

LOS PATRONES DE FRAGMENTACIÓN DE ADN LIBRE DE CÉLULAS DERIVADOS DE DELFI DIFERENCIAN SUBTIPOS HISTOLÓGICOS DE CÁNCERES DE PULMÓN DE FORMA NO INVASIVA

REFERENCIA CRUZADA A SOLICITUDES RELACIONADAS

[0001] La presente solicitud reivindica el beneficio de prioridad conforme al artículo 119(e) del título 35 del U.S.C. de la solicitud provisional de patente estadounidense n.º de serie 63/445,284 presentada el 13 de febrero de 2023, la solicitud provisional de patente estadounidense n.º de serie 63/470,101 presentada el 31 de mayo de 2023, y la solicitud provisional de patente estadounidense n.º de serie 63/528,237 presentada el 21 de julio de 2023. Las descripciones de las solicitudes anteriores se consideran parte de la descripción de la presente solicitud y se incorporan mediante referencia a la presente solicitud en su totalidad.

CAMPO DE LA INVENCION

[0002] La presente invención se refiere en general al diagnóstico no invasivo y la subtipificación del cáncer de pulmón de células pequeñas (SCLC) y, más específicamente, con el análisis de patrones genómicos de ADN libre de células fragmentado (cfDNA) junto con características clínicas y demográficas de pacientes individuales. La presente invención también se refiere al perfil de fragmentación del ADN libre de células como método para la evaluación de la fracción tumoral y la monitorización del tratamiento en el cáncer de pulmón no microcítico (NSCLC). Dadas las dificultades logísticas para realizar biopsias de tumores SCLC y NSCLC, se necesitan nuevos enfoques para obtener métodos viables y rápidos de subtipificación de SCLC y NSCLC de forma no invasiva.

ANTECEDENTES

[0003] El cáncer de pulmón microcítico es una neoplasia agresiva con pronóstico desfavorable. Aunque los cánceres de pulmón microcítico se gestionan clínicamente como un único tipo de cáncer, nuevas pruebas apoyan que estos subtipos (neuroendócrino de alto grado frente a neuroendócrino de bajo grado) de cáncer de pulmón microcítico adquieren diversos estados transcripcionales y epigenéticos. Además, los distintos subtipos de cáncer de pulmón microcítico responden a tratamientos específicos, como la inmunoterapia en el

caso del subtipo de cáncer de pulmón microcítico neuroendócrino de bajo grado y muy inflamado. El cáncer de pulmón no microcítico es cualquier tipo de cáncer de pulmón epitelial distinto del cáncer de pulmón microcítico. Los tipos más comunes de cáncer de pulmón no microcítico son el carcinoma de células escamosas, el carcinoma de células grandes y el adenocarcinoma, pero existen otros tipos que se dan con menor frecuencia, y todos los tipos pueden presentarse en variantes histológicas inusuales. Como clase, el cáncer de pulmón no microcítico suele ser menos sensible a la quimioterapia y la radioterapia que el cáncer de pulmón microcítico. Los pacientes con enfermedad resecable pueden curarse mediante cirugía o cirugía seguida de quimioterapia, así como quimioterapia seguida de cirugía. En muchos pacientes con enfermedad irresecable puede lograrse el control local con radioterapia, pero la curación se observa en relativamente pocos pacientes. Los pacientes con enfermedad localmente avanzada irresecable pueden lograr una supervivencia a largo plazo con radioterapia combinada con quimioterapia. Los pacientes con enfermedad metastásica avanzada pueden lograr una mayor supervivencia y paliar los síntomas con quimioterapia, agentes dirigidos y otras medidas de apoyo. Se necesitan nuevos métodos no invasivos para identificar y subtipificar el cáncer de pulmón no microcítico y microcítico con el fin de mejorar la evolución de los pacientes.

COMPENDIO DE LA INVENCION

[0004] La presente invención se basa en el descubrimiento seminal de que la caracterización de los patrones de fragmentación del ADN libre de células (cfDNA) en plasma a nivel de todo el genoma mediante secuenciación del genoma completo de baja cobertura puede mejorar el diagnóstico del cáncer cuando se analiza junto con determinadas características clínicas y demográficas de los pacientes individuales.

[0005] En una modalidad, la presente invención proporciona métodos para subtipificar el cáncer de pulmón microcítico en un sujeto como cáncer de pulmón microcítico neuroendócrino de bajo grado o neuroendócrino de alto grado que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de 30× a 0,1×; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de

transcripción especificados; analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de cáncer de pulmón microcítico neuroendócrino de alto grado.

[0006] En una modalidad, la presente invención proporciona métodos para subtipificar el cáncer de pulmón microcítico en un sujeto como cáncer de pulmón microcítico neuroendócrino de bajo grado o neuroendócrino de alto grado que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de $10\times$ a $0,1\times$; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados; analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de cáncer de pulmón microcítico neuroendócrino de alto grado.

[0007] En una modalidad, la presente invención proporciona métodos para subtipificar el cáncer de pulmón microcítico en un sujeto como cáncer de pulmón microcítico neuroendócrino de bajo grado o neuroendócrino de alto grado que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de

alrededor de $9\times$ a $0,1\times$; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados; analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de cáncer de pulmón microcítico neuroendócrino de alto grado. En algunos aspectos, el método utiliza el factor de transcripción especificado ASCL1, NEUROD1, POUF23, YAP1, o cualquier combinación de estos.

[0008] En algunos aspectos, el método utiliza el aprendizaje automático para subtipificar el cáncer de pulmón microcítico en un sujeto.

[0009] En algunos aspectos, la subtipificación se realiza al calcular $\log_2$ (relación de profundidad de lectura) a través de ventanas de 20 pb que comienzan a 2 kb de los sitios de unión del factor de transcripción especificados.

[0010] En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) es la cobertura total sobre un factor de corrección de cobertura calculado mediante el uso de la cobertura mediana de anclajes de 5' y 3' 500pb en los extremos de la ventana de 2kb (es decir, $\log_2(\text{profundidad de lectura}/\text{factor de corrección})$).

[0011] En algunos aspectos, el método utiliza todos los cálculos de cobertura se desplazan por 1 para evitar divisiones por 0. En algunos aspectos, este ajuste se expresa como $f(x) = \log_2((\text{profundidad de la ventana} + 1)/(\text{mediana}(\text{anclajes}) + 1)$.

[0012] En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) se calcula para cada sitio de unión del factor de transcripción y luego se agrega al calcular la mediana de cada ventana para obtener una puntuación de cobertura del fragmento.

[0013] Algunos aspectos comprenden además el uso de un modelo aditivo general para suavizar la cobertura y corregir el sesgo de GC.

[0014] En algunos aspectos, los intervalos genómicos no se solapan.

[0015] En algunos aspectos, cada uno de los intervalos genómicos comprende de miles a millones de pares de bases.

[0016] En algunos aspectos, se determina un perfil de fragmentación del cfDNA dentro de cada intervalo genómico.

[0017] En algunos aspectos, el perfil de fragmentación del cfDNA comprende un tamaño de fragmento mediano.

[0018] En algunos aspectos, el perfil de fragmentación del cfDNA comprende una distribución del tamaño de fragmento.

[0019] Algunos aspectos comprenden además administrarle al sujeto que se identificó que tiene un subtipo de cáncer de pulmón microcítico neuroendócrino de alto grado un agente terapéutico adecuado para el tratamiento del tipo de cáncer.

[0020] En algunos aspectos, el agente terapéutico es una inmunoterapia.

[0021] En una modalidad, la presente invención proporciona métodos para la subtipificación no invasiva del cáncer de pulmón no microcítico en un sujeto como adenocarcinoma o carcinoma escamoso que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de $30\times$ a $0,1\times$; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados; analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de adenocarcinoma o carcinoma de células escamosas.

[0022] En una modalidad, la presente invención proporciona métodos para la subtipificación no invasiva del cáncer de pulmón no microcítico en un sujeto como adenocarcinoma o carcinoma escamoso que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de $10\times$ a $0,1\times$;

mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados; analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de adenocarcinoma o carcinoma de células escamosas.

[0023] En una modalidad, la presente invención proporciona métodos para la subtipificación no invasiva del cáncer de pulmón no microcítico en un sujeto como adenocarcinoma o carcinoma escamoso que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de $9\times$ a $0,1\times$; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados; analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de adenocarcinoma o carcinoma de células escamosas.

[0024] En algunos aspectos, el factor de transcripción especificado ASCL1, NEUROD1, POUF23, YAP1, o cualquier combinación de estos.

[0025] En algunos aspectos, el método utiliza el aprendizaje automático para subtipificar el cáncer de pulmón microcítico en el sujeto.

[0026] En algunos aspectos, la subtipificación se realiza al calcular $\log_2$ (relación de profundidad de lectura) a través de ventanas de 20 pb que comienzan a 2 kb de los sitios

de unión del factor de transcripción especificados.

[0027] En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) es la cobertura total sobre un factor de corrección de cobertura calculado mediante el uso de la cobertura mediana de anclajes de 5' y 3' 500pb en los extremos de la ventana de 2kb (es decir, $\log_2(\text{profundidad de lectura}/\text{factor de corrección})$).

[0028] En algunos aspectos, todos los cálculos de cobertura se desplazan por 1 para evitar divisiones por 0. En algunos aspectos, este ajuste se expresa como $f(x) = \log_2((\text{profundidad de la ventana} + 1)/(\text{mediana}(\text{anclajes}) + 1))$.

[0029] En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) se calcula para cada sitio de unión del factor de transcripción y luego se agrega al calcular la mediana de cada ventana para obtener una puntuación de cobertura del fragmento.

[0030] Algunos aspectos comprenden además el uso de un modelo aditivo general para suavizar la cobertura y corregir el sesgo de GC.

[0031] En algunos aspectos, los intervalos genómicos no se solapan.

[0032] En algunos aspectos, cada uno de los intervalos genómicos comprende de miles a millones de pares de bases.

[0033] En algunos aspectos, se determina un perfil de fragmentación del cfDNA dentro de cada intervalo genómico.

[0034] En algunos aspectos, el perfil de fragmentación del cfDNA comprende un tamaño de fragmento mediano.

[0035] En algunos aspectos, el perfil de fragmentación del cfDNA comprende una distribución del tamaño de fragmento.

[0036] Algunos aspectos comprenden además administrarle al sujeto que se identificó que tiene un subtipo de adenocarcinoma o carcinoma escamoso un agente terapéutico adecuado para el tratamiento del tipo de cáncer.

[0037] En algunos aspectos, el agente terapéutico es una inmunoterapia.

[0038] Algunos aspectos comprenden además la cuantificación de una fracción tumoral circulante, que refleja el rendimiento de la fracción alélica mayor (MAF) en relación con la evaluación de los criterios de evaluación de la respuesta en tumores sólidos (RECIST) tanto en adenocarcinoma como en carcinoma escamoso.

BREVE DESCRIPCIÓN DE LAS FIGURAS

[0039] La **Figura 1** ilustra las puntuaciones DELFI de los pacientes diagnosticados con SCLC en el estudio del Ejemplo 1.

[0040] La **Figura 2** ilustra que los perfiles de fragmentación de todo el genoma fueron significativamente consistentes entre los puntos temporales de pretratamiento, postratamiento, progresión y respuesta, lo que sugiere que el tamaño de los fragmentos de ADN tumoral circulante de todo el genoma es un método potente para detectar cambios durante el monitoreo del tratamiento de inmunoterapia del cáncer SCLC.

[0041] La **Figura 3** ilustra los perfiles de fragmentación de todo el genoma (y las puntuaciones DELFI correspondientes) que muestran diferencias parciales entre los dos subtipos principales de casos de SCLC: Neuroendócrino de alto grado y neuroendócrino de bajo grado; lo que sugiere que el tamaño de los fragmentos de ADN tumoral circulante en todo el genoma puede ser un método potente para detectar la subtipificación durante el seguimiento del tratamiento con inmunoterapia del cáncer SCLC.

[0042] La **Figura 4** ilustra los genes diferenciales entre los casos de SCLC neuroendócrino de alto grado y de bajo grado, mediante el uso de datos disponibles públicamente (Lissa et al. 2022) y un Análisis de Componentes Principales (sobre datos WGS de plasma DELFI 30X) para agrupar los dos subtipos. Los datos del análisis PCA revelaron grupos distinguibles entre las muestras pretratamiento de SCLC NE alto frente a NE bajo.

[0043] Las **Figuras 5A-5B** ilustran un análisis Eigencor que revela una correlación significativa entre los componentes principales 1 y 2 con el estado NE de las muestras.

[0044] Las **Figuras 6A-6B** ilustran el potencial del ensayo DELFI para subtipificar el SCLC mediante el uso de TFBS, se realizaron análisis de fragmentación del cfDNA de todo el genoma en los sitios de unión de ASCL1 en las muestras de NCI pretratamiento. La Figura 5A ilustra que se observaron distintos grupos de muestras de SCLC correspondientes a diferentes niveles de activación de ASCL1. Es importante destacar que los dos grupos correspondían perfectamente a los diferentes subtipos de SCLC. El principal factor de diferenciación entre las muestras fueron, de hecho, las diferencias en la cobertura de los fragmentos identificados en el centro de los sitios de unión a ASCL1. La Figura 6B ilustra que otras características clínicas o de la muestra tienen un impacto insignificante sobre la agrupación de estas muestras. En general, estos datos muestran cómo el análisis DELFI

puede distinguir perfectamente entre los subtipos de SCLC mediante el uso de la cobertura de fragmentos en TFBS.

[0045] La **Figura 7** ilustra un paciente ilustrativo que apoya la hipótesis de que la señal detectada procede de linfocitos infiltrantes de tumores. El paciente NCI-0422, diagnosticado de SCLC con subtipo inflamatorio fue tratado con Durvalumab más Olaparib. NCI-0422 respondió bien al tratamiento, pero posteriormente se le diagnosticó enfermedad progresiva. La unión genómica variable es detectable antes del tratamiento y durante la progresión. Los picos se centran en una mezcla de TFBS y TSS.

[0046] Las **Figuras 8A-8B** ilustran ASCL1 sitio de unión entre respondedores frente a no respondedores. La Figura 8A muestra 500 bins específicos de tipo celular de los datos de PMD para determinar la fracción de glóbulos blancos.

[0047] La **Figura 9** ilustra un modelo de aprendizaje automático para predecir un SCLC neuroendócrino de alto grado frente a de bajo grado.

[0048] Las **Figuras 10A-10B** ilustra que el método descrito en la presente es capaz de diferenciar neuroendócrino frente a no neuroendócrino mediante el cálculo de relación de profundidad de lectura en coordenadas genómicas específicas de SCLC (Figura 10A). La Figura 10B muestra que el método descrito en la presente parece subtipificar mejor los casos de SCLC.

[0049] La **Figura 11** ilustra que la relación de profundidad de lectura en coordenadas genómicas específicas de SCLC se transforman en un PCA bidimensional para separar las muestras en grupos neuroendócrinos frente a no neuroendócrinos. La mayoría de las muestras están claramente divididas en distintos grupos.

[0050] La **Figura 12** ilustra PCA derivado de relación de profundidad de lectura en coordenadas genómicas específicas de SCLC es capaz de separar las muestras pretratamiento en subtipos específicos de SCLC (A,N,P,Y). La verdad fundamental de los subtipos de SCLC se calculó mediante el análisis diferencial de la expresión génica de secuenciación de ARN tisular.

[0051] La **Figura 13** ilustra representa cómo los perfiles de fragmentación informan sobre el estado del cáncer de pulmón y los patrones de respuesta al tratamiento.

[0052] La **Figura 14** ilustra que DELFI-TF derivado del modelo muestra una fuerte correlación con la Frecuencia Alélica Mutante en todas las muestras.

[0053] La **Figura 15** ilustra que DELFI-TF cuantifica con precisión la fracción tumoral

circulante, que refleja el rendimiento de MAF en relación con la evaluación RECIST.

[0054] La **Figura 16** ilustra que el cfDNA derivado de tumores presenta una fragmentación alterada.

[0055] Las **Figuras 17A-17C** ilustran que DELFI-TF detecta con precisión la fracción tumoral circulante sin confundirse por la hematopoyesis clonal.

[0056] La **Figura 18** muestra los patrones de fragmentación del cfDNA para diferenciar con precisión los subtipos de NSCLC.

[0057] La **Figura 19** muestra el desarrollo del modelo DELFI-TF de prueba de concepto.

[0058] La **Figura 20** es una computadora 800 ilustrativa que puede utilizarse para implementar los métodos descritos en la presente.

DESCRIPCIÓN DETALLADA

[0059] La presente invención se basa en el descubrimiento seminal de que la caracterización de los patrones de fragmentación del ADN libre de células (cfDNA) en plasma a nivel de todo el genoma mediante secuenciación del genoma completo de baja cobertura mejora el diagnóstico del cáncer cuando se analiza junto con determinadas características clínicas y demográficas de los pacientes individuales.

[0060] En la presente se describen métodos para subtipificar el cáncer de pulmón microcítico en un sujeto como cáncer de pulmón microcítico neuroendócrino de bajo grado o neuroendócrino de alto grado que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de $30\times$ a $0,1\times$; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados; analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de

cáncer de pulmón microcítico neuroendócrino de alto grado.

[0061] En la presente se describen métodos para subtipificar el cáncer de pulmón microcítico en un sujeto como cáncer de pulmón microcítico neuroendócrino de bajo grado o neuroendócrino de alto grado que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de $10\times$ a $0,1\times$; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados; analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de cáncer de pulmón microcítico neuroendócrino de alto grado.

[0062] En la presente se describen métodos para subtipificar el cáncer de pulmón microcítico en un sujeto como cáncer de pulmón microcítico neuroendócrino de bajo grado o neuroendócrino de alto grado que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de $9\times$ a $0,1\times$; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados; analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de

cáncer de pulmón microcítico neuroendócrino de alto grado.

[0063] En algunos aspectos, el método utiliza el factor de transcripción especificado ASCL1, NEUROD1, POUF23, YAP1, o cualquier combinación de estos.

[0064] En algunos aspectos, el método utiliza el aprendizaje automático para subtipificar el cáncer de pulmón microcítico en el sujeto.

[0065] En algunos aspectos, la subtipificación se realiza al calcular $\log_2$ (relación de profundidad de lectura) a través de ventanas de 20 pb que comienzan a 2 kb de los sitios de unión del factor de transcripción especificados.

[0066] En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) es la cobertura total sobre un factor de corrección de cobertura calculado mediante el uso de la cobertura mediana de anclajes de 5' y 3' 500pb en los extremos de la ventana de 2kb (es decir, $\log_2(\text{profundidad de lectura}/\text{factor de corrección})$).

[0067] En algunos aspectos, el método utiliza todos los cálculos de cobertura se desplazan por 1 para evitar divisiones por 0. En algunos aspectos, este ajuste se expresa como $f(x) = \log_2((\text{profundidad de la ventana} + 1)/(\text{mediana}(\text{anclajes}) + 1)$.

[0068] En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) se calcula para cada sitio de unión del factor de transcripción y luego se agrega al calcular la mediana de cada ventana para obtener una puntuación de cobertura del fragmento.

[0069] Algunos aspectos comprenden además el uso de un modelo aditivo general para suavizar la cobertura y corregir el sesgo de GC.

[0070] En algunos aspectos, los intervalos genómicos no se solapan.

[0071] En algunos aspectos, cada uno de los intervalos genómicos comprende de miles a millones de pares de bases.

[0072] En algunos aspectos, se determina un perfil de fragmentación del cfDNA dentro de cada intervalo genómico.

[0073] En algunos aspectos, el perfil de fragmentación del cfDNA comprende una mediana del tamaño de fragmento; en otros aspectos, el perfil de fragmentación del cfDNA comprende una distribución del tamaño de fragmento.

[0074] Algunos aspectos comprenden además administrarle al sujeto que se identificó que tiene un subtipo de cáncer de pulmón microcítico neuroendócrino de alto grado un agente terapéutico adecuado para el tratamiento del tipo de cáncer.

[0075] En algunos aspectos, el agente terapéutico es una inmunoterapia.

[0076] En la presente se describen métodos no invasivos para subtipificar el cáncer de pulmón microcítico en un sujeto como cáncer de pulmón microcítico neuroendócrino de bajo grado o neuroendócrino de alto grado que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de $30\times$ a $0,1\times$; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados;

[0077] analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de cáncer de pulmón microcítico neuroendócrino de alto grado. En algunos aspectos, el método utiliza el factor de transcripción especificado ASCL1, NEUROD1, POUF23, YAP1, o cualquier combinación de estos. En algunos aspectos, el método utiliza el aprendizaje automático para subtipificar el cáncer de pulmón microcítico en el sujeto. En algunos aspectos, la subtipificación se realiza al calcular $\log_2$ (relación de profundidad de lectura) a través de ventanas de 20 pb que comienzan a 2 kb de los sitios de unión del factor de transcripción especificados. En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) es la cobertura total sobre un factor de corrección de cobertura calculado mediante el uso de la cobertura mediana de anclajes de 5' y 3' 500pb en los extremos de la ventana de 2kb (es decir, $\log_2(\text{profundidad de lectura}/\text{factor de corrección})$). En algunos aspectos, el método utiliza todos los cálculos de cobertura se desplazan por 1 para evitar divisiones por 0. En algunos aspectos, este ajuste se expresa como $f(x) = \log_2((\text{profundidad de la ventana} + 1)/(\text{mediana}(\text{anclajes}) + 1))$. En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) se calcula para cada sitio de unión del factor de transcripción y luego se agrega al calcular la mediana de cada ventana para obtener una puntuación de cobertura del fragmento. Algunos aspectos comprenden además el uso de un modelo aditivo general para suavizar la cobertura y corregir el sesgo de GC. En algunos

aspectos, los intervalos genómicos no se solapan. En algunos aspectos, cada uno de los intervalos genómicos comprende de miles a millones de pares de bases. En algunos aspectos, se determina un perfil de fragmentación del cfDNA dentro de cada intervalo genómico. En algunos aspectos, el perfil de fragmentación del cfDNA comprende un tamaño de fragmento mediano. En algunos aspectos, el perfil de fragmentación del cfDNA comprende una distribución del tamaño de fragmento. Algunos aspectos comprenden además administrarle al sujeto que se identificó que tiene un subtipo de cáncer de pulmón microcítico neuroendócrino de alto grado un agente terapéutico adecuado para el tratamiento del tipo de cáncer. En algunos aspectos, el agente terapéutico es una inmunoterapia.

[0078] En la presente se describen métodos no invasivos para subtipificar el cáncer de pulmón microcítico en un sujeto como cáncer de pulmón microcítico neuroendócrino de bajo grado o neuroendócrino de alto grado que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de $10\times$ a $0,1\times$; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados;

[0079] analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de cáncer de pulmón microcítico neuroendócrino de alto grado. En algunos aspectos, el método utiliza el factor de transcripción especificado ASCL1, NEUROD1, POUF23, YAP1, o cualquier combinación de estos. En algunos aspectos, el método utiliza el aprendizaje automático para subtipificar el cáncer de pulmón microcítico en el sujeto. En algunos aspectos, la subtipificación se realiza al calcular $\log_2$ (relación de profundidad de lectura) a través de ventanas de 20 pb que comienzan a 2 kb de los sitios de unión del factor de transcripción especificados. En algunos aspectos, el $\log_2$ (relación de

profundidad de lectura) es la cobertura total sobre un factor de corrección de cobertura calculado mediante el uso de la cobertura mediana de anclajes de 5' y 3' 500pb en los extremos de la ventana de 2kb (es decir, $\log_2(\text{profundidad de lectura}/\text{factor de corrección})$). En algunos aspectos, el método utiliza todos los cálculos de cobertura se desplazan por 1 para evitar divisiones por 0. En algunos aspectos, este ajuste se expresa como $f(x) = \log_2((\text{profundidad de la ventana} + 1)/(\text{mediana}(\text{anclajes}) + 1))$. En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) se calcula para cada sitio de unión del factor de transcripción y luego se agrega al calcular la mediana de cada ventana para obtener una puntuación de cobertura del fragmento. Algunos aspectos comprenden además el uso de un modelo aditivo general para suavizar la cobertura y corregir el sesgo de GC. En algunos aspectos, los intervalos genómicos no se solapan. En algunos aspectos, cada uno de los intervalos genómicos comprende de miles a millones de pares de bases. En algunos aspectos, se determina un perfil de fragmentación del cfDNA dentro de cada intervalo genómico. En algunos aspectos, el perfil de fragmentación del cfDNA comprende un tamaño de fragmento mediano. En algunos aspectos, el perfil de fragmentación del cfDNA comprende una distribución del tamaño de fragmento. Algunos aspectos comprenden además administrarle al sujeto que se identificó que tiene un subtipo de cáncer de pulmón microcítico neuroendócrino de alto grado un agente terapéutico adecuado para el tratamiento del tipo de cáncer. En algunos aspectos, el agente terapéutico es una inmunoterapia.

[0080] En la presente se describen métodos no invasivos para subtipificar el cáncer de pulmón microcítico en un sujeto como cáncer de pulmón microcítico neuroendócrino de bajo grado o neuroendócrino de alto grado que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de $9\times$ a $0,1\times$; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados;

[0081] analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación

del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de cáncer de pulmón microcítico neuroendócrino de alto grado. En algunos aspectos, el método utiliza el factor de transcripción especificado ASCL1, NEUROD1, POUF23, YAP1, o cualquier combinación de estos. En algunos aspectos, el método utiliza el aprendizaje automático para subtipificar el cáncer de pulmón microcítico en el sujeto. En algunos aspectos, la subtipificación se realiza al calcular $\log_2$ (relación de profundidad de lectura) a través de ventanas de 20 pb que comienzan a 2 kb de los sitios de unión del factor de transcripción especificados. En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) es la cobertura total sobre un factor de corrección de cobertura calculado mediante el uso de la cobertura mediana de anclajes de 5' y 3' 500pb en los extremos de la ventana de 2kb (es decir, $\log_2(\text{profundidad de lectura}/\text{factor de corrección})$). En algunos aspectos, el método utiliza todos los cálculos de cobertura se desplazan por 1 para evitar divisiones por 0. En algunos aspectos, este ajuste se expresa como $f(x) = \log_2((\text{profundidad de la ventana} + 1)/(\text{mediana}(\text{anclajes}) + 1))$. En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) se calcula para cada sitio de unión del factor de transcripción y luego se agrega al calcular la mediana de cada ventana para obtener una puntuación de cobertura del fragmento. Algunos aspectos comprenden además el uso de un modelo aditivo general para suavizar la cobertura y corregir el sesgo de GC. En algunos aspectos, los intervalos genómicos no se solapan. En algunos aspectos, cada uno de los intervalos genómicos comprende de miles a millones de pares de bases. En algunos aspectos, se determina un perfil de fragmentación del cfDNA dentro de cada intervalo genómico. En algunos aspectos, el perfil de fragmentación del cfDNA comprende un tamaño de fragmento mediano. En algunos aspectos, el perfil de fragmentación del cfDNA comprende una distribución del tamaño de fragmento. Algunos aspectos comprenden además administrarle al sujeto que se identificó que tiene un subtipo de cáncer de pulmón microcítico neuroendócrino de alto grado un agente terapéutico adecuado para el tratamiento del tipo de cáncer. En algunos aspectos, el agente terapéutico es una inmunoterapia.

[0082] En una modalidad, la presente invención proporciona métodos para la subtipificación no invasiva del cáncer de pulmón no microcítico en un sujeto como

adenocarcinoma o carcinoma escamoso que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de $30\times$ a $0,1\times$; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados; analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de adenocarcinoma o carcinoma de células escamosas.

[0083] En una modalidad, la presente invención proporciona métodos para la subtipificación no invasiva del cáncer de pulmón no microcítico en un sujeto como adenocarcinoma o carcinoma escamoso que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de $10\times$ a $0,1\times$; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados; analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de adenocarcinoma o carcinoma de células escamosas.

[0084] En una modalidad, la presente invención proporciona métodos para la subtipificación no invasiva del cáncer de pulmón no microcítico en un sujeto como

adenocarcinoma o carcinoma escamoso que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de $9\times$ a $0,1\times$; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados; analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de adenocarcinoma o carcinoma de células escamosas.

[0085] En algunos aspectos, el factor de transcripción especificado ASCL1, NEUROD1, POUF23, YAP1, o cualquier combinación de estos.

[0086] En algunos aspectos, el método utiliza el aprendizaje automático para subtipificar el cáncer de pulmón microcítico en el sujeto.

[0087] En algunos aspectos, la subtipificación se realiza al calcular $\log_2$ (relación de profundidad de lectura) a través de ventanas de 20 pb que comienzan a 2 kb de los sitios de unión del factor de transcripción especificados.

[0088] En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) es la cobertura total sobre un factor de corrección de cobertura calculado mediante el uso de la cobertura mediana de anclajes de 5' y 3' 500pb en los extremos de la ventana de 2kb (es decir, $\log_2(\text{profundidad de lectura}/\text{factor de corrección})$).

[0089] En algunos aspectos, todos los cálculos de cobertura se desplazan por 1 para evitar divisiones por 0. En algunos aspectos, este ajuste se expresa como $f(x) = \log_2((\text{profundidad de la ventana} + 1)/(\text{mediana}(\text{anclajes}) + 1))$.

[0090] En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) se calcula para cada sitio de unión del factor de transcripción y luego se agrega al calcular la mediana de cada ventana para obtener una puntuación de cobertura del fragmento.

[0091] Algunos aspectos comprenden además el uso de un modelo aditivo general para

suavizar la cobertura y corregir el sesgo de GC.

[0092] En algunos aspectos, los intervalos genómicos no se solapan.

[0093] En algunos aspectos, cada uno de los intervalos genómicos comprende de miles a millones de pares de bases.

[0094] En algunos aspectos, se determina un perfil de fragmentación del cfDNA dentro de cada intervalo genómico.

[0095] En algunos aspectos, el perfil de fragmentación del cfDNA comprende un tamaño de fragmento mediano.

[0096] En algunos aspectos, el perfil de fragmentación del cfDNA comprende una distribución del tamaño de fragmento.

[0097] Algunos aspectos comprenden además administrarle al sujeto que se identificó que tiene un subtipo de adenocarcinoma o carcinoma escamoso un agente terapéutico adecuado para el tratamiento del tipo de cáncer.

[0098] En algunos aspectos, el agente terapéutico es una inmunoterapia.

[0099] Algunos aspectos comprenden además la cuantificación de una fracción tumoral circulante, que refleja el rendimiento de la frecuencia del alelo mayor (MAF) en relación con la evaluación de los criterios de evaluación de la respuesta en tumores sólidos (RECIST) tanto en adenocarcinoma como en carcinoma escamoso.

[0100] En una modalidad, la presente invención proporciona métodos para la subtipificación no invasiva del cáncer de pulmón no microcítico en un sujeto como adenocarcinoma o carcinoma escamoso que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de 30× a 0,1×; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados; analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en

los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de adenocarcinoma o carcinoma de células escamosas. En algunos aspectos, el factor de transcripción especificado ASCL1, NEUROD1, POUF23, YAP1, o cualquier combinación de estos. En algunos aspectos, el método utiliza el aprendizaje automático para subtipificar el cáncer de pulmón microcítico en el sujeto. En algunos aspectos, la subtipificación se realiza al calcular $\log_2$ (relación de profundidad de lectura) a través de ventanas de 20 pb que comienzan a 2 kb de los sitios de unión del factor de transcripción especificados. En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) es la cobertura total sobre un factor de corrección de cobertura calculado mediante el uso de la cobertura mediana de anclajes de 5' y 3' 500pb en los extremos de la ventana de 2kb (es decir, $\log_2(\text{profundidad de lectura}/\text{factor de corrección})$). En algunos aspectos, todos los cálculos de cobertura se desplazan por 1 para evitar divisiones por 0. En algunos aspectos, este ajuste se expresa como $f(x) = \log_2((\text{profundidad de la ventana} + 1)/(\text{mediana}(\text{anclajes}) + 1))$. En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) se calcula para cada sitio de unión del factor de transcripción y luego se agrega al calcular la mediana de cada ventana para obtener una puntuación de cobertura del fragmento. Algunos aspectos comprenden además el uso de un modelo aditivo general para suavizar la cobertura y corregir el sesgo de GC. En algunos aspectos, los intervalos genómicos no se solapan. En algunos aspectos, cada uno de los intervalos genómicos comprende de miles a millones de pares de bases. En algunos aspectos, se determina un perfil de fragmentación del cfDNA dentro de cada intervalo genómico. En algunos aspectos, el perfil de fragmentación del cfDNA comprende un tamaño de fragmento mediano. En algunos aspectos, el perfil de fragmentación del cfDNA comprende una distribución del tamaño de fragmento. Algunos aspectos comprenden además administrarle al sujeto que se identificó que tiene un subtipo de adenocarcinoma o carcinoma escamoso un agente terapéutico adecuado para el tratamiento del tipo de cáncer. En algunos aspectos, el agente terapéutico es una inmunoterapia. Algunos aspectos comprenden además la cuantificación de una fracción tumoral circulante, que refleja el rendimiento de la frecuencia del alelo mayor (MAF) en relación con la evaluación de los criterios de evaluación de la respuesta en tumores sólidos (RECIST) tanto en adenocarcinoma como en carcinoma escamoso.

[0101] En una modalidad, la presente invención proporciona métodos para la subtipificación no invasiva del cáncer de pulmón no microcítico en un sujeto como

adenocarcinoma o carcinoma escamoso que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de $10\times$ a $0,1\times$; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados; analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de adenocarcinoma o carcinoma de células escamosas. En algunos aspectos, el factor de transcripción especificado ASCL1, NEUROD1, POUF23, YAP1, o cualquier combinación de estos. En algunos aspectos, el método utiliza el aprendizaje automático para subtipificar el cáncer de pulmón microcítico en el sujeto. En algunos aspectos, la subtipificación se realiza al calcular $\log_2$ (relación de profundidad de lectura) a través de ventanas de 20 pb que comienzan a 2 kb de los sitios de unión del factor de transcripción especificados. En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) es la cobertura total sobre un factor de corrección de cobertura calculado mediante el uso de la cobertura mediana de anclajes de 5' y 3' 500pb en los extremos de la ventana de 2kb (es decir, $\log_2(\text{profundidad de lectura}/\text{factor de corrección})$). En algunos aspectos, todos los cálculos de cobertura se desplazan por 1 para evitar divisiones por 0. En algunos aspectos, este ajuste se expresa como $f(x) = \log_2((\text{profundidad de la ventana} + 1)/(\text{mediana}(\text{anclajes}) + 1))$. En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) se calcula para cada sitio de unión del factor de transcripción y luego se agrega al calcular la mediana de cada ventana para obtener una puntuación de cobertura del fragmento. Algunos aspectos comprenden además el uso de un modelo aditivo general para suavizar la cobertura y corregir el sesgo de GC. En algunos aspectos, los intervalos genómicos no se solapan. En algunos aspectos, cada uno de los intervalos genómicos comprende de miles a millones de pares de bases. En algunos aspectos, se determina un perfil de fragmentación del cfDNA dentro de cada

intervalo genómico. En algunos aspectos, el perfil de fragmentación del cfDNA comprende un tamaño de fragmento mediano. En algunos aspectos, el perfil de fragmentación del cfDNA comprende una distribución del tamaño de fragmento. Algunos aspectos comprenden además administrarle al sujeto que se identificó que tiene un subtipo de adenocarcinoma o carcinoma escamoso un agente terapéutico adecuado para el tratamiento del tipo de cáncer. En algunos aspectos, el agente terapéutico es una inmunoterapia. Algunos aspectos comprenden además la cuantificación de una fracción tumoral circulante, que refleja el rendimiento de la frecuencia del alelo mayor (MAF) en relación con la evaluación de los criterios de evaluación de la respuesta en tumores sólidos (RECIST) tanto en adenocarcinoma como en carcinoma escamoso.

[0102] En una modalidad, la presente invención proporciona métodos para la subtipificación no invasiva del cáncer de pulmón no microcítico en un sujeto como adenocarcinoma o carcinoma escamoso que comprenden: procesar fragmentos de cfDNA de una muestra obtenida del sujeto y generar bibliotecas de secuenciación; someter las bibliotecas de secuenciación a la secuenciación del genoma completo para obtener fragmentos secuenciados, donde la cobertura del genoma es de alrededor de $9\times$ a $0,1\times$; mapear los fragmentos secuenciados a un genoma para obtener intervalos genómicos de secuencias mapeadas en sitios de unión de factores de transcripción especificados; analizar los intervalos genómicos de secuencias mapeadas para determinar las longitudes y cantidades de fragmentos de cfDNA para establecer una puntuación de cobertura de fragmentos de cfDNA en sitios específicos de unión a factores de transcripción mediante el uso de las longitudes y cantidades de fragmentos de cfDNA; y subtipificación del cáncer de pulmón microcítico en el sujeto basada en la activación de factores de transcripción; donde una disminución de las puntuaciones agregadas de cobertura de fragmentos de cfDNA en los sitios de unión a factores de transcripción especificados es indicativa de un subtipo de adenocarcinoma o carcinoma de células escamosas. En algunos aspectos, el factor de transcripción especificado ASCL1, NEUROD1, POUF23, YAP1, o cualquier combinación de estos. En algunos aspectos, el método utiliza el aprendizaje automático para subtipificar el cáncer de pulmón microcítico en el sujeto. En algunos aspectos, la subtipificación se realiza al calcular $\log_2$ (relación de profundidad de lectura) a través de ventanas de 20 pb que comienzan a 2 kb de los sitios de unión del factor de transcripción especificados. En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) es la cobertura total sobre un factor

de corrección de cobertura calculado mediante el uso de la cobertura mediana de anclajes de 5' y 3' 500pb en los extremos de la ventana de 2kb (es decir, $\log_2(\text{profundidad de lectura/factor de corrección})$). En algunos aspectos, todos los cálculos de cobertura se desplazan por 1 para evitar divisiones por 0. En algunos aspectos, este ajuste se expresa como $f(x) = \log_2((\text{profundidad de la ventana} + 1)/(\text{mediana(anclajes)} + 1))$. En algunos aspectos, el $\log_2$ (relación de profundidad de lectura) se calcula para cada sitio de unión del factor de transcripción y luego se agrega al calcular la mediana de cada ventana para obtener una puntuación de cobertura del fragmento. Algunos aspectos comprenden además el uso de un modelo aditivo general para suavizar la cobertura y corregir el sesgo de GC. En algunos aspectos, los intervalos genómicos no se solapan. En algunos aspectos, cada uno de los intervalos genómicos comprende de miles a millones de pares de bases. En algunos aspectos, se determina un perfil de fragmentación del cfDNA dentro de cada intervalo genómico. En algunos aspectos, el perfil de fragmentación del cfDNA comprende un tamaño de fragmento mediano. En algunos aspectos, el perfil de fragmentación del cfDNA comprende una distribución del tamaño de fragmento. Algunos aspectos comprenden además administrarle al sujeto que se identificó que tiene un subtipo de adenocarcinoma o carcinoma escamoso un agente terapéutico adecuado para el tratamiento del tipo de cáncer. En algunos aspectos, el agente terapéutico es una inmunoterapia. Algunos aspectos comprenden además la cuantificación de una fracción tumoral circulante, que refleja el rendimiento de la frecuencia del alelo mayor (MAF) en relación con la evaluación de los criterios de evaluación de la respuesta en tumores sólidos (RECIST) tanto en adenocarcinoma como en carcinoma escamoso.

[0103] Antes de describir las composiciones y métodos de la presente, debe entenderse que la presente invención no se limita a los métodos y sistemas particulares descritos, ya que dichos métodos y sistemas pueden variar. También debe entenderse que la terminología que se utiliza en la presente es únicamente a los efectos de describir modalidades específicas y no pretende ser limitante ya que el alcance de la presente invención se encontrará limitado únicamente a las reivindicaciones adjuntas.

[0104] Tal como se usa en la presente memoria descriptiva y en las reivindicaciones adjuntas, las formas singulares “un”, “una” y “el/la”, incluyen referencias plurales a menos que el contexto indique claramente lo contrario. Por lo tanto, por ejemplo, las referencias a “el método” incluyen uno o más métodos y/o etapas del tipo que se describió en la presente,

lo que resultará evidente para aquellos expertos en la técnica tras leer la presente descripción y en lo sucesivo.

[0105] A menos que se defina lo contrario, todos los términos técnicos y científicos utilizados en la presente tienen el mismo significado que entiende comúnmente un experto en la técnica a la que pertenece la presente invención. Si bien en la práctica o análisis de la invención se puede utilizar cualquier método y material similar o equivalente a los descritos en la presente, los métodos, dispositivos y materiales preferidos se describen ahora.

[0106] En la presente se describen métodos para subtipificar el cáncer de pulmón no microcítico (NSCLC) de forma no invasiva mediante el uso de fragmentomas de ADN libre de células (cfDNA). El NSCLC suele presentarse en diferentes subtipos, como el adenocarcinoma y el carcinoma de células escamosas. Este novedoso enfoque no invasivo basado en DELFI distingue los subtipos de adenocarcinoma y carcinoma escamoso de NSCLC investigando fragmentos libres de células que reflejan los estados epigenéticos únicos de los NSCLC. Se desconoce si otras biopsias líquidas basadas en cfDNA pueden diferenciar subtipos de NSCLC sin tener acceso a datos clínicos de tejidos.

[0107] Los subtipos de NSCLC pueden predecirse a partir del enfoque basado en DELFI mediante el uso de datos de cfDNA LC-WGS. En primer lugar, el clasificador de aprendizaje automático DELFI detecta la presencia de cáncer en pacientes con NSCLC. Se investigan tanto los perfiles de fragmentación de todo el genoma como las correspondientes puntuaciones de fracción tumoral (TF) DELFI para identificar diferencias preliminares entre los dos principales subtipos clínicos de casos de NSCLC: adenocarcinoma y carcinoma de células escamosas. DELFI-TF detecta con precisión la fracción tumoral circulante sin confundirse por la hematopoyesis clonal. Además, DELFI-TF cuantifica con precisión la fracción tumoral circulante, reflejando el rendimiento de MAF en relación con la evaluación RECIST tanto en adenocarcinoma como en carcinoma escamoso. Se utilizaron datos disponibles públicamente para ejecutar un análisis de componentes principales con el fin de agrupar los dos subtipos mediante el uso de datos de fragmentación del genoma completo. La **Figura 19** muestra el desarrollo del modelo DELFI-TF de prueba de concepto. En segundo lugar, se aplica una lista de sitios de inicio de la transcripción (TSS) accesibles de forma diferencial, capaz de diferenciar entre adenocarcinoma y escamoso. Se puede obtener una lista de sitios de inicio de transcripción accesibles de forma diferencial a partir de bases de datos públicas como UCSC o Ensembl, entre otras. La creación de estas listas

es similar a lo que se haría para los experimentos de secuenciación de ARN, salvo que en lugar de transcripto por millón se utiliza la relación de profundidad de lectura como sustituto de la expresión. En algunos aspectos, se seleccionan los sitios con los mayores y más significativos cambios logarítmicos de proporción del 20 % de una cohorte de prueba y se aplican al resto de la cohorte en el mismo locus, seguido de una agrupación jerárquica para determinar si los subtipos permanecen juntos. En otro aspecto, se seleccionan los transcriptos más expresados de TCGA para un tipo de cáncer dado y se filtran de la lista aquellos transcriptos que también se expresan a cualquier nivel en la LMA (cáncer hematológico como sustituto imperfecto de la sangre normal). En tercer lugar, se investigó si los DELFI-TSS más diferenciales mostraban cambios cortos/largos, confirmando aún más que estas moléculas de cfDNA derivadas de tumores. Por último, se utiliza una característica operativa del receptor (ROC) que representa la sensibilidad y especificidad del enfoque DELFI-fragmentoma para identificar subtipos de NSCLC con el fin de identificar la precisión diagnóstica de nuestro enfoque entre las muestras de grupo identificadas. Dadas las dificultades logísticas para realizar biopsias tumorales de NSCLC, creemos que este enfoque podría ser un método viable y rápido para subtipificar el NSCLC de forma no invasiva.

[0108] Este documento proporciona métodos y materiales para determinar un perfil de fragmentación de cfDNA en un mamífero (por ejemplo, en una muestra obtenida de un mamífero). Tal como se usan en la presente, las expresiones "perfil de fragmentación", "diferencias en los patrones de fragmentación en función de la posición" y "diferencias en el tamaño y la cobertura de los fragmentos en función de la posición en el genoma" son equivalentes y pueden utilizarse indistintamente. En algunos casos, la determinación de un perfil de fragmentación de cfDNA en un mamífero puede utilizarse para identificar que un mamífero tiene cáncer. Por ejemplo, los fragmentos de cfDNA obtenidos de un mamífero (por ejemplo, de una muestra obtenida de un mamífero) pueden someterse a una secuenciación de todo el genoma de baja cobertura, y los fragmentos secuenciados pueden mapearse en el genoma (por ejemplo, en ventanas no solapadas) y evaluarse para determinar un perfil de fragmentación de cfDNA. Como se describe en la presente, el perfil de fragmentación del cfDNA de un mamífero que tiene cáncer es más heterogéneo (por ejemplo, en longitudes de fragmentos) que el perfil de fragmentación del cfDNA de un mamífero sano (por ejemplo, un mamífero que no tiene cáncer). En sí, este documento

también proporciona métodos y materiales para evaluar, monitorear y/o tratar mamíferos (por ejemplo, seres humanos) que tienen o se sospecha que tienen cáncer. En algunos casos, este documento proporciona métodos y materiales para identificar que un mamífero tiene cáncer. Por ejemplo, una muestra (por ejemplo, una muestra sanguínea) obtenida de un mamífero puede evaluarse para determinar la presencia y, opcionalmente, el tejido de origen del cáncer en el mamífero basándose, al menos en parte, en el perfil de fragmentación del cfDNA del mamífero. En algunos casos, este documento proporciona métodos y materiales para monitorear si un mamífero tiene cáncer. Por ejemplo, una muestra (por ejemplo, una muestra sanguínea) obtenida de un mamífero puede evaluarse para determinar la presencia del cáncer en el mamífero basándose, al menos en parte, en el perfil de fragmentación del cfDNA del mamífero. En algunos casos, este documento proporciona métodos y materiales para identificar que un mamífero tiene cáncer y administrarle uno o más tratamientos contra el cáncer al mamífero para tratarlo. Por ejemplo, una muestra (por ejemplo, una muestra de sangre) obtenida de un mamífero puede evaluarse para determinar si el mamífero tiene cáncer basándose, al menos en parte, en el perfil de fragmentación del cfDNA del mamífero, y pueden administrarse al mamífero uno o más tratamientos contra el cáncer.

[0109] Un perfil de fragmentación del cfDNA puede incluir uno o más patrones de fragmentación del cfDNA. Un patrón de fragmentación de cfDNA puede incluir cualquier patrón de fragmentación de cfDNA apropiado. Los ejemplos de patrones de fragmentación del cfDNA incluyen, sin limitación, la mediana de tamaño de fragmento, la distribución del tamaño de fragmento, la relación entre fragmentos pequeños de cfDNA y fragmentos grandes de cfDNA, y la cobertura de los fragmentos de cfDNA. En algunos casos, un patrón de fragmentación de cfDNA incluye dos o más (por ejemplo, dos, tres o cuatro) de la mediana de tamaño de fragmento, la distribución del tamaño de fragmento, la relación entre fragmentos pequeños de cfDNA y fragmentos grandes de cfDNA, y la cobertura de los fragmentos de cfDNA. En algunos casos, el perfil de fragmentación del cfDNA puede ser un perfil de cfDNA de todo el genoma (por ejemplo, un perfil de cfDNA de todo el genoma en ventanas de todo el genoma). En algunos casos, el perfil de fragmentación del cfDNA puede ser un perfil de región objetivo. Una región objetivo puede ser cualquier parte apropiada del genoma (por ejemplo, una región cromosómica). Los ejemplos de regiones cromosómicas para las que puede determinarse un perfil de fragmentación de cfDNA como se describe en

la presente incluyen, sin limitación, una parte de un cromosoma (por ejemplo, una parte de 2q, 4p, 5p, 6q, 7p, 8q, 9q, 10q, 11q, 12q, y/o 14q) y un brazo cromosómico (por ejemplo, un brazo cromosómico de 8q, 13q, 11q, y/o 3p). En algunos casos, un perfil de fragmentación de cfDNA puede incluir dos o más perfiles de regiones objetivo.

[0110] En algunos casos, puede utilizarse un perfil de fragmentación del cfDNA para identificar cambios (por ejemplo, alteraciones) en las longitudes de los fragmentos de cfDNA. Una alteración puede ser una alteración de todo el genoma o una alteración en una o más regiones o locus específicos. Una región objetivo puede ser cualquier región que contenga una o más alteraciones específicas del cáncer. Los ejemplos de alteraciones específicas del cáncer, y sus localizaciones cromosómicas, incluyen, sin limitación, las mostradas en la Tabla 3 (Apéndice C) y las mostradas en la Tabla 6 (Apéndice F). En algunos casos, un perfil de fragmentación del cfDNA puede utilizarse para identificar (por ejemplo, identificar simultáneamente) desde alrededor de 10 alteraciones a alrededor de 500 alteraciones (por ejemplo, de alrededor de 25 a alrededor de 500, de alrededor de 50 a alrededor de 500, de alrededor de 100 a alrededor de 500, de alrededor de 200 a alrededor de 500, de alrededor de 300 a alrededor de 500, de alrededor de 10 a alrededor de 400, de alrededor de 10 a alrededor de 300, de alrededor de 10 a alrededor de 200, de alrededor de 10 a alrededor de 100, de alrededor de 10 a alrededor de 50, de alrededor de 20 a alrededor de 400, de alrededor de 30 a alrededor de 300, de alrededor de 40 a alrededor de 200, de alrededor de 50 a alrededor de 100, de alrededor de 20 a alrededor de 100, de alrededor de 25 a alrededor de 75, de alrededor de 50 a alrededor de 250, o de alrededor de 100 a alrededor de 200, alteraciones).

[0111] En algunos casos, puede utilizarse un perfil de fragmentación del cfDNA para detectar ADN derivado de tumores. Por ejemplo, se puede utilizar un perfil de fragmentación de cfDNA para detectar ADN derivado de tumores al comparar un perfil de fragmentación de cfDNA de un mamífero que tiene o se sospecha que tiene cáncer con un perfil de fragmentación de cfDNA de referencia (por ejemplo, un perfil de fragmentación de cfDNA de un mamífero sano y/o un perfil de fragmentación de ADN nucleosómico de células sanas del mamífero que tiene o se sospecha que tiene cáncer). En algunos casos, un perfil de fragmentación de cfDNA de referencia es un perfil generado previamente a partir de un mamífero sano. Por ejemplo, los métodos proporcionados en la presente pueden utilizarse para determinar un perfil de fragmentación de cfDNA de referencia en un mamífero sano, y

ese perfil de fragmentación de cfDNA de referencia puede almacenarse (por ejemplo, en una computadora u otro medio de almacenamiento electrónico) para su comparación futura con un perfil de fragmentación de cfDNA de prueba en un mamífero que tiene o se sospecha que tiene cáncer. En algunos casos, se determina un perfil de fragmentación de cfDNA de referencia (por ejemplo, un perfil de fragmentación de cfDNA almacenado) de un mamífero sano en todo el genoma. En algunos casos, se determina un perfil de fragmentación de cfDNA de referencia (por ejemplo, un perfil de fragmentación de cfDNA almacenado) de un mamífero sano en un intervalo subgenómico.

[0112] En algunos casos, un perfil de fragmentación de cfDNA puede utilizarse para identificar que un mamífero (por ejemplo, un ser humano) tiene cáncer (por ejemplo, un cáncer colorrectal, un cáncer de pulmón, un cáncer de mama, un cáncer gástrico, un cáncer de páncreas, un cáncer de vías biliares y/o un cáncer de ovario).

[0113] Un perfil de fragmentación de cfDNA puede incluir un patrón de tamaño de fragmento de cfDNA. Los fragmentos de cfDNA pueden ser de cualquier tamaño apropiado. Por ejemplo, el fragmento de cfDNA puede tener una longitud de alrededor de 50 pares de bases (pb) a alrededor de 400 pb. Como se describe en la presente, un mamífero con cáncer puede tener un patrón de tamaño de fragmento de cfDNA que contenga una mediana de tamaño de fragmento de cfDNA menor que la mediana de tamaño de fragmento de cfDNA en un mamífero sano. Un mamífero sano (por ejemplo, un mamífero que no tiene cáncer) puede tener tamaños de fragmentos de cfDNA con una mediana de tamaño de fragmento de cfDNA de alrededor de 166,6 pb a alrededor de 167,2 pb (por ejemplo, alrededor de 166,9 pb). En algunos casos, un mamífero con cáncer puede tener tamaños de fragmentos de cfDNA que son, en promedio, de alrededor de 1,28 pb a alrededor de 2,49 pb (por ejemplo, 1,88 pb) menores que los tamaños de fragmento de cfDNA de un mamífero sano. Por ejemplo, un mamífero con cáncer puede tener tamaños de fragmento de cfDNA con una mediana de tamaño de fragmento de cfDNA de alrededor de 164,11 pb a alrededor de 165,92 pb (por ejemplo, alrededor de 165,02 pb).

[0114] Un perfil de fragmentación de cfDNA puede incluir una distribución del tamaño de los fragmentos de cfDNA. Como se describe en la presente, un mamífero con cáncer puede tener una distribución de tamaños de cfDNA más variable que una distribución de tamaños de fragmentos de cfDNA en un mamífero sano. En algunos casos, una distribución de tamaño puede estar dentro de una región objetivo. Un mamífero sano (por ejemplo, un

mamífero que no tiene cáncer) puede tener una distribución de tamaño de fragmento de cfDNA de región objetivo de alrededor de 1 o inferior a alrededor de 1. En algunos casos, un mamífero con cáncer puede tener una distribución de tamaños de fragmentos de CfDNA de región objetivo mayor (por ejemplo, 10, 15, 20, 25, 30, 35, 40, 45, 50 o más pb mayor, o cualquier número de pares de bases entre estos números) que una distribución de tamaños de fragmentos de cfDNA de región objetivo en un mamífero sano. En algunos casos, un mamífero con cáncer puede tener una distribución de tamaños de fragmentos de CfDNA de región objetivo menor (por ejemplo, 10, 15, 20, 25, 30, 35, 40, 45, 50 o más pb menor, o cualquier número de pares de bases entre estos números) que una distribución de tamaños de fragmentos de cfDNA de región objetivo en un mamífero sano. En algunos casos, un mamífero con cáncer puede tener una distribución del tamaño de los fragmentos de cfDNA de la región objetivo entre 47 pb menor y 30 pb mayor que la distribución del tamaño de los fragmentos de cfDNA de la región objetivo en un mamífero sano. En algunos casos, un mamífero con cáncer puede tener una distribución del tamaño de los fragmentos de cfDNA de la región objetivo de, por término medio, una diferencia de 10, 11, 12, 13, 14, 15, 15, 17, 18, 19, 20 o más pb en las longitudes de los fragmentos de cfDNA. Por ejemplo, un mamífero con cáncer puede tener una distribución del tamaño de los fragmentos de cfDNA de la región objetivo de, por término medio, alrededor de 13 pb de diferencia en las longitudes de los fragmentos de cfDNA. En algunos casos, una distribución de tamaño puede ser una distribución de tamaño de todo el genoma. Un mamífero sano (por ejemplo, un mamífero que no tiene cáncer) puede tener distribuciones muy similares de fragmentos de cfDNA cortos y largos en todo el genoma. En algunos casos, un mamífero con cáncer puede tener, en todo el genoma, una o más alteraciones (por ejemplo, aumentos y disminuciones) en los tamaños de los fragmentos de cfDNA. La o las alteraciones pueden ser cualquier región cromosómica apropiada del genoma. Por ejemplo, una alteración puede producirse en una parte de un cromosoma. Ejemplos de porciones de cromosomas que pueden contener una o más alteraciones en los tamaños de fragmentos de cfDNA incluyen, sin limitación, partes de 2q, 4p, 5p, 6q, 7p, 8q, 9q, 10q, 11q, 12q y 14q. Por ejemplo, una alteración puede ser a través de un brazo cromosómico (por ejemplo, un brazo cromosómico entero).

[0115] Un perfil de fragmentación de cfDNA puede incluir una relación entre fragmentos pequeños de cfDNA y fragmentos grandes de cfDNA y una correlación de las relaciones de fragmentos con las relaciones de fragmentos de referencia. Tal como se utiliza en la

presente, con respecto a las relaciones de fragmentos pequeños de cfDNA con respecto a fragmentos grandes de cfDNA, un fragmento pequeño de cfDNA puede tener una longitud de alrededor de 100 pb a alrededor de 150 pb. Tal como se utiliza en la presente, con respecto a las relaciones de fragmentos grandes de cfDNA con respecto a fragmentos grandes de cfDNA, un fragmento pequeño de cfDNA puede tener una longitud de alrededor de 151 pb a 220 pb. Como se describe en la presente, un mamífero con cáncer puede tener una correlación de relaciones de fragmentos (por ejemplo, una correlación de relaciones de fragmentos de cfDNA con respecto a relaciones de fragmentos de ADN de referencia, como relaciones de fragmentos de ADN de uno o más mamíferos sanos) que es inferior (por ejemplo, 2 veces inferior, 3 veces inferior, 4 veces inferior, 5 veces inferior, 6 veces inferior, 7 veces inferior, 8 veces inferior, 9 veces inferior, 10 veces inferior o más) que en un mamífero sano. Un mamífero sano (por ejemplo, un mamífero que no tiene cáncer) puede tener una correlación de relaciones de fragmentos (por ejemplo, una correlación de relaciones de fragmentos de cfDNA con relaciones de fragmentos de ADN de referencia, tales como relaciones de fragmentos de ADN de uno o más mamíferos sanos) de alrededor de 1 (por ejemplo, alrededor de 0,96). En algunos casos, un mamífero con cáncer puede tener una correlación de relaciones de fragmentos (por ejemplo, una correlación de relaciones de fragmentos de cfDNA con respecto a relaciones de fragmentos de ADN de referencia tales como relaciones de fragmentos de ADN de uno o más mamíferos sanos) que es, en promedio, de alrededor de 0,19 a alrededor de 0,30 (por ejemplo, alrededor de 0,25) inferior a una correlación de relaciones de fragmentos (por ejemplo, una correlación de relaciones de fragmentos de cfDNA con respecto a relaciones de fragmentos de ADN de referencia, como relaciones de fragmentos de ADN de uno o más mamíferos sanos) en un mamífero sano.

[0116] Un perfil de fragmentación de cfDNA puede incluir la cobertura de todos los fragmentos. La cobertura de todos los fragmentos puede incluir ventanas (por ejemplo, ventanas no solapadas) de cobertura. En algunos casos, la cobertura de todos los fragmentos puede incluir ventanas de fragmentos pequeños (por ejemplo, fragmentos de alrededor de 100 pb a alrededor de 150 pb de longitud). En algunos casos, la cobertura de todos los fragmentos puede incluir ventanas de fragmentos grandes (por ejemplo, fragmentos de alrededor de 151 pb a alrededor de 220 pb de longitud).

[0117] Se puede obtener un perfil de fragmentación del cfDNA mediante el uso de

cualquier método apropiado. En algunos casos, el cfDNA de un mamífero (por ejemplo, un mamífero que tiene o se sospecha que tiene cáncer) puede procesarse en bibliotecas de secuenciación que pueden someterse a secuenciación del genoma completo (por ejemplo, secuenciación del genoma completo de baja cobertura), mapearse en el genoma y analizarse para determinar las longitudes de los fragmentos de cfDNA. Las secuencias mapeadas pueden analizarse en ventanas no solapadas que cubren el genoma. Las ventanas pueden ser de cualquier tamaño. Por ejemplo, las ventanas pueden tener una longitud de miles a millones de bases. Como ejemplo no taxativo, una ventana puede tener una longitud de alrededor de 5 megabases (Mb). Se puede mapear cualquier número apropiado de ventanas. Por ejemplo, se pueden mapear de decenas a miles de ventanas en el genoma. Por ejemplo, se pueden mapear de centenas a miles de ventanas en el genoma. Dentro de cada ventana puede determinarse un perfil de fragmentación del cfDNA. En algunos casos, puede obtenerse un perfil de fragmentación del cfDNA como se describe en el Ejemplo 1. En algunos casos, se puede obtener un perfil de fragmentación del cfDNA como se muestra en la FIG. 1.

[0118] En algunos casos, los métodos y materiales descritos en la presente también pueden incluir aprendizaje automático. Por ejemplo, puede utilizarse el aprendizaje automático para identificar un perfil de fragmentación alterado (por ejemplo, mediante el uso de la cobertura de los fragmentos de cfDNA, el tamaño de los fragmentos de cfDNA, la cobertura de los cromosomas y el ADNmt).

[0119] En algunos casos, los métodos y materiales descritos en la presente pueden ser el único método utilizado para identificar que un mamífero (por ejemplo, un ser humano) tiene cáncer (por ejemplo, un cáncer colorrectal, un cáncer de pulmón, un cáncer de mama, un cáncer gástrico, un cáncer de páncreas, un cáncer de vías biliares y/o un cáncer de ovario). Por ejemplo, la determinación de un perfil de fragmentación del cfDNA puede ser el único método utilizado para identificar que un mamífero tiene cáncer.

[0120] En algunos casos, los métodos y materiales descritos en la presente pueden utilizarse junto con uno o más métodos adicionales utilizados para identificar que un mamífero (por ejemplo, un ser humano) tiene cáncer (por ejemplo, un cáncer colorrectal, un cáncer de pulmón, un cáncer de mama, un cáncer gástrico, un cáncer de páncreas, un cáncer de vías biliares y/o un cáncer de ovario). Los ejemplos de métodos utilizados para identificar que un mamífero tiene cáncer incluyen, de modo no taxativo, la identificación de

una o más alteraciones de secuencia específicas del cáncer, la identificación de una o más alteraciones cromosómicas (por ejemplo, aneuploidías y reordenamientos) y la identificación de otras alteraciones del cfDNA. Por ejemplo, la determinación de un perfil de fragmentación del cfDNA puede utilizarse junto con la identificación de una o más mutaciones específicas del cáncer en el genoma de un mamífero para identificar que un mamífero tiene cáncer. Por ejemplo, la determinación de un perfil de fragmentación del cfDNA puede utilizarse junto con la identificación de una o más aneuploidías en el genoma de un mamífero para identificar que un mamífero tiene cáncer.

[0121] En algunos aspectos, este documento también proporciona métodos y materiales para evaluar, monitorear y/o tratar mamíferos (por ejemplo, seres humanos) que tienen o se sospecha que tienen cáncer. En algunos casos, este documento proporciona métodos y materiales para identificar que un mamífero tiene cáncer. Por ejemplo, una muestra (por ejemplo, una muestra sanguínea) obtenida de un mamífero puede evaluarse para determinar si el mamífero tiene cáncer basándose, al menos en parte, en el perfil de fragmentación del cfDNA del mamífero. En algunos casos, este documento proporciona métodos y materiales para identificar la localización (por ejemplo, el lugar anatómico o el tejido de origen) de un cáncer en un mamífero. Por ejemplo, una muestra (por ejemplo, una muestra sanguínea) obtenida de un mamífero puede evaluarse para determinar el tejido de origen del cáncer en el mamífero basándose, al menos en parte, en el perfil de fragmentación del cfDNA del mamífero. En algunos casos, este documento proporciona métodos y materiales para identificar que un mamífero tiene cáncer y administrarle uno o más tratamientos contra el cáncer al mamífero para tratarlo. Por ejemplo, una muestra (por ejemplo, una muestra de sangre) obtenida de un mamífero puede evaluarse para determinar si el mamífero tiene cáncer basándose, al menos en parte, en el perfil de fragmentación del cfDNA del mamífero, y administrarle uno o más tratamientos contra el cáncer al mamífero. En algunos casos, este documento proporciona métodos y materiales para tratar a un mamífero que tiene cáncer. Por ejemplo, uno o más tratamientos contra el cáncer pueden administrarse a un mamífero que se identificó que tiene cáncer (por ejemplo, basándose, al menos en parte, en el perfil de fragmentación del cfDNA del mamífero) para tratar al mamífero. En algunos casos, durante o después del curso de un tratamiento contra el cáncer (por ejemplo, cualquiera de los tratamientos contra el cáncer descritos en la presente), un mamífero puede someterse a monitoreo (o ser seleccionado para un mayor monitoreo) y/o

a pruebas diagnósticas adicionales. En algunos casos, el monitoreo puede incluir la evaluación de mamíferos que tienen o se sospecha que tienen cáncer mediante, por ejemplo, la evaluación de una muestra (por ejemplo, una muestra sanguínea) obtenida del mamífero para determinar el perfil de fragmentación del cfDNA del mamífero como se describe en la presente, y los cambios en los perfiles de fragmentación del cfDNA a lo largo del tiempo pueden utilizarse para identificar la respuesta al tratamiento y/o identificar al mamífero como portador de cáncer (por ejemplo, un cáncer residual).

[0122] Cualquier mamífero adecuado puede ser evaluado, monitoreado o tratado como se describe en la presente. Un mamífero puede ser un mamífero con cáncer. Un mamífero puede ser un mamífero que se sospecha que tiene cáncer. Los ejemplos de mamíferos que pueden ser evaluados, monitoreados y/o tratados como se describe en la presente incluyen, sin limitación, seres humanos, primates como monos, perros, gatos, caballos, vacas, cerdos, ovejas, ratones y ratas. Por ejemplo, un ser humano que tenga o se sospeche que tenga cáncer puede evaluarse para determinar un perfil de fragmentación del cfDNA como se describe en la presente y, opcionalmente, puede tratarse con uno o más tratamientos contra el cáncer como se describe en la presente.

[0123] Cualquier muestra apropiada de un mamífero puede evaluarse como se describe en la presente (por ejemplo, evaluar un patrón de fragmentación del ADN). En algunos casos, una muestra puede incluir ADN (por ejemplo, ADN genómico). En algunos casos, una muestra puede incluir cfDNA (por ejemplo, ADN circulante tumoral [ADNct]). En algunos casos, una muestra puede ser una muestra líquida (por ejemplo, una biopsia líquida). Los ejemplos de muestras que pueden contener ADN y/o polipéptidos incluyen, sin limitación, sangre (por ejemplo, sangre entera, suero o plasma), amnios, tejidos, orina, líquido cefalorraquídeo, saliva, esputo, lavado bronco-alveolar, bilis, líquido linfático, líquido de quiste, heces, ascitis, frotis de Papanicolaou, leche materna y condensado de aliento exhalado. Por ejemplo, puede evaluarse una muestra de plasma para determinar un perfil de fragmentación de cfDNA como se describe en la presente.

[0124] Una muestra de un mamífero que vaya a evaluarse como se describe en la presente (por ejemplo, evaluarse para determinar un patrón de fragmentación del ADN) puede incluir cualquier cantidad adecuada de cfDNA. En algunos casos, una muestra puede incluir una cantidad limitada de ADN. Por ejemplo, se puede obtener un perfil de fragmentación de cfDNA a partir de una muestra que incluya menos ADN del que se requiere

normalmente para otros métodos de análisis de cfDNA, como los descritos en, por ejemplo, Phallen et al., 2017 Sci Transl Med 9; Cohen et al., 2018 Science 359:926; Newman et al., 2014 Nat Med 20:548; y Newman et al., 2016 Nat Biotechnol 34:547).

[0125] En algunos casos, una muestra puede procesarse (por ejemplo, para aislar y/o purificar ADN y/o polipéptidos a partir de la muestra). Por ejemplo, el aislamiento y/o purificación del ADN puede incluir la lisis celular (por ejemplo, mediante el uso de detergentes y/o tensioactivos), la eliminación de proteínas (por ejemplo, mediante el uso de una proteasa), y/o la eliminación de ARN (por ejemplo, mediante el uso de una RNasa). A modo de otro ejemplo, el aislamiento y/o purificación de polipéptidos puede incluir lisis celular (por ejemplo, usando detergentes y/o surfactantes), eliminación de ADN (por ejemplo, mediante el uso de una DNasa), y/o eliminación de ARN (por ejemplo, mediante el uso de una RNasa).

[0126] Se describen métodos adicionales en las patentes estadounidenses n.º 10,982,279 y n.º 10,975,431, cuya descripción se considera parte de la presente y se incorpora en la presente solicitud en su totalidad mediante referencia.

Ejemplo de implementación de hardware

[0127] La **Figura 20** ilustra una computadora 800 ilustrativa que puede utilizarse para implementar los métodos descritos en la presente. Por ejemplo, la computadora 800 puede incluir un sistema de aprendizaje automático que entrene un modelo de aprendizaje automático para subtipificar el cáncer de pulmón microcítico o el cáncer de pulmón no microcítico como se ha descrito anteriormente o una parte o combinación de estos en algunas modalidades. La computadora 800 puede ser cualquier dispositivo electrónico que ejecute aplicaciones de software derivadas de instrucciones compiladas, incluidas, sin limitación, computadoras personales, servidores, teléfonos inteligentes, reproductores multimedia, tabletas electrónicas, consolas de juegos, dispositivos de correo electrónico, etc. En algunas implementaciones, la computadora 800 puede incluir uno o más procesadores 802, uno o más dispositivos de entrada 804, uno o más dispositivos de visualización 806, una o más interfaces de red 808, y uno o más medios legibles por computadora 812. Cada uno de estos componentes puede estar acoplado por el bus 810, y en algunas modalidades, estos componentes pueden estar distribuidos entre múltiples ubicaciones físicas y acoplados por una red.

[0128] El dispositivo de visualización 806 puede ser cualquier tecnología de visualización conocida, que incluye, de modo no taxativo, dispositivos de visualización que utilicen tecnología de pantalla de cristal líquido (LCD) o diodos emisores de luz (LED). El procesador o procesadores 802 pueden utilizar cualquier tecnología de procesador conocida, incluidos, de modo no taxativo, procesadores gráficos y procesadores multinúcleo. El dispositivo de entrada 804 puede ser cualquier tecnología de dispositivo de entrada conocida, que incluye, de modo no taxativo, un teclado (incluido un teclado virtual), ratón, bola de seguimiento, cámara y almohadilla o pantalla sensible al tacto. El bus 810 puede ser cualquier tecnología de bus interna o externa conocida, lo que incluye, de modo no taxativo, ISA, EISA, PCI, PCI Express, USB, Serial ATA o FireWire. El medio legible por computadora 812 puede ser cualquier medio no transitorio que participe en el suministro de instrucciones al procesador 804 para su ejecución, que incluye, sin limitación, medios de almacenamiento no volátiles (por ejemplo, discos ópticos, discos magnéticos, unidades flash, etc.), o medios volátiles (por ejemplo, SDRAM, ROM, etc.).

[0129] El medio legible por computadora 812 puede incluir varias instrucciones 814 para implementar un sistema operativo (por ejemplo, Mac OS®, Windows®, Linux). El sistema operativo puede ser multiusuario, multiproceso, multitarea, multihilo, en tiempo real y similares. El sistema operativo puede realizar tareas básicas, lo que incluye, de modo no taxativo: reconocer la entrada desde el dispositivo de entrada 804; enviar la salida al dispositivo de visualización 806; mantener un registro de archivos y directorios en el medio legible por computadora 812; controlar los dispositivos periféricos (por ejemplo, unidades de disco, impresoras, etc.) que pueden ser controlados directamente o a través de un controlador de E/S; y gestionar el tráfico en el bus 810. Las instrucciones de comunicaciones de red 816 pueden establecer y mantener conexiones de red (por ejemplo, software para implementar protocolos de comunicación, como TCP/IP, HTTP, Ethernet, telefonía, etc.).

[0130] Las instrucciones de aprendizaje automático 818 pueden incluir instrucciones que permiten al computadora 800 funcionar como un sistema de aprendizaje automático y/o entrenar modelos de aprendizaje automático para generar valores DMS como se describe en la presente. Las aplicaciones 820 pueden ser una aplicación que utilice o implemente los procesos descritos en la presente y/u otros procesos. Los procesos también pueden implementarse en el sistema operativo 814. Por ejemplo, la aplicación 820 y/o el sistema operativo pueden crear tareas en las aplicaciones como se describe en la presente.

[0131] Las características descritas pueden implementarse en uno o más programas informáticos que son ejecutables en un sistema programable que incluye al menos un procesador programable, acoplado para recibir datos e instrucciones a un sistema de almacenamiento de datos, y para transmitir datos e instrucciones a partir de este, al menos un dispositivo de entrada y al menos un dispositivo de salida. Un programa informático es un conjunto de instrucciones que pueden utilizarse, directa o indirectamente, en un computadora para realizar una determinada actividad o producir un determinado resultado. Un programa informático puede escribirse en cualquier forma de lenguaje de programación (por ejemplo, Objective-C, Java), incluso en lenguajes compilados o interpretados, y puede implementarse de cualquier forma, inclusive como un programa que se ejecuta de forma autónoma o como módulo, componente, subrutina u otra unidad adecuada para su uso en un entorno informático.

[0132] Los procesadores adecuados para la ejecución de un programa de instrucciones pueden incluir, a modo de ejemplo, microprocesadores de uso general y especial, y el procesador único o uno de varios procesadores o núcleos, de cualquier tipo de ordenador. Por lo general, un procesador puede recibir instrucciones y datos de una memoria de solo lectura y/o una memoria de acceso aleatorio. Los elementos esenciales de un ordenador pueden incluir un procesador para ejecutar instrucciones y una o más memorias para almacenar instrucciones y datos. Generalmente, un computadora también puede incluir, o estar acoplado operativamente para comunicarse con, uno o más dispositivos de almacenamiento masivo para almacenar archivos de datos; tales dispositivos incluyen discos magnéticos, como discos duros internos y discos extraíbles; discos magneto-ópticos; y discos ópticos. Los dispositivos de almacenamiento adecuados para incorporar de forma tangible instrucciones y datos de programas informáticos pueden incluir todas las formas de memoria no volátil, incluidos, a modo de ejemplo, dispositivos de memoria semiconductora, tales como EPROM, EEPROM y dispositivos de memoria flash; discos magnéticos, tales como discos duros internos y discos extraíbles; discos magneto-ópticos; y discos CD-ROM y DVD-ROM. El procesador y la memoria pueden complementarse con circuitos integrados de aplicación específica (ASIC) o incorporarse a estos.

[0133] Para permitir la interacción con el usuario, las funciones pueden implementarse en un ordenador que disponga de un dispositivo de visualización, como un monitor LED o LCD, para mostrar información al usuario, y un teclado y un dispositivo señalador, como un

ratón o una bola de seguimiento, mediante los cuales el usuario pueda introducir datos en el ordenador.

[0134] Las funciones pueden implementarse en un sistema informático que incluya un componente back-end, como un servidor de datos, o que incluya un componente middleware, como un servidor de aplicaciones o un servidor de Internet, o que incluya un componente front-end, como un computadora cliente con una interfaz gráfica de usuario o un navegador de Internet, o cualquier combinación de estos. Los componentes del sistema pueden estar conectados por cualquier forma o medio de comunicación de datos digitales, tal como una red de comunicación. Los ejemplos de redes de comunicación son, por ejemplo, una red telefónica, una LAN, una WAN y los computadoras y redes que forman la Internet.

[0135] El sistema informático puede incluir clientes y servidores. Un cliente y servidor son generalmente remotos entre sí e interactúan típicamente a través de una red. La relación del cliente y servidor puede surgir en virtud de programas informáticos que se ejecutan en las computadoras respectivas y que tienen una relación cliente-servidor entre sí.

[0136] Una o más características o etapas de las modalidades descritas pueden implementarse mediante el uso de una Interfaz de Programación de Aplicaciones (API). Una API puede definir uno o más parámetros que se pasan entre una aplicación de llamada y otro código de software (por ejemplo, un sistema operativo, rutina de biblioteca, función) que proporciona un servicio, que proporciona datos, o que realiza una operación o un cálculo.

[0137] La API puede implementarse como una o más llamadas en el código del programa que envían o reciben uno o más parámetros a través de una lista de parámetros u otra estructura basada en una convención de llamada definida en un documento de especificación de la API. Un parámetro puede ser una constante, una clave, una estructura de datos, un objeto, una clase de objeto, una variable, un tipo de datos, un puntero, una matriz, una lista u otra llamada. Las llamadas a la API y los parámetros pueden implementarse en cualquier lenguaje de programación. El lenguaje de programación puede definir el vocabulario y la convención de llamada que un programador empleará para acceder a las funciones que soportan la API.

[0138] En algunas implementaciones, una llamada a la API puede informar a una aplicación de las capacidades de un dispositivo que ejecuta la aplicación, como la capacidad

de entrada, la capacidad de salida, la capacidad de procesamiento, la capacidad de alimentación, la capacidad de comunicación, etc.

[0139] Mientras que varias modalidades se han descrito anteriormente, se debe entender que han sido planteados a modo de ejemplo y de forma no taxativa. Será evidente para los expertos en la técnica pertinente que se pueden hacer varios cambios en la forma y detalles en la presente sin apartarse del espíritu y alcance. De hecho, después de leer la descripción anterior, será evidente para un experto en la técnica pertinente cómo implementar modalidades alternativas. Por ejemplo, pueden proporcionarse otras etapas, o pueden eliminarse etapas, de los flujos descritos, y pueden añadirse o eliminarse otros componentes de los sistemas descritos. Por consiguiente, otras implementaciones están comprendidas por el alcance de las siguientes reivindicaciones.

[0140] Además, debe entenderse que las figuras que destacan la funcionalidad y las ventajas se presentan únicamente a modo de ejemplo. La metodología y el sistema descritos son lo suficientemente flexibles y configurables como para que puedan utilizarse de formas distintas a la mostrada.

[0141] Aunque el término "al menos uno" puede utilizarse a menudo en la memoria descriptiva, las reivindicaciones y las figuras, los términos "un", "uno", "el", "dicho", etc. también significan "al menos uno" en la memoria descriptiva, las reivindicaciones y las figuras.

[0142] Por último, la intención del solicitante es que solamente se interpreten con de conformidad con el artículo 112(f) del título 35 del U.S.C. las reivindicaciones que incluyan la redacción expresa "medios para" o "etapas para". Las reivindicaciones que no incluyan expresamente la frase "medios para" o "etapa para" no deben interpretarse de conformidad con el artículo 112(f) del título 35 del U.S.C.

[0143] Los métodos y sistemas descritos actualmente son útiles para subtipificar el cáncer de pulmón no microcítico o el cáncer de pulmón microcítico en un sujeto y, opcionalmente, para tratar el subtipo de cáncer en el sujeto. Cualquier sujeto apropiado, tal como un mamífero, puede ser evaluado y/o tratado como se describe en la presente. Los ejemplos de algunos mamíferos que pueden ser evaluados y/o tratados como se describe en la presente incluyen, sin limitación, seres humanos, primates como monos, perros, gatos, caballos, vacas, cerdos, ovejas, ratones y ratas. Por ejemplo, un ser humano que tenga o se sospeche que tenga cáncer puede evaluarse mediante el uso de un método descrito en

la presente y, opcionalmente, puede tratarse con uno o más tratamientos contra el cáncer como se describe en la presente. Los métodos descritos en la presente pueden incluir administrarle al sujeto que se identificó que tenía el tipo de cáncer un agente terapéutico adecuado para el tratamiento del tipo de cáncer.

[0144] Cuando se trata a un sujeto que tiene o se sospecha que tiene cáncer como se describe en la presente, al sujeto se le puede administrar uno o más tratamientos contra el cáncer. Un tratamiento contra el cáncer puede ser cualquier tratamiento adecuado contra el cáncer. Uno o más tratamientos contra el cáncer descritos en la presente pueden administrarse a un sujeto con cualquier frecuencia adecuada (por ejemplo, una o varias veces durante un periodo de tiempo que oscila entre días y semanas). Los ejemplos de tratamientos contra el cáncer se incluyen, sin limitación, intervención quirúrgica, quimioterapia adyuvante, quimioterapia neoadyuvante, radioterapia, terapia hormonal, terapia citotóxica, inmunoterapia, terapia de linfocitos T adoptivos (por ejemplo, receptores de antígenos quiméricos y/o linfocitos T con receptores de linfocitos T de tipo salvaje o modificados), terapia dirigida como la administración de inhibidores de cinasa (por ejemplo, inhibidores de cinasa dirigidos a una lesión genética concreta, como una translocación o mutación) (por ejemplo, inhibidor de la cinasa, un anticuerpo, un anticuerpo biespecífico), inhibidores de la transducción de señales, anticuerpos biespecíficos o fragmentos de anticuerpos (por ejemplo, BiTE), anticuerpos monoclonales, inhibidores del punto de control inmunitario, cirugía (por ejemplo, resección quirúrgica), o cualquier combinación de los anteriores. En algunos aspectos, un tratamiento contra el cáncer puede reducir la gravedad del cáncer, reducir un síntoma del cáncer y/o reducir el número de células cancerosas presentes en el sujeto.

[0145] En algunos aspectos, un tratamiento para el cáncer puede ser un agente quimioterapéutico. Los ejemplos no taxativos de agentes quimioterapéuticos incluyen: amsacrina, azacitidina, axatioprina, bevacizumab (o un fragmento de unión a antígeno de este), bleomicina, busulfán, carboplatino, capecitabina, clorambucilo, cisplatino, ciclofosfamida, citarabina, dacarbazina, daunorrubicina, docetaxel, doxifluridina, doxorubicina, epirubicina, clorhidratos de erlotinib, etopósido, fludarabina, floxuridina, fludarabina, fluorouracilo, gemcitabina, hidroxiurea, idarubicina, ifosfamida, irinotecán, lomustina, mecloretamina, melfalán, mercaptopurina, metotrxato, mitomicina, mitoxantrona, oxaliplatino, paclitaxel, pemetrexed, procarbazona, ácido transretinoico total, estreptozocina,

tafluposida, temozolomida, tenipósido, tioguanina, topotecán, uramustina, valrubicina, vinblastina, vincristina, vindesina, vinorelbina y combinaciones de estos. Otros ejemplos de terapias contra el cáncer son conocidos en la técnica; véanse, por ejemplo, las directrices terapéuticas de la American Society of Clinical Oncology (ASCO), la European Society for Medical Oncology (ESMO) o la National Comprehensive Cancer Network (NCCN).

[0146] En varios aspectos, el ADN está presente en una muestra biológica tomada de un sujeto y utilizada en la metodología de la invención. La muestra biológica puede ser prácticamente cualquier tipo de muestra biológica que incluya ADN. La muestra biológica suele ser un líquido, como sangre total o una parte de ésta con cfDNA circulante. En algunas modalidades, la muestra incluye ADN de un tumor o de una biopsia líquida, como por ejemplo, de modo no taxativo, líquido amniótico, humor acuoso, humor vítreo, sangre, sangre entera, sangre fraccionada, plasma, suero, leche materna, líquido cefalorraquídeo (LCR), cerumen, quilo, limo, endolinfa, perilinfa, heces, aliento, ácido gástrico, jugo gástrico, linfa, mucosidad (incluidos el drenaje nasal y la flema), líquido pericárdico, líquido peritoneal, líquido pleural, pus, reuma, saliva, condensados de aliento exhalado, sebo, semen, esputo, sudor, líquido sinovial, lágrimas, vómito, líquido prostático, líquido de aspirado de pezón, líquido lacrimal, transpiración, hisopos de mejilla, lisado celular, líquido gastrointestinal, tejido de biopsia y orina u otro fluido biológico. En un aspecto, la muestra incluye ADN de una célula tumoral circulante.

[0147] Tal como se describió anteriormente, la muestra biológica puede ser una muestra sanguínea. La muestra sanguínea puede obtenerse mediante métodos conocidos en la técnica, como la punción digital o la flebotomía. Convenientemente, la muestra de sangre es de alrededor de 0,1 a 20 ml, o alternativamente de alrededor de 1 a 15 ml, siendo el volumen de sangre de aproximadamente 10 ml. También pueden utilizarse cantidades menores, así como ADN libre circulante en sangre. El micromuestreo y la toma de muestras por biopsia con aguja, catéter, excreción o producción de fluidos corporales que contengan ADN también son fuentes potenciales de muestras biológicas.

[0148] Los métodos y sistemas de la descripción utilizan información de secuencias de ácidos nucleicos y, por lo tanto, pueden incluir cualquier método o dispositivo de secuenciación para realizar la secuenciación de ácidos nucleicos, incluida la amplificación de ácidos nucleicos, reacción en cadena de la polimerasa (PCR), secuenciación nanopore, secuenciación 454, secuenciación marcada por inserción. En algunos aspectos, la

metodología o los sistemas de la descripción utilizan sistemas como los proporcionados por Illumina, Inc, (incluidos, de modo no taxativo, sistemas HiSeq™ X10, HiSeq™ 1000, HiSeq™ 2000, HiSeq™ 2500, Genome Analyzers™, MiSeq™, NextSeq, NovaSeq 6000), Applied Biosystems Life Technologies (sistema SOLiD™, secuenciador iónico PGM™, secuenciador iónico Proton™) o Genapsys o BGI MGI y otros sistemas. El análisis de ácidos nucleicos también puede realizarse mediante sistemas proporcionados por Oxford Nanopore Technologies (GridiON™, MiniON™) o Pacific Biosciences (Pacbio™ RS II o Sequel I o II).

[0149] La presente invención incluye sistemas para realizar pasos de los métodos descritos y se describe parcialmente en términos de componentes funcionales y varios pasos de procesamiento. Dichos componentes funcionales y pasos de procesamiento pueden realizarse mediante cualquier número de componentes, operaciones y técnicas configuradas para realizar las funciones especificadas y lograr los diversos resultados. Por ejemplo, la presente invención puede emplear diversas muestras biológicas, biomarcadores, elementos, materiales, computadoras, fuentes de datos, sistemas y medios de almacenamiento, técnicas y procesos de recopilación de información, criterios de procesamiento de datos, análisis estadísticos, análisis de regresión y similares, que pueden llevar a cabo diversas funciones.

[0150] En consecuencia, la invención proporciona además un sistema no invasivo para subtipificar el cáncer de pulmón microcítico o el cáncer de pulmón no microcítico. En varios aspectos, el sistema incluye: (a) un secuenciador configurado para generar un conjunto de datos de secuenciación de todo el genoma de baja cobertura para una muestra; y (b) un sistema informático y/o procesador con funcionalidad para realizar un método de la invención.

[0151] En algunos aspectos, el sistema informático incluye además uno o más módulos adicionales. Por ejemplo, el sistema puede incluir uno o más de una unidad de extracción y/o aislamiento operable para seleccionar análisis de componentes genéticos adecuados, por ejemplo, fragmentos de cfDNA de un tamaño particular.

[0152] En algunos aspectos, el sistema informático incluye además un dispositivo de visualización. El dispositivo de visualización puede ser operable para mostrar una línea de ajuste de curva, una línea de ajuste de curva de referencia, y/o una comparación de ambas.

[0153] Los métodos para la subtipificación no invasiva del cáncer de pulmón microcítico o del cáncer de pulmón no microcítico según diversos aspectos de la presente invención

pueden implementarse de cualquier manera adecuada, por ejemplo, mediante el uso de un programa informático que funcione en el sistema informático. Como se analiza en la presente, un sistema ilustrativo, de acuerdo con varios aspectos de la presente invención, se puede implementar en conjunción con un sistema informático, por ejemplo, un sistema informático convencional que comprende un procesador y una memoria de acceso aleatorio, como un servidor de aplicaciones de acceso remoto, servidor de red, computadora personal o estación de trabajo. El sistema informático también incluye de manera adecuada también dispositivos de memoria o sistemas de almacenamiento de información adicionales, tal como un sistema de almacenamiento en masa y una interfaz de usuario, por ejemplo, un monitor convencional, teclado y dispositivo de rastreo. No obstante, el sistema informático puede incluir cualquier sistema informático adecuado y equipo asociado, y puede configurarse de cualquier manera adecuada. En una modalidad, el sistema informático comprende un sistema independiente. En otra modalidad, el sistema informático forma parte de una red informática que incluye un servidor y una base de datos.

[0154] El software necesario para recibir, procesar y analizar la información puede implementarse en un único dispositivo o en varios. El software puede ser accesible a través de una red de forma que el almacenamiento y el procesamiento de la información se realice de forma remota con respecto a los usuarios. El sistema según diversos aspectos de la presente invención y sus diversos elementos proporcionan funciones y operaciones para facilitar la detección y/o el análisis, como la recopilación de datos, el procesamiento, el análisis, la elaboración de informes y/o el diagnóstico. Por ejemplo, en el presente aspecto, el sistema informático ejecuta el programa informático, que puede recibir, almacenar, buscar, analizar y comunicar información relativa al genoma humano o a una región de este. El programa informático puede comprender múltiples módulos que realizan diversas funciones u operaciones, como un módulo de procesamiento para procesar datos sin procesar y generar datos complementarios y un módulo de análisis para analizar datos sin procesar y datos complementarios para generar evaluaciones cuantitativas de un modelo de estado de la enfermedad y/o información de diagnóstico.

[0155] Los procedimientos realizados por el sistema pueden comprender cualquier proceso adecuado para facilitar el análisis y/o la subtipificación del cáncer de pulmón microcítico o del cáncer de pulmón no microcítico. En una modalidad, el sistema está configurado para establecer un modelo de subtipo de enfermedad y/o determinar el subtipo

de enfermedad en un paciente. Determinar o identificar el subtipo de enfermedad puede incluir la generación de cualquier información útil sobre el estado del paciente en relación con la enfermedad, como realizar un diagnóstico, proporcionar información útil para un diagnóstico, evaluar la etapa o el progreso de una enfermedad, identificar una condición que pueda indicar una susceptibilidad a la enfermedad, identificar si se pueden recomendar más pruebas, predecir y/o evaluar la eficacia de uno o más programas de tratamiento, o evaluar de otro modo el estado de la enfermedad, la probabilidad de enfermedad, u otro aspecto de la salud del paciente.

[0156] Los siguientes ejemplos se proporcionan para ilustrar adicionalmente las ventajas y características de la invención, pero no se pretende que limiten el alcance de la invención. Si bien este ejemplo es típico de los que podrían utilizarse, también pueden utilizarse otros procedimientos, metodologías o técnicas conocidos por los expertos en la técnica.

EJEMPLO 1

[0157] Disección de subtipos de cáncer de pulmón microcítico mediante fragmentomas de ADN libre de células

[0158] ANTECEDENTES: El cáncer de pulmón microcítico (SCLC) es una forma agresiva de cáncer de pulmón, fuertemente asociada al tabaquismo y a la exposición a otras sustancias químicas ambientales. Los pacientes con SCLC presentan una tasa de supervivencia a 5 años inferior al 8 % (Gay, C. M. *et al.* Patterns of transcription factor programs and immune pathway activation define four major subtypes of SCLC with distinct therapeutic vulnerabilities. *Cancer Cell* **39**, 346-360.e7 (2021)). El SCLC es un tipo de tumor heterogéneo formado por células tumorales con características neuroendocrinas y no neuroendocrinas. El SCLC presenta dos subtipos principales: Neuroendócrino de alto grado (NE alto) frente a neuroendócrino de bajo grado (NE bajo) (Gay *et al.*). Los subtipos NE alto se caracterizan por la activación de factores de transcripción específicos de linaje: *ASCL1* y *NEUROD1*. NE bajo se caracteriza por factores no neuroendócrinos como el *POU2F3* (SCLC-P) o la elevada presencia de linfocitos T inflamatorios (SCLC-I). A pesar de estas heterogeneidades moleculares y clínicas, el SCLC se trata como una entidad única con resultados previsiblemente desfavorables. Resultados recientes mostraron que el grupo inflamado (SCLC-I) tiene una biología distinta y responde mejor a la inmunoterapia en comparación con los otros subtipos de SCLC (Gay *et al.* y Lissa, D. *et al.* Heterogeneity of

neuroendocrine transcriptional states in metastatic small cell lung cancers and patient-derived models. *Nat Commun* **13**, 2023 (2022)). Los fragmentomas libres de células basados en DELFI diferencian con precisión los subtipos de SCLC NE alto frente a NE bajo de forma no invasiva. El objetivo final de la subtipificación DELFI del SCLC es guiar la selección del tratamiento óptimo para cada paciente diagnosticado de SCLC avanzado.

[0159] MÉTODOS: El ADN libre de células circulantes (cfDNA) se aisló a partir de muestras de plasma de pacientes diagnosticados de SCLC recidivante y tratados con durvalumab más olaparib en un ensayo de fase II (NCT02484404). Los datos inmunohistoquímicos y genómicos de las biopsias tisulares previas al tratamiento se utilizaron para clasificar los subtipos de SCLC en NE alto (n=10) frente a NE bajo (n=5). Para inferir los perfiles de expresión génica tumoral en el cfDNA, investigamos las señales genómicas de los factores de transcripción específicos del tejido regulados de forma diferencial en el SCLC y aplicamos un nuevo enfoque basado en DELFI para informar sobre los subtipos moleculares del SCLC. En detalle, DELFI calcula la distribución de fragmentos para cada TFBS dentro de una ventana de 2kb y escala entre 0-1 independientemente para cada muestra. A continuación, se realizó un análisis de componentes principales en R y se definieron grupos basados en los sitios TFBS de ASCL1 (Mathios, D. et al. Detection and characterization of lung cancer using cell-free DNA fragmentomes. *Nat Commun* **12**, 5060 (2021)). La información clínica se examinó en análisis ortogonales.

[0160] RESULTADOS: La plataforma fragmentómica interna de DELFI detecta casos de SCLC (N=47) con alta sensibilidad, con una puntuación DELFI mediana de 1,0 (IC 95 % 0,99-1), coincidiendo con datos previos de muestras internas de SCLC de DELFI. Los análisis de fragmentación del cfDNA en todo el genoma en los sitios de unión a ASCL1 (~12 000 coordenadas genómicas) en los pacientes con SCLC revelan una disminución de la cobertura cerca de los sitios de unión a factores de transcripción de los pacientes con SCLC en comparación con las muestras de pacientes sin cáncer y NSCLC. La diferencia en el centro de los sitios de unión de ASCL1 fue el principal factor de diferenciación entre las muestras de SCLC también pretratadas. La investigación de la distribución de los fragmentos en los sitios de unión de ASCL1 fue capaz de diferenciar las muestras de NE alto de las de NE bajo. Los pacientes (tratados con I/O) clasificados como NE bajo por el clasificador de subtipificación DELFI tuvieron una mejor respuesta al tratamiento que los clasificados como NE alto por DELFI. La expresión diferencial de pseudogenes identificó

miles de sitios de inicio de la transcripción (TSS) altamente enriquecidos en los casos de NE alto en comparación con NE bajo. Las muestras LOW-NE mostraron un fuerte enriquecimiento para el factor KLF4 específico de los linfocitos T, lo que refleja la relación de neutrófilos con respecto a linfocitos de este subtipo de SCLC. Estos datos reflejan el fenotipo inflamado, lo que posiblemente explica por qué estos subtipos de cáncer tienen más probabilidades de responder al tratamiento de I/O.

[0161] Tabla 1. Características de los pacientes e información clínica asociada

Característica	Conjunto de validación, n=20
Edad, mediana de años (intervalo)	63 (41, 75)
Sexo, n (%)	
Masculino	9 (45)
Femenino	11 (55)
TIL, n (%)	
Desertó	2 (10)
Excluido	9 (45)
Inflamado	3 (15)
No realizado	6 (30)
Metástasis, n (%)	
0	7 (35)
1	13 (65)
Desconocida	0 (0)
PFS	
Sin progresión ni muerte, n (%)	0 (0)
Progresión o muerte, n (%)	20 (100)
Mediana de PFS, meses (intervalo)	16,7 (7,4, 46,2)
Mediana de la duración del tratamiento, días (intervalo)	55 (20, 494)

EJEMPLO 2

[0162] Uso del enfoque basado en cfDNA circulante DELFI para subtipificar el cáncer de pulmón microcítico (SCLC) de forma no invasiva mediante el uso de fragmentomas de ADN libre de células.

[0163] El SCLC es una neoplasia agresiva de pronóstico desfavorable. Aunque los SCLC se gestionan clínicamente como un único tipo de cáncer, nuevas evidencias apoyan que estos subtipos (neuroendócrino de alto grado frente a de bajo grado) de SCLC adquieren diversos estados transcripcionales y epigenéticos. Además, los distintos subtipos de SCLC responden a tratamientos específicos, como la inmunoterapia en el caso del subtipo de SCLC neuroendócrino de bajo grado y alta inflamación.

[0164] El novedoso enfoque no invasivo basado en DELFI descrito en la presente distingue entre subtipos de SCLC neuroendócrino de alto grado y de bajo grado al investigar fragmentos libres de células que reflejan los estados epigenéticos únicos de los SCLC. Este novedoso método es capaz de diferenciar a los pacientes que probablemente responderán a la inmunoterapia, como el bloqueo de puntos de control inmunitarios.

[0165] Mediante el uso de datos de WGS de cfDNA derivado de plasma, se investigó la cobertura de fragmentos en coordenadas genómicas específicas para revelar el subtipo específico de casos de SCLC. El análisis dirigido basado en DELFI es capaz de diferenciar entre los casos de SCLC neuroendócrino de alto grado frente a neuroendócrino de bajo grado.

[0166] Los subtipos de SCLC pueden predecirse a partir del enfoque basado en DELFI mediante el uso de datos de cfDNA LC-WGS.

[0167] En primer lugar, el clasificador de aprendizaje automático DELFI detecta la presencia de cáncer en pacientes con SCLC. Se investigaron tanto los perfiles de fragmentación del genoma completo como las puntuaciones DELFI correspondientes para identificar diferencias preliminares entre los dos subtipos clínicos principales de casos de SCLC: neuroendócrino de alto grado y neuroendócrino de bajo grado. Posteriormente, se utilizaron los datos públicos disponibles para realizar un análisis de componentes principales con el fin de agrupar los dos subtipos.

[0168] En segundo lugar, mediante el uso de datos disponibles públicamente, se determinó que estos grupos de muestras de SCLC presentan una disminución en la cobertura de fragmentos agregados en los sitios de unión del factor de transcripción ASCL1,

lo que clasifica estos casos como SCLC neuroendócrinos de alto grado.

[0169] En tercer lugar, también se investigó si uno de estos grupos de muestras de SCLC presentaba una reducción de la cobertura de fragmentos en los sitios de unión genómica regulados por factores de transcripción hematopoyéticos (SCLC neuroendócrino de bajo grado).

[0170] En cuarto lugar, los casos neuroendócrinos de alto grado frente a de bajo grado se distinguieron en función de la distribución de la longitud de los fragmentos en los dominios parcialmente metilados específicos de linfocitos T.

[0171] Por último, se utilizó una característica operativa del receptor (ROC) que representa la sensibilidad y especificidad del enfoque DELFI-fragmentoma para identificar subtipos de SCLC con el fin de identificar la precisión diagnóstica de nuestro enfoque entre las muestras de grupo identificadas.

[0172] Dado el limitado acceso a las biopsias de tumores de SCLC, se cree que este enfoque podría ser un método viable para subtipificar el SCLC de forma no invasiva. Un ensayo prometedor tanto para las empresas farmacéuticas durante los ensayos clínicos para evaluar nuevos fármacos de inmunoterapia como para los clínicos que buscan seleccionar qué pacientes diagnosticados de SCLC tienen más posibilidades de responder a la inmunoterapia.

[0173] Las puntuaciones DELFI de los pacientes diagnosticados de SCLC en este estudio se muestran en la Figura 1.

[0174] La **Figura 2** ilustra que los perfiles de fragmentación de todo el genoma fueron significativamente consistentes entre los puntos temporales de pretratamiento, postratamiento, progresión y respuesta, lo que sugiere que el tamaño de los fragmentos de ADN tumoral circulante de todo el genoma es un método potente para detectar cambios durante el monitoreo del tratamiento de inmunoterapia del cáncer SCLC.

[0175] La **Figura 3** ilustra los perfiles de fragmentación de todo el genoma (y las puntuaciones DELFI correspondientes) que muestran diferencias parciales entre los dos subtipos principales de casos de SCLC: neuroendócrino de alto grado y neuroendócrino de bajo grado; lo que sugiere que el tamaño de los fragmentos de ADN tumoral circulante en todo el genoma puede ser un método potente para detectar la subtipificación durante el seguimiento del tratamiento con inmunoterapia del cáncer SCLC.

[0176] La **Figura 4** ilustra los genes diferenciales entre los casos de SCLC

neuroendócrino de alto grado y de bajo grado, mediante el uso de datos disponibles públicamente (Lissa et al. 2022) y un Análisis de Componentes Principales (sobre datos WGS de plasma DELFI 30X) para agrupar los dos subtipos. Los datos del análisis PCA revelaron grupos distinguibles entre las muestras de pretratamiento de SCLC NE alto frente a NE bajo.

[0177] La **Figura 5** ilustran un análisis Eigencor que revela una correlación significativa entre los componentes principales 1 y 2 con el estado NE de las muestras.

[0178] Las **Figuras 6 A-B** ilustran el potencial del ensayo DELFI para subtipificar el SCLC mediante el uso de TFBS, se realizaron análisis de fragmentación del cfDNA de todo el genoma en los sitios de unión de ASCL1 en las muestras de NCI pretratamiento. La Figura 5A ilustra que se observaron distintos grupos de muestras de SCLC correspondientes a diferentes niveles de activación de ASCL1. Es importante destacar que los dos grupos correspondían perfectamente a los diferentes subtipos de SCLC. El principal factor de diferenciación entre las muestras fueron, de hecho, las diferencias en la cobertura de los fragmentos identificados en el centro de los sitios de unión a ASCL1. La Figura 6B ilustra que otras características clínicas o de la muestra tienen un impacto insignificante sobre la agrupación de estas muestras. En general, estos datos muestran cómo el análisis DELFI puede distinguir perfectamente entre los subtipos de SCLC mediante el uso de la cobertura de fragmentos en TFBS.

[0179] La **Figura 7** ilustra un paciente ilustrativo que apoya la hipótesis de que la señal detectada procede de linfocitos infiltrantes de tumores. El paciente NCI-0422, diagnosticado de SCLC con subtipo inflamatorio fue tratado con Durvalumab más Olaparib. NCI-0422 respondió bien al tratamiento, pero posteriormente se le diagnosticó enfermedad progresiva. La unión genómica variable es detectable antes del tratamiento y durante la progresión. Los picos se centran en una mezcla de TFBS y TSS.

[0180] Las **Figuras 8 A-B** ilustran ASCL1 sitio de unión entre respondedores frente a no respondedores. La Figura 8A muestra 500 bins específicos de tipo celular de los datos de PMD para determinar la fracción de glóbulos blancos. La Figura 8B muestra los bins específicos de tejido linfocitario.

[0181] La **Figura 9** ilustra un modelo de aprendizaje automático para predecir un SCLC neuroendócrino de alto grado frente a de bajo grado.

[0182] La subtipificación se realiza al calcular el log2 (relación de profundidad de lectura)

a través de ventanas de 20 pb comenzando a 2kb de los sitios ASCL1 TFBS. La relación de profundidad de lectura es la cobertura total sobre un factor de corrección de cobertura calculado mediante el uso de la cobertura mediana de anclajes de 5' y 3' 500pb en los extremos de la ventana de 2kb (es decir, $\log_2(\text{profundidad de lectura}/\text{factor de corrección})$). A continuación, se utiliza un modelo aditivo general para suavizar la cobertura y corregir el sesgo de la GC. Para validar el enfoque descrito en la presente, el estado neuroendócrino se deriva de la puntuación NE50 media por sujeto/estado de tratamiento, un valor positivo se clasifica como NE+. El estado del subtipo se deriva de la agregación de la media para cada uno de los 4 subtipos (ASCL1, NEUROD1, POUF23, YAP1) de tal forma que cada estado de sujeto/tratamiento tiene un valor para cada subtipo. A continuación, se selecciona como verdadero el subtipo con la puntuación más alta. Los fragmentos se calculan como se ha descrito anteriormente (Cristiano et al.), las puntuaciones corresponden a los últimos modelos disponibles.

[0183] Las **Figuras 10 A-B** ilustra que el método descrito en la presente es capaz de diferenciar neuroendócrino frente a no neuroendócrino mediante el cálculo de relación de profundidad de lectura en coordenadas genómicas específicas de SCLC (Figura 10A). La Figura 10B muestra que el método descrito en la presente parece subtipificar mejor los casos de SCLC.

[0184] La **Figura 11** ilustra que la relación de profundidad de lectura en coordenadas genómicas específicas de SCLC se transforman en un PCA bidimensional para separar las muestras en grupos neuroendócrinos frente a no neuroendócrinos. La mayoría de las muestras están claramente divididas en distintos grupos.

[0185] La **Figura 12** ilustra PCA derivado de relación de profundidad de lectura en coordenadas genómicas específicas de SCLC es capaz de separar las muestras pretratamiento en subtipos específicos de SCLC (A,N,P,Y). La verdad fundamental de los subtipos de SCLC se calculó mediante el análisis diferencial de la expresión génica de secuenciación de ARN tisular.

[0186] La plataforma fragmentómica descrita en la presente detecta el cáncer de pulmón microcítico (SCLC) con alta sensibilidad. El ensayo de subtipificación de SCLC DELFI es capaz de subtipificar muestras de SCLC, sin necesidad de conocimientos clínicos. El ensayo de subtipificación de SCLC DELFI identificó cuatro grupos de muestras de SCLC pertenecientes a diferentes niveles de activación de ASCL1. Se confirmó que las muestras

con alta activación de ASCL1 pertenecían a subtipos neuroendócrinos (SCLC-A y SCLC-N), mientras que las muestras con baja señal de ASCL1 pertenecían a subtipos no neuroendócrinos (SCLC-P y SCLC-Y).

EJEMPLO 3

[0187] Perfil de fragmentación del ADN libre de células como método para la evaluación de la fracción tumoral y el monitoreo del tratamiento en NSCLC

[0188] Antecedentes fragmentómicos de cfDNA: Tradicionalmente, el monitoreo se realiza mediante técnicas de imagenología, pero muchas personas no tienen fácil acceso a hospitales que cuenten con equipos de imagenología y conocimientos adecuados, y deben recorrer distancias considerables para someterse a estos procedimientos, lo que dificulta el monitoreo continuo. No es necesario saber *a priori* dónde se encuentra un tumor. Además, no es necesario conocer de antemano las mutaciones somáticas que alberga un tumor. Los costos asociados a las biopsias líquidas son baratos en comparación con el diagnóstico por imagenología, lo que debería mejorar a medida que los costos de secuenciación sigan disminuyendo.

[0189] Las mutaciones del cfDNA tienen una señal limitada y pueden confundirse por la hematopoyesis clonal. DELFI capta muchas características del universo del cfDNA. Los patrones de fragmentación del cfDNA están determinados por la organización basal de la cromatina. Los perfiles de fragmentación del cfDNA son muy consistentes en personas sanas y están alterados en pacientes con cáncer.

[0190] Descripción general del modelo DELFI-TF y aplicación al seguimiento de pacientes con NSCLC durante la terapia: se analizaron alícuotas de cfDNA de muestras de plasma de pacientes con CCR mediante ddPCR para determinar el estado de mutación RAS, así como secuenciación WGS de paso bajo. Se alinearon los datos WGS y se obtuvieron distribuciones del tamaño de los fragmentos para 504 intervalos de 5 Mb en todo el genoma. Se entrenó un modelo de regresión bayesiana y se validó de forma cruzada con muestras de RAS MT mediante el uso de características fragmentómicas. La Figura 20 muestra el desarrollo del modelo DELFI-TF de prueba de concepto.

[0191] Se aplicó un modelo DELFI-TF entrenado para CRC a una cohorte de NSCLC del mundo real.

Tabla 2 - Características clínicas de la cohorte de NSCLC

Características clínicas	
Mediana de edad (intervalo)	71,5 (53-88)
Sexo	
Masculino	11
Femenino	10
Mediana de IMC (intervalo)	24,5 (17-37)
Etapa	
Etapa III	7
Etapa IV	14
Expresión de PD-L1	
Positiva fuerte	5
Positiva débil	7
Negativa	4
NC	5
Tipo de tratamiento	
Inmunoterapia	3
Inmunoterapia + quimioterapia	9
Quimioterapia	9
Mejor respuesta general	
Respuesta completa	1
Respuesta parcial	2
Enfermedad estable	1
Enfermedad progresiva	10
Tratamiento en curso	7

[0192] Resultados: La **Figura 13** ilustra representando cómo los perfiles de fragmentación informan sobre el estado del cáncer de pulmón y los patrones de respuesta al tratamiento.

[0193] La **Figura 14** ilustra que DELFI-TF derivado del modelo muestra una fuerte correlación con la Frecuencia Alélica Mutante en todas las muestras.

[0194] La **Figura 15** ilustra que DELFI-TF cuantifica con precisión la fracción tumoral circulante, que refleja el rendimiento de MAF en relación con la evaluación RECIST.

[0195] La **Figura 16** ilustra que el cfDNA derivado de tumores presenta una fragmentación alterada.

[0196] Las **Figuras 17 A-C** ilustran que DELFI-TF detecta con precisión la fracción tumoral circulante sin confundirse por la hematopoyesis clonal.

[0197] La **Figura 18** muestra los patrones de fragmentación del cfDNA para diferenciar con precisión los subtipos de NSCLC.

[0198] Conclusiones: Los perfiles de fragmentos de cfDNA de todo el genoma son anómalos en pacientes con cáncer

[0199] Las puntuaciones de fragmentación del cfDNA (DELFITF) están altamente correlacionadas con las frecuencias alélicas de mutación conocidas. La fragmentación del cfDNA predice el estado RECIST. La fragmentación del cfDNA no se ve afectada por la hematopoyesis clonal. Las características de fragmentación del cfDNA pueden distinguir de forma no invasiva subtipos histológicos de cánceres de pulmón.