

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 731 460**

51 Int. Cl.:

C40B 20/04 (2006.01)

C07H 21/04 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **11.10.2011** **PCT/US2011/055803**

87 Fecha y número de publicación internacional: **12.04.2012** **WO12048341**

96 Fecha de presentación y número de la solicitud europea: **11.10.2011** **E 11831761 (9)**

97 Fecha y número de publicación de la concesión europea: **27.03.2019** **EP 2625320**

54 Título: **Alto rendimiento de células individuales con código de barra**

30 Prioridad:

08.10.2010 US 391364 P

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

15.11.2019

73 Titular/es:

**PRESIDENT AND FELLOWS OF HARVARD
COLLEGE (100.0%)
17 Quincy Street
Cambridge, MA 02138, US**

72 Inventor/es:

**CHURCH, GEORGE M. y
VIGNEAULT, FRANCOIS**

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 731 460 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

Alto rendimiento de células individuales con código de barra

5 Campo

La presente invención se refiere a métodos y composiciones para obtener y analizar secuencias de ácido nucleico derivadas de varias células individuales al mismo tiempo.

10 Antecedentes de la invención

El análisis clásico de una célula individual se realiza mediante el aislamiento de una célula individual en un único pocillo de una placa de procesamiento desde el cual puede amplificarse el ADN y/o el ARN o donde la célula puede subcultivarse en una población más grande, con ambos enfoques realizados hasta que se logre suficiente material genómico para su procesamiento posterior subsecuente. Una limitación de tales enfoques es que no siempre es posible aislar células individuales de una sección de tejido o de una mezcla o población celular compleja. Además, en una población de células amplificadas clonalmente en cultivo, incluso si las células deben presentar el mismo genoma exacto, lo que debe ser en teoría, la información transcriptómica es variable de una célula a otra. Además, durante el cultivo las células modifican sus patrones de expresión, por lo que a menudo es preferible capturar la información transcriptómica cuando las células están en su entorno original. Además, las cantidades extremadamente bajas de ADN y/o ARN que se obtienen cuando se aísla una célula individual hacen que las etapas de procesamiento posteriores sean bastante difíciles. Además, los procesos mediante los cuales el ADN y/o el ARN se amplifican hasta cantidades lo suficientemente grandes para permitir tales análisis causa un sesgo significativo en el material resultante y, por lo tanto, no es representativo de los ácidos nucleicos en la célula. Finalmente, los enfoques clásicos están limitados en la cantidad de células individuales que pueden evaluarse en un análisis. Por ejemplo, debe estudiarse una población compleja de 10 000 células; será necesario clasificar y separar 10 000 células (mediante el uso de, por ejemplo, aproximadamente 100 placas de 96 pocillos), lo que requiere una inversión sustancial en equipos de automatización costosos, así como también un tiempo de procesamiento significativo y costos adicionales.

Los primeros enfoques incluyeron la síntesis de ADN agrupado dividido. Si bien la síntesis de ADN agrupado dividido sobre perlas puede usarse potencialmente para lograr perlas con código de barras único (Brenner y otros (2000) Proc. Natl Acad Sci. USA 97: 1665), las dificultades técnicas asociadas con tal enfoque y la incorporación ineficiente del nucleótido durante la síntesis química de la secuencia, resultan en perlas que tienen muy pocas secuencias de oligonucleótidos con secuencias y/o longitudes correctas. Incluso cuando la síntesis química de nucleótidos es bastante eficiente, hay, en promedio, un 1 % de no incorporación en cada ciclo de nucleótidos. En consecuencia, los intentos de sintetizar un código de barras clonal sobre las perlas de síntesis de ADN agrupado dividido de longitud adecuada no tuvieron éxito. Por ejemplo, para un oligonucleótido típico de 50-60 nucleótidos, esta tasa de error resultaría en menos del 40 % de los oligos sobre las perlas que tienen la secuencia correcta. Además, debido a que los oligonucleótidos se sintetizan sobre un soporte sólido, es imposible identificar el correcto, mediante el uso de métodos de purificación tal como la purificación por HPLC o PAGE. La síntesis agrupada dividida se desarrolló originalmente por Linx Therapeutics, que fue adquirida por Solexa, que fue adquirida por Illumina en base al trabajo inicial sobre síntesis agrupada dividida, pero la tecnología se abandonó debido a estos problemas. Por lo tanto, no se ha logrado el uso eficiente de las perlas con código de barras. Las perlas con un núcleo interno de gradiente de tinte (tal como el que se usó por Luminex Corporation) pueden usarse en aplicaciones donde se usa la señal global del código de barras de la perla. Si bien ese enfoque es aceptable cuando se desea una intensidad de señal promedio, es inadecuado cuando el uso posterior de estas moléculas requiere una identificación única de la célula. Además, las "perlas de luminex" solamente pueden generarse en una cantidad limitada que resulta en una capacidad limitada para examinar más de unos pocos cientos de células.

El presente enfoque ofrece ventajas particulares sobre los enfoques anteriores, tal como la síntesis de ADN agrupado dividido sobre las perlas.

Resumen

El presente enfoque produce de manera eficiente perlas con código de barras recubiertas con copias clonales de los oligonucleótidos con código de barras que tienen la secuencia correcta. Además, la velocidad, la facilidad y el costo de producción también son ventajosos. Y, a diferencia de la síntesis de ADN agrupado dividido sobre perlas, pueden generarse millones de perlas con código de barras únicamente para el análisis de células individuales.

En un aspecto, la invención consiste en un enfoque para codificar con barras varias células individuales en una mezcla compleja de células. Cada célula se proporciona con un código de barras individual, único para cada celda. El código de barras único permite que los ácidos nucleicos de cada célula (genoma o transcriptoma) se asocien con la célula original. Por lo tanto, para cualquier célula individual dada, pueden identificarse múltiples genes y transcritos diferentes y correlacionarlos con la misma célula porque las secuencias comparten el mismo código de barras único.

El código de barras único se inserta en cada célula individual de manera que cada célula recibe un código de barras único y está presente en una cantidad lo suficientemente grande como para permitir el direccionamiento genómico o

transcriptómico subsecuente. Una vez que se inserta el código de barras, se llevan a cabo manipulaciones posteriores para capturar y después secuenciar todos estos códigos de barras únicos y las secuencias de interés del genoma o transcriptoma en una reacción simultánea. El enfoque actual, cuando se combina con la tecnología de secuenciación de alto rendimiento, permite analizar un gran número de células individuales y lograr el análisis en un solo ensayo de reacción.

5 En principio, puede secuenciarse cualquier número de célula y cualquier número de regiones seleccionadas por célula. El número de células individuales que pueden procesarse está limitado solamente por restricciones prácticas, tal como la velocidad de la secuenciación de alto rendimiento; por ejemplo. En algunas modalidades, se usan tecnologías de secuenciación de alto rendimiento, tal como las realizadas en plataformas de secuenciación tal como Illumina HiSeq o

10 analizador de genoma, Roche 454, Pacific Bioscience, Ion Torrents, Harvard Polonator, ABI Solid u otros instrumentos similares en el campo. Pueden usarse enfoques de secuenciación clásicos, tal como la secuenciación de Sanger; sin embargo, el verdadero poder en la tecnología es ser capaz de secuenciar un número mayor de secuencias de células individuales simultáneamente. Las plataformas de secuenciación de alto rendimiento son, por lo tanto, más adecuadas para la mayoría de las modalidades. Si una plataforma de secuenciación genera 10 millones de lecturas por corrida, entonces puede secuenciarse un transcrito único en 1 millón de células para lograr una cobertura de 10x. En otras

15 modalidades, un transcriptoma parcial, por ejemplo, dirigido a 10 000 transcritos únicos, requiere solamente que 100 células sean seleccionadas para captura y secuenciación.

En algunas modalidades, se realiza un análisis del transcriptoma de ARN completo o seleccionado. Por lo tanto, en una célula individual, solamente pueden secuenciarse los transcritos seleccionados. En otras aplicaciones, todos o

20 sustancialmente todos los transcritos pueden capturarse y secuenciarse. Aún en otras modalidades, se realiza un análisis de ADN genómico completo o parcial.

Los análisis de múltiples células en poblaciones heterogéneas de células son particularmente útiles cuando se estudian muestras complejas o mezclas. Las muestras complejas o mezclas de células incluyen, por ejemplo, muestras metagenómicas, secciones de tejidos normales y cancerosos, colonias de células madre y embrionarias. La secuenciación del genoma y del transcriptoma es conveniente cuando las secuencias son altamente divergentes; por ejemplo, en determinados tipos de células o en células en determinadas etapas. Las aplicaciones particularmente adecuadas incluyen tipificación molecular de haplotipos, tipificación de HLA, y la caracterización del perfil de los receptores de células T y B.

25 Las muestras metagenómicas se refieren a muestras que contienen genomas de múltiples orígenes, tales como las especies. Por ejemplo, el presente enfoque puede aplicarse a mezclas de especies bacterianas para permitir la secuenciación de ácidos nucleicos de múltiples bacterias en un ensayo seguido de la correlación de las secuencias con la misma célula bacteriana. De manera similar, las secuencias de ácido nucleico de células extrañas que viven en la boca pueden determinarse y correlacionarse con la misma célula.

35 Otras características y ventajas de determinadas modalidades de la presente invención se harán más evidentes en la siguiente descripción de las modalidades y figuras de estas, y a partir de las reivindicaciones.

Breve descripción de las figuras

40 El archivo de la patente o la solicitud contiene al menos un dibujo a color. Las copias de esta publicación de patente o de la solicitud de patente con la (las) figura(s) a color se proporcionarán por la Oficina bajo petición y pago de la tarifa necesaria. Las anteriores y otras características y ventajas de la presente invención se entenderán más completamente a partir de la siguiente descripción detallada de las modalidades ilustrativas tomadas junto con las figuras adjuntas, en las que:

45 *Las Figuras 1A-1E* representan esquemáticamente un método para amplificar cada molécula única compuesta de un código de barras degenerado sobre una perla de acuerdo con determinados aspectos de la invención. (A) Unión del oligonucleótido molde al código de barras. (B) Saturación del soporte sólido con cebador de anclaje. (C) Realización la PCR en emulsión. (D) Se completa la PCR en emulsión. (E) Se recuperan las perlas con código de barras.

50 *Las Figuras 2A-2D* representan esquemáticamente la PCR en emulsión de una célula individual. (A) Captura de células y perlas con código de barra en una emulsión. (B) Lisis celular. (C) Hibridación del ADN y/o el ARN en una perla con código de barras seguido de la extensión del cebador y/o la transcripción inversa.

55 *Las Figuras 3A-3D* representan esquemáticamente el procesamiento posterior de las perlas recuperadas unidas a fragmentos de código de barras. (A) Ejemplo de un molde de ARN. (B) Síntesis de la segunda cadena. (C) Cebador específico del gen, digestión por restricción o ligación del adaptador universal. (D) El ADN con código de barras recuperado, está listo para la secuenciación de alto rendimiento.

60 *La Figura 4* representa las perlas de acuerdo con determinados aspectos de la invención. Cy 5 muestra la presencia de un nucleótido de adenina en la posición uno del código de barras. Cy3 muestra la presencia de un nucleótido de timina en la posición uno del código de barras, Texas Red (TxRed) muestra la presencia de un nucleótido de citosina en la posición uno del código de barras. El isotiocianato de fluoresceína (FITC) muestra la presencia de una guanina en la posición uno del código de barras. La secuenciación de manera consecutiva de cada posición del código de barras proporciona el identificador único del código de barras. Cada transcrito capturado por las perlas puede correlacionarse

65 con una célula de inicio única porque cada célula está representada por un código de barras único.

Las Figuras 5A-5G representan un método para generar múltiples copias de un código de barras degenerado de forma única para el análisis de una célula individual de acuerdo con determinados aspectos de la invención. (A) Molde inverso de ADN (es decir, de inicio). (B) Circularización mediante ligación. (C) Amplificación de círculo rodante mediante el uso de una polimerasa de la cadena que se desplaza y cebador complementario. (D) Inserción dentro de liposoma o emulsión con secuencia complementaria de restricción y enzima de restricción. (E) Oligonucleótidos codificados en barras resultantes. Cada liposoma o emulsión contiene un código de barras único y degenerado. (F) Cada liposoma puede fusionarse directamente con una sola célula (o cada emulsión con código de barras puede fusionarse con una célula en emulsión). (G) El llamado de la secuenciación de la región del código de barras de la amplificación de círculo rodante (Rolony) demostró clonalidad. Los productos del Rolony se ordenaron en una cuadrícula de características de tamaño de 250 nanómetros.

Las Figuras 6A-6B representan un método para generar múltiples copias de un código de barras degenerado de forma única para el análisis de una célula individual de acuerdo con determinados aspectos de la invención con relación a la selección de más de una secuencia de ácido nucleico de interés. El panel izquierdo en las Figuras 6A y 6B muestra un cebador de hibridación de la secuencia oligo-dT, que puede dirigirse a las colas poliA de los ARNm que se encuentran en una célula. El panel derecho en las Figuras 6A y 6B demuestra mediante el uso de un cebador de "secuencia universal", que tiene una secuencia complementaria a un extremo saliente común a numerosas secuencias de cebadores de hibridación, para generar una perla que tiene oligonucleótidos que hibridan con múltiples dianas de ácidos nucleicos diferentes de interés (mostrados en rojo y azul en el extremo 3' del oligonucleótido.)

Las Figuras 7A-7E muestra la clonalidad de las perlas mediante el uso de diferentes concentraciones de cebador. (Fig. 7A: 0,1 pM; Fig. 7B: 1pM, Fig. 7C: 10 pM, y Fig. 7D: 100 pM). La Figura 7E representa la muestra de 100 pM superpuesta con las perlas. Ver Ejemplo 1.

Las Figuras 8A-8H muestra la clonalidad de las perlas en las emulsiones. La Figura 8A muestra el recubrimiento de las perlas con un código de barras único sobre luz blanca, que muestra la clonalidad de las perlas con una cantidad óptima del molde de inicio. La Figura. 8B muestra un ciclo de secuenciación del código de barras sobre los canales de fluorescencia solamente. La Figura 8C muestra luz blanca solamente. La Figura 8D muestra la captura de una perla individual en la emulsión. La Figura 8E muestra perlas con código de barras en presencia de células lisadas en la emulsión posterior a la amplificación. Las Figuras 8F y G son ampliaciones de la Figura 8E. La Figura 8H muestra la introducción de códigos de barras fluorescentes en células individuales.

Las Figuras 9A-9C Muestra las secuencias usadas en los aspectos de la invención. La Figura 9A muestra las secuencias 5' y 3' de un cebador (SEC ID NO: 1) usado en el sistema Illumina. La Figura 9B muestra la secuencia 5' y 3' de un cebador (SEQ ID NO: 2) usado en los aspectos de la invención, que incluye el cebador de secuencia de anclaje, la posición del código de barras de 20 nucleótidos (--BC(N20)--), y la secuencia oligo dT. Las secuencias agrupadas facilitan la secuenciación en el sistema Illumina. La Figura 9C muestra un oligonucleótido de muestra unido a una perla que tiene un cebador de secuencia de anclaje, un código de barras de 20 nucleótidos (--BC(N20)--) y una secuencia oligo dT (SEQ ID NO:3).

Descripción detallada

En determinados aspectos, los métodos y composiciones descritos en la presente descripción son útiles para el análisis de células individuales, tales como, por ejemplo, para el estudio de genomas, transcriptomas, proteomas, rutas metabólicas y similares de muestras complejas de células. En otros aspectos, los métodos y composiciones descritos en la presente descripción pueden usarse para el descubrimiento de anticuerpos mediante el apareamiento de cadenas pesadas y ligeras en células B y T individuales, así como también para la tipificación de HLA, y la tipificación de haplotipos de grandes intervalos. Aún en otros aspectos, los métodos y composiciones descritos en la presente descripción pueden usarse para controlar el impacto de las moléculas y fármacos pequeños y su efecto en muestras complejas normales o cancerosas para el descubrimiento de nuevos fármacos. Aún en otro aspecto, los métodos y la composición pueden usarse para detectar y analizar patógenos tales como bacterias o virus en muestras biológicas.

En determinadas modalidades ilustrativas, se proporcionan métodos para crear copias clonales de secuencias de códigos de barras (por ejemplo, códigos de barras degenerados) y suministrar las secuencias de códigos de barras dentro de una pluralidad de células individuales. De acuerdo con un aspecto de la invención, una pluralidad de secuencias de ácido nucleico únicas que comprenden un código de barras degenerado se amplifican sobre un soporte (por ejemplo, una perla) de manera que cada área discreta del soporte (por ejemplo, cada perla) se recubrirá con una copia clonal de una secuencia de ácido nucleico inicial (Figura 1). En consecuencia, cada área discreta de un soporte; la perla, por ejemplo, será un código de barras único con una pluralidad de oligonucleótidos de código de barras seleccionados. En determinadas modalidades ilustrativas, se realiza una PCR en emulsión, en donde las secuencias de oligonucleótidos degenerados se unen a una perla mediante el uso de una dilución máxima equivalente a una molécula por perla. La longitud del oligonucleótido del código de barras se relaciona con el tamaño de la muestra celular de interés. En general, los códigos de barras tienen al menos 3 nucleótidos de longitud. A menudo, son unos 20 nucleótidos. Por lo tanto, por ejemplo, un oligonucleótido unido al soporte que tiene una longitud total de aproximadamente 50-60 nucleótidos, incluye nucleótidos que codifican un cebador de secuenciación, 20 nucleótidos para el código de barras y un cebador de hibridación.

En algunas modalidades, el soporte es una perla. El oligonucleótido de molde inicial cargado sobre las perlas tiene una región de cebador de secuenciación (que se utilizará para facilitar la secuenciación del código de barras), una región degenerada (el código de barras real) y una región de cebador de hibridación, que tiene una secuencia complementaria a la secuencia de ácido nucleico diana o secuencias de interés. El cebador de hibridación puede ser ADN o ARN (Figura 1A). Algunas perlas pueden contener oligonucleótidos que se unen a más de un ácido nucleico diana de interés.

Después las perlas se saturan con un cebador de anclaje. (Figura 1B). El cebador de anclaje tiene la misma secuencia que la región del cebador de secuenciación del oligonucleótido molde. El cebador de anclaje sirve como el segundo extremo de cebado de la PCR, que permite la unión del producto generado durante la PCR en emulsión, a las perlas. Después, las perlas se amplifican en PCR en emulsión (Figura 1C) mediante el uso de un cebador complementario a la sección de hibridación de la molécula inicial. Cuando se completa la PCR en emulsión, el cebador de anclaje se extiende y contiene una copia del código de barras y del cebador de hibridación. Posteriormente, la perla puede purificarse a partir de las emulsiones y usarse en aplicaciones posteriores.

Una vez que se preparan las perlas con código de barras, se usan en una segunda PCR en emulsión en presencia de una célula individual. La célula está contenida dentro de su propia emulsión única, lo que permite la PCR simultánea en un ensayo individual que contiene varias células. (Figura 2). Las perlas y las células pueden introducirse entre sí de cualquier manera adecuada. Por ejemplo, por transfección mediante el uso de liposomas, o por emulsificación. Las muestras que contienen múltiples perlas y múltiples células se diluyen para lograr un máximo de una perla y una célula por reacción de PCR en emulsión. En la Figura 2, se muestra un ejemplo de un evento célula-perla.

En algunas modalidades, desde miles hasta millones de los eventos mostrados en la Figura 2 pueden realizarse en un ensayo individual, tal como un ensayo realizado en un pocillo individual. Cada célula individual es secuestrada en su propia emulsión única en presencia de una perla con código de barras. Las reacciones múltiples están en el mismo volumen de reacción para todas las células. Debido a que se analizan tantas células en un ensayo individual, el enfoque es equivalente a mezclar millones de pocillos de placas de PCR. Por lo tanto, un ensayo individual no se limita en cuanto a la cantidad de células individuales con respecto a las diana, o la cantidad de transcrito con respecto a la diana por células individuales, siempre que cada célula tenga un código de barras único tanto en una emulsión individual por célula o mediante la transfección de liposomas de una perla individual o sistema de código de barras. Ver la Figura 2A. Millones de emulsiones pueden estar presentes en un ensayo individual; es decir, en un solo pocillo.

Después de la lisis celular, la diana de ácido nucleico de interés se hibrida a las secuencias complementarias sobre el molde en la perla con código de barras. Figura 2B. La transcripción inversa, para una diana de ARN, o la extensión del cebador, para una diana de ADN, se realiza, y se adiciona un código de barras a la diana de ARN o ADN de la célula. Figura 2D. Dentro de una célula, se adiciona el mismo código de barras a todas las secuencias diana. Por lo tanto, como se muestra en la Figura 2D, la célula número 1, la perla con el código de barras número 1 ha capturado cuatro ejemplos de la secuencia diana (verde, amarillo, púrpura y rojo). Cada célula independiente en la reacción tiene un código de barras diferente. Figura 2D.

Los fragmentos de ADN de interés de las perlas con código de barras se recuperan y procesan en ensayos posteriores. Cuando la perla tiene unido ARN, se realiza la síntesis de ADNc, seguida de amplificación por PCR mediante el uso de un cebador específico del gen (o escisión de restricción, y/o ligación del adaptador, seguido por PCR) de manera similar a lo que se describió anteriormente (Kim y otros (2007) Science 316: 1481). Ver la Figura 3. Después se realiza la secuenciación del ADN mediante el uso de la tecnología de alto rendimiento. El cebador de secuenciación se usa para secuenciar el código de barras, a través del cebador de hibridación en la secuencia diana. La secuenciación de la diana informa acerca de los niveles de identidad y expresión del transcrito, u otra secuencia genómica o transcriptómica de interés. La secuencia del código de barras permite que cada secuencia diana se correlacione con la célula individual de la que se originaron las secuencias. Si bien cada transcrito que se origina en una célula tendrá la misma secuencia de código de barras, la variación en la información genómica o transcriptómica a través de la población celular se determina mediante el análisis de varias células individuales al mismo tiempo. Debido a que cada célula individual contiene un código de barras único diferente de las otras células individuales, las secuencias identificadas que tienen el mismo código de barras pueden correlacionarse con la misma célula de origen.

En determinadas modalidades, pueden obtenerse y analizarse múltiples ARNm de cada célula individual. Por ejemplo, puede usarse oligo-dT (o cebadores similares) como cebador de hibridación. Ver las Figuras 6A y 6B. La secuencia oligo-dT hibrida con las colas de poliA de los ARNm y, por lo tanto, captura simultáneamente múltiples ARN mensajeros de una célula individual. Esto permite el análisis completo o sustancialmente completo del transcriptoma de múltiples células individuales en una mezcla compleja. La caracterización de los transcriptomas de múltiples células en función de cada célula tiene una aplicación particular en los estudios que investigan cuáles son células malignas en comparación con otras muestras de cáncer. Además, en pacientes sometidos a terapia contra el cáncer, el presente enfoque permite monitorear la mutación de cada genoma y transcriptoma de cada célula antes y después del tratamiento; por ejemplo, con un fármaco, o después de la cirugía. Esta información es particularmente útil cuando se combina con medicamentos que se sabe que están afectados por la secuencia de una proteína. Por ejemplo, el inhibidor del EGFR Erbitux® (cetuximab) es inefectivo cuando se usa con determinadas mutaciones de K-ras. El presente enfoque puede usarse de forma diagnóstica para determinar, hasta el nivel de una célula individual, cuántas células en una muestra de tumor portan las mutaciones que hacen que las células sean resistentes a Erbitux®. La información acerca de las secuencias de ácido nucleico de múltiples

proteínas en cada célula tumoral es valiosa para determinar si continuar o interrumpir el tratamiento con un fármaco dado, o cambiar a un fármaco alternativo.

En otra modalidad, al menos dos oligonucleótidos que tienen diferentes cebadores de hibridación se unen a la misma perla, lo que permite capturar y secuenciar numerosos ácidos nucleicos diana en la misma célula. Para producir perlas que contienen diferentes cebadores de hibridación, se une una secuencia universal en dirección 3' del cebador de código de barras. Ver las Figuras 6A y 6B. La secuencia universal es complementaria a una región del extremo saliente sobre un segundo cebador que contiene el cebador de hibridación, que se dirige al gen de interés. Pueden usarse múltiples cebadores de hibridación, cada uno dirigido a un gen de interés diferente. La secuencia universal, común a los extremos salientes de todos los cebadores de hibridación, permite la incorporación de múltiples secuencias del cebador de hibridación sobre las perlas mediante la PCR. Ver las Figuras 6A y 6B.

Las perlas con múltiples cebadores de hibridación que se dirigen a diferentes ácidos nucleicos de interés tienen un uso particular en las aplicaciones de células inmunitarias. En una modalidad, los conjuntos específicos de oligonucleótidos complementarios dirigidos al gen que codifica las cadenas pesada y ligera de los anticuerpos de células B o su ARN se usan para capturar el apareamiento de cada una de las cadenas pesada y ligera de cada célula individual que definen cada anticuerpo específico. En otra modalidad, las secuencias que codifican los componentes del receptor de células T pueden seleccionarse y secuenciarse. Ver (Embleton y otros (1992) *Nucleic Acids Res.* 20(15):3831; Chapal y otros (1997) *Biotechniques* 23(3):518).

Aún en otras modalidades, los cebadores de hibridación se seleccionan para analizar los polimorfismos de nucleótidos simples (SNP), y para la tipificación de grandes intervalos de haplotipos (Zhang y otros, "Long-range polony haplotyping of individual human chromosome molecules," *Nat Genet.* 2006 Mar;38(3):382-7). Estos enfoques proporcionan información específica para cada célula en muestras biológicas de múltiples células.

En determinados ejemplos relacionados con el sistema inmunitario, el código de barras no es necesario si se usan estrategias para unir las cadenas pesada y ligera antes de la PCR o la escisión de las moléculas de las perlas, tal como la ligación, el acoplamiento CRE-LOX o los fragmentos de cada perla única, como se describió por Embleton y otros, (1992) *Nucleic Acids Res.* 20(15): 3831; Chapal y otros, (1997) *Biotechniques* 23(3):518, pero de tal manera que varias células individuales a la vez pueden tratarse como se describió en la presente invención.

Como se usa en la presente descripción, el término "código de barras" se refiere a una secuencia oligonucleotídica única que permite identificar una base de ácido nucleico correspondiente y/o una secuencia de ácido nucleico. En determinados aspectos, la base de ácido nucleico y/o la secuencia de ácido nucleico se ubican en una posición específica en una secuencia de polinucleótido más grande (por ejemplo, un polinucleótido unido covalentemente a una perla). En determinadas modalidades, los códigos de barras pueden tener cada uno una longitud dentro de un intervalo de 4 a 36 nucleótidos, o de 6 a 30 nucleótidos, o de 8 a 20 nucleótidos. En determinados aspectos, las temperaturas de fusión de los códigos de barras dentro de un conjunto están dentro de 10 °C entre sí, dentro de 5 °C entre sí, o dentro de 2 °C entre sí. En otros aspectos, los códigos de barras son miembros de un conjunto de hibridación cruzada mínima. Es decir, la secuencia de nucleótidos de cada miembro de tal conjunto es lo suficientemente diferente de la de cada otro miembro del conjunto de manera que ningún miembro pueda formar un dúplex estable con el complemento de cualquier otro miembro en condiciones rigurosas de hibridación. En un aspecto, la secuencia de nucleótidos de cada miembro de un conjunto de hibridación cruzada mínima difiere de las de todos los demás miembros en al menos dos nucleótidos. Las tecnologías de códigos de barras son conocidas en la técnica y se describen en Winzeler y otros, (1999) *Science* 285:901; Brenner (2000) *Genoma Biol.* 1:1 Kumar y otros, (2001) *Nature Rev.* 2: 302; Giaever y otros, (2004) *Proc. Natl. Acad. Sci. USA* 101:793; Eason y otros, (2004) *Proc. Natl. Acad. Sci. USA* 101:11046; y Brenner (2004) *Genoma Biol.* 5:240.

"Complementario" o "sustancialmente complementario" se refiere a la hibridación o apareamiento de bases o a la formación de un dúplex entre nucleótidos o ácidos nucleicos, tal como, por ejemplo, entre las dos cadenas de una molécula de ADN de doble cadena o entre un oligonucleótido cebador y un sitio de unión del cebador sobre un ácido nucleico monocatenario. Los nucleótidos complementarios son, generalmente, A y T/U, o C y G. Se dice que dos moléculas de ARN o ADN monocatenarias son sustancialmente complementarias cuando los nucleótidos de una cadena están alineados de manera óptima y comparados con las inserciones o delecciones de nucleótidos apropiadas se aparean con al menos aproximadamente el 80 % de los nucleótidos de la otra cadena, generalmente al menos aproximadamente 90 % a 95 %, y con mayor preferencia de aproximadamente 98 al 100 %. Alternativamente, existe una complementariedad sustancial cuando una cadena de ARN o ADN hibridará en condiciones de hibridación selectiva con su complemento. Típicamente, la hibridación selectiva ocurrirá cuando haya al menos aproximadamente 65 % de complementariedad en un tramo de al menos 14 a 25 nucleótidos, al menos aproximadamente 75 %, o al menos aproximadamente 90 % de complementariedad. Ver Kanehisa (1984) *Nucl. Acids Res.* 12:203.

Como se usa en la presente descripción, el término "hibridación" se refiere al proceso en el que dos polinucleótidos de cadena simple se unen de manera no covalente para formar un polinucleótido de cadena doble estable. El término "hibridación" puede referirse, además, a la hibridación de cadena triple. El polinucleótido de doble cadena resultante (generalmente) es un "híbrido" o "dúplex." Las "condiciones de hibridación", típicamente, incluirán concentraciones de sales de menos de aproximadamente 1 M, más generalmente de menos de aproximadamente 500 mM e incluso más generalmente de menos de aproximadamente 200 mM. Las temperaturas de hibridación pueden ser tan bajas como 5 °C,

pero, típicamente son mayores que 22 °C, más típicamente mayores que aproximadamente 30 °C, y con frecuencia por encima de aproximadamente 37 °C. Las hibridaciones se realizan generalmente en condiciones muy rigurosas, es decir, condiciones en las que una sonda se hibridará con su subsecuencia diana. Las condiciones rigurosas dependen de la secuencia y son diferentes en diversas circunstancias. Los fragmentos más largos pueden requerir temperaturas de hibridación mayores para la hibridación específica. Como otros factores pueden afectar la rigurosidad de la hibridación, que incluye la composición de bases y la longitud de las cadenas complementarias, la presencia de solventes orgánicos y la extensión de los apareamientos incorrecto de bases, la combinación de parámetros es más importante que la medida absoluta de uno solo. Generalmente, las condiciones rigurosas se seleccionan para que sean aproximadamente 5 °C más bajas que la T_m para la secuencia específica en condiciones de fuerza iónica y el pH definidos. Las condiciones rigurosas ilustrativas incluyen una concentración de sales de al menos 0.01 M a no más de 1 M de concentración del ion Na (u otras sales) a un pH de 7,0 a 8,3 y una temperatura de al menos 25 °C. Por ejemplo, las condiciones de 5XSSPE (NaCl 750 mM, fosfato de Na 50 mM, EDTA 5 mM, pH 7,4) y una temperatura de 25-30 °C son adecuadas para hibridaciones de sondas específicas de alelos. Para condiciones rigurosas, ver por ejemplo, Sambrook, Fritsche y Maniatis, *Molecular Cloning A Laboratory Manual*, 2ª Ed. Cold Spring Harbor Press (1989) y Anderson *Nucleic Acid Hybridization*, 1ra Ed., BIOS Scientific Publishers Limited (1999). Las frases "Hibridación específica a" o "que específicamente hibrida a" o expresiones similares se refieren a la unión, formación de dúplex, o hibridación de una molécula sustancialmente a, o solamente a una secuencia de nucleótidos particular o secuencias en condiciones rigurosas cuando esa secuencia está presente en una mezcla compleja (por ejemplo, celular total) de ADN o ARN.

"Nucleósido" como se usa en la presente descripción incluye los nucleósidos naturales, que incluyen las formas 2'-deoxi y 2'-hidroxilo, por ejemplo, como se describe en Komberg y Baker, *DNA Replication*, 2da Ed. (Freeman, San Francisco, 1992). Los "análogos", en referencia a nucleósidos, incluye nucleósidos sintéticos que tienen porciones de bases modificadas y/o porciones de azúcar modificados, por ejemplo, descritos por Scheit, *Nucleotide Analogs* (John Wiley, New York, 1980); Uhlman and Peyman, *Chemical Reviews*, 90:543-584 (1990), o similares, con la condición de que sean capaces de hibridar de manera específica. Tales análogos incluyen nucleósidos sintéticos diseñados para mejorar las propiedades de unión, reducir la complejidad, aumentar la especificidad, y similares. Los polinucleótidos que comprenden análogos con propiedades mejoradas de hibridación o resistencia a las nucleasas se describen en Uhlman y Peyman (citado anteriormente); Crooke y otros, *Exp. Opin. Ther. Patentes*, 6: 855-870 (1996); Mesmaeker y otros, *Current Opinion in Structural Biology*, 5:343-355 (1995); y similares. Los tipos ilustrativos de polinucleótidos que son capaces de mejorar la estabilidad del dúplex incluyen oligonucleótidos fosforamidatos (a los que se hace referencia en la presente descripción como "amidatos"), ácidos nucleicos peptídicos (a los que se hace referencia en la presente descripción como "PNA"), oligo-2'-O-alkilribonucleótidos, polinucleótidos que contienen C-5 propinilpirimidinas, ácidos nucleicos bloqueados (LNA) y compuestos similares. Tales oligonucleótidos están disponibles comercialmente o pueden sintetizarse mediante el uso de los métodos descritos en la literatura.

Como se usa en la presente descripción, los términos "molécula de ácido nucleico", "secuencia de ácido nucleico", "fragmento de ácido nucleico", "oligonucleótido", "fragmento de oligonucleótido" y "polinucleótido" se usan indistintamente y pretenden incluir, pero no se limitan a, una forma polimérica de nucleótidos que pueden tener varias longitudes, ya sea desoxirribonucleótidos o ribonucleótidos, o análogos de estos. Las moléculas de ácido nucleico incluyen ADN monocatenario (ADNss), ADN bicatenario (ADNds), ARN monocatenario (ARNss) y ARN bicatenario (ARNds). Las diferentes moléculas de ácido nucleico pueden tener diferentes estructuras tridimensionales, y pueden realizar diversas funciones, conocidas o desconocidas. Los ejemplos no limitantes de moléculas de ácido nucleico incluyen un gen, un fragmento de gen, un espacio genómico, un exón, un intrón, ADN intergénico (que incluye, sin limitación, ADN heterocromático), ARN mensajero (ARNm), ARN de transferencia, ARN ribosomal, ribozimas, ARN de interferencia pequeño (siRNA), miRNA, ARN nucleolar pequeño (snoRNA), ADNc, polinucleótidos recombinantes, polinucleótidos ramificados, plásmidos, vectores, ADN aislado de una secuencia, ARN aislado de una secuencia, sondas de ácido nucleico, y cebadores. Las moléculas de ácido nucleico útiles en los métodos descritos en la presente descripción pueden comprender secuencias de ácido nucleico natural y variantes de estas, secuencias de ácido nucleico artificial, o una combinación de tales secuencias.

Una secuencia de oligonucleótidos se refiere a un polímero lineal de monómeros nucleosídicos naturales o modificados unidos mediante enlaces fosfodiéster o análogos de estos. El término "oligonucleótido" generalmente se refiere a un polímero más corto, por ejemplo, que comprende de aproximadamente 3 a aproximadamente 100 monómeros, y el término "polinucleótido" generalmente se refiere a polímeros más largos, por ejemplo, que comprende de aproximadamente 100 monómeros a muchos miles de monómeros, por ejemplo, 10,000 monómeros, o más. Un "fragmento de oligonucleótido" se refiere a una secuencia de oligonucleótidos que se ha dividido en dos o más secuencias de oligonucleótidos más pequeñas. Los oligonucleótidos que comprenden sondas o cebadores usualmente tienen longitudes en el intervalo de 12 a 60 nucleótidos, y más generalmente, de 18 a 40 nucleótidos. Los oligonucleótidos y polinucleótidos pueden ser naturales o sintéticos. Los oligonucleótidos y polinucleótidos incluyen desoxirribonucleósidos, ribonucleósidos y análogos no naturales de estos, tal como sus formas anoméricas, ácidos nucleicos peptídicos (PNA) y similares, siempre y cuando sean capaces de unirse específicamente a un genoma diana por medio de un patrón regular de interacciones monómero a monómero, tales como el tipo de apareamiento de bases de Watson-Crick, apilamiento de bases, tipos de apareamiento de bases Hoogsteen o Hoogsteen inverso, o similares.

Usualmente, los monómeros nucleosídicos se unen mediante enlaces fosfodiéster. Cuando un oligonucleótido se representa mediante una secuencia de letras, tal como "ATGCCTG", se entenderá que los nucleótidos están en el orden

- 5' a 3' de izquierda a derecha y que "A" significa desoxiadenosina, "C" significa desoxicitidina, "G" significa desoxiguanosina, "T" significa desoxitimidina, y "U" significa el ribonucleósido, uridina, a menos que se indique lo contrario. Usualmente los oligonucleótidos comprenden los cuatro desoxinucleótidos naturales; sin embargo, pueden comprender, además, ribonucleósidos o análogos de nucleótidos no naturales. Está claro para los expertos en la técnica cuando los oligonucleótidos que tienen nucleótidos naturales o no naturales pueden emplearse en los métodos y procedimientos descritos en la presente descripción. Por ejemplo, cuando se requiere el procesamiento mediante una enzima, generalmente se requieren oligonucleótidos que consisten únicamente en nucleótidos naturales. Igualmente, cuando una enzima tiene requisitos de sustrato específicos de oligonucleótidos o polinucleótidos para la actividad, por ejemplo, ADN monocatenario, dúplex de ARN/ADN o similares, la selección de la composición apropiada para los sustratos de oligonucleótidos o polinucleótidos está bien dentro del conocimiento de un experto en la técnica, especialmente con la orientación de los tratados, tales como Sambrook y otros, *Molecular Cloning*, Segunda Edición (Cold Spring Harbor Laboratory, Nueva York, 1989), y referencias similares. Los oligonucleótidos y polinucleótidos pueden ser monocatenarios o bicatenarios.
- Las moléculas de ácido nucleico pueden incluir opcionalmente uno o más nucleótidos no estándar, análogos de nucleótidos y/o nucleótidos modificados. Los ejemplos de nucleótidos modificados incluyen, pero no se limitan a, diaminopurina, S²T, 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-iodouracilo, hipoxantina, xantina, 4-acetilcitosina, 5-(carboxihidroxi)metiluracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidrouracilo, beta-D-galactosilqueosina, inosina, N6-isopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-adenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta-D-manosilqueosina, 5'-metoxycarboximetiluracilo, 5-metoxiuracilo, 2-metil-D46-isopenteniladenina, ácido uracil-5-oxiacético (v), wibutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, ácido uracil-5-oxiacético metiléster, ácido uracil-5-oxiacético (v), 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil) uracilo, (acp3)w, 2,6-diaminopurina y similares. Las moléculas de ácido nucleico pueden, además, modificarse en la porción de la base (por ejemplo, en uno o más átomos que típicamente están disponibles para formar un enlace de hidrógeno con un nucleótido complementario y/o en uno o más átomos que típicamente no son capaces de formar un enlace de hidrógeno con un nucleótido complementario), la porción de azúcar o esqueleto de fosfato.
- En ciertas modalidades ilustrativas, se proporcionan polinucleótidos grandes. En determinados aspectos, se usan técnicas de aislamiento que maximizan las longitudes de polinucleótidos que se obtienen (por ejemplo, moléculas de ADN). Por ejemplo, para obtener los polinucleótidos puede usarse la lisis o desproteización *in situ* (por ejemplo, con EDTA, detergente, proteasa, cualquiera de sus combinaciones y similares) después de la inclusión en agarosa (como se realiza de forma rutinaria para la electroforesis en gel de campo pulsado).
- Las moléculas de ácido nucleico pueden aislarse de fuentes naturales o adquirirse de fuentes comerciales. Las secuencias de oligonucleótidos pueden, además, prepararse mediante cualquier método adecuado, por ejemplo, métodos estándar de fosforamida tal como los descritos por Beaucage y Carruthers ((1981) *Tetrahedron Lett.* 22: 1859) o el método del triéster de acuerdo con Matteucci y otros, (1981) *J. Am. Chem. Soc.* 103:3185), o mediante otros métodos químicos que utilizan un sintetizador de oligonucleótidos automatizado comercial o métodos de matriz de alta densidad y alto rendimiento conocidos en la técnica (ver las patentes de los EE.UU. núms. 5,602,244, 5,574,146, 5,554,744, 5,428,148, 5,264,566, 5,141,813, 5,959,463, 4,861,571 y 4,659,774). Los oligonucleótidos pre-sintetizados pueden, además, obtenerse comercialmente de una diversidad de proveedores.
- Las moléculas de ácido nucleico pueden obtenerse de una o más muestras biológicas. Como se usa en la presente descripción, una "muestra biológica" puede ser una célula individual o varias células. Una muestra biológica puede comprender un solo tipo de célula o una combinación de dos o más tipos de células. Una muestra biológica incluye, además, una colección de células que realizan una función similar, tal como las que se encuentran, por ejemplo, en un tejido. En consecuencia, determinados aspectos de la invención se dirigen a muestras biológicas que contienen uno o más tejidos. Como se usa en la presente descripción, un tejido incluye, pero no se limita a, tejido epitelial (por ejemplo, piel, el revestimiento de glándulas, intestino, piel y órganos tales como el hígado, pulmón, riñón), endotelio (por ejemplo, el revestimiento de los vasos sanguíneos y linfáticos), mesotelio (por ejemplo, el revestimiento de los espacios pleurales, peritoneales y pericárdicos), mesénquima (por ejemplo, células que llenan los espacios entre los órganos, que incluyen células adiposas, músculo, hueso, cartílago y tendón), células sanguíneas (por ejemplo, eritrocitos, granulocitos, neutrófilos, eosinófilos, basófilos, monocitos, linfocitos T (también conocidos como células T), linfocitos B (también conocidos como células B), células plasmáticas, megacariocitos y similares), neuronas, células germinales (por ejemplo, espermatozoides, ovocitos), células de líquido amniótico, placenta, células madre y similares. Una muestra de tejido incluye muestras microscópicas, así como también muestras macroscópicas. En determinados aspectos, puede obtenerse una muestra de una o más de una célula individual en cultivo, muestras metagenómicas, células madre embrionarias, células madre pluripotentes inducidas, muestras de cáncer, secciones de tejido, biopsias y similares, y cualquier combinación de éstas.
- En determinados aspectos, se proporcionan secuencias de ácido nucleico que se derivan u obtienen de uno o más organismos. Como se usa en la presente descripción, el término "organismo" incluye, pero no se limita a, un humano, un primate no humano, una vaca, un caballo, una oveja, una cabra, un cerdo, un perro, un gato, un conejo, un ratón, una rata, un jerbo, una rana, un sapo, un pez (por ejemplo, *Danio rerio*) un gusano redondo (por ejemplo, *C. elegans*) y

cualquier especie transgénica de estos. El término "organismo" incluye, además, pero no se limita a, una levadura (por ejemplo, una célula *S. cerevisiae*), una tétrada de levadura, una colonia de levadura, una bacteria, una colonia bacteriana, un virión, virosoma, partículas similares a virus y/o cultivos de estos, y similares.

El aislamiento, extracción o derivación de secuencias de ácido nucleico puede realizarse mediante cualquier método adecuado. El aislamiento de secuencias de ácido nucleico de una muestra biológica generalmente incluye el tratamiento de una muestra biológica de tal manera que las secuencias de ácido nucleico presentes en la muestra se extraen y se hacen disponibles para el análisis. Cualquier método de aislamiento que resulte en secuencias de ácido nucleico extraídas puede usarse en la práctica de la presente invención. Se entenderá que el método particular usado para extraer secuencias de ácido nucleico dependerá de la naturaleza de la fuente.

Los métodos de extracción de ADN son bien conocidos en la técnica. Un protocolo de aislamiento de ADN clásico se basa en la extracción mediante el uso de solventes orgánicos, tal como una mezcla de fenol y cloroformo, seguido de precipitación con etanol (J. Sambrook y otros, "Molecular Cloning: A Laboratory Manual," 1989, 2da Ed., Cold Spring Harbour Laboratory Press: Nueva York, N.Y.). Otros métodos incluyen: extracción de ADN mediante desalado (P. Sunnucks y otros, Genetics, 1996, 144: 747-756; SM Aljanabi y I. Martínez, Nucl. Acidos Res. 1997, 25: 4692-4693.), extracción de ADN mediante las sales de bromuro de trimetilamonio (S. Gustincich y otros, BioTechniques, 1991, 11: 298-302) y la extracción de ADN mediante tiocianato de guanidinio (JBW Hammond y otros, Biochemistry, 1996, 240: 298-300). Una variedad de kits está disponible comercialmente para extraer ADN de muestras biológicas (por ejemplo, BD Biosciences Clontech (Palo Alto, CA); Epicentre Technologies (Madison, WI); Gentra Systems, Inc. (Minneapolis, MN); MicroProbe Corp. (Bothell, WA); Organon Teknika (Durham, NC); y Qiagen Inc. (Valencia, CA)).

Además, los métodos de extracción de ARN son bien conocidos en la técnica (ver, por ejemplo, J. Sambrook y otros, "Molecular Cloning: A Laboratory Manual" 1989, 2ª Ed., Cold Spring Harbour Laboratory Press: Nueva York) y numerosos kits para la extracción de ARN de fluidos corporales están disponibles comercialmente (por ejemplo, Ambion, Inc. (Austin, TX); Amersham Biosciences (Piscataway, NJ); BD Biosciences Clontech (Palo Alto, CA); BioRad Laboratories (Hercules, CA); Dynal Biotech Inc. (Lake Success, NY); Epicentre Technologies (Madison, WI); Gentra Systems, Inc. (Minneapolis, MN); GIBCO BRL (Gaithersburg, MD); Invitrogen Life Technologies (Carlsbad, CA); MicroProbe Corp. (Bothell, WA); Organon Teknika (Durham, NC); Promega, Inc. (Madison, WI); and Qiagen Inc. (Valencia, CA)).

En determinadas modalidades ilustrativas, las secuencias de oligonucleótidos se inmovilizan sobre un soporte sólido. El soporte puede ser simples rejillas cuadradas, rejillas de ajedrez (por ejemplo, desplazadas), matrices hexagonales y similares. Los soportes adecuados incluyen, pero no se limitan a, láminas, perlas, chips, partículas, hebras, geles, láminas, tubos, esferas, recipientes, capilares, almohadillas, rodajas, películas, placas de cultivo, placas (por ejemplo, de 96 pocillos, de 48 pocillos, de 24 pocillos, de 12 pocillos, de ocho pocillos, de seis pocillos, de cuatro pocillos, un solo pocillo y similares), superficies celulares (por ejemplo, células de *S. aureus*) y similares. En diversas modalidades, un soporte sólido puede ser biológico, no biológico, orgánico, inorgánico, o cualquiera de sus combinaciones.

En determinadas modalidades ilustrativas, se proporcionan perlas y matrices basadas en perlas. Como se usa en la presente descripción, el término "perla" se refiere a una partícula discreta que puede ser esférica (por ejemplo, microesferas) o tener una forma irregular. Las perlas pueden ser tan pequeñas como de aproximadamente 0.1 µm de diámetro o tan grandes como de aproximadamente varios milímetros de diámetro. Las perlas pueden comprender una variedad de materiales que incluyen, pero no se limitan a, materiales paramagnéticos, cerámica, plástico, vidrio, poliestireno, metilestireno, polímeros acrílicos, titanio, látex, sefarosa, celulosa, nailon y similares.

De acuerdo con determinados ejemplos, un soporte (por ejemplo, una perla) puede tener grupos funcionales unidos a su superficie que pueden usarse para unir a la perla uno o más reactivos descritos en la presente descripción. Pueden unirse uno o más reactivos a un soporte (por ejemplo, una perla) mediante hibridación, unión covalente, unión magnética, unión por afinidad y similares. Las perlas recubiertas con una variedad de aditamentos están disponibles comercialmente (Dynabeads, Invitrogen). Los soportes (por ejemplo, perlas) pueden, además, funcionalizarse mediante el uso de, por ejemplo, químicas en fase sólida conocidas en la técnica (ver, por ejemplo, Patente de los EE.UU. núm. 5,919,523).

Como se usa en la presente descripción, el término "unir" se refiere tanto a interacciones covalentes como a interacciones no covalentes. Una interacción covalente es un enlace químico entre dos átomos o radicales formados porque comparten un par de electrones (es decir, un enlace simple), dos pares de electrones (es decir, un doble enlace) o tres pares de electrones (es decir, un triple enlace). Además, las interacciones covalentes son conocidas en la técnica como interacciones de pares de electrones o enlaces de pares de electrones. Las interacciones no covalentes incluyen, pero no se limitan a, interacciones de van der Waals, enlaces de hidrógeno, enlaces químicos débiles (es decir, a través de fuerzas no covalentes de corto alcance), interacciones hidrofóbicas, enlaces iónicos y similares. Una revisión de las interacciones no covalentes puede encontrarse en Alberts y otros, en Molecular Biology of the Cell, 3ra Ed., Garland Publishing, 1994.

En ciertas modalidades ilustrativas, se proporcionan métodos para amplificar secuencias de ácido nucleico. Los métodos ilustrativos para amplificar ácidos nucleicos incluyen la reacción en cadena de la polimerasa (PCR) (ver, por ejemplo, Mullis y otros, (1986) Cold Spring Harb. Simp. Cuant. Biol. 51 Pt 1:263 y Cleary y otros (2004) Nature Methods 1:241; y las patentes de los EE.UU. núms. 4.683.195 y 4.683.202), PCR de anclaje, PCR RACE, reacción en cadena de ligación (LCR)

(ver, por ejemplo, Landegran y otros, (1988) Science 241:1077-1080; y Nakazawa y otros, (1994) Proc. Natl Acad Sci. USA 91:360-364), replicación de secuencia auto sostenida (Guatelli y otros, (1990) Proc. Natl Acad Sci. USA 87:1874), sistema de amplificación transcripcional (Kwoh y otros, (1989) Proc. Natl Acad Sci. USA 86:1173), Replicasa Q-Beta (Lizardi y otros, (1988) BioTechnology 6:1197), PCR recursiva (Jaffe y otros, (2000) J. Biol. Chem. 275:2619; y Williams y otros, (2002) J. Biol. Chem. 277:7790), los métodos de amplificación descritos en las patentes de los EE.UU. Nos. 6.391.544, 6.365.375, 6.294.323, 6.261.797, 6.124.090 y 5.612.199, amplificación isotérmica (por ejemplo, amplificación de círculo rodante (RCA), amplificación de círculo rodante hiperramificado (HRCa), amplificación por desplazamiento de cadena (SDA), amplificación dependiente de helicasa (HDA), PWGA) o cualquier otro método de amplificación de ácido nucleico mediante el uso de técnicas bien conocidas por los expertos en la técnica.

La "Reacción en cadena de la polimerasa" o "PCR", se refiere a una reacción para la amplificación *in vitro* de secuencias de ADN específicas mediante la extensión simultánea del cebador de las cadenas complementarias de ADN. En otras palabras, la PCR es una reacción para hacer múltiples copias o réplicas de un ácido nucleico diana flanqueado por sitios de unión del cebador, tal reacción comprende una o más repeticiones de las siguientes etapas: (i) desnaturalización del ácido nucleico diana, (ii) hibridación de los cebadores a los sitios de unión del cebador, y (iii) extensión de los cebadores mediante una polimerasa de ácido nucleico en presencia de nucleósidos trifosfato. Usualmente, la reacción se realiza en ciclos a diferentes temperaturas optimizadas para cada etapa en un instrumento de termociclador. Las temperaturas particulares, las duraciones en cada etapa y las velocidades de cambio entre las etapas dependen de muchos factores bien conocidos por los expertos en la técnica, por ejemplo, ejemplificados mediante las referencias: McPherson y otros, Editores, PCR: A Practical Approach and PCR2: A Practical Approach (IRL Press, Oxford, 1991 y 1995, respectivamente). Por ejemplo, en una PCR convencional que utiliza Taq ADN polimerasa, un ácido nucleico diana bicatenario puede desnaturalizarse a una temperatura superior a 90 °C, los cebadores se hibridan a una temperatura en el intervalo de 50-75 °C y los cebadores se extienden a una temperatura en el intervalo de 72-78 °C.

El término "PCR" abarca formas derivadas de la reacción, que incluyen, pero no se limitan a, RT-PCR, PCR en tiempo real, PCR anidada, PCR cuantitativa, PCR multiplexada, PCR de ensamblaje y similares. Los volúmenes de reacción varían desde unos pocos cientos de nanolitros, por ejemplo, 200 nL, hasta unos pocos cientos de microlitros, por ejemplo, 200 microlitros. La "PCR de transcripción inversa" o "RT-PCR" significa una PCR que está precedida por una reacción de transcripción inversa que convierte un ARN diana en un ADN monocatenario complementario, que después se amplifica, por ejemplo, Tecott y otros, Patente de Estados Unidos núm. 5.168.038. La "PCR en tiempo real" significa una PCR para la cual se controla la cantidad de producto de reacción, es decir, el amplicón, a medida que avanza la reacción. Hay muchas formas de PCR en tiempo real que difieren principalmente en los químicos de detección usados para monitorear el producto de reacción, por ejemplo, Gelfand y otros, Patente de Estados Unidos núm. 5.210.015 ("Taqman"); Wittwer y otros, Patentes de EE.UU. núms 6.174.670 y 6.569.627 (colorantes intercalantes); Tyagi y otros, Patente de Estados Unidos núm. 5.925.517 (balizas moleculares). Las químicas de detección para PCR en tiempo real se revisaron en Mackay y otros, Nucleic Acids Research, 30: 1292-1305 (2002). La "PCR anidada" significa una PCR de dos etapas en donde el amplicón de una primera PCR se convierte en la muestra para una segunda PCR mediante el uso de un nuevo conjunto de cebadores, al menos uno de los cuales se une a una ubicación interior del primer amplicón. Como se usa en la presente descripción, "cebadores iniciales" en referencia a una reacción de amplificación anidada se refiere a los cebadores usados para generar un primer amplicón, y "cebadores secundarios" significa el uno o más cebadores usados para generar un segundo amplicón, o anidado. La "PCR multiplexada" significa una PCR en donde múltiples secuencias diana (o una única secuencia diana y una o más secuencias de referencia) se llevan a cabo simultáneamente en la misma mezcla de reacción, por ejemplo Bernard y otros, (1999) Anal. Biochem., 273:221-228 (PCR a tiempo real de dos colores). Normalmente, se emplean distintos conjuntos de cebadores para cada secuencia que se amplifica. La "PCR cuantitativa" significa una PCR diseñada para medir la abundancia de una o más secuencias diana específicas en una muestra o espécimen. Las técnicas para la PCR cuantitativa son bien conocidas por los expertos en la técnica, como se ejemplifica en las siguientes referencias: Freeman y otros, Biotechniques, 26: 112-126 (1999); Becker-Andre y otros, Nucleic Acids Research, 17: 9437-9447 (1989); Zimmerman y otros, Biotechniques, 21: 268-279 (1996); Diviacco y otros, Gene, 122: 3013-3020 (1992); Becker-Andre y otros, Nucleic Acids Research, 17:9437-9446 (1989); y similares.

En determinadas modalidades ilustrativas, se proporcionan métodos para determinar las identidades de las secuencias de las secuencias de ácido nucleico. La determinación de la secuencia de una secuencia de ácido nucleico de interés (por ejemplo, secuencias de ácido nucleico de células inmunitarias) puede realizarse mediante el uso de una variedad de métodos de secuenciación conocidos en la técnica que incluyen, pero no se limitan a, secuenciación mediante hibridación (SBH), secuenciación mediante ligación (SBL), secuenciación cuantitativa de adición de nucleótidos fluorescentes incrementales (QIFNAS), ligación y escisión escalonadas, transferencia de energía de resonancia de fluorescencia (FRET), balizas moleculares, digestión de la sonda reportera TaqMan, pirosecuenciación, secuenciación por fluorescencia *in situ* (FISSEQ), perlas FISSEQ (Patente de los EE.UU. núm. 7.425.431), secuenciación oscilante (PCT/US05/27695), secuenciación de multiplex (EE.UU. número de serie 12/027,039, presentada el 6 de febrero de 2008; Porreca y otros (2007) Nat. Methods 4:931), secuenciación de colonias polimerizadas (POLONY) (patentes de los EE.UU. núms. 6.432.360, 6.485.944 y 6.511.803, y PCT/US05/06425); secuenciación en círculo rodante en nanocuatriculas (ROLONY) (EE.UU. número de serie 12/120,541, presentada el 14 de mayo de 2008), ensayos de ligación de oligos alelo específicos (por ejemplo, ensayo de ligación de oligos (OLA), OLA en molécula de molde único mediante el uso de una sonda lineal ligada y lecturas de la amplificación en círculo rodante (RCA), sondas candados ligadas, y/o OLA en molécula de molde único mediante el uso de una sonda candado circular ligada y lecturas de la amplificación en círculo rodante (RCA) y similares. Además pueden usarse los métodos de secuenciación de alto rendimiento, por ejemplo, en secuenciación de

matriz cíclica mediante el uso de plataformas tales como Roche 454, Illumina Solexa, ABI-SOLiD, ION Torrents, Complete Genomics, Pacific Bioscience, Helicos, plataformas Polonator (Worldwide Web Site: Polonator.org) y similares. Los métodos de secuenciación de alto rendimiento se describen en EE.UU. número de serie 61/162.913. presentada el 24 de marzo de 2009. Una variedad de tecnologías de secuenciación basadas en la luz son conocidas en la técnica (Landegren y otros (1998) Genome Res. 8:769-76; Kwok (2000) Pharmacogenomics 1:95-100; y Shi (2001) Clin. Chem. 47:164-172).

En determinadas modalidades ilustrativas, se proporcionan métodos de pronóstico, diagnóstico y/o monitoreo de uno o más trastornos o enfermedades asociadas con un agente infeccioso. Los agentes infecciosos incluyen, pero no se limitan a, virus, bacterias, hongos, parásitos, proteínas infecciosas y similares.

Los virus incluyen, pero no se limitan a, virus animales de ADN o ARN. Como se usa en la presente descripción, los virus de ARN incluyen, pero no se limitan a, familias de virus tales como *Picornaviridae* (por ejemplo, poliovirus), *Reoviridae* (por ejemplo, rotavirus), *Togaviridae* (por ejemplo, virus de la encefalitis, virus de la fiebre amarilla, virus de la rubeola), *Orthomyxoviridae* (por ejemplo, virus de la influenza), *Paramyxoviridae* (por ejemplo, virus sincitial respiratorio, virus del sarampión, virus de las paperas, virus parainfluenza), *Rhabdoviridae* (por ejemplo, rabies virus), *Coronaviridae*, *Bunyaviridae*, *Flaviviridae*, *Filoviridae*, *Arenaviridae*, *Bunyaviridae* y *Retroviridae* (por ejemplo, virus linfotrópico de las células T humana (HTLV), virus de la inmunodeficiencia humana (HIV)). Como se usa en la presente descripción, los virus de ADN incluyen, pero no se limitan a, familias de virus tales como *Papovaviridae* (por ejemplo, virus del papiloma), *Adenoviridae* (por ejemplo, adenovirus), *Herpesviridae* (por ejemplo, virus del herpes simple), y *Poxviridae* (por ejemplo, virus de la viruela).

Las bacterias incluyen, pero no se limitan a, bacterias gram positivas, bacterias gram negativas, bacterias ácido resistentes y similares.

Como se usa en la presente descripción, las bacterias gram positivas incluyen, pero no se limitan a, *Actinomedurae*, *Actinomyces israelii*, *Bacillus anthracis*, *Bacillus cereus*, *Clostridium botulinum*, *Clostridium difficile*, *Clostridium perfringens*, *Clostridium tetani*, *Corynebacterium*, *Enterococcus faecalis*, *Listeria monocytogenes*, *Nocardia*, *Propionibacterium acnes*, *Staphylococcus aureus*, *Staphylococcus epiderm*, *Streptococcus mutans*, *Streptococcus pneumoniae* y similares.

Como se usa en la presente descripción, las bacterias gram negativas incluyen, pero no se limitan a, *Afpia felis*, *Bacteriodes*, *Bartonella bacilliformis*, *Bordetella pertussis*, *Borrelia burgdorferi*, *Borrelia recurrentis*, *Brucella*, *Calymatobacterium granulomatis*, *Campylobacter*, *Escherichia coli*, *Francisella tularensis*, *Gardnerella vaginalis*, *Haemophilus aegyptius*, *Haemophilus ducreyi*, *Haemophilus influenzae*, *Helicobacter pylori*, *Legionella pneumophila*, *Leptospira interrogans*, *Neisseria meningitidis*, *Porphyromonas gingivalis*, *Providencia stuartii*, *Pseudomonas aeruginosa*, *Salmonella enteridis*, *Salmonella typhi*, *Serratia marcescens*, *Shigella boydii*, *Streptobacillus moniliformis*, *Streptococcus pyogenes*, *Treponema pallidum*, *Vibrio cholerae*, *Yersinia enterocolitica*, *Yersinia pestis* y similares.

Como se usa en la presente descripción, las bacterias ácido resistentes incluyen, pero no se limitan a, *Mycobacterium avium*, *Mycobacterium leprae*, *Mycobacterium tuberculosis* y similares.

Como se usa en la presente descripción, otras bacterias que no entran en las otras tres categorías incluyen, pero no se limitan a, *Bartonella henselae*, *Chlamydia psittaci*, *Chlamydia trachomatis*, *Coxiella burnetii*, *Mycoplasma pneumoniae*, *Rickettsia akari*, *Rickettsia prowazekii*, *Rickettsia rickettsii*, *Rickettsia tsutsugamushi*, *Rickettsia typhi*, *Ureaplasma urealyticum*, *Diplococcus pneumoniae*, *Ehrlichia chaffeensis*, *Enterococcus faecium*, *Meningococci* y similares.

Como se usa en la presente descripción, los hongos incluyen, pero no se limitan a, *Aspergilli*, *Candidae*, *Candida albicans*, *Coccidioides immitis*, *Cryptococci*, y combinaciones de estos.

Como se usa en la presente descripción, los microbios parásitos incluyen, pero no se limitan a, *Balantidium coli*, *Cryptosporidium parvum*, *Cyclospora cayatanensis*, *Encephalitozoa*, *Entamoeba histolytica*, *Enterocytozoon bieneusi*, *Giardia lamblia*, *Leishmaniae*, *Plasmodii*, *Toxoplasma gondii*, *Trypanosomae*, ameba trapezoidal y similares.

Como se usa en la presente descripción, los parásitos incluyen gusanos (por ejemplo, helmintos), particularmente gusanos parásitos que incluyen, pero no se limitan a, Nematoda (gusanos redondos, por ejemplo, tricocéfalos, anquilostomas, oxiuros, ascárides, filaridas y similares), Cestoda (por ejemplo, tenias)

Como se usa en la presente descripción, las proteínas infecciosas incluyen priones. Los trastornos causados por priones incluyen, pero no se limitan a, trastornos humanos tales como la enfermedad de Creutzfeldt-Jakob (ECJ) que incluye, por ejemplo, la enfermedad iatrogénica de Creutzfeldt-Jakob (ICJD), la variante de la enfermedad de Creutzfeldt-Jakob (vCJD), la enfermedad de Creutzfeldt-Jakob familiar (fCJD), y la enfermedad de Creutzfeldt-Jakob esporádica (sCJD), síndrome de Gerstmann-Strausler-Scheinker (GSS), insomnio familiar fatal (FFI), insomnio fatal esporádico (SFI), kuru, y similares, así como también trastornos en animales tales como la tembladera (ovejas y cabras), encefalopatía espongiforme bovina (EEB) (ganado bovino), encefalopatía de visón transmisible (TME) (visón), enfermedad de desgaste crónico (CWD) (alce, ciervo mula), encefalopatía espongiforme felina (gatos), encefalopatía ungulada exótica (EUE) (nyala, oryx, kudu mayor), encefalopatía espongiforme del avestruz y similares.

En determinadas modalidades ilustrativas, se proporcionan métodos de pronóstico, diagnóstico y/o monitoreo de uno o más trastornos de proliferación celular. Los trastornos de proliferación celular pretenden incluir trastornos asociados con la proliferación rápida. Como se usa en la presente descripción, el término "trastorno proliferativo celular" incluye trastornos caracterizados por una proliferación no deseada o inapropiada de uno o más subconjuntos de células en un organismo multicelular. El término "cáncer" se refiere a varios tipos de neoplasias malignas, la mayoría de las cuales pueden invadir los tejidos circundantes, y pueden hacer metástasis en diferentes sitios (ver, por ejemplo, PDR Medical Dictionary 1ª edición (1995)). Los términos "neoplasma" y "tumor" se refieren a un tejido anormal que crece mediante la proliferación celular más rápidamente de lo normal. *Id.* Tal tejido anormal muestra una falta parcial o completa de organización estructural y coordinación funcional con el tejido normal que puede ser benigno (es decir, tumor benigno) o maligno (es decir, tumor maligno).

El lenguaje "tratamiento de trastornos de proliferación celular" está destinado a incluir la prevención de la inducción, el inicio, el establecimiento o el crecimiento de neoplasmas en un sujeto o una disminución en el crecimiento de neoplasmas preexistentes en un sujeto. El lenguaje puede describir, además, la inhibición de la invasión de células neoplásicas en tejidos vecinos o la metástasis de un neoplasma de un sitio a otro. Los ejemplos de los tipos de neoplasmas que pretenden abarcar la presente invención incluyen, pero no se limitan a, los neoplasmas asociados con cánceres de mama, piel, huesos, próstata, ovarios, útero, cuello uterino, hígado, pulmón, cerebro, laringe, vesícula biliar, páncreas, recto, paratiroides, tiroides, glándulas suprarrenales, sistema inmunológico, tejido nervioso, cabeza y cuello, colon, estómago, bronquios y/o riñones.

Los trastornos proliferativos celulares pueden incluir, además, trastornos asociados con la hiperproliferación de las células del músculo liso vascular, tales como trastornos cardiovasculares proliferativos, por ejemplo, aterosclerosis y reestenosis. Los trastornos de proliferación celular pueden incluir, además, trastornos tales como trastornos proliferativos de la piel, por ejemplo, ictiosis ligada al X, psoriasis, dermatitis atópica, dermatitis alérgica de contacto, hiperqueratosis epidermolítica, y dermatitis seborreica. Los trastornos de proliferación celular pueden incluir, además, trastornos tales como enfermedad poliquística renal autosómica dominante (ADPKD), mastocitosis, y trastornos de proliferación celular causados por agentes infecciosos tales como virus.

En determinadas modalidades ilustrativas, se proporcionan métodos de pronóstico, diagnóstico y/o monitoreo de uno o más trastornos autoinmunitarios. Como se usa en la presente descripción, el término "trastorno autoinmunitario" es una enfermedad o trastorno causado por un sujeto que produce una respuesta inmunitaria inadecuada contra sus propios tejidos. Como se usa en la presente descripción, un trastorno autoinmunitario incluye, pero no se limita a, trastornos como la enfermedad de Addison, alopecia areata, espondilitis anquilosante, síndrome antifosfolípido, anemia hemolítica autoinmunitaria, hepatitis autoinmunitaria, enfermedad autoinmunitaria del oído interno (EAED), síndrome linfoproliferativo autoinmunitario (ALPS), púrpura trombocitopénica autoinmunitaria (ATP), enfermedad de Balo, enfermedad de Bechet, bulous penfigoide, cardiomiopatía, dermatitis herpetiforme de la enfermedad celiaca, síndrome de fatiga crónica por disfunción inmunitaria (CFIDS), polineropatía desmielinizante inflamatoria crónica, penfigoide cicatricial, enfermedad de aglutinación por frío, síndrome de CREST, enfermedad de Crohn, enfermedad de Degos, dermatomiositis, dermatomiositis juvenil, lupus discoide, crioglobulinemia esencial mixta, fibromialgia-fibromiositis, enfermedad de Graves, Guillain-Barre, tiroiditis de Hashimoto, fibrosis pulmonar idiopática, púrpura trombocitopenia idiopática (ITP), nefropatía por IgA, diabetes dependiente de insulina, artritis juvenil, liquen plano, lupus, enfermedad de Ménière, enfermedad del tejido conectivo mixta, esclerosis múltiple, miastenia gravis, pénfigo vulgar, anemia perniciosa, poliarteritis nudosa, policondritis, síndromes poliglandular, polimialgia reumática, polimiositis y dermatomiositis, agammaglobulinemia primaria, cirrosis biliar primaria, psoriasis, fenómeno de Raynaud, síndrome de Reiter, fiebre reumática, artritis reumatoide, sarcoidosis, escleroderma, síndrome de Sjögren, síndrome de la persona rígida, artritis de Takayasu, arteritis temporal/arteritis de células gigantes, colitis ulcerativa, vasculitis, vitiligo, granulomatosis de Wegener y similares (Ver la American Autoimmune Related Diseases Association, Inc. sitio de internet: aarda.org).

Debe entenderse que las modalidades de la presente invención que se han descrito son meramente ilustrativas de algunas de las aplicaciones de los principios de la presente invención. Los expertos en la técnica pueden realizar numerosas modificaciones basadas en las enseñanzas presentadas en la presente descripción sin apartarse del verdadero espíritu y alcance de la invención.

Ejemplo 1

Preparación de Perlas con Código de Barras

a. Carga del Oligonucleótido Molde con Código de Barras sobre las Perlas

Las perlas (1 μ M; perlas de carboxílico C1 de 1 micra) se suspendieron mediante agitación con vortex y se transfirieron en un volumen de 80 μ l a un tubo de silicona de 1,5 mL (Ambion). Las perlas se lavaron dos veces con 2X (el tampón de unión y lavado contiene Tris-HCl 10 mM pH7,5, EDTA 1 mM, NaCl 2 M; "B&W"). Las perlas se aislaron mediante el uso de imanes entre los lavados. Las secuencias de nucleótidos se enumeran en la Tabla 4. Las perlas lavadas se suspendieron en 100 μ l de B&W al que se añadió el oligonucleótido molde oligo dT con código de barras (HSCT_BC_anclaje1) a las concentraciones que se muestran en la Tabla 1.

Tabla 1.

Tubo	Concentración de Reserva del Cebador	Volumen del cebador (μl)
1	100 pM	80
2	10 pM	80
3	1 pM	80
4	0.1 pM	80

El oligonucleótido molde y las perlas se incubaron en un rotador durante 20 minutos, después se lavaron dos veces con 200 μl de B&W 1x y posteriormente se suspendieron en 100 μl de B&W 2x.

b. Saturación de las Perlas con el Cebador de Anclaje

Las perlas, precargadas con el oligonucleótido molde como en el Ejemplo 1 (a) anterior, se incubaron en un rotador durante 20 minutos con una mezcla del cebador de anclaje (HSCT_Perla_anclaje1 1 mM y Tampón B&W 1x) para recubrir las perlas con el cebador de anclaje y después se lavaron dos veces con 200 μl de B&W 1x, una vez en 200 μl de TE, y después se suspendieron en 100 μl de TE. El cebador de anclaje tiene una biotina en el extremo 5' que se une a las perlas recubiertas con estreptavidina. Típicamente, el 30 % de las perlas tienen un oligonucleótido. En esas perlas, el 100 % o sustancialmente todos los cebadores de anclaje se extienden típicamente.

c. PCR en emulsión para la Síntesis del oligonucleótido a partir del Cebador de Anclaje

La mezcla acuosa y la mezcla oleosa se prepararon como se describe en la Tabla 2.

Tabla 2.

MEZCLA ACUOSA		
Componente	Volumen para 1 tubo (μl)	Volumen para 4.5 tubos (μl)
10 x tampón de PCR (Enzymatics)	96	432
MgCl ₂ 50 mM	242	1089
Mezcla dNTP 25 mM	135	607
HSCT_dA-rev_emulsión cebador 2 mM	6	27
MEZCLA OLEOSA		
Componente	Volumen para 1 tubo (μl)	Volumen para 4.5 tubos (μl)
Tegosoft DEC	4.4	19.8
Aceite mineral	1.2	5.4
ABIL WE09	425	1.9

Ambas soluciones se mezclaron mediante agitación con vortex. La mezcla de aceite se dejó desgasificar y después se colocaron porciones de 5,5 mL en tubos de ensayo de aluminio de 50 mL recubiertos con teflón.

Las emulsiones se hicieron mediante la adición de 800 μl de la mezcla de PCR, 100 μl de Enzymatics Taq (5U/μl), mediante la agitación con vortex y centrifugación rápida e inmediatamente después se adicionó 60 mL de perlas ancladas con código de barras, seguido de agitación con vortex y centrifugación. La mezcla de 960 mL se transfirió a un tubo de aceite y se agitó con vortex durante 2,25 min a 2200 rpm, lo que fue seguido por una PCR en emulsión mediante el uso de las siguientes etapas del protocolo de PCR:

- 94 °C durante 5 min
- 94 °C durante 15 segundos
- 58 °C durante 30 segundos
- 70 °C durante 75 segundos
- Se repiten los ciclos a partir de la etapa b 119 veces
- 72 °C durante 2 min
- Incubar a ~ 10 °C hasta que esté listo para usar.

La formación de una emulsión se confirmó mediante la comprobación bajo un microscopio de que se obtuvo una consistencia cremosa blanca cuando se añadió una mezcla de emulsionante/aceite (240 µl de emulsionante:960 µl de aceite, o 480 µl de emulsionante:720 µl de aceite) a una capa acuosa (384 µl) y se agitó con vortex a 4 °C durante 5 minutos. Los resultados se muestran en las Figuras 7A-7E.

En un experimento similar, se usaron perlas de 3 micras Dynal M270 en condiciones similares y se lograron resultados similares.

Además, el código de barras se logró de la siguiente manera.

Tabla 3.

MEZCLA ACUOSA		
Componente	Concentración final	Volumen por tubo (µl)
dH2O	-	520.4
10 x tampón de PCR (Enzymatics)	1x	80
Mezcla dNTP 25 mM	2mM	64
HSCT_dA_rev_emulsión cebador 2 mM	10 µM	4
BSA (Sigma) 30 % (w/v)	0.06	1.6
MEZCLA OLEOSA		
Componente	Volumen para 1 tubo (µl)	Volumen para 4.5 tubos (µl)
Tegosoft DEC	4.4	19.8
Aceite mineral	1.2	5.4
ABIL WE09	425	1.9

La mezcla acuosa se agitó con vortex, después de adicionaron 0,6 mL de mezcla por cada tubo de 1,5 mL (Ambion; antiadherente). Se adicionaron 50 µl de perlas de anclaje M280 HSCT por tubo, después los tubos se sonicaron durante 3 ciclos de 10 segundos. Después de la sonicación, los tubos se colocaron en hielo y se adicionaron 80 µl de Taq Polimerasa (5U/µl) por tubo. Los tubos se agitaron con vortex nuevamente y se colocaron en hielo. Se adicionaron 800 µl de la mezcla a la fase oleosa, los tubos se agitaron con vortex y se realizó la PCR como se describió en el Ejemplo 1, parte c. En experimentos similares se usaron placas de 96 pocillos. Cada pocillo contenía 55 µl/pocillo de la mezcla.

Nombre de la Secuencia	SEQ ID NO:	Secuencia
HSCT_BC_anclaje1	4	/52-Bio/ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT NNN NNN NNN NNN NNN NNN NNC AGC TTT TTT TTT TTT TTT TTT TTT T
HSCT_Perla_anclaje1	5	/52-Bio/ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT
HSCT_clonaltest_B C_seq	6	/5Phos/AGA TCG GAA GAG CGT CGT GTA
HSCT_dA_rev_emulsión cebador	7	AAA AAA AAA AAA AAA AAA AAA AAA ACG AC
HSCT_BC_anclaje_r ev (no-bio)	8	AAA AAA AAA AAA AAA AAA AAA AAA AGC TGN NNN NNN NNN NNN NNN NNN NAG ATC GGA AGA GCG TCG TGT AGG GAA AGA GTG T
Perlas unidas a un oligonucleótido con Código de Barras	9	PERLA/52-Bio/ACACTCTTTCCCT ACACGACGCTCTTCCGA TCT NNN NNN NNN NNN NNN NNN NNC AGC TTT TTT TTT TTT TTT TTT TTT T ATGTGCTGCGAGAAGGCTAGA/5Phos/

La Tabla 4 muestra las secuencias usadas en el Ejemplo 1.

La secuencia final unida a la perla en el Ejemplo 1 se muestra en la SEQ ID NO:9. La perla está conectada en dirección 5' a 3' al oligonucleótido que codifica la secuencia del cebador de anclaje, el código de barras (N20) y un cebador oligo dT.

Ejemplo 5

Introducción de una Perla Única con Código de Barras por Célula

5 La Figura 4 demuestra la introducción de perlas que portan oligonucleótidos con código de barras únicos en células individuales. Aquí, después de la PCR en emulsión las perlas se secuenciaron para una base de su código de barras con el propósito de mostrar que cada perla tiene un código de barras único y demostrar la clonalidad. Cada nucleótido es rastreado por diferentes fluoróforos como se describió anteriormente (Porreca y otros, (2006) Curr. Protoc. Mol. Biol. Capítulo 7: Unidad 78). Cy 5 muestra la presencia de un nucleótido de adenina en la posición uno del código de barras. Cy3 muestra la presencia de un nucleótido de timina en la posición uno del código de barras, Texas Red (Txred) muestra la presencia de un nucleótido de citosina en la posición uno del código de barras. El isotiocianato de fluoresceína (FITC) muestra la presencia de una guanina en la posición uno del código de barras. Se muestra la superposición de imágenes de los cuatro fluoróforos para una posición única en los códigos de barras y se demuestra la clonalidad. La clonalidad se refiere a cada perla única que alberga un único código de barras, que se amplificó con éxito sobre la perla. Si las perlas hubieran contenido múltiples códigos de barras; es decir, que hubieran sido no clonal (por ejemplo, que tienen múltiples moldes de código de barras cargados sobre la perla por accidente), la superposición habría demostrado más de un color de fluoróforo por perla cuando se rastrea una posición única en el código de barras durante la secuenciación. La secuenciación completa del código de barras, que permite la correlación con la célula, se realiza mediante múltiples ciclos sucesivos de secuenciación para cada posición de nucleótido.

20 El análisis con microscopía óptica de luz blanca de las perlas y la reacción de la emulsión muestra que el molde de inicio y la perla en la emulsión se diluyeron correctamente para lograr un máximo de una perla o menos por emulsión y un molde o menos por perla.

Ejemplo 6

Introducción de los Oligonucleótidos con Código de Barras únicos sobre un Soporte de Rejilla

30 Se hicieron múltiples copias del mismo código de barras único para el análisis de células individuales mediante el producto de amplificación en círculo rodante (RCA) (Rolony) a partir de un oligonucleótido inicial circularizado con código de barras único (Figura 5). Ver la Solicitud Publicada en los EE.UU. núm. 20090018024. El Rolony con código de barras único se escinde en oligonucleótidos con código de barras dirigidos cuando se incuba en presencia de un fragmento de ADN complementario compatible con restricción y enzima de restricción. Además, la escisión puede realizarse, por ejemplo, en liposomas o dentro de las emulsiones. Los liposomas que contenían oligonucleótidos con códigos de barras se fusionaron después con las células, lo que permitió que el cebador de hibridación se hibridara con el ácido nucleico diana de interés en cada célula, como se describió en el enfoque basado en perlas. La Figura 5 muestra el rastreo de Rolony (similar al rastreo de las perlas con código de barras, pero ordenado en una rejilla) para demostrar la eficiencia en la generación de Rolony únicamente clonal con código de barras. La Figura 5 demuestra que el rolony se amplifica clonalmente, porque para cada rastreo de una sola posición solo se superpone un fluoróforo para esa posición. La secuenciación posterior de las otras posiciones de nucleótidos puede realizarse para identificar los códigos de barras completos (usados para correlacionarse con la célula individual de origen) e identificar los transcritos capturados.

REIVINDICACIONES

1. Un método para analizar una o más secuencias de ácido nucleico diana de una célula individual que comprende: suministrar una perla mediante el uso de una emulsión o liposomas a una célula individual, en donde la perla comprende un oligonucleótido unido esta, el oligonucleótido comprende una región de cebador de secuenciación, un código de barras, y una región de cebador de hibridación, en donde el código de barras en el oligonucleótido es único para la célula individual, y en donde la región del cebador de hibridación comprende una secuencia complementaria a una secuencia de ácido nucleico diana e hibrida con la secuencia de ácido nucleico diana de la célula individual.
2. El método de conformidad con la reivindicación 1, en donde la región del cebador de hibridación comprende una secuencia oligo-d(t).
3. El método de conformidad con la reivindicación 1 o 2, en donde la región del cebador de hibridación comprende, además, una secuencia de unión universal.
4. El método de conformidad con cualquiera de las reivindicaciones 1-3, en donde en el oligonucleótido, la región del cebador de secuenciación está en sentido 5' con respecto al código de barras y la región del cebador de hibridación está en sentido 3' con respecto al código de barras.
5. El método de conformidad con cualquiera de las reivindicaciones anteriores, en donde el código de barras comprende de 4 nucleótidos a 36 nucleótidos, preferentemente, de 8 a 20 nucleótidos.
6. El método de conformidad con cualquiera de las reivindicaciones anteriores, en donde la perla se aísla con la célula individual, y en donde la célula individual aislada es una célula B, una célula T o una célula bacteriana.
7. El método de conformidad con la reivindicación 6, en donde la célula individual es una célula inmunitaria, una célula de un tumor, o una célula de un sujeto con un tumor.
8. El método de conformidad con cualquiera de las reivindicaciones anteriores, en donde la perla y la célula individual se aíslan en una emulsión de aceite y agua, preferentemente, en donde la célula individual se lisa después del aislamiento en la emulsión de aceite y agua.
9. El método de conformidad con cualquiera de las reivindicaciones anteriores, que comprende además transcribir de forma inversa o extender una o más secuencias de ácido nucleico diana hibridadas para producir cadenas complementarias que incluyen el código de barras.
10. El método de conformidad con la reivindicación 9, que comprende, además, amplificar las cadenas complementarias que incluyen el código de barras.
11. El método de conformidad con la reivindicación 10, que comprende amplificar y secuenciar las cadenas complementarias que incluyen el código de barras.
12. El método de conformidad con la reivindicación 11, en donde la secuenciación se realiza sobre las cadenas complementarias que incluyen el código de barras de una pluralidad de células individuales, la pluralidad que varía desde miles de células a millones de células.
13. El método de conformidad con la reivindicación 12, que comprende además correlacionar al menos una de las cadenas secuenciadas que incluye el código de barras a una célula individual.
14. El método de conformidad con cualquiera de las reivindicaciones anteriores, en donde la una o más secuencias de ácido nucleico diana comprenden ARNm, un polinucleótido que codifica un componente del receptor de células T, un polinucleótido que tiene una secuencia de ácido nucleico de células inmunitarias, un polinucleótido que codifica un componente de HLA, un polinucleótido que comprende una secuencia indicativa de haplotipo, un polinucleótido que codifica un componente del receptor de células B, un polinucleótido de cadena pesada o un polinucleótido de cadena ligera.
15. El método de conformidad con cualquiera de las reivindicaciones anteriores, que comprende analizar una pluralidad de células, donde cada célula de la pluralidad es una célula individual secuestrada con una perla.
16. El método de conformidad con la reivindicación 15, en donde cada célula de la pluralidad es una célula individual secuestrada con una perla que tiene un código de barras único.
17. El método de conformidad con cualquiera de las reivindicaciones anteriores, en donde la perla comprende una pluralidad de oligonucleótidos unidos a ella.

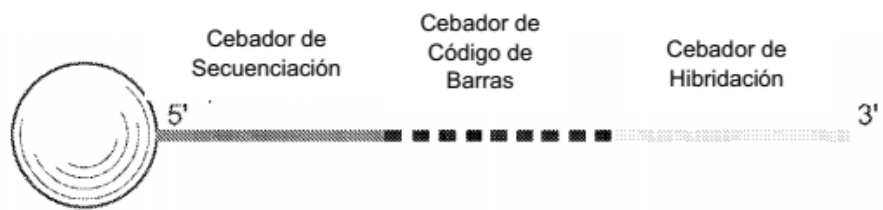


Figura 1A

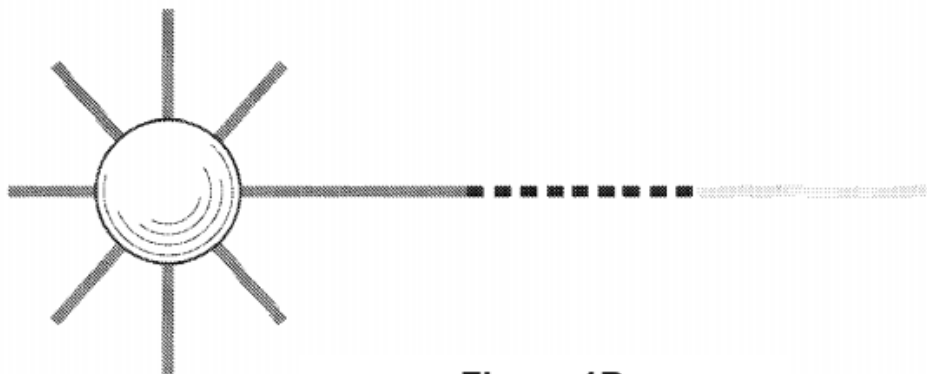


Figura 1B

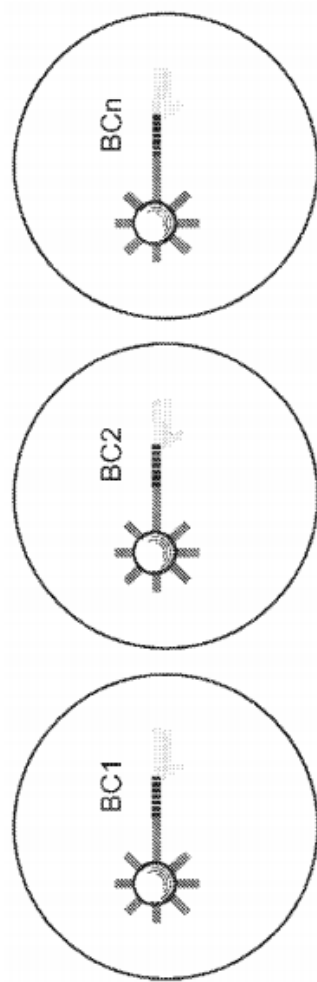


Figura 1C

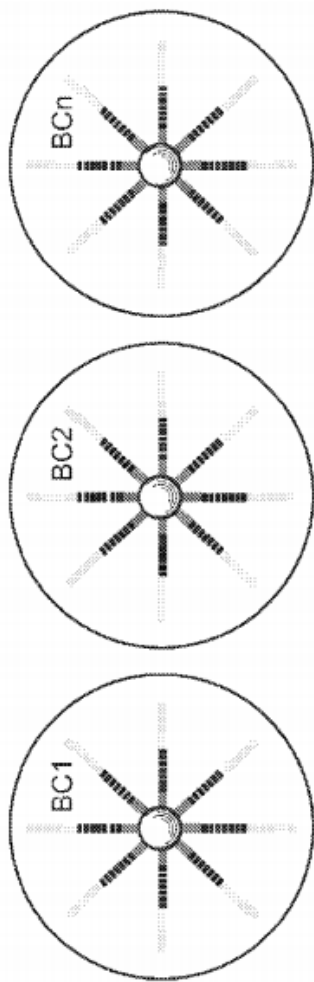


Figura 1D

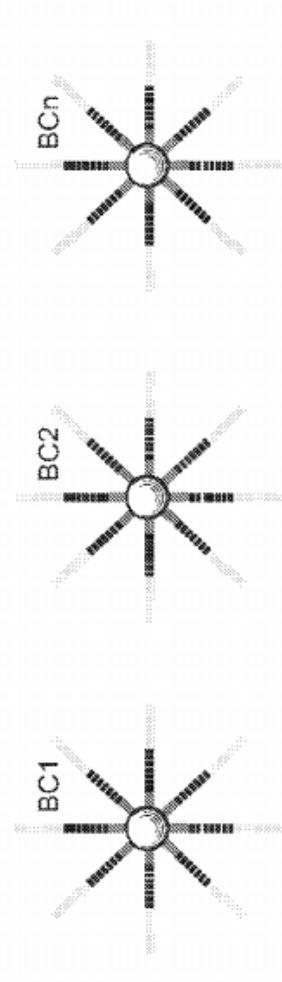


Figura 1E

Figura 2A

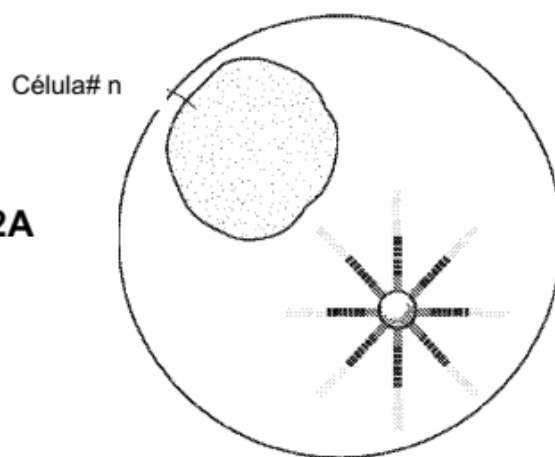


Figura 2B

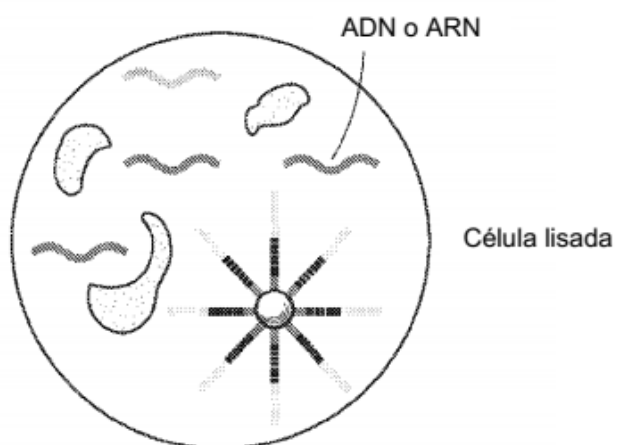
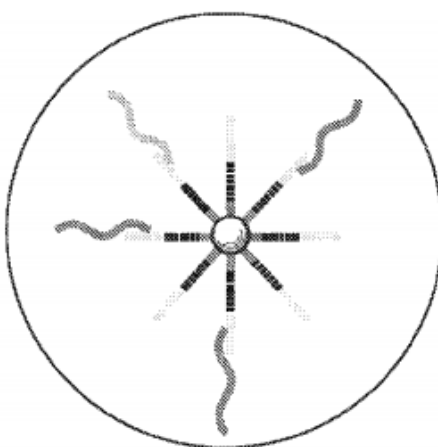
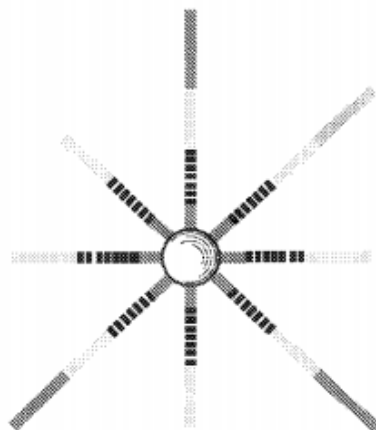


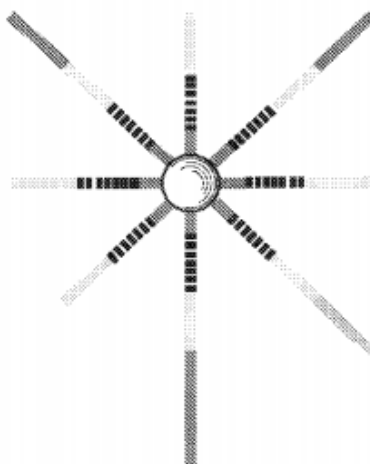
Figura 2C



Célula # 1 n
Perla con código de barra #n



Célula # 2
Perla con código de barra #2



Célula # 1
Perla con código de barra #1

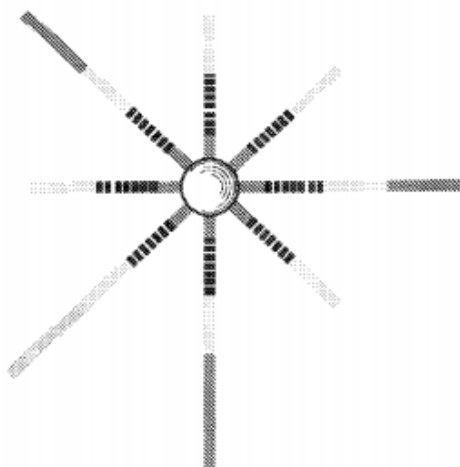


Figura 2D



Figura 3A



Figura 3B

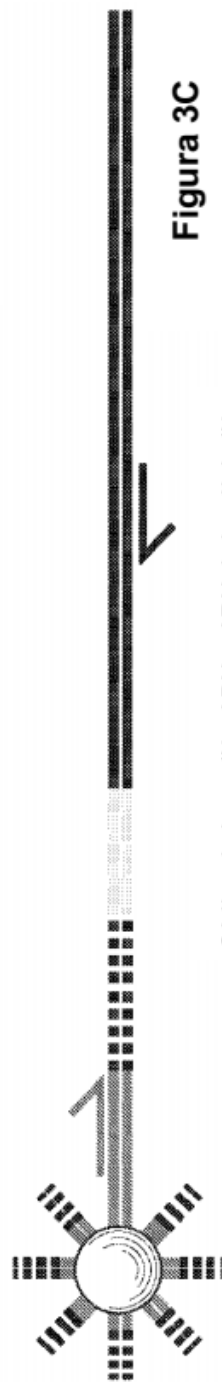


Figura 3C

Código de barra #1 = ADN o ARN de la célula #1

Código de barra #2 = ADN o ARN de la célula #2

Código de barra #n = ADN o ARN de la célula #n

Figura 3D

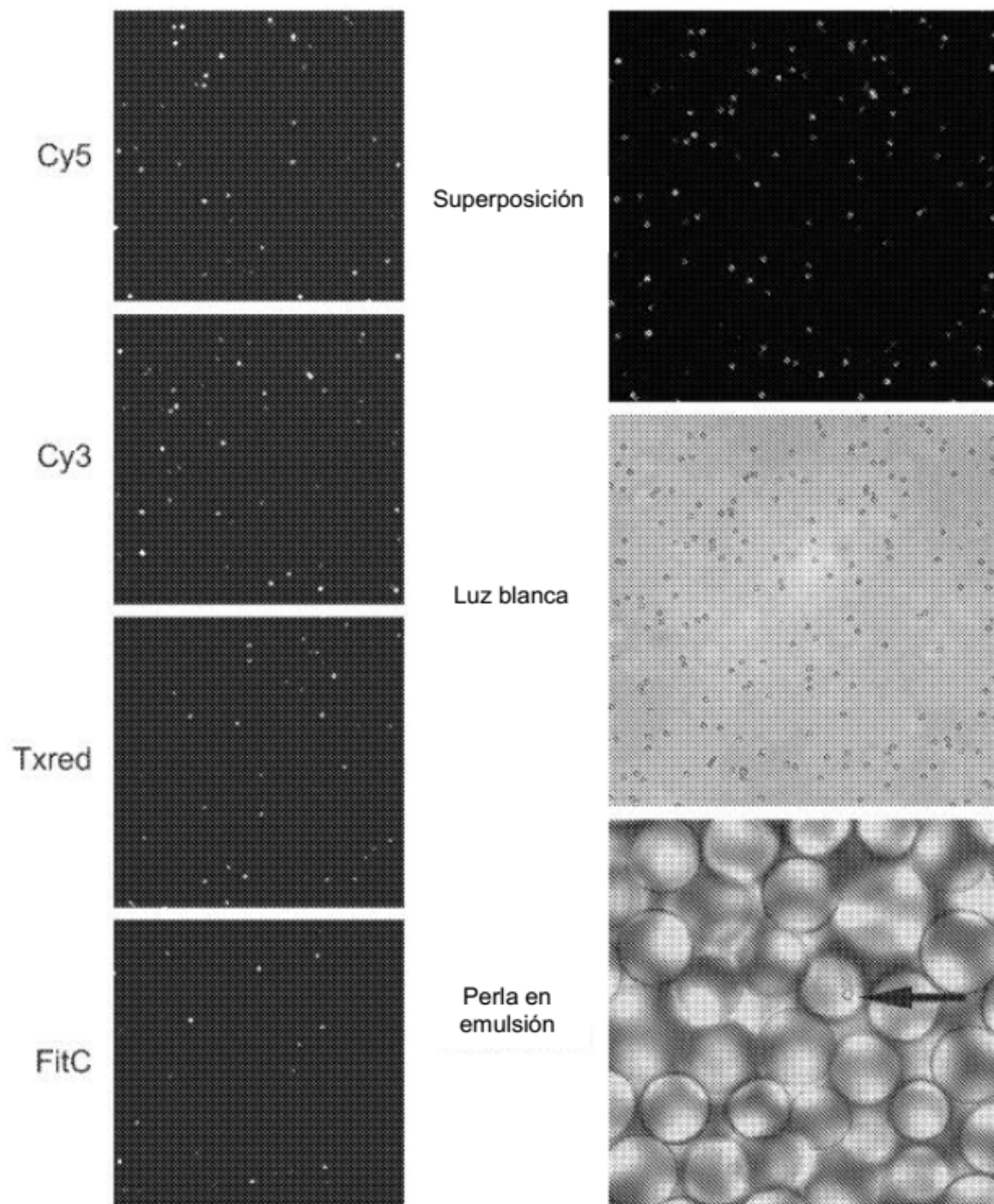


Figura 4

Figura 5A

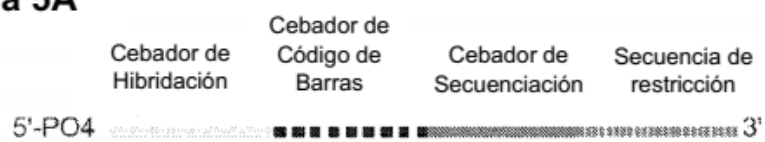


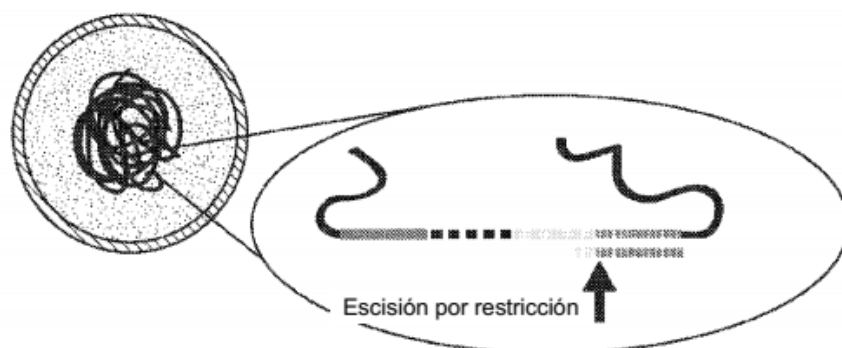
Figura 5B



Figura 5C



Figura 5D



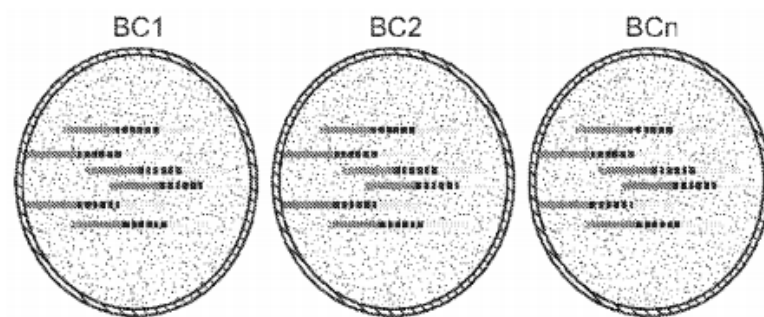


Figura 5E

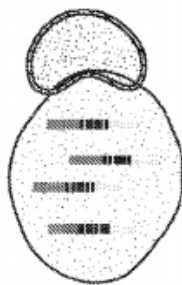


Figura 5F

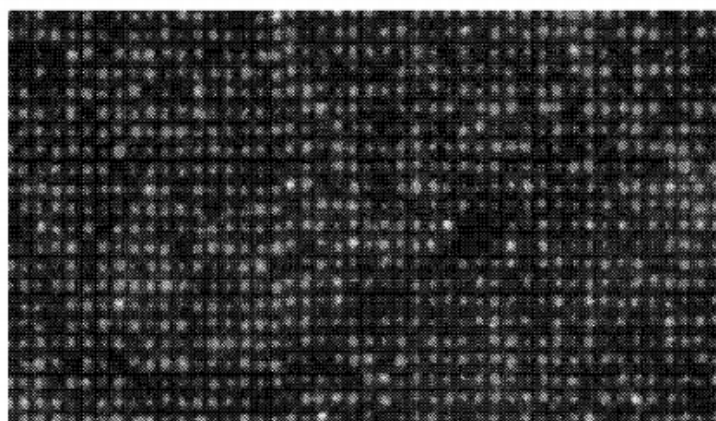


Figura 5G

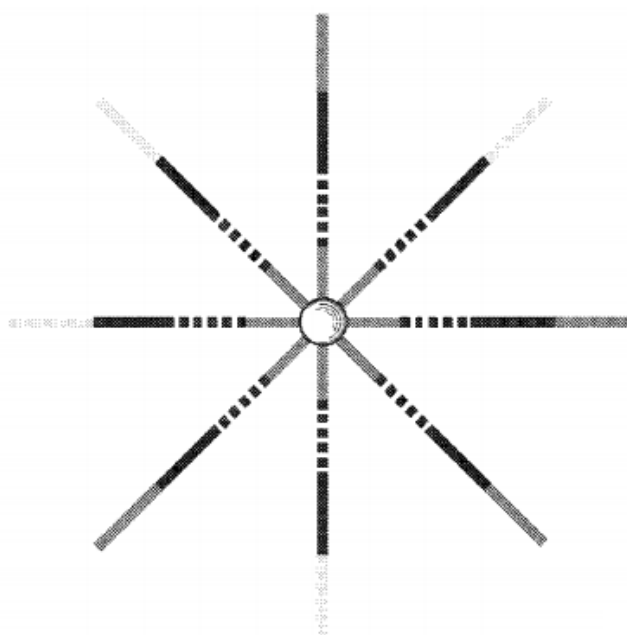
Oligonucleótido molde unido al código de barras



Figura 6A

Perla resultante con código de barras

Diversos cebadores de hibridación



Un cebador de hibridación

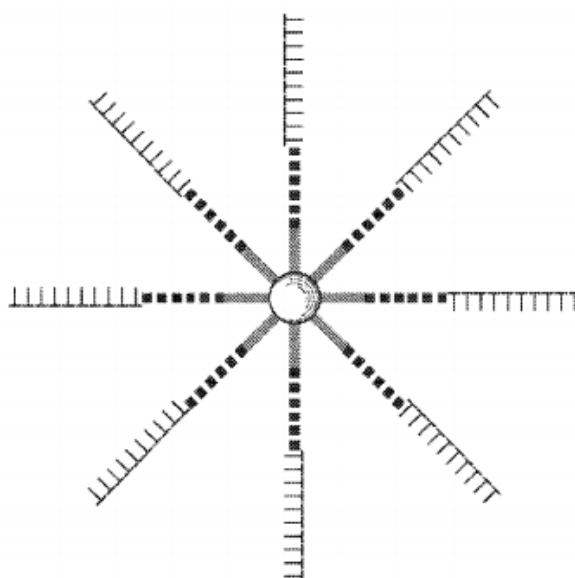


Figura 6B

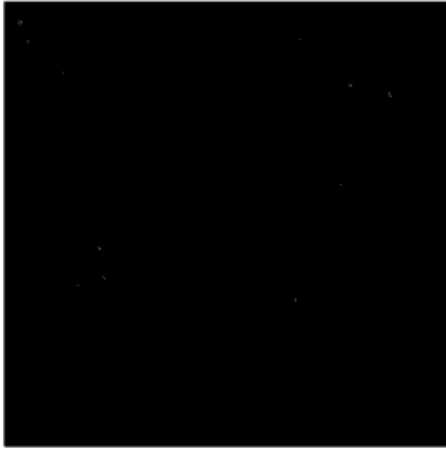


Figura 7B



Figura 7D



Figura 7A



Figura 7C

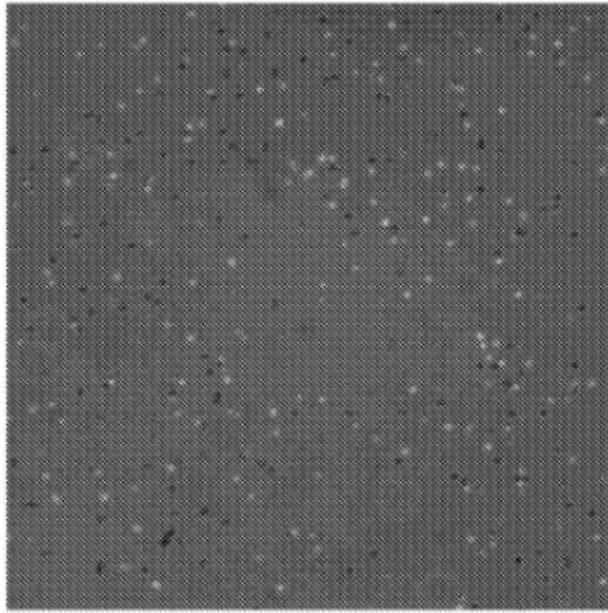


Figura 7E

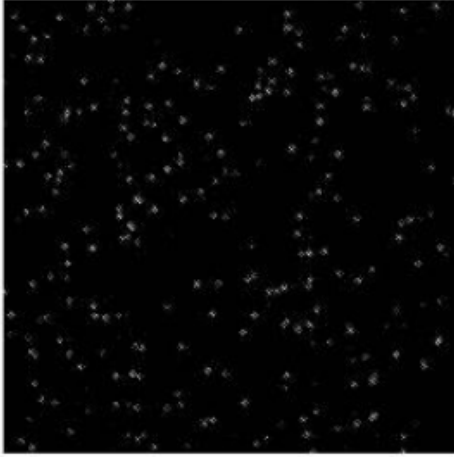


Figura 8B

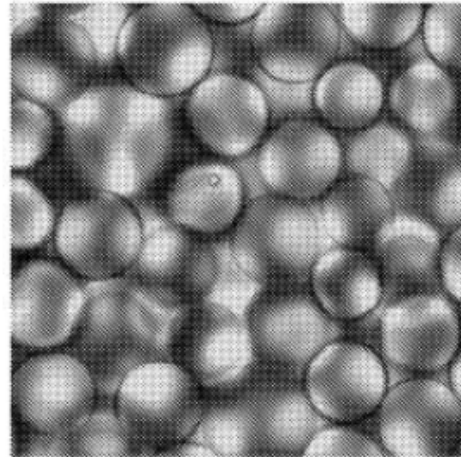


Figura 8D

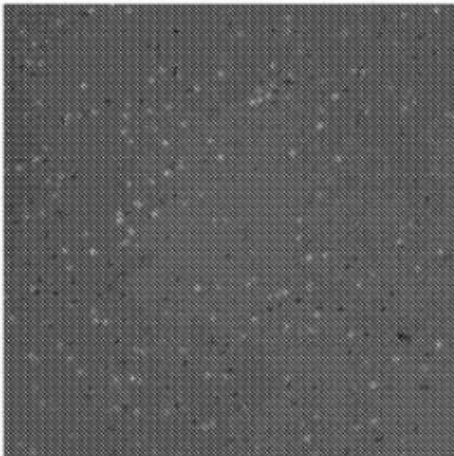


Figura 8A

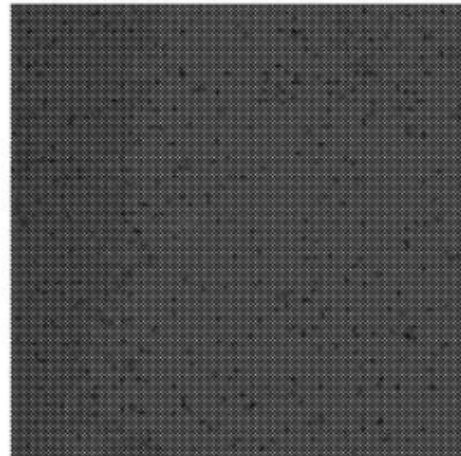


Figura 8C

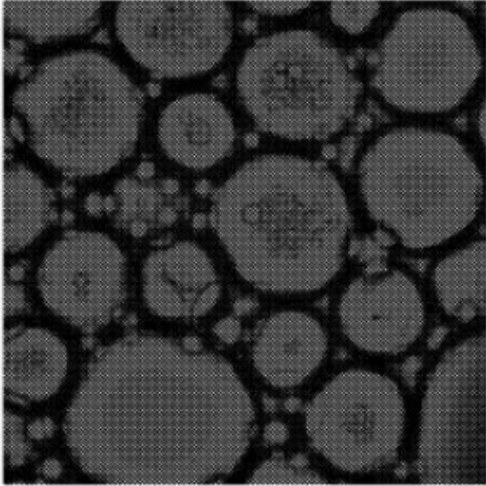


Figura 8F

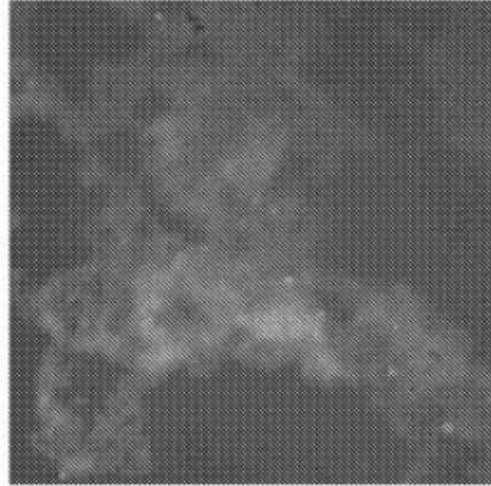


Figura 8H

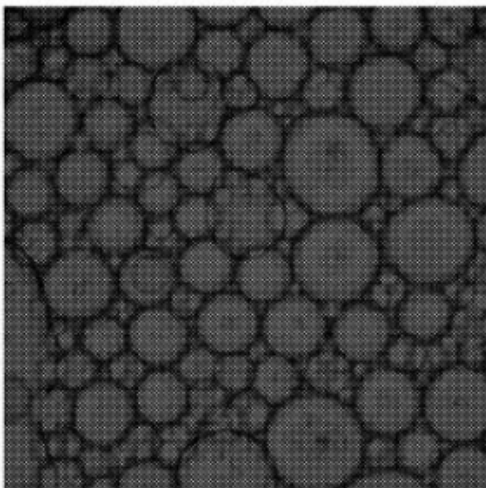


Figura 8E

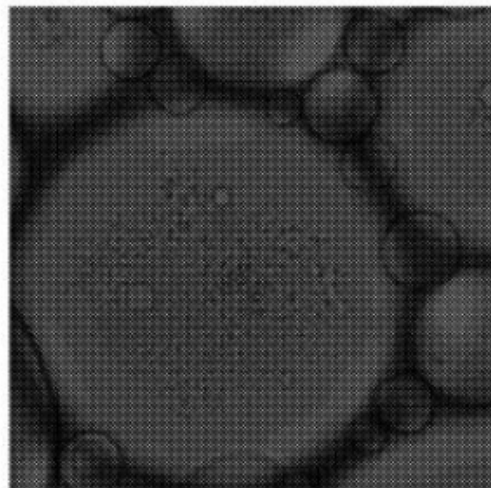


Figura 8G



Figura 9A

Librería del transcriptoma de células individuales:

5'	Cebador grupo	Cebador A sec. anclaje	
5'	(AATGATACGGCGACCA	CCGAGATCT) [A	CACTCTTTCCCTACACGACGCTCTTCCGATCT] [--BC(N20) --] CAGC[oligodT]
	Cebador B sec. inversa	Grupo B	
(ARN)	[AGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCG] (ATCTCGTATGCCGCTCTTCTGCTTG) 3'		
3'	Cebador grupo	Cebador A sec. anclaje	
3'	(TTA	CTATGCCGCTGGTGGCTCTAGA) [TGTGAGAAAGGGATGTGCTGCGAGAAGGCTAGA] [--BC(N20) --] CAGC[oligodT]	
	Cebador B sec. inversa	Grupo B	
(ARN)	[TCTAGCCTTCTCGCCAAGTCGTCCTTACGGCTCTGGC] (TAGAGCATACGGCAGAGACGAAC) 5'		

Figura

Figura 9B

Muestra del Oligonucleótido con código de barras

BEad- [ACACTCTTTCCCTACACGACGCTCTTCCGATCT][--BC(N20)--]CAGC[oligodT]

BC=código de barras

Figura 9C