

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 774 165**

51 Int. Cl.:

C12Q 1/6806 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **17.06.2013 PCT/US2013/046130**

87 Fecha y número de publicación internacional: **19.12.2013 WO13188872**

96 Fecha de presentación y número de la solicitud europea: **17.06.2013 E 13733180 (7)**

97 Fecha y número de publicación de la concesión europea: **01.01.2020 EP 2861760**

54 Título: **Secuenciación de alto rendimiento de múltiples transcritos de una única célula**

30 Prioridad:

15.06.2012 US 201261660370 P

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

17.07.2020

73 Titular/es:

**THE BOARD OF REGENTS OF THE UNIVERSITY
OF TEXAS SYSTEM (100.0%)**

**210 West 7th Street
Austin, TX 78701 , US**

72 Inventor/es:

**HUNICKE-SMITH, SCOTT;
DEKOSKY, BRANDON;
ELLINGTON, ANDY y
GEORGIU, GEORGE**

74 Agente/Representante:

MORENO NOGALES, Ángeles

ES 2 774 165 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

Secuenciación de alto rendimiento de múltiples transcritos de una única célula

5 **ANTECEDENTES DE LA INVENCION****1. Campo de la invención**

10 La presente invención se refiere en general al campo de la biología molecular y la inmunología. Más en particular, se refiere a procedimientos para el aislamiento con alto rendimiento de ADNc que codifican receptores y anticuerpos de células inmunitarias.

2. Descripción de la técnica relacionada

15 Existe la necesidad de identificar la expresión de dos o más transcritos de células individuales con alto rendimiento. En particular, para numerosas aplicaciones biotecnológicas y médicas, es importante identificar y secuenciar los pares de genes que codifican las dos cadenas que comprenden receptores inmunitarios adaptativos de células individuales con un rendimiento muy alto para determinar con exactitud los repertorios completos de receptores inmunitarios expresados en pacientes o en animales de laboratorio. Los receptores inmunitarios expresados por los linfocitos B y T se codifican respectivamente por los genes VH y VL de anticuerpos y por los genes de cadena α/β o γ/δ de TCR. Los humanos tienen muchas decenas de miles o millones de linfocitos B y T distintos clasificados en diferentes subconjuntos en base a la expresión de marcadores de superficie (proteínas CD) y factores de transcripción (por ejemplo, FoxP3 en el subconjunto de linfocitos T Treg). Se han usado tecnologías de secuenciación de ADN de alto rendimiento para determinar los repertorios de cadenas VH o VL o, de forma alternativa, de TCR α y β en subconjuntos de linfocitos de relevancia para estados de enfermedad particulares o, más en general, para estudiar la función del sistema inmunitario adaptativo (Wu *et al.*, 2011). Los investigadores en inmunología tienen una necesidad especialmente grande de analizar con alto rendimiento múltiples transcritos a la vez.

30 Los procedimientos disponibles actualmente para la secuenciación de repertorios inmunitarios implican el aislamiento de ARNm de una población celular de interés, por ejemplo, linfocitos B de memoria o células plasmáticas de la médula ósea, seguido de RT-PCR en masa para sintetizar ADNc para secuenciación de ADN de alto rendimiento (Reddy *et al.*, 2010; Krause *et al.*, 2011). Sin embargo, las cadenas de anticuerpos pesadas y ligeras (o receptores de linfocitos T α y β) se codifican en cadenas de ARNm separadas y se deben secuenciar por separado. Por tanto, estos procedimientos disponibles tienen el potencial de revelar todos los repertorios inmunitarios de cadena pesada y ligera individualmente, pero aún no pueden resolver los emparejamientos de cadena pesada y ligera con alto rendimiento. Sin un análisis de múltiple transcritos a nivel de una única célula para recopilar datos de emparejamientos de cadena pesada y ligera, el receptor inmunitario adaptativo completo, que incluye ambas cadenas, no se puede secuenciar o reconstruir y expresar para su posterior estudio.

40 **SUMARIO DE LA INVENCION**

45 La presente invención proporciona un procedimiento que comprende a) separar células individuales en un compartimento con microesferas conjugadas con oligonucleótidos poli(T); b) lisar las células; c) permitir que los transcritos de ARNm liberados de las células se hibriden con oligonucleótidos conjugados con las microesferas; d) agrupar, lavar y resuspender las microesferas en solución con reactivos para retrotranscripción asociada a reacción en cadena de la polimerasa con extensión de solapamiento, que incluye cebadores diseñados para crear un único producto de PCR que comprende ADNc de al menos dos transcritos de interés unidos covalentemente; e) realizar una retrotranscripción asociada a reacción en cadena de la polimerasa con extensión de solapamiento para formar una molécula de ADN individual que comprende al menos dos ADNc de los al menos dos transcritos derivados de una única célula, en el que los al menos dos ADNc comprenden secuencias de TCR emparejadas o secuencias de VH y VL emparejadas; y f) secuenciar el ADNc asociado con las microesferas.

55 De acuerdo con la invención, en algunos aspectos se proporciona un procedimiento para obtener una pluralidad de secuencias de VH y VL de anticuerpos emparejadas que comprende (a) aislar linfocitos B individuales en compartimentos individuales con oligonucleótidos inmovilizados para cebar la retrotranscripción; (b) lisar los linfocitos B y permitir que los transcritos de ARNm se asocien con los oligonucleótidos inmovilizados; (c) realizar la retrotranscripción seguida de la amplificación por PCR para obtener los ADNc correspondientes a los transcritos de ARNm de linfocitos B individuales; (d) secuenciar los ADNc; y (e) identificar las secuencias de VH y VL de anticuerpos emparejadas para una pluralidad de linfocitos B individuales. En otros aspectos, se proporciona un procedimiento para obtener una pluralidad de secuencias de receptores de linfocitos T emparejadas que comprende (a) aislar linfocitos T individuales en compartimentos individuales con oligonucleótidos inmovilizados para cebar la retrotranscripción; (b) lisar los linfocitos T y permitir que los transcritos de ARNm se asocien con los oligonucleótidos inmovilizados; (c) realizar la retrotranscripción seguida de la amplificación por PCR para obtener los ADNc correspondientes a los transcritos de ARNm de linfocitos T individuales; (d) secuenciar los ADNc; y (e) identificar las

secuencias de receptores de linfocitos T emparejadas para una pluralidad de linfocitos T individuales en base a la secuenciación.

5 En otros aspectos, el procedimiento comprende obtener secuencias de al menos 10.000, 100.000 o 1.000.000 células individuales (por ejemplo, entre aproximadamente 100.000 y 10 millones o 100 millones de células individuales). Por tanto, en algunos aspectos, un procedimiento comprende obtener al menos 5.000, 10.000 o 100.000 secuencias de VH y VL de anticuerpos emparejadas individuales (por ejemplo, entre aproximadamente 10.000 y 100.000, 1 millón o 10 millones de secuencias emparejadas individuales). De acuerdo con la invención, la obtención de secuencias emparejadas, tales como las secuencias de VH y VL, comprende unir ADNc (por ejemplo, ADNc VH y VL) realizando una retrotranscripción asociada a reacción en cadena de la polimerasa con extensión de solapamiento para unir los ADNc en moléculas individuales.

15 En un aspecto, el procedimiento puede comprender además determinar transcritos emparejados de forma natural usando análisis de probabilidad. En este aspecto, la identificación de los transcritos emparejados puede comprender comparar recuentos de lectura de secuenciación sin procesar. Por ejemplo, un análisis de probabilidad puede comprender realizar las etapas de la FIG. 9. En un aspecto específico, un procedimiento puede comprender identificar las secuencias de VH y VL de anticuerpos emparejadas realizando un análisis de probabilidad de las secuencias. En determinados aspectos, el análisis de probabilidad se puede basar en las secuencias CDR-H3 y/o CDR-L3. En algunos casos, identificar las secuencias de VH y VL de anticuerpos emparejadas puede comprender comparar recuentos de lectura de secuenciación sin procesar. En otro aspecto, el análisis de probabilidad puede comprender realizar las etapas de la FIG. 9.

25 Determinados aspectos de los presentes modos de realización se refieren a agentes de captura de ARNm. El agente de captura de ARNm es una microesfera que comprende oligonucleótidos inmovilizados. En un aspecto, la microesfera es una microesfera de sílice o una microesfera magnética. El agente de captura de ARNm comprende oligonucleótidos que hibridan ARNm; en particular, los oligonucleótidos comprenden al menos un poli(T). En determinados aspectos, una microesfera de los modos de realización es más pequeña que las células individuales que se están aislando (por ejemplo, linfocitos B).

30 En algunos aspectos, los compartimentos individuales de los modos de realización pueden ser pocillos en una placa de gel o microtitulación. En un aspecto, los compartimentos individuales pueden tener un volumen de menos de 5 nl. En algunos aspectos, los pocillos se pueden sellar con una membrana permeable antes de la lisis de las células o antes de realizar la RT-PCR. En aún otro aspecto, los compartimentos individuales pueden ser microvesículas en una emulsión.

35 En otros aspectos de la presente divulgación, la fijación de células individuales (y un agente de captura de ARNm) y la lisis de las células (etapas (a) y (b)) se pueden realizar simultáneamente. Por tanto, en algunos aspectos, un procedimiento puede comprender aislar células individuales y un agente de captura de ARNm en microvesículas individuales en una emulsión y en presencia de una solución de lisis celular.

40 El procedimiento de los modos de realización comprende unir ADNc realizando una retrotranscripción asociada a reacción en cadena de la polimerasa con extensión de solapamiento para unir al menos 2 transcritos en una molécula de ADN individual (por ejemplo, en la etapa (e)).

45 En aún otros aspectos, la fijación de las células individuales puede comprender introducir células en un dispositivo que comprende una pluralidad de micropocillos de modo que la mayoría de las células se capturen como células individuales (junto con un agente de captura de ARNm, tal como una microesfera). Los procedimientos comprenden la secuenciación de dos o más transcritos unidos covalentemente a la misma microesfera.

50 Por tanto, en algunos modos de realización se proporciona un procedimiento para obtener una pluralidad de secuencias de VH y VL de anticuerpos emparejadas en las que las células son linfocitos B. En un aspecto, el procedimiento es un procedimiento para obtener secuencias de VH y VL de anticuerpos emparejadas para un anticuerpo que se une a un antígeno de interés. En determinados aspectos, las microesferas se pueden conjugar con el antígeno de interés y los oligonucleótidos solo se pueden conjugar con las microesferas en presencia de un anticuerpo que se une al antígeno de interés. Por ejemplo, las microesferas se pueden recubrir con un antígeno de interés y el agente de captura de ARNm, a saber, el oligo-T se puede asociar con la microesfera solo en presencia de un anticuerpo que se une al antígeno (véase, por ejemplo, la FIG. 10). Por ejemplo, el agente de captura de ARNm se puede asociar con proteína A o funcionalizar de otro modo para unirse a un anticuerpo si está presente.

60 Se puede obtener una muestra de un sujeto para su uso en los procedimientos reivindicados (por ejemplo, una muestra que comprende células para su uso en los procedimientos de los modos de realización). Las muestras se pueden tomar directamente de un sujeto o se pueden obtener de un tercero. Las muestras incluyen, pero no se limitan a, muestras de suero, mucosa (por ejemplo, saliva), linfa, orina, heces y tejido sólido. De forma similar, determinados aspectos de los modos de realización se refieren a líquidos biológicos y anticuerpos y/o ácidos nucleicos a partir de los mismos. Por ejemplo, el líquido biológico puede ser sangre (por ejemplo, suero), líquido cefalorraquídeo, líquido sinovial, leche materna, sangre del cordón umbilical, líquido peritoneal, secreciones de la

mucosa, lágrimas, secreciones nasales, saliva, leche o secreciones genitourinarias. En determinados aspectos, las células para su uso de acuerdo con los modos de realización son células de mamífero, tales como células de ratón, rata o mono. En aspectos preferentes, las células son células humanas.

5 En algunos aspectos, las células para su uso en los modos de realización son linfocitos B, tales como linfocitos B de un órgano seleccionado, tal como médula ósea. Por ejemplo, los linfocitos B pueden ser linfocitos B maduros, tales como células plasmáticas de médula ósea, células plasmáticas de bazo o células plasmáticas de ganglios linfáticos, o células de sangre periférica o de un órgano linfático. En determinados aspectos, los linfocitos B se seleccionan o enriquecen en base a la expresión diferencial de los marcadores de la superficie celular (por ejemplo, Blimp-1, 10 CD138, CXCR4 o CD45). En algunos casos se obtienen secuencias de una clase seleccionada de anticuerpos, tales como secuencias de IgE, IgM, IgG o IgA.

No de acuerdo con la invención, un procedimiento puede comprender inmunizar al sujeto (por ejemplo, antes de 15 obtener una muestra celular). El procedimiento puede comprender además el aislamiento de un tejido linfático. El aislamiento del tejido linfático se puede realizar al menos o aproximadamente 1, 2, 3, 4, 5, 6, 6, 8, 9, 10 días o cualquier intervalo intermedio después de la inmunización. El procedimiento puede comprender además obtener una población de ácidos nucleicos de tejido linfático, preferentemente sin separar los linfocitos B del tejido linfático. El tejido linfático puede ser un tejido linfático primario, secundario o terciario, tal como médula ósea, bazo o ganglios linfáticos. El sujeto puede ser cualquier animal, tal como mamíferos, peces, anfibios o aves. El mamífero puede ser 20 un ser humano, ratón, primate, conejo, oveja o cerdo.

Para determinar las secuencias de ácido nucleico (por ejemplo, en los linfocitos B o en los tejidos linfáticos), se puede usar cualquier procedimiento de secuenciación de ácidos nucleicos conocido en la técnica, incluyendo la 25 secuenciación de ADN de alto rendimiento. Ejemplos no limitantes de procedimientos de secuenciación de alto rendimiento comprenden secuenciación por síntesis (por ejemplo, secuenciación 454), secuenciación por fijación, secuenciación por hibridación, secuenciación de ADN de molécula individual, secuenciación de múltiples colonias, secuenciación de nanoporos o una combinación de las mismas.

También se describe en el presente documento un sistema que comprende (a) una salida de fase fluida acuosa 30 dispuesta dentro de una fase oleosa que fluye de forma anular; y (b) una fase fluida acuosa, en la que el fluido de fase acuosa comprende una suspensión de células y se dispersa dentro de la fase oleosa que fluye, dando como resultado gotitas emulsionadas con baja dispersión de tamaño que comprenden una suspensión acuosa de células. En un aspecto, la salida de la fase fluida acuosa es una aguja. En otro aspecto, la salida de la fase fluida acuosa es un tubo de vidrio. En determinados aspectos, la fase oleosa fluye a través de un tubo de vidrio o tubo polimérico. En 35 determinados aspectos, la fase acuosa fluye a través de un tubo polimérico. En todavía otro aspecto, la concentración de células, el caudal de la fase fluida acuosa y el caudal de la fase oleosa permiten la formación de gotitas, en las que cada gotita contiene una única célula. En algunos aspectos, las celdas se seleccionan del grupo que consiste en: linfocitos B, linfocitos T, células NKT y células cancerosas. En determinados aspectos, la fase fluida acuosa comprende microesferas para reactivos de retrotranscripción para captura de ácidos nucleicos, reactivos de 40 reacción en cadena de polimerasa y/o combinaciones de los mismos.

Se describe además una composición que comprende (a) una microesfera; (b) un oligonucleótido que se puede unir a ARNm; y (c) dos o más cebadores específicos para un transcrito de interés.

45 Se describe además una composición que comprende una emulsión que tiene una pluralidad de microvesículas individuales, comprendiendo dichas microvesículas una microesfera con oligonucleótidos inmovilizados para cebar la retrotranscripción y linfocitos B individuales, que se han alterado para liberar transcritos de ARNm.

También se describe en el presente documento un procedimiento que comprende (a) añadir una secuencia común a 50 la región 5' de dos o más oligonucleótidos que son específicos de un conjunto de dianas génicas; (b) realizar la amplificación de ácidos nucleicos del conjunto de dianas génicas cebando la secuencia común; y (c) incluir en la amplificación de ácidos nucleicos oligonucleótidos que comprenden la secuencia común inmovilizada sobre una superficie de modo que los oligonucleótidos inmovilizados ceban la amplificación de ácidos nucleicos y dando como resultado la captura en superficie de secuencias amplificadas.

55 Como se usa en el presente documento en la memoria descriptiva, "uno" o "una" pueden significar uno o más. Como se usa en el presente documento en la(s) reivindicación(es), cuando se usa junto con la palabra "que comprende", las palabras "uno" o "una" pueden significar uno o más de uno.

60 El uso del término "o" en las reivindicaciones se usa para significar "y/o" a menos que se indique explícitamente que se refiere solo a alternativas o que las alternativas sean mutuamente excluyentes, aunque la divulgación respalda una definición que se refiere solo a alternativas e "y/o." Como se usa en el presente documento, "otro" puede significar al menos un segundo o más.

A lo largo de la presente solicitud, el término "aproximadamente" se usa para indicar que un valor incluye la variación inherente de error para el dispositivo, siendo empleado el procedimiento para determinar el valor, o la variación que existe entre los sujetos del estudio.

- 5 Otros objetos, rasgos característicos y ventajas de la presente invención serán evidentes a partir de la siguiente descripción detallada, mientras que la invención se define en las reivindicaciones adjuntas.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

10 Los siguientes dibujos forman parte de la presente memoria descriptiva y se incluyen para demostrar adicionalmente determinados aspectos de la presente invención. La invención se puede entender mejor por referencia a uno o más de estos dibujos en combinación con la descripción detallada de modos de realización específicos presentados en el presente documento.

15 **FIG. 1** muestra células aisladas en pocillos sellados individuales. Los objetos pequeños y esféricos dentro de los pocillos son microesferas. Esta imagen se toma a través de la membrana de diálisis. El diámetro del pocillo es de aproximadamente 56 μm .

20 **FIG. 2** muestra *Izquierda*: una única célula aislada inmediatamente antes de la lisis; *Centro*: la célula durante el proceso de lisis; y *Derecha*: el micropocillo inmediatamente después de la lisis, usando microscopia de lapso de tiempo. El diámetro del pocillo es de aproximadamente 56 μm .

25 **FIG. 3** muestra el producto de RT-PCR OE unido. Las letras indican localizaciones aproximadas de regiones constantes, variables, de unión y de diversidad, mientras que los números indican localizaciones aproximadas de regiones determinantes de la complementariedad.

30 **FIG. 4** muestra una visión general del proceso de RT-PCR (con extensión de solapamiento) de unión. a) Los cebadores de la región V con una región de solapamiento pesada/ligera complementaria 5' se hibridan al ADNc de primera hebra. b) El ADNc de segunda hebra se forma por una extensión 5' a 3'; la región de solapamiento se incorpora a todo el ADNc. c) Después de la desnaturalización, las cadenas pesada y ligera con el sentido de primera hebra se hibridan para generar un producto completo de 850 pb a través de una extensión 5' a 3'. La CDR-H3 y la CDR-L3 se localizan cerca del exterior de la construcción unida final, lo que permite el análisis de CDR3 mediante secuenciación de ambos extremos 2x250 con Illumina.

35 **FIG. 5** muestra células MOPC-21 encapsuladas de manera viable en gotitas formadas por medio de *flow focusing* (enfoque de flujo). Las dos corrientes de entrada al dispositivo de *flow focusing* estaban compuestas de partes iguales de células MOPC-21 en PBS (100.000 células/ml, corriente celular) y azul tripano al 0,4 % en PBS (corriente de tinte), y la corriente celular y la corriente de tinte se mezclaron entre sí inmediatamente antes del punto de formación de gotitas de emulsión. Se demostró que las células MOPC-21 excluían el azul tripano, lo que demuestra la encapsulación viable de células individuales dentro de las gotitas de emulsión.

40 **FIG. 6** presenta una visión general de la tecnología de secuenciación de alto rendimiento para múltiples transcritos aplicada a la secuenciación de ARNm de VH y VL de anticuerpos naturales a partir de poblaciones de linfocitos B. i) Las poblaciones de linfocitos B se separan según el fenotipo deseado (por ejemplo, linfocitos B de memoria (LBm), linfocitos B indiferenciados (LB indiferenciados)). ii) Las células individuales se aíslan por sedimentación aleatoria en una matriz de micropocillos; a los pocillos se añaden también microesferas de poli(dT). iii) Los pocillos se sellan con una membrana de diálisis y se equilibran con tampón de lisis para lisar las células e hibridar los ARNm de VH y VL a microesferas de poli(dT) (la mancha representa una célula lisada, los círculos representan microesferas magnéticas, las líneas negras representan hebras de ARNm). iv) Las microesferas se recuperan y se emulsionan para síntesis de ADNc y PCR de unión para generar un producto de ADNc de VH:VL de ~850 pares de bases. v) Se realiza una secuenciación de nueva generación para secuenciar las hebras unidas. vi) El procesamiento bioinformático se usa para analizar el repertorio de emparejamientos VH:VL.

55 **FIG. 7** muestra la amplificación de ADN de cadena pesada y ligera en microesferas magnéticas con oligonucleótidos inmovilizados para la secuenciación de alto rendimiento. a) Las microesferas muestran una mezcla de 3 oligonucleótidos inmovilizados: poli(T) para la captura de ARNm, AHX89 para la amplificación de la cadena pesada y BRH06 para la amplificación de la cadena ligera. b) La retrotranscripción se inicia a partir del ARNm capturado (representado por líneas discontinuas grises) que se ha hibridado a oligonucleótidos poli(T) inmovilizados. Los cebadores de retrotranscripción de región constante de inmunoglobulina especialmente diseñados tienen AHX89 en el extremo 5' (para la cadena pesada) o BRH06 (para la cadena ligera). La retrotranscripción asociada a reacción en cadena de la polimerasa se produce dentro de gotitas de emulsión. c) Los cebadores directos de la región V tienen una secuencia <F3> en el extremo 5' (cadena pesada) o una secuencia <F5> (cadena ligera) que se usará para iniciar la pirosecuenciación. Las hebras de ADNc se muestran como líneas negras.

65 **FIG. 8** muestra un diagrama del aparato de corriente portadora y boquilla. Un tubo capilar de vidrio suministra una corriente portadora de fase oleosa externa (flechas) que rodea una salida de aguja. La aguja inyecta la fase acuosa

que contiene las células, y las fuerzas de cizallamiento del flujo anular de la fase oleosa generan gotitas monodispersas.

FIG. 9 muestra un algoritmo de árbol de decisiones general para el emparejamiento de secuencias de VH y VL.

FIG. 10 muestra esquemáticamente un proceso ejemplar de captura de ARNm de células individuales aisladas que codifican anticuerpos de alta afinidad para un antígeno particular. (a) Los linfocitos B secretores de anticuerpos (arriba a la izquierda) se aíslan en compartimentos que contienen microesferas con antígeno inmovilizado. Las microesferas capturan el anticuerpo secretado (gris) si el linfocito B codifica un anticuerpo de alta afinidad para el antígeno. (b) Cualquier anticuerpo secretado por las células no unido se elimina por lavado y se añade un anticuerpo anti-IgG (blanco) con ADNmc poli(dT) unido (hebras negras) al compartimento. La construcción anti-IgG:poli(dT) (u otro resto de captura de ARNm) se inmoviliza en microesferas que contienen el anticuerpo capturado. El ADNmc poli(dT) se coloca solo con células que secretan el anticuerpo de alta afinidad para el antígeno deseado. (c) Los compartimentos se sellan y las células se lisan. Las hebras de ARNm (círculos pequeños) liberadas de las células que secretaron el anticuerpo de alta afinidad se capturan por medio de hibridación al poli(dT) de las construcciones poli(dT):anticuerpo:microesfera. A continuación, las microesferas se pueden recuperar para el análisis de transcritos de ARNm de una única célula.

DESCRIPCIÓN DE MODOS DE REALIZACIÓN ILUSTRATIVOS

La presente divulgación se refiere en general a la secuenciación de dos o más genes expresados en una única célula con alto rendimiento. Más en particular, la presente divulgación proporciona un procedimiento para la secuenciación de alto rendimiento de pares de transcritos coexpresados en células individuales para determinar los pares de cadenas polipeptídicas que comprenden receptores inmunitarios (por ejemplo, secuencias de VH y VL de anticuerpos).

Los procedimientos de la presente divulgación permiten determinar el repertorio de receptores inmunitarios y anticuerpos en un organismo individual o población de células. En particular, los procedimientos de la presente divulgación pueden ayudar a determinar pares de cadenas polipeptídicas que constituyen receptores inmunitarios. Los linfocitos B y los linfocitos T expresan cada uno receptores inmunitarios; los linfocitos B expresan inmunoglobulinas y los linfocitos T expresan receptores de linfocitos T (TCR). Ambos tipos de receptores inmunitarios consisten en dos cadenas polipeptídicas. Las inmunoglobulinas consisten en cadenas pesada variable (VH) y ligera variable (VL). Los TCR son de dos tipos: uno que consiste en una cadena α y una cadena β , y uno que consiste en una cadena γ y una cadena δ . Cada uno de los polipéptidos de un receptor inmunitario tiene una región constante y una región variable. Las regiones variables son el resultado de la recombinación y el reordenamiento de fragmentos de genes en el cromosoma de un linfocito B o T. En los linfocitos B se produce una diversificación adicional de regiones variables por hipermutación somática. Por tanto, el sistema inmunitario tiene un gran repertorio de receptores, y cualquier par de receptores dado expresado por un linfocito se codifica por un par de transcritos separados y únicos. Solo conociendo la secuencia de ambos transcritos del par se puede estudiar el receptor como un todo. Conocer las secuencias de los pares de cadenas de receptores inmunitarios expresadas en una única célula también es esencial para determinar el repertorio inmunitario de un individuo o población de células dados.

Los procedimientos actualmente disponibles para analizar múltiples transcritos en células individuales, tales como los dos transcritos que comprenden receptores inmunitarios adaptativos, están limitados por el bajo rendimiento y los muy altos costes de instrumentación y reactivos. Actualmente no existe ninguna tecnología para analizar rápidamente cuántas células expresan un conjunto de transcritos de interés o, más específicamente, para secuenciar pares de cadenas de receptores de linfocitos naturales con un rendimiento muy alto (más de 10.000 células por ejecución). La presente divulgación tiene como objetivo corregir estas deficiencias proporcionando una nueva técnica para secuenciar múltiples transcritos simultáneamente a nivel de una única célula con un rendimiento dos a tres órdenes de magnitud mayor que el estado de la técnica actual.

Una ventaja de los procedimientos de la presente divulgación es que los procedimientos dan como resultado un rendimiento que es varios órdenes de magnitud mayor que el estado de la técnica. Además, la presente divulgación permite la capacidad de unir dos transcritos para poblaciones de células grandes con alto rendimiento, más rápidamente y a un coste mucho menor que las tecnologías de la competencia.

En determinados modos de realización, la presente divulgación proporciona procedimientos que comprenden separar células individuales en un compartimento con microesferas conjugadas con oligonucleótidos; lisar las células; permitir que los transcritos de ARNm liberados de las células se hibriden con los oligonucleótidos; realizar una retrotranscripción asociada a reacción en cadena de la polimerasa con extensión de solapamiento para unir covalentemente el ADN a partir de al menos dos transcritos derivados de una única célula; y secuenciar el ADN unido. En determinados modos de realización, las células pueden ser células de mamífero. En determinados modos de realización, las células pueden ser linfocitos B, linfocitos T o células NKT.

5 En otros modos de realización, la presente divulgación proporciona procedimientos que comprenden separar células individuales en un compartimento con microesferas conjugadas con oligonucleótidos; lisar la célula; permitir que los transcritos de ARNm liberados de las células se hibriden con los oligonucleótidos conjugados con las microesferas; realizar una retrotranscripción asociada a reacción en cadena de la polimerasa para formar al menos dos ADNc a partir de al menos dos transcritos derivados de una única célula; y secuenciar el ADNc unido a las microesferas.

10 En otro modo de realización, la presente divulgación proporciona un procedimiento que comprende mezclar células con microesferas que tienen un diámetro menor que el diámetro de las células, en el que las microesferas se conjugan con oligonucleótidos, fijar las células y las microesferas dentro de compartimentos que tienen un volumen de menos de 5 nl, lisar las células y permitir que los transcritos de ARNm se asocien con las microesferas, aislar las microesferas y el ARNm asociado de los compartimentos, realizar una retrotranscripción seguida de amplificación por PCR en el ARNm asociado a microesferas, y secuenciar el producto de ADN de cada microesfera para identificar el ADNc asociado con cada microesfera

15 La presente divulgación también describe un sistema que comprende una salida de fase fluida acuosa dispuesta dentro de una fase oleosa que fluye de forma anular, en la que el fluido de fase acuosa comprende una suspensión de células y se dispersa dentro de la fase oleosa que fluye, dando como resultado gotitas emulsionadas con baja dispersión de tamaño que comprende una suspensión acuosa de células.

20 La presente divulgación también describe una composición que comprende una microesfera, un oligonucleótido que se puede unir a ARNm, y dos o más cebadores específicos para un transcrito de interés.

25 La presente divulgación también describe un dispositivo que comprende conjuntos ordenados de micropocillos, cada uno con dimensiones diseñadas para acomodar un único linfocito. Los micropocillos pueden ser pocillos circulares de 56 µm de diámetro y 50 µm de profundidad, para un volumen total de 125 pl. El volumen de dichos micropocillos normalmente oscilaría entre 20 y 3.000 pl, aunque se puede usar una amplia variedad de tamaños, formas y dimensiones de pocillos para acomodar una única célula. El micropocillo puede ser un nanopocillo. El dispositivo puede ser un chip. El dispositivo de la presente divulgación permite el atrapamiento directo de decenas de miles de células individuales, con cada célula en su propio micropocillo, en un solo chip. El chip puede ser del tamaño de un portaobjetos de microscopio. Se puede usar un chip de micropocillos para capturar células individuales en sus propios micropocillos individuales (FIG. 6). El chip de micropocillos se puede preparar en polidimetilsiloxano (PDMS); sin embargo, también se pueden usar otros materiales adecuados conocidos en la técnica, tales como poliacrilimida, silicio y vidrio esmerilado para crear el chip de micropocillos.

35 También se pueden capturar varias microesferas u otras partículas conjugadas con oligonucleótidos en los micropocillos con las células individuales de acuerdo con los procedimientos de la presente divulgación. En determinados modos de realización, las microesferas pueden comprender oligonucleótidos inmovilizados en la superficie de las microesferas. En otros modos de realización, las microesferas pueden ser magnéticas. En otros modos de realización, las microesferas se pueden recubrir con uno o más oligonucleótidos. Los oligonucleótidos comprenden poli(T). Una membrana de diálisis cubre los micropocillos, manteniendo las células y las microesferas en los micropocillos mientras se dializan los reactivos de lisis en los micropocillos. Los reactivos de lisis provocan la liberación de los transcritos de ARNm de las células en el micropocillo con las microesferas. Los oligonucleótidos poli(T) en las microesferas capturan las colas poli(A) de ARNm. Por tanto, cada microesfera se recubre con moléculas de ARNm de una única célula. A continuación, las microesferas se agrupan, lavan y resuspenden en solución con reactivos para la retrotranscripción asociada a reacción en cadena de la polimerasa (RT-PCR) con extensión de solapamiento (OE). Esta mezcla de reacción incluye cebadores diseñados para crear un único producto de PCR que comprende ADNc de dos transcritos de interés unidos covalentemente entre sí. Antes del termociclado, la suspensión de microesferas y solución de reactivos se emulsiona en la fase oleosa para crear gotitas con no más de una microesfera por gotita. Los productos de ADNc unidos de RT-PCR OE se recuperan y usan como molde para la PCR con cebadores internos, que amplifica los transcritos de interés unidos. Los productos purificados de la PCR con cebadores internos se secuencian a continuación y se analiza la información de emparejamiento (FIG. 6).

55 La presente divulgación también proporciona un procedimiento para atrapar ARNm de células individuales en microesferas, realizar síntesis de ADNc, unir las secuencias de dos o más ADNc deseados de células individuales para crear una molécula individual, y finalmente revelar la secuencia de los transcritos unidos por secuenciación (NextGen) de alto rendimiento. De acuerdo con la presente divulgación, una manera de incrementar el rendimiento en ensayos biológicos es usar una emulsión que genere una gran cantidad de microrreactores paralelos tridimensionales. Los protocolos de emulsión en biología molecular a menudo producen 109-1011 gotitas por ml (volumen inferior a pl). Los procedimientos basados en emulsión para la reacción en cadena de la polimerasa (PCR) de células individuales han encontrado una amplia aceptación, y la PCR en emulsión es un procedimiento sólido y fiable que se encuentra en muchos protocolos de secuenciación de nueva generación. Sin embargo, la RT-PCR de muy alto rendimiento en gotitas de emulsión aún no se ha implementado debido a que los lisados celulares dentro de la gotita inhiben la reacción de la retrotranscriptasa. La inhibición del lisado celular de la RT-PCR se puede mitigar mediante dilución a un volumen adecuado.

En otro modo de realización, las células se lisan en gotitas de emulsión que contienen microesferas para la captura de ácidos nucleicos. Las microesferas se conjugan con oligonucleótido, y el oligonucleótido comprende poli(T). En determinados modos de realización, la microesfera puede ser magnética. Una solución acuosa con una suspensión de células y microesferas se emulsiona en fase oleosa inyectando una suspensión acuosa de células y microesferas en una corriente de fase oleosa en rápido movimiento. Las fuerzas de cizallamiento generadas por la fase oleosa en movimiento crean gotitas a medida que la suspensión acuosa se inyecta en la corriente, creando una emulsión con una baja dispersión de tamaño de las gotitas. Cada célula está en su propia gotita junto con varias microesferas conjugadas con oligonucleótidos. La uniformidad del tamaño de las gotitas ayuda a garantizar que las gotitas individuales no contengan más de una célula. A continuación, las células se lisan térmicamente y la mezcla se enfría para permitir que las microesferas capturen ARNm. La emulsión se rompe y se recogen las microesferas. Las microesferas se resuspenden en una solución para RT-PCR OE en emulsión para unir entre sí los ADNc de los transcritos de interés. La PCR con cebadores internos y la secuenciación de los transcritos unidos se realizan de acuerdo con la presente divulgación. En determinados modos de realización, la suspensión acuosa de células comprende reactivos de retrotranscripción. En determinados otros modos de realización, la suspensión acuosa de células comprende al menos uno de los reactivos de reacción en cadena de la polimerasa y los reactivos de retrotranscripción asociada a reacción en cadena de la polimerasa.

En otro modo de realización, las gotitas de emulsión que contienen células individuales y reactivos de RT-PCR se forman mediante inyección en una fase oleosa en rápido movimiento. El termociclado se realiza a continuación en estas gotitas directamente. En determinados modos de realización se puede usar una retrotranscripción asociada a reacción en cadena de la polimerasa con extensión de solapamiento para unir ADNc de múltiples transcritos de interés.

En otro modo de realización, los ADNc de interés de una única célula se unen por medio de RT-PCR a las microesferas como se describe a continuación, y los transcritos en las microesferas se secuencian directamente usando secuenciación de alto rendimiento. Una mezcla equitativa de tres especies de cebadores oligonucleotídicos funcionalizados se puede conjugar con microesferas funcionalizadas. Uno de los oligonucleótidos puede ser poli(T) para capturar la cola poli(A) de los ARNm. Los otros dos oligonucleótidos pueden ser cebadores específicos para amplificar los transcritos de interés. Las microesferas preparadas de esta manera se mezclan con las células en una solución acuosa, y la suspensión de células y microesferas se emulsiona de modo que cada célula esté en su propia gotita junto con un exceso de microesferas. En determinados modos de realización, cada gotita puede contener un promedio de 55 microesferas. Las células se lisan térmicamente, y los oligonucleótidos poli(T) de las microesferas se unen a los ARNm. La emulsión se rompe y las microesferas se recogen, lavan y resuspenden en una solución con reactivos y cebadores para RT-PCR que dará como resultado la amplificación de los transcritos de interés de tal manera que los transcritos se unan a las microesferas. La suspensión de microesferas se emulsiona y se realiza la RT-PCR. Las microesferas se recogen y someten a secuenciación de alto rendimiento, que secuenciará directamente los dos transcritos unidos a las microesferas iniciando lecturas de secuencias múltiples usando al menos dos cebadores diferentes, donde cada cebador de iniciación es específico para un transcrito de interés. Los dos transcritos se emparejan mediante localización de las microesferas en la rejilla de secuenciación de alto rendimiento, revelando secuencias que se expresan juntas a partir de una única célula. La secuenciación se puede realizar, por ejemplo, en la plataforma SOLiD de Applied Biosystem, el Proton Torrent de Life Technologies o la plataforma de secuenciación HiSeq de Illumina.

El diseño del cebador para RT-PCR OE determina qué transcritos de interés expresados por una célula dada se unen entre sí. Por ejemplo, en determinados modos de realización, los cebadores se pueden diseñar para que los respectivos ADNc de los transcritos de la cadena VH y VL se unan covalentemente entre sí. La secuenciación de los ADNc unidos revela los pares de secuencias de VH y VL expresados por células individuales. En otros modos de realización, los conjuntos de cebadores también se pueden diseñar de modo que se puedan determinar las secuencias de pares de TCR expresados en células individuales o de modo que se pueda determinar si una población de células coexpresa cualquiera de los dos genes de interés.

El sesgo puede ser un problema importante en las reacciones de PCR que usan múltiples cebadores de amplificación porque las pequeñas diferencias en la eficacia del cebador generan grandes disparidades de producto debido a la naturaleza exponencial de la PCR. Una manera de aliviar el sesgo del cebador es amplificando múltiples genes con el mismo cebador, lo que normalmente no es posible con un conjunto de múltiples cebadores. Al incluir una región de amplificación común en el extremo 5' de múltiples cebadores únicos de interés, la región de amplificación común se añade al extremo 5' de todos los productos de PCR durante el primer acontecimiento de duplicación. Después del acontecimiento de duplicación inicial, la amplificación se logra cebando solo en la región común para reducir el sesgo del cebador y permitir que la distribución final del producto de PCR siga siendo representativa de la distribución de molde original.

Dicha región común se puede explotar de diversas maneras. Una aplicación clara es añadir el cebador de amplificación común a una concentración mayor y los cebadores únicos (con una región común 5') a una concentración baja, de modo que la mayoría de la amplificación de ácidos nucleicos se produzca por medio de la secuencia común para reducir el sesgo de amplificación. Otra aplicación es la captura en superficie de productos de amplificación, por ejemplo, para capturar el producto de PCR en una microesfera durante la PCR en emulsión. Si los

oligonucleótidos d secuencia común se inmovilizan sobre la superficie de una microesfera, los productos de PCR de interés se unirán covalentemente a la microesfera durante la amplificación. De esta manera, se puede capturar un conjunto muy diverso de transcritos en una superficie usando una única secuencia de oligonucleótido inmovilizado.

5 Por ejemplo, dos regiones comunes diferentes se pueden inmovilizar sobre la superficie de una microesfera a igual concentración (por ejemplo, una secuencia común para la cadena pesada y una secuencia común diferente para la cadena ligera). Después de la amplificación por PCR, la microesfera se recubrirá con aproximadamente un 50 % de producto de amplificación de cadena pesada y un 50 % de producto de amplificación de cadena ligera. Este equilibrio entre la representación de la cadena pesada y ligera en la superficie de la microesfera ayuda a garantizar una señal suficiente de las cadenas pesada y ligera cuando la microesfera se somete a secuenciación de alto rendimiento.

15 En consecuencia, la presente divulgación describe procedimientos que comprenden añadir una secuencia común a la región 5' de dos o más oligonucleótidos que son específicos para un conjunto de dianas génicas; y realizar la amplificación de ácido nucleico del conjunto de dianas génicas cebando la secuencia común. En determinados ejemplos, la secuencia común n se inmoviliza sobre una superficie. En otros ejemplos, la secuencia común se puede usar para capturar productos de amplificación.

20 Los procedimientos de la presente divulgación permiten obtener información con respecto a múltiples transcritos expresados a partir de una única célula. En determinados modos de realización se pueden usar análisis probabilísticos para identificar pares naturales con recuentos de lectura o frecuencias superiores a los recuentos de lectura o frecuencias de pares no naturales. La información se puede usar, por ejemplo, para estudiar patrones de coexpresión de genes en diferentes poblaciones de células cancerosas. En determinados modos de realización, los tratamientos se pueden adaptar en base a la información de expresión obtenida usando los procedimientos de la presente divulgación. Otros modos de realización se pueden centrar en el descubrimiento de nuevos receptores de linfocitos.

EJEMPLOS

30 Los siguientes ejemplos se incluyen para demostrar modos de realización preferentes de la invención. Los expertos en la técnica deben apreciar que las técnicas divulgadas en los ejemplos que siguen representan técnicas descubiertas por el autor de la invención para funcionar bien en la práctica de la invención y, por tanto, se pueden considerar modos preferentes para su práctica.

Ejemplo 1 - Construcción de una placa de micropocillos de alta densidad

35 Una rejilla de micropilares (56 µm de diámetro, 50 µm de altura) se modelan fotolitográficamente en una oblea de sílice usando un SU-8 fotorresistente (Fisher Scientific) y la oblea de sílice se usa como molde para imprimir chips de polidimetilsiloxano (PDMS) (Sylgard 184, Dow Corning) con las dimensiones de un portaobjetos de microscopio estándar y que contiene aproximadamente 170.000 pocillos por chip. Las dimensiones del micropilar pueden variar de aproximadamente 5 µm a aproximadamente 300 µm de ancho y de aproximadamente 5 µm a aproximadamente 300 µm de altura. Los chips de PDMS moldeados se silanizan en una cámara de plasma de oxígeno durante 5 minutos para generar una superficie hidrófila. Los chips de PDMS se bloquean a continuación en seroalbúmina bovina (BSA) al 1 % durante 30 minutos y se lavan con agua desionizada y solución salina tamponada con fosfato (PBS) para preparar la siembra celular.

Ejemplo 2 - Procedimiento para unir dos transcritos de una única célula con alto rendimiento

50 El procedimiento para unir físicamente dos o más transcritos derivados de una única célula con alto rendimiento usa el dispositivo de micropocillos de PDMS sellado del Ejemplo 1 para atrapar las células individuales en pocillos separados. También se produce la lisis celular, y las microesferas magnéticas de tamaño en micrómetros de poli(T) para la captura de ARNm también se introducen en los micropocillos. Tras cargar las células y las microesferas, el dispositivo se sella con una membrana de diálisis y se introduce una solución de lisis. Posteriormente, las microesferas se recuperan, resuspenden en solución con reactivos, cebadores y enzima polimerasa para RT-PCR con extensión de solapamiento (OE), y la solución se emulsiona a continuación de modo que cada microesfera se encapsule dentro de una gotita de emulsión individual. La emulsión se somete a termociclado para unir físicamente los dos transcritos (por ejemplo, ADNc de cadena pesada y ligera de inmunoglobulina), y los productos unidos se recuperan de la emulsión después del ciclado. Se realiza una amplificación por PCR con cebadores internos y, a continuación, el ADN resultante se secuencia usando Illumina o cualquier otra tecnología de secuenciación NextGen que pueda proporcionar lecturas de longitud apropiada para interpretar inequívocamente la información de emparejamiento de los transcritos (FIG. 6).

65 El procedimiento explicado anteriormente se empleó para unir la cadena pesada variable (VH) y la cadena ligera variable (VL) de inmunoglobulina en mezclas que comprenden las líneas celulares de hibridoma de ratón MOPC-21 y MOPC-315. Se conocen las secuencias de VH y VL expresadas por cada una de estas líneas celulares y, por tanto, estos experimentos sirvieron para la validación del procedimiento. Se retiraron por separado 5 ml de cada una de las

células MOPC-21 y MOPC-315 del cultivo dos días después del paso (las células se cultivaron en matraces de cultivo T-25 ventilados Falcon de 10 ml de volumen en RPMI-1640, FBS al 10 %, P/S al 1 %) y se colocaron en tubos de 15 ml. Se determinó la densidad celular de 150.000 células viables/ml con >98 % de viabilidad, medida con un hemocitómetro y exclusión con azul tripano. Se añadió RNasa A a cada tubo a una concentración de 30 µg/ml y las células se incubaron a 37 °C durante 30 minutos. A continuación, las células se lavaron tres veces con medio de cultivo completo y dos veces con PBS (pH 7,4). Los lavados se realizaron por centrifugación a 250 x g a temperatura ambiente durante 5 minutos, seguida de aspiración y resuspensión. Las concentraciones celulares se determinaron nuevamente con un hemocitómetro, y las células MOPC-21 y MOPC-315 se mezclaron para formar una suspensión celular con una concentración total de 35.000 células/ml en PBS, compuesta de 17.500 células MOPC-21/ml y 17.500 células MOPC-315/ml.

Se coloraron 500 µl de la mezcla de células MOPC-21 y MOPC-315 a un dispositivo de micropocillos de PDMS que se había incubado con BSA para bloquear la adsorción no específica. Se añadieron 17.500 células totales a cada chip. Se usaron cuatro chips en paralelo (70.000 células totales distribuidas en cuatro chips de PDMS), y se permitió que las células se depositaran en los pocillos por gravedad durante 5 minutos con agitación suave. Como cada chip de PDMS contiene aproximadamente 120.000 pocillos y la carga de células se estima eficazmente en aproximadamente un 70 %, aproximadamente 1 de cada 10 pocillos contiene células aisladas. La incidencia de dos células por pocillo se puede estimar con exactitud con las estadísticas de Poisson y, en estas condiciones, >95 % de los pocillos que contienen células contenían una única célula.

A continuación, las superficies de los dispositivos de micropocillos se lavaron con PBS para eliminar las células no adsorbidas de las superficies del chip, y se resuspendieron 25 µl de microesferas magnéticas de poli(T) (kit mRNA Direct, 2,8 µm de diámetro, Invitrogen Corp.) en 50 µl de PBS y se aplicaron a cada superficie del dispositivo de micropocillos, para un promedio de 55 microesferas de poli(T) por pocillo. Después de permitir que las microesferas magnéticas se depositaran en los pocillos por gravedad, sobre cada superficie de chip se colocó una membrana de diálisis bloqueada con BSA (celulosa regenerada con un MWCO de 12.000-14.000, ancho plano de 25 mm, Fisher Scientific) que se enjuagó con PBS. Se eliminó el PBS de las superficies del chip y de la membrana usando una pipeta de 200 µl. A continuación, se cortó el extremo cónico de una punta de pipeta de 1000 µl para formar un cilindro plano que se arrastró a través de las membranas, presionando las membranas hacia los chips de PDMS y eliminando el exceso de PBS entre los dispositivos de micropocillos de PDMS y la membrana de diálisis, sellando así los micropocillos y atrapando las células y las microesferas en su interior (FIG. 1).

La lisis celular y la unión de ARNm a las microesferas magnéticas de poli(T) atrapadas dentro de los micropocillos se realizaron mediante diálisis. Se aplicaron 500 µl de solución de lisis celular (LiCl 500 mM en tampón Tris 100 mM (pH 7,5) con desoxicolato de sodio al 0,1 % y complejo de vanadil ribonucleósido 10 mM) a las membranas de diálisis, y la lisis se produjo a temperatura ambiente mientras se dializaban los reactivos en los micropocillos. La lisis celular se completó completamente en <5 minutos, como se determinó por microscopia de lapso de tiempo (FIG. 2).

Los chips de micropocillos de PDMS se mantuvieron durante 20 minutos a temperatura ambiente dentro de una placa de Petri y, a continuación, se colocaron en una sala fría a 4 °C durante 10 minutos adicionales. Se colocó un imán Dynal MPC-S debajo del dispositivo de micropocillos de PDMS para mantener las microesferas magnéticas dentro de los micropocillos mientras se retiraba la membrana de diálisis con unas pinzas y se desechaba. El imán se colocó a continuación debajo de otra placa de Petri con 4 subdivisiones, una de las cuales contenía 2 ml de tampón de lisis/unión frío del kit mRNA Direct (Tris 100 mM, pH 7,5, LiCl 500 mM, EDTA 10 mM, LiDS al 1 %, DTT 5 mM). Los cuatro dispositivos de micropocillos de PDMS se invirtieron y resuspendieron secuencialmente en los 2 ml de solución para permitir que el imán arrastrara las microesferas fuera de los micropocillos y los introdujera en la solución tampón de lisis/unión del kit mRNA Direct. Las microesferas magnéticas se resuspendieron en los 2 ml de tampón de lisis/unión del kit mRNA Direct y la solución se dividió en dos tubos Eppendorf y se colocó en el separador magnético Dynal MPC-S. Las microesferas se lavaron una vez sin resuspensión usando 1 ml por tubo de tampón de lavado 1 (Tris 100 mM, pH 7,5, LiCl 500 mM, EDTA 1 mM, 4 °C). A continuación, las microesferas se lavaron de inmediato nuevamente en el tampón de lavado 1 con resuspensión. A continuación, las microesferas se resuspendieron de inmediato en el tampón de lavado 2 (Tris 20 mM, pH 7,5, KCl 50 mM, MgCl 3 mM) y se volvieron a colocar en el separador magnético. Finalmente, las microesferas se suspendieron en 2,85 ml de mezcla fría de RT-PCR (Quanta OneStep Fast, VWR) que contenía BSA al 0,05 % en peso (Invitrogen Ultrapure BSA, 50 mg/ml) y las concentraciones de cebador enumeradas en la Tabla 1. La amplificación se realizó con dos cebadores comunes (CHrev-AHX89 y CLrev-BRH06) a alta concentración que se hibridan al complemento reverso del extremo 5' de los cebadores específicos CLrev y CHrev. Los cebadores de la región V también contienen secuencias conectoras en el extremo 5' para efectuar el enlace VH-VL. Se reservaron previamente 25 µl de la mezcla fría de RT-PCR para ciclar sin microesferas o emulsionar como control sin molde. La mezcla fría de RT-PCR que contenía las microesferas magnéticas de poli(T) se añadió gota a gota a un tubo de dispersión Ika en agitación (DT-20, VWR) que contenía 9 ml de fase oleosa enfriada (aceite mineral de calidad para biología molecular con Span-80 al 4,5 %, Tween 80 al 0,4 %, Triton X-100 al 0,05 %, % en v/v, siendo los reactivos de la fase oleosa de Sigma Aldrich Corp.), y la mezcla se agitó durante 5 minutos a baja velocidad. La emulsión resultante se añadió a placas de PCR de 96 pocillos, con 100 µl de emulsión por pocillo, y se colocaron en un termociclador. La etapa de RT se realizó en las siguientes condiciones: 30 minutos a 55 °C, seguido de 2 min a 95 °C. La amplificación por PCR se realizó a continuación en las siguientes condiciones: tres ciclos de 94 °C durante 30 s de desnaturalización, 57 °C durante 1 min de

hibridación y 72 °C durante 3 min de extensión; a continuación veintisiete ciclos de 94 °C durante 30 s de desnaturalización, 59 °C durante 30 s de hibridación y 72 °C durante 3 min de extensión; a continuación, una etapa de extensión final durante 7 min a 72 °C. La FIG. 3 muestra un diagrama de los productos unidos finales.

5 **Tabla 1:** Cebadores para RT-PCR de unión en emulsión de MOPC-21/MOPC-315

Conc.	ID del cebador
400	CLrev-BRH06
400	CHrev-AHX89
40	MOPC21-CHrev-AHX89
40	MOPC21-CLrev-BRH06
40	MOPC315-CLrev-BRH06
40	MOPC315-CHrev-AHX89
40	MOPC21-VH-OE2
40	MOPC21-VL-OE2
40	MOPC315-VH-OE
40	MOPC315-VL-OE

Después del termociclado, la emulsión se recogió y se dividió en tres tubos Eppendorf y se centrifugó a temperatura ambiente durante 10 minutos a 16.000 x g. La fase superior del aceite mineral se desechó y se añadieron 1,5 ml de éter dietílico para extraer la fase oleosa restante y romper la emulsión. Se retiró la fase superior de éter y se realizaron dos extracciones adicionales con éter. A continuación, se desechó la fase de éter y se eliminó el disolvente de éter residual en un SpeedVac durante 25 minutos a temperatura ambiente. La fase acuosa restante se diluyó 5:1 en tampón de unión a ADN; a continuación, se dividió en tres partes y se pasó a través de tres columnas de centrifugación de sílice (DNA Clean & Concentrator, Zymo Research Corp.) para capturar el producto de ADNc de RT-PCR. Después de lavar cada columna con 300 µl de tampón de lavado (Zymo Research Corp), el ADNc se eluyó con 20 µl en cada columna, y se realizó una reacción de PCR con cebadores internos (tampón de PCR ThermoPol con polimerasa Taq, New England Biosciences) en un volumen total de 200 µl usando 4 µl de ADNc eluido como molde. Después de una etapa de desnaturalización de 2 min a 94 °C, el ciclo se realizó a 94 °C durante 30 s de desnaturalización, 62 °C durante 30 s de hibridación, 72 °C durante 20 s de extensión, durante 30 ciclos. Se usaron 400 nM de cada cebador interno (Tabla 2) para amplificar las cadenas pesada y ligera unidas, lo que generó un producto unido de aproximadamente 800 pb.

10 **Tabla 2:** Cebadores para PCR con cebadores internos de MOPC-21/MOPC-315

Conc.	ID del cebador
400	MOPC21-CHrev-seq
400	MOPC21-CLrev-seq
400	MOPC315-CHrev-seq
400	MOPC315-CLrev-seq

El producto de PCR con cebadores internos se sometió a electroforesis en un gel de agarosa al 1 %, y la banda de 800 pb se cortó y se disolvió en tampón de disolución de agarosa durante 10 minutos a 50 °C; a continuación, se capturó en y se eluyó de una columna de centrifugación de sílice de acuerdo con los protocolos del fabricante (Zymo Research Corp.) para obtener un producto de PCR con cebadores internos purificado. El ADNc purificado se sometió a lecturas de pares de bases de ambos extremos con la plataforma de secuenciación Illumina HiSeq. También se puede usar para este propósito otra tecnología de secuenciación NextGen (por ejemplo, Roched 454, Pacific Biosciences, etc.) que pueda proporcionar lecturas adecuadas para identificar el transcrito unido. (FIG. 3). Los datos de HiSeq obtenidos se asignaron a secuencias MOPC-21 y MOPC-315 conocidas usando el programa informático SHort Read Mapping Package (SHRIMP) y se filtraron para lecturas de alta calidad con >90 % de identidad con secuencias de transcritos conocidas. De esta manera, se obtuvieron aproximadamente 18.000 secuencias de cadena pesada y ligera unidas (Tabla 3).

25 **Tabla 3:** Recuentos de lectura sin procesar para pares VH-VL secuenciados

		Ligera	
		MOPC-21	MOPC-315
Pesada	MOPC-21	9.689	426
	MOPC-315	1.042	6.591

Los emparejamientos de transcritos correctos se determinaron adicionalmente a partir del grado de asimetría de emparejamiento de los datos de secuenciación de ADN sin procesar. Para dos transcritos cualesquiera dados, por ejemplo, la cadena pesada H_i y la cadena ligera L_j de inmunoglobulina, con frecuencias globales asignadas de cadena pesada o ligera f_{Hi} y f_{Lj} , se calcula una medida de la asimetría de emparejamiento, s :

$$s = \frac{\text{Lecturas observadas}}{\text{Lecturas esperadas (emparejamiento aleatorio)}} = \frac{[(n.^{\circ} \text{ de pares } H_i)L_j]}{f_{Hi} \times f_{Lj} \times (n.^{\circ} \text{ total de pares})}$$

El valor de s calculado compara la frecuencia con la que VL se empareja con una VH particular con la frecuencia de VL en todo el conjunto de secuencias. Un valor de $s > 1$ indica que se observa un emparejamiento cadena pesada-cadena ligera a una frecuencia superior a la correspondiente al emparejamiento aleatorio. Los emparejamientos naturales se restan de las entradas con un valor máximo de s (Tabla 4). La asimetría de emparejamiento, s , para pares de cadena pesada y ligera secuenciados, calculada a partir de aproximadamente 18.000 pares VH-VL secuenciados se muestra en la Tabla 4. Los emparejamientos naturales de cadena pesada y ligera se predicen por el valor máximo de s para cada cadena pesada y se resaltan en verde. Esta tabla demuestra la capacidad de nuestro procedimiento para resolver emparejamientos naturales de cadena pesada y ligera en una mezcla heterogénea de células.

Tabla 4: Asimetría de emparejamiento calculada

		Ligera	
		MOPC-21	MOPC-315
Pesada	MOPC-21	1,58	0,11
	MOPC-315	0,23	2,18

Ejemplo 3 - Análisis de emparejamiento de transcritos de alto rendimiento usando mezclas definidas de 5 líneas celulares

Se mezclaron cinco líneas de linfocitos B inmortalizadas en diferentes proporciones y se usaron para examinar la eficacia del emparejamiento de los productos unidos generados por PCR OE. Las cinco líneas de linfocitos B usadas en este experimento fueron: MOPC-21, MOPC-315, IM-9, ARH-77 y DB (véase la Tabla 5). DB expresa niveles extremadamente bajos de transcritos de VH y VL y se usó como control negativo.

Todas las líneas celulares se obtuvieron de ATCC y se cultivaron en RPMI-1640 complementado con FBS al 10 % y penicilina/estreptomicina al 1 % (véase el Ejemplo 2). Después de un tratamiento con RNasa durante 30 minutos y posterior lavado, las células se sembraron en micropocillos a una densidad de 17.500 células totales por chip junto con microesferas magnéticas de poli(T) de acuerdo con el Ejemplo 2. Los pocillos se sellaron con una membrana de diálisis, se lisaron las células y se permitió que el ARNm se hibridara con las microesferas (Ejemplo 2). A continuación, se recuperaron las microesferas, se resuspendieron en una mezcla de RT-PCR OE y se colocaron en una emulsión (Ejemplo 2). Las concentraciones de cebador de RT-PCR OE usadas se dan en la Tabla 6, y las condiciones de termociclado se presentan en la Tabla 7.

Tabla 5: Una visión general de las 5 líneas celulares

% en mezcla	Línea celular	ID de ATCC	Organismo	Clase Ig	Expresión relativa de Ig
65	IM-9	CCL-159	<i>Homo sapiens</i>	IgG / IgK	Baja
35	MOPC-21	63035	<i>Mus musculus</i>	IgG / IgK	Media
6	ARH-77	CRL-1621	<i>Homo sapiens</i>	IgG / IgK	Alta
3	MOPC-315	TIB-23	<i>Mus musculus</i>	IgA / IgL	Media
1	DB	CRL-2289	<i>Homo sapiens</i>	IgG / IgL	Muy baja

Tabla 6: Cebadores de RT-PCR OE para la mezcla de líneas celulares

Conc.	ID del cebador
400	CLrev-BRH06
400	CHrev-AHX89
40	MOPC21-CHrev-AHX89
40	MOPC21-CLrev-BRH06

ES 2 774 165 T3

40	MOPC315-CLrev-BRH06
40	MOPC315-CHrev-AHX89
40	MOPC21-VH-OE2
40	MOPC21-VL-OE2
40	MOPC315-VH-OE
40	MOPC315-VL-OE
40	hlgG-rev-OE-AHX89
40	hlgKC-rev-OE-BRH06
40	hlgLC-rev-OE-BRH06
40	hVH1-fwd-OE
40	hVH157-fwd-OE
40	hVH2-fwd-OE
40	hVH3-fwd-OE
40	hVH4-fwd-OE
40	hVH4-DP63-fwd-OE
40	hVH6-fwd-OE
40	hVH3N-fwd-OE
40	hVK1-fwd-OE
40	hVK2-fwd-OE
40	hVK3-fwd-OE
40	hVK5-fwd-OE
40	hVLI-fwd-OE
40	hVL1459-fwd-OE
40	hVL15910-fwd-OE
40	hVL2-fwd-OE
40	hVL3-fwd-OE
40	hVL-DPL16-fwd-OE
40	hVL3-38-fwd-OE
40	hVL6-fwd-OE
40	hVL78-fwd-OE

Tabla 7: Condiciones de termociclado de RT-PCR OE

Nº de ciclos	Temp. (°C)	Tiempo (min)
1	55	30
	94	2
4	94	0,5
	50	0,5
	72	3
4	94	0,5
	55	0,5
	72	3
22	94	0,5
	60	0,5
	72	3
1	72	7

- 5 El producto de RT-PCR OE en emulsión se recuperó por extracción con éter dietílico, seguido de captura en y elución de una columna de centrifugación de sílice (Ejemplo 2) para su uso como molde en una PCR con cebadores internos en las siguientes condiciones: 94 °C durante 2 min de desnaturalización inicial, 94 °C durante 30 s de

desnaturalización, 62 °C durante 30 s de hibridación, 72 °C durante 20 s de extensión, 40 ciclos en total. Las secuencias y concentraciones de cebadores internos se informan en las Tablas 2 y 8.

Tabla 8: Cebadores de PCR con cebadores internos para generar productos unidos de aproximadamente 800 pb

Conc. (nM)	ID del cebador
400	hlgG-all-rev-OEnested
400	hlgKC-rev-OEnested
400	hlgLC-rev-OEnested

El producto de PCR con cebadores internos se sometió a electroforesis en un gel de agarosa al 1 %, se cortó una región de 650 a 1000 pb y se purificó con una columna de centrifugación de sílice (Ejemplo 2). El ADNc recuperado se sometió a secuenciación de ambos extremos de 100 pb Illumina HiSeq. Los datos de HiSeq se asignaron a un archivo de referencia que contenía secuencias de cadena pesada y ligera para los cinco clones, y los datos se filtraron para obtener lecturas de ambos extremos con >90 % de coincidencia con las secuencias de referencia, como en el Ejemplo 2. Los emparejamientos naturales se identificaron consultando los datos de asimetría de emparejamiento. Para cualquier cadena pesada H_i y la cadena ligera L_j de inmunoglobulina, con frecuencias globales asignadas de cadena pesada o ligera f_{Hi} y f_{Lj} , se calculó una medida de la asimetría de emparejamiento, s :

$$s = \frac{\text{Lecturas observadas}}{\text{Lecturas esperadas (emparejamiento aleatorio)}} = \frac{[(n.^{\circ} \text{ de pares } H_i)L_j]}{f_{Hi} \times f_{Lj} \times (n.^{\circ} \text{ total de pares})}$$

El valor de s calculado compara la frecuencia con la que VL se empareja con una VH particular con la frecuencia de VL en todo el conjunto de secuencias. Un valor de $s > 1$ indica que se observa un emparejamiento cadena pesada-cadena ligera a una frecuencia superior a la correspondiente al emparejamiento aleatorio. Los emparejamientos naturales se restan de las entradas con un valor máximo de s para cada cadena pesada. La Tabla 9 muestra los emparejamientos naturales identificados y la asimetría de emparejamiento, s , para pares de cadena pesada y ligera secuenciados, calculada a partir de aproximadamente 66.000 pares VH-VL secuenciados. Los emparejamientos naturales de cadena pesada y ligera se predijeron por el valor máximo de s para cada cadena pesada y se resaltan en verde. La Tabla 9 demuestra la capacidad de nuestro procedimiento para resolver emparejamientos naturales de cadena pesada y ligera en una mezcla heterogénea de células con alto rendimiento.

Tabla 9: Resolución de pares naturales de cadena pesada y ligera

		Ligera			
		IM-9	MOPC-21	ARH-77	MOPC-315
Pesada	IM-9	9,11	0,29	0,48	0,00
	MOPC-21	0,87	1,48	0,38	0,93
	ARH-77	0,86	0,08	2,25	0,21
	MOPC-315	1,91	0,17	1,89	23,84

Ejemplo 4 - Procedimiento para unir dos transcritos de linfocitos B individuales atrapados dentro de placas de micropocillos de alta densidad

Se deja que una población de linfocitos B se depositen por gravedad en placas de micropocillos de PDMS, construidas como se describe en el Ejemplo 1. En este ejemplo, cada portaobjetos de PDMS contiene $1,7 \times 10^5$ pocillos, de modo que cuatro portaobjetos procesados simultáneamente admiten 68.000 linfocitos en una ocupación $\geq 1:10$ célula/pocillo, lo que da al menos un 95 % de probabilidad de que haya solo una célula por pocillo en base a estadísticas de Poisson. Las microesferas magnéticas de poli(dT) con un diámetro de 2,8 μm se depositan en los micropocillos a un promedio de 55 microesferas/pocillo y los portaobjetos se cubren con una membrana de diálisis. Posteriormente, los portaobjetos cubiertos con membrana se incuban con una solución de lisis celular optimizada que contiene dodecilsulfato de litio al 1 % que da como resultado una lisis celular completa en <1 min. El ARNm se hibrida a las microesferas magnéticas de poli(dT) que se recogen, lavan y resuspenden en solución con reactivos, cebadores, enzima retrotranscriptasa y enzima polimerasa para RT-PCR con extensión de solapamiento (OE). De esta manera, las microesferas se aíslan dentro de las gotitas que comprenden la emulsión de agua en aceite. La emulsión se somete a termociclado para unir físicamente los dos transcritos (por ejemplo, ADNc de cadena pesada y ligera de inmunoglobulina), y los productos unidos se recuperan de la emulsión después del ciclado. Se realiza una amplificación por PCR con cebadores internos y, a continuación, el ADN resultante se secuencia usando Illumina o cualquier otra tecnología de secuenciación NextGen que pueda proporcionar lecturas

de longitud apropiada para interpretar inequívocamente la información de emparejamiento de los transcritos. Una visión general del procedimiento se presenta en la FIG. 6.

5 El procedimiento explicado anteriormente se empleó para unir la cadena pesada variable (VH) y la cadena ligera variable (VL) de inmunoglobulina en mezclas de células primarias humanas.

10 Se vacunó a un varón sano de 30 años de edad con la vacuna antigripal trivalente FluVirin de 2010-2011 (Novartis) y se extrajo sangre el día 14 después de la vacunación después de obtener el consentimiento informado. Las PBMC se aislaron y se resuspendieron en DMSO/FCS al 10 % para su crioconservación. Las PBMC congeladas se descongelaron y las suspensiones celulares se tiñeron en PBS/BSA al 0,2 % con anti-CD19 humano (HIB19, BioLegend, San Diego, CA), anti-CD27 humano (0323, BioLegend), anti-CD38 humano (HIT2, BioLegend) y anti-CD3 humano (7D6, Invitrogen, Grand Island, NY). Los linfocitos B de memoria CD19⁺CD3⁺CD27⁺CD38^{int} se separaron usando un sistema separador FACSaria II (BD Biosciences, San Diego, CA). Las células se crioconservaron en DMSO/FCS al 10 % para el posterior emparejamiento VH:VL de alto rendimiento o se separaron como células individuales en placas de 96 pocillos que contenían un cóctel de inhibidores de RNasa (Promega, Madison, WI) y Tris-HCl 10 mM a pH 8,0 para el análisis por RT-PCR de células individuales. El ADNc se sintetizó a partir de células separadas como células individuales usando el kit de síntesis de ADNc Maxima First Strand (Fermentas, Waltham, MA), seguido de la amplificación de los genes variables de inmunoglobulina usando conjuntos de cebadores y condiciones de PCR descritas previamente (Smith *et al.*, 2009). Los genes variables se determinaron con un programa informático de análisis interno usando el motor de búsqueda IMGT (Brochet *et al.*, 2008).

25 Los linfocitos B de memoria congelados para emparejamiento VH:VL de alto rendimiento se descongelaron y recuperaron por centrifugación a 250 x g durante 10 min. Las células se resuspendieron en 200 µl de RPMI-1640 complementado con 1 x GlutaMAX, 1 x aminoácidos no esenciales, 1 x piruvato de sodio y 1 x penicilina/estreptomina (Life Technologies) y se incubaron a 37 °C durante 13 h en una placa de 96 pocillos. Las células recuperadas se centrifugaron nuevamente a 250 x g durante 10 min y se resuspendieron en 400 µl de PBS, y se extrajeron 6 µl para el recuento celular con un hemocitómetro. Se recuperaron aproximadamente 8.800 células de las reservas congeladas. Los linfocitos B de memoria se enriquecieron a continuación con ~880 células IM-9 (número ATCC: CCL-159) como control interno. Las células se resuspendieron sobre dos portaobjetos de micropocillos de PDMS (340.000 pocillos) y se permitió que se depositaran por gravedad en los pocillos durante 5 min con agitación suave. Se ha calculado que el procedimiento de siembra celular tiene una eficacia de un 90 % midiendo la concentración celular en los tampones de siembra, tanto antes como después de la siembra; por tanto, en este experimento se analizaron 8.000 células primarias. La fracción de células aisladas en los estados de célula única y de múltiples células por pocillo se calculó usando estadísticas de Poisson:

35

$$P(k, \mu) = \frac{\mu^k e^{-\mu}}{k!}$$

40 donde k es igual al número de células en un único micropocillo y μ es el número promedio de células por pocillo, de modo que la proporción de 1:39 célula:pocillo usada en este experimento corresponde a que un 98,7 % de las células se depositaron a una ocupación de una célula/pocillo. Se resuspendieron 25 µl de microesferas magnéticas de poli(dT) (kit mRNA Direct de Invitrogen) en 50 µl de PBS y se distribuyeron sobre cada superficie de portaobjetos de PDMS (media de 55 microesferas de poli(dT) por pocillo). Se permitió que las microesferas magnéticas se depositaran en los pocillos por gravedad durante ~5 min; a continuación, sobre cada superficie de portaobjetos se colocó una membrana de diálisis bloqueada con BSA (celulosa regenerada con un MWCO de 12.000-14.000, ancho plano de 25 mm, Fisher Scientific) que se enjuagó con PBS, sellando los micropocillos y atrapando las células y las microesferas en su interior (FIG. 1). Se eliminó el exceso de PBS de las superficies del portaobjetos y de la membrana usando una pipeta de 200 µl. Se aplicaron 500 µl de solución de lisis celular (LiCl 500 mM en tampón TRIS 100 mM (pH 7,5) con dodecilsulfato de litio al 1 %, EDTA 10 mM y DTT 5 mM) a las membranas de diálisis durante 20 min a temperatura ambiente. La microscopía de lapso de tiempo reveló que todas las células se lisaron completamente en un periodo de 1 min (FIG. 2). Posteriormente, los portaobjetos se incubaron a 4 °C durante 10 min, momento en el cual se colocó un imán Dynal MPC-S debajo del dispositivo de micropocillos de PDMS para mantener las microesferas magnéticas dentro de los micropocillos mientras se retiraba la membrana de diálisis con unas pinzas y se desechaba. Los portaobjetos de PDMS se invirtieron rápidamente en una placa de Petri que contenía 2 ml de solución de lisis fría y se colocó el imán debajo de la placa de Petri para sacar las microesferas de los micropocillos. Posteriormente, partes alícuotas de 1 ml de la solución de lisis que contenía microesferas resuspendidas se colocaron en tubos Eppendorf y las microesferas se dejaron sedimentar en un separador magnético Dynal MPC-S y se lavaron una vez sin resuspensión usando 1 ml por tubo de tampón de lavado 1 (Tris 100 mM, pH 7,5, LiCl 500 mM, EDTA 1 mM, 4 °C). Las microesferas se resuspendieron en el tampón de lavado 1, se dejaron sedimentar y se resuspendieron en tampón de lavado 2 (Tris 20 mM, pH 7,5, KCl 50 mM, MgCl 3 mM) y se dejaron sedimentar nuevamente. Finalmente, las microesferas se suspendieron en 2,85 ml de mezcla fría de RT-PCR (Quanta OneStep Fast, VWR) que contenía BSA al 0,05 % en peso (Invitrogen Ultrapure BSA, 50 mg/ml) y conjuntos de cebadores para la amplificación de la unión de VH y VL (FIG. 4 y Tablas 6 y 10). La suspensión que contenía las microesferas magnéticas de poli(dT) se añadió gota a gota a un tubo de dispersión IKA en agitación (DT-20, VWR) que contenía 9 ml de fase oleosa enfriada (aceite mineral de calidad para biología molecular con Span-80 al 4,5 %, Tween 80 al 0,4 %, Triton X-100 al 0,05 %, en % v/v, Sigma-Aldrich, St. Louis, MO), y la mezcla

65

se agitó durante 5 min a baja velocidad. La emulsión resultante se añadió a placas de PCR de 96 pocillos, con 100 µl de emulsión por pocillo, y se colocaron en un termociclador. La etapa de RT se realizó en las siguientes condiciones: 30 min a 55 °C, seguido de 2 min a 94 °C. La amplificación por PCR se realizó en las siguientes condiciones: cuatro ciclos de 94 °C durante 30 s de desnaturalización, 50 °C durante 30 s de hibridación, 72 °C durante 2 min de extensión; cuatro ciclos de 94 °C durante 30 s de desnaturalización, 55 °C durante 30 s de hibridación, 72 °C durante 2 min de extensión; 22 ciclos de 94 °C durante 30 s de desnaturalización, 60 °C durante 30 s de hibridación, 72 °C durante 2 min de extensión; a continuación, una etapa de extensión final durante 7 min a 72 °C. Después del termociclado, la emulsión se inspeccionó visualmente para garantizar la ausencia de una fase acuosa voluminosa, que es un indicador clave de la estabilidad de la emulsión. Después de la verificación visual, la emulsión se recogió y se centrifugó a temperatura ambiente durante 10 min a 16.000 x g, se descartó la fase superior del aceite mineral y se añadieron 1,5 ml de éter dietílico para extraer la fase oleosa restante y romper la emulsión. Se desechó la fase superior de éter, se realizaron dos extracciones adicionales con éter y se eliminó el éter residual en un SpeedVac durante 25 min a temperatura ambiente. La fase acuosa se diluyó 5:1 en tampón de unión a ADN y se pasó a través de una columna de centrifugación de sílice (DNA Clean & Concentrator, Zymo Research, Irvine, CA) para capturar el producto de ADNc. La columna se lavó dos veces con 300 µl de tampón de lavado (Zymo Research Corp) y el ADNc se eluyó en 40 µl de agua sin nucleasas. Finalmente, se realizó una amplificación por PCR con cebadores internos (tampón de PCR ThermoPol con polimerasa Taq, New England Biosciences, Ipswich, MA) en un volumen total de 200 µl usando 4 µl de ADNc eluido como molde con cebadores 400 nM (Tablas 8 y 11) en las siguientes condiciones: desnaturalización inicial a 94 °C durante 2 min, desnaturalización a 94 °C durante 30 s durante 39 ciclos, hibridación a 62 °C durante 30 s y extensión a 72 °C durante 20 s, extensión final a 72 °C durante 7 min. El producto unido de aproximadamente 850 pb (FIG. 3) se extrajo mediante electroforesis en gel de agarosa y se secuenció usando la plataforma MiSeq NextGen de secuenciación de ambos extremos 2x250 (Illumina, San Diego, CA).

Tabla 10: Conjuntos de cebadores para amplificación por RT-PCR de la unión de VH y VL humanas

Conc. (nM)	ID del cebador
40	hlgA-rev-OE-AHX89
40	hlgM-rev-OE-AHX89

Tabla 11: Conjuntos de cebadores para amplificación por PCR con cebadores internos de VH y VL humanas

Conc. (nM)	ID del cebador
400	hlgA-all-rev-OEnested
400	hlgM-rev-OEnested

Para el análisis bioinformático, se filtraron datos sin procesar de MiSeq 2x250 para obtener una puntuación mínima de calidad de Phred de 20 sobre el 50 % de los nucleótidos para garantizar una alta calidad de lectura en la región que contiene CDR3 (aproximadamente nt 65-115 de HC o nt 55-100 de LC). Los datos de secuenciación se sometieron al Sistema Internacional de Información ImMunoGeneTics (IMGT) para mapear los genes de la estirpe germinal V(D)J (Brochet *et al.*, 2008). Los datos de secuenciación se filtraron para los puntos de unión V(D)J sin cambio de pauta de lectura, y las secuencias VH y Vκ, λ productivas se emparejaron con el ID de lectura de Illumina. Las secuencias de nucleótidos de CDR-H3 se extrajeron y agruparon a una identidad de nucleótidos del 96 % con omisiones terminales ignoradas, para generar una lista de CDR-H3 únicas en el conjunto de datos. Se descubrió que el punto de corte de identidad de nucleótidos del 96 % era el punto de corte óptimo para agrupar los errores de secuenciación en clones de control enriquecidos; el número de secuencias únicas de CDR-H3 y, por tanto, el número de genes V únicos informados se refieren al número de grupos recuperados de la muestra (Tabla 12). El recuento de lectura superior de CDR-L3 para cada grupo de CDR-H3 se asignó como un par afín y se generó una lista de pares VH:VL recuperados. La proporción de exactitud observada de 942:1 demostró la conservación de los emparejamientos correctos de cadena pesada y ligera en la línea celular de control enriquecida IM-9 (Tabla 12).

Tabla 12: Estadísticas experimentales clave para el Ejemplo 4

Inmunización	Gripe (Fluvirin 2010-11)
Tipo de célula	Día 14, linfocitos B de memoria
Células frescas vs. Congelación/descongelación	Congelación/descongelación
Proporción célula:pocillo	1:39
% de células como células individuales	98,7 %
CDR-H3 individuales recuperadas	240
Enriquecimiento con células de control	IM-9
Proporción de exactitud¹	942:1

¹ Para células enriquecidas conocidas (lecturas VL correctas):(lecturas VL incorrectas superiores)

Los emparejamientos VH:VL identificados usando este enfoque de alto rendimiento se compararon con los identificados usando el procedimiento de separación de células individuales establecido (Smith *et al.*, 2009; Wrammert *et al.*, 2008); este análisis se realizó con enmascaramiento doble. Los linfocitos B de memoria CD19⁺CD3⁺CD27⁺CD38^{int} se aislaron de un voluntario sano 14 días después de la vacunación con la vacuna antigripal Fluvirin trivalente de 2010-2011 (Smith *et al.*, 2009). Para el análisis scRT-PCR, 168 linfocitos B individuales se distribuyeron en cuatro placas de 96 pocillos y se llevaron a cabo 168 reacciones de RT-PCR y 504 reacciones de PCR con cebadores internos individualmente para amplificar por separado los genes de VH y VL (κ y λ). Los productos de ADN se resolvieron mediante electroforesis en gel y se secuenciaron para proporcionar un total de 51 pares VH:VL, de los cuales 50 eran únicos. Se recuperaron un total de 240 pares CDR-H3:CDR-L3 únicos. En el análisis de RT-PCR de células individuales también se observaron cuatro secuencias de CDR-H3 detectadas en el conjunto de emparejamiento de alto rendimiento. Un análisis con enmascaramiento reveló que los pares CDR-H3:CDR-L3 aislados por los dos enfoques estaban en total concordancia (DeKosky *et al.*, 2013). La concordancia entre los procedimientos de secuenciación por RT-PCR de células individuales establecidos y los procedimientos de secuenciación de alto rendimiento demostró una alta exactitud en las secuencias VH:VL recuperadas de acuerdo con los procedimientos descritos en la presente divulgación.

Ejemplo 5 - Aislamiento de anticuerpos de alta afinidad después del emparejamiento VH:VL de alto rendimiento

Este ejemplo describe el aislamiento de anticuerpos antitetánicos de alta afinidad a partir de linfocitos B periféricos humanos después de la inmunización de refuerzo. Una donante femenina se sometió a inmunización de refuerzo contra TT/toxoide diftérico (TD, 20 I.E. TT y 2 I.E. toxoide diftérico, Sanofi Pasteur Merck Sharpe & Dohme GmbH, Leimen, Alemania) después de obtener el consentimiento informado de la Charité Universitätsmedizin Berlin (las muestras se codificaron anónimamente y el estudio se aprobó por la junta de aprobación ética del hospital, número EA1/178/11, y la Junta de Revisión Institucional de la Universidad de Texas en Austin, IRB n.º 2011-11-0095). Siete días después de la inmunización contra toxoide tetánico (TT), se extrajo sangre EDTA y se aislaron PBMC mediante separación por gradiente de densidad como se describe (Mei *et al.*, 2009). Las PBMC se tiñeron en PBS/BSA a 4 °C durante 15 min con anti-CD3/CD14-PacB humano (clones UCHT1 y M5E2, respectivamente, Becton Dickinson, BD), anti-CD19-PECy7 humano (clon SJ25Cl, BD), anti-CD27-Cy5 humano (clon 2E4, amable obsequio de René van Lier, Academic Medical Center, Universidad de Ámsterdam, Países Bajos, etiquetado en Deutsches Rheumaforschungszentrum (DRFZ), Berlín), anti-CD20-PacO humano (clon HI47, Invitrogen), anti-IgD-PerCpCy5.5 humano (clon L27, BD), anti-CD38-PE humano (clon HIT2, BD) y anti-TT-digoxigenina humana (marcada en DRFZ) durante 15 min a 4 °C. Las células se lavaron y se realizó una segunda tinción con anti-digoxigenina-FITC (Roche, marcada en DRFZ) y se añadió DAPI antes de la separación. Los plasmoblastos CD19⁺CD3⁺CD14⁻CD38⁺⁺CD27⁺⁺CD20⁻TT⁺ se separaron usando un sistema separador FACSAria II (BD Biosciences). Una porción de las células separadas se lavaron y se crioconservaron en DMSO/FCS al 10 % para un emparejamiento VH:VL de alto rendimiento.

Un vial que contenía aproximadamente 2.000 plasmoblastos TT⁺ congelados se descongeló y se recuperaron por centrifugación a 250 x g durante 10 min; se anticipa que aproximadamente un 20-30 % de las células son viables (Kyu *et al.*, 2009). Las células se resuspendieron en 300 μ l de RPMI-1640 complementado con FBS al 10 %, 1 x GlutaMAX, 1 x aminoácidos no esenciales, 1 x piruvato de sodio y 1 x penicilina/estreptomicina (todo de Life Technologies) y se incubaron a 37 °C durante 13 h en una placa de 96 pocillos. Las células recuperadas se centrifugaron nuevamente a 250 x g durante 10 min y se resuspendieron en 400 μ l de PBS, y se extrajeron 6 μ l para el recuento celular con un hemocitómetro. Las células se enriquecieron con aproximadamente 30 células ARH-77 como control interno (número ATCC CRL-1621) y los transcritos VH:VL se unieron como se describe en el Ejemplo 4, omitiendo los cebadores IgM y usando una PCR con cebadores internos de 38 ciclos; el producto resultante se sometió a secuenciación MiSeq 2x250. Las cadenas VH y VL también se amplificaron individualmente para obtener secuencias de VH y VL completas para la expresión de anticuerpos. El producto de PCR con cebadores internos se diluyó 1:9 y se usaron 0,5 μ l como molde en una reacción de PCR con las siguientes condiciones: cebadores 400 nM (Tablas 8, 11 y 13), desnaturalización inicial durante 2 min a 94 °C, desnaturalización a 94 °C durante 30 s durante 12 ciclos, hibridación a 62 °C durante 30 s y extensión a 72 °C durante 15 s, extensión final a 72 °C durante 7 min. Los productos resultantes de VH de 450 pb o VL de ~400 pb se purificaron por electroforesis en gel de agarosa y se sometieron a secuenciación MiSeq 2x250. Los datos de secuenciación se procesaron como se describe anteriormente; adicionalmente, se seleccionaron diez pares de VH y VL de los emparejamientos de plasmoblastos TT⁺ para someterlos a expresión de anticuerpos y someterlos a prueba. Para la secuenciación de anticuerpos completos de estos diez genes, las lecturas de 2 x 250 pb que contenían el gen V 5' FR1-CDR2 y 3' CDR2-FR4 se emparejaron por ID de lectura en Illumina y se construyeron secuencias consenso a partir de lecturas que contenían el CDR3 exacto de interés. Los genes de anticuerpos se clonaron a continuación en los vectores de expresión de IgG humana pMAZ-VH y pMAZ-VL, respectivamente (Mazor *et al.*, 2007). Se cotransfectaron 40 μ g de cada producto de conexión circularizado en células HEK293F (Invitrogen, NY, EE. UU.). El medio se recogió 6 d después de la transfección por centrifugación y la IgG se purificó mediante una columna de cromatografía de agarosa con proteína A (Pierce, IL, EE. UU.).

Tabla 13: Conectores para cebadores de amplificación separados de VH y VL

Conc. (nM)	ID del cebador
400	Conector-VHfwd
400	Conector-VLfwd

5 Las afinidades antigénicas se determinaron mediante ELISA competitivo (Friguet *et al.*, 1985) usando diferentes concentraciones de IgG en una dilución en serie del antígeno, que varía de 100 nM a 0,05 nM en presencia de leche al 1 % en PBS. Las placas se recubrieron durante la noche a 4 °C con 10 µg/ml de TT en tampón carbonato 50 mM, pH 9,6, se lavaron tres veces en PBST (PBS con Tween 20 al 0,1 %) y se bloquearon con leche al 2 % en PBS durante 2 h a temperatura ambiente. Se añadieron muestras preequilibradas de IgG con antígeno TT a la placa de ELISA bloqueada, se incubaron durante 1 h a temperatura ambiente, y las placas se lavaron 3 veces con PBST y se incubaron con 50 µl de anticuerpo secundario anti-cadena ligera kappa humano conjugado con HRP (1:5.000, leche al 2 % en PBS) durante ~2 min a 25 °C. Las placas se lavaron 3 veces con PBST; a continuación, se añadieron 50 µl de sustrato Ultra TMB (Thermo Scientific, Rockford, IL) a cada pocillo y se incubaron a 25 °C durante 5 min. Las reacciones se detuvieron usando un volumen igual de H₂SO₄ 1 M y se leyó la absorbancia a 450 nm (BioTek, Winooski, VT). Cada réplica de ELISA competitiva se ajustó usando una ecuación logística de cuatro parámetros (4PL), con el error representado como la D.E. de 2-3 réplicas para cada IgG analizada. Los diez anticuerpos mostraron especificidad para TT y TT unido con alta afinidad (0,1 nM ≤ KD ≤ 18 nM; Tabla 14) (DeKosky *et al.*, 2013). La alta afinidad de los anticuerpos anti-TT recuperados demuestra la aplicación de los procedimientos de secuenciación de VH:VL de alto rendimiento en la presente divulgación para el descubrimiento de anticuerpos de donantes de células humanas.

Tabla 14: Afinidades de unión a toxoide tetánico (TT) de IgG aisladas por secuenciación de alto rendimiento de pares VH:VL. Las afinidades se calcularon a partir de curvas de dilución de ELISA competitiva.

ID de anticuerpo	Asignación a familia génica ¹	Afinidad (K _D)
TT1	HV3-HD1-HJ6: KV3-KJ5	1,6 ± 0,1 nM
TT2	HV3-HD3-HJ4: LV3-LJ1	14 ± 3 nM
TT3	HV1-HD2-HJ4: KV3-KJ5	3,6 ± 1,8 nM
TT4	HV2-HD2-HJ4: KV1-KJ1	2,7 ± 0,3 nM
TT5	HV4-HD2-HJ6: KV2-KJ3	18 ± 4 nM
TT6	HV1-HD3-HJ4: KV1-KJ2	0,57 ± 0,03 nM
TT7	HV4-HD3-HJ4: KV1-KJ2	0,46 ± 0,01 nM
TT8	HV3-HD3-HJ4: LV8-LJ3	2,8 ± 0,3 nM
TT9	HV4-HD2-HJ4: KV1-KJ1	0,10 ± 0,01 nM
TT10	HV1-HD3-HJ5: KV3-KJ5	1,6 ± 0,1 nM

¹ Cada una de las cadenas pesada y ligera eran distintas.

25 Ejemplo 6 - Identificación bioinformática de secuencias de VH:VL mediante concordancia de emparejamiento mutuo

30 Los ejemplos 4 y 5 divulgan la identificación de pares de secuencias VH:VL correctos a partir de la secuenciación de alto rendimiento, con lo que la secuencia de VL de mayor recuento de lectura para una secuencia de VH dada reveló los pares VH:VL asociados naturales codificados por linfocitos B individuales. De forma alternativa, este ejemplo describe un procedimiento para identificar pares VH:VL correctos entre los datos de amplicón VH:VL de alto rendimiento mediante emparejamiento de consenso de secuencias de VH y de VL.

35 Se recopilan datos de emparejamientos sin procesar y se tabula la máxima frecuencia de VL para cada secuencia de VH en el Archivo 1. La máxima frecuencia de VH para cada VL se tabuló por separado en el Archivo 2. Se pueden usar muchas técnicas computacionales para lograr la etapa de tabulación; por ejemplo, "grep -m 1 CDR3 filename" en el shell Bash/Linux puede seleccionar el par afín mejor clasificado para una secuencia de CDR-H3 o CDR-L3 (CDR3) de un archivo (filename) que contiene datos de emparejamiento sin procesar que se han preclasificado para contener secuencias ordenadas por recuentos de lectura descendentes. Otras soluciones para la tabulación de datos incluyen el uso de un hash para recopilar secuencias y recuentos de lectura de secuencias (por ejemplo, lenguaje informático Perl), o el uso de un diccionario para recopilar secuencias y recuentos de lectura (por ejemplo, Python) u otras estructuras de almacenamiento de datos (por ejemplo, memorias asociativas o matrices asociativas). El Archivo 1 y el Archivo 2 se compararon y cualquier par VH:VL que aparecía en ambos archivos
45 mostró "consenso" en que el par descrito por la VL mejor clasificada para una VH dada coincidía con el de la VH mejor clasificada para una VL dada. Se pueden aplicar muchas técnicas computacionales para lograr

comparaciones de archivos; una solución para la comparación de archivos usa el comando "join" en Bash/Linux, donde las líneas que contienen los campos deseados que coinciden entre documentos se imprimen en la salida estándar. El algoritmo descrito en el presente ejemplo fue eficaz tanto en la identificación de los pares VH:VL correctos como en la reducción de errores de secuenciación menores, ya que los pares VH:VL que contenían errores de secuenciación a menudo se filtran por criterios de concordancia mutuos. Un árbol de decisiones general del algoritmo usado para el emparejamiento se proporciona en la FIG. 9.

Ejemplo 7 - Emparejamiento VH:VL de linfocitos B de memoria expandidos

Los linfocitos B de memoria se aislaron y expandieron *in vitro*, y dos partes alícuotas de las células expandidas se procesaron para emparejamiento de alto rendimiento. La expansión clonal *in vitro* da como resultado múltiples copias de células que contienen los mismos pares VH:VL, aumentando por tanto la probabilidad de secuenciar el mismo par VH:VL en partes alícuotas separadas derivadas de la misma muestra de linfocitos B.

Se aislaron PBMC de sangre humana donada y se tiñeron con CD20-FITC (clon 2H7, BD Biosciences, Franklin Lakes, NJ, EE. UU.), CD3-PerCP (HIT3a, BioLegend, San Diego, CA, EE. UU.), CD19-v450 (HIB19, BD) y CD27-APC (M-T271, BD). Los linfocitos B de memoria CD3⁻ CD19⁺CD20⁺CD27⁺ se incubaron durante cuatro días en presencia de RPMI-1640 complementado con FBS al 10 %, 1 x GlutaMAX, 1 x aminoácidos no esenciales, 1 x piruvato de sodio y 1 x penicilina/estreptomicina (todo de Life Technologies) junto con 10 µg/ml de anticuerpo anti-CD40 (5C3, BioLegend), 1 µg/ml de cPg ODN 2006 (Invivogen, San Diego, CA, EE. UU.), 100 unidades/ml de IL-4, 100 unidades/ml de IL-10 y 50 ng/ml de IL-21 (PeproTech, Rocky Hill, NJ, EE. UU.). Se sembraron 91.000 linfocitos B expandidos en 12 chips y, después de alcanzar una proporción de eficacia de siembra en pocillos estimada del 90 %, se analizaron aproximadamente 41.000 linfocitos B expandidos por grupo (proporción 1:25 célula:pocillo) de acuerdo con los procedimientos descritos en el Ejemplo 4. El análisis bioinformático se realizó como se describe en el Ejemplo 6. Se secuenciaron 1.033 secuencias de CDR-H3 con >1 lectura en ambos grupos, y 972/1.033 mostraron pares de CDR-L3 coincidentes para proporcionar una fracción de coincidencia del 94,09 %. La exactitud de emparejamiento, A_P, se puede estimar a partir de la fracción de coincidencia de CDR-L3, f_{coincidencia}, de los dos grupos independientes:

$$f_{coincidencia} = A_{P,Grupo1} \times A_{P,Grupo2} = A_P^2$$

$$A_P = f_{coincidencia}^{1/2}$$

que proporcionó una exactitud general de un 97,0 %. El límite teórico de exactitud de la tasa de células individuales por pocillo por distribución de Poisson (98 % para la proporción 1:25 célula:pocillo utilizada en este experimento) se correlacionó muy estrechamente con la exactitud de los emparejamientos VH:VL determinada experimentalmente.

Ejemplo 8 - Uso de cebadores de péptidos líder para emparejamiento VH:VL

En este ejemplo se usaron cebadores que hibridan a la región del péptido líder de los ADNc de anticuerpos (a diferencia de los cebadores específicos para el marco 1 de los dominios de VH y VL, divulgados en el Ejemplo 4) para secuenciar los pares VH:VL de anticuerpos. Los linfocitos B de memoria se aislaron de PBMC humanas donadas, y las células se dividieron en dos grupos: El Grupo 1 consistió en 29.000 células y se analizó inmediatamente (usando un total de 510.000 pocillos, proporción 1:16 célula:pocillo), mientras que el Grupo 2 se expandió como se describe en el Ejemplo 7 y se analizaron 28.000 células después de la expansión *in vitro* (usando un total de 680.000 pocillos, proporción 1:24 célula:pocillo). Ambos experimentos se llevaron a cabo como se describe en el Ejemplo 7 usando cebadores para extensión de solapamiento de péptidos líder informados en la Tabla 15 y un ciclado por RT-PCR de unión en emulsión con las siguientes condiciones: 30 min a 55 °C, seguido de 2 min a 94 °C; cuatro ciclos de 94 °C durante 30 s de desnaturalización, 54 °C durante 30 s de hibridación, 72 °C durante 2 min de extensión; 29 ciclos de 94 °C durante 30 s de desnaturalización, 60 °C durante 30 s de hibridación, 72 °C durante 2 min de extensión; a continuación, una etapa de extensión final durante 7 min a 72 °C. También se incluyó una región con código de barras adicional en los cebadores de unión a VL (región 16N) que se usó para identificar lecturas de secuencias múltiples de acontecimientos de unión individuales (Tabla 15). La PCR con cebadores internos se realizó como en el Ejemplo 5, con 25 ciclos de PCR para cada grupo.

Tabla 15: Cebadores de RT-PCR con extensión de solapamiento dirigidos a la región del péptido líder del ARNm de anticuerpo

Conc. (nM)	ID del cebador
400	CHrev-AHX89
400	CLrev-BRH06
40	hIgG-rev-OE-AHX89
40	hIgA-rev-OE-AHX89

40	hIgM-rev-OE-AHX89
40	hIgKC-rev-OE-BRH06
40	hIgLC-rev-OE-BRH06
40	VH_L
40	VH3_L
40	VH4/6_L
40	VH5_L
40	hVλ1for_L
40	hVλ2for_L
40	hVλ3for_L
40	hVλ3for-2_L
40	hVλ3for-3_L
40	hVλ4/5for_L
40	hVλ6for_L
40	hVλ7for_L
40	hVλ8for_L
40	hVκ1/2for_L
40	hVκ3for_L
40	hVκ4for_L

Después de la secuenciación de alto rendimiento con Illumina de 2 x 250 pb de productos de PCR con cebadores internos, 23 de las 23 CDR-H3 observadas con ≥ 2 lecturas en ambos grupos de péptidos líder mostraron una CDR-L3 coincidente. Este ejemplo demuestra que se pueden usar diversos conjuntos de cebadores para secuenciar múltiples transcritos usando los procedimientos de la presente divulgación.

Ejemplo 9 - Formación de gotitas de agua en aceite de una única célula de baja dispersión usando una boquilla y una corriente portadora anular

En este ejemplo, las líneas de linfocitos B inmortalizados MOPC-21 se encapsularon de manera viable en gotitas de emulsión de tamaño controlado que consistían en una mezcla de células en PBS y tinte de azul tripano para la visualización de la viabilidad celular. Este ejemplo demuestra el aislamiento de células individuales en gotitas de emulsión de distribución de tamaño controlada; estando además las gotitas compuestas de dos corrientes acuosas diferentes que se mezclan inmediatamente antes de la formación de las gotitas (FIG. 7).

Las células MOPC-21 se resuspendieron a una concentración de 500.000 células/ml de PBS. Se construyó un aparato de emulsificación coaxial insertando una aguja de calibre 26 (Hamilton Company, Reno, NV, EE. UU.) dentro de un tubo hipodérmico de calibre 19 (Hamilton) y la aguja se ajustó de modo que la punta de la aguja quedara al ras del extremo del tubo hipodérmico. Se colocaron agujas concéntricas en el interior de un tubo de vidrio con un diámetro externo de 3/8 de pulgada (Wale Apparatus, Hellertown, PA, EE. UU.) con un orificio de 140 μm de modo que la salida de la aguja quedara aproximadamente a 2 mm del orificio de la boquilla. La solución acuosa de PBS/células se inyectó a través de la aguja a una velocidad de 500 $\mu\text{l}/\text{min}$, mientras que se inyectó una solución de PBS/azul tripano al 0,4 % (Sigma-Aldrich, St. Louis, MO, EE. UU.) a través del tubo hipodérmico de calibre 19, y se pasó una fase oleosa (aceite mineral de calidad para biología molecular con Span-80 al 4,5 %, Tween 80 al 0,4 %, Triton X-100 al 0,05 %, en % v/v, Sigma Aldrich Corp.) a través del tubo de vidrio a una velocidad de 3 ml/min. Las gotitas suspendidas en la fase oleosa se recogieron en un tubo Eppendorf de 2 ml. Se usó una bomba de jeringa (KD Scientific Legato 200, Holliston, MA, EE. UU.) para controlar los caudales acuosos y una bomba de engranajes (M-50, Valco Instruments, Houston, TX, EE. UU.) para controlar los caudales de aceite, y las emulsiones resultantes se analizaron mediante microscopía óptica. Se generaron gotitas con un diámetro medio de aproximadamente 85 μm y las células individuales encapsuladas mostraron una alta viabilidad medida por exclusión del azul tripano (FIG. 5).

Ejemplo 10 - Secuenciación de múltiples transcritos en linfocitos B mediante encapsulación en gotitas de emulsión

En este ejemplo, la lisis celular y la hibridación de ARNm a microesferas de poli(T) se realizaron dentro de una emulsión generada usando el procedimiento explicado en el Ejemplo 9. Se aisló una población de linfocitos B de memoria y las células se expandieron como en el Ejemplo 7. Los linfocitos B de memoria se resuspendieron en PBS a una concentración de 100.000 células/ml y se pasaron a través de la aguja de calibre 26 más interna del dispositivo generador de emulsión del Ejemplo 8 a una velocidad de 500 $\mu\text{l}/\text{min}$. Se dejaron sedimentar 450 μl de microesferas magnéticas de poli(dT) (1,0 μm de diámetro, New England Biosciences, Ipswich, MA, EE. UU.) con un imán y se resuspendieron en 5 ml de tampón de lisis/unión celular (Tris 100 mM, pH 7,5, LiCl 500 mM, EDTA 10 mM,

dodecilsulfato de litio al 0,5 %, DTT 5 mM), y la mezcla resultante se pasó a través del tubo hipodérmico de calibre 19 a una velocidad de 500 µl/min, mientras que la fase oleosa (aceite mineral de calidad para biología molecular con Span-80 al 4,5 %, Tween 80 al 0,4 %, Triton X-100 al 0,05 %, en % v/v, Sigma Aldrich Corp.) se pasó a través del tubo de vidrio más externo a una velocidad de 3 ml/min para generar una emulsión que consistía en gotitas acuosas de aproximadamente 85 µm de diámetro que contenían células individuales. La corriente de emulsión se recogió en tubos Eppendorf de 2 ml, y las células se lisaron con detergente a medida que se generaron las gotitas para permitir la captura de ARNm en microesferas magnéticas de poli(dT) encapsuladas dentro de las gotitas de emulsión.

Cada uno de los tubos de emulsión de 2 ml se mantuvo a temperatura ambiente durante tres minutos antes de colocarse en hielo durante un mínimo de diez minutos. A continuación, los tubos se centrifugaron a 16.000 x g durante 5 minutos a 4 °C, y la fase superior de aceite mineral se retiró y desechó. Se añadieron 200 µl de éter dietílico frío para romper químicamente la emulsión y los tubos se centrifugaron a 16.000 x g durante 2,5 minutos para permitir la sedimentación de las microesferas magnéticas. Se extrajeron microesferas magnéticas usando una pipeta, se dejaron sedimentar y se resuspendieron en 2 ml de tampón de lisis/unión (Tris 100 mM, pH 7,5, LiCl 500 mM, EDTA 10 mM, LiDS al 0,5 %, DTT 5 mM). Las microesferas se lavaron y se resuspendieron en una mezcla de RT-PCR OE como en el Ejemplo 8. Se usaron cebadores de péptidos líder, cuyas concentraciones se dan en la Tabla 15. La suspensión de microesferas en la mezcla de RT-PCR OE se emulsionó y se sometió a termociclado, se extrajo el ADNc y se realizó una PCR con cebadores internos (véase el Ejemplo 8). El producto de PCR con cebadores internos se sometió a electroforesis para purificar los transcritos unidos, que a continuación se secuenciaron como en el Ejemplo 8 anterior.

Después de la secuenciación de alto rendimiento con Illumina de 2 x 250 pb de los productos de PCR con cebadores internos, se recuperaron 14.121 pares VH:VL con ≥ 2 lecturas de acuerdo con el algoritmo descrito en el Ejemplo 6 (7.367 pares VH:VL en el Grupo 1 y 6.754 pares en el Grupo 2). Se observaron 3.935 CDR-H3 con ≥ 2 lecturas en ambos grupos. 3.899 de las 3.935 CDR-H3 observadas en ambos grupos mostraron una CDR-L3 coincidente, lo que indica una exactitud general de un 99,5 % de acuerdo con la fórmula explicada en el Ejemplo 7. El presente ejemplo demuestra la secuenciación de múltiples transcritos a través de la captura de ARNm de células individuales aisladas dentro de gotitas de emulsión.

30 **Ejemplo 11 - Secuenciación paralela de ADNc de cadena pesada y ligera de células individuales**

Los ejemplos anteriores demostraron el uso de microesferas magnéticas para capturar ARNm y la unión covalente de los ADNc deseados de una única célula (por ejemplo, ADNc de VH y VL) para crear un único amplicón. Los amplicones VH-VL individuales generados de este modo se secuenciaron mediante secuenciación de ADN de alto rendimiento para revelar el repertorio de secuencias de VH y VL emparejadas naturalmente.

En el ejemplo, los ADNc capturados en microesferas se secuenciaron directamente sin unirse (es decir, sin crear un amplicón VH-VL unido). De esta manera, se reveló la identidad de los transcritos deseados de una única célula sin necesidad de realizar una PCR con extensión de solapamiento. En primer lugar, se conjugó una mezcla equitativa de tres cebadores 5'-amino funcionalizados (Tabla 17) con microesferas magnéticas funcionalizadas de modo que los oligonucleótidos inmovilizados en cada microesfera magnética estuvieran en la siguiente proporción: 1/3 de poli(T) para la captura de ARNm, 1/3 de cebador específico para el transcrito 1 deseado (por ejemplo, el cebador AHX89 de la Tabla 1) y 1/3 de cebador específico para el transcrito 2 deseado (Tabla 17). Estas microesferas magnéticas conjugadas con cebador cumplieron un doble propósito: en primer lugar, tras la lisis, los cebadores de poli(T) capturaron el ARNm de cadena pesada y ligera de células individuales, como en los Ejemplos 4-6; en segundo lugar, en la etapa de RT-PCR en emulsión, los cebadores AHX89 y BRH06 provocaron que el ADNc de cadena pesada y ligera se amplificara en la superficie de las microesferas. Después de la RT-PCR, las microesferas magnéticas se usaron como molde de secuenciación para la secuenciación de alto rendimiento. El procedimiento se explica en la FIG. 7.

Se inmovilizó una mezcla equitativa de tres oligonucleótidos 5'-amina (Tabla 16) en microesferas magnéticas funcionalizadas de acuerdo con los protocolos del fabricante (microesferas de ácido carboxílico Dynal MyOne, 1,0 µm de diámetro, Invitrogen Corp.). A continuación, una mezcla de células inmortalizadas MOPC-21 y MOPC-315 se lavó y suspendió a 100.000 células/ml en PBS (pH 7,4). Se añadieron $1,2 \times 10^8$ microesferas magnéticas funcionalizadas por ml de solución de lisis celular/unión a ARNm, como se explica en el Ejemplo 10. La suspensión de células/microesferas se emulsionó como en el Ejemplo 10, las células se lisaron y el ARNm se hibridó con las microesferas. A continuación, se recuperaron las microesferas rompiendo la emulsión, se lavaron como se describe en el Ejemplo 10 y se realizó la RT-PCR en emulsión. Las concentraciones de los cebadores de RT-PCR se dan en la Tabla 17. Las condiciones de ciclado fueron como sigue: 30 min a 55 °C, seguido de 2 min a 94 °C; cuatro ciclos de 94 °C durante 30 s de desnaturalización, 57 °C durante 1 min de hibridación, 72 °C durante 2 min de extensión; 29 ciclos de 94 °C durante 30 s de desnaturalización, 59 °C durante 30 s de hibridación, 72 °C durante 2 min de extensión; a continuación, una etapa de extensión final durante 7 min a 72 °C.

Tabla 16: Cebadores conjugados con la superficie de las microesferas magnéticas

Conc.	ID del cebador
33%	oligodT(25)-5'amina
33%	CHrev-AHX89-5'amina
33%	CLrev-BRH06-5'amina

Tabla 17: Cebadores en la mezcla de RT-PCR de MOPC-21/MOPC-315

Conc.	ID del cebador
400	CHrev-AHX89
400	CLrev-BRH06
40	MOPC21-CHrev-AHX89
40	MOPC21-CLrev-BRH06
40	MOPC315-CLrev-BRH06
40	MOPC315-CHrev-AHX89
400	MOPC21-VH-OE-5'<F3>
400	MOPC21-VL-OE-5'<F5>
400	MOPC315-VH-OE-5'<F3>
400	MOPC315-VL-OE-5'<F5>

Después de la RT-PCR en emulsión, la emulsión se rompió con n-butanol de acuerdo con los protocolos del fabricante de secuenciación génica SOLiD (Applied Biosystems), y las microesferas magnéticas se usaron como molde directo en la plataforma de secuenciación Ion Torrent (Life Technologies). La secuenciación se inició primero con el cebador de cadena pesada <F3> para obtener secuencias de ADNc de cadena pesada, seguido de la secuenciación con el cebador de cadena ligera <F5> para obtener secuencias de ADNc de cadena ligera. Las secuencias de la cadena pesada y ligera se hicieron coincidir según su localización en la plataforma de secuenciación Ion Torrent para obtener los emparejamientos naturales de cadena pesada y ligera.

15 **Ejemplo 12 - Secuenciación de transcritos con emparejamientos VH:VL de células que codifican anticuerpos de alta afinidad**

Los ejemplos previos detallaron el uso de diversas técnicas para secuenciar múltiples transcritos a partir de una variedad de poblaciones celulares. El presente ejemplo describe un procedimiento para la secuenciación de alto rendimiento de secuencias de anticuerpos de VH:VL emparejados de forma natural a partir de células individuales que codifican anticuerpos de alta afinidad específicos para un antígeno de interés particular usando la captura de poli(dT) dependiente de antígeno y la secuenciación de VH:VL posterior.

Se prepararon microesferas magnéticas recubiertas de antígeno mediante acoplamiento covalente del toxoide tetánico (TT) libre de vacunas (oligonucleótido 1 mM, TT 40 nM, Statens Serum Institut, Copenhagen, Dinamarca) a microesferas magnéticas funcionalizadas con ácido carboxílico (microesferas Dynal MyOne COOH de 1 µm de diámetro, Life Technologies) de acuerdo con los protocolos del fabricante.

Se recogieron PBMC de sangre donada 14 días después de la administración de la vacuna de refuerzo con toxoide tetánico (TT)/toxoide diftérico (TD; 20 I.E. TT y 2 I.E. toxoide diftérico, Sanofi Pasteur MSD GmbH, Leimen, Alemania) y se separaron mediante tinción con anticuerpos marcados y separación FACS, como en el Ejemplo 7. Se sembraron linfocitos B de memoria en portaobjetos de PDMS estériles como se describe en el Ejemplo 4 junto con microesferas recubiertas de antígeno (aproximadamente 40 microesferas/pocillo), y las células se sellaron dentro de los pocillos usando una membrana de diálisis y se cultivaron en el interior de los portaobjetos de micropocillos de PDMS durante cuatro días en un medio de estimulación de linfocitos B de memoria: RPMI-1640 complementado con FBS deficitario en inmunoglobulina al 10 %, 1 x GlutaMAX, 1 x aminoácidos no esenciales, 1 x piruvato de sodio y 1 x penicilina/estreptomicina (todo de Life Technologies) junto con 10 µg/ml de anticuerpo anti-CD40 (5C3, BioLegend), 500 U/ml de IL-4 y 5 ng/ml de IL-5 (PeproTech, Rocky Hill, NJ, EE. UU.). Durante este tiempo, se estimularon las células para secretar anticuerpos (Taubenheim *et al.*, 2012), y cualquier anticuerpo secretado específico para TT se unió a las microesferas magnéticas que contenían antígeno inmovilizado.

Se mezcló una solución de oligonucleótidos poli(dT)₂₅ marcados con estreptavidina en 5' (Integrated DNA Technologies, EE. UU.) en una proporción equimolar con conjugado anti-IgG-biotina humano de cabra (B1140, Sigma-Aldrich, EE. UU.). La estreptavidina y la biotina se asociaron en solución para formar anticuerpos anti-IgG con oligonucleótidos poli(dT)₂₅ anclados para la captura de ARNm. Después de cuatro días de cultivo, el sellado se

rompió y la superficie del portaobjetos se lavó suavemente tres veces con 400 µl de PBS para eliminar los anticuerpos secretados sin alterar las células y las microesferas del interior de los pocillos. Se eliminó el exceso de PBS y se añadieron 350 µl de medio RPMI-1640 que contenía conjugado anticuerpo anti-IgG/poli(dT)₂₅ 10 nM a la superficie del portaobjetos de micropocillos y el portaobjetos se incubó a temperatura ambiente durante 45 minutos. Durante los 45 minutos de incubación, cualquier microesfera marcada con antígeno que se había recubierto por anticuerpos anti-TT después de la fase de secreción de 4 días (es decir, microesferas marcadas con antígeno colocalizadas en un pocillo con una célula secretora que codificaba un anticuerpo específico para TT) se adornó con poli(dT)₂₅ para la captura de ARNm. Posteriormente, los portaobjetos se lavaron suavemente tres veces con 400 µl de PBS para eliminar el exceso de conjugado anticuerpo/oligonucleótido y los micropocillos se sellaron con una mezcla de diálisis, se lisaron las células, se recuperaron las microesferas con un imán y se realizó una RT-PCR de unión en emulsión como en el NUEVO Ejemplo 3, con la excepción de que se usó dodecilsulfato de litio al 0,1 % en el tampón de lisis celular en lugar de dodecilsulfato de litio al 1 %. Se realizó una PCR con cebadores internos y los transcritos unidos se secuenciaron usando una plataforma de secuenciación Next Generation de lectura larga, como en el NUEVO Ejemplo 5.

El procedimiento explicado en el presente procedimiento enriqueció el conjunto de secuencias para los pares VH:VL específicos de antígeno de alta afinidad, ya que solo las microesferas marcadas con antígeno con inmunoglobulina IgG unida contenían la secuencia de poli(dT)₂₅ requerida para la captura de ARNm después de la lisis celular. Por tanto, el procedimiento explicado en este ejemplo demuestra la aplicación de la técnica de emparejamiento VH:VL de alto rendimiento para la secuenciación de un gran número de pares VH:VL específicos de antígeno en un solo experimento sin necesidad de expresión de inmunoglobulina en superficie.

Ejemplo 13 - RT-PCR en células individuales emulsionadas usando una emulsión de gotitas de baja dispersión

Como en el Ejemplo 6, se formó una emulsión inyectando una corriente acuosa fuera de una boquilla en una fase oleosa anular en rápido movimiento. Las fuerzas de cizallamiento generadas por la corriente portadora indujeron la formación de gotitas acuosas con una distribución de tamaños estrechamente controlada, y el procedimiento de boquilla/corriente portadora generó emulsiones de tamaños de gotitas monodispersas que reducen la incidencia de múltiples células por gotita de emulsión que se generan cuando existe un intervalo de tamaños de gotitas. En este ejemplo se usó una mezcla de dos líneas celulares inmortalizadas (MOPC-21 y MOPC-315) para demostrar la encapsulación celular y la RT-PCR de unión directamente en gotitas de emulsión de aproximadamente 4 nl de volumen sin etapas intermedias de lisis celular o captura de ARNm.

Una mezcla equitativa de células MOPC-21 y MOPC-315 tratadas con RNasa y lavadas (como en el Ejemplo 2) se resuspendió a una concentración de 50.000 células totales/ml en PBS, mientras que se preparó otra fase acuosa que consistía en 2x mezcla concentrada de RT-PCR (Quanta OneStep Fast qRT-PCR) con BSA al 0,1 % (Invitrogen Ultrapure BSA, 50 mg/ml), inhibidor SuperAse In RNase al 4 % (Invitrogen, EE. UU.) y detergente NP-40 al 0,1 %. Se preparó un aparato de emulsión como en el Ejemplo 10. Todas las agujas y los tubos de suministro de agujas se prebloquearon en BSA al 1 % durante 30 minutos y se enjuagaron con PBS, y las células en PBS se administraron a través de la aguja interior (calibre 26) mientras que la mezcla de RT-PCR y el detergente se administraron por medio de la aguja exterior (calibre 19), siendo ambas fases acuosas de 500 µl/min. La fase portadora de aceite (aceite mineral de calidad para biología molecular con Span-80 al 4,5 %, Tween 80 al 0,4 %, Triton X-100 al 0,05 %, % en v/v, siendo los reactivos de la fase oleosa de Sigma Aldrich Corp.) fluyó a través del tubo de vidrio exterior en un velocidad de 3 ml/min y las muestras se recogieron como en el Ejemplo 10. Se emulsionó un total de 2 ml de la mezcla células/RT-PCR mezclada con 2 ml de diluyente NP-40 para aproximadamente 100.000 células analizadas. Las concentraciones de cebadores para la mezcla de RT-PCR se dan en la Tabla 1, usándose las mismas condiciones de termociclado que en el Ejemplo 11.

La emulsión de células para RT-PCR se colocó a continuación en placas de 96 pocillos y se sometió a un termociclado, se extrajo el ADNc y se realizó una PCR con cebadores internos (véase el Ejemplo 4). Los cebadores de PCR con cebadores internos se dan en la Tabla 2, y las condiciones de termociclado para la PCR fueron como sigue: una etapa de desnaturalización de 2 min a 94 °C, seguida de termociclado a 94 °C durante 30 segundos de desnaturalización, 62 °C durante 30 s de hibridación, 72 °C durante 20 s de extensión, durante 30 ciclos. El producto de PCR con cebadores internos se sometió a electroforesis para purificar el ADNc de VH-VL unido, que se usó como molde en secuenciación NextGen.

Por lo tanto, la presente invención se adapta bien para alcanzar los fines y ventajas mencionados, así como aquellos que son inherentes a la misma. Los modos de realización particulares divulgados anteriormente son solo ilustrativos, ya que la presente invención se puede modificar y practicar de maneras diferentes pero equivalentes que son evidentes para los expertos en la técnica que tienen el beneficio de las enseñanzas del presente documento. Además, no se prevén limitaciones a los detalles de construcción o diseño mostrados en el presente documento, aparte de las descritas en las reivindicaciones a continuación. Si bien las composiciones y los procedimientos se describen en términos de "que comprende", "que contiene" o "que incluye" diversos componentes o etapas, las composiciones y los procedimientos también pueden "consistir esencialmente en" o "consistir en" los diversos componentes y etapas. Todos los números e intervalos divulgados anteriormente pueden variar en cierta cantidad.

Siempre que se divulgue un intervalo numérico con un límite inferior y un límite superior, se divulga específicamente cualquier número y cualquier intervalo incluido que se encuentre dentro del intervalo. En particular, se debe entender que todos los intervalos de valores (de la forma "de aproximadamente a a aproximadamente b", o, de manera equivalente, "de aproximadamente a a b", o, de manera equivalente, "de aproximadamente a-b") divulgados en el presente documento describen todos los números e intervalos englobados dentro del intervalo más amplio de valores. Además, los términos de las reivindicaciones tienen su significado simple y habitual, a menos que el titular de la patente los defina explícita y claramente de otro modo. Además, los artículos indefinidos "un", "uno" o "una", como se usan en las reivindicaciones, se definen en el presente documento como uno o más de uno de los elementos que introducen.

Todos los procedimientos divulgados y reivindicados en el presente documento se pueden preparar y ejecutar sin experimentación excesiva a la luz de la presente divulgación.

Además, los agentes que están relacionados tanto química como fisiológicamente se pueden sustituir por los agentes descritos en el presente documento, siempre que se obtengan los mismos resultados o resultados similares. Se considera que todos los sustitutos y modificaciones similares evidentes para los expertos en la técnica están dentro del alcance y el concepto de la invención como se define por las reivindicaciones adjuntas.

REFERENCIAS

Brochet, X., Lefranc, M.-P. & Giudicelli, V. IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res.* 36, W503-W508 (2008).

Chan, M. *et al.* Evaluation of Nanofluidics Technology for High-Throughput SNP Genotyping in a Clinical Setting. *J Mol Diagn* 13, 305-312 (2011).

Citri, A. *et al.* Comprehensive qPCR profiling of gene expression in single neuronal cells. *Nature Protocols* 7, 118-127 (2012).

DeKosky, B.J. *et al.* High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nat Biotech* 31, 166-169 (2013).

Friguet, B., Chaffotte, A.F., Djavadi-Ohanian, L. & Goldberg, M.E. Measurements of the true affinity constant in solution of antigen-antibody complexes by enzyme-linked immunosorbent assay. *Journal of Immunological Methods* 77, 305-319 (1985).

Kojima, T. *et al.* PCR amplification from single DNA molecules on magnetic beads in emulsion: application for high-throughput screening of transcription factor targets. *Nucleic Acids Res.* 33 (2005).

Krause, J.C. *et al.* Epitope-Specific Human Influenza Antibody Repertoires Diversify by B Cell Intracлонаl Sequence Divergence and Interclonal Convergence. *The Journal of Immunology* 187, 3704-3711 (2011).

Kyu, S.Y. *et al.* Frequencies of human influenza-specific antibody secreting cells or plasmablasts post vaccination from fresh and frozen peripheral blood mononuclear cells. *Journal of Immunological Methods* 340, 42-47 (2009).

Mar, J.C. *et al.* Inferring steady state single-cell gene expression distributions from analysis of mesoscopic samples. *Genome Biol* 7 (2006).

Mary, P. *et al.* Analysis of gene expression at the single-cell level using microdroplet-based microfluidic technology. *Biomicrofluidics* 5 (2011).

Mazor, Y., Barnea, I., Keydar, I. & Benhar, I. Antibody internalization studied using a novel IgG binding toxin fusion. *Journal of Immunological Methods* 321, 41-59 (2007).

Mei, H.E. *et al.* Blood-borne human plasma cells in steady state are derived from mucosal immune responses. *Blood* 113, 2461-2469 (2009).

Meijer, P. *et al.* Isolation of human antibody repertoires with preservation of the natural heavy and light chain pairing. *Journal of molecular biology* 358, 764-772 (2006).

Novak, R. *et al.* Single-Cell Multiplex Gene Detection and Sequencing with Microfluidically Generated Agarose Emulsions. *Angew. Chem.-Int. Edit.* 50, 390-395 (2011).

Reddy, S.T. *et al.* Monoclonal antibodies isolated without screening by analyzing the variable-gene repertoire of plasma cells. *Nature biotechnology* 28, 965-U920 (2010).

- Sanchez-Freire, V. *et al.* Microfluidic single-cell real-time PCR for comparative analysis of gene expression patterns. *Nat. Protocols* 7, 829-838 (2012).
- 5 Smith, K. *et al.* Rapid generation of fully human monoclonal antibodies specific to a vaccinating antigen. *Nat. Protocols* 4, 372-384 (2009).
- Taubenheim, N. *et al.* High Rate of Antibody Secretion Is not Integral to Plasma Cell Differentiation as Revealed by XBP-1 Deficiency. *The Journal of Immunology* 189, 3328-3338 (2012).
- 10 Toriello, N.M. *et al.* Integrated microfluidic bioprocessor for single-cell gene expression analysis. *Proc Natl Acad Sci USA* 105, 20173-20178 (2008).
- White, A.K. *et al.* High-throughput microfluidic single-cell RT-qPCR. *Proc Natl Acad Sci USA* (2011).
- 15 Wrammert, J. *et al.* Rapid cloning of high-affinity human monoclonal antibodies against influenza virus. *Nature* 453, 667-671 (2008).
- Wu, X. *et al.* Focused Evolution of HIV-1 Neutralizing Antibodies Revealed by Structures and Deep Sequencing. *Science* 333, 1593-1602 (2011).

REIVINDICACIONES

1. Un procedimiento que comprende:
- 5 a) separar células individuales en un compartimento con microesferas conjugadas con oligonucleótidos que comprenden poli(T);
- b) lisar las células;
- 10 c) permitir que los transcritos de ARNm liberados de las células se hibriden con oligonucleótidos conjugados con las microesferas;
- d) agrupar, lavar y resuspender las microesferas en solución con reactivos para retrotranscripción asociada a reacción en cadena de la polimerasa con extensión de solapamiento, que incluye cebadores diseñados para crear un único producto de PCR que comprende ADNc de al menos dos transcritos de interés unidos covalentemente;
- 15 e) la suspensión de microesferas y solución de reactivos se emulsiona en la fase oleosa para crear gotitas con no más de una microesfera por gotita;
- 20 f) realizar una retrotranscripción asociada a reacción en cadena de la polimerasa con extensión de solapamiento para formar una molécula de ADN individual que comprende al menos dos ADNc de los al menos dos transcritos derivados de una única célula, en el que los al menos dos ADNc comprenden secuencias de TCR emparejadas o secuencias de VH y VL emparejadas; y
- 25 g) secuenciar el ADNc asociado con las microesferas.
2. El procedimiento de la reivindicación 1, en el que la microesfera es magnética.
3. El procedimiento de la reivindicación 1, en el que los compartimentos individuales son micropocillos.
- 30 4. El procedimiento de la reivindicación 1, en el que los compartimentos individuales son microvesículas en una emulsión.
5. El procedimiento de la reivindicación 3, en el que los pocillos se sellan con una membrana permeable antes de la etapa (d).
- 35 6. El procedimiento de la reivindicación 1, que comprende obtener secuencias de al menos 10.000 células individuales.
- 40 7. El procedimiento de la reivindicación 1, en el que la etapa (g) comprende la secuenciación de dos o más transcritos unidos covalentemente a la misma microesfera.
8. El procedimiento de la reivindicación 1, en el que las células se seleccionan del grupo que consiste en: linfocitos B, linfocitos T y células NKT.

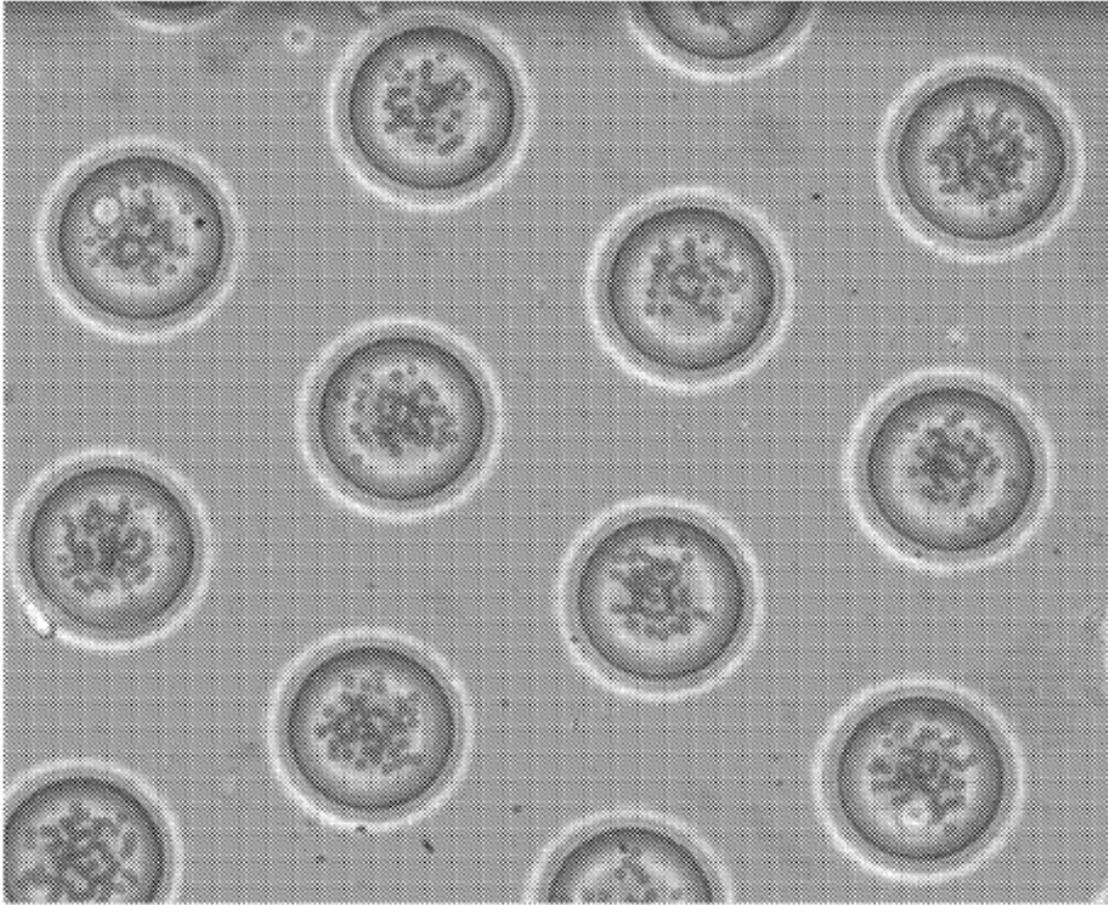


FIG. 1

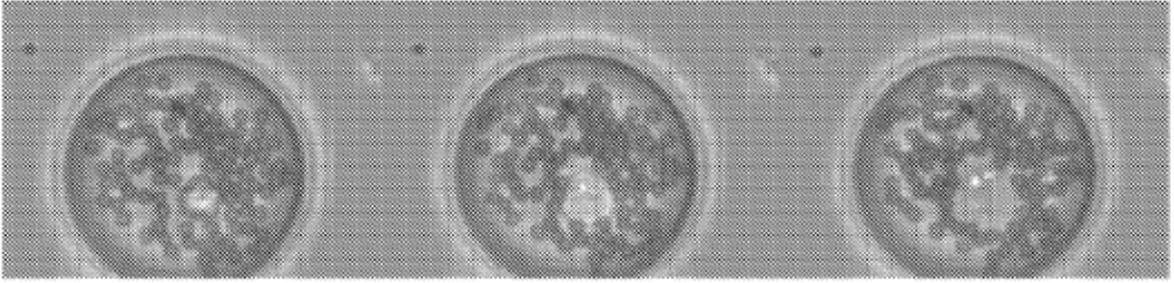


FIG. 2

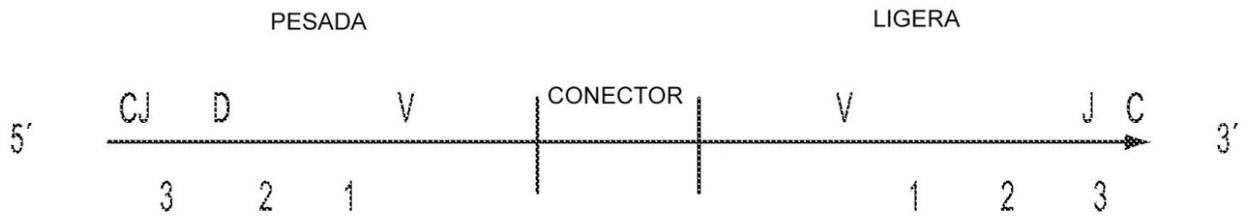


FIG. 3

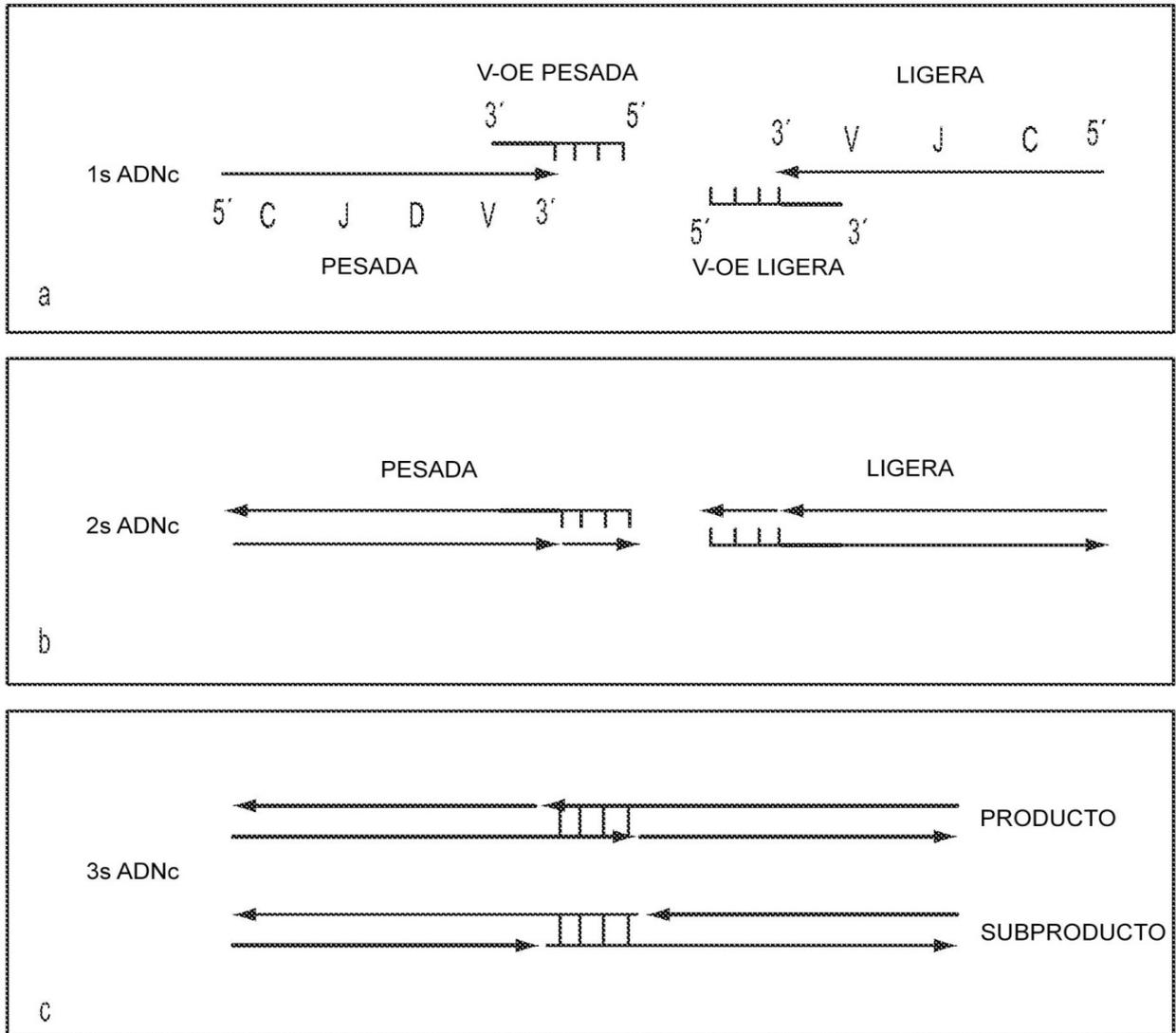


FIG. 4

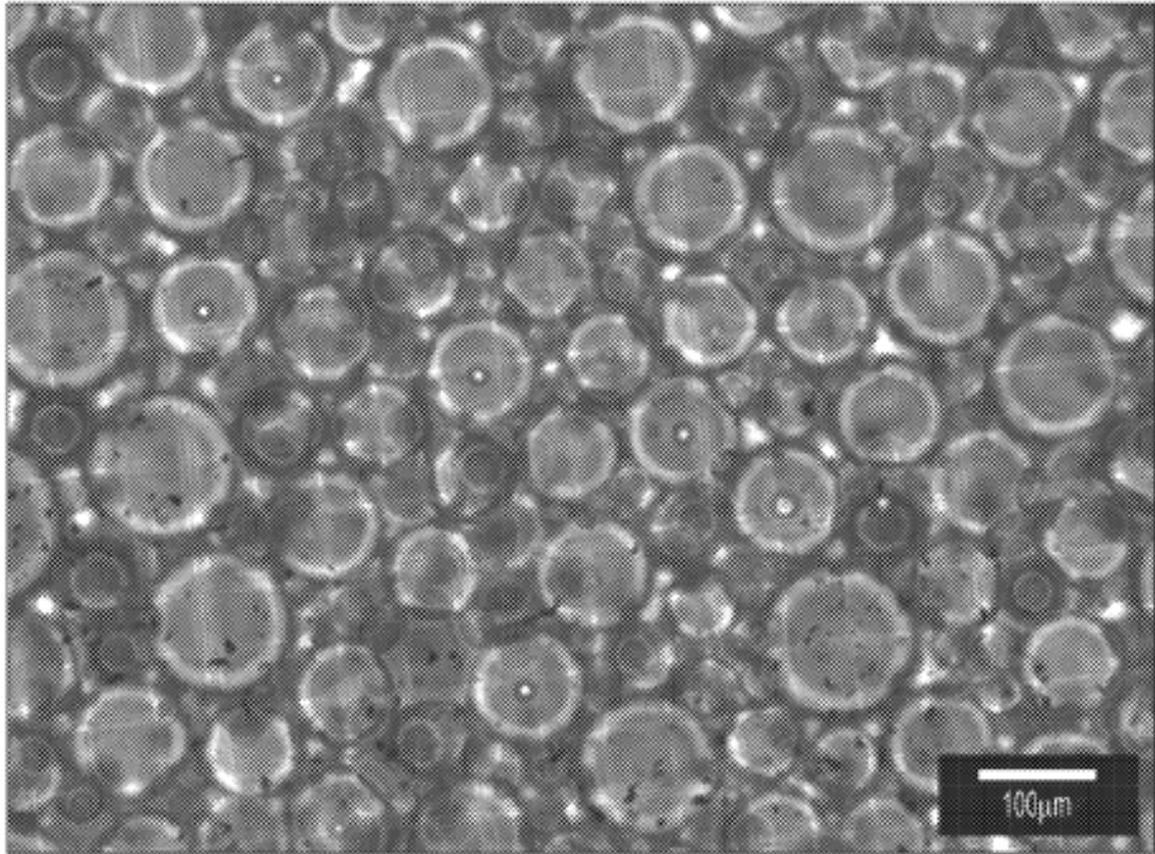


FIG. 5

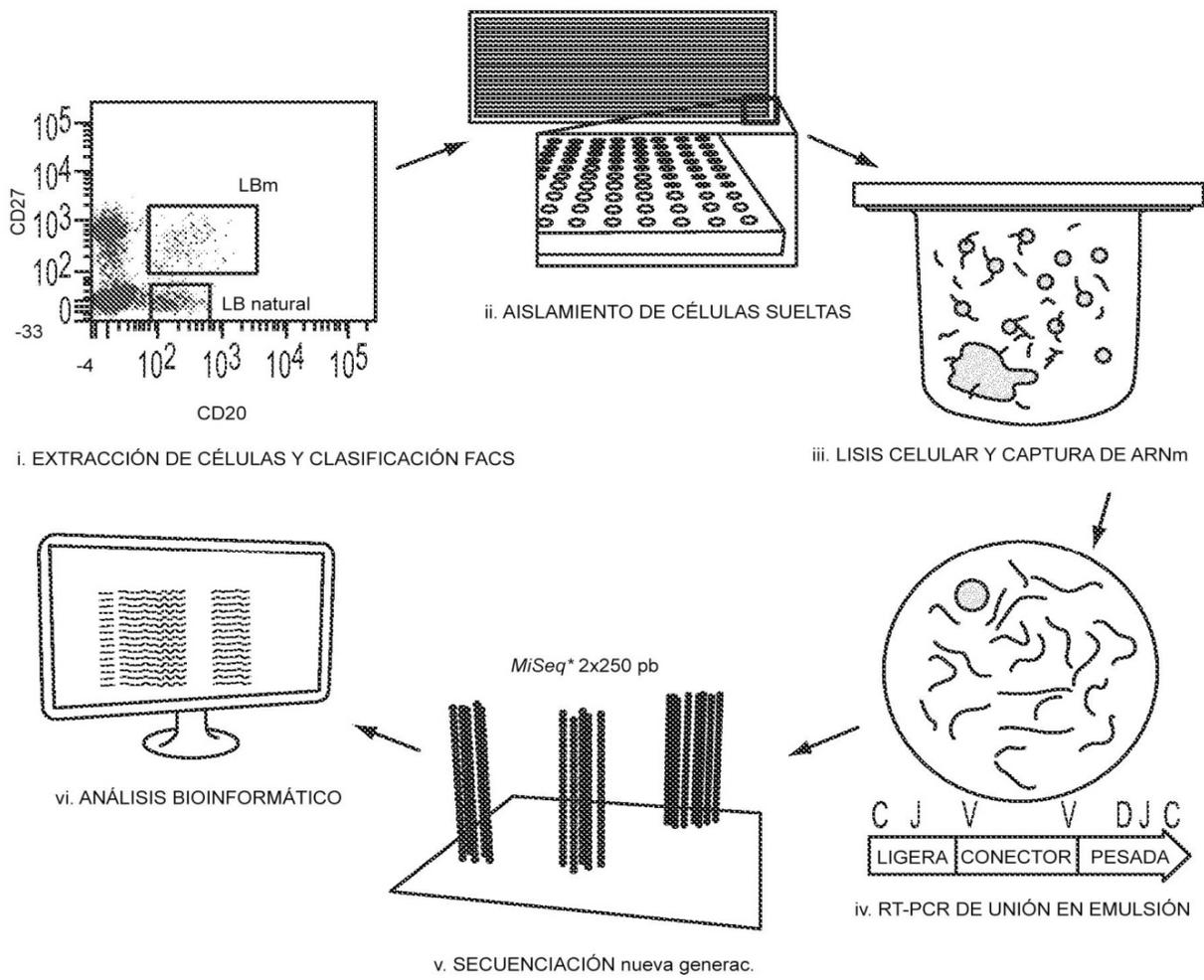


FIG. 6

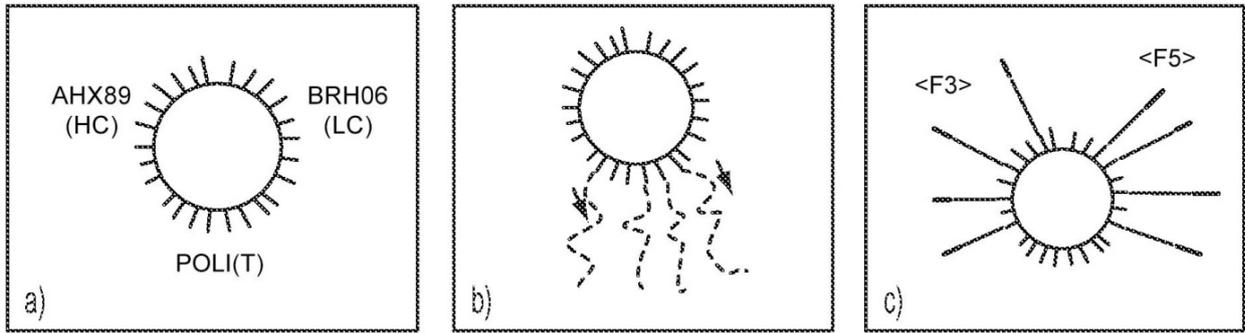


FIG. 7

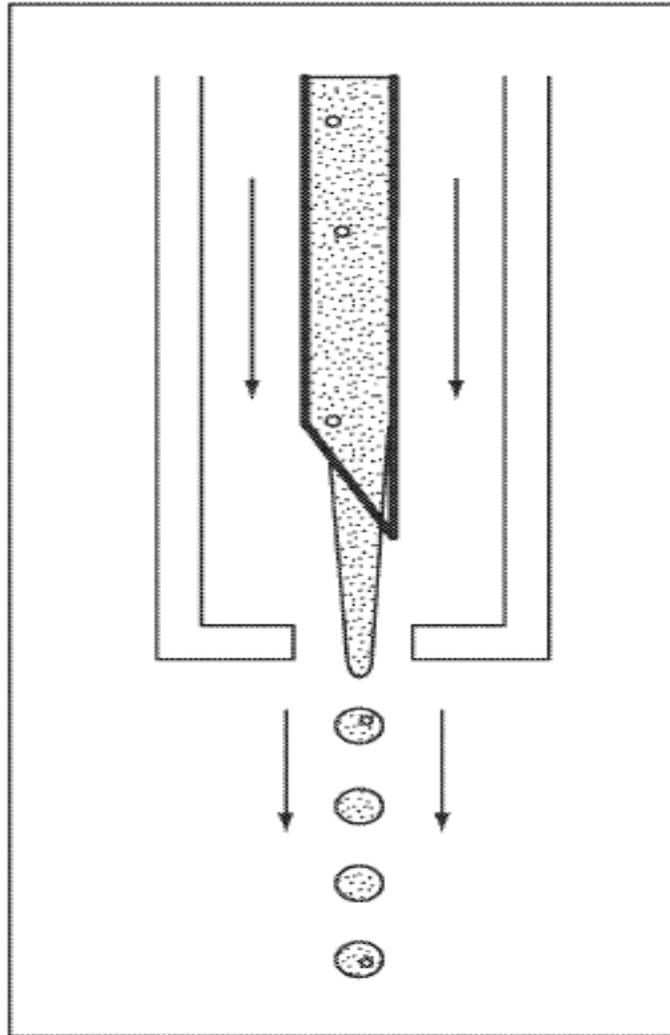


FIG. 8

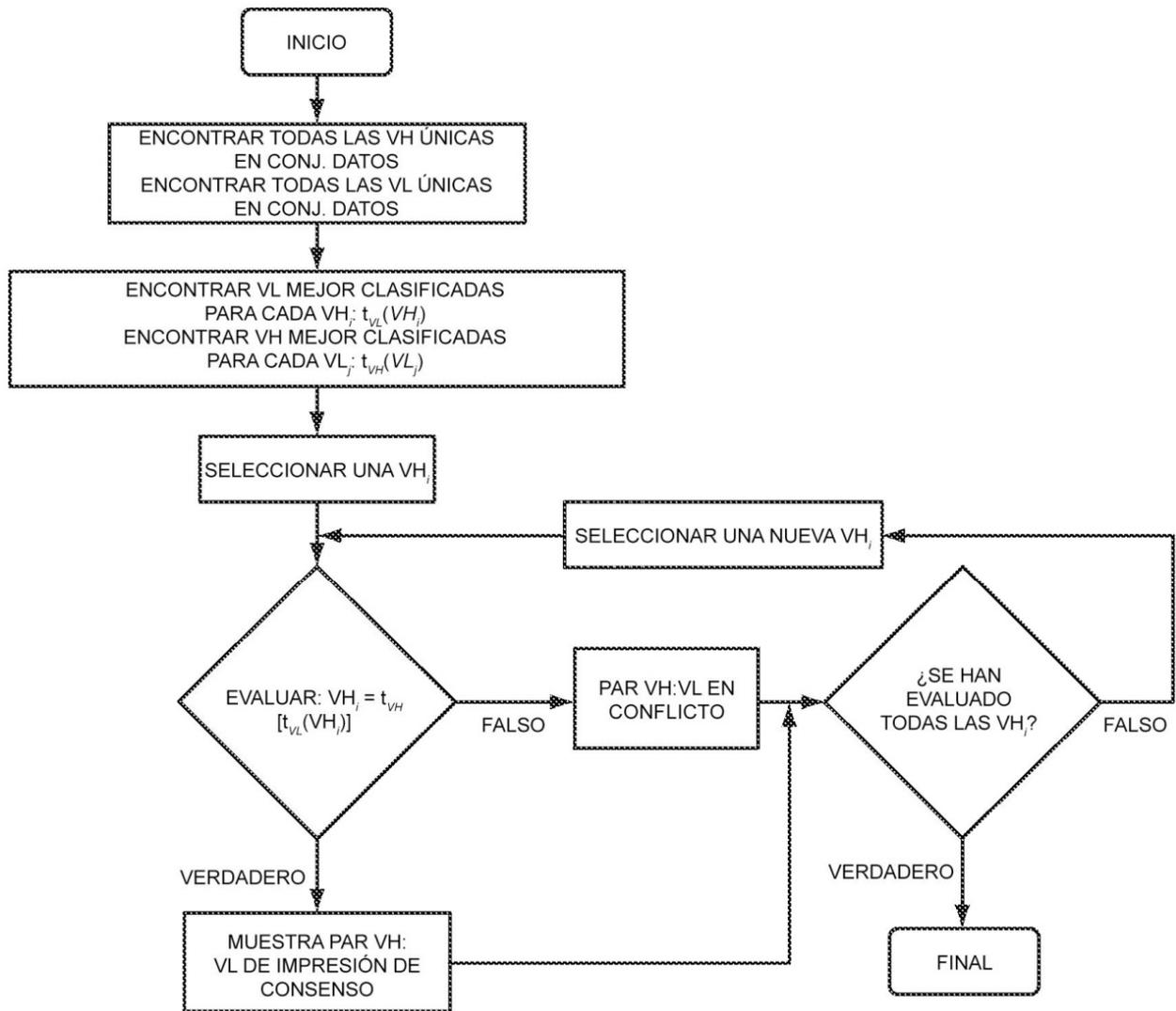


FIG. 9

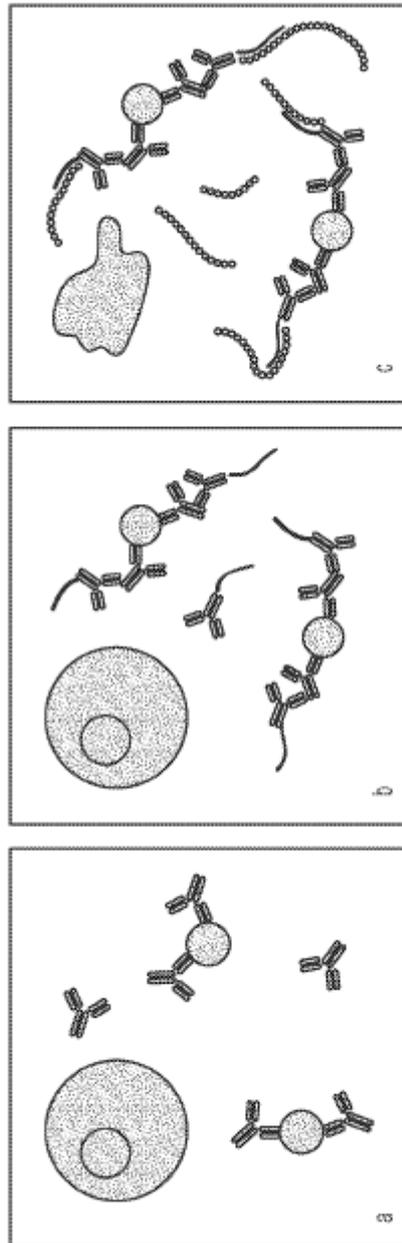


FIG. 10