



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 281 138**

51 Int. Cl.:
C12N 15/10 (2006.01)
C12Q 1/68 (2006.01)
C07H 21/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **98948499 .3**
86 Fecha de presentación : **24.09.1998**
87 Número de publicación de la solicitud: **1017853**
87 Fecha de publicación de la solicitud: **12.07.2000**

54 Título: **Bibliotecas de ácido nucleico normalizadas y métodos de producción de las mismas.**

30 Prioridad: **24.09.1997 US 59817 P**
23.09.1998 US 159496

45 Fecha de publicación de la mención BOPI:
16.09.2007

45 Fecha de la publicación del folleto de la patente:
16.09.2007

73 Titular/es: **INVITROGEN CORPORATION**
1600 Faraday Avenue
Carlsbad, California 92008, US

72 Inventor/es: **Li, Wu-Bo;**
Jessee, Joel y
Nisson, Paul E.

74 Agente: **Arizti Acha, Mónica**

ES 2 281 138 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

Bibliotecas de ácido nucleico normalizadas y métodos de producción de las mismas.

5 **Campo de la invención**

La presente invención pertenece al campo de la biología molecular y la genética. La invención se refiere generalmente a métodos para producir bibliotecas de ácido nucleico normalizadas, tales que la variación en la abundancia de las moléculas de ácido nucleico individuales disminuye sustancialmente (por ejemplo, hasta no más de aproximadamente dos órdenes de magnitud). La invención también se refiere a bibliotecas normalizadas producidas por estos métodos, a las moléculas de ácido nucleico aisladas a partir de estas bibliotecas, a los constructos genéticos (por ejemplo, vectores) que comprenden estas moléculas de ácido nucleico, y a las células huésped que comprenden dichas bibliotecas normalizadas.

15 **Antecedentes de la invención**

La aclaración de los mecanismos que dictan el funcionamiento normal de las células vivas requiere un entendimiento detallado de la información codificada en todos los genes (también denominado en el presente documento a modo de sinónimo, el genoma). Para mapear y secuenciar los genes contenidos en los genomas de diferentes organismos, las secuencias del ARN mensajero (ARNm), que son representativas de los genes del genoma, se usan normalmente para evaluar la creación de una célula particular u organismo de interés. Sin embargo, los ARNm (que se calcula que alcanzan un número de 100.000 en los seres humanos) se producen a diferentes niveles dentro de los distintos tipos de célula a diferentes puntos en desarrollo (por ejemplo, hay menos de una copia por célula de algunos ARNm y hay millones de copias por célula de otros). Estos ARNm, su desarrollo y su expresión regulada específica de tipo celular, y su traducción en proteínas es lo que produce el carácter único de un tipo de célula particular. Por ejemplo, las células de músculo adulto producen altos niveles de ARNm de mioglobina, mientras que los glóbulos rojos maduros contienen altos niveles de hemoglobina. En el feto, la hemoglobina se produce por el hígado; sin embargo, tras el nacimiento, el tipo de hemoglobina producida y la fuente tisular cambian, debido a cambios en la expresión génica.

Un entendimiento de los detalles moleculares del funcionamiento normal de las células es esencial para entender y tratar enfermedades heredadas en las que la regulación y la expresión de uno o más genes pueden haber cambiado. En este objetivo se integra la producción de bibliotecas de ácidos nucleicos clonados, a partir de las cuales pueden aislarse todos o sustancialmente todos los miembros de las bibliotecas con aproximadamente la misma probabilidad.

Una biblioteca normalizada con un menor intervalo de concentraciones relativas de sus miembros, por ejemplo tan bajo como aproximadamente 2-4 veces, tendrá la ventaja de permitir que esencialmente todos los ARNm estén disponibles para el aislamiento y posterior análisis. Este tipo de biblioteca ampliaría el entendimiento de la función normal de genes individuales y del genoma en general. Sin embargo, ninguno de los métodos notificados hasta el momento han dado como resultado la producción de bibliotecas de ácido nucleico normalizadas en las que esencialmente se representan todas las moléculas de ácido nucleico o genes expresados en un tipo de célula o tejido particular y pueden aislarse con una alta probabilidad. Aunque algunos investigadores han intentado normalizar (es decir, reducir la variación en la abundancia relativa de los componentes de la población de moléculas de ácido nucleico), ninguno ha tenido éxito al llevar la abundancia relativa de la población total hasta un intervalo de dos órdenes de magnitud (Bonaldo, M., Lennon, G., Soares, M.B., *Genome Res.* 6:791-866 (1996); Ko, M.S.H., *Nucl. Acids Res.* 18:5705-5711 (1990); Pantanjali, S.R., *et al.*, *Proc. Natl. Acad. Sci. USA* 88:1943-1947 (1991); Soares, M.B., *Proc. Natl. Acad. Sci. USA* 91:9228-9232 (1994)). Las bibliotecas resultantes "normalizadas" no han podido proporcionar la cantidad de información novedosa necesaria para entender la expresión de la mayoría de los genes. Por tanto, existe una necesidad actual de métodos de producción de bibliotecas de ácido nucleico normalizadas, y de bibliotecas de ácido nucleico normalizadas producidas mediante tales métodos.

El documento WO 95/08647 describe un método para la construcción de bibliotecas de ADNc normalizadas, y usos de las mismas.

Bonaldo Fatina de M *et al.*, *Genome Research*, Vol. 6, Nº 9, 1 de septiembre de 1996, Págs. 791-806, describe la "normalización" y "sustracción": dos enfoques para facilitar el descubrimiento de genes.

Ko, *Nucleic Acids Research*, Vol. 18, Nº 19, 1990, págs. 5705-5711, describe una "biblioteca de ADNc equiparada" mediante la reasociación de ADNc de doble cadena cortos.

El documento WO 95/11986 describe un método para generar una biblioteca de ADNc sustraída, y usos de la biblioteca.

Hongping Jiang *et al.*, *Molecular and Cellular Differentiation*, Vol. 1, Nº 3, 1993, págs. 285-299, describe el uso de un protocolo de hibridación de sustracción sensible y eficaz para la identificación de genes regulados diferencialmente durante la inducción de la diferenciación en células de melanoma humanas.

Sankhavaram *et al.*, *P.N.A.S. (USA)*, Vol. 88, Nº 5, 1 de marzo de 1991, págs. 1943-1947, describe la construcción de una biblioteca de ADNc (normalizada) de abundancia uniforme.

ES 2 281 138 T3

La presente invención proporciona un método para la normalización de una biblioteca de ácido nucleico que comprende:

- 5 (a) incubar una biblioteca de ácido nucleico que va normalizarse con moléculas de ácido nucleico haptenilado sintéticas complementarias a todas o a una parte de las moléculas de ácido nucleico de dicha biblioteca en condiciones que favorecen la hibridación de las moléculas más abundantes de dicha biblioteca con dichas moléculas de ácido nucleico haptenilado; y
- 10 (b) eliminar dichas moléculas hibridadas mediante interacciones ligando-hapteno, produciendo así una biblioteca normalizada.

La presente invención, por tanto, satisface esta necesidad proporcionando métodos para producir bibliotecas de ácido nucleico normalizadas (por ejemplo, bibliotecas de moléculas de ácido nucleico clonadas de las cuales cada molécula de ácido nucleico miembro puede aislarse con aproximadamente una probabilidad equivalente). En particular, 15 la invención se refiere a métodos para normalizar una biblioteca de ácido nucleico, que puede ser una biblioteca de ADNc de cadena sencilla o cadena doble, tal como se proporciona en las reivindicaciones.

En un aspecto preferido de la invención, la concentración relativa de todos los miembros de la biblioteca normalizada se encuentra dentro de uno a dos órdenes de magnitud. En otro aspecto preferido, la invención permite la supresión o eliminación de la molécula de ácido nucleico contaminante de la biblioteca normalizada. Tal contaminación puede 20 incluir vectores dentro de la biblioteca que no contienen insertos (por ejemplo, ruido de fondo). De esta manera, toda o una parte sustancial de la biblioteca normalizada comprenderá vectores que contienen moléculas de ácido nucleico insertadas de la biblioteca.

La invención también hace referencia a tales métodos en los que las condiciones que favorecen la hibridación de las moléculas más abundantes de la biblioteca con las moléculas hapteniladas se seleccionan del grupo que consiste en: (a) una Cot igual a o superior a 25; (b) una Cot igual a o superior a 50; (c) una Cot igual o superior a 100; (d) una Cot igual a o superior a 1.000; (e) una Cot igual a o superior a 2.000; (f) una Cot igual a o superior a 5.000; 25 (g) una Cot desde aproximadamente 10 hasta 10.000; (h) una Cot desde aproximadamente 25 hasta 10.000; (i) una Cot desde aproximadamente 50 hasta 10.000; (j) una Cot desde aproximadamente 1.000 hasta 10.000; (k) una Cot desde aproximadamente 5.000 hasta 10.000; (l) una Cot desde aproximadamente 500 hasta 5.000; (m) una Cot desde aproximadamente 100 hasta 1000; y (n) una Cot de menos de 10.000.

En un aspecto preferido de la invención, se incuba una población de ARNm en condiciones suficientes para producir una población de moléculas de ADNc complementarias a todas o a una parte de dichas moléculas de ARNm. Preferiblemente, tal población de moléculas de ADNc (por ejemplo ADNc de cadena sencilla) se produce mezclando la población de moléculas de ARNm (moléculas molde) con uno o más polipéptidos que tienen actividad transcriptasa inversa que incuban dicha mezcla en condiciones suficientes para producir una población de moléculas de ADNc de 30 cadena sencilla complementarias a todas o a una parte de dichas moléculas de ARNm. Las moléculas de ADNc de cadena sencilla pueden usarse entonces como moléculas molde para crear moléculas de ADNc de cadena doble incubando la mezcla en las condiciones apropiadas en presencia de una o más ADN polimerasas. La población resultante de bibliotecas de ADNc de cadena doble o cadena sencilla puede normalizarse según la invención. Preferiblemente, tales bibliotecas de ADNc se insertan en uno o más vectores antes de la normalización. Alternativamente, las bibliotecas de ADNc pueden normalizarse antes de la inserción en uno o más vectores y tras la normalización pueden clonarse 45 en uno o más vectores.

En un aspecto particularmente preferido de la invención, la biblioteca que va a normalizarse está contenida en (insertada en) uno o más vectores, que pueden ser un plásmido, un cósmido, un fagémido y similar. Tales vectores preferiblemente comprenden uno o más promotores que permiten la síntesis de al menos una molécula de ARN a 50 partir de todas o una parte de las moléculas de ácido nucleico (preferiblemente moléculas de ADNc) insertadas en el vector. Así, mediante el uso de los promotores, pueden crearse y usarse moléculas de ARN hapteniladas complementarias a todas o a una parte de las moléculas de ácido nucleico de la biblioteca para normalizar la biblioteca según la invención. Tales moléculas de ARN sintetizadas (que se han haptenilado) serán complementarias a todos o a una parte de los insertos de vector de la biblioteca. Las moléculas más abundantes en la biblioteca pueden eliminarse entonces preferentemente mediante la hibridación de las moléculas de ARN hapteniladas con la biblioteca, produciendo de esta 55 manera la biblioteca normalizada de la invención. Sin ser limitadas, se piensa que las moléculas de ARN sintetizadas son representativas de la biblioteca; es decir, la especie más abundantes en la biblioteca da como resultado un ARN haptenilado más abundante usando el método anterior. La abundancia relativa de las moléculas dentro de la biblioteca, y por tanto, dentro del ARN haptenilado determina la tasa de eliminación de una especie particular de la biblioteca; si la abundancia de una especie particulares es alta, tal especie altamente abundante se eliminará más fácilmente mientras que una especie poco abundante se eliminará menos fácilmente de la población. La normalización mediante este proceso permite por tanto equiparar sustancialmente el nivel de cada especie dentro de la biblioteca.

En otro aspecto preferido de la invención, la biblioteca que va a normalizarse no necesita insertarse en uno o más 65 vectores antes de la normalización. En tal aspecto de la invención, las moléculas de ácido nucleico de la biblioteca pueden usarse para sintetizar moléculas de ácido nucleico hapteniladas usando técnicas bien conocidas. Por ejemplo, moléculas de ácido nucleico hapteniladas pueden sintetizarse en presencia de una o más ADN polimerasas, uno o más cebadores o sondas apropiados o uno o más nucleótidos (los nucleótidos y/o cebadores o sondas pueden ser

haptencilados). De esta manera, se producirán moléculas de ADN haptenciladas y podrán usarse para normalizar la biblioteca según la invención. Alternativamente, pueden añadirse uno o más promotores a (o ligarse a) las moléculas de la biblioteca, permitiendo de esta manera la síntesis de moléculas de ARN haptenciladas para usarse para normalizar la biblioteca según la invención. Por ejemplo, los adaptadores que contienen uno o más promotores se añaden a
 5 (o se ligan a) uno o más extremos de moléculas de cadena doble de la biblioteca (por ejemplo, una biblioteca de ADNc preparada a partir de una población de moléculas de ARNm). Tales promotores pueden usarse entonces para preparar moléculas de ARN haptenciladas complementarias a todas o a una parte de las moléculas de ácido nucleico de la biblioteca. Según la invención, la biblioteca puede normalizarse entonces y, si se desea, insertarse en uno o más vectores.

10 Mientras el ARN haptencilado se usa preferiblemente para normalizar bibliotecas, otras moléculas de ácido nucleico haptencilado pueden usarse según la invención. Por ejemplo, puede sintetizarse ADN haptencilado a partir de la biblioteca y usarse según la invención.

15 Haptencios adecuados para usarse en los métodos de la invención incluyen, pero no se limitan a, avidina, estreptavidina, proteína A, proteína G, un receptor Fc de la superficie celular, un antígeno específico de anticuerpo, un sustrato específico de enzima, polimixina B, proteína neutralizante de endotoxina (ENP), Fe⁺⁺⁺, un receptor de transferrina, un receptor de insulina, un receptor de citocina, CD4, espectrina, fodrina, ICAM-1, ICAM-2, C3bi, fibrinógeno, Factor X, anquirina, una integrina, vitronectina, fibronectina, colágeno, laminina, glicoforina, Mac-1, LFA-1, β -actina,
 20 gp120, una citocina, insulina, ferrotansferrina, apotransferrina, lipopolisacárido, una enzima, un anticuerpo, biotina y combinaciones de los mismos. Un haptencio particularmente preferido es biotina.

Según la invención, pueden aislarse moléculas híbridadas producidas por los métodos anteriormente descritos, por ejemplo, mediante extracción o mediante interacciones ligando-haptencio. Preferiblemente, se usan métodos de extracción (por ejemplo, usando disolventes orgánicos. Puede lograrse el aislamiento mediante interacciones ligando-haptencio incubando las moléculas haptenciladas con un soporte sólido que comprende al menos un ligando que se une al haptencio. Los ligandos preferidos para usar en tales métodos de aislamiento corresponden al haptencio particular usado, e incluyen, pero no se limitan a, biotina, un anticuerpo, una enzima, lipopolisacárido, apotransferrina, ferrotansferrina, insulina, una citocina, gp120, β -actina, LFA-1, Mac-1, glicoforina, laminina, colágeno, fibronectina, vitronectina, una integrina, anquirina, C3bi, fibrinógeno, Factor X, ICAM-1, ICAM-2, espectrina, fodrina, CD4, un receptor de citocina, un receptor de insulina, un receptor de transferrina, Fe⁺⁺⁺, polimixina B, proteína que neutraliza endotoxina (ENP), un sustrato específico de enzima, proteína A, proteína G, un receptor Fc de la superficie celular, un antígeno específico de anticuerpo, avidina, estreptavidina o combinaciones de los mismos. El soporte sólido usado en estos métodos de aislamiento puede ser nitrocelulosa, diazocelulosa, vidrio, poliestireno, poli(cloruro de vinilo), polipropileno, polietileno, dextrano, Sepharose, agar, almidón, nylon, una perla de látex, una perla magnética, una perla paramagnética, una perla supramagnética o una placa de microtitulación. Los soportes sólidos preferidos son perlas magnéticas, perlas paramagnéticas y perlas supramagnéticas, y particularmente preferidas son las perlas que comprenden una o más moléculas de estreptavidina o avidina.

40 En otro aspecto de la invención, se somete a las bibliotecas normalizadas a un aislamiento adicional o etapas de selección que permiten la eliminación de contaminación o ruido de fondo no deseado. Tal contaminación o ruido de fondo puede incluir ácidos nucleicos no deseados. Por ejemplo, cuando una biblioteca que va a normalizarse se construye en uno o más vectores, puede estar presente un bajo porcentaje de vector (sin inserto) en la biblioteca. Con la normalización, tal baja abundancia de moléculas (por ejemplo, ruido de fondo de vector) puede volverse un constituyente más significativo como resultado del proceso de normalización. Es decir, puede aumentarse el nivel relativo de tal baja abundancia de ruido de fondo como parte del proceso de normalización.

Puede lograrse la eliminación de tales ácidos nucleicos contaminantes incubando una biblioteca normalizada con una o más sondas haptenciladas que son específicas para las moléculas de ácido nucleico de la biblioteca (por ejemplo sondas específicas de diana). En particular, la eliminación de secuencias contaminantes puede lograrse seleccionando aquellos ácidos nucleicos que tienen la secuencia de interés o eliminando aquellas moléculas que no contienen secuencias de interés. Según la invención, la eliminación de moléculas de ácido nucleico contaminantes puede realizarse sobre cualquier biblioteca normalizada (ya esté o no la biblioteca construida en un vector). Por tanto, las sondas se diseñarán de manera que no reconozcan o hibriden los ácidos nucleicos contaminantes (como en la realización preferida que usa sonda de biotina oligodA-*NotI* 3'). Con la hibridación de la sonda haptencilada con moléculas de ácido nucleico de la biblioteca, las sondas haptenciladas se unirán a y seleccionarán las secuencias deseadas en la biblioteca normalizada y dejarán las moléculas de ácido nucleico contaminantes, dando como resultado una biblioteca normalizada. La biblioteca normalizada seleccionada puede entonces aislarse. En un aspecto preferido, tales bibliotecas normalizadas seleccionadas aisladas son de cadena sencilla, y pueden convertirse en cadena doble tras la selección incubando la biblioteca de cadena sencilla en condiciones suficientes como para hacer que las moléculas de ácido nucleico sean de cadena doble. Las moléculas de cadena doble pueden transformarse entonces en una o más células huésped. Alternativamente, la biblioteca normalizada puede convertirse en cadena doble usando la sonda o cebador haptencilado (preferiblemente específico de diana) y luego seleccionado mediante extracción o interacciones ligando-haptencio. Tales moléculas de cadena doble seleccionadas pueden transformarse en una o más células huésped.

65 En otro aspecto de la invención, pueden reducirse o eliminarse los ácidos nucleicos contaminantes incubando la biblioteca normalizada en presencia de uno o más cebadores específicos para secuencias de biblioteca (específicos para clones que contienen insertos, por ejemplo oligodA-*NotI*). Este aspecto de la invención puede comprender incubar

la biblioteca normalizada de cadena sencilla con uno o más nucleótidos (preferiblemente nucleótidos que confieren resistencia a nucleasa a las moléculas de ácido nucleico sintetizadas), y uno o más polipéptidos que tienen actividad polimerasa, en condiciones suficientes como para hacer que las moléculas de ácido nucleico sean de cadena doble. Las moléculas de cadena doble resultantes pueden transformarse entonces en una o más células huésped. Alternativamente, las moléculas de cadena doble resultantes que contienen nucleótidos que confieren resistencia a nucleasa pueden digerirse con tal nucleasa y transformarse en una o más células huésped.

En todavía otro aspecto, la eliminación o supresión del ácido nucleico contaminante puede lograrse antes de la normalización de la biblioteca, dando como resultado de esta manera la biblioteca normalizada seleccionada de la invención. En tal método, la biblioteca que va a normalizarse puede someterse a cualquiera de los métodos descritos en el presente documento para eliminar moléculas de ácido nucleico no deseadas y luego la biblioteca puede normalizarse mediante el proceso de la invención para proporcionar las bibliotecas normalizadas seleccionadas de la invención.

Según la invención, las moléculas de ácido nucleico de cadena doble se convierten en cadena sencilla antes de la hibridación. Así, los métodos de la invención pueden comprender además tratar las moléculas de ácido nucleico de cadena doble de la biblioteca anteriormente descritas en condiciones suficientes como para hacer que las moléculas de ácido nucleico sean de cadena sencilla. Tales condiciones pueden comprender la degradación de una cadena de las moléculas de ácido nucleico de cadena doble (preferiblemente usando proteína de gen II y exonucleasa III), o la desnaturalización de las moléculas de ácido nucleico de doble cadena usando calor, condiciones alcalinas y similares.

La invención también hace referencia a bibliotecas de ácido nucleico normalizadas, bibliotecas de ácido nucleico normalizadas seleccionadas y células huésped transformadas mediante los métodos anteriormente descritos.

La figura 1 es una representación esquemática del fagémido que se ha usado para construir una biblioteca de ADNc clonada direccionalmente.

La figura 2 es un diagrama esquemático de la producción de bibliotecas de fagémido normalizadas usando hibridación sustractiva con un controlador de ARN de biblioteca total biotinilada denominada a modo de sinónimo molécula de ácido nucleico haptenilado.

La figura 3 es un diagrama que muestra como pueden usarse las sondas específicas de diana biotiniladas en 3' para producir bibliotecas de fagémido normalizadas de ruido de fondo bajo también denominadas en el presente documento como bibliotecas normalizadas seleccionadas.

La figura 4 es un diagrama que muestra como puede usarse una sonda específica de diana biotinilada en 5' para reducir el ruido de fondo en bibliotecas de fagémido normalizadas también denominadas en el presente documento bibliotecas normalizadas seleccionadas.

La figura 5 es un diagrama que muestra como los nucleótidos resistentes a nucleasa y una nucleasa dan bibliotecas de fagémido normalizadas de ruido de fondo bajo, también denominadas en el presente documento bibliotecas normalizadas seleccionadas.

La figura 6 es una fotografía de un gel teñido con bromuro de etidio del enriquecimiento de varios ADNc de TGF β , que están presentes en abundancias considerablemente diferentes en una biblioteca de ADNc no normalizada, a diferentes Cot de sustracción en una biblioteca de ADNc de cerebro fetal humano normalizada para la que se han aplicado dos métodos de eliminación de ruido de fondo diferentes.

La figura 7 es una representación esquemática de la normalización de una biblioteca usando promotores que comprenden adaptadores. Tras la normalización, puede clonarse la biblioteca en un vector. En este método, la eliminación de secuencias de vector contaminantes puede ser innecesaria, dado que la selección de secuencias de ruido de fondo puede realizarse antes de la clonación.

Descripción detallada de la invención

Definiciones

En la siguiente descripción, se usan ampliamente varios términos empleados en la tecnología de ADN recombinante. Para proporcionar un entendimiento más claro y consistente de la memoria descriptiva y reivindicaciones, incluyendo el alcance que va a darse a tales términos, se proporcionan las siguientes definiciones.

Biblioteca. Tal como se usa en el presente documento, el término “biblioteca” o “biblioteca de ácido nucleico” significa un conjunto de moléculas de ácido nucleico (circulares o lineales) representativas de todo o una parte significativa del contenido de ADN de un organismo (una “biblioteca genómica”), o un conjunto de moléculas de ácido nucleico representativas de todos o una parte significativa de los genes expresados (una “biblioteca de ADNc”) en una célula, tejido, órgano u organismo. Tales bibliotecas pueden o no estar contenidas en uno o más vectores.

Normalizado. Tal como se usa en el presente documento, el término “normalizado” o “biblioteca normalizada” significa una biblioteca de ácido nucleico que se ha manipulado, preferiblemente usando los métodos de la invención,

ES 2 281 138 T3

para reducir la variación relativa en abundancia entre moléculas de ácido nucleico miembro en la biblioteca hasta un intervalo no superior a aproximadamente 25 veces, no superior a aproximadamente 20 veces, no superior a aproximadamente 15 veces, no superior a aproximadamente 10 veces, no superior a aproximadamente 7 veces, no superior a aproximadamente 6 veces, no superior a aproximadamente 5 veces, no superior a aproximadamente 4 veces, no superior a aproximadamente 3 veces o no superior a aproximadamente 2 veces.

Controlador. Tal como se usa en el presente documento, el término “controlador” se refiere a una población de moléculas de ácido nucleico (preferiblemente ARN) que son complementarias a todas o a una parte de moléculas de ácido nucleico de una biblioteca. Tal controlador comprende preferiblemente uno o más haptenos y preferiblemente están en exceso molar (superior a 10, preferiblemente superior a 20 veces) comparado con la biblioteca de interés. Según la invención, el controlador se sintetiza preferiblemente a partir de la biblioteca que va a normalizarse y luego se usa el controlador para normalizar esa biblioteca.

Ruido de fondo. Tal como se usa en el presente documento, ruido de fondo hace referencia a moléculas de ácido nucleico contaminantes que pueden estar presentes en una biblioteca construida. Las moléculas de ácido nucleico contaminantes típicas son vectores en los que se ha construido la biblioteca pero que han perdido la molécula de ácido nucleico insertada (mediante delección o de otra manera) o que no contienen insertos de ácido nucleico. Las sondas o cebadores específicos de diana descritos en el presente documento no se hibridarán con las secuencias de ruido de fondo o contaminantes.

Vector. Tal como se usa en el presente documento, un “vector” es un ADN de plásmido, cósmido, fagémido o fago u otra molécula de ADN que puede replicarse de manera autónoma en una célula huésped, y que se caracteriza por uno o un número pequeño de sitios de reconocimiento de endonucleasa de restricción en los que tales secuencias de ADN pueden cortarse de una manera determinada sin pérdida de una función biológica esencial del vector, y en los que puede insertarse ADN para provocar su replicación y clonación. El vector también puede contener un marcador adecuado para su uso en la identificación de células transformadas con el vector. Los marcadores, por ejemplo, incluyen, pero no se limitan a, resistencia a tetraciclina o resistencia a ampicilina.

Cebador. Tal como se usa en el presente documento, “cebador” se refiere a un oligonucleótido de cadena sencilla que se extiende mediante unión covalente de monómeros de nucleótido durante la amplificación o polimerización de una molécula de ADN.

Sonda. Tal como se usa en el presente documento, “sonda” se refiere a un oligonucleótido de cadena sencilla que puede usarse para hibridar y/o aislar una o más moléculas de ácido nucleico de interés. Tales sondas pueden o no comprender uno o más haptenos.

Molde. El término “molde” tal como se usa en el presente documento se refiere a moléculas de ácido nucleico de cadena doble o de cadena sencilla que van a amplificarse, sintetizarse o secuenciarse. En el caso de moléculas de cadena doble, la desnaturalización de sus cadenas para formar una primera y segunda cadena se realiza preferiblemente antes de que estas moléculas puedan amplificarse, sintetizarse o secuenciarse, o la molécula de cadena doble molécula puede usarse directamente como molde. Para moldes de cadena sencilla, un cebador, complementario a una parte del molde, se hibrida en condiciones apropiadas y entonces una o más polimerasas pueden sintetizar una molécula de ácido nucleico complementaria a todo o a una parte de dicho molde. Alternativamente, para los moldes de cadena doble, pueden usarse uno o más promotores (por ejemplo promotor) en combinación con una o más polimerasas para crear moléculas de ácido nucleico complementarias a todo o a una parte del molde. Las moléculas recientemente sintetizadas, según la invención, pueden ser iguales o más cortas en longitud que el molde original.

Que incorpora. El término “que incorpora” tal como se usa en el presente documento significa formar parte de un cebador o molécula de ADN y/o ARN.

Amplificación. Tal como se usa en el presente documento “amplificación” se refiere a cualquier método *in vitro* para aumentar el número de copias de una secuencia de nucleótidos con el uso de una polimerasa. La amplificación de ácido nucleico da como resultado la incorporación de nucleótidos en un cebador o molécula de ADN y/o ARN formando así una molécula nueva complementaria a un molde. La molécula de ácido nucleico formada y su molde pueden usarse como moldes para sintetizar moléculas de ácido nucleico adicionales. Tal como se usa en el presente documento, una reacción de amplificación puede consistir en muchas rondas de replicación. Las reacciones de amplificación de ADN incluyen, por ejemplo, reacciones en cadena de la polimerasa (PCR). Una reacción de PCR puede consistir en de 5 a 100 “ciclos” de desnaturalización y síntesis de una molécula de ADN.

Oligonucleótido. “Oligonucleótido” se refiere a una molécula sintética o natural que comprende secuencias de nucleótidos unidas covalentemente que están unidas mediante un enlace fosfodiéster entre la posición 3’ de la desoxirribosa o ribosa de un nucleótido y la posición 5’ de la desoxirribosa o ribosa del nucleótido adyacente. Un oligonucleótido bloqueante se refiere a oligonucleótidos que se usan para evitar la hibridación de moléculas de ácido nucleico (por ejemplo sonda o cebador) con moléculas no queridas o no deseadas. Por ejemplo, el oligonucleótido bloqueante puede que evitar las secuencias de extremo 5’ y en algunos casos 3’ de los componentes controladores se hibriden con el vector de la biblioteca.

Nucleótido. Tal como se usa en el presente documento “nucleótido” se refiere a una combinación de base-azúcar-fosfato. Los nucleótidos son unidades monoméricas de una secuencia de ácido nucleico (ADN y ARN). El término nucleótido incluye ribonucleósido trifosfato ATP, UTP, CTP, GTP y desoxirribonucleósidos trifosfato tales como dATP, dCTP, dITP, dUTP, dGTP, dTTP, o derivados de los mismos. Tales derivados incluyen, por ejemplo, [α S]dATP, 7-desaza-dGTP y 7-desaza-dATP, y derivados de nucleótido que confieren resistencia a nucleasa sobre la molécula de ácido nucleico que los contiene. El término nucleótido tal como se usa en el presente documento también se refiere a dideoxirribonucleósidos trifosfato (ddNTPs) y sus derivados. Ejemplos ilustrados de dideoxirribonucleósidos trifosfato incluyen, pero no se limitan a, ddATP, ddCTP, ddGTP, ddITP, y ddTTP. Según la presente invención, un “nucleótido” puede no estar marcado con etiqueta o estar marcado con etiqueta detectable mediante técnicas bien conocidas. Etiquetas detectables incluyen, por ejemplo, isótopos radiactivos, etiquetas fluorescentes, etiquetas quimioluminiscentes, etiquetas bioluminiscentes y etiquetas de enzima.

Hibridación. Los términos “hibridación” e “hibridante” se refieren a apareamiento de bases de dos moléculas de ácido nucleico de cadena sencilla complementarias (ARN y/o ADN) para dar una molécula de cadena doble. Tal como se usa en el presente documento, pueden hibridarse dos moléculas de ácido nucleico, aunque el apareamiento de bases no es completamente complementario. Consecuentemente, bases con apareamiento erróneo no evitan la hibridación de dos moléculas de ácido nucleico siempre que se usen condiciones apropiadas, bien conocidas en la técnica. En la presente invención, el término “hibridación” se particularmente refiere a hibridación de un controlador con la biblioteca que va a normalizarse. Otros términos usados en los campos de tecnología de ADN recombinante y biología celular y molecular tal como se usan en el presente documento se entenderán generalmente por un experto en la técnica aplicable.

Resumen

La presente invención se refiere generalmente a métodos para producir bibliotecas de ácido nucleico normalizadas, y a bibliotecas normalizadas producidas por estos métodos. En una realización preferida de la invención, la biblioteca normalizada producida es una biblioteca de ADNc, que puede ser de cadena sencilla o de cadena doble. Según la invención, la normalización de una biblioteca de ácido nucleico se logra usando moléculas de ácido nucleico haptenilado (es decir, moléculas de ácido nucleico que tienen acopladas covalentemente a las mismas una o más moléculas de hapteno, tal como las descritas a continuación) que se hibridarán más rápidamente con las moléculas de ácido nucleico más sumamente abundantes de la biblioteca. Tales moléculas de ácido nucleico haptenilado se denominan controlador. Esta hibridación forma complejos de moléculas de ácido nucleico que pueden entonces eliminarse (reduciendo de esta manera la abundancia de moléculas de ácido nucleico unidas en la biblioteca), preferiblemente por medio de interacciones ligando-hapteno o mediante técnicas de extracción. Se ha descubierto que, mediante los métodos de la invención, pueden producirse bibliotecas de ácido nucleico normalizadas que tienen una variación máxima en la abundancia de moléculas de ácido nucleico miembro no superior a de aproximadamente 2 a aproximadamente 10 veces. Además, los métodos de la invención proporcionan bibliotecas normalizadas que tiene ruido de fondo significativamente reducido. Por tanto, la invención proporciona métodos para producir bibliotecas de ácido nucleico, particularmente bibliotecas de ADNc, de las que cada molécula de ácido nucleico miembro puede aislarse con probabilidad aproximadamente equivalente, independientemente de su número de copia en la biblioteca original.

Fuentes de bibliotecas de ácido nucleico

Usando los métodos de la invención, pueden prepararse bibliotecas de ácido nucleico normalizadas, particularmente bibliotecas de ADNc normalizadas, a partir de una variedad de bibliotecas de de ácido nucleico. Tales bibliotecas que van a normalizarse pueden prepararse usando técnicas convencionales o pueden obtenerse comercialmente (Life Technologies, Inc., Rockville, MD). Bibliotecas de ácido nucleico para su uso en la presente invención incluyen aquellas que comprenden poblaciones de moléculas de ácido nucleico de cadena sencilla o cadena doble, o preferiblemente poblaciones de moléculas de ADN de cadena sencilla o de cadena doble. Bibliotecas de ácido nucleico más preferidas que van a normalizarse según la invención incluyen aquellas que comprenden bibliotecas de ADN (ADNc) complementario. Tales bibliotecas de ADNc (de cadena doble o de cadena sencilla) pueden prepararse usando técnicas bien conocidas que usan ARN mensajero o ARN poliA+ o pueden obtenerse comercialmente, por ejemplo de Life Technologies, Inc. (Rockville, Maryland), u otras fuentes comerciales que serán familiares para un experto en la técnica. Bibliotecas de ADNc usadas según la invención se preparan preferiblemente con transcriptasa inversa con actividad ARNasa H sustancialmente reducida (véase más adelante). Se prefieren los vectores pCMVSPORT para la construcción de biblioteca y las bibliotecas de ADNc de Life Technologies, Inc. (Rockville, MD) se almacenan en estos vectores. En un aspecto preferido de la invención, las moléculas de ácido nucleico de la biblioteca pueden estar contenidas en uno o más vectores, tales como plásmidos, cósmidos o fagos.

Según la invención, pueden prepararse las bibliotecas de ácido nucleico a partir de poblaciones de moléculas de ácido nucleico obtenidas a partir de fuentes naturales, tales como una variedad de células, tejidos, órganos u organismos. Las células que pueden usarse como fuentes de moléculas de ácido nucleico pueden ser procariontes (células bacterianas, incluyendo las de la especie del género *Escherichia*, *Bacillus*, *Serratia*, *Salmonella*, *Staphylococcus*, *Streptococcus*, *Clostridium*, *Chlamydia*, *Neisseria*, *Treponema*, *Mycoplasma*, *Borrelia*, *Legionella*, *Pseudomonas*, *Mycobacterium*, *Helicobacter*, *Erwinia*, *Agrobacterium*, *Rhizobium*, y *Streptomyces*) o eucariotas (incluyendo hongos (especialmente levaduras), plantas, protozoos y otros parásitos), y animales incluyendo insectos (particularmente células de *Drosophila* spp.), nematodos (particularmente células de *Caenorhabditis elegans*), y mamíferos (particularmente células humanas)).

Las células somáticas de mamífero que pueden usarse como fuentes de poblaciones o bibliotecas de ácidos nucleicos incluyen células sanguíneas (reticulocitos y leucocitos), células endoteliales, células epiteliales, células neuronales (de los sistemas nervioso central o periférico), células de músculo (incluyendo miocitos y mioblastos de músculo esquelético, liso o cardíaco), células de tejido conectivo (incluyendo fibroblastos, adipocitos, condrocitos, condroblastos, osteocitos y osteoblastos) y otras células de estroma (por ejemplo, macrófagos, células dendríticas, células Schwann). Las células germinales de mamíferos (espermatoцитos y ovocitos) también pueden usarse como fuentes de ácidos nucleicos o bibliotecas para su uso en la invención, como también pueden las células progenitoras, precursoras y madre que dan lugar a las células somáticas y germinales anteriores. También son adecuados para usar como fuentes de ácido nucleico son tejidos u órganos de mamíferos tales como los derivados de fuentes de cerebro, riñón, hígado, páncreas, sangre, médula ósea, músculo, nervios, piel, tejido genitourinario, circulatorio, linfático, gastrointestinal y conjuntivo, así como los derivados de un embrión o feto de mamífero (incluyendo seres humanos).

Cualquiera de las células procariotas o eucariotas, tejidos y órganos anteriores puede ser normales, enfermos, transformados, establecidos, precursores, fetales o embrionarios. Células enfermas pueden incluir, por ejemplo, las implicadas en enfermedades infecciosas (causadas por bacterias, hongos o levaduras, virus (incluyendo VIH) o parásitos), en patologías genéticas o bioquímicas (por ejemplo, fibrosis quística, hemofilia, enfermedad de Alzheimer, distrofia muscular o esclerosis múltiple) o en procesos cancerosos. Líneas celulares animales transformadas o establecidas pueden incluir, por ejemplo, células COS, células CHO, células VERO, células BHK, células HeLa, células HepG2, células K562, células F9 y similares. Otras células, líneas celulares, tejidos, órganos y organismos adecuados como fuentes de ácidos nucleicos para su uso en la presente invención resultarán evidentes para un experto en la técnica. Estas células, tejidos, órganos y organismos pueden obtenerse a partir de sus fuentes naturales o pueden obtenerse comercialmente a partir de fuentes tales como Colección Americana de Cultivos Tipo (Rockville, Maryland) y otras que son conocidas para el experto en la técnica.

Una vez se obtienen las células, tejidos, órganos u otras muestras de partida, pueden aislarse moléculas de ácido nucleico (tales como ARNm o ARN poliA+), y prepararse bibliotecas de ácido nucleico (tales como bibliotecas de ADNc) a partir de las mismas, mediante métodos que se conocen bien en la técnica (véase, por ejemplo, Maniatis, T., *et al.*, Cell 15: 687-701 (1978); Okayama, H., y Berg, P., Mol. Cell. Biol. 2:161-170 (1982); Gubler, U., y Hoffman, B.J., Gene 25:263-269 (1983)). Tal como se indicó anteriormente, bibliotecas de ácido nucleico preparadas de tal manera contendrán normalmente un amplio intervalo de abundancias de moléculas de ácido nucleico miembro, dependiendo de la fuente de célula, tejido u organismo, y de la fase de desarrollo o ciclo celular de la fuente. Los métodos de la invención pueden usarse entonces para normalizar, o estrechar o reducir las abundancias relativas de moléculas de ácido nucleico en la biblioteca de ácido nucleico.

35 *Producción de bibliotecas de ácido nucleico normalizadas*

En la práctica de la invención, se normalizan las bibliotecas de ácido nucleico, para producir bibliotecas de ácido nucleico normalizadas, mediante métodos que pueden comprender una o más etapas. Un método preferido de la invención puede comprender, por ejemplo:

- (a) sintetizar una o más moléculas de ácido nucleico complementarias a todas o una parte de las moléculas de ácido nucleico de la biblioteca, en las que las moléculas de ácido nucleico sintetizadas comprenden al menos un hapteno, produciendo de esta manera moléculas de ácido nucleico haptenilado (por ejemplo controlador);
- (b) incubar una biblioteca de ácido nucleico que va a normalizarse con las moléculas de ácido nucleico haptenilado en condiciones que favorecen la hibridación de las moléculas más sumamente abundantes de la biblioteca con las moléculas de ácido nucleico haptenilado; y
- (c) eliminar las moléculas hibridadas, produciendo de esta manera una biblioteca normalizada.

Según la invención, pueden producirse moléculas de ácido nucleico haptenilado complementarias a todas o una parte de moléculas de ácido nucleico de la biblioteca, por ejemplo, incubando las moléculas de ácido nucleico de la biblioteca con al menos un polipéptido que tiene actividad ácido nucleico polimerasa y con al menos un nucleótido que comprende al menos un hapteno. Si se usan uno o más cebadores para la síntesis, los cebadores pueden comprender uno o más haptenos para producir las moléculas de ácido nucleico haptenilado (sin o con el uso de nucleótidos haptenilados durante la síntesis). Polipéptidos preferidos que tienen actividad ácido nucleico polimerasa para su uso en este aspecto de la invención incluyen aquellas que tienen actividad transcriptasa inversa y aquellas que tienen actividad ADN polimerasa o ARN polimerasa.

Polipéptidos preferidos que tienen actividad transcriptasa inversa (es decir, aquellos polipéptidos que pueden catalizar la síntesis de una molécula de ADN a partir de un molde de ARN) incluyen, pero no se limitan a, transcriptasa inversa del virus de leucemia murina Moloney (VLM-M), transcriptasa inversa del virus de sarcoma de Rous (VSR), transcriptasa inversa del virus de mieloblastosis aviar (VMA), transcriptasa inversa del virus asociado de Rous (VAR), transcriptasa inversa del virus asociado de mieloblastosis (VAM), transcriptasa inversa del virus de inmunodeficiencia humano (VIH), transcriptasa inversa retroviral, transcriptasa inversa de retrotransposón, transcriptasa inversa de la hepatitis B, transcriptasa inversa del virus de mosaico de la coliflor y transcriptasa inversa bacteriana. Se prefieren particularmente aquellos polipéptidos que tienen actividad transcriptasa inversa que también tienen actividad ARNasa

H sustancialmente reducida (es decir, polipéptidos "ARNasa H-"). Por un polipéptido que tiene "actividad ARNasa H sustancialmente reducida" se quiere decir que el polipéptido tiene menos de aproximadamente el 20%, más preferiblemente menos de aproximadamente el 15%, el 10% o el 5%, y lo más preferiblemente menos de aproximadamente 2%, de actividad ARNasa H de una enzima de tipo natural o ARNasa H+ tal como una transcriptasa inversa de VLM-M de tipo natural. La actividad ARNasa H puede determinarse mediante una variedad de ensayos, tales como los descritos, por ejemplo, en la patente de los EE.UU. N° 5.244.797, en Kotewicz, M.L., *et al.*, Nucl. Acids Res. 16: 265 (1988) y en Gerard, G.F., *et al.*, FOCUS 14 (5):91 (1992), las descripciones de todos los cuales se incorporan en el presente documento como referencia. Polipéptidos ARNasa H- adecuados para su uso en la presente invención incluyen, pero no se limitan a, transcriptasa inversa H- de VLM-M, transcriptasa inversa H- de VSR, transcriptasa inversa H- de VAM, transcriptasa inversa H- de VAR, transcriptasa inversa H- de VAM, transcriptasa inversa H- de VIH, y transcriptasa inversa SUPERSCRIPT™ I y transcriptasa inversa SUPERSCRIPT™ II que están disponibles comercialmente, por ejemplo de Life Technologies, Inc. (Rockville, Maryland).

Otros polipéptidos que tienen actividad ácido nucleico polimerasa adecuados para su uso en los presentes métodos incluyen ADN polimerasas termófilas tales como ADN polimerasa I, ADN polimerasa III, fragmento de Klenow, polimerasa de T7, y polimerasa de T5, y ADN polimerasas termoestables que incluyen, pero no se limitan a, ADN polimerasa de *Thermus thermophilus* (*Tth*), ADN polimerasa de *Thermus aquaticus* (*Taq*), ADN polimerasa de *Thermotoga neopolitana* (*Tne*), ADN polimerasa de *Thermotoga maritima* (*Tma*), ADN polimerasa de *Thermococcus litoralis* (*Tli* o VENT®), ADN polimerasa de *Pyrococcus furiosus* (*Pfu* o DEEPVENT®), ADN polimerasa de *Pyrococcus woosii* (*Pwo*), ADN polimerasa de *Bacillus sterothophilus* (*Bst*), ADN polimerasa de *Sulfolobus acidocaldarius* (*Sac*), ADN polimerasa de *Thermoplasma acidophilum* (*Tac*), ADN polimerasa de *Thermus flavus* (*Tf/Tub*), ADN polimerasa de *Thermus ruber* (*Tru*), ADN polimerasa de *Thermus brockianus* (DYNAZYME®), ADN polimerasa de *Methanobacterium thermoautotrophicum* (*Mth*), y mutantes, variantes y derivados de las mismas.

Las ARN polimerasas usadas preferiblemente en la invención pueden incluir ARN polimerasa de SP6, ARN polimerasa de T7, ARN polimerasa de T3 y similares. Con el uso de ARN polimerasas, se usan normalmente uno o más promotores (por ejemplo promotor SP6, promotor T7, etc.). Por ejemplo, las moléculas de ADN de cadena doble (o biblioteca de cadena doble) que contienen uno o más promotores se usan en combinación con una o más ARN polimerasas para crear moléculas de ARN haptenilado complementarias a toda o una parte del molde de biblioteca de cadena doble. Preferiblemente, tales moléculas de ARN se encuentran en gran exceso molar comparado con los moldes. Según la invención, tales promotores pueden proporcionarse por el vector en el que se clonan las moléculas de biblioteca o mediante moléculas adaptadoras (por ejemplo oligonucleótidos de cadena doble) que se añaden a las moléculas de biblioteca. Cuando se usan tales moléculas adaptadoras, los adaptadores, (que comprenden preferiblemente uno o más promotores) se añaden a las moléculas de biblioteca. Preferiblemente, las moléculas de biblioteca son moléculas lineales de cadena doble (por ejemplo ADNc lineal de cadena doble producido tras la primera y segunda síntesis), y los adaptadores pueden añadirse usando técnicas convencionales (por ejemplo ligasas) a uno o ambos extremos de tales moléculas.

Nucleótidos preferidos para su uso en los métodos de la presente invención incluyen, pero no se limitan a, ribonucleósido trifosfatos tales como ATP, UTP, CTP, GTP y derivados de los mismos, y desoxirribonucleótido trifosfatos tales como dATP, dCTP, dTTP, dUTP, dGTP, dTTP, o derivados de los mismos. Tales derivados incluyen [α S]dATP, 7-desaza-dGTP y 7-desaza-dATP, o los ribonucleósido trifosfatos correspondientes en los que la desoxirribosa se ha sustituido por ribosa. Según la invención, los nucleótidos o derivados de los mismos preferiblemente comprenden una o más moléculas de hapteno unidas covalentemente a los mismos.

Moléculas de hapteno preferidas para su uso en estos métodos incluyen, sin limitación: (i) biotina; (ii) un anticuerpo; (iii) una enzima; (iv) lipopolisacárido; (v) apotransferrina; (vi) ferrotransferrina; (vii) insulina; (viii) citocinas (factores de crecimiento, interleucinas o factores estimulantes de colonia); (ix) gp120; (x) β -actina; (xi) LFA-1; (xii) Mac-1; (xiii) glicoforina; (xiv) laminina; (xv) colágeno; (xvi) fibronectina; (xvii) vitronectina; (xviii) integrinas $\alpha_v\beta_1$ y $\alpha_v\beta_3$; (xix) integrinas $\alpha_3\beta_1$, $\alpha_4\beta_1$, $\alpha_4\beta_7$, $\alpha_5\beta_1$, $\alpha_v\beta_1$, $\alpha_{10}\beta_3$, $\alpha_v\beta_3$ y $\alpha_v\beta_6$; (xx) integrinas $\alpha_1\beta_1$, $\alpha_2\beta_1$, $\alpha_3\beta_1$ y $\alpha_v\beta_3$; (xxi) integrinas $\alpha_1\beta_1$, $\alpha_2\beta_1$, $\alpha_3\beta_1$, $\alpha_6\beta_1$, $\alpha_7\beta_1$ y $\alpha_6\beta_5$; (xxii) anquirina; (xxiii) C3bi, fibrinógeno o factor X; (xxiv) ICAM-1 o ICAM-2; (xxv) espectrina o fodrina; (xxvi) CD4; (xxvii) un receptor de citocina (por ejemplo, factor de crecimiento, interleucina, o factor estimulante de colonia); (xxviii) un receptor de insulina; (xxix) un receptor de transferrina; (xxx) Fe⁺⁺⁺; (xxxi) polimixina B o proteína neutralizante de endotoxina (ENP); (xxxii) un sustrato específico de enzima; (xxxiii) proteína A, proteína G, un receptor Fc de superficie celular o un antígeno específico de anticuerpo; (xxxiv) avidina y estreptavidina; y combinaciones de los mismas. Un hapteno particularmente preferido para su uso en los métodos de la invención es biotina. Las moléculas de ácido nucleico haptenilado, en las que una o más moléculas de hapteno están unidas (preferiblemente de manera covalente) a uno o más nucleótidos de la molécula de ácido nucleico, pueden producirse usando métodos de síntesis orgánica convencionales que serán familiares para el experto en la técnica. Por ejemplo, la molécula de ácido nucleico puede biotinilarse en el extremo 5' produciendo en primer lugar grupos amino (NH₂) en 5' seguido de la adición de éster Cab-NHS (Langer, P.R., *et al.*, Proc. Natl. Acad. Sci. USA 78: 6633 (1981)). En un aspecto particularmente preferido de la invención, se prepara una molécula de ácido nucleico haptenilado que puede ser una molécula de ARN o una molécula de ADN, que comprende una o más, dos o más, tres o más, o cuatro o más moléculas de hapteno, lo más preferiblemente moléculas de biotina.

Una vez se han producido moléculas de ácido nucleico haptenilado que son complementarias a las moléculas de ácido nucleico de la biblioteca, se usan para normalizar la biblioteca de ácido nucleico mediante hibridación. Espe-

cíficamente, la biblioteca de ácido nucleico que va a normalizarse se incuba preferiblemente con un exceso molar de la población de moléculas de ácido nucleico haptenilado (por ejemplo superior o igual a 10 veces o preferiblemente superior o igual a 20 veces el exceso molar), preparado tal como se describió anteriormente, en condiciones que favorecen la hibridación más rápida de las moléculas de ácido nucleico haptenilado con las moléculas de ácido nucleico más sumamente abundantes y se hibridan menos rápidamente con las moléculas de ácido nucleico menos abundantes presentes en la biblioteca. Tales condiciones que favorecen la hibridación pueden comprender, por ejemplo, incubar la biblioteca que va a normalizarse con las moléculas de ácido nucleico haptenilado en un intervalo de Cot. Cot es el producto de la concentración inicial de ácido nucleico (moles de nucleótido por litro, Co) y tiempo (segundos, t). La Cot se obtiene convirtiendo la concentración de nucleótidos que reaccionan y el tiempo de hibridación en unidades convencionales ($\text{mol}\cdot\text{s}^{-1}$ o M·s). Tal como se describe en detalle en los ejemplos a continuación, Cot particularmente preferidas para su uso en los presentes métodos incluyen, pero no se limitan a: una Cot igual a o superior a 25; una Cot igual a o superior a 50; una Cot igual a o superior a 100, una Cot igual a o superior a 200; una Cot igual a o superior a 250; una Cot igual a o superior a 500; una Cot igual a o superior a 1000; y una Cot inferior a aproximadamente 10.000. Alternativamente, pueden usarse las condiciones de hibridación que consisten en un intervalo de Cot, incluyendo una Cot de desde aproximadamente 10 hasta aproximadamente 10.000: una Cot de desde aproximadamente 25 hasta aproximadamente 10.000; una Cot de desde aproximadamente 50 hasta aproximadamente 10.000; una Cot de desde aproximadamente 100 hasta aproximadamente 10.000; una Cot de desde aproximadamente 200 hasta aproximadamente 10.000; una Cot de desde aproximadamente 250 hasta aproximadamente 10.000; y una Cot de desde aproximadamente 500 hasta aproximadamente 10.000. Otras condiciones de hibridación adecuadas para su uso con los presentes métodos resultarán evidentes para un experto en la técnica y pueden determinarse con sólo experimentación rutinaria.

En estas condiciones, las moléculas de ácido nucleico haptenilado se hibridan más rápidamente con las moléculas de ácido nucleico más sumamente abundantes presentes en la biblioteca y menos rápidamente con los miembros menos abundantes. Los complejos de hibridación formados entre la biblioteca y las moléculas de ácido nucleico haptenilado pueden eliminarse entonces mediante una variedad de métodos, dando como resultado la reducción en número de copia de las moléculas de ácido nucleico sumamente abundantes en la biblioteca y produciendo de esta manera una biblioteca de ácido nucleico normalizada.

Según la invención, la eliminación de complejos se logra mediante interacciones ligando-hapteno usando un ligando que se une específicamente al hapteno que está unido a las moléculas de ácido nucleico haptenilado. En un método preferido de este tipo, el ligando puede estar unido, preferiblemente de manera covalente, a un soporte sólido tal como nitrocelulosa, diazocelulosa, vidrio, poliestireno (incluyendo placas de microtitulación), poli(cloruro de vinilo), polipropileno, polietileno, dextrano, Sepharose, agar, almidón, nylon, o perlas, que pueden ser perlas de látex, perlas magnéticas, perlas paramagnéticas, superparamagnéticas o perlas de vidrio. Soportes sólidos particularmente preferidos son perlas magnéticas, perlas paramagnéticas, y perlas superparamagnéticas, que están disponibles comercialmente, por ejemplo de Life Technologies, Inc. (Rockville, MD), Dynal A.S. (Oslo, Noruega), o de Sigma (St. Louis, Missouri).

Acoplados a estos soportes sólidos puede estar cualquier ligando que pueda unir el hapteno usado para haptenar las moléculas de ácido nucleico. Ejemplos de ligandos adecuados para su uso en los presentes métodos (que corresponden en orden a las moléculas de hapteno enumeradas anteriormente) incluyen sin limitación: (i) avidina y estreptavidina; (ii) proteína A, proteína G, un receptor Fc de superficie celular, o un antígeno específico de anticuerpo; (iii) un sustrato específico de enzima; (iv) polimixina B o proteína neutralizante de endotoxina (ENP); (v) Fe^{+++} ; (vi) un receptor de transferrina; (vii) un receptor de insulina; (viii) un receptor de citocina (por ejemplo, factor de crecimiento, interleucina o factor estimulante de colonia); (ix) CD4; (x) espectrina o fodrina; (xi) ICAM-1 o ICAM-2; (xii) C3bi, fibrinógeno o factor X; (xiii) anquirina; (xiv) integrinas $\alpha_1\beta_1$, $\alpha_2\beta_1$, $\alpha_3\beta_1$, $\alpha_6\beta_1$, $\alpha_7\beta_1$ y $\alpha_6\beta_5$; (xv) integrinas $\alpha_1\beta_1$, $\alpha_2\beta_1$, $\alpha_3\beta_1$ y $\alpha_v\beta_3$; (xvi) integrinas $\alpha_3\beta_1$, $\alpha_4\beta_1$, $\alpha_4\beta_7$, $\alpha_5\beta_1$, $\alpha_v\beta_1$, $\alpha_{\text{IIB}}\beta_3$, $\alpha_v\beta_3$ y $\alpha_v\beta_6$; (xvii) integrinas $\alpha_v\beta_1$ y $\alpha_v\beta_3$; (xviii) vitronectina; (xix) fibronectina; (xx) colágeno; (xxi) laminina; (xxii) glicoforina; (xxiii) Mac-1; (xxiv) LFA-1; (xxv) β -actina; (xxvi) gp 120; (xxvii) citocinas (factores de crecimiento, interleucinas o factores estimulantes de colonia); (xxviii) insulina; (xxix) ferrotansferrina; (xxx) apotransferrina; (xxxi) lipopolisacárido; (xxxii) una enzima; (xxxiii) un anticuerpo; (xxxiv) biotina; y combinaciones de los mismos. Ligandos preferidos incluyen avidina y estreptavidina. Por supuesto, la elección del ligando dependerá de la elección del hapteno usado para la producción de molécula de ácido nucleico haptenilado; ligandos apropiados para usar en los métodos de la invención serán por tanto evidentes para un experto de la técnica. La unión de la(s) molécula(s) de ligando al soporte sólido puede realizarse mediante cualquier método de acoplamiento de ligando tal como acoplamiento covalente, hidrófobo o iónico (incluyendo recubrimiento) que resultarán familiares para un experto en la técnica. Por ejemplo, en un aspecto preferido de la invención en el que las moléculas de ácido nucleico haptenilado comprenden biotina, un ligando de unión a biotina tal como avidina o estreptavidina puede estar unido al soporte sólido. En un aspecto particularmente preferido, el soporte sólido usado son perlas magnéticas, paramagnéticas o superparamagnéticas acopladas con avidina o estreptavidina.

Normalmente, las condiciones que favorecen las interacciones ligando-hapteno incluyen la incubación en una solución salina tamponada, preferiblemente una disolución de cloruro de sodio tamponada con TRIS, fosfato HEPES o carbonato, más preferiblemente una disolución de cloruro de sodio tamponada con TRIS, todavía más preferiblemente una disolución que comprende TRIS-HCl aproximadamente 10-100 mM y NaCl aproximadamente 300-2000 mM, y lo más preferiblemente una disolución que comprende TRIS-HCl aproximadamente 10 mM y NaCl aproximadamente 1 M, a un pH de aproximadamente 6-9, más preferiblemente un pH de aproximadamente 7-8, todavía más preferible-

mente un pH de aproximadamente 7,2-7,6, y lo más preferiblemente un pH de aproximadamente 7,5. La incubación se realiza preferiblemente a de 0°C hasta aproximadamente 25°C, y lo más preferiblemente a aproximadamente 25°C, durante aproximadamente 30-120 minutos, de manera preferible aproximadamente 45-90 minutos, y de la manera más preferible aproximadamente 60 minutos, para permitir la unión de moléculas de ácido nucleico haptencilado (y por tanto las moléculas de ácido nucleico de biblioteca complementaria a las que están hibridadas) al soporte sólido acoplado a ligando.

Una vez los complejos haptencilados se han unido al soporte en fase sólida, la biblioteca de ácido nucleico normalizado, que comprende moléculas de ácido nucleico de un menor intervalo de abundancias que la biblioteca de entrada, puede recogerse a partir de los sobrenadantes o eluatos (es decir, los materiales no unidos en disolución). Por ejemplo, en un aspecto preferido en el que las moléculas de ácido nucleico están unidas a avidina o estreptavidina; o una fase sólida acoplada a avidina o estreptavidina, las moléculas de ácido nucleico que comprenden la biblioteca de ácido nucleico normalizada, tal como una biblioteca de ADNc normalizada, pueden obtenerse aspirando con cuidado y recogiendo los sobrenadantes. En un aspecto particularmente preferido en el que se usan perlas magnéticas, paramagnéticas o superparamagnéticas acopladas a avidina o estreptavidina como soporte sólido, las perlas que contienen ácido nucleico biotencilado pueden segregarse a partir de los sobrenadantes usando un imán (tal como un separador de partículas magnéticas Magna-Sep; Life Technologies, Inc.) y pueden retirarse los sobrenadantes usando una pipeta. La eliminación de complejos haptencilados se logra preferiblemente extrayendo con un disolvente orgánico (por ejemplo fenol, cloroformo etc.). Los enfoques anteriormente descritos dan como resultado la producción de una biblioteca de ácido nucleico normalizada, que puede ser de cadena sencilla o de cadena doble y que puede usarse inmediatamente, almacenarse hasta su uso, o tratarse y luego purificarse según la invención o mediante técnicas que se conocen bien en la bibliografía (véase, por ejemplo, Gubler, U., y Hoffman, B.J., *Gene* 25: 263-269 (1983); Krug, M.S., y Berger, S.L., *Meth. Enzymol.* 152:316-325 (1987); Sambrook, J., *et al.*, *Molecular Cloning: A Laboratory Manual*, 2ª ed., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, págs. 8.60-8.63 (1987)), y otras que serán familiares para un experto en la técnica.

Eliminación o reducción de ruido de fondo

La invención también proporciona métodos para la producción de biblioteca de ácido nucleico normalizada seleccionada con un ruido de fondo de clones no recombinantes y transpuestos muy bajo. Tal como se usa en el presente documento, una biblioteca normalizada seleccionada es una biblioteca en la que una o más moléculas de ácido nucleico específicas o conjuntos de moléculas de ácido nucleico se han enriquecido en la biblioteca normalizada y otras moléculas de ácido nucleico de menos interés se han eliminado mediante uno o varios enfoques descritos en el presente documento. Por tanto, la invención se refiere adicionalmente a la eliminación de moléculas de ácido nucleico contaminantes o de fondo de la biblioteca normalizada. Según la invención, tal supresión o eliminación de ácidos nucleicos contaminantes puede realizarse antes o después de la normalización. Moléculas de ácido nucleico contaminantes típicas en una biblioteca son moléculas de vector que no contienen moléculas de ácido nucleico de la biblioteca (en las que el vector no pudo recibir un inserto o el vector perdió el inserto mediante delección durante la propagación de la biblioteca fuente).

Según la invención, pueden usarse sondas específicas de diana (por ejemplo oligodA-*NotI*) en varios métodos para reducir o eliminar ácidos nucleicos contaminados de la biblioteca de interés. Tales sondas son específicas de diana porque reconocen y se hibridan con moléculas de las moléculas de biblioteca pero no con secuencias de ácido nucleico contaminantes (tales como vectores sin insertos de biblioteca). Uno de tales medios implica usar una o más sondas específicas de diana haptenciladas para capturar o aislar la biblioteca de interés. En tales métodos, la biblioteca normalizada es preferiblemente de cadena sencilla (o, si es de cadena doble, se convierte en cadena sencilla mediante métodos descritos en el presente documento). Mediante hibridación de las sondas haptenciladas con la biblioteca normalizada, la biblioteca normalizada hibridada puede seleccionarse separada de ácido nucleico contaminante usando, por ejemplo, interacciones hapteno/ligando o extracción. La biblioteca normalizada seleccionada de cadena sencilla resultante puede convertirse entonces en cadena doble incubando la biblioteca con uno o más polipéptidos que tienen actividad polimerasa en condiciones suficientes para sintetizar biblioteca normalizada seleccionada de cadena doble.

Alternativamente, la biblioteca normalizada se hibrida con un cebador haptencilado, específico de diana, y las moléculas pueden convertirse entonces en cadena doble incubándolas con uno o más polipéptidos que tienen actividad polimerasa en condiciones suficientes para sintetizar biblioteca normalizada de cadena doble. Para convertir tales moléculas en cadena doble, pueden usarse uno o más nucleótidos resistentes a nucleasa. Entonces pueden seleccionarse las moléculas de cadena doble separadas de las moléculas de ácido nucleico contaminantes usando, por ejemplo, interacciones hapteno/ligando o extracción.

En ambos casos, la biblioteca normalizada seleccionada de cadena doble resultante de la invención puede transformarse entonces en una o más células huésped en una etapa de selección adicional. Según la invención, se transforman moléculas de cadena sencilla con una frecuencia muy baja mientras que las moléculas de cadena doble se transforman con una frecuencia muy alta. Por tanto, la transformación permite una etapa de selección adicional en la que se eliminan o suprimen moléculas contaminantes de cadena sencilla. Por ejemplo, cuando se usa un cebador o sonda específica de diana en la etapa de síntesis de cadena doble, no se ceban ácidos nucleicos no específicos y por tanto no se convierten en cadena doble y no estarán presentes en la biblioteca normalizada seleccionada.

En otro aspecto de la invención, se convierte una biblioteca normalizada seleccionada de cadena sencilla seleccionada con sondas hapteniladas en cadena doble con cebadores (preferiblemente cebadores específicos de diana) y uno o más nucleótidos que confieren resistencia a la nucleasa a la molécula de cadena doble sintetizada. La digestión con tal nucleasa de este tipo permite la eliminación de moléculas de cadena sencilla que no se han convertido en cadena doble por los cebadores. Tales moléculas de cadena doble pueden transformarse entonces en una o más células huésped como una etapa de selección adicional.

Aún en otro aspecto, la biblioteca normalizada seleccionada puede prepararse incubando la biblioteca normalizada de cadena sencilla con uno o más cebadores específicos de diana que no están haptenilados en combinación con uno o más nucleótidos que confieren resistencia a la nucleasa. La digestión de la mezcla prevé la selección de las moléculas de ácido nucleico deseadas y, como etapa de selección adicional, las moléculas de cadena doble resultantes pueden transformarse en una o más células huésped.

Según la invención, pueden prepararse moléculas de cadena sencilla a partir de cadena doble tratando moléculas de cadena doble en condiciones suficientes para convertirlas en cadena sencilla. Tales condiciones pueden comprender, por ejemplo, degradación de una cadena de las moléculas de ácido nucleico de cadena doble en la biblioteca, tal como usando una endonucleasa, una exonucleasa, y similares, y preferiblemente usando proteína de gen II y exonucleasa III (disponible de Life Technologies, Inc., Rockville, MD). Alternativamente, tales condiciones pueden comprender desnaturalizar las moléculas de cadena doble con calor, condiciones iónicas, pH (por ejemplo base) y similares.

Nucleótidos que confieren resistencia a la nucleasa usados según la invención son preferiblemente análogos de nucleótido. Tales análogos de nucleótido incluyen, pero no se limitan a, nucleótidos metilados tales como 5-metildesoxicitosina, 3-metildesoxiadenosina, 7-metilguanina y similares. Los expertos en la técnica reconocerán otros análogos de nucleótido que inhiben o bloquean exonucleasas o endonucleasas de restricción (nucleasas). También se conocen en la técnica combinaciones de análogos de nucleótido y enzimas adecuadas que pueden usarse según la invención (véase Life Technologies 1997-1998 Catalog y Reference Guide, Capítulo 6).

Kits

La presente invención también proporciona kits para su uso en la producción y aislamiento de bibliotecas normalizadas seleccionadas y normalizadas. Los kits según este aspecto de la invención comprenden un medio transportador, tal como una caja, caja de cartón, tubo o similar, que tiene en confinamiento cerrado en el mismo uno o más recipientes, tales como viales, tubos, ampollas, botellas y similares. El kit de la invención puede comprender el controlador para normalizar una biblioteca o los componentes necesarios para preparar el controlador usado para normalizar una biblioteca (por ejemplo, una o más polimerasas, uno o más adaptadores que comprenden promotores, uno o más vectores que comprenden promotores, uno o más nucleótidos haptenilados y/o uno o más cebadores o sondas haptenilados). Tales kits pueden comprender una o más sondas o cebadores específicos de diana (que están haptenilados o no). En aspectos adicionales, los kits de la invención pueden comprender uno o más nucleótidos (por ejemplo, nucleótidos que confieren resistencia a la nucleasa y/o una o más endonucleasas, exonucleasas o enzimas de restricción, tales como proteína de gen II o exonucleasa III o *HhaI*, usadas para la digestión de moléculas de ácido nucleico).

Kits adicionales proporcionados por la invención comprenden uno o más recipientes que contienen una o más de las bibliotecas de ácido nucleico normalizadas o bibliotecas de ácido nucleico normalizadas seleccionadas anteriormente descritas de la invención. Las bibliotecas en estos kits de la invención pueden ser de cadena sencilla o de cadena doble, y son preferiblemente bibliotecas de ADNc.

Los kits abarcados por este aspecto de la presente invención pueden comprender además uno o más reactivos adicionales (por ejemplo, tampones adecuados) y compuestos necesarios para usar las bibliotecas normalizadas y bibliotecas normalizadas seleccionadas de la invención.

Usos

La presente invención puede usarse en una variedad de aplicaciones que requieren la producción rápida y aislamiento de bibliotecas de ácido nucleico normalizadas y normalizadas seleccionadas, particularmente bibliotecas de ADNc. El uso principal para tales bibliotecas es para el descubrimiento de genes y para preparar bases de datos de genes. Las bibliotecas preparadas por los métodos de la invención pueden usarse como fuentes de moléculas de ácido nucleico molde para reacciones de amplificación (tal como mediante PCR), para identificar rápidamente y/o clonar moléculas de ácido nucleico de bajo número de copias, y para producir polipéptidos mediante técnicas de ingeniería genética.

Por tanto, la invención también se refiere a métodos para la amplificación de una molécula de ácido nucleico, y a moléculas de ácido nucleico amplificadas mediante estos métodos. Según este aspecto de la invención, una molécula de ácido nucleico puede amplificarse (es decir, prepararse copias adicionales de la molécula de ácido nucleico) amplificando una molécula de ácido nucleico (por ejemplo, una molécula de ADNc) contenida en una biblioteca normalizada o biblioteca normalizada seleccionada de la invención según cualquier método de amplificación conocido en la técnica. Métodos de amplificación particularmente preferidos según este aspecto de la invención incluyen la PCR

ES 2 281 138 T3

(patentes de los EE.UU. números 4.683.195 y 4.683.202), amplificación por desplazamiento de cadena (SDA; patente de los EE.UU. número 5.455.166; documento EP 0 684 315), y amplificación basada en secuencia de ácido nucleico (NASBA; patente de los EE.UU. número 5.409.818; documento EP 0 329 822). Los más preferidos son los métodos que comprenden una o más amplificaciones por PCR.

5 La invención también se refiere a métodos que pueden usarse para preparar vectores que comprenden las bibliotecas normalizadas o normalizadas seleccionadas de la presente invención, a las células huésped que comprenden esos vectores, a los métodos para la producción de un polipéptido recombinante usando estos vectores y células huésped, y a polipéptidos recombinantes producidos usando estos métodos. Según este aspecto de la invención, un polipéptido
10 recombinante puede producirse cultivando cualquiera de las células huésped recombinantes anteriores en condiciones que favorecen la producción de un polipéptido de las mismas, y el aislamiento del polipéptido. Métodos para cultivar células huésped recombinantes, y para la producción y aislamiento de polipéptidos de las mismas, se conocen bien por un experto en la técnica.

15 Se producen vectores según la invención mediante inserción, usando métodos que se conocen en la técnica, una o más de las moléculas de ácido nucleico de interés en un vector. El vector usado en este aspecto de la invención puede ser, por ejemplo, un plásmido, un cósmido o un fago. Se prefieren vectores que comprenden regiones de control de acción en cis con respecto al ácido nucleico que codifica para el polipéptido de interés. Pueden suministrarse factores de acción en trans apropiados por el huésped, suministrarse por un vector complementario o suministrarse por el
20 propio vector tras la introducción en el huésped.

En ciertas realizaciones preferidas, los vectores son vectores de expresión que proporcionan expresión específica de las moléculas de ácido nucleico contenidas en las bibliotecas normalizadas o bibliotecas normalizadas seleccionadas de la invención, vectores que pueden ser inducibles y/o específicos del tipo celular. Se prefieren particularmente entre
25 tales vectores los que son pueden inducirse por factores ambientales que se manipulan fácilmente, tales como la temperatura y aditivos de nutrientes.

Los vectores de expresión útiles en la presente invención incluyen vectores derivados de cromosomas, epitomas y virus, por ejemplo, vectores derivados de plásmidos bacterianos o bacteriófagos, y vectores derivados de combinaciones de los mismos, tales como cósmidos y fagémidos, e incluirán preferiblemente al menos un marcador seleccionable tal como gen de resistencia frente a tetraciclina o ampicilina para cultivar en una célula huésped bacteriana. Antes de la inserción en un vector de expresión de este tipo, las moléculas de ácido nucleico contenidas en las bibliotecas de la invención pueden unirse operativamente a un promotor apropiado, tal como el promotor de fago lambda PL, los promotores *lac*, *trp* y *tac* de *E. coli*. El experto en la técnica conocerá otros promotores adecuados, Entre los vectores preferidos para su uso en la presente invención se incluyen pQE70, pQE60 y pQE-9, disponibles de Qiagen; vectores pBS, vectores Phagescript, vectores Bluescript, pNH8A, pNH16a, pNH18A, pNH46A, disponibles de Stratagene; pcDNA3 disponible de Invitrogen; pGEX, pTrxfus, pTrc99a, pET-5, pET-9, pKK223-3, pKK233-3, pDR540, pRIT5 disponibles de Pharmacia; y pSPORT1, pSPORT2, pCMVSPORT 2.0 y pSV-SPORT1, disponibles de Life Technologies, Inc. Otros vectores adecuados resultarán fácilmente evidentes para el experto en la
40 técnica.

Células huésped representativas que pueden usarse según la invención incluyen, pero no se limitan a, células bacterianas, células de levadura, células vegetales y células animales. Células huésped bacterianas preferidas incluyen células de *Escherichia spp.* (particularmente células *E. coli* y lo más particularmente cepas DH10B y Stbl2 de *E. coli*),
45 células de *Bacillus spp.* (particularmente células *B. subtilis* y *B. megaterium*), células de *Streptomyces spp.*, células de *Erwinia spp.*, células de *Klebsiella spp.* y células de *Salmonella spp.* (particularmente células *S. typhimurium*). Células huésped animales preferidas incluyen células de insectos (lo más particularmente células de *Spodoptera frugiperda* Sf9 y Sf21 y células de *Trichoplusa High-Five*) y células de mamíferos (lo más particularmente células CHO, COS, VERO, BHK y de seres humanos). Éstas y otras células huésped adecuadas están comercialmente disponibles, por ejemplo, de Life Technologies, Inc. (Rockville, Maryland), Colección Americana de Cultivos Tipo (Rockville, Maryland) e Invitrogen (San Diego, California).

Ejemplos

55 Ejemplo 1

Producción de bibliotecas de ADNc normalizadas a partir de bibliotecas de ADNc clonadas direccionalmente

En este ejemplo se describe el procedimiento para construir una biblioteca de ADNc normalizada en el vector
60 pCMVSPORT 2.0 (figuras 1 y 2). Consiste en i) aislar ADN fagémido de una biblioteca de ADNc clonada direccionalmente, ii) convertir el ADN de la biblioteca de ADNc circular de cadena doble (cd) en a) un molde de cd lineal para la producción con ARN polimerasa de controlador de ARN biotilado y b) ADN circular de cadena sencilla (cs) usando Gen II y Exonucleasa III, iii) combinar el controlador y ADN de biblioteca circular de cs con dos oligonucleótidos bloqueantes en una hibridación de sustracción, iv) reparar el ADN circular de cs no sustraído y v) transformarlo
65 en células de *E. coli* produciendo así una biblioteca de ADNc normalizada primaria.

La producción de ADN de cs circular a partir de una biblioteca de ADN de cd circular se realiza de la siguiente manera. Digerir 10 μ g de ADNc de cd circular en 1 X tampón Gen II Tris-HCl 20 mM (pH = 8), NaCl 80 mM, MgCl₂

ES 2 281 138 T3

25 mM, β -mercaptoetanol 2 mM, glicerol al 5%, BSA 5 mg/ml con 8 μ l de Gen II a 30°C durante 40 min en un volumen final de 200 μ l. Terminar la reacción mediante incubación a 65°C durante 5 min. Añadir 12 μ l de exonucleasa III, e incubar a 37°C durante 30 min. Añadir 8 μ l (10 U/ μ l) de *NotI* e incubar la mezcla durante 1 h a 37°C. Añadir 2 μ l de exonucleasa III, y seguir incubando durante 1 hora a 37°C. Extraer dos veces con fenol/cloroformo/alcohol isoamílico (25:24:1) y precipitar en etanol. Resuspender el ADNc de cs circular en 10 μ l de TE libre de ARNasa. Se convirtió la biblioteca de ADNc de cerebro fetal (Life Technologies, Inc., Rockville, MD) en cadena sencilla mediante este procedimiento.

La producción de ADNc de cd linealizado a partir de ADNc de cd circular es tal como sigue. Digerir 50 μ g de ADNc de cd circular con 200 unidades de *NotI* (LTI) en 300 μ l de 1 X tampón de reacción [Tris-HCl 5 mM, pH 8,0; MgCl₂ 1 mM; NaCl 10 mM] durante 3 horas a 37°C. Añadir 100 unidades de *NotI*, e incubar durante otras 3 horas a 37°C. Extraer dos veces con fenol/cloroformo/alcohol isoamílico (25:24:1 v/v) y precipitar en etanol. Resuspender el ADNc de cd linealizado en 30 μ l de tampón TE libre de ARNasa. Se linealizó la biblioteca de ADNc de cerebro fetal humano (Life Technologies, Inc., Rockville, MD) de esta manera.

La producción de controlador de ARN biotinilado a partir de ADN de biblioteca de ADNc de cd circular se realiza de la siguiente manera. Preparar una mezcla de los siguientes componentes: 1,214 ml de agua tratada con DEPC, 400 μ l de 5 X tampón de transcripción [Tris-HCl 200 mM (pH 7,9), MgCl₂ 30 mM, espermidina-(HCl)₃ 10 mM], 200 μ l de mezcla de rNTP (10 μ M de cada ATP, GTP y UTP, CTP 5 μ M, biotina-14-CTP 20 μ M), 16 μ l (20 μ g) de cd linealizada de biblioteca de ADNc de cerebro fetal humano (véase anteriormente), 100 μ l de DTT 0,1 M, y 70 μ l de ARN polimerasa de SP6 (350 unidades/ μ l). Se construyó la biblioteca de ADNc de cerebro fetal humano (Life Technologies, Inc., Rockville, MD) en vector pCMV-SPORT que contiene un promotor de CMV, un promotor de polimerasa de SP6 y T7 que flanquea el sitio de clonación múltiple (MSC) para la síntesis de controlador de ARN. Mezclar e incubar a 37°C durante 13 horas. Añadir 1 ml de acetato de amonio 7,5 M y 8 ml de etanol. Enfriar sobre anhídrido carbónico sólido durante 30 min, microcentrifugar durante 25 min a 4°C y resuspender el sedimento en 1 ml de TE. Calentar la disolución a 65°C y volver a precipitar de nuevo. Lavar el sedimento en etanol al 70%, secar y resuspender en 1,92 ml de agua, 40 μ l de Tris-HCl 1 M, [pH 7,5], calentar a 65°C para la resuspensión. Añadir 20 μ l de MgCl₂ 1 M, 20 μ l de ADNasa I (2.660 unidades) al ARN resuspendido e incubar a 37°C durante 1 h. Transferir el ARN tratado a un tubo nuevo y añadir 40 μ l de EDTA 0,5 M, incubar a 65°C durante 10 min. y precipitar con 1 ml de acetato de amonio 7,5 M más 8 ml de etanol. Resuspender el sedimento en 300 μ l de TE, calentar a 65°C para ayudar a dicha resuspensión y cargar sobre una columna de 1 cm x 18 cm (Sephadex G-50) y recoger el primer pico detectado mediante absorbancia UV a 260 nm. Precipitar el material recogido (~4 ml) con 2 ml de acetato de amonio 7,5 M y 16 ml de etanol. Resuspender el sedimento en 120 μ l de TE, lavar el tubo con 20 μ l de TE y combinar las 2 muestras. Este procedimiento proporciona un controlador haptenilado de la biblioteca de ADNc de cerebro fetal humano para su uso en la normalización de la biblioteca de ADNc de cerebro fetal humano.

La hibridación sustractiva se lleva a cabo usando el siguiente procedimiento. Desnaturalizar una mezcla de los siguientes componentes a 80°C durante 1 min: 1 μ g de biblioteca de ADNc de cs circular (véase anteriormente), 0,5 μ g del oligonucleótido de oligodA 5'(A)₄₀3' (oligodA), 3 μ g de oligonucleótido de sentido promotor de SP6-SalI 5' GAA GGT ACG CCT GCA GGT ACC GGT CCG GAA TTC CCG GGT CGA CCC ACG 3' (SEQ ID NO: 1) (SP6-SalI), NaCl 0,25 M en 22 μ l de tampón de hibridación 1x [HEPES 50 mM (pH 7,5), EDTA 1 mM y SDS al 0,1%]. Tras la desnaturalización, incubar la mezcla a temperatura ambiente durante 30 min.

Para la biblioteca de Cot = 500, se desnaturalizan 85 μ g del controlador de ARN biotinilado (véase anteriormente) en 22 μ l de 1 X tampón de hibridación a 90°C durante 2 min, enfriar sobre hielo durante 1 min, y añadir 1 μ l de NaCl 5 M. Transferir el ADN cs circular prehibridado al controlador de ARN biotinilado e incubar a 42°C durante 24 h. Para la biblioteca de Cot = 5, se hibridan 10,5 μ g de controlador de ARN durante 2 h; para la biblioteca de Cot = 50, se hibridan 41 μ g de controlador de ARN durante 5 h; para la biblioteca de Cot = 0, no se añade ningún controlador de ARN y se incuba la mezcla durante 24 h.

Tras la incubación, transferir la mezcla a un tubo nuevo, añadir 25 μ g de estreptavidina e incubar a temperatura ambiente 5 min. Extraer la disolución con un volumen igual de PCIA (fenol/cloroformo/alcohol isoamílico, 25:24:1). Realizar una extracción inversa de la fase orgánica con 15 μ l de TE que contiene NaCl 1 M y combinar las extracciones acuosas. Repetir la unión a estreptavidina y extracción con PCIA otras dos veces. Precipitar la fase acuosa con acetato de sodio 0,3 M y etanol. Resuspender el sedimento en 15 μ l de TE y dializar frente a TE (10 mM:0,5 mM) durante 30 min. Transferir el ADN a un tubo nuevo y medir el volumen. Este ADNc resultante es una biblioteca de ADNc de cadena sencilla normalizada.

El análisis de clones tras la sustracción se realiza de la siguiente manera. Cuando el ADN de cs circular que permanece tras la sustracción se convierte en ADNc de cd usando un cebador oligodA-*NotI*, dNTPs, como polimerasa de reparación y se transforma en células *E. coli*, una gran fracción de los transformantes contienen plásmidos que no contienen insertos (tabla 1).

TABLA 1

Porcentaje de clones de ADNc recombinante y tamaño de inserto promedio tras la sustracción de una biblioteca de ADNc de cerebro fetal humano total

Biblioteca de ADNc de cerebro fetal humano	% de recombinantes (24 clones independientes)	Tamaño de inserto promedio (kb)
Cot = 0	92	1,3
Cot = 5	79	1,2
Cot = 50	67	1,4
Cot = 500	45	1,1

Tras el análisis de los clones no contienen insertos, se determinó que estaban presentes en la biblioteca original con una frecuencia inferior al 1%, pero se enriquecieron tras la sustracción ya que no tienen una molécula controladora correspondiente para sustraerlos (figura 2). Se desarrollaron dos enfoques para eliminar esto del ruido de fondo y se describen en los ejemplos 2 y 3.

Ejemplo 2

Eliminación del ruido de fondo de una biblioteca de ADNc normalizada usando la selección con una sonda de OligodA-NotI biotinilada específica de diana

Como resultado del procedimiento de sustracción descrito en el ejemplo 1, hay una tendencia a aumentar el ruido de fondo que depende directamente de la Cot de la etapa de sustracción (tabla 1). Ya que se usa un controlador de biblioteca total, se enriquecerán los clones que no contienen un homólogo en el controlador. Esto se observó en el procedimiento descrito en el ejemplo 1 (figura 2). Para tratar este problema, se desarrollaron dos métodos y se describe un tercero para eliminar prácticamente el ruido de fondo. En el primer caso, descrito en este ejemplo, se usó la selección de clones recombinantes usando una sonda biotinilada de oligodA-NotI (figura 3) tal como sigue.

Tras la sustracción, reparación y transformación, el 45% de los clones derivados del protocolo de Cot = 500 eran recombinantes (tabla 1), sin embargo, usando selección con sonda con un cebador oligodA-NotI biotinilado (5'(A)₁₅GGG CGG CCG C 3') (SEQ ID NO: 2), se seleccionaron los clones recombinantes separados de los no recombinantes permitiendo la construcción de una biblioteca de ADNc normalizada sin cambio significativo en el tamaño de inserto promedio y la práctica eliminación de clones no recombinantes (tabla 2).

TABLA 2

Porcentaje de clones de ADNc recombinantes y tamaño de inserto promedio tras la sustracción de la biblioteca de ADNc de cerebro fetal humano total y selección de GENETRAPPER™ con sonda de oligodA-NotI biotinilada

Biblioteca de ADNc de cerebro fetal humano	% de recombinantes (96 clones independientes)	Tamaño de inserto promedio (kb)
Cot = 500	99	1,25

Más del 98% de los clones tomados aleatoriamente contienen insertos que son, de promedio, tan grandes como la biblioteca de ADNc no normalizada de la que se derivan. Además, el análisis mediante PCR de amplicones de TGF-β raros y abundantes indica que se ha logrado la normalización sustancial (figura 6). Obsérvese que aunque el producto de PCR de TGF-β1 no puede detectarse en las bibliotecas no normalizadas y de baja Cot, se detecta en las bibliotecas de Cot superior.

Se calentó el ADNc de cs circular normalizado del ejemplo 1 a 70°C durante 1 min y se enfrió sobre hielo durante 1 min. Se hibridaron 200 ng del cebador oligodA-NotI biotinilado (véase anteriormente) a 37°C durante 1 h. Se incubó

ES 2 281 138 T3

la mezcla de hibridación con 80 μg de perlas magnéticas de estreptavidina. Se marcaron las perlas tres veces con 100 μl de tampón de lavado (Tris-HCl 10 mM [pH 7,5], EDTA 1 mM). Se resuspendieron las perlas en 20 μl 1 X tampón de elución, glicerina 10 mM y se guardó el eluato. Se repitió la etapa de elución con 15 μl de 1 X tampón de elución y se combinaron los eluatos. Se repitió este protocolo tres veces y los eluatos.

5 Se reparó el ADNc de cadena sencilla capturado tal como sigue: realizar una reparación combinando 4 μl de 10 X tampón de reparación [Tris-HCl 100 mM (pH 8,8 a 25°C), MgCl_2 15 mM, KCl 500 mM, Triton X-100 al 1%], 1 μl de dNTP 10 mM, 1 μl de enzima de reparación Dynazyme (2 m/ μl) (*Thermus brockianus* de Finnzymes) y 34 μl de agua. Se mezcló esta mezcla y se almacenó sobre hielo húmedo. Se preparó una mezcla de cebador de ADN
10 añadiendo lo siguiente a un tubo de microcentrífuga nuevo: 4 μl de 10 X tampón de reparación, 35 μl de ADNc capturado de la etapa anterior y 1 μl (50 ng) de cebador oligodA-*NotI* no biotinilado. Se centrifugó la mezcla de cebador a temperatura ambiente durante 2 seg a 14.000 x g y se incubó a 95°C durante 1 min. Al mismo tiempo, se incubó la mezcla de reparación a 70°C. Se transfirió la mezcla de cebador de ADN al baño a 70°C y se incubó durante 1 min. Se añadieron 40 μl de la mezcla de reparación previamente calentada al tubo que contenía la mezcla de cebador
15 de ADN. Se mezclaron los contenidos pipeteando y luego se incubó la mezcla a 70°C durante 15 min para permitir la extensión con cebador (síntesis de ADNc de cadena doble). Se retiraron los tubos del baño de agua y se centrifugaron a temperatura ambiente durante 2 s a 14.000 x g. Se precipitó el ADN reparado añadiendo 1 μl de glucógeno, 41 μl de acetato de sodio 7,5 M, y 320 μl de etanol a -20°C a cada tubo. Se agitaron con vórtex los tubos y se colocaron en hielo durante 10 min o a 4°C durante toda la noche. Entonces se centrifugaron los tubos a 4°C durante 30 min a 14.000
20 x g. Se eliminó cuidadosamente el etanol del sedimento pequeño y se cubrió con 100 μl de etanol al 70% (-20°C). Se centrifugaron los tubos a 4°C durante 2 min a 14.000 x g y se eliminó todo el etanol y se secaron los sedimentos a temperatura ambiente durante 10 min o hasta sequedad. Se disolvieron los sedimentos en 10 μl de tampón TE y se almacenaron a 4°C. Se electroporaron 2 μl de alícuotas del ADN reparado por 20 μl de alícuotas de *E. coli* competente DH10B ElectroMax.

25 Ejemplo 3

Eliminación de ruido de fondo de una biblioteca de ADNc normalizada usando síntesis de reparación con OligodA-NotI con análogos de nucleótido que confieren resistencia a la nucleasa.

30 Usando el enfoque en el ejemplo 2 para eliminar el ruido de fondo, construir una biblioteca de ADNc normalizada con más de 1×10^6 clones primarios requiere como mínimo tres selecciones independientes y 15 electroporaciones (tabla 3).

35 TABLA 3

Comparación de diversos métodos para eliminar el ruido de fondo

Método	Número de electroporaciones	N° total de clones	% de recombinantes
Selección con sonda biotinilada	15. 3 selecciones	$1,2 \times 10^6$	> 95%
Selección de reparación resistente a nucleasa	5	$4,8 \times 10^6$	> 95%

60 Para tratar este problema, se desarrolló un enfoque alternativo para reducir el ruido de fondo en bibliotecas normalizadas. En este método, denominado síntesis de reparación resistente a nucleasa, se usa la misma sonda descrita en el ejemplo 2, oligodA-*NotI*, pero en este caso no está biotinilada (figura 5). Sin embargo, pueden usarse sondas biotiniladas tal como se usan en el ejemplo 2 para incluir la etapa de selección adicional del ejemplo 2. Cuando se compara con el método de selección del ejemplo 2, puede construirse una biblioteca que es cuatro veces más compleja y requiere un tercio del número de electroporaciones (tabla 3). Además, se elimina prácticamente el ruido de fondo de
65 la biblioteca y no se cambia el tamaño de inserto de la biblioteca (tabla 4). Finalmente, cuando se examinaron genes sumamente abundantes mediante hibridación de colonias, su abundancia disminuyó de 15 a 18 veces (tabla 5) y la abundancia de genes raros aumentó sustancialmente (figura 6).

ES 2 281 138 T3

TABLA 4

Porcentaje de clones de ADNc recombinantes y tamaño de inserto promedio tras la sustracción de biblioteca de ADNc de cerebro fetal humano total y tratamiento con 5-metilcitosina/HhaI

Biblioteca de ADNc de cerebro fetal humano	% de recombinantes (80 clones independientes)	Tamaño de inserto promedio (kb)
Cot = 500	> 95	1

TABLA 5

Análisis de biblioteca de ADNc normalizada: la depleción de ADNc abundantes depende directamente del grado de sustracción

Gen	Cot = 0	Cot = 5	Cot = 50	Cot = 500
α -tubulina	0,78%	0,62%	0,24%	0,043%
EF-1 α	0,42%	0,28%	0,13%	0,029%
Hibridación en colonia usando sondas de oligonucleótido marcadas con ³² P directamente en la α -tubulina y factor de elongación 1 (EF-1 α)				

Se reparó una biblioteca de ADNc de cadena sencilla normalizada generada mediante sustracción (véase el ejemplo 1) tal como sigue: se preparó una mezcla de reparación combinando 3 μ l de 10 X tampón de reparación [Tris-HCl 100 mM (pH 8,8 a 25°C), MgCl₂ 15 mM, KCl 500 mM, Triton X-100 al 1%], 1 μ l de dNTP 10 mM (que contenía 5-metil-dCTP 10 mM), 1 μ l de enzima de reparación Dynazyme (2 u/ μ l) (*Thermus brockianus* de Finnzymes) y 25 μ l de agua, mezclar y almacenar sobre hielo húmedo. Se preparó una mezcla de cebador de ADN para cada reacción añadiendo lo siguiente a un tubo de microcentrífuga nuevo: 11 μ l de agua destilada en autoclave, 3 μ l de 10 X tampón de reparación, 15 μ l de ADN dializado de la etapa anterior, y 1 μ l (50 ng) de oligoA-NotI no biotinilado. Se centrifugó la mezcla a temperatura ambiente durante 2 seg a 14.000 x g. Se incubó la mezcla de cebador de ADN a 95°C durante 1 min. Al mismo tiempo, se incubó la mezcla de reparación a 70°C. Se transfirió la mezcla de cebador de ADN al baño a 70°C y se incubó durante 1 min. Se añadieron 30 μ l de la mezcla de reparación previamente calentada al tubo que contenía la mezcla de cebador. Se mezclaron los contenidos pipeteando y se incubaron a 70°C durante 15 min para permitir la extensión de cebador (síntesis de ADN de cadena doble). Se retiraron los tubos del baño de agua y se centrifugaron a temperatura ambiente durante 2 seg a 14.000 x g. Se precipitó el ADN reparado añadiendo 1 μ l de glucógeno, 32 μ l de acetato de amonio 7,5 M, y 250 μ l de etanol a -20°C a cada tubo. Se agitaron con vórtex los tubos y se colocaron en hielo durante 10 min o a 4°C durante toda la noche. Entonces se centrifugaron los tubos a 4°C durante 30 min a 14.000 x g. Se eliminó cuidadosamente el etanol a partir del pequeño sedimento y se recubrió con 100 μ l de etanol al 70% (-20°C). Se centrifugó el tubo a 4°C durante 2 min a 14.000 x g. Se eliminó todo el etanol y se dejaron los sedimentos a temperatura ambiente durante 10 min o hasta sequedad. Se disolvieron los sedimentos en 10 μ l de tampón TE y se almacenaron a 4°C. Se digirió el reparado con 0,5 unidades de HhaI en 20 μ l de 1 X tampón (Tris-HCl 5 mM, pH 8,0; MgCl₂ 1 mM; NaCl 5 mM) a 37°C durante 30 min. Se precipitó en etanol el ADN y se resuspendió el sedimento seco resuspendido en 8 μ l de TE. Se electroporaron alícuotas de 2 μ l de ADN reparado por alícuota de 20 μ l de *E. coli* competente DH10B ElectroMax.

REIVINDICACIONES

1. Método para la normalización de una biblioteca de ácido nucleico que comprende:

- 5 (a) incubar una biblioteca de ácido nucleico que va normalizarse con moléculas de ácido nucleico haptenilado sintéticas complementarias a todas o a una parte de las moléculas de ácido nucleico de dicha biblioteca en condiciones que favorecen la hibridación de las moléculas más abundantes de dicha biblioteca con dichas moléculas de ácido nucleico haptenilado; y
- 10 (b) eliminar dichas moléculas hibridadas mediante interacciones ligando-hapteno, produciendo así una biblioteca normalizada.

2. Método según la reivindicación 1, en el que dicha biblioteca de ácido nucleico es una biblioteca de ADNc.

15 3. Método según la reivindicación 2, en el que las moléculas de ácido nucleico de dicha biblioteca de ADNc son de cadena sencilla.

20 4. Método según la reivindicación 2, en el que las moléculas de ácido nucleico de dicha biblioteca de ADNc son de cadena doble.

5. Método según la reivindicación 2, en el que dicha biblioteca de ADNc se produce mediante un método que comprende una población de moléculas de ARNm en condiciones suficientes para producir una biblioteca de ADNc a partir de dicha población de moléculas de ARNm.

25 6. Método según la reivindicación 1, en el que dichas moléculas de ácido nucleico haptenilado son moléculas de ARN.

30 7. Método según la reivindicación 1, que además comprende la reducción o eliminación de moléculas de ácido nucleico contaminante de dicha biblioteca.

8. Método según la reivindicación 7, en el que dicha reducción o eliminación se realiza antes o después de la normalización de dicha biblioteca.

35 9. Método según la reivindicación 7, en el que las moléculas de ácido nucleico contaminante son uno o más vectores.

10. Método según la reivindicación 7, en el que dicha reducción o eliminación comprende incubar dicha biblioteca con al menos una sonda haptenilada.

40 11. Método según la reivindicación 10, en el que dicha sonda se hibrida con moléculas de ácido nucleico de dicha biblioteca.

45 12. Método según la reivindicación 11, en el que dicha sonda no puede hibridarse con secuencias de vector de dicha biblioteca.

13. Método según la reivindicación 10, en el que los haptenos de dicha sonda haptenilada se usan para aislar una biblioteca normalizada que tiene moléculas de ácido nucleico contaminante sustancialmente reducidas, produciendo así una biblioteca normalizada seleccionada.

50 14. Método según la reivindicación 13, en el que dichos haptenos se seleccionan del grupo que consiste en avidina, estreptavidina, proteína A, proteína G, un receptor Fc de la superficie celular, un antígeno específico para anticuerpo, un sustrato específico para enzima, polimixina B, proteína neutralizante de endotoxina (ENP), Fe⁺⁺⁺, un receptor de transferrina, un receptor de insulina, un receptor de citocina, CD4, espectrina, fodrina, ICAM-1, ICAM-2, C3bi, fibrinógeno, factor X, anquirina, una integrina, vitronectina, fibronectina, colágeno, laminina, glicoforina, Mac-1, LFA-1, β -actina, gp120, una citocina, insulina, ferrotransferrina, apotransferrina, lipopolisacárido, una enzima, un anticuerpo, biotina y combinaciones de los mismos.

55 15. Método según la reivindicación 14, en el que dicho hapteno es biotina.

60 16. Método según la reivindicación 13, en el que dicha biblioteca normalizada seleccionada se aísla mediante interacciones hapteno-ligando y/o extracción.

65 17. Método según la reivindicación 16, en el que el aislamiento comprende el uso de un soporte sólido que comprende al menos un ligando que se une a dicho hapteno.

18. Método según la reivindicación 13, en el que dicha biblioteca normalizada seleccionada es de cadena sencilla.

ES 2 281 138 T3

19. Método según la reivindicación 18, que además comprende incubar dicha biblioteca normalizada seleccionada de cadena sencilla en condiciones suficientes para convertir dichas moléculas en cadena doble.

20. Método según la reivindicación 19, en el que dichas condiciones comprenden incubar dicha biblioteca normalizada seleccionada de cadena sencilla con uno o más nucleótidos, uno o más polipéptidos que tienen actividad polimerasa y uno o más cebadores.

21. Método según la reivindicación 20, en el que dichos uno o más nucleótidos son análogos de nucleótido que confieren resistencia a la nucleasa sobre dichas moléculas de cadena doble.

22. Método según la reivindicación 21, que además comprende digerir una muestra que comprende dichas moléculas de cadena doble con dicha nucleasa.

23. Método según la reivindicación 22, que además comprende transformar dichas moléculas de cadena doble en una o más células huésped.

24. Método según la reivindicación 19, que además comprende transformar dichas moléculas de cadena doble en una o más células huésped.

25. Método según la reivindicación 20, en el que dichos cebadores son cebadores específicos de diana.

26. Método según la reivindicación 25, que además comprende transformar dichas moléculas de cadena doble en una o más células huésped.

27. Método según la reivindicación 6, en el que dichas moléculas de ARN se producen mediante una o más ARN polimerasas.

28. Método según la reivindicación 27, en el que dichas ARN polimerasas se seleccionan del grupo que consiste en ARN polimerasas de SP6, T7 y T3.

29. Método según la reivindicación 6, en el que dichas moléculas de ARN se producen con uno o más promotores.

30. Método según la reivindicación 29, en el que dichos promotores se proporcionan por uno o más vectores o mediante uno o más adaptadores.

31. Método según la reivindicación 30, en el que dichos promotores permiten la síntesis de al menos una molécula de ARN a partir de todas o parte de las moléculas de ácido nucleico de dicha biblioteca.

32. Método según la reivindicación 1, en el que dichas moléculas hibridadas se eliminan mediante interacciones hapteno-ligando y/o extracción.

33. Método según la reivindicación 32, en el que dicha eliminación comprende el uso de un soporte sólido que comprende al menos un ligando.

34. Método según la reivindicación 4, que además comprende tratar dicha biblioteca de ADNc de cadena doble en condiciones suficientes para hacer que dichas moléculas sean de cadena sencilla.

35. Método según la reivindicación 34, en el que dichas condiciones comprenden la degradación de una cadena de dichas moléculas de cadena doble.

36. Método según la reivindicación 34, en el que dichas condiciones comprenden desnaturalizar dichas moléculas de cadena doble.

37. Método según la reivindicación 35, en el que dicha degradación se logra con gen II y exonucleasa III.

38. Método según la reivindicación 1, en el que dichas condiciones de hibridación se seleccionan del grupo que consiste en:

(a) una Cot igual o superior a 25;

(b) una Cot igual o superior a 50;

(c) una Cot igual o superior a 100;

(d) una Cot desde aproximadamente 10 hasta 10.000;

(e) una Cot desde aproximadamente 25 hasta 10.000;

ES 2 281 138 T3

(f) una Cot desde aproximadamente 50 hasta 10.000;

(g) una Cot desde aproximadamente 100 hasta 10.000; y

5 (h) una Cot inferior a 10.000;

en el que Cot es el producto de la concentración inicial de ácido nucleico (moles de nucleótidos por litro, Co) y el tiempo (t).

10 39. Método según la reivindicación 7, en el que dicha reducción o eliminación comprende incubar dicha biblioteca con al menos un cebador y al menos un nucleótido que confiere resistencia a la nucleasa en condiciones suficientes para preparar moléculas de ácido nucleico de cadena doble.

15 40. Método según la reivindicación 39, en el que dicho cebador se hibrida con moléculas de ácido nucleico de dicha biblioteca.

41. Método según la reivindicación 39, en el que dicho cebador no puede hibridarse con secuencias de vector de dicha biblioteca.

20 42. Método según la reivindicación 39, en el que dicho nucleótido es un análogo de nucleótido.

43. Método según la reivindicación 42, en el que dicho nucleótido es un nucleótido metilado.

25 44. Método según la reivindicación 43, en el que dicho nucleótido metilado es 5-metil-desoxicitosina.

45. Método según la reivindicación 39, que además comprende digerir dichas moléculas de ácido nucleico de cadena doble con una o más nucleasas.

30 46. Método según la reivindicación 45, que además comprende transformar dichas moléculas digeridas en una o más células huésped.

35

40

45

50

55

60

65

Figura 1

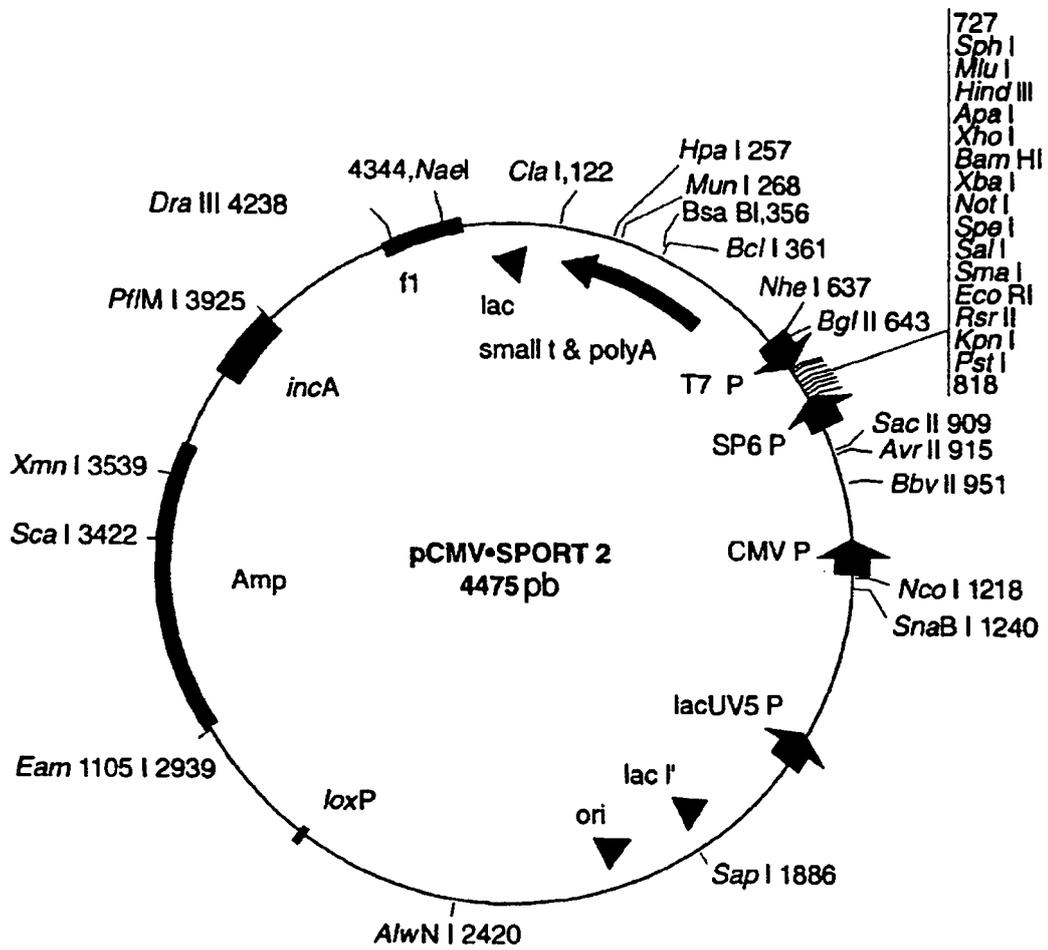


Figura 2

Producción de una biblioteca de ADNc normalizada mediante sustracción con un controlador de ARN biotinilado

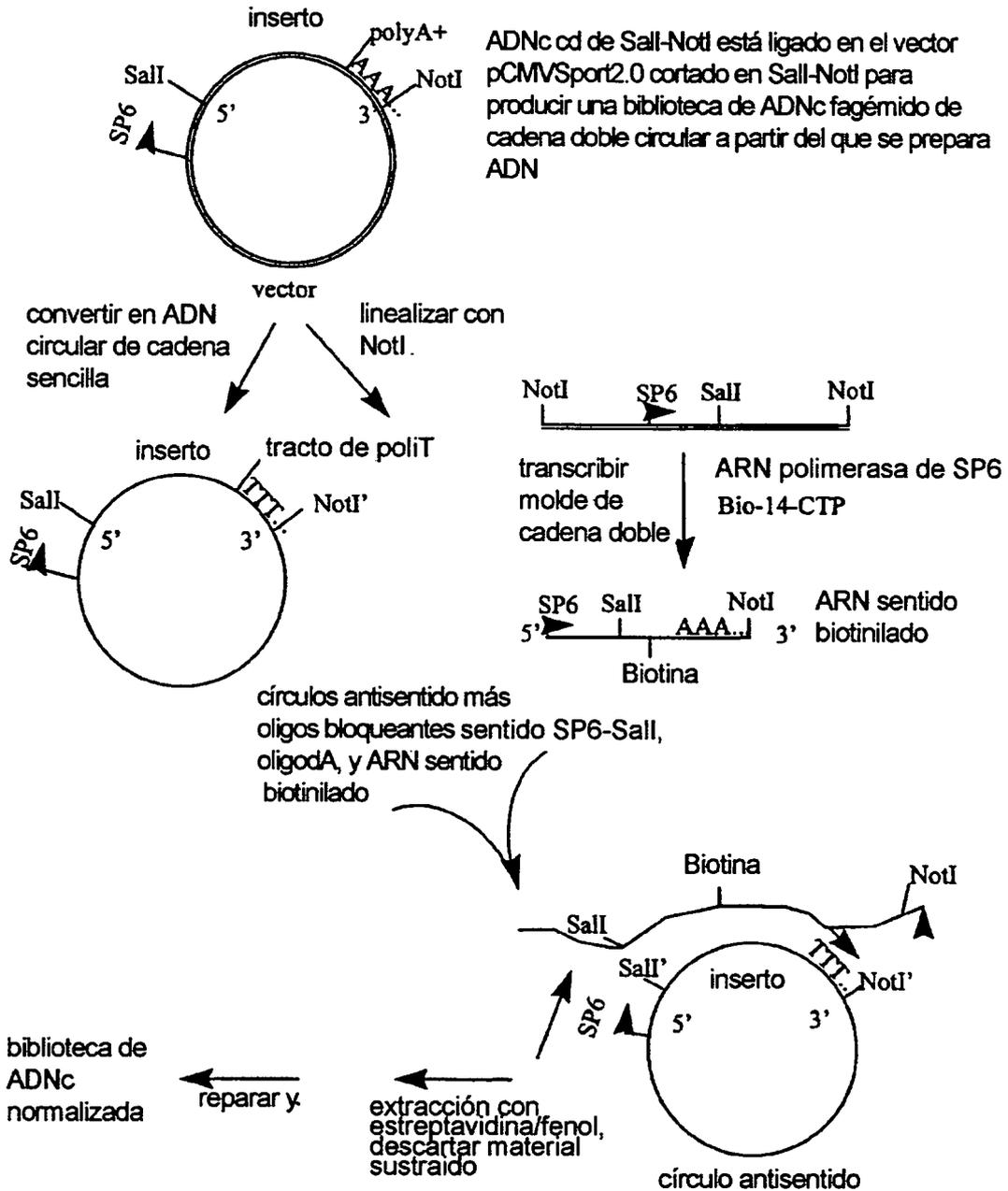


Figura 3.

Bibliotecas de fagémido normalizadas seleccionadas: eliminación de ruido de fondo usando una selección en sonda oligodA-NotI biotinilada en 3'

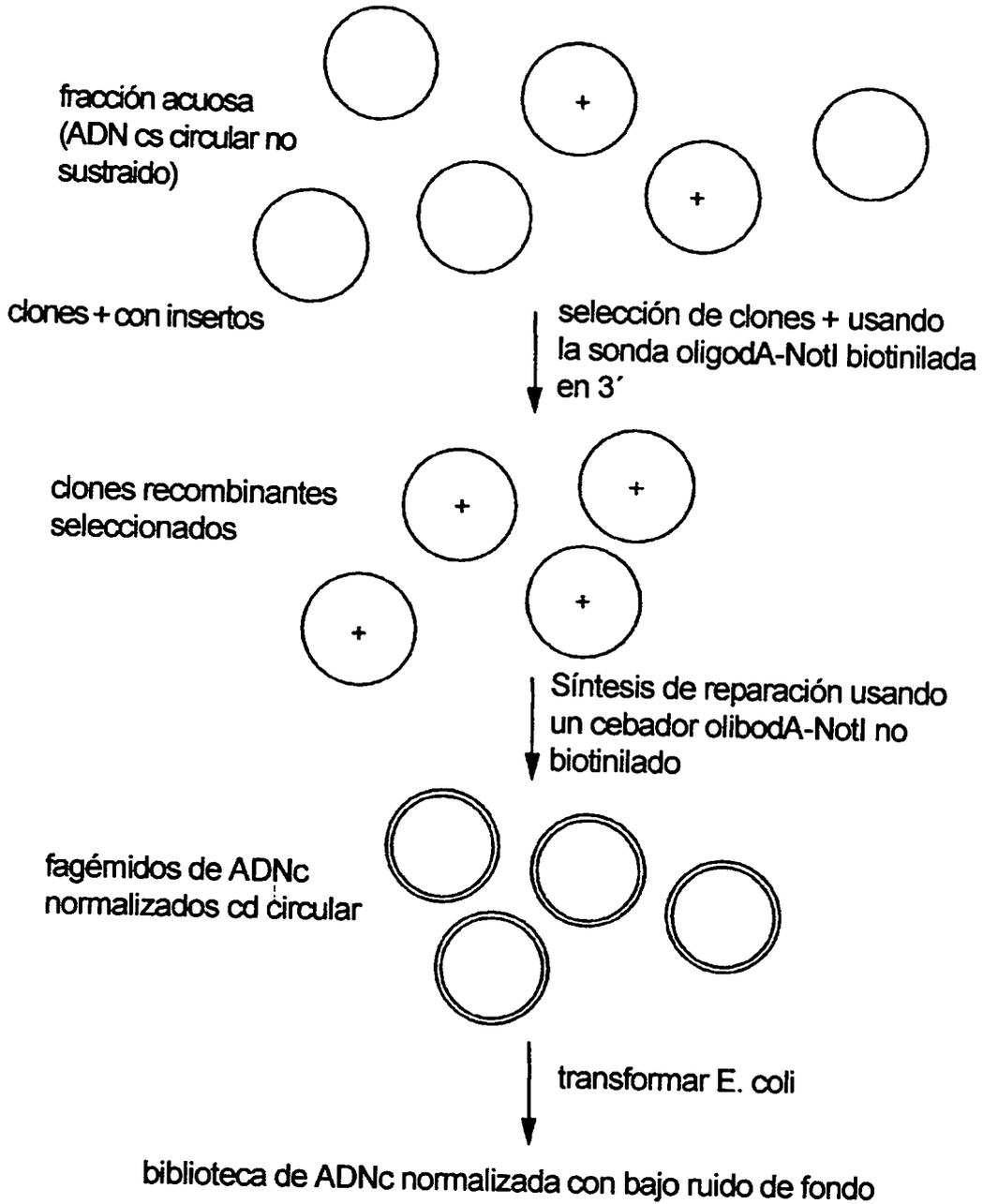


Figura 4.

Bibliotecas de fagémido normalizadas seleccionadas: eliminación de ruido de fondo usando un cebador oligodA-NotI biotinilado en 5'

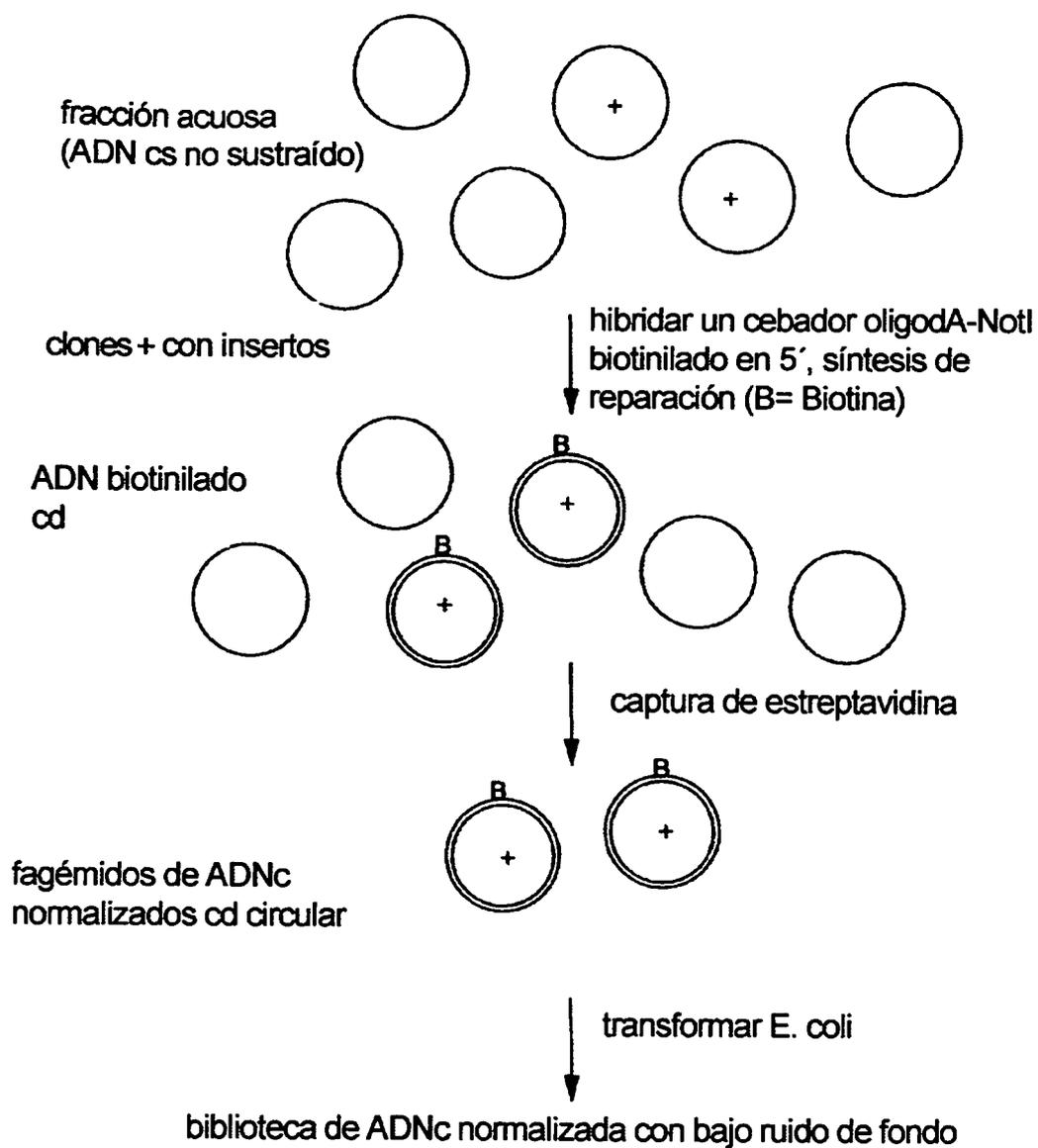


Figura 5

Bibliotecas de fagémidos normalizadas seleccionadas: eliminación de ruido de fondo usando 5-metilcitosina y HhaI

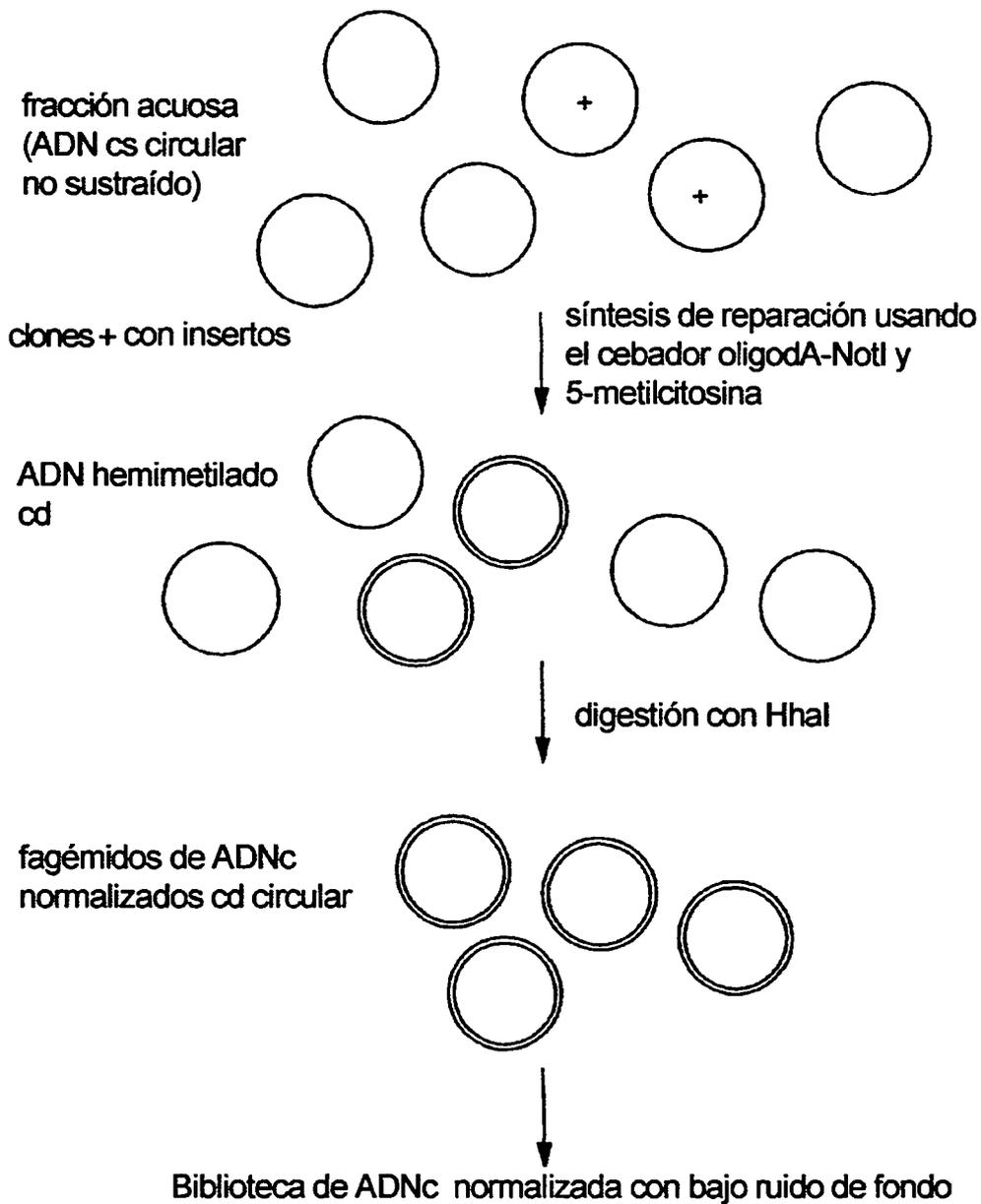


Figura 6

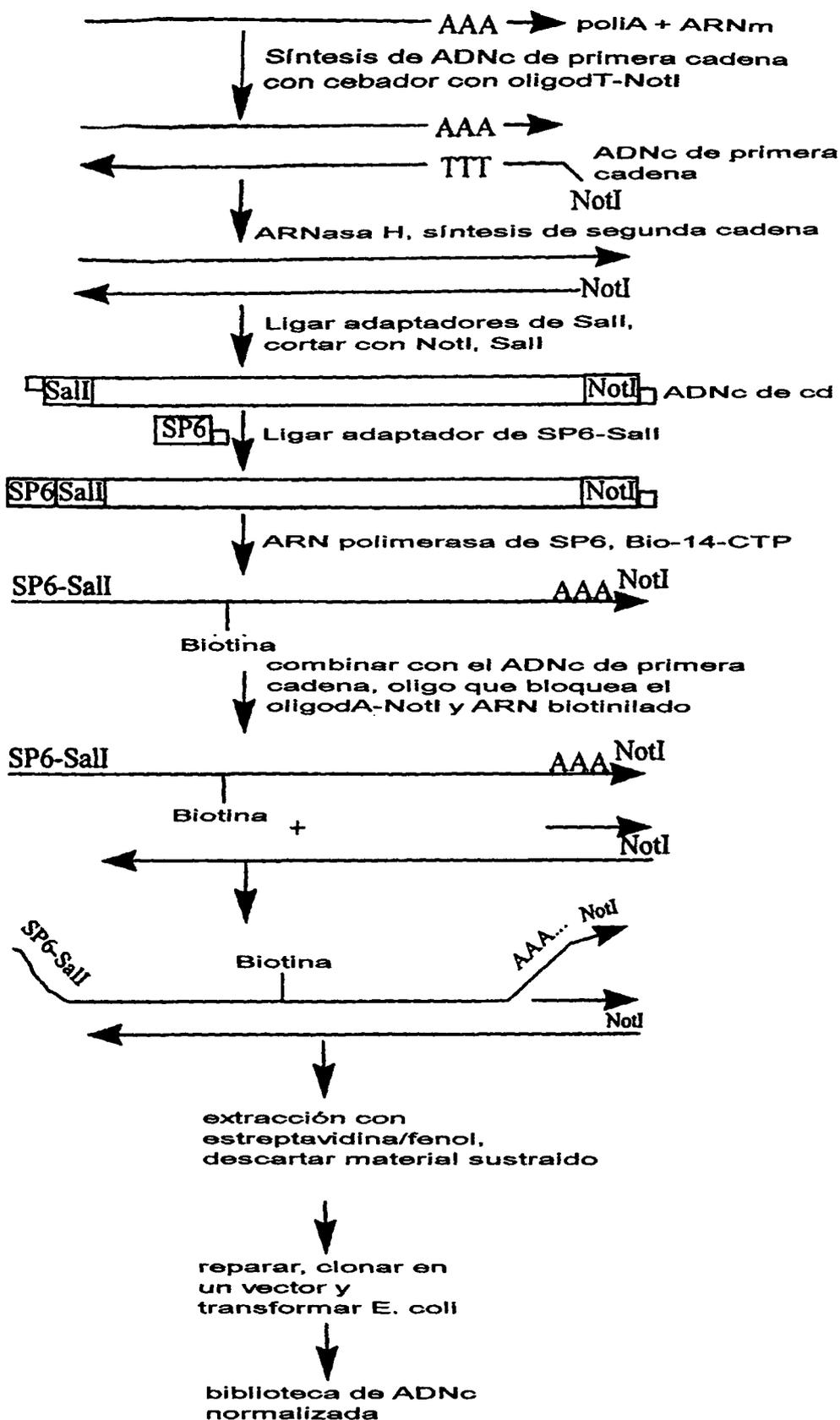
Enriquecimiento de clon de ADNc raro es una función de la Cot de sustracción:
familia de gen TGF β



1 Selección con 3'-biotina-oligodA-NotI

2 Síntesis de reparación 5-metilcitosina/digestión con HhaI

Figura 7



ES 2 281 138 T3

LISTA DE SECUENCIAS

- <110> Life Technologies, Inc.
9800 Medical Center Drive
5 Rockville, Maryland 20850-3321 EE.UU.
- <120> Bibliotecas de ácido nucleico normalizadas y métodos de producción de las mismas
<130> 0942,447PC01
<140>
10 <141>
<150> US (falta por asignar)
<151> 23-09-1998
15 <150> 60/059,817
<151> 24-09-1997
<160> 2
<170> PatentIn Ver. 2,0
20 <210> 1
<211> 48
<212> ADN
25 <213> Secuencia artificial
<220>
<223> Descripción de secuencia artificial: oligonucleótido sintético
30 <400> 1

gaaggtagc ctgcaggtac cgggccgaa ttccggggtc gaccacg 48
- <210> 2
35 <211> 25
<212> ADN
<213> Secuencia artificial
40 <220>
<223> Descripción de la secuencia artificial: oligonucleótido sintético
45 <400> 2

aaaaaaaaa aaaagggcg gccgc 25
- 50
55
60
65