a 21
 <210> 76
 <211> 24
 <212> DNA
 <213> artificial DNA
 <220>
 40 <223> primer
 <400> 76
 tgccagcaaa gcgctcccct ttag 24
 <210> 77
 <211> 24
 45 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 <400> 77
 50 gcccgcaagt gcttctttt agag 24
 <210> 78
 <211> 23
 <212> DNA
 <213> artificial DNA
 55 <220>
 <223> primer
 <400> 78
 gcccgcttac aaaggaagc act 23
 <210> 79
 60 <211> 23
 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 65 <400> 79
 tgccaggaag gcgcttcct ttg 23

ES 2 631 458 T3

<210> 80
 <211> 16
 <212> DNA
 <213> artificial DNA
 5 <220>
 <223> primer
 <400> 80
 gtgcagggtc cgaggt 16
 <210> 81
 10 <211> 16
 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 15 <400> 81
 gtcatccttg cgcagg 16
 <210> 82
 <211> 20
 <212> DNA
 20 <213> artificial DNA
 <220>
 <223> primer
 <400> 82
 cgcttcggca gcacatac 20
 25 <210> 83
 <211> 20
 <212> DNA
 <213> artificial DNA
 <220>
 30 <223> primer
 <400> 83
 aggggccatg ctaatctt 20
 <210> 84
 <211> 23
 35 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 <400> 84
 40 cagcgtggtt tgacctcat aga 23
 <210> 85
 <211> 18
 <212> DNA
 <213> artificial DNA
 45 <220>
 <223> primer
 <400> 85
 cactgacggg caccggag 18
 <210> 86
 <211> 28
 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 55 <400> 86
 gacctcata gaagattcta gagctagc 28
 <210> 87
 <211> 7
 <212> RNA
 60 <213> artificial
 <220>
 <223> seed sequence
 <400> 87
 aaggcac 7
 65 <210> 88
 <211> 7

<212> RNA
 <213> artificial
 <220>
 <223> seed sequence
 5 <400> 88
 guguuca 7
 <210> 89
 <211> 7
 <212> RNA
 10 <213> artificial
 <220>
 <223> seed sequence
 <400> 89
 acauuca 7
 15 <210> 90
 <211> 7
 <212> RNA
 <213> artificial
 <220>
 20 <223> seed sequence
 <400> 90
 ccaucga 7
 <210> 91
 <211> 7
 25 <212> RNA
 <213> artificial
 <220>
 <223> seed sequence
 <400> 91
 30 aacacug 7
 <210> 92
 <211> 7
 <212> RNA
 <213> artificial
 35 <220>
 <223> sed sequence
 <400> 92
 aucuucc 7
 <210> 93
 40 <211> 7
 <212> RNA
 <213> artificial
 <220>
 <223> seed sequence
 45 <400> 93
 aacacug 7
 <210> 94
 <211> 7
 <212> RNA
 50 <213> artificial
 <220>
 <223> seed sequence
 <400> 94
 aucuuac 7
 55 <210> 95
 <211> 7
 <212> RNA
 <213> artificial
 <220>
 60 <223> seed sequence
 <400> 95
 aauacug 7
 <210> 96
 <211> 7
 65 <212> RNA
 <213> artificial

<220>
 <223> seed sequence
 <400> 96
 5 aucuac 7
 <210> 97
 <211> 7
 <212> RNA
 <213> artificial
 <220>
 10 <223> seed sequence
 <400> 97
 aauacug 7
 <210> 98
 <211> 7
 15 <212> RNA
 <213> artificial
 <220>
 <223> seed sequence
 <400> 98
 20 gucuac 7
 <210> 99
 <211> 7
 <212> RNA
 <213> artificial
 25 <220>
 <223> seed sequence
 <400> 99
 ccucau 7
 <210> 100
 30 <211> 7
 <212> RNA
 <213> artificial
 <220>
 <223> seed sequence
 <400> 100
 35 auuucag 7
 <210> 101
 <211> 7
 <212> RNA
 40 <213> artificial
 <220>
 <223> seed sequence
 <400> 101
 aauacug 7
 45 <210> 102
 <211> 7
 <212> RNA
 <213> artificial
 <220>
 50 <223> seed sequence
 <400> 102
 ggaaugu 7
 <210> 103
 <211> 7
 55 <212> RNA
 <213> artificial
 <220>
 <223> seed sequence
 <400> 103
 60 aaagcgc 7
 <210> 104
 <211> 7
 <212> RNA
 <213> artificial
 65 <220>
 <223> seed sequence

ES 2 631 458 T3

<400> 104
 agugcuu 7
 <210> 105
 <211> 7
 5 <212> RNA
 <213> artificial
 <220>
 <223> seed sequence
 <400> 105
 10 aagugcu 7
 <210> 106
 <211> 7
 <212> RNA
 <213> artificial
 15 <220>
 <223> seed sequence
 <400> 106
 uacaaag 7
 <210> 107
 20 <211> 7
 <212> RNA
 <213> artificial
 <220>
 <223> seed sequence
 25 <400> 107
 aaggcgc 7
 <210> 108
 <211> 25
 <212> RNA
 30 <213> macaca mulatta
 <400> 108
 aaagugcuuc cuuuuagagg guuac 25
 <210> 109
 <211> 85
 35 <212> RNA
 <213> macaca mulatta
 <400> 109
cucaggcugu gaccucua aggggaagcgc uuucuguggu cugaaagaaa agaaagugcu 60

uccuuuuaga ggguuaccgu uugag 85
 <210> 110
 40 <211> 23
 <212> RNA
 <213> macaca mulatta
 <400> 110
 aaagugcuuc cuuuuagagg guu 23
 45 <210> 111
 <211> 87
 <212> RNA
 <213> macaca mulatta
 <400> 111
ucucaggcug ugaccucua gaggggaagcgc cuuucugugg ucugaaagaa aagaaagugc 60

 50 **uuccuuuuag aggguuaccg uuugaga 87**
 <210> 112
 <211> 25
 <212> DNA
 <213> artificial DNA
 55 <220>
 <223> primer
 <400> 112
 ctcaactta actggaaga atgtc 25
 <210> 113
 60 <211> 19

<212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 5 <400> 113
 tcctttcac cagcaagct 19
 <210> 114
 <211> 20
 <212> DNA
 10 <213> artificial DNA
 <220>
 <223> primer
 <400> 114
 ggctcagatt caggaacagc 20
 15 <210> 115
 <211> 20
 <212> DNA
 <213> artificial DNA
 <220>
 20 <223> primer
 <400> 115
 gcttcaacgg caaagtctc 20
 <210> 116
 <211> 20
 25 <212> DNA
 <213> artificial DNA
 <220>
 <223> primer
 <400> 116
 30 ggctggaac cagttctcg 20
 <210> 117
 <211> 20
 <212> DNA
 <213> artificial DNA
 35 <220>
 <223> primer
 <400> 117
 ggcatgaatg tccctgatt 20
 <210> 118
 40 <211> 21
 <212> RNA
 <213> isomiR
 <400> 118
 aagugcuucc uuuuagaggg u 21
 45 <210> 119
 <211> 22
 <212> RNA
 <213> isomiR
 <400> 119
 50 caagugcuuc cuuuuagagg gu 22
 <210> 120
 <211> 22
 <212> RNA
 <213> isomiR
 55 <400> 120
 aagugcuucc uuuuagaggg uu 22
 <210> 121
 <211> 22
 <212> RNA
 60 <213> isomiR
 <400> 121
 caaagcguc cccuuuagag gu 22
 <210> 122
 <211> 21
 65 <212> RNA
 <213> isomiR

	<400> 122	
	caaagcguc cccuuagag g	21
	<210> 123	
	<211> 20	
5	<212> RNA	
	<213> isomiR	
	<400> 123	
	caaagcguc cccuuagag	20
	<210> 124	
10	<211> 22	
	<212> RNA	
	<213> isomiR	
	<400> 124	
	gaaggcgcu cccuuaggag ug	22
15	<210> 125	
	<211> 21	
	<212> RNA	
	<213> isomiR	
	<400> 125	
20	gaaggcgcu cccuuaggag u	21
	<210> 126	
	<211> 21	
	<212> RNA	
	<213> isomiR	
25	<400> 126	
	cuacaaagg aagcacuuc u	21
	<210> 127	
	<211> 22	
	<212> RNA	
30	<213> isomiR	
	<400> 127	
	cuacaaagg aagcacuuc uc	22
	<210> 128	
	<211> 20	
35	<212> RNA	
	<213> isomiR	
	<400> 128	
	cuacaaagg aagcacuuc	20
	<210> 129	
40	<211> 22	
	<212> RNA	
	<213> isomiR	
	<400> 129	
	uaaggcacgc ggugaugcc aa	22
45	<210> 130	
	<211> 21	
	<212> RNA	
	<213> isomiR	
	<400> 130	
50	uaaggcacgc ggugaugcc a	21
	<210> 131	
	<211> 19	
	<212> RNA	
	<213> isomiR	
55	<400> 131	
	uaaggcacgc ggugaugc	19
	<210> 132	
	<211> 20	
	<212> RNA	
60	<213> isomiR	
	<400> 132	
	uaaggcacgc ggugaugcc	20
	<210> 133	
	<211> 22	
65	<212> RNA	
	<213> isomiR	

	<400> 133	
	uggaauguaa ggaagugugu gg	22
	<210> 134	
	<211> 23	
5	<212> RNA	
	<213> isomiR	
	<400> 134	
	uggaauguaa ggaagugugu ggu	23
	<210> 135	
10	<211> 21	
	<212> RNA	
	<213> isomiR	
	<400> 135	
	uggaauguaa ggaagugugu g	21
15	<210> 136	
	<211> 25	
	<212> RNA	
	<213> isomiR	
	<400> 136	
20	aacauucaac gcugucggug aguuu	25
	<210> 137	
	<211> 23	
	<212> RNA	
	<213> isomiR	
25	<400> 137	
	aacauucaac gcugucggug agu	23
	<210> 138	
	<211> 22	
	<212> RNA	
30	<213> isomiR	
	<400> 138	
	aacauucaac gcugucggug ag	22
	<210> 139	
	<211> 24	
35	<212> RNA	
	<213> isomiR	
	<400> 139	
	aacauucaac gcugucggug aguu	24
40	<210> 140	
	<211> 18	
	<212> RNA	
	<213> isomiR	
	<400> 140	
45	aacauucaac gcugucgg	18
	<210> 141	
	<211> 19	
	<212> RNA	
	<213> isomiR	
	<400> 141	
50	aacauucaac gcugucggu	19
	<210> 142	
	<211> 20	
	<212> RNA	
	<213> isomiR	
55	<400> 142	
	aacauucaac gcugucggug	20
	<210> 143	
	<211> 21	
	<212> RNA	
60	<213> isomiR	
	<400> 143	
	aacauucaac gcugucggug a	21
	<210> 144	
	<211> 22	
65	<212> RNA	
	<213> isomiR	

	<400> 144	
	uaacacuguc ugguaaagau gg	22
	<210> 145	
	<211> 21	
5	<212> RNA	
	<213> isomiR	
	<400> 145	
	uaacacuguc ugguaaagau g	21
	<210> 146	
10	<211> 23	
	<212> RNA	
	<213> isomiR	
	<400> 146	
	uaacacuguc ugguaaagau ggc	23
15	<210> 147	
	<211> 20	
	<212> RNA	
	<213> isomiR	
	<400> 147	
20	uaacacuguc ugguaaagau	20
	<210> 148	
	<211> 19	
	<212> RNA	
	<213> isomiR	
25	<400> 148	
	uaacacuguc ugguaaaga	19
	<210> 149	
	<211> 24	
	<212> RNA	
30	<213> isomiR	
	<400> 149	
	uaacacuguc ugguaaagau ggcu	24
	<210> 150	
	<211> 23	
35	<212> RNA	
	<213> isomiR	
	<400> 150	
	uaacacuguc ugguaacgau guu	23
40	<210> 151	
	<211> 22	
	<212> RNA	
	<213> isomiR	
	<400> 151	
45	uaacacuguc ugguaacgau gu	22
	<210> 152	
	<211> 23	
	<212> RNA	
	<213> isomiR	
	<400> 152	
50	uaauacugcc ugguaaugau gac	23
	<210> 153	
	<211> 22	
	<212> RNA	
	<213> isomiR	
55	<400> 153	
	uaauacugcc ugguaaugau ga	22
	<210> 154	
	<211> 22	
	<212> RNA	
60	<213> isomiR	
	<400> 154	
	caucuacug ggcagcauug ga	22
	<210> 155	
	<211> 21	
65	<212> RNA	
	<213> isomiR	

ES 2 631 458 T3

<400> 155
 caucuuacug ggcagcauug g 21
 <210> 156
 <211> 23
 5 <212> RNA
 <213> isomiR
 <400> 156
 uaaucugcc ggguaaugau gga 23
 <210> 157
 10 <211> 22
 <212> RNA
 <213> isomiR
 <400> 157
 uaaucugcc ggguaaugau gg 22
 15 <210> 158
 <211> 21
 <212> RNA
 <213> isomiR
 <400> 158
 20 uccuucuuac caccggaguc u 21
 <210> 159
 <211> 22
 <212> RNA
 <213> isomiR
 25 <400> 159
 uccuucuuac caccggaguc ug 22
 <210> 160
 <211> 23
 <212> RNA
 30 <213> isomiR
 <400> 160
 uccuucuuac caccggaguc ugu 23
 <210> 161
 <211> 22
 <212> RNA
 35 <213> isomiR
 <400> 161
 uaaucuguc ugguaaaacc gu 22
 <210> 162
 40 <211> 21
 <212> RNA
 <213> isomiR
 <400> 162
 uaaucuguc ugguaaaacc g 21
 45 <210> 163
 <211> 24
 <212> RNA
 <213> artificial
 <220>
 50 <223> mimic
 <400> 163
 ucacaaccuc cuagaagag uaga 24

REIVINDICACIONES

- 5 1. Molécula de ARNmi o un equivalente o un mimético o un isomiR de la misma donde una fuente de dicha molécula de ARNmi o equivalente de la misma comprende al menos 80 nucleótidos y comprende una unidad con al menos 98% de identidad con SEQ ID NO: 1 o una fuente de la misma, o una composición que comprende dicha molécula de ARNmi, equivalente o mimético o isomiR de la misma o fuente de la misma para su uso como un medicamento para tratar, revertir, curar y/o retrasar un cáncer asociado a la EMT induciendo un proceso de MET, donde dicha molécula de ARNmi es un ARNmi-518b, ARNmi-520f y/o ARNmi-524.
- 10 2. Molécula de ARNmi, un equivalente o un mimético o un isomiR de la misma o una fuente de la misma o la composición que comprende dicha molécula de ARNmi, equivalente o fuente de la misma según la reivindicación 1 para su uso según la reivindicación 1, donde el cáncer asociado a la EMT es un carcinoma.
- 15 3. Molécula de ARNmi, un equivalente o un mimético o un isomiR de la misma o una fuente de la misma o la composición que comprende dicha molécula de ARNmi, equivalente o fuente de la misma según la reivindicación 2 para su uso según la reivindicación 1 o 2, donde el carcinoma es un cáncer de vejiga o próstata.
- 20 4. Composición según cualquiera de las reivindicaciones 1 a 3 para el uso según cualquiera de las reivindicaciones 1 a 3, que comprende además otra molécula de ARNmi, un mimético o isomiR o fuente de la misma seleccionado de: al menos uno de ARNmi-124-1, ARNmi-206, ARNmi-181a-1, ARNmi-141, ARNmi-200a, ARNmi-200b, ARNmi-200c, ARNmi-429 y ARNmi-205 y/o una fuente de la misma.
- 25 5. Método para la identificación de una sustancia capaz de tratar, revertir, curar y/o retrasar un cáncer asociado a la EMT en una célula mediante la inducción de un proceso de MET, el método comprende las etapas de:
- 30 (a) proporcionar una población de células de prueba capaces de expresar una molécula de ARNmi o equivalente o un isomiR de la misma tal y como se define en la reivindicación 1, preferiblemente, la población de prueba comprende células de vejiga, más preferiblemente la población de células de prueba comprende células de mamíferos, aún más preferiblemente células humanas;
- 35 (b) poner en contacto o incubar la población de células de prueba con la sustancia;
- (c) determinar el nivel de expresión de dicha molécula de ARNmi o equivalente o isomiR de la misma o la actividad o el nivel de estado estable de dicha molécula de ARNmi o equivalente o isomiR de la misma en la población de células de prueba puesta en contacto o incubada con la sustancia;
- 40 (d) comparar la expresión, la actividad o el nivel de estado estable determinado en (c) con la expresión, la actividad o el nivel de estado estable de dicha molécula o equivalente o isomiR en una población de células de prueba que no está en contacto con la sustancia; e,
- 45 (e) identificar una sustancia que produce una diferencia en el nivel de expresión, la actividad o el nivel de estado estable de dicha molécula de ARNmi o equivalente o isomiR, entre la población de células de prueba que está en contacto con la sustancia y la población de células de prueba que no está en contacto con la sustancia.
- 50 6. Método según la reivindicación 5, en el cual los niveles de expresión, las actividades o los niveles de estado estable de al menos otra molécula de ARNmi, o isomiR o fuente de la misma se comparan, preferiblemente donde la otra molécula de ARNmi, o isomiR o fuente de la misma se selecciona de: al menos uno de ARNmi-124-1, ARNmi-206, ARNmi-181a-1, ARNmi-141, ARNmi-200a, ARNmi-200b, ARNmi-200c, ARNmi-429 y ARNmi-205 y/o un isomiR y/o una fuente del mismo.

Fig 1a

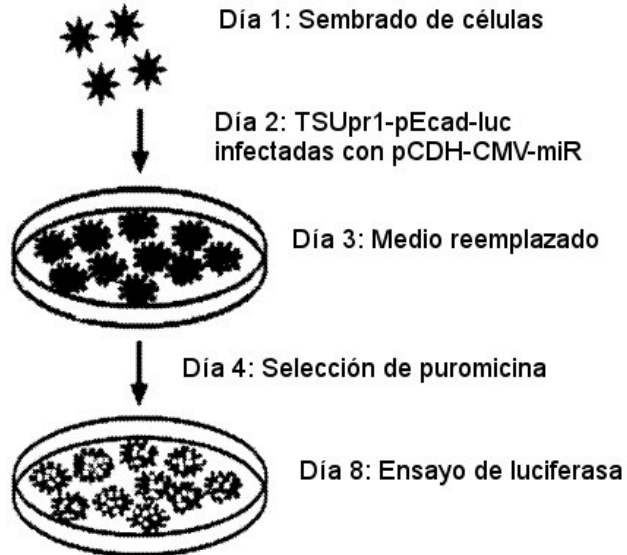


Fig 1b

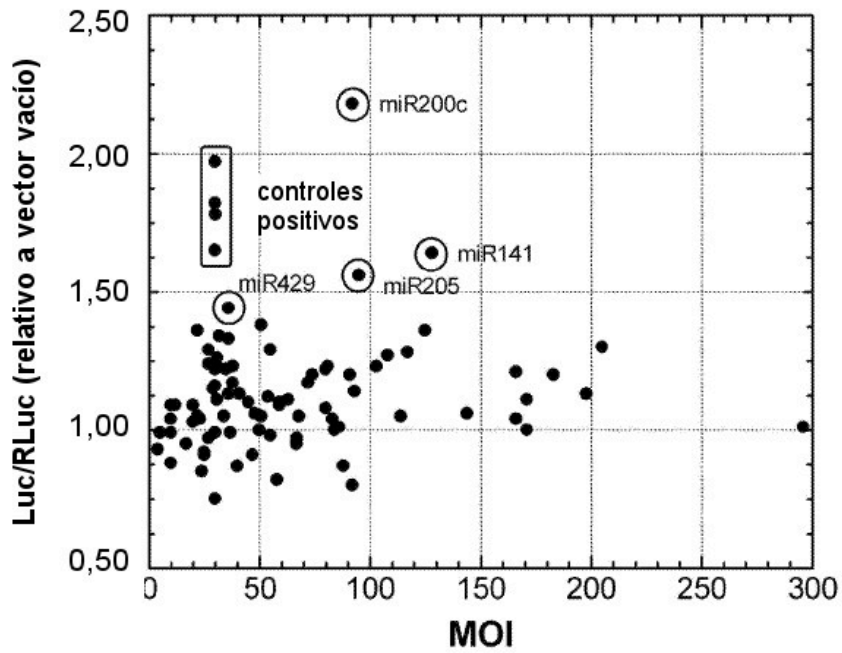


Fig 2

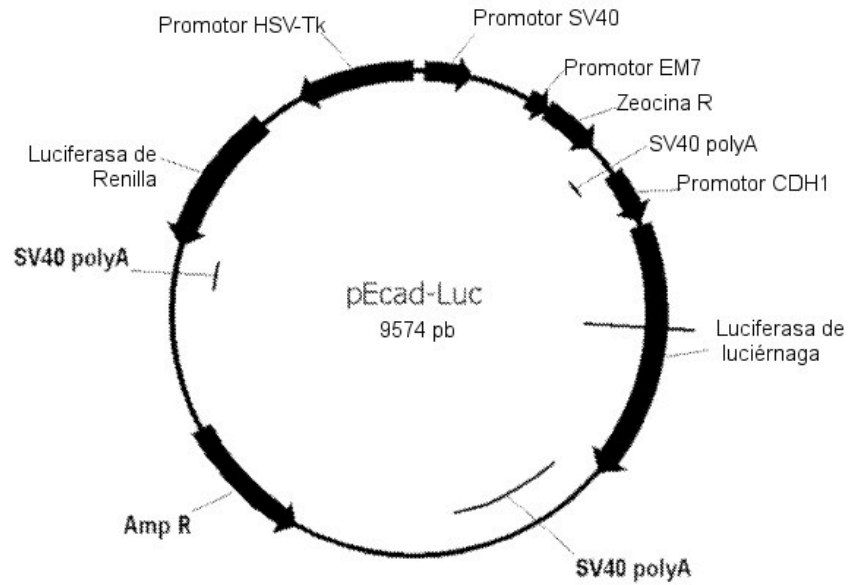


Fig 3

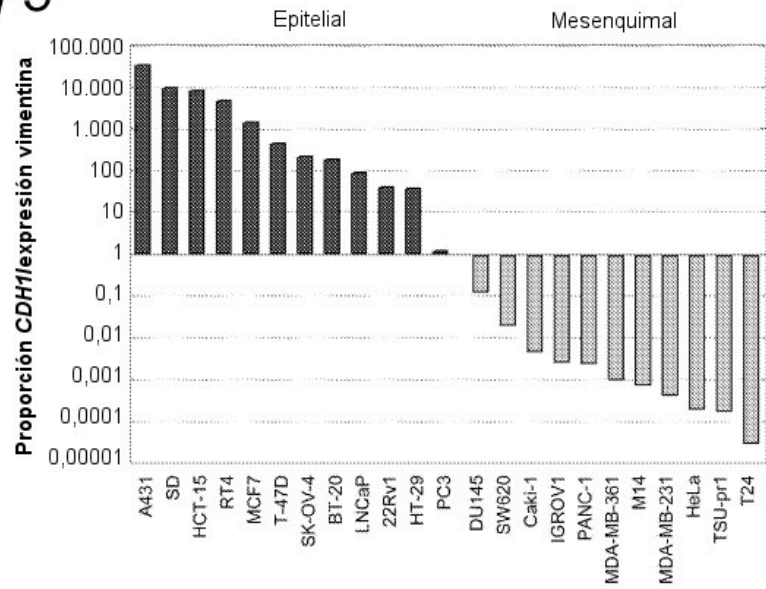


Fig 4a

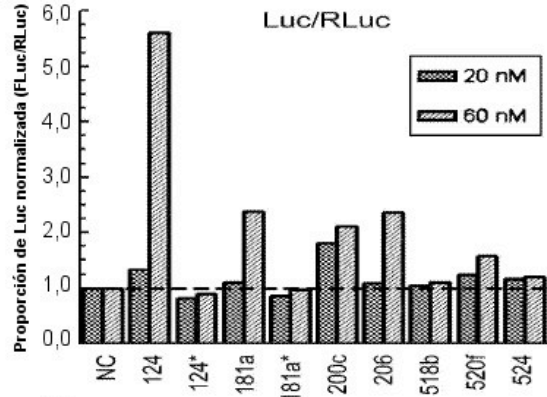


Fig 4b

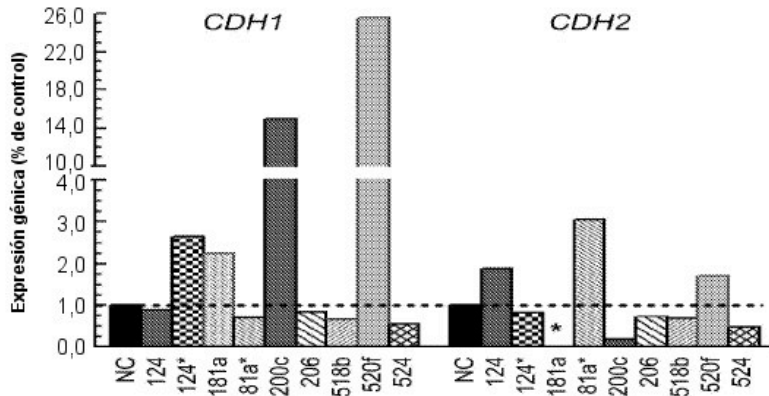


Fig 4c

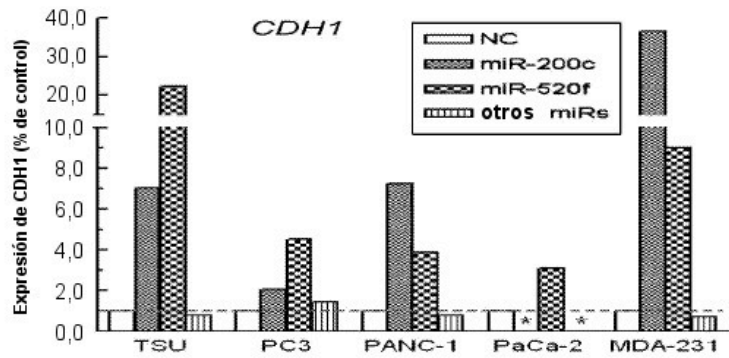


Fig 5a

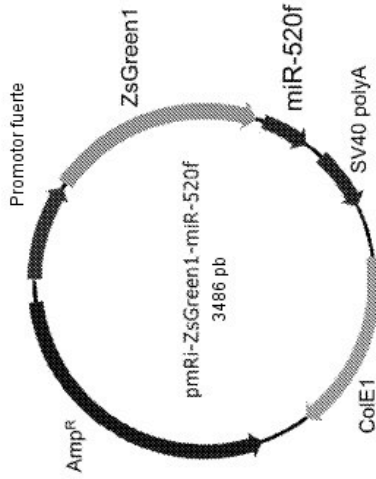


Fig 5b

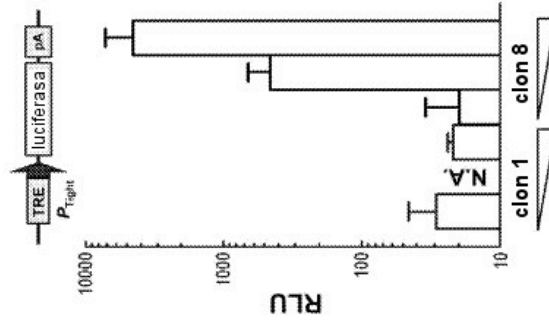


Fig 5c

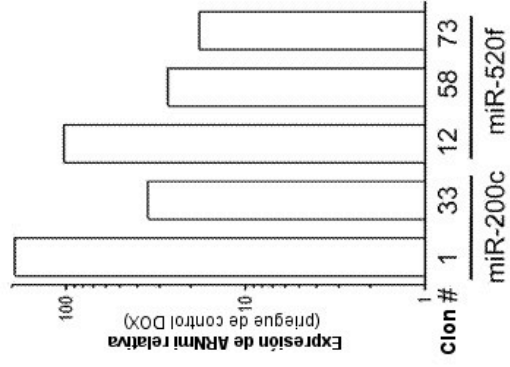


Fig 6b

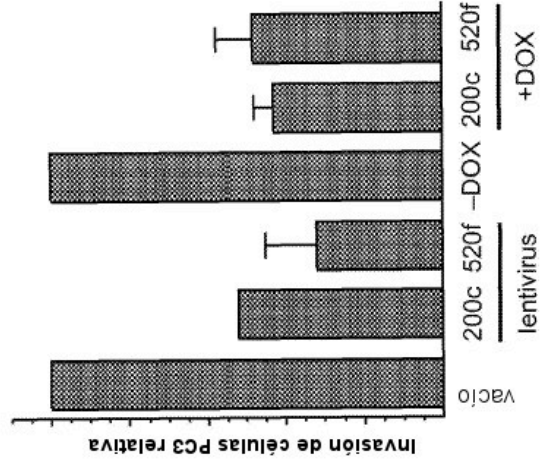


Fig 6a

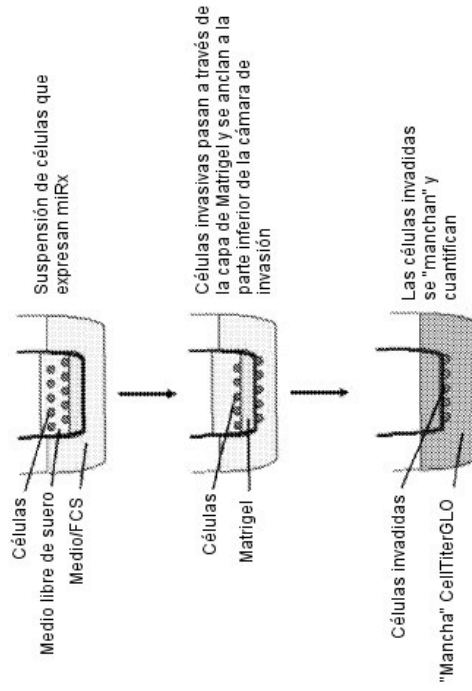


Fig 6d

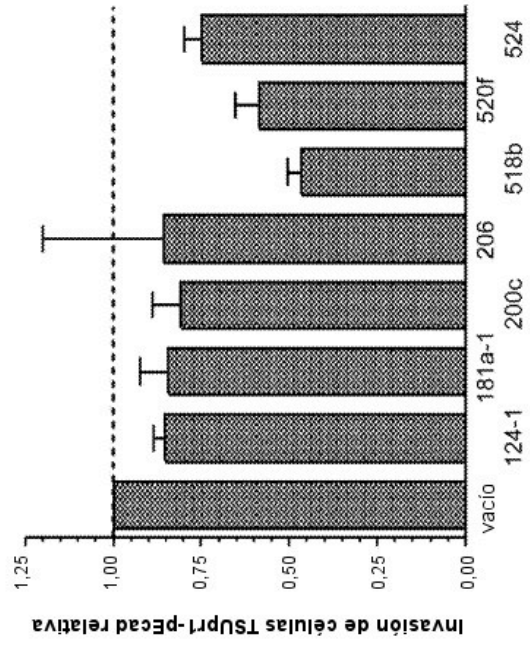


Fig 6c

