

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 379 634**

51 Int. Cl.:
C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **05713739 .0**
96 Fecha de presentación: **17.02.2005**
97 Número de publicación de la solicitud: **1725573**
97 Fecha de publicación de la solicitud: **29.11.2006**

54 Título: **Métodos y materiales que usan sondas de señalización**

30 Prioridad:
18.02.2004 US 546075 P

45 Fecha de publicación de la mención BOPI:
30.04.2012

45 Fecha de la publicación del folleto de la patente:
30.04.2012

73 Titular/es:
**CHROMOCELL CORPORATION
685 U.S. HIGHWAY ONE
NORTH BRUNSWICK, NJ 08902, US**

72 Inventor/es:
**SHEKDAR, Kambiz;
SAWCHUK, Dennis J. y
MONTEZ, Jason M.**

74 Agente/Representante:
de Elizaburu Márquez, Alberto

ES 2 379 634 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

Métodos y materiales que usan sondas de señalización

Antecedentes de la Invención

5 Las sondas de ácido nucleico que reconocen e informan de la presencia de una secuencia específica de ácido nucleico se han usado para detectar ácidos nucleicos específicos principalmente en reacciones *in vitro*. Véase, por ejemplo, Patente U.S. 5.925.517. Un tipo de sonda se diseña para tener una estructura en forma de horquilla, con una cadena central de nucleótidos complementaria a la secuencia diana y extremos que comprenden secuencias cortas mutuamente complementarias. Véase, por ejemplo, Tyagi y Kramer, Nature Biotechnology, 14, 303-308 (1996).

10 Un extremo de la sonda en forma de tallo-bucle se une covalentemente a un fluoróforo y el otro a un resto apantallador. Cuando están en su estado nativo con los extremos hibridados, la proximidad del fluoróforo y el apantallador es tal que se produce una fluorescencia relativamente baja o esencialmente no se produce fluorescencia. La sonda tallo-bucle experimenta un cambio conformacional cuando hibrida con su ácido nucleico diana que resulta en el cambio detectable en la producción de fluorescencia por el fluoróforo. Los investigadores han usado la sonda en forma de horquilla para realizar la visualización *in situ* de ARN mensajero (Matsuo, 1998, Biochim. Biophys. Acta 1379: 178-184) en células vivas. La Patente U.S. 6.692.965 se refiere al uso de balizas moleculares que reconocen un ARNm diana. Dichas balizas pueden usarse conjuntamente con FACS para separar células.

Resumen de la Invención

20 La presente invención describe métodos y composiciones que comprenden sondas nuevas de señalización para analizar o aislar células o generar líneas celulares que expresan uno o más ARN. El método se basa en la detección de la señal producida por las sondas después de su hibridación con la secuencia diana. El ARN puede introducirse en las células mediante una construcción de ADN, o puede sospecharse que las células expresan el ARN endógenamente. La construcción de ADN puede codificar además una secuencia etiqueta y la sonda de señalización es complementaria a la secuencia etiqueta. En una realización, las células aisladas o las líneas celulares generadas son funcionalmente nulas para la expresión o tienen una expresión reducida de una o más proteínas o ARN preseleccionados. La invención también describe un método para generar animales transgénicos usando células que se aíslan según los métodos descritos.

30 La invención describe sondas de señalización usadas en un método para cuantificar el nivel de expresión de uno o más transcritos de ARN. Además, las sondas de señalización pueden usarse en un método para identificar un compuesto o secuencia de ARN que modula la transcripción de al menos un ARN preseleccionado. En otra realización, las sondas de señalización pueden usarse en un método para identificar eventos de recombinación genética en células vivas. La sonda de señalización comprende una o más cadenas de nucleótidos, en el que la sonda de señalización comprende nucleótidos que son complementarios a un ácido nucleico diana (por ejemplo, ARN) de interés y en el que la sonda de señalización comprende además una pareja de interacción que comprende dos restos. La estructura de la sonda de señalización es tal que cuando la sonda de señalización no está hibridada con la secuencia diana, los dos restos de la pareja de interacción están físicamente localizados de manera tal que no se produce señal o se produce señal de fondo. Cuando la sonda de señalización hibrida con la secuencia diana, los dos restos están de tal manera que se produce una señal. Alternativamente, los restos de la sonda de señalización pueden ser tales que cuando la sonda no está hibridada con la diana, se produce una señal particular y se produce una señal diferente después de la hibridación de la sonda con la secuencia diana. Los dos restos de la pareja de interacción pueden estar unidos a uno o más extremos de una o más cadenas de la sonda de señalización. Alternativamente, los restos pueden estar incorporados internamente en una o más cadenas de la sonda de señalización. Los nucleótidos de la sonda de señalización también pueden estar modificados.

45 La presente invención también describe sondas proteasa. En una realización, la sonda de señalización o proteasa comprende dos cadenas separadas de ácido nucleico o ácido nucleico modificado, una o más partes de las cuales hibridan entre sí, y al menos un extremo de una cadena es adyacente a un extremo de la otra cadena. El ácido nucleico puede ser ADN o ARN. Para la sonda de señalización con dos cadenas separadas, en una realización, una cadena tiene al menos un resto apantallador en un extremo y la otra cadena tiene al menos un fluoróforo en el extremo adyacente. Para la sonda proteasa, en una realización, una cadena tiene al menos una enzima proteolítica en un extremo y la otra cadena tiene al menos un inhibidor de la enzima proteolítica en el extremo adyacente.

50 En otra realización, la sonda de señalización o proteasa se diseña para comprender al menos una región mutuamente complementaria y al menos una región no complementaria. En una realización, al menos una región no complementaria puede diseñarse para formar una región bucle. En una realización, la sonda se diseña para formar al menos una estructura tallo-bucle. En otra realización, la sonda forma una estructura "dumbbell" o una estructura de unión de tres brazos. En una realización, la sonda de señalización tiene al menos un fluoróforo y al menos un resto apantallador en

cada extremo de la cadena. En una realización, la sonda proteasa tiene una enzima proteolítica y un inhibidor de la enzima proteolítica en cada extremo de la cadena.

5 En otra realización, la sonda de señalización o proteasa está químicamente modificada. Uno o más de los núcleos de tipo azúcar-fosfodiéster, 2'OH y base purina o pirimidina está modificado. En una realización, el núcleo de desoxirribosa se reemplaza por ácido nucleico peptídico.

10 En una realización, la secuencia etiqueta es un ARN estructural, es decir, el ARN tiene estructura secundaria, preferiblemente una estructura de unión de tres brazos. En una realización, la secuencia etiqueta comprende la estructura o secuencia según la Figura 42 A, B o C. La presente invención también proporciona una construcción de ADN que comprende al menos un ADN que codifica al menos un ARN de interés y la secuencia etiqueta. La invención también describe vectores y células que comprenden la construcción de ADN.

Particularmente, la presente invención proporciona:

[1] Un método para aislar una pluralidad de células, en el que al menos un subconjunto de las células expresa un ARN que no se expresa por otro subconjunto de las células, que comprende las etapas de:

15 introducir en las células una pluralidad de ADN que codifican una pluralidad de ARN diferentes, en el que cada ADN codifica además una secuencia etiqueta de ácido nucleico, y

en el que al menos un subconjunto de la pluralidad de los ADN codifica la misma secuencia etiqueta de ácido nucleico;

exponer dichas células a una misma sonda de señalización que produce una señal detectable después de hibridar con dicha misma secuencia etiqueta de ácido nucleico; y

aislar dichas células que producen la señal.

20 [2] El método de [1], en el que la pluralidad de ARN diferentes forma una biblioteca de expresión.

[3] El método de [2], que comprende además crecer separadamente células aisladas individualmente para generar una pluralidad de líneas celulares separadas.

[4] El método de [2], que comprende además combinar las células aisladas.

[5] El método de [4], que comprende además crecer las células combinadas.

25 [6] El método de una cualquiera de [1] a [5], en el que la pluralidad de ARN diferentes, o proteínas codificadas por la pluralidad de ARN diferentes, se seleccionan del grupo que consiste en ARN o proteínas en la misma ruta biológica o relacionada, ARN o proteínas que actúan aguas arriba o abajo entre sí, ARN o proteínas que tienen una función moduladora, activadora o represora entre sí, ARN o proteínas que son dependientes entre sí para función o actividad, ARN o proteínas que son componentes del mismo complejo y proteínas de la misma familia de proteínas.

30 [7] El método de [1] para aislar células que expresan dos o más bibliotecas de expresión de ARN, que comprende las etapas de:

introducir en las células una pluralidad de ADN que codifican una primera biblioteca de expresión de ARN, en el que cada ADN codifica además una primera secuencia etiqueta de ácido nucleico,

35 y en el que al menos un subconjunto de la pluralidad de ADN codifica la misma primera secuencia etiqueta de ácido nucleico;

introducir en las células una pluralidad de ADN que codifican una segunda biblioteca de expresión de ARN, en el que cada ADN codifica además una segunda secuencia etiqueta de ácido nucleico y en el que al menos un subconjunto de la pluralidad de ADN codifica la misma segunda secuencia etiqueta de ácido nucleico;

40 exponer las células a una primera sonda de señalización que produce una señal detectable después de hibridar con dicha primera secuencia etiqueta de ácido nucleico;

exponer las células a una segunda sonda de señalización que produce una señal detectable después de hibridar con dicha segunda secuencia etiqueta de ácido nucleico; y aislar las células que producen ambas señales.

[8] El método de una cualquiera de [1] a [7], en el que la secuencia etiqueta de ácido nucleico comprende múltiples secuencias diana, en el que una sonda de señalización hibrida con cada secuencia diana.

45 [9] El método de una cualquiera de [1] a [8], en el que el ADN codifica múltiples secuencias etiqueta de ácido nucleico.

- [10] El método de una cualquiera de [1] a [9], en el que el ADN que codifica dicha secuencia etiqueta de ácido nucleico está
- (a) en marco con el ADN que codifica dicho ARN; o
 - (b) fuera de marco con el ADN que codifica dicho ARN.
- 5 [11] El método de una cualquiera de [1] a [10], en el que dicho ADN codifica un ARN antisentido, un ARNsh o un ARNsi.
- [12] El método de una cualquiera de [1] a [11], en el que el ADN codifica además un marcador de selección y en el que el método comprende además la etapa de seleccionar las células usando el marcador de selección después de introducir el ADN en las células pero antes de exponer dichas células a la sonda de señalización.
- 10 [13] El método de una cualquiera de [1] a [12], en el que dicho ADN está unido de manera operativa a un promotor condicional.
- [14] El método de [13], en el que el ARN codificado por el ADN, o la proteína codificada por el ARN, es letal o perjudicial para la célula cuando se expresa.
- [15] El método de [13] o [14] que comprende además la etapa de añadir a las células un compuesto que modula la expresión de dicho ARN o pluralidad de ARN antes de la etapa de exposición.
- 15 [16] El método de una cualquiera de [1], [2] y [4] a [15] que comprende además la etapa de
- (a) cultivar las células aisladas; o
 - (b) generar una pluralidad de líneas celulares por cultivo de las células aisladas.
- [17] El método de [1], en el que cada ADN de dicha pluralidad de ADN codifica además una segunda secuencia de ADN que codifica un ARN preseleccionado.
- 20 [18] El método de [17], en el que dicha pluralidad de ADN codifica una pluralidad de ARN de ensayo variables.
- [19] Un método para identificar un compuesto que activa un promotor condicional, que comprende las etapas de:
- añadir un compuesto de ensayo a las células aisladas por el método de [14] o [15];
 - ensayar para la presencia del ARN expresado bajo el control del promotor condicional; e
 - identificar el compuesto de ensayo como un compuesto que activa el promotor condicional si la célula expresa el ARN bajo el control del promotor condicional.
- 25 [20] Un método para identificar un ARN de ensayo que activa un promotor condicional que comprende las etapas de:
- ensayar las células aisladas por el método de [19] para la presencia del ARN bajo el control del promotor condicional;
 - obtener las células que expresan el ARN bajo el control del promotor condicional; e
 - identificar el ARN de ensayo que activa el promotor condicional.
- 30 En la presente memoria se describe un método para aislar células que expresan un ARN que comprende las etapas de proporcionar células que expresan el ARN, exponer las células a una sonda de señalización que produce una señal detectable después de hibridar con el ARN y aislar las células que producen la señal. En una realización, el ARN es un ARN endógeno. En otra realización, el ARN endógeno se expresa como resultado de ADN que se introduce en la célula,
- 35 por ejemplo, el ADN comprende secuencias amplificadoras o promotoras u otras que inducen la expresión de un ARN endógeno. Por ejemplo, el ADN puede codificar una proteína, por ejemplo, un factor de transcripción que induce la expresión de un ARN. En otra realización más, el ARN está codificado por un ácido nucleico que se introduce en las células. El ADN que se introduce en las células puede codificar además una secuencia etiqueta en la que la sonda de señalización puede estar dirigida opcionalmente a la secuencia etiqueta y/o ARN. La secuencia etiqueta puede estar en marco o fuera de marco con el marco de lectura abierto del ARN. En una realización, el método incluye detectar la
- 40 expresión tanto de un ARN exógeno o heterólogo como de un ARN endógeno. Estos métodos pueden ponerse en práctica para detectar múltiples ARN y/o etiquetas al mismo tiempo. Los ARN pueden ser ARN iguales o diferentes. En aquellas realizaciones en las que el método se usa para detectar más de un ARN usando más de una sonda de señalización, las señales producidas por las diferentes sondas de señalización pueden ser iguales o pueden ser diferentes entre sí.

En la presente memoria también se describe un método para aislar células que comprende más de una copia de un ADN exógeno o heterólogo. Dicho método incluye las etapas de introducir un ADN que codifica un ARN y una secuencia etiqueta e introducir un ADN que codifica el mismo ARN y una secuencia etiqueta diferente; exponer las células a sondas de señalización que producen una señal detectable a ambas etiquetas, y aislar las células que producen ambas señales.

5 En cualquiera de los métodos descritos en la presente memoria en los que hay más de una etapa de exposición, una o más de las etapas de exposición (es decir, la etapa en la que las células se exponen a la sonda de señalización) pueden realizarse simultáneamente o secuencialmente.

10 En cualquiera de los métodos descritos en la presente memoria, los ARN o proteínas pueden estar en la misma ruta biológica o una relacionada, actuar aguas arriba o abajo entre sí, tener una función moduladora, activadora o represora uno respecto al otro, depender entre sí para función o actividad, ser componentes del mismo complejo, miembros de la misma familia de proteínas, etc.

15 En la presente memoria se describe un método para aislar células que comprende una construcción de ADN que codifica un primer ARN que está bajo el control de un promotor condicional, que comprende las etapas de introducir en las células una construcción de ADN que codifica un ARN bajo el control de un promotor constitutivo, en el que la construcción de ADN codifica además un segundo ARN bajo el control de un promotor condicional, bajo condiciones en las que el segundo ARN no se expresa o se expresa a un nivel bajo; exponer las células a una sonda de señalización que produce una señal detectable después de hibridar con el primer ARN; y aislar las células que producen la señal. La construcción de ADN puede codificar además un ARN de ensayo, en el que, por ejemplo, el ARN de ensayo es variable, por ejemplo, obtenido de una biblioteca de expresión. Estas células pueden usarse para obtener o identificar ARN o compuestos que son capaces de activar el promotor condicional que dirige la expresión del segundo ARN.

20 Esta invención también proporciona métodos para aislar una pluralidad de células en los que un subconjunto de células expresa un ARN que no es expresado por otro subconjunto de células, que comprende las etapas de introducir en las células una pluralidad de ADN que codifica una pluralidad de ARN, en los que al menos un subconjunto de la pluralidad de ARN son diferentes entre sí, exponer las células a una pluralidad de diferentes sondas de señalización, en las que las sondas de señalización producen una señal detectable después de hibridar con uno o más ARN codificados por la pluralidad de ADN y aislar las células que producen la señal. El subconjunto de células puede ser una célula o más de una célula. En una realización, el ADN no codifica ARN sino que resulta en la expresión de ARN endógeno, por ejemplo, el ADN comprende secuencias amplificadoras o promotoras, por ejemplo, el ADN comprende secuencias que inducen la expresión del ARN endógeno. Por ejemplo, el ADN puede codificar una proteína, por ejemplo, un factor de transcripción que induce la expresión de un ARN. El ADN puede codificar además una secuencia etiqueta y la sonda de señalización puede estar dirigida a la secuencia etiqueta y/o la secuencia de ARN. En una realización, la pluralidad de ARN forma una biblioteca de expresión. En otra realización, al menos un subconjunto de los ADN codifica la misma secuencia etiqueta.

35 En la presente memoria también se describe un método para aislar dos o más bibliotecas de ARN de células que comprende las etapas de introducir en las células ADN que codifica una primera biblioteca de expresión de ARN, en la que cada ADN codifica además una primera secuencia etiqueta, introducir en las células ADN que codifica una segunda biblioteca de expresión de ARN, en la que cada ADN codifica además una segunda secuencia etiqueta, exponer las células a una primera sonda de señalización que produce una señal detectable después de hibridar con la primera etiqueta y una segunda sonda de señalización que produce una señal detectable después de hibridar con la segunda secuencia etiqueta y aislar las células que producen ambas señales. Este método también puede realizarse con ADN que codifica bibliotecas de expresión de ARN adicionales en las que se usa una tercera etiqueta, etc.

Cualquiera de los métodos descritos en la presente memoria puede usarse para identificar un compuesto que modula la expresión de un ARN o una pluralidad de ARN añadiendo el compuesto a las células y ensayando para un cambio (incremento o disminución) de la señal producida por la o las sondas de señalización.

45 En la presente memoria se describen métodos para reducir la expresión de una proteína que comprenden las etapas de introducir en las células un ADN que codifica un ARN antisentido o un ARNsh que reduce la expresión de la proteína, exponer las células a una primera sonda de señalización que produce una señal detectable después de hibridar con el ARN antisentido o ARNsh y aislar las células que producen la señal. Este método puede comprender además la etapa de exponer las células a una segunda sonda de señalización que produce una señal detectable después de hibridar con el ARN que codifica la proteína, en el que la ausencia de señal de la segunda sonda de señalización indica que la expresión de la proteína está reducida. También se puede ensayar para expresión reducida de la proteína usando otros métodos, por ejemplo, usando un anticuerpo que se une específicamente a la proteína, usando un ensayo funcional, por ejemplo, ensayando una actividad biológica conocida de la proteína, etc. En una realización de este método, la etapa de exponer las células a la primera sonda de señalización se omite. Cualquiera de los métodos que usan ARNsi también puede realizarse usando ARNsh.

En cualquiera de los métodos descritos en la presente memoria, el ADN que se introduce en las células puede estar unido de manera operativa a un promotor condicional. En una realización, el ARN codificado por el ADN es letal o perjudicial para la célula. El ADN puede comprender además un marcador seleccionable que puede usarse para seleccionar las células que comprenden el ADN.

5 En la presente memoria también se describe un método para cuantificar el nivel de expresión de un ARN en una muestra biológica que comprende las etapas de exponer la muestra biológica a una primera sonda de señalización que produce una señal detectable después de hibridar con el ARN, cuantificar el nivel de la señal en la muestra biológica y correlacionar el nivel de la señal con el nivel de expresión del ARN.

10 En la presente memoria se describe además un método para identificar un compuesto que modula la expresión de un ARN que comprende añadir un compuesto de ensayo a células que expresan el ARN, exponer las células a una sonda de señalización que produce una señal detectable después de hibridar con el ARN y comparar la señal producida por las células expuestas al compuesto de ensayo con la señal producida por las células no expuestas al compuesto de ensayo, en el que un incremento o disminución de la señal producida por las primeras células comparado con la señal producida por las últimas células indica que el compuesto modula la expresión del ARN. En una realización, el ARN está codificado por ADN que se introduce en las células. En una realización, el compuesto es un ARN o proteína.

En la presente memoria se describe un método para identificar un evento de recombinación genética en células vivas que comprende las etapas de exponer las células a una sonda de señalización que produce una señal detectable después de hibridar con un ARN transcrito a partir de una secuencia recombinada, en el que la detección de una célula que produce la señal indica que la célula comprende el evento de recombinación genética.

20 En la presente memoria también se describen células producidas por cualquiera de los métodos. Estas células pueden cultivarse y pueden usarse para generar líneas celulares o una pluralidad de líneas celulares. Las células pueden usarse para varios propósitos, por ejemplo, en un ensayo basado en células o en el que la célula se implanta en un animal, animal no humano, o mamífero. La célula puede ser una célula madre embrionaria, una célula primaria, germinal o madre. La célula también puede ser una célula inmortalizada. Las células pueden ser células endoteliales, epidérmicas, mesenquimales, neurales, renales, hepáticas, hematopoyéticas o inmunes. Las células pueden ser eucariotas, procariotas, de mamífero, levadura, planta, humanas, de primates, bovinas, porcinas, felinas, de roedores, marsupiales, murinas u otras células.

30 Las secuencias etiqueta pueden comprender múltiples secuencias diana, en las que una sonda de señalización hibrida con cada secuencia diana. Las secuencias etiqueta pueden ser un ARN que tiene estructura secundaria. La estructura puede ser una estructura de unión de tres brazos. El ADN puede comprender múltiples secuencias etiqueta. La secuencia etiqueta puede transcribirse como el mismo ARN que el ARN codificado por el ADN o la secuencia etiqueta puede transcribirse como un ARN separado. También se describe una construcción de ADN que comprende una secuencia de ADN que codifica un ARN y una secuencia etiqueta. La secuencia etiqueta puede ser una cualquiera de las descritas en la presente memoria. También se proporcionan células y vectores que comprenden la construcción de ADN.

35 En la presente memoria también se describen bibliotecas de líneas celulares de mamífero que comprenden al menos 1.000, al menos 800, al menos 600, al menos 500, al menos 400, al menos 200, al menos 100 o al menos 50 líneas celulares, en las que cada línea celular comprende una secuencia expresada integrada de manera estable. También se describen bibliotecas de líneas celulares de mamífero que comprenden al menos 500, al menos 400, al menos 300, al menos 200, al menos 100, al menos 50 líneas celulares, en las que cada línea celular comprende al menos dos secuencias integradas de manera estable. También se describen bibliotecas de líneas celulares de mamífero que comprenden al menos 100, al menos 50, al menos 25, al menos 10 líneas celulares, en las que cada línea celular comprende al menos tres secuencias integradas de manera estable. También se describen bibliotecas de líneas celulares de mamífero que comprenden al menos al menos 50, al menos 25, al menos 20, al menos 10 líneas celulares, en las que cada línea celular comprende al menos cuatro secuencias integradas de manera estable. Las secuencias integradas de manera estable en estas líneas celulares pueden carecer además de un marcador de selección, por ejemplo, un gen de resistencia a fármacos. Las secuencias integradas de manera estable pueden tener una identidad de secuencia conocida o desconocida. Estas secuencias pueden tener una homología de secuencia compartida, significancia funcional u origen relacionado. Estas bibliotecas pueden usarse para varios propósitos, por ejemplo, en un ensayo de cribado basado en células.

40 45 50 En la presente memoria también se describe un método para identificar un compuesto que aumenta la detección de dianas en las células usando sondas de señalización que comprende las etapas de introducir una sonda de señalización en células que comprenden una secuencia diana, en el que la sonda de señalización produce una señal detectable después de hibridar con la secuencia diana, exponer las células a un compuesto de ensayo y detectar la señal producida por las células, en el que un incremento de la señal producida por las células expuestas al compuesto de ensayo

comparado con la señal producida por las células no expuestas al compuesto de ensayo indica que el compuesto de ensayo es un compuesto que aumenta la detección de dianas en las células usando sondas de señalización.

En la presente memoria también se describe un método para identificar un compuesto que media o mejora la introducción de sondas de señalización en células que comprende las etapas de exponer células a una sonda de señalización en presencia de un compuesto de ensayo, en el que las células comprenden una secuencia diana y en el que la sonda de señalización produce una señal después de hibridar con la secuencia diana; y detectar la señal producida por las células, en el que un incremento en la señal producida por las células expuestas al compuesto de ensayo comparado con las células no expuestas al compuesto de ensayo indica que el compuesto de ensayo es un compuesto que media o mejora la introducción de sondas de señalización en las células.

Éstos y otros aspectos descritos en la presente memoria se apreciarán a partir de la descripción detallada siguiente.

Descripción Breve de los Dibujos

La Figura 1 y 2 representan sondas de señalización o proteasa con dos cadenas separadas. Los grupos químicos que interaccionan se muestran como óvalos. Los diferentes óvalos presentan una realización de la invención, en la que el óvalo oscuro indica un resto apantallador y los óvalos blancos y grises indican diferentes fluoróforos.

Las Figuras 3, 4 y 5 representan sondas de señalización o proteasa diseñadas para tener una estructura de tallo-bucle. Los grupos químicos que interaccionan se muestran como óvalos. Los diferentes óvalos presentan una realización de la invención, en la que el óvalo oscuro indica un resto apantallador y el óvalo blanco indica un fluoróforo.

La Figura 6 representa sondas de señalización o proteasa con una estructura de unión de tres brazos. Los grupos químicos que interaccionan se muestran como óvalos. Los diferentes óvalos presentan una realización de la invención, en la que el óvalo oscuro indica un resto apantallador y el óvalo blanco indica un fluoróforo.

La Figura 7 representa sondas de señalización o proteasa con una estructura "dumbbell". Los grupos químicos que interaccionan se muestran como óvalos. Los diferentes óvalos presentan una realización de la invención, en la que el óvalo oscuro indica un resto apantallador y el óvalo blanco indica un fluoróforo.

La Figura 8 muestra la secuencia y la conformación nativa predicha de la sonda fluorescente FP1. La secuencia de FP1 comprende bases que están diseñadas para ser complementarias a la secuencia diana y bases flanqueantes adicionales. Las bases flanqueantes están subrayadas. El Panel A muestra la estructura predicha de la secuencia usando programas de plegamiento de ADN según Nucleic Acids Res. 31: 3429-3431 (2003). El Panel B muestra la auto dimerización predicha de la secuencia de FP1 según el programa informático oligoanalyzer 3.0 disponible en <http://biotools.idtdna.com/analyzer/oligocalc.asp>. En los dos Paneles A y B, las bases flanqueantes están sombreadas en gris, los óvalos blancos y negros indican restos de fluoróforo y apantalladores, respectivamente.

La Figura 9 muestra la secuencia y la conformación nativa predicha de la sonda fluorescente FP2. La secuencia comprende bases que están diseñadas para ser complementarias a la secuencia diana y bases flanqueantes adicionales. Las bases flanqueantes están subrayadas. El Panel A muestra la estructura predicha de la secuencia usando programas de plegamiento de ADN según Nucleic Acids Res. 31: 3429-3431 (2003). Es probable que toda o parte de la región sombreada forme pares de bases de Watson-Crick, formando de esta manera una unión de tres brazos. El Panel B muestra la auto dimerización predicha de la secuencia de FP2 según el programa informático oligoanalyzer 3.0 disponible en <http://biotools.idtdna.com/analyzer/oligocalc.asp>. En los dos Paneles A y B, las bases flanqueantes están sombreadas en gris, los óvalos blancos y negros indican restos de fluoróforo y apantalladores, respectivamente.

La Figura 10 muestra la secuencia y la conformación nativa predicha de la sonda fluorescente FP3. La secuencia de FP3 comprende bases que están diseñadas para ser complementarias a la secuencia diana y bases flanqueantes adicionales. Las bases flanqueantes están subrayadas. La figura muestra la auto dimerización predicha de la secuencia de FP3 según el programa informático oligoanalyzer 3.0 disponible en <http://biotools.idtdna.com/analyzer/oligocalc.asp>. Las bases flanqueantes están sombreadas en gris, los óvalos blancos y negros indican restos de fluoróforo y apantalladores, respectivamente.

La Figura 11 muestra la secuencia y la conformación nativa predicha de la sonda fluorescente FP4. La secuencia de FP4 comprende bases que están diseñadas para ser complementarias a la secuencia diana y bases flanqueantes adicionales. Las bases flanqueantes están subrayadas. La figura muestra la estructura predicha de la secuencia usando programas de plegamiento de ADN según Nucleic Acids Res. 31: 3429-3431 (2003). Las bases flanqueantes están sombreadas en gris, los óvalos blancos y negros indican restos de fluoróforo y apantalladores, respectivamente.

Las Figuras 12 a 13 muestran las secuencias de las sondas fluorescentes FP5 a FP6. Las secuencias comprenden bases que están diseñadas para ser complementarias a la secuencia diana y bases flanqueantes adicionales. Las bases flanqueantes están subrayadas.

5 La Figura 14 muestra la secuencia y la conformación nativa predicha de la sonda fluorescente FP7. La secuencia de FP7 comprende bases que están diseñadas para ser complementarias a la secuencia diana y bases flanqueantes adicionales. Las bases flanqueantes están subrayadas. Se muestra la auto dimerización predicha de la secuencia de FP7 según el programa informático oligoanalyzer 3.0 disponible en <http://biotools.idtdna.com/analyzer/oligocalc.asp>. Las bases flanqueantes están sombreadas en gris, los óvalos blancos y negros indican restos de fluoróforo y apantalladores, respectivamente.

10 La Figura 15 muestra la secuencia y la conformación nativa predicha de la sonda fluorescente FP8. La secuencia de FP8 comprende bases que están diseñadas para ser complementarias a la secuencia diana y bases flanqueantes adicionales. Las bases flanqueantes están subrayadas. El Panel A muestra la estructura predicha de la secuencia usando programas de plegamiento de ADN según Nucleic Acids Res. 31: 3429-3431 (2003). El Panel B muestra la auto dimerización predicha de la secuencia de FP1 según el programa informático oligoanalyzer 3.0 disponible en <http://biotools.idtdna.com/analyzer/oligocalc.asp>. En los dos Paneles A y B, las bases flanqueantes están sombreadas en gris, los óvalos blancos y negros indican restos de fluoróforo y apantalladores, respectivamente.

Las Figuras 16 a 20 muestran las secuencias de las sondas fluorescentes FP9 a FP13. Las secuencias comprenden bases que están diseñadas para ser complementarias a la secuencia diana y bases flanqueantes adicionales. Las bases flanqueantes están subrayadas.

20 La Figura 21 muestra la secuencia y la conformación nativa predicha de la sonda fluorescente FP14. La secuencia de FP14 comprende bases que están diseñadas para ser complementarias a la secuencia diana y bases flanqueantes adicionales. Las bases flanqueantes están subrayadas. La figura muestra la estructura predicha de la secuencia usando programas de plegamiento de ADN según Nucleic Acids Res. 31: 3429-3431 (2003). Las bases flanqueantes están sombreadas en gris, los óvalos blancos y negros indican restos de fluoróforo y apantalladores, respectivamente.

25 Las Figuras 22 a 24 muestran la secuencia y la conformación nativa predicha de las sondas fluorescentes FP15 a FP17, respectivamente. Las secuencias comprenden bases que están diseñadas para ser complementarias a la secuencia diana y bases flanqueantes adicionales. Las bases flanqueantes están subrayadas. El Panel A muestra la estructura predicha de la secuencia usando programas de plegamiento de ADN según Nucleic Acids Res. 31: 3429-3431 (2003). El Panel B muestra la auto dimerización predicha de la secuencia de FP15 según el programa informático oligoanalyzer 3.0 disponible en <http://biotools.idtdna.com/analyzer/oligocalc.asp>. En los dos Paneles A y B, las bases flanqueantes están sombreadas en gris, los óvalos blancos y negros indican restos de fluoróforo y apantalladores, respectivamente.

30 La Figura 25 muestra el número de células observado respecto a la intensidad de fluorescencia durante el proceso FACS. La Figura 25A muestra el perfil para células NIH3T3 control expuestas a la sonda de señalización FP1. El fluoróforo Alexafluor633 se usó en FP1. Las células NIH3T3 control no estaban transfectadas con plásmido y no contienen secuencia diana. La Figura 25B muestra el perfil para células NIH3T3 transfectadas expuestas a la sonda de señalización FP1. Las células NIH3T3 transfectadas contenían ADN que codifica el ARN de interés y una secuencia etiqueta1 mostrada en la Figura 42A. La Figura 25C es una superposición de las Figuras 25A y 25B. La Figura 25C muestra que FACS distingue las células transfectadas con el plásmido que codifica las secuencias diana de las células control no transfectadas.

40 La Figura 26 muestra el número de células observado respecto a la intensidad de fluorescencia durante el proceso FACS. La Figura 26A muestra el perfil para células NIH3T3 control expuestas a la sonda de señalización FP2. El fluoróforo Alexafluor680 se usó en FP2. Las células NIH3T3 control no estaban transfectadas con plásmido y no contienen secuencia diana. La Figura 26B muestra el perfil para células NIH3T3 transfectadas expuestas a la sonda de señalización FP2. Las células NIH3T3 transfectadas contenían ADN que codifica el ARN de interés y una secuencia etiqueta2 mostrada en la Figura 42B. La Figura 26C es una superposición de las Figuras 26A y 26B. La Figura 26C muestra que FACS distingue las células transfectadas con el plásmido que codifica las secuencias diana de las células control no transfectadas.

45 La Figura 27 muestra el número de células observado respecto a la intensidad de fluorescencia durante el proceso FACS. La Figura 27A muestra el perfil para células NIH3T3 control expuestas a la sonda de señalización FP3. El fluoróforo fluoresceína se usó en FP3. Las células NIH3T3 control no estaban transfectadas con plásmido y no contienen secuencia diana. La Figura 27B muestra el perfil para células NIH3T3 transfectadas expuestas a la sonda de señalización FP3. Las células NIH3T3 transfectadas contenían ADN que codifica el ARN de interés y una secuencia etiqueta3 mostrada en la Figura 42C. La Figura 27C es una superposición de las Figuras 27A y 27B. La Figura 27C muestra que FACS distingue las células transfectadas con el plásmido que codifica las secuencias diana de las células control no transfectadas.

La Figura 28 muestra el número de células observado respecto a la intensidad de fluorescencia durante el proceso FACS. La Figura 28A muestra el perfil para células HeLa control expuestas a la sonda de señalización FP1. El fluoróforo fluoresceína se usó en FP1. Las células HeLa control no estaban transfectadas con plásmido y no contienen secuencia diana. La Figura 28B muestra el perfil para células HeLa transfectadas expuestas a la sonda de señalización FP1. Las células HeLa transfectadas contenían ADN que codifica el ARN vav inverso. La Figura 28C es una superposición de las Figuras 28A y 28B. La Figura 28C muestra que FACS distingue las células transfectadas con el plásmido que codifica las secuencias diana de las células control no transfectadas.

La Figura 29 muestra el número de células observado respecto a la intensidad de fluorescencia durante el proceso FACS. La Figura 29A muestra el perfil para células HeLa control expuestas a la sonda de señalización FP8. El fluoróforo fluoresceína se usó en FP8. Las células HeLa control no estaban transfectadas con plásmido y no contienen secuencia diana. La Figura 29B muestra el perfil para células HeLa transfectadas expuestas a la sonda de señalización FP8. Las células HeLa transfectadas contenían ADN que codifica el ARN vav inverso. La Figura 29C es una superposición de las Figuras 29A y 29B. La Figura 29C muestra que FACS distingue las células transfectadas con el plásmido que codifica las secuencias diana de las células control no transfectadas.

La Figura 30 muestra el número de células observado respecto a la intensidad de fluorescencia durante el proceso FACS. La Figura 30A muestra el perfil para células HeLa control expuestas a la sonda de señalización FP5. El fluoróforo fluoresceína se usó en FP5. Las células HeLa control no estaban transfectadas con plásmido y no contienen secuencia diana. La Figura 30B muestra el perfil para células HeLa transfectadas expuestas a la sonda de señalización FP5. Las células HeLa transfectadas contenían ADN que codifica el ARN vav inverso. La Figura 30C es una superposición de las Figuras 30A y 30B. La Figura 30C muestra que FACS distingue las células transfectadas con el plásmido que codifica las secuencias diana de las células control no transfectadas.

La Figura 31 muestra el número de células observado respecto a la intensidad de fluorescencia durante el proceso FACS. La Figura 31A muestra el perfil para células HeLa control expuestas a la sonda de señalización FP9. El fluoróforo fluoresceína se usó en FP9. Las células HeLa control no estaban transfectadas con plásmido y no contienen secuencia diana. La Figura 31B muestra el perfil para células HeLa transfectadas expuestas a la sonda de señalización FP9. Las células HeLa transfectadas contenían ADN que codifica el ARN vav inverso. La Figura 31C es una superposición de las Figuras 31A y 31B. La Figura 31C muestra que FACS distingue las células transfectadas con el plásmido que codifica las secuencias diana de las células control no transfectadas.

La Figura 32 muestra las imágenes de fluorescencia de células HeLa seleccionadas con fármaco transfectadas con un plásmido de expresión que codifica una parte de la secuencia de vav clonada en orientación inversa (referida como r-vav) así como un gen de resistencia a fármaco. Las células se expusieron a sondas fluorescentes (FP) diseñadas para reconocer la misma secuencia diana en r-vav (5' GTTCTTAAGGCACAGGAACTGGGA 3'). Las imágenes se obtuvieron usando un microscopio de fluorescencia y filtros diseñados para detectar la fluorescencia de Fam. Todas las FP usadas aquí se marcaron usando FAM excepto FP1, que se marcó usando fluoresceína. Cada uno de los Paneles A, B, C, D se expuso a FP10, FP11 FP12 y FP13.

La Figura 33 muestra las imágenes de fluorescencia de células transfectadas con construcciones que codifican ARN y secuencias etiqueta diseñadas para ser reconocidas por FP15 (Paneles A, B, D) o transfectadas con construcciones control que codifican el mismo ARN pero no la secuencia etiqueta (Panel C o E). Las secuencias etiqueta usadas se denominaron 6-5, 6-7 ó 6B10, que contenían 1, 2 y 3 copias de la secuencia diana, respectivamente, para la que FP15 se diseñó para reconocer. Las células transfectadas con construcciones que comprenden las secuencia etiqueta diseñadas para ser reconocidas por FP15 (Paneles A, B y D) presentaron una fluorescencia mayor que las células control (Panel C o E).

La Figura 34 muestra las imágenes de fluorescencia de células transfectadas con construcciones que codifican ARN y la secuencia etiqueta 6CA4 diseñadas para ser reconocidas por FP16 (Panel A) o transfectadas con construcciones control que codifican el mismo ARN pero no la secuencia etiqueta (Panel B). Las células transfectadas con construcciones que comprenden las secuencia etiqueta (Panel A) presentaron una fluorescencia mayor que las células control (Panel B).

La Figura 35A muestra una parte (subrayada) del complemento inverso de la secuencia de ADN vav (ADN r-vav) seleccionada para formar la secuencia etiqueta. Esta secuencia se clonó en un plásmido de expresión diseñado para expresar el ARNm de r-vav. La secuencia indicada en negrita es la secuencia diana para determinadas sondas fluorescentes. La Figura 35B muestra las secuencias subrayadas en la Figura 35A después de que se hayan combinado para formar una secuencia etiqueta (secuencia etiqueta1).

La Figura 36A muestra la estructura predicha de parte del ARN de r-vav usando programas de plegamiento de ARN en Nucleic Acids Res. 31: 3429-3431 (2003). La Figura 36B es la estructura predicha de la secuencia etiqueta1 mostrada en la Figura 35B. El sombreado indica la secuencia diana diseñada para ser reconocida por algunas de las sondas fluorescentes.

Las Figuras 37A, B y C muestran las estructuras predichas para la secuencia etiqueta 1, 2 y 3 como se describe en la Figura 42. Estas estructuras se parecen entre sí pero presentan una secuencia diferente para el reconocimiento por las sondas fluorescentes. La predicción se generó usando programas de plegamiento de ARN en Nucleic Acids Res. 31: 3429-3431 (2003).

5 La Figura 38 muestra la señal de fluorescencia emitida por las FP en disolución en presencia de secuencia de oligo diana o control. Las muestras se iluminaron con UV y se fotografiaron. Todas las FP usadas aquí incorporaron Fluoresceína. Cada uno de los tubos contenía 16ul en total que consistían en 5ul de una preparación madre de FP 20uM, 1,5ul 25mM MgCl₂, 8u 120uM oligo y 1,5ul de agua, teniendo una concentración final de magnesio de aproximadamente 2,34 mM. Aquí se usaron FP1 y FP18 y se sintetizaron incorporando uniones sulfuro entre las bases
10 de la secuencia diseñada para reconocer los oligos diana TO-FP1 y TO-FP18, respectivamente. FP1 está dirigida frente a la secuencia del oligo diana 1 (TO-FP1 5' GTTCTTAAGGCACAGGAACTGGGA 3') y FP18 está dirigida frente a la secuencia del oligo diana FP18 (TO-FP18 5' TCCCAGTTCCTGTGCCTTAAGAAC 3'). Las secuencia de TO-FP1 y TO-FP18 eran complementos inversos entre sí. TO-FP18 tiene una secuencia que no es diana de FP1 y sirvió como un oligo control para FP1. TO-FP1 tiene una secuencia que no es diana de FP18 y sirvió como un oligo control para FP18.

15 En todos los Paneles, las composiciones de los tubos son como se indica a continuación:

tubo	FP	oligo
1	FP18	TO-FP18
2	FP18	TO-FP1
3	FP1	TO-FP18
20 4	FP1	TO-FP1

Esta figura muestra que cada una de las FP ensayadas indicaba específicamente la presencia de secuencias diana mediante la emisión de una señal mayor en los tubos que contienen oligos que tienen la secuencia diana comparado con los tubos control que contienen oligos que no tienen secuencia diana. Los tubos que contienen FP en presencia de oligos que comprenden la secuencia diana se indican con un asterisco.

25 La Figura 39 muestra la señal de fluorescencia emitida por las FP en disolución en presencia de secuencia de oligo diana o control. Las muestras se iluminaron con UV y se fotografiaron. Todas las FP usadas aquí incorporaron FAM. Cada uno de los Paneles A, B, C muestra cuatro tubos a los que se añadió la misma FP. Cada tubo contiene un total de 10ul que contiene 2ul de una preparación madre de FP 20uM y 1ul de una preparación madre de oligo 100um en PBS suplementado con 4mM MgCl₂. En cada Panel, el tubo 1 no contenía preparación madre de oligo y en su lugar contenía
30 1ul de agua, el tubo 2 contenía oligo TO-M1, el tubo 3 contenía oligo TO-M2 y el tubo 4 contenía oligo TO-M3.

Las FP ensayadas en cada Panel se listan a continuación, junto con el oligo que incluye la secuencia que está diseñada para ser reconocida por la FP:

Panel	FP	oligo que comprende la secuencia diana
A	FP4	TO-M3 (tubo 4)
35 B	FP6	TO-M2 (tubo 3)
C	FP7	TO-M3 (tubo 4)

Esta figura muestra que cada una de las FP ensayadas indicaba específicamente la presencia de secuencias diana mediante la emisión de una señal mayor en los tubos que contienen oligos que tienen la secuencia diana comparado con los tubos control que no contienen oligo o contienen oligo que no tienen secuencia diana. Los tubos que contienen FP en presencia de oligos que comprenden la secuencia diana se indican con un asterisco.
40

Las secuencias en la dirección 5' a 3' para TO-M1, TO-M2 y TO-M3 se listan a continuación:

TO-M1:

TTTCTCTGTGATCCGGTACAGTCCTTCTGCGCAGGTGGACAGGAA
 GGTTCCTAATGTTCTTAAGGCACAGGAAGTGGGACATCTGGGCCCCG
 GAAAGCCTTTTTCTCTGTGATCCGGTACAGTCCTTCTGCGCAGGT
 GGACAGGAAGGTTCTAATGTTCTT

TO-M2:

TTTAACTGATGGATGGAACAGTCCTTCTGCGCAGGTGGACAGCTT
 GGTTCCTAATGAAGTTAACCCTGTCGTTCTGCGACATCTGGGCCCCG
 GAAAGCGTTTAACTGATGGATGGAACAGTCCTTCTGCGCAGGTGG

ACAGCTTGGTTCTAATGAAGTT

TO-M3:

GTAAAGTCAGACATCCGGTACAGTCCTTCTGCGCAGGTGGACAGG
 AAGGTTCTAATGTTCTATAGGGTCTGCTTGTGCTCATCTGGGCC
 CGGAGATGCGTAAAGTCAGACATCCGGTACAGTCCTTCTGCGCAG
 GTGGACAGGAAGGTTCTAATGTTCTAT

5 La Figura 40 muestra la señal de fluorescencia emitida por las FP en disolución en presencia de secuencia de oligo diana o control. Las muestras se iluminaron con UV y se fotografiaron. Todas las FP usadas aquí incorporaron FAM. Las FP 1, 2 y 3 se ensayaron en los Paneles A, B y C, respectivamente, y cada una se diseñó para reconocer secuencias diana relacionadas incorporadas en las etiquetas 1, 2 y 3, respectivamente, como se describe en la Figura 42. Los tubos que contienen las FP en presencia de oligos que comprenden la secuencia diana se indican con un asterisco.

El protocolo para el panel A, B, C fue según el protocolo para la FIG 39 y las muestras se describen a continuación:

Panel	FP	oligo que comprende la secuencia diana
A	FP1	TO-M1 (tubo 2)
10 B	FP2	TO-M2 (tubo 3)
C	FP3	TO-M3 (tubo 4)

15 La Figura 41 muestra la secuencia y la conformación nativa predicha de la sonda fluorescente FP18. La secuencia comprende bases que están diseñadas para ser complementarias a la secuencia diana y bases flanqueantes adicionales. Las bases flanqueantes están subrayadas. El Panel A muestra la estructura predicha de la secuencia usando programas de plegamiento de ADN según Nucleic Acids Res. 31: 3429-3431 (2003). El Panel B muestra la auto dimerización predicha de la secuencia de FP2 según el programa informático oligoanalyzer 3.0 disponible en <http://biotools.idtdna.com/analyzer/oligoanalyzer.asp>. En los dos Paneles A y B, las bases flanqueantes están sombreadas en gris, los óvalos blancos y negros indican restos de fluoróforo y apantalladores, respectivamente.

20 Las Figuras 42A, B y C muestran las tres secuencias etiqueta reconocidas por sondas fluorescentes. En las Figuras 42A, B y C, las secuencias diana se indican en negrita y también se muestran en la Figura 42D. La primera secuencia

(etiqueta1, 42A) es la misma que la secuencia indicada en la Figura 35B. Las dos secuencias siguientes (etiqueta2, 42B y etiqueta3, 42C) son versiones alteradas de etiqueta1. Las diferencias en la secuencia de etiqueta2 y etiqueta3 comparada con etiqueta1 están subrayadas en el Panel D. Se hicieron cambios adicionales en la secuencia en el resto de las secuencias etiqueta para compensar los cambios hechos en las partes mostradas.

5 La Figura 43 muestra el diseño de la secuencia etiqueta2 a partir de la secuencia etiqueta1. A-F indica los cambios secuenciales de bases hechos durante el diseño.

La Figura 44 muestra el diseño de la secuencia etiqueta3 a partir de la secuencia etiqueta1. A-F indica los cambios secuenciales de bases hechos durante el diseño.

10 La Figura 45 muestra que las células aisladas según los métodos de esta invención son viables. El Panel A muestra una célula aislada usando FACS después de que la célula se transfectó con tres construcciones de ADN cada una codificando un ARN de interés etiquetado con etiqueta1, 2 y 3, respectivamente. Las células se seleccionaron con fármaco y se expusieron a FP1, 2 y 3 y se aislaron. La intensidad de fluorescencia de las células para cada una de las tres sondas estuvo por encima de las intensidades de fondo comparadas con las células control no transfectadas con ninguna de las construcciones de ADN. Las células se sembraron en placas individualmente en un pocillo de una placa de 96 pocillos directamente por FACS y de una se obtuvo imagen justo después de su aislamiento. El Panel B muestra la misma célula una hora más tarde, después de que se adhirió a la superficie del pocillo. El Panel C muestra la misma célula al día siguiente, después de que haber experimentado división celular.

15 Cada uno de los tres paneles muestra que las células aisladas según los métodos permanecen viables a pesar de los efectos previamente desconocidos de los reactivos usados para exponer las células a las sondas. Estos efectos incluyen el comprometer la membrana plasmática de la célula y posiblemente el someter además a las células a mayores presiones durante FACS. El Panel A muestra que no se encontró que la membrana celular estuviera comprometida y los Paneles B y C demuestran además que la célula es viable ya que puede adherirse a la superficie de la placa de cultivo y dividirse, siendo estas dos características propiedades de las células viables.

20 La Figura 46 muestra los resultados de análisis FACS de células 293T transfectadas con mcon1 comparado con células control.

La Figura 47 muestra los resultados de análisis FACS de células 293T transfectadas con mcon2 comparado con células control.

La Figura 48 muestra los resultados de análisis FACS de células 293T transfectadas con mcon3 comparado con células control.

30 La Figura 49 muestra los resultados de análisis FACS de células 293T transfectadas con mcon4 comparado con células control.

La Figura 50 muestra los resultados de análisis FACS de células 293T transfectadas con mcon5 comparado con células control.

35 La Figura 51 muestra los resultados de análisis FACS de células 293T transfectadas con mcon6 comparado con células control.

La Figura 52 muestra los resultados de análisis FACS de células 293T transfectadas con mcon7 comparado con células control.

La Figura 53 muestra los resultados de análisis FACS de células 293T transfectadas con mcon8 comparado con células control.

40 La Figura 54 muestra los resultados de análisis FACS de células 293T transfectadas con mcon9 comparado con células control.

La Figura 55 muestra los resultados de análisis FACS de células 293T transfectadas con mcon10 comparado con células control.

45 La Figura 56 muestra los resultados de análisis FACS de células 293T transfectadas con mcon11 comparado con células control.

La Figura 57 muestra los resultados de análisis FACS de células 293T transfectadas con mcon12 comparado con células control.

La Figura 58 muestra los resultados de análisis FACS de células 293T transfectadas con mcon13 comparado con células control.

La Figura 59 muestra los resultados de análisis FACS de células 293T transfectadas con mcon14 comparado con células control.

5 La Figura 60 muestra los resultados de análisis FACS de células 293T transfectadas con mcon15 comparado con células control.

Descripción Detallada de la Invención

10 A no ser que se defina otra cosa, todos los términos técnicos y científicos usados en la presente memoria tienen el mismo significado que el entendido comúnmente por un experto en la técnica a la que pertenece esta invención. En caso de conflicto, la presente especificación, incluyendo las definiciones, prevalecerá.

El término "adyacente" tal y como se usa en el contexto de las sondas se refiere a una condición de proximidad para permitir que una pareja de interacción interaccione funcionalmente entre sí. Por ejemplo, la condición de proximidad permite que un fluoróforo sea apantallado o parcialmente apantallado por un resto apantallador o que un inhibidor de proteasa inhiba o inhiba parcialmente una proteasa. La distancia requerida para que un fluoróforo y un apantallador actualmente conocido interaccionen es aproximadamente 20-100 Å.

El término "par de bases" se refiere a pares de bases de Watson-Crick.

El término "región protuberante" se refiere a una región monocatenaria de un nucleótido o nucleótido modificado que no está emparejado por bases. El nucleótido protuberante puede estar entre regiones mutuamente complementarias (por ejemplo, Figura 4A).

20 El término "estructura dumbbell" se refiere a una cadena de ácido nucleico o ácido nucleico modificado que tiene la conformación de dos estructuras tallo-bucle unidas a través del extremo de un brazo de cada una de las regiones tallo (por ejemplo, Figura 7). La unión puede ser una región no complementaria o una unión fosfodiéster con o sin modificación.

25 El término "pareja de interacción" se refiere a dos grupos químicos que interaccionan funcionalmente cuando están adyacentes entre sí y cuando no están adyacentes entre sí, producen una señal detectable comparado con la ausencia de señal o señal de fondo producida por los grupos químicos que interaccionan o producen una señal diferente que la señal producida por los grupos químicos que interaccionan. Una pareja que interacciona incluye pero no está limitada a un fluoróforo y un apantallador, un marcaje quimioluminiscente y un apantallador o aducto, un dímero marcador y donante y aceptor FRET o una combinación de éstos. Una sonda de señalización puede comprender más de una pareja de interacción. Por ejemplo, una sonda de señalización que desplaza la longitud de onda tiene un primer fluoróforo y un segundo fluoróforo en el que ambos interaccionan con el apantallador y los dos fluoróforos son parejas donante y aceptora de FRET.

35 El término "región bucle" se refiere a una región monocatenaria de más de un nucleótido o nucleótido modificado que no está emparejada por pares de bases (por ejemplo, Figura 4B y Figura 23A). El bucle también puede estar entre una región mutuamente complementaria (Figura 8A).

40 El término "sonda de señalización" se refiere a una sonda que comprende una secuencia complementaria a una secuencia de ácido nucleico diana y al menos una región mutuamente complementaria y que comprende además al menos una pareja de interacción. Cuando la sonda de señalización no está unida a su secuencia diana, los restos de la pareja de interacción están adyacentes entre sí de manera que no se produce señal o se produce poca o se produce una señal diferente. Cuando la sonda de señalización está unida a la secuencia diana, los restos de la pareja de interacción no están ya adyacentes entre sí y se produce una señal detectable o una señal diferente de la señal producida por la sonda en su estado no unido. En una realización, la sonda de señalización es una sonda fluorogénica o fluorescente que comprende un fluoróforo y un resto apantallador y se produce un cambio en la fluorescencia después de la hibridación con la secuencia diana. Los restos de la pareja de interacción pueden estar unidos al extremo de la sonda de señalización o pueden estar unidos en la secuencia de ácido nucleico. Los ejemplos de restos que pueden incorporarse internamente en la secuencia de la sonda de señalización incluyen los apantalladores: dabcil dT, BHQ2 dT y BHQ1 dT y los fluoróforos: fluoresceína dT, Alexa dT y Tamra dT.

50 El término "sonda proteasa" se refiere a una sonda que comprende una secuencia complementaria a una secuencia diana y al menos una región mutuamente complementaria y que comprende además al menos una enzima proteolítica y al menos un inhibidor de la enzima proteolítica u otra molécula capaz de inactivar reversiblemente la enzima. Cuando la sonda no hibrida con una secuencia diana, la proximidad de la enzima proteolítica y el inhibidor de la enzima proteolítica les permite interaccionar, inhibiendo la actividad proteolítica. Después de la hibridación de la sonda con la secuencia

diana, la enzima proteolítica y su inhibidor se separan, activando la enzima proteolítica. La enzima proteolítica y el inhibidor pueden estar unidos covalentemente o no covalentemente a la sonda.

5 El término "región de emparejamiento erróneo" se refiere a una región bicatenaria en una molécula de ácido nucleico o molécula de ácido nucleico modificada, en el que las bases o bases modificadas no forman emparejamiento de bases de Watson-Crick (por ejemplo, Figura 4B y C). La región de emparejamiento erróneo está entre dos regiones emparejadas por pares bases. La región bicatenaria puede no estar unida por hidrógeno o unida por hidrógeno para formar pares de bases Hoogsteen, etc, o ambos.

El término "región mutuamente complementaria" se refiere a una región en una molécula de ácido nucleico o molécula de ácido nucleico modificada que está emparejada por pares de bases Watson-Crick.

10 El término "región no complementaria" se refiere a una región en una molécula de ácido nucleico o molécula de ácido nucleico modificada que no está emparejada por pares de bases Watson-Crick. Por ejemplo, la región no complementaria puede diseñarse para tener nucleótidos protuberantes, un bucle monocatenario, nucleótidos salientes en los extremos 5' ó 3' o regiones con emparejamiento erróneo.

15 El término "región tallo" se refiere a una región en una molécula de ácido nucleico o molécula de ácido nucleico modificada que tiene al menos dos pares de bases Watson-Crick. Por ejemplo, la región tallo puede diseñarse para tener más de una región mutuamente complementaria unida por regiones no complementarias o formar una región mutuamente complementaria continua.

20 El término "estructura tallo-bucle" se refiere a una molécula de ácido nucleico o molécula de ácido nucleico modificada con una secuencia de bucle monocatenario flanqueada por una pareja de brazos 5' y 3' de oligonucleótido u oligonucleótido modificado (por ejemplo, Figura 4). Los brazos 5' y 3' forman la región tallo.

25 El término "estructura de unión de tres brazos" se refiere a una cadena de ácido nucleico o ácido nucleico modificado que tiene una conformación de una región tallo, una primera región tallo-bucle y una segunda región tallo-bucle unidas entre sí mediante brazos de las regiones tallo (por ejemplo, Figura 6). La primera región tallo-bucle está en 5' respecto a la segunda región tallo-bucle. Las tres regiones pueden estar conectadas mediante una región no complementaria, una unión fosfodiéster o una unión fosfodiéster modificada o una combinación de éstas.

Sonda de Señalización

Pareja de Interacción

30 La sonda de señalización puede tener más de una pareja de interacción o tener diferentes parejas de interacción. En una realización, la sonda de señalización es una sonda fluorogénica. En una realización, la sonda fluorogénica no emite o emite un nivel de fondo de fluorescencia en su estado no hibridado, pero emite fluorescencia después de o emite fluorescencia por encima del nivel de fondo después de unirse a su diana. Pueden usarse múltiples fluoróforos para incrementar la señal o proporcionar fluorescencia a diferentes rangos de color. Pueden usarse múltiples apantalladores para disminuir o eliminar la señal en ausencia de la secuencia diana. Los ejemplos de apantalladores incluyen pero no están limitados a DABCYL, EDAC, Cesio, bromuro de p-xileno-bis-piridinio, nanopartículas de Talio y Oro. Los ejemplos de fluoróforos incluyen pero no están limitados a sulforodamina 101, acridina, ácido 5-(2'-aminoetil)aminoftalino-1-sulfónico (EDANS), Rojo Texas, Eosina, y Biodipy y Alexa Fluor 350, Alexa Fluor 405, Alexa Fluor 430, Alexa Fluor 488, Alexa Fluor 500, Alexa Fluor 514, Alexa Fluor 532, Alexa Fluor 546, Alexa Fluor 555, Alexa Fluor 568, Alexa Fluor 594, Alexa Fluor 610, Alexa Fluor 633, Alexa Fluor 635, Alexa Fluor 647, Alexa Fluor 660, Alexa Fluor 680, Alexa Fluor 700, Alexa Fluor 750, Alofocianina, Aminocumarina, Bodipyl-FL, Cy2, Cy3, Cy3.5, Cy5, Cy5.5, carboxifluoresceína (FAM), Cascada Azul, APC-Cy5, APC-Cy5.5, APC-Cy7, Cumarina, ECD (Rojo613), Fluoresceína (FITC), Hexaclorofluoresceína (HEX), Hidroxycumarina, Lisamina Rodamina B, Lucifer amarillo, Metoxicumarina, Oregón Verde 488, Oregón Verde 514, Pacific Azul, conjugados PE-Cy7, PerC, PerCP-Cy5.5, R-Ficoeritrina (PE), Rodamina, Rodamina Verde, Rodamina Roja-X, Tetraclorofluoresceína (TET), TRITC, Tetrametilrodamina, Rojo Texas-X, TRITC, XRITC y Puntos cuánticos. Véase, por ejemplo, Tyagi et al. Nature Biotechnology 16: 49-53, (1998) y Dubertret et al., Nature Biotechnology, 19: 36-370 (2001).

50 En la presente memoria también se describen sondas de señalización que desplazan la longitud de onda. En una realización, un extremo de la sonda tiene al menos un fluoróforo recolector y un fluoróforo emisor, un extremo adyacente de la sonda tiene al menos un resto apantallador. Véase, por ejemplo, Tyagi et al., Nature Biotechnology, 18, 1191-1196 (2000). En una realización, el fluoróforo recolector y el fluoróforo emisor están en el mismo extremo, en el que el fluoróforo emisor está en el extremo distal y un resto apantallador está en un extremo opuesto respecto al fluoróforo recolector. El fluoróforo emisor puede estar separado del fluoróforo recolector por un brazo espaciador de unos pocos nucleótidos. El fluoróforo recolector absorbe fuertemente en el intervalo de longitud de onda de la fuente de luz monocromática. En ausencia de la secuencia diana, ambos fluoróforos están apantallados. En presencia de dianas, la

sonda emite fluorescencia en el intervalo de emisión del fluoróforo emisor. El desplazamiento en el espectro de emisión se debe a la transferencia de energía absorbida desde el fluoróforo recolector al fluoróforo emisor por transferencia de energía resonante de fluorescencia. Estos tipos de sondas de señalización pueden proporcionar una señal más fuerte que las sondas de señalización que contienen un fluoróforo que no puede absorber energía eficazmente de fuentes de luz monocromáticas, En una realización, el fluoróforo recolector es fluoresceína y el fluoróforo emisor es 6-carboxirodamina 6G, tetrametilrodamina o rojo Texas.

En otra realización, un extremo de la sonda tiene al menos un fluoróforo F1 y otro extremo adyacente tiene al menos otro fluoróforo F2. Los dos fluoróforos se eligen de manera que la transferencia de energía resonante de fluorescencia (FRET) ocurrirá cuando están muy cerca. Cuando la sonda no está unida a su secuencia diana, después de excitación en la banda de absorción de F1, la fluorescencia de F1 se apantalla por F2, y se observa la fluorescencia de F2. Cuando la sonda está unida a su secuencia diana, FRET se reduce o elimina y la fluorescencia de F1 se elevará mientras que la de F2 disminuirá o desaparecerá. Esta diferencia en las intensidades de fluorescencia puede monitorizarse y puede calcularse una proporción entre la fluorescencia de F1 y F2. Como algunas veces se observa una fluorescencia residual en los sistemas de fluoróforo-apantallador, este sistema puede ser más ventajoso en la detección cuantitativa de la secuencia diana. Véase, Zhang et al., *Angew. Chem. Int. Ed.*, 40, 2, p. 402-405 (2001).

Los ejemplos de parejas donante-aceptor FRET incluyen pero no están limitados al grupo de cumarinas y grupo de 6-carboxifluoresceína, respectivamente.

En una realización, la sonda de señalización comprende una pareja de marcaje luminiscente y aducto. La interacción del aducto con el marcaje luminiscente disminuye la señal producida por el marcaje. Véase Becker y Nelson, Patente U.S. 5.731.148.

En otra realización, la sonda de señalización comprende al menos un dímero de marcaje. Cuando la sonda está unida a la secuencia diana, la señal de los marcajes es diferente de la señal del marcaje en la conformación de dímero.

En otra realización más, la pareja de interacción puede ser una enzima y un inhibidor de esa enzima, por ejemplo, una nucleasa y un inhibidor de nucleasa, una quinasa y un inhibidor de la quinasa, una proteasa y un inhibidor de la proteasa, una fosfatasa y un inhibidor de la fosfatasa, una caspasa y un inhibidor de la caspasa, o una ribozima y un inhibidor de la ribozima o un antígeno y un anticuerpo que se une específicamente al antígeno de manera que la diana detectada de la sonda puede trasladarse a una localización celular específica del antígeno, por ejemplo, a la sinapsis de una neurona, etc.

Conformación de las Sondas de Señalización y Sondas Proteasa u otras Sondas

Estructura Bicatenaria

En la presente memoria se describen sondas de señalización o proteasa u otras sondas que comprenden al menos dos cadenas separadas de ácido nucleico que están diseñadas para hibridar entre sí o formar al menos una región mutuamente complementaria. Al menos un extremo de una cadena es adyacente a un extremo de la otra cadena (Figura 1). El ácido nucleico puede ser ADN, ARN o ADN o ARN modificado. Las dos cadenas pueden ser cadenas idénticas que forman un auto dímero (Figura 8B). Las cadenas también pueden no tener una secuencia idéntica.

Las dos cadenas separadas pueden diseñarse para ser totalmente complementarias o comprender regiones complementarias y regiones no complementarias. En una realización, las dos cadenas separadas se diseñan para ser totalmente complementarias entre sí. En una realización, las dos cadenas forman una región mutuamente complementaria de 4 a 9, 5 a 6, 2 a 10, 10 a 40 ó 40 a 400 pares de bases continuos en cada extremo (véanse, por ejemplo, las Figuras 8, 9, 15, 22, 24 ó 41). Las cadenas pueden contener 5-7, 8-10, 11-15, 16-22, más de 30, 3-10, 11-80, 81-200 o más de 200 nucleótidos o nucleótidos modificados. Las dos cadenas pueden tener el mismo o un número diferente de nucleótidos (Figura 2). Por ejemplo una cadena puede ser más larga que la otra (Figura 2C). En una realización, el extremo 5' de una cadena está separado de la otra cadena o el extremo 3' de esa cadena está separado de la otra cadena, o ambos, en el que la separación es de hasta 10, hasta 20 o hasta 30 nucleótidos o nucleótidos modificados.

La región que hibrida con la secuencia diana puede estar en las regiones complementarias, regiones no complementarias de una o ambas cadenas o una combinación de éstas. Más de una secuencia de ácido nucleico diana puede tomarse como diana por la misma sonda de señalización. La una o más dianas pueden estar en la misma o diferentes secuencias y pueden ser exactamente complementarias a la parte de la sonda diseñada para unirse a la diana o al menos suficientemente complementarias. En una realización, las dos cadenas forman una región mutuamente complementaria en cada extremo y la secuencia complemento diana reside en las regiones distintas de las regiones mutuamente complementarias en los extremos (Figura 8B).

En una realización, la sonda de señalización con al menos dos cadenas separadas es una sonda fluorogénica. En una realización, una cadena tiene al menos un resto apantallador en un extremo y un fluoróforo en un extremo adyacente de la otra cadena (Figura 1). En una realización, cada uno de los extremos 5' y 3' de una cadena tiene el mismo o un fluoróforo diferente y cada uno de los extremos 5' y 3' de la otra cadena tiene el mismo o un resto apantallador diferente (Figura 1B y 2A). En una realización, el extremo 5' de una cadena tiene un fluoróforo y el extremo 3' tiene un resto apantallador y el extremo 3' de la otra cadena tiene el mismo o un resto apantallador diferente y el extremo 5' tiene el mismo o un fluoróforo diferente (Figura 1C y 2B).

Para la sonda proteasa, en una realización, una cadena tiene al menos una enzima proteolítica en un extremo, y un inhibidor de la enzima proteolítica en un extremo adyacente de la otra cadena. En una realización, cada uno de los extremos 5' y 3' de una cadena tiene una enzima proteolítica y cada uno de los extremos 5' y 3' de la otra cadena tiene un inhibidor de la enzima proteolítica. En una realización, el extremo 5' de una cadena tiene una enzima proteolítica y el extremo 3' tiene un inhibidor de la enzima proteolítica y el extremo 3' de la otra cadena tiene un inhibidor de la enzima proteolítica y el extremo 5' tiene una enzima proteolítica.

Estructura Tallo-bucle

En otra realización, la sonda de señalización o proteasa u otra sonda es una cadena de ácido nucleico o ácido nucleico modificado que comprende al menos una región mutuamente complementaria y al menos una región no complementaria. En una realización, la sonda forma una estructura tallo-bucle. La región tallo puede ser mutuamente complementaria o comprender regiones mutuamente complementarias y regiones no complementarias (Figura 4). Por ejemplo, la región tallo puede tener nucleótidos protuberantes que no están emparejados por pares de bases (Figura 4). La región tallo también puede contener nucleótidos salientes en los extremos 5' ó 3' que no están emparejados por pares de bases (Figuras 3B y 3C).

Cuando la región tallo es totalmente complementaria, la región tallo puede incluir 3-4, 5-6, 7-8, 9-10, 2-6, 7-10 ó 11-30 pares de bases (véanse, por ejemplo las Figuras 3 y 5). La región bucle puede contener 10-16, 17-26, 27-36, 37-45, 3-10, 11-25 ó 25-60 nucleótidos. En una realización, la región tallo forma 4-10, 4, ó 5 pares de bases continuos (véase, por ejemplo, la Figura 23).

En una realización, la estructura tallo-bucle comprende al menos una pareja interactiva que comprende dos grupos químicos y un grupo químico está en cada extremo de la cadena. En una realización, la sonda de señalización tiene al menos un fluoróforo y un resto apantallador en cada extremo de la cadena (Figuras 3, 4 y 5). La sonda proteasa tiene al menos una enzima proteolítica y un inhibidor de la enzima proteolítica en cada extremo de la cadena.

En una realización, la región tallo comprende dos regiones mutuamente complementarias conectadas mediante una región no complementaria, la región mutuamente complementaria adyacente a la pareja interactiva forma 5 a 9 pares de bases y la región mutuamente complementaria adyacente a la región bucle forma 4 a 5 pares de bases (Figura 8, 15, 21, 22, 24 ó 41). En una realización, la región no complementaria es una región bucle monocatenaria (Figura 8), una región de emparejamiento erróneo (Figura 15) o ambas. En otra realización, la región tallo comprende tres regiones mutuamente complementarias conectadas mediante dos regiones no complementarias, la primera región mutuamente complementaria adyacente a la pareja interactiva forma 4 a 5 pares de bases, la segunda región mutuamente complementaria forma 2 a 3 pares de bases y la tercera región mutuamente complementaria adyacente a la región bucle forma 2 a 3 pares de bases.

En la estructura tallo-bucle, la región que es complementaria a la secuencia diana puede estar en una o más regiones tallo o regiones bucle, o ambas. La región en el tallo que hibrida con la diana puede estar en las regiones mutuamente complementarias, regiones no complementarias o ambas. En una realización, la secuencia complemento diana está en la región bucle monocatenaria. En una realización, las regiones distintas de la región tallo adyacentes a la pareja interactiva es la secuencia complemento diana (Figura 8). Más de una secuencia de ácido nucleico diana puede ser la diana de la misma sonda. La una o más dianas pueden estar en la misma o diferentes secuencias y pueden ser exactamente complementarias a la parte de la sonda diseñada para unirse a la diana o al menos suficientemente complementarias.

El incremento de la longitud del tallo puede incrementar la estabilidad de las sondas de señalización en su conformación cerrada y, así, puede incrementar la relación señal a ruido de la señal detectable. La exposición de estas sondas de señalización a las células puede realizarse a temperaturas ligeramente elevadas que todavía son seguras para la célula seguido de un retorno a temperaturas normales. A temperaturas mayores, las sondas de señalización se abrirían y se unirían a su diana si está presente. Una vez enfriadas, las sondas de señalización no unidas a la diana revertirían a sus estados cerrados, lo que se ve apoyado por la estabilidad incrementada del tallo. De manera similar, pueden usarse otras fuerzas para conseguir el mismo resultado, por ejemplo DMSO que se piensa que relaja el emparejamiento de bases.

Estructura de Unión de Tres brazos

5 En otra realización, la sonda de señalización o proteasa u otra sonda es una cadena de ácido nucleico que forma una estructura de unión de tres brazos (Figuras 6A y 6B). En esta estructura, una región tallo y dos regiones tallo-bucle están conectadas para formar una unión de tres vías. Las regiones tallo pueden contener 2-5, 7-9, 10-12 pares de bases. El bucle de las regiones tallo-bucle puede contener 3-7, 8-10, 11-13 nucleótidos o nucleótidos modificados.

En una realización, la estructura de unión de tres brazos comprende al menos una pareja interactiva que comprende dos grupos químicos y un grupo químico está en cada extremo de la cadena. En una realización, la sonda tiene al menos un fluoróforo y un resto apantallador en cada extremo de la cadena. La sonda proteasa tiene al menos una enzima proteolítica y un inhibidor de la enzima proteolítica en cada extremo de la cadena.

10 En una realización, la región tallo adyacente a la pareja interactiva forma 3 a 4 ó 3 a 6 pares de bases continuos, la región tallo de la primera estructura tallo-bucle forma 4 a 5 pares de bases continuos y la región tallo de la segunda estructura tallo-bucle forma 2 a 3 pares de bases continuos (Figura 9). En una realización, las tres regiones están conectadas por una unión fosfodiéster o unión fosfodiéster modificada mediante los brazos de las regiones tallo. En una
15 realización, las tres regiones están conectadas por 1 ó 2 nucleótidos o nucleótidos modificados mediante los brazos de las regiones tallo.

La región en el tallo que hibrida con la diana puede estar en las regiones mutuamente complementarias, regiones no complementarias o ambas. En una realización, la secuencia complemento diana está en la región bucle monocatenaria. En una realización, las regiones distintas de la región tallo adyacente a la pareja interactiva es la secuencia complemento diana (Figura 9). Más de una secuencia de ácido nucleico diana puede tomarse como diana por la misma sonda. La una o más dianas pueden estar en la misma o diferentes secuencias y pueden ser exactamente
20 complementarias a la parte de la sonda diseñada para unirse a la diana o al menos suficientemente complementarias.

Estructura "dumbbell"

25 En otra realización, la sonda de señalización o proteasa u otra sonda es una cadena de ácido nucleico que forma una estructura con forma de "dumbbell" (Figura 7 u 11). La estructura es dos regiones tallo-bucle conectadas mediante un brazo de las dos regiones tallo. Las regiones tallo pueden contener 3-5, 7-9, 10-12 pares de bases. El bucle de las regiones tallo-bucle puede contener 5-7, 8-10, 11-13 nucleótidos. En una realización, la estructura "dumbbell" tiene una región tallo de 3 pares de bases continuos y una región tallo de 4 pares de bases continuos. En una realización, las dos regiones tallo están conectadas por 1 ó 2 nucleótidos o nucleótidos modificados. En otra realización, las dos regiones tallos están conectadas por una unión fosfodiéster o unión fosfodiéster modificada. En una realización, la estructura tallo-bucle, estructura "dumbbell" o estructura de unión de tres brazos tiene más de 30 nucleótidos o nucleótidos modificados.
30

En una realización, la sonda de señalización tiene al menos un fluoróforo y un resto apantallador en cada extremo de la cadena. La sonda proteasa tiene al menos una enzima proteolítica y un inhibidor de la enzima proteolítica en cada extremo de la cadena.

35 La región en el tallo que hibrida con la diana puede estar en las regiones mutuamente complementarias, regiones no complementarias o una combinación de éstas. En una realización, la secuencia complemento diana está en la región bucle monocatenaria. En una realización, la secuencia complemento diana es la región distinta de las dos regiones tallo. Más de una secuencia de ácido nucleico diana puede tomarse como diana por la misma sonda. La una o más dianas pueden estar en la misma o diferentes secuencias y pueden ser exactamente complementarias a la parte de la sonda diseñada para unirse a la diana o al menos suficientemente complementarias.

40 En la técnica están disponibles programas de plegamiento de ADN o ARN para predecir la conformación de un ácido nucleico o ácido nucleico modificado dado. Dichos programas de plegamiento incluyen pero no están limitados a los programas descritos en Nucleic Acids Res. 31: 3429-3431 (2003) y el programa informático oligoanalyzer 3.0 disponible en <http://biotools.idtdna.com/analyzer/oligocalc.asp>. Dichos programas de plegamiento predicen habitualmente un número de estructuras energéticamente más favorables. En otras realizaciones, se describen en la presente memoria las
45 estructuras energéticamente más favorables de las sondas FP1-18 (Figuras 8-24 y 41) que se predicen por los programas de plegamiento. Si la energía de la conformación se mide por energía libre, el valor más bajo de energía libre (negativo) indica que la conformación es energéticamente más favorable.

Modificación Química de las Sondas de Señalización y Proteasa u otra Sonda

50 En la presente memoria también se describen sondas de señalización o proteasa u otras sondas que están modificadas químicamente. Puede modificarse una o más de las bases 2'OH del núcleo de tipo azúcar-fosfodiéster. La sustitución de la unión fosfodiéster incluye pero no está limitada a -OP(OH)(O)O-, -OP(O-M⁺)(O)O-, -OP(SH)(O)O-, -OP(SM⁺)(O)O-, -NHP(O)₂O-, -OC(O)₂O-, -OCH₂C(O)₂NH-, -OCH₂C(O)₂O-, -OP(CH₃)(O)O-, -OP(CH₂C₆H₅)(O)O-, -P(S)(O)O- y

-OC(O)₂NH-. M⁺ es un catión inorgánico u orgánico. El núcleo también puede ser un ácido nucleico peptídico, en el que el núcleo de desoxirribosa fosfato se reemplaza por un núcleo pseudo peptídico. El ácido nucleico peptídico se describe en Hyrup y Nielsen, *Bioorganic & Medicinal Chemistry* 4: 5-23, 1996, y Hydig-Hielsen y Godskesen, WO 95/32305.

5 La posición 2' del azúcar incluye pero no está limitada a H, OH, alcoxi C₁-C₄, OCH₂-CH=CH₂, OCH₂-CH=CH-CH₃, OCH₂-CH=CH-(CH₂)_nCH₃ (n=0, 1 ... 30), halógeno (F, Cl, Br, I), alquilo C₁-C₆ y OCH₃. Alcoxi C₁-C₄ y alquilo C₁-C₆ pueden ser o pueden incluir grupos que tienen cadena lineal, ramificada o cíclica.

10 Las bases del nucleótido pueden ser una cualquiera de adenina, guanina, citosina, timina uracilo, inosina o las modificaciones siguientes. Las bases modificadas incluyen pero no están limitadas a N4-metil desoxiguanosina, deaza o aza purinas y pirimidinas. Los nitrógenos del anillo tales como N1 de adenina, N7 de guanina, N3 de citosina pueden estar alquilados. Las bases de pirimidina pueden estar sustituidas en la posición 5 ó 6 y las bases de purina pueden estar sustituidas en la posición 2, 6 u 8. Véanse, por ejemplo, Cook, WO 93/13121; Sanger, *Principles of Nucleic Acid Structure*, Springer-Verlag, Nueva York (1984).

15 Los derivados de los nucleótidos convencionales son muy conocidos en la técnica e incluyen, por ejemplo moléculas que tienen un tipo de azúcar diferente. La posición O4' del azúcar puede estar sustituida con S o CH₂. Por ejemplo, una secuencia de reconocimiento de base de nucleótido puede tener restos ciclobutilo conectados por restos de unión, en los que los restos de ciclobutilo tienen bases heterocíclicas unidas a ellos. Véase, por ejemplo, Cook et al., *Publicación Internacional* WO 94/19023.

Otras modificaciones químicas de las sondas útiles para facilitar la administración de las sondas a las células incluyen, pero no están limitadas a, colesterol, péptidos de transducción (por ejemplo, TAT, penetratina, etc.).

20 Métodos

25 Los métodos descritos en la presente memoria se basan en la capacidad de las sondas de señalización de producir una señal detectable después de hibridar con secuencias de ARN diana en células vivas. La señal producida debe ser más detectable que la producida en células control (por ejemplo, fluorescencia de fondo). Así, no es necesario que las células control no produzcan nada de fluorescencia. En una realización, el método es para detectar o cuantificar ARN. Un método es para aislar las células o generar líneas celulares que expresan al menos un ARN. En cualquiera de los métodos descritos en la presente memoria que implican el aislamiento de células, las células pueden cultivarse y también pueden cultivarse para generar líneas celulares. Una construcción de ADN que codifica un ARN o un ARN y una secuencia etiqueta se introduce en las células. La construcción de ADN puede integrarse en diferentes localizaciones del genoma de la célula. La integración en uno o más loci específicos también puede conseguirse. Después, las células transfectadas se exponen a la sonda de señalización, que genera una señal detectable después de unirse a la secuencia ARN diana o etiqueta. Las células que producen la señal detectable se aíslan. Las células pueden aislarse y cultivarse por cualquier método de la técnica, por ejemplo, las células pueden aislarse y sembrarse en placas individualmente o en lote. Las líneas celulares pueden generarse creciendo las células aisladas.

35 Cualquiera de los métodos descritos en la presente memoria puede realizarse usando un marcador de selección. Aunque la selección con fármaco (o selección usando cualquier otro marcador de selección adecuado) no es una etapa requerida, puede usarse para enriquecer la población de células transfectadas para células transfectadas de manera estable, siempre que las construcciones transfectadas se diseñen para conferir resistencia al fármaco. Si la selección usando sondas de señalización se realiza demasiado pronto después de la transfección, algunas células positivas pueden estar transfectadas sólo temporalmente y no de manera estable. Sin embargo, esto puede minimizarse proporcionando un paso de células suficiente lo que permite la dilución o pérdida del plásmido transfectado en las células transfectadas no de manera estable. Algunos plásmidos integrados de manera estable pueden no generar ningún ARN correspondiente a los insertos de ADNc clonados. Otros pueden generar ARN que pueden no ser detectados o detectados de manera ineficaz por las sondas de señalización.

45 Los ARN pueden tener uno o más de los diferentes papeles siguientes: ARN mensajeros que codifican proteínas, proteínas de fusión, péptidos fusionados a proteínas, señales de exportación, señales de importación, señales de localización intracelular u otras señales, que pueden estar fusionadas a proteínas o péptidos; ARN antisentido, ARNsi, ARN estructurales, ARN celulares incluyendo pero no limitado a tales como ARN ribosomales, ARNt, ARNhn, ARNsn; ARN aleatorios, ARN correspondientes a ADNc o EST; ARN de diversas especies, ARN correspondientes a oligonucleótidos, ARN correspondientes a preparaciones de ADNc de célula completa, tejido u organismo; ARN que tienen alguna actividad de unión a otros ácidos nucleicos, proteínas, otros componentes celulares o moléculas de fármaco; ARN que pueden incorporarse en varios complejos macromoleculares; ARN que pueden afectar alguna función celular; o ARN que no tienen la función o actividad mencionada anteriormente pero que sin embargo pueden ser expresados por células; ARN correspondientes a ARN virales o extraños, ARN conector o secuencia que une uno o más ARN; o ARN que sirven como etiquetas o una combinación o recombinación de secuencias no modificadas mutagenizadas, aleatorizadas o con transposición de uno o más de cualquiera de los anteriores. Los ARN pueden estar

55

bajo el control de promotores constitutivos o condicionales incluyendo pero no limitado a promotores inducibles, reprimibles, específicos de tejido, de choque térmico, de desarrollo, específico de linaje celular o temporales o una combinación o recombinación de secuencias no modificadas o mutagenizadas, aleatorizadas, con transposición de uno o más de cualquiera de los anteriores.

5 En una realización, las sondas de señalización son sondas fluorogénicas. Puede usarse tecnología de separación celular por fluorescencia o relacionada con sondas fluorogénicas para identificar y/o separar células que presentan un determinado nivel o niveles de fluorescencia a una o más longitudes de onda. La capacidad de detectar de manera fiable y eficaz ARNm así como otros ARN en células vivas permite su uso para identificar y, si se desea, separar células tomando como base sus características deseadas, por ejemplo, usando un Separador de Células Activado por Fluorescencia (FACS). La tecnología FACS permite actualmente la separación de hasta 70.000 células por segundo. 10 5.000.000 de células pueden separarse en menos de 2 minutos.

1. Generación de Líneas Celulares que Expresan Proteínas

Algunas de las etapas más tediosas implicadas en la generación de líneas celulares se eliminan por la aplicación de sondas de señalización como se describe en la presente memoria. En una realización, después de la transfección con una construcción de ADN que codifica un gen deseado, se introducen en estas células sondas fluorogénicas diseñadas para reconocer el mensaje del gen de interés. Esta etapa puede realizarse después de la selección usando un marcador de selección, por ejemplo, selección por fármaco siempre que la construcción de ADN transfectada también codifique resistencia a fármaco. Aquellas células que transcriban el gen emitirán fluorescencia. El análisis FACS posterior resulta en el aislamiento de las células fluorescentes que pueden crecerse para dar lugar a líneas celulares que expresan el gen elegido. 15 20

En una realización, las sondas de señalización se diseñan para ser complementarias a una parte del ARN que codifica la proteína de interés o a partes de sus regiones no traducidas en 5' ó 3'. Si la sonda de señalización diseñada para reconocer un ARN mensajero de interés es capaz de detectar secuencias diana presentes endógenamente, la proporción de éstas en comparación con la proporción de la secuencia diana producida por las células transfectadas es tal que el separador es capaz de discriminar los dos tipos de células. El gen de interés puede etiquetarse con una secuencia etiqueta y la sonda de señalización puede diseñarse de manera que reconozca la secuencia etiqueta. La secuencia etiqueta puede estar en marco con la parte codificadora de la proteína del mensajero del gen o fuera de marco con ésta, dependiendo de si se desea etiquetar la proteína producida. 25 30

Además, el nivel de expresión del gen de interés en cualquier célula dada puede variar. Esto puede deberse a una variedad de factores que pueden influir en el nivel de la expresión del ARN incluyendo pero no limitado a, la cantidad o número de copias de ADN que se transfectan en una célula, el sitio de cualquier integración genómica resultante del ADN y la integridad del ADN y la expresión resultante de éste después de la integración genómica. Se puede aplicar FACS para evaluar los niveles de expresión y seleccionar diferencialmente células individuales que expresan el mismo gen. 35

2. Generación de Líneas Celulares que Regulan Genes a la Baja

Hay varios estudios que describen la generación de líneas celulares que no expresan ARN que codifica una proteína sino uno que es el antisentido de un gen o parte de un gen. Dichos métodos pretenden reducir la cantidad de un ARN o proteína específico en una célula dada. Las etapas descritas anteriormente para la generación de líneas celulares que expresan proteínas son aplicables igualmente aquí y son virtualmente idénticas excepto en que aquí la sonda de señalización se diseña para detectar un ARN que es un ARN antisentido. 40

No todos los intentos para preparar líneas celulares transfectadas de manera estable que expresan antisentido resultan en líneas celulares en las que la expresión de la proteína diana se ve afectada suficientemente. Esta dificultad ha hecho que merezca menos la pena continuar con la producción de dichas líneas celulares. Debido a la facilidad del procedimiento descrito aquí, se ensaya fácilmente la eficacia de numerosas secuencias genéticas diferentes para su capacidad para rendir líneas celulares que expresan antisentido activas, es decir, eficaces. Se puede analizar éstas para determinar cuáles presenta perfiles de expresión apropiados en las que puede analizarse la regulación a la baja de genes diana. 45

La interferencia de ARN es una estrategia alternativa que también pretende disminuir los niveles de transcripción de genes específicos. La interferencia de ARN puede inducirse temporalmente usando construcciones de ARNsi o ADN sintetizadas químicamente que codifican ARNsi cortos o puede inducirse de manera estable si se generan células estables que expresan apropiadamente los ARNsi cortos. Los métodos descritos aquí también pueden usarse para analizar o aislar células o líneas celulares tomando como base su expresión de dichos ARNsi. 50

La aplicación de los métodos descritos aquí para obtener células que expresan ARN que inducen interferencia de ARN es importante adicionalmente porque ayudará a superar algunas complejidades encontradas actualmente en dichas células. Por ejemplo, aunque la interferencia de ARN se realiza con el fin de reducir los niveles transcripcionales de ARN diana y puede causar efectos aguas abajo en los niveles transcripcionales de otros ARN, también se ha mostrado para reducir los niveles transcripcionales de ARN que no se pretende que sean dianas. La identidad de estas dianas no pretendidas varía dependiendo de la secuencia del ARN que se usa para inducir la interferencia de ARN. Ésta es una característica que complica el uso de interferencia de ARN en células que pueden resultar comprometidas por dichas consecuencias no pretendidas. Como los métodos descritos aquí permiten la generación eficaz de múltiples células estables que expresa cada una, por ejemplo, una secuencia de ARN diferente usada para inducir interferencia de ARN para el mismo gen, cada una de estas secuencias puede ensayarse en la célula para sus efectos en los niveles transcripcionales de otros genes. El análisis de esta información puede usarse para distinguir ARN que tienen niveles transcripcionales disminuidos, en los que la disminución no se debe a un nivel de expresión disminuido del ARN diana. La interferencia de ARN ha ganado popularidad rápidamente ya que se ha usado con éxito para superar las dificultades asociadas con el logro de regulación a la baja específica usando ARN antisentido. Nuestros métodos pueden ayudar en la determinación de las secuencias más eficaces para una interferencia de ARN más específica que tienen una actividad no específica reducida. Como los métodos también proporcionan un método para la selección del antisentido más eficaz y específico, también representan un método para identificar ARN antisentido eficaces.

3. Diferenciación Entre Células Basada en Antígenos Localizados en la Superficie Celular

Los inmunólogos y otros han usado mucho FACS para separar células. Generalmente, este método se basa en el marcaje de proteínas localizadas en la superficie celular con sondas marcadas diferencialmente, habitualmente sondas de anticuerpo marcadas con fluoróforo. Por ejemplo, puede realizarse el de células positivas para la expresión de proteínas localizadas en la superficie celular. Este método se usa lo más comúnmente en condiciones diseñadas para conservar la integridad de la célula y mantener su viabilidad.

Según la presente invención, para detectar la presencia de proteínas localizadas en la superficie celular, puede prepararse una sonda de señalización dirigida al ARNm que codifica la proteína de interés. La sonda de señalización se introduce en las células por transfección sin suprimir la viabilidad celular. Después, puede usarse el separador celular para aislar las células que sean positivas. Además, si se usa una combinación de sondas de señalización, cada una dirigida al ARNm de una de las proteínas de interés, cada una puede marcarse diferencialmente. Si las células tienen un número mayor de dianas del que puede detectarse en una única aplicación de FACS, pueden realizarse múltiples ciclos de separación para separar las células.

Los métodos descritos aquí también permiten el análisis o separación de células para otros ARN celulares, por ejemplo, ARNm que codifican proteínas que están localizadas internamente o secretadas por la célula o ARN que no codifican proteínas. Como resultado, pueden detectarse uno o más de los ARN celulares descritos previamente como dianas, incluyendo ARN que codifican proteínas que son inaccesibles para las sondas de anticuerpo usadas comúnmente o para las que todavía no se han desarrollado sondas. Éstas pueden incluir proteínas asociadas a membrana, que se extienden en la membrana, ancladas a la membrana, citoplásmicas o nucleoplásmicas.

4. Ensayo de Células para la Expresión de ARN Específicos y Cuantificación del Nivel de la Expresión de ARN en las Células

Si el ARN diana de una sonda fluorogénica que se introduce en una célula está presente, la célula emitirá fluorescencia. Esta información puede evaluarse cualitativamente por el uso de Microscopía de Fluorescencia (incluyendo, confocal, de escaneo láser u otros tipos de microscopía) o FACS y también es cuantificable por cualquiera de éstos. Por ejemplo, en lugar de realizar la reacción en cadena de la polimerasa con transcripción inversa (RT-PCR) *in situ* en secciones de tejido para determinar un patrón de expresión para un ARN particular, se usa una sonda de señalización para llevar a cabo el mismo experimento. Además, usando una combinación de sondas fluorogénicas que emiten fluorescencia diferencialmente, cada una dirigida a un ARN específico, se ensaya la presencia o cantidad de varios ARN de interés en una etapa. La detección de ARN (en muestras fijadas) puede realizarse a temperaturas determinadas empíricamente para imitar la generación de señal no específica de diana. Es una práctica común establecer condiciones óptimas de temperatura para dirigir la detección cuando se usan sondas de ácido nucleico. (Localización de Antígenos en Combinación con la Detección de ARN en Células y Tejidos).

Usando células o secciones de tejido fijadas, se usa la inmunocitoquímica para describir la localización de los antígenos de proteína reconocidos y usando sondas de señalización dirigidas a ARN específicos, se co-localizan en las mismas muestras los ARN de interés. Se ha mostrado que las sondas fluorogénicas dirigidas a ARN funcionan en células fijadas.

5. Generación de Líneas Celulares que Expresan Múltiples ARN o Proteínas

Usando los métodos descritos en la presente memoria, se generan rápidamente líneas celulares transfectadas de manera estable que expresan cualquier número de ARN o proteínas, incluso sin la necesidad de mantener estas células en presencia de una mezcla de varios fármacos selectivos (o usando otros agentes selectivos). Después de la transfección génica y opcionalmente selección con fármaco, se introduce una combinación de sondas de señalización, una para el mensajero de cada proteína, en las células. Mediante el diseño de la secuencia complementaria diana de cada sonda fluorogénica para que hibride con el ARNm sólo de uno de los genes o con las secuencias etiqueta con las que los mensajeros pueden estar asociados, cada sonda de señalización se diseña para reconocer el ARNm codificado sólo por uno de los genes. En una realización, las células se separan por FACS. Mediante la selección de una o más señales, se genera una variedad de líneas celulares en una única aplicación.

Se puede tener la necesidad de producir una línea celular que expresa un número de ARN de interés que es superior al número que puede identificarse en una única aplicación de FACS. Por ejemplo, sería muy informativo tener una línea celular en la que se sobreexpresen todas las secuencias de proteínas y ARN que se piensa que están implicadas en la formación de un complejo particular o implicadas en una ruta biológica. Por ejemplo, los ARN o proteínas en la misma ruta biológica o una relacionada, ARN o proteínas que actúan aguas arriba o abajo entre sí, ARN o proteínas que tienen una función moduladora, activadora o represora entre sí, ARN o proteínas que dependen entre sí para su función o actividad, ARN o proteínas que forman un complejo o se unen entre sí, o ARN o proteínas que comparten homología (por ejemplo, homología de secuencia, estructural o funcional). Si el número de ARN requerido es mayor que el que se puede analizar en una aplicación de FACS, para conseguir esto, se repiten las etapas descritas anteriormente usando células que ya expresan una combinación de algunos de los ARN de las células huésped en las que se transfectarán construcciones adicionales que codifican ARN adicionales. Pueden usarse múltiples ciclos de los métodos descritos para obtener células que expresan todos o un subconjunto de los ARN que se requieren.

Si los múltiples ARN que se van a expresar se clonan todos en construcciones que confieren a las células resistencia al mismo fármaco, en una realización, puede usarse FACS para aislar las células que expresan todos los ARN deseados. En el caso en el que las secuencias se integren de manera estable en el genoma, se desea que las células no pierdan la expresión de ninguna de las secuencias. Sin embargo, es posible que una o más de las secuencias se pierdan. Si éste es el caso, se incrementa la concentración del fármaco selectivo o agente de selección en el medio en el que se crecen estas células, haciendo que esta posibilidad sea menos probable. Alternativamente, se usan construcciones cada una de las cuales confiere resistencia a un fármaco diferente y se mantienen las células en una mezcla de los fármacos apropiados. También, puede elegirse un subconjunto de las construcciones que se van a transfectar de manera estable en las células de manera que codifiquen un gen de resistencia para un fármaco y otro subconjunto para codificar un gen de resistencia para otro fármaco.

Además, si algunas células de una línea celular pierden la expresión de un ARN de interés, entonces como un recurso, el primer experimento para aislar la línea celular como se ha descrito anteriormente se repite y se obtienen nuevas células. Alternativamente, la mezcla de células descrita se analiza por FACS, con el objetivo de reaislar las células que expresan todos los ARN deseados. Éste es un procedimiento muy útil ya que rinde de nuevo células que dan lugar a una línea celular con la misma composición genética de la línea celular seleccionada originalmente.

Las estrategias descritas anteriormente rinden un suministro ilimitado de células que expresan cualquier combinación de proteínas y secuencias de ARN, susceptibles para métodos de análisis virtualmente ilimitados. Además es posible que una proteína que se sobreexpresa pueda ser tóxica para la célula y, como se discutirá más adelante, esta posibilidad puede tratarse fácilmente.

La facilidad con la que es posible reaislar las células que expresan todos los ARN deseados de las células que no expresan más todos los ARN hace posible mantener líneas celulares en ausencia de fármaco o en presencia de concentraciones mínimas de fármaco. Los métodos descritos en la presente memoria también permiten la reaplicación de sondas de señalización a células o líneas celulares generadas previamente. Por ejemplo, para determinar si y en qué medida las células todavía son positivas para cualquiera de uno o más de los ARN para los que se aislaron originalmente.

6. Generación de Líneas Celulares que Sobreexpresan Notablemente Uno o Más ARN o Proteínas

Para cada gen que se quiere sobreexpresar mucho, por ejemplo, dos o más secuencias para el mismo gen se clonan en primer lugar en construcciones de ADN que opcionalmente también confieren resistencia a fármacos u otro marcador seleccionable. Cada una de las múltiples secuencias para cada gen se diseña para incluir la secuencia que codifica una secuencia etiqueta diferente. En una realización, después de la transfección de las construcciones de ADN en las células y selección posterior, se introducen en las células sondas fluorogénicas, cada una de las cuales está dirigida sólo a una secuencia etiqueta y marcadas con fluorescencia diferencialmente, y se usa el separador de células para aislar las células positivas para sus señales. Dichas células han integrado en sus genomas al menos una copia de cada una de las

secuencias etiquetadas diferencialmente y así la expresión de la secuencia de interés se produce a partir de un número incrementado de copias de esencialmente la misma secuencia de interés. La secuencia de interés puede integrarse en diferentes localizaciones del genoma de la célula. Este método se usa conjuntamente con el uso de FACS para seleccionar aquellas células que tengan mayor intensidad para la señal de cada fluoróforo. Una parte o todas las diferentes etiquetas pueden ser idénticas de manera que puede usarse una sonda de señalización común dirigida frente a esta secuencia común para detectar todas las diferentes etiquetas en las células.

7. Generación de Líneas Celulares que Expresan Múltiples ARN Antisentido

Se crean líneas celulares transfectadas de manera estable que producen múltiples mensajes antisentido como sigue. Dichos mensajes antisentido están dirigidos frente a ARNm u otros ARN. Se seleccionan células que expresan a diferentes niveles una cualquiera de las secuencia antisentido transfectadas. Mediante ciclos repetidos de transfecciones estables, se seleccionan fácilmente células que darán lugar a líneas celulares transfectadas de manera estable que expresan el mensaje antisentido de un número ilimitado de ARN.

Por supuesto, las células que expresan otros ARN distintos de ARN antisentido pueden prepararse por los métodos descritos en la presente memoria. Dichos ARN incluyen pero no están limitados a uno o más de ARNm, ARNsi, ARNsh, otros ARN estructurales tales como ARNhn, ARNt o ARNsn, ARN que tienen actividad de interferencia de ARN, ARN que sirven como etiquetas, etc.

8. Generación de Bibliotecas de Líneas Celulares

Una pluralidad de células se transfectan con construcciones de ADN para formar una biblioteca de expresión. Las bibliotecas de expresión de secuencias de ADN pueden incluir cualquiera de las clases que son conocidas en la técnica o una mezcla de éstas, incluyendo pero no limitado a bibliotecas de ADNc o EST generadas a partir, por ejemplo, de organismos completos, tejidos, células o líneas celulares y bibliotecas sintéticas incluyendo pero no limitado a oligonucleótidos o secuencias que codifican péptidos. De manera similar, pueden usarse bibliotecas de construcciones de ADN no diseñadas específicamente como bibliotecas de expresión. Por ejemplo, una biblioteca de construcción de ADN puede incluir secuencias de ADN que pueden tener funciones reguladoras tales como elementos promotores, represores o amplificadores y éstos pueden ser constitutivos, inducibles o reprimibles. Asimismo, las bibliotecas de construcción de ADN pueden comprender segmentos grandes de ADN tal como loci genéticos enteros o parciales o ADN genómico, a partir de los cuales se puede producir la transcripción. Cualquiera de estas bibliotecas de expresión de construcciones de ADN puede estar cada una totalmente o parcialmente mutagenizada, aleatorizada, recombinada, con transposición, alterada o tratada en cualquier combinación de éstas. Además, cualquiera de estos tipos de bibliotecas puede comprender además al menos una etiqueta que se expresa y que puede usarse como una diana para las sondas de señalización.

Las bibliotecas de expresión de secuencias de ADN para clases específicas de proteínas pueden prepararse y usarse para generar bibliotecas de líneas celulares. Por ejemplo, los ADNc para proteínas quinasas pueden clonarse en construcciones de expresión que comprenden una secuencia etiqueta. Las células que expresan de manera estable diferentes quinasas pueden obtenerse y usarse para generar una biblioteca de líneas celulares limitada a las líneas celulares que expresan quinasas. La clase de secuencias puede elegirse para cumplir con aplicaciones adicionales tales como cribado de fármacos.

Las bibliotecas de líneas celulares también pueden expresar clases de proteínas que tienen homología de secuencia, que pertenecen a la misma familia de proteínas o en una familia funcional y también proteínas definidas por su papel en una ruta o complejo o sistema de proteínas dado. Por ejemplo, en un cribado de fármacos para compuestos que pueden unirse a varias proteínas de VIH, puede prepararse una biblioteca de líneas celulares en la que cada línea celular expresa una proteína diferente de VIH y usarse para cribar compuestos fármacos.

Las bibliotecas de líneas celulares pueden usarse para ensayar péptidos o proteínas secretados que tienen una actividad deseada. En primer lugar, una biblioteca de líneas celulares se prepara usando una biblioteca de expresión (para péptidos, proteínas, EST, ADNc, etc) que codifica además una señal de exportación traducida en marco con los péptidos/proteínas. Los péptidos/proteínas expresados se secretan en el medio de crecimiento. El efecto de estos péptidos en una actividad particular puede determinarse en un ensayo diseñado apropiadamente. El sobrenadante del cultivo tisular de las líneas celulares puede recogerse y aplicarse a células de ensayo para ensayar la actividad.

Además, puede generarse una biblioteca de líneas celulares por ejemplo transfectando una biblioteca de expresión en células, y opcionalmente, seleccionando en primer lugar las células transfectadas con la biblioteca de expresión (por ejemplo, usando una sonda de señalización dirigida frente a una etiqueta que está incluida en la biblioteca de expresión) y exponiendo las células a una o más sondas de señalización dirigidas frente a uno o más ARN de interés. Esto permitiría expresar varios ARN en células y determinar cuál de estos ARN resulta en la regulación transcripcional al alza o a la baja aguas abajo de uno o más ARN de interés.

9. Generación de Líneas Celulares Que son Inactivaciones Funcionales para Una o Más Proteínas

Los métodos descritos en la presente memoria proporcionan los medios para preparar inactivaciones funcionales en células cultivadas. Se generan líneas celulares que son inactivaciones funcionales de una proteína cualquiera de interés mediante la generación de células que expresan a partir de múltiples loci virtualmente el mismo ARN antisentido o ARNsi que tiene actividad de interferencia de ARN para una única secuencia de ARN. Cualquiera de los métodos descritos en la presente memoria que usa ARNsi también puede realizarse usando ARNsh. Por ejemplo, se transfectan en células múltiples construcciones cada una de las cuales codificará el ARN antisentido para un gen particular o ARNsi que tiene actividad de interferencia de ARN para el gen, o ambos. Aquí cada secuencia de ARN antisentido sólo se diferencia en que cada una estará etiquetada con la secuencia de nucleótidos de una única secuencia etiqueta. Se seleccionan aquellas células que expresan uno o más o todos los ARN antisentido etiquetados diferencialmente. De manera similar, la presencia de cada ARNsi o ARNsh con actividad de interferencia de ARN se determinará por la detección de una etiqueta con la que cada ARNsi o ARNsh está asociado. Como FACS se usa para cuantificar la fluorescencia como se ha descrito anteriormente, esta característica permite seleccionar aquellas células que expresan más fuertemente una cualquiera o más de las secuencias antisentido. Se podrían aislar las células que presentan los niveles de expresión deseados del ARN diana o poca o ninguna expresión de éste debido a la expresión de cualquier número o combinación de ARN expresados, que actúan para disminuir el nivel de expresión del ARN diana.

De manera importante, una o más de las diferentes secuencias antisentido y ARNsi que tiene actividad de interferencia de ARN dirigidos al mismo gen pueden usarse en esta estrategia. Por ejemplo, algunos de los ARN antisentido, cuya expresión se selecciona usando sondas de señalización y FACS, se diseñan de manera que estén dirigidos a una región particular del ARN mensajero para el gen, mientras que otros se diseñan de manera que estén dirigidos a una parte alternativa del mismo mensajero. Con el fin de generar líneas celulares que son inactivaciones funcionales de una proteína de interés, se transfecta de manera estable en células tantas secuencias genéticas que codifican ARN antisentido similares o diferentes o ARNsi que tienen actividad de interferencia de ARN frente al mismo gen de interés como sea necesario para la producción de una línea celular que no presenta niveles detectables de expresión de la proteína de interés o, alternativamente, niveles aceptablemente bajos de expresión.

Además, se generan líneas celulares en las que múltiples proteínas están inactivadas funcionalmente o tienen niveles de expresión reducidos repitiendo el procedimiento descrito anteriormente mientras se toman como diana cualquier número de secuencias para inactivarse funcionalmente por antisentido o ARNsi. Por ejemplo, para estudiar la función de un complejo de proteínas, se inactivan o reducen los niveles de expresión de una, todas o cualquier combinación de las proteínas que forman el complejo.

10. Generación de Líneas Celulares Que son Inactivaciones Funcionales Sólo para Formas Seleccionadas de Corte y Empalme Alternativo de Uno o Más Genes

Frecuentemente se traducen en proteínas versiones de corte y empalme diferentes de un único gen con diferentes funciones. Usando los métodos descritos en la presente memoria, se generan líneas celulares en las que sólo formas de corte y empalme alternativo seleccionadas de una o más proteínas son inactivaciones funcionales o tienen niveles de expresión reducidos. Por ejemplo, diseñando antisentidos o ARNsi que sólo están dirigidos a aquellas versiones de corte y empalme alternativo del ARN mensajero del gen que se quiere eliminar de la célula, se inactivan funcionalmente todos los ARN de corte y empalme alternativo del gen de interés o se reducen suficientemente sus niveles de expresión hasta niveles deseados excepto para aquellos mensajes de corte y empalme alternativo que son de interés.

11. Generación de Líneas Celulares que Expresan Uno o Más ARN o Proteínas Mientras están Inactivadas Funcionalmente para Una o Más Proteínas Diferentes

Por ejemplo, para un grupo dado de proteínas que se piensa que interactúan entre sí, se pueden estudiar sus interacciones generando líneas celulares transfectadas de manera estable en las que una o más de las proteínas de interés están inactivadas funcionalmente o tienen niveles de expresión reducidos por la expresión celular de antisentido o ARNsi (o ARNsh). La función de las proteínas de interés restantes en la célula puede estudiarse pero quizá de manera más interesante, dicha célula podría alterarse adicionalmente manipulándola además de manera que sobreexpresen una o más de las proteínas de interés restantes. Además, se puede sobreexpresar o eliminar o reducir la expresión de proteínas adicionales en células. De nuevo, es posible que la sobreexpresión de determinadas proteínas o una inactivación funcional de determinadas proteínas pueda ser letal para las células. Esto es un problema que se planteará más adelante.

Pueden usarse métodos análogos para regular aleatoriamente al alza o a la baja genes por la introducción de ADN con papeles directos o indirectos en la regulación transcripcional. Por ejemplo, pueden transfectarse en células secuencias de ADN incluyendo pero no limitadas a una o una combinación de secuencias promotoras, amplificadoras o represoras, o una secuencia que tiene alguna otra actividad de unión o funcional que resulta en la modulación de los niveles transcripcionales para uno o más ARN. La actividad de estos elementos puede ser constitutiva, inducible o reprimible.

Las sondas de señalización para uno o más ARN específicos pueden usarse para identificar o aislar células en las que los niveles de expresión de estos ARN se han incrementado o disminuido. La integración estable de algunas secuencias de ADN usando los métodos descritos puede activar o inactivar aleatoriamente suficientemente la transcripción de loci genéticos. Estas bibliotecas de líneas celulares pueden cribarse para células que sobreexpresan o están inactivadas eficazmente para genes específicos. Las células con los niveles deseados de uno o más ARN pueden seleccionarse de esta manera. Pueden realizarse múltiples ciclos de este procedimiento, si es necesario, para aislar células que tienen los perfiles de expresión deseados para múltiples ARN de interés.

12. Generación de Ratones Transgénicos

Para algunos propósitos, el estudio de células en cultivo no es suficiente. La metodología descrita anteriormente, sin embargo, también se presta a la manipulación de células madre embrionarias. Pueden obtenerse células madre embrionarias que podrían expresar múltiples ARN o proteínas o actuar como inactivaciones funcionales de múltiples proteínas o un subconjunto de las formas de corte y empalme alternativo de múltiples proteínas, etc, según los procedimientos anteriores. Dichas células madre embrionarias se usan como la base para la generación de animales transgénicos.

Las células aisladas según los métodos descritos pueden implantarse en organismos directamente o sus núcleos pueden transferirse a otras células receptoras y éstas pueden implantarse. Un uso podría ser generar animales transgénicos y otros usos pueden incluir pero no estar limitados a introducir células que sintetizan o secretan productos celulares en el organismo y células que se modifican por ingeniería para realizar los papeles deseados en el organismo.

13. Generación de Líneas Celulares Transfectadas de Manera Estable Inducibles

La sobreexpresión o la ausencia de expresión de determinadas proteínas o ARN en células puede ser letal o perjudicial. Aún así puede tener una importancia crítica estudiar una célula que sobreexpresa una proteína o ARN tóxico, o una que es una inactivación funcional de una proteína o ARN, sin la que la célula es incapaz de sobrevivir. Para este objetivo, se generan células transfectadas de manera estable en las que ARN seleccionados que tienen dichos efectos perjudiciales en la célula están bajo el control de promotores inducibles o condicionales. Para aislar dichas líneas celulares, en una realización, las células transfectadas y opcionalmente seleccionadas por fármaco se inducen en primer lugar mínimamente para afectar la transcripción de los genes inducibles y las células se someten a análisis FACS después de la transfección en ellas de sondas de señalización diseñadas para reconocer los ARN apropiados. Las células obtenidas se mantienen de manera que los ARN tóxicos se inducen y transcriben sólo cuando es necesario.

Los sistemas inducibles pueden ser ventajosos para aplicaciones distintas de la expresión de ARN tóxicos. Por ejemplo, se induce la expresión de secuencias genéticas transfectadas de manera estable en las células en un determinado punto durante el ciclo celular de una línea celular sincronizada. Alternativamente, si los productos expresados de un conjunto de una o más secuencias genéticas transfectadas de manera estable se piensa que actúa en los productos expresados de otro conjunto, es interesante clonar las secuencias genéticas del primer conjunto bajo el control de un promotor inducible y aquellas del segundo conjunto bajo el control de un segundo promotor inducible. Mediante inducciones variadas, se estudia los productos expresados codificados por cualquier conjunto de secuencias genéticas en ausencia o presencia de los productos expresados del otro conjunto. En general, las secuencias de ADN que se incorporan en las células como se ha descrito en los métodos anteriores pueden ponerse cada una bajo el control de un promotor u otro regulador de la transcripción con la actividad deseada. Por ejemplo, pueden usarse promotores, amplificadores o represores inducibles, específicos de tejido, específicos de tiempo o temporales, así como elementos reguladores que se modulan, activan o reprimen debido a señales celulares o extracelulares, incluyendo pero no limitado a uno o más compuestos o químicos, otras células, proteínas, péptidos, hormonas, moléculas de señalización, factores secretados por las células, extractos completos o fraccionados de organismos, tejidos o células, o muestras medioambientales. En estos casos, las células se exponen en primer lugar a los niveles apropiados de los agentes que regulan la transcripción antes de su exposición a la sonda de señalización.

14. Detección de Eventos de Recombinación Genéricos en Células Vivas y el Aislamiento Posterior de Células No Recombinadas o Recombinadas Diferencialmente

De forma paralela al uso de sondas de señalización para detectar células que han experimentado los eventos de recombinación implicados en la creación de líneas celulares estables, está el uso de sondas de señalización para detectar y aislar a partir de una mezcla de células vivas aquellas células que han experimentado otros eventos de recombinación específicos. El mismo principio puede usarse para ensayar recombinación VDJ, translocación e integración del genoma viral, por ejemplo.

En los eventos de recombinación celular, por ejemplo, una secuencia de ADN genómico se intercambia por otra. Si una secuencia de ADN que codifica una región en la que ocurre un evento de recombinación se transcribe en ARN, la presencia de dicho evento se detecta por una sonda de señalización diseñada para reconocer el ARN transcrito a partir

de la secuencia de ADN no recombinada o el que se transcribe a partir de la secuencia recombinada. Dicho ensayo también puede realizarse usando Microscopio de Fluorescencia (u otro equipo que pueda cuantificar la señal resultante). Si se quiere separar las células que se han recombinado de aquellas que no, se somete las células a FACS y se separan. Además, FACS puede usarse para separar las células tomando como base la presencia o ausencia en ellas de numerosos eventos de recombinación.

15. Separación de Células Tomando como Base los ARN Expresados

El uso de sondas de señalización como se describe en la presente memoria permite separar las células tomando como base su expresión de ARN que codifican proteínas localizadas internamente, localizadas en la superficie celular o secretadas así como otros ARN que pueden estar presentes en la célula. Por ejemplo, empezando a partir de una población mixta de células, se aíslan aquellas células que expresan proteínas localizadas internamente de interés diseñando sondas de señalización que reconocen los ARNm que dan lugar a estas proteínas. Estas sondas de señalización se transfectan en la mezcla de células y puede usarse FACS para separarlas según sea apropiado. Pueden realizarse múltiples ciclos de separación.

Además, un investigador puede estar interesado, por ejemplo, en aislar células que expresan el ARNm de una o más proteínas específicas o ARN que se transcriben en respuesta a un factor añadido dado, o en determinar qué factor añadido induce o reprime la transcripción de una o más proteínas o ARN. Los factores añadidos que pueden ensayarse de esta manera pueden incluir pero no están limitados a uno o más de ácidos nucleicos, proteínas, péptidos, hormonas, moléculas de señalización, compuestos químicos, químicos inorgánicos u orgánicos, células, extractos completos o fraccionados o derivados de organismos, tejidos o células, productos purificados o aislados de células u organismos, muestras del medioambiente u otras fuentes.

Para aislar las células que se inducen para expresar uno o más ARN específicos en respuesta a una citoquina, por ejemplo, una mezcla de células se induce en primer lugar por la citoquina y se transfectan con sondas de señalización, cada una de las cuales se diseña para reconocer el ARNm que dará lugar a una de las proteínas de interés. En una realización, se usa FACS para aislar aquellas células que sean positivas para el ARNm de interés. En una realización alternativa, también se ensayan las células infectadas con un virus, por ejemplo, para su expresión de un gen particular. Alternativamente, un conjunto o biblioteca de compuestos tales como una biblioteca de compuestos químicos, o una biblioteca de expresión de ARN, puede aplicarse a las células para determinar cuál si hay alguno de los compuestos o mezcla de compuestos, o ARN, induce, reprime o modula la transcripción de una o más proteínas o ARN específicos.

Es posible con la metodología descrita anteriormente en la presente memoria detectar células positivas para la presencia de ARN con uno o más de los diferentes papeles siguientes: ARN mensajeros que codifican proteínas, proteínas de fusión, péptidos fusionados con proteínas, señales de exportación, señales de importación, señales de localización intracelular u otras señales, que pueden estar fusionados con proteínas o péptidos; ARN antisentido, ARNsi, ARN cortos que forman estructuras de horquilla que tienen una actividad similar a ARNsi; ARN estructurales, ARN celulares incluyendo pero no limitado a tales como ARN ribosomales, ARNt, ARNhn, ARNsn; ARN aleatorios, ARN correspondientes a ADNc o EST; ARN de diversas especies, ARN correspondientes a oligonucleótidos, ARN correspondientes a preparaciones de ADNc de células, tejidos u organismos completos; ARN que tienen alguna actividad de unión con otros ácidos nucleicos, proteínas, otros componentes celulares o moléculas de fármaco; ARN que pueden estar incorporados en varios complejos macromoleculares; ARN que pueden influir en alguna función celular; o ARN que no tienen la función o actividad mencionada anteriormente pero que pueden expresarse por las células; ARN correspondientes a ARN virales o extraños, ARN conector o secuencia que conecta uno o más ARN; o ARN que sirven de etiquetas o una combinación o recombinación de secuencias no modificadas mutagenizadas, aleatorizadas o translocadas de uno cualquiera o más de los anteriores. Los ARN pueden estar bajo el control de promotores constitutivos o condicionales incluyendo pero no limitado a promotores inducibles, reprimibles, específicos de tejido o temporales o una combinación o recombinación de secuencias no modificadas o mutagenizadas, aleatorizadas, translocadas de uno cualquiera o más de los anteriores. La expresión de los ARN descritos anteriormente puede resultar de la introducción en células de construcciones de ADN, vectores u otros métodos de administración que administran ácidos nucleicos que resultan en su expresión.

16. Detección *In Vivo* de Ácidos Nucleico y Selección Posterior de Células usando Sondas Proteasa

En la presente memoria también se describe una nueva forma de sonda proteasa que, a diferencia de la sonda de señalización, presenta actividad proteolítica después de unirse a sus ácidos nucleicos diana. Dicha actividad proteolítica puede usarse para propósitos de detección pero también para degradar secuencias de proteína particulares en una célula si el ácido nucleico diana está presente en la célula. Por ejemplo, una proteasa que escinde específicamente una proteína viral puede activarse cuando se activa la transcripción de una secuencia viral, tal como en una infección latente y cuando, por ejemplo, la sonda proteasa está dirigida frente al mensajero viral.

5 En la presente memoria también se describe una sonda de hibridación unitaria que genera actividad proteolítica, referida en la presente memoria como una sonda proteasa. Dichas sondas proteasa también comprenden nucleótidos o nucleótidos modificados complementarios a un ARN diana y nucleótidos o nucleótidos modificados mutuamente complementarios. Estas sondas proteasa operan de una manera similar a las sondas mencionadas anteriormente pero en lugar de una producción de o cambio en la señal fluorescente después de que la sonda de señalización interacciona con su secuencia de ácido nucleico diana, la sonda proteasa se vuelve proteolítica en presencia de la diana.

10 Se pueden sustituir las sondas proteasa en lugar de las sondas de señalización en los métodos anteriores, rindiendo nuevas posibilidades. Después de la transfección de las sondas proteasa en células que expresan el ARN que es reconocido por una sonda proteasa, la sonda proteasa hibrida con su diana. Esto causa la activación de la proteasa ya que en su estado hibridado la proteasa no está ya en la proximidad de su inhibidor de proteasa. Una célula en la que está presente la diana de dicha sonda proteasa y es reconocida, resulta dañada y se selecciona así. Las células que no expresan este ARNm resultan más o menos afectadas. A la inversa, la sonda proteasa puede diseñarse para catalizar una reacción proteolítica que estimula o confiere de otra manera un efecto beneficioso al crecimiento o viabilidad celular o confiere una ventaja de crecimiento a las células en las que su diana está presente y es reconocida.

15 Además, las sondas proteasa son útiles en varias aplicaciones adicionales. La actividad de la proteasa puede medirse fácilmente y además la proteasa activa en presencia de una secuencia diana de ácido nucleico particular puede emplearse no sólo para propósitos de detección sino también para propósitos terapéuticos, en los que, por ejemplo, una célula a la que se administra la sonda proteasa se proteoliza y se vuelve no viable si se transcribe un gen particular, por ejemplo, uno relacionado con la transformación celular, oncogénesis, desproliferación, y semejantes. Por ejemplo, dada
20 una mezcla de células en las que algunas de las células se infectan con un virus particular, se introduce en las células una sonda proteasa dirigida a un ARNm viral específico. Las células que portan dicho ARNm activan la actividad proteolítica de la sonda proteasa que contienen y esto destruye estas células.

25 Preferiblemente, el inhibidor de la enzima proteolítica es un péptido o compuesto químico pequeño, aunque también son útiles otras moléculas incluyendo pero no limitado a metales y quelantes de metales, para proporcionar una inhibición reversible de la enzima después de la interacción con el inhibidor. Los ejemplos de parejas útiles de enzimas proteolíticas e inhibidores de la enzima proteolítica incluyen pero no están limitados a aminopeptidasa y amastatina, proteasas de cisteína semejantes a tripsina y antipaína, aminopeptidasa y bestatina, proteasas de cisteína semejantes a quimiotripsina y quimioestatina, aminopeptidasa y diprotina A o B, carboxipeptidasa A y EDTA, proteasas de serina semejantes a elastasa y elastina y termolisina o aminopeptidasa M y 1,10-fenantrolina.

30 Además, las sondas que incorporan otras parejas de interacción pueden usarse cuando un miembro de la pareja de interacción tiene una actividad deseada y el segundo actúa para inhibir o disminuir esta actividad cuando las sondas no están unidas a la diana. Después de la unión a sus dianas, la actividad de la sonda se presenta ya que el miembro inhibidor de la pareja de interacción no está ya en la proximidad del miembro que tiene la actividad deseada.

Tomando como base la descripción anterior, pueden realizarse los métodos siguientes.

35 Un método para aislar células que expresan al menos un ARN, que comprende las etapas de:

- a) introducir en células ADN que codifica dicho al menos un ARN;
- b) exponer dichas células al menos a una sonda de señalización que produce una señal detectable después de hibridar con dicho al menos un ARN: y
- c) aislar dichas células que producen dicha señal.

40 Este método puede comprender además la etapa de crecer las células aisladas para generar una línea celular que expresa el ARN. Puede generarse una pluralidad de líneas celulares si la construcción de ADN se integra en diferentes localizaciones en el genoma de las células transfectadas. A no ser que la integración genómica de una construcción transfectada esté dirigida a una localización particular en el genoma, la integración se piensa que ocurre aleatoriamente, de manera que cada célula positiva puede ser diferente de otra y habrá múltiples líneas celulares diferentes todas
45 positivas para los ARN para los que se han seleccionado. También se puede generar una pluralidad de líneas celulares si la construcción de ADN se introduce en una población mixta de células, por ejemplo, células o líneas celulares inmortalizadas, primarias, madre y germinales. Las células también pueden ser de cualquier línea celular establecida, incluyendo pero no limitado a HeLa, HEK 293T, Vero, Caco, Caco-2, MDCK, COS-1, COS-7, K562, Jurkat, CHO-K1, Huvec, CV-1, HuH-7, NIH3T3, HEK293, 293, A549, HepG2, IMR-90, MCF-7, U-2 OS o CHO. Opcionalmente, la
50 construcción de ADN puede codificar además al menos un marcador de resistencia a fármaco u otro marcador seleccionable y el método puede comprender además la etapa de seleccionar las células usando el marcador seleccionable después de la etapa a). Las células aisladas pueden crecerse separadamente o combinadas. Siempre que

las células se aislen, ya sea después de la transfección con una o más construcciones o una o más bibliotecas de expresión, las células aisladas pueden crecerse separadamente entre sí o combinadas.

Las células aisladas pueden prepararse además para expresar un segundo ARN. Ya sea de manera simultánea o secuencial, las etapas adicionales incluyen transfectar las células o línea celular con una segunda construcción de ADN que codifica un segundo ARN; exponer dichas células a una segunda sonda de señalización que produce una señal detectable después de hibridar con dicho segundo ARN; y aislar las células que presentan la señal de al menos uno o ambos de dicho ARN y segundo ARN. La primera sonda de señalización puede producir la misma o una señal diferente que la segunda sonda de señalización, por ejemplo, puede tener el mismo o diferentes fluoróforos. Las células o líneas celulares que expresan más de dos ARN pueden proporcionarse mediante la repetición de las etapas simultáneamente o secuencialmente. La segunda construcción de ADN también puede contener el mismo o un marcador de resistencia u otro marcador seleccionable diferente. Si el primer y segundo marcadores de resistencia a fármaco son el mismo, puede conseguirse la selección simultánea incrementando el nivel del fármaco. Pueden generarse una pluralidad de líneas celulares mediante la repetición de las etapas anteriores de una manera simultánea o secuencial usando construcciones de ADN que forman una biblioteca de expresión, en la que al menos una parte de las células expresa un ARN diferente.

Se describe una estrategia relacionada en la que una secuencia etiqueta asociada con el gen transfectado se usa como la diana para la sonda de señalización, de la que una aplicación es permitir la selección de células cuyo ARN puede ser difícil de identificar por encima del fondo, por ejemplo, si la sonda de señalización detecta una especie de ARN muy relacionada. De acuerdo con esto, se describe un método para aislar células que expresan al menos un ARN que comprende las etapas de:

- a) introducir en células ADN que codifica dicho al menos un ARN y al menos una secuencia etiqueta;
- b) exponer dichas células al menos a una sonda de señalización que produce una señal detectable después de hibridar con la secuencia etiqueta; y
- c) aislar dichas células que producen la señal.

Este método es esencialmente el mismo que el descrito anteriormente, excepto en que la sonda de señalización usada se diseña para reconocer la secuencia etiqueta en lugar de dicho ARN. Un beneficio de este procedimiento sobre el anterior es que sólo se necesita un pequeño número de sondas de señalización, correspondiente al número de diferentes secuencias etiqueta, para preparar un gran número de diferentes líneas celulares que expresan uno o más ARN. Opcionalmente, la construcción de ADN puede codificar además al menos un marcador de resistencia a fármaco u otro marcador seleccionable y el método comprende además la etapa de seleccionar células resistentes al menos a un fármaco u otro agente selectivo para el que dicho marcador confiere resistencia después de la etapa a). Las células aisladas pueden crecerse separadamente o combinadas. Siempre que las células se aislen, ya sea después de la transfección con una o más construcciones o una o más bibliotecas de expresión, las células aisladas pueden crecerse separadamente entre sí o combinadas.

Las secuencias etiqueta se refieren a una secuencia de ácido nucleico que se expresa como parte de un ARN que se va a detectar con la sonda de señalización. Las sondas de señalización pueden dirigirse frente a la etiqueta mediante el diseño de las sondas para incluir una parte que es complementaria a la secuencia de la etiqueta. Los ejemplos de secuencias etiqueta que pueden usarse en la presente memoria y frente a las que pueden prepararse las sondas de señalización incluyen pero no están limitadas al transcrito de ARN de etiquetas epítipo que incluyen pero no están limitadas a HA (proteína de hemaglutinina de influenza), myc, his, proteína C, VSV-G, FLAG o FLU. Éstas y otras secuencias etiqueta son conocidas para un experto en la técnica y corresponden típicamente a secuencias de aminoácidos que pueden incorporarse en productos proteicos expresados y seleccionarse habitualmente tomando como base la disponibilidad de anticuerpos robustos o reactivos de detección de proteínas que pueden usarse para informar de su presencia. No se quiere que las secuencias etiqueta descritas en la presente memoria se refieran solamente a secuencias que pueden usarse para modificar a nivel de aminoácidos productos proteicos codificados por los ARN que se etiquetan o para ayudar en la detección posterior de cualquiera de estos productos proteicos modificados mediante el uso del anticuerpo o reactivos de detección de proteínas correspondientes. Tal y como se usa en la presente memoria, la secuencia etiqueta proporciona al menos una única secuencia de ácido nucleico para reconocimiento por una sonda de señalización. Las sondas de señalización se han descrito para usarse en la detección de una variedad de ARN. Cualquiera de estos ARN puede usarse como etiqueta. La parte de ADN de la construcción que codifica la secuencia etiqueta puede estar en marco o fuera de marco con la parte de la construcción de ADN que codifica la parte codificadora de proteína del al menos un ARN. Así, la secuencia etiqueta no necesita ser traducida para la detección por la sonda de señalización.

Una secuencia etiqueta puede comprender una secuencia repetida de forma múltiple que se diseña para actuar como sitios diana para una sonda de señalización. Dicha secuencia etiqueta proporcionará múltiples sitios diana para la sonda de señalización. Como resultado, puede unirse un número mayor de sondas de señalización. Esto incrementará la señal

total que puede generarse a partir de una molécula de ácido nucleico cualquiera que se quiere detectar por una sonda de señalización y así se incrementan las proporciones señal a ruido.

Además de las secuencias diana para las sondas, las etiquetas pueden comprender una o más secuencias adicionales (referidas como "auxiliar" o "secuencia auxiliar") diseñadas, identificadas o seleccionadas para mejorar la detección de las células que expresan las etiquetas. Por ejemplo, los auxiliares pueden tener un número de efectos incluyendo pero no limitados a efectos en el plegamiento, localización o estructura secundaria, terciaria o cuaternaria de las etiquetas en las que cualquiera o una combinación de estos efectos actúa para mejorar o incrementar la detección de la secuencia diana en las células. Los auxiliares pueden influir en el plegamiento o estructura de la etiqueta de manera que la secuencia diana se presenta de manera que es más accesible para la unión de la sonda. Esto podría resultar de emparejamientos de bases alterados o podría deberse a interacciones de unión entre los auxiliares y las proteínas u otros componentes celulares o introducidos. Los auxiliares pueden funcionar para estabilizar o para hacer más dinámico el plegamiento o estructura de las secuencias que los comprenden. También, los auxiliares pueden actuar para estabilizar respecto a la degradación las secuencias que los comprenden bien antes, durante o después de la unión de la sonda, y dichos efectos podrían resultar directamente de cambios en el plegamiento o estructura de secuencias o como consecuencia de la unión de proteínas o componentes celulares o introducidos en las secuencias auxiliares. También, los auxiliares pueden actuar para incrementar la transcripción de secuencias que los comprenden, por ejemplo aumentando la eficacia del inicio transcripcional o procesamiento, disminuyendo la finalización prematura de la transcripción o incrementando la eficacia del procesamiento post-transcripcional.

Pueden usarse estrategias funcionales para identificar auxiliares independientemente de cómo ejerzan su efecto y pueden existir diferentes secuencias auxiliares para cualquier etiqueta dada y sonda correspondiente. Las secuencias auxiliares que funcionan para múltiples conjuntos de una etiqueta y sonda correspondiente también pueden identificarse funcionalmente.

Pueden ensayarse secuencias variables para identificar cuáles actúan como auxiliares por ejemplo construyendo una biblioteca de expresión que comprende una secuencia génica y una secuencia etiqueta en la que las secuencias variables se insertan en medio de las secuencias génicas y etiqueta. Si el gen tiene un codón de parada, la secuencia variable puede insertarse en 3' respecto al codón de parada. Pueden insertarse secuencias variables adicionales en diferentes sitios. Después, se introduce la biblioteca de expresión en las células y las células se ensayan posteriormente por introducción de una sonda de señalización dirigida frente a la secuencia diana. Las células que presentan una señal incrementada por encima del control se detectan (en la que la señal control es la señal presentada por células control o células en las que se ha introducido una construcción de expresión control, por ejemplo, una que comprende el gen y la etiqueta pero no la secuencia variable adicional). Dichas células pueden aislarse por ejemplo por FACS y las secuencias variables representadas por ellas pueden aislarse y caracterizarse adicionalmente, si se desea. Esta estrategia puede usarse para detectar o identificar secuencias que actúan como secuencias auxiliares para las combinaciones específicas de un gen, etiqueta y sonda de señalización correspondiente usada. Esencialmente puede usarse la misma estrategia para encontrar secuencias auxiliares para cualquier secuencia que comprenda al menos una secuencia diana para una sonda de señalización correspondiente (es decir, una biblioteca de expresión que comprenda una secuencia variable y una secuencia diana pero por ejemplo ningún gen o una biblioteca de expresión que comprenda un gen, comprendiendo él mismo la secuencia diana y una secuencia variable pero ninguna etiqueta pueden usarse cada una para detectar auxiliares adecuados). Las células que se aíslan pueden crecerse y pueden generarse líneas celulares.

El beneficio de las secuencias auxiliares que se detectan o identifican de esta manera puede ser específico para el contexto de secuencia que se usó (es decir, para la etiqueta, gen, secuencia diana o sonda correspondiente específico). Las secuencias auxiliares más versátiles que son beneficiosas para varios contextos de secuencia más amplios pueden determinarse experimentalmente por ejemplo según los métodos descritos anteriormente. Por ejemplo, pueden realizarse ciclos iterativos de selección de secuencias variables en los que las secuencias variables que actúan como secuencias auxiliares pueden aislarse en cada ciclo y ensayarse en ciclos posteriores. En este caso por ejemplo cada ciclo posterior de ensayo se realizaría creando vectores de expresión en los que las secuencias aisladas del ciclo previo se usarían para crear una construcción de expresión que comprende un gen o etiqueta o secuencia diana que es diferente de la usada en la construcción de expresión del primer ciclo. Los métodos descritos también pueden usarse para confirmar la versatilidad de las secuencias auxiliares en diversos contextos de secuencia. La identificación de secuencias auxiliares versátiles o universales puede ser útil para ayudar en la detección de diversas secuencias en células.

Una fuente de secuencias variables podría ser una secuencia genómica. La secuencia genómica puede digerirse con enzimas de restricción para rendir fragmentos de varios tamaños que pueden obtenerse para clonación para crear las bibliotecas de expresión.

Las secuencias etiqueta pueden elegirse o diseñarse para presentar una determinada cantidad de estructura secundaria predicha o determinada experimentalmente, con el objetivo de presentar de manera óptima la secuencia etiqueta para la

- unión con la sonda de señalización. Como tales, las secuencias etiqueta comprenden al menos una secuencia frente a la que está dirigida al menos una sonda de señalización y las secuencias etiqueta pueden comprender además secuencia adicional que no se elige o diseña para interactuar directamente con la sonda de señalización. Los algoritmos de predicción del plegamiento de ácidos nucleicos pueden usarse para diseñar secuencias etiqueta potenciales según sus adaptaciones estructurales. Véase, por ejemplo, *Nucleic Acids Res.* 31: 3429-3431 (2003). Los algoritmos de predicción del plegamiento de ácidos nucleicos habitualmente predicen un número de estructuras energéticamente más favorables de una secuencia dada. Alternativamente, pueden ensayarse bibliotecas de secuencias que representan secuencias variables de ácidos nucleicos, por ejemplo incluyendo pero no limitado a ADN genómico digerido, para determinar o identificar qué secuencias actúan para ayudar en la detección de las secuencias etiqueta por las sondas de señalización. Esto representa una estrategia funcional para la identificación o aislamiento de secuencias etiqueta. Esto puede conseguirse creando bibliotecas de expresión que comprenden al menos una secuencia común elegida o diseñada para ser reconocida por la sonda de señalización y al menos una secuencia variable. Dicha biblioteca de expresión puede transfectarse en células, las células pueden exponerse a la sonda de señalización y las células más positivas pueden aislarse por FACS. Las secuencias variables representadas por estas células pueden aislarse. Por ejemplo, las secuencias pueden amplificarse directamente a partir de las células aisladas por técnicas de PCR seguido de clonación de los productos amplificados. Alternativamente, las células aisladas pueden lisarse para resultar en la liberación de las construcciones de ADN correspondientes a las secuencias variables expresadas en las células aisladas y estas construcciones o vectores pueden aislarse y propagarse a partir del material resultante. Por ejemplo, en el caso en el que las construcciones son vectores plasmídicos, el material lisado puede usarse para transformar células bacterianas competentes, seguido de aislamiento y amplificación del plásmido usando huéspedes bacterianos. Para secuencias etiqueta que incorporan múltiples unidades de secuencia repetidas, cada una de estas unidades puede no adoptar necesariamente la misma estructura debido a las interacciones potenciales entre las unidades repetidas y/o otra secuencia presente en la molécula que incorpora la secuencia etiqueta. La estructura de cualquier secuencia etiqueta dada podría estar influida por su contexto de secuencia.
- En una realización, la secuencia etiqueta se obtiene de ARN inverso *vav*. En una realización, la secuencia etiqueta forma una estructura de unión de tres brazos (Figura 36). En una realización, la secuencia etiqueta tiene una longitud de 10-100, 80-100, 90-100, 80-120, 100-2Kb, 2Kb-15Kb nucleótidos. En una realización, la secuencia diana es la región de todo o parte del extremo 3' del tallo de la primera región tallo-bucle, hasta la unión entre la primera y segunda región tallo-bucle, hasta todo o parte del extremo 5' del tallo de la segunda región tallo-bucle (Figura 36).
- En una realización, la región tallo comprende 8-9 pares de bases, la primera región tallo-bucle comprende 4-6 pares de bases y la segunda región tallo-bucle comprende 13-17 pares de bases (Figura 37). En una realización, las regiones tallo de los tres brazos comprenden además regiones no complementarias. En una realización, el tallo de la región tallo y la primera región tallo-bucle comprenden además una región de emparejamiento erróneo, la segunda región tallo-bucle comprende además 2-7 emparejamientos erróneos o regiones de protuberancia. En una realización, la unión entre las regiones tallo tiene un total de 8-12 nucleótidos (Figura 37). En una realización, la secuencia etiqueta comprende la estructura o secuencia según la Figura 42 A, B o C. En otra realización, la secuencia etiqueta tiene las estructuras energéticamente más favorables predichas a partir de la secuencia según la Figura 42 A, B o C.
- En la presente memoria también se describe una construcción de ADN que comprende al menos un ADN que codifica al menos un ARN de interés y una secuencia etiqueta como se ha descrito anteriormente. En la presente memoria también se describen vectores y células que comprenden la construcción de ADN.
- En una realización adicional, la línea celular puede prepararse para expresar al menos un segundo ARN; incluyendo las etapas además transfectar la línea celular con una segunda construcción de ADN que codifica el segundo ARN y una segunda secuencia etiqueta y, opcionalmente, un segundo marcador seleccionable, por ejemplo, un marcador de resistencia a fármaco, opcionalmente, seleccionar para células que transcriben el segundo marcador; exponer las células a una segunda sonda de señalización que produce una señal detectable después de hibridar con la segunda secuencia etiqueta y aislar las células que presentan la señal de ambas o al menos una de la primera y segunda secuencia etiqueta. En el caso de dos ARN, la parte de la secuencia de ADN que codifica la segunda secuencia etiqueta también puede estar en marco o fuera de marco con la parte de la secuencia de ADN que codifica la parte codificadora de proteína del segundo ARN. El segundo ARN puede transfectarse bien simultáneamente o secuencialmente con el primero. Si el método se realiza simultáneamente y se usa el mismo marcador seleccionable, por ejemplo, un marcador de resistencia a fármaco para ambas construcciones, puede usarse un nivel más alto de fármaco o agente selectivo apropiado para seleccionar para células que expresan ambas construcciones. Además, más de dos ARN pueden proporcionarse en la línea celular repitiendo las etapas mencionadas anteriormente.
- Una pluralidad de líneas celulares puede generarse repitiendo las etapas anteriores de una manera simultánea o secuencial usando construcciones de ADN que forman una biblioteca de expresión. En una realización, la biblioteca de expresión usa una única secuencia etiqueta y las células se exponen a la misma sonda de señalización que es complementaria a la secuencia etiqueta. En otra realización, diferentes bibliotecas de expresión pueden usar diferentes

secuencias etiqueta. En una realización, cada línea celular expresa un ARN. En otra realización, cada línea celular expresa más de un ARN. Esto puede conseguirse transfectando la célula con múltiples construcciones de ADN secuencialmente o simultáneamente. La probabilidad de obtener una célula transfectada de manera estable con múltiples construcciones de ADN puede incrementarse introduciendo una mayor concentración de construcciones de ADN usadas para la transfección. Alternativamente, las células de partida pueden haberse transfectado ya con una primera construcción de ADN para obtener una célula con múltiples construcciones de ADN. También pueden usarse múltiples bibliotecas de expresión diferentes para transfectar las células. Cada biblioteca puede incorporar una secuencia etiqueta distinta que se detecta usando las sondas de señalización correspondientes. Cada biblioteca de expresión puede comprender un marcador seleccionable, por ejemplo, un gen de resistencia a fármaco.

- 5
- 10 Las células aisladas pueden crecerse individualmente o combinadas. Las células aisladas individualmente o combinadas pueden crecerse para proporcionar poblaciones de células. Las líneas celulares pueden generarse creciendo células aisladas individualmente. Las líneas celulares individuales o múltiples pueden crecerse separadamente o combinadas. Si una combinación de líneas celulares produce una actividad deseada, puede fraccionarse adicionalmente hasta que se identifica la línea celular o conjunto de líneas celulares que tiene este efecto. Esto puede facilitar el mantenimiento de
- 15 altos números de líneas celulares sin los requerimientos de mantener cada una separadamente.

Se describe otro método más para generar una línea celular que sobreexpresa un ARN que comprende las etapas de:

- a) introducir en las células un primer ADN que codifica dicho ARN y una primera secuencia etiqueta; y al menos un segundo ADN que codifica dicho ARN y una segunda secuencia etiqueta, en el que la primera y segunda secuencias etiqueta son diferentes;
- 20 b) exponer dichas células al menos a una sonda de señalización que produce una señal detectable después de hibridar con dicha primera secuencia etiqueta y al menos a una sonda de señalización que produce una señal detectable después de hibridar con dicha segunda secuencia etiqueta; y
- c) aislar las células que presentan la señal de al menos una de dichas sondas de señalización.

Este método puede comprender además la etapa de crecer las células aisladas para generar una línea celular que expresa o sobreexpresa el ARN. Una pluralidad de líneas celulares puede generarse si la construcción de ADN se integra en diferentes localizaciones en el genoma de la célula transfectada. A no ser que la integración genómica de una construcción transfectada este dirigida a una localización particular en el genoma, la integración se produce aleatoriamente, de manera que cada célula positiva puede ser diferente de las demás y habrá múltiples líneas celulares diferentes todas positivas para los ARN para los que se seleccionan. Opcionalmente, la construcción de ADN puede codificar además al menos un marcador seleccionable, por ejemplo, un marcador de resistencia a fármaco y el método puede comprender además la etapa de seleccionar las células usando el marcador seleccionable, por ejemplo, seleccionar para resistencia al menos a un fármaco para el que dicho marcador confiere resistencia después de la etapa a). Siempre que las células se aíslan, ya sea después de la transfección con una o más construcciones o una o más bibliotecas de expresión, las células aisladas pueden crecerse separadamente entre sí o combinadas.

35 En una realización, las células expresan un antisentido o ARNsi o ARNsh o proteína. Además, las células que se han hecho que expresen una proteína o proteínas particulares pueden usarse como punto de partida para crear células que expresan proteínas y moléculas de ARN antisentido. Por supuesto, las células que expresan las moléculas antisentido o ARNsi o ARNsh pueden usarse como punto de partida para añadir ARN adicionales que codifican proteínas adicionales, usando los métodos de la presente memoria. La transfección simultánea de ARN que codifican proteínas y moléculas antisentido o ARNsi o ARNsh, con las sondas de señalización correspondientes y, si se desea, secuencias etiqueta, también puede realizarse. Las diferentes combinaciones de los procedimientos mencionados anteriormente están englobadas por la presente memoria.

Asimismo, se describen métodos para aislar células que expresan al menos un ARN que comprenden las etapas de:

- a) proporcionar células que se sospecha que expresan dicho al menos un ARN;
- 45 b) exponer dichas células al menos a una sonda de señalización que produce una señal detectable después de hibridar con dicho al menos un ARN;
- c) aislar dichas células que producen la señal.

El método también puede usarse para identificar células que también expresan un segundo ARN, usando una segunda sonda de señalización que produce una señal detectable después de hibridar con su ARN diana, aislándose las células que tienen fluorescencia de ambas o de cada una de la primera y segunda sondas de señalización. La expresión simultánea de más de dos ARN también se puede conseguir. Siempre que las células se aíslan, ya sea después de

50

transfección con una o más construcciones o una o más bibliotecas de expresión, las células aisladas pueden crecerse separadamente una de otra o combinadas.

Se describe otro método para aislar células que expresan al menos un ARN exógeno y un ARN endógeno, que comprende las etapas de:

- 5 a) introducir en las células el ADN que codifica dicho al menos un ARN exógeno, en el que dichas células expresan potencialmente al menos un ARN endógeno;
- b) exponer dichas células al menos a una primera sonda de señalización que produce una señal detectable después de hibridar con dicho al menos un ARN exógeno;
- 10 c) exponer dichas células al menos a una segunda sonda de señalización que produce una señal detectable después de hibridar con dicho al menos un ARN endógeno, en el que dicha segunda sonda de señalización produce una señal diferente de la de la primera sonda de señalización; y
- d) aislar dichas células que producen al menos una de dichas señales después de hibridar dichas sondas de señalización con sus ARN respectivos.

15 Los dos métodos anteriores pueden comprender además la etapa de generar una línea celular o una pluralidad de líneas celulares que expresan dicho al menos un ARN exógeno y al menos un ARN endógeno creciendo dichas células aisladas.

Estos métodos son útiles para un ARN que expresa una proteína, por ejemplo, una proteína localizada en la superficie celular, proteína intracelular, proteína secretada u otra proteína. Estos métodos no requieren el uso de sondas para las proteínas en sí mismas, lo que puede ser más difícil o podrá afectar a la célula de manera que no pueden realizarse experimentos adicionales. Más de un ARN que codifica una proteína puede identificarse usando una pluralidad de sondas de señalización, hasta el número detectable simultáneamente por la tecnología usada para el aislamiento. Opcionalmente, la construcción de ADN puede codificar además al menos un marcador seleccionable, por ejemplo, un marcador de resistencia a fármaco y el método puede comprender además la etapa de seleccionar las células usando el marcador seleccionable, por ejemplo, seleccionar células que son resistentes al menos a un fármaco para el que dicho marcador confiere resistencia después de la etapa a). Las células aisladas pueden crecerse separadamente o combinadas. Siempre que las células se aíslan, ya sea después de la transfección con una o más construcciones o una o más bibliotecas de expresión, las células aisladas pueden crecerse separadamente entre sí o combinadas.

Para los métodos anteriores, en una realización, las células pueden implantarse en un animal. En una realización, la sonda de señalización es una sonda fluorogénica y las células que expresan dicho ARN emiten fluorescencia. El aislamiento de dichas células que emiten fluorescencia puede realizarse usando tecnología de separación de células activada por fluorescencia o cualquier tecnología que pueda usarse para aislar células basada en fluorescencia. En una realización, se usan dos sondas de señalización para dirigirse a la secuencia de ARN o etiqueta. El fluoróforo de la primera sonda puede ser el mismo o diferente del de la segunda sonda. En el caso de dos fluoróforos diferentes, pueden tener longitudes de onda de emisión similares o diferentes. Actualmente, la tecnología FACS puede permitir la detección de hasta siete fluoróforos diferentes durante un procedimiento de separación. Los métodos anteriores pueden repetirse simultáneamente para obtener la expresión de hasta siete proteínas diferentes. Si se desea, una línea celular que expresa siete proteínas diferentes puede usarse como punto de partida para la introducción de más proteínas según el procedimiento. Al avanzar la tecnología FACS, será capaz de resolver un mayor número de señales y se podrán seleccionar células que tienen una mayor número de ARN en una aplicación de FACS.

40 Naturalmente, los procedimientos mencionados anteriormente pueden usarse para cuantificar el nivel de expresión de al menos un transcrito de ARN en una muestra biológica que comprende las etapas de:

- a) exponer la muestra biológica a una primera sonda de señalización que produce una señal detectable después de hibridar con dicho transcrito de ARN;
- b) cuantificar el nivel de la señal en la muestra biológica; y
- 45 c) correlacionar el nivel de la señal con dicho nivel del al menos un transcrito de ARN.

La muestra biológica puede ser una muestra celular, una muestra de tejido o preparaciones obtenidas de éstas; éstas pueden estar congeladas y/o fijadas, por ejemplo, con formaldehído, glutaraldehído o cualquiera de los numerosos fijadores celulares que no interfieren con la detección de ARN usando sondas de señalización. Las preparaciones de muestras celulares incluyen pero no están limitadas a orgánulos o compartimentos subcelulares, mitocondria, orgánulos de células, fracciones subcelulares, membrana plasmática de células. material intracelular o extracelular, preparaciones

de membrana, preparaciones de ácido nucleico de una cualquiera o más de éstas, preparaciones de cualquier virus u otro virus u organismo presente en una cualquiera o más de éstas o cualquier combinación de una o más de éstas.

5 Para la realización de sondas fluorogénicas, la fluorescencia puede cuantificarse por microscopía de fluorescencia o tecnología de separación celular activada por fluorescencia. Pueden cuantificarse especies de ARN adicionales (es decir, cuantificarse) simultáneamente usando una segunda sonda de señalización que emiten fluorescencia después de
 10 hibridar con un segundo transcrito de ARN. El método anterior puede usarse simultáneamente con ensayos que utilizan un informador fluorogénico para la detección de eventos, estados o composiciones intracelulares. Dichos ensayos fluorescentes incluyen pero no están limitados a TUNEL, Apoptosis, necrosis, flujo Ca^{2+} /ión, flujo pH, inmunofluorescencia, marcaje de orgánulos, adhesión celular, ciclo celular, contenido de ADN y ensayos usados para
 15 detectar las interacciones entre: proteína-proteína, proteína-ADN, proteína-ARN. Los reactivos que pueden marcarse con fluorescencia para usarse en estos ensayos incluyen pero no están limitados a Proteínas (marcadas con moléculas fluorescentes o proteínas autofluorescentes); indicadores metabólicos fluorescentes (C12 resazurina, CFSE para divisiones celulares); sustratos o subproductos fluorescentes; lectinas marcadas con fluorescencia; químicos fluorescentes; compuestos fluorescentes enjaulados; marcajes de ácidos nucleicos fluorescentes; polímeros, lípidos, residuos de aminoácidos y/o análogos de nucleótidos fluorescentes.

La introducción de moléculas antisentido o ARNsi en las células es útil para eliminar o reducir funcionalmente los niveles de una o más proteínas o ARN de la célula. Según los métodos anteriores, se proporciona un método para aislar células o generar células que funcionalmente carecen o tienen reducida la expresión de al menos una proteína o ARN
 20 preseleccionado que comprende las etapas de proporcionar en dichas células una pluralidad de antisentido o ARNsi para dicha proteína o ARN preseleccionado, cada uno proporcionado según los métodos mencionados anteriormente, en el que dicha pluralidad de antisentido o ARNsi se une esencialmente a todo o a un nivel suficiente de transcritos de ARNm de dicho al menos una proteína o ARN preseleccionado. La proteína preseleccionada puede ser una forma de corte y empalme alternativo de un producto génico.

Según líneas similares, se proporciona un método para generar un animal transgénico que funcionalmente es un mutante que carece de la expresión de al menos una proteína o ARN preseleccionado o que expresa dicha al menos una proteína o ARN
 25 preseleccionado a niveles reducidos, que comprende realizar las etapas descritas anteriormente en la presente memoria utilizando células madre embrionarias y usando dichas células madre embrionarias viables para producir dicho animal transgénico.

Asimismo, se proporciona un método para aislar células o generar una línea celular que funcionalmente carece o está reducida para la expresión de al menos una proteína o ARN y sobreexpresa al menos una proteína o ARN distinto, que
 30 comprende realizar los métodos de la presente memoria en las mismas células. De manera similar, se describe un método para generar una línea celular que expresa un antisentido o ARNsi letal bajo el control de un promotor inducible o una secuencia que tiene alguna otra actividad de unión o funcional que resulta en la modulación de los niveles de transcripción. Esto puede conseguirse realizando el método de la presente memoria, en el que la etapa de transfección se realiza en presencia de una cantidad mínima de un inductor o compuesto.

Por lo tanto, en la presente memoria se describe un método para aislar células que sobreexpresan al menos una primera proteína y que funcionalmente carecen de la expresión o tienen expresión reducida para al menos una segunda proteína, que comprende las etapas de:

40 a) introducir en las células al menos un primer ADN que codifica al menos un ARN que codifica dicha al menos primera proteína y al menos una primera secuencia etiqueta; y al menos un segundo ADN que codifica dicho al menos un ARN y al menos una segunda secuencia etiqueta, en el que dicha primera y segunda secuencias etiqueta son diferentes;

b) introducir en las células al menos un ADN que codifica al menos un ARN antisentido o ARNsi que se une a o interfiere con el transcrito de ARNm de dicha al menos segunda proteína;

45 c) exponer dichas células al menos a una primera sonda de señalización que produce una señal detectable después de hibridar con dicha al menos primera secuencia etiqueta y al menos a una segunda sonda de señalización que produce una señal detectable después de hibridar con dicha al menos segunda secuencia etiqueta;

d) exponer dichas células al menos a una sonda de señalización que produce una señal detectable después de hibridar con dicho al menos un ARN antisentido o ARNsi; y

50 e) aislar las células que producen al menos una de dichas señales después de hibridar dichas sondas de señalización con sus ARN respectivos.

En otra realización más, la presente invención proporciona un método para identificar un compuesto que modula la transcripción de al menos un ARN preseleccionado, que comprende las etapas de:

- a) añadir un compuesto individual o conjunto de compuestos a las células;
- b) exponer dichas células al menos a una sonda de señalización que produce una señal detectable después de hibridar con dicho al menos un ARN preseleccionado;
- c) cuantificar el nivel de la señal en dichas células;
- 5 d) identificar las células que tienen un incremento o disminución en la señal comparado con la señal de las células sin compuesto añadido; y opcionalmente
- e) identificar los compuestos que modulan la transcripción de dicho al menos un ARN preseleccionado.
- En una realización, dicho ARN preseleccionado está codificado por el genoma de la célula. En una realización, dicho ARN preseleccionado está codificado por una construcción de ADN que se transfecta en las células antes de la etapa a).
- 10 En una realización, las células transfectadas se exponen a una sonda de señalización diseñada para reconocer dicho ARN y las células expresan dicho ARN. En una realización, la construcción de ADN comprende un promotor u operador y codifica un represor, amplificador o una secuencia que modula la transcripción. En una realización, el ARN preseleccionado está unido a una secuencia etiqueta y la sonda de señalización produce una señal detectable después de hibridar con la secuencia etiqueta.
- 15 En otra realización, en la presente memoria se describe un método para identificar una secuencia de ARN que modula la transcripción de al menos un ARN preseleccionado, que comprende las etapas de:
- a) introducir en las células al menos una construcción que codifica una secuencia de ARN de ensayo que potencialmente modula la transcripción de dicho al menos un ARN preseleccionado;
- b) exponer dichas células al menos a una sonda de señalización que produce una señal detectable después de hibridar con dicho al menos un ARN preseleccionado;
- 20 c) cuantificar el nivel de la señal en dichas células:
- d) seleccionar las células que tienen un incremento o disminución en la señal comparado con la señal de las células sin secuencia de ARN de ensayo; y opcionalmente
- e) identificar una secuencia de ARN de ensayo que modula la transcripción de dicho al menos un ARN preseleccionado.
- 25 Estas células pueden aislarse y crecerse para proporcionar líneas celulares. Pueden crecerse separadamente o combinadas. La modulación de la transcripción puede ser una regulación al alza o a la baja aguas abajo del ARN preseleccionado. En una realización, dicho ARN preseleccionado está codificado por el genoma. Dicho ARN preseleccionado está codificado por una construcción de ADN que se transfecta en las células antes de la etapa a). En una realización las células transfectadas se exponen a una sonda de señalización diseñada para reconocer dicho ARN y las células expresan dicho ARN. En una realización, el ARN preseleccionado está unido a una secuencia etiqueta y la sonda de señalización produce una señal detectable después de hibridar con la secuencia etiqueta. En una realización,
- 30 se usa una biblioteca de expresión de secuencias de ARN para identificar las secuencias de ARN que modulan la transcripción. En una realización, la secuencia de ARN de ensayo está unida a una secuencia etiqueta. La etapa e) se facilita exponiendo las células después de la etapa a) al menos a una sonda de señalización que produce una señal detectable después de hibridar con dichas secuencias de ARN o secuencias etiqueta, seguido de la etapa b). Las señales producidas por las sondas de señalización dirigidas a la secuencia de ARN de ensayo o secuencia etiqueta pueden ser diferentes de la sonda de señalización dirigida al ARN preseleccionado y por lo tanto las células pueden exponerse a diferentes sondas de señalización simultáneamente.
- 35 En la presente memoria también se describe un método para identificar eventos de recombinación genética en células vivas que comprende las etapas de:
- 40 a) exponer una célula a una sonda de señalización que produce una señal detectable después de hibridar con una secuencia de ARN seleccionada del grupo que consiste en la transcrita a partir de una secuencia recombinada y la transcrita a partir de la secuencia no recombinada; y
- b) detectar dichas células que expresan dicha secuencia de ARN. La detección y/o separación de las células puede realizarse por FACS o microscopía de fluorescencia.
- 45

La presente invención puede entenderse mejor por referencia a los Ejemplos siguientes no limitativos, que se proporcionan como ejemplares de la invención. Los ejemplos siguientes se presentan con el fin de ilustrar más completamente las realizaciones preferidas de la invención. Sin embargo, esto no deben considerarse en ningún caso

como limitantes del amplio alcance de la invención. Los métodos y materiales similares o equivalentes a los descritos en la presente memoria también pueden usarse en la práctica o ensayo de la presente invención.

17. Identificación de reactivos o compuestos para introducir o mejorar la introducción de sondas de señalización en células o para aumentar la detección de dianas

5 Los reactivos que pueden usarse para introducir sondas de señalización en células pueden identificarse ensayando varios químicos incluyendo pero sin limitarse a proteínas, lípidos, polímeros, extractos o compuestos o mezclas de éstos. Los químicos pueden estar en forma de gas, líquido o sólido. Esto puede hacerse mezclando los químicos con la sonda de señalización en varias proporciones que van de 1:1.000.000.000 a 1.000.000.000:1 en varios disolventes incluyendo pero sin limitarse a disolventes orgánicos y acuosos, disoluciones o medios tamponados o cualquier combinación de éstos, en los que la mezcla se incuba durante varios periodos de tiempo que van de 1 minuto a 48 horas y en los que la incubación se realiza bajo varias condiciones de temperatura que van de por debajo de 0 a por encima de 100 grados Celsius y con varios grados de agitación o mezclado de nada hasta agitación suave u oscilación a agitación con vórtex o pipeteo constante y en los que la incubación se realiza en luz o en oscuridad o bajo varias condiciones ambientales diferentes. Después, las mezclas se aplicarán a las células y la administración de la sonda se ensayará usando microscopía de fluorescencia, tecnología de lectura de placas de fluorescencia o por FACS u otro método de detección de fluorescencia. Esto puede hacerse, por ejemplo, en placas de 96, 384 ó 1.536 pocillos o en lechos compatibles con el análisis posterior de alto rendimiento. Obsérvese que cuando se usan lechos, los lechos pueden analizarse o aislarse usando FACS y si hay lechos aislados codificados adicionalmente pueden usarse para determinar la identidad de las mezclas con las que se han tratado.

20 Las células usadas pueden estar vivas o fijadas y pueden presentarse adheridas a una superficie sólida incluyendo pero no limitada a placas de cultivo tisular o lechos o pueden estar en suspensión. Las células pueden lavarse usando una variedad de tampones o disoluciones antes de la adición de las mezclas. Una vez que los reactivos se han aplicado a las células la reacción puede incubarse durante varios periodos de tiempo con grados variables de agitación o mezclado en luz u oscuridad o bajo varias condiciones ambientales distintas y a varias temperaturas, todo como se ha descrito anteriormente, y en la que estos parámetros pueden limitarse por ejemplo si se quiere mantener la viabilidad celular. Después de la incubación, las células pueden analizarse directamente o lavarse como se ha descrito anteriormente antes del análisis.

30 Los métodos anteriores pueden realizarse usando una sonda de señalización o una sonda constitutivamente activa (por ejemplo en el caso de sondas fluorogénicas, puede usarse una sonda que carece de un apantallador). También, cada una de las mezclas que se va a ensayar puede ensayarse usando más de una muestra de células. Por ejemplo, puede ensayarse una mezcla usando tanto células que se sabe que comprenden la diana de la sonda de señalización usada así como células control. Los múltiples tipos celulares pueden presentarse mezclados o separadamente. Las mezclas preferidas serán aquellas que resultan en una señal a ruido incrementada, por ejemplo aquellas que resultan en una señal incrementada en células que comprenden la diana comparada con células control.

35 De manera similar, podrían realizarse esencialmente los mismos métodos excepto cuando las células se añaden a las mezclas. Por ejemplo, las mezclas que se van a ensayar podrían aplicarse en primer lugar a las cámaras de ensayo por ejemplo a pocillos de placas de 96, 384 ó 1.536 pocillos o lechos, y las células podrían aplicarse a continuación. Las células pueden añadirse directamente a las mezclas sembradas en las placas o pueden añadirse una vez que las mezclas se han procesado, por ejemplo, por secado, calentamiento o evaporación.

40 De una manera relacionada, los compuestos pueden ensayarse para su capacidad de aumentar la detección de dianas en células usando sondas de señalización. En este caso, las sondas de señalización se introducirán en primer lugar en las células en el que las células podrían incluir tanto células que se sabe que comprenden la diana como células control. Después, podrían ensayarse los químicos como se ha descrito anteriormente y usando condiciones y parámetros variables como se ha descrito anteriormente para su capacidad de mejorar la señal. Obsérvese que los químicos usados en este caso no se mezclarán con las sondas de señalización. Los químicos preferidos serán aquellos detectados o identificados para incrementar las proporciones señal a ruido, por ejemplo aquellos que resultan en una señal incrementada en células que comprenden la diana comparado con las células control. Los químicos identificados usando esta estrategia pueden actuar mediante varios mecanismos incluyendo pero no limitado a incrementar la administración de la sonda de señalización en el citoplasma de la célula, influir en el plegamiento o estructura de la diana de manera que su detección por la sonda se mejora (por ejemplo causando una mejor accesibilidad de la diana) o reducir la señal de fondo no específica (por ejemplo reduciendo la cantidad de sonda de señalización unida a la superficie exterior de las células).

55 Los químicos que actúan incrementando las proporciones señal a ruido que pueden usarse para introducir las sondas de señalización en células o que pueden usarse para aumentar la detección de la diana también pueden determinarse usando sondas de señalización de ensayo y control cada una aplicada a células que se sabe que comprenden la diana y

analizando los químicos para una señal incrementada cuando se usa una sonda de señalización de ensayo comparado con una sonda de señalización control.

Ejemplo 1

5 Protocolo General. Material de Partida: Las sondas de señalización pueden introducirse en las células que no expresan ningún ARN de la construcción de ADN o pueden usarse para detectar mensajes de ARN codificados por la construcción de ADN. El método de introducción de las sondas de señalización en cualquiera de estos dos tipos de células es idéntico. El protocolo siguiente sólo requiere que las células que se van a analizar se puedan separar entre sí y que se puedan analizar por FACS.

10 1) Como se describe más detalladamente en la descripción de la invención, las sondas de señalización pueden usarse conjuntamente con FACS para separar células de un tejido tomando como base la expresión o ausencia de expresión en las células de ARN específicos. Para este objetivo, las células se separan en primer lugar entre sí por métodos estándar y bien establecidos tales como homogeneización y tratamiento químico adicional. Las sondas de señalización apropiadas se introducen en dichas células según el protocolo siguiente.

15 2) En segundo lugar, se pueden usar las sondas de señalización para seleccionar células que expresan ARN particulares codificados por la construcción de ADN que se ha transfectado en una población de células. Para este objetivo, se transfecta en primer lugar en un cultivo de células una construcción de ADN o construcciones de ADN que codifican los ARN deseados. Las sondas de señalización pueden generarse para reconocer estos ARN, como se describe con más detalle en la descripción de la invención. La transfección de la construcción de ADN en las células puede conseguirse mediante una gran variedad de métodos incluyendo pero no imitados al uso de cualesquiera de los reactivos o kits obtenidos de empresas biotécnicas (Qiagen, Promega, Gene Therapy Systems, Invitrogen, Stratagene, etc.) según las instrucciones de los fabricantes. Las construcciones de ADN se eligen de manera que cada una confiera resistencia a un antibiótico. Después de la transfección de estas construcciones de ADN en las células y de un periodo breve para la recuperación de las células (habitualmente 24 horas), las células se someten a los antibióticos apropiados de manera que sólo aquellas células a las que las construcciones de ADN han conferido resistencia al antibiótico sobrevivan. Esto dura generalmente tres a cuatro días y algunas veces más tiempo, dependiendo tanto del tipo de célula como del antibiótico usado.

El resultado es que un conjunto de células permanece y que todas éstas serán resistentes a los antibióticos, pero sólo una pequeña fracción de éstas expresa los ARN de interés. Para seleccionar las células que expresan los ARN deseados, puede seguirse el protocolo siguiente.

30 Ejemplo 2 Selección de las Células Usando Sondas de Señalización

1) Transfección de las sondas de señalización en las células: las sondas de señalización deben diseñarse de manera que reconozcan el ARN deseado mediante hibridación a una secuencia endógena en el ARN o mediante hibridación a una etiqueta que se añade a la secuencia nativa de ARN. El diseño de las sondas de señalización se elabora según la descripción de la invención.

35 La transfección puede realizarse por una gran variedad de métodos, de manera similar a la transfección de las construcciones de ADN en las células. El método empleado debe elegirse tomando como base el tipo celular que se usa ya que algunas células responden mejor a algunos métodos de transfección respecto a otros métodos. La transfección debe realizarse según las instrucciones del fabricante del reactivo de transfección usado y puede ser necesario optimizarlo. La optimización puede incluir el tratamiento con químicos para aumentar la administración del material transfectado en la célula o citoplasma de la célula.

La transfección de sondas de señalización en células puede realizarse bien en células en suspensión o en células que crecen en superficies sólidas, dependiendo de las células y del reactivo de transfección usados.

45 2) Después de la transfección de sondas fluorogénicas en células, las células pueden someterse a análisis FACS. FACS puede usarse para separar las células positivas para una cualquiera o más de las sondas fluorogénicas usadas. También puede usarse para separar células tomando como base la intensidad de la señal de las sondas fluorogénicas, permitiendo de esta manera al investigador seleccionar las células que expresan los ARN a diferentes niveles.

Ejemplo 3

Generación de Líneas Celulares que Expresan Uno o Más ARN

50 Después de la selección por FACS, las células positivas pueden mantenerse en un medio apropiado como se describe con más detalle en la descripción de la invención. Estas células proporcionarán líneas celulares que expresan los ARN de interés.

Concentración de la sonda de señalización: La concentración de la sonda de señalización que se va a usar depende de varios factores. Por ejemplo, se debe considerar la abundancia en las células del ARN que se va a detectar y la accesibilidad de este ARN para la sonda de señalización. Por ejemplo, si el ARN que se va a detectar está presente en cantidades muy bajas o si se encuentra en una parte del ARN que no es fácilmente accesible tomando como base el plegamiento tridimensional del ARN o debido a la unión de proteínas al ARN, entonces se usará más sonda de señalización aquí que en los casos en los que el ARN que se va a detectar es muy abundante y en el que el sitio reconocido por la sonda de señalización es muy accesible. La cantidad exacta de la sonda de señalización que se va a usar tendrá que determinarse empíricamente para cada aplicación.

Esto puede conseguirse introduciendo diferentes cantidades de las sondas de señalización en diferentes grupos de células y seleccionando la condición en la que la fluorescencia de fondo es baja y en la que la señal es alta (la condición en la que no todas sino algunas de las células son positivas para la sonda de señalización).

Ejemplo 4

Exposición de las Células a las Sondas de Señalización

Las células adheridas, parcialmente adheridas a o establecidas en superficies, o en disolución pueden exponerse a las FP usando varios métodos para introducir moléculas en células, incluyendo pero sin limitarse a métodos conocidos en la técnica tales como microinyección, fuerzas de cizalla mecánica tales como vórtex o mezclado, paso a través de agujas o técnicas de carga de células incluyendo raspado, permeabilización usando reactivos tales como determinados antibióticos o detergentes o una combinación de reactivos o disolventes o mediante el uso de varios reactivos de transfección con distintas propiedades químicas (por ejemplo basados en liposomas, basados en químicos o proteínas) o mediante una combinación de uno cualquiera o más de estos métodos.

A continuación se muestra un protocolo de muestra más detallado que describe el uso de reactivos de transfección basados en lípidos:

1. Preparación de las Células

Las células se sembraron en placas en pocillos de cultivo tisular antes de (método de siembra de células en placas 1) o en el mismo día de su exposición a (método de siembra de células en placas 2) las FP o las células se transfirieron a tubos de microcentrífuga en el mismo día de su exposición a las FP (método de siembra de células en placas 3).

Estos tres métodos preparativos se usaron con éxito. El método de siembra de células en placas 1 permite un tiempo suficiente para que las células se adhieran a la superficie del pocillo de cultivo, mientras que el métodos de siembra de células en placas 2 permite que las células se asienten en la placa o se adhieran en varios grados para ser procesadas dependiendo de cuánto tiempo pase antes del procesamiento adicional de las células y el método de siembra de células en placas 3 permite que las células sean procesadas directamente después de haber sido transferidas a los tubos sin dejar ningún tiempo para que se asienten o se adhieran a ninguna superficie, aunque el procesamiento también puede realizarse después de que haya pasado un tiempo dado.

Las células se lavan una vez o más con tampón tal como medio sin suero o PBS, aunque pueden usarse otros tampones. Generalmente, los tampones incluyeron $MgCl_2$ a diferentes concentraciones milimolares. La etapa de lavado puede omitirse dependiendo del método de exposición.

2. Preparación de los Reactivos de FP y Exposición a las Células

Las FP se prepararon para adición a las preparaciones de células usando reactivos de transfección disponibles comercialmente y según los protocolos que describen los fabricantes. Los protocolos de los fabricantes instruyen que hay múltiples parámetros que necesitan determinarse empíricamente para múltiples variables incluyendo qué tipos de células se usan, a qué confluencia o concentración deben procesarse las células, qué moléculas específicas deben introducirse, en qué proporción y cantidades absolutas deben combinarse los diferentes reactivos y durante cuánto tiempo deben realizarse o incubarse las diferentes etapas, qué etapas pueden omitirse y otras variables. En el protocolo de los fabricantes, aunque se proporcionan determinados intervalos, los protocolos también sugieren que éstos pueden excederse u otros parámetros pueden tener que optimizarse para el uso exitoso de su reactivo. En general, la exposición de las células a las FP se realizó usando los parámetros para estas condiciones que estuvieron en los intervalos sugeridos como los intervalos preliminares en los protocolos de los fabricantes.

En general, las FP se añadirán a un tubo que contiene medio sin suero y el reactivo de transfección se añadirá a un segundo tubo que también contiene medio sin suero usando volúmenes y concentraciones sugeridas por el fabricante. Los contenidos de cada tubo se mezclarán, combinarán e incubarán durante un periodo de tiempo todo como indican los protocolos de los fabricantes. Después, se aplicarán las FP a la preparación de células y las células se ensayarán después de varios periodos de incubación.

El protocolo para exponer las FP a las células en las Figura 32 es como sigue:

Las células se sembraron en placas un día antes de la exposición a las FP, a aproximadamente 1×10^5 células/ml y a 0,5 ml por pocillo de una placa de 24 pocillos. Las células usadas fueron células transfectadas con y seleccionadas con fármaco para una construcción de expresión que codifica r-vav.

- 5 Los reactivos de FP se prepararon incubando 0,625 a 2,5 ul de una preparación madre 20 uM de FP en 50 ul a 200 ul de medio sin suero que contiene de 1 a 4 mM $MgCl_2$ en un tubo y 0,625 a 2,5 ul de Tfx50 (Promega) en un volumen igual de medio sin suero que tiene la misma concentración de $MgCl_2$ en un segundo tubo. Los contenidos de cada tubo se mezclarán y se combinarán e incubarán durante 15 a 45 minutos a temperatura ambiente.

- 10 Las células se lavaron una o más veces con medio sin suero suplementado con la misma concentración de $MgCl_2$ usada anteriormente. La preparación de FP se aplicará a las células y se puede añadir opcionalmente medio sin suero extra más $MgCl_2$ dependiendo del volumen de la preparación de FP añadido de manera que las células estén cubiertas.

- 15 Las células se incubaron durante 3 a 5 horas en un incubador de cultivo tisular y se ensayaron y opcionalmente puede añadirse DMSO antes de la observación. El DMSO puede ayudar en la administración de la sonda de señalización a las células o en el plegamiento o presentación de la diana que se va a detectar. También pueden usarse otros disolventes o químicos con el objetivo de incrementar la administración de la sonda de señalización en la célula o el citoplasma de la célula o con el objetivo de mejorar la detección de la diana por la sonda de señalización. Los disolventes y químicos pueden ensayarse para su conveniencia o idoneidad determinando si resultan en una eficacia incrementada de la administración o señal a ruido incrementada para la detección de la diana en células que expresan la diana comparado con células que no expresan la diana, en el que ambos tipos de células se exponen a la sonda de señalización.

- 20 El protocolo para la Figura 34 se realizó esencialmente como se describe para la Figura 32 excepto que FP16 se añadió a las células que se habían transfectado o no transfectado y se seleccionaron con fármaco para una construcción de expresión que codifica un ARN que comprende la secuencia diana 6CA4.

El protocolo para exponer las FP a las células en la Figura 33 (A,B,C) es como sigue:

- 25 Aproximadamente 50 a 100 ul de células sembradas en placas a aproximadamente $2,5 \times 10^5$ células/ml en PBS suplementado con 4 mM $MgCl_2$ (PBS+4) se sembraron en placas en pocillos de una placa de 96 pocillos y las células se expusieron a las FP después de que hubiera pasado tiempo suficiente para que se asentaran en la superficie de la placa pero antes de que se hubieran extendido.

- 30 Los reactivos de FP se prepararon mezclando 2,5 ul de una preparación madre 20 uM de FP en 100 ul de PBS+4 en un tubo y se añadieron 7 ul de Lipofectamina (InVitrogen) a 100 ul de PBS+4 en un segundo tubo. Los contenidos de cada tubo se mezclaron y las disoluciones se incubaron durante 30 minutos a temperatura ambiente antes de combinarlas, mezclarlas e incubarlas durante 15 minutos adicionales a temperatura ambiente. 50 ul de la preparación de FP se añadieron a cada pocillo. Un lavado de las células con medio sin suero u otro tampón es opcional. Las células se ensayaron después de incubación en un incubador de cultivo tisular durante 2 a 3 horas y se ensayaron.

- 35 El protocolo para exponer las FP a las células en la Figura 33 (D,E) fue esencialmente el mismo que para la Figura 32 excepto que se añadieron 4 ul de Reactivo Plus (In Vitrogen) al tubo que contiene la FP.

Ejemplo 5 Análisis y Aislamiento de las Células mediante FACS

Las células se expusieron a las FP como se ha descrito anteriormente y se procesaron para despegarlas de las superficies si estaban adheridas y separarlas entre sí (por ejemplo usando tripsina, aunque pueden usarse otros procedimientos enzimáticos o no enzimáticos) y se aplicarán a FACS.

- 40 Después de la exposición a las FP, la disolución que contiene FP se eliminará de las células y se aplicará directamente tripsina a las células. Puede usarse el lavado de las células antes de esta etapa usando varios tampones o reactivos, incluyendo reactivos designados para eliminar cualquier reactivo de FP u otro reactivo que puede haberse asociado con las superficies celulares durante la exposición de las células a las FP. Las células se resuspenderán en tampón (por ejemplo medio que contiene suero y magnesio) y las células se dispersarán adicionalmente por ejemplo por pipeteo.

- 45 Las células se analizarán por FACS según métodos FACS estándar. Para el análisis de las células, comparando las intensidades de fluorescencia de las células control y las células que potencialmente expresan las secuencias diana para la presencia de estas secuencias usando las FP, se puede determinar la fluorescencia de fondo de las células que han experimentado este procedimiento para determinar si las células que potencialmente expresan las secuencias diana muestran algún cambio en la fluorescencia. Las células pueden aislarse bien individualmente o en lote tomando como base esta información. Por ejemplo, la tecnología existente permite el aislamiento directo de las células deseadas en pocillos únicos de placas de 96 pocillos o pueden obtenerse múltiples células deseadas.
- 50

Mediante el incremento o disminución del umbral de intensidad de fluorescencia en FACS, pueden aislarse las células que presentan diferentes niveles de señal de fluorescencia. Se puede desear ajustar el umbral muy alto para tener una gran seguridad de que las células son de hecho positivas o se puede ajustar un umbral más bajo si la pureza de las células positivas en la población que se obtiene no es crítica, por ejemplo en el caso en el que se quiere simplemente enriquecer las células aisladas con células positivas.

Ejemplo 6 Determinación de la Estabilidad de la Expresión

Se pueden monitorizar los niveles de expresión de uno o más genes en las células durante el tiempo. Por ejemplo, dada una línea celular obtenida de células aisladas como positivas para la expresión de una secuencia usando una FP diseñada para reconocer la secuencia, se pueden exponer las células y las células control a la FP y determinar la proporción de las intensidades de fluorescencia de los dos tipos de células para la señal emitida por la FP. Este procedimiento puede repetirse en el tiempo y los cambios en la proporción reflejarán los cambios en los niveles relativos de expresión para la secuencia que se está detectando por la FP. Este procedimiento puede realizarse usando múltiples FP.

Ejemplo 7 Diseño de la Secuencia Etiqueta

La secuencia para etiqueta1 se usó para generar dos secuencias diferentes adicionales de manera que cada una de las secuencias pudiera ser reconocida por una única sonda de señalización, con la intención de generar secuencias que tienen la misma o una estructura predicha similar particularmente respecto a las regiones de la estructura más directamente implicadas en la unión de la sonda de señalización de manera que la secuencia diana de la sonda de señalización esté presentada de manera similar para la unión. Para cada una de las dos secuencias etiqueta diferentes (etiqueta2 y 3) que se generaron, la secuencia diana en la secuencia original para etiqueta1 se cambió en primer lugar y se hicieron cambios compensatorios en algunas bases adicionales que se había predicho que interaccionan con las secuencias cambiadas en un intento de conservar las interacciones en estas mismas posiciones (Figura 43 y 44). Se obtuvo la estructura predicha de la nueva secuencia y si no era muy parecida a la estructura predicha de etiqueta 1, la estructura se usó para predecir qué cambios adicionales serían necesarios. Esto fue un proceso iterativo realizado hasta que se obtuvieron las nuevas secuencias que tienen estructuras predichas similares a la de etiqueta1. Los cambios realizados a estas secuencias incluyen sustituciones, deleciones y adiciones de bases.

Ejemplo 8 Exposición de Sondas de Señalización a Temperaturas Elevadas

Se predice que FP17 forma una región mutuamente complementaria de 7 pares de bases adyacente a la pareja interactiva. Se requirieron temperaturas elevadas para que FP17 proporcionara una señal de fluorescencia fuerte cuando se incubó en presencia de oligo diana comparado con oligo control. Por ejemplo, poniéndolas brevemente en agua caliente de 90 a 95°C. Los tubos usados aquí contenían 16 ul en total que consistían en 5 ul de una preparación madre 20 uM de FP, 1,5 ul 25 mM MgCl₂, 8 ul 20 uM oligo y 1,5 ul de agua, que tiene una concentración final de magnesio de aproximadamente 2,34 mM. El oligo diana usado fue TO-FP1 y el oligo control usado fue TO-FP18 como se ha descrito anteriormente.

Ejemplo 9 Sondas de Señalización Químicamente Modificadas

Se sintetizaron quince sondas diferentes químicamente modificadas basadas en la secuencia de la sonda FP1. Todas estas sondas tienen la misma secuencia y están dirigidas a la misma secuencia diana. Las sondas se introdujeron en células 293T. Estas células expresan una secuencia etiqueta que incluye la secuencia diana de FP1. Se usó FACS para analizar la fluorescencia de estas células. Las quince sondas diferentes se describen a continuación. Los resultados de los análisis FACS se muestran en las Figs. 46-60. Los resultados muestran que todas las quince sondas son capaces de detectar células que expresan la secuencia diana. Todas las sondas eran idénticas respecto a concentración, método de administración, fluoróforo, apantallador y secuencia excepto por la modificación química.

ES 2 379 634 T3

Sonda	Modificación	Secuencia	Comentarios
Mcon 1	2-Amino-dA y 5-Metil-dC	GCCAGTCCCAGTTCTGTGCCTTAAGAA <u>C</u> CTCGC	C= 5-Metil dC <u>A</u> = 2-Amino dA
Mcon 2	Oligonucleótidos unidos 2'-5'	GCCAGTCCCAGTTCTGTGCCTTAAGAACCTCGC	Negrita= unido 2'-5'
Mcon 3	Citosina Arabinósido (Ara-C)	GCCAGT CCC AGTT CT GTGCCTTAAGAACCTCGC	Negrita= Ara-C
Mcon 4	Espaciador Fosforamidita 9	GCCAG S TCCCAGTTCTGTGCCTTAAGAAC S CTCGC	S = Espaciador Fosforamidita 9
Mcon 5	2'-desoxi-2'-fluoro-ARN (2'-F-ARN)	GCCAG <u>ucccAGuuccuGuGccuuAAGAAc</u> CTCGC	Subrayado= unión fósforotioato Minúsculas= 2'-F-ARN
Mcon 6	2'-desoxi-2'-fluoro-ARN (2'-F-ARN)	GCCAG <u>uCCcAGuTCcTGuGCcTuAAGAAc</u> CTCGC	Subrayado= unión fósforotioato Minúsculas= 2'-F-ARN
Mcon 7	2-amino-A	GCCAGTCCCAGTTCTGTGCCTTAAGAACCTCGC	A = 2-amino-A
Mcon 8	2'-O-metil-5-Metil-C (2'-OMe-5-Me-C)	GCCAGT CCC AGTT CT GTGCCTTAAGAACCTCGC	C = 2'-OMe-5-Me-C
Mcon 9	Ácido Nucleico Bloqueado (LNA)	GCCAGTCCCAGTTCTGTGCCTTAAGAACCTCGC	Negrita= LNA
Mcon 10	Ácido Nucleico Bloqueado (LNA)	GCCAGTCCCAGTTCT G TGCCTTAAGAACCTCGC	Negrita= LNA
Mcon 11	Uniones fósforotioato	GCCAGT <u>CCCAGTTCTGTGCCTTAAGAACCTCGC</u>	Subrayado= unión fósforotioato
Mcon 12	Uniones fósforotioato	GCCAGT <u>CCCAGTTCTGTGCCTTAAGAACCTCGC</u>	Subrayado= unión fósforotioato
Mcon 13	2'-O-metil-ARN (2'-OMe-ARN)	GCCAG <i>TCCCAGTTCTGTGCCTTAAGAACCTCGC</i>	Itálica negrita= 2'-O-metil-ARN
Mcon 14	2'-O-metil-ARN (2'-OMe-ARN)	GCCAGTCCCAGTT CT GTGCCT <i>TAAGAACCTCGC</i>	Itálica negrita= 2'-O-metil-ARN
Mcon 15	Análogos de Pirimidina propinilo C-5	GCCAG <u>ucccAGuuccuGuGccuuAAGAAc</u> CTCGC	Minúsculas= análogo C5-propino

Nota: todas las sondas tienen un 5' Cy5 y un 3' BHQ-3

REIVINDICACIONES

1. Un método para aislar una pluralidad de células, en el que al menos un subconjunto de las células expresa un ARN que no es expresado por otro subconjunto de las células, que comprende las etapas de:
 - 5 introducir en las células una pluralidad de ADN que codifican una pluralidad de ARN diferentes, en el que cada ADN codifica además una secuencia etiqueta de ácido nucleico, y en el que al menos un subconjunto de la pluralidad de los ADN codifica la misma secuencia etiqueta de ácido nucleico;
 - exponer dichas células a una misma sonda de señalización que produce una señal detectable después de hibridar con dicha misma secuencia etiqueta de ácido nucleico; y
 - aislar dichas células que producen la señal.
- 10 2. El método de la reivindicación 1, en el que la pluralidad de ARN diferentes forma una biblioteca de expresión.
3. El método de la reivindicación 2, que comprende además crecer separadamente células aisladas individualmente para generar una pluralidad de líneas celulares separadas.
4. El método de la reivindicación 2, que comprende además combinar las células aisladas.
5. El método de la reivindicación 4, que comprende además crecer las células combinadas.
- 15 6. El método de una cualquiera de las reivindicaciones 1 a 5, en el que la pluralidad de ARN diferentes, o proteínas codificadas por la pluralidad de ARN diferentes, se seleccionan del grupo que consiste en ARN o proteínas en la misma ruta biológica o relacionada, ARN o proteínas que actúan aguas arriba o abajo entre sí, ARN o proteínas que tienen una función moduladora, activadora o represora entre sí, ARN o proteínas que son dependientes entre sí para función o actividad, ARN o proteínas que son componentes del mismo complejo y proteínas de la misma familia de proteínas.
- 20 7. El método de la reivindicación 1 para aislar células que expresan dos o más bibliotecas de expresión de ARN, que comprende las etapas de:
 - introducir en las células una pluralidad de ADN que codifican una primera biblioteca de expresión de ARN, en el que cada ADN codifica además una primera secuencia etiqueta de ácido nucleico y en el que al menos un subconjunto de la pluralidad de ADN codifica la misma primera secuencia etiqueta de ácido nucleico;
 - 25 introducir en las células una pluralidad de ADN que codifican una segunda biblioteca de expresión de ARN, en el que cada ADN codifica además una segunda secuencia etiqueta de ácido nucleico y en el que al menos un subconjunto de la pluralidad de ADN codifica la misma segunda secuencia etiqueta de ácido nucleico;
 - exponer las células a una primera sonda de señalización que produce una señal detectable después de hibridar con dicha primera secuencia etiqueta de ácido nucleico;
 - 30 exponer las células a una segunda sonda de señalización que produce una señal detectable después de hibridar con dicha segunda secuencia etiqueta de ácido nucleico; y
 - aislar las células que producen ambas señales.
8. El método de una cualquiera de las reivindicaciones 1 a 7, en el que la secuencia etiqueta de ácido nucleico comprende múltiples secuencias diana, en el que una sonda de señalización hibrida con cada secuencia diana.
- 35 9. El método de una cualquiera de las reivindicaciones 1 a 8, en el que el ADN codifica múltiples secuencia etiqueta de ácido nucleico.
10. El método de una cualquiera de las reivindicaciones 1 a 9, en el que el ADN que codifica dicha secuencia etiqueta de ácido nucleico está
 - (a) en marco con el ADN que codifica dicho ARN; o
 - 40 (b) fuera de marco con el ADN que codifica dicho ARN.
11. El método de una cualquiera de las reivindicaciones 1 a 10, en el que dicho ADN codifica un ARN antisentido, un ARNsh o un ARNsi.

12. El método de una cualquiera de las reivindicaciones 1 a 11, en el que el ADN codifica además un marcador de selección y en el que el método comprende además la etapa de seleccionar las células usando el marcador de selección después de introducir el ADN en las células pero antes de exponer dichas células a la sonda de señalización.
- 5 13. El método de una cualquiera de las reivindicaciones 1 a 12, en el que dicho ADN está unido de manera operativa a un promotor condicional.
14. El método de la reivindicación 13, en el que el ARN codificado por el ADN, o la proteína codificada por el ARN, es letal o perjudicial para la célula cuando se expresa.
15. El método de la reivindicación 13 ó 14 que comprende además la etapa de añadir a las células un compuesto que modula la expresión de dicho ARN o pluralidad de ARN antes de la etapa de exposición.
- 10 16. El método de una cualquiera de las reivindicaciones 1, 2 y 4 a 15 que comprende además la etapa de
- (a) cultivar las células aisladas; o
- (b) generar una pluralidad de líneas celulares por cultivo de las células aisladas.
17. El método de la reivindicación 1, en el que cada ADN de dicha pluralidad de ADN codifica además una segunda secuencia de ADN que codifica un ARN preseleccionado.
- 15 18. El método de la reivindicación 17, en el que dicha pluralidad de ADN codifica una pluralidad de ARN de ensayo variables.
19. Un método para identificar un compuesto que activa un promotor condicional, que comprende las etapas de:
- aislar las células por el método de la reivindicación 14 ó 15;
- añadir un compuesto de ensayo a dichas células aisladas;
- 20 ensayar para la presencia del ARN expresado bajo el control del promotor condicional; e
- identificar el compuesto de ensayo como un compuesto que activa el promotor condicional si la célula expresa el ARN bajo el control del promotor condicional.
20. Un método para identificar un ARN de ensayo que activa un promotor condicional que comprende las etapas de:
- aislar las células por el método de la reivindicación 14 ó 15;
- 25 ensayar para la presencia del ARN bajo el control del promotor condicional;
- obtener las células que expresan el ARN bajo el control del promotor condicional; e
- identificar el ARN de ensayo que activa el promotor condicional.

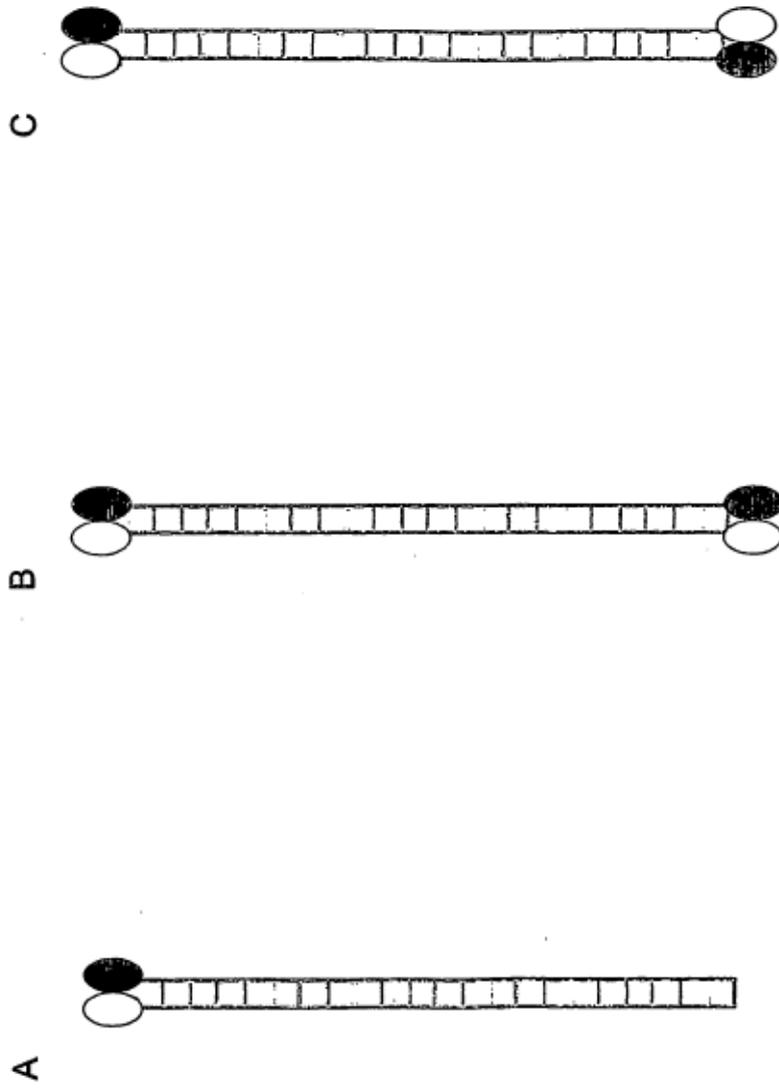


FIG. 1

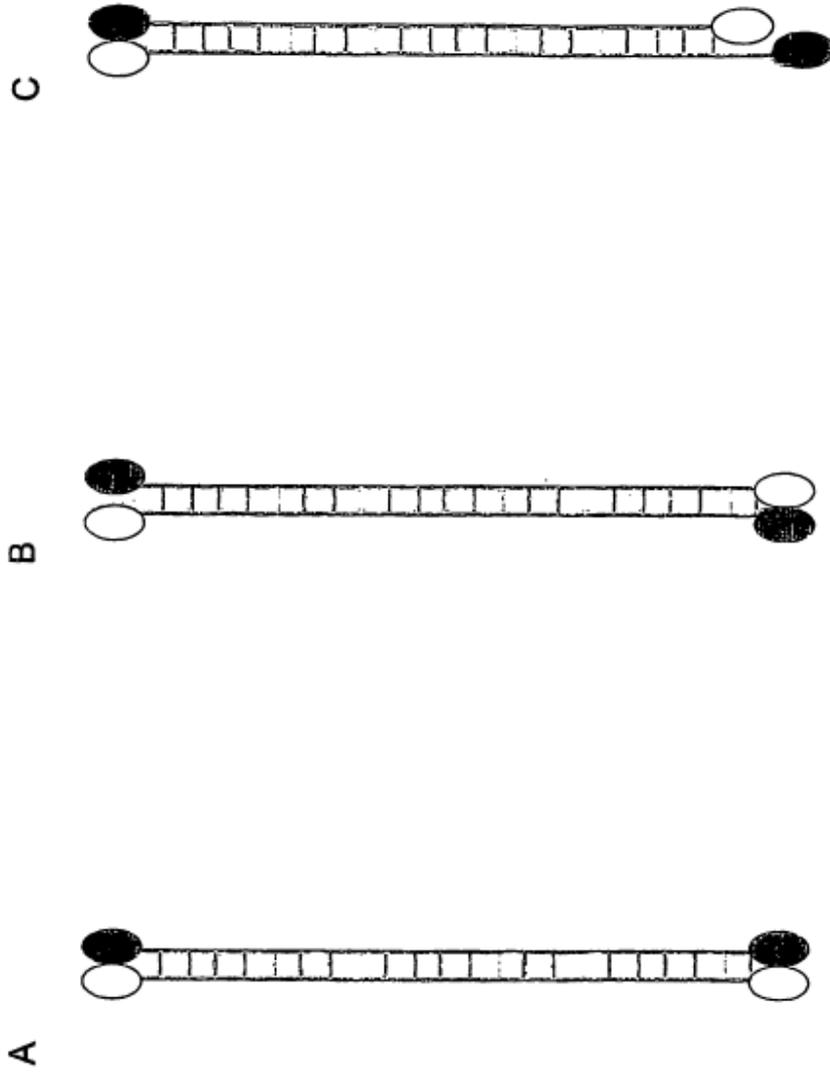


FIG. 2

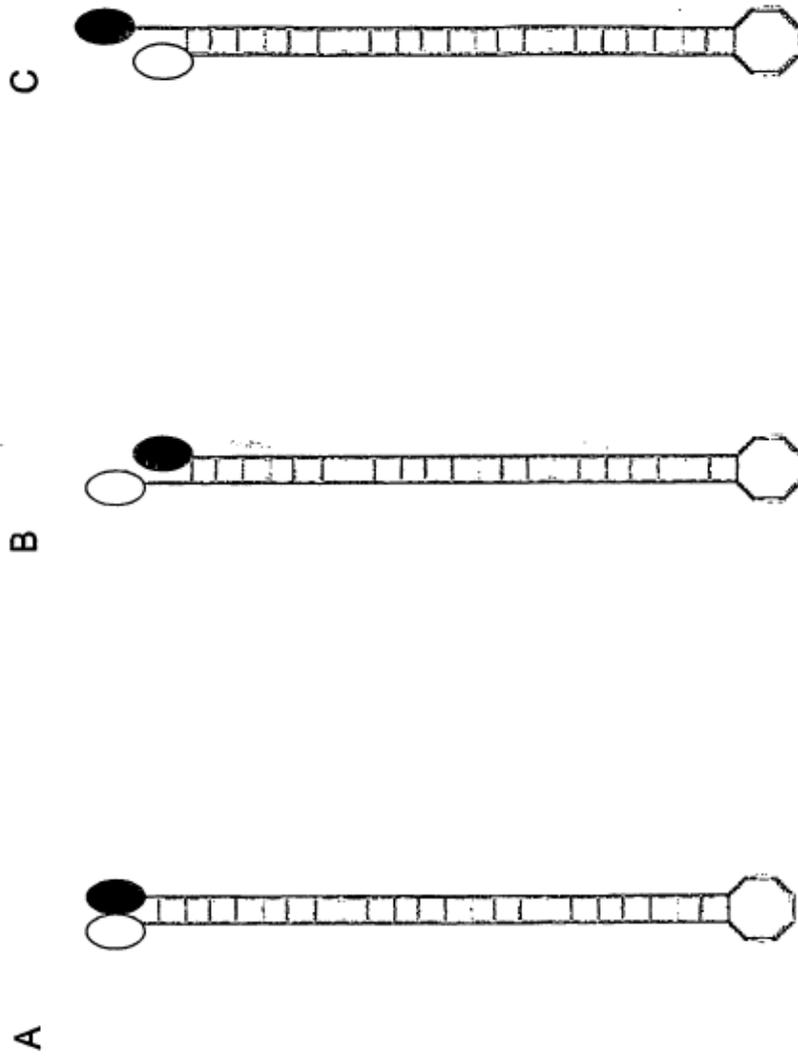


FIG. 3

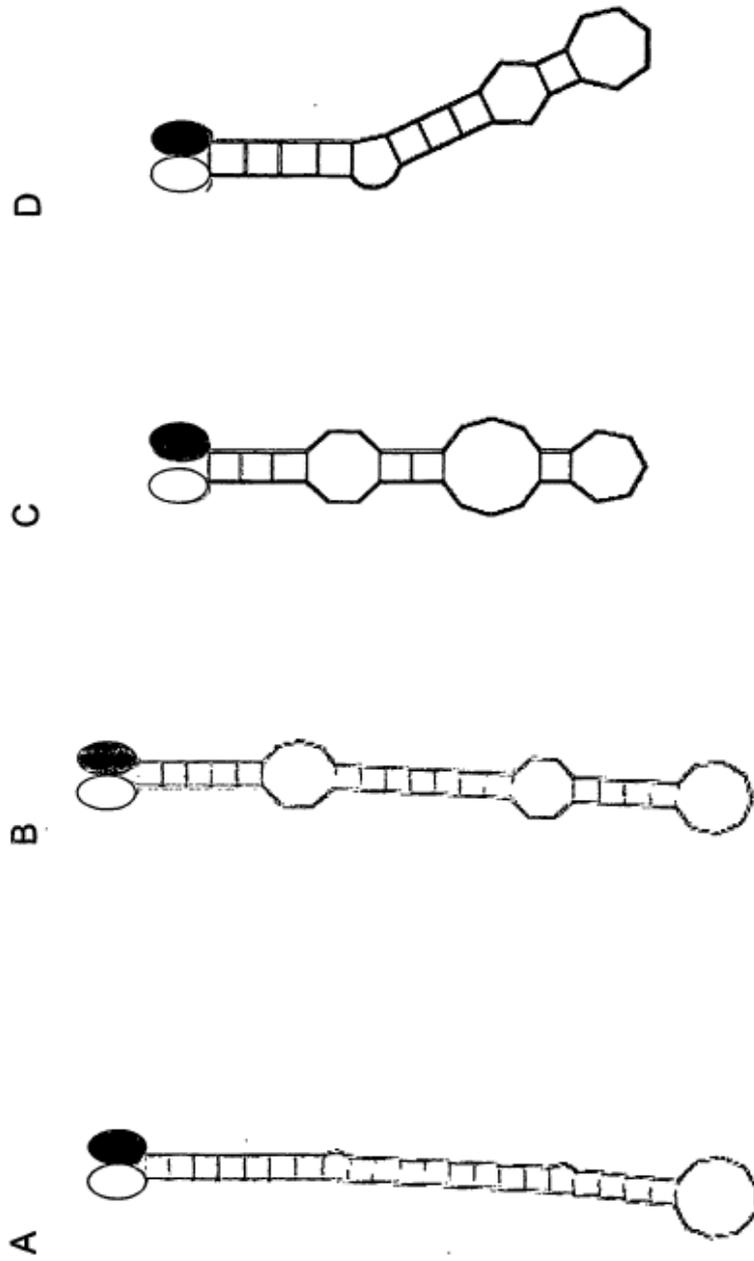


FIG. 4

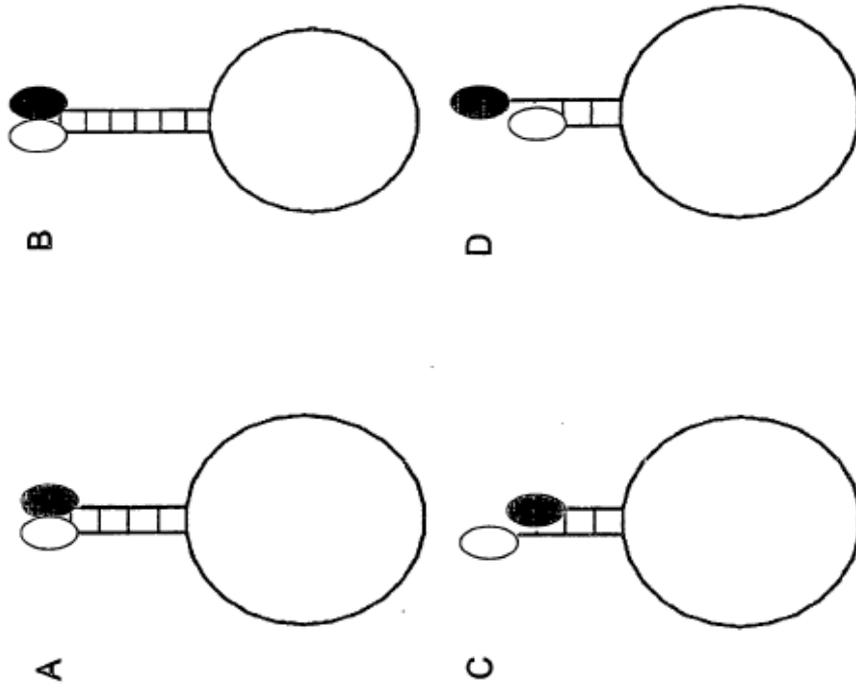


FIG. 5

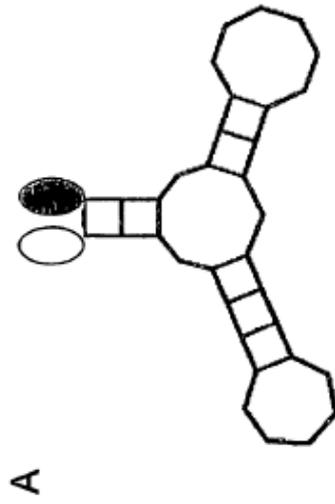
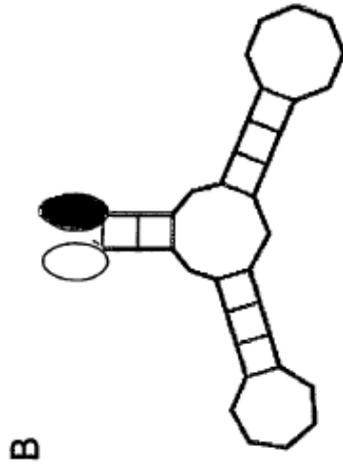


FIG. 6

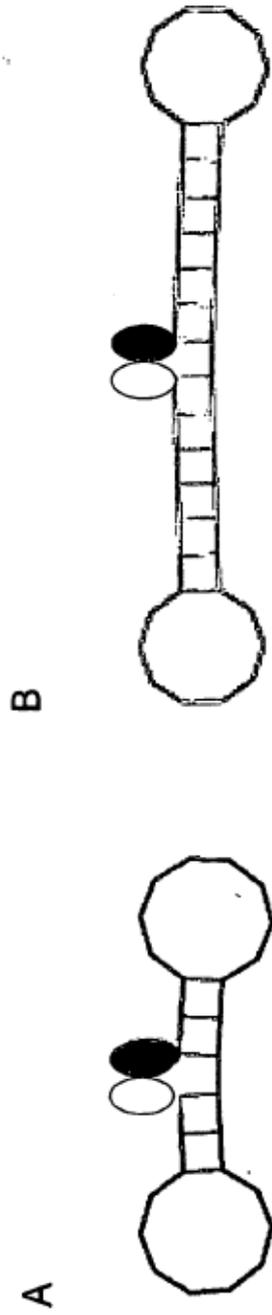


FIG. 7

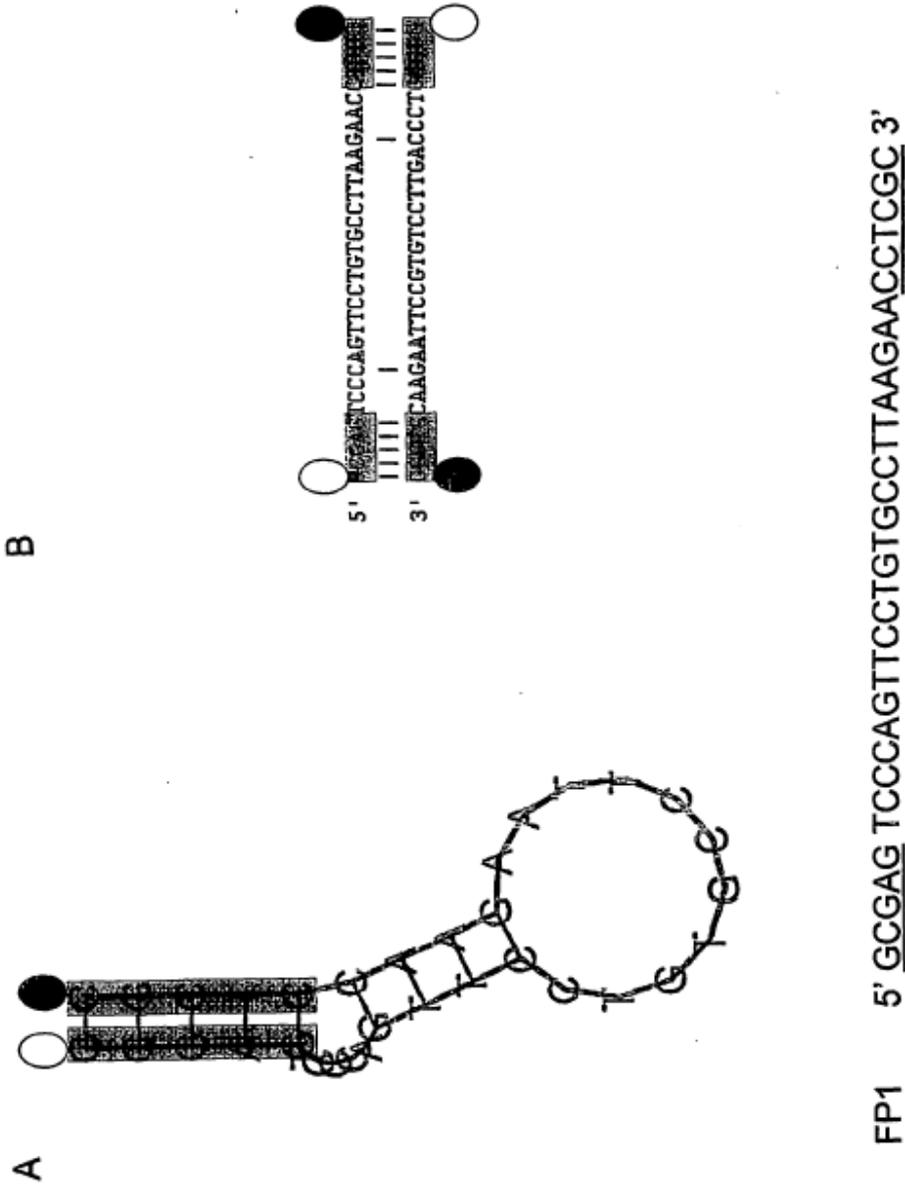
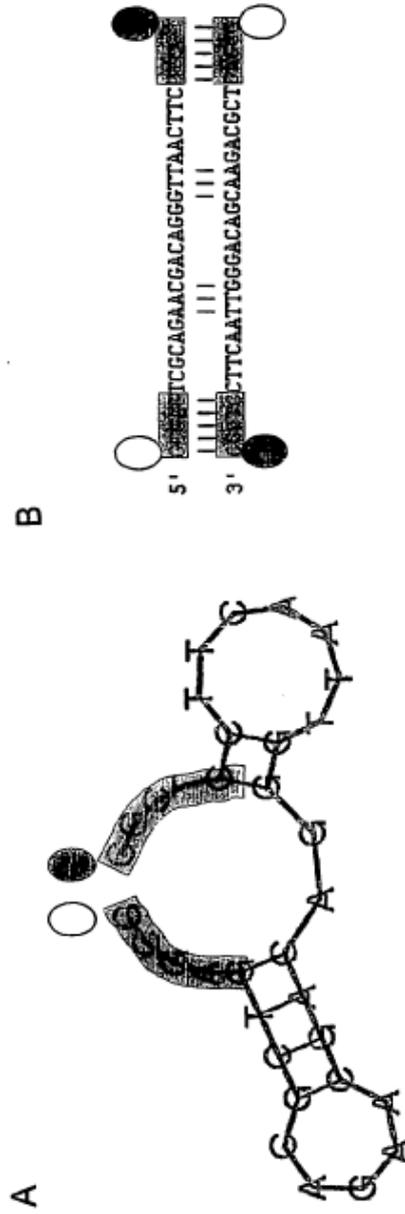
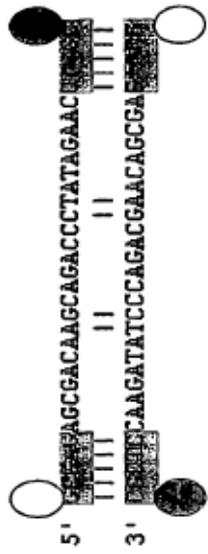


FIG. 8



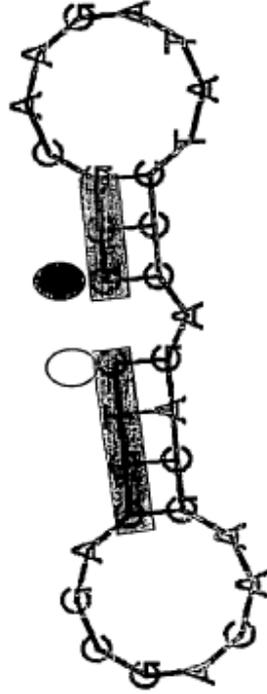
FP2 5' GCGAG TCGCAGAACGACAGGGTTAACTTCCTCGC 3'

FIG. 9



FP3 5' GCGAG AGCGACAAGCAGACCCTATAGAAC CTCGC 3'

FIG. 10



FP4 5' CTGC AGCGACAAGCAGACCCTATAGAAC GGG 3'

FIG. 11

FP5 5' ggaa TCCCAGTTCCTGCGCCTTAAGAAC ttcc 3'

FIG. 12

FP6 5' gga TCGCAGAACGACAGGGTTAACTTC tcc 3'

FIG. 13

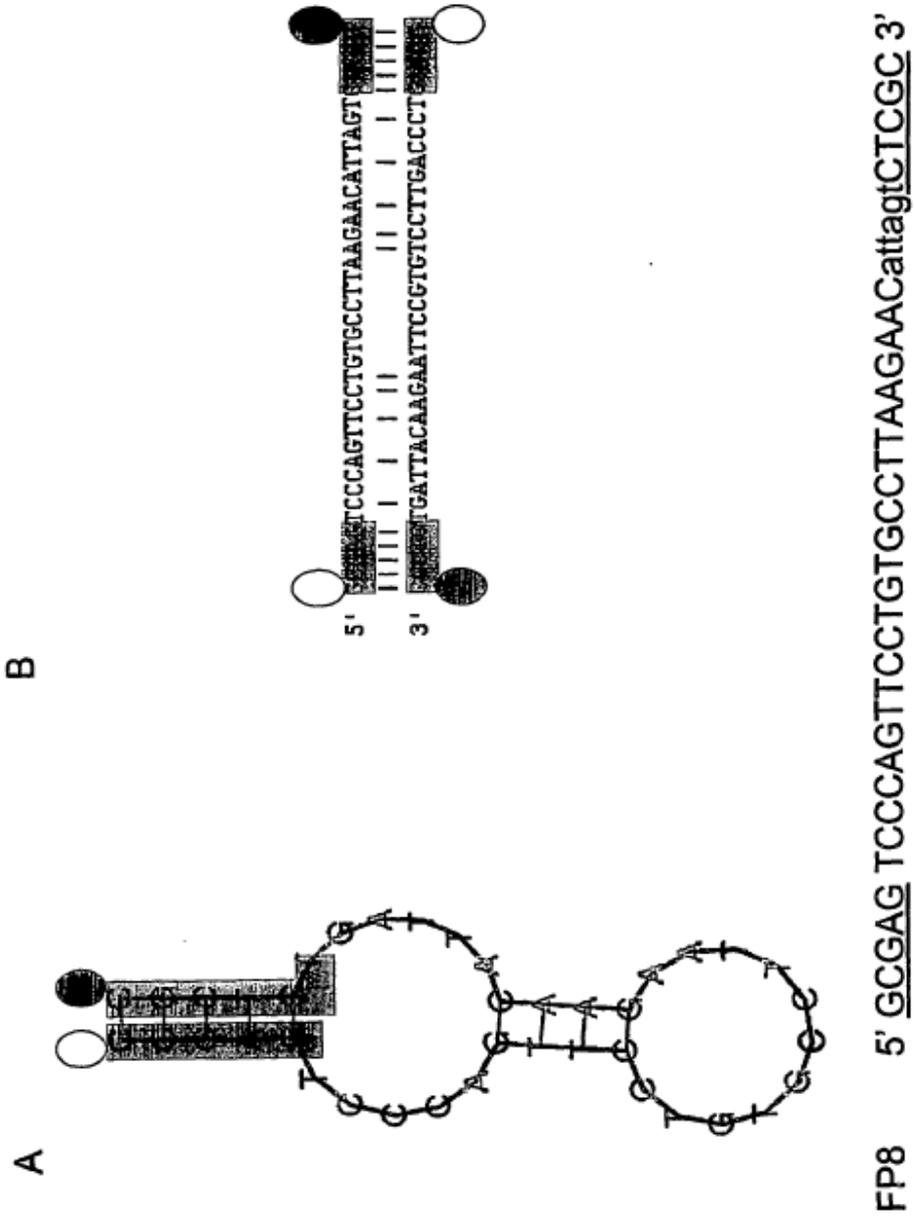


FIG. 15

FP9 5' ggaa TCCCAGTTCCTGTGCCCTTAAGAAACattagttcc 3'

FIG. 16

FP10 5' gtgat TCCCAGTTCCTGTGCCTTAAGAAC atcac 3'

FIG. 17

FP11 5' ggac TCCCAGTTCCTGTGCCCTTAAGAAC gtcc 3'

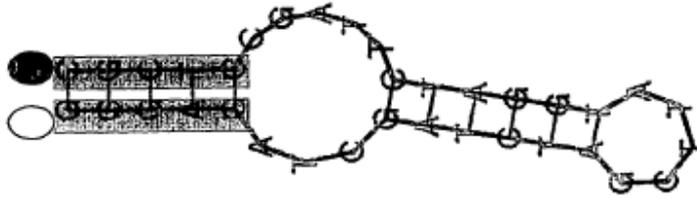
FIG. 18

FP12 5' ggg TCCCAGTTCCTGTGCCCTTAAGAAC ccc 3'

FIG. 19

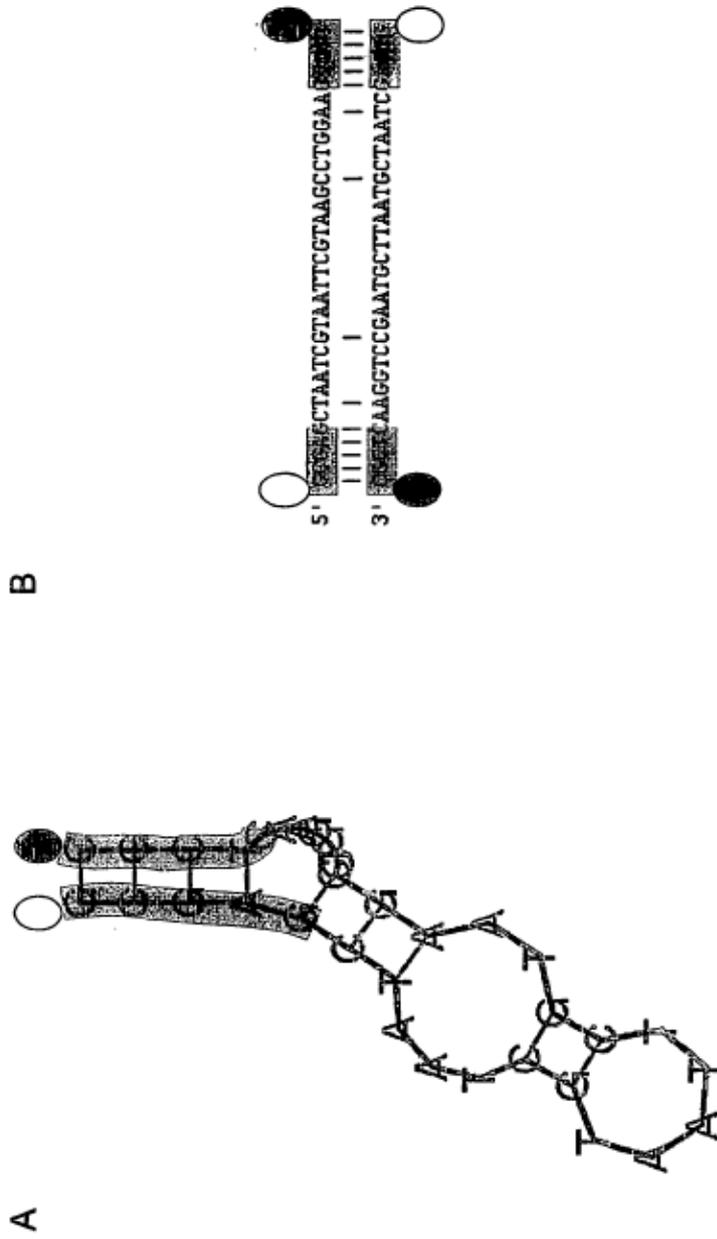
FP13 5' ggg TCCCAGTTCCTGTGCCCTTAAGAAC ggc 3'

FIG. 20



FP14 5' gcgag ATCGATCTAGCTTATGGATCTTAGC CIGGC 3'

FIG. 21



FP15 5' gagag CTAATCGTAATTCGTAAGCCTGGAA CICGC 3'

FIG. 22

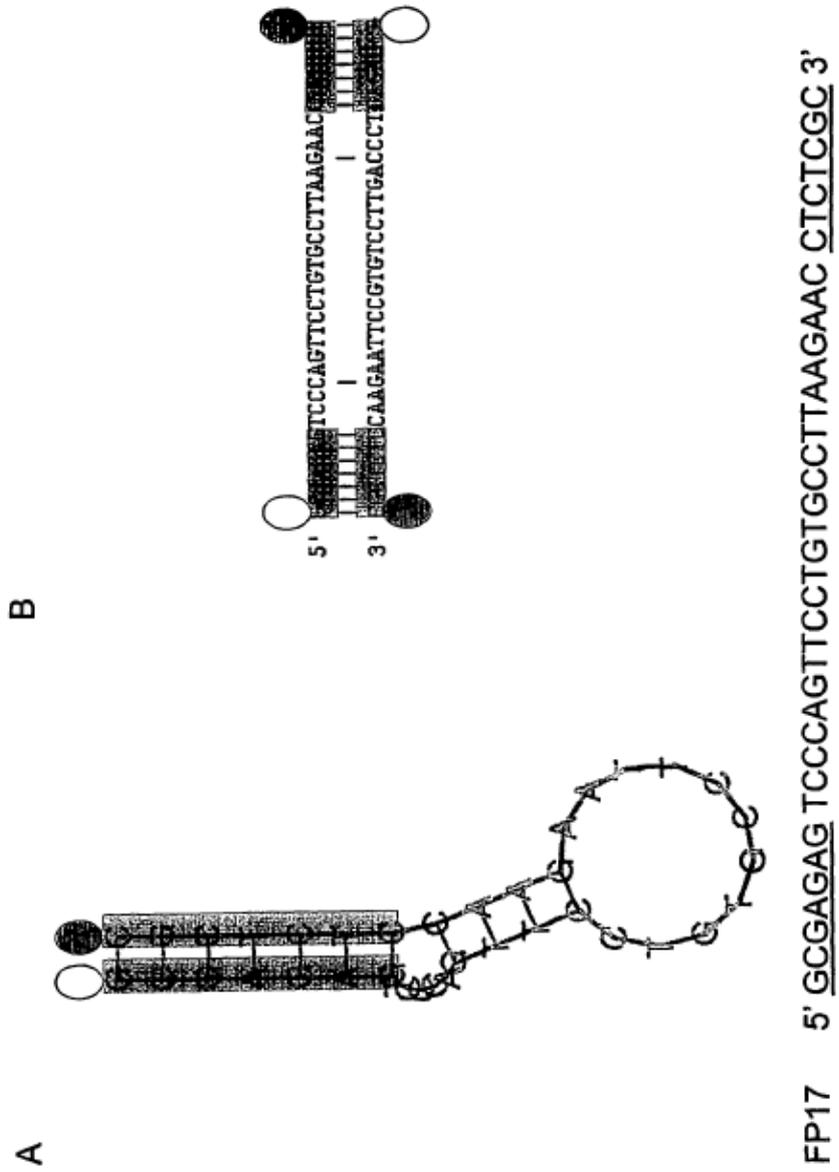


FIG. 24

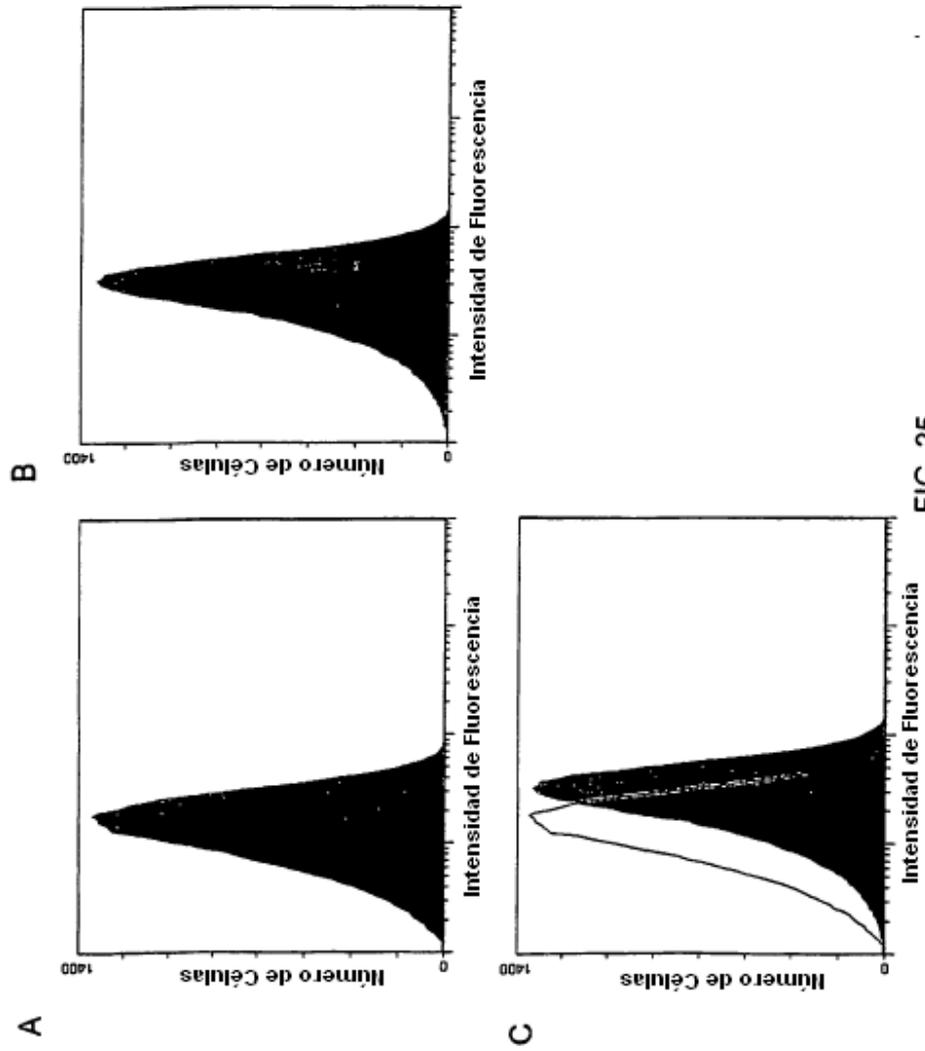


FIG. 25

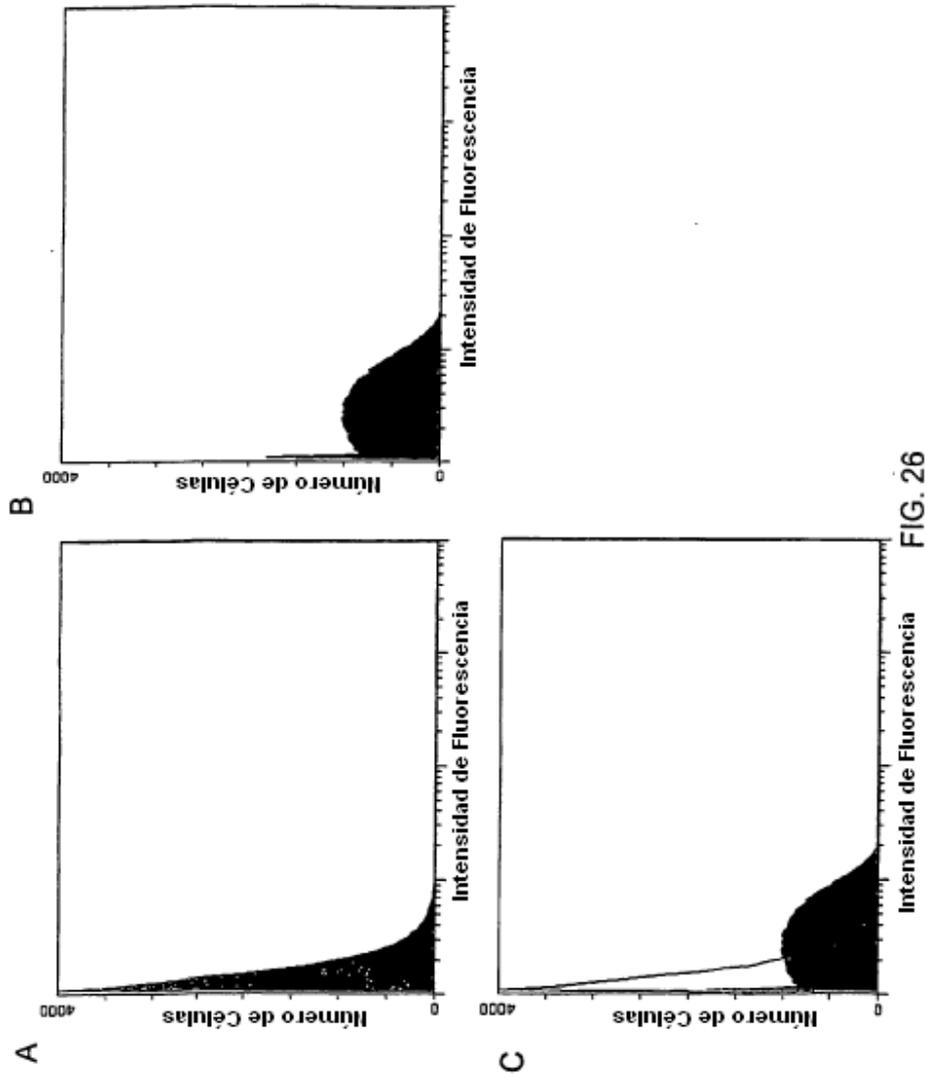


FIG. 26

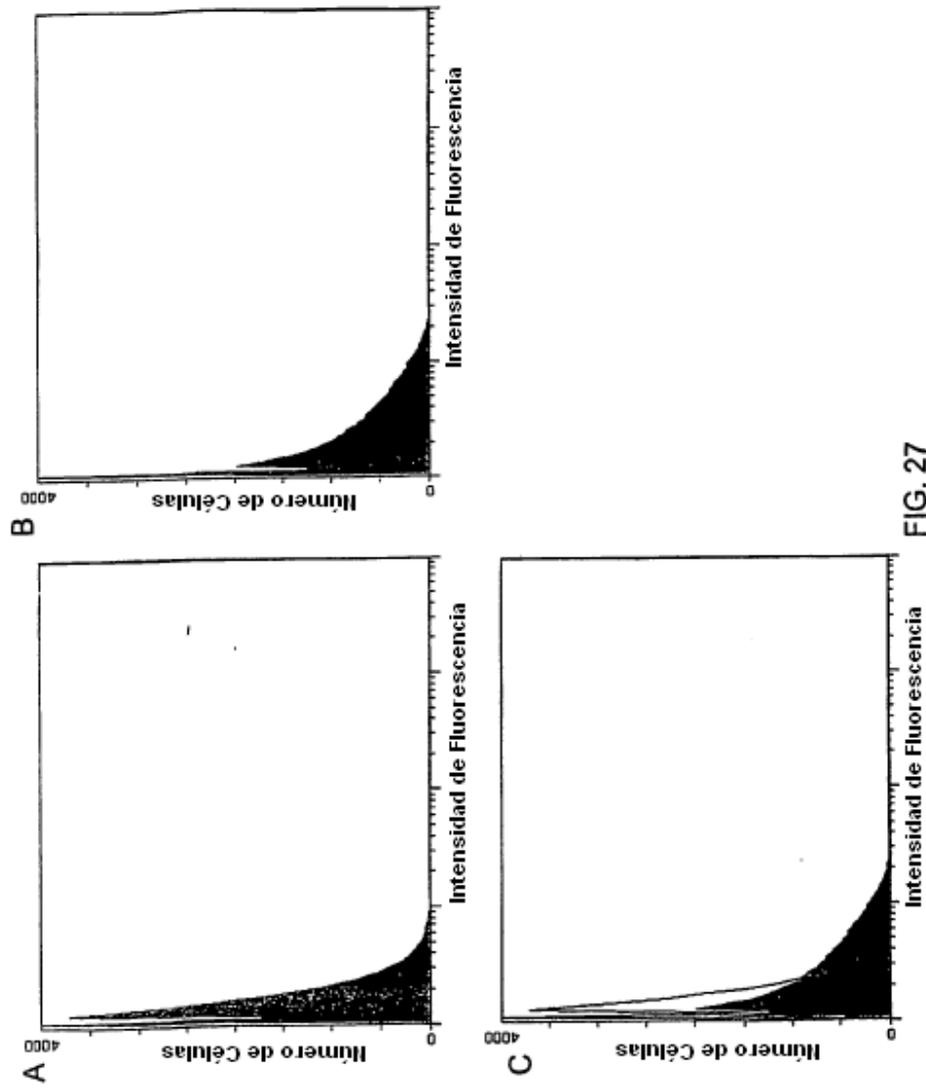


FIG. 27

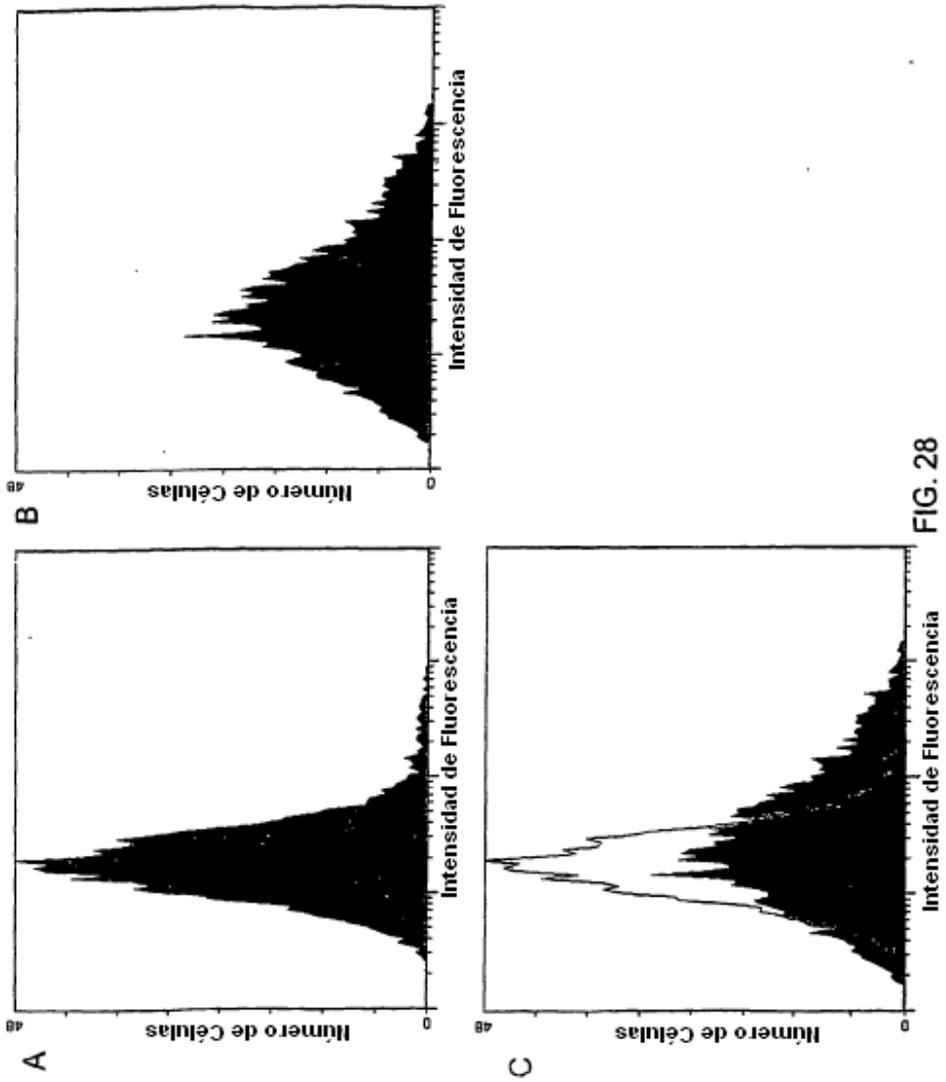


FIG. 28

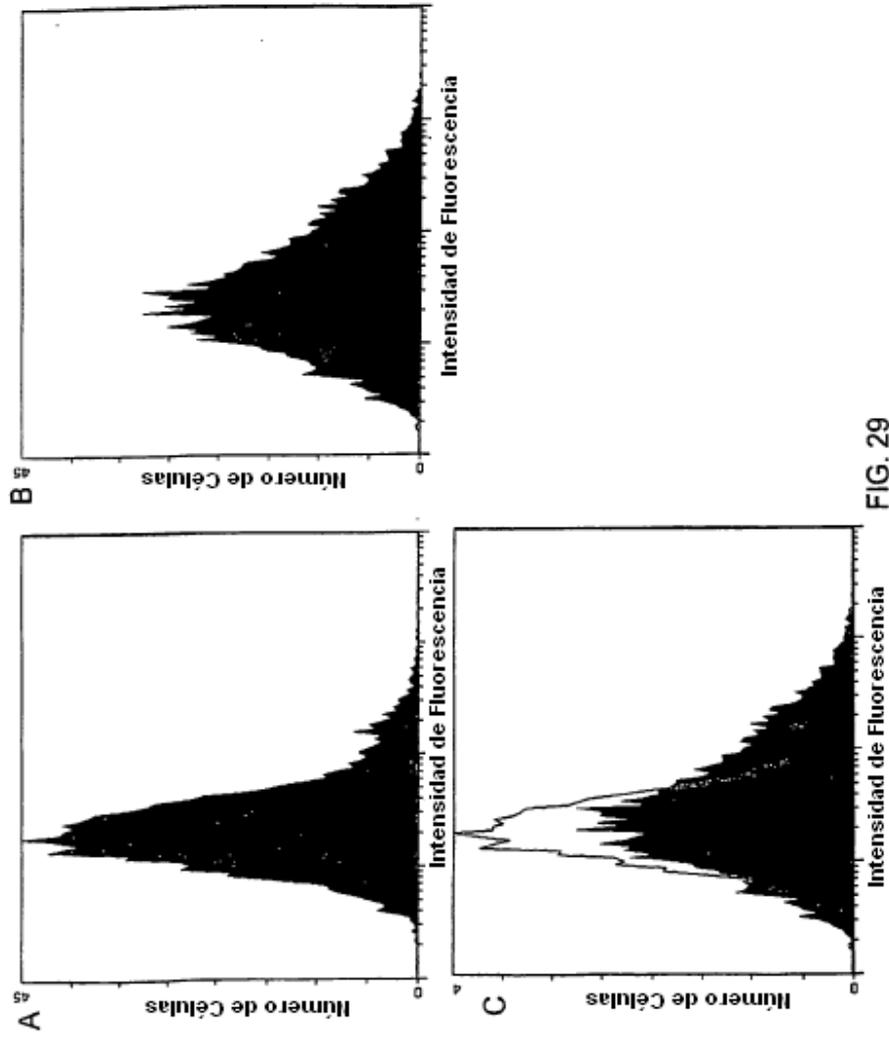


FIG. 29

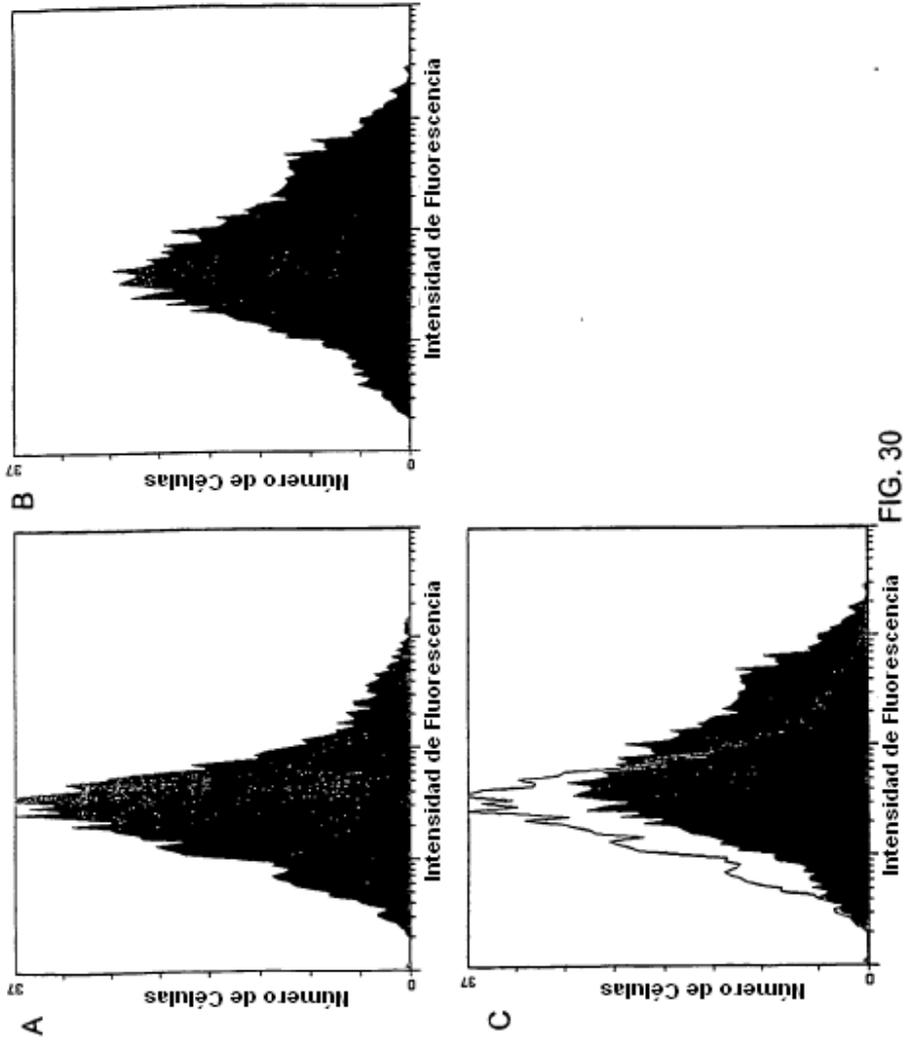


FIG. 30

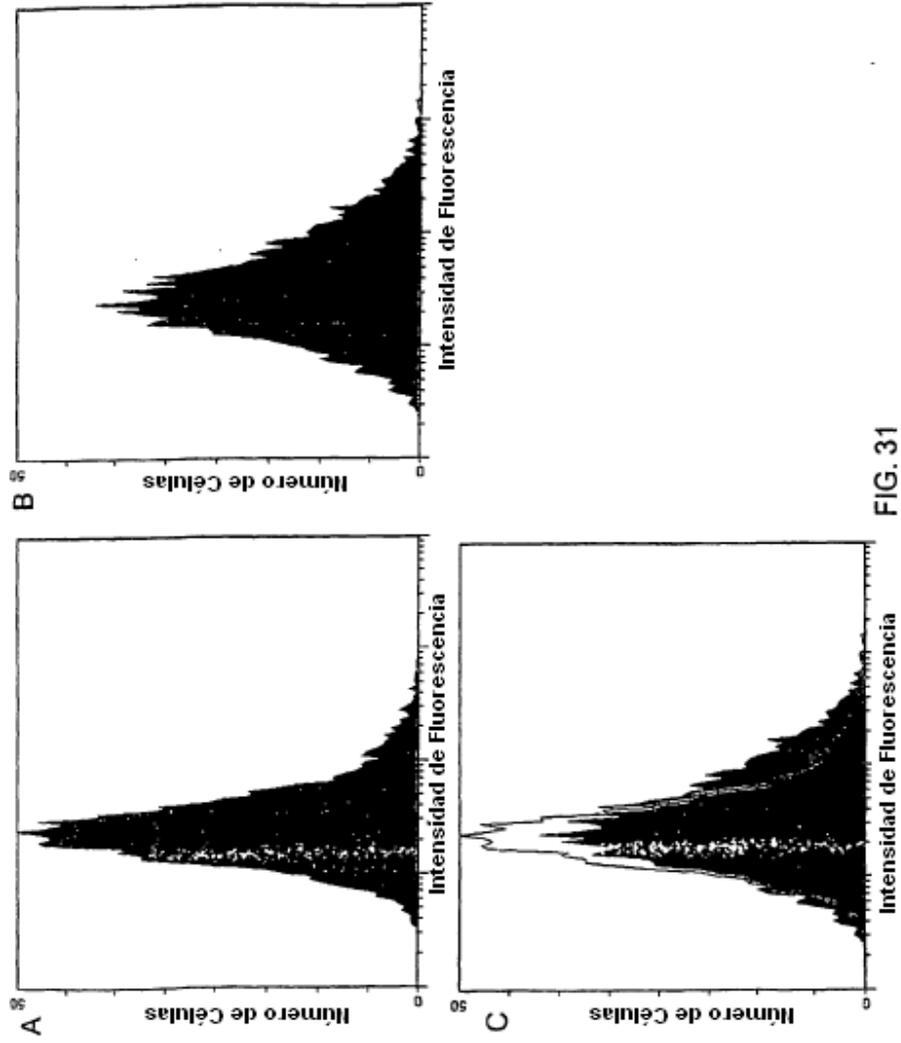


FIG. 31

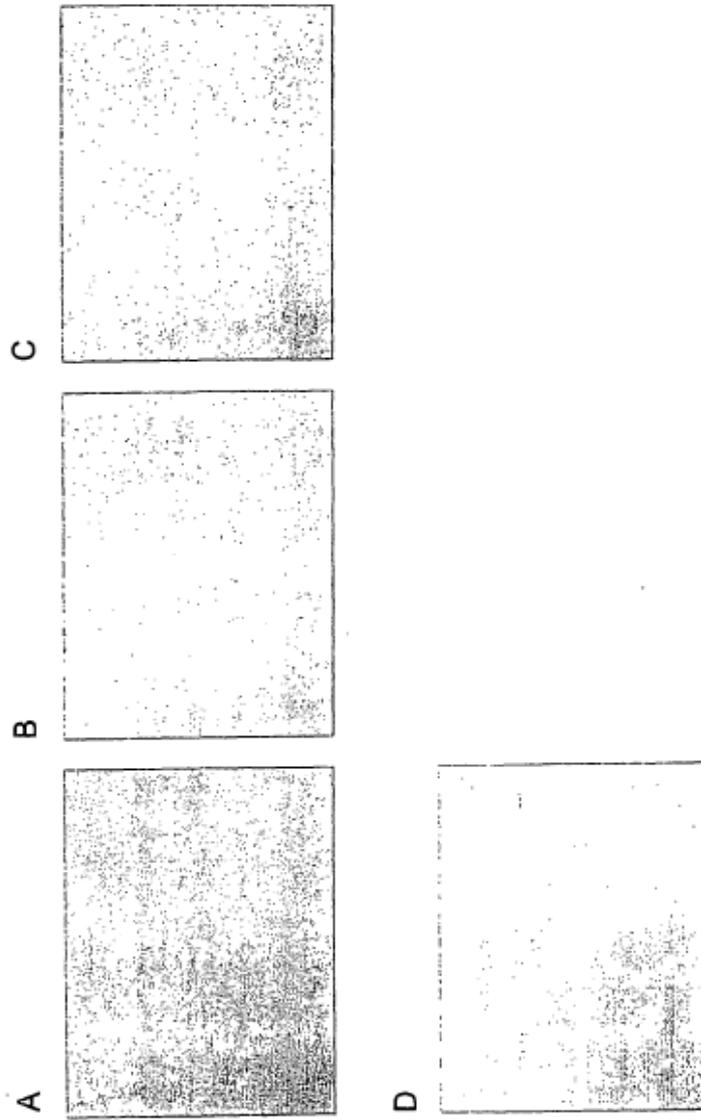


FIG. 32

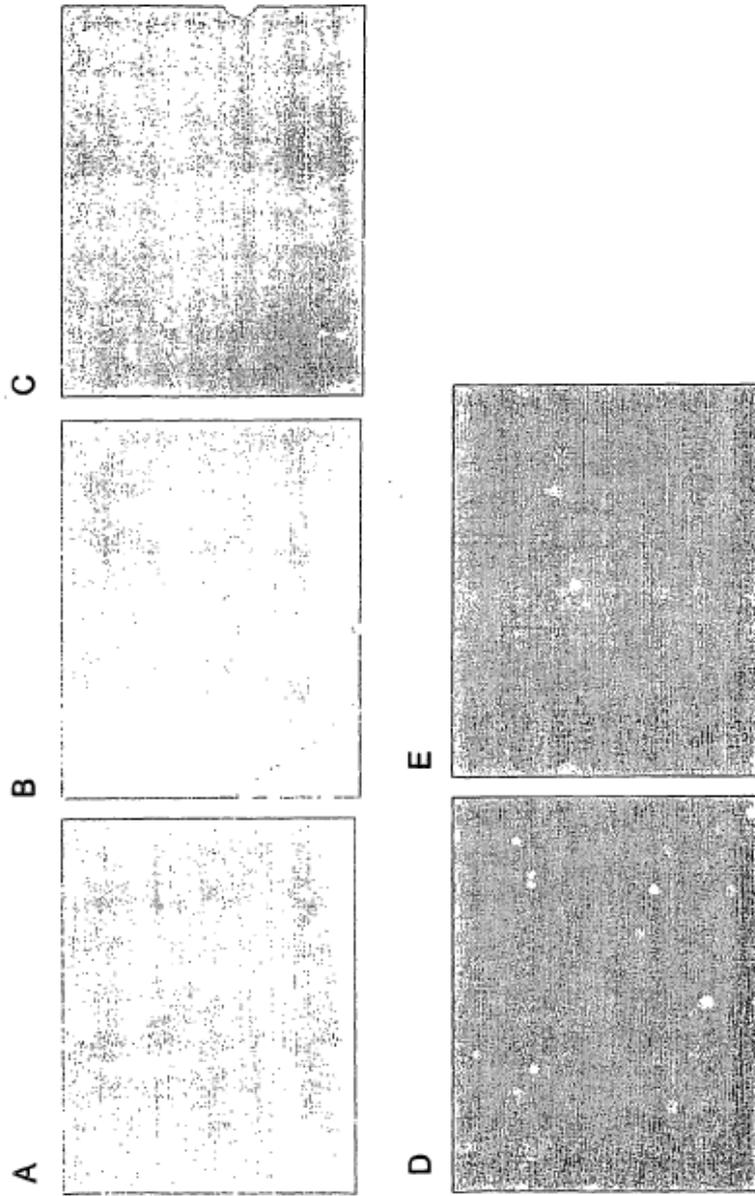


FIG. 33

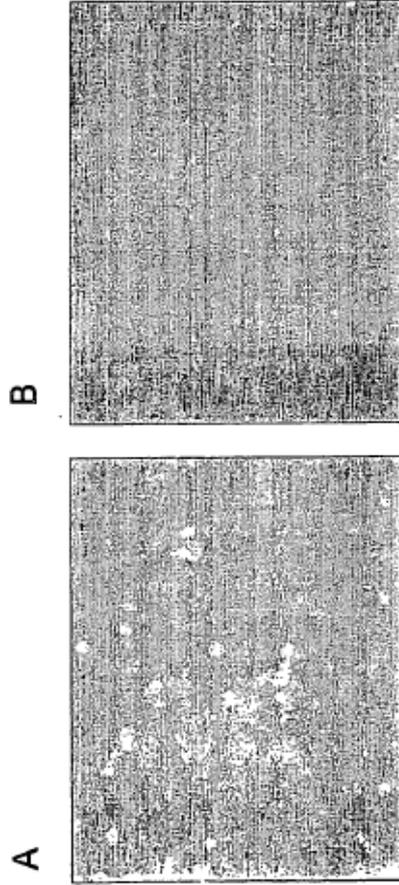


FIG. 34

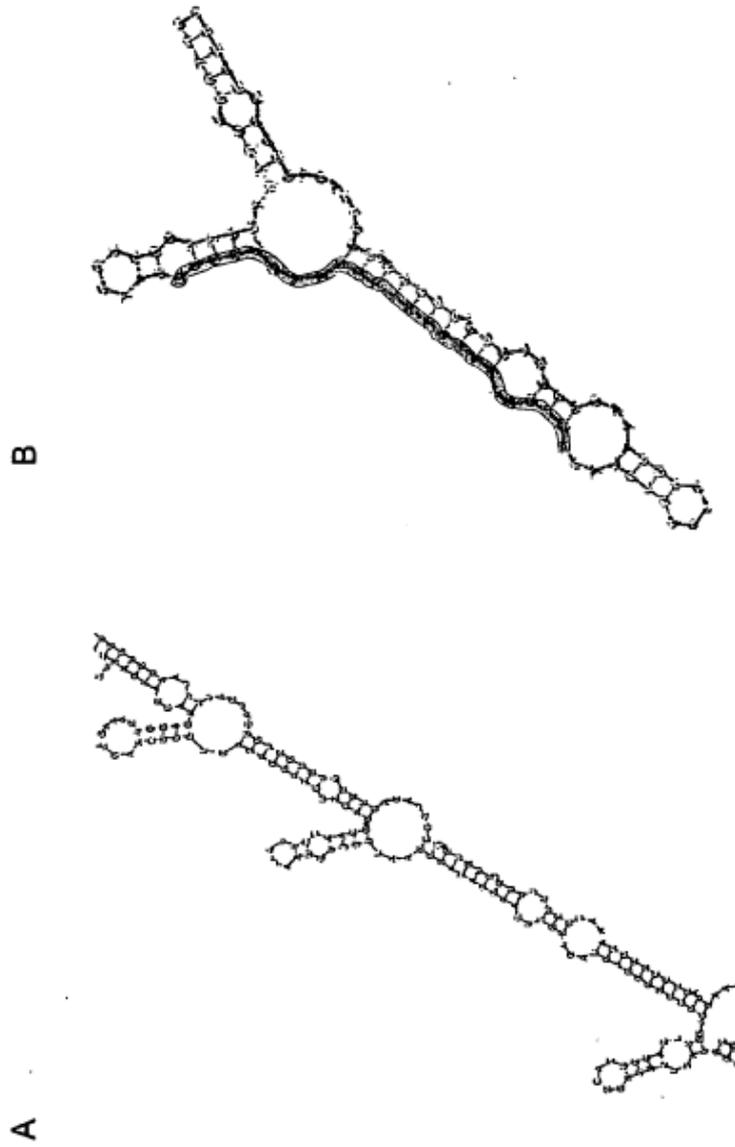


FIG. 36

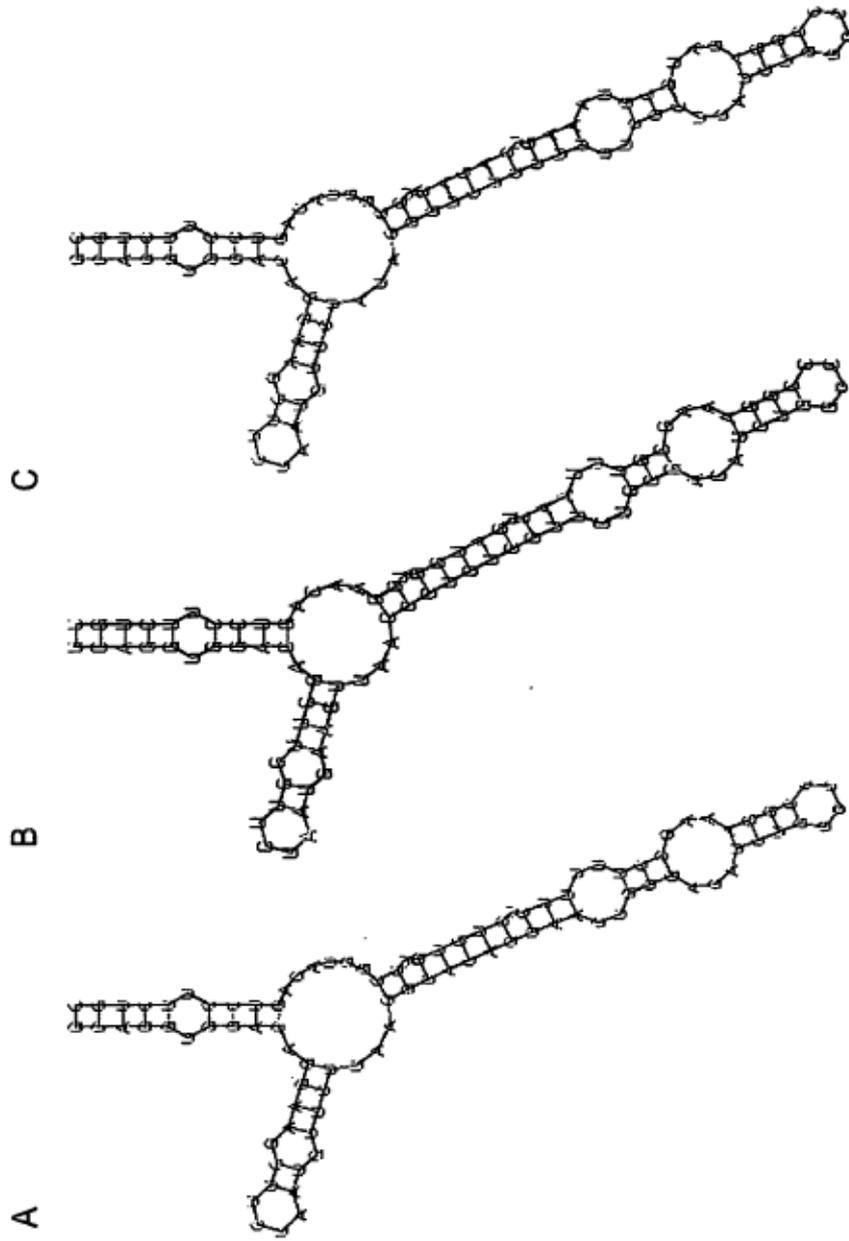


FIG. 37

5



FIG. 38

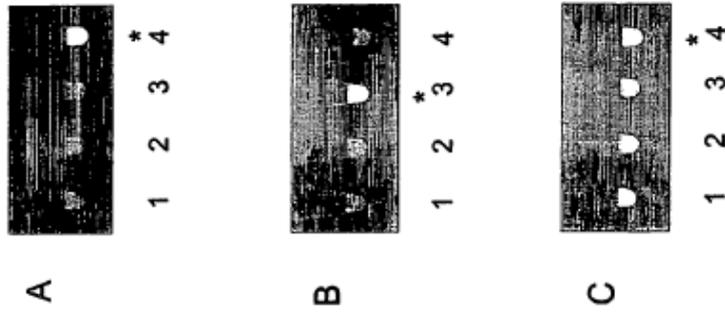


FIG. 39

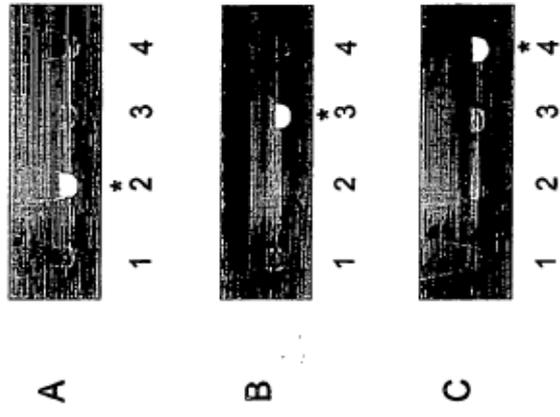
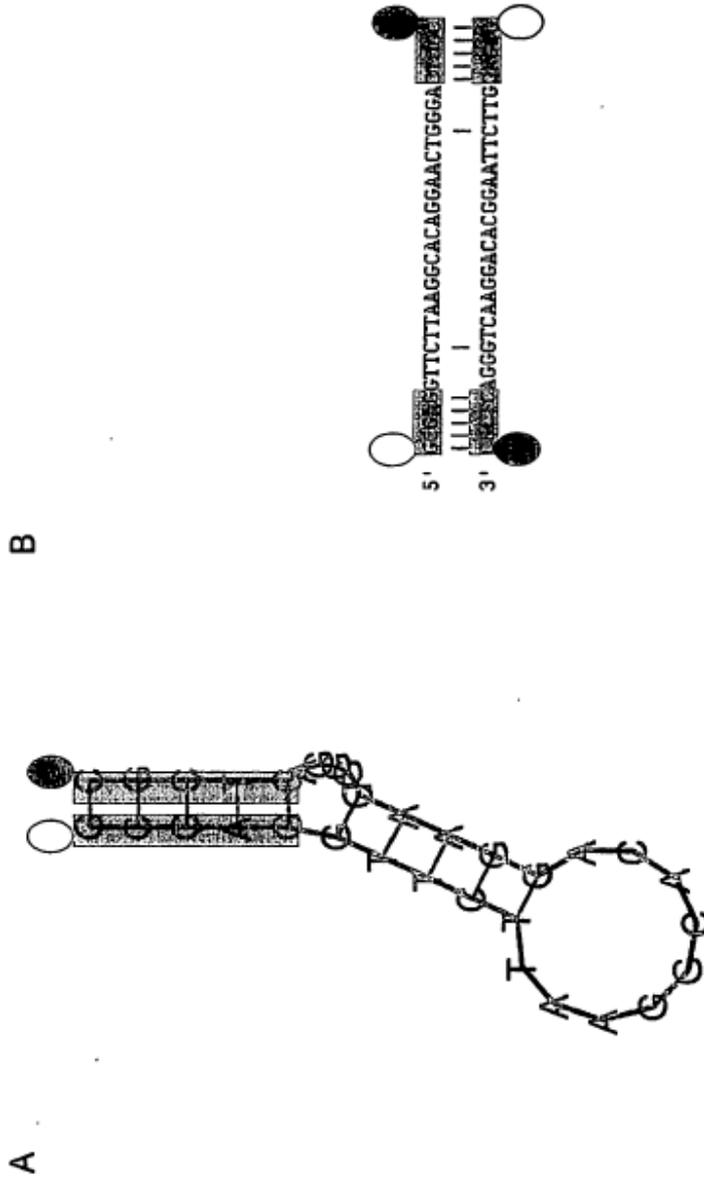


FIG. 40



FP18 5' GCGAG GTTCTTAAGGCACAGGAACTGGGA CTCGC 3'

FIG. 41

A
 GCAGGTGGACAGGAAGGTTCTAATGTTCTTAAGGCACA
 GGAAGTGGGACATCTGGGCCCGGAAAGCCTTTTCTCT
 GTGATCCGGTACAGTCCTTCTGC

B
 GCAGGTGGACAGCTTGGTTCTAATGAAGTTAACCCCTGT
 CGTTCTGGACATCTGGGCCCGGAAAGCGTTAACUGA
 TGGAUUGGAACAGTCCTTCTGC

C
 GCAGGTGGACAGGAAGGTTCTAATGTTCTATAGGGTCT
 GCTTGTGCTCATCTGGGCCCGGAGATGCGT
 AAAGTCAGACATCCGGTACAGTCCTTCTGC

D

1	GTTCTTAAGGCACAGGAACTGGGA	
2	GAAGTTAACCCCTGICGTTCTGCCGA	50% diferente de la diana 1;
3	GTTCTATAGGGTICGCTTGTGCT	66% diferente de la diana 3
		50% diferente de la diana 1

FIG. 42

A inicio
GCAGGTGGACAGGAGTTCTAATGTTCTTAAGCCACAGAACTGGGACATCTGGCCCGGAAAGCCTTTTTCTCTGTGATCCGGTACAGTCCTTC
TGC

B 1
GCAGGTGGACAGGCTTGGTTCTAATGAAGTTAACCCCTGCTGTTCTGGGACATCTGGCCCGGAAAGCCGTTAACTGACAGATGGGGTACAGTCCTTC
TGC

C 2
GCAGGTGGACAGGCTTGGTTCTAATGAAGTTAACCCCTGCTGTTCTGGGACATCTGGCCCGGAAAGCCGTTAAC[TGACA]GATGGGGTACAGTCCT
TCTGC

D 3
GCAGGTGGACAGGCTTGGTTCTAATGAAGTTAACCCCTGCTGTTCTGGGACATCTGGCCCGGAAAGCCGTTAAC[[T]]GATGGGGTACAGTCCTTC
TGC

E 4
GCAGGTGGACAGGCTTGGTTCTAATGAAGTTAACCCCTGCTGTTCTGGGACATCTGGCCCGGAAAGCCGTTAAC[TAT]]GGTACAGTCCT
TCTGC

F 5
GCAGGTGGACAGGCTTGGTTCTAATGAAGTTAACCCCTGCTGTTCTGGGACATCTGGCCCGGAAAGCCGTTAAC[TATGATGGAA]ACAGTCCTTC
C

G modificaciones de secuencia comparado con la secuencia previa:
1 subrayado = cambio en bases
2 [un paréntesis] = bases cortadas
3 [[doble paréntesis]] = bases añadidas
4 [[doble paréntesis]] = bases añadidas
5 subrayado = cambio en bases

FIG. 43

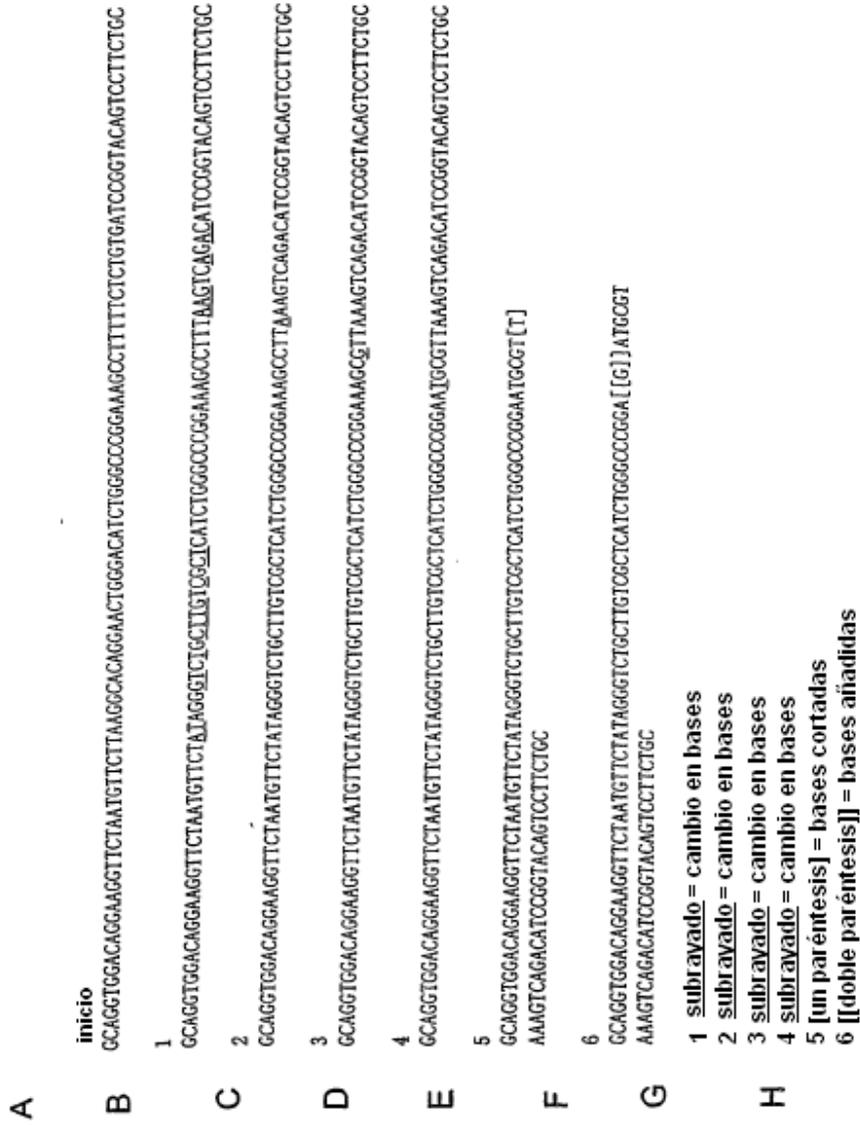


FIG. 44

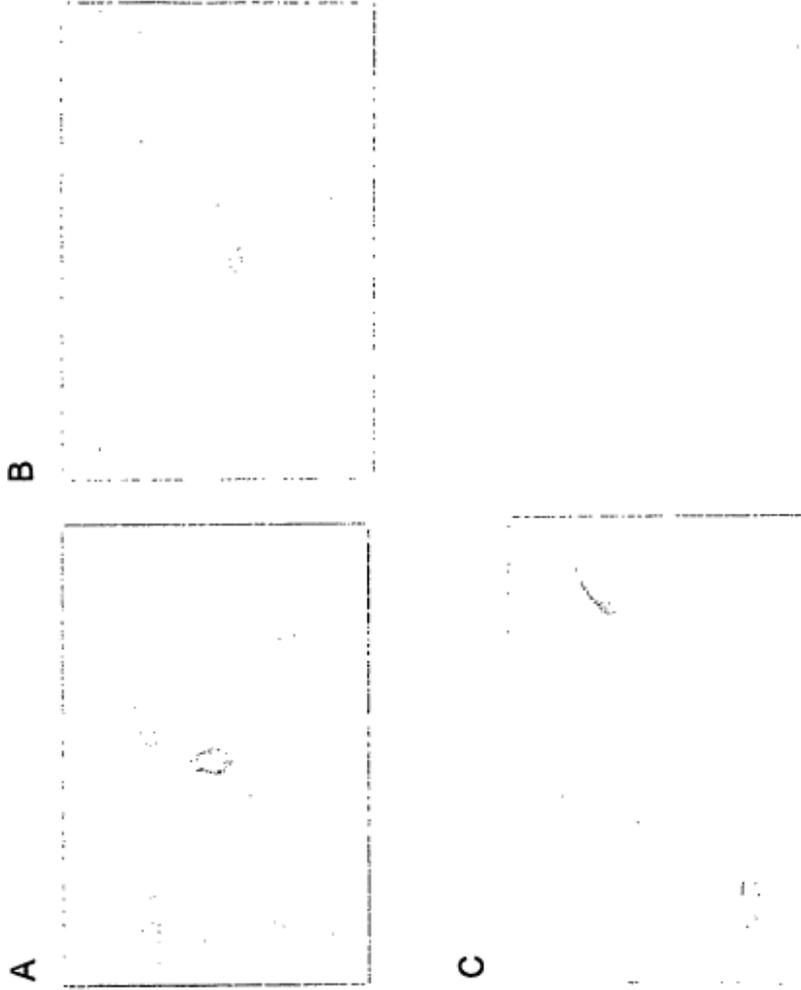


FIG. 45

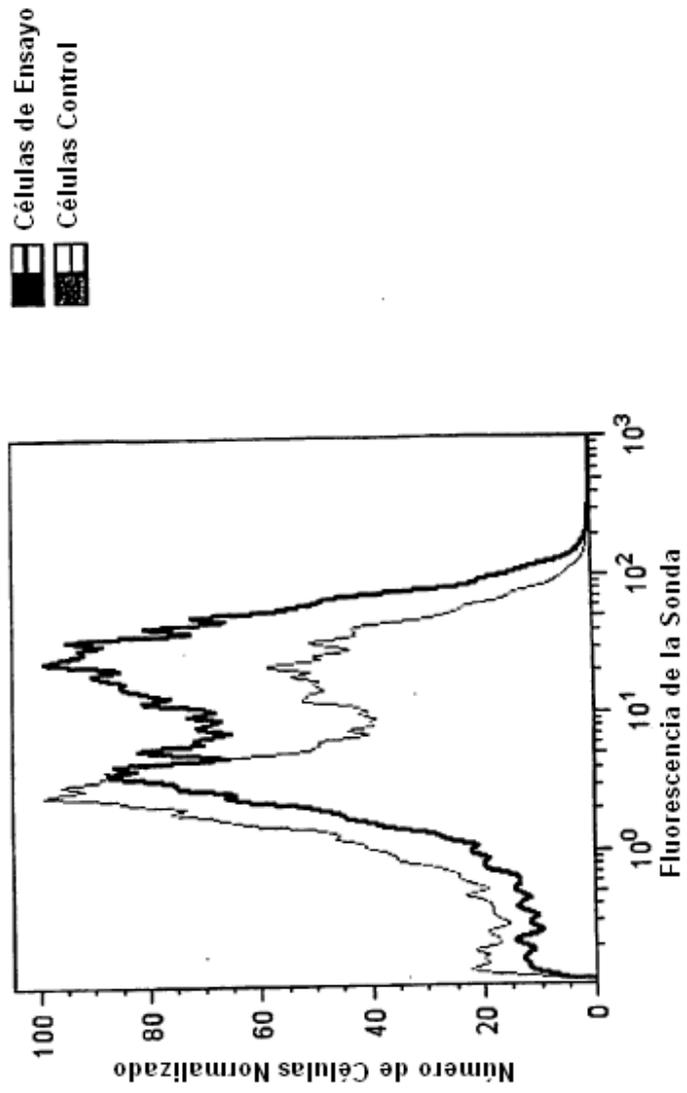


FIG. 46 mcon1

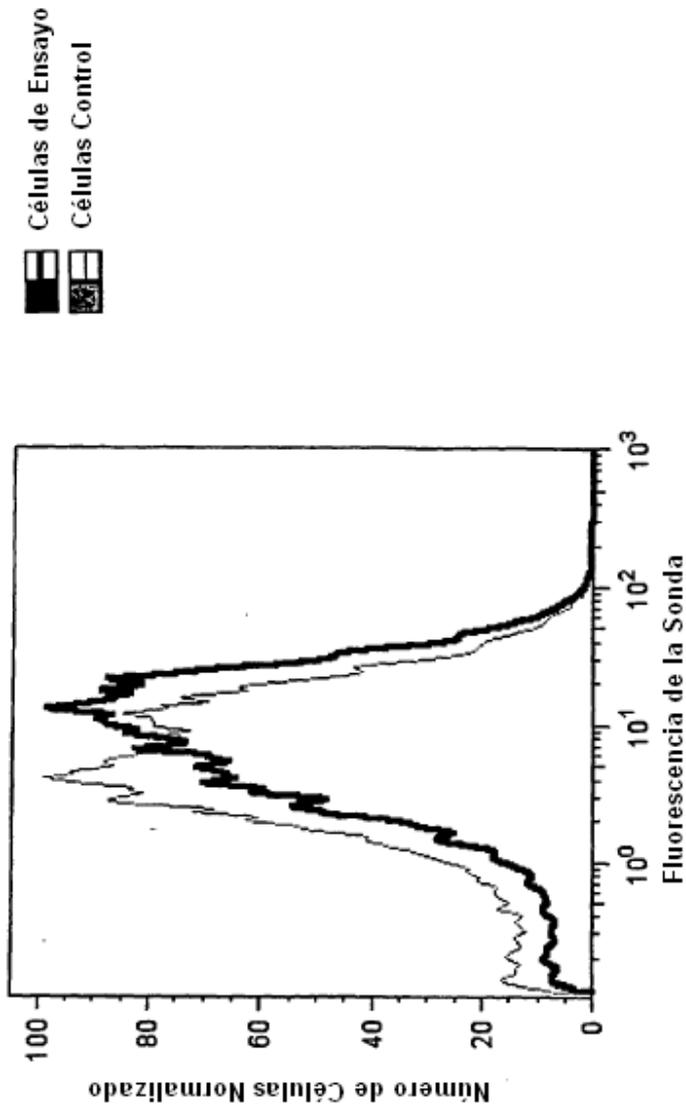


FIG. 47 mcon2

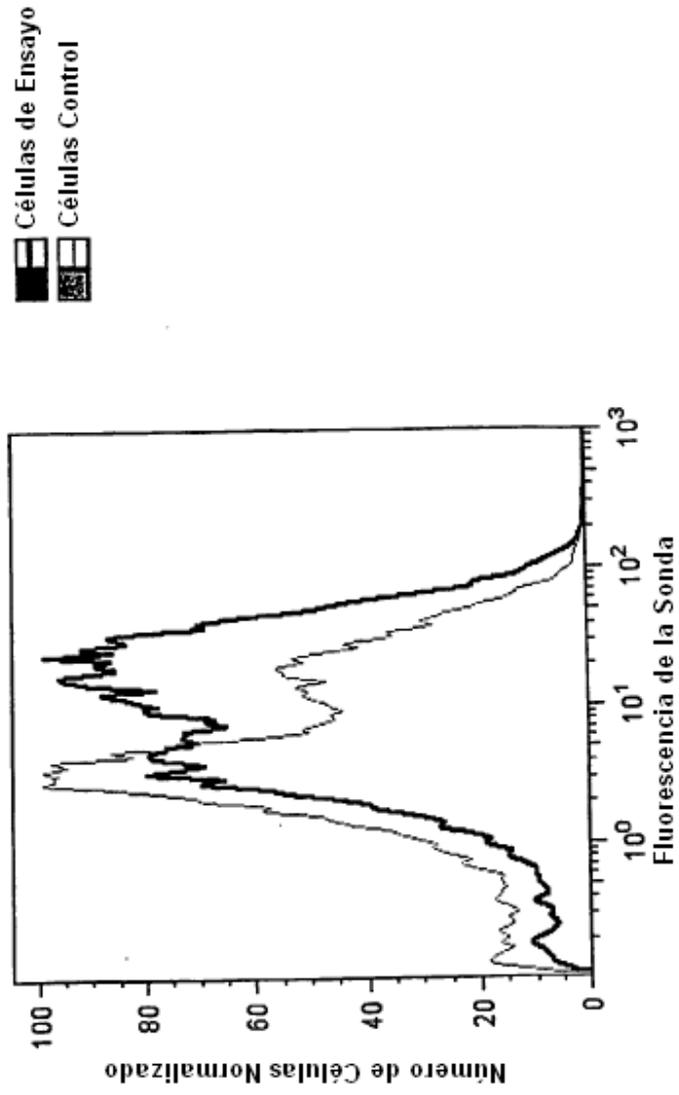


FIG. 48 mcon3

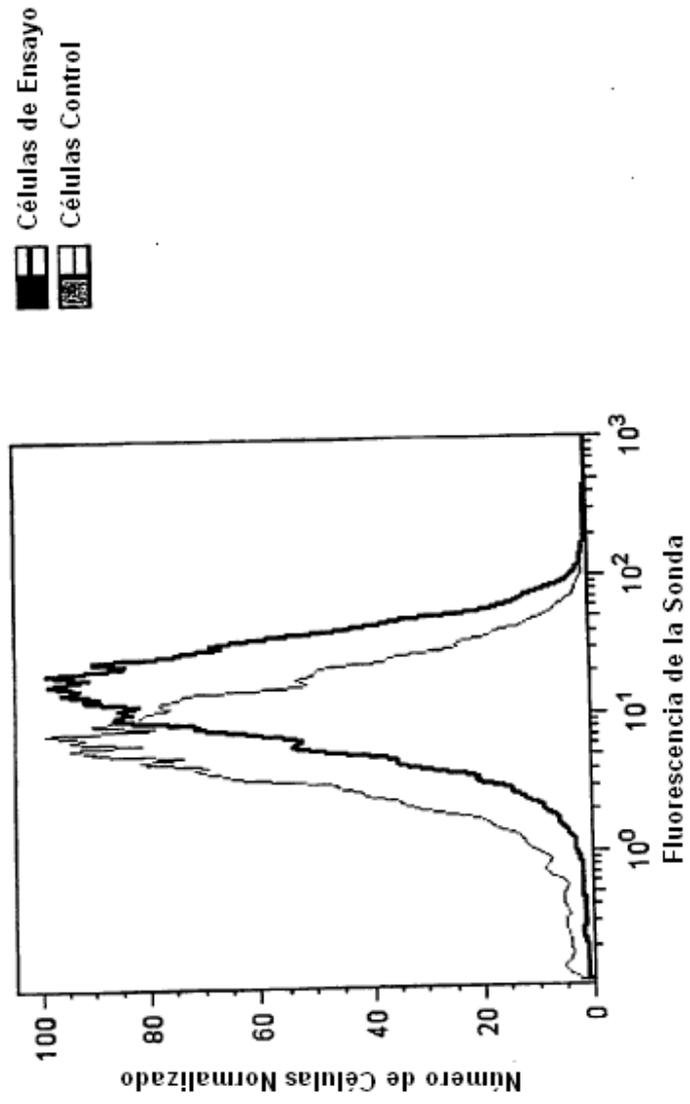


FIG. 49 mcon4

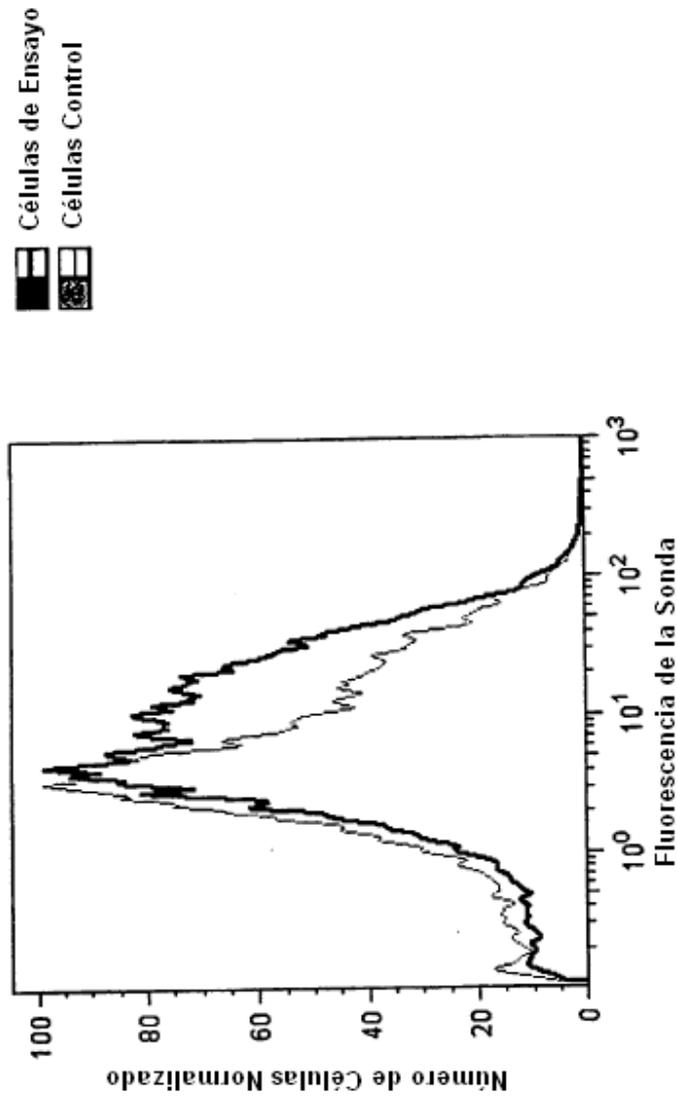


FIG. 50 mcon5

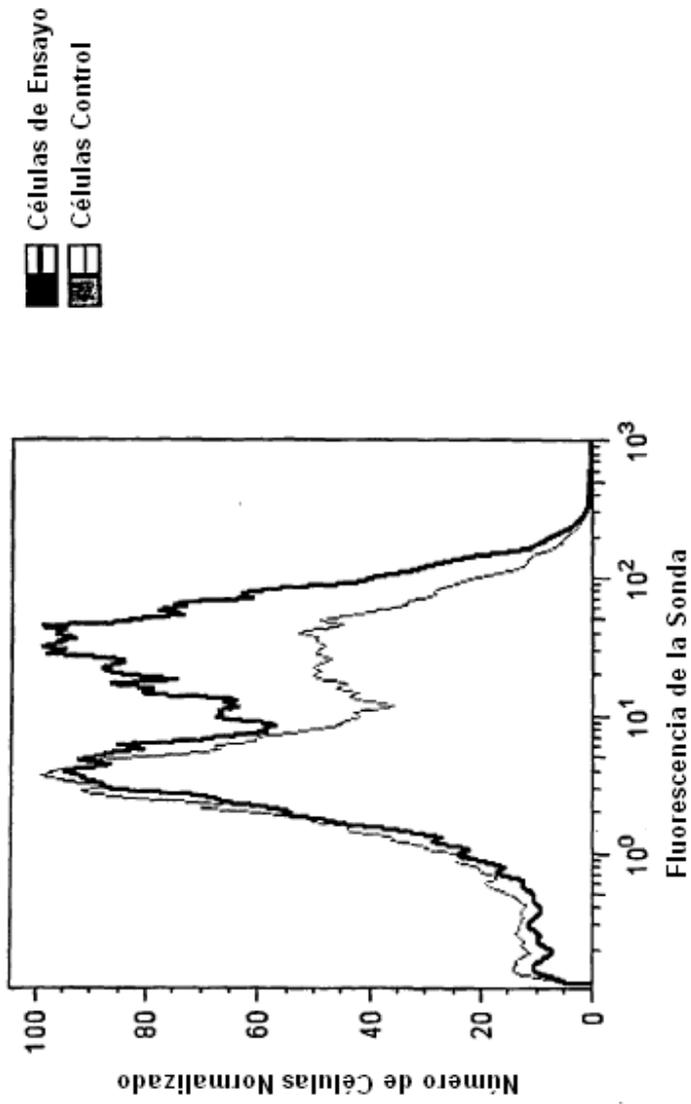


FIG. 51 mcon6

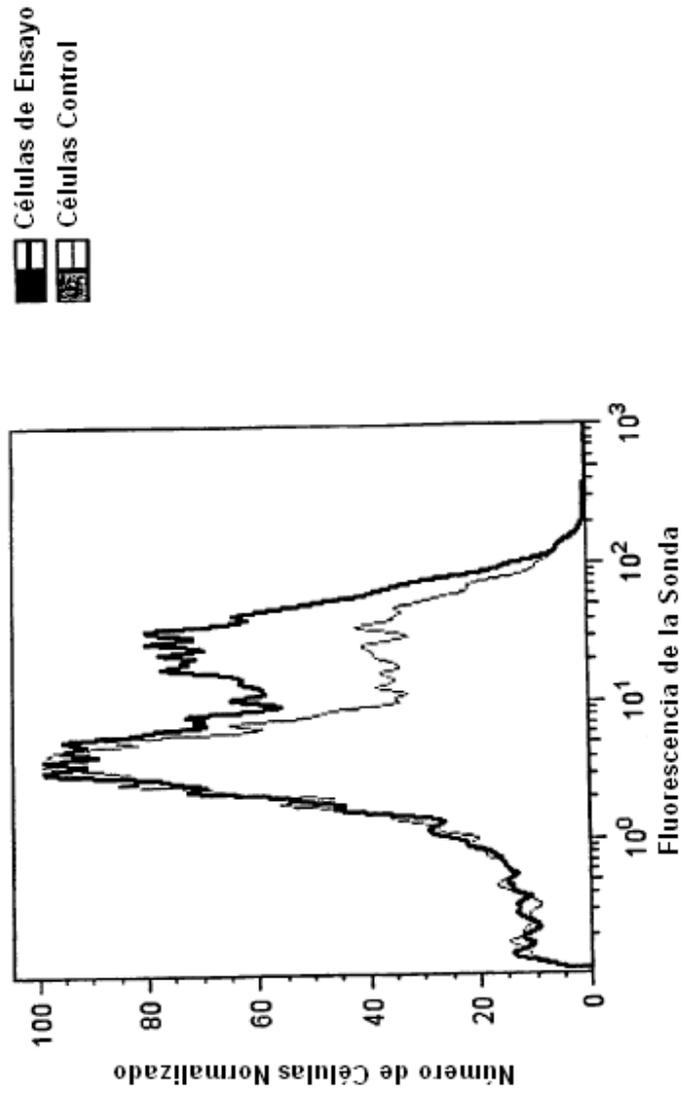


FIG. 52 mcon7

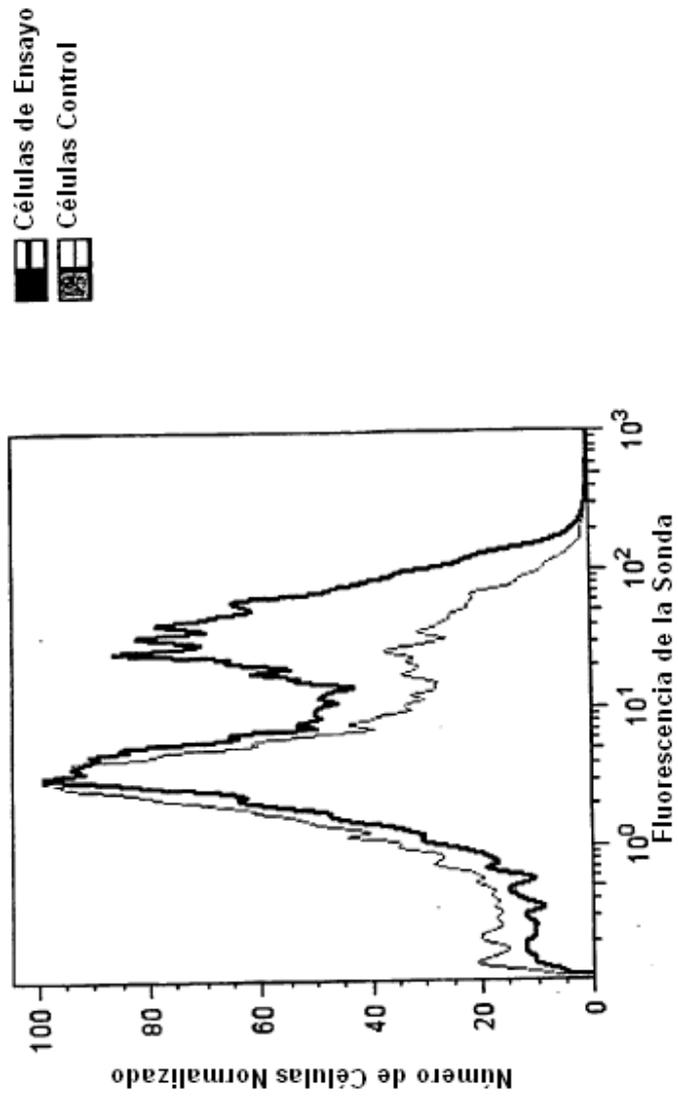


FIG. 53 mcon8

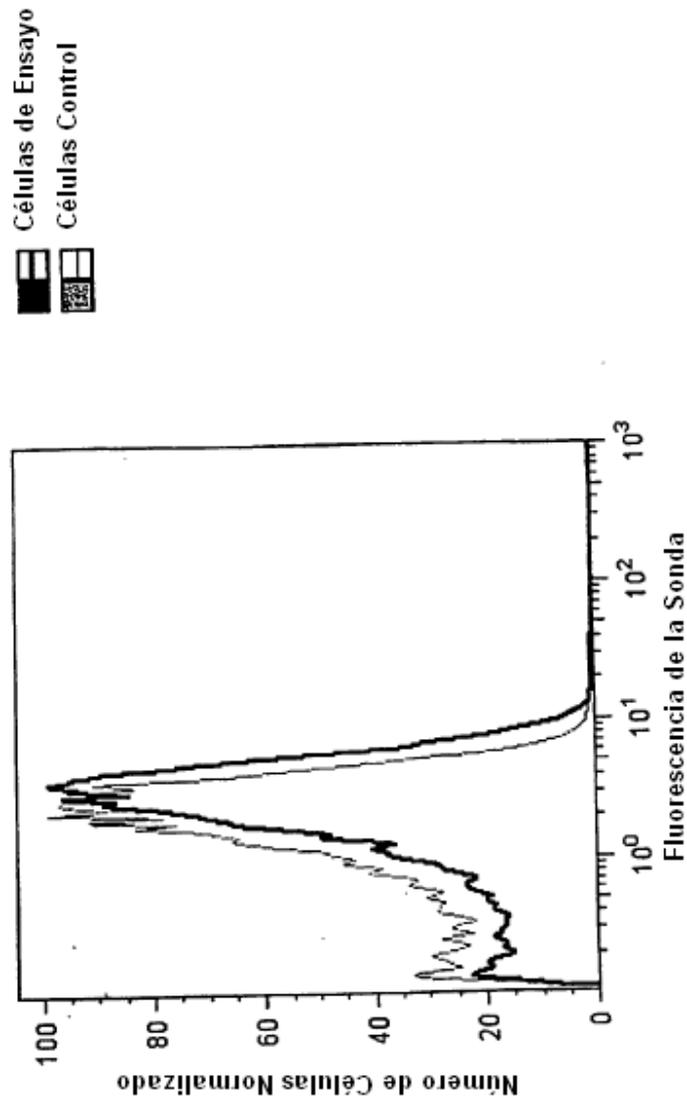


FIG. 54 mcon9

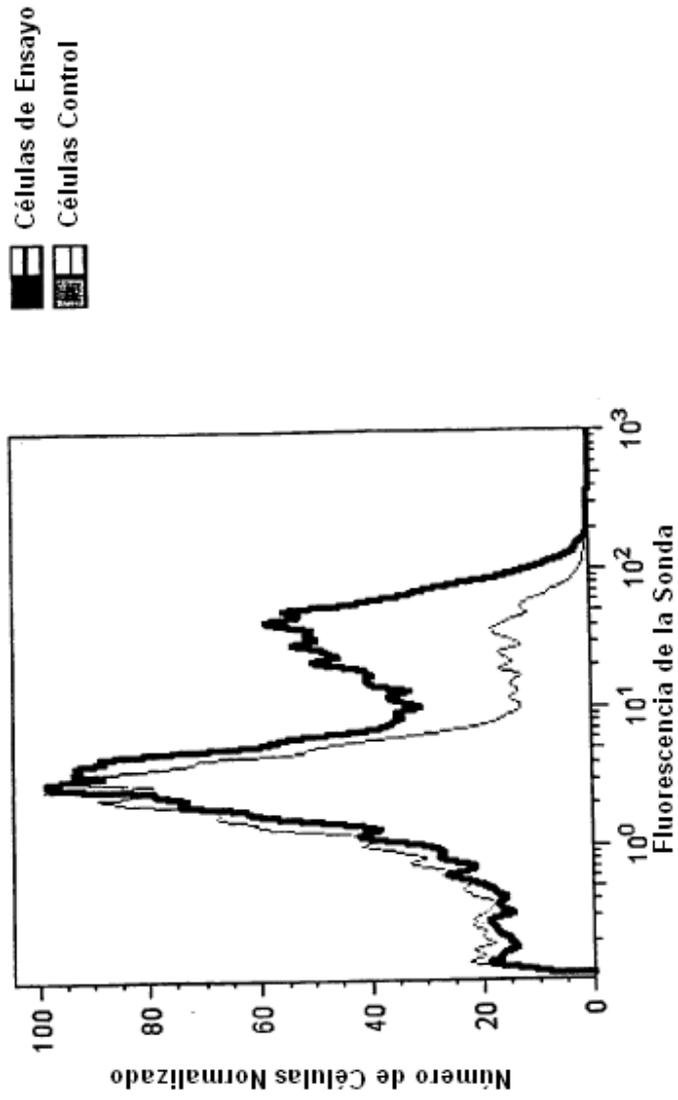


FIG. 55 mcon10

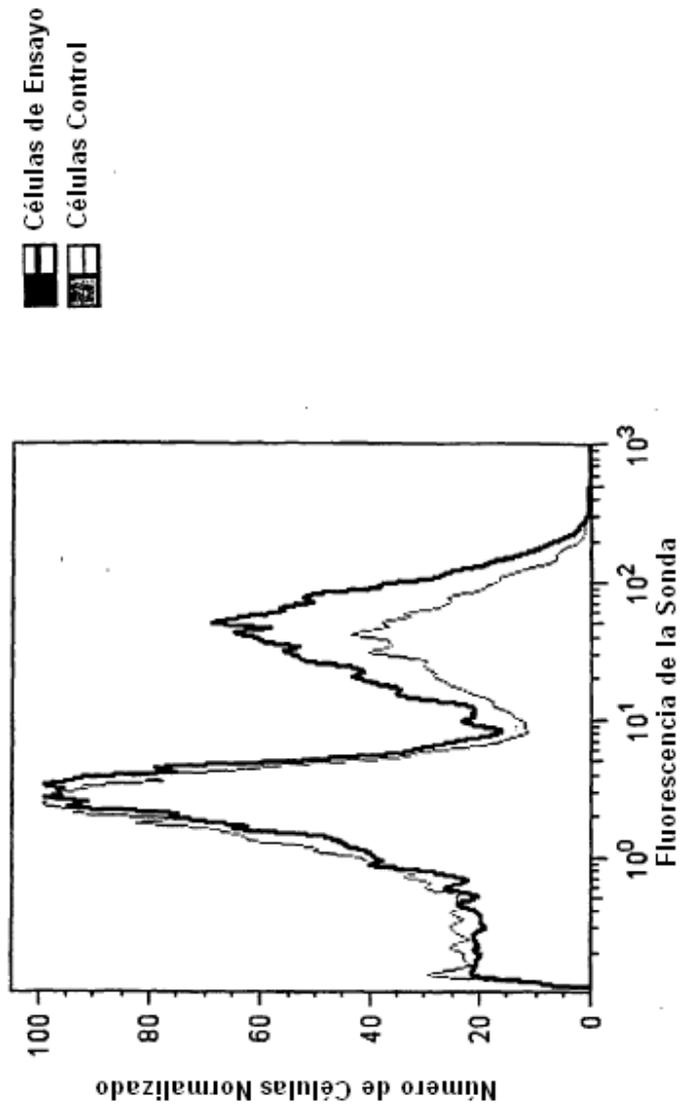


FIG. 56 mcon11

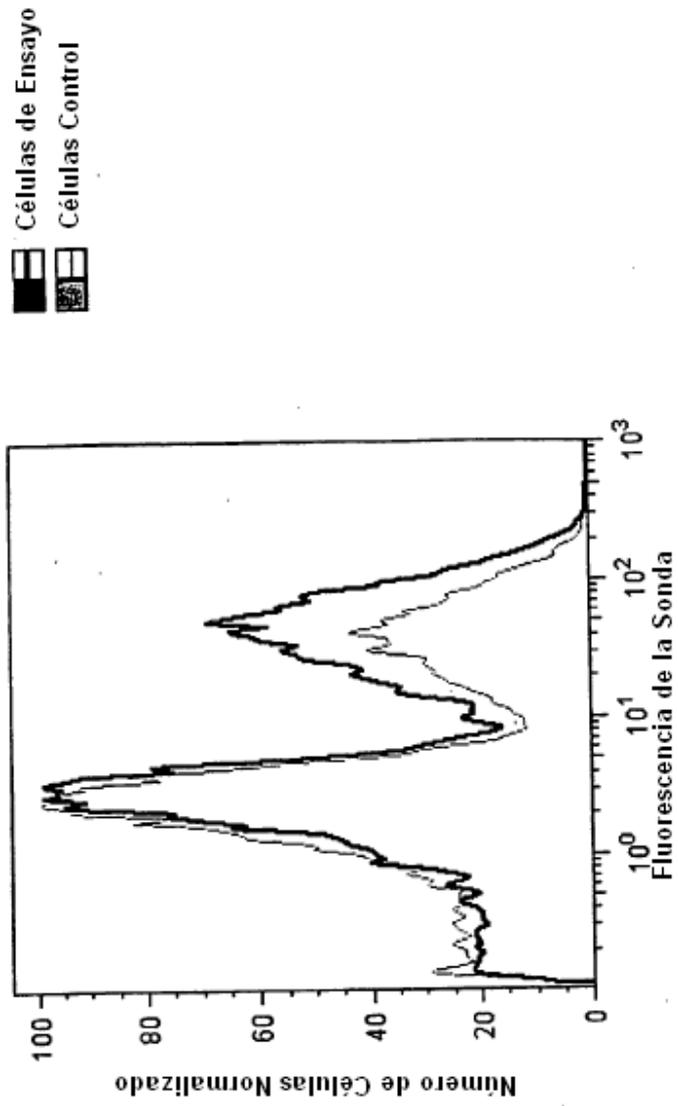


FIG. 56 mcon11

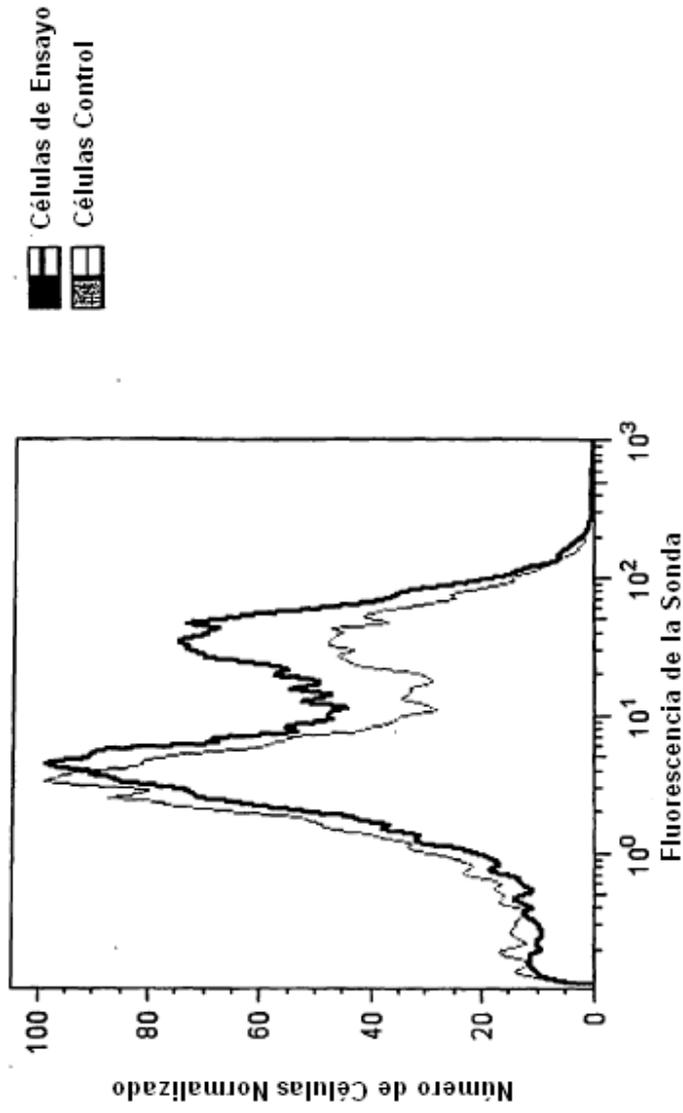


FIG. 57 mcon12

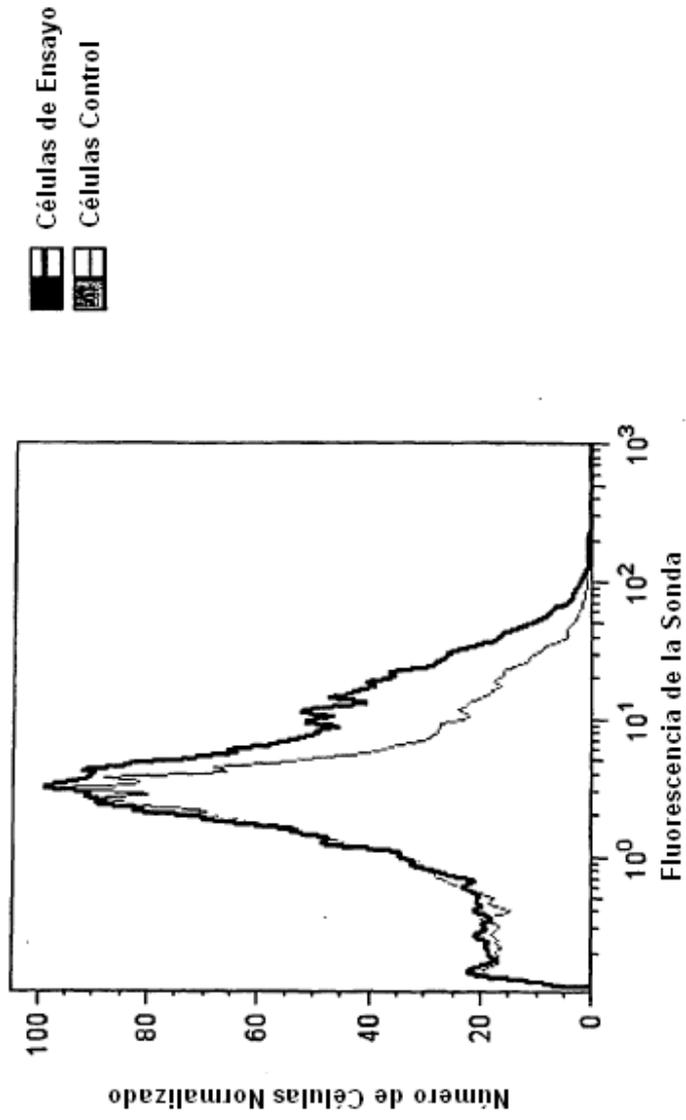


FIG. 58 mcon13

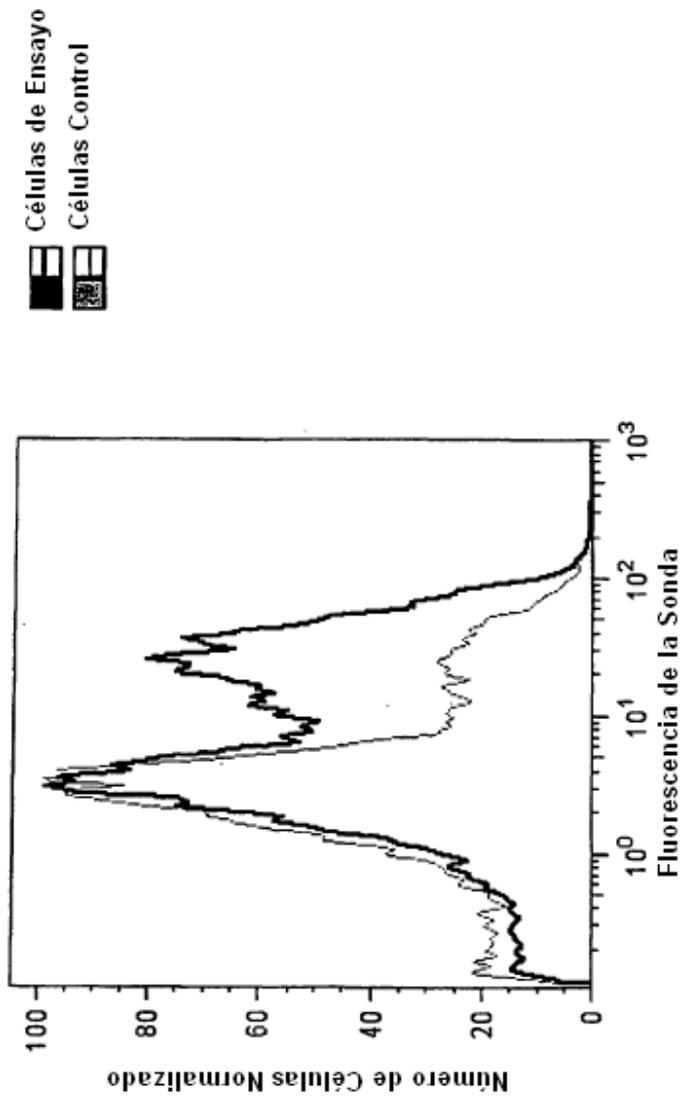


FIG. 59 mcon14

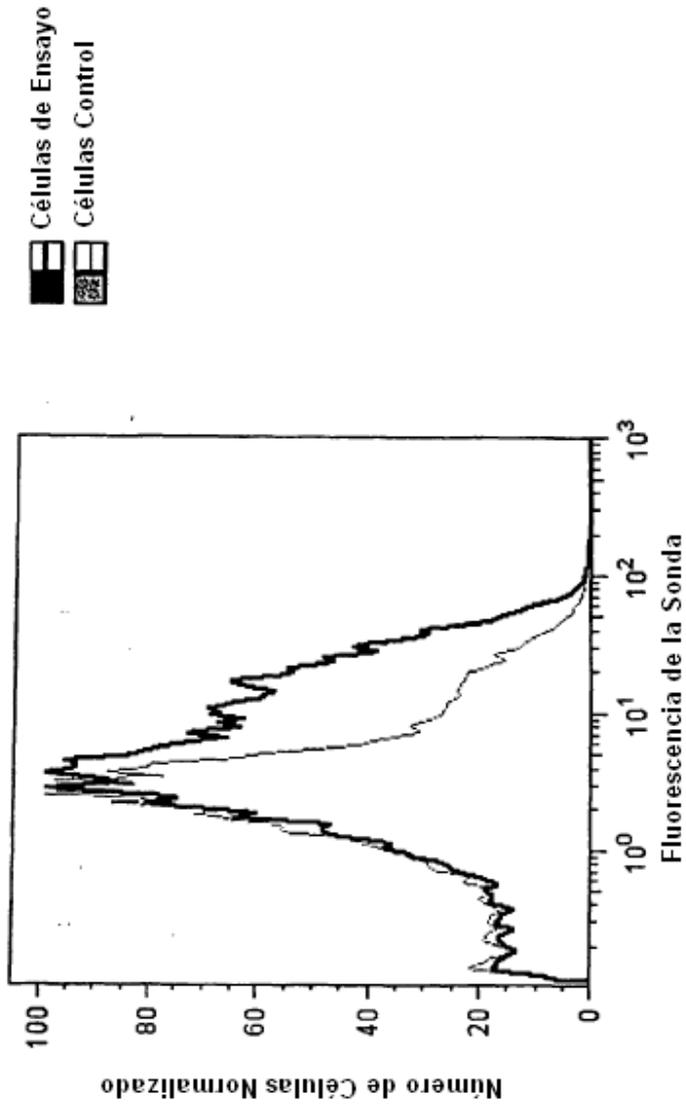


FIG. 60 mcon15