

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 699 630**

51 Int. Cl.:

C12N 15/113 (2010.01)

C07H 21/02 (2006.01)

C12N 15/63 (2006.01)

A01K 67/027 (2006.01)

A61K 48/00 (2006.01)

A61K 31/7105 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **22.04.2011** **PCT/US2011/033615**

87 Fecha y número de publicación internacional: **27.10.2011** **WO11133889**

96 Fecha de presentación y número de la solicitud europea: **22.04.2011** **E 11772783 (4)**

97 Fecha y número de publicación de la concesión europea: **19.09.2018** **EP 2561078**

54 Título: **Novedosos ARNhc diseñados estructuralmente**

30 Prioridad:

23.04.2010 US 327510 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
12.02.2019

73 Titular/es:

COLD SPRING HARBOR LABORATORY (100.0%)
1 Bungtown Road
Cold Spring Harbor, NY 11724, US

72 Inventor/es:

HANNON, GREGORY, J. y
CHELOUFI, SIHEM

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 699 630 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

Novedosos ARNhc diseñados estructuralmente

La presente solicitud reivindica la prioridad de la Solicitud Provisional de los Estados Unidos 61/327.510, presentada el 23 de abril de 2010.

5 Esta divulgación de patente contiene material que está sujeto a la protección de derechos de autor, el titular de los derechos de autor no tiene ninguna objeción a la reproducción por fax por cualquier persona del documento de patente o la divulgación de patente, tal como aparece en el archivo o registro de patentes de la Oficina de Marcas y Patentes de EE. UU., pero se reserva cualquier y todos los derechos de autor.

10 A lo largo de esta solicitud, se citan solicitudes de patentes, solicitudes de patentes publicadas, patentes emitidas y otorgadas, textos y referencias bibliográficas.

1. Antecedentes de la invención

La presente invención se refiere en parte a mejoras dirigidas a la utilización de la tecnología de interferencia por ARN (iARN) que explota una ruta de biogénesis de ARN pequeño recién identificada.

15 La tecnología tradicional de iARN en mamíferos aprovecha la ruta canónica de los microARN (miARN). Comenzando con los ARN precursores transcritos de Pol 11, la ruta de la biogénesis implica dos pasos: DROSHA/DGCR8 divide el transcrito precursor en un ARN de horquilla corta que se exporta al citoplasma y luego se procesa con la enzima ARNasaIII DICER para producir ARN dúplex pequeño maduro (21-22 nt), que luego se carga en una de las cuatro proteínas Argonauta (AGO1 hasta AGO4) para formar un complejo de silenciamiento inducido por ARN activo (RISC, de sus siglas en inglés). La ruta de iARN endógena convencional, por lo tanto, comprende tres intermediarios de ARN: un transcrito de miARN primario largo, en gran parte monocatenario (ARNmpri); un transcrito de miARN precursor que tiene una estructura de tallo y bucle y procede del ARNmpri (pre-miARN); y un miARN maduro.

25 Las proteínas argonautas son los efectores clave de las rutas reguladoras mediadas por ARN pequeño que modulan la expresión génica, regulan la estructura y función de los cromosomas y proporcionan una defensa inmune innata contra virus y transposones (Hutvagner, G. & Simard, M. J. Nat Rev Mol Cell Biol 9, 22-32 (2008)). La estructura de las proteínas Ago está bien conservada, y consiste en un dominio amino-terminal, el dominio medio y sus dominios distintivos PAZ y Piwi. Las relaciones estructura-función en esta familia son cada vez más comprendidas (Joshua-Tor, L. Cold Spring Harb Symp Quant Biol 71, 67-72 (2006)). Los dominios PAZ y Mid ayudan a anclar la guía de ARN pequeños, ya que PAZ se une al extremo 3' utilizando una serie de restos aromáticos conservados y el dominio Mid proporciona una bolsa de unión al extremo 5'. El dominio Piwi contiene un motivo ARNasa H que era críptico en la secuencia primaria pero fácilmente reconocible en la estructura terciaria. La carga de una diana altamente complementaria en un Ago lleva el fosfato escindible, los nucleótidos opuestos 10 y 11 de la guía de ARN pequeños, al sitio activo de la enzima, lo que permite la escisión del ARN para dejar los extremos 5' P y 3' OH (Elbashir, S. y col. Genes Dev 15, 188-200 (2001), Elbashir, S. M., y col. EMBO J 20, 6877-88 (2001), Yuan, Y. R. y col. Mol Cell 19, 405-19 (2005), Martinez, J. & Tuschl, T. Genes Dev 18, 975-80 (2004), Schwarz, D. S., y col. Curr Biol 14, 787-91 (2004)).

35 Las proteínas Ago se pueden dividir en tres clados. El clado Piwi es específico para animales, y forma parte de un elegante sistema inmune innato que controla la actividad de los elementos genéticos móviles (Malone, C. D. & Hannon, G. J. Cell 136, 656-68 (2009)). El clado Wago es específico para gusanos y actúa en una variedad de diferentes procesos biológicos (Yigit, E. y col. Cell 127, 747-57 (2006)). El clado Ago se define por la similitud con Arabidopsis Ago1 (Bohmert, K. y col. EMBO J 17, 170-80 (1998)). Las proteínas del clado Ago se encuentran tanto en plantas como en animales, en las que un hilo unificador es su papel en la regulación de genes. En plantas, algunos miembros de la familia Ago se unen a los microARN y se dirigen así a reconocer y escindir los ARNm diana complementarios (Baumberger, N. y Baulcombe, D. C. Proc Natl Acad Sci U S A 102, 11928-33 (2005), Qi, Y., Denli, A. M. & Hannon, G. J. Mol Cell 19, 421-8 (2005)).

45 Los microARN animales funcionan de manera diferente a sus homólogos de plantas, con casi todas las interacciones diana-microARN que proporcionan una complementariedad insuficiente para orientar adecuadamente el fosfato escindible para la escisión. En este caso, el reconocimiento de dianas se basa principalmente en una secuencia "semilla" que corresponde a los nucleótidos de miARN (Joshua-Tor, L. Cold Spring Harb Symp Quant Biol 71, 67-72 (2006), Malone, C. D. & Hannon, G. J. Cell 136, 656-68 (2009)). Si bien el emparejamiento de la diana con otras partes del miARN puede contribuir al reconocimiento, el emparejamiento con semillas parece ser el factor dominante en la determinación de la regulación (Yekta, S. y col. Science 304, 594-6 (2004)). Unas pocas interacciones extensas entre microARN y diana pueden conducir a la escisión de dianas en mamíferos (Davis, E. y col. Curr Biol 15, 743-9 (2005), Harfe, B. D. y col., Proc Natl Acad Sci U S A 102, 10898-903 (2005)). Sin embargo, todavía no se ha demostrado que ninguno de estos sea crítico para la regulación de las dianas (Sekita, Y. y col. Nat Genet 40, 243-8 (2008), Hornstein, E. y col. Nature 438, 671-4 (2005), Tolia, N. H. & Joshua-Tor, L. Nat Chem Biol 3, 36-43 (2007)).

A pesar del hecho de que los microARN animales regulan las dianas sin escisión mediada por Ago, el centro catalítico de las Argonautas está profundamente conservado. Esto consiste en una tríada DDH catalítica que sirve como un sitio de coordinación de metales (Liu, J. y col. *Science* 305, 1437-41 (2004)). De las cuatro proteínas del clado Ade en mamíferos, solo Ago2 ha conservado tanto el motivo DDH como la actividad endonucleasa demostrable (Rivas, F. V. y col. *Nat Struct Mol Biol* 12, 340-9 (2005), Song, J. y col. *Science* 305, 1434-7 (2004), Azuma-Mukai, A. y col. *Proc Natl Acad Sci U S A* 105, 7964-9 (2008)). Ago1, Ago3 y Ago4 están unidas dentro de un único locus de ~ 190 kb y han perdido competencia catalítica. Un análisis de las células mutantes Ago2 ha indicado que las proteínas codificadas por el locus Ago 1/3/4 pueden soportar el silenciamiento mediado por miARN (Rivas, F. V. y col. *Nat Struct Mol Biol* 12, 340-9 (2005)). Esto deja sin una explicación clara para el mantenimiento de un miembro de la familia Ago catalíticamente competente, ya que los miARN son los socios exclusivos de estas proteínas en casi todos los tipos de células (Babiarz, J. E., Ruby, J. G., Wang, Y., Bartel, D. P. & Blelloch, R., *Genes Dev* 22, 2773-85 (2008); Ender, C. y col. *Mol Cell* 32, 519-28 (2008) Tam, O.C. y col. *Nature*, 453:534-538(2008); Kaneda, M. y col., *Epigenetics Chromatin*, 2:9 (2009)).

2. Sumario de la invención

Cada realización desvelada en el presente documento se contempla como aplicable a cada una de las otras realizaciones desveladas. Por lo tanto, todas las combinaciones de los diversos elementos descritos en el presente documento están dentro del ámbito de la invención.

Se proporciona un diseño mejorado de ARNhc basado en miméticos estructurales de precursores de miR-451. Estos miméticos de ARNhc de miR-451 se canalizan a través de una nueva ruta de biogénesis de ARN pequeño, requieren catálisis por AGO2 y son procesados por Drosha, pero son independientes del procesamiento por DICER. Esta ruta de miARN alimenta elementos activos solo en Ago2 debido a su actividad catalítica única. Estos datos demuestran que esta ruta de biogénesis de ARN pequeño recién identificada puede explotarse in vivo para producir moléculas activas.

La utilización de miméticos de ARNhc de miR-451 proporciona una clara ventaja sobre los ARNhc convencionales en que solo se genera una cadena activa, eliminando así los efectos fuera de la diana que podrían resultar de la incorporación de la cadena con sentido del dúplex en un RISC activo. El diseño de un mimético de ARNhc de miR-451 es muy simple y no requiere la utilización de secuencias de la molécula precursora de miR-451, ya que es solo un mimético estructural. Sin embargo, partes de la secuencia primaria de miARN-451, o de la secuencia primaria de otro miARN puede utilizarse en algunas realizaciones, por ejemplo, en aspectos de la invención relacionados con miméticos de ARNhc primarios. En algunos aspectos de la invención, los miméticos de miR-451 primarios se procesan mediante la etapa drosha del procesamiento de miR-451 mientras se pasa por alto la ruta canónica, por ejemplo, en ciertos ARNhc de la invención que se cargan en Ago2 directamente. En un primer aspecto se proporciona una molécula de ARN de horquilla corta (ARNhc) que comprende

(i) una primera secuencia de 21, 22 o 23 nucleótidos complementaria a una secuencia en un gen diana, que tiene una secuencia diferente a la secuencia madura de miR-451, y

(ii) una segunda secuencia que sigue directamente a la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 o 18 nucleótidos contados desde el extremo 5' de la primera secuencia,

en la que los últimos 3 o los últimos 4 nucleótidos de la primera secuencia forman una región de bucle en el ARNhc.

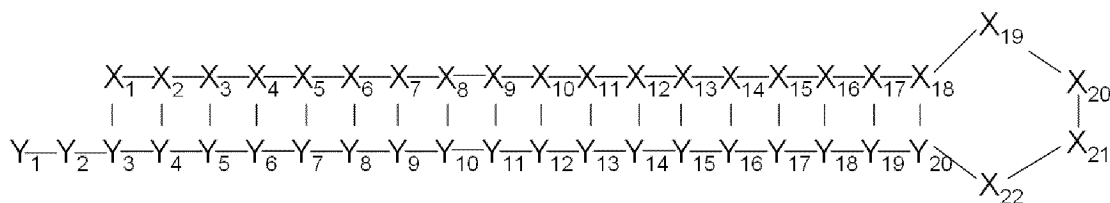
De acuerdo con la invención, la invención proporciona el diseño y utilización de miméticos de ARNhc de miR-451 basados en moléculas de ARNip existentes. En una realización de la invención, la invención proporciona el diseño y utilización de miméticos de ARNhc de miR-451 basados en cualquier secuencia de 21-23 nt en la región codificante de un gen diana. En otra realización de la invención, la invención proporciona el diseño y utilización de miméticos de ARNhc de miR-451 basados en cualquier secuencia de 21-23 nt en la región no codificante de un gen diana. En particular, el mimético de ARNhc de miR-451 comprende una secuencia que es completamente complementaria a una secuencia larga de 21 a 23 nucleótidos en el gen diana, o a la secuencia diana de 21 a 23 nucleótidos del ARNip. Al diseñar el mimético de ARNhc de miR-451, esta secuencia completamente complementaria se posiciona dentro del ARNhc, de manera que el procesamiento por Ago2 del ARNhc y el recorte adicional dentro del complejo RISC genera una molécula de silenciamiento activa que comprende dicha secuencia complemente complementaria.

En algunas realizaciones, el ARNhc de la invención es un ARNhc sintético.

En un ejemplo no limitante, el diseño de un mimético de ARNhc de miR-451 dirigido a p53 se representa en (Figura 3). El ARNhc ~ 40 nt resultante tiene un tallo corto y un bucle hermético y no puede ser procesado por DICER. En cambio, se escinde por AGO2 y luego se recorta para generar las cadenas activas dirigidas al ARNm de p53. En otro ejemplo no limitante descrito en el presente documento a continuación en el Ejemplo 6, p53 puede ser atenuado utilizando la cadena principal de la secuencia primaria de miR-451 mediante el injerto de una secuencia de ARNhc dirigida a p53 en la secuencia primaria de miR-451.

En una realización de la invención, se proporciona un ARNhc que comprende una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en un gen diana, que tiene una secuencia diferente a la secuencia madura de miR-451, y una segunda secuencia seguida directamente de la primera secuencia, en la que la segunda secuencia es complementaria a la secuencia de los primeros 17 nucleótidos contados desde el extremo 5' de la primera secuencia.

En una realización de la invención, se proporciona un ARNhc que tiene la estructura



en la que X₂ a X₂₂ son nucleótidos complementarios a una secuencia en un gen diana, y están en una secuencia diferente a la secuencia madura de miR451; Y₄ a Y₂₀ son nucleótidos complementarios de X₂ a X₁₈; y X₁, Y₁, Y₂, e Y₃, son nucleótidos que pueden estar presentes o ausentes, en el que, X₁ e Y₃, cuando están presentes, pueden ser complementarios o no complementarios.

En una realización de la invención, se proporciona un ARNhc que comprende una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en la región codificante de un gen diana, y una segunda secuencia que sigue directamente a la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 o 18 nucleótidos contados desde el extremo 5' de la primera secuencia.

De acuerdo con la invención y el presente ARNhc, los últimos 3 nucleótidos o, alternativamente, los últimos 4 nucleótidos, de la primera secuencia forman una región de bucle en la molécula de horquilla corta.

En varias formas de realización del presente ARNhc, el ARNhc tiene un saliente de 1 nucleótido en su extremo 3' o, alternativamente, un saliente de 2, 3 o más de 3 nucleótidos en su extremo 3'.

En varias formas de realización del presente ARNhc, el ARNhc tiene un saliente de 1 nucleótido en su extremo 5' o, alternativamente, un saliente de 2, 3 o más de 3 nucleótidos en su extremo 5'.

En una realización del presente ARNhc, el ARNhc no tiene un saliente de 3' o 5'.

En una realización del presente ARNhc, el ARNhc consiste en una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en la región codificante de un gen diana, y una segunda secuencia que sigue directamente a la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 o 18 nucleótidos contados desde el extremo 5' de la primera secuencia.

En una realización de la invención, se proporciona un ARNhc que comprende una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en un intrón u otra región no codificante de un gen diana, y una segunda secuencia que sigue directamente a la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 o 18 nucleótidos contados desde el extremo 5' de la primera secuencia.

En algunas realizaciones de la invención, un ARNhc comprende una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en un gen diana no codificante.

En algunas realizaciones de la invención, un ARNhc comprende una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en un ARN largo no codificante diana.

De acuerdo con la invención y el presente ARNhc, los últimos 3 nucleótidos o, alternativamente, los últimos 4 nucleótidos, de la primera secuencia forman una región de bucle en la molécula de horquilla corta.

En varias formas de realización del presente ARNhc, el ARNhc tiene un saliente de 1 nucleótido en su extremo 3' o, alternativamente, un saliente de 2, 3 o más de 3 nucleótidos en su extremo 3'.

En varias formas de realización del presente ARNhc, el ARNhc tiene un saliente de 1 nucleótido en su extremo 5' o, alternativamente, un saliente de 2, 3 o más de 3 nucleótidos en su extremo 5'.

En una realización del presente ARNhc, el ARNhc no tiene un saliente de 3' o 5'.

En una realización del presente ARNhc, el ARNhc consiste en una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en la región codificante de un gen diana, y una segunda secuencia

que sigue directamente a la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 o 18 nucleótidos contados desde el extremo 5' de la primera secuencia.

Otros aspectos de la invención incluyen un vector de expresión que comprende una secuencia que codifica un ARNhc como se describe en el presente documento, unido operativamente a un promotor de la ARN polimerasa, y una biblioteca de dichos vectores de expresión. El vector de expresión o biblioteca de vectores de expresión se puede introducir en una célula de mamífero *in vitro* o *in vivo* en un procedimiento para atenuar la expresión del gen diana. El ARNhc se expresa en una cantidad suficiente para atenuar la expresión del gen diana de una manera específica de secuencia. En una realización preferida, el ARNhc se expresa de manera estable en la célula de mamífero.

En un aspecto de la invención, se proporciona un procedimiento *in vitro* para atenuar la expresión de un gen diana en una célula de mamífero, comprendiendo el procedimiento introducir en la célula de mamífero un vector de expresión que comprende una secuencia que codifica una molécula de ARN de horquilla corta (ARNhc) unido operativamente a un promotor de la ARN polimerasa, comprendiendo el ARNhc:

- (i) una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en la región codificante del gen diana, que tiene una secuencia diferente a la secuencia madura de miR-451,
- (ii) una segunda secuencia que sigue directamente a la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 o 18 nucleótidos contados desde el extremo 5' de la primera secuencia,

en la que los últimos 3 o últimos 4 nucleótidos de la primera secuencia forman una región de bucle en el ARNhc, en la que la molécula de ARNhc se expresa en la célula de mamífero en una cantidad suficiente para atenuar la expresión del gen diana de una manera específica de secuencia, por lo que se inhibe la expresión del gen diana.

En determinadas realizaciones, el presente vector de expresión comprende una secuencia que codifica el ARNhc acorde unido operativamente a un promotor de ARN polimerasa. En determinadas realizaciones, la invención proporciona para su uso una biblioteca de vectores de expresión, en la que cada vector de expresión comprende una secuencia que codifica el ARNhc unido operativamente a un promotor de ARN polimerasa.

El ARNhc de la invención se puede introducir en la célula directamente, o puede formar un complejo con lípidos catiónicos, empaquetarse dentro de liposomas, o administrarse de otro modo a la célula. En determinadas realizaciones, el ARNhc puede ser un ARNhc sintético, que incluye los ARNhc que incorporan nucleótidos modificados, tales como aquellos con modificaciones químicas en el grupo 2'-OH en la cadena principal de la azúcar ribosa, tales como el 2'-O-metilo (2'OMe), sustituciones 2'-flúor (2'F), y aquellas que contienen 2'OMe, o 2'F, o 2'-desoxi, o modificaciones de "ácido nucleico bloqueado" (LNA, de sus siglas en inglés). En algunas realizaciones, un ARNhc de la invención contiene nucleótidos modificados que aumentan la estabilidad o la semivida de la molécula de ARNhc *in vivo* y/o *in vitro*. Como alternativa, el ARNhc puede comprender uno o más aptámeros, que interactúan con una diana de interés para formar un complejo aptámero:diana. El aptámero puede estar en el extremo 5' o 3' del ARNhc. Los aptámeros se pueden desarrollar a través del procedimiento de selección SELEX y sintetizarse químicamente. Un aptámero se elige generalmente para unirse preferentemente a una diana. Las dianas adecuadas incluyen pequeñas moléculas orgánicas, polinucleótidos, polipéptidos y proteínas. Las proteínas pueden ser proteínas de la superficie celular, proteínas extracelulares, proteínas de membrana o proteínas séricas, como la albúmina. Dichas moléculas diana pueden ser internalizadas por una célula, efectuando así la absorción celular del ARNhc. Otras dianas potenciales incluyen orgánulos, virus y células.

También se incluye en la invención una célula de mamífero aislada que comprende los ARNhc descritos en el presente documento. En una realización preferida, la célula de mamífero es una célula humana. Otro aspecto de la invención proporciona un mamífero no humano que comprende la célula descrita anteriormente. En una de las enseñanzas, el mamífero no humano puede ser un mamífero quimérico, algunas de cuyas células somáticas o germinales que comprenden los ARNhc se describen en el presente documento. Como alternativa, el mamífero no humano puede ser un mamífero transgénico, cuyas células somáticas o germinales comprenden los ARNhc descritos en el presente documento. Por lo tanto, también se proporcionan mamíferos transgénicos cuyos genomas comprenden una secuencia que codifica los ARNhc de la invención. En una realización, el mamífero transgénico es un ratón.

También se incluye en la invención una célula no mamífera aislada que comprende los ARNhc descritos en el presente documento. Las células pueden ser las de organismos vertebrados u organismos no vertebrados, como los insectos. Las células pueden ser las de peces (por ejemplo, las del género *Fugu* o el género *Danio*), ranas (por ejemplo, las del género *Xenopus*), gusanos redondos (por ejemplo, los del género *Caenorhabditis*), moscas (como el género *Drosophila*) u otras. Otro aspecto de la invención proporciona un animal no humano que comprende la célula descrita anteriormente. En determinadas enseñanzas, el animal no humano puede ser un animal quimérico, algunas de cuyas células somáticas o germinales que comprenden los ARNhc se describen en el presente documento. Como alternativa, el animal no humano puede ser un animal transgénico, cuyas células somáticas o germinales comprenden los ARNhc descritos en el presente documento. Por lo tanto, también se proporcionan animales transgénicos cuyos genomas comprenden una secuencia que codifica los ARNhc de la invención.

Otra enseñanza de la invención proporciona el diseño de miARN basados en miméticos estructurales de precursores de miR-451. En determinadas enseñanzas, tales miméticos estructurales de miARN de miR-451, o un vector de expresión o biblioteca de vectores de expresión que codifican tales miméticos estructurales se pueden introducir en diferentes patrimonios genéticos de células de mamíferos, en particular en células deficientes de las enzimas de la ruta canónica, tales como *dicer*, para detectar e identificar dichos miARN capaces de rescatar el fenotipo nulo y las funciones funcionales de dichos miARN para contribuir al fenotipo.

En una enseñanza de la invención, se proporciona un miARN de origen no natural que comprende una primera secuencia de 21, 22 o 23 nucleótidos que corresponden a la secuencia madura completa, o una parte de esa secuencia, de un miARN de mamífero distinto de miR-451, y una segunda secuencia que sigue directamente a la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 o 18 nucleótidos contados desde el extremo 5' de la primera secuencia. El experto en la materia apreciará que la secuencia madura de un miARN de mamífero puede identificarse alternativamente como la secuencia de la cadena guía, o secuencia guía para ese miARN.

En línea con la presente invención, el miARN de origen no natural presente, los últimos 3 nucleótidos o, alternativamente, los últimos 4 nucleótidos, de la primera secuencia forman una región de bucle en la molécula de horquilla corta.

En varias enseñanzas del presente miARN de origen no natural, el miARN de origen no natural tiene un saliente de 1 nucleótido en su extremo 3', o alternativamente un saliente de 2, 3 o más de 3 nucleótidos en su extremo 3'.

En varias enseñanzas del presente miARN de origen no natural, el miARN de origen no natural tiene un saliente de 1 nucleótido en su extremo 5' o, alternativamente, un saliente de 2, 3 o más de 3 nucleótidos en su extremo 5'.

En una enseñanza del presente miARN de origen no natural, el miARN de origen no natural no tiene un saliente en 3' o 5'.

Una enseñanza de la invención proporciona una composición farmacéutica que comprende un ARNhc que comprende una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en la región codificante de un gen diana, que tiene una secuencia diferente a la secuencia madura de miR-451, y una segunda secuencia seguida directamente de la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 o 18 nucleótidos contados desde el extremo 5' de la primera secuencia, en la que los últimos 3 o los últimos 4 nucleótidos de la primera secuencia forman una región de bucle en el ARNhc

Una enseñanza de la invención proporciona un ARNhc que comprende una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en la región codificante de un gen diana, que tiene una secuencia diferente a la secuencia madura de miR-451, y una segunda secuencia seguida directamente de la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 o 18 nucleótidos contados desde el extremo 5' de la primera secuencia para la fabricación de un medicamento en la que los últimos 3 o los últimos 4 nucleótidos de la primera secuencia forman una región de bucle en el ARNhc.

3. Breve descripción de los dibujos

Las **Figuras 1A-D** muestran la utilización de miméticos de ARNhc de miR-451 para iARN en células de mamíferos. Representaciones esquemáticas de la horquilla mimética pre-let-7c-miR-451 (SEQ ID NO: 1) en comparación con las pre-let-7c nativas (SEQ ID NO: 2). Cadena guía en rojo y cadena secundaria (*passenger*) en azul. (**Figura 1A**) Histogramas de GFP superpuestos que indican la actividad de las moléculas let-7c utilizando células ME del sensor GFP let-7c. Las células se cotransfectaron con control PE-ARNip. Se apilaron 105 células PE positivas y se analizaron para determinar la expresión de GFP. (**Figura 1B**) Ensayo de luciferasa dual que indica la actividad let-7 madura. Parte superior: Esquema de la construcción indicador let-7 MSCV-luciferasa que contiene dos sitios let-7c perfectamente coincidentes en la 3'UTR. Parte inferior: Histograma que muestra valores de luminiscencia de las relaciones luciferasa/renilla. Los datos son la media de tres repeticiones técnicas $\pm$ DE. (**Figura 1C**) Panel superior: esquema del diseño de la horquilla de p53 en la cadena principal mir-30 (SEQ ID NO: 3) o que imita el pliegue de miR-451 (SEQ ID NO: 4). Panel inferior: Análisis de transferencia Western que muestra la reducción de la expresión de p53 en células ME tras la transfección de las horquillas de p53 y la inducción de p53 con adriamicina. (**Figura 1D**)

Las **Figuras 2A-B** muestran etapas en la biogénesis de ARNhc de miR-451. Modelo de la biogénesis de miR-451 que utiliza un plásmido de expresión de MSCV artificial. La transcripción primaria, impulsada por el promotor LTR, se procesa mediante *drosha* para liberar el pre-miR-451 de 40 nt (SEQ ID NO: 5) que Ago2 procesa para generar el complejo RISC activo. (**Figura 2A**) IP-northerns que muestran el procesamiento del miR-451 maduro solo en los inmunoprecipitados Ago2. Los complejos Ago1 podrían cargar pre-miR-451 pero no fueron capaces de procesarlo a su forma madura. (**Figura 2B**).

Las **Figuras 3A-B** muestran un ejemplo esquemático para generar una molécula mimética miR-451. Aquí se muestra p53-ARNhc-1224 como un ejemplo dirigido al siguiente sitio en la señal de p53: UCCACUACAAGUACAUGUGUAA (SEQ ID NO: 6). La **Figura 3A** representa la ruta de procesamiento de miARN canónica que utiliza la cadena principal mir-30 de la horquilla de p53 (SEQ ID NO: 7). La secuencia del bucle mir-30 en verde. Las cadenas están codificadas por color: la cadena antisentido en rojo y cadena con sentido que corresponde al sitio diana en azul. Los sitios de corte de la ARNasIII DICER se representan utilizando flechas. La **Figura 3B** muestra la generación del ARNhc p53-1224 que imita la estructura de miR-451 que se puede canalizar a través de la ruta de biogénesis de miARN mediada por Ago2. La cadena antisentido (roja) que abarca el tallo está diseñada para extenderse en el bucle. El brazo secundario (*passenger*) se resalta en azul. La catálisis por Ago2 del enlace fosfodiéster previsto se muestra con unas tijeras. Esta ruta genera solo RISC activo Ago2.

Las **Figuras 4A-C** muestran que la expresión de miR-451 maduro depende de la catálisis por Ago2. Diagrama de dispersión de las lecturas de miARN en el hígado fetal de tipo silvestre frente al mutante. (**Figura 4A**). RT-PCR cuantitativa de los niveles de transcripción primaria de mir-144 (SEQ ID NO: 8) y miR-451 (SEQ ID NO: 9) en muestras de hígado ts y mutantes. Los datos son la media de tres réplicas biológicas +/- DE * prueba t con igual varianza $p > 0,05$. (**Figura 4B**). La estructura única de la horquilla de miR-451 en comparación con mir-144 con la anotación mirbase de miR-451 maduro y mir-144 mapeado a la estructura de horquilla secundaria predicha que se muestra. Cadena guía en rojo y cadena secundaria (*passenger*) en azul. (**Figura 4C**).

Las **Figuras 5A-5F** muestran biogénesis no canónica de miR-451. Efecto sobre los niveles de miARN maduros de la ablación condicional de Drosha en las MEF de Drosha flox/flox Cre-ER. (**Figura 5A**). Procesamiento in vitro de los transcritos primarios miR-451 y mir-144 por inmunoprecipitados de Drosha. pre-miR144 y pre-miR451 se indican con sus tamaños esperados correspondientes. Los fragmentos adicionales liberados por el posible procesamiento Drosha del flanco 5' miR-451 se indican con asteriscos. Los flancos están indicados con puntas de flecha. (**Figura 5B**). Transferencias Northern para la confirmación de los ensayos in vitro de procesamiento de Drosha. (**Figura 5C**). Efecto sobre los niveles de miARN maduros de la ablación condicional de Dicer en las células ME de Dicer flox/flox Cre-ER. (**Figura 5D**) Efectos sobre los niveles de miARN maduros en células ME estables Dicer nulas. (**Figura 5E**). Expresión de miR-451 madura en células ME estables dicer-nulas. U6 se utiliza como control de carga. (**Figura 5F**).

Las **Figuras 6A-6D** muestran que se requiere catálisis por Ago2 para la biogénesis de miR-451. Paneles de la izquierda: Transferencia Northern en ARN total de hígados ts y mutantes que sondan la cadena guía miR-451 y los brazos secundarios (*passenger*) de la horquilla (indicado). Let-7 se utiliza como control de carga. Paneles de la derecha: Transferencias Northern de inmunoprecipitados de control de Ago2 e IgG de extractos de hígado ts y mutantes con las sondas indicadas. (**Figura 6A**). Distribución de la longitud de lectura de miARN para el miARN indicado de la secuenciación profunda de hígados TS y mutantes. (**Figura 6B**). Predicción de un sitio de escisión de Ago2 de miR-451. Parte superior: heterogeneidad del extremo 3' de miR-451. Parte inferior: sitio de escisión previsto en el 30° enlace fosfodiéster de pre-miR-451. (**Figura 6C**). Gel izquierdo: ensayo de escisión in vitro de pre-miR-451 por un inmunoprecipitado Ago2. Gel derecho: confirmación del carácter del extremo 3' del producto de escisión Ago2 mediante la eliminación beta y las reacciones de ligadura en el extremo 3'. (**Figura 6D**).

La **Figura 7** muestra la expresión de miméticos estructurales de ARNhc de miR-451. Un fragmento mir-144-451 clonado en el sitio MluI/BglII del vector y que abarca la secuencia agrupada mir-144-455 se amplifica fuera del genoma humano. Del fragmento amplificado, se genera un casete miR-451 mediante la subclonación de un fragmento de la secuencia agrupada mir-144-451 que comprende secuencias flanqueantes de miR-451 5' y 3', sitios de restricción de ingeniería en cada uno de los extremos 5' y 3' de ese fragmento, y la subclonación del casete resultante en una cadena principal del plásmido de expresión de MSCV. Se genera una construcción de expresión de MSCV que codifica un mimético de ARNhc de miR-451 dirigido contra el ARNm de p53 mediante el reemplazo de la secuencia precursora de miR-451 nativa (parte sombreada; AAACCGTTACCATTACTGAGTTTAGTAATGGTAATGGTTCT) (SEQ ID NO: 11) con una secuencia que codifica el mimético de ARNhc de miR-451

La **Figura 8** muestra que los precursores de ARNhc basados en MicroARN-451 (productos drosha) son funcionales en células madre embrionarias (ME) de ratón y manifiestan una respuesta de dosis diferente en comparación con los miméticos de precursores de ARNhc basados en miR-30.

La **Figura 9** muestra que el ARNhc primario basado en MicroARN-451 es funcional, permitiendo la expresión estable de los miméticos de miR-451 que utilizan una cadena principal miR-451.

La **Figura 10** muestra que los ARNhc primarios basados en Micro-451 se procesan a través de la ruta miR-451.

La **Figura 11** muestra la generación de una cadena principal más larga de miR-451 primaria injertable a través de la introducción de sitios de restricción en la secuencia endógena de miR-451 para fines de clonación. (**Figura 11A**). Diseño de sitios de restricción en la cadena principal clonada de miR-144-451 en el vector de expresión retroviral MSCV. Hay que tener en cuenta que los sitios de restricción se encuentran fuera de la región de reconocimiento del complejo Drosha prevista. Las secuencias flanqueantes están en minúsculas con la barra gris

que cubre la extensión del precursor procesado por droscha. En el presente caso, el producto procesado con droscha puede incluir ahora nucleótidos flanqueantes de 3'. (**Figura 11B**) Análisis del ensayo del sensor de luciferasa miR-451 que muestra que no hay interrupción de la eficacia de silenciamiento de microARN maduro. Las titulaciones de la construcción sensor de luciferasa de luciérnaga MiR-451 y las construcciones de microARN primarios endógenos que albergan miR-451 de tipo silvestre o mutante con los correspondientes sitios de restricción de (**Figura 11A**) cotransfectados con construcciones de renilla. El nivel de luciferasa se midió mediante un ensayo doble de indicador de luciferasa. MSCV-PIG se utiliza como control negativo.

La **Figura 12** muestra un ejemplo de diseño de mimético de miR-451. Estructura secundaria prevista de la secuencia primaria endógena mínima de miR-451 utilizada para la expresión estable. Los miméticos se injertan dentro de la región variable de ARNhc resaltada en verde de acuerdo con la secuencia diana en el gen. Hay que tener en cuenta que el oscilante GU se convierte en un par de bases perfecto en los miméticos, se conserva la protuberancia en el tallo basal y se demuestra experimentalmente que la eliminación de la T en la posición 41 produce miméticos no funcionales, lo que sugiere un requisito estructural para el procesamiento con droscha.

La **Figura 13** muestra la reducción de la expresión de ARNhc HOTAIR en líneas MCG7 estables que expresan horquillas de miméticos de miR-451 u horquillas de miméticos de miR30 (**Figura 13A**). Los experimentos también se llevaron a cabo utilizando la reducción de la expresión de ARNip HOTAIR en la línea celular MCF-7 (**Figura 13B**).

4. Descripción detallada de la invención

4.1 Definiciones generales

Como se utiliza en el presente documento, el término "secuencia" puede significar una cadena o parte de una cadena de nucleótidos, o el orden de los nucleótidos dentro de una cadena o parte de una cadena, dependiendo del contexto apropiado en el que se utiliza el término. A menos que se especifique lo contrario en el contexto, el orden de los nucleótidos se recita desde la dirección 5' a la 3' de una cadena.

Una "secuencia codificante" o una secuencia "que codifica" una molécula particular es un ácido nucleico que se transcribe (en el caso del ADN) o se traduce (en el caso del ARNm) en un ARN inhibidor (por ejemplo, un ARNhc o un antisentido) o polipéptido, *in vitro* o *in vivo*, cuando está unido operativamente a una secuencia reguladora apropiada. Los límites de la secuencia codificante se determinan mediante un codón de inicio en el extremo 5' (amino) y un codón de parada de la traducción en el extremo 3' (carboxilo). Una secuencia codificante puede incluir, pero no se limita a, ADNc de ARNm procariota o eucariota, secuencias de ADN genómico de ADN procariota o eucariota, y secuencias de ADN sintético. Una secuencia de terminación de la transcripción generalmente se ubicará en 3' a la secuencia codificante.

Como se utiliza en el presente documento, una secuencia "que sigue directamente" a la primera secuencia, al describir un ARNhc, significa una secuencia que se extiende desde, por ejemplo, el extremo 3' de, la primera secuencia sin ningún nucleótido que intervenga entre ellos.

Como se utiliza en el presente documento, la expresión "completamente complementaria" con respecto a una secuencia se refiere a un complemento de la secuencia mediante el emparejamiento de bases de Watson-Crick, mediante el cual la guanina (G) se empareja con la citosina (C) y la adenina (A) se empareja con el uracilo (U) o la timina (T). Una secuencia puede ser completamente complementaria a la longitud completa de otra secuencia, o puede ser completamente complementaria a una porción o longitud específica de otra secuencia. Un experto en la materia reconocerá que el U puede estar presente en el ARN y que la T puede estar presente en el ADN. Por lo tanto, una A dentro de cualquiera de una secuencia de ARN o ADN puede emparejarse con una U en una secuencia de ARN o una T en una secuencia de ADN.

Como se utiliza en el presente documento, la expresión "apareamiento de bases oscilante" con respecto a dos secuencias de ácido nucleico complementarias se refiere al emparejamiento de bases de G a uracilo U en lugar de C, cuando una o ambas cadenas de ácido nucleico contienen la ribonucleobase U.

Como se utiliza en el presente documento, la expresión "completamente complementaria de forma sustancial" con respecto a una secuencia se refiere al complemento inverso de la secuencia que permite tanto el emparejamiento de bases de Watson-Crick como el emparejamiento de bases oscilante, por lo que G se empareja con C o U y A se empareja con U o T. Una secuencia puede ser complementaria de forma sustancial a la longitud completa de otra secuencia, o puede ser complementaria de forma sustancial a una porción o longitud específica de otra secuencia. Un experto en la materia reconocerá que el U puede estar presente en el ARN y que la T puede estar presente en el ADN. Por lo tanto, un U dentro de una secuencia de ARN puede emparejarse con A o G en una secuencia de ARN o una secuencia de ADN, mientras que una A dentro de una secuencia de ARN o ADN puede emparejarse con un U en una secuencia de ARN o una T en una secuencia de ADN.

Como se utiliza en el presente documento, el término "canónico" con respecto a iARN significa, que requiere escisión por DICER para la maduración de una molécula de ARNhc. Por lo tanto, un "ARNhc canónico" es un ARNhc que requiere escisión por DICER antes de convertirse en un ARNhc maduro, y una "ruta canónica" en relación con la

iARN mediada por ARNhc es una ruta que implica la escisión de un ARNhc no maduro por DICER.

El término "gen" se refiere a un ácido nucleico que comprende un marco de lectura abierto que codifica un polipéptido, que incluye secuencias tanto de exones como (opcionalmente) de intrones. El ácido nucleico también puede incluir opcionalmente secuencias no codificantes tales como secuencias promotoras y/o potenciadoras.

- 5 Como se utiliza en el presente documento, la expresión "ARN no codificante largo" se refiere a un transcrito de ARN codificante no proteico más largo que 200 nucleótidos.

- 10 El término "ARNm" se refiere a un ácido nucleico transcrito de un gen a partir del cual se traduce un polipéptido, y puede incluir regiones no traducidas tales como una 5'UTR y/o una 3'UTR. Se entenderá que un ARNhc de la invención puede comprender una secuencia de nucleótidos que es complementaria a cualquier secuencia de una molécula de ARNm, que incluye las regiones traducidas, la 5'UTR, la 3'UTR y las secuencias que incluyen tanto una región traducida como una porción de 5'UTR o 3'UTR. Un ARNhc de la invención puede comprender una secuencia de nucleótidos que es complementaria a una región de una molécula de ARNm que abarca el codón de inicio o el codón de terminación.

- 15 "Biblioteca" se refiere a una colección de moléculas de ácido nucleico (circular o lineal). En una realización preferida, una biblioteca (denominada alternativamente biblioteca de ADNc) es representativa de todos los genes expresados en una célula, tejido, órgano u organismo. Una biblioteca también puede comprender secuencias aleatorias creadas por síntesis de novo, mutagénesis u otra modificación o alteración de una o más secuencias. Una biblioteca puede estar contenida en uno o más vectores.

- 20 "Ácido nucleico" se refiere a polinucleótidos tales como ácido desoxirribonucleico (ADN) y ácido ribonucleico (ARN). El término puede incluir polinucleótidos monocatenarios y bicatenarios.

"Unido operativamente" significa que la secuencia codificante está unida a una secuencia reguladora de una manera que permite la expresión de la secuencia codificante. Las secuencias reguladoras incluyen promotores, potenciadores y otros elementos de control de la expresión que son reconocidos en la materia y se seleccionan para dirigir la expresión de la secuencia codificante.

- 25 Las moléculas de ARN "recombinante" son aquellas producidas mediante técnicas de ADN recombinante; es decir, producidas a partir de células transformadas mediante una construcción de ADN exógeno que codifica el ARN deseado. Las moléculas de ARN "sintéticas" son aquellas preparadas mediante síntesis química.

Un "sujeto" o "paciente" puede ser un animal humano o no humano.

- 30 Una "célula transducida" es una que ha sido modificada genéticamente. La modificación genética puede ser estable o transitoria. Los procedimientos de transducción (es decir, la introducción de vectores o construcciones en las células) incluyen, pero no se limitan a, fusión de liposomas (transposomas), infección viral y procedimientos habituales de transfección de ácido nucleico como la electroporación la precipitación con fosfato de calcio y la microinyección. La transducción exitosa tendrá un efecto deseado en la célula transducida, como la expresión de genes, el silenciamiento de genes, la mejora de una diana genética o la activación de un evento fisiológico diana.

- 35 En una realización, "tratar" significa ralentizar, detener o revertir la progresión de una enfermedad o trastorno. "Tratar" también puede significar una mejora de los síntomas asociados con una enfermedad o trastorno.

- 40 "Vector" se refiere a un vehículo para introducir un ácido nucleico en una célula. Los vectores incluyen, pero no se limitan a, plásmidos, fagémidos, virus, bacterias y vehículos procedentes de fuentes virales o bacterianas. Los vectores también pueden incluir aptámeros, en los que el aptámero forma parte o se conjuga con la molécula de ARNi (Dassie y col., Nature Biotechnology 27, 839 - 846 (2009), Zhou y Rossi, Silence, 1:4 (2010), McNamera y col., Nature Biotechnology 24, 1005 - 1015 (2006)). Un "plásmido" es una molécula de ADN de doble cadena circular. Un tipo útil de vector para su utilización en la presente invención es un vector viral, en el que se insertan secuencias de ADN heterólogas en un genoma viral que se puede modificar para eliminar uno o más genes virales o partes de los mismos. Determinados vectores son capaces de replicación autónoma en una célula hospedadora (por ejemplo, vectores que tienen un origen de replicación que funciona en la célula hospedadora). Otros vectores se pueden integrar de manera estable en el genoma de una célula hospedadora y, por lo tanto, se replican junto con el genoma hospedador.

4.2 Moléculas de ARNi

- 50 Las moléculas de ARN de interferencia o de ARN inhibidor pequeño (ARNi) incluyen ARN de interferencia pequeños (ARNip), ARNip asociados a repetición (ARNipar) y microARN (miARN) en todas las etapas del procesamiento, incluidos ARNhc, pri-miARN y pre-miARN. Estas moléculas tienen diferentes orígenes: los ARNip se procesan a partir de precursores bicatenarios (ARNbc) con dos cadenas distintas de ARN emparejado de bases; los ARNip que proceden de secuencias repetitivas en el genoma se denominan ARNipar; los miARN proceden de un sol transcrito que forma horquillas emparejadas. El emparejamiento de bases de los ARNip y los miARN puede ser perfecto (es decir, completamente complementario) o imperfecto, que incluye protuberancias en la región dúplex.

El diseño de los miméticos de ARNhc de miR-451 es útil en la presente invención y, en particular, la elección de las secuencias diana para los miméticos de ARNhc de miR-451 se puede basar en los existentes ARNhc, ARNip, ARN que interactúan con piwi (ARNpi), micro ARN (miARN), ARN de bicatenario (ARNbc), ARN antisentido o cualquier otra especie de ARN que se pueda escindir dentro de una célula para formar ARN interferentes, con modificaciones compatibles descritas en el presente documento. Como se utiliza en el presente documento, los ARNhc útiles en la presente invención incluyen, sin limitación, los ARNhc modificados, que incluyen los ARNhc con estabilidad mejorada *in vivo*. Los ARNhc modificados incluyen moléculas que contienen análogos de nucleótidos, incluidas aquellas moléculas que tienen adiciones, deleciones y/o sustituciones en la nucleobase, el azúcar o la cadena principal; y moléculas que están entrecruzadas o, por otro lado, modificadas químicamente. El (los) nucleótido(s) modificado(s) puede(n) estar dentro de porciones de la molécula de ARNhc, o en toda su extensión. Por ejemplo, la molécula de ARNhc puede modificarse o contener ácidos nucleicos modificados en regiones en su extremo 5', su extremo 3' o ambos, y/o dentro de la cadena guía, cadena secundaria (*passenger*) o ambos, y/o dentro de nucleótidos que sobresalen del extremo 5', del extremo 3' o de ambos. (Véase Crooke, Patentes de los Estados Unidos N.º 6.107.094 y 5.898.031; Elmen y col., las Publicaciones de los Estados Unidos N.º 2008/0249039 y 2007/0191294; Manoharan y col., la Publicación de los Estados Unidos N.º 2008/0213891; MacLachlan y col., la Publicación de los Estados Unidos N.º 2007/0135372; y Rana, la Publicación de los Estados Unidos N.º 2005/0020521.)

Como se utiliza en el presente documento, una "molécula de ARNhc" incluye un ARNhc de tallo-bucle convencional, que forma un miARN precursor (pre-miARN). "ARNhc" también incluye ARNhc incorporados en micro-ARN (ARNhc basados en miARN), en los que la cadena guía y la cadena secundaria (*passenger*) del dúplex de miARN se incorporan en un miARN existente (o natural) o en un miARN modificado o sintético (diseñado). Cuando se transcribe, un ARNhc convencional (es decir, no un mimético de ARNhc de miR-451) forma un miARN primario (pri-miARN) o una estructura muy similar a un pri-miARN natural. El pri-miARN es procesado posteriormente por Drosha y sus cofactores en pre-miARN. Por lo tanto, el término "ARNhc" incluye moléculas de pri-miARN (ARNhc-mir) y moléculas de pre-miARN.

Una "estructura de tallo-bucle" se refiere a un ácido nucleico que tiene una estructura secundaria que incluye una región de nucleótidos que se sabe o se predice que forman una doble cadena o dúplex (porción de tallo) que está unida en un lado por una región de nucleótidos monocatenarios predominantemente (porción de bucle). Las expresiones estructuras de "horquilla" y "plegado hacia atrás" también se utilizan en el presente documento para referirse a estructuras de tallo-bucle. Dichas estructuras son bien conocidas en la materia y el término se utiliza de manera consistente con su significado conocido en la materia. Como se sabe en la materia, la estructura secundaria no requiere emparejamiento de bases exacto. Por lo tanto, el tallo puede incluir uno o más desajustes de bases o protuberancias. Como alternativa, el emparejamiento de bases puede ser exacto, es decir, no incluir ningún desajuste.

"Construcción que expresa iARN" o "construcción de iARN" es una expresión genérica que incluye preparaciones de ácido nucleico diseñadas para lograr un efecto de interferencia por ARN. Una construcción que expresa iARN comprende una molécula de ARNi que puede escindirse *in vivo* para formar un ARNip o un ARNhc maduro. Por ejemplo, una construcción de iARN es un vector de expresión capaz de dar lugar a un ARNip o un ARNhc maduro *in vivo*. Los ejemplos no limitativos de vectores que pueden utilizarse de acuerdo con la presente invención se describen en el presente documento, por ejemplo, en la sección 4.6. Procedimientos ejemplares para hacer y administrar construcciones de iARN largos o cortos se pueden encontrar, por ejemplo, en los documentos WO01/68836 y WO01/75164.

4.3 Utilización de ARNi

El ARNi es una herramienta poderosa para estudios *in vitro* e *in vivo* de la función de los genes en células de mamíferos y para terapia en contextos tanto humanos como veterinarios. La inhibición de un gen diana es específica de la secuencia, ya que las secuencias de genes correspondientes a una parte de la secuencia de ARNi, y el propio gen diana, se dirigen específicamente a la inhibición genética. Se han descrito tres mecanismos de utilización de ARNi en células de mamíferos. El primero es la administración citoplásmica de moléculas de ARNip, que se sintetizan químicamente o se generan mediante la digestión con DICER del ARNbc. Estos ARNip se introducen en las células utilizando procedimientos de transfección estándar. Los ARNip entran en el RISC para silenciar la expresión del ARNm diana.

El segundo mecanismo es la administración nuclear, a través de vectores virales, de casetes de expresión génica que expresan un ARN de horquilla corta (ARNhc). El ARNhc se basa en el micro ARN interferente (miARN), un desencadenante endógeno de la ruta de iARN (Lu y col., 2005, *Advances in Genetics* 54: 117-142, Fewell y col., 2006, *Drug Discovery Today* 11: 975-982). Los ARNhc convencionales, que imitan al pre-miARN, se transcriben mediante la ARN Polimerasa II o III como moléculas monocatenarias que forman estructuras tallo-bucle. Una vez producidos, salen del núcleo, son escindidos por DICER y entran en el RISC como ARNip.

El tercer mecanismo es idéntico al segundo mecanismo, excepto que el ARNhc se modela en miARN primario (ARNhcmir), en lugar de en los transcritos de pre-miARN (Fewell y col., 2006). Un ejemplo es la construcción de miARN de miR-30. La utilización de este transcrito produce un ARNhc más fisiológico que reduce los efectos tóxicos.

El ARNhc-mir se escinde primero para producir ARNhc, y luego se escinde nuevamente por DICER para producir ARNip. El ARNip se incorpora luego en el RISC para la degradación del ARNm diana.

Para la degradación del ARNm, la represión de la traducción o la desadenilación, los ARNip o miARN maduros se cargan en el Complejo de silenciamiento inducido por ARN (RISC) mediante el complejo de carga de RISC (RLC, de sus siglas en inglés). Posteriormente, la cadena guía lleva al RISC a identificar los ARNm diana de una manera específica de secuencia y el componente cortador de RISC hidroliza el enlace fosfodiéster acoplando los nucleótidos del ARNm diana emparejados con los nucleótidos 10 y 11 de la cadena de guía del ARN. El cortador forma, junto con distintas clases de ARN pequeños, el complejo efector de iARN, que es el núcleo del RISC. Por lo tanto, la "cadena guía" es la parte del ARN bicatenario que se asocia con RISC, en oposición a la "cadena secundaria (*passenger*)", que no está asociada con RISC.

4.4 Identificación de una ruta alternativa de biogénesis de miARN

En el presente documento se desvela que el miARN de miR-451 específico para eritroides, se canaliza a través de una nueva ruta de biogénesis de ARN pequeño que requiere catálisis por AGO2 y es independiente del procesamiento por DICER. miR-451 es procesado mediante Drosha, su maduración no requiere Dicer. En cambio, el pre-miARN se carga en AGO2 y se escinde por el centro catalítico AGO para generar un extremo 3' intermedio, que luego se recorta más (Cheloufi y col., Nature 465 (7298): 584-9 (2010)) (Figura 3).

4.5 Diseño de miméticos de ARNhc de miR-451

Se puede diseñar y expresar miméticos de ARNhc de miR-451 según las características del gen nativo que codifica el ARNhc de miR-451. En particular, la arquitectura de miR-451 se puede utilizar para expresar miméticos de ARNhc de miR-451 de plásmidos de expresión basados en el promotor pol II mediante la utilización de una variedad de vectores de expresión basados en ARN pol II o incluso de plásmidos de expresión basados en el promotor pol III que utilizan promotores dependientes de pol III. En determinadas realizaciones, los vectores de expresión pueden emplear la utilización de casetes de expresión que comprenden el mimético de ARNhc de miR-451. En determinadas realizaciones, los vectores de expresión que codifican miméticos de ARNhc de miR-451 pueden basarse en cadenas principales de vectores basados en CMV o basados en MSCV. En determinadas realizaciones, los vectores de expresión que codifican miméticos de ARNhc de miR-451 pueden basarse en cadenas principales de vectores de lentivirus autoinactivadores (SIN, de sus siglas en inglés). De manera general, las cadenas principales de vectores apropiados incluyen cadenas principales de vectores utilizados en la construcción de vectores de expresión para ARNhc convencionales, incluidos los ARNhc basados en mir-30. La utilización ejemplar de casetes de expresión en la construcción de vectores de expresión de ARNhc también útil en la construcción de casetes de expresión que codifican los miméticos de ARNhc de miR-451 de la invención se encuentra en las siguientes referencias: Gottwein E. y Cullen B. Meth. Enzymol. 427:229-243, 2007; Dickens y col., Nature Genetics, 39:914-921, 2007; Chen y col., Science 303: 83-86, 2004; Zeng y Cullen, ARN 9: 112-123, 2003.

En determinadas realizaciones, utilizar los miméticos de ARNhc de miR-451 puede emplear moléculas precursoras que comprenden secuencias flanqueantes. La molécula precursora está compuesta de cualquier tipo de molécula basada en ácido nucleico capaz de acomodar dichas secuencias flanqueantes y la secuencia mimética de ARNhc de miR-451. En determinadas realizaciones, los procedimientos para la expresión eficaz de los miméticos de ARNhc de miR-451 implican la utilización de vectores de expresión que comprenden secuencias que codifican una molécula precursora, en los que la molécula precursora codificada es un mimético de ARNhc de miR-451 en el contexto de secuencias flanqueantes. En algunas realizaciones, las secuencias flanqueantes comprenden secuencias primarias de miR-451. En algunas realizaciones, las secuencias flanqueantes comprenden secuencias primarias de un miARN o miARN distintos de miR-451. En algunas realizaciones, las secuencias de miARN primarias utilizadas, o como parte de las secuencias flanqueantes, pueden dirigir la escisión por drosha del miARN. En general, este tipo de enfoque en la utilización de miARN precursores y los componentes individuales del precursor (secuencias flanqueantes y secuencia de microARN) se proporcionan en la publicación de EE.UU. N° 2008/0226553.

Para investigar la secuencia frente a los requisitos estructurales de miR-451 para entrar a la ruta alternativa de la biogénesis de los miARN, se diseñaron miméticos estructurales de miR-451 que podrían producir let-7c. En la concentración probada, estos miméticos estructurales fueron tan potentes como el pre-let-7c nativo en la supresión de GFP o indicador de luciferasa que contiene sitios complementarios let-7 perfectos (Figura 1A-C). El precursor de miR-451 también podría ser remodelado para expresar un ARNhc que reprime eficazmente p53 (Figura 1D). Se demostró que el transcrito primario codificado a partir de un plásmido de MSCV de miR-451 transfectado transitoriamente se procesa a su forma madura en células de riñón embrionario humano 293 y solo se carga AGO2 con la forma madura de la molécula (Figura 2). El plásmido de expresión miR-451 también se procesa en fibroblastos embrionarios de ratón y células madre embrionarias de ratón (Cheloufi y col., Nature 465 (7298): 584-9 (2010)).

Estos resultados demuestran que mediante el diseño de moléculas de ARNhc que imitan la estructura de miR-451, el procesamiento de estos ARNhc en células de mamíferos se canaliza hacia la ruta alternativa de biogénesis de los miARN. En una realización de la invención, los miméticos de ARNhc de miR-451 se pueden utilizar para la supresión o el silenciamiento de genes particulares expresados a través de un mecanismo postranscripcional dirigiendo el

ARNhc contra el ARNm expresado. En otra realización de la invención, los miméticos de miR-451 pueden utilizarse para la supresión o el silenciamiento de genes particulares expresados a través de un mecanismo transcripcional, dirigiendo el ARNhc contra intrones u otras regiones no codificantes del gen diana.

En una realización de la invención, el diseño de un mimético de ARNhc de miR-451 puede basarse en la secuencia de un ARNip dirigido contra el gen diana. De acuerdo con la invención, el diseño de un mimético de ARNhc de miR-451 puede basarse en cualquier secuencia de 21-23 nt en la secuencia codificante de un gen diana. En otra realización, el diseño de un mimético de ARNhc de miR-451 puede basarse en cualquier secuencia de 21-23 nt en un intrón u otra región no codificante de un gen diana. En particular, el mimético de ARNhc de miR-451 comprende una secuencia que es completamente complementaria a una secuencia larga de 21 a 23 nucleótidos en el gen diana, o a la secuencia diana de 21 a 23 nucleótidos del ARNip. Al diseñar el mimético de ARNhc de miR-451, esta secuencia completamente complementaria se posiciona dentro del ARNhc, de manera que el procesamiento por Ago2 del ARNhc y el recorte adicional dentro del complejo RISC genera una molécula de silenciamiento activa que comprende dicha secuencia completamente complementaria.

En un ejemplo no limitante, el diseño de un mimético de ARNhc de miR-451 dirigido a p53 se representa en (Figura 3). El ARNhc ~ 40 nt resultante tiene un tallo corto y un bucle hermético y no puede ser procesado por DICER. En cambio, se escinde por AGO2 y luego se recorta para generar las cadenas activas.

En algunas realizaciones de la invención, un ARNhc comprende una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en un gen diana no codificante.

En algunas realizaciones de la invención, un ARNhc comprende una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en un ARN largo no codificante diana.

En un aspecto de la invención, se proporciona un ARNhc que comprende una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en la región codificante de un gen diana, y una segunda secuencia que sigue directamente a la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 o 18 nucleótidos contados desde el extremo 5' de la primera secuencia.

En algunas realizaciones, un ARNhc comprende una primera secuencia de 21, 22 o 23 nucleótidos que es complementaria a una secuencia en una molécula de ARNm, gen o ARN largo no codificante diana y una segunda secuencia que sigue directamente a la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 o 18 nucleótidos contados desde el extremo 5' de la primera secuencia.

En algunas realizaciones, un ARNhc de la invención puede ser un ARNhc aislado.

De acuerdo con la invención y el presente ARNhc, los últimos 3 nucleótidos o, alternativamente, los últimos 4 nucleótidos, de la primera secuencia forman una región de bucle en la molécula de horquilla corta.

En varias formas de realización del presente ARNhc, el ARNhc tiene un saliente de 1 nucleótido en su extremo 3' o, alternativamente, un saliente de 2, 3 o más de 3 nucleótidos en su extremo 3'.

En varias formas de realización del presente ARNhc, el ARNhc tiene un saliente de 1 nucleótido en su extremo 5' o, alternativamente, un saliente de 2, 3 o más de 3 nucleótidos en su extremo 5'.

En una realización del presente ARNhc, el ARNhc no tiene un saliente de 3' o 5'.

En una realización del presente ARNhc, el ARNhc consiste en una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en la región codificante de un gen diana, y una segunda secuencia que sigue directamente a la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 o 18 nucleótidos contados desde el extremo 5' de la primera secuencia.

Otra enseñanza de la invención proporciona el diseño de miARN de origen no natural basados en miméticos estructurales de precursores de miR-451. En determinadas enseñanzas, tales miméticos estructurales de miARN de miR-451, o un vector de expresión o biblioteca de vectores de expresión que codifican tales miméticos estructurales se pueden introducir en diferentes patrimonios genéticos de células de mamíferos, en particular en células deficientes de las enzimas de la ruta canónica, tales como dicer, para detectar e identificar dichos miARN capaces de rescatar el fenotipo nulo y las funciones funcionales de dichos miARN para contribuir al fenotipo.

Otra enseñanza de la invención proporciona el diseño de miméticos de miARN de origen no natural de precursores de miR-451 que comprenden la secuencia guía de un ARNhc canónico de origen natural. Determinadas realizaciones de la invención comprenden un vector de expresión o bibliotecas de vectores de expresión que codifican dichos ARNhc. En algunas realizaciones, un ARNhc de la invención puede diseñarse para dirigirse a una secuencia normalmente dirigida por un miARN canónico. En algunas realizaciones, el ARNhc, o una biblioteca de ARNhc puede expresarse en células deficientes en una o más de las enzimas de la ruta canónica como el dicer,

para detectar y/o identificar dichos miARN capaces de rescatar el fenotipo nulo, o parte del fenotipo nulo y/o los roles funcionales de dichos miARN para contribuir a un fenotipo o una parte de un fenotipo. Un experto en la materia entenderá que los miARN no siempre son completamente complementarios a sus secuencias diana, incluidas las de la 3'UTR de un ARNm diana.

- 5 En una enseñanza de la invención, se proporciona un miARN de origen no natural que comprende una primera secuencia de 21, 22 o 23 nucleótidos que corresponden a la secuencia madura completa, o una parte de esa secuencia, de un miARN de mamífero distinto de miR-451, y una segunda secuencia que sigue directamente a la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 o 18 nucleótidos contados desde el extremo 5' de la primera secuencia. El experto en la materia
10 apreciará que la secuencia madura de un miARN de mamífero puede identificarse alternativamente como la secuencia de la cadena guía, o secuencia guía para ese miARN.

En una realización de la invención, se proporciona un ARNhc que comprende una primera secuencia de 21, 22 o 23 nucleótidos complementaria a una secuencia en un gen diana, y una segunda secuencia que sigue directamente a la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los
15 primeros 17 nucleótidos contados desde el extremo 5' de la primera secuencia.

En una realización de la invención, se proporciona un ARNhc que comprende una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en un gen diana, que tiene una secuencia diferente a la secuencia madura de miR-451, y una segunda secuencia seguida directamente de la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 nucleótidos contados
20 desde el extremo 5' de la primera secuencia.

En algunas realizaciones, la primera secuencia de 21, 22 o 23 nucleótidos es completamente complementaria a una secuencia en un gen diana.

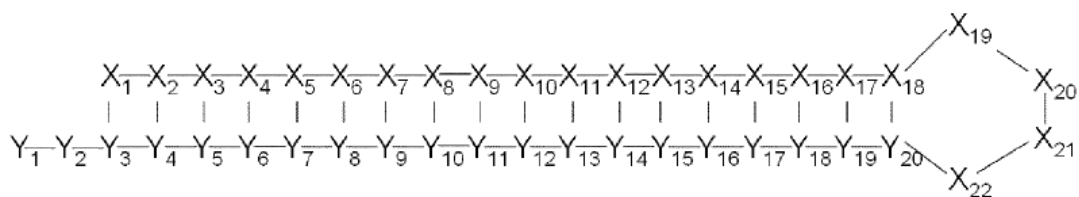
En algunas realizaciones, la primera secuencia de 21, 22 o 23 nucleótidos es complementaria a una región codificante del gen diana.

- 25 En algunas realizaciones, la primera secuencia de 21, 22 o 23 nucleótidos es complementaria a una secuencia en una molécula de ARNm codificada por el gen, en la que la secuencia en la molécula de ARNm está presente en la secuencia del gen diana.

En algunas realizaciones, la primera secuencia de 21, 22 o 23 nucleótidos es complementaria a una secuencia de región no traducida (UTR) de 3' en una molécula de ARNm codificada por el gen, en la que la secuencia 3'UTR en la
30 molécula de ARNm está presente en la secuencia del gen diana.

En algunas realizaciones, la segunda secuencia que sigue directamente a la primera secuencia es completamente complementaria a la secuencia de los primeros 18 nucleótidos contados desde el extremo 5' de la primera secuencia. En algunas realizaciones, el ARNhc está formado por 38 a 50 nucleótidos.

En una realización de la invención, se proporciona un ARNhc que tiene la estructura



- 35 en la que X₂ a X₂₂ son nucleótidos complementarios a una secuencia en un gen diana, y están en una secuencia diferente a la secuencia madura de miR451; Y₄ a Y₂₀ son nucleótidos complementarios de X₂ a X₁₈; y X₁, Y₁, Y₂, e Y₃, son nucleótidos que pueden estar presentes o ausentes, en el que, X₁, e Y₃, cuando están presentes, pueden ser complementarios o no complementarios

- 40 De acuerdo con la enseñanza de la invención y el presente miARN de origen no natural, los últimos 3 nucleótidos o, alternativamente, los últimos 4 nucleótidos, de la primera secuencia forman una región de bucle en la molécula de horquilla corta.

En varias enseñanzas del presente miARN de origen no natural, el miARN de origen no natural tiene un saliente de 1 nucleótido en su extremo 3', o alternativamente un saliente de 2, 3 o más de 3 nucleótidos en su extremo 3'.

- 45 En varias enseñanzas del presente miARN de origen no natural, el miARN de origen no natural tiene un saliente de 1 nucleótido en su extremo 5' o, alternativamente, un saliente de 2, 3 o más de 3 nucleótidos en su extremo 5'.

En una enseñanza del presente miARN de origen no natural, el miARN de origen no natural no tiene un saliente en 3' o 5'.

En algunas realizaciones, el ARNhc de la invención es un ARNhc sintético.

Como ejemplos no limitantes, en ciertas realizaciones, el presente miARN de origen no natural puede comprender una secuencia, seleccionada de la secuencia madura de miR-18b humano (uaaggugcaucuagugcaguuag) (SEQ ID NO: 12) de 21, 22 o 23 nucleótidos: por ejemplo, uaaggugcaucuagugcaguu (SEQ ID NO: 13), uaaggugcaucuagugcaguuu (SEQ ID NO: 14), uaaggugcaucuagugcaguuag (SEQ ID NO: 15), aaggugcaucuagugcaguuag (SEQ ID NO: 16), aggugcaucuagugcaguuag (SEQ ID NO: 17). En ciertas realizaciones, el presente miARN de origen no natural puede comprender una secuencia, seleccionada de la secuencia madura de miR-21 humano (uagcuuaucaucagacugauguuga) (SEQ ID NO: 18) de 21 o 22 nucleótidos: por ejemplo, uagcuuaucaucagacugauguug (SEQ ID NO: 19), uagcuuaucaucagacugauguuga (SEQ ID NO: 20), agcuuaucaucagacugauguuga (SEQ ID NO: 21). En otras realizaciones, el miARN de origen no natural puede comprender una secuencia de 21, 22 o 23 nucleótidos seleccionados de la secuencia madura de cualquier otro miARN humano o de mamífero, en el que dichas secuencias están fácilmente disponibles a través de bases de datos públicas, como miRBase, (Griffiths-Jones y col., miR-Base: tools for microARN genomics, NAR, 2008, Vol. 36, base de datos número D154-D158), accesible en <http://www.mirbase.org/>, dichas secuencias miARN maduras disponibles en miRBase específicamente incorporadas en el presente documento como referencia.

En determinadas realizaciones preferidas, el presente miARN de origen no natural puede comprender una secuencia, seleccionada de la secuencia madura de un miARN humano o de mamífero con participación conocida en cáncer y otras enfermedades, o aquellas con implicación conocida en el desarrollo y la diferenciación, por ejemplo, de dichos miARN y secuencias maduras de los mismos, identificadas en las Patentes y Publicaciones de Patentes de los Estados Unidos N.º 7.232.806, 7.307.067, 7.670.840, el documento US 2010/0048674 (25 de febrero de 2010), el documento US 2007/0072204 (29 de marzo de 2007), el documento US 2009/0124566 (14 de mayo de 2009), el documento US 2009/0176723 (9 de julio de 2009), el documento US 2009/0186353 (23 de julio de 2009) y el documento US 2009/0286242 (19 de noviembre de 2009).

En determinadas realizaciones, se proporciona un miARN de origen no natural que comprende una primera secuencia de 21, 22 o 23 nucleótidos correspondientes a la secuencia completa de la cadena secundaria (*passenger*) de un miARN de mamífero, o una parte de esa secuencia, y una segunda secuencia que sigue directamente a la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 o 18 nucleótidos contados desde el extremo 5' de la primera secuencia. En determinadas realizaciones, el presente miARN de origen no natural puede comprender una secuencia de 21, 22 o 23 nucleótidos seleccionados de la secuencia de la cadena secundaria (*passenger*) de cualquier miARN humano o de mamífero, en el que dichas secuencias de la cadena secundaria (*passenger*), o anotaciones que definen dichas secuencias, están fácilmente disponibles a través de bases de datos públicas, como miRBase, (Griffiths-Jones y col., miRBase: tools for microARN genomics, NAR, 2008, Vol. 36, base de datos número D154-D158), accesible en <http://www.mirbase.org/>, en la que dichas secuencias o anotaciones que identifican dichas secuencias, y están disponibles en miRBase. También se incluye en la invención una célula no mamífera aislada que comprende los ARNhc descritos en el presente documento. Las células pueden ser las de organismos vertebrados u organismos no vertebrados, como los insectos. Las células pueden ser las de peces (por ejemplo, las del género *Fugu* o el género *Danio*), ranas (por ejemplo, las del género *Xenopus*), gusanos redondos (por ejemplo, los del género *Caenorhabditis*), moscas (como el género *Drosophila*) u otras. Otra realización de la invención proporciona un animal no humano que comprende la célula descrita anteriormente. En determinadas enseñanzas, el animal no humano puede ser un animal quimérico, algunas de cuyas células somáticas o germinales que comprenden los ARNhc se describen en el presente documento. Como alternativa, el animal no humano puede ser un animal transgénico, cuyas células somáticas o germinales comprenden los ARNhc descritos en el presente documento. Por lo tanto, también se proporcionan animales transgénicos cuyos genomas comprenden una secuencia que codifica los ARNhc de la invención.

4.6 Vectores

En determinadas realizaciones, los vectores de expresión que codifican miméticos de ARNhc de miR-451 pueden basarse en cadenas principales de vectores basados en CMV o basados en MSCV. En determinadas realizaciones, los vectores de expresión que codifican miméticos de ARNhc de miR-451 pueden basarse en cadenas principales de vectores de lentivirus autoinactivadores (SIN, de sus siglas en inglés). De manera general, las cadenas principales de vectores apropiados incluyen cadenas principales de vectores utilizados en la construcción de vectores de expresión para ARNhc convencionales, incluidos los ARNhc basados en mir-30. Ejemplos de cadenas principales de vectores y metodologías para la construcción de vectores de expresión adecuados para su utilización con los miméticos de ARNhc de miR-451 de la presente invención, y los procedimientos para introducir dichos vectores de expresión en diversas células de mamíferos se encuentran en las siguientes referencias: Premsrurit PK. y col., Cell, 145(1):145-158, 2011, Gottwein E. y Cullen B. Meth. Enzymol. 427:229-243, 2007, Dickens y col., Nature Genetics, 39:914-921, 2007, Chen y col., Science 303: 83-86, 2004; Zeng y Cullen, ARN 9: 112-123, 2003.

Los vectores pueden ser vectores diana, como los que utilizan la recombinación de flp en el locus colA permitiendo

la integración de copia única. Otros sitios diana en el genoma del ratón incluyen, pero no se limitan a ROSA26 y HPRT. Además, la transposasa puede utilizarse para introducir miméticos en el genoma de un animal o en la célula de un animal. Véase, Premsrurit PK. y col., Cell, 145(1):145-158, (2011).

5 Los vectores descritos en la solicitud internacional n.º PCT/US2008/081193 (documento WO 09/055724) y los procedimientos para hacer y utilizar los vectores pueden utilizarse en la presente invención. La divulgación provista en el presente documento ilustra los principios generales de la construcción de vectores y la expresión de secuencias de construcciones de vectores, y no pretende limitar la presente invención.

10 Los ARNhc se pueden expresar a partir de vectores para proporcionar un silenciamiento sostenido y una administración de alto rendimiento en casi cualquier tipo de célula. En una determinada realización, el vector es un vector viral. Los ejemplos de vectores virales incluyen retrovirales, incluyendo vectores virales lentivirales, adenovirales, baculovirales y aviares. La utilización de la administración de ARNi basada en vectores virales no solo permite integraciones genómicas estables de una sola copia, sino que también evita la respuesta no específica de secuencia a través del receptor 3 de tipo peaje en la superficie celular (TLR3, de sus siglas en inglés), lo que ha planteado muchas preocupaciones por la especificidad de los efectos mediados por ARNi. En una realización de la presente invención, se introduce un conjunto de ARNhc en HSC murinas utilizando un vector conocido en la materia.

15 Los retrovirus de los que pueden proceder los vectores de plásmidos retrovirales incluyen, pero no se limitan a, virus de la leucemia murina de Moloney, virus de la necrosis del bazo, virus del sarcoma de Rous, virus del sarcoma de Harvey, virus de la leucosis aviar, virus de la leucemia del mono gibbon, virus de inmunodeficiencia humana, virus del sarcoma mieloproliferativo y virus del tumor mamario. Se puede emplear un vector de plásmido retroviral para transducir líneas celulares de empaquetamiento para formar líneas celulares productoras. Los ejemplos de células de empaquetamiento que pueden transfectarse incluyen, pero no se limitan a, las PE501, PA317, R-2, R-AM, PA12, T19-14x, VT-19-17-H2, RCRE, RCRIP, GP+E-86, GP+envAm12, y las líneas celulares DAN como se describe en Miller, Human Gene Therapy 1:5-14 (1990). El vector puede transducir las células de empaquetamiento a través de cualquier medio conocido en la materia. Una línea celular productora genera partículas de vectores retrovirales infecciosas que incluyen polinucleótidos que codifican una proteína de replicación de ADN. Dichas partículas de vector retroviral pueden emplearse luego, para transducir células eucariotas, ya sea in vitro o in vivo. Las células eucariotas transducidas expresarán una proteína de replicación de ADN.

20 En determinadas realizaciones, las células pueden diseñarse utilizando un virus adenoasociado (AAV, de sus siglas en inglés). Los AAV son virus defectuosos de origen natural que requieren virus auxiliares para producir partículas infecciosas (Muzyczka, N., Curr. Topics in Microbiol. Immunol. 158:97 (1992)). También es uno de los pocos virus que pueden integrar su ADN en células no divisorias. Los vectores que contienen tan poco como 300 pares de bases de AAV se pueden empaquetar e integrar, pero el espacio para el ADN exógeno se limita a aproximadamente 4,5 kb. Los procedimientos para producir y utilizar tales AAV son conocidos en la materia. Véase, por ejemplo, las Patentes de Estados Unidos N.º 5.139.941, 5.173.414, 5.354.678, 5.436.146, 5.474.935, 5.478.745 y 5.589.377. Por ejemplo, un vector de AAV puede incluir todas las secuencias necesarias para la replicación del ADN, la encapsidación y la integración de la célula hospedadora. El vector de AAV recombinante puede transfectarse en células de empaquetamiento que están infectadas con un virus auxiliar, utilizando cualquier técnica estándar, incluida la lipofección, electroporación, precipitación con fosfato cálcico, etc. Los virus auxiliares apropiados incluyen adenovirus, citomegalovirus, virus de la variolovacuna o virus herpes. Una vez que las células de empaquetamiento se transfectan e infectan, producirán partículas virales de AAV infecciosas que contienen la construcción polinucleotídica. Estas partículas virales se utilizan luego para transducir células eucariotas.

40 En determinadas realizaciones, las células se pueden diseñar utilizando un lentivirus y vectores basados en lentivirus. Dicho enfoque es ventajoso porque permite la expresión específica de tejido en animales mediante la utilización de promotores pol II específicos de tipo celular, la transducción eficaz de una amplia gama de tipos de células, incluidas las células no divisorias y las células que son difíciles de infectar mediante retrovirus, y la reducción de la expresión de genes inducibles y reversibles mediante la utilización de promotores tet sensibles y otros inducibles. Los procedimientos para expresar ARNhc mediante la producción y utilización de células diseñadas por lentivirus son conocidos en la materia. Para descripciones ejemplares de dichos procedimientos, véase, por ejemplo, Stegmeier F. y col., Proc Natl Acad Sci USA 2005, 102(37):13212-13217, Klinghoffer y col., ARN 2010, 16:879-884, los contenidos de los cuales se incorporan específicamente en el presente documento. Se puede lograr una producción eficaz de lentivirus recombinante incompetente con la replicación, por ejemplo, mediante la co-transfección de vectores de expresión y plásmidos de empaquetamiento que utilizan líneas celulares de empaquetamiento disponibles comercialmente, como TLA-HEK293 TM, y plásmidos de empaquetamiento, disponibles de Thermo Scientific/Open Biosystems, Huntsville, Alabama.

55 Esencialmente, se puede emplear cualquier procedimiento para introducir una construcción de ácido nucleico en células. Los procedimientos físicos para introducir ácidos nucleicos incluyen la inyección de una solución que contiene la construcción, el bombardeo por partículas cubiertas por la construcción, remojar una célula, muestra de tejido u organismo en una solución del ácido nucleico, o la electroporación de las membranas celulares en presencia de la construcción. Se puede utilizar una construcción viral empaquetada en una partícula viral para lograr una introducción eficaz de una construcción de expresión en la célula y la transcripción del ARNhc codificado. Se pueden utilizar otros procedimientos conocidos en la materia para introducir ácidos nucleicos en las células, tales como el

transporte de vehículos mediado por lípidos, el transporte mediado por químicos, como el fosfato de calcio, y similares. Por lo tanto, la construcción de ácido nucleico que codifica ARNhc se puede introducir junto con componentes que realizan una o más de las siguientes actividades: la absorción mejorada de ARN por la célula, promover el recocido de las cadenas dúplex, estabilizar las cadenas recocidas, o aumentar la inhibición del gen diana.

La expresión de los miARN endógenos está controlada por los promotores de la ARN polimerasa II (Pol II). Se ha demostrado que los ARNhc también son impulsados de manera más eficaz por los promotores Pol II, en comparación con los promotores de la ARN polimerasa III (Dickins y col., 2005, Nat. Genet. 39: 914-921). Por lo tanto, en una determinada realización, la secuencia codificante de la molécula de ARNi se controla mediante un promotor inducible o un sistema de expresión condicional, que incluye, sin limitación, los promotores de ARN polimerasa de tipo II. Los ejemplos de promotores útiles en el contexto de la invención son los promotores inducibles por tetraciclina (incluyendo los TRE-tight), los promotores inducibles por IPTG, los sistemas transactivadores de tetraciclina y los sistemas de transactivador inverso de tetraciclina (rtTA, de sus siglas en inglés). También pueden utilizarse promotores constitutivos, al igual que promotores específicos de células o tejidos. Muchos promotores serán ubicuos, de manera que se expresan en todos los tipos de células y tejidos. Una determinada realización utiliza promotores sensibles a la tetraciclina, uno de los sistemas de expresión de genes condicionales más eficaces en estudios *in vitro* e *in vivo*. Véase la solicitud de Patente internacional PCT/US2003/030901 (Publicación N.º WO 2004-029219 A2) y Fewell y col., 2006, Drug Discovery Today 11: 975-982, para una descripción ejemplar de ARNhc inducible.

Para facilitar el monitoreo de la reducción de la expresión del gen diana, las células que albergan la construcción que expresa iARN pueden comprender además una construcción marcador o indicador, tal como una construcción fluorescente. La construcción indicador puede expresar un marcador, como la proteína fluorescente verde (GFP, de sus siglas en inglés), la proteína fluorescente verde mejorada (EGFP, de sus siglas en inglés), la proteína verde fluorescente Renilla Reniformis, GFPmut2, GFPuv4, la proteína fluorescente amarilla (YFP, de sus siglas en inglés), tal como VENUS, la proteína fluorescente amarilla mejorada (EYFP, de sus siglas en inglés), la proteína fluorescente cian (CFP, de sus siglas en inglés), la proteína fluorescente cian mejorada (ECFP, de sus siglas en inglés), la proteína fluorescente azul (BFP, de sus siglas en inglés), la proteína fluorescente azul mejorada (EBFP, de sus siglas en inglés), la proteína citrina y fluorescente roja del discosoma (dsRED). Otros marcadores detectables adecuados incluyen cloranfenicol acetiltransferasa (CAT), proteínas luminiscentes como la luciferasa lacZ (β -galactosidasa) y peroxidasa de rábano picante (HRP, de sus siglas en inglés), nopalina sintasa (NOS), octopina sintasa (OCS) y alcalina fosfatasa. El gen marcador puede introducirse por separado en la célula que alberga la construcción de ARNhc (*por ejemplo*, cotransfectado, etc.). Como alternativa, el gen marcador puede estar en la construcción de ARNhc, y la expresión del gen marcador se puede controlar por la misma unidad de traducción o por separado, por ejemplo, por un SERI (sitio de entrada ribosomal interno). En un aspecto de la invención, los genes marcadores se pueden incorporar en vectores de expresión de "sensor" para su utilización en procedimientos de alto rendimiento para determinar la eficacia de la reducción de la expresión de miméticos de ARNhc de miR-451 dirigidos contra genes particulares y para identificar las secuencias diana más potentes para un gen diana particular. Dichos procedimientos, que incluyen el diseño y utilización de plásmidos y construcciones indicadoras para probar la potencia de moléculas de ARNhc particulares, aquí útiles para probar la potencia de los miméticos de ARNhc de miR-451 se describen en la publicación PCT Fellman y col., documento WO/2009/055724.

Los indicadores también pueden ser aquellos que confieren resistencia a un medicamento, tales como neomicina, ampicilina, bleomicina, cloranfenicol, gentamicina, higromicina, kanamicina, lincomicina, metotrexato, fosfotricina, puromicina, doxiciclina y tetraciclina. Los indicadores también pueden ser genes letales, como las secuencias de la timidina quinasa del virus del herpes simple (HSV-TK, de sus siglas en inglés), así como las secuencias que codifican diversas toxinas que incluyen la toxina diftérica, la toxina tetánica, la toxina del cólera y la toxina pertussis. Otro marcador de selección negativa adicional es el gen de la hipoxantina-guanina fosforribosil transferasa (HPRT, de sus siglas en inglés) para la selección negativa en 6-tioguanina.

Para facilitar la cuantificación de ARNhc específicos en una población compleja de células infectadas con una biblioteca de ARNhc, cada construcción de ARNhc puede comprender además un código de barras. Un código de barras es una secuencia de nucleótidos única (generalmente un 19-mer), unida a cada ARNhc. El código de barras se puede utilizar para controlar la abundancia de cada ARNhc a través de la hibridación de micromatrices (Fewell y col., 2006, Drug Discovery Today 11: 975-982). En una determinada realización, cada construcción de ARNhc también comprende un código de barras único. Para obtener más información sobre la utilización de códigos de barras en los análisis agrupados de ARNhc, véase el documento WO 04/029219, Bernards y col., 2006, Nature Methods 3: 701-706, y Chang y col., 2006, Nature Methods 3: 707-714.

4.7 Moléculas de ARNhc para su utilización en procedimientos de tratamiento

En determinadas realizaciones, la invención proporciona una composición formulada para la administración de miméticos de ARNhc de miR-451 *in vivo* a un sujeto, tal como un sujeto humano o veterinario. Una composición así formulada puede comprender una célula madre que comprende una construcción de ácido nucleico que codifica un mimético de ARNhc de miR-451 diseñado para disminuir la expresión de un gen diana. Una composición también puede comprender un excipiente farmacéuticamente aceptable.

Por ejemplo, el mimético de ARNhc de miR-451 se puede expresar de manera fiable *in vivo* en una variedad de tipos de células. En ciertas realizaciones, las células se administran para tratar una afección. Existe una variedad de mecanismos por los cuales las células que expresan ARNhc pueden ser útiles para tratar el cáncer y otras enfermedades. Por ejemplo, una afección puede ser causada, en parte, por una población de células que expresan un gen no deseado. Estas células pueden extirparse y reemplazarse con células administradas que comprenden ARNhc que disminuye la expresión del gen no deseado. Un ARNhc se puede dirigir esencialmente a cualquier gen, cuya expresión disminuida puede ser útil en el tratamiento del cáncer u otra enfermedad.

Se puede utilizar cualquier célula adecuada. Por ejemplo, las células a transfectar pueden ser esencialmente cualquier tipo de célula para su implantación en un sujeto. La célula que tiene el gen diana puede ser de línea germinal o somática, totipotente o pluripotente, divisoria o no divisoria, parenquimatosa o epitelial, inmortalizada o transformada, o similar. La célula puede ser una célula madre o una célula diferenciada. Después de la transfección, las células madre se pueden administrar a un sujeto o cultivarse para formar células madre diferenciadas (por ejemplo, células madre embrionarias cultivadas para formar células madre neurales, hematopoyéticas o pancreáticas) o cultivarse para formar células diferenciadas.

Las células madre pueden ser células madre obtenidas recientemente de un donante, y en ciertas realizaciones, las células madre son células madre autólogas. Las células madre también pueden ser de una línea de células madre establecida que se propaga *in vitro*. Las células madre adecuadas incluyen células madre embrionarias y células madre adultas, ya sean totipotentes, pluripotentes, multipotentes o de menor capacidad de desarrollo. Las células madre pueden proceder de mamíferos, como roedores (por ejemplo, ratón o rata), primates (por ejemplo, monos, chimpancés o seres humanos), cerdos o rumiantes (por ejemplo, vacas, ovejas y cabras). Los ejemplos de células madre embrionarias de ratón incluyen: la línea celular ME JM1 descrita en M. Qiu y col., Genes Dev 9, 2523 (1995), y la línea ROSA descrita en G. Friedrich, P. Soriano, Genes Dev 5, 1513 (1991) y células ME de ratón descritas en la Patente de los Estados Unidos N.º 6.190.910. Muchas otras líneas ME de ratón están disponibles en Jackson Laboratories (Bar Harbor, Maine). Los ejemplos de células madre embrionarias humanas incluyen aquellos disponibles a través de los siguientes proveedores: Arcos Bioscience, Inc., Foster City, California; CyThera, Inc., San Diego, California; ES Cell International, Melbourne, Australia; Geron Corporation, Menlo Park, California; University of California, San Francisco, California; y Wisconsin Alumni Research Foundation, Madison, Wisconsin. Además, se describen ejemplos de células madre embrionarias en las siguientes patentes y solicitudes de patente publicadas de los Estados Unidos: 6.245.566; 6.200.806; 6.090.622; 6.331.406; 6.090.622; 5.843.780; 20020045259; 20020068045. Los ejemplos de células madre adultas humanas incluyen las descritas en las siguientes patentes y solicitudes de patente de los Estados Unidos: 5.486.359; 5.766.948; 5.789.246; 5.914.108; 5.928.947; 5.958.767; 5.968.829; 6.129.911; 6.184.035; 6.242.252; 6.265.175; 6.387.367; 20020016002; 20020076400; 20020098584; y, por ejemplo, en la publicación PCT WO 01/11011. En determinadas realizaciones, una célula madre adecuada puede proceder de un proceso de fusión o desdiferenciación celular, tal como se describe en la solicitud de patente de los Estados Unidos 20020090722, en las publicaciones PCT WO 02/38741, WO 01/51611, WO 99/63061 y WO 96/07732.

Las células transfectadas también se pueden utilizar en la fabricación de un medicamento para el tratamiento de sujetos. Los ejemplos de excipientes farmacéuticamente aceptables incluyen matrices, soportes u otros sustratos a los que se pueden unir las células (formadas opcionalmente como perlas, tubos o membranas sólidas o huecas), así como reactivos que son útiles para facilitar la administración (por ejemplo, tampones y sales), conservar las células (por ejemplo, quelantes como sorbatos, EDTA, EGTA o aminos cuaternarias u otros antibióticos), o promover el injerto. Las células se pueden encapsular en una membrana o en una microcápsula. Las células se pueden colocar en microcápsulas compuestas de alginato o poliácridatos. (Sugamori y col. (1989) Trans. Am. Soc. Artif. Intern. Organs 35:791; Levesque y col. (1992) Endocrinology 130:644; y Lim y col. (1992) Transplantation 53:1180).

Los procedimientos adicionales para encapsular células son conocidos en la materia. (Aebischer y col. Patente de los Estados Unidos n.º 4.892.538; Aebischer y col. Patente de los Estados Unidos n.º 5.106.627; Hoffman y col. (1990) Expt. Neurobiol. 110:39-44; Jaeger y col. (1990) Prog. Brain Res. 82:4146; y Aebischer y col. (1991) J. Biomech. Eng. 113:178-183, Patente de los Estados Unidos n.º 4.391.909; Patente de los Estados Unidos n.º 4.353.888; Sugamori y col. (1989) Trans. Am. Artif. Intern. Organs 35:791-799; Sefton y col. (1987) Biotechnol. Bioeng. 29:1135-1143; y Aebischer y col. (1991) Biomaterials 12:50-55).

El sitio de implantación de las composiciones celulares se puede seleccionar por un experto en la materia dependiendo del tipo de célula y el objeto terapéutico. Los sitios de implantación ejemplares incluyen administración intravenosa o intraarterial, administración al hígado (a través de la inyección en la vena porta), la cavidad peritoneal, la cápsula renal o la médula ósea.

En determinadas enseñanzas, la invención proporciona la modificación y la administración *in vivo* de miméticos de ARNhc de miR-451 como moléculas sintéticas de ARNi. La modificación y administración *in vivo* de moléculas sintéticas de ARNi, que incluye los ARNhc que incorporan nucleótidos modificados, tales como aquellos con modificaciones químicas en el grupo 2'-OH en la cadena principal de la azúcar ribosa, tales como el 2'-O-metilo (2'OMe), sustituciones 2'-flúor (2'F) y aquellas que contienen 2'OMe, o 2'F, o 2'-deoxi, o modificaciones del "ácido nucleico bloqueado" (LNA) pueden llevarse a cabo como se describe en las patentes de los Estados Unidos N.º 6.627.616, 6.897.068, 6.379.966; en las publicaciones de la solicitud de patente de los Estados Unidos n.º US.

2005/0107325 (19 de mayo de 2005), US 2007/0281900 (6 de diciembre de 2007) y US 2007/0293449 (20 de diciembre de 2007); y en Vorhies y Nemunaitis JJ, *Methods Mol Biol.* 2009;480:11-29, López-Fraga M y col., *Infect Disord Drug Targets.* Diciembre de 2008; 8(4):262-73, Watts y col., *Drug Discov Today.* Octubre de 2008; 13(19-20):842-55, Lu y Woodlee, *Methods Mol Biol.* 2008; 437:93-107, de Fougerolles y col., *Hum Gene Ther.* Febrero de 2008; 19(2):125-32, Rossi JJ, *Hum Gene Ther.* Abril de 2008;19(4):313-7, Belting M y Wittrup A. *Methods Mol Biol.* 2009;480:1-10, Pushparaj y col., *J. Dent. Res.* 2008; 87: 992-1003, Shrivastava y Srivastava, *Biotechnol J.* 2008 Mar; 3(3):339-53, y Raemdonck K. y col., *Drug Discov Today.* Noviembre de 2008; 13(21-22):917-31, CastanottoD & Rossi JJ, *Nature Enero de 2009; 457:426-433*, Davis M y col., la publicación en línea de *Nature Advance* (21 de marzo de 2010) doi:10.1038/nature08956.

4.8 Procedimientos de selección

Las construcciones que codifican miméticos de ARNhc de miR-451 o las bibliotecas de dichas construcciones se pueden introducir en células/organismos intactos y se pueden utilizar en la selección, como la selección de alto rendimiento (HTS, de sus siglas en inglés). Por ejemplo, mediante la utilización de miméticos de ARNhc de miR-451 o bibliotecas de tales miméticos para reducir la expresión de genes diana, se puede evaluar la función de esos genes diana, por ejemplo, en una enfermedad. De forma similar, las dianas potenciales para los productos farmacéuticos se pueden identificar o estudiar utilizando dichos procedimientos. Se puede utilizar un panel de miméticos de ARNhc de miR-451 que afectan la expresión del gen diana en diversos grados. En particular, puede ser útil medir cualquier correlación entre el grado de disminución de la expresión génica y un fenotipo particular.

Las bibliotecas de miméticos de ARNhc de miR-451 se pueden producir utilizando procedimientos conocidos en la materia. Por ejemplo, las bibliotecas de miméticos de ARNhc de miR-451 pueden basarse en bibliotecas existentes, tales como bibliotecas de ARNhc existentes. Los materiales y procedimientos existentes para el diseño y la construcción de casetes de expresión, la selección y modificación de vectores y cadenas principales de vectores, la construcción de bibliotecas, el diseño de secuencias diana y la validación de bibliotecas, tal como se aplican a las bibliotecas de ARNhc convencionales, se pueden aplicar en la construcción de bibliotecas compuestas por miméticos de ARNhc de miR-451 de la presente invención. Como ejemplos no limitativos, tales materiales y procedimientos se describen en Chang y col., *Nature Meth.* 3:707-714 (2006), la publicación PCT Fellman y col., documento WO/2009/055724.

En determinadas enseñanzas, la invención proporciona procedimientos para seleccionar/evaluar la función génica in vivo. Una célula que contiene una construcción para la expresión de un mimético de ARNhc de miR-451 puede introducirse en un animal y puede evaluarse un fenotipo para determinar el efecto de la expresión génica disminuida. Puede generarse un animal completo a partir de dichas células (por ejemplo, células ME) que contienen la construcción mimética de ARNhc de miR-451. Puede evaluarse un fenotipo del animal. El animal puede ser esencialmente cualquier animal tratable experimentalmente, como un primate no humano, un roedor, un canino, un felino, etc. También pueden generarse poblaciones de animales que expresan diferentes miembros de una biblioteca de miméticos de ARNhc de miR-451. Los fenotipos de dichos animales pueden evaluarse para determinar, por ejemplo, el efecto de un gen diana sobre un fenotipo de enfermedad (por ejemplo, iniciación o crecimiento de tumores), diferenciación de células madre, sensibilidad a medicamentos (por ejemplo, sensibilidad de células tumorales a medicamentos quimioterapéuticos), susceptibilidad a infecciones virales, bacterianas u otras, o cualquier otro fenotipo de interés, incluidos los fenotipos de enfermedades.

A menos que se indique lo contrario, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que un experto en la materia a la que la presente invención pertenece entiende comúnmente. Los procedimientos y materiales ejemplares se describen a continuación, aunque los procedimientos y materiales similares o equivalentes a los descritos en el presente documento también se pueden utilizar en la práctica o el ensayo de la presente invención.

5. Ejemplos de la invención

A continuación se proporcionan ejemplos para facilitar una comprensión más completa de la invención. Los siguientes ejemplos ilustran los modos ejemplares de hacer y practicar la invención. Sin embargo, el ámbito de la invención no se limita a las realizaciones específicas desveladas en estos Ejemplos, que son solo para fines de ilustración, ya que pueden utilizarse procedimientos alternativos para obtener resultados similares.

5.1 Ejemplo 1: Identificación de una ruta de biogénesis de microARN no canónica

5.1.1 La expresión de miR-451 maduro depende de la catálisis por Ago2

Para investigar la presión evolutiva para conservar la actividad enzimática de Argonauta, se diseñó un ratón con alelos Ago2 catalíticamente inactivos. Los mutantes homocigotos murieron poco después del nacimiento con una anemia evidente. (Cheloufi y col., *Nature* 465(7298): 584-9 (2010)). Los resultados sugirieron que la escisión de la diana dirigida por miARN podría ser importante para la maduración de los eritrocitos. Como paso hacia la identificación de dicha diana, se perfilaron los miARN expresados en el hígado, uno de los sitios hematopoyéticos fetales. La secuenciación profunda de animales de tipo silvestre y homocigotos Ago2^{ADH} reveló que prácticamente todos los microARN estaban presentes en niveles casi idénticos. Sin embargo, un miARN, miR-451, representó el 11

% de todas las lecturas de miARN en el hígado fetal normal, pero se redujo drásticamente en los mutantes (**Figura 4**).

Estudios anteriores han demostrado una fuerte dependencia del desarrollo de pro-E en eritroblastos basófilos en la expresión de miR-45137. Conjuntamente, miR-451 y miR-144 forman un grupo de microARN con expresión robusta en células eritroides. Este patrón se puede explicar, en parte, en función de la presencia de sitios reguladores para el factor de transcripción de dedo de zinc GATA-1, que actúa como un regulador maestro de la diferenciación 38 de eritroides. El circuito regulador parece estar intacto en los animales Ago2ADH, ya que no se observan cambios en los niveles de pri-miR144-451 en mutantes homocigotos (**Figura 4B**). Esto apuntó fuertemente a un impacto de la catálisis en la maduración de miR-451 en lugar de la expresión de miR-451.

La biogénesis de los microARN se produce a través de una ruta de procesamiento de dos pasos en la que Drosha escinde inicialmente el transcrito del microARN primario para liberar una horquilla de pre-miARN 39. Esto se exporta al citoplasma y es reconocido y escindido por Dicer para producir el dúplex maduro, que se carga en Ago. La cadena secundaria (*passenger*) se elimina mediante mecanismos desconocidos para obtener un complejo listo para el reconocimiento de dianas.

Un examen del precursor de miR-451 y su cadena madura reveló una característica inusual. Como se anotó, los 6 nucleótidos terminales de miR-451 maduro largo de 23 nt abarcan la región del bucle y se extienden en la cadena complementaria del precursor de la horquilla. Esta disposición parece incompatible con las actividades enzimáticas bien estudiadas de Drosha y Dicer, que normalmente liberarían el mapeo de microARN maduro solo al tallo (**Figura 4C**). Por lo tanto, se exploró la posibilidad de que miR-451 pueda adoptar un modo inusual de biogénesis.

5.1.3 *Biogénesis no canónica de miR-451*

Se comenzó por la evaluación de la dependencia de miR-451 en Drosha. Se creó una construcción, que impulsa la expresión del precursor de miR-144/451 a partir de un promotor viral fuerte y se introdujo en MEF homocigotas para un alelo de Drosha condicional. Tras la activación de Cre-ER y la pérdida de función de Drosha, se observó una reducción de 20 veces en los niveles de miR-451 maduro. Esto fue incluso más dramático que el efecto en un miARN, let-7c, con una dependencia bien establecida de los factores de procesamiento canónicos (**Figura 5A**). También se evaluó la capacidad de Drosha para liberar pre-miR-451 in vitro. Se purificaron los complejos de Drosha por afinidad a partir de células 293T humanas y se mezclaron con fragmentos sintetizados in vitro de pri-miR-451 o pri-miR-144. En ambos casos, se observaron bandas del tamaño apropiado para el pre-miARN (**Figura 5B**). En el caso de pri-miR-451, el procesamiento del flanco 5' de la transcripción se pliega en una horquilla adicional, que puede ser liberada por Drosha para dar fragmentos adicionales. Como resultado, solo se observó un flanco. Se confirmaron las identidades de las bandas pre-miARN mediante transferencia Northern con sondas de oligonucleótidos correspondientes a la especie predicha (**Figura 5C**). Considerados en conjunto, estos experimentos proporcionan apoyo genético y bioquímico para Drosha que cataliza la escisión de pre-miR-451 de su transcripción primaria.

Pre-miR-451 tiene una región de tallo inusualmente corta de 17 nt. Estudios anteriores indican que esto es demasiado corto para ser reconocido y procesado eficazmente por Dicer (Siolas y col., 2004). Por lo tanto, se examinó el papel de Dicer en la maduración de miR-451. Se introdujo el vector de expresión pre-miR-451 en células ME que son homocigotas para los alelos condicionales de Dicer y expresan Cre-ER. Mientras que la pérdida aguda de Dicer causó una reducción de aproximadamente 80 veces en un microARN de células ME de control (miR-294), los niveles de miR-451 no cambiaron (**Figura 5D**). Una población pura de células ME Dicer nulas continuas mostró una reducción de más de 500 veces en el microARN convencional, mientras que los niveles de miR-451 no se vieron afectados (**Figura 2B**). También se confirmaron estos resultados utilizando el análisis de transferencia Northern de las células ME dicer nulas que expresan de forma transitoria el precursor de miR-451 (Figura 1A-B). Finalmente, se incubó pre-miARN de miR-451 sintético con Dicer recombinante y no se observaron productos de escisión maduros, aunque el pre-let-7 se procesó eficazmente. Por lo tanto, la conversión de pre-miR-451 en un miARN maduro procede independientemente de Dicer. Por lo tanto, existe un esfuerzo por identificar una ruta alternativa de maduración.

5.1.4 *Se requiere catálisis por Ago2 para la biogénesis de miR-451*

A Mediante transferencia Northern, se examinaron las especies de miR-451 en hígados de tipo silvestre y mutantes Ago2ADH. Esto confirmó la pérdida del miARN maduro en los animales mutantes. Sin embargo, se observó la aparición de una banda de ~ 40 nt que migró conjuntamente con un pre-miR-451 sintético y se hibridó con las sondas en sus brazos 5' y 3' (**Figura 6A**). Esto indicó la acumulación del producto de escisión de Drosha en animales mutantes. Notablemente, las mismas bandas observadas en el ARN total también se detectaron en inmunoprecipitados de Ago2 (**Figura 6A**). Esto demostró la carga directa del pre-miARN en Ago2 y planteó la posibilidad de que el centro catalítico Ago2 pudiera ayudar a catalizar la maduración de este microARN.

Las propiedades bioquímicas bien establecidas de Ago2 predicen que se escindirán un pre-miR-451 cargado después de su 30ª base. Se buscaron pruebas para evidenciar de dicho intermedio en bibliotecas de ARN pequeño de hígado fetal que abarcan un intervalo de tamaño ampliado. Al trazar una distribución de tamaños de lecturas

correspondientes a un miARN convencional, miR-144, se obtuvo el patrón esperado, un pico agudo a ~ 20 nt. En cambio, el miR-451 mostró una distribución de tamaño heterogénea, exclusivamente debido a la variación en su extremo 3'. Una especie abundante correspondió precisamente al producto de escisión Ago previsto (**Figura 6B-C**).

Se confirmó la capacidad de Ago2 para cargar y escindir pre-miR-451 utilizando ensayos in vitro (**Figura 5D**). Los complejos Ago2 de tipo silvestre o catalíticamente inactivos, o los complejos Ago1 (**Figuras 2B, 6D**) se purificaron por afinidad a partir de células 293T y se mezclaron con pre-miR-451 marcado en el extremo 5'. Solo el Ago2 de tipo silvestre produjo el producto esperado, y esto dependió de la presencia de Mg²⁺. No se produjo ningún producto si se proporcionó una versión mutante del precursor en la que una mutación puntual interrumpió el emparejamiento en el sitio de escisión. Las reacciones de eliminación beta y de ligación confirmaron que la escisión dejó un extremo 3' OH libre como se esperaba de las proteínas Argonauta. Estos datos apoyan firmemente un papel para el centro catalítico Ago2 en la maduración de miR-451. Esto es quizás similar al papel propuesto de la escisión de la cadena secundaria (*passenger*) en la maduración del RISCip.

Considerados en conjunto, los resultados sugieren un modelo en el que miR-451 ingresa al RISC a través de una ruta alternativa de biogénesis. Aunque la escisión por Drosha prosigue normalmente, el paso de Dicer se omite y el pre-miARN se carga directamente en Argonauta. Esto es sorprendente, considerando estudios previos que indican un acoplamiento de escisión de Dicer y RISC cargando Chendrimada, T. P. y col. *Nature* 436, 740-4 (2005), Wang, H. W. y col. *Nat Struct Mol Biol* 16, 1148-53 (2009). Dicho complejo también carecería de interacciones entre el dominio PAZ y el extremo 3' de un producto Dicer convencional. Song, J. J. y col. *Nat Struct Biol* 10, 1026-32 (2003), Wang, Y. y col., *Nature* 456:209-13 (2008). Un informe anterior indicó la capacidad del RISC para adaptarse a dichas especies y postuló un potencial para la escisión de Ago en la maduración de los microARN canónicos. Diederichs, S. & Haber, D. A. *Cell* 131, 1097-108 (2007). Sin embargo, no se demostró ningún papel fisiológico para dicha actividad, y no se detectaron defectos medibles en el procesamiento de los miARN canónicos en mutantes Ago2ADH. La maduración de miR-451 continua con una escisión mediada por Ago que produce un intermedio que se recorta aún más. Si bien esto podría ocurrir a través de la digestión endo o exonucleolítica, la distribución observada de los extremos 3', muchos de los cuales tienen restos de U sin plantilla, parece más consistente con este último modelo. Aunque la enzimología precisa de este paso permanece oscura, los estudios preliminares no apoyan las funciones de Eri-1 o el complejo de exosomas.

5.1.5 Procedimientos

Cepas de ratón

Las cepas de ratones mutantes de inserción Ago2, generadas previamente se utilizaron para el análisis de mutantes, derivación de células ME y análisis indicador. Liu, J. y col. *Science* 305, 1437-41 (2004). La cepa trampa del gen Ago1 se generó a través de la transmisión de la línea germinal de las células ME trampa del gen Ago1 de Bay Genomics (RRR031). Se generaron ratones mutantes knock-in catalíticamente inactivos para Ago2 a través de la transmisión de la línea germinal de clones de células ME positivas dirigidos con un cromosoma artificial bacteriano (RP23-56M12) que se modificó con una mutación puntual D598A en el dominio PIWI de Ago2.

Tinción de beta-galactosidasa

Para la tinción de todo el montaje, se diseccionaron embriones de diferentes etapas junto con sus compartimentos extraembrionarios en PBS. Se realizó la tinción con beta galactosidasa utilizando los reactivos de tinción de millipore. Se realizó tinción con X-gal durante la noche a temperatura ambiente. Para las secciones placentarias, se tiñeron primero las placentas completas para B-gal, se seccionaron y se contrastaron con hematoxilina y eosina.

Cruces mutantes Ago2 y derivación de células madre embrionarias (ME)

El fenotipo mutante Ago2 se reexaminó combinando dos alelos de inserción para facilitar la genotipificación de la progenie homocigota y para aprovechar el alelo indicador beta gal Ago2. Las células ME Ago2 nulas se derivaron como describió previamente Nagy, A. y col., *Manipulating the Mouse Embryo: A Laboratory Manual* (CSHL press, 2003). Se genotiparon las células nulas utilizando cebadores específicos para ambos alelos de inserción. Se genotiparon las células nulas utilizando cebadores específicos para ambos alelos de inserción. Alelo Ago2mc: directo (GACGGTGAAGAAGCACAGGAA) (SEQ ID NO: 22), inverso (GGTCCGATGGGAAAGTGTAGC) (SEQ ID NO: 23). Alelo Ago2gt: directo (ATGGGATCGGCCATTGAA) (SEQ ID NO: 24), inverso (GAACGTCGAAGAAGGCG) (SEQ ID NO: 25).

RT-PCR, Transferencia Western e inmunoprecipitación

Se diseñaron cebadores RT-PCR Ago2 corriente abajo de ambos alelos de inserción: Ago2F: TGTTCCAGCAACCTGTTCATC (SEQ ID NO: 26), Ago2R: GATGATCTCCTGTCTGGTGCT (SEQ ID NO: 27) Se utilizaron cebadores de actina como control de normalización. ActinaF: ATGCTCCCCGGGCTGTAT (SEQ ID NO: 28), ActinaR: CATAGGAGTCCTTCTGACCCATTC (SEQ ID NO: 29). Se realizó QRT-PCR utilizando el superíndice III de invitrogen y el reactivo verde de PCR cibernético de Applied Biosystems. Se midieron los niveles de miARN utilizando ensayos de miARN primarios o maduros de Applied Biosystems. Se realizaron la transferencia Western y el análisis de inmunoprecipitación de Ago2 utilizando el anticuerpo abnova eif2c2 (M01). Se realizó Western de p53

utilizando el anticuerpo monoclonal de ratón de santa cruz (Pab240).

Agregación ME-tetraploide

Las células ME nulas de Ago2 se inyectaron en blastocitos tetraploides como describió previamente Nagy, A. & Rossant, J. en Gene Targeting A Practical Approach (ed. Joiner, A. L.) 189-192 (Oxford University Press, 2000). Se transfirieron los embriones a madres de acogida y se diseccionaron en E12.5. Se realizó tinción Beta gal como se describe anteriormente.

Recolección de sangre periférica y análisis FACS

Se extrajo sangre de fetos decapitados (pre-mortem) utilizando microcapilares heparanizados y el hemograma se realizó utilizando hemavet. Para el análisis de FACS, se aislaron células individuales de hígado, bazo y médula ósea neonatales y se tiñeron conjuntamente con los anticuerpos Ter119 y CD71 (BD) y se analizaron en un citómetro de flujo (BD) LSRII como describió anteriormente Socolovsky, M. y col. Blood 98, 3261-73 (2001). El mismo número de eventos de cada muestra se recolectó de acuerdo con la clasificación de eliminación de dobletes y se analizó de la siguiente manera: se definió la población de células ProE mediante eventos positivos medianos CD71 altos/ter119. La alta población ter119 se subdividió adicionalmente en poblaciones de células basófilas, basófilas tardías/policromáticas y ortocromáticas/reticulocitos de acuerdo con los parámetros CD71 y FSC para definir los eritroblastos diferenciadores subsiguientes Liu, Y. y col. Blood 108, 123-33 (2006).

Clonación de ARN pequeño y anotación bioinformática

Se extrajo ARN total de hígados E18.5 utilizando trizol. Se generaron dos bibliotecas de ARN pequeño con un intervalo de tamaño de 19-30 nt y 30-40 nt utilizando una estrategia de clonación de ARN pequeño modificada Aravin, A. & Tuschl, T. FEBS Lett 579, 5830-40 (2005), Pfeffer, S. y col. Nat Methods 2, 269-76 (2005). En resumen, la fracción de ARN pequeño se ligó secuencialmente en 3'OH y 5'fosfatos con conectores sintéticos, se transcribió de manera inversa y se amplificó utilizando cebadores de secuenciación de solexa. Se generaron alrededor de 7 millones de lecturas para cada biblioteca de ARN pequeño. Luego se recortaron las secuencias del conector en 3', se colapsaron y se asignaron al genoma del ratón sin desajustes mediante la utilización de múltiples pistas de anotación, a saber: genes UCSC, miARN y repeticiones. Para este estudio, se utilizó la base de datos mirbase para anotar los miARN clonados.

Cultivo celular, plásmidos, transfecciones y ensayos con sensores

El vector de expresión Mir-144-451 se construyó mediante la clonación del grupo genómico en un vector retroviral pMSCV. Se cultivaron células Cre-ER MEF y ME como se describió 48 anteriormente. La escisión del alelo dier y drosha fue mediada a través del tratamiento con tamoxifeno (100 nM) durante 5 días, seguido de una transfección transitoria del plásmido que expresa miR-451 utilizando lipofectamina (Invitrogen). Para los ensayos de procesamiento in vitro y las transferencias Northern, se cultivaron células 293T en DMEM + FBS al 10 % y se cotransfectaron utilizando reactivo Mirus LT-1 con drosha marcado con flag y construcciones de DGCR8, Ago2 o Ago1 marcado con myc con el vector de expresión MSCV-miR144-451 o solo Ago2 marcado con myc. Los ensayos de luciferasa duales se realizaron tal como se describió anteriormente. Para la validación de las células ME Ago2 nulas, se cotransfectó un plásmido luciferasa sin sitio artificial con un dúplex de ARNip perfectamente combinado (dharmacon).

Ensayos de procesamiento in vitro de Drosha

El mapeo del fragmento de PCR a miR-451 y mir-144 se amplificó fuera del genoma humano con la secuencia del promotor T7. Se generaron los transcritos de ARN Pri-451 y Pri-144 utilizando el producto genómico de PCR y el kit de transcripción in vitro T7 de Ambion. Se purificaron las transcripciones en gel y se utilizaron en un ensayo de procesamiento in vitro de drosha como describió anteriormente Lee, Y. y col. Nature 425, 415-9 (2003), Denli, A. M., y col. Nature 432, 231-5 (2004).

Análisis de transferencia Northern de ARN

Se extrajo ARN de homogeneizados de hígado e inmunoprecipitados de Ago2 (IP) utilizando el reactivo trizol. Se ejecutaron 10-15 µg de ARN total y © del ARN de IPed en un gel de acrilamida al 20 % y se transfirieron a una membrana de nailon cargada positivamente (Hybond). Se entrecruzaron las membranas, se prehibridaron en solución ultra-hyb (ambión) y se hibridaron con sondas de ADN marcadas con P32 complementarias a miR-451 y let-7c. Se lavaron las membranas con 2XSSC, SDS al 0,1 % y 1XSSC, SDS al 0,1 % y se expusieron en un tamiz de phosphoimager durante la noche.

Ensayos de escisión de Ago2 y eliminación beta

Las construcciones marcadas con myc de Ago2 (ts y D797A) se transfectaron en células 293T. Se recogieron los lisados después de 48 horas, se inmunoprecipitaron utilizando perlas de agarosa myc. Se llevó a cabo la reacción de catálisis en perlas que utilizan pre-miR-451 sintético marcado en el extremo 5' P32 (dharmacon) como se describió

anteriormente (Liu, J. y col. Science 305, 1437-1441 (2004)). Se realizó la eliminación beta a través del tratamiento del ARN purificado de las perlas de Ago2 con peryodato de sodio durante 30 minutos a temperatura ambiente seguido de precipitación con etanol. Se resuspendió el ARN en un tampón de carga que contenía TBE y se ejecutó en un gel de acrilamida al 20 % en el que se produce la reacción de eliminación beta.

5 5.2 Ejemplo 2: Diseño de miméticos estructurales de ARNhc de miR-451

Para investigar la secuencia frente a los requisitos estructurales para entrar a la ruta alternativa de la biogénesis de los miARN, se diseñaron ARNhc como miméticos estructurales del precursor miR-451 para producir let-7c. Este ARNhc diseñado estructuralmente fue al menos tan eficaz como el pre-let-7c nativo en la supresión de una GFP o indicador de luciferasa que contiene sitios complementarios de let-7 perfectos (**Figura 1A-C**). También se diseñó un mimético de ARNhc de miR-451 dirigido a p53, dirigido en el presente documento, al siguiente sitio en el ARNm de p53: UCCACUACAAGUACAUGUGUAA (SEQ ID NO: 6). (**Figura 1D, Figura 3**). Por consiguiente, se puede utilizar, por tanto, una construcción de expresión para reprimir eficazmente p53 en células mediante la expresión del mimético de ARNhc de miR-451.

5.2.1 Procedimientos

La prueba de la funcionalidad de los miméticos de miR-451 se realizó mediante tres estrategias: (1) cotransfección de miméticos let-7-miR-451, pre-let-7 o dúplex let-7 o ARN de control (dharmacon) a una concentración de 100 nM con construcción indicador de luciferasa let-7c que contiene dos sitios de coincidencia perfecta en la 3' UTR en células HEK293, (2) De manera similar, las células ME del sensor GFP de Let-7 inducible por tetraciclina que contienen dos sitios perfectamente coincidentes cotransfectados con ARNip marcado con PE y miméticos let-7-miR-451 (50 nM) seguidos de un análisis de GFP de la población de células PE positivas utilizando un citómetro de flujo LSRII (BD). El sensor de GFP se indujo mediante dox (1 µg/ml), (3) Para la reducción de la expresión de p53, se transfectaron células ME con ARNhc de p53 y miméticos de p53-miR-451 seguido de inducción de p53 con adriamicina (0,5 µg/ml) en las últimas 8 horas antes de la recolección. Todas las células se recogieron 48 horas después de la transfección.

5.3 Ejemplo 3: Expresión de miméticos estructurales de ARNhc de miR-451

Los miméticos de ARNhc de miR-451 descritos en la presente solicitud proporcionan una herramienta con amplia aplicabilidad para su utilización en aplicaciones basadas en ARNi, tanto para investigación como en aplicaciones médicas. Como ejemplo no limitante, los miméticos de ARNhc de miR-451 diseñados a través de los procedimientos de la presente solicitud y dirigidos contra genes particulares se pueden utilizar para reprimir de manera eficaz la expresión de estos genes en células de mamíferos, tanto en cultivo como en animales completos, incluyendo animales transgénicos, mediante la expresión de miméticos de ARNhc de miR-451 o una molécula precursora para dicho ARNhc. Como ejemplo no limitante adicional, un plásmido de expresión de MSCV (véase, por ejemplo, **Figura 2, Ej. 1**), se utiliza para ilustrar cómo hacer construcciones de expresión que codifican miméticos de ARNhc de miR-451 dirigidos contra p53, basado en una cadena principal de miR-451. El enfoque se describe en la **Figura 7**.

Un fragmento de mir-144-451 clonado en el sitio MluI/BglII del vector de expresión Mir-144-451 (**Figuras 2, 7, Ej. 1**) y que abarca la secuencia del grupo mir-144-455 se amplifica fuera del genoma humano. Del fragmento amplificado, se genera un casete miR-451 mediante la subclonación de un fragmento de la secuencia agrupada mir-144-451 que comprende secuencias flanqueantes de miR-451 5' y 3', sitios de restricción de ingeniería en cada uno de los extremos 5' y 3' de ese fragmento, y la subclonación del casete resultante en una cadena principal del plásmido de expresión de MSCV. Se genera una construcción de expresión de MSCV que codifica un mimético de ARNhc de miR-451 dirigido contra el ARNm de p53 mediante el reemplazo de la secuencia precursora de miR-451 nativa (Figura 7, parte sombreada; AAACCGTTACCATTACTGAGTTTAGTAATGGTCT) (SEQ ID NO: 11) con una secuencia que codifica el mimético de ARNhc de miR-451. Por conveniencia, puede ser deseable diseñar de restricción en tal construcción en los extremos 5' y 3' de la secuencia que codifica el mimético de miARN de miR-451, de tal manera que un mimético de ARNhc de miR-451 alternativo pueda integrarse fácilmente en la construcción mediante la eliminación de la secuencia mimética de ARNhc de miR-451 dirigida a p53 y el reemplazo de esa secuencia con la del mimético de ARNhc de miR-451 dirigida deseada.

De manera general, para que se retengan suficientes secuencias que actúen en cis (y determinantes estructurales) en el mimético de ARNhc miR-451 expresado para permitir un procesamiento eficaz de Drosha, es apropiado incluir 20 o más pb de la secuencia flanqueante de miR-451. En algunos casos, puede ser deseable alterar la longitud de la secuencia flanqueantes para optimizar la expresión del mimético de miR-451 maduro. Se puede seleccionar una cualquiera de las diversas longitudes de una u otra secuencia flanqueante de 5' y 3' de 5 a 60 pb y la construcción se puede diseñar para integrar la longitud deseada de la secuencia flanqueante en el casete de construcción de expresión. Un experto en la materia apreciará que dichas longitudes de secuencia flanqueante incluyen 5 pb y 60 pb y cada valor entero intermedio entre 5 y 60.

Los ejemplos se proporcionan para ilustrar la utilidad general de la invención y no pretenden limitar la implementación de este enfoque. El enfoque ilustrado en el presente documento ofrece una considerable flexibilidad en la utilización de varias construcciones de expresión, vectores alternativos y procedimientos de administración,

todos los cuales pueden optimizarse rutinariamente para su utilización en particular en células, tejidos, órganos o animales. Por ejemplo, las construcciones de expresión que emplean una cadena principal de miR-451 para la expresión de los miméticos de ARNhc de miR-451 de la invención pueden basarse en cualquier construcción de expresión análoga basada en ARN pol II utilizada para la expresión de ARNhc convencionales, incluyendo construcciones que incorporan promotores inducibles/reprimibles, específicos de tejido o regulados por el desarrollo, sitios IRES para la expresión bicistónica, marcadores seleccionables, marcadores fluorescentes y sensores de ARNi.

5.4 Ejemplo 4: Los precursores de ARNhc basados en microARN-451 (productos drosha) son funcionales

Los precursores de ARNhc basados en microARN-451 (productos drosha) son funcionales en células ME de ratón y manifiestan una respuesta a la dosis diferente en comparación con los miméticos de precursores de ARNhc basados en miR-30. Sin desear estar ligado o limitado por, ninguna teoría científica, esto puede sugerir reglas diferentes en el reconocimiento de dianas entre las dos rutas. Se generaron curvas de titulación que muestran la eficacia de los ARNhc específicos de p53 en miméticos basados en miR-451 y miR-30 (**Fig. 8**). Se transfectaron tres estructuras precursoras de ARN sintético de ARNhc de p53 de ratón (shp53.1224, shp53.279 y shp53.1404) en miR-451 (40 nt de largo) y miR-30 (61 nt de largo) en células ME de ratón. La potencia de la horquilla de p53 se clasifica principalmente como superior, intermedia y débil de acuerdo con los datos del sensor descritos en los ARNhc basados en miR-30 en el sistema de expresión de vectores primarios (Fellmann, C. y col. Mol Cell 41, 733-746 (2011)). Las concentraciones de los ARNhc de p53 se valoraron utilizando un ARNhc de control de longitud similar como control para asegurar una cantidad igual del ARN transfectado en cada concentración del ARNhc de direccionamiento. Se trataron las células con doxorrubicina a una concentración final de 500 ng/ml durante 8 horas antes de la recolección. El nivel de expresión de p53 se detectó mediante transferencia Western y se cuantificó basándose en el control negativo.

5.5 Ejemplo 5: El ARNhc primario basado en microARN-451 es funcional

La expresión estable de los miméticos de miR-451 que utilizan una cadena principal de miR-451 se logró como se ha descrito para miR-30. La ruta miR-451 depende del procesamiento de Drosha y Ago2 independientemente solo de Dicer. Se realizó la medición de la eficacia de reducción de la expresión de miméticos primarios basados en miR-451 y miR-30. (**Figuras 9A y B**) La transferencia Western de p53 (panel izquierdo) seguida de cuantificación (panel derecho) en MEF primarios infectados con una "copia única" de Multiplicidad de Infección (MOI, de sus siglas en inglés) baja y células NIH3T3 infectadas a MOI baja o alta con ARNhc de p53 de ratón en cadenas principales de miR-451 o miR-30 retrovirales (MSCV), respectivamente. Se trataron las células con doxorrubicina a una concentración final de 500 ng/ml durante 8 horas. (**Figura 9C**) La reducción de la expresión de la luciferasa de Renilla (panel izquierdo) que utiliza cuatro ARNhc de luciferasa de Renilla en cadenas principales de miR-451 o miR-30 retrovirales (MSCV) se infectó en células indicadoras NIH3T3-renilla. La luminiscencia de la luciferasa de Renilla se normalizó a la absorbancia total de proteínas utilizando el ensayo de medición de BCA. El nivel de expresión de GFP (eficacia de la infección) se mostró en el gráfico inferior.

5.6 Ejemplo 6: Los ARNhc primarios basados en Micro-451 se procesan a través de la ruta miR-451

El análisis de transferencia Northern a partir de muestras de ARN coincidentes de los experimentos de p53 en el Ejemplo 5 (**Figura 9**) se realizó utilizando sondas radiomarcadas complementarias de las secuencias de 22 nt maduras de ARNhc procesados. Se detectaron ARNhc de p53 maduros en células NIH3T3 infectadas con miméticos de pri-shp53-miR-451 y miméticos de pri-shp53-miR-30 a baja (**Figura 10A**) o alta (**Figura 10B**) MOI. Las cuantificaciones de densitometría se muestran en los paneles de la derecha. (**Figura 10C y D**) Análisis de inmunoprecipitación Ago2-northern que utiliza sondas específicas de las secuencias de 22 nt maduras de ARNhc primarios procesados. Se transfectaron los miméticos de pri-shp53-miR-451 y pri-shp53-miR-30 solos (**Figura 10C**) o cotransfectaron con Ago2 de tipo silvestre o construcciones muertas catalíticas de Ago2 (**Figura 10D**) en células HEK293T. Los miméticos de miR-451 se cargaron con éxito en complejos Ago2, pero solo se procesaron hasta su forma madura en la proteína de tipo silvestre. El mimético precursor de 40-mero se acumula en el Ago2 catalíticamente inactivo (**Figura 10**).

5.7 Ejemplo 7: Pasos del diseño del mimético estructural de ARNhc de miR-451

1. Elegir la secuencia diana de 22-mero, por ejemplo, Renilla-ARNhc-1 (la primera que se probó)
TAGGAATTATAATGCTTATCTA (SEQ ID NO: 30)
2. Invertir el complemento de la secuencia diana de Renilla de 22-mer
TAGATAAGCATTATAATTCCTA (SEQ ID NO: 31)
3. Recortar 1 a 18 nt en el complemento inverso
TAGATAAGCATTATAATT (SEQ ID NO: 32)
4. Invertir el complemento del recortado 1-18 para hacer el tallo
AATTATAATGCTTATCTA (SEQ ID NO: 33)

5. Unir las secuencias del paso 2 y el paso 4 (el tallo) en este orden para generar el ARNhc de 40-mero
TAGATAAGCATTATAATTCCTA **AATTATAATGCTTATCTA** (SEQ ID NO: 34)

6. Asegurarse de que hay una protuberancia en la primera posición: Si el primer nucleótido del ARNhc es A o T, asegurarse de que la posición 40ª sea C (como miR-451 endógeno), si la primera posición es C o G, la posición 40ª debe cambiarse a A. En este caso, sustituir A en la posición 40 con C:

TAGATAAGCATTATAATTCCTA **AATTATAATGCTTATCTC** (SEQ ID NO: 35)

7. Añadir las regiones flanqueantes de miR-451 endógeno y los sitios de restricción para la clonación en el vector de destino (la cadena principal mínima representada aquí en letras minúsculas es de aproximadamente 61 nt y 63 nt de longitud para los flancos 5' y 3' respectivamente)

gaagctctgtcagcctgtcacaacactgactgactgccaggcactgggaatggcaaggTAGATAAGCATTATAATTCCTAA**AATTATAATGCTTATCT**Ctcttgctataccagaaaacgtgccaggagaagagaactcaggaccctgaagcagact actggaa (SEQ ID NO: 36)

Los sitios de restricción se introducen con cebadores de PCR para generar el dúplex de ADN para la clonación o dos oligos complementarios se ordenan y se reasocian para formar el dúplex (**Figuras 11 y 12**).

El ARNhc dirigido a otros genes se puede hacer de una manera análoga, que utiliza una secuencia diferente de nucleótidos para dar como resultado el ARNhc como se describe en las diversas realizaciones en el presente documento.

5.8 Ejemplo 8: Reducción de la expresión del ARN largo no codificante

Un ejemplo de la utilización del diseño de miméticos ARNhc de miR-451 y miR-30 HOTAIR en el vector de expresión MSCV. Se eligieron las secuencias diana a partir de los ARNip indicadores en la literatura para dirigirse a HOTAIR de manera eficaz (Wan, Y. y Chang, H.Y. Cell Cycle 9, 3391-3392 (2010)). Cuando se prueban en cultivo, los ARNhc y los ARNip se comportan de manera diferente. Los miméticos basados en miR-451 lograron reducir exitosamente los niveles de HOTAIR en un caso más eficazmente que los miméticos basados en miR-30.

5.9 Ejemplo 9: Chip en mosaico de miR-451 para generar una biblioteca de ARNhc basada en miR-451

Para comprender las reglas subyacentes de procesamiento y reconocimiento de la diana de los miméticos de miR-451, se generó un chip de mosaico de ARNhc en un paso de nucleótido para 10 genes diferentes (p53, bcl2, mcl1, myc, rpa3, kras, PCNA, GFP, mkate2 y mcherry). Se generó una biblioteca de oligos sintéticos de 164-mero utilizando la plataforma de síntesis de oligos de Agilent (Santa Clara, CA, EE. UU.). Se amplificó luego la biblioteca utilizando los flancos constantes y se clonó en un vector de destino MSCV. Luego se transfectó la biblioteca en células 293T y se analizó el procesamiento de los ARNhc a través de la generación de bibliotecas de ARN pequeño seguidas de la secuenciación de solexa. Se analizó la calidad de estas bibliotecas y su eficacia de procesamiento. La Tabla 1 muestra la eficacia de las secuencias recuperadas en la biblioteca de ARN pequeño según la entrada.

Tabla 1. Eficacia de las secuencias recuperadas en la biblioteca de ARN pequeño

		número total de lecturas	lecturas de mapeo del genoma	lecturas de mapeo de la biblioteca de ARNhc	miARN
alineación de miR-451 con CHIP-ARNhc previstos	ARN total maduro de 19-33 nt	11.761.255	67,59 %	1,16 %	55,12 %
	IP de Ago2 maduros de 19-33 nt	15.293.024	79,35 %	1,67 %	77,62 %
	pre-fracción de ARN total de 40 nt	2.285.916	19,64 %	0,66 %	0,08 %
Alineación de miR-30 en las cadenas guía previstas de 22-mer	ARN total maduro	13.824.773	62,81 %	7,42 %	53,08 %
	IP de Ago2 maduros	15.640.560	64,02 %	9,49 %	61,74 %

NOVEDOSOS ARNhc DISEÑADOS ESTRUCTURALMENTE

RESUMEN DE LA DIVULGACIÓN

Se proporciona un diseño mejorado de ARNhc basado en miméticos estructurales de precursores de miR-451. Estos miméticos de ARNhc de miR-451 se canalizan a través de una nueva ruta de biogénesis de ARN pequeño, requieren catálisis por AGO2 y son procesados por Drosha, pero son independientes del procesamiento por DICER. Esta ruta de miARN alimenta elementos activos solo en Ago2 debido a su actividad catalítica única. Estos datos demuestran que esta ruta de biogénesis de ARN pequeño recién identificada puede explotarse in vivo para producir moléculas activas.

LISTADO DE SECUENCIAS

	<110> Cold Spring Harbor Laboratory Hannon, Gregory J. Cheloufi, Sihem	
	<120> NOVEDOSOS SHARN DISEÑADOS ESTRUCTURALMENTE	
	<130> 82764-A-PCT	
5	<150> 61/327.510 <151> 23/04/2010 <160> 36 <170> PatentIn versión 3.3	
10	<210> 1 <211> 40 <212> ARN <213> Artificial <220> <223> horquilla mimética de pre-let-7c-miR-451	
15	<400> 1 ugagguagua gguuguaugg uuguacaacc uacuaccuca 40	
20	<210> 2 <211> 67 <212> ARN <213> Mus musculus <400> 2 ugagguagua gguuguaugg uuuagaguua caccugggga guuaacugua caaccuucua 60 gcuuucc 67	
25	<210> 3 <211> 63 <212> ARN <213> Artificial <220> <223> horquilla de p53 en la cadena principal de mir-30 <400> 3 cccacuacaa guacaugugu aaugugaag ccacagaugu auuacacaug uacuuguagu 60 gga 63	
30	<210> 4 <211> 40 <212> ARN <213> Artificial	
35	<220> <223> horquilla de p53 que imita el pliegue de mir-451 <400> 4 uuacacaugu acuuguagug gacuacaagu acauguguaa 40	
40	<210> 5 <211> 41 <212> ARN <213> Mus musculus <400> 5 aaaccguuac cauucagag uuuagaaug guaaccguuc u 41	

	<210> 6		
	<211> 22		
	<212> ARN		
	<213> Mus musculus		
5	<400> 6		
	uccacuacaa guacaugugu aa	22	
	<210> 7		
	<211> 61		
10	<212> ARN		
	<213> Artificial		
	<220>		
	<223> mir-30 según pre-1224-p53hc		
	<400> 7		
	cacuacaagu acauguguaa uagugaagcc acagauguau uacacaugua cuuguagugg	60	
	a	61	
15	<210> 8		
	<211> 66		
	<212> ARN		
	<213> Mus musculus		
	<400> 8		
	ggcuggggaug ucaucauaua cuguaaguuu gugaugagac acuacaguau agaugaugua	60	
20	cuaguc	66	
	<210> 9		
	<211> 72		
	<212> ARN		
	<213> Mus musculus		
25	<400> 9		
	cuuggggaug gcgaggaaac cguuaccuu acugaguuaa guaaugguua cgguucucu	60	
	gcugcuccca ca	72	
	<210> 10		
	<211> 462		
30	<212> ADN		
	<213> Artificial		
	<220>		
	<223> Fragmento de la secuencia agrupada mir-144-451 que comprende secuencias flanqueantes de miR-451 5' y 3' con sitios de restricción diseñados		
	<400> 10		

ES 2 699 630 T3

```

agatctgagc agggagcagg aagctgtgtg tgtccagccc tgacctgtcc tgttctgccc      60
ccagcccctc acagtgtttt tcaagccatg cttcctgtgc cccagtgagg gccctggctg      120
ggatatcatc atatactgta agtttgcatg gagacactac agtatagatg atgtactagt      180
ccggggcacc ccagctctgg agcctgacaa ggaggacagg agagatgctg caagcccaag      240
aagctctctg ctcagcctgt cacaacctac tgactgccag ggcacttggg aatggcaagg      300
aaaccgttac cattactgag tttagtaatg gtaatgggtc tcttgctata cccagaaaac      360
gtgccaggaa gagaactcag gaccctgaag cagactactg gaaggagac tccagctcaa      420
acaaggcagg ggtgggggag tgggattggg ggtaggacgc gt                          462

```

```

5  <210> 11
   <211> 41
   <212> ADN
   <213> Mus musculus

   <400> 11
   aaaccgttac cattactgag tttagtaatg gtaatgggtc t      41

10 <210> 12
   <211> 23
   <212> ARN
   <213> Homo sapiens

   <400> 12
   uaaggugcau cuagugcagu uag      23

15 <210> 13
   <211> 21
   <212> ARN
   <213> Homo sapiens

   <400> 13
   uaaggugcau cuagugcagu u      21

20 <210> 14
   <211> 22
   <212> ARN
   <213> Homo sapiens

   <400> 14
25 uaaggugcau cuagugcagu ua      22

   <211> 23
   <212> ARN
   <213> Homo sapiens

   <400> 15
30 uaaggugcau cuagugcagu uag      23

   <210> 16
   <211> 22
   <212> ARN
   <213> Homo sapiens

35 <400> 16
   aaggugcauc uagugcaguu ag      22

   <210> 17
   <211> 21
   <212> ARN
40 <213> Homo sapiens

```

<400> 17
 aggugcaucu agugcaguua g 21

 <210> 18
 <211> 22
 5 <212> ARN
 <213> Homo sapiens

 <400> 18
 uagcuuauca gacugauguu ga 22

 <210> 19
 10 <211> 21
 <212> ARN
 <213> Homo sapiens

 <400> 19
 uagcuuauca gacugauguu g 21
 15 <210> 20
 <211> 22
 <212> ARN
 <213> Homo sapiens

 <400> 20
 20 uagcuuauca gacugauguu ga 22

 <210> 21
 <211> 21
 <212> ARN
 <213> Homo sapiens
 25 <400> 21
 agcuuauca gacugauguu a 21

 <210> 22
 <211> 21
 30 <212> ADN
 <213> Artificial

 <220>
 <223> Cebador para el genotipado de alelos de inserción Ago2 en células de ratón Ago2 nulas.

 <400> 22
 gacggtgaag aagcacagga a 21
 35 <210> 23
 <211> 21
 <212> ADN
 <213> Artificial

 <220>
 40 <223> Cebador para el genotipado de alelos de inserción Ago2 en células de ratón Ago2 nulas.

 <400> 23
 ggtccgatgg gaaagtgtag c 21

 <210> 24
 <211> 18
 45 <212> ADN
 <213> Artificial

 <220>
 <223> Cebador para el genotipado de alelos de inserción Ago2 en células de ratón Ago2 nulas.

 <400> 24
 50 atgggacgg ccattgaa 18

 <210> 25

<211> 18
 <212> ADN
 <213> Artificial

5

<220>
 <223> Cebador para el genotipado de alelos de inserción Ago2 en células de ratón Ago2 nulas.

<400> 25
 gaactcgtca agaaggcg 18

10

<210> 26
 <211> 20
 <212> ADN
 <213> Artificial

<220>
 <223> Cebador de RT-PCR para alelos de inserción Ago2 en ratones.

15

<400> 26
 tgttcagca acctgtcatc 20

<210> 27
 <211> 20
 <212> ADN
 <213> Artificial

20

<220>
 <223> Cebador de RT-PCR para alelos de inserción Ago2 en ratones.

<400> 27
 gatgatctcc tgcggtgct 20

25

<210> 28
 <211> 18
 <212> ADN
 <213> Artificial

<220>
 <223> Cebador de RT-PCT para actina en ratones.

30

<400> 28
 atgctccccg ggctgtat 18

35

<210> 29
 <211> 24
 <212> ADN
 <213> Artificial

<220>
 <223> Cebador de RT-PCR para actina en ratones.

<400> 29
 cataggagtc cttctgaccc attc 24

40

<210> 30
 <211> 22
 <212> ADN
 <213> Artificial

45

<220>
 <223> Secuencia diana para Renilla-ARNhc-1

<400> 30
 taggaattat aatgcttatc ta 22

50

<210> 31
 <211> 22
 <212> ADN
 <213> Artificial

	<220>		
	<223> Complemento inverso de la secuencia Renilla-ARNhc-1		
	<400> 31		
	tagataagca ttataattcc ta	22	
5	<210> 32		
	<211> 18		
	<212> ADN		
	<213> Artificial		
	<220>		
10	<223> Complemento inverso recortado de la secuencia diana Renilla-ARNhc-1		
	<400> 32		
	tagataagca ttataatt	18	
	<210> 33		
	<211> 18		
15	<212> ADN		
	<213> Artificial		
	<220>		
	<223> Complemento inverso del complemento inverso recortado de la secuencia diana Renilla-ARNhc-1		
	<400> 33		
20	aattataatg cttatcta	18	
	<210> 34		
	<211> 40		
	<212> ADN		
	<213> Artificial		
25	<220>		
	<223> Secuencia Renilla-ARNhc-1		
	<400> 34		
	tagataagca ttataattcc taaattataa tgcttatcta	40	
	<210> 35		
30	<211> 40		
	<212> ADN		
	<213> Artificial		
	<220>		
	<223> Secuencia Renilla-ARNhc-1 con sustitución de secuencia		
35	<400> 35		
	tagataagca ttataattcc taaattataa tgcttatctc	40	
	<210> 36		
	<211> 164		
40	<212> ADN		
	<213> Artificial		
	<220>		
	<223> Secuencia Renilla-ARNhc-1 con secuencias flanqueantes de miR451		
	<400> 36		
	gaagctctct gctcagcctg tcacaaccta ctgactgccca gggcacttgg gaatggcaag	60	
	gtagataagc attataattc ctaaattata atgcttatct ctcttgctat acccagaaaa	120	
	cgtgccagga agagaactca ggaccctgaa gcagactact ggaa	164	
45			

REIVINDICACIONES

1. Una molécula de ARN de horquilla corta (ARNhc) que comprende

- (i) una primera secuencia de 21, 22 o 23 nucleótidos completamente complementaria a una secuencia en un gen diana, que tiene una secuencia diferente a la secuencia madura de miR-451, y
 (ii) una segunda secuencia que sigue directamente a la primera secuencia, en la que la segunda secuencia es completamente complementaria a la secuencia de los primeros 17 o 18 nucleótidos contados desde el extremo 5' de la primera secuencia,

en la que los últimos 3 o últimos 4 nucleótidos de la primera secuencia forman una región de bucle en el ARNhc.

2. El ARNhc de la reivindicación 1, en el que la primera secuencia de 21, 22 o 23 nucleótidos es complementaria a una región codificante del gen diana.

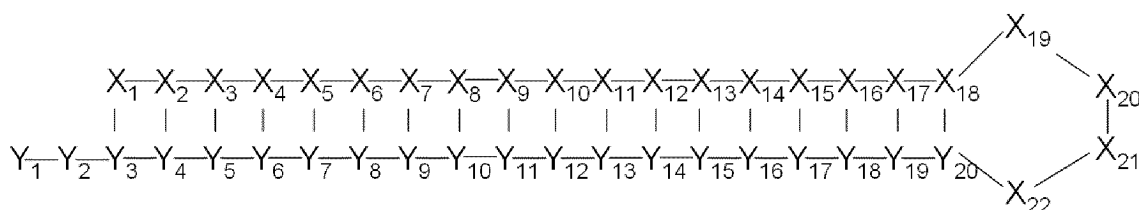
3. El ARNhc de la reivindicación 1 o 2, en el que la primera secuencia de 21, 22 o 23 nucleótidos es complementaria a una secuencia en una molécula de ARNm codificada por el gen, en el que la secuencia en la molécula de ARNm está presente en la secuencia del gen diana.

4. El ARNhc de la reivindicación 3, en el que la primera secuencia de 21, 22 o 23 nucleótidos es complementaria a una secuencia que incluye tanto una región traducida como una porción de la secuencia de la región 3' no traducida (UTR) en una molécula de ARNm codificada por el gen, en el que la secuencia UTR 3' en la molécula de ARNm está presente en la secuencia del gen diana.

5. El ARNhc de una cualquiera de las reivindicaciones 1 a 4, en el que la segunda secuencia que sigue directamente a la primera secuencia es completamente complementaria a la secuencia de los primeros 18 nucleótidos contados desde el extremo 5' de la primera secuencia.

6. El ARNhc de una cualquiera de las reivindicaciones 1 a 5, en el que el ARNhc consiste en 38 a 50 nucleótidos.

7. El ARNhc de una cualquiera de las reivindicaciones 1 a 6, que tiene la estructura



en la que X₂ a X₂₂ son nucleótidos complementarios a una secuencia en un gen diana, y están en una secuencia diferente a la secuencia madura de miR-451;

Y₄ a Y₂₀ son nucleótidos complementarios de X₂ a X₁₈; y

X₁, Y₁, Y₂, e Y₃, son nucleótidos que pueden estar presentes o ausentes,

en el que, X₁ e Y₃, cuando están presentes, pueden ser complementarios o no complementarios.

8. Un vector de expresión que comprende una secuencia que codifica el ARNhc de acuerdo con una cualquiera de las reivindicaciones 1 a 7 unido operativamente a un promotor de ARN polimerasa.

9. Una biblioteca de vectores de expresión, cada vector de expresión que codifica el ARNhc de acuerdo con una cualquiera de las reivindicaciones 1 a 7, está unido operativamente a un promotor de ARN polimerasa.

10. Una célula de mamífero aislada que comprende el ARNhc de acuerdo con una cualquiera de las reivindicaciones 1 a 7.

11. Un mamífero no humano transgénico cuyo genoma comprende una secuencia que codifica el ARNhc de acuerdo con una cualquiera de las reivindicaciones 1 a 7.

12. Un procedimiento *in vitro* para atenuar la expresión de un gen diana en una célula de mamífero, comprendiendo el procedimiento introducir en la célula de mamífero un vector de expresión que comprende una secuencia que codifica la molécula de ARN de horquilla corta (ARNhc) de una cualquiera de las reivindicaciones 1 a 7, en el que la molécula de ARNhc se expresa en la célula de mamífero en una cantidad suficiente para atenuar la expresión del gen diana de una manera específica de secuencia, por lo que se inhibe la expresión del gen diana.

13. El vector de expresión de la reivindicación 8 para su uso en la atenuación de la expresión de un gen diana en una célula de mamífero.

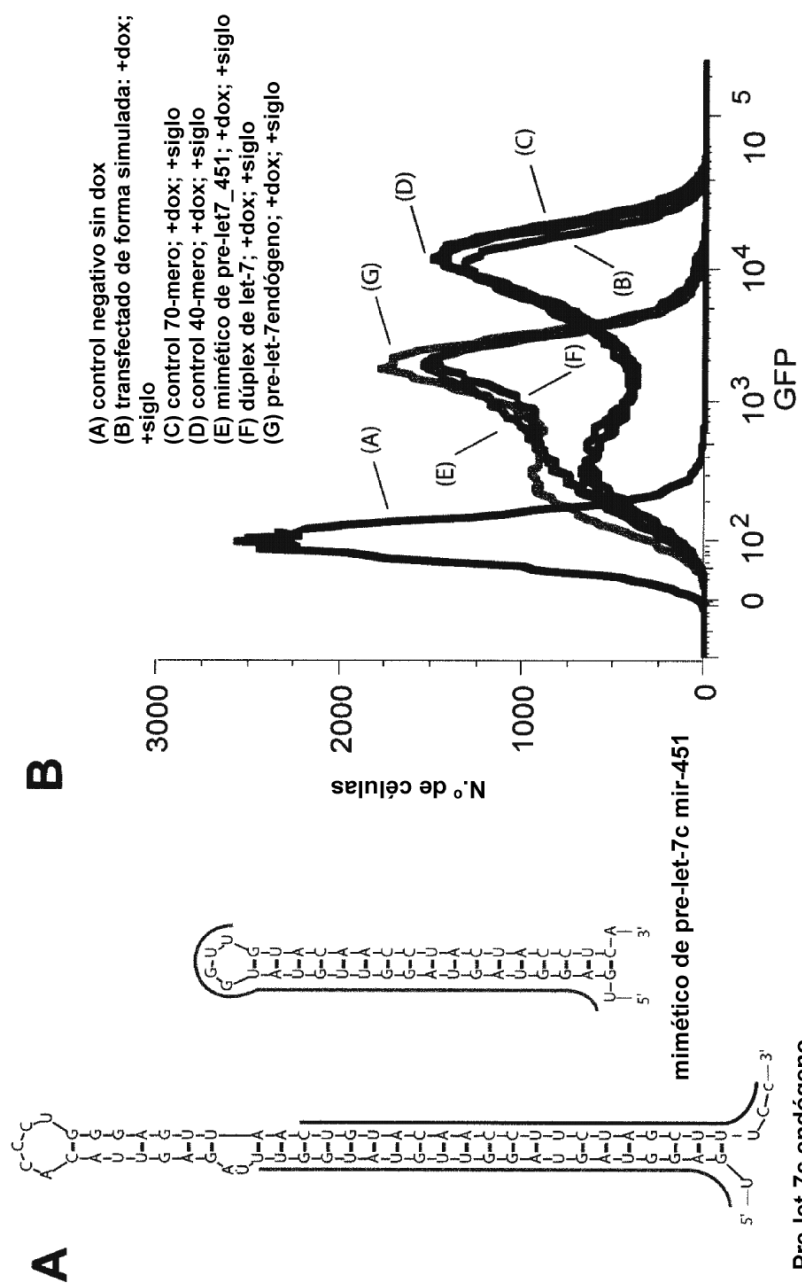
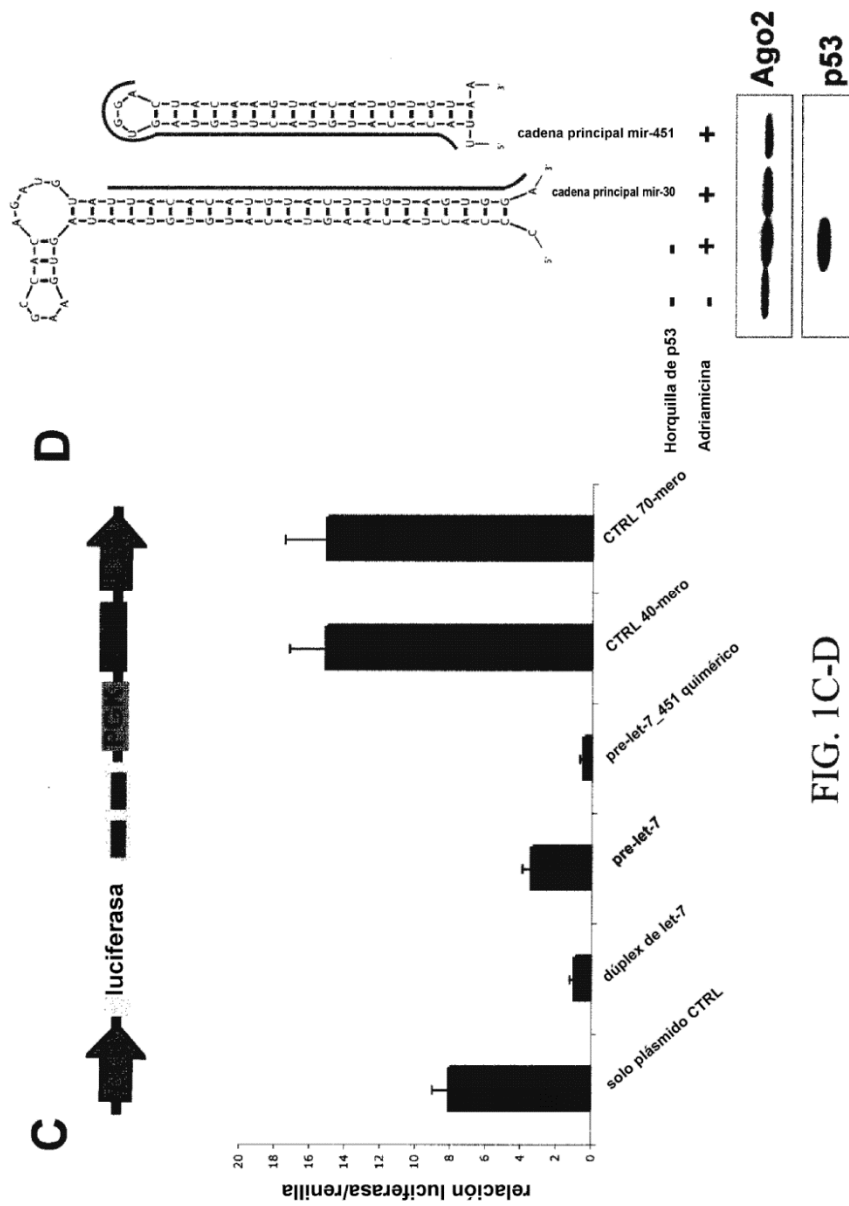


FIG. 1A-B



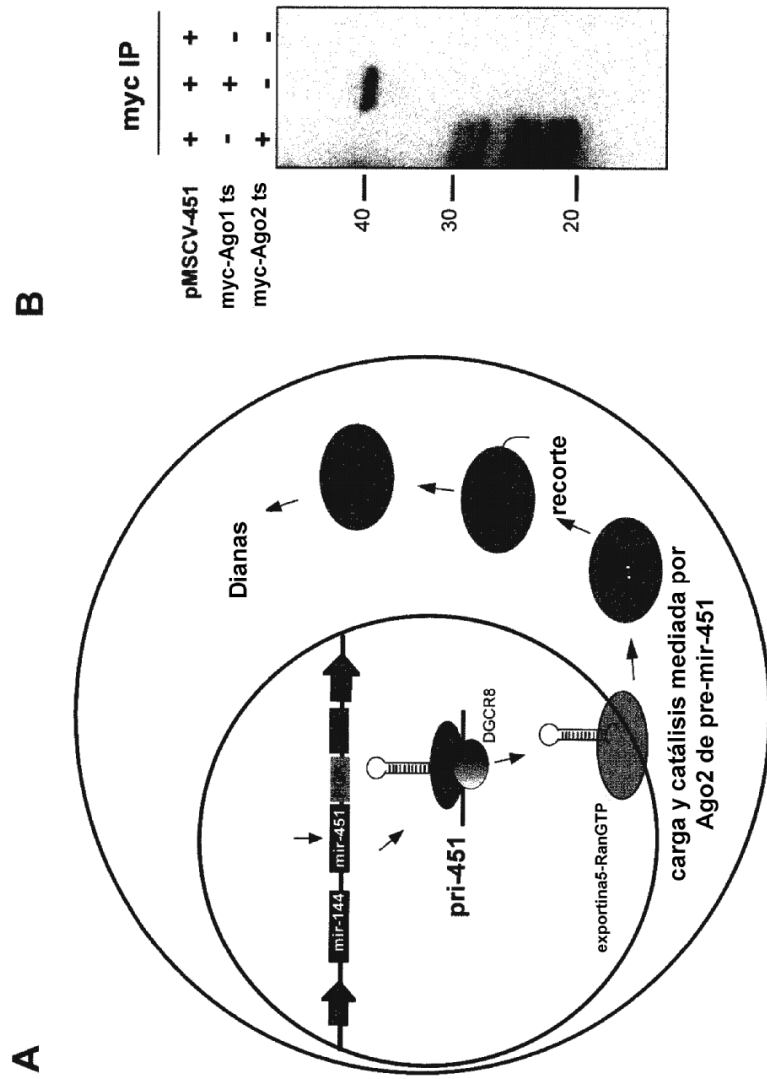


FIG. 2

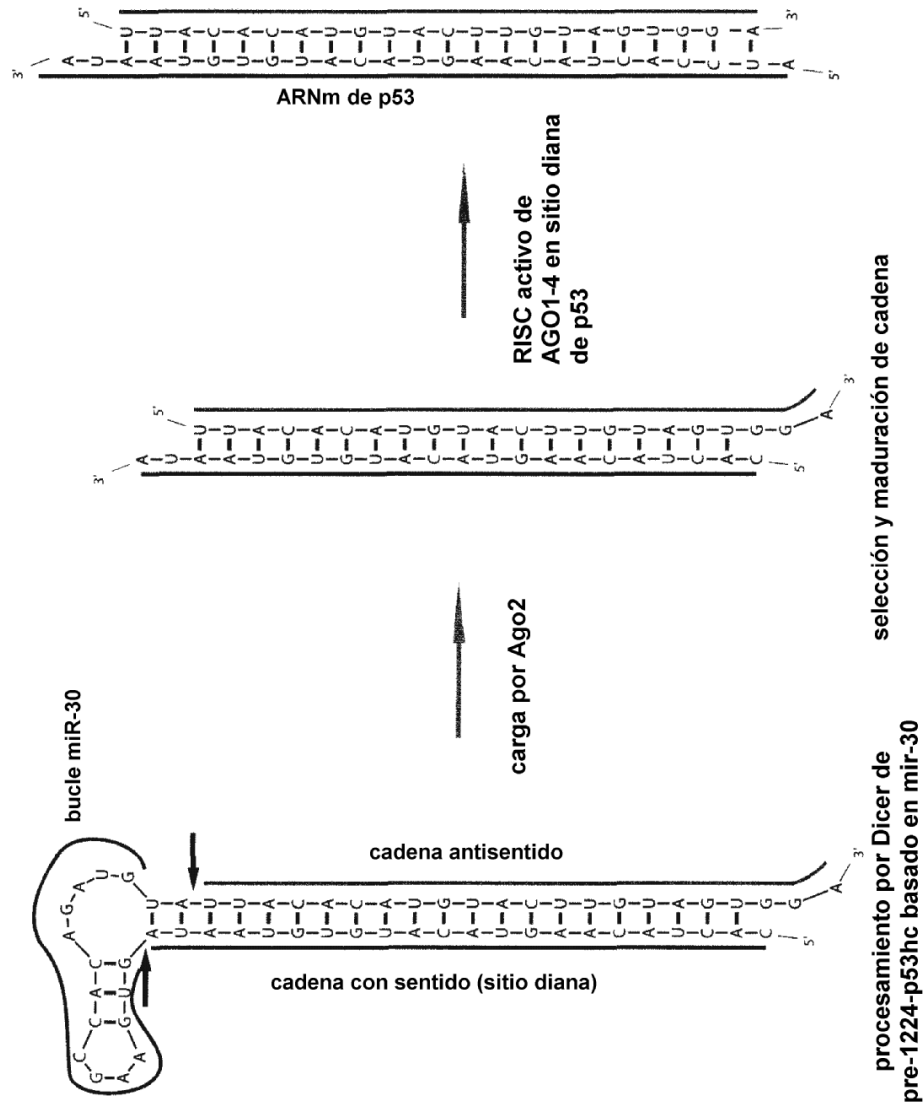


FIG. 3A

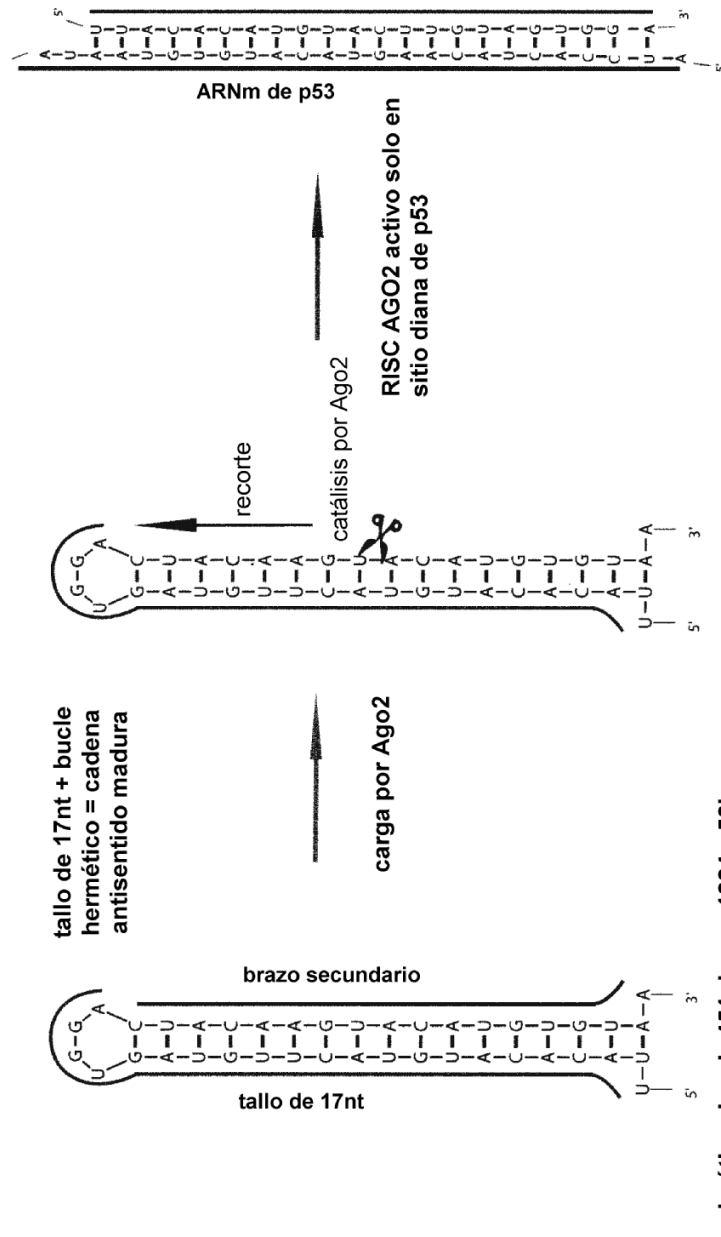
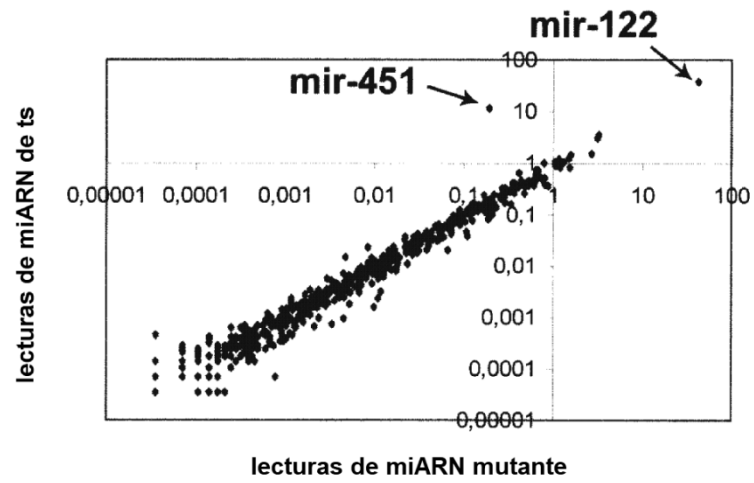


FIG. 3B

a



b

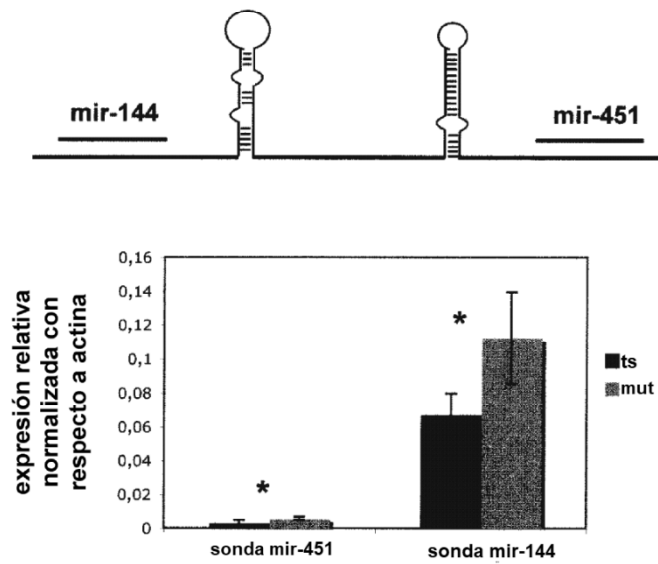


FIG. 4A-B

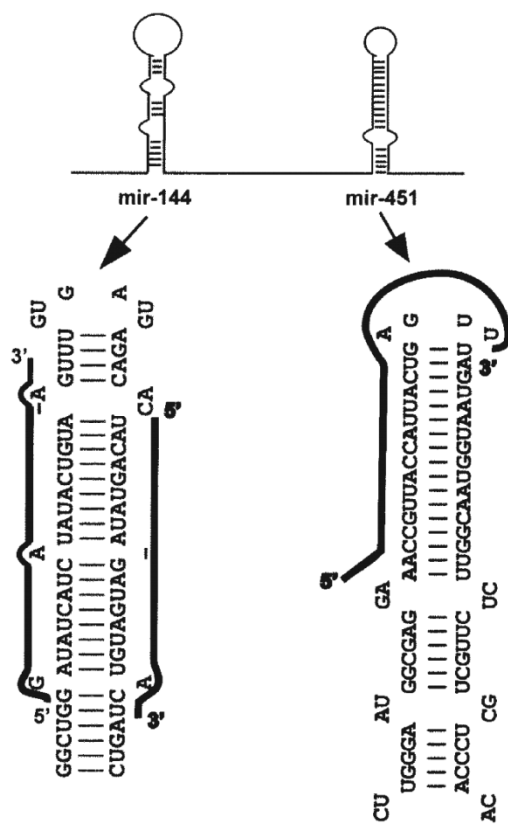


FIG. 4C

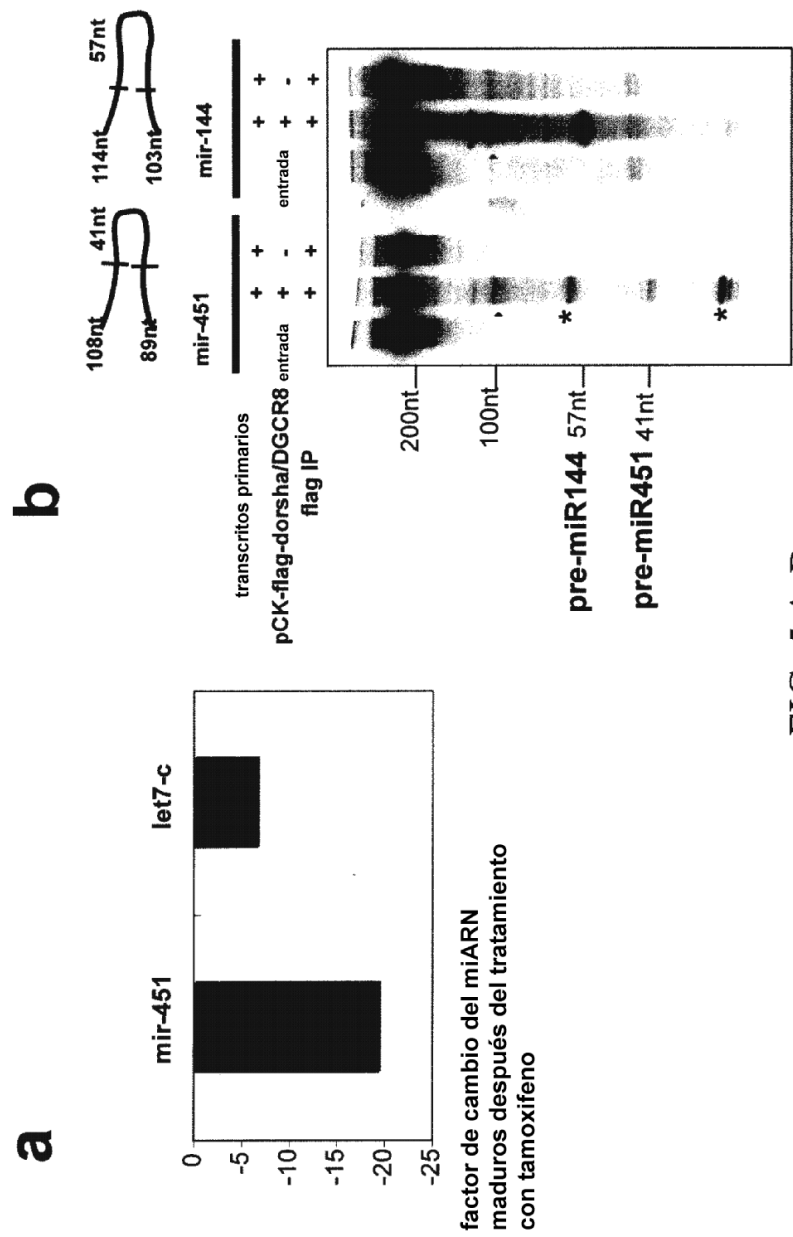


FIG. 5 A-B

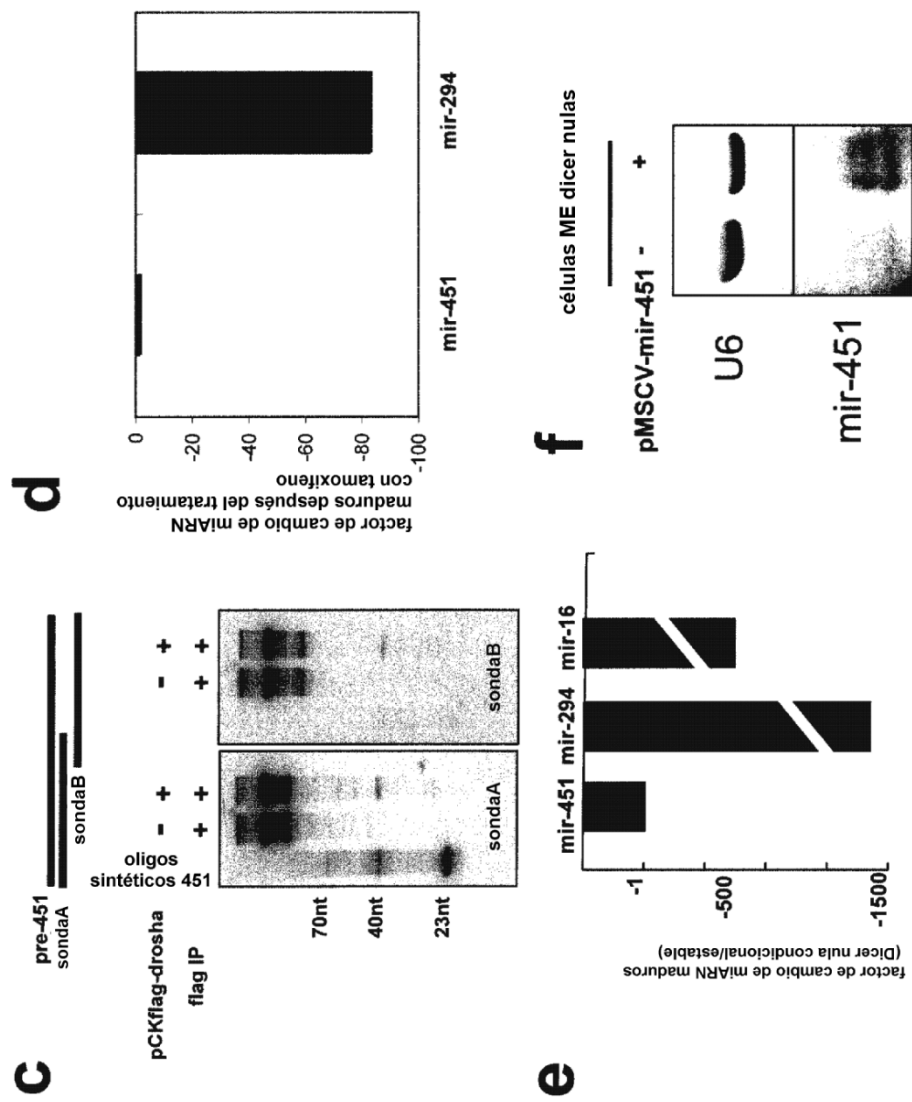
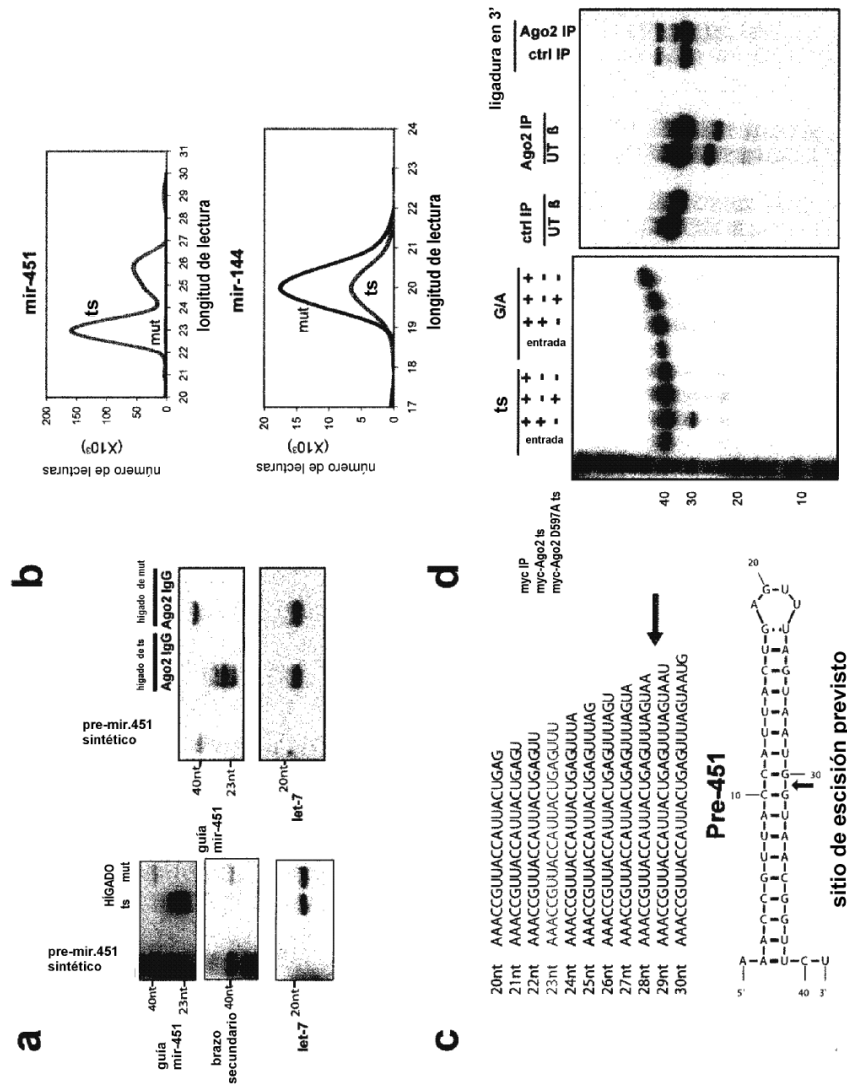


FIG. 5 C-F



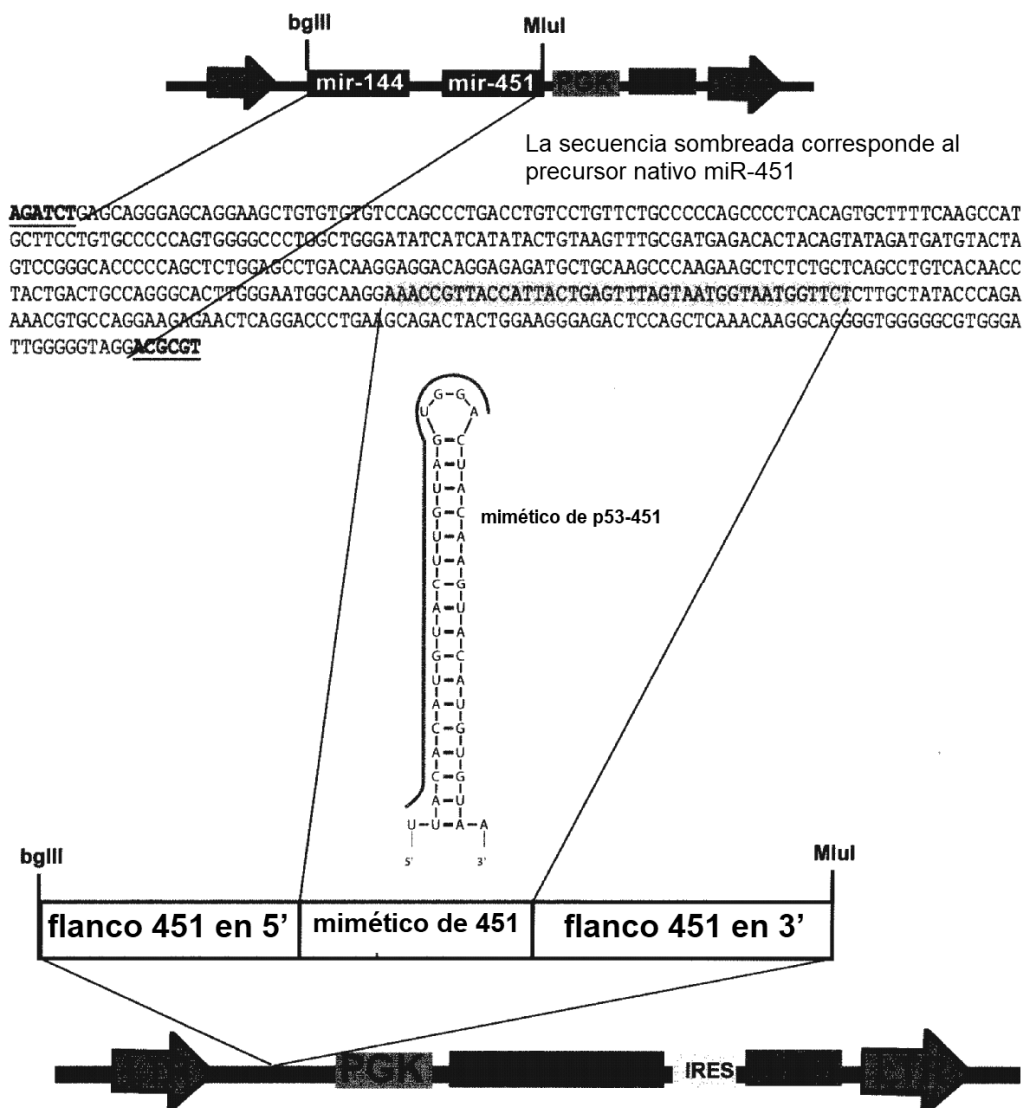


FIG. 7

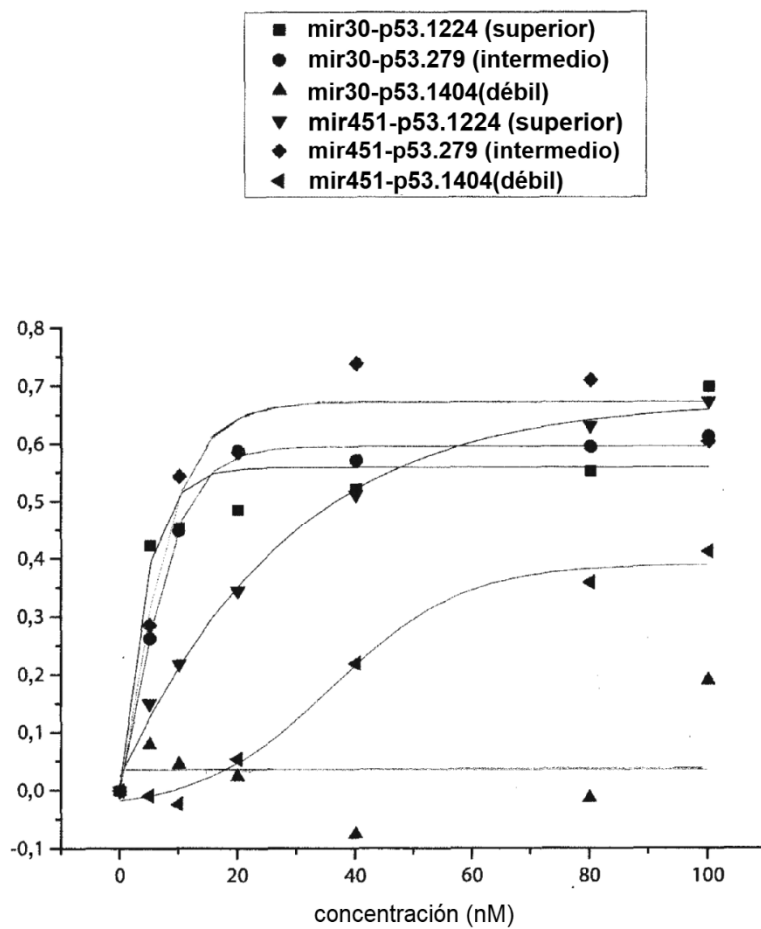
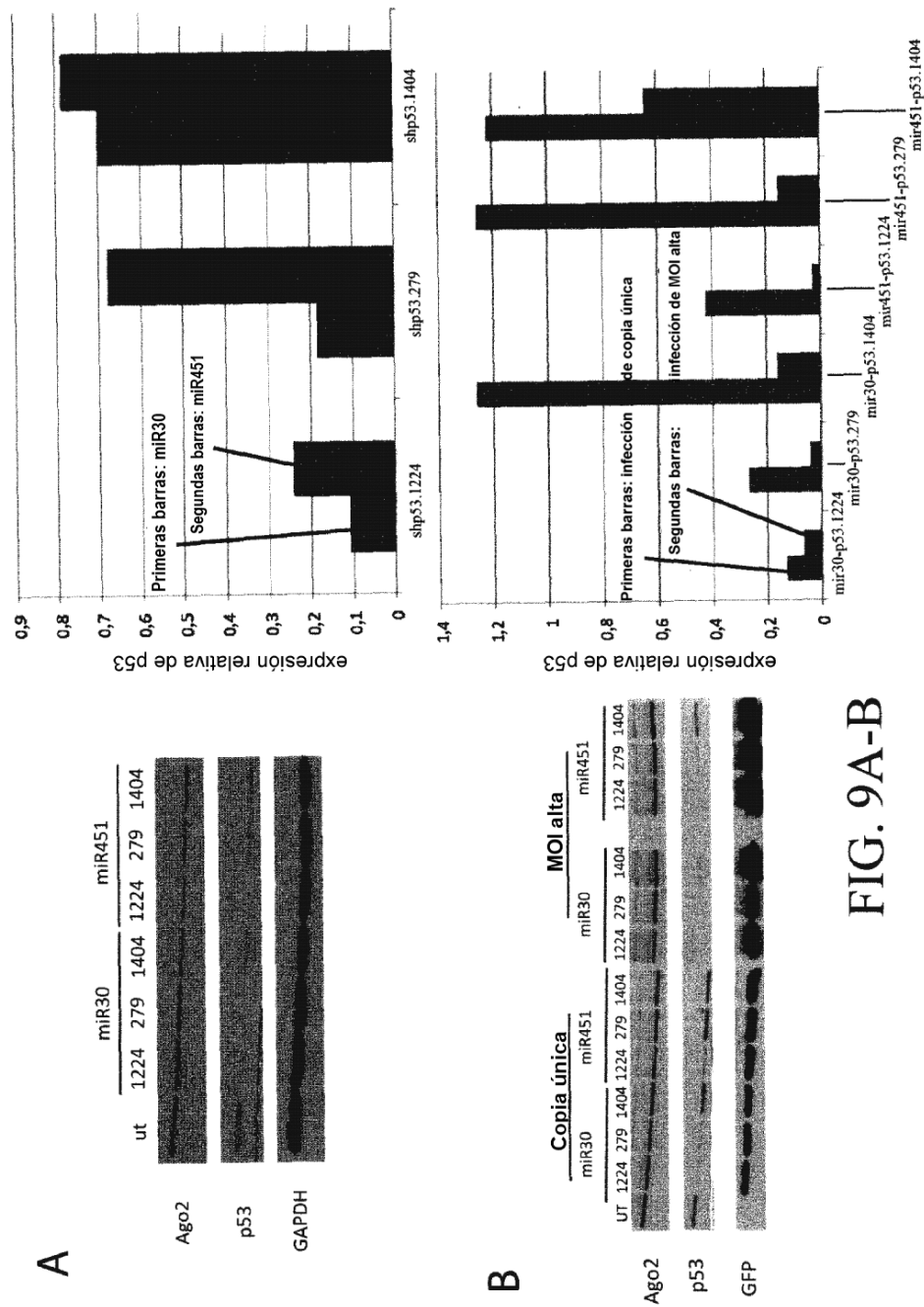


FIG. 8



C

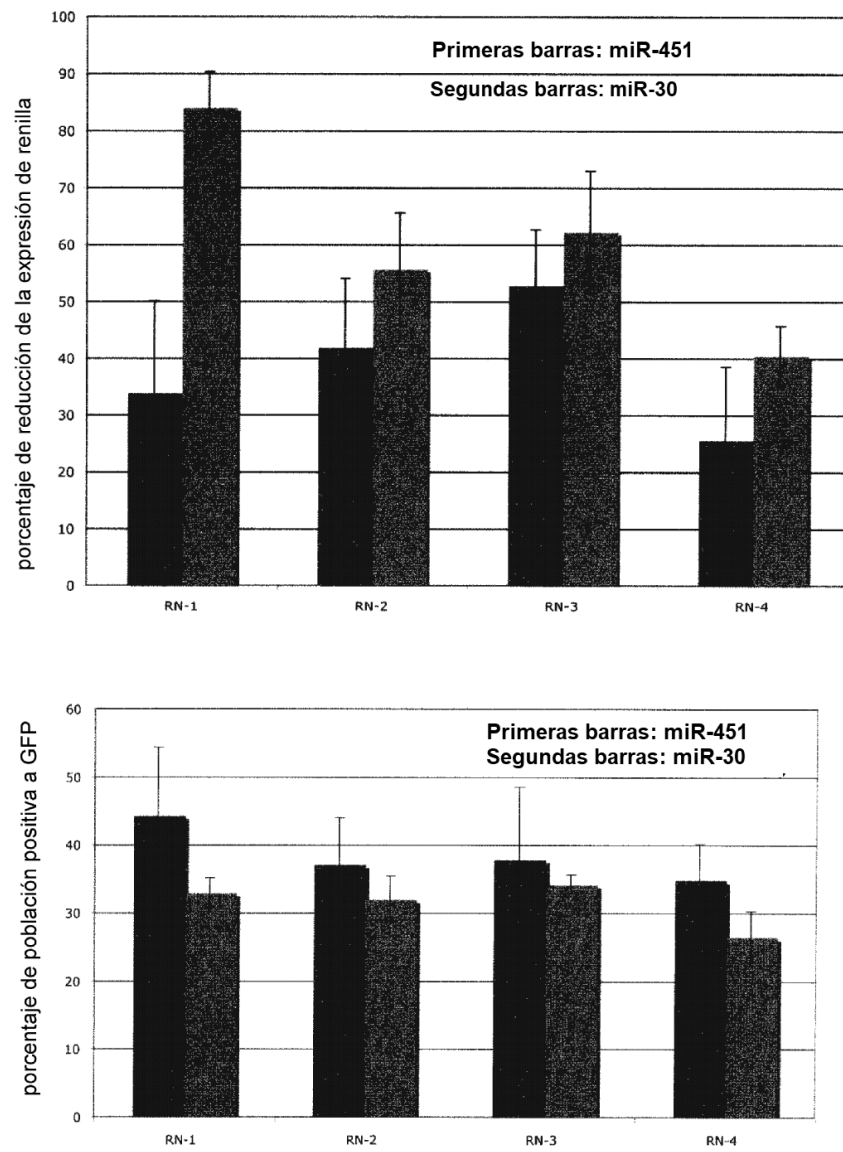


FIG. 9C

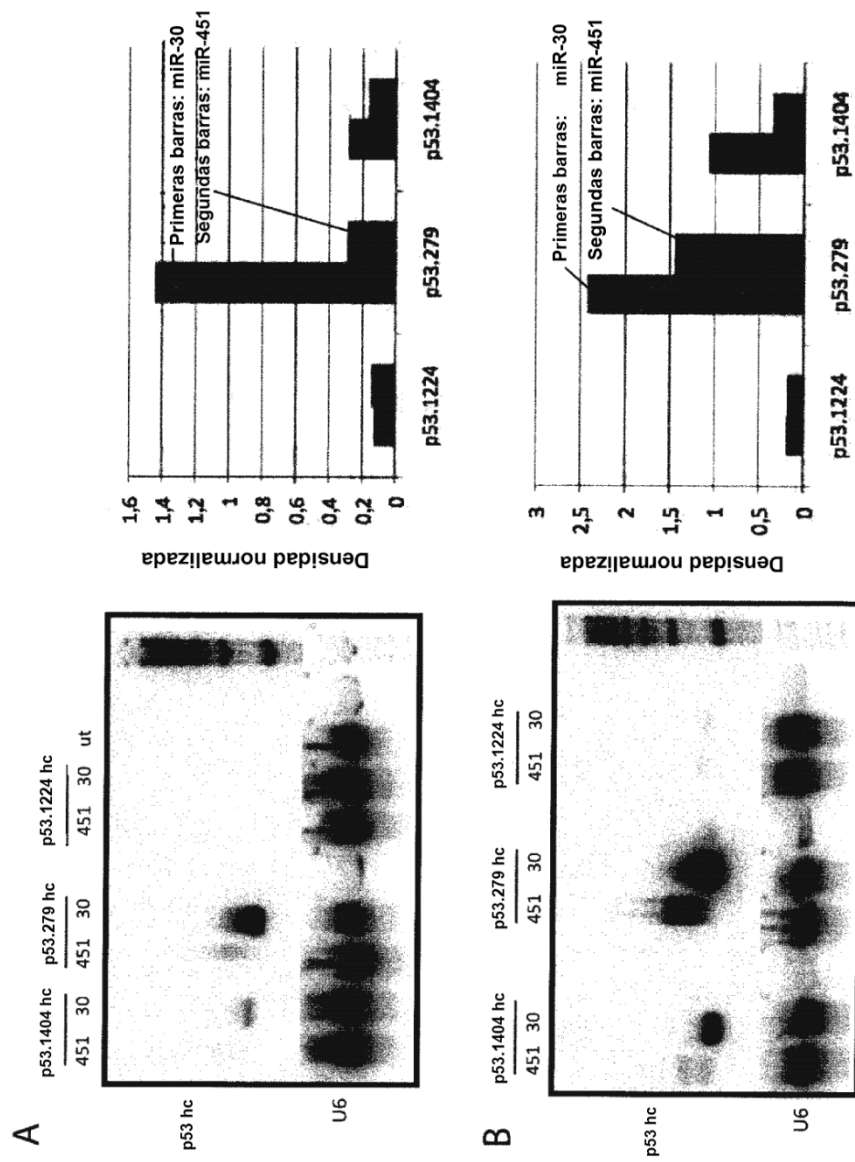


FIG. 10 A-B

C

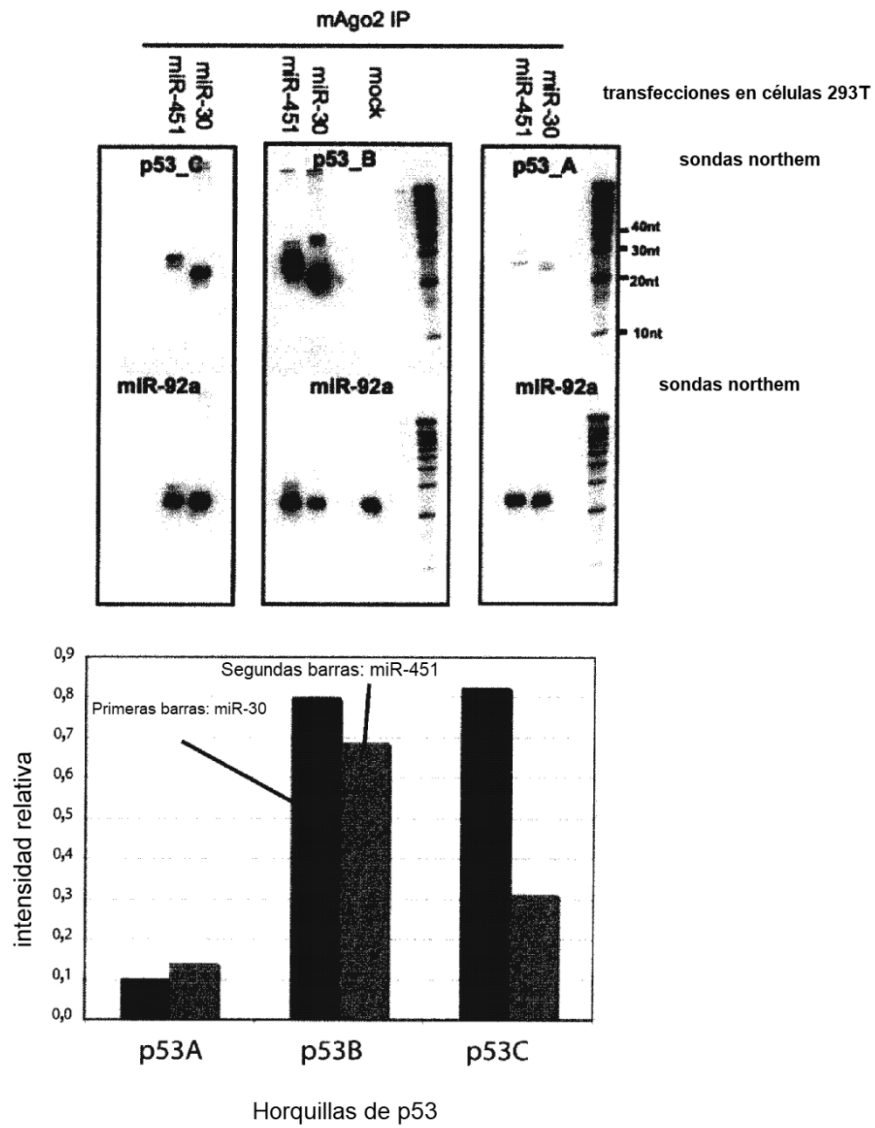


FIG. 10C

D

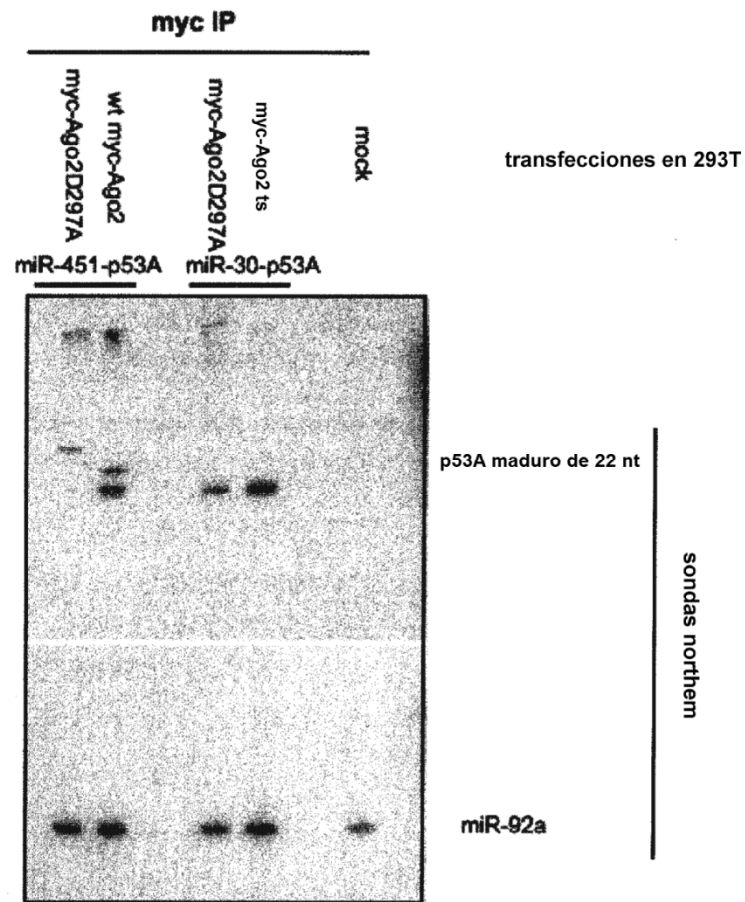


FIG. 10D

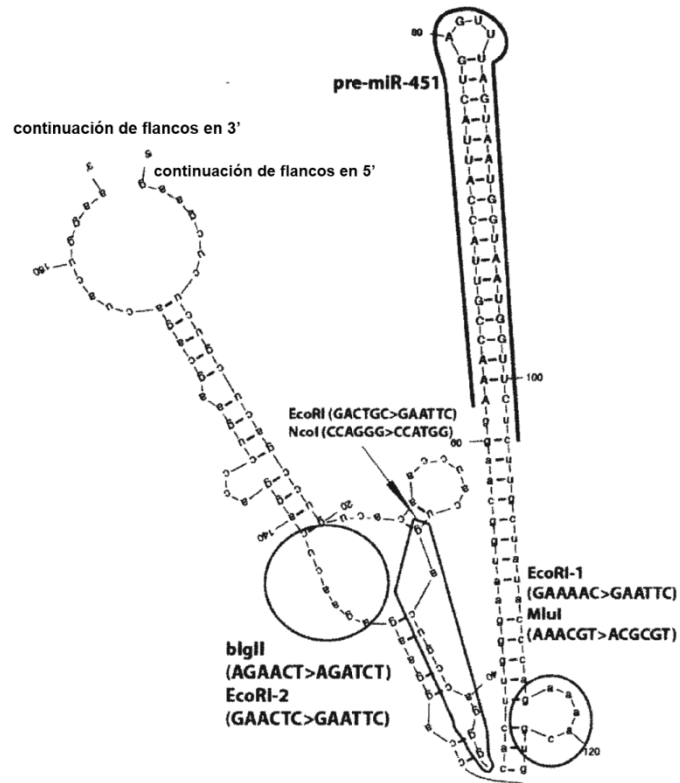


FIG. 11A

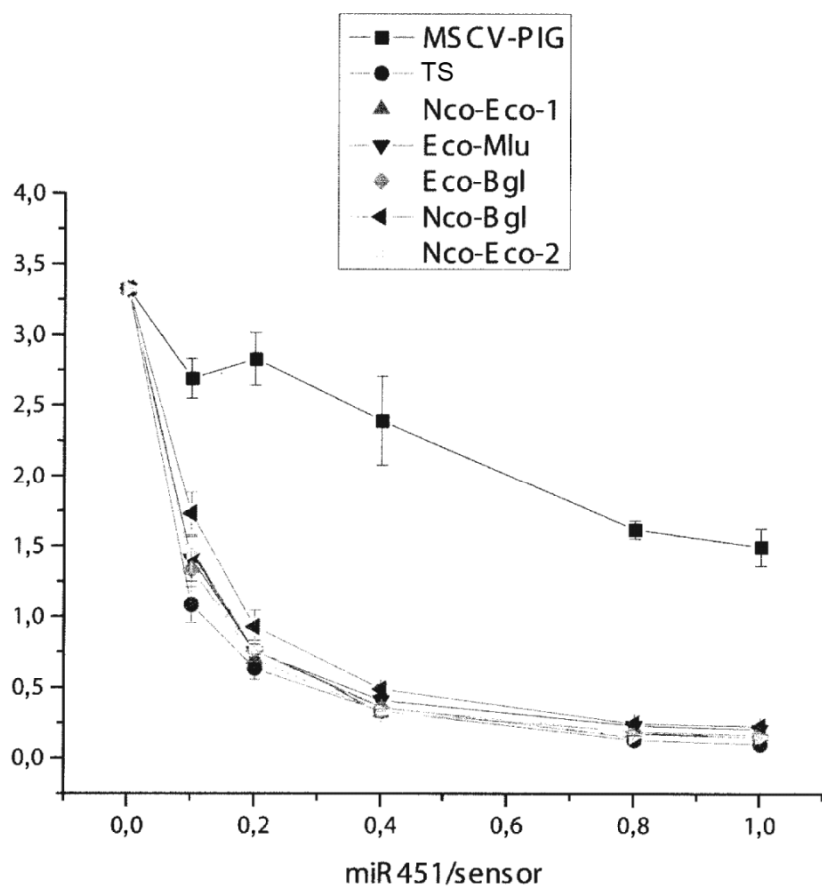
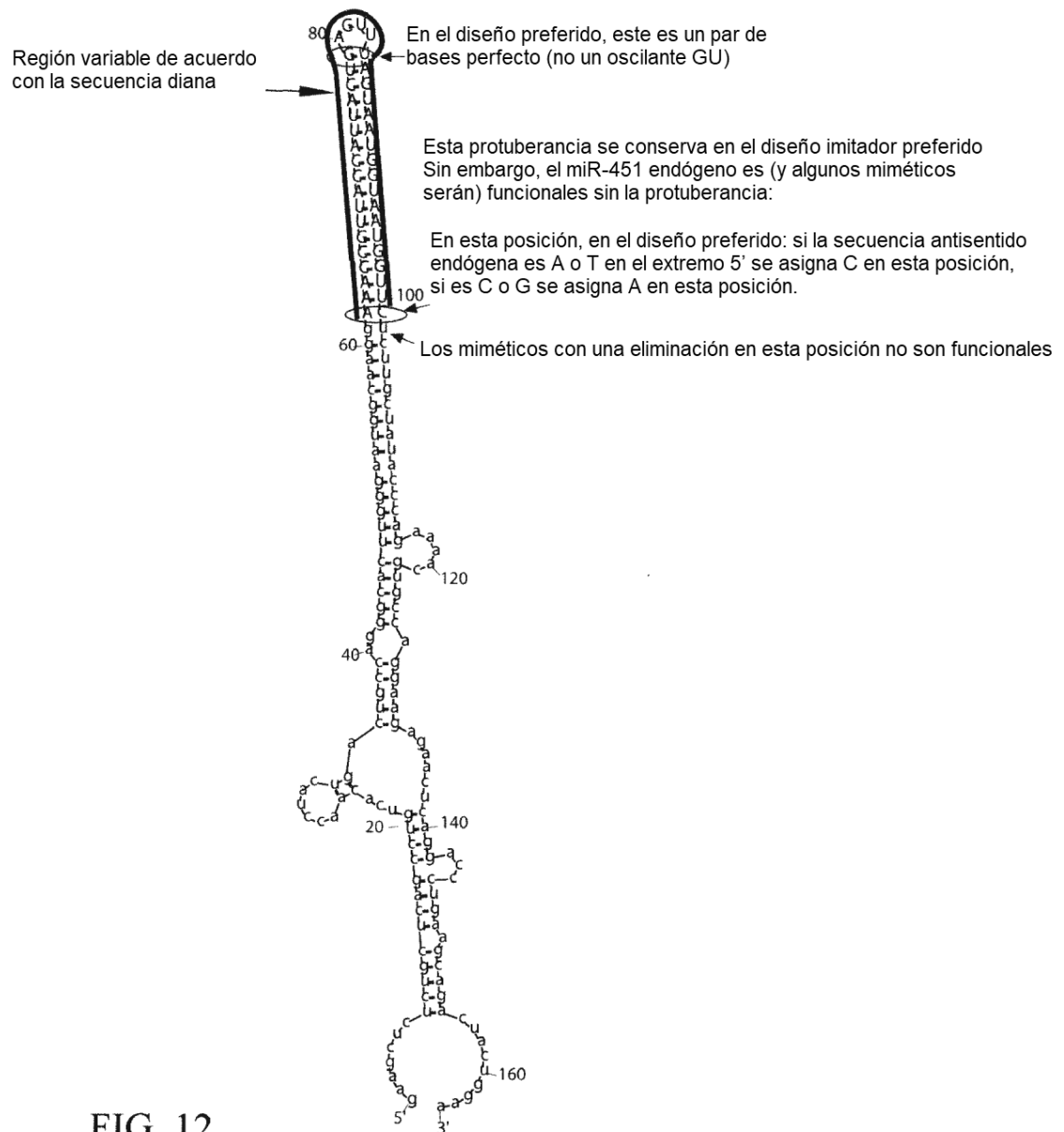


FIG. 11B



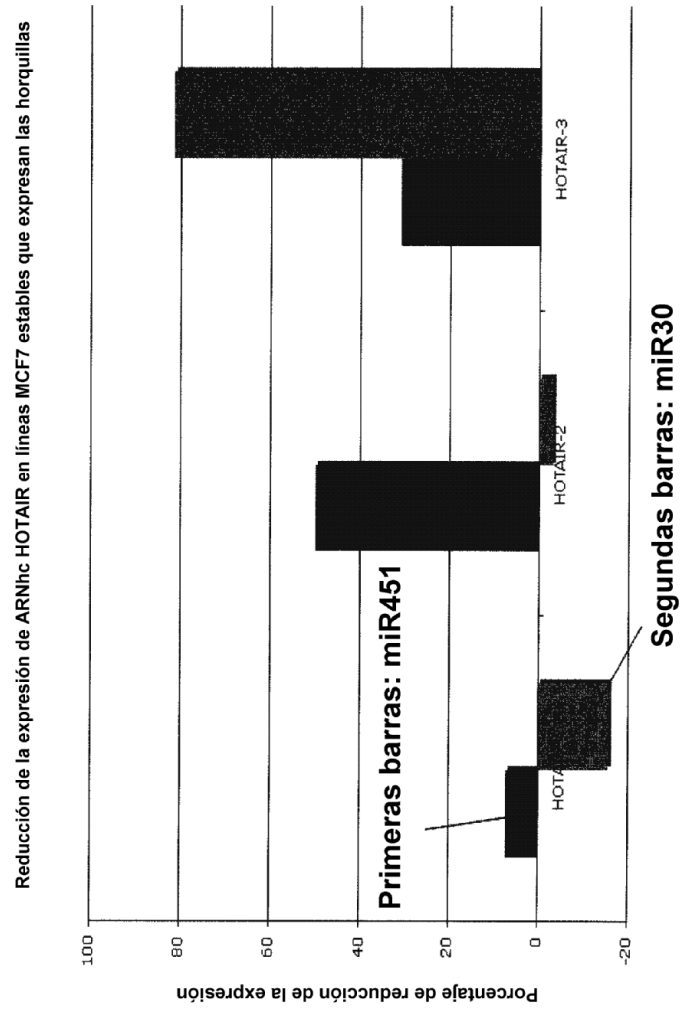


FIG. 13A

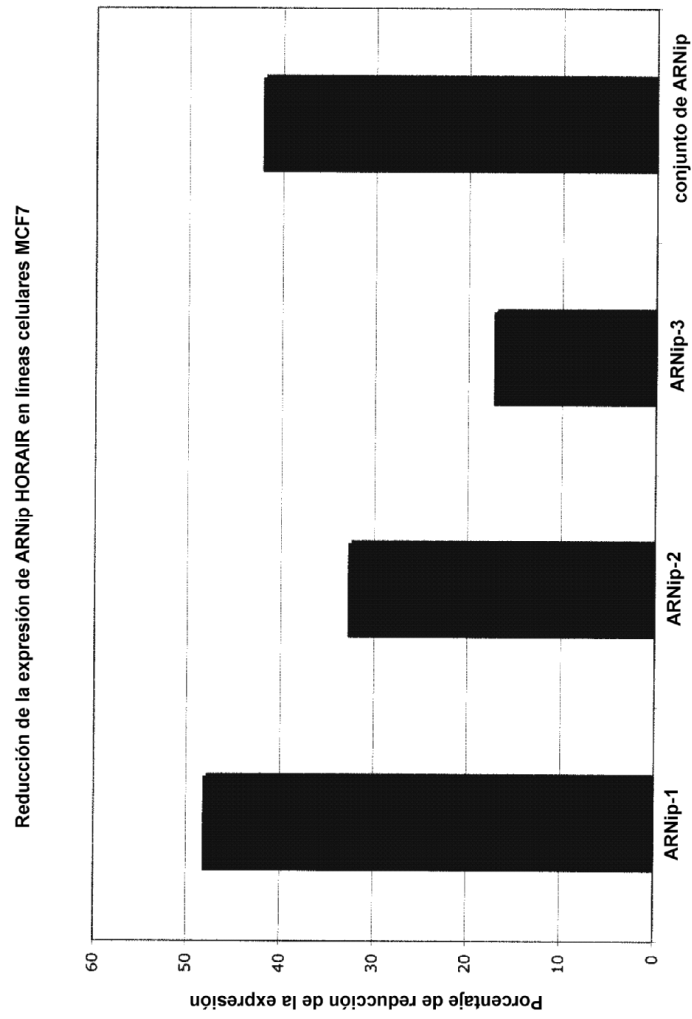


FIG. 13B