

MÉTODO DE ANÁLISIS DE SANGRE ALMACENADA PARA EL ANÁLISIS POSTERIOR DE ADN LIBRE DE CÉLULAS

1. CAMPO DE LA INVENCION

[0002] La presente invención se refiere al almacenamiento y prueba para detectar anomalías en ADN libre de células (ADNcf) del plasma, tal como plasma en sangre materna o en sangre de un paciente con cáncer o en un paciente que está siendo examinado para detectar cáncer.

2. ANTECEDENTES DE LA INVENCION

[0003] Desde su introducción en la atención clínica en 2011 (Ver Ehrich, 2011 y Palomaki, 2011), la prueba prenatal no invasiva (NIPT, por sus siglas en inglés) ha proporcionado a muchas mujeres embarazadas en diferentes países información sobre su riesgo de aneuploidías fetales. Cualquiera que sea la plataforma utilizada o el protocolo seguido, la mayoría de los laboratorios informarán una tasa de fracaso de entre 0 y 25 %. (Badeau, 2017.) Los fracasos suelen deberse a una fracción fetal demasiado baja para ser confiable. (Blais, 2018.) Se conocen algunas razones para una fracción fetal baja, como la edad gestacional temprana, un IMC (índice de masa corporal) alto, presencia de alguna anomalía fetal o placentaria, un T18 y la edad materna. (Hou, 2019; Ashoor, 2013.)

[0004] El aumento de la edad gestacional se asocia con un aumento de la fracción fetal; sin embargo, las pruebas de detección son útiles al comienzo del embarazo y el aumento de la fracción fetal se observa principalmente al final del embarazo. (Wang, 2013.) A nivel molecular, el ADNcf fetal se puede distinguir del ADNcf materno por el hecho de que los fragmentos son más cortos (Fan, 2010; Shi, 2020), hipermetilados y enriquecidos en ubicaciones genómicas específicas (Wang, 2013). Recientemente, varios estudios informaron que el uso de electroforesis en gel de agarosa de alta concentración para seleccionar los fragmentos más cortos de una biblioteca preparada a partir de ADNcf antes de la secuenciación aumentó significativamente la fracción fetal (Welker, 2021; Qiao, 2019; Xue, 2020; Qiao, 2019; Hu, 2019). El mismo procedimiento pareció mejorar también la detección de alelos mutantes raros en el contexto del cáncer. (Underhill, 2021.)

[0005] También se demostró que el ADN tumoral circulante libre de células es más corto que el ADNcf de células sanas. (Jiang, 2015.) Para ambas aplicaciones, se utilizaron diferentes sistemas para aislar o enriquecer estos fragmentos cortos con electroforesis utilizando geles prefabricados como E-gel EX al 2% de Invitrogen (Thermofisher), casetes de gel de agarosa al 2-3% de

BluePippin (Sage Bioscience) y casetes de gel de agarosa al 3 % de Yourgene Health Canada, Inc. (Underhill, 2021).

[0006] Para obtener una proporción adecuada de ADNcf para una secuenciación significativa, el plasma debe prepararse con cierta precaución para evitar la contaminación con el ADN genómico del huésped. Las muestras de sangre deben procesarse entre 4 y 8 horas después de su recolección y algunos prefieren un paso adicional de centrifugación a alta velocidad para la preparación del plasma (Barrett, 2011.) Sin embargo, esto no siempre es posible porque los sitios de recolección no siempre están equipados para llevar a cabo una centrifugación de alta velocidad de muestras de plasma antes de congelarlas para enviarlas a un laboratorio de pruebas central. La solución adoptada por la mayoría de los sitios de flebotomía es utilizar tubos especializados diseñados para mantener la integridad del ADNcf, por ejemplo, mediante el uso de agentes reticulantes para estabilizar la membrana celular. Un ejemplo de dichos tubos especializados son los tubos de extracción de sangre Streck®. No utilizar un tubo especializado y almacenar sangre durante un período prolongado provoca un aumento del ADNcf con el tiempo debido a la descomposición del material genómico, (Fernando 2018.) Los tubos especializados estabilizan la muestra de sangre, lo que da tiempo para que la muestra se transfiera a un laboratorio con el equipo para procesarla. Sin embargo, estos tubos especializados son costosos.

[0007] Recientemente se informó que el uso de tubos Vacutainer K₂EDTA con gel para recolectar plasma para diagnóstico molecular podría producir ADNcf adecuado para la prueba prenatal no invasiva (Giroux, 2021). Se demostró que el uso de estos tubos seguido de filtración era rentable y eficiente para reducir pasos que requieren tiempo en el laboratorio. Sin embargo, a pesar de que la mayoría de los sitios pueden centrifugar un tubo de sangre, muchos sitios no tienen equipos ni personal para abrir el tubo de sangre y filtrar el plasma.

[0008] Por lo tanto, existe una necesidad de alternativas para almacenar muestras de sangre completa y procesar sangre para un análisis de secuencia significativo de ADNcf.

3. RESUMEN DE LA INVENCION

[0009] Las realizaciones de la presente divulgación se refieren a un método para analizar secuencias de ácido nucleico obtenidas a partir de una muestra de sangre para proporcionar información de secuencia que comprende: recolectar una muestra de sangre en un dispositivo de recolección de sangre, en donde dicho dispositivo de recolección de sangre no comprende un

fijador; almacenar la sangre sin fijador antes de aislar el ADN libre de células del plasma a una temperatura mayor que -20°C y menor que 35°C (tal como temperatura ambiente o temperaturas de refrigeración); separar el plasma de las células sanguíneas presentes en la muestra de sangre antes o después del almacenamiento; aislar ADN libre de células del plasma más de 24 horas después de la recolección de sangre de un sujeto; separar el ADN libre de células por tamaño del ADN libre de células y aislar el ADN libre de células que tiene menos de 300 pb; y analizar secuencias de ácido nucleico del ADN libre de células aislado para detectar información de secuencia. La sangre se puede almacenar durante un período prolongado sin fijador antes de aislar el ADNcf del plasma. Este período de tiempo puede ser de hasta 3, 4, 5, 6, 7, 8 o 9 días después de la recolección de sangre.

[0010] Un dispositivo de recolección adecuado para su uso con la invención puede comprender un recipiente que tiene solo una o más composiciones en el mismo, en el que la una o más composiciones consisten en uno o más anticoagulantes y un ión metálico seleccionado entre potasio, litio y sodio. Un ejemplo de tales dispositivos es un tubo K_2EDTA . Alternativamente, el dispositivo de recolección de sangre puede comprender un recipiente que tiene solo una o más composiciones en su interior, en el que la una o más composiciones consisten en uno o más anticoagulantes y un ión metálico seleccionado entre potasio, litio y sodio; y un gel separador. En otras realizaciones, el dispositivo de recolección puede comprender un recipiente que tiene solo una o más composiciones en el mismo, en donde la una o más composiciones comprenden o consisten en un coagulante.

[0011] Como se mencionó anteriormente, la etapa de separación se puede utilizar para aislar ADNcf que tiene menos de 300 pb en la muestra de ADNcf aislado. En algunas realizaciones, el ADNcf que se aísla en la etapa de separación por tamaños tiene 185 pb o menos o 165 pb o menos. En otras realizaciones, el ADNcf que se aísla en la etapa de separación por tamaños tiene 155 pb o menos, o 150 pb o menos. En realizaciones adicionales, la porción del ADNcf aislado de la muestra de plasma tiene 50 pb o más o 80 pb o más. La etapa de separación se puede realizar mediante cualquier número de técnicas adecuadas de selección de tamaño de polinucleótidos conocidas en el campo, incluyendo electroforesis en gel, cromatografía, separación basada en perlas o un filtro de membrana con un tamaño de poro que impida el paso de polinucleótidos de tamaño mayor que el umbral deseado.

[0012] Después de la selección de tamaño, las secuencias de ácido nucleico del ADN libre de células aislado pueden secuenciarse y analizarse para determinar la presencia de una anomalía genética en el ADN fetal o para determinar información de secuencia acerca de o confirmar la presencia de células cancerosas en un sujeto con base en la presencia de uno más marcadores genéticos asociados con tumorigenicidad o malignidad. Para obtener información de secuencia sobre el ADNcf aislado, es posible que sea necesario preparar una biblioteca de secuencias y amplificar el ADNcf. Una vez que se completa la selección de tamaño y la preparación de la biblioteca, la muestra se puede analizar para obtener información de secuencia. La información de secuencia se puede analizar para detectar mutaciones y anomalías cromosómicas que incluyen, entre otras, translocación, transversión, monosomía, trisomía y otras aneuploidías, delección, inserción, metilación, amplificación, fragmento, translocación y reordenamiento o cromatrispsis. La información de secuencia también se puede utilizar para detectar cualquier alteración en la secuencia genética en comparación con una secuencia de referencia.

[0013] Otro aspecto de esta divulgación es una prueba de validez de secuencia que puede detectar un desempeño anormal del paso de selección de tamaño o una degradación inaceptable de la muestra. En un método que implica la selección de tamaño de una muestra de ADN que comprende ADNcf y la secuenciación del ADN de la muestra después de la selección de tamaño; el método puede comprender una prueba de validez de secuencia de la información de secuencia obtenida del paso de secuenciación. La prueba de validez comprende crear un primer perfil de tamaño de fragmento normalizado a partir de la información de secuenciación asociada con la muestra, comparar uno o más valores dentro del perfil de tamaño de fragmento con uno o más de los rangos de valores correspondientes obtenidos de un conjunto de parámetros de referencia que comprende una pluralidad de perfiles de tamaño de fragmentos válidos; y aceptando la información de secuencia y los resultados del análisis de la misma si uno de los uno o más valores dentro del perfil de tamaño de fragmento asociado con la muestra está dentro del rango de valores correspondiente o rechazando la información de secuencia si uno de los uno o más valores dentro del perfil de tamaño de fragmento asociado con la muestra está fuera del rango de valores correspondiente. La prueba de validez también puede comprender calcular un primer valor de frecuencia de tamaño de fragmento relativo a partir de la información de secuenciación asociada con la muestra; comparar el primer valor de frecuencia de tamaño de fragmento relativo con un conjunto de valores de frecuencia de tamaño de fragmento relativo de referencia que comprende

una pluralidad de frecuencias de tamaño de fragmento relativo válidas, y aceptando la información de secuencia y los resultados del análisis de la misma si la primera frecuencia de tamaño de fragmento asociada con la muestra está dentro de un rango especificado de valores dentro del conjunto de referencia o rechazando la información de secuencia y los resultados del análisis de la misma si la primera frecuencia de tamaño de fragmento asociada con la muestra está fuera de un rango específico de valores dentro del conjunto de referencia.

4. BREVE DESCRIPCIÓN DE LAS FIGURAS

[0014] La **FIG. 1** representa un gráfico que muestra la fracción fetal de varias muestras en varios puntos de tiempo para muestras almacenadas en un tubo de preparación de plasma o un tubo Streck® y almacenadas durante 2 a 9 días.

[0015] La **FIG. 2A** representa un gráfico que muestra la fracción fetal calculada a partir de la reducción de fragmentos del cromosoma X en el embarazo masculino. La **FIG. 2B** representa la fracción fetal evaluada con el número de fragmentos del cromosoma Y solo para machos.

[0016] La **FIG. 3** representa un gráfico que muestra el tamaño de los fragmentos secuenciados calculados a partir del mapeo de secuenciación de extremos pares. En fragmentos grises antes de la selección de tamaño y en azul después de la selección de tamaño.

[0017] La **FIG. 4** representa un esquema del estudio descrito en el Ejemplo 3. Al momento de la inscripción, cada participante donó 20 ml de sangre. La sangre se recolectó en contenedores Vacutainer con EDTA de 5 ml y se transportó para su almacenamiento y procesamiento a un laboratorio de destino. La sangre se almacenó dentro de sus Vacutainers a temperatura ambiente durante 0 a 8 horas, 72 horas, 120 horas o 168 horas antes de procesarla mediante doble centrifugación para obtener plasma con EDTA. El plasma se almacenó a -80 °C y se procesó a intervalos una vez que se habían recolectado suficientes muestras para una serie.

[0018] La **FIG. 5** representa un ANOVA unidireccional de la estimación de la fracción fetal sin procesar para las 268 muestras válidas después de la secuenciación, mostrando los datos de estimación de la fracción fetal por punto de tiempo, donde la media para todos los puntos de tiempo se muestra mediante la línea horizontal, el verde representa el ANOVA y el azul representa $\pm$ desviación estándar, y el análisis de media y desviación estándar del ANOVA. Punto de tiempo 1 = 0-8 horas, Punto de tiempo 2 = 72 horas, Punto de tiempo 3 = 120 horas y Punto de tiempo 4 = 168 horas.

[0019] FIG. 6. Medias, desviaciones estándar y ANOVA de Welch de las 275 bibliotecas válidas para este estudio, más la muestra M056 (Punto de tiempo 1; 2,49 ng/μL) de la Serie 5. Los datos de cuantificación de la biblioteca posterior a la PCR (ng/μL) se presentan por Punto de tiempo. Nota: las cuatro pruebas de varianzas iguales demuestran que las varianzas no son iguales para los datos en todos los puntos de tiempo; por lo tanto, se realizó un ANOVA de Welch en lugar de un ANOVA. Punto de tiempo 1 = 0-8 horas, Punto de tiempo 2 = 72 horas, Punto de tiempo 3 = 120 horas y Punto de tiempo 4 = 168 horas. Un valor $P < 0,0001$ en ANOVA de Welch muestra una variación significativa en los datos de concentración de la biblioteca en los cuatro Puntos de tiempo del estudio ($P < 0,05$).

[0020] FIG. 7A. Distribución del perfil de tamaño y frecuencia para cada Punto de tiempo para la muestra de paciente M047 donde: Punto de tiempo 1 = Rojo/TP1 (YGL013-6), Punto de tiempo 2 = Naranja/TP2 (YGL015-6), Punto de tiempo 3 = Azul/TP3 (YGL038 -6) y punto de tiempo 4 = Verde/TP4 (YGL012-6).

[0021] FIG. 7B. Distribución del perfil de tamaño y frecuencia para cada Punto de tiempo para la muestra de paciente M100 donde: Punto de tiempo 1 = Rojo/TP1 (YGL026R-12), Punto de tiempo 2 = Naranja/TP2 (YGL003-12), Punto de tiempo 3 = Azul/TP3 (YGL021 -12) y Punto de tiempo 4 = Verde/TP4 (YGL012-12).

5. DESCRIPCIÓN DETALLADA

[0022] En el presente documento se describe un método para almacenar sangre que contiene ADNcf a secuenciar. De acuerdo con la presente divulgación, las muestras de sangre completa se almacenan en un dispositivo de recolección de sangre que no comprende un fijador durante un período de tiempo prolongado, superior a 24 horas, antes de que se aísle el ADNcf de la sangre o plasma para obtener información de secuencia. El método comprende realizar un paso de selección de tamaño del ADN aislado para seleccionar ADNcf que sea menor que un cierto tamaño, por ejemplo, menor o igual a 185 pb, antes de obtener información de secuencia del ADNcf seleccionado por tamaño. Al seleccionar ADNcf que tiene una longitud de nucleótidos menor o igual a, por ejemplo, 185 pb, se aumenta la fracción de ADNcf fetal que está en la muestra de ADNcf, aumentando así la calidad de la muestra para un análisis de secuencia significativo.

5.1. Definiciones

[0023] “Muestra de sangre” en el presente documento se refiere a una muestra de sangre completa que no ha sido fraccionada ni separada en sus partes componentes.

[0024] "Fijación" es una técnica que ayuda a mantener la estructura de las células y/o componentes subcelulares tales como orgánulos celulares (por ejemplo, núcleo) en una muestra de sangre. La fijación modifica la estructura química o biológica de los componentes celulares, por ejemplo, reticulándolos. La fijación impide la lisis de células enteras y orgánulos celulares, inhibiendo así la liberación de ácidos nucleicos celulares al medio circundante. Por ejemplo, la fijación puede inhibir la liberación del ADN nuclear de los glóbulos blancos en una fracción de plasma durante la centrifugación de la sangre completa.

[0025] "Fijador" es un agente utilizado para fijar una muestra de sangre.

[0026] “Secuencia de interés” se refiere a una secuencia de ácido nucleico que está asociada con una diferencia en la representación de secuencia en individuos sanos versus enfermos. Una secuencia de interés puede ser una secuencia de un cromosoma que está mal representada, es decir, sobrerrepresentada o subrepresentada, en una enfermedad o condición genética. Una secuencia de interés puede ser una porción de un cromosoma, es decir, un segmento cromosómico o un cromosoma completo. Por ejemplo, una secuencia de interés puede ser un cromosoma que está sobrerrepresentado en una condición de aneuploidía, o un gen que codifica un supresor de tumores que está subrepresentado en un cáncer. Las secuencias de interés incluyen secuencias que están sobrerrepresentadas o subrepresentadas en la población total, o una subpoblación de células, de un sujeto. También puede denominarse "marcador" de un genotipo genético asociado con un fenotipo particular, como una enfermedad.

[0027] "Etiqueta de secuencia" se refiere a una secuencia de al menos 30 pb leída de una cadena de ADNcf. En algunas realizaciones, la etiqueta se ha asignado específicamente, es decir, se ha mapeado, a una secuencia más grande, por ejemplo, un genoma de referencia, mediante alineación.

5.2. Método para aumentar la fracción de una porción de ADNcf y analizar la misma

[0028] Las realizaciones de la presente divulgación se dirigen a un método para aumentar la fracción de una porción de ADNcf de plasma, tal como de una muestra de sangre materna o una muestra de sangre para un diagnóstico de cáncer, y analizar las secuencias de ácido nucleico de la fracción para proporcionar información de secuencia. El método comprende recolectar una

muestra de sangre en un dispositivo de recolección de sangre. El dispositivo de recolección de sangre no comprende un fijador. En algunas realizaciones, el dispositivo de recolección de sangre consta de un recipiente con uno o más anticoagulantes y un ión metálico seleccionado entre potasio, sodio y litio. En algunas realizaciones, el dispositivo de recolección de sangre consta de un recipiente con uno o más anticoagulantes y un ión metálico, tal como potasio, sodio o litio, y un gel separador.

[0029] La muestra de sangre se centrifuga antes o después del almacenamiento sin fijador para separar el plasma de las células sanguíneas. Antes o después de que se separe el plasma, la muestra de sangre se puede almacenar a una temperatura en un rango de -20°C a 35°C, tal como a temperatura ambiente, a temperaturas de refrigeración, o entre temperatura ambiente y temperaturas de refrigeración hasta que la muestra se pueda seguir procesando. Cuando la muestra se puede procesar entre 24 y 216 horas, por ejemplo, 48/72 a 216 horas o 48/72 a 192 horas o 48/72 a 168 horas o 48/72 a 144 horas después de la recolección de sangre, el ADNcf es aislado del plasma. Para aumentar la fracción de una porción del ADNcf, el ADNcf se separa por tamaño, tal como mediante electroforesis en gel, y se aísla el ADNcf que tiene menos de 300 pb (por ejemplo, menos de 185, 165, 160, 155 o 150 pb). Ya sea antes o después de la selección del tamaño, la muestra de ADNcf se puede preparar para la secuenciación opcionalmente amplificando y generando una biblioteca de secuencias que comprende el ADNcf. Luego se puede analizar el tamaño de ADNcf seleccionado y aislado para detectar información de secuencia. Los pasos descritos en este documento pueden ser realizados por una entidad o varias entidades.

5.2.1 Dispositivo de recolección de sangre

[0030] Se recolecta una muestra de sangre en un dispositivo de recolección. El dispositivo de recolección puede ser un recipiente de recolección al vacío, normalmente un tubo. Según la presente divulgación, el dispositivo de recolección de sangre no contiene fijador. Ejemplos de dispositivos de recolección de sangre adecuados incluyen tubos de recolección de sangre con EDTA K2 o K3 con o sin gel separador.

[0031] Como se definió anteriormente, un fijador es un agente utilizado para fijar una muestra de sangre. Los ejemplos de fijadores incluyen aldehídos (p. ej., formaldehído, glutaraldehído y derivados de los mismos), alcoholes de formalina, reticulantes reactivos con sulfhidrilo, reticulantes reactivos con carbohidrato, reticulantes reactivos con carboxilo, reticulantes

fotorreactivos, reticulantes escindibles, AEDP (ácido 3-[(2- aminoetil)ditio]propiónico·HCl) , APG (monohidrato de p-azidofenil glioxal), BM(PEO)3 (1,8-bis-MaleiMido-PEO3), BM(PEO)4 (1,8-bis-MaleiMido-PEO4), BMB (1,4-bismaleimidobutano), BMDB (1,4-Bis-Maleimidil-2,3-dihidroxibutano), BMH (Bis-Maleimidohexano), BMOE (Bis-Maleimidoetano) , BS3 (bis(sulfosuccinimidil)suberato), BSOCOES (bis[2-(succinimidoxycarbonilo)etil]sulfona), DFDNB (1-5-difluoro-2,4-dinitrobenzoceno), DMA (adipimidato de dimetilo·2HCl), DMP(dimetil pimelimidato·2HCl), DMS (suberimidato de dimetilo·2HCl), DPDPB(1,4-Di-(3'-[2'piridilditio]propionamido)butano), DSG (Glutarato de disuccinimidilo), DSP ditiobis(sulfosuccinimidilpropionato) (Reactivo de Lomant), DSS (suberato de disuccinimidilo), DST (tartrato de disuccinimidilo), DTBP (Dimetil 3,3'-ditiobispropionimidato·2HC), DTM(Ditiobis-maleimidoetano), DTSSP (3,3'-ditiobis(sulfosuccinimidil) propionato), EGS (etilen glicolbis(succinimidilsuccinato), HBVS (1,6-hexano-bis-vinilsulfona), sulfo-BSOCOES, sulfo-DST y sulfo-EGS.

[0032] Otros fijadores pueden ser un estabilizador de la membrana celular tal como DMAE (dimetilaminoetanol), colesterol, derivados del colesterol, altas concentraciones de magnesio, vitamina E y derivados de la vitamina E, calcio, gluconato de calcio, taurina, niacina, derivados de hidroxilamina, bimoclomol, sacarosa, astaxantina, glucosa, amitriptilina, isómero A fenilacetato de hopano tetral, isómero B fenilacetato de hopano tetral, citicolina, inositol, vitamina B, complejo de vitamina B, hemisuccinato de colesterol, sorbitol, calcio, coenzima Q, ubiquinona, vitamina K, complejo de vitamina K, menaquinona, Zonegran, zinc, extracto de ginkgo biloba, difenilhidantoína, Perftoran, polivinilpirrolidona, fosfatidilserina, Tegretol, PABA, cromoglicato disódico, nedocromil sódico, feniloína, citrato de zinc, Mexitil, Dilantin, hialuronato de sodio o polaxámero 188.

[0033] Otros fijadores son agentes liberadores de formaldehído tales como uno seleccionado del grupo que consiste en: diazolidinil urea, imidazolidinil urea, dimetol-5,5-dimetilhidantoína, dimetilol urea, 2-bromo-2-nitropropano-1, 3-diol, oxazolidinas, hidroximetilglicinato de sodio, 5-hidroximetoximetil-1-laza-3,7-dioxabicyclo[3.3.0]octano, 5-hidroximetil-1-laza-3,7dioxabicyclo[3.3.0]octano, 5-hidroxipoli[metileno]metil-1-laza-3,7dioxabicyclo[3.3.0]octano, adamantina cuaternaria y cualquier combinación de los mismos.

[0034] En otra realización, además de EDTA y/o heparina, el dispositivo de recolección de sangre puede no contener agentes que reduzcan la degradación del ADN, como un inhibidor de

nucleasas. Los ejemplos de inhibidores de nucleasas incluyen pirocarbonato de dietilo, etanol, ácido aurintricarboxílico (ATA), formamida, complejos de vanadilo-ribonucleósido, macaloide, proteinasa K, hidroxilamina-oxígeno-ion cúprico, bentonita, sulfato de amonio, ditioneitol (DTT), beta-mercaptoetanol, cisteína, ditioeritritol, clorhidrato de tris(2-carboxietil)fosfeno y un catión divalente como Mg^{+2} , Mn^{+2} , Zn^{+2} , Fe^{+2} , Ca^{+2} , Cu^{+2} . También se pueden excluir del dispositivo otros agentes que reducen la degradación del ADN, como el cloruro de zinc, el clorhidrato de guanidina, el isotiocianato de guanidina, la N-lauroilsarcosina y el dodecilsulfato de Na.

[0035] La presente divulgación contempla la exclusión de cualquier combinación específica de los fijadores anteriores o cualquier combinación específica de los fijadores y/o agentes anteriores que reducen la degradación del ADN, tales como inhibidores de nucleasas.

[0036] En algunas realizaciones, el dispositivo de recolección contiene un anticoagulante. Un anticoagulante puede seleccionarse entre ácido etilendiaminotetraacético (EDTA), heparina, citrato, oxalato y cualquier combinación de los mismos. En algunas realizaciones, el anticoagulante puede secarse por aspersión sobre la pared del dispositivo de recolección. En otras realizaciones, el anticoagulante es una solución líquida en el dispositivo. Los disolventes adecuados incluyen agua, solución salina, dimetilsulfóxido, alcohol y mezclas de los mismos. En realizaciones, el dispositivo de recolección comprende un recipiente que tiene solo una o más composiciones, donde la una o más composiciones consisten en un anticoagulante y un ion metálico, como los seleccionados entre metales alcalinos o alcalinotérreos. En particular, el ion metálico puede seleccionarse entre calcio, sodio, litio y potasio.

[0037] En realizaciones con EDTA, el EDTA puede estar quelado con un ion metálico seleccionado entre potasio, litio y sodio, y normalmente es potasio. En algunas realizaciones, el dispositivo de recolección de sangre comprende un recipiente que tiene solo una o más composiciones, donde la una o más composiciones consisten en EDTA quelado.

[0038] En algunas realizaciones, el dispositivo de recolección contiene un gel separador. Dichos geles se utilizan comúnmente en tubos de separación de suero (SST, por sus siglas en inglés) para la preparación de plasma (véase, por ejemplo, BD Vacutainer® SST™ y PST™, Becton, Dickinson, and Company). El gel tendrá una densidad intermedia entre la de las células en la sangre y la fase líquida de la sangre (densidades de 1,09 y 1,03 g/cm³ respectivamente). En virtud de su densidad, el gel separador separa físicamente el componente líquido (plasma) de la

sangre de los eritrocitos y glóbulos blancos después de la centrifugación y actúa como barrera entre la fase líquida y las células. El gel separador es un gel inerte, como un gel de poliéster o un gel de silicona. Otros geles adecuados incluyen un gelificante a base de sorbitol en un oligómero de diacrilato.

[0039] En algunas realizaciones, el dispositivo de recolección de sangre comprende un recipiente que tiene solo una o más composiciones, donde la una o más composiciones consisten en un anticoagulante y un ion metálico, tal como EDTA quelado, y un gel separador.

[0040] En algunas realizaciones, el dispositivo de recolección contiene un coagulante.

5.2.2 Almacenamiento del dispositivo de recolección de sangre y extracción de moléculas grandes del plasma

[0041] En algunas realizaciones, después de que se recolecta la muestra de sangre de un paciente, la sangre se puede almacenar a temperaturas entre -20 °C y 35 °C, como la temperatura ambiente estándar (25 °C), la temperatura ambiente (20-24 °C), las temperaturas de refrigeración (1-8 °C) o cualquier temperatura intermedia durante un período de hasta 7-9 días antes de la centrifugación para separar el plasma de la fracción celular y la extracción de ADN. El tiempo de almacenamiento antes de la centrifugación puede ser de 2, 3, 4, 5, 6, 7, 8 o 9 días. La temperatura de almacenamiento puede ser de 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C o algún valor intermedio o rango de los mismos. En otras realizaciones, la temperatura de almacenamiento puede estar dentro del rango de 9-23 °C, 20-25 °C, 25-30 °C o 30-35 °C. En algunas realizaciones, el tiempo de almacenamiento no excede los 7 días o está entre 2-5 días, 2-6 días, 2-7 días, 3-5 días, 3-6 días, 4-6 días, 3-7 días o 4-7 días. En algunas realizaciones, el tiempo de almacenamiento está entre 2-7 días o 3-7 días y se almacena a temperatura de refrigeración, temperatura ambiente o alguna temperatura intermedia.

[0042] Una vez retirada del almacenamiento, como el almacenamiento refrigerado, la muestra se puede fraccionar para eliminar las partículas de mayor tamaño. Las técnicas para ello se describen en Pös et al., Int. J. Mol. Sci. 2020, 21, 8634, titulado "Technical and Methodological Aspects of Cell-Free Nucleic Acids Analyzes", que se incorpora aquí por referencia. En una realización, esto se realiza mediante centrifugación. La centrifugación puede ser un proceso de un solo paso o de dos pasos. En algunas realizaciones, la centrifugación se lleva a cabo a temperaturas de refrigeración. En otras realizaciones, es a temperatura ambiente. Para el primer paso de centrifugación, la muestra se somete a una centrifugación a baja velocidad. El plasma se extrae de

la muestra fraccionada a baja velocidad. Para el segundo paso, el plasma se somete a un paso de centrifugación a baja o alta velocidad. La práctica estándar es un paso de alta velocidad. Los parámetros de centrifugación a baja velocidad son aquellos que se utilizan comúnmente para separar el plasma de los componentes celulares en una muestra de sangre. Por ejemplo, para el paso de baja velocidad, la velocidad de centrifugado puede estar entre 1.000 y 2.000 x g, y el tiempo de centrifugado puede ser de 8, 9, 10, 11, 12 minutos o algún tiempo intermedio o un rango de los mismos. En algunas realizaciones, la velocidad de centrifugado puede ser de 1.500 a 1.900 x g o de 1.600 a 1.800 x g y el tiempo de centrifugado es de 9 a 11 minutos o 10 minutos. Para el paso de alta velocidad, la velocidad de centrifugado puede ser aproximadamente de 3 a 10 veces la del paso de baja velocidad durante la misma cantidad de tiempo. Por ejemplo, la centrifugación a alta velocidad puede ser de 3.000 a 18.000 x g.

[0043] En lugar de un segundo paso de centrifugación, también se puede utilizar un paso de microfiltración para eliminar partículas más grandes del plasma. En algunas realizaciones, el paso de microfiltración comprende un proceso de separación por membrana para eliminar las partículas que tienen un peso molecular promedio >200 kDa, >100 kDa, >75 kDa o >65 kDa utilizando membranas con un tamaño de poro inferior a 3, 2 o 1 μm . En algunas realizaciones, el filtro de membrana de microfiltración tiene un tamaño de poro de 0,55 μm , 0,50 μm o 0,45 μm o menos.

[0044] En otras realizaciones, la muestra de sangre se recolecta en un dispositivo de recolección, tal como con un dispositivo de recolección con un gel separador, y se centrifuga poco después de la recolección a baja velocidad para separar el plasma de la fracción celular de la muestra de sangre. En realizaciones, este paso se puede realizar en un plazo de 6 a 8 horas desde la recolección de la muestra y el dispositivo de recolección con la muestra fraccionada en su interior (es decir, no se requiere el aislamiento inmediato del plasma) se puede almacenar a partir de entonces. Como alternativa, el plasma se puede retirar del dispositivo de recolección y almacenar en un segundo dispositivo de recolección que no comprende un fijador y almacenar. Los parámetros de centrifugación a baja velocidad son los mismos que los descritos anteriormente. La temperatura de almacenamiento y el tiempo de almacenamiento son los mismos que los descritos anteriormente. Una vez retirado del almacenamiento, el plasma se puede retirar del dispositivo de recolección, si no se retiró antes del almacenamiento, y se puede someter a un segundo paso de centrifugación o un paso de filtración de acuerdo con los parámetros descritos anteriormente.

5.2.3 Aislamiento de ADNcf

[0045] Después de que se recolecta una fracción de plasma como se ha descrito anteriormente, el ADNcf se aísla/extrae del plasma. La extracción es en realidad un proceso de varios pasos que implica separar el ADN del plasma en una columna u otra matriz de unión en fase sólida. El ADNcf extraído normalmente incluye tanto ADNcf materno como fetal. En algunas realizaciones, el aislamiento se puede realizar mediante las siguientes operaciones.

A. Desnaturalizar y/o degradar las proteínas en el plasma (p. ej., contacto con proteasas) y añadir clorhidrato de guanidina u otro reactivo caotrópico a la solución (para facilitar la extracción del ADNcf de la solución)

B. Poner en contacto el plasma tratado con una matriz de soporte, como una membrana o perlas en una columna o partículas superparamagnéticas con un recubrimiento para unir selectivamente los ácidos nucleicos. Dichos recubrimientos pueden comprender sílice, carboxilos, aminas, imidazoles o combinaciones de los mismos.

C. Lavar la matriz de soporte.

D. Liberar el ADNcf de la matriz con una solución de liberación para recuperar el ADNcf para el proceso posterior (p. ej., preparación de la biblioteca indexada) y los análisis estadísticos.

[0046] La primera parte de este procedimiento de aislamiento de ADNcf implica desnaturalizar o degradar las proteínas del nucleosoma y tomar otras medidas para liberar el ADN del nucleosoma. Una mezcla de reactivos típica utilizada para lograr este aislamiento incluye un detergente, proteasa y un agente caotrópico como el clorhidrato de guanina. La proteasa sirve para degradar las proteínas del nucleosoma, así como las proteínas de fondo en el plasma, como la albúmina y las inmunoglobulinas. El agente caotrópico altera la estructura de las macromoléculas al interferir con las interacciones intramoleculares mediadas por fuerzas no covalentes, como los enlaces de hidrógeno. El agente caotrópico también hace que los componentes del plasma, como las proteínas, tengan carga negativa. La carga negativa hace que el medio sea algo incompatible energéticamente con el ADN cargado negativamente. El uso de un agente caotrópico para facilitar la purificación del ADN se describe en Boom et al., “Rapid and Simple Method for Purification of Nucleic Acids”, J. Clin. Microbiology, vol. 28, núm. 3, 1990.

[0047] Después de este tratamiento de degradación de proteínas, que libera, al menos parcialmente, las hebras de ADN de las proteínas del nucleosoma, la solución resultante se hace pasar a través de una columna o se expone de otro modo a una matriz de soporte que adhiere selectivamente el ADNcf en el plasma tratado. La carga negativa impartida a los componentes del medio facilita la adsorción de ADN en los poros de una matriz de soporte.

[0048] Después de pasar el plasma tratado a través de la matriz de soporte, la matriz de soporte con el ADNcf unido se lava para eliminar las proteínas y otros componentes no deseados de la muestra. Después del lavado, el ADNcf se libera de la matriz y se recupera. A modo de ejemplo, en el caso de perlas funcionalizadas con amina o imidazol, en una solución ácida, el ADNcf se adsorbe a la fracción de amina o imidazol cargada positivamente. En una solución básica, las fracciones de amina e imidazol están cargadas negativamente y el ADNcf no se adsorbe. En consecuencia, para eliminar los componentes no deseados de la muestra, las perlas magnéticas se inmovilizan con un imán y las perlas se lavan con un lavado ácido. Después del lavado, las perlas se separan del imán y el ADNcf se puede recuperar mediante el lavado con una solución de pH más alto. Se utiliza un proceso similar para perlas recubiertas de carboxilo. En el caso de las perlas recubiertas de carboxilo, el ADNcf se une reversiblemente a las perlas en presencia de polietilenglicol (PEG) dependiendo de la concentración de sal en ellas.

[0049] Estos procesos de extracción de ADN suelen perder una fracción significativa del ADN disponible en el plasma. Las matrices de soporte tienen una alta afinidad por el ADNcf, lo que limita la cantidad de ADNcf que se puede separar fácilmente de la matriz. Como consecuencia, el rendimiento del paso de extracción de ADNcf puede ser bastante bajo. Normalmente, la eficiencia está muy por debajo del 50% (por ejemplo, se ha descubierto que el rendimiento típico de ADNcf es de 4-12 ng/ml de plasma de los ~30 ng/ml de plasma disponibles).

[0050] Se pueden utilizar kits disponibles comercialmente para la separación manual y automatizada de ADNcf, como el kit QIAamp DNA Micro (Qiagen de Valencia, CA) o el kit NucleoSpin Plasma (Macherey–Nagel o Duren, Alemania). Un proceso de aislamiento de ADNcf se describe en la Patente de EE. UU. N.º 10.837.055, que se incorpora aquí por referencia en su totalidad.

5.2.4 Preparación de la biblioteca de secuencias

[0051] El ADNcf purificado obtenido a partir del procedimiento de aislamiento/extracción anterior se puede utilizar para preparar una biblioteca para secuenciación. Para secuenciar una población de fragmentos de ADN de doble cadena utilizando sistemas de secuenciación masivamente paralelos, los fragmentos de ADN deben estar flanqueados por secuencias adaptadoras conocidas. Una colección de dichos fragmentos de ADN con adaptadores en cada extremo se denomina biblioteca de secuenciación. Dos ejemplos de métodos adecuados para generar bibliotecas de secuenciación a partir de ADN purificado son (1) la unión basada en ligadura de adaptadores conocidos a cada extremo del ADN fragmentado y (2) la inserción mediada por transposasa de secuencias adaptadoras.

[0052] Los métodos para generar una biblioteca de secuencias se describen en la publicación de EE.UU. n.º US2007/0128624 titulada "Methods for preparing libraries of template polynucleotides" de Illumina Cambridge, Ltd., la publicación de EE.UU. n.º US2013/0123120 titulada "Highly Multiplex PCR Methods and Compositions" de Natera, Inc., J.F. Hess, et al., *Library preparation for next generation sequencing: A review of automation strategies*, Biotechnology Advances, volumen 41, 2020, 107537, ISSN 0734-9750, <https://doi.org/10.1016/j.biotechadv.2020.107537>, y Head, Steven R et al. *Library construction for next-generation sequencing: overviews and challenges*, BioTechniques vol. 56,2 61-4, 66, 68, passim. 1 de febrero de 2014, doi:10.2144/000114133, todos los cuales se incorporan aquí por referencia en su totalidad. El experto apreciaría que existen muchas técnicas adecuadas de secuenciación paralela masiva.

[0053] Las bibliotecas de secuenciación se pueden normalizar para facilitar la agrupación de muestras volumétricas utilizando técnicas conocidas por el experto en la materia.

5.2.5 Amplificación del ADNcf

[0054] En algunas realizaciones, el ADNcf presente en la muestra puede amplificarse, ya sea antes o después de preparar la biblioteca de secuencias. La amplificación puede realizarse con cualquier técnica adecuada, como métodos basados en la reacción en cadena de la polimerasa (PCR) o la amplificación por polimerasa recombinasa (RPA). Dichas técnicas son conocidas por los expertos en el campo. La RPA se describe en Lobato, et al., *Recombinase polymerase amplification: Basics, applications and recent advances*, Trends in analytical chemistry: TRAC

vol. 98 (2018): 19-35. doi:10.1016/j.trac.2017.10.015, que se incorpora aquí por referencia en su totalidad.

[0055] Para amplificar un segmento de ADN mediante PCR, primero se calienta la muestra para que el ADN se desnaturalice, o se separe en dos fragmentos de ADN de hebra sencilla. A continuación, una enzima denominada "polimerasa Taq" sintetiza - crea - dos nuevas hebras de ADN, utilizando las hebras originales como moldes. Este proceso da como resultado la duplicación del ADN original, con cada una de las nuevas moléculas que contienen una hebra antigua y una nueva de ADN. A continuación, cada una de estas hebras se puede utilizar para crear dos nuevas copias, y así sucesivamente. El ciclo de desnaturalización y síntesis de ADN nuevo se repite. En algunas realizaciones, se realizan 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 ciclos de amplificación. La PCR se describe con más detalle en Lorenz, Todd C., *Polymerase chain reaction: basic protocol plus troubleshooting and optimization strategies*, Journal of Visualized Experiments, 63 e3998. 22 de mayo de 2012, doi:10.3791/3998, que se incorpora aquí por referencia en su totalidad.

5.2.6 Selección de tamaño

[0056] Antes o después de la amplificación, se realiza una selección de tamaño para aumentar el porcentaje de la fracción fetal en la muestra, aumentando así su señal cuando se secuencian. La longitud máxima de secuencia deseada de ADNcf para aislar mediante selección de tamaño puede ser 185 pb, 180 pb, 170 pb, 175 pb, 170 pb, 165 pb, 160 pb, 155 pb, 150 pb o 145 pb, o algún valor intermedio. En algunas realizaciones, los tamaños deseados de ADNcf están en un rango de 80-160 pb, 80-155 pb, 80-150 pb, 90-160 pb, 90-155 pb, 90-150 pb, 100-150 pb, 100-155 pb, 100-160 pb, 110-150 pb, 110-155 pb, 110-160 pb, 80-170 pb o 80-185 pb. En algunas realizaciones, los tamaños deseados de ADNcf están en un rango de 100-200 pb, 150-250 pb, 200-300 pb y 250-300 pb.

[0057] La selección de tamaño se puede realizar mediante cualquier técnica para separar macromoléculas por tamaño, como electroforesis en gel, cromatografía, matriz de unión de perlas, o filtración por membrana. En el caso de la electroforesis en gel, el rango de tamaño deseado se puede recuperar manualmente o mediante automatización utilizando un instrumento, como LightBench o NIMBUS Select de Yourgene Health Canada Inc.

[0058] Los métodos manuales para recuperar ADN de un gel electroforético incluyen: (1) cortar una porción del gel correspondiente a una banda de moléculas de ADN de una longitud particular, seguido de la extracción de ADN de la porción de gel y la precipitación del ADN; y (2) recolectar una alícuota o una serie de alícuotas directamente del gel electroforético en un punto del gel que sea accesible a una pipeta, conocido en la técnica como un "pozo de extracción", seguido de la precipitación del ADN.

[0059] Un sistema de electroforesis automatizado que puede utilizarse para la selección de tamaño se describe en la patente de EE. UU. N.º 10.775.344 propiedad de Yourgene Health Canada Inc. y que se incorpora aquí por referencia en su totalidad. El sistema de electroforesis puede configurarse para distribuir potencia a cada canal de gel y puede modular dicha potencia en función de un control de la retroalimentación de corriente en cada canal. En una realización, el sistema comprende un casete interconectado que comprende una pluralidad de canales, comprendiendo el sistema: una estación de trabajo robótica para recibir al menos una señal de control de un dispositivo informático externo para controlar la operación de electroforesis en el mismo, comprendiendo la estación de trabajo robótica: un brazo que comprende al menos una pipeta; y un módulo de potencia incorporado; un primer pedestal modular acoplado eléctricamente y recibido en la estación de trabajo robótica, estando alojado el casete interconectado dentro o posicionado sobre el pedestal, comprendiendo el casete interconectado la pluralidad de canales, estando asociados un par de contactos eléctricos con cada uno de la pluralidad de canales, incluyendo el primer pedestal modular una pluralidad de cables o alambres conductores eléctricos para distribuir una señal de potencia controlable independientemente desde el módulo de potencia a cada canal, comprendiendo el pedestal un procesador y una memoria, estando configurado el procesador para: recibir una señal de potencia desde el módulo de potencia; recibir dicha al menos una señal de control designada para el casete interconectado del primer pedestal modular; y modular la señal de potencia en dependencia de la señal de control para generar la señal de potencia controlable independientemente definida para cada canal del casete interconectado del primer pedestal modular. En una realización, el procesador está configurado para modular la señal de potencia adicionalmente en función de los parámetros de electroforesis almacenados en la memoria y asociados con cada uno de dichos canales de casete. El sistema puede comprender además al menos un segundo pedestal modular y un módulo de transferencia de potencia acoplado eléctricamente al módulo de potencia, estando configurado el módulo de transferencia de potencia para comunicar

la señal de potencia entre el módulo de potencia y dicho primer pedestal modular y dicho al menos segundo pedestal modular como una transmisión. La señal de control es identificable de forma única para cada uno de dichos canales de casete mediante multiplexación. La modulación de la señal de potencia puede comprender la modulación utilizando al menos uno de los siguientes métodos: control de voltaje analógico; modulación de ancho de pulso; control de ciclo de trabajo para modulación de ancho de pulso; y control de frecuencia para modulación de ancho de pulso. En realizaciones particulares, el procesador está configurado además para (i.) monitorear una retroalimentación de corriente de cada uno de dichos canales; y (ii.) ajustar la señal de potencia modulada para cada uno de dichos canales en función de dicho monitoreo.

[0060] Otras técnicas de electroforesis se describen a continuación: la electroforesis capilar en una matriz polimérica de baja viscosidad y auto-recubrimiento se describe en Du M, Flanagan J H Jr, Lin B, Ma Y, *Electrophoresis* 2003 septiembre; 24 (18): 3147-53), que se incorpora aquí por referencia en su totalidad; la extracción selectiva en un dispositivo de electroforesis microfabricado se describe en Lin R, Burke D T, Burn M A, *J. Chromatogr. A.* 2003 agosto. 29; 1010(2): 255-68, que se incorpora aquí por referencia en su totalidad; y la electroforesis de microchip en matrices poliméricas de viscosidad reducida se describe en Xu F, Jabasini M, Liu S, Baba Y, *Analyst.* 2003 junio; 128(6): 589-92), que se incorpora aquí por referencia en su totalidad.

[0061] La selección de tamaño de ADNcf también se puede realizar mediante cromatografía, como cromatografía en geles de agarosa o poliacrilamida, cromatografía líquida de alto rendimiento de fase inversa de pares iónicos (IP RP HPLC, véase Hecker K H, Green S M, Kobayashi K, *J. Biochem. Biophys. Methods* 20 de noviembre de 2000; 46(1-2): 83-93, que se incorpora aquí por referencia en su totalidad), cromatografía de membrana de adsorción (véase Teeters M A, Conrardy S E, Thomas B L, Root T W, Lightfoot E N, *J. Chromatogr. A.* 7 de marzo de 2003; 989(1): 165-73 que se incorpora aquí por referencia en su totalidad); centrifugación en gradiente de densidad (véase Raptis L, Menard H A, *J. Clin. Invest.* 1980 diciembre; 66(6): 1391-9, que se incorpora aquí por referencia en su totalidad); y métodos que utilizan medios nanotecnológicos como matrices de trampas entrópicas microfabricadas (véase Han J, Craighead H G, *Analytical Chemistry*, vol. 74, n.º 2, 15 de enero de 2002, que se incorpora aquí por referencia en su totalidad). Una técnica para la selección de tamaño basada en perlas para ADNcf que también se puede utilizar se describe en Raymond et al., *UltraPrep is a scalable, cost-effective, bead-based*

method for purifying cell-free DNA. PLoS ONE 15(6) (2020): e0231854. <https://doi.org/10.1371/journal.pone.0231854>, que se incorpora aquí por referencia en su totalidad.

[0062] Si el paso de selección de tamaño se realiza después de la preparación de la biblioteca de secuencias, la longitud máxima deseada de ADNcf (por ejemplo, menos de 160 pb) o el rango de longitudes es el mismo, pero con el fin de aislar la banda apropiada que sale del dispositivo de selección de tamaño, el máximo objetivo o el rango a aislar tendrían que ajustarse para tener en cuenta los adaptadores u otros polinucleótidos ligados o insertados en el ADNcf. Por ejemplo, si el uso de adaptadores agrega 100 pb a la longitud de un polinucleótido de ADNcf y el máximo objetivo fue menor a 160 pb, entonces los parámetros de selección de tamaño serían menores a 260 pb.

[0063] Al realizar el paso de selección de tamaño en una muestra de una mujer embarazada, el porcentaje de ADNcf fetal en la muestra que se va a analizar para obtener información de secuencia aumenta en un 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190% o 200% en comparación con el porcentaje anterior a la selección de tamaño. Como se muestra en el Ejemplo 1 y la Figura 3, el porcentaje de ADNcf fetal estimado por SeqFF es aproximadamente igual o mayor que el porcentaje de ADNcf fetal en una muestra almacenada en un tubo Streck® durante el mismo período. (SeqFF es la fracción fetal expresada como porcentaje del ADNcf total secuenciado, calculado a partir de los recuentos de lecturas de secuencias. Consulte Kim SK, et al.: *Determination of fetal DNA fraction from the plasma of pregnant women using sequence read counts*. Prenat Diagn 2015, 35:810-815, que se incorpora aquí por referencia en su totalidad). De manera similar, al observar solo muestras masculinas, la muestra de tamaño seleccionado tiene un mayor porcentaje de ADNcf para todas las muestras.

[0064] De manera similar, al realizar el paso de selección de tamaño en una muestra para un diagnóstico de cáncer, el porcentaje de ADNcf de células cancerosas en la muestra que se va a analizar para obtener información de secuencia aumenta en un 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190% o 200% en comparación con el porcentaje anterior a la selección de tamaño.

5.2.7 Análisis del ADNcf para obtener información de secuencia

[0065] Una vez que se completa la selección de tamaño y la preparación de la biblioteca, la muestra se puede analizar para obtener información de secuencia. La información de secuencia se

puede analizar para detectar mutaciones y anomalías cromosómicas, incluidas, entre otras, translocación, transversión, monosomía, trisomía y otras aneuploidías, delección, inserción, metilación, amplificación, fragmento, translocación y reordenamiento o cromatrispsis. La información de secuencia también se puede utilizar para detectar cualquier alteración en la secuencia genética en comparación con una secuencia de referencia.

[0066] En una realización, la muestra se secuencia como parte de un procedimiento para determinar una secuencia de interés y para evaluar variaciones en el número de copias (CNV, por sus siglas en inglés). Se puede utilizar cualquiera de varias tecnologías de secuenciación, incluidas la secuenciación de "primera generación", la secuenciación de próxima generación, la secuenciación de lectura larga o la secuenciación de nanoporos.

[0067] Algunas tecnologías de secuenciación están disponibles comercialmente, como la plataforma de secuenciación por hibridación de Affymetrix Inc. (Sunnyvale, California), las plataformas de secuenciación por síntesis de Illumina, Inc. (Hayward, California) y la secuenciación basada en nanobolas de Complete Genomics (San José, CA). Otras tecnologías de secuenciación de moléculas individuales incluyen, entre otras, la tecnología SMRT™ de Pacific Biosciences (véase Anthony Rhoads et al., *PacBio Sequencing and Its Applications, Genomics, Proteomics & Bioinformatics*, Volumen 13, Número 5, 2015, página 278-289, ISSN 1672-0229, <https://doi.org/10.1016/j.gpb.2015.08.002>, que se incorpora aquí por referencia en su totalidad), la tecnología ION TORRENT™ y la secuenciación de nanoporos fueron desarrolladas, por ejemplo, por Oxford Nanopore Technologies (véase Soni G V y Meller A. *Clin Chem* 53: 1996-2001 [2007], que se incorpora aquí por referencia en su totalidad).

[0068] Si bien el método Sanger automatizado se considera una tecnología de "primera generación", la secuenciación Sanger, incluida la secuenciación Sanger automatizada, también se puede emplear en los métodos descritos en este documento. Los métodos de secuenciación adecuados adicionales incluyen, entre otros, tecnologías de obtención de imágenes de ácidos nucleicos, por ejemplo, microscopía de fuerza atómica (AFM, por sus siglas en inglés) o microscopía electrónica de transmisión (TEM, por sus siglas en inglés).

[0069] A modo de ejemplo, la tecnología de secuenciación de Illumina se puede utilizar para secuenciar el ADNcf. La técnica de Illumina se basa en la unión del ADNcf a una superficie plana, ópticamente transparente sobre la que se unen los anclajes de oligonucleótidos. El ADN molde se repara en sus extremos para generar extremos fosforilados romos en 5', y la actividad de la

polimerasa del fragmento Klenow se utiliza para añadir una única base A al extremo 3' de los fragmentos de ADN fosforilados romos. Esta adición prepara los fragmentos de ADN para la ligadura a adaptadores de oligonucleótidos, que tienen un saliente de una única base T en su extremo 3' para aumentar la eficiencia de la ligadura. Los oligonucleótidos adaptadores son complementarios a los anclajes de la celda de flujo. En condiciones de dilución limitante, el ADN molde de hebra sencilla modificado por adaptador se añade a la celda de flujo y se inmoviliza por hibridación a los anclajes. Los fragmentos de ADN unidos se extienden y se amplifican mediante puentes para crear una celda de flujo de secuenciación de densidad ultraalta con cientos de millones de grupos, cada uno de los cuales contiene 1.000 copias del mismo molde. En una modalidad, el ADNcf se amplifica mediante PCR antes de someterse a la amplificación de grupos. Como alternativa, se utiliza una preparación de biblioteca genómica sin amplificación y el ADNcf se enriquece utilizando solo la amplificación de grupos (Kozarewa et al., *Nature Methods* 6:291-295 (2009)). Los moldes se secuencian utilizando una tecnología robusta de secuenciación de ADN por síntesis de cuatro colores que emplea terminadores reversibles con colorantes fluorescentes removibles. La detección de fluorescencia de alta sensibilidad se logra utilizando excitación láser y óptica de reflexión interna total. Las lecturas de secuencia corta de aproximadamente 20-300 pb, por ejemplo, 150 pb, se pueden alinear con un genoma de referencia de repeticiones enmascaradas y se identifica el mapeo único de las lecturas de secuencia corta con el genoma de referencia utilizando un software especialmente desarrollado de canalización de análisis de datos. También se pueden utilizar genomas de referencia sin repeticiones enmascaradas. Ya se utilicen genomas de referencia con o sin repeticiones enmascaradas, solo se cuentan las lecturas que se asignan de forma única al genoma de referencia. Después de completar la primera lectura, los moldes se pueden regenerar in situ para permitir una segunda lectura desde el extremo opuesto de los fragmentos. Por lo tanto, se puede utilizar la secuenciación de un solo extremo o de extremos emparejados de los fragmentos de ADN. Se realiza la secuenciación parcial de los fragmentos de ADN presentes en la muestra y se cuentan las etiquetas de secuencia que comprenden lecturas de longitud predeterminada, por ejemplo, 36 pb, que se asignan a un genoma de referencia conocido. En una realización, la secuencia del genoma de referencia es la secuencia NCBI36/hg18, que está disponible en genome.ucsc.edu/cgi-bin/hgGateway?org=Human&db=hg18&hgsid=166260105). Como alternativa, la secuencia genómica de referencia es GRCh37/hg19, que está disponible en genome.ucsc.edu/cgi-bin/hgGateway. Otras fuentes de información pública sobre secuencias

incluyen GenBank, dbEST, dbSTS, EMBL (el Laboratorio Europeo de Biología Molecular) y DDBJ (el Banco de Datos de ADN de Japón). Hay varios algoritmos informáticos disponibles para alinear secuencias, incluidos, entre otros, BLAST (Altschul et al., 1990), BLITZ (MPsrch) (Sturrock y Collins, 1993), FASTA (Person y Lipman, 1988) o BOWTIE (Langmead et al., Genome Biology 10:R25.1-R25.10 (2009)), todos los cuales se incorporan aquí por referencia en su totalidad.

[0070] A modo de otro ejemplo, la información de secuencia de los ácidos nucleicos en la muestra de prueba, por ejemplo, ADNcf en una muestra de prueba materna o ADNcf en un sujeto que se somete a un examen de detección de cáncer, se puede obtener utilizando la secuenciación 454 (Roche) (por ejemplo, como se describe en Margulies, M. et al. Nature 437:376-380 [2005], que se incorpora aquí por referencia en su totalidad). La secuenciación 454 implica típicamente dos pasos. En el primer paso, el ADN se corta en fragmentos de aproximadamente 300-800 pares de bases, y los fragmentos se vuelven romos en sus extremos. A continuación, se ligan adaptadores de oligonucleótidos a los extremos de los fragmentos. Los adaptadores sirven como cebadores para la amplificación y secuenciación de los fragmentos. Los fragmentos se pueden unir a perlas de captura de ADN, por ejemplo, perlas recubiertas de estreptavidina utilizando, por ejemplo, el Adaptador B, que contiene una etiqueta de biotina 5'. Los fragmentos unidos a las perlas se amplifican por PCR dentro de gotitas de una emulsión de aceite-agua. El resultado son múltiples copias de fragmentos de ADN amplificados clonalmente en cada perla. En el segundo paso, las perlas se capturan en pocillos (por ejemplo, pocillos del tamaño de un microlitro). La pirosecuenciación se realiza en cada fragmento de ADN en paralelo. La adición de uno o más nucleótidos genera una señal luminosa que se registra mediante una cámara CCD en un instrumento de secuenciación. La intensidad de la señal es proporcional al número de nucleótidos incorporados. La pirosecuenciación utiliza pirofosfato (PPi) que se libera tras la adición de nucleótidos. El PPi se convierte en ATP mediante la ATP sulfurilasa en presencia de adenosina 5' fosfosulfato. La luciferasa utiliza ATP para convertir la luciferina en oxiluciferina, y esta reacción genera luz que se mide y analiza.

[0071] En otra realización ilustrativa, pero no limitante, la información de secuencia puede obtenerse realizando una secuenciación por ligadura, que corta el ADN genómico en fragmentos y une adaptadores a los extremos 5' y 3' de los fragmentos para generar una biblioteca de fragmentos. Alternativamente, los adaptadores internos pueden introducirse ligando adaptadores a

los extremos 5' y 3' de los fragmentos, circularizando los fragmentos, digiriendo el fragmento circularizado para generar un adaptador interno y uniendo adaptadores a los extremos 5' y 3' de los fragmentos resultantes para generar una biblioteca de pares de parejas. A continuación, se preparan poblaciones de perlas clonales en microrreactores que contienen perlas, cebadores, molde y componentes de PCR. Después de la PCR, los moldes se desnaturalizan y las perlas se enriquecen para separar las perlas con moldes extendidos. Los moldes en las perlas seleccionadas se someten a una modificación 3' que permite la unión a un portaobjetos de vidrio. La secuencia se puede determinar mediante hibridación secuencial y ligadura de oligonucleótidos parcialmente aleatorios con una base determinada central (o par de bases) que se identifica mediante un fluoróforo específico. Después de registrar un color, el oligonucleótido ligado se escinde y se elimina y luego se repite el proceso.

[0072] En otra realización, la información de secuencia para los ácidos nucleicos en la muestra de prueba, por ejemplo, ADNcf en una muestra de prueba materna, ADNcf en un sujeto que se está examinando para detectar un cáncer, se puede obtener utilizando la tecnología de secuenciación de molécula única en tiempo real (SMRT™) de Pacific Biosciences. En la secuenciación SMRT, la incorporación continua de nucleótidos marcados con colorante se visualiza durante la síntesis de ADN. Moléculas de ADN polimerasa individuales se unen a la superficie inferior de detectores de longitud de onda de modo cero individuales (detectores ZMW) que obtienen información de secuencia mientras los nucleótidos fosfoenlazados se incorporan a la cadena de cebador en crecimiento. Un detector ZMW comprende una estructura de confinamiento que permite la observación de la incorporación de un único nucleótido por la ADN polimerasa contra un fondo de nucleótidos fluorescentes que se difunden rápidamente dentro y fuera del ZMW (por ejemplo, en microsegundos). Por lo general, se necesitan varios milisegundos para incorporar un nucleótido a una cadena en crecimiento. Durante este tiempo, la etiqueta fluorescente se excita y produce una señal fluorescente, y la etiqueta fluorescente es escindida. La medición de la fluorescencia correspondiente del colorante indica qué base se incorporó. El proceso se repite para proporcionar una secuencia.

[0073] En otra realización, la información de secuencia de los ácidos nucleicos en la muestra de prueba, por ejemplo, ADNcf en una muestra de prueba materna o ADNcf en un sujeto que se está examinando para detectar un cáncer, se obtiene mediante el uso de secuenciación de nanoporos. Las técnicas de análisis de ADN de secuenciación de nanoporos son desarrolladas por varias

empresas, incluidas, por ejemplo, Oxford Nanopore Technologies (Oxford, Reino Unido), Sequenom y NABsys. La secuenciación de nanoporos es una tecnología de secuenciación de una sola molécula mediante la cual una sola molécula de ADN se secuencia directamente a medida que pasa a través de un nanoporo. Un nanoporo es un orificio pequeño, típicamente del orden de 1 nanómetro de diámetro. La inmersión de un nanoporo en un fluido conductor y la aplicación de un potencial (voltaje) a través de éste da como resultado una ligera corriente eléctrica debido a la conducción de iones a través del nanoporo. La cantidad de corriente que fluye es sensible al tamaño y la forma del nanoporo. A medida que una molécula de ADN pasa a través de un nanoporo, cada nucleótido de la molécula de ADN obstruye el nanoporo en un grado diferente, modificando la magnitud de la corriente a través del nanoporo en diferentes grados. Por lo tanto, este cambio en la corriente a medida que la molécula de ADN pasa a través del nanoporo proporciona una lectura de la secuencia de ADN.

[0074] En otra realización, la información de secuencia para los ácidos nucleicos en la muestra de prueba, por ejemplo, ADNcf en una muestra de prueba materna o ADNcf en un sujeto que se está examinando para detectar un cáncer, se puede obtener utilizando la matriz de transistores de efecto de campo sensibles químicamente (chemFET) (por ejemplo, como se describe en la publicación de solicitud de patente de EE. UU. N.º 2009/0026082, que se incorpora aquí por referencia en su totalidad). En un ejemplo de esta técnica, las moléculas de ADN se pueden colocar en cámaras de reacción y las moléculas de molde se pueden hibridar con un cebador de secuenciación unido a una polimerasa. La incorporación de uno o más trifosfatos en una nueva cadena de ácido nucleico en el extremo 3' del cebador de secuenciación se puede discernir como un cambio en la corriente mediante un chemFET. Una matriz puede tener múltiples sensores chemFET. En otro ejemplo, se pueden unir ácidos nucleicos individuales a perlas, y los ácidos nucleicos se pueden amplificar en la perla, y las perlas individuales se pueden transferir a cámaras de reacción individuales en una matriz chemFET, donde cada cámara tiene un sensor chemFET, y los ácidos nucleicos se pueden secuenciar.

[0075] En otra realización, la información de secuencia se obtiene utilizando la secuenciación de moléculas individuales de Ion Torrent, que combina la tecnología de semiconductores con una química de secuenciación simple para traducir directamente la información codificada químicamente (A, C, G, T) en información digital (0, 1) en un chip semiconductor. En la naturaleza, cuando un nucleótido se incorpora a una cadena de ADN mediante una polimerasa, se

libera un ion de hidrógeno como subproducto. Ion Torrent utiliza una matriz de alta densidad de pocillos micromaquinados para realizar este proceso bioquímico de una manera masivamente paralela. Cada pocillo contiene una molécula de ADN diferente. Debajo de los pocillos hay una capa sensible a iones y debajo de ésta un sensor de iones. Cuando se añade un nucleótido, por ejemplo, un nucleótido C, a un molde de ADN y luego se incorpora a una cadena de ADN, se liberará un ion de hidrógeno. La carga de ese ion cambiará el pH de la solución, que puede detectarse mediante el sensor de iones de Ion Torrent. El secuenciador (que es básicamente el medidor de pH de estado sólido más pequeño del mundo) lee la base y pasa directamente de la información química a la información digital. A continuación, el secuenciador Ion Personal Genome Machine Machine (PGM™) inunda secuencialmente el chip con un nucleótido tras otro. Si el siguiente nucleótido que inunda el chip no coincide, no se registrará ningún cambio de voltaje y no se leerá ninguna base. Si hay dos bases idénticas en la cadena de ADN, el voltaje será el doble y el chip registrará dos bases idénticas. La detección directa permite registrar la incorporación de nucleótidos en segundos.

[0076] En otra realización, el presente método comprende obtener información de secuencia para los ácidos nucleicos en la muestra de prueba, por ejemplo, ADNcf en una muestra de prueba materna, utilizando secuenciación por hibridación. La secuenciación por hibridación comprende poner en contacto la pluralidad de secuencias de polinucleótidos con una pluralidad de sondas de polinucleótidos, en donde cada una de la pluralidad de sondas de polinucleótidos puede estar opcionalmente unida a un sustrato. El sustrato puede ser una superficie plana que comprende una matriz de secuencias de nucleótidos conocidas. El patrón de hibridación con la matriz puede utilizarse para determinar las secuencias de polinucleótidos presentes en la muestra. En otras realizaciones, cada sonda está unida a una perla, por ejemplo, una perla magnética o similar. La hibridación con las perlas puede determinarse y utilizarse para identificar la pluralidad de secuencias de polinucleótidos dentro de la muestra.

[0077] En otra realización, el presente método comprende obtener información de secuencia para los ácidos nucleicos en la muestra de prueba, por ejemplo, ADNcf en una muestra de prueba materna, mediante la secuenciación masivamente paralela de millones de fragmentos de ADN utilizando la química de secuenciación basada en terminadores reversibles y secuenciación por síntesis de Illumina (por ejemplo, como se describe en Bentley et al., Nature 6:53-59 [2009], que se incorpora aquí por referencia en su totalidad).

[0078] El mapeo de las etiquetas de secuencia se logra comparando la secuencia de la etiqueta con la secuencia de la referencia para determinar el origen cromosómico de la molécula de ácido nucleico secuenciada (en este caso, ADNcf), y no se necesita información de secuencia genética específica. Se puede permitir un pequeño grado de desajuste (0-2 desajustes por etiqueta de secuencia) para explicar polimorfismos menores que pueden existir entre el genoma de referencia y los genomas en la muestra mixta.

[0079] Normalmente se obtiene una pluralidad de etiquetas de secuencia por muestra. Las etiquetas de secuencia comprenden lecturas de entre 20 y 185 pb, por ejemplo, 100 pb. En una realización, todas las lecturas de secuencia se mapean en todas las regiones del genoma de referencia. En una realización, se cuentan las etiquetas que se han mapeado en todas las regiones, por ejemplo, todos los cromosomas, del genoma de referencia, y se determina la CNV, es decir, la sobrerrepresentación o subrepresentación de una secuencia de interés, por ejemplo, un cromosoma o una porción del mismo, en la muestra de ADN mixto. El método no requiere la diferenciación entre los dos genomas.

[0080] La precisión requerida para determinar correctamente si una CNV, por ejemplo, aneuploidía, poliploidía y delección, está presente o ausente en una muestra, se basa en la variación del número de etiquetas de secuencia que se mapean en el genoma de referencia entre muestras dentro de una serie de secuenciación (variabilidad intercromosómica), y la variación del número de etiquetas de secuencia que se mapean en el genoma de referencia en diferentes series de secuenciación (variabilidad intersecuencial). Por ejemplo, las variaciones pueden ser particularmente pronunciadas para las etiquetas que se mapean en secuencias de referencia ricas en GC o pobres en GC. Otras variaciones pueden resultar del uso de diferentes protocolos para la extracción y purificación de los ácidos nucleicos, la preparación de las bibliotecas de secuenciación y el uso de diferentes plataformas de secuenciación. El presente método utiliza dosis de secuencia (dosis de cromosomas o dosis de segmentos) basadas en el conocimiento de secuencias normalizadoras (secuencias de cromosomas normalizadora o secuencias de segmentos normalizadora), para tener en cuenta intrínsecamente la variabilidad acumulada que surge de la variabilidad intercromosómica (intra-series), inter-secuenciación (inter-series) y dependiente de la plataforma. Las dosis de cromosomas se basan en el conocimiento de una secuencia de cromosomas normalizadora, que puede estar compuesta por un solo cromosoma, o por dos o más cromosomas seleccionados de los cromosomas 1-22, X, y Y. Alternativamente, las secuencias de

cromosomas normalizadoras pueden estar compuestas por un solo segmento de cromosoma, o por dos o más segmentos de un cromosoma o de dos o más cromosomas. Las dosis de segmentos se basan en el conocimiento de una secuencia de segmentos normalizadora, que puede estar compuesta por un solo segmento de cualquier cromosoma, o por dos o más segmentos de cualesquiera dos o más cromosomas 1-22, X, y Y.

[0081] En una realización, el ADNcf se secuenciar para determinar la secuencia de alelos de un locus de interés dentro de la muestra. A partir de las secuencias, se identifica un locus de interés heterocigoto y se cuantifica una proporción de los alelos. La proporción indica la presencia o ausencia de una anomalía cromosómica.

[0082] En algunas realizaciones, los métodos también implican determinar si el feto tiene una enfermedad genética a partir de la al menos una secuencia de interés del feto. En algunas realizaciones, esto incluye determinar si el feto es homocigoto en un alelo causante de la enfermedad dentro de la secuencia de interés cuando la madre es heterocigota del mismo alelo. En algunas realizaciones, el alelo causante de la enfermedad es un alelo de un polimorfismo de un solo nucleótido (SNP, por sus siglas en inglés), en la secuencia de interés. En algunas realizaciones, el alelo causante de la enfermedad es un alelo de una repetición corta en tándem (STR, por sus siglas en inglés) en la secuencia de interés.

[0083] En algunas realizaciones, la al menos una secuencia de interés comprende un sitio de un alelo asociado con una enfermedad. La al menos una secuencia de interés puede incluir uno o más de los siguientes: polimorfismo de un solo nucleótido, repetición en tándem, microdelección, inserción, e indel. En algunas realizaciones, los métodos implican además determinar si el feto es homocigoto o heterocigoto para el alelo asociado con la enfermedad.

[0084] En algunas realizaciones, el método implica además contar las etiquetas de secuencia para determinar una variación en el número de copias (CNV) de una secuencia cromosómica, y/o examinar secuencias de las etiquetas de secuencia para detectar variaciones en el número de copias de una secuencia cromosómica. Algunas realizaciones implican además determinar para el feto la presencia o ausencia de un polimorfismo o anomalía cromosómica. En algunas realizaciones, el método implica además determinar para el feto la presencia o ausencia de un SNP, repetición en tándem, inserción, delección, indel, translocación, duplicación, inversión, CNV, aneuploidía parcial y/o aneuploidía completa. La aneuploidía cromosómica completa puede ser una duplicación, una multiplicación o una pérdida de un cromosoma completo. En algunas realizaciones, el método

implica además localizar la aneuploidía parcial. En algunas realizaciones, el SNP se selecciona de un SNP único y un SNP en tándem. En algunas realizaciones, la repetición en tándem se selecciona entre una repetición de dinucleótido y una repetición de trinucleótido.

[0085] En algunas realizaciones, el método implica además determinar un valor de secuencia cromosómica normalizada para una secuencia cromosómica de interés y comparar el valor de secuencia con un valor umbral superior e inferior, en donde el valor de secuencia que excede los valores umbral superior o inferior respectivamente determina la presencia o ausencia de la CNV.

[0086] En algunas realizaciones, el análisis de la pluralidad de etiquetas de secuencia implica determinar la presencia o ausencia de una anomalía cromosómica de los cromosomas 1-22, X, y Y. Algunas realizaciones implican además determinar si el feto es homocigoto en un alelo dentro de la secuencia de interés, o determinar que la madre es heterocigota y el feto es homocigoto. Algunas realizaciones implican además determinar la fracción fetal del ADNcf.

[0087] Los métodos y sistemas para analizar ADNcf se describen en la Patente de EE. UU. N.º 10.837.055, que se incorpora aquí por referencia en su totalidad.

[0088] Otro aspecto de esta divulgación es una prueba de validez de secuencia que puede detectar un rendimiento anormal del paso de selección de tamaño o una degradación inaceptable de la muestra. Una realización de esta prueba se describe en el Ejemplo 4 y se utilizó para evaluar la idoneidad de una muestra en el Ejemplo 3. Si bien el Ejemplo 3 describe el uso del método en el contexto del uso con un tubo K2EDTA, se entiende que esta prueba de validez se puede utilizar para evaluar la integridad de una muestra en la que se realiza un paso de selección de tamaño, independientemente del tipo de tubo de recolección o el tiempo o las condiciones de almacenamiento. En un método que implica la selección de tamaño de una muestra de ADN que comprende ADNcf y la secuenciación del ADN de la muestra después de la selección de tamaño; el método puede comprender una prueba de validez de secuencia de la información de secuencia obtenida del paso de secuenciación. La prueba de validez comprende crear un primer perfil de tamaño de fragmento normalizado a partir de la información de secuenciación asociada con la muestra, comparar uno o más valores dentro del perfil de tamaño de fragmento con uno o más de los rangos de valores correspondientes obtenidos a partir de un conjunto de parámetros de referencia que comprende una pluralidad de perfiles de tamaño de fragmento válidos; y aceptando la información de secuencia y los resultados de análisis de la misma si uno de los uno o más valores dentro del perfil de tamaño de fragmento asociado con la muestra está dentro del rango de valores

correspondiente o rechazando la información de secuencia si uno de los uno o más valores dentro del perfil de tamaño de fragmento asociado con la muestra está fuera del rango de valores correspondiente. La prueba de validez también puede comprender calcular un primer valor de frecuencia relativa del tamaño de fragmento a partir de la información de secuenciación asociada con la muestra; comparar el primer valor de frecuencia relativa del tamaño de fragmento con un conjunto de valores de frecuencia relativa del tamaño de fragmento de referencia que comprende una pluralidad de frecuencias relativas del tamaño de fragmento válidas, y aceptar la información de secuencia y los resultados del análisis de la misma si el primer valor de frecuencia relativa del tamaño de fragmento asociado con la muestra está dentro de un rango especificado de valores dentro del conjunto de referencia o rechazar la información de secuencia y los resultados del análisis de la misma si el primer valor de frecuencia relativa del tamaño de fragmento asociado con la muestra está fuera de un rango especificado de valores dentro del conjunto de referencia.

[0089] .

EJEMPLOS

5.3. Ejemplo 1 –

[0090] Este estudio es un estudio comparativo que evalúa la calidad del ADNcf recuperado de la sangre almacenada y transportada en un tubo K2EDTA con gel (después de una centrifugación baja) y en un tubo Streck sin centrifugar. Trabajos anteriores han demostrado que el tubo EDTA con gel era tan bueno como el tubo Streck para recolectar sangre y preparar el ADNcf para la prueba prenatal no invasiva (NIPT) sin reducción de la fracción fetal si la sangre se procesaba el mismo día (Giroux, 2021). El presente estudio prueba la posibilidad de mantener el tubo centrifugado a 4 °C durante cinco días y enviarlo frío al laboratorio de secuenciación a 4.800 km de distancia.

[0091] Mujeres embarazadas fueron reclutadas exclusivamente en Vancouver después de consentir en participar en el estudio. Todas las mujeres embarazadas reclutadas tenían entre 10 y 14 semanas de embarazo.

En la Tabla 1 se muestran los detalles de la población del estudio.

[0092] Tabla 1.

Número de mujeres embarazadas	61
Embarazo (semanas)	
Media	10,9
rango	10-13,9
Peso (kg)	
Media	63,9
rango	41,3-136
Índice de Masa Corporal	
Media	23,8
rango	17-42
Número de mujeres con un peso > 80 kg	5
Número de mujeres con un IMC > 30	5

[0093] Para cada participante, se recogieron dos muestras de sangre periférica de 10 ml en tubos Streck Cat# 218997 (Streck, NB, EE. UU.) y se recogieron 8 ml en tubos de sangre K2EDTA con gel para diagnóstico molecular Cat# GR-455040 (Greiner, Austria). Los tubos EDTA con gel se centrifugaron a 1.800 g durante 10 min dentro de las 6 horas posteriores a la recolección de la muestra (según lo recomendado por el fabricante) y luego se mantuvieron a 4 °C hasta que se enviaron fríos una vez a la semana a un laboratorio en la ciudad de Quebec. La sangre recolectada en tubos Streck se envió a temperatura ambiente dos veces a la semana al mismo laboratorio. En total, se recogieron 61 muestras en Vancouver y se enviaron a más de 4.800 km a un laboratorio de pruebas en la ciudad de Quebec.

[0094] El plasma se recogió cuidadosamente utilizando una pipeta de transferencia de 1 ml de los tubos Streck de 2 a 5 días después de la recolección de sangre. El plasma se filtró a través de un HPF-Millex-PVDF-Durapore de 0,45 µM de Millipore (Merck, MA, EE. UU.) utilizando una jeringa de 5 ml. El plasma en el tubo K2EDTA se decantó en el cilindro de la jeringa de 5 ml y se filtró utilizando el mismo tipo de filtro que el indicado anteriormente (Giroux, 2021). Los tubos EDTA con gel pasaron de 2 a 9 días entre la recolección de sangre y la filtración.

[0095] El ADNcf se preparó a partir de 2 a 5 ml de plasma utilizando el kit QIAamp Circulating Nucleic Acid Cat# 55114 (Qiagen, Alemania). El ADN se recuperó en 50 o 70 µl y se cuantificó por fluorometría utilizando el kit Qubit dsDNA HS Assay Kit Cat# Q32851 (ThermoFisher, MA, EE. UU.). Se utilizaron volúmenes de 15 a 25 µl que contenían entre 2 y 15 ng de ADNcf para

preparar bibliotecas utilizando el kit KAPA Hyper Prep Cat# 07962363001 (Roche, Suiza) siguiendo las instrucciones del fabricante con algunas modificaciones. Los volúmenes se redujeron a la mitad en cada paso. Se utilizaron adaptadores Kapa Unique Dual-indexed Cat # 08861919702 (Roche, Suiza) a una concentración final de 0,5 μ M. Se realizaron diez ciclos de amplificación y se realizó una limpieza posterior a la amplificación con perlas Kapa 0,6X-1X Cat #07983298001 (Roche, Suiza).

[0096] Para un grupo de muestras, las bibliotecas se cuantificaron mediante fluorometría utilizando el kit Qubit dsDNA HS Assay Kit como se indicó anteriormente. La molaridad se calculó utilizando un tamaño promedio de 325 pares de bases (pb). Se preparó un conjunto de hasta 17 bibliotecas equimolares para secuenciar extremos emparejados durante 150 ciclos en el instrumento Illumina Next-Seq 550 con una celda de flujo de salida media Cat#20024909 (Illumina, CA, EE. UU.).

[0097] Para un segundo grupo de muestras, se seleccionaron bibliotecas por tamaño utilizando el LightBench de Yourgene Health Canada Inc. (Manchester, Reino Unido). Se procesaron 25 μ l de muestra más tampón de carga y marcadores etiquetados de 200 pb/300 pb juntos en un gel prefabricado al 3%. Se programó un software para extraer ADN que oscilaba entre 229 pb y 286 pb. El ADN recuperado se cuantificó con Qubit y la masa se convirtió a molaridad utilizando un tamaño promedio de 275 pb. Se preparó un grupo de hasta 25 bibliotecas y se secuenció en una celda de flujo de salida media como se indicó anteriormente.

RESULTADOS

[0098] Mediante el procedimiento estándar con tubo Streck, las muestras llegaron en un plazo de 5 días a temperatura ambiente. Siete mujeres tenían una fracción fetal inferior al 4% y seis de ellas tenían un peso superior a 70 kg o un IMC superior a 25.

[0099] Con los tubos EDTA con gel, las muestras llegaron también en un plazo de 5 días, pero tuvieron que permanecer frías. La mayoría de las muestras llegaron en un plazo de dos días con un envío prioritario con el transportista elegido. Sin embargo, dos envíos que incluían 14 muestras se retrasaron 4 y 5 días, y estas muestras ya no se mantuvieron frías. Estas 14 muestras sufrieron durante el transporte y tuvieron una fracción fetal mucho menor de lo esperado. Su reducción media de la fracción fetal fue del 4,6% en comparación con la obtenida con el tubo Streck. Los tubos EDTA con gel que llegaron fríos en dos días o menos (N=47) tuvieron una fracción fetal

similar a la observada con el tubo Streck, pero se redujeron en un promedio de 1,1%. Hubo una tendencia entre el tiempo transcurrido a 4°C (los tubos se mantuvieron en el frigorífico hasta el envío) y el grado de reducción de la fracción fetal. Observamos una mayor reducción de la fracción fetal con el aumento del tiempo antes del tratamiento, desde una reducción del 0,75% durante 2-3 días hasta una reducción del 1,41% durante 6 días o más.

[00100] Todas las bibliotecas preparadas a partir de ADN aislado de plasma de tubos EDTA con gel se procesaron en una electroforesis en gel al 3% utilizando el instrumento LightBench™ de Yourgene Health Canada Inc. El software se configuró para recolectar fragmentos de 229 pares de bases a 286 pares de bases, lo que tiene en cuenta adaptadores que agregan 136 pb a una cadena ADNcf. El material recolectado se secuenció y la fracción fetal se evaluó con SeqFF como de costumbre. Con la excepción de 5 muestras, se observó un aumento de la fracción fetal. Las 5 que no exhibieron un aumento tuvieron fracciones fetales más altas inicialmente. Las muestras retrasadas en el transporte durante cinco días fueron las que aumentaron menos en comparación con los resultados del tubo Streck. Dos tuvieron una fracción fetal menor en 0,4% y 3%, respectivamente, en comparación con el tubo Streck, pero la fracción fetal mejoró en comparación con el tubo EDTA con gel antes de la selección de tamaño y las otras doce aumentaron entre 0,1 y 1,5% (Fig. 1. Muestras a la izquierda). Las 47 muestras restantes aumentaron en promedio un 5,4%, con un rango de 0,8 a 10,2%, excepto 3 muestras que permanecieron por encima del 10%, pero disminuyeron en comparación con Streck e incluso con tubo EDTA antes de la selección por tamaño (Fig. 1. Muestras con estrellas). La determinación de la fracción fetal con SeqFF produce una estimación aproximada, pero sin embargo es muy útil para el embarazo femenino. También se determinó la fracción fetal calculada a partir de la reducción de fragmentos ChrX en el embarazo masculino y se encontró un aumento importante para cada muestra analizada (Fig. 2A). Incluso una muestra con una fracción fetal más baja con SeqFF apareció mucho más alta con este método para calcular (estrella azul en las Figs. 1 y 2A). El tubo utilizado y la selección en agarosa al 3% no aumentaron de manera artificial la fracción fetal al interferir con el recuento de SeqFF, ya que también se observó un aumento en las lecturas del cromosoma Y en los embarazos con fetos masculinos (n = 25). (Fig. 2B complementaria).

[00101] La Figura 3 muestra el tamaño de los fragmentos secuenciados calculados a partir del mapeo de secuenciación de extremos emparejados. En gris, los fragmentos antes de la selección de tamaño y en azul, después de la selección de tamaño.

Tabla 2

	Streck	EDTA - TODAS	EDTA - 4°C	EDTA + ST
FF media (%)	8,6	6,8	7,6	12,6
FF Rango (%)	2,3-26,8	0,4-25,2	0,4-25,2	6,4-22,6
Número de muestras con FF <4%	7	16	9	0

Ejemplo 2

[00102] El objetivo de este segundo estudio es evaluar el impacto de extender el tiempo de almacenamiento antes de la centrifugación de sangre completa recolectada en un tubo K2EDTA sobre la detección de ADNcf fetal. Se incubaron tubos separados de una única extracción de sangre de cada participante durante un intervalo de puntos de tiempo, se centrifugaron para obtener plasma y luego se procesaron para aislar y amplificar el ADN, realizar una selección de tamaño, crear una biblioteca de ADN y secuenciar el ADNcf. Se utilizaron el porcentaje de fracción fetal y las proporciones cromosómicas para determinar el efecto del tiempo de almacenamiento antes de la centrifugación. Esto permite medir el impacto del tiempo variable de almacenamiento de sangre sobre el porcentaje de fracción fetal.

[00103] Las muestras se procesaron en los siguientes puntos de tiempo: Punto de tiempo 1 – 0-8 horas, Punto de tiempo 2 – 72 horas, Punto de tiempo 3 – 120 horas y Punto de tiempo 4 – 168 horas. Durante este tiempo, una parte de las muestras se almacenó a temperaturas de refrigeración y una parte se almacenó a temperatura ambiente. En el punto de tiempo apropiado, el plasma se separó de la sangre completa mediante centrifugación. Todos los tubos se centrifugaron a 1.600 g x 10 min dentro de las 6 h posteriores a la recolección de la muestra. El plasma de los tubos EDTA se vertió en un cilindro de jeringa y se filtró a través de un filtro Millipore de 0,45 µm. Las muestras de plasma aisladas se almacenaron a -80 °C hasta que se recopilaron al menos noventa y seis muestras, momento en el que comenzó la extracción de ADN. El almacenamiento a esta temperatura preservó la muestra hasta la extracción. La extracción del ADNcf se realizó con el kit QIAamp Circulating Nucleic Acid Cat# 55114 (Qiagen, Alemania). Se utilizó un LightBench® de Yourgene Health Canada Inc. para seleccionar un tamaño determinado de ADNcf en la muestra. Los parámetros de selección de tamaño son los mismos que en el Ejemplo 1. Las bibliotecas de

secuenciación NGS se prepararon y secuenciaron en un sistema Illumina. Las fracciones fetales se calcularon utilizando SeqFF.

Ejemplo 3

[00104] El objetivo de este tercer estudio también fue evaluar el impacto de extender el tiempo de almacenamiento antes de la centrifugación de sangre completa recolectada en tubos de sangre K2EDTA (también conocidos como Vacutainers K2EDTA) en la detección de ADNcf fetal. Se incubaron tubos separados de una única extracción de sangre de cada participante durante los puntos de tiempo 1 a 4, se centrifugaron dos veces para obtener plasma, luego se procesaron para aislar y amplificar el ADN, crear una biblioteca de secuenciación de ADN, agrupar las bibliotecas de secuenciación, realizar la selección de tamaño y secuenciar el ADNcf. El porcentaje de fracción fetal y las proporciones cromosómicas se utilizaron para determinar el efecto del tiempo de almacenamiento antes de la centrifugación. Esto permitió medir el impacto del tiempo de almacenamiento de sangre variable en el porcentaje de fracción fetal y la calidad de las pruebas NIPT en tubos de sangre estándar con aislamiento de plasma retrasado.

[00105] Las muestras se almacenaron a temperatura ambiente y se procesaron en los siguientes puntos de tiempo: Punto de tiempo 1 – 0-8 horas, Punto de tiempo 2 – 72 horas, Punto de tiempo 3 – 120 horas y Punto de tiempo 4 – 168 horas. Las temperaturas ambiente oscilaron entre 20 °C y 35 °C. En el punto de tiempo apropiado, el plasma se separó de la sangre completa mediante centrifugación doble. El proceso de centrifugación de aislamiento de plasma comprende un primer paso de centrifugación durante 10 minutos a 1.600 rpm y un segundo paso de centrifugación durante 10 minutos a 3.000 rpm. Después de la primera centrifugación, el plasma del tubo K2EDTA se transfirió a un segundo tubo y se sometió al segundo paso de centrifugación. Después del segundo paso, el plasma se retiró del segundo tubo y se transfirió a un tercer tubo y se almacenó a -80 °C hasta que se recogieron suficientes muestras para comenzar una serie de ≤ 48 muestras, momento en el que comenzó la extracción de ADN. Las muestras de plasma aisladas se almacenaron a -80°C hasta que se recogieron aproximadamente 48 muestras, momento en el que comenzó el proceso de extracción de ADN. El almacenamiento a esta temperatura preservó la muestra hasta la extracción.

[00106] La extracción del ADNcf se realizó con el kit MagMAX™ Cell-Free DNA Isolation Kit. A pesar de no utilizar tubos Streck, el paso de lisis se llevó a cabo como parte del guión y no tiene ningún efecto perjudicial sobre las muestras recogidas en tubos EDTA. Se utilizó un QS250TM

de Yourgene Health Canada Inc. para seleccionar un tamaño determinado de ADNcf en la muestra. Los parámetros de selección de tamaño fueron 106-146 pb (el valor no incluye la longitud del adaptador) y el rango recuperado fue ~100-186 pb (la longitud media del fragmento es ~143-146 pb). Las bibliotecas de secuenciación NGS se prepararon y secuenciaron en un sistema Illumina. Las fracciones fetales se estimaron utilizando SeqFF.

[00107] *Detalles del conjunto de muestra*

[00108] Se recogieron 105 conjuntos coincidentes de muestras clínicas para este estudio. Se recogieron muestras con edad gestacional, altura, peso e IMC conocidos. Se incubaron muestras de sangre individuales extraídas de una única extracción de sangre de cada participante en los puntos de tiempo 1 a 4, se centrifugaron para obtener plasma y luego se procesaron mediante IONA® Nx. Se utilizaron el porcentaje de fracción fetal y las proporciones cromosómicas para determinar el efecto del tiempo de almacenamiento antes de la centrifugación.

[00109] De casi todos los participantes, se recogieron cuatro muestras de sangre de 5 ml (denominadas muestras coincidentes) de cada participante en BD Vacutainer™ Plastic K2EDTA Tubes de ciento cinco pacientes voluntarios (mujeres embarazadas >10+0 semanas de gestación). Para cada participante, alícuotas coincidentes de 5 ml se sometieron a centrifugación en cada uno de los puntos de tiempo 1 a 4 después de la incubación a temperatura ambiente. La siguiente Tabla 3 proporciona un resumen de las muestras que se recogieron. Aquellos puntos de tiempo que se desvían de 105 muestras son menos de 105 muestras debido a la obtención de menos de 20 ml del participante.

[00110] Tabla 3.

Número total de pacientes		77			
N.º de Muestras Punto de tiempo 1	de	N.º de Muestras Punto de tiempo 2	N.º de Muestras Punto de tiempo 3	N.º de Muestras Punto de tiempo 4	No. Total de muestras
77		77	75	73	302

[00111] Los participantes fueron designados como M029 a M105. Las siguientes muestras se sometieron por error a una única centrifugación en lugar de a una doble centrifugación, como es habitual en el protocolo IONA Nx:

M029-M30 Punto de tiempo 1-3

M031-M032 Punto de tiempo 1-2

Estas muestras se han excluido de este informe debido a la utilización de un centrifugado único en la preparación del plasma.

[00112] Las muestras restantes se sometieron a doble centrifugación, las cuales se resumen a continuación.

M029-M030 Punto de tiempo 4

M031-M032 Punto de tiempo 3 y 4

M033-M105 Punto de tiempo 1,2,3 y 4

Las muestras anteriores son las que se incluyeron en el análisis.

[00113] **Criterios de exclusión e inclusión de participantes**

[00114] Los criterios de inclusión y exclusión fueron los siguientes:

CRITERIOS DE INCLUSIÓN

Mujer embarazada ≥ 18 años

Mujeres que pudieron comprender y dar su consentimiento informado

Embarazo único $\geq 10+0$ semanas de gestación sin límite superior de corte para la gestación, es decir, incluso hasta la fecha prevista del parto

CRITERIOS DE EXCLUSIÓN

Aneuploidía materna conocida o cáncer materno

Casos de gemelos o gemelos evanescentes

Enfermedad infecciosa conocida, por ejemplo, infección actual por COVID-19 o dentro del período de aislamiento por COVID-19, hepatitis B

Cualquier persona considerada perteneciente a un grupo vulnerable

Cualquier persona que se sepa que ha recibido un trasplante de órgano o células madre

Cualquier persona que se sepa que ha recibido una transfusión de sangre en los últimos 12 meses

[00115] **Resultados**

[00116] **Resumen de las muestras**

A continuación, en la Tabla 4, se muestra un resumen del estado de validez del control de calidad de todas las muestras de pacientes centrifugadas dos veces en los cuatro puntos de tiempo utilizados en este estudio.

Tabla 4.

Número de Paciente	Punto de tiempo 1	Punto de tiempo 2	Punto de tiempo 3	Punto de tiempo 4
M029	ND	ND	ND	Válida
M030	ND	ND	ND	Válida
M031	ND	ND	Válida	Válida
M032	ND	ND	Válida	Válida
M033	Válida	Válida	Válida	Válida
M034	Válida	Válida	Válida	Válida
M035	Válida	Válida	Válida	Válida
M036	Válida	Válida	Válida	Válida
M037	Válida	Válida	Válida	Válida
M038	Válida	Válida	Válida	Válida
M039	Válida	Válida	Válida	Válida
M040	Validez de PB Fallida	Válida	Muestra No Disponible	Válida
M041	Válida	Válida	Válida	Válida
M042	Válida	Válida	Válida	Válida
M043	Válida	Válida	Válida	Válida
M044	Válida	Válida	Válida	Válida
M045	Válida	Válida	Válida	Válida
M046	Válida	Válida	Válida	Válida
M047	Válida	Válida	Válida	Válida
M048	Válida	Válida	Válida	Válida
M049	Válida	Válida	Válida	Válida
M050	Válida	Válida	Válida	Válida
M051	Válida	Válida	Válida	Válida
M052	Válida	Válida	Válida	Válida
M053	Válida	Válida	Válida	Válida
M054	Válida	Válida	Válida	Válida
M055	Válida	Válida	Válida	Válida
M056	Validez de PB Fallida *	Válida	Válida	Válida

Número de Paciente	Punto de tiempo 1	Punto de tiempo 2	Punto de tiempo 3	Punto de tiempo 4
M057	Válida	Válida	Válida	Válida
M058	Válida	Válida	Válida	Válida
M059	Válida	Válida	Válida	Válida
M060	Válida	Válida	Válida	Válida
M061	Válida	Válida	Válida	Válida
M062	Válida	Válida	Válida	Válida
M063	Válida	Válida	Válida	Válida
M064	Válida	Válida	Válida	Válida
M065	Válida	Válida	Válida	Válida
M066	Válida	Válida	Válida	Válida
M067	Válida	Válida	Válida	Válida
M068	Válida	Válida	Validez de Secuenciación Fallida	Válida
M069	Válida	Válida	Válida	Válida
M070	Válida	Válida	Válida	Válida
M071	Válida	Válida	Válida	Muestra No Disponible
M072	Válida	Válida	Válida	Válida
M073	Válida	Válida	Válida	Válida
M074	Válida	Válida	Válida	Válida
M075	Válida	Válida	Válida	Válida
M076	Válida	Válida	Validez de Secuenciación Fallida	Válida
M077	Válida	Válida	Válida	Válida
M078	Válida	Válida	Válida	Válida
M079	Válida	Muestra No Disponible	Válida	Válida
M080	Válida	Válida	Válida	Validez de Secuenciación Fallida
M081	Válida	Válida	Válida	Muestra No Disponible

Número de Paciente	Punto de tiempo 1	Punto de tiempo 2	Punto de tiempo 3	Punto de tiempo 4
M082	Válida	Muestra No Disponible	Válida	Válida
M083	Válida	Válida	Válida	Válida
M084	Muestra No Disponible	Válida	Válida	Válida
M085	Válida	Válida	Válida	Validez de Secuenciación Fallida
M086	Válida	Válida	Válida	Válida
M087	Válida	Válida	Válida	Válida
M088	<i>Omitida</i>	<i>Omitida</i>	<i>Omitida</i>	<i>Omitida</i>
M089	Válida	Validez de PB Fallida	Válida	Validez de Secuenciación Fallida
M090	Validez de PB Fallida	Válida	Válida	Válida
M091	Válida	Válida	Válida	Válida
M092	Validez de PB Fallida	Validez de PB Fallida	Válida	Válida
M093	Válida	Válida	Válida	Válida
M094	Válida	Válida	Válida	Válida
M095	Válida	Válida	Válida	Válida
M096	Válida	Válida	Válida	Válida
M097	Válida	Válida	Válida	Válida
M098	Válida	Válida	Válida	Válida
M099	Validez de PB Fallida	Válida	Válida	Válida
M100	Válida	Válida	Válida	Válida
M101	Validez de PB Fallida	Válida	Validez de Secuenciación Fallida	Validez de Secuenciación Fallida
M102	Válida	Válida	Válida	Validez de Secuenciación Fallida
M103	Válida	Válida	Válida	Válida

Número de Paciente	Punto de tiempo 1	Punto de tiempo 2	Punto de tiempo 3	Punto de tiempo 4
M104	Validez de PB Fallida	Válida	Válida	Válida
M105	Válida	Válida	Válida	Válida

*Tabla 1: Resumen de todas las muestras de pacientes, puntos de tiempo y estado de validez de la muestra, incluida la validez de la muestra de una nueva prueba, de todas las muestras centrifugadas dos veces utilizadas en este estudio. PB = Preparación de la Biblioteca; * = M056 (Punto de tiempo 1) se agrupó y se seleccionó el tamaño en la Serie 5 a pesar de no ser válida como se detalla en la Sección 8.4; Muestra No Disponible = Muestra no recibida para procesar; Omitida = Muestras centrifugadas una sola vez; Omitida por otro motivo.*

[00117] Validez de la Biblioteca

[00118] Se calcularon las tasas de validez de preparación de la biblioteca (:PB) para cada serie y para cada punto de tiempo respectivo, donde las bibliotecas con un valor de cuantificación posterior a la PCR de $\geq 2,5$ ng/ μ L se consideraron válidas.

[00119] La validez inicial general de la biblioteca y la tasa de validez después de una nueva prueba para todas las muestras en cada punto de tiempo del estudio se muestran a continuación en la Tabla 5.

	Punto de tiempo 1	Punto de tiempo 2	Punto de tiempo 3	Punto de tiempo 4	General
Bibliotecas válidas iniciales	60/70	65/68	71/72	74/74	270/284
Tasa de validez inicial	85,71%	95,59%	98,61%	100,00%	95,07%
Bibliotecas válidas después de una nueva prueba	63/70	66/68	72/72	-	275/284

	Punto de tiempo 1	Punto de tiempo 2	Punto de tiempo 3	Punto de tiempo 4	General
Tasa de validez después de una nueva prueba	90,00%	97,06%	100,00%	100,00%	96,83%

[00120]

[00121] *Tabla 5: Validez inicial de la biblioteca y tasas de validez después de una nueva prueba para cada punto de tiempo y la tasa de validez general como parte de este estudio. Punto de tiempo 1 = 0-8 horas, Punto de tiempo 2 = 72 horas, Punto de tiempo 3 = 120 horas y Punto de tiempo 4 = 168 horas*

[00122] Se realizaron estadísticas de resumen (medias y desviaciones estándar) y una prueba ANOVA de Welch para las 275 bibliotecas válidas después de una nueva prueba (mencionada en 5) más la muestra M056 (Punto de tiempo 1) de la Serie 5 que se incluyó en el grupo de muestras para la selección de tamaño debido a que estaba cerca de cumplir con los criterios de validez de biblioteca (2,49 ng/μL). Los datos para las 276 muestras se muestran en la Figura 6.

[00123] Se realizó una prueba ANOVA de Welch en lugar de ANOVA ya que la prueba de varianzas desiguales demostró que las desviaciones estándar para los datos de cada punto de tiempo no eran iguales (ver figura 6); como tal, ANOVA no fue un método de análisis apropiado ya que esta prueba supone varianzas iguales.

[00124] Visualmente, los datos demuestran una concentración de biblioteca más alta desplazada concomitantemente con un tiempo de almacenamiento de muestra extendido en tubos K2EDTA antes de la centrifugación. El Punto de tiempo 4 demostró la concentración de biblioteca media más alta (11,31 ng/μL, Figura 6), mientras que el Punto de tiempo 1 demostró la concentración media más baja (6,12 ng/μL, Figura 6). El Punto de tiempo 2 presentó la desviación estándar más grande (DE = 2,90, Figura 6), y el Punto de tiempo 4 presentó la desviación estándar más pequeña (DE = 1,48, Figura 6). El Punto de tiempo 1, el Punto de tiempo 2 y el Punto de tiempo 3 tuvieron una desviación estándar de 2,48, 2,90 y 2,35 respectivamente, lo que sugiere que hay una varianza comparable entre estos Puntos de tiempo. Estas observaciones son consistentes con la hipótesis de que el tiempo de almacenamiento extendido de las muestras en los tubos EDTA causa un aumento

en la entrada de ADN en la reacción de preparación de la biblioteca; por lo tanto, se esperaba que los Puntos de tiempo posteriores devuelvan concentraciones de biblioteca más altas.

[00125] El valor $p < 0,0001$ en la ANOVA de Welch en la Figura 6 indica que hay una variación estadísticamente significativa en los datos de cuantificación de la biblioteca entre los Puntos de tiempo.

[00126] Cabe señalar que, en teoría, cualquier fragmento de ADN más grande (>1 kb) de la muestra de partida puede conservarse durante los pasos de limpieza de la muestra en el flujo de trabajo de preparación de la biblioteca basándose únicamente en el tamaño e independientemente del éxito de la ligadura/amplificación del adaptador. Por lo tanto, dicho material puede contribuir a la señal de fluorescencia en el ensayo de cuantificación fluorescente de la biblioteca, lo que contribuye a una mayor concentración de la biblioteca para Puntos de tiempo posteriores y a un tiempo de almacenamiento de la muestra más prolongado en el estudio (Figura 6).

[00127] Los datos de concentración de la biblioteca para los 276 puntos de datos indicados anteriormente se procesaron a través de un análisis de efectos principales utilizando la evaluación de calidad de datos del indicador de variabilidad en JMP para estimar qué proporción de la variabilidad en los datos de concentración de la biblioteca era atribuible a la variable de estudio (tiempo de almacenamiento en EDTA) y las series realizadas. Los resultados de los componentes de varianza de esta prueba se muestran a continuación en la Tabla 6.

Tabla 6.

Componentes de Varianza				
Componente	Componente de Var	% del total	20 40 60 80	RC(Comp VAR)
Serie	1.140335	10.0		1.0679
Punto de tiempo	5.668879	50.0		2.3809
Dentro de	4.537492	40.0		2.1301
Total	11.346706	100.0		3.3685

Tabla 6. Análisis de componentes de varianza de los datos de concentración de la biblioteca por número de serie y Punto de tiempo para las 276 muestras analizadas en el estudio.

[00128] El análisis de los componentes de varianza mostró que la variable de estudio – tiempo de almacenamiento en EDTA (Punto de tiempo) – contribuyó con el 50% de la varianza observada en los datos de concentración de la biblioteca (Tabla 6). Por lo tanto, el 50% de la varianza observada en los datos de concentración de la biblioteca proviene de otras variables del flujo de

trabajo; el 10% de esta varianza se atribuye a la variabilidad de una serie a otra, y el 40% restante se atribuye a variables no investigadas específicamente en este análisis/estudio (Tabla 6). Las posibles variables pueden incluir, entre otras, múltiples operadores que realizan la centrifugación de la muestra y el aislamiento del plasma en los diferentes puntos de tiempo, la variabilidad del laboratorio/flujo de trabajo en las Series 1-9 (p. ej., Series del Laboratorio de Servicio 1-3 frente a Series de I+D 4-9), así como la variabilidad relacionada con los propios pacientes de los que se recolectaron las muestras.

[00129] Validez de la secuenciación

[00130] Para evaluar la calidad de la muestra de ADN y omitir cualquier muestra que tuviera un problema con el proceso de secuenciación que afectara la calidad de los resultados de la secuencia, cada muestra secuenciada se analizó para varios parámetros, principalmente, control de ejecución y control de calidad de secuenciación de toda la serie, recuento de fragmentos, perfil de tamaño de fragmentos (en comparación con lo que se espera para la biblioteca de secuenciación de ADNcf seleccionada por tamaño), perfil GC, fracción fetal y controles de consistencia. Todas las muestras deben tener al menos $\geq 2\%$ de fracción fetal. Además, todas las muestras con riesgo de un resultado falso negativo o falso positivo se evalúan utilizando nuestra evaluación de fracción fetal dinámica propietaria. Esto adapta el nivel de fracción fetal requerido para la muestra a la calidad y cantidad de los datos de soporte de la secuenciación.

[00131] La tasa de validez de secuenciación general para todas las muestras analizadas (11 muestras tuvieron una nueva prueba) fue de 272/274 (99,27%). Once (11) muestras que inicialmente no superaron la validez de muestra se volvieron a analizar en una serie posterior; nueve (9) de estas muestras arrojaron un resultado válido en la nueva prueba.

[00132] Las tasas de validez de la secuencia de prueba inicial y de la prueba repetida 1x en cada punto de tiempo se presentan a continuación en la Tabla 7.

Tabla 7.

	Punto de tiempo 1	Punto de tiempo 2	Punto de tiempo 3	Punto de tiempo 4	General
Validez de la secuenciación	64/64	66/66	69/72	60/74	259/276
Tasa de validez de la secuenciación	100%	100%	95,83%	81,08%	93,84%

	Punto de tiempo 1	Punto de tiempo 2	Punto de tiempo 3	Punto de tiempo 4	General
Validez de la secuenciación después de una nueva prueba	64/64	66/66	69/72	69/74	268/276
Tasa de validez de la secuenciación después de una nueva prueba	100%	100%	95,83%	93,24%	97,10%

Tabla 7: Validez de la secuenciación inicial y tasas de repetición de la prueba 1x para cada Punto de tiempo y tasa de validez general como parte de este estudio. Punto de tiempo 1 = 0-8 horas, Punto de tiempo 2 = 72 horas, Punto de tiempo 3 = 120 horas y Punto de tiempo 4 = 168 horas

[00133] Después de una nueva prueba, un total de ocho (8) muestras fallaron las verificaciones de validez de secuenciación.

[00134] De las seis (6) muestras que no pasaron la secuenciación y que no se volvieron a analizar:

2 fallos en la secuenciación se debieron a un fallo en el perfil de tamaño de fragmento en la posición de 160 pb (M080, punto de tiempo 4 y M085, punto de tiempo 4)

4 fallos en la secuenciación se debieron a un fallo en el perfil de tamaño de fragmento en la posición de 180 pb (M089, punto de tiempo 4; M101, Punto de tiempo 3; M101, punto de tiempo 4 y M102, Punto de tiempo 4)

De las dos (2) muestras que no superaron la validez de secuenciación después de una nueva prueba:

1 fallo en una nueva prueba de secuenciación se debió a un fallo en el perfil de tamaño de fragmentos en la posición de 180 pb (M068, Punto de tiempo 3)

1 fallo en una nueva prueba de secuenciación se debió a un fallo en el perfil de tamaño de fragmentos tanto en la posición de 120 pb como en la de 180 pb (M076, Punto de tiempo 3)

[00135] Es probable que un mayor número de muestras que no superan las comprobaciones de control de calidad de secuenciación en Puntos de tiempo posteriores (en particular, la comprobación del "Perfil de tamaño de fragmento anormal") se deba al aumento de la entrada de ADN como resultado del tiempo de almacenamiento prolongado de la muestra en los tubos EDTA, lo que da como resultado una meseta de PCR y la formación heterodúplex de fragmentos de la

biblioteca en ciclos de PCR posteriores. Por lo tanto, se espera que la secuenciación de fragmentos heterodúplex de la biblioteca después del enriquecimiento de la muestra mediante la selección de tamaño (en el QS250) aumente el riesgo de que una muestra no supere las comprobaciones de control de calidad del software IONA; de hecho, el Punto de tiempo 4 tiene la mayor incidencia de muestras que no superan la validez inicial de la secuenciación y después de una nueva prueba, más comúnmente para el "Perfil de tamaño de fragmento anormal". Sin embargo, repetir la muestra a través de la preparación y secuenciación de la biblioteca pudo generar un resultado válido.

[00136] Un ejemplo del cambio en la frecuencia del tamaño de los fragmentos concomitante con un tiempo de almacenamiento prolongado antes de la centrifugación se puede observar en

[00137] dos pacientes que arrojaron un resultado válido para cada Punto de tiempo.

[00138] El cambio representativo en los perfiles de fragmentos hacia la derecha (es decir, tamaños de inserción de biblioteca de ADN más grandes) demostrado en las Figuras 7A y 7B para las muestras de pacientes M047 y M100 en los cuatro Puntos de tiempo es el resultado esperado de la formación heterodúplex; en sí mismo una consecuencia del aumento de la entrada de ADN en la preparación de la biblioteca a partir del tiempo de almacenamiento extendido de la muestra en tubos EDTA. Debe notarse que todos los fragmentos de ADN alineados representados por los perfiles en las Figuras 7A y 7B están dentro del tamaño esperado de las unidades de mononucleosoma de ADNcf y no son fragmentos de ADN esperados resultantes de la lisis celular como resultado de tiempos de almacenamiento extendidos antes de la centrifugación y el aislamiento del plasma.

[00139] *Porcentaje de fracción fetal*

[00140] El método "IONA FF" aplica uno de dos algoritmos para producir la estimación, dependiendo de los valores medidos a partir del análisis del recuento de cromosomas. IONA FF analiza las proporciones de los cromosomas X y Y. Luego...

- Si existe un **muy** alto grado de confianza en que una muestra es de feto masculino, se utiliza el algoritmo estimador 'E1', que utiliza un modelo lineal para estimar FF a partir de los datos de los cromosomas X y Y, siendo que el modelo ha sido calibrado durante el desarrollo.
- De lo contrario (es decir, si no se puede estar seguro de que se trata de un feto masculino), se utiliza el algoritmo "E2" para una muestra. Este método realiza estimaciones a partir de un modelo que relaciona las frecuencias relativas de los fragmentos largos y cortos con la

FF. También es un modelo lineal, pero se calibra por serie utilizando las estimaciones E1 (basadas en X/Y) ya calculadas. Esto aumenta la precisión de las estimaciones basadas en el tamaño en comparación con el uso de un modelo estático, al eliminar la variación entre series en el proceso de medición.

La estimación se envía luego al resto del sistema, junto con una estimación interna de la incertidumbre de la medición, que también utilizan los comprobadores de validez. Otra forma de calcular la fracción fetal es con el algoritmo conocido en el campo como SeqFF.

[00141] El porcentaje de fracción fetal se determinó a partir de los datos de las 268 muestras válidas después del análisis con el software IONA. También se realizó un ANOVA unidireccional sobre los datos de fracción fetal sin procesar (las estimaciones de fracción fetal enriquecidas obtenidas a partir de los datos de secuenciación sin procesar) para evaluar la comparación entre cada Punto de tiempo. El ANOVA de todos los datos de fracción fetal se muestra en la Figura 5.

[00142] Visualmente, los datos demuestran una tendencia a un porcentaje de fracción fetal inferior desplazado concomitantemente con un tiempo de almacenamiento prolongado antes de la centrifugación. Todos los Puntos de tiempo presentaron desviaciones estándar similares al analizar las 268 estimaciones de fracción fetal, respaldadas por la prueba de varianza desigual indicada anteriormente, lo que sugiere que existe una varianza comparable en todos estos Puntos de tiempo.

[00143] El valor p de $<0,0001$ en el ANOVA de la Figura 5 indica que hay una variación significativa en el porcentaje de fracción fetal entre los Puntos de tiempo.

[00144] No obstante, 267 de las 268 muestras tienen una fracción fetal mayor al 4%, un umbral común aplicado para las pruebas NIPT, y todas pasaron el corte IONA Nx FF.

[00145] Análisis de aneuploidías autosómicas raras (AAR)

[00146] Las 272 (100%) muestras válidas arrojaron un resultado de complemento AAR válido de "No detectada", lo que indica que no se detectó aneuploidía.

[00147] Análisis de concordancia

[00148] El análisis de concordancia en cada Punto de tiempo para muestras válidas dentro de este estudio se realizó solo cuando se obtuvo un resultado de referencia válido (Punto de tiempo 1), incluso en ausencia de un resultado para otro Punto de tiempo respectivo (es decir, resultante de un fallo de la biblioteca o de la secuenciación), donde un resultado no válido en este contexto se refiere a un resultado "No calculado" en referencia a la determinación de aneuploidía de cromosomas sexuales/ distribución del tamaño de fragmento. También se realizó un análisis de

concordancia para muestras de pacientes que obtuvieron un resultado válido para los cuatro Puntos de tiempo.

[00149] Análisis de concordancia para aneuploidías autosómicas raras

[00150] El análisis de concordancia para AAR en comparación con el valor inicial (Punto de tiempo 1) se resume a continuación en la Tabla 11. Se consideraron las muestras en las que había disponible un resultado válido de AAR en el Punto de tiempo 1 junto con al menos otro Punto de tiempo.

Tabla 11.

	Punto de tiempo 1	Punto de tiempo 2	Punto de tiempo 3	Punto de tiempo 4
Resultado válido para AAR	64	60/64	62/64	59/64
# Válido para análisis de concordancia		60	62	59
AAR concordante vs. Punto de tiempo 1	N/D	60/60	62/62	59/59
Concordante	N/D	100%	100%	100%

Tabla 11. Análisis de concordancia para AAR en comparación con el resultado de línea de base válido (Punto de tiempo 1).

[00151] Punto de tiempo 2 – 60/60 (100%) de las muestras de pacientes arrojaron un resultado de AAR “No detectada” concordante con el valor de línea de base.

[00152] Punto de tiempo 3 – 62/62 (100%) de las muestras de pacientes arrojaron un resultado de AAR “No detectada” concordante con el valor de línea de base.

[00153] Punto de tiempo 4 – 59/59 (100%) de las muestras de pacientes arrojaron un resultado de AAR “No detectada” concordante con el valor de línea de base.

[00154] El análisis de concordancia para AAR en comparación con la línea base de cincuenta y cinco (55) muestras de pacientes que arrojaron un resultado válido para los cuatro Puntos de tiempo se resume a continuación en la Tabla 12.

Tabla 12.

	Punto de tiempo 1	Punto de tiempo 2	Punto de tiempo 3	Punto de tiempo 4
Resultado válido para AAR	55	55/55	55/55	55/55
Concordancia	N/A	100%	100%	100%

Tabla 12. Análisis de concordancia para AAR en comparación con resultados de línea de base válidos de cincuenta y cinco (55) muestras de pacientes que arrojaron un resultado válido para los cuatro Puntos de tiempo

[00155] Punto de tiempo 2 – 55/55 (100%) de las muestras arrojaron un resultado de ARR “No detectada” concordante con la línea base.

[00156] Punto de tiempo 3 – 55/55 (100%) de las muestras arrojaron un resultado de ARR “No detectada” concordante con la línea base.

[00157] Punto de tiempo 4 – 55/55 (100%) de las muestras arrojaron un resultado de ARR “No detectada” concordante con la línea base.

[00158] T13/18/21 Análisis de Concordancia

[00159] El análisis de concordancia para aneuploidía T13/18/21 en comparación con la línea de base (Punto de tiempo 1) se resume a continuación en la Tabla 13.

[00160]

Tabla 13.

	Punto de tiempo 1	Punto de tiempo 2	Punto de tiempo 3	Punto de tiempo 4
Resultado T13/18/21 Válido	64	60/60	62/62	59/59

	Punto de tiempo 1	Punto de tiempo 2	Punto de tiempo 3	Punto de tiempo 4
# Válido para análisis de concordancia		60	62	59
T13/18/21 Concordante vs. Punto de tiempo 1	N/A	60/60	62/62	59/59
Concordancia	N/A	100%	100%	100%

Tabla 13. Análisis de concordancia para aneuploidía T13/18/21 en comparación con la línea de base (Punto de tiempo 1)

[00161] Punto de tiempo 2 – 60/60 (100%) las muestras de pacientes arrojaron un resultado de “Riesgo bajo” concordante con el valor de línea de base.

[00162] Punto de tiempo 3 – 62/62 (100%) las muestras de pacientes arrojaron un resultado de “Riesgo bajo” concordante con el valor de línea de base.

[00163] Punto de tiempo 4 – 59/59 (100%) las muestras de pacientes arrojaron un resultado de “Riesgo bajo” concordante con el valor de línea de base.

[00164] El análisis de concordancia para aneuploidía T13/18/21 comparado con la línea base de cincuenta y cinco (55) muestras de pacientes que arrojaron un resultado válido para los cuatro Puntos de tiempo se resume a continuación en la Tabla 14.

Tabla 14

	Punto de tiempo 1	Punto de tiempo 2	Punto de tiempo 3	Punto de tiempo 4
Resultado T13/18/21 Válido	55/55	55/55	55/55	55/55

	Punto de tiempo 1	Punto de tiempo 2	Punto de tiempo 3	Punto de tiempo 4
Válida para análisis de concordancia	N/A	55	55	55
Concordancia	N/A	100%	100%	100%

Table 14. Análisis de concordancia para aneuploidía T13/18/21 comparado con la línea base de cincuenta y cinco (55) muestras de pacientes que arrojaron un resultado válido todos los Puntos de tiempo.

[00165] Punto de tiempo 2 – 55/55 (100%) de las muestras arrojaron un resultado de “Riesgo bajo” concordante con el valor de línea de base.

[00166] Punto de tiempo 3 – 55/55 (100%) de las muestras arrojaron un resultado de “Riesgo bajo” concordante con el valor de línea de base.

[00167] (100%) de las muestras arrojaron un resultado de “Riesgo bajo” concordante con el valor de línea de base.

Ejemplo 4

[00168] Este ejemplo proporciona una realización de un parámetro de una comprobación de validez de secuencia que comprueba el perfil de tamaño de fragmento y determina si el perfil es normal o anormal. El propósito de esta comprobación es reducir el riesgo de degradación del rendimiento en una prueba NIPT debido a la degradación de la muestra y/o la eficacia reducida de un paso de selección del tamaño del ADN fetal (también denominado enriquecimiento del ADN fetal), detectando casos de enriquecimiento insuficiente e invalidando los análisis de muestra correspondientes de modo que no produzcan un resultado de prueba. La prueba también se puede utilizar para comprobar secuencias de células cancerosas en el ADN libre de células.

[00169] Se sabe que el ADN fetal libre de células está presente en un rango de tamaño específico que es diferente del de ADN materno libre de células. Por lo tanto, es posible detectar casos de enriquecimiento ineficaz y/o degradación excesiva de la muestra calculando un perfil de tamaño de fragmento normalizado a partir de datos de secuenciación asociados con una muestra y comparando sus parámetros con un conjunto de parámetros de referencia precalculados que

pueden distinguir el enriquecimiento efectivo del ineficaz y/o la degradación excesiva de la muestra.

[00170] La comprobación de validez del perfil de tamaño está diseñada para garantizar, por ejemplo, que:

1. El enriquecimiento del componente de ADN fetal libre de células del ADN materno libre de células es eficaz; el enriquecimiento se puede aplicar a través de la selección de tamaño como se describe en este documento, asegurando que los fragmentos en rangos de tamaño con más probabilidades de ser fetales se seleccionen preferentemente y/o
2. Los datos de secuenciación de la muestra poseen un perfil de tamaño de fragmentos de ADN consistente con las características biológicas del material de ADN libre de células extraído del plasma de una mujer embarazada.

[00171] Se definen una o más “*envolturas de tamaño*” para valores de tamaño de fragmentos particulares de interés. (En el Ejemplo 3, los valores de tamaño de fragmentos de interés fueron fragmentos de 120 pb y 180 pb). Esencialmente, cada envoltura encierra un perfil de tamaño de fragmento en aquellos puntos que se considera válido. Por lo tanto, si un perfil de tamaño de fragmento entrante cumple las condiciones para ser encerrado por al menos una de las envolturas así definidas, se considera que la Comprobación de Validez de Perfil de Tamaño ha sido superada para la muestra correspondiente.

[00172] Parámetros de configuración: Los siguientes parámetros son necesarios para configurar esta realización del parámetro de comprobación de validez. Deben almacenarse en una representación estructurada jerárquica eficiente (por ejemplo, utilizando un formato basado en XML).

- **Número de envolturas de tamaño definido** – N_{env} (entero, obligatorio). 108,
- **Matriz de tamaños de envolturas** – $SizeEnv[.]$. Para cada entrada:
 - **Número de puntos de control de tamaño** – N_{chk} (entero, obligatorio).
 - **Matriz de puntos de control de tamaño** – $SizeChk[.]$. Para cada entrada:
 - Tamaño del fragmento para el cual se definió – s_{chk} (entero, obligatorio).
 - Rango de frecuencia relativa válido – límite inferior – z_L (real, opcional).
Nota: Siempre que se proporcione un valor para z_U , este valor puede omitirse, en cuyo caso el límite inferior se tomará como cero.
 - Rango de frecuencia relativa válido – límite superior – z_U (real, opcional).

Nota: Siempre que se proporcione un valor para z_L , este valor puede omitirse, en cuyo caso el límite superior se tomará como uno.

[00173] *Datos de entrada de prueba:* Los datos de entrada son una matriz de valores de tamaño de fragmentos medidos en datos de secuenciación.

[00174] *Salida de prueba:* El método produce una única salida, de la siguiente manera: Estado de Validez del Perfil de Tamaño – V_{SP} (booleano, obligatorio).

[00175] *Cálculo del perfil de tamaño de muestra:* En primer lugar, se forma una distribución de frecuencia relativa a partir de la matriz de valores de medición de tamaño de fragmento producidos para la muestra, tal como se genera mediante el proceso de análisis central a través de un método apropiado para la plataforma de secuenciación en uso. Esta distribución se debe calcular como una matriz compuesta por la frecuencia de cada tamaño que se produce, normalizada de acuerdo con el número total de fragmentos incluidos en la distribución, de este modo:

$$Z[s] = \frac{n_s}{\sum_{j \in S} n_j}$$

en donde n_s es el número de fragmentos presentes de tamaño s para todos los tamaños de fragmentos $s \in S$, y S es el conjunto de todos los tamaños de fragmentos medidos a partir de los datos de secuenciación, donde se considera que los fragmentos contribuyentes son solo aquellos que (después de la deduplicación) se han alineado de forma única con un cromosoma distinto de 13, 18, 21, X y Y. Se anota que cuando se encontró que N fragmentos se alineaban en la misma ubicación genómica y, por lo tanto, en el mismo lugar de la herramienta de deduplicación, cada fragmento de tamaño s contribuirá con $1/N$ al recuento de fragmentos n_s .

[00176] Para cualquier valor de tamaño no representado (es decir, $s \notin S$), el valor de distribución de frecuencia relativa correspondiente $Z[s]$ se toma como cero.

[00177] *Cálculo del estado de validez de la muestra:* La distribución de frecuencias relativas Z calculada para la muestra se compara con cada una de las envolturas definidas en el conjunto de parámetros de configuración. Para cada envoltura, esta comparación se realiza de la siguiente manera:

1. En cada punto de control de tamaño, se utiliza el valor de tamaño correspondiente s_{chk} para buscar un valor de frecuencia relativa en ese tamaño en la distribución calculada para la muestra, es decir: $Z[s_{chk}]$.
2. El valor de distribución así obtenido se compara con los parámetros de rango de frecuencia válidos, z_L y z_U . Si no cae dentro del intervalo $[z_L, z_U)$, es decir: $Z[s_{chk}] < z_L$ o $Z[s_{chk}] \geq z_U$, entonces la comparación del punto de control se marca como fallida. Si el valor de distribución sí cae dentro del intervalo $[z_L, z_U)$, la comparación está marcada como superada.

[00178] La comparación continúa hasta el siguiente punto de control de tamaño en la envoltura actual hasta que no quede ninguno.

[00179] A continuación, se calcula un resultado de comparación para la envoltura como el "y" lógico de todos los resultados de los puntos de control constituyentes. Es decir, se considera que una comparación de envolturas ha sido superada solo si ha sido superada para todos los puntos de control constituyentes.

[00180] Si se determina que una sola comparación de envolturas ha sido superada, se genera un resultado válido para la validez de la muestra a través de la Comprobación de Validez del Perfil de Tamaño, y finaliza con un valor $V_{SP} = \text{verdadero}$. Solo se requiere que una comparación de envolturas sea superada para que la comprobación de validez en su conjunto dé un resultado válido.

[00181] Si, después de que se hayan completado todas las comparaciones de envolturas, se encuentra que ninguna ha superado todos los puntos de control constituyentes, se genera un resultado no válido para la validez de la muestra a través de la Comprobación de Validez del Perfil de Tamaño, y finaliza con un valor $V_{SP} = \text{falso}$. Esto equivale a un perfil de tamaño de fragmento anormal y la muestra no supera la verificación de validez de la secuencia y se rechaza. Es necesario volver a recolectar sangre.

Equivalentes

[00182] Aunque la invención se describe en detalle con referencia a realizaciones específicas de la misma, se entenderá que las variaciones que son funcionalmente equivalentes están dentro del alcance de esta invención. De hecho, varias modificaciones de la invención además de las mostradas y descritas en este documento resultarán evidentes para los expertos en la materia a partir de la descripción anterior y los dibujos adjuntos. Se pretende que dichas modificaciones

estén dentro del alcance de las reivindicaciones adjuntas. Los expertos en la materia reconocerán, o podrán determinar utilizando únicamente la experimentación rutinaria, muchos equivalentes a las realizaciones específicas de la invención descritas en este documento. Se pretende que dichos equivalentes estén abarcados por las siguientes reivindicaciones.

[00183] Todas las publicaciones mencionadas en esta especificación se incorporan aquí por referencia a la especificación, en la misma medida que si cada publicación individual se hubiera indicado específica e individualmente para ser incorporada aquí por referencia en su totalidad.

[00184] Se pueden realizar muchas modificaciones y variaciones de esta invención sin apartarse de su espíritu y alcance, como será evidente para los expertos en la materia. Las realizaciones específicas descritas en este documento se ofrecen únicamente a modo de ejemplo, y la invención debe estar limitada únicamente por los términos de las reivindicaciones adjuntas, junto con el alcance completo de los equivalentes a los que dichas reivindicaciones tienen derecho.