

(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS
ESPAÑA



(11) Número de publicación: **2 993 503**

(51) Int. Cl.:

B01J 19/00 (2006.01)
G01N 33/48 (2006.01)
G01N 33/483 (2006.01)
G01N 33/543 (2006.01)
G01N 33/58 (2006.01)
G01N 33/68 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Fecha de presentación y número de la solicitud europea: **18.12.2017 E 21216039 (4)**

(97) Fecha y número de publicación de la concesión europea: **28.08.2024 EP 4056263**

(54) Título: **Sistema y método para matriz de sensores de corona de proteínas para la detección temprana de enfermedades**

(30) Prioridad:

16.12.2016 US 201662435409 P

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

02.01.2025

(73) Titular/es:

**THE BRIGHAM AND WOMEN'S HOSPITAL, INC.
(100.00%)
75 Francis Street
Boston, MA 02115, US**

(72) Inventor/es:

**FAROKHZAD, OMID;
MAHMOUDI, MORTEZA y
CORBO, CLAUDIA**

(74) Agente/Representante:

SÁEZ MAESO, Ana

ES 2 993 503 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Sistema y método para matriz de sensores de corona de proteínas para la detección temprana de enfermedades

Antecedentes de la invención

5 El campo de la invención se relaciona con matrices de sensores para la detección y diagnóstico de diferentes estados de enfermedad, en particular, la invención se relaciona con la capacidad de diagnosticar o pronosticar enfermedades o trastornos.

10 Cuanto antes se diagnostique una enfermedad, más probable es que la enfermedad se pueda curar o controlar con éxito, lo que conduce a un mejor pronóstico para el paciente. Cuando trata una enfermedad a tiempo, es posible que pueda prevenir o retrasar los problemas de la enfermedad y puede mejorar los resultados para el paciente, incluida la prolongación de la vida del paciente y/o la calidad de vida.

15 El diagnóstico temprano del cáncer es crucial, ya que muchos tipos de cáncer pueden tratarse con éxito en sus etapas iniciales. Por ejemplo, la supervivencia a cinco años después del diagnóstico y tratamiento tempranos de los cánceres de mama, ovario y pulmón es del 90 %, 90 % y 70 %, respectivamente, en comparación con el 15 %, 5 % y 10 % para los pacientes diagnosticados en la etapa más avanzada de la enfermedad. Una vez que las células cancerosas abandonan su tejido de origen, es muy poco probable que el tratamiento exitoso utilizando las terapias establecidas disponibles. Aunque reconocer las señales de advertencia de los cánceres y tomar medidas inmediatas puede conducir a un diagnóstico temprano, la mayoría de los cánceres (por ejemplo, de pulmón) muestran síntomas solo después de que las células cancerosas ya han invadido los tejidos circundantes y han hecho metástasis en todo el cuerpo. Por ejemplo, más del 60 % de los pacientes con cáncer de mama, pulmón, colon y ovario tienen colonias ocultas o incluso metastásicas en el momento en que se detecta el cáncer. Por lo tanto, existe una necesidad urgente de desarrollar un enfoque eficaz para la detección temprana del cáncer. Tal enfoque debe tener la sensibilidad para identificar un cáncer en varias etapas y la especificidad para dar un resultado negativo cuando la persona que se somete a la prueba no tiene cáncer. Ha habido grandes esfuerzos para desarrollar métodos para la detección temprana de cánceres; aunque se han introducido una gran cantidad de factores de riesgo y biomarcadores, sigue siendo difícil de alcanzar una plataforma ampliamente relevante para la detección temprana de un amplio rango de cánceres.

30 Como diversos tipos de cáncer pueden cambiar la composición del plasma sanguíneo, incluso en sus primeras etapas, un enfoque prometedor para la detección temprana es el análisis molecular de sangre para biomarcadores. Aunque esta estrategia ya ha funcionado para algunos tipos de cáncer (como el PSA para el cáncer de próstata), aún no existen biomarcadores específicos para la detección temprana de la mayoría de los cánceres. Para tales cánceres (por ejemplo, pulmón), ninguno de los biomarcadores circulantes candidatos definidos ha sido validado clínicamente, y muy pocos han alcanzado el desarrollo clínico en etapa tardía. Por lo tanto, existe una necesidad urgente de novedosos enfoques para mejorar nuestra capacidad para detectar el cáncer en etapas muy tempranas. M. Rahman, M. Mahmoudi, "Disease specific protein corona," Proc. SPIE 9338, Colloidal Nanoparticles for Biomedical Applications X, 93380V (12 de marzo de 2015) divulga el uso de patrones de corona de proteínas en una nanopartícula para el diagnóstico y tratamiento de diferentes enfermedades. No divulgá huellas biomoleculares de acuerdo con las presentes reivindicaciones.

Sumario de la invención

40 La presente invención se refiere a un método para detectar un estado de enfermedad tal como se define en las reivindicaciones adjuntas. La singularidad de la presente invención es la combinación del reconocimiento de una huella biomolecular a partir de una muestra de un sujeto y la capacidad de determinar un estado de enfermedad para ese sujeto en un continuo de salud.

45 Lo anterior y otros aspectos y ventajas de la invención aparecerán a partir de la siguiente descripción. En la descripción, se hace referencia a los dibujos adjuntos que forman parte de la misma, y en los que se muestra a modo de ilustración una realización preferida de la invención. Sin embargo, dicha realización no representa necesariamente el alcance completo de la invención y, por lo tanto, se hace referencia a las reivindicaciones y en el presente documento para interpretar el alcance de la invención.

Breve descripción de los dibujos

50 Figura 1. Esquema de una realización que muestra un ejemplo de diseño de estudio del enfoque de patrón de corona de proteínas para la detección de cáncer. Se incuban tres tipos de liposomas con plasma de personas sanas y pacientes con cáncer, y el patrón de corona de proteínas que se forma en cada liposoma en el plasma de cada sujeto (sano y con diferentes tipos de cáncer) se caracteriza por cromatografía líquida-espectrometría de masas en tandem (LC-MS/MS). La formación de coronas de proteínas en los tres liposomas da como resultado el enriquecimiento de un conjunto superpuesto pero distinto de proteínas plasmáticas seleccionadas, y las proteínas enriquecidas son la base para el análisis multivariante posterior. A través de enfoques de clasificación, las proteínas importantes en los patrones de corona se identificaron y utilizaron para predecir

cánceres utilizando plasmas ciegos y muestras de cohortes para probar la precisión del nanosistema de corona de proteínas de múltiples nanopartículas. Se utilizaron 29 sujetos humanos (25 pacientes cancerosos, es decir, 5 pacientes por 5 tipos de cáncer y 4 sujetos sanos) que representan 261 ejecuciones distintas de LC-MS/MS (3 liposomas, 29 sujetos, 3 réplicas) para entrenar un modelo de clasificación. Se usaron 16 sujetos humanos (3 pacientes por 5 tipos de cáncer y 1 sujeto sano) de plasma ciego y 144 ejecuciones distintas de LC-MS/MS (3 liposomas, 16 sujetos, 3 réplicas) para la predicción del cáncer, es decir, para probar el modelo de clasificación. También se utilizaron 15 sujetos humanos (5 pacientes por 3 tipos de cáncer) de plasma de cohorte que representan 135 ejecuciones distintas de LC-MS/MS (3 liposomas, 15 sujetos, 3 réplicas) para la predicción muy temprana del cáncer.

10 Figura 2A. Imágenes TEM de liposomas con sus perfiles de distribución de tamaño.

Figura 2B. Propiedades fisicoquímicas de diferentes liposomas antes y después de interacciones con plasma humano de pacientes con diferentes enfermedades. Datos de DLS y Zeta-Potencial en varios liposomas antes de las interacciones con plasma humano y complejos corona libres de exceso de plasma, obtenidos después de la incubación con plasma de pacientes sanos y con cáncer (Pdi: Índice de polidispersidad del ajuste 15 cumulante).

Figura 2C. Clasificación de las coronas de proteína identificadas a partir de liposomas de acuerdo con sus funciones fisiológicas en plasma humano de individuos sanos y de pacientes con diferentes tipos de cáncer (los datos presentados reflejan un cálculo a partir de cinco plasmas biológicos por grupo y tres réplicas técnicas por plasma)

20 Figura 3A. Clasificación de coronas identificadas por elementos de matriz de sensores de acuerdo con sus funciones fisiológicas, incluida la fase aguda, en plasma humano de sujetos sanos y pacientes con diferentes tipos de cáncer.

Figura 3B. Clasificación de coronas identificadas por elementos de matriz de sensores de acuerdo con sus 25 funciones fisiológicas, incluida la coagulación en plasma humano de sujetos sanos y pacientes con diferentes tipos de cáncer.

Figura 3C. Clasificación de coronas identificadas por elementos de matriz de sensores de acuerdo con sus funciones fisiológicas, incluidas las inmunoglobulinas en plasma humano de sujetos sanos y pacientes con diferentes tipos de cáncer.

30 Figura 3D. Clasificación de coronas identificadas por elementos de matriz de sensores de acuerdo con sus funciones fisiológicas, incluidas las lipoproteínas en plasma humano de sujetos sanos y pacientes con diferentes tipos de cáncer.

Figura 3E. Clasificación de coronas identificadas por elementos de matriz de sensores de acuerdo con sus 35 funciones fisiológicas, incluida la fuga de tejido en plasma humano de sujetos sanos y pacientes con diferentes tipos de cáncer.

Figura 3F. Clasificación de coronas identificadas por elementos de matriz de sensores de acuerdo con sus funciones fisiológicas, incluidas las proteínas del complemento en plasma humano de sujetos sanos y pacientes con diferentes tipos de cáncer.

40 Figura 3G. Clasificación de coronas identificadas por elementos de matriz de sensores de acuerdo con sus funciones fisiológicas, incluidas otras proteínas plasmáticas en plasma humano de sujetos sanos y pacientes con diferentes tipos de cáncer.

Figura 4A. Descubrimiento de predictores y contribución de cada predictor individual a la separación de cada clase mediante análisis de discriminación PLS. La exploración de predictores por VIP ponderado se realizó agregando las variables clasificadas al modelo PLS-DA una por una y calculando el error de clasificación para la validación cruzada interna (10 veces). La disminución del error de clasificación condujo al descubrimiento de 45 un conjunto mínimo de 69 predictores con la mayor importancia posible para separar cada clase de las demás. La contribución de cada marcador individual a la separación de cada clase según el análisis de discriminación PLS. Marcadores de clasificación de gráficos VIP de 69 variables seleccionadas por su contribución a la separación de cada clase del análisis de discriminación PLS. La puntuación $VIP > 1$ indica una proteína importante que conduce a una buena predicción de la pertenencia a una clase, mientras que las variables con puntuaciones $VIP < 1$ indican proteínas sin importancia para cada clase.

50 Figura 4B representa los resultados de las 69 variables para el glioblastoma.

Figura 4C. Representa los valores VIP para las 69 variables de meningioma.

Figura 4D. Representa los valores VIP para las 69 variables para el cáncer de pulmón.

Figura 4E. Representa los valores VIP para las 69 variables del mieloma.

Figura 4F. Representa los valores VIP para las 69 variables del cáncer de páncreas.

Figura 5A. Gráfico de PLS-DA que muestra la separación de diferentes muestras cancerosas entre sí y de los controles. Gráfico de puntuación de PLS obtenido usando PLS-toolbox, proyectando los objetos en el subespacio creado por la primera, segunda y tercera variables latentes del modelo.

5 Figura 5B. Gráfico de PLS-DA que muestra la separación de diferentes muestras cancerosas entre sí y de los controles. Objetos visualizados donde se muestran la 4^a y 5^a variables latentes del modelo. Como se puede observar, los casos de meningioma y glioblastoma no se separaron en tres dimensiones adecuadamente, pero se pueden separar en la cuarta y quinta dimensiones del modelo PLS.

10 Figura 5C. Mapa de asignación obtenido por CPANN con todas las variables y variables seleccionadas. (C) El mapa de asignación obtenido mediante el entrenamiento de una red CPANN (8 × 8 neuronas) utilizando un conjunto de datos completo (1823 variables). La calidad del mapeo no es buena y existen conflictos de diferentes tipos de cáncer en términos de mapeo.

15 Figura 5D. Mapa de asignación obtenido mediante el entrenamiento de una red CPANN (8×8 neuronas) utilizando 69 variables. Los vectores de entrada de alta dimensión (muestras) se mapean en una red bidimensional de neuronas, preservando la similitud y la topología. Los colores indican la similitud de una neurona con un tipo específico de vector de entrada (tipo de clase). Este mapa también demuestra la importancia del paso de selección del predictor y el efecto de la eliminación de predictores no informativos e irrelevantes en la calidad del modelo.

20 Figura 5E. Dendrograma que representa las 51 proteínas identificadas como capaces de distinguir entre los seis grupos de cáncer.

25 Figura 5F. Las 51 proteínas identificadas como capaces de distinguir entre los seis grupos se presentan en un 'mapa de calor' generado con un algoritmo de agrupación no supervisado (HCA aglomerativo con enlace de vecino más lejano). Inspección visual tanto del dendrograma (Figura 5E) y mapa de calor (Figura 5F) demuestra la firma de la corona de proteína específica del cáncer y la agrupación clara de seis grupos de muestras (cinco grupos de muestras cancerosas más muestras normales) y también las similitudes esperadas entre cinco pacientes de cada grupo. El mapa de calor también indica diferencias sustanciales en los patrones de variables (marcadores) de diferentes tipos de cáncer (cada columna representa un paciente y cada fila representa una proteína). Los niveles de proteína más altos y más bajos se indican en rojo y verde, respectivamente.

30 Figura 6A. Gráfico de puntuación de PLS obtenido al considerar 69 marcadores importantes, proyectando los objetos de la cohorte en el subespacio creado por la primera y segunda variables latentes del modelo.

FG. 6B. El modelo PLS-DA se genera utilizando 8 variables que proyectan los objetos de la cohorte en el subespacio creado por la primera y segunda variables latentes del modelo, con excelentes estadísticas.

Figura 6C. El mapa de asignación obtenido mediante el entrenamiento de una red CPANN (8x8 neuronas) utilizando 69 marcadores importantes.

35 Figura 6D. El mapa de asignación obtenido mediante el entrenamiento de una red CPANN (8x8 neuronas) usando solo 8 marcadores sin errores de clasificación. Los números de muestra se indican en cada neurona.

Figura 7A. Representación esquemática del esquema de estudio. Construcción de modelos informativos de selección y clasificación de variables.

40 Figura 7B. Se enumeran el nombre de la proteína y la ID de 69 variables seleccionadas. Algunas de las proteínas estaban presentes en la corona de proteínas de más de un liposoma (DOPG, DOTAP y CHOL se indican con fuentes: cursiva y subrayado, negrita y sin formato, respectivamente).

Figura 7C. Los biomarcadores específicos de la enfermedad cubiertos como variables significativas por los modelos propuestos.

45 Figura 8A. Gráfico de características operativas del receptor (ROC) derivado de PLS-DA basado en las 69 variables mejor clasificadas para el control. Gráfica ROC de sensibilidad (tasa de verdaderos positivos, eje Y) versus 1 - especificidad (tasa de falsos positivos, eje X) basada en un PLS-DA basado en los 69 marcadores con la contribución más alta para seis clases (control, glioblastoma, meningioma, mieloma, páncreas, pulmón).

50 Figura 8B. Gráfico de características operativas del receptor (ROC) derivado de PLS-DA basado en las 69 variables mejor clasificadas para el glioblastoma. Gráfica ROC de sensibilidad (tasa de verdaderos positivos, eje Y) versus 1 - especificidad (tasa de falsos positivos, eje X) basada en un PLS-DA basado en los 69 marcadores con la contribución más alta para seis clases (control, glioblastoma, meningioma, mieloma, páncreas, pulmón).

Figura 8C. Gráfico de característica operativa del receptor (ROC) derivado de PLS-DA basado en las 69 variables mejor clasificadas para el meningioma. Gráfica ROC de sensibilidad (tasa de verdaderos positivos, eje Y) versus 1 - especificidad (tasa de falsos positivos, eje X) basada en un PLS-DA basado en los 69 marcadores con la contribución más alta para seis clases (control, glioblastoma, meningioma, mieloma, páncreas, pulmón).

Figura 8D. Gráfico de características operativas del receptor (ROC) derivado de PLS-DA basado en las 69 variables mejor clasificadas para el mieloma. Gráfica ROC de sensibilidad (tasa de verdaderos positivos, eje Y) versus 1 - especificidad (tasa de falsos positivos, eje X) basada en un PLS-DA basado en los 69 marcadores con la contribución más alta para seis clases (control, glioblastoma, meningioma, mieloma, páncreas, pulmón).

Figura 8E. Gráfico de característica operativa del receptor (ROC) derivado de PLS-DA basado en las 69 variables mejor clasificadas para el páncreas. Gráfica ROC de sensibilidad (tasa de verdaderos positivos, eje Y) versus 1 - especificidad (tasa de falsos positivos, eje X) basada en un PLS-DA basado en los 69 marcadores con la contribución más alta para seis clases (control, glioblastoma, meningioma, mieloma, páncreas, pulmón).

Figura 8F. Gráfico de característica operativa del receptor (ROC) derivado de PLS-DA basado en las 69 variables mejor clasificadas para pulmón. Gráfica ROC de sensibilidad (tasa de verdaderos positivos, eje Y) versus 1 - especificidad (tasa de falsos positivos, eje X) basada en un PLS-DA basado en los 69 marcadores con la contribución más alta para seis clases (control, glioblastoma, meningioma, mieloma, páncreas, pulmón).

Figura 9A. Representación esquemática del despliegue de una matriz de datos de tres vías en una matriz de dos vías.

Figura 9B. Mapa de asignación obtenido por CPANN (14×14) entrenado usando 90 muestras (réplicas) con todas las 1823 variables. Los números de muestra se indican en cada neurona. El color de la neurona (etiqueta asignada) se decide en función de la similitud entre la etiqueta de clase (un vector binario de 6×1) y el vector de peso en la capa de salida de la neurona correspondiente. A pesar de utilizar todos los biomarcadores, existen algunas similitudes distintas entre las muestras de la misma clase de cáncer. Las muestras replicadas también se mapean en neuronas adyacentes o en las mismas

Figura 9C. El error de clasificación del mapa CPANN se calculó en diferentes tamaños de mapa utilizando una validación cruzada de 10 veces.

Figura 9D. La red CPANN tiene 69 capas de peso, que es igual al número de variables utilizadas para entrenar el modelo. La $i^{\text{ésima}}$ capa de peso refleja el efecto de la $i^{\text{ésima}}$ variable (biomarcador) sobre el patrón del mapa de asignación.

Figura 9E La correlación del mapa de asignación y 69 capas de peso (mapas de peso) se puede calcular y podría ayudar a identificar los biomarcadores relacionados con cada clase de cáncer. También se puede decidir visualmente; la similitud puede ser monitoreada por el valor absoluto del Coeficiente de Correlación de dos mapas.

Figura 9F. La correlación del mapa de asignación y 69 capas de peso (mapas de peso) se puede calcular y podría ayudar a identificar los biomarcadores relacionados con cada clase de cáncer. También se puede decidir visualmente; la similitud puede ser monitoreada por el valor absoluto del Coeficiente de Correlación de dos mapas.

Figura 10A. Importancia de las proteínas para la clasificación frente al porcentaje de proteínas adsorbidas en el nanosistema de corona de proteínas. Los paneles (a)-(c) ilustran la importancia de las interacciones proteína-liposoma observadas ('variables') en la predicción de cánceres específicos. Las proteínas se agrupan por sus funciones fisiológicas. Los paneles (d)-(f) ilustran el porcentaje de proteínas adsorbidas en cada liposoma. Los grupos de proteína-liposoma que emergen como relevantes para la predicción de un cáncer son muy distintos entre los cánceres (paneles (a)-(c)). Además, esta distinción es sustancialmente más pronunciada que la variación en el porcentaje de proteínas adsorbidas en los liposomas en esos cánceres (paneles (d)-(f)).

Figura 10B. Diagrama de Venn que muestra el número de proteínas únicas identificadas en la composición de la corona de cada liposoma y sus combinaciones (la tabla de la derecha presenta los mismos datos numéricamente).

Figura 10C. Importancia de la variable para la clasificación. Cada fila indica la importancia de una proteína específica. Los tres puntos de cada fila corresponden a la importancia de la interacción observada con cada uno de los tres liposomas. Las líneas horizontales que cruzan los puntos indican los percentiles 25 y 75 de la importancia entre los clasificadores entrenados en 1000 sorteos aleatorios del conjunto de entrenamiento de los datos. Estos intervalos de confianza indican la "estabilidad" del modelo entrenado en términos de las interacciones proteína-liposoma en las que se basa de manera crucial, con respecto a las extracciones aleatorias de datos.

5 Figura 11 Los promedios ponderados de las interacciones proteína-liposoma clasifican los cánceres. Distribución de las puntuaciones z absolutas para cada grupo de pacientes, histogramadas en las 100 proteínas más abundantes (gris) y biomarcadores previamente identificados (blanco con puntos negros). La barra negra larga corresponde a la puntuación z para una combinación lineal de las interacciones proteína-liposoma. Una puntuación z grande para una interacción proteína-liposoma específica indica que el grupo se 'separa' del resto de los pacientes en esa interacción en particular. En consecuencia, la figura indica que mientras que ninguna interacción proteína-liposoma individual es suficiente para clasificar ninguno de los cánceres, sus combinaciones ponderadas inducen una separación de entre 2 y 2.5 desviaciones estándar.

10 Figura 12 Los multiliposomas concentran proteínas raras y poco abundantes. Contribución corona de proteína frente a concentración plasmática conocida, representada en escala logarítmica. Cada punto representa una única proteína y liposoma, con aporte corona para cada enfermedad e individuos sanos, y tanto el aporte corona como la concentración plasmática están normalizados con respecto a la albúmina. Las concentraciones de plasma varían en 10 órdenes de magnitud, mientras que la matriz de liposomas detecta estas mismas proteínas en 4-5 órdenes de magnitud. Las contribuciones de corona de proteínas cuya concentración plasmática se 15 desconoce/no se informa se representan en el área roja a la derecha.

20 Figura 13 Ejemplos de los tipos de elementos sensores a nanoscala que se pueden usar para algunas realizaciones de las matrices de sensores. Se pueden utilizar como elementos sensores diferentes tipos de nanopartículas (por ejemplo, nanopartículas orgánicas, inorgánicas y poliméricas) con diversas propiedades fisicoquímicas (por ejemplo, diferentes propiedades superficiales, tamaños y formas). La matriz de sensores se puede crear desde un mínimo de dos elementos hasta un número ilimitado de elementos.

Figura 14 Ejemplo de un método de recolección de partículas recubiertas de corona en condiciones *in vitro*, *ex vivo*, e *in vivo*. Las partículas se incuban con fluidos biológicos (por ejemplo, plasma de pacientes con diferentes tipos de enfermedades) y las partículas recubiertas de corona se recolectan y almacenan para su análisis.

25 Figura 15A. Ejemplos de conjugación de materiales de nanoobjetos (con diferentes propiedades fisicoquímicas) a sustratos (con diferentes propiedades fisicoquímicas) para hacer un chip de matriz de sensor de corona de proteína (A) antes de las interacciones con la fuente de proteína (por ejemplo, plasma humano de diversas enfermedades). La corona de proteína específica se formará en la superficie de los nanoobjetos, con diferentes propiedades fisicoquímicas. Los sustratos también pueden quedar recubiertos por varios tipos de proteínas, que tienen efectos insignificantes sobre la eficacia de detección del chip.

30 Figura 15B. Ejemplos de conjugación de materiales de nanoobjetos (con diferentes propiedades fisicoquímicas) a sustratos (con diferentes propiedades fisicoquímicas) para hacer un chip de matriz de sensor de corona de proteína después de interacciones con la fuente de proteína (por ejemplo, plasma humano de diversas enfermedades).

35 Figura 16A. Ejemplos de chip de matriz de sensor de corona de proteína con nanocurvaturas (producidas mediante una amplia variedad de enfoques disponibles, como litografía y fundición en molde), con diferentes propiedades fisicoquímicas, antes de las interacciones con la fuente de proteína (por ejemplo, plasma humano de diversas enfermedades). La corona de proteína específica se formará en la superficie de los nanoobjetos, con diferentes propiedades fisicoquímicas. Los sustratos también pueden quedar recubiertos por varios tipos de proteínas, que tienen efectos insignificantes sobre la eficacia de detección del chip.

40 Figura 16B. Ejemplos de chip de matriz de sensor de corona de proteína con nanocurvaturas (producidas mediante una amplia variedad de enfoques disponibles, como litografía y fundición en molde), con diferentes propiedades fisicoquímicas después de interacciones con la fuente de proteína (por ejemplo, plasma humano de diversas enfermedades). La corona de proteína específica se formará en la superficie de los nanoobjetos, con diferentes propiedades fisicoquímicas. Los sustratos también pueden quedar recubiertos por varios tipos de proteínas, que tienen efectos insignificantes sobre la eficacia de detección del chip.

45 Figura 17A. Ejemplos de conjugación de materiales de nanoobjetos (con diferentes propiedades fisicoquímicas) a sustratos (con diferentes propiedades fisicoquímicas) para hacer un micro/nano chip fluídico de matriz de sensor de corona de proteína antes y (b) después de las interacciones con la fuente de proteína (por ejemplo, plasma humano de diversas enfermedades). La corona de proteína específica se formará en la superficie de los nanoobjetos, con diferentes propiedades fisicoquímicas. Los sustratos también pueden quedar recubiertos por varios tipos de proteínas, que tienen efectos insignificantes sobre la eficacia de detección del chip.

50 Figura 17B. Ejemplos de conjugación de materiales de nanoobjetos a sustratos para hacer un micro/nanochip fluídico de matriz de sensor de corona de proteína después de interacciones con la fuente de proteína (por ejemplo, plasma humano de diversas enfermedades). La corona de proteína específica se formará en la superficie de los nanoobjetos, con diferentes propiedades fisicoquímicas.

55 Figura 18A. Ejemplos de chips micro/nano fluídicos de matriz de sensores de corona de proteínas con nanocurvaturas (producidas mediante una amplia variedad de enfoques disponibles, como litografía y fundición

en molde), con diferentes propiedades fisicoquímicas, antes de las interacciones con la fuente de proteínas (por ejemplo, plasma humano de diversas enfermedades).

Figura 18B. Ejemplos de chips micro/nano fluídicos de matriz de sensor de corona de proteína con nanocurvaturas (producidas mediante una amplia variedad de enfoques disponibles, como litografía y fundición en molde), con diferentes propiedades fisicoquímicas, después de interacciones con fuentes de proteínas (por ejemplo, plasma humano de diversas enfermedades).

Figura 19A. Ejemplos de conjugación de materiales de nanoobjetos ordenados aleatoriamente (con diferentes propiedades fisicoquímicas) a sustratos (con diferentes propiedades fisicoquímicas) para hacer un chip de matriz de sensor de corona de proteína (A) antes de las interacciones con la fuente de proteína (por ejemplo, plasma humano de diversas enfermedades).

Figura 19B. Ejemplos de conjugación de materiales de nanoobjetos ordenados aleatoriamente (con diferentes propiedades fisicoquímicas) a sustratos (con diferentes propiedades fisicoquímicas) para hacer un chip de matriz de sensor de corona de proteína (b) después de interacciones con la fuente de proteína (por ejemplo, plasma humano de diversas enfermedades).

Figura 20A. Ejemplos de chip de matriz de sensor de corona de proteína con nanocurvaturas ordenadas aleatoriamente (producidas mediante una amplia variedad de enfoques disponibles, como litografía y fundición en molde), con diferentes propiedades fisicoquímicas, (A) antes de las interacciones con la fuente de proteína (por ejemplo, plasma humano de diversas enfermedades).

Figura 20B. Ejemplos de chip de matriz de sensor de corona de proteína con nanocurvaturas ordenadas aleatoriamente con diferentes propiedades fisicoquímicas (b) después de interacciones con la fuente de proteína (por ejemplo, plasma humano de diversas enfermedades).

Figura 21A. Ejemplos de conjugación de materiales de nanoobjetos ordenados aleatoriamente (con diferentes propiedades fisicoquímicas) a sustratos (con diferentes propiedades fisicoquímicas) para hacer un micro/nanochip fluídico (A) de matriz de sensor de corona de proteína antes de las interacciones con la fuente de proteína (por ejemplo, plasma humano de diversas enfermedades).

Figura 21B. Ejemplos de conjugación de materiales de nanoobjetos ordenados aleatoriamente a sustratos para hacer un micro/nanochip fluídico de matriz de sensor de corona de proteína (b) después de interacciones con la fuente de proteína (por ejemplo, plasma humano de diversas enfermedades).

Figura 22A. Ejemplos de chips micro/nanofluídicos de matriz de sensores de corona de proteínas con nanocurvaturas ordenadas aleatoriamente (producidas mediante una amplia variedad de enfoques disponibles, como litografía y fundición en molde), con diferentes propiedades fisicoquímicas, (A) antes de las interacciones con la fuente de proteínas (por ejemplo plasma de diversas enfermedades).

Figura 22B. Ejemplos de chips micro/nano fluídicos de matriz de sensor de corona de proteína con nanocurvaturas ordenadas aleatoriamente con diferentes propiedades fisicoquímicas después de interacciones con la fuente de proteína (por ejemplo, plasma humano de diversas enfermedades).

Figura 23 Ejemplo de conjugación de nanopartículas a la superficie del sustrato de sílice (como sustrato representativo) a través de la reacción de amidación entre los grupos amino en la superficie del sustrato de sílice y los grupos de ácido carboxílico en la superficie de la nanopartícula.

Figura 24 Ejemplo de conjugación de nanopartículas con la superficie del sustrato de sílice (como sustrato representativo) mediante la reacción de apertura del anillo entre los grupos epoxi de la superficie del sustrato de sílice y los grupos amino de la superficie de la nanopartícula.

Figura 25 Ejemplo de conjugación de nanopartículas a la superficie del sustrato de sílice (como sustrato representativo) a través de la reacción de adición de Michael entre los grupos maleimida en la superficie del sustrato de sílice y los grupos tiol o amino en la superficie de la nanopartícula.

Figura 26 Ejemplo de conjugación de nanopartículas a la superficie del sustrato de sílice (como sustrato representativo) a través de la reacción de uretano entre los grupos isocianato en la superficie del sustrato de sílice y los grupos hidroxilo o amino en la superficie de la nanopartícula.

Figura 27 Ejemplo de conjugación de nanopartículas a la superficie del sustrato de sílice (como sustrato representativo) mediante la reacción de oxidación entre los grupos tiol de la superficie del sustrato de sílice y los de la superficie de la nanopartícula.

Figura 28 Ejemplo de conjugación de nanopartículas a la superficie del sustrato de sílice (como sustrato representativo) a través de la química "Click" entre grupos azida en la superficie del sustrato de sílice y grupos alquino en la superficie de la nanopartícula.

- Figura 29 Ejemplo de conjugación de nanopartículas a la superficie del sustrato de sílice (como sustrato representativo) a través de la reacción de intercambio de tiol entre grupos 2-piridilditio en la superficie del sustrato de sílice y grupos tiol en la superficie de la nanopartícula.
- 5 Figura 30 Ejemplo de conjugación de nanopartículas a la superficie del sustrato de sílice (como sustrato representativo) a través de la reacción de coordinación entre los grupos de ácido borónico en la superficie del sustrato de sílice y los grupos diol en la superficie de la nanopartícula.
- Figura 31 Ejemplo de conjugación de nanopartículas con la superficie del sustrato de sílice (como sustrato representativo) a través de la reacción de adición irradiada con luz ultravioleta entre los enlaces C=C en la superficie del sustrato de sílice y los enlaces C=C en la superficie de las nanopartículas.
- 10 Figura 32. Ejemplo de conjugación de nanopartículas a la superficie del sustrato de oro (como sustrato representativo) a través de enlaces Au-tiol.
- Figura 33 Ejemplo de conjugación de nanopartículas a la superficie del sustrato de oro (como sustrato representativo) a través de la reacción de amidación entre los grupos de ácido carboxílico en la superficie del sustrato de oro y los grupos amino en la superficie de la nanopartícula.
- 15 Figura 34. Ejemplo de conjugación de nanopartículas a la superficie del sustrato de oro (como sustrato representativo) a través de la química "Click" entre los grupos azida en la superficie del sustrato de oro y los grupos alquino en la superficie de la nanopartícula.
- 20 Figura 35. Ejemplo de conjugación de nanopartículas a la superficie del sustrato de oro (como sustrato representativo) a través de la reacción de uretano entre los grupos NHS en la superficie del sustrato de oro y los grupos amino en la superficie de la nanopartícula.
- Figura 36. Ejemplo de conjugación de nanopartículas a la superficie del sustrato de sílice (como sustrato representativo) a través de la reacción de apertura del anillo entre los grupos epoxi en la superficie del sustrato de oro y los grupos amino en la superficie de las nanopartículas.
- 25 Figura 37. Ejemplo de conjugación de nanopartículas a la superficie del sustrato de oro (como sustrato representativo) a través de la reacción de coordinación entre los grupos de ácido borónico en la superficie del sustrato de sílice y los grupos diol en la superficie de la nanopartícula.
- Figura 38. Ejemplo de conjugación de nanopartículas con la superficie del sustrato de oro (como sustrato representativo) a través de la reacción de adición irradiada con luz ultravioleta entre los enlaces C=C en la superficie del sustrato de oro y los enlaces C=C en la superficie de la nanopartícula.
- 30 Figura 39. Ejemplo de conjugación de nanopartículas a la superficie del sustrato de oro (como sustrato representativo) a través de la interacción "Ligando-Receptor" entre la biotina en la superficie del sustrato de oro y la avidina en la superficie de la nanopartícula.
- Figura 40 Ejemplo de conjugación de nanopartículas a la superficie del sustrato de oro (como sustrato representativo) a través de la interacción "Anfitrión-Huésped" entre a-ciclodextrina (a-CD) en la superficie del sustrato de oro y adamantina (Ad) en la superficie de la nanopartícula.
- 35 Figura 41. Disociación de proteínas de la superficie de nanopartículas y análisis de su composición corona. Análisis de los datos de la matriz de sensores de corona de proteínas con enfoques supervisados y no supervisados para identificar y discriminar enfermedades.
- Figura 42 Ejemplo de matriz de sensores con elementos sensores a nanoescala de orden aleatorio para lectura de fluorescencia o luminiscencia.
- 40 Figura 43. Ejemplo de matriz de sensores con elementos sensores de orden nanométrico para lectura de fluorescencia o luminiscencia.
- Figura 44A. Caracterización de nanopartículas desnudas de poliestireno y sílice con diferente funcionalización (ninguna, modificación de amina (NH₂) y modificación de carboxilo (COOH)), mostrando las tres diferentes nanopartículas de poliestireno (no funcionalizadas, P-NH₂ y P-COOH) utilizadas, sus tamaños, DLS y potencial zeta de las partículas desnudas.
- 45 Figura 44B. Caracterización de nanopartículas desnudas de poliestireno y sílice con diferente funcionalización (ninguna, modificación de amina (NH₂) y modificación de carboxilo (COOH)), mostrando las tres diferentes nanopartículas de sílice (no funcionalizadas, S-NH₂ y S-COOH) utilizadas, sus tamaños, DLS y potencial zeta.
- Figura 44C. Caracterización de nanopartículas de poliestireno y sílice desnudas con diferente funcionalización (ninguna, modificación de amina (NH₂) y modificación de carboxilo (COOH)), mostrando TEM de nanopartículas desnudas de poliestireno.

Figura 44D. Caracterización de nanopartículas desnudas de poliestireno y sílice con diferente funcionalización (ninguna, modificación de amina (NH₂) y modificación de carboxilo (COOH)), mostrando TEM de nanopartículas desnudas de sílice.

Figura 45A. Caracterización de nanopartículas de sílice y poliestireno recubiertas de proteína con diferente funcionalización (ninguna, modificación de amina (NH₂) y modificación de carboxilo (COOH)), mostrando los tamaños, DLS y potencial zeta de las nanopartículas de poliestireno cargadas con corona de proteínas.

Figura 45B. Caracterización de nanopartículas de sílice y poliestireno recubiertas de proteína con diferente funcionalización (ninguna, modificación de amina (NH₂) y modificación de carboxilo (COOH)), mostrando los tamaños, DLS y potencial zeta de las nanopartículas de sílice cargadas con corona de proteínas.

Figura 45C. Caracterización de nanopartículas de sílice y poliestireno recubiertas de proteína con diferente funcionalización (ninguna, modificación de amina (NH₂) y modificación de carboxilo (COOH)), mostrando TEM de nanopartículas de poliestireno cargadas con corona de proteínas.

Figura 45D. Caracterización de nanopartículas de sílice y poliestireno recubiertas de proteína con diferente funcionalización (ninguna, modificación de amina (NH₂) y modificación de carboxilo (COOH)), mostrando TEM de nanopartículas de sílice cargadas con corona de proteínas.

Figura 46. Un diagrama del tipo de muestras de plasma de cáncer analizadas con nanopartículas de poliestireno y sílice.

Figura 47. Perfiles de corona de proteína de nanopartículas de poliestireno y sílice (100 nm) con superficies simples, modificadas con amina y modificadas con carboxilo después de la incubación con plasma de pacientes con diferentes tipos de cáncer, analizados por SDS PAGE.

Figura 48. Perfiles de corona de proteína de nanopartículas de poliestireno y sílice (100 nm) con superficies simples, modificadas con amina y modificadas con carboxilo después de la incubación con plasma de individuos sanos, según lo analizado por SDS-Page

Figura 49. Gráfico que representa la separación de los pacientes con cáncer de los individuos sanos utilizando una matriz de sensores de la presente invención.

Figura 50A. Caracterización de nanopartículas de poliestireno y sílice utilizadas para el cribado CAD, mostrando el perfil de nanopartículas desnudas, tratadas con CAD, NO CAD y CONTROL.

Figura 50B. Caracterización de nanopartículas de poliestireno y sílice utilizadas para el cribado CAD, mostrando potencial zeta para los diferentes grupos de nanopartículas.

Figura 50C. Caracterización de nanopartículas de poliestireno y sílice utilizadas para el cribado CAD, mostrando TEM de las diferentes nanopartículas en el cribado CAD.

Figura 51A. Concentraciones de proteínas de diferentes coronas de proteínas a partir del análisis de nanopartículas CAD, NO CAD y sin riesgo de CAD (CONTROL), que muestra el ensayo de Bradford de concentraciones de proteínas de las diferentes coronas de proteínas.

Figura 51B. Las concentraciones de proteínas de diferentes coronas de proteínas del análisis de nanopartículas CAD, NO CAD y sin riesgo de CAD (CONTROL), que muestran perfiles de corona de proteínas personalizados, se analizaron y compararon a través de 1D-SDS-PAGE.

Figura 51C. Las concentraciones de proteína de diferentes coronas de proteínas a partir del análisis de nanopartículas CAD, NO CAD y sin riesgo de CAD (CONTROL), muestran que el análisis de gel por densitometría determinó diferencias en la cantidad de proteína en las PC CAD, NO CAD y CONTROL.

Figura 52. Barras que representan las diferencias en la contribución porcentual de las 20 proteínas más abundantes en las PC.

Figura 53. Gráfico que representa la clasificación del sujeto en CAD, NO CAD y CONTROL mediante el análisis de las huellas producidas por las nanopartículas de corona recubiertas de proteína.

Figura 54A. Identidad sintética y biológica de nanopartículas después de la incubación en plasma de la enfermedad de Alzheimer. Análisis de seguimiento de nanopartículas Nanosight (tamaño). Nanopartículas de poliestireno antes y después del recubrimiento con coronas de proteína AD. Las nanopartículas desnudas tienen entre 90 y 100 nm y un tamaño homogéneo. Las nanopartículas recubiertas de AD PC son más grandes y de tamaño menos homogéneo. Se informan los perfiles de intensidad y el diagrama de dispersión de cada medición. Los valores son promedio ± SD (n=3).

Figura 54B. Identidad sintética y biológica de nanopartículas después de la incubación en plasma de la enfermedad de Alzheimer. Análisis de seguimiento de nanopartículas Nanosight (tamaño). Nanopartículas de sílice antes y después del recubrimiento con coronas de proteína AD. Las nanopartículas desnudas tienen entre 90 y 100 nm y un tamaño homogéneo. Las nanopartículas recubiertas de AD PC son más grandes y de tamaño 5 menos homogéneo. Se informan los perfiles de intensidad y el diagrama de dispersión de cada medición. Los valores son promedio ± SD (n=3).

Figura 55. análisis TEM. Las nanopartículas antes y después del recubrimiento con coronas de proteína AD se analizaron mediante microscopía electrónica de transmisión para evaluar posibles cambios en la morfología y el tamaño. P: poliestireno; PN: poliestireno-NH₂; PC: poliestireno-COOH; S: sílice; SN: sílice-NH₂; SC: sílice-COOH. Todas las nanopartículas muestran un aumento de tamaño tras la incubación en plasma. 10

Figura 56. Geles SDS-PAGE y análisis densitométrico de las bandas. Orden de carga: P, P-NH₂, P-COOH, S, S-NH₂, S-COOH donde P: poliestireno; PN: poliestireno-NH₂; PC: poliestireno-COOH; S: sílice; SN: sílice-NH₂; 15 SC: sílice-COOH. Los perfiles corona de proteínas personalizados se han analizado y comparado a través de SDS-PAGE. Se muestran cuatro geles representativos de corona de proteína de Alzheimer y una corona de proteína saludable. ImageJ analizó la intensidad de las bandas en relación con las proteínas plasmáticas adsorbidas en nanopartículas (eje y: intensidad, eje x: peso molecular).

Figura 57. Clasificación de sanos y enfermos de AD. Los puntos blancos son AD y los puntos negros son muestras sanas. 20

Figura 58. Análisis de gel SDS-Page de nanopartículas de sílice de diferentes diámetros utilizando el mismo volumen cargado (10 ul, izquierda) o la misma cantidad (10 ug, derecha).

Figura 59. Esquema que muestra los experimentos realizados para investigar la existencia de ácidos nucleicos en la corona biomolecular. 25

Figura 60 Análisis en gel de agarosa de la unión de ácidos nucleicos a tres nanopartículas diferentes.

Figura 61. Análisis del contenido de ácidos nucleicos en plasma.

Figura 62. Análisis de las cantidades de ácido nucleico asociadas con la corona de biomoléculas de una nanopartícula cuando la urea disoció la proteína de la corona. 30

Figura 63. Análisis de las cantidades de ácido nucleico asociadas con la corona de biomoléculas de una nanopartícula cuando las coronas de proteína no estaban disociadas de la superficie de las partículas.

Figura 64. Análisis de las cantidades de ácido nucleico asociadas con la corona de biomoléculas de una nanopartícula cuando los ácidos nucleicos se purifican primero del plasma con un kit de purificación y luego se incuban con nanopartículas. 35

Figura 65. Diagrama esquemático de un método para distinguir estados de una muestra biológica compleja.

Figura 66. Diagrama esquemático de un sistema informático.

Descripción detallada de la invención

35 La presente invención se ha descrito en términos de una o más realizaciones preferidas, y debe apreciarse que muchos equivalentes, alternativas, variaciones y modificaciones, además de las indicadas expresamente, son posibles y están dentro del alcance de la invención.

La presente invención proporciona un método para detectar un estado de enfermedad en un sujeto. El método en su sentido general se define en la reivindicación independiente 1. La matriz de sensores utilizada difiere de las matrices de sensores conocidas que involucran sensores individuales que detectan biomoléculas específicas. En la presente matriz de sensores, no es necesario conocer las biomoléculas, ya que el sistema no se basa en la presencia o ausencia de biomoléculas específicas o cantidades de marcadores de enfermedades específicas. Esta nueva matriz de sensores es capaz de detectar cambios en las composiciones de la corona de biomoléculas asociadas con los diferentes elementos sensores. Esta capacidad de detectar cambios o patrones relativos (bien las biomoléculas reales asociadas con los diferentes elementos sensores o en las cantidades y/o conformaciones de cada biomolécula asociada con cada elemento sensor) permite determinar una huella de biomolécula única para cada matriz. Esta huella de biomolécula puede estratificar diferentes estados de salud y enfermedad de los sujetos. En algunas realizaciones, la huella biomolecular no solo es capaz de diferenciar entre sujetos sanos y sujetos en varias etapas diferentes de una enfermedad o trastorno, sino también para determinar un estado previo a la enfermedad en un sujeto en el que el sujeto desarrollará la enfermedad o trastorno en un momento dado tiempo más tarde. Esto es significativamente diferente y novedoso con respecto a los sistemas en la técnica que miden o detectan biomarcadores específicos asociados con una enfermedad o trastorno para proporcionar una predisposición (por ejemplo, una posibilidad o probabilidad) de desarrollar la enfermedad. El sensor y los métodos presentes pueden detectar una 40 45 50

enfermedad antes de que aparezcan signos o síntomas, en otras palabras, pueden prediagnosticar la enfermedad antes de que aparezcan signos o síntomas específicos.

La singularidad de la presente invención es la combinación de este reconocimiento de una huella biomolecular de una muestra de un sujeto y la capacidad de determinar un estado de enfermedad para ese sujeto en un continuo de salud.

La presente invención se basa en el trabajo de los inventores que han demostrado que la superficie de los elementos sensores, por ejemplo, las nanopartículas, se cubre rápidamente con una capa de diferentes biomoléculas, incluidas las proteínas, para formar una "corona de biomoléculas" cuando se pone en contacto con una muestra biológica. El tipo, la cantidad y las categorías de las biomoléculas que componen esta corona de biomoléculas están fuertemente relacionados con las propiedades fisicoquímicas de los propios elementos sensores y las complejas interacciones entre las diferentes biomoléculas y los elementos sensores. Estas interacciones conducen a la producción de una firma de corona de biomoléculas única para cada elemento sensor. En otras palabras, dependiendo de qué biomoléculas interactúen con el elemento sensor, no solo influye en la composición de la corona de biomoléculas, sino que también puede alterar qué otras biomoléculas diferentes también pueden interactuar con ese elemento sensor específico.

Se pueden poner en contacto diferentes elementos sensores, cada uno con su propia firma de corona de biomolécula, con una muestra para producir una huella de biomolécula única para esa muestra. Esta huella puede usarse luego para determinar un estado de enfermedad de un sujeto. Las realizaciones de la invención se discutirán con más detalle a continuación.

20 La presente invención involucra matrices de sensores que comprenden una pluralidad de elementos sensores en los que la pluralidad de elementos sensores difieren entre sí en al menos una propiedad fisicoquímica. Cada elemento sensor es capaz de unir una pluralidad de biomoléculas en una muestra para producir una pluralidad de firmas de corona de biomoléculas. Cada elemento sensor tiene una firma de corona de biomolécula distinta.

25 La pluralidad de elementos sensores cuando se ponen en contacto con una muestra produce una pluralidad de firmas de corona de biomoléculas que juntas forman una huella de biomolécula. La "huella de biomolécula" es la composición o patrón combinado de biomoléculas de al menos dos firmas de corona de biomoléculas para la pluralidad de elementos sensores.

Como se usa aquí, el término "elemento sensor" se refiere a elementos que pueden unirse a una pluralidad de biomoléculas cuando están en contacto con una muestra y abarca el término "elemento sensor a nanoescala".

30 En una realización, el elemento sensor es un elemento de aproximadamente 5 nanómetros a aproximadamente 50000 nanómetros en al menos una dirección. Los elementos sensores adecuados incluyen, por ejemplo, pero no se limitan a, un elemento sensor de aproximadamente 5 nm a aproximadamente 50,000 nm en al menos una dirección, incluyendo aproximadamente 5 nm a aproximadamente 40,000 nm, alternativamente aproximadamente 5 nm a aproximadamente 30,000 nm, alternativamente aproximadamente 5 nm. nm a aproximadamente 20,000 nm, alternativamente de aproximadamente 5 nm a aproximadamente 10,000 nm, alternativamente de aproximadamente 5 nm a aproximadamente 5000 nm, alternativamente de aproximadamente 5 nm a aproximadamente 1000 nm, alternativamente de aproximadamente 5 nm a aproximadamente 500 nm, alternativamente de aproximadamente 5 nm a 50 nm, alternativamente

40 aproximadamente 300 nm, alternativamente de aproximadamente 3 nm a 30 nm, alternativamente aproximadamente 10 nm a 100 nm, alternativamente aproximadamente 20 nm a 200 nm, alternativamente aproximadamente 30 nm a 300 nm, alternativamente aproximadamente 40 nm a 400 nm, alternativamente aproximadamente 50 nm a 500 nm, alternativamente aproximadamente 60 nm a 600 nm, alternativamente aproximadamente 70 nm a 700 nm, alternativamente aproximadamente 80 nm a 800 nm, alternativamente aproximadamente 90 nm a 900 nm, alternativamente aproximadamente 100 nm a 1000 nm, alternativamente aproximadamente 1000 nm a 10000 nm, alternativamente aproximadamente 10000 nm a 50000 nm y cualquier

45 combinación o cantidad intermedia (por ejemplo, 5nm, 10nm, 15nm, 20 nm, 25 nm, 30 nm, 35 nm, 40 nm, 45 nm, 50 nm, 55 nm, 60 nm, 65 nm, 70 nm, 80 nm, 90 nm, 100 nm, 125 nm, 150 nm, 175 nm, 200 nm, 225 nm, 250 nm, 275 nm, 300 nm, 350 nm, 400 nm, 450 nm, 500 nm, 550 nm, 600 nm, 650 nm, 700 nm, 750 nm, 800 nm, 850 nm, 900 nm, 1000 nm, 1200 nm, 1300 nm, 1400 nm, 1500 nm, 1600 nm, 1700 nm, 1800 nm, 1900 nm, 2000 nm, 2500 nm, 3000 nm, 3500 nm, 4000 nm, 4500 nm, 5000 nm, 5500 nm, 6000 nm, 6500 nm, 7 nm, 7500 nm, 8000 nm, 8500 nm, 9000 nm, 10000 nm, 11000nm, 12000 nm, 13000 nm, 14000 nm, 15000 nm, 16000 nm, 17000 nm, 18000 nm, 19000 nm, 20000nm, 25000 nm, 30000 nm, 35000 nm, 40000 nm, 45000 nm, 50000 nm y cualquier número intermedio). Un elemento sensor a nanoscala se refiere a un elemento sensor que tiene menos de 1 micra en al menos una dirección. Ejemplos adecuados de rangos de elementos sensores a nanoscala incluyen, pero no se limitan a, por ejemplo, elementos de aproximadamente 5 nm a

55 nanoescala incluyen, pero no se limitan a, por ejemplo, elementos de aproximadamente 5 nm a
aproximadamente 1000 nm en una dirección, incluyendo, por ejemplo, aproximadamente 5 nm a
aproximadamente 500 nm, alternativamente aproximadamente 5 nm a aproximadamente 400 nm,
alternativamente de aproximadamente 5 nm a aproximadamente 300 nm, alternativamente de
aproximadamente 5 nm a aproximadamente 200 nm, alternativamente de aproximadamente 5 nm a
aproximadamente 100 nm, alternativamente de aproximadamente 5 nm a aproximadamente 50 nm,
alternativamente de aproximadamente 10 nm a aproximadamente 1000 nm, alternativamente
aproximadamente 10 nm a aproximadamente 750 nm, alternativamente aproximadamente 10 nm a

aproximadamente 500 nm, alternativamente aproximadamente 10 nm a aproximadamente 250 nm, alternativamente aproximadamente 10 nm a aproximadamente 200 nm, alternativamente aproximadamente 10 nm a aproximadamente 100 nm, alternativamente aproximadamente 50 nm a aproximadamente 1000 nm, alternativamente de aproximadamente 50 nm a aproximadamente 500 nm, alternativamente de 5
aproximadamente 50 nm a aproximadamente 250 nm, alternativamente de aproximadamente 50 nm a aproximadamente 200 nm, alternativamente de aproximadamente 50 nm a aproximadamente 100 nm, y cualquier combinación, rango o cantidad intermedia (por ejemplo, 5 nm, 10 nm, 15nm, 20nm, 25nm, 30nm, 35nm, 40n m, 45 nm, 50 nm, 55 nm, 60 nm, 65 nm, 70 nm, 80 nm, 90 nm, 100 nm, 125 nm, 150 nm, 175 nm, 200 nm, 225 nm, 250 nm, 275 nm, 300 nm, 350 nm, 400 nm, 450 nm, 500 nm, 550 nm, 600 nm, 650 nm, 700 nm, 750 nm, 800 nm, 850 nm, 900 nm, 1000 nm, etc.). En referencia a las matrices de sensores descritas en 10 el presente documento, el uso del término elemento sensor incluye el uso de un elemento sensor a nanoscala para el sensor y los métodos asociados.

El término "pluralidad de elementos sensores" se refiere a más de uno, por ejemplo, al menos dos elementos sensores. En algunas realizaciones, la pluralidad de elementos sensores incluye al menos dos elementos sensores a al menos 1000 elementos sensores, preferiblemente aproximadamente dos elementos sensores a 15 aproximadamente 100 elementos sensores. En realizaciones adecuadas, la matriz comprende al menos dos a al menos 100 elementos sensores, alternativamente al menos dos a al menos 50 elementos sensores, alternativamente al menos 2 a 30 elementos sensores, alternativamente al menos 2 a 20 elementos sensores, alternativamente al menos 2 a 10 elementos sensores, alternativamente al menos 3 a al menos 50 elementos 20 sensores, alternativamente al menos 3 a al menos 30 elementos sensores, alternativamente al menos 3 a al menos 20 elementos sensores, alternativamente al menos 3 a al menos 10 elementos sensores, alternativamente al menos 4 a al menos 50 elementos sensores, alternativamente al menos 4 a al menos 30 elementos sensores, alternativamente al menos 4 a al menos 10 elementos sensores, e incluyendo cualquier número de elementos 25 sensores contemplado en el medio (por ejemplo, al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 225, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, etc.). En algunas realizaciones, la matriz de sensores comprende al menos 6 30 elementos sensores a al menos 20 elementos sensores, alternativamente al menos 6 elementos sensores a al menos 10 elementos sensores.

El término "pluralidad de elementos sensores a nanoscala" se refiere a más de uno, por ejemplo, al menos dos elementos sensores a nanoscala. En algunas realizaciones, la pluralidad de elementos sensores a nanoscala incluye al menos dos elementos sensores a nanoscala hasta al menos 1000 elementos sensores a nanoscala, preferiblemente entre dos elementos sensores a nanoscala y aproximadamente 100 elementos 35 sensores a nanoscala. En realizaciones adecuadas, la matriz comprende al menos dos a al menos 100 elementos sensores a nanoscala, alternativamente al menos dos a al menos 50 elementos sensores a nanoscala, alternativamente al menos 2 a 30 elementos sensores a nanoscala, alternativamente al menos 2 a 20 elementos sensores a nanoscala, alternativamente al menos 2 a 10 elementos sensores a nanoscala, alternativamente al menos 3 a al menos 50 elementos sensores a nanoscala, alternativamente al menos 3 a al menos 30 elementos sensores a nanoscala, alternativamente al menos 3 a al menos 20 elementos sensores a 40 nanoscala, alternativamente al menos 3 al menos 10 elementos sensores a nanoscala, alternativamente al menos 4 a al menos 50 elementos sensores a nanoscala, alternativamente al menos 4 a al menos 30 elementos sensores a nanoscala, alternativamente al menos 4 a al menos 10 elementos sensores a nanoscala, e incluyendo cualquier número de elementos sensores a nanoscala contemplados en el medio (por ejemplo, al 45 menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 225, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, etc.).

Como se usa en el presente documento, el término "corona de biomoléculas" se refiere a la pluralidad de 50 diferentes biomoléculas que pueden unirse a un elemento sensor. El término "corona de biomoléculas" abarca "corona de proteínas", que es un término utilizado en la técnica para referirse a las proteínas, lípidos y otros componentes del plasma que se unen a nanopartículas cuando entran en contacto con muestras biológicas o sistemas biológicos. Para su uso en el presente documento, el término "corona de biomoléculas" también 55 abarca tanto la corona de proteína blanda como la dura como se hace referencia en la técnica, véase, por ejemplo, Milani et al. "Reversible versus Irreversible Binding of Transferring to Polystyrene Nanoparticles: Soft and Hard Corona" ACS NANO, 2012, 6(3), pp. 2532-2541; Mirshafiee et al. "Impact of protein pre-coating on the protein corona composition and nanoparticle cellular uptake" Biomaterials vol. 75, Jan 2016 pp. 295-304, Mahmoudi et al. "Emerging understanding of the protein corona at the nanobio interfaces" Nanotoday 11(6) Dec. 2016, pp. 817-832, and Mahmoudi et al. "Protein-Nanoparticle Interactions: Opportunities and Challenges" Chem. Rev., 2011, 111(9), pp. 5610-5637. Como se describe en la técnica, la curva de adsorción muestra la 60 acumulación de una monocapa fuertemente unida hasta el punto de saturación de la monocapa (en una relación de proteína a NP definida geométricamente), más allá del cual se forma una capa secundaria débilmente unida. Mientras que la primera capa está unida de forma irreversible (corona dura), la capa

secundaria (corona blanda) muestra un intercambio dinámico. Las proteínas que se adsorben con alta afinidad forman lo que se conoce como corona "dura", que consiste en proteínas fuertemente unidas que no se desorben fácilmente, y las proteínas que se adsorben con baja afinidad forman la corona "blanda", que consiste en proteínas poco unidas. La corona blanda y dura también se puede definir en función de sus tiempos de intercambio. La corona dura generalmente muestra tiempos de intercambio mucho más grandes, del orden de varias horas. Véase, por ejemplo, M. Rahman et al. Protein-Nanoparticle Interactions, Spring Series in Biophysics 15, 2013.

El término "firma de corona de biomolécula" se refiere a la composición, firma o patrón de diferentes biomoléculas que están unidas a cada elemento sensor por separado. La firma no solo se refiere a las diferentes biomoléculas sino también a las diferencias en la cantidad, nivel o cantidad de la biomolécula unida al elemento sensor, o diferencias en el estado conformacional de la biomolécula que está unida al elemento sensor. Se contempla que las firmas de corona de biomoléculas de cada elemento sensor pueden contener algunas de las mismas biomoléculas, pueden contener biomoléculas distintas con respecto a los otros elementos sensores y/o pueden diferir en nivel o cantidad, tipo o confirmación de la biomolécula. La firma de corona de la biomolécula puede depender no solo de las propiedades fisicoquímicas del elemento sensor, sino también de la naturaleza de la muestra y la duración de la exposición. En algunos casos, la firma de corona biomolécula es una firma de corona proteína. En otro caso, la firma de corona de biomolécula es una firma de corona de polisacárido. En otro caso más, la firma de corona de biomolécula es una firma de corona de metabolito. En algunos casos, la firma de corona biomolécula es una firma de corona lipídica.

En algunas realizaciones, la firma de corona de biomoléculas comprende las biomoléculas que se encuentran en una corona blanda y una corona dura. En algunas realizaciones, la corona blanda es una corona de proteína blanda. En algunas realizaciones, la corona dura es una corona de proteína dura.

El término "biomolécula" se refiere a componentes biológicos que pueden estar involucrados en la formación de corona, incluidos, pero no limitados a, proteínas, polipéptidos, polisacáridos, un azúcar, un lípido, una lipoproteína, un metabolito, un oligonucleótido, un metaboloma o combinación de los mismos. Se contempla que las firmas de corona de biomoléculas de cada elemento sensor pueden contener algunas de las mismas biomoléculas, pueden contener biomoléculas distintas con respecto a los otros elementos sensores y/o pueden diferir en nivel o cantidad, tipo o confirmación de la biomolécula que se une a cada elemento sensor. En una realización, la biomolécula se selecciona del grupo de proteínas, ácidos nucleicos, lípidos y metabolomas.

En algunas realizaciones, la matriz de sensores comprende, consiste esencialmente en un primer elemento sensor que produce una primera firma de corona de biomolécula y al menos un segundo elemento sensor que produce al menos una segunda firma de corona de biomolécula cuando la matriz de sensores se pone en contacto con un muestra biológica. Una huella de biomoléculas es la combinación de la primera firma de biomoléculas y la al menos una segunda firma de biomoléculas. Se contempla que la firma de biomoléculas se pueda hacer a partir de al menos dos firmas de corona de biomoléculas hasta tantas firmas de biomoléculas diferentes que se ensayan, por ejemplo al menos 1000 firmas de corona de biomoléculas diferentes. La corona de biomoléculas puede analizarse por separado para cada elemento sensor para determinar la firma de corona de biomoléculas para cada elemento y combinarse para determinar la huella de biomoléculas o las dos o más coronas de biomoléculas pueden analizarse al mismo tiempo para desarrollar la huella de biomoléculas a la vez.

En algunas realizaciones, la huella de biomoléculas incluye al menos dos firmas de corona de biomoléculas. En algunas realizaciones, la huella de la biomoléculas incluye desde dos firmas de corona de biomoléculas hasta al menos 1000 firmas de corona de biomoléculas, preferiblemente entre dos firmas de corona de biomoléculas y aproximadamente 100 firmas de corona de biomoléculas. En realizaciones adecuadas, la huella de biomoléculas comprende al menos dos a al menos 100 firmas de corona de biomoléculas, alternativamente al menos dos a al menos 50 firmas de corona de biomoléculas, alternativamente al menos 2 a 30 firmas de corona de biomoléculas, alternativamente al menos 2 a 20 firmas de corona de biomoléculas, alternativamente al menos 2 a 10 firmas de corona de biomoléculas, alternativamente al menos 3 a al menos 50 firmas de corona de biomoléculas, alternativamente al menos 3 a al menos 30 firmas de corona de biomoléculas, alternativamente al menos 3 a al menos 20 firmas de corona de biomoléculas, alternativamente al menos 3 a al menos 10 firmas de corona de biomoléculas, alternativamente al menos 4 a al menos 50 firmas de corona de biomoléculas, alternativamente al menos 4 a al menos 30 firmas de corona de biomoléculas, alternativamente al menos 4 a al menos 20 firmas de corona de biomoléculas, alternativamente al menos 4 a al menos 10 firmas de corona de biomoléculas, e incluyendo cualquier número de firmas de corona de biomoléculas naturalezas contempladas en el medio (por ejemplo, al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 225, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, etc.).

Los avances en los análisis proteómicos que utilizan espectrometría de masas han ofrecido nuevos conocimientos sobre los cambios que tienen lugar en todo el espectro de la salud y la enfermedad, incluido el cáncer en etapa temprana. Sin embargo, la sensibilidad y la especificidad de los enfoques de espectrometría

de masas no han sido adecuadas para la detección temprana sólida de cánceres, en parte debido al alto ruido 5 creado por las 10000 proteínas que componen el proteoma humano, con concentraciones estimadas que oscilan entre 35 y 50 mg/ml de albúmina para 1-10 pg/ml para algunas citoquinas. Las tecnologías existentes han requerido un compromiso entre la profundidad de cobertura y el rendimiento del procesamiento de 10 proteínas plasmáticas. Se han realizado varios intentos para aumentar sustancialmente los bajos niveles actuales de detección de proteínas, incluidas las deplecciones de proteínas muy abundantes, el marcaje isobárico a nivel de péptidos para la cuantificación relativa multiplexada, las estrategias de fraccionamiento de plasma posteriores a la deplección, las técnicas de recolección de biomarcadores, los enfoques matemáticos 15 para analizar los niveles altos de -conjuntos de datos de calidad y flujo de trabajo multiplexado (es decir, una combinación de enfoques). A pesar de tales esfuerzos, los enfoques de espectrometría de masas para la proteómica del plasma no han tenido un gran éxito en la detección temprana de cánceres. De hecho, ningún estudio previo, proteómico o de otro tipo, ha informado una predicción y clasificación precisas de una variedad de cánceres, incluidas las etapas presintomáticas más tempranas. La presente matriz de sensores ha 20 proporcionado el primer sistema de detección que predice y clasifica con precisión un estado de enfermedad, incluido el estado de enfermedad presintomático para un número de enfermedades diferentes.

Los intentos anteriores han tratado de usar la "corona de proteína específica de la enfermedad" para identificar 25 un tipo de cáncer usando electroforesis en gel y cambios en el tamaño de agregación de las nanopartículas. Sin embargo, como mostramos en los Ejemplos a continuación, las diferencias sutiles en la corona de proteína en la superficie de un tipo de nanopartícula no fueron suficientes para una identificación y discriminación sólidas y precisas de los cánceres con una precisión de predicción aceptable, principalmente debido al problema 30 persistente de la inadecuada cobertura proteómica. La matriz sensorial descrita en este documento es capaz de clasificar con precisión los estados de enfermedad. No solo es capaz de predecir el estado de la enfermedad, sino que también puede clasificar a los pacientes que son presintomáticos de la enfermedad (por ejemplo, Alzheimer) o clasificar a los pacientes según el tipo de enfermedad (por ejemplo, tipo de cáncer).

25 Para potenciar materialmente la capacidad de la corona de proteína para la detección robusta y precisa del cáncer con una excelente capacidad de predicción, los inventores desarrollaron una matriz de sensores (a veces denominada en el presente documento nanosistema de corona de proteína o nanosistema de matriz de sensores). En comparación con los enfoques anteriores limitados a la superficie de una sola nanopartícula, proporciona datos proteómicos significativamente más completos en un rango dinámico más amplio de 30 concentraciones de proteínas plasmáticas. La matriz de sensores permite el muestreo de una muestra biológica compleja (por ejemplo, muestra de plasma humano) utilizando multinanopartículas con diferentes propiedades fisicoquímicas para aumentar significativamente el número y el rango de proteínas de baja y alta abundancia identificadas sin agotamiento de proteínas. Esto reduce efectivamente el ruido en la vasta información 35 proteómica disponible, lo que produce una diferenciación temprana más precisa de la firma proteómica que es característica de una enfermedad. Además, debido a la combinación de la interacción proteína-nanopartícula y proteína-proteína que se deriva de manera única usando la matriz de sensores, cada tipo de proteína puede estar presente en diferentes concentraciones en la superficie de diferentes nanopartículas, proporcionando información proteómica adicional. El uso de multinanopartículas con diferentes propiedades fisicoquímicas se 40 debe principalmente a nuestros hallazgos recientes de que incluso pequeñas alteraciones en las propiedades fisicoquímicas de las nanopartículas pueden provocar cambios dramáticos pero reproducibles en la composición de la corona de proteínas.

45 La presente matriz de sensores tiene una sensibilidad y un rango dinámico de diez (10) órdenes de magnitud en términos de detección de proteínas utilizando enfoques de espectroscopia de masas. El presente ensayo es capaz de detectar proteínas que se encuentran en el rango sub-*ng* dentro de una muestra. Este ensayo o enfoque tiene un rango dinámico mucho mayor que los ensayos actuales para medir proteínas dentro de una muestra. Por ejemplo, la espectrometría de masas solo tiene un rango dinámico de 4-6 órdenes de magnitud. Esta nueva matriz de sensores tiene la capacidad de muestrear un rango dinámico mayor que el que se había logrado anteriormente. La presente matriz de sensores permite la detección y determinación de proteínas poco 50 abundantes y raras que antes no podíamos detectar.

50 El término "muestra" se refiere a una muestra biológica o una muestra biológica compleja obtenida de un sujeto. Las muestras biológicas adecuadas incluyen, pero no se limitan a, fluidos biológicos, incluidos, pero no limitados a, sangre sistémica, plasma, suero, lavado pulmonar, lisados celulares, sangre menstrual, orina, muestras de tejido procesadas, líquido amniótico, líquido cefalorraquídeo, lágrimas, saliva, semen y similares. En una realización preferida, la muestra es una muestra de sangre o suero. El plasma sanguíneo contiene 55 varios miles de proteínas diferentes con diferencias de doce órdenes de magnitud en las concentraciones de estas proteínas. La presente matriz de sensores es capaz de detectar cambios dentro de estas muestras de sangre a lo largo del tiempo o sobre los estados de enfermedad del sujeto.

60 En algunas realizaciones, los fluidos biológicos o las muestras biológicas complejas se preparan mediante métodos y kits conocidos en la técnica. Por ejemplo, algunas muestras biológicas (por ejemplo, sangre menstrual, muestras de sangre, semen, etc.) pueden centrifugarse primero a baja velocidad para eliminar restos de células, coágulos de sangre y otros componentes celulares que puedan interferir con la matriz. En otras realizaciones, por ejemplo, las muestras de tejido pueden procesarse, por ejemplo las muestras de tejido

pueden triturarse u homogeneizarse, tratarse con enzimas para desmenuzar el tejido y/o centrifugarse para eliminar los restos celulares que permitan el ensayo y la extracción de biomoléculas dentro de las muestras de tejido. Los métodos adecuados para aislar y/o preparar y almacenar adecuadamente muestras de sangre son conocidos en la técnica y pueden incluir, pero no limitados a, la adición de un agente anticoagulante.

- 5 Los elementos sensores adecuados incluyen, pero no se limitan a, por ejemplo, partículas, tales como partículas orgánicas, partículas no orgánicas o combinaciones de las mismas. En algunas realizaciones, las partículas son, por ejemplo, nanopartículas, micropartículas, micelas, liposomas, óxido de hierro, grafeno, sílice, partículas basadas en proteínas, poliestireno, partículas de plata y oro, puntos cuánticos, paladio, platino, titanio y combinaciones de los mismos. En algunas realizaciones, las nanopartículas son liposomas. Un experto 10 en la técnica podría seleccionar y preparar partículas adecuadas. En algunas realizaciones preferidas, los elementos sensores son elementos sensores a nanoescala. Los elementos sensores a nanoescala adecuados tienen menos de 1 micrómetro en al menos una dirección. En algunos aspectos, los elementos sensores a nanoescala tienen menos de aproximadamente 100 nm en al menos una dirección.

Visión general

- 15 La presente divulgación proporciona un método para detectar un estado de enfermedad usando un nanosistema de corona de biomoléculas. En una realización, el método comprende detectar una corona de proteína específica de la enfermedad.

Matrices de sensores

- 20 La Figura 65 muestra un esquema ejemplar del sistema de matriz actualmente divulgado. Como se muestra en el paso 1 de la Figura 65, se puede recolectar una muestra biológica compleja (por ejemplo, sangre 704) de un sujeto 702 que exprese un estado biológico 703 (por ejemplo, un estado de enfermedad, por ejemplo antes de cualquier síntoma físico de la enfermedad y/o durante las etapas temprana e intermedia de una enfermedad). Las muestras biológicas 704 adecuadas incluyen, pero no se limitan a, fluidos biológicos, incluidos, pero no limitados a, sangre sistémica, plasma, suero, lavado pulmonar, lisados celulares, sangre menstrual, orina, 25 muestras de tejido procesadas, líquido amniótico, líquido cefalorraquídeo, lágrimas, saliva, semen y similares. En una realización preferida, la muestra es una muestra de sangre o suero. En algunas realizaciones, el plasma 706 puede separarse de las células sanguíneas 708 de sujetos que expresan un estado biológico 703 (por ejemplo, personas sanas (estado sin enfermedad) y pacientes con cáncer (estado de enfermedad), como se muestra en el paso 2 de la Figura 65).

- 30 A continuación, como se muestra en el paso 3 de la Figura 65, una muestra biológica compleja (por ejemplo, plasma 706) se puede incubar con una matriz de sensores 710 que comprende una pluralidad de partículas 712 con diferentes propiedades fisicoquímicas. La pluralidad de partículas 712 se puede incubar con el plasma 706 para permitir que las biomoléculas en el plasma 706 (por ejemplo, proteínas en el plasma 706) se unan a una o más de las partículas 712. Posteriormente, como se muestra en el paso 4 de la Figura 65, las 35 biomoléculas (por ejemplo, proteínas) unidas a las partículas 712 se pueden aislar en una solución de proteína 714 para su posterior análisis, por ejemplo, para determinar las composiciones de las proteínas unidas a cada tipo de partícula 712 (por ejemplo, partículas aniónicas, neutras y catiónicas 712).

- 40 La solución de proteína 714 se puede caracterizar, por ejemplo, como se muestra en el paso 5 de la Figura 65, cromatografía líquida-espectrometría de masas en tandem (LC-MS/MS) 716. Las proteínas identificadas mediante LC-MS/MS 716 se pueden analizar en el paso 6 de la Figura 65 para determinar una huella de biomolécula 718 (por ejemplo, representativa de proteínas, ácidos nucleicos, lípidos y polisacáridos que se unen a una o más de las partículas 712) asociada con el estado biológico 703.

- 45 En el paso 7 de la Figura 65, un ordenador 720 (por ejemplo, un sistema informático 101 en la Figura 66) puede usarse para asociar una huella de biomolécula 718 (por ejemplo, huella de proteína) con un estado biológico 703 (por ejemplo, estado de salud, estado de enfermedad). Por ejemplo, el análisis de una huella de biomolécula 718 (por ejemplo, huella de proteína) de al menos dos muestras (por ejemplo, muestras biológicas complejas como el plasma 706) se puede realizar con un sistema informático 720 para generar una asociación 722 entre el estado biológico 703 y la huella de biomolécula 718 (en el paso 8 de la Figura 65). La generación 50 de asociación 722 puede ser mediante un análisis de asociación o clasificación estadística utilizando métodos conocidos en la técnica, que incluyen, pero no se limitan a, una amplia variedad de análisis de datos supervisados y no supervisados y enfoques de agrupamiento tales como análisis de conglomerados jerárquicos (HCA), principal análisis de componentes (PCA), análisis discriminante de mínimos cuadrados parciales (PLS-DA), aprendizaje automático (también conocido como bosque aleatorio), regresión logística, árboles de decisión, máquina de vectores de soporte (SVM), k-vecinos más cercanos, bayes ingenuos, regresión lineal, regresión polinomial, SVM para regresión, agrupamiento de K-medias y modelos ocultos de Markov, entre otros. En otras palabras, las huellas de biomoléculas 718 de cada muestra (por ejemplo, plasma 706) se comparan/analizan (por ejemplo, utilizando un ordenador 720) entre sí para determinar con significancia estadística qué patrones son comunes entre las huellas individuales para determinar un estado biológico que es asociado con la biomolécula (por ejemplo, proteína) huella 718.

La asociación 722 puede vincular la huella de la biomolécula 718 (por ejemplo, la huella de la proteína) con una amplia variedad de estados biológicos 703. Por ejemplo, la comparación de huellas de biomoléculas 718 entre un sujeto 702 diagnosticado con una enfermedad (es decir, el estado biológico 703 es un estado de enfermedad) y un sujeto 702 no diagnosticado con la enfermedad (es decir, el estado biológico 703 es un estado sin enfermedad) puede dar lugar a una asociación 722 entre la huella de biomolécula 718 del sujeto 702 con la enfermedad y el estado de enfermedad. Tal asociación 722 entre la huella del biomarcador 718 y el estado de la enfermedad (es decir, el estado biológico 703) puede determinarse en algunas realizaciones muy temprano durante la progresión de una enfermedad (es decir, antes de que se manifieste cualquier síntoma físico de la enfermedad y/o antes del diagnóstico de la enfermedad) o en momentos posteriores durante la progresión de la enfermedad.

Otros ejemplos de estados biológicos 703 que se pueden asociar con una huella de biomolécula 718 incluyen la respuesta o falta de respuesta a un fármaco o producto farmacéutico, nivel de activación del sistema inmunitario (por ejemplo, debido a la exposición de un sujeto a un antígeno exógeno), susceptibilidad de un sujeto a los efectos adversos asociados con la administración de un fármaco, y la identificación del potencial de un sujeto para exhibir una reacción alérgica a la administración de una composición o sustancia particular.

Sistemas de control por ordenador (no reivindicados *per se* pero que pueden ser útiles para la práctica de la invención)

La presente divulgación proporciona sistemas de control informático que están programados para implementar métodos de la divulgación. La Figura 66 muestra un sistema informático 100 que está programado o configurado de otro modo para asociar una huella de biomolécula 718 con un estado biológico 703. Esta determinación, análisis o clasificación estadística se realiza mediante métodos conocidos en la técnica, que incluyen, pero no se limitan a, una amplia variedad de análisis de datos supervisados y no supervisados y enfoques de agrupamiento tales como análisis de conglomerados jerárquicos (HCA), análisis de componentes principales (PCA), análisis discriminante de mínimos cuadrados parciales (PLS-DA), aprendizaje automático (también conocido como bosque aleatorio), regresión logística, árboles de decisión, máquina de vectores de soporte (SVM), k-vecinos más cercanos, bayes ingenuos, regresión lineal, regresión polinomial, SVM para regresión, agrupamiento de K-medias y modelos ocultos de Markov, entre otros. El sistema informático 100 puede realizar varios aspectos del análisis de las huellas de biomoléculas 718 de la presente descripción, como, por ejemplo, comparar/analizar la corona de biomoléculas de varias muestras para determinar con significación estadística qué patrones son comunes entre las coronas de biomoléculas individuales para determinar una huella de biomolécula 718 que está asociada con el estado biológico 703. El sistema informático puede usarse para desarrollar clasificadores para detectar y discriminar diferentes huellas de biomoléculas 718 (por ejemplo, características de la composición de una corona de proteína). Los datos recopilados de la matriz de sensores divulgada actualmente se pueden usar para entrenar un algoritmo de aprendizaje automático, específicamente un algoritmo que recibe mediciones de matriz de un paciente y genera composiciones de corona de biomoléculas específicas de cada paciente. Antes de entrenar el algoritmo, los datos sin procesar de la matriz se pueden eliminar primero para reducir la variabilidad en las variables individuales.

El aprendizaje automático se puede generalizar como la capacidad de una máquina de aprendizaje para realizar con precisión ejemplos/tareas nuevas e invisibles después de haber experimentado un conjunto de datos de aprendizaje. El aprendizaje automático puede incluir los siguientes conceptos y métodos. Los conceptos de aprendizaje supervisado pueden incluir AODE; Red neuronal artificial, tal como Backpropagation, Autoencoders, redes Hopfield, máquinas Boltzmann, máquinas Boltzmann restringidas y redes neuronales Spiking; Estadísticas bayesianas, como la red bayesiana y la base de conocimiento bayesiana; Razonamiento basado en casos; regresión del proceso gaussiano; Programación de expresión génica; método de grupo de manejo de datos (GMDH); programación lógica inductiva; aprendizaje basado en instancias; Aprendizaje perezoso; aprendizaje de autómatas; Cuantificación de vectores de aprendizaje; árbol modelo lógístico; Longitud mínima del mensaje (árboles de decisión, gráficos de decisión, etc.), tal como el algoritmo del vecino más cercano y el modelado analógico; Probablemente aprendizaje aproximadamente correcto (PAC); reglas de onda descendente rules, una metodología de adquisición de conocimiento; Algoritmos de aprendizaje automático simbólico; Máquinas de vectores de soporte; bosques aleatorios; Conjuntos de clasificadores, tales como la agregación Bootstrap (empaquetado) y Boosting (metaalgoritmo); clasificación ordinal; Redes difusas de información (IFN); campo aleatorio condicional; ANOVA; Clasificadores lineales, como el discriminante lineal de Fisher, la regresión lineal, la regresión logística, la regresión logística multinomial, el clasificador Naïve Bayes, el perceptrón, las máquinas de vectores de soporte; clasificadores cuadráticos; k-vecino más cercano; Boosting; Árboles de decisión, tales como C4.5, bosques aleatorios, ID3, CART, SLIQ, SPRINT; redes bayesianas, tales como Naïve Bayes; y modelos ocultos de Markov. Los conceptos de aprendizaje no supervisado pueden incluir; Algoritmo de maximización de expectativas; Cuantificación de vectores; Mapa topográfico generativo; método de cuello de botella de información; Red neuronal artificial, tal como el mapa de autoorganización; Aprendizaje de reglas de asociación, como el algoritmo Apriori, el algoritmo Eclat y el algoritmo de crecimiento de FP; Agrupamiento jerárquico, como el agrupamiento de enlace único y el agrupamiento conceptual; Análisis de conglomerados, como algoritmo K-medias, agrupamiento difuso, DBSCAN y algoritmo OPTICS; y detección de valores atípicos, como el factor local de valores atípicos. Los conceptos de aprendizaje semisupervisado pueden incluir; modelos generativos; separación de baja densidad;

métodos basados en gráficos; y Co-formación. Los conceptos de aprendizaje por refuerzo pueden incluir; Aprendizaje de diferencias temporales; Q-aprendizaje; aprendizaje de autómatas; y SARASA. Los conceptos de aprendizaje profundo pueden incluir; Redes de creencias profundas; máquinas profundas de Boltzmann; redes neuronales convolucionales profundas; redes neuronales recurrentes profundas; y memoria temporal jerárquica.

5 El sistema informático 100 representado en la Figura 66 está adaptado para implementar un método descrito en este documento. El sistema 100 incluye un servidor informático central 101 que está programado para implementar métodos ejemplares descritos en este documento. El servidor 101 incluye una unidad de procesamiento central (CPU, también "procesador") 105 que puede ser un procesador de un solo núcleo, un procesador de múltiples núcleos o una pluralidad de procesadores para procesamiento en paralelo. El servidor 101 también incluye memoria 110 (por ejemplo, memoria de acceso aleatorio, memoria de solo lectura, memoria flash); unidad de almacenamiento electrónico 115 (por ejemplo, disco duro); interfaz de comunicaciones 120 (por ejemplo, adaptador de red) para comunicarse con uno o más sistemas; y dispositivos periféricos 125 que pueden incluir caché, otra memoria, almacenamiento de datos y/o adaptadores de pantalla electrónica. La memoria 110, la unidad de almacenamiento 115, la interfaz 120 y los dispositivos periféricos 125 están en comunicación con el procesador 105 a través de un bus de comunicaciones (líneas sólidas), como una placa base. La unidad de almacenamiento 115 puede ser una unidad de almacenamiento de datos para almacenar datos. El servidor 101 está acoplado operativamente a una red informática ("red") 130 con la ayuda de la interfaz de comunicaciones 120. La red 130 puede ser Internet, una intranet y/o una extranet, una intranet y/o una extranet que está en comunicación con Internet, una red de telecomunicaciones o de datos. La red 130 en algunos casos, con la ayuda del servidor 101, puede implementar una red de igual a igual, que puede permitir que los dispositivos acoplados al servidor 101 se comporten como un cliente o un servidor.

10 La unidad de almacenamiento 115 puede almacenar archivos, como informes de sujetos y/o comunicaciones con los datos sobre individuos, o cualquier aspecto de los datos asociados con la presente divulgación.

15 20 25 El servidor informático 101 puede comunicarse con uno o más sistemas informáticos remotos a través de la red 130. El uno o más sistemas informáticos remotos pueden ser, por ejemplo, ordenadores personales, portátiles, tabletas, teléfonos, teléfonos inteligentes o asistentes digitales personales.

La unidad de almacenamiento 115 puede almacenar archivos, como informes de sujetos y/o comunicaciones con los datos sobre individuos, o cualquier aspecto de los datos asociados con la presente divulgación.

20 25 30 En algunas aplicaciones, el sistema informático 100 incluye un solo servidor 101. En otras situaciones, el sistema incluye múltiples servidores en comunicación entre sí a través de una intranet, extranet y/o internet.

30 35 40 45 50 55 60 El servidor 101 se puede adaptar para almacenar datos de medición o una base de datos como se proporciona aquí, información del paciente del sujeto, tal como, por ejemplo, historial médico, historial familiar, datos demográficos y/u otra información clínica o personal de relevancia potencial para un aplicación particular. Dicha información puede almacenarse en la unidad de almacenamiento 115 o el servidor 101 y dichos datos pueden transmitirse a través de una red.

Los métodos que se describen en este documento pueden implementarse por medio de un código ejecutable (o software) de máquina (o procesador de computadora) almacenado en una ubicación de almacenamiento electrónico del servidor 101, tal como, por ejemplo, en la memoria 110 o la unidad de almacenamiento electrónico 115. Durante el uso, el código puede ser ejecutado por el procesador 105. En algunos casos, el código puede recuperarse de la unidad de almacenamiento 115 y almacenarse en la memoria 110 para que el procesador 105 pueda acceder fácilmente a él. En algunas situaciones, la unidad de almacenamiento electrónico 115 se puede excluir y las instrucciones ejecutables por máquina se almacenan en la memoria 110. Alternativamente, el código se puede ejecutar en un segundo sistema informático 140.

Los aspectos de los sistemas y métodos proporcionados en este documento, como el servidor 101, pueden incorporarse en la programación. Varios aspectos de la tecnología pueden considerarse como "productos" o "artículos de fabricación", normalmente en forma de código ejecutable por máquina (o procesador) y/o datos asociados que se transportan o incorporan en un tipo de medio legible por máquina. El código ejecutable por máquina se puede almacenar en una unidad de almacenamiento electrónico, tal memoria (por ejemplo, memoria de solo lectura, memoria de acceso aleatorio, memoria flash) o un disco duro. Los medios de tipo "almacenamiento" pueden incluir parte o la totalidad de la memoria tangible de los ordenadores, procesadores o similares, o módulos asociados de los mismos, como varias memorias de semiconductores, unidades de cinta, unidades de disco y similares, que pueden proporcionar almacenamiento no transitorio, en cualquier momento para la programación del software. En ocasiones, todo el software o partes del mismo pueden comunicarse a través de Internet o de otras redes de telecomunicaciones. Dichas comunicaciones, por ejemplo, pueden permitir la carga del software desde un ordenador o procesador a otro, por ejemplo, desde un servidor de gestión o un ordenador central a la plataforma informática de un servidor de aplicaciones. Por lo tanto, otro tipo de medios que pueden soportar los elementos de software incluyen ondas ópticas, eléctricas y electromagnéticas, como las que se utilizan a través de interfaces físicas entre dispositivos locales, a través de redes fijas alámbricas y ópticas y a través de varios enlaces aéreos. Los elementos físicos que transportan tales ondas, tales como enlaces alámbricos o inalámbricos, enlaces ópticos o similares, también se pueden considerar como soportes del software. Como se usa en este documento, a menos que se limite a medios de

"almacenamiento" tangibles y no transitorios, los términos como "medio legible" de computadora o máquina pueden referirse a cualquier medio que participe en proporcionar instrucciones a un procesador para su ejecución.

5 Los sistemas informáticos descritos en el presente documento pueden comprender código ejecutable por ordenador para realizar cualquiera de los algoritmos o métodos basados en algoritmos descritos en el presente documento. En algunas aplicaciones, los algoritmos aquí descritos harán uso de una unidad de memoria que se compone de al menos una base de datos.

10 Los datos relacionados con la presente divulgación pueden transmitirse a través de una red o conexiones para su recepción y/o revisión por parte de un receptor. El receptor puede ser, pero no limitarse a, el sujeto a quien se refiere el informe; o a un cuidador del mismo, por ejemplo, un proveedor de atención médica, gerente, otro profesional de atención médica u otro cuidador; una persona o entidad que realizó y/u ordenó el análisis. El receptor también puede ser un sistema local o remoto para almacenar dichos informes (por ejemplo, servidores u otros sistemas de una arquitectura de "computación en la nube"). En una realización, un medio legible por ordenador incluye un medio adecuado para la transmisión de un resultado de un análisis de una muestra 15 biológica usando los métodos descritos en este documento.

20 Aspectos de los sistemas y métodos proporcionados en este documento, como el sistema informático 101 en la Figura 66 puede incorporarse en la programación. Varios aspectos de la tecnología pueden considerarse como "productos" o "artículos de fabricación", normalmente en forma de código ejecutable por máquina (o procesador) y/o datos asociados que se transportan o incorporan en un tipo de medio legible por máquina. El código ejecutable por máquina se puede almacenar en una unidad de almacenamiento electrónico, como la memoria (por ejemplo, memoria de solo lectura, memoria de acceso aleatorio, memoria flash) o un disco duro. Los medios de tipo "almacenamiento" pueden incluir parte o la totalidad de la memoria tangible de las computadoras, procesadores o similares, o módulos asociados de los mismos, como varias memorias de semiconductores, unidades de cinta, unidades de disco y similares, que pueden proporcionar almacenamiento 25 no transitorio, en cualquier momento para la programación del software. En ocasiones, todo el software o partes del mismo pueden comunicarse a través de Internet o de otras redes de telecomunicaciones. Dichas comunicaciones, por ejemplo, pueden permitir la carga del software desde un ordenador o procesador a otro, por ejemplo, desde un servidor de gestión o un ordenador central a la plataforma informática de un servidor de aplicaciones. Por lo tanto, otro tipo de medios que pueden soportar los elementos de software incluyen ondas 30 ópticas, eléctricas y electromagnéticas, como las que se utilizan a través de interfaces físicas entre dispositivos locales, a través de redes fijas alámbricas y ópticas y a través de varios enlaces aéreos. Los elementos físicos que transportan dichas ondas, tales como enlaces alámbricos o inalámbricos, enlaces ópticos o similares, también pueden ser considerados medios portadores del software. Tal como se utiliza en el presente documento, a menos que se limite a medios de "almacenamiento" tangibles y no transitorios, los términos como 35 "medio legible" por ordenador o máquina se refieren a cualquier medio que participe en proporcionar instrucciones a un procesador para su ejecución.

40 Por lo tanto, un medio legible por una máquina, tal como un código ejecutable por ordenador, puede tomar muchas formas, que incluyen, pero no se limitan a, un medio de almacenamiento tangible, un medio de onda portadora o un medio de transmisión física. Los medios de almacenamiento no volátiles incluyen, por ejemplo, discos ópticos o magnéticos, como cualquiera de los dispositivos de almacenamiento en cualquier ordenador o similar, como los que se pueden usar para implementar las bases de datos, etc., que se muestran en los dibujos. Los medios de almacenamiento volátiles incluyen la memoria dinámica, como la memoria principal de dicha plataforma informática. Los medios de transmisión tangibles incluyen cables coaxiales; cable de cobre y fibra óptica, incluidos los cables que componen un bus dentro de un sistema informático. Los medios de 45 transmisión de ondas portadoras pueden adoptar la forma de señales eléctricas o electromagnéticas, u ondas acústicas o luminosas como las generadas durante las comunicaciones de datos por radiofrecuencia (RF) e infrarrojos (IR). Por lo tanto, las formas comunes de medios legibles por ordenador incluyen, por ejemplo: un disquete, un disco flexible, un disco duro, una cinta magnética, cualquier otro medio magnético, un CD-ROM, DVD o DVD-ROM, cualquier otro medio óptico, papel de tarjetas perforadas cinta, cualquier otro medio de almacenamiento físico con patrones de agujeros, una RAM, una ROM, una PROM y una EPROM, una FLASH- 50 EPROM, cualquier otro chip o cartucho de memoria, una onda portadora que transporta datos o instrucciones, cables o enlaces que transportan dicha portadora onda, o cualquier otro medio desde el cual un ordenador pueda leer código de programación y/o datos. Muchas de estas formas de medios legibles por ordenador pueden estar involucradas en llevar una o más secuencias de una o más instrucciones a un procesador para su ejecución.

Propiedades fisicoquímicas

55 En algunas realizaciones, la pluralidad de elementos sensores comprende, consiste esencialmente en, o consiste en una pluralidad de partículas, donde cada partícula se diferencia entre sí por al menos una propiedad fisicoquímica tal que cada elemento sensor tiene una firma de corona de biomolécula única cuando se coloca en contacto con la misma muestra.

La propiedad fisicoquímica del elemento sensor que se encuentra en una matriz se refiere, por ejemplo, a la composición, tamaño, carga superficial, hidrofobicidad, hidrofilicidad, funcionalidad superficial (grupos funcionales superficiales), topografía superficial, curvatura superficial y forma. El término composición abarca el uso de diferentes tipos de materiales y las diferencias en las propiedades químicas y/o físicas de los materiales, por ejemplo, la conductividad del material elegido entre los elementos sensores.

5 La curvatura de la superficie generalmente está determinada por el tamaño de las nanopartículas. Por lo tanto, a escala nanométrica, a medida que cambia el tamaño de las nanopartículas, cambia la curvatura de la superficie de la partícula, y este cambio de la curvatura de la superficie afecta la selectividad de unión de la superficie. Por ejemplo, en cierta curvatura, la superficie de la partícula puede tener una afinidad de unión por 10 un tipo específico de biomolécula donde una curvatura diferente tendrá una afinidad de unión diferente y/o una afinidad de unión por una biomolécula diferente. La curvatura se puede ajustar para crear una pluralidad de elementos sensores con afinidad alterada por diferentes biomoléculas. Se puede crear una matriz de sensores que incluya una pluralidad de elementos sensores que tengan diferentes curvaturas (por ejemplo, diferentes tamaños) lo que da como resultado una pluralidad de elementos sensores, cada uno con una firma de corona 15 de biomolécula diferente.

20 La morfología de la superficie también puede modificarse mediante métodos tales como modelado de la superficie para proporcionar diferentes afinidades, ingeniería de curvaturas de la superficie en múltiples escalas de longitud y similares. El patrón de la superficie se proporciona, por ejemplo, formando los elementos sensores mediante polimerización en bloque en la que los al menos dos bloques tienen químicas diferentes, formando las nanopartículas usando mezclas de al menos dos polímeros diferentes y separando las fases de los polímeros durante la polimerización, y/o reticular los polímeros separados después de la separación de fases. La curvatura de superficie diseñada en escalas de longitud múltiple se proporciona, por ejemplo, mediante el empleo de emulsiones de Pickering (Sacanna et al. 2007) estabilizado por partículas finamente divididas para 25 la síntesis de nanopartículas. En algunas realizaciones, las partículas finamente divididas se seleccionan de, por ejemplo, silicatos, aluminatos, titanatos, óxidos metálicos tales como óxidos de aluminio, silicio, titanio, níquel, cobalto, hierro, manganeso, cromo o vanadio, negros de carbono y nitruros o carburos, por ejemplo, nitruro de boro, carburo de boro, nitruro de silicio o carburo de silicio, entre otros.

30 Por ejemplo, los elementos sensores que incluyen elementos sensores a nanoescala pueden funcionalizarse cada uno para tener diferentes propiedades fisicoquímicas. Los métodos adecuados para funcionalizar los elementos sensores se conocen en la técnica y dependen de la composición del elemento sensor (por ejemplo, oro, óxido de hierro, sílice, plata, etc.), e incluyen, pero no se limitan a, por ejemplo, funcionalizados con aminopropilo, funcionalizados con amina, funcionalizado con ácido borónico, funcionalizado con ácido carboxílico, funcionalizado con metilo, funcionalizado con N-succinimidil éster, funcionalizado con PEG, funcionalizado con estreptavidina, funcionalizado con éter metílico, funcionalizado con 35 trietoxipropilaminosilano, funcionalizado con tiol, funcionalizado con PCP, funcionalizado con citrato, funcionalizado con ácido lipoico, funcionalizado con BPEI, funcionalizado con carboxilo, funcionalizados con hidroxilo, y similares. En una realización, las nanopartículas pueden funcionalizarse con un grupo amina (-NH₂) o un grupo carboxilo (COOH). En algunas realizaciones, los elementos sensores a nanoescala se funcionalizan con un grupo funcional polar. Los ejemplos no limitantes del grupo funcional polar comprenden un grupo carboxilo, un grupo hidroxilo, un grupo tiol, un grupo ciano, un grupo nitrógeno, un grupo amonio, un grupo imidazolio, 40 un grupo sulfonio, un grupo piridinio, un grupo pirrolidinio, un grupo fosfonio o cualquier combinación de los mismos. En algunas realizaciones, el grupo funcional es un grupo funcional ácido (por ejemplo, un grupo ácido sulfónico, un grupo carboxilo y similares), un grupo funcional básico (por ejemplo, un grupo amino, un grupo amino secundario cíclico (como un grupo pirrolidilo y un grupo piperidilo), grupo piridilo, grupo imidazol, 45 grupo guanidina, etc.), un grupo carbamoilo, un grupo hidroxilo, un grupo aldehído y similares. En algunas realizaciones, el grupo funcional polar es un grupo funcional iónico. Los ejemplos no limitantes del grupo de función iónica comprenden un grupo amonio, un grupo imidazolio, un grupo sulfonio, un grupo piridinio, un grupo pirrolidinio, un grupo fosfonio. En algunas realizaciones, los elementos sensores se funcionalizan con un grupo funcional polímerizable. Los ejemplos no limitativos del grupo funcional polímerizable incluyen un grupo 50 vinilo y un grupo (met)acrílico. En algunas realizaciones, el grupo funcional es acrilato de pirrolidilo, ácido acrílico, ácido metacrílico, acrilamida, metacrilato de 2-(dimetilamino)etilo, metacrilato de hidroxietilo y similares.

55 En otras realizaciones, las propiedades fisicoquímicas de los elementos sensores pueden modificarse modificando la carga superficial. Por ejemplo, la superficie se puede modificar para proporcionar una carga neutra neta, una carga superficial positiva neta, una carga superficial negativa neta o una carga zwitteriónica. La carga de la superficie se puede controlar durante la síntesis del elemento o mediante la modificación posterior a la síntesis de la carga a través de la funcionalización de la superficie. Para las nanopartículas poliméricas, las diferencias de carga se pueden obtener durante la síntesis utilizando diferentes procedimientos de síntesis, diferentes comonomeros cargados y en sustancias inorgánicas al tener estados de oxidación mixtos.

60 Nanopartículas

En algunas realizaciones, las partículas son nanopartículas. En algunas realizaciones, las partículas son liposomas. Los liposomas pueden comprender cualquier lípido capaz de formar una partícula. El término "lípido" se refiere a un grupo de compuestos orgánicos que son ésteres de ácidos grasos y se caracterizan por ser insolubles en agua pero solubles en muchos disolventes orgánicos. Los lípidos generalmente se dividen en al menos tres clases: (1) "lípidos simples" que incluyen grasas y aceites así como ceras; (2) "lípidos compuestos" que incluyen fosfolípidos y glicolípidos; y (3) "lípidos derivados" tales como esteroideos. En una realización, el liposoma comprende uno o más lípidos catiónicos o lípidos aniónicos y uno o más lípidos estabilizadores. Los liposomas adecuados son conocidos en la técnica e incluyen, pero no se limitan a, por ejemplo, DOPG (1,2-dioleoil-sn-glicero-3-fosfo-(1'-rac-glicerol), DOTAP(1,2-Dioleil-3 trimetilamonio-propano)-DOPE (dioleoilfosfatidiletanolamina), CHOL (DOPC-colesterol) y combinaciones de los mismos.

La superficie a base de lípidos de un liposoma puede ponerse en contacto con un subconjunto de biomoléculas (por ejemplo, proteínas) de una muestra biológica compleja (por ejemplo, plasma o cualquier muestra que tenga una mezcla compleja de biomoléculas como proteínas y ácido nucleico y al menos una de una polisacárido y lípido) en una interfase lípido-biomolécula (por ejemplo, proteína), uniendo así el subconjunto de proteínas para producir un patrón de unión de biomolécula (por ejemplo, proteína).

En una realización, el liposoma comprende un lípido catiónico. Como se usa aquí, el término "lípido catiónico" se refiere a un lípido que es catiónico o se vuelve catiónico (protonado) cuando el pH desciende por debajo del pK del grupo ionizable del lípido, pero es progresivamente más neutral a valores de pH más altos. A valores de pH por debajo del pK, el lípido puede asociarse con ácidos nucleicos cargados negativamente. En ciertas realizaciones, el lípido catiónico comprende un lípido zwitteriónico que asume una carga positiva al disminuir el pH. En ciertas realizaciones, los liposomas comprenden lípidos catiónicos. En algunas realizaciones, el lípido catiónico comprende cualquiera de varias especies de lípidos que llevan una carga neta positiva a un pH selectivo, tal como el pH fisiológico. Tales lípidos incluyen, pero no se limitan a, cloruro de N,N-dioleil-N,N-dimetilamonio (DODAC); cloruro de N-(2,3-dioleiloxi)propil)-N,N,N-trimetilamonio (DOTMA); bromuro de N,N-diestearil-N,N-dimetilamonio (DDAB); cloruro de N-(2,3-dioleiloxi)propil)-N,N,N-trimetilamonio (DOTAP); 3-(N-(N',N'-dimetilaminoetano)-carbamoil) colesterol (DC-Chol), N-(1-(2,3-dioleiloxi)propil)-N-2-(esperminacarboxamido)etil)-N,N-dimetil-trifluoroacetato de amonio (DOSPA), dioctadecilamidoglicil carboxiespermina (DOGS), 1,2-dioleoil-3-dimetilamonio propano (DODAP), N,N-dimetil-2,3-dioleiloxi)propilamina (DODMA), Bromuro de N-(1,2-dimiristiloxyprop-3-il)-N,N-dimetil-N-hidroxietilamonio (DMRIE), 1,2-dioleoil-sn-3-fosfoetanolamina (DOPE), trifluoroacetato de N-(1-(2,3-dioleiloxi)propil)-N-(2-(esperminacarboxamido)etil)-N,N-dimetilamonio (DOSPA), dioctadecilamidoglicilcarboxiespermina (DOGS) y 1,2-ditetradecanoil-sn-glicero-3-fosfocolina (DMPC). Los siguientes lípidos son catiónicos y tienen una carga positiva por debajo del pH fisiológico: DODAP, DODMA, DMDMA, 1,2-dilinoleiloxi-N,N-dimetilaminopropano (DLnDMA), 1,2-dilinoleniloxi-N,N-dimetilaminopropano (DLenDMA). En alguna realización, el lípido es un aminolípido.

En ciertas realizaciones, el liposoma comprende uno o más lípidos adicionales que estabilizan la formación de partículas durante su formación. Los lípidos estabilizadores adecuados incluyen lípidos neutros y lípidos aniónicos. El término "lípido neutro" se refiere a cualquiera de una serie de especies de lípidos que existen en forma zwitteriónica neutra o sin carga a pH fisiológico. Los lípidos neutros representativos incluyen diacilfosfatidilcolinas, diacilfosfatidiletanolaminas, ceramidas, esfingomielinas, dihidroesfingomielinas, cefalinas y cerebrósidos. Los lípidos neutros ejemplares incluyen, por ejemplo, diestearoilfosfatidilcolina (DSPC), dioleoilfosfatidilcolina (DOPC), dipalmitoilfosfatidilcolina (DPPC), dioleoilfosfatidilglicerol (DOPG), dipalmitoilfosfatidilglicerol (DPPG), dioleoilfosfatidiletanolamina (DOPE), palmitoiloleoilfosfatidilcolina (POPC), palmitoiloleoilfosfatidilcolina (POPC), y dioleoilfosfatidiletanolamina 4-(N-maleimidometil)-ciclohexano-1-carboxilato (DOPE-mal), dipalmitoilfosfatidiletanolamina (DPPE), dimiristilofoetanolamina (DMPE), diestearoil-fosfatidiletanolamina (DSPE), 16-O-monometil PE, 16-O-dimetil PE, 18-1-trans PE, 1-estearoil-2-oleoil-fosfatidietanolamina (SOPE) y 1,2-dielaidoil-sn-glicero-3-fosfoetanolamina (transDOPE). En una realización, el lípido neutro es 1,2-diestearoil-sn-glicero-3-fosfocolina (DSPC).

El término "lípido aniónico" se refiere a cualquier lípido que esté cargado negativamente a pH fisiológico. Estos lípidos incluyen fosfatidilglicerol, cardiolipina diacilfosfatidilserina, ácido diacilfosfatídico, N-dodecanoilfosfatidiletanolaminas, N-succinilfosfatidiletanolaminas, N-glutarilfosfatidiletanolaminas, lisilfosfatidigliceroles, palmitoiloleoilfosfatidilglicerol (POPG) y otros grupos modificadores aniónicos unidos a lípidos neutros. En ciertas realizaciones, el liposoma comprende glicolípidos (por ejemplo, monosialogangliósido GM1). En ciertas realizaciones, el liposoma comprende un esterol, tal como el colesterol. En ciertas realizaciones, el liposoma comprende un lípido estabilizador adicional que es un lípido de polietilenglicol. Los lípidos de polietilenglicol adecuados incluyen fosfatidiletanolamina modificada con PEG, ácido fosfatídico modificado con PEG, ceramidas modificadas con PEG (por ejemplo, PEG-CerC14 o PEG-CerC20), dialquilaminas modificadas con PEG, diacilgliceroles modificados con PEG, dialquigliceroles modificados con PEG. Los lípidos de polietilenglicol representativos incluyen PEG-c-DOMG, PEG-c-DMA y PEG-s-DMG. En una realización, el polietilenglicol-lípido es N-[(metoxipoli(etylenglicol).sub.2000)carbamil]-1,2-dimiristiloxypropil-3-amino (PEG-c-DMA). En una realización, el polietilenglicol-lípido es PEG-c-DOMG).

Los liposomas adecuados pueden ser nanopartículas lipídicas sólidas (SLN) que pueden estar hechas de lípidos sólidos, emulsionantes y/o agua/disolvente. SLN puede incluir, pero no limitarse a, una combinación de los siguientes ingredientes: triglicéridos (triestearina), glicéridos parciales (Imwitor), ácidos grasos (ácido esteárico, ácido palmítico) y esteroides (colesterol) y ceras (palmitato de cetilo). Se han utilizado varios emulsionantes y su combinación (Pluronic F 68, F 127) para estabilizar la dispersión de lípidos. Los ingredientes adecuados para el uso en la preparación de elementos sensores SNL incluyen, pero no se limitan a, por ejemplo, fosfolípidos, glicerol, poloxámero 188, fosfatidilcolina de soja, compritol, palmitato de cetilo, PEG 2000, PEG 4500, Tween 85, oleato de etilo, alginato de Na, etanol/butanol, glicérido de triestearina, PEG 400, miristato de isopropilo, Pluronic F68, Tween 80, trimiristina, triestearina, trilaurina, ácido esteárico, caprato de glicerilo como Capmu®MCM C10, aceite de teobroma, aceite de coco triglicérido, 1-octadecanol, behenato de glicerol como Compritol® 888 ATO, palmitostearato de glicerol como Precirol® ATO 5, y cera de palmitato de cetilo y similares.

En algunas realizaciones, la pluralidad de elementos sensores comprende, consiste esencialmente en, o consiste en una pluralidad de medias partículas de diferentes formas geométricas que pueden fabricarse mediante tecnología de moldeo, impresión 3D o impresión 4D. Las medias partículas adecuadas son conocidas en la técnica e incluyen, pero no se limitan a, medias partículas y partículas parciales en cualquier forma geométrica, por ejemplo, esferas, varillas, triángulos, cubos y combinaciones de los mismos. De manera adecuada, en una realización, la pluralidad de medias partículas tienen diferentes propiedades fisicoquímicas fabricadas mediante impresión 3D.

En algunas realizaciones, los elementos sensores, incluidos los elementos sensores a nanoescala, se fabrican mediante impresión 3D o 4D. Se conocen en la técnica métodos adecuados de impresión 3D y 4D de elementos sensores, incluidos elementos sensores a nanoescala. El material adecuado para la impresión 3D y 4D incluye, pero no se limita a, por ejemplo, plásticos y polímeros sintéticos (por ejemplo, diacrilato de polietilenglicol (PEG-DA), polí(e-caprolactona) (PCL), polí(óxido de propileno) (PPO), polí(óxido de etileno) (PEO), etc.), metales, polvos, vidrio, cerámica e hidrogeles. Las formas adecuadas hechas por impresión 3D o 4D incluyen, pero no se limitan a, por ejemplo, esferas completas o parciales (por ejemplo, 3/4 o medias esferas), varillas, cubos, triángulos u otras formas geométricas o no geométricas.

Las técnicas de impresión 3D incluyen, pero no se limitan a, la impresión por microextrusión, la bioimpresión por inyección de tinta, la bioimpresión asistida por láser, la estereolitografía, la impresión omnidireccional y la impresión de sellos.

En algunas realizaciones, los elementos sensores a nanoescala son nanopartículas. Las nanopartículas adecuadas son conocidas en la técnica e incluyen, pero no se limitan a, por ejemplo, polímeros, copolímeros, terpolímeros naturales o sintéticos (con los núcleos compuestos de metales u óxidos inorgánicos, incluidos los núcleos magnéticos). Las nanopartículas poliméricas adecuadas incluyen, pero no se limitan a, por ejemplo, poliestireno; polí(lisina), quitosano, dextrano, polí(acrilamida) y sus derivados, tales como N-isopropilacrilamida, N-tertbutilacrilamida, N,N-dimetilacrilamida, polietilenglicol, polí(alcohol vinílico), gelatina, almidón, (bio)polímeros degradables, sílice y similares.

En diversas realizaciones, el núcleo de las nanopartículas puede incluir una partícula orgánica, una partícula inorgánica o una partícula que incluya tanto materiales orgánicos como inorgánicos. Por ejemplo, las partículas pueden tener una estructura de núcleo que sea o incluya una partícula de metal, una partícula de punto cuántico, una partícula de óxido de metal o una partícula de núcleo-envoltura. Por ejemplo, la estructura central puede ser o incluir una partícula polimérica o una partícula basada en lípidos, y los enlazadores pueden incluir un lípido, un tensioactivo, un polímero, una cadena de hidrocarburo o un polímero anfifílico. Por ejemplo, los enlazadores pueden incluir polietilenglicol o polialquilenglicol, por ejemplo, los primeros extremos de los enlazadores pueden incluir un lípido unido a polietilenglicol (PEG) y los segundos extremos pueden incluir grupos funcionales unidos al PEG. En estos métodos, el primer o segundo grupo funcional puede incluir un grupo amina, un grupo maleimida, un grupo hidroxilo, un grupo carboxilo, un grupo piridiltiolo o un grupo azida.

En ciertas realizaciones, las nanopartículas pueden comprender polímeros que incluyen, por ejemplo, poliestireno sulfonato de sodio (PSS), óxido de polietileno (PEO), polioxietilenglicol, polietilenglicol (PEG), polietilenimina (PEI), ácido poliláctico, policaprolactona, ácido poliglicólico, polímero de polí(láctido-co-glicólico) (PLGA), polímero de éter de celulosa, polivinilpirrolidona, acetato de vinilo, copolímero de polivinilpirrolidona-acetato de vinilo, alcohol polivinílico (PVA), acrilato, ácido poliacrílico (PAA), acetato de vinilo, copolímeros de ácido crotónico, poliacrilamida, polietilenfosfonato, polibutenofosfonato, poliestireno, polivinilfosfonato, polialquíleno, polímero de carboxivinilo, alginato de sodio, carragenano, goma xantana, goma arábiga, goma arábiga, goma guar, pululano, agar, quitina, quitosano, pectina, goma karaya, goma de algarrobo, maltodextrina, amilosa, almidón de maíz, almidón de patata, almidón de arroz, almidón de tapioca, almidón de guisantes, almidón de batata, almidón de cebada, almidón de trigo, almidón hidroxipropilado con alto contenido de amilosa, dextrina, leván, elsinan, gluten, colágeno, aislado de proteína de suero, caseína, proteína de leche, proteína de soja, queratina o una gelatina, o un copolímero, derivado o mezcla de los mismos.

En otras realizaciones, el polímero puede ser o incluir un polietileno, policarbonato, polianhídrido, polihidroxíácido, polipropilfumerato, policaprolactona, poliamida, poliacetal, poliéster, poliéster, poli(ortoéster), policianoacrilato, alcohol polivinílico, poliuretano, polifosfaceno, poliacrilato, polimetacrilato, policianoacrilato, poliurea, poliestireno o una poliamina, o un copolímero, derivado o mezcla de los mismos.

- 5 En algunas realizaciones, la presente divulgación proporciona nanopartículas que comprenden polímeros biodegradables. Los polímeros biodegradables ejemplares no limitativos pueden ser poli-β-amino-ésteres (PBAE), poli(amidoaminas), poliésteres que incluyen ácido poliláctico-co-glicólico (PLGA), polianhídridos, polímeros biorreducibles y otros polímeros biodegradables. En algunas realizaciones, el polímero biodegradable comprende poli(1,4-butanodiol diacrilato-co-4-amino-1-butanol) modificado en el extremo con 2-(3-aminopropilamino)etanol, poli(1,4-butanodiol diacrilato-co-4-amino-1-butanol) modificado en el extremo con (1-(3-aminopropil)-4- metilpiperazina, poli(1,4-butanodiol diacrilato-co-5-amino-1-pentanol) modificado en el extremo con 2-(3-aminopropilamino)etanol, poli(1,4-butanodiol diacrilato-co-5-amino-1-pentanol) modificado en el extremo con (1-(3-aminopropil)-4-metilpiperazina, diacrilato-co-3-amino-1-propanol) modificado en el extremo con 2-(3-aminopropilamino)etanol poli(1,5 pentanodiol y poli(1,5 pentanodiol diacrilato-co-3-amino-1-propanol) modificado en el extremo con (1-(3-aminopropil)-4-metilpiperazina.
- 10
- 15

Sustratos de matriz

- En algunas realizaciones, la matriz de sensores comprende un sustrato. Independientemente de la identidad del elemento sensor, esta invención puede materializarse mediante una matriz de elementos sensores inmovilizados, conectados y/o acoplados a un sustrato sólido. El sustrato puede comprender, consistir esencialmente en o consistir en polidimetilsiloxano (PDMS), sílice, oro o sustrato recubierto de oro, plata o sustrato recubierto de plata, platino o sustrato recubierto de platino, zinc o sustrato recubierto de zinc, sustrato recubierto de carbono y similares. Un experto en la materia podría seleccionar un sustrato apropiado para la matriz de sensores. En algunas realizaciones, los elementos sensores y el sustrato están hechos del mismo elemento, por ejemplo, oro. En algunas realizaciones, el sustrato y los elementos sensores (por ejemplo, nanopartículas) forman un chip.

- 20
- 25
- 30
- 35
- 40
- 45
- 50
- 55
- 60
- En algunas realizaciones, la pluralidad de elementos sensores comprende una única superficie, placa o chip que contiene dos o más elementos (regiones) sensores discretos con diferencias topológicas que permiten la formación de corona de biomoléculas discretas en cada elemento (región) discreto. La placa de superficie o chip puede fabricarse para incluir los dos o más elementos discretos (regiones) mediante los métodos descritos en este documento. Las regiones discretas pueden ser superficies elevadas de diferentes formas geométricas, diferentes tamaños o diferentes cargas u otras diferencias topológicas que dan como resultado elementos sensores discretos con capacidad para formar coronas de biomoléculas discretas.

En algunas realizaciones, los elementos sensores están unidos de forma no covalente al sustrato. Los métodos adecuados de unión no covalente se conocen en la técnica e incluyen, pero no se limitan a, por ejemplo, coordinación de metales, interacción de carga, interacción hidrófobo-hidrófobo, quelación y similares. En otras realizaciones, los elementos sensores se unen covalentemente al sustrato. Los métodos adecuados para unir covalentemente los elementos sensores y los sustratos incluyen, pero no se limitan a, por ejemplo, química de click, radiación y similares.

Solo con fines ilustrativos, los métodos para unir los elementos sensores, por ejemplo, elementos sensores a nanoscalas, a los sustratos se muestran en las FIGURAS 23-40. Por ejemplo, los elementos sensores pueden conjugarse con un sustrato (por ejemplo, sustrato de sílice) a través de la reacción de amidación entre los grupos amino en la superficie del sustrato de sílice y los grupos de ácido carboxílico en la superficie de las nanopartículas. (Figura 23), a través de la reacción de apertura del anillo entre los grupos epoxi en la superficie del sustrato de sílice y los grupos amino en la superficie de las nanopartículas (Figura 24), a través de la reacción de adición de Michael entre los grupos maleimida en la superficie del sustrato de sílice y los grupos tiol o amino en la superficie de las nanopartículas (Figura 25), a través de la reacción de uretano entre los grupos isocianato en la superficie del sustrato de sílice y los grupos hidroxilo o amino en la superficie de las nanopartículas (Figura 26), a través de la reacción de oxidación entre los grupos tiol en la superficie del sustrato de sílice y los de la superficie de las nanopartículas (Figura 27), a través de la química "Click" entre los grupos azida en la superficie del sustrato de sílice y los grupos alquino en la superficie de las nanopartículas (Figura 28), a través de la reacción de intercambio de tiol entre los grupos 2-piridilditio en la superficie del sustrato de sílice y los grupos tiol en la superficie de las nanopartículas (Figura 29), a través de la reacción de coordinación entre los grupos de ácido borónico en la superficie del sustrato de sílice y los grupos diol en la superficie de las nanopartículas (Figura 30), a través de la reacción de adición irradiada con luz ultravioleta entre los enlaces C=C en la superficie del sustrato y los enlaces C=C en la superficie de las nanopartículas (Figura 31) y similares. Los métodos adecuados para conjugar elementos sensores con sustrato de oro son conocidos en la técnica e incluyen, por ejemplo, la conjugación a través de enlaces Au-tiol. (Figura 32), a través de la reacción de amidación entre los grupos de ácido carboxílico en la superficie del sustrato de oro y los grupos amino en la superficie de las nanopartículas (Figura 33), a través de la química "Click" entre los grupos azida en la superficie del sustrato de oro y los grupos alquino en la superficie de las nanopartículas (Figura 34), a través de la reacción de uretano entre los grupos NHS en la superficie del sustrato de oro y los grupos amino en la superficie de las nanopartículas (Figura 35).

nanopartículas (Figura 35), a través de la reacción de apertura del anillo entre los grupos epoxi en la superficie del sustrato de oro y los grupos amino en la superficie de las nanopartículas (Figura 36), a través de la reacción de coordinación entre los grupos de ácido borónico en la superficie del sustrato de sílice y los grupos diol en la superficie de las nanopartículas (Figura 37), a través de la reacción de adición irradiada con luz ultravioleta entre los enlaces C=C en la superficie del sustrato de oro y los enlaces C=C en la superficie de las nanopartículas (Figura 38), a través de la interacción "Ligando-Receptor" entre la biotina en la superficie del sustrato de oro y la avidina en la superficie de las nanopartículas (Figura 39), a través de la interacción "Anfitrón-Huésped" entre la a-ciclodextrina (a-CD) en la superficie del sustrato de oro y la adamantina (Ad) en la superficie de las nanopartículas (Figura 40), y similares.

En otro ejemplo, se puede utilizar la llamada "química Click" para unir los grupos superficiales funcionales a las estructuras centrales de las nanopartículas (véase, por ejemplo, el catálogo de Sigma Aldrich y Patente de EE.UU. Nº 7,375,234). De las reacciones que comprenden el campo de la química Click, un ejemplo es la cicloadición 1,3-dipolar de Huisgen de alquinos a azidas para formar 1,4-disustituidos-1,2,3-triazoles. La reacción catalizada por cobre (I) es suave y muy eficiente, no requiere grupos protectores y no requiere purificación en muchos casos. Los grupos funcionales azida y alquino son generalmente inertes a las moléculas biológicas y ambientes acuosos. El triazol tiene similitudes con la fracción amida omnipresente que se encuentra en la naturaleza, pero a diferencia de las amidas, no es susceptible de escisión. Además, son casi imposibles de oxidar o reducir.

La pluralidad de elementos sensores se puede unir al sustrato al azar o en un patrón distinto. Los elementos sensores pueden colocarse sustancialmente de manera uniforme. El patrón de los elementos sensores dispuestos puede variar según el patrón en el que los elementos sensores están unidos al sustrato. Cada elemento sensor está separado por una distancia. La distancia entre los elementos sensores (por ejemplo, nanopartículas) dispuestos en el sustrato puede variar según la longitud del conector utilizado para unir u otras condiciones de fabricación. De acuerdo con diversas realizaciones, la pluralidad de elementos sensores en la matriz puede fabricarse con una distancia y un patrón deseados entre elementos. Se conocen en la técnica distintos patrones adecuados, que incluyen, pero no se limitan a, líneas paralelas, cuadrados, círculos, triángulos y similares. Además, los elementos sensores pueden disponerse en filas o columnas. En algunas realizaciones, el sustrato es un sustrato plano, en otras realizaciones, el sustrato tiene forma de microcanales o nanocanales. Solo con fines ilustrativos, las realizaciones adecuadas se describen en la Figura 15-22. Los elementos sensores pueden estar contenidos dentro de microcanales o nanocanales que restringen o controlan el flujo de la muestra a través de la matriz de sensores. Los microcanales adecuados pueden oscilar entre 10 μm y aproximadamente 100 μm de tamaño.

En algunas realizaciones, los ejemplos no limitantes de la pluralidad de elementos sensores incluyen, pero no se limitan a, (a) una pluralidad de elementos sensores hechos del mismo material pero que difieren en propiedades fisicoquímicas, (b) una pluralidad de elementos sensores donde uno o más elementos sensores están hechos de un material diferente con propiedades fisicoquímicas iguales o diferentes, (c) una pluralidad de elementos sensores hechos del mismo material que difieren en tamaño, (d) una pluralidad de elementos sensores hechos de diferentes materiales con relativamente el mismo tamaño; (e) una pluralidad de elementos sensores hechos de diferentes materiales y de diferentes tamaños, (f) una pluralidad de elementos sensores en los que cada elemento está hecho de un material diferente, (g) una pluralidad de elementos sensores que tienen diferentes cargas, entre otros. La pluralidad de elementos sensores puede estar en cualquier combinación adecuada de dos o más elementos sensores en la que cada elemento sensor proporciona una firma de corona de biomolécula única. Por ejemplo, la pluralidad de elementos sensores puede incluir uno o más liposomas y una o más nanopartículas descritas en este documento. En una realización, la pluralidad de elementos sensores puede ser una pluralidad de liposomas con contenido variable de lípidos y/o cargas variables (catiónico/aniónico/neutro). En otra realización, la pluralidad de sensores puede contener una o más nanopartículas hechas del mismo material pero de diferentes tamaños y propiedades fisicoquímicas. En otra realización, la pluralidad de sensores puede contener una o más nanopartículas hechas de diferentes materiales (por ejemplo, sílice y poliestireno) con tamaños y/o propiedades fisicoquímicas similares o variables (por ejemplo, modificaciones, por ejemplo, $-\text{NH}_2$, -funcionalización COOH). Estas combinaciones se proporcionan únicamente como ejemplos y no limitan el alcance de la invención.

El ángulo de curvatura en la superficie de las partículas puede cambiar dependiendo del tamaño de las partículas. Este cambio en el ángulo de curvatura, a su vez, cambia el área superficial a la que las proteínas pueden unirse e interactuar entre sí en las partículas. Como se muestra en la Figura 58, el aumento del tamaño de la partícula da como resultado un cambio en la cantidad de proteína unida y también en el patrón de proteínas unidas a las nanopartículas de diferentes tamaños (en este ejemplo, el análisis SDS-PAGE de proteínas en nanopartículas de diámetros de 0.1 μm , se muestran 3 μm y 4 μm).

La novedad de la matriz de sensores es que no solo puede detectar diferentes proteínas entre los diferentes elementos sensores, sino también la capacidad de comparar los niveles de la misma proteína entre los diferentes elementos sensores. Por ejemplo, sin estar ligado a ninguna teoría pero con el fin de ilustrar la unicidad de la presente matriz sensorial, se describe un ejemplo teórico. En algunas realizaciones, la muestra se pone en contacto con un primer elemento sensor A, un segundo elemento sensor B y un tercer elemento

sensor C, donde cada elemento sensor produce una firma de corona de proteína distante (es decir, A', B' y C'). Las composiciones de cada firma de corona de proteína A', B' y C' pueden ser diferentes entre sí. En algunas realizaciones, A' B' y C' pueden comprender la misma proteína pero en una cantidad diferente, lo que puede proporcionar información proteómica adicional que no se puede obtener mediante la caracterización de la muestra con enfoques previamente conocidos. En otras palabras, cada información de corona de proteína única de cada nanopartícula sirve como variables únicas y, por lo tanto, proporciona más datos proteómicos. Por ejemplo, la albúmina se puede encontrar en la firma de corona de proteínas de un solo elemento sensor A, en la firma de dos sensores, por ejemplo A y B, B y C o A y C, o en las tres firmas de biomoléculas del sensor (A, B, C). Además, la matriz de sensores no solo determina la presencia o ausencia de la proteína, por ejemplo, albúmina, sino que también puede determinar la comparación de la cantidad de proteína de un sensor a otro. Por ejemplo, la albúmina se puede encontrar en una concentración X en la firma de A, una concentración de 1/3X en el sensor B y una concentración de 2X en el sensor C para una huella de biomolécula específica. En otra huella de biomolécula, la misma proteína, la albúmina, podría encontrarse en 3 concentraciones diferentes, por ejemplo, 1/8X en el sensor D, 3X en el sensor E y 1/4X en el sensor F. Así, una pluralidad de sensores no da solo un punto de datos para la concentración de una proteína, pero puede haber una comparación de la concentración de la proteína entre dos o más sensores. Además, la concentración de proteínas raras o poco abundantes se puede comparar con la concentración de una proteína conocida, lo que proporciona más datos sobre las coronas de proteínas. Por ejemplo, con fines ilustrativos, la concentración de una proteína desconocida, por ejemplo proteína Z, puede compararse con la cantidad de una proteína conocida, por ejemplo albúmina en las diferentes coronas de biomoléculas. Por ejemplo, Z puede encontrarse en una relación de albúmina de 1:8 en el sensor A, 1:50 en el sensor B y no estar presente en el sensor C. La Figura 12 proporciona un análisis de la comparación de proteínas raras con la concentración de albúmina dentro de las coronas de proteínas analizadas. Por lo tanto, el análisis estadístico puede tomar tanto la presencia de una proteína, la concentración relativa entre cada elemento sensor y la concentración de una proteína rara o poco abundante en comparación con una proteína conocida de una concentración particular al analizar los datos.

En algunas realizaciones, un canal se forma mediante litografía, grabado, grabado en relieve o moldeado de una superficie polimérica. En general, el proceso de fabricación puede implicar uno o más de cualquiera de estos procesos, y las diferentes partes de la matriz pueden fabricarse utilizando diferentes métodos y ensamblarse o unirse entre sí.

La litografía implica el uso de luz u otra forma de energía, tal como un haz de electrones, para cambiar un material. Por lo general, un material polimérico o precursor (por ejemplo, fotoprotección, un material resistente a la luz) se recubre sobre un sustrato y se expone selectivamente a la luz u otra forma de energía. Dependiendo de la fotoprotección, las regiones expuestas de la fotoprotección permanecen o se disuelven en etapas de procesamiento posteriores conocidas generalmente como "revelado". Este proceso da como resultado un patrón de la fotoprotección sobre el sustrato. En algunas realizaciones, la fotoprotección se utiliza como maestro en un proceso de moldeo. En algunas realizaciones, se vierte un precursor polimérico sobre el sustrato con fotoprotector, se polimeriza (es decir, se cura) y se despega.

En algunas realizaciones, la fotoprotección se utiliza como máscara para un proceso de grabado. Por ejemplo, después de modelar la fotoprotección en un sustrato de silicio, los canales se pueden grabar en el sustrato utilizando un proceso de grabado de iones reactivos profundos (DRIE) u otro proceso de grabado químico conocido en la técnica (por ejemplo, grabado con plasma, grabado con KOH, grabado con HF, etc.). Se retira la fotoprotección y el sustrato se une a otro sustrato utilizando cualquiera de los procedimientos de unión conocidos en la técnica (por ejemplo, unión anódica, unión adhesiva, unión directa, unión eutéctica, etc.). Se pueden incluir múltiples pasos litográficos y de grabado y pasos de mecanizado, como taladrado, según se requiera.

En algunas realizaciones, un sustrato polimérico puede calentarse y presionarse contra un molde maestro para un proceso de estampado. El molde maestro se puede formar mediante una variedad de procesos, que incluyen litografía y maquinado. A continuación, el sustrato polimérico se une con otro sustrato para formar canales y/o un aparato mezclador. Los procesos de mecanizado pueden incluirse si es necesario.

En algunas realizaciones, se inyecta un polímero, metal o aleación fundidos en un molde adecuado y se deja enfriar y solidificar para un proceso de moldeo por inyección. El molde normalmente consiste en dos partes que permiten retirar el componente moldeado. Las piezas así fabricadas pueden unirse para dar como resultado el sustrato.

En algunas realizaciones, el ataque químico puede usarse para formar canales. Pueden usarse técnicas litográficas para modelar un material sobre un sustrato. Este material está recubierto por otro material de diferente naturaleza química. Este material puede someterse a procesos de litografía y grabado, u otro proceso de mecanizado. Luego, el sustrato se expone a un agente químico que elimina selectivamente el primer material. Se forman canales en el segundo material, dejando vacíos donde estaba presente el primer material antes del proceso de grabado.

En algunas realizaciones, los microcanales se mecanizan directamente en un sustrato mediante mecanizado láser o mecanizado CNC. Varias capas así mecanizadas pueden unirse entre sí para obtener el sustrato final.

5 En algunas realizaciones, el ancho o la altura de cada canal oscila entre aproximadamente 1 μm y aproximadamente 1000 μm . En algunas realizaciones, el ancho o la altura de cada canal oscila entre aproximadamente 5 μm y aproximadamente 500 μm . En algunas realizaciones, el ancho o la altura de cada canal oscila entre aproximadamente 10 μm y aproximadamente 100 μm . En algunas realizaciones, la anchura o la altura de cada canal oscila entre aproximadamente 25 μm y aproximadamente 100 μm . En algunas realizaciones, el ancho o la altura de cada canal oscila entre aproximadamente 50 μm y aproximadamente 100 μm . En algunas realizaciones, el ancho o la altura de cada canal oscila entre aproximadamente 75 μm y aproximadamente 100 μm . En algunas realizaciones, el ancho o la altura de cada canal oscila entre aproximadamente 10 μm y aproximadamente 75 μm . En algunas realizaciones, la anchura o la altura de cada canal oscila entre aproximadamente 10 μm y aproximadamente 50 μm . En algunas realizaciones, el ancho o la altura de cada canal oscila entre aproximadamente 10 μm y aproximadamente 25 μm .

10 15 En algunas realizaciones, el ancho o alto máximo de un canal es de aproximadamente 1 μm , aproximadamente 5 μm , aproximadamente 10 μm , aproximadamente 20 μm , aproximadamente 30 μm , aproximadamente 40 μm , aproximadamente 50 μm , aproximadamente 60 μm , aproximadamente 70 μm , aproximadamente 80 μm , aproximadamente 90 μm , aproximadamente 100 μm , aproximadamente 250 μm , aproximadamente 500 μm o aproximadamente 1000 μm .

20 25 En algunas realizaciones, el ancho de cada canal oscila entre aproximadamente 5 μm y aproximadamente 100 μm . En algunas realizaciones, el ancho de un canal es de aproximadamente 5 μm , aproximadamente 10 μm , aproximadamente 15 μm , aproximadamente 20 μm , aproximadamente 25 μm , aproximadamente 30 μm , aproximadamente 35 μm , aproximadamente 40 μm , aproximadamente 45 μm , aproximadamente 50 μm , aproximadamente 60 μm , aproximadamente 70 μm , aproximadamente 80 μm , aproximadamente 90 μm o aproximadamente 100 μm .

30 35 En algunas realizaciones, la altura de cada canal oscila entre aproximadamente 10 μm y aproximadamente 1000 μm . En algunas realizaciones, la altura de un canal es de aproximadamente 10 μm , aproximadamente 100 μm , aproximadamente 250 μm , aproximadamente 400 μm , aproximadamente 500 μm , aproximadamente 600 μm , aproximadamente 750 μm o aproximadamente 1000 μm . En realizaciones específicas, la altura del canal o canales por los que fluye la muestra es de aproximadamente 500 μm . En realizaciones específicas, la altura del canal o canales por los que fluye la muestra es de aproximadamente 500 μm .

40 45 En algunas realizaciones, la longitud de cada canal oscila entre aproximadamente 100 μm y aproximadamente 10 cm. En algunas realizaciones, la longitud de un canal es de aproximadamente 100 μm , aproximadamente 1.0 mm, aproximadamente 10 mm, aproximadamente 100 mm, aproximadamente 500 mm, aproximadamente 600 mm, aproximadamente 700 mm, aproximadamente 800 mm, aproximadamente 900 mm, aproximadamente 1.0 cm, aproximadamente 1.1 cm, aproximadamente 1.2 cm, aproximadamente 1.3 cm, aproximadamente 1.4 cm, aproximadamente 1.5 cm, aproximadamente 5 cm o aproximadamente 10 cm. En realizaciones específicas, la longitud del canal o canales por los que fluye la muestra es de aproximadamente 1.0 cm. En realizaciones específicas, la longitud del canal o canales por los que fluye la muestra es de aproximadamente 1.0 cm.

50 55 40 Nanosistema de corona de biomoléculas (no reivindicado *per se* pero puede ser útil para poner en práctica la invención)

En el presente documento se proporciona un nanosistema de corona de biomoléculas o matrices de sensores que comprenden una pluralidad de elementos sensores en los que la pluralidad de elementos sensores difieren entre sí en al menos una propiedad fisicoquímica. En algunas realizaciones, una pluralidad de elementos sensores son una pluralidad de nanopartículas. En algunas realizaciones, una pluralidad de nanopartículas son una pluralidad de liposomas. En algunas realizaciones, cada elemento sensor puede unirse a una pluralidad de biomoléculas en una muestra biológica compleja para producir una firma de corona de biomoléculas. En algunas realizaciones, cada elemento sensor tiene una firma de corona de biomoléculas distinta.

La firma de corona de biomoléculas se refiere a la composición, firma o patrón de diferentes biomoléculas que están unidas a cada elemento sensor separado o cada nanopartícula. En algunos casos, la firma de corona biomoléculas es una firma de corona proteína. En otro caso, la firma de corona de biomoléculas es una firma de corona de polisacárido. En otro caso más, la firma de corona de biomoléculas es una firma de corona de metabolito. En algunos casos, la firma de corona biomoléculas es una firma de corona lipidómica. La firma no solo se refiere a las diferentes biomoléculas, sino también a las diferencias en la cantidad, el nivel o la cantidad de la biomolécula unida al elemento sensor o a la nanopartícula, o a las diferencias en el estado conformacional de la biomolécula unida al elemento sensor o a la nanopartícula. nanopartícula Se contempla que las firmas de corona de biomoléculas de cada elemento sensor pueden contener algunas de las mismas biomoléculas, pueden contener biomoléculas distintas con respecto a los otros elementos sensores o nanopartículas, y/o pueden diferir en nivel o cantidad, tipo o confirmación de la biomolécula. La firma de corona de la biomolécula

puede depender no solo de las propiedades fisicoquímicas del elemento sensor o de la nanopartícula, sino también de la naturaleza de la muestra y la duración de la exposición. En algunas realizaciones, la firma de corona de biomolécula comprende las biomoléculas que se encuentran en una corona blanda y una corona dura.

5 En algunas realizaciones, la matriz de sensores comprende, consiste esencialmente en o consiste en un primer elemento sensor que produce una primera firma de corona de biomolécula y al menos un segundo elemento sensor o al menos una nanopartícula que produce al menos una segunda firma de corona de biomolécula cuando la matriz de sensor se pone en contacto con una muestra biológica compleja. Una huella de biomolécula es la combinación de la primera firma de biomolécula y la al menos una segunda firma de biomolécula. Se contempla que la firma de biomoléculas se pueda hacer a partir de al menos dos firmas de corona de biomoléculas hasta tantas firmas de biomoléculas diferentes que se ensayan, por ejemplo al menos 1000 firmas de corona de biomoléculas diferentes. La corona de biomoléculas se puede analizar por separado para cada elemento sensor a fin de determinar la firma de corona de biomoléculas para cada elemento, cada nanopartícula o cada liposoma y combinarse para determinar la huella de la biomolécula o las dos o más 10 coronas de biomoléculas se pueden analizar al mismo tiempo para desarrollar la huella de la biomolécula a la vez.

La huella de la biomolécula puede distinguir entre diferentes estados biológicos posibles (por ejemplo, estados de enfermedad) de un sujeto. En algunas realizaciones, la huella de la biomolécula se asocia con el desarrollo de una enfermedad o trastorno y/o se puede asociar con un estado de enfermedad del sujeto.

20 En algunas realizaciones, la huella de la biomolécula puede determinar un estado de enfermedad para un sujeto. El término "estado de enfermedad" para un sujeto como se usa en el presente documento se refiere a la capacidad de la matriz de sensores de la presente tecnología para poder diferenciar entre los diferentes estados de una enfermedad dentro de un sujeto. Este término abarca un estado previo a la enfermedad o un estado precursor de una enfermedad o trastorno (un estado en el que el sujeto puede no tener signos o 25 síntomas externos de la enfermedad o trastorno, pero desarrollará la enfermedad o trastorno en el futuro) y un estado de enfermedad en el que el sujeto tiene una etapa de la enfermedad o trastorno (por ejemplo, una etapa temprana, intermedia o tardía de la enfermedad o trastorno). En otras palabras, el estado de enfermedad es un espectro que abarca un continuo en cuanto a la salud de un sujeto con respecto a una enfermedad o trastorno. La matriz de la presente invención es capaz de distinguir diferentes estados de enfermedad para un 30 sujeto determinando una huella de biomolécula que se puede comparar con diferentes huellas de biomolécula que están asociadas con diferentes estados de enfermedad en el espectro y con sujetos sanos. En otro ejemplo, la huella de la biomolécula puede asociarse con un estado anterior a la enfermedad o un estado de enfermedad precursora, en el que el sujeto parece sano en ese momento sin signos o síntomas externos de una enfermedad, pero desarrollará la enfermedad en el futuro. En otro ejemplo, la huella de la biomolécula puede 35 indicar que el sujeto tiene la enfermedad, y es capaz de distinguir si la enfermedad está en las etapas temprana, intermedia o tardía por la huella de la biomolécula única asociada con cada etapa.

Como se discutió anteriormente, el estado de enfermedad también incluye un estado precursor de una enfermedad o trastorno. Este estado precursor es un estado en el que el sujeto no tiene ningún signo o síntoma externo de la enfermedad o trastorno (aunque puede haber cambios submacro dentro de las biomoléculas del 40 sujeto que se encuentran en su sangre u otros fluidos biológicos), pero desarrollará la enfermedad. o desorden en el futuro. Este estado precursor también se puede describir como un estado en el que se observan los primeros cambios patológicos de una enfermedad, por ejemplo cambios en la huella de la biomolécula de una muestra biológica del paciente.

Otro ejemplo de un estado biológico es un estado saludable, sin enfermedad. La matriz también puede 45 determinar una huella de biomolécula específica asociada con un estado sano, sin enfermedad, en el que el sujeto no desarrollará la enfermedad en el futuro. En este caso, el sujeto no tiene evidencia en el momento de la prueba de que tiene la enfermedad o la desarrollará en el futuro.

En el presente documento, el término "estado biológico" abarca cualquier característica biológica de un sujeto que pueda manifestarse en una muestra biológica como se define en el presente documento. Se puede detectar 50 un estado biológico usando los métodos descritos en este documento donde dos sujetos que difieren en el estado biológico manifiestan esas diferencias en la composición de una muestra. Por ejemplo, un estado biológico incluye un estado de enfermedad de un sujeto. Un estado de enfermedad puede detectarse cuando el estado de enfermedad da lugar a cambios en la composición molecular (por ejemplo, el nivel de una o más proteínas) de una muestra de un sujeto que expresa el estado de enfermedad en relación con una muestra de un sujeto que no tiene el estado de enfermedad (es decir, cuando el estado biológico es un estado saludable o 55 un estado sin enfermedad).

Otro ejemplo de un estado biológico es un nivel de respuesta de un sujeto a un tratamiento terapéutico particular (por ejemplo, la administración de uno o una combinación de fármacos o productos farmacéuticos). En algunas realizaciones, un estado biológico es la capacidad de respuesta (por ejemplo, con respecto a un umbral particular de análisis) de un sujeto a un fármaco particular. En otra realización, un estado biológico es la falta 60

de respuesta (por ejemplo, con respecto a un umbral particular de análisis) de un sujeto a un fármaco particular. En algunas realizaciones, el nivel de respuesta de un sujeto a un fármaco (es decir, el estado biológico de respuesta al fármaco o el estado biológico de falta de respuesta al fármaco) está asociado con factores tales como la variabilidad en el metabolismo o la farmacocinética del fármaco entre sujetos.

- 5 Otro ejemplo de un estado biológico es el nivel de respuesta inmune exhibido por un sujeto. En algunas realizaciones, el estado biológico puede aumentar la respuesta inmunitaria. En otras realizaciones, el estado biológico puede ser una respuesta inmunitaria disminuida. La respuesta inmune puede diferir entre sujetos como resultado de una serie de variables. Por ejemplo, la respuesta inmunitaria puede diferir entre sujetos como resultado de la diferente exposición a un antígeno introducido exógenamente (por ejemplo, asociado con 10 un virus o bacteria), como resultado de diferencias en su susceptibilidad a una enfermedad o trastorno autoinmune, o secundariamente como resultado de una respuesta a otros estados biológicos en un sujeto (por ejemplo, estados de enfermedad como el cáncer).

15 Otros ejemplos de estados biológicos que pueden asociarse con una huella de biomolécula incluyen la susceptibilidad de un sujeto a los efectos adversos asociados con la administración de un fármaco y la identificación del potencial de un sujeto para mostrar una reacción alérgica a la administración de una composición o sustancia particular.

20 La innovación en estas matrices de sensores y los métodos asociados difieren de los métodos actuales de detección o medición de la presencia o ausencia o niveles de ciertos biomarcadores para predecir si un sujeto puede tener predisposición o probabilidad de desarrollar una enfermedad o trastorno en etapas muy tempranas de la enfermedad, antes de que se pueda controlar cualquier signo o síntoma. De manera adecuada, la matriz 25 es capaz de diferenciar la salud del sujeto de la ausencia de enfermedad o trastorno, que tiene un precursor de una enfermedad o trastorno y que tiene la enfermedad o trastorno. Sin embargo, la invención no se limita a esas realizaciones y cubre el espectro de otros estados de enfermedad que pueden ocurrir dentro del continuo de salud y enfermedad de un sujeto.

25 Además, la innovación de las presentes matrices de sensores de la invención puede demostrarse mediante los siguientes ejemplos. La matriz de sensores de la presente invención es muy sensible y capaz de detectar no solo cambios en pequeñas cantidades de biomoléculas dentro de una muestra, sino también depender de las interacciones entre las biomoléculas. Por ejemplo, sin estar ligado a ninguna teoría pero con el fin de ilustrar la unicidad de la presente matriz sensorial, se describe un ejemplo teórico. Si, por ejemplo, se recolectan muestras 30 biológicas de un sujeto a lo largo del tiempo (por ejemplo, antes de cualquier signo y síntoma de la enfermedad, estado previo a la enfermedad y durante las etapas temprana e intermedia de una enfermedad). En estas muestras, simplemente midiendo un nivel de una biomolécula X en la muestra (por ejemplo, cuantificación de la cantidad) se puede encontrar que la concentración no se altera en los diferentes estados de enfermedad. Por lo tanto, medir la biomolécula X no sería un marcador útil para la enfermedad. Sin embargo, usando la 35 presente matriz de sensores, aunque el nivel del biomarcador X puede ser el mismo, la interacción de la biomolécula X con otras biomoléculas, Y y Z, puede cambiar la composición de la firma de corona de proteína asociada con la muestra a lo largo del tiempo y las muestras. Por ejemplo, el biomarcador X puede cambiar su asociación de asociarse con la biomolécula Y a la biomolécula Z, o con el tiempo la interacción de X con Y y Z conduce a un cambio conformacional en X. Este tipo de cambios, que no cambia la concentración general, pero 40 sí cambia la huella única de la biomolécula asociada con el estado de enfermedad (que incluye la biomolécula X) permitiría la distinción y asociación con otros estados patológicos. Esto no se descubriría si utiliza métodos del arte de medir la cuantificación de la biomolécula X dentro de la muestra, ya que este nivel permaneció igual durante toda la enfermedad. Este es un ensayo inesperadamente altamente específico y útil que no se había visto antes. Con este ensayo, se pueden determinar biomarcadores de una enfermedad que antes no se 45 caracterizaban como biomarcadores, incluidos patrones de biomarcadores, ya que es posible que estos marcadores no cambien en cantidades absolutas dentro de las muestras, pero tendrán un cambio constante y medible con respecto a la interacción con los elementos sensores en la matriz de sensores. La capacidad de distinguir patrones mediante el análisis de la corona de biomoléculas permite asociar estos patrones con diferentes estados de enfermedad.

50 Como se discutió anteriormente, la presente matriz de sensores tiene la capacidad de detectar proteínas de más de diez órdenes de magnitud, que es una sensibilidad más alta que cualquier método descrito anteriormente. La singularidad de la presente matriz de sensores se encuentra en la capacidad de la matriz para detectar proteínas independientemente del nivel de concentración dentro de la muestra. El ensayo se basa en la capacidad de la biomolécula (por ejemplo, proteína), conocida o desconocida, dentro de la muestra para interactuar con los elementos sensores, y para interactuar de manera diferente y en diferentes cantidades con diferentes elementos sensores y también para interactuar con las otras biomoléculas asociadas con el elemento sensor. Esto, a su vez, permite que exista la capacidad de detectar proteínas raras y de baja abundancia incluso en presencia de proteínas muy abundantes (tales como la albúmina) dentro de la muestra. 55 Como se demuestra en la Figura 12 del Ejemplo 1, se ha demostrado la capacidad de la matriz de sensores para detectar concentraciones superiores a 10 órdenes de magnitud incluso en presencia de proteínas muy abundantes. Además, la matriz de sensores también puede detectar proteínas con concentraciones plasmáticas desconocidas/no informadas, que no se han encontrado previamente con los métodos actuales.

Una característica única entre otras de la presente matriz de sensores es la capacidad del presente sistema para poder detectar proteínas poco abundantes y proteínas raras. La presente matriz de sensores se puede utilizar no solo para determinar un estado de enfermedad (sin enfermedad, preenfermedad o enfermedad en etapa temprana y tardía), sino también, en algunos casos, para distinguir entre subtipos de enfermedades (por ejemplo, distinguir entre diferentes tipos de cáncer, por ejemplo, cáncer de pulmón, cáncer de mama, mieloma, etc.).

5 Métodos de detección

La presente matriz de sensores se puede usar en una variedad de métodos descritos en este documento. Nuevamente, la invención en su sentido general es el método de acuerdo con la reivindicación independiente

10 1. La capacidad de determinar una huella de biomolécula única para muestras proporciona un medio novedoso e innovador para medir un estado de enfermedad específico en un sujeto. Estas huellas de biomoléculas se pueden usar para determinar el estado de enfermedad de un sujeto, diagnosticar o pronosticar una enfermedad en un sujeto o identificar patrones únicos de biomarcadores que están asociados con un estado de enfermedad o una enfermedad o trastorno. Por ejemplo, los cambios en la huella de una biomolécula en un sujeto a lo largo del tiempo (días, meses, años) permiten la capacidad de rastrear una enfermedad o trastorno en un sujeto (por ejemplo que puede estar asociado con la etapa temprana de una enfermedad o cualquier otro estado de enfermedad. Como se discutió anteriormente, la capacidad de detectar una enfermedad en una etapa temprana, por ejemplo, el cáncer, incluso antes de que se desarrolle por completo o haga metástasis, permite un aumento significativo en los resultados positivos para esos pacientes y la capacidad de aumentar la esperanza de vida y reducir la mortalidad asociada con esa enfermedad. Por lo tanto, la matriz de sensores de la presente invención brinda una oportunidad única para poder desarrollar huellas de biomoléculas asociadas con las etapas previas o estados precursores de la enfermedad.

25 Se entiende que incluso antes de que una enfermedad haya progresado y muestre signos o síntomas mensurables, a nivel macroscópico se están produciendo cambios en el cuerpo y los sistemas biológicos de un sujeto. Ser capaz de reconocer estos primeros signos previos a la enfermedad contemplados mediante el uso de la matriz de sensores de la presente invención. Los inventores han descubierto que al comparar la huella de biomolécula de un sujeto durante diferentes momentos de un estado de enfermedad (por ejemplo, antes de que una enfermedad haya mostrado síntomas y se haya desarrollado, después de los primeros signos o síntomas de la enfermedad, y/o en las últimas etapas de la enfermedad) proporciona una huella de biomolécula única vinculada a los diferentes estados de enfermedad relacionados con la progresión de la enfermedad. En otras palabras, se puede determinar una huella de biomolécula que permitiría poder identificar sujetos que van a desarrollar una enfermedad en un momento posterior. Esto permitiría un seguimiento temprano y un tratamiento temprano, mejorando en gran medida el resultado del sujeto diagnosticado con la enfermedad.

30 35 En algunas realizaciones, se proporciona un método para detectar una enfermedad o trastorno en un sujeto. El método comprende los pasos definidos en las reivindicaciones adjuntas.

40 El paso de determinar una huella de biomolécula asociada con la muestra comprende detectar la firma de corona de biomolécula para al menos dos elementos sensores, donde la combinación de al menos dos firmas de corona de biomoléculas produce la huella de biomoléculas. En algunas realizaciones, las firmas de corona de biomoléculas de los al menos dos elementos sensores se analizan por separado y los resultados se combinan para determinar la huella de la biomolécula. En algunas realizaciones, las firmas de corona de biomoléculas de al menos dos elementos se analizan al mismo tiempo o en la misma muestra.

45 50 55 El método para determinar la huella de la biomolécula comprende detectar y determinar las firmas de corona biomoleculares de los al menos dos elementos sensores. En algunas realizaciones, este paso se puede realizar separando la pluralidad de biomoléculas unidas a cada elemento sensor (por ejemplo, separando la corona de biomolécula del elemento sensor) y ensayando la pluralidad de biomoléculas para determinar la composición de la pluralidad de coronas de biomolécula para determinar un huella de biomoléculas. Según el diseño de la matriz, en algunos casos, la composición de la firma de corona de cada biomolécula de cada elemento sensor se analiza de forma independiente, y los resultados se combinan para producir la huella de la biomolécula (por ejemplo, cada elemento sensor está en un canal o compartimento separado en el que el La composición específica de la corona de biomoléculas para ese elemento sensor específico puede analizarse por separado (por ejemplo, separando las biomoléculas y analizándolas mediante espectrometría de masas y/o cromatografía o detectando la pluralidad de biomoléculas aún unidas al elemento sensor mediante fluorescencia, luminiscencia u otros significados). En otra realización, los al menos dos elementos sensores están en la misma matriz y la composición de la corona de biomoléculas para los al menos dos elementos sensores se analiza al mismo tiempo disociando la corona de biomoléculas de ambos elementos sensores en una solución y analizando esa solución para determinar la firma de una biomolécula. Este último método sería el método de elección si se utiliza una tecnología de matriz de chips.

60 Los métodos para ensayar la pluralidad de biomoléculas que componen la firma de corona de biomolécula o la huella de biomolécula son conocidos en la técnica e incluyen, pero no se limitan a, por ejemplo, electroforesis

en gel, cromatografía líquida, espectrometría de masas, espectroscopía de resonancia magnética nuclear (RMN), espectroscopía infrarroja por transformada de Fourier (FTIR), dicroísmo circular, espectrometría Raman y una combinación de los mismos. En una realización preferida, el ensayo es por cromatografía líquida, espectrometría de masas o una combinación de los mismos.

- 5 En una realización preferida, el ensayo del sensor es una matriz sin etiquetas.

En algunas realizaciones, se contempla que se pueden usar matrices marcadas, en las que se puede determinar una firma de corona como una señal de cambio (por ejemplo, fluorescencia, luminiscencia, carga, tintes colorimétricos). Un ejemplo adecuado se muestra en la Figura 42 y 43. Por ejemplo, una matriz puede incluir colorantes químicamente sensibles en una formulación imprimible para la detección e identificación de 10 biomoléculas basándose en las interacciones de equilibrio con las biomoléculas y los colorantes sensibles.

En algunas realizaciones, el elemento sensor comprende un complejo con un primer componente y un 15 fluoróforo polimérico u otro componente desactivador químicamente complementario al primer componente donde dicho complejo tiene un fondo inicial o fluorescencia de referencia. Una vez que el primer componente entra en contacto con la corona de la biomolécula, efectuará la extinción del fluoróforo y se podrá medir este cambio en la fluorescencia. Después de irradiar y/o excitar el sensor con un láser, el efecto y/o el cambio en la 20 fluorescencia de cada elemento del sensor se puede medir y comparar con la fluorescencia de fondo para producir la huella de la biomolécula.

Las matrices de sensores y los métodos descritos en el presente documento se usan para determinar un estado 25 de enfermedad y/o pronosticar o diagnosticar una enfermedad o trastorno. Las enfermedades o trastornos contemplados incluyen, pero no se limitan a, por ejemplo, cáncer, enfermedad cardiovascular, enfermedad endocrina, enfermedad inflamatoria, enfermedad neurológica y similares.

En una realización, la enfermedad o trastorno es cáncer. En realizaciones adecuadas, la matriz de sensores y 30 los métodos descritos en el presente documento no solo pueden diagnosticar el cáncer (por ejemplo, determinar si un sujeto (a) no tiene cáncer, (b) está en una etapa de desarrollo precanceroso, (c) está en etapa temprana etapa de cáncer, (d) está en una etapa tardía de cáncer) pero en algunas realizaciones es capaz de determinar el tipo de cáncer. Como se demuestra en los ejemplos a continuación, una matriz de sensores que comprende seis elementos sensores fue capaz de determinar con precisión el estado de la enfermedad de la presencia o ausencia de cáncer. Además, los Ejemplos demuestran que una matriz de sensores que comprende seis elementos sensores fue capaz de distinguir entre diferentes tipos de cáncer (por ejemplo, cáncer de pulmón, glioblastoma, meningioma, mieloma y cáncer de páncreas).

Los términos "cáncer" y "tumor", tal como se utilizan aquí de forma intercambiable, pretenden abarcar cualquier cáncer, enfermedad neoplásica y preneoplásica que se caracterice por un crecimiento anormal de células. El cáncer se puede seleccionar, por ejemplo, del grupo que consiste en cáncer de pulmón, cáncer de páncreas, mieloma, leucemia mieloide, meningioma, glioblastoma, cáncer de mama, carcinoma de células escamosas de 35 esófago, adenocarcinoma gástrico, cáncer de próstata, cáncer de vejiga, cáncer de ovario, cáncer de tiroides, cáncer neuroendocrino, carcinoma de colon, cáncer de ovario, cáncer de cabeza y cuello, enfermedad de Hodgkin, linfomas no Hodgkin, cáncer de recto, cánceres urinarios, cánceres uterinos, cánceres orales, cánceres de piel, cáncer de estómago, tumores cerebrales, cáncer de hígado, cáncer de laringe, cáncer de esófago, tumores mamarios, fibrosarcoma, mixosarcoma, liposarcoma, condrosarcoma, sarcoma osteogénico, 40 cordoma, angiosarcoma, endoteliosarcoma, sarcoma de Ewing, carcinoma de células escamosas, carcinoma de células basales, adenocarcinoma, carcinoma de glándulas sudoríparas, carcinoma de glándulas sebáceas, carcinoma papilar, adenocarcinomas papilares, cistaneocarcinoma, carcinoma medular, carcinoma broncogénico, carcinoma de células renales, hepatoma, carcinoma de vías biliares, coriocarcinoma, seminoma, carcinoma embrionario, tumor de Wilms, cáncer de cuello uterino, tumor testicular, cáncer de endometrio, 45 carcinoma de pulmón, carcinoma de pulmón de células pequeñas, carcinoma de vejiga, carcinoma epitelial, glioblastomas, neuronomas, craneofaringiomas, schwannomas, glioma, astrocitoma, meningioma, melanoma, neuroblastoma, retinoblastoma, leucemias y linfomas, leucemia linfocítica aguda y policitemia vera mielocítica aguda, mieloma múltiple, macroglobulinemia de Waldenstrom y enfermedad de cadenas pesadas, leucemias no 50 linfocíticas agudas, leucemia linfocítica crónica, leucemia mielógena crónica, leucemia linfoide aguda nula en la infancia (ALL) ALL tímica, ALL de células B, leucemia megacariocítica aguda, linfoma de Burkitt y leucemia de células T, carcinoma de pulmón de células no pequeñas pequeñas y grandes, leucemia granulocítica aguda, tumores de células germinales, cáncer de endometrio, cáncer gástrico, leucemia de células pilosas, cáncer de tiroides y otros cánceres conocidos en la técnica. En una realización preferida, el cáncer se selecciona del 55 grupo que consiste en cáncer de pulmón, cáncer de páncreas, mieloma, leucemia mieloide, meningioma, glioblastoma, cáncer de mama, carcinoma de células escamosas de esófago, adenocarcinoma gástrico, cáncer de próstata, cáncer de vejiga, cáncer de ovario, cáncer de tiroides y cáncer neuroendocrino.

Como se usa en este documento, los términos "enfermedad cardiovascular" (CVD) o "trastorno cardiovascular" se usan para clasificar numerosas afecciones que afectan el corazón, las válvulas cardíacas y la vasculatura (por ejemplo, venas y arterias) del cuerpo y abarcan enfermedades y afecciones que incluyen, pero se limitan a, a aterosclerosis, infarto de miocardio, síndrome coronario agudo, angina, insuficiencia cardíaca congestiva,

aneurisma aórtico, disección aórtica, aneurisma ilíaco o femoral, embolia pulmonar, fibrilación auricular, accidente cerebrovascular, ataque isquémico transitorio, disfunción sistólica, disfunción diastólica, miocarditis, taquicardia auricular, fibrilación ventricular, endocarditis, enfermedad vascular periférica y enfermedad arterial coronaria (CAD). Además, el término enfermedad cardiovascular se refiere a sujetos que finalmente tienen un evento cardiovascular o complicación cardiovascular, refiriéndose a la manifestación de una condición adversa en un sujeto provocada por una enfermedad cardiovascular, tal como muerte súbita cardíaca o síndrome coronario agudo, que incluye, pero no limitado a, infarto de miocardio, angina inestable, aneurisma, accidente cerebrovascular, insuficiencia cardíaca, infarto de miocardio no fatal, accidente cerebrovascular, angina de pecho, ataques isquémicos transitorios, aneurisma aórtico, disección aórtica, cardiomiopatía, cateterismo cardíaco anormal, generación de imágenes cardíacas anormales, revascularización por stent o injerto, riesgo de experimentar una prueba de esfuerzo anormal, riesgo de experimentar una perfusión miocárdica anormal y muerte.

Como se usa en el presente documento, la capacidad para detectar, diagnosticar o pronosticar enfermedades cardiovasculares, por ejemplo, aterosclerosis, puede incluir determinar si el paciente se encuentra en una etapa previa de enfermedad cardiovascular, ha desarrollado formas tempranas, moderadas o graves de enfermedad cardiovascular, o ha sufrido uno o más eventos cardiovasculares o complicaciones asociadas con la enfermedad cardiovascular.

La aterosclerosis (también conocida como enfermedad vascular arteriosclerótica o ASVD) es una enfermedad cardiovascular en la que la pared de una arteria se engrosa como resultado de la invasión y acumulación y depósito de placas arteriales que contienen glóbulos blancos en la capa más interna de las paredes de las arterias, lo que resulta en la estrechamiento y endurecimiento de las arterias. La placa arterial es una acumulación de macrófagos o restos, y contiene lípidos (colesterol y ácidos grasos), calcio y una cantidad variable de tejido conjuntivo fibroso. Las enfermedades asociadas con la aterosclerosis incluyen, pero no se limitan a, aterotrombosis, cardiopatía coronaria, trombosis venosa profunda, enfermedad de la arteria carótida, angina de pecho, enfermedad arterial periférica, enfermedad renal crónica, síndrome coronario agudo, estenosis vascular, infarto de miocardio, aneurisma o accidente cerebrovascular.

Con fines ilustrativos, en una realización, las matrices de sensores pueden distinguir las diferentes etapas de la aterosclerosis, incluidos, pero no limitados a, los diferentes grados de estenosis en un sujeto.

Además, solo con fines ilustrativos, los ejemplos siguientes demuestran el uso de una serie de sensores para detectar los diferentes estados de la enfermedad de las arterias coronarias. Una matriz de sensores que contenía seis elementos sensores pudo distinguir sujetos con CAD diagnosticados por angiografía coronaria, pacientes con síntomas que tenían vasos coronarios sanos (NO CAD), pacientes con reestenosis (recurrencia de CAD después del tratamiento) y sujetos sanos sin factores de riesgo (Figura 53). La matriz de sensores actual fue lo suficientemente sensible como para detectar la diferencia entre las personas que presentan síntomas de enfermedad de las arterias coronarias pero que no tienen estenosis de las arterias (por ejemplo, grupos SIN CAD frente a grupos con CAD). Esto proporciona una nueva prueba de CAD de diagnóstico que se puede utilizar como una prueba de detección no invasiva para pacientes en riesgo.

El término "enfermedad endocrina" se usa para referirse a un trastorno asociado con la desregulación del sistema endocrino de un sujeto. Las enfermedades endocrinas pueden resultar de una glándula que produce demasiada o muy poca hormona endocrina, lo que provoca un desequilibrio hormonal, o debido al desarrollo de lesiones (como nódulos o tumores) en el sistema endocrino, que pueden o no afectar los niveles hormonales. Las enfermedades endocrinas adecuadas que pueden tratarse incluyen, pero no se limitan a, acromegalía, enfermedad de Addison, cáncer suprarrenal, trastornos suprarrenales, cáncer anaplásico de tiroides, síndrome de Cushing, tiroiditis de De Quervain, diabetes, cáncer folicular de tiroides, diabetes gestacional, bocio, Enfermedad de Graves, trastornos del crecimiento, deficiencia de la hormona del crecimiento, tiroiditis de Hashimoto, cáncer de tiroides de células de Hurthle, hiperglucemia, hiperparatiroidismo, hipertiroidismo, hipoglucemia, hipoparatiroidismo, hipotiroidismo, testosterona baja, cáncer medular de tiroides, MEN 1, MEN 2A, MEN 2B, menopausia, síndrome metabólico, obesidad, osteoporosis, cáncer papilar de tiroides, enfermedades de las paratiroides, feocromocitoma, trastornos hipofisarios, tumores hipofisarios, síndrome de ovario poliquístico, prediabetes, silente, tiroiditis, cáncer de tiroides, enfermedades de la tiroides, nódulos tiroideos, tiroiditis, síndrome de Turner, diabetes tipo 1, diabetes tipo 2, y similares.

Tal como se menciona en el presente documento, enfermedad inflamatoria se refiere a una enfermedad provocada por una inflamación descontrolada en el cuerpo de un sujeto. La inflamación es una respuesta biológica del sujeto a un estímulo nocivo que puede ser externo o interno, como patógenos, células y tejidos necrosados, irritantes, etc. Sin embargo, cuando la respuesta inflamatoria se vuelve anormal, da como resultado una lesión del tejido propio y puede provocar diversas enfermedades y trastornos. Las enfermedades inflamatorias pueden incluir, pero no se limitan a, asma, glomerulonefritis, enfermedad inflamatoria intestinal, artritis reumatoide, hipersensibilidad, enfermedad pélvica inflamatoria, enfermedades autoinmunes, artritis; enterocolitis necrotizante (NEC), gastroenteritis, enfermedad inflamatoria pélvica (PID), enfisema, pleuresia, pielitis, faringitis, angina, acné vulgar, infección del tracto urinario, apendicitis, bursitis, colitis, cistitis, dermatitis, flebitis, rinitis, tendinitis, amigdalitis, vasculitis, enfermedades autoinmunes; enfermedad celíaca; prostatitis

crónica, hipersensibilidades, lesión por reperfusión; sarcoidosis, rechazo de trasplantes, vasculitis, cistitis intersticial, fiebre del heno, periodontitis, aterosclerosis, psoriasis, espondilitis anquilosante, artritis idiopática juvenil, enfermedad de Behcet, espondiloartritis, uveítis, lupus eritematoso sistémico y cáncer. Por ejemplo, la artritis incluye artritis reumatoide, artritis psoriásica, osteoartritis o artritis idiopática juvenil y similares.

- 5 Los trastornos neurológicos o enfermedades neurológicas se usan indistintamente y se refieren a enfermedades del cerebro, la columna vertebral y los nervios que los conectan. Las enfermedades neurológicas incluyen, pero no se limitan a, tumores cerebrales, epilepsia, enfermedad de Parkinson, enfermedad de Alzheimer, ALS, malformación arteriovenosa, enfermedad cerebrovascular, aneurismas cerebrales, epilepsia, esclerosis múltiple, neuropatía periférica, neuralgia posherpética, accidente cerebrovascular, demencia frontotemporal, enfermedad desmielinizante (que incluyen, pero no se limitan a, esclerosis múltiple, enfermedad de Devic (es decir, neuromielitis óptica), mielinolisis pontina central, leucoencefalopatía multifocal progresiva, leucodistrofias, síndrome de Guillain-Barré, neuropatía inflamatoria progresiva, enfermedad de Charcot-Marie-Tooth, polineuropatía desmielinizante inflamatoria crónica, y neuropatía periférica anti-MAG) y similares. Los trastornos neurológicos también incluyen trastornos neurológicos inmunomediados (IMND), que incluyen enfermedades en las que al menos un componente del sistema inmunológico reacciona contra las proteínas del huésped presentes en el sistema nervioso central o periférico y contribuye a la patología de la enfermedad. Las IMND pueden incluir, pero no se limitan a, enfermedades desmielinizantes, síndromes neurológicos paraneoplásicos, encefalomielitis inmunomediada, neuropatía autonómica inmunomediada, miastenia grave, encefalopatía asociada a autoanticuerpos y encefalomielitis aguda diseminada.
- 10 20 En un ejemplo no limitativo, los Ejemplos siguientes proporcionan un método para diagnosticar la enfermedad de Alzheimer en un paciente usando la matriz de sensores y los métodos descritos en el presente documento. La matriz de sensores no solo pudo distinguir con precisión entre pacientes con o sin enfermedad de Alzheimer, sino que también pudo detectar pacientes que eran presintomáticos y desarrollaron la enfermedad de Alzheimer varios años después de la detección (según lo determinado por plasmas de cohorte). Esto proporciona las ventajas de poder tratar una enfermedad en una etapa muy temprana, incluso antes del desarrollo de la enfermedad.
- 15 25 30 35 Las matrices de sensores y métodos de la presente invención en algunas realizaciones pueden detectar una etapa previa a la enfermedad de una enfermedad o trastorno. Una etapa previa a la enfermedad es una etapa en la que el paciente no ha desarrollado ningún signo o síntoma de la enfermedad. Una etapa precancerosa sería una etapa en la que no se han identificado cáncer, tumor o células cancerosas dentro del sujeto. Una etapa de enfermedad preneurológica sería una etapa en la que una persona no ha desarrollado uno o más síntomas de la enfermedad neurológica. La capacidad de diagnosticar una enfermedad antes de que se presenten uno o más signos o síntomas de la enfermedad permite un seguimiento estrecho del sujeto y la capacidad de tratar la enfermedad en una etapa muy temprana, lo que aumenta la posibilidad de poder detener la progresión o reducir la gravedad de la enfermedad.

- 30 35 40 45 Las matrices de sensores y métodos de la presente invención en algunas realizaciones pueden detectar las primeras etapas de una enfermedad o trastorno. Las primeras etapas de la enfermedad se refieren a cuando los primeros signos o síntomas de una enfermedad pueden manifestarse en un sujeto. Por lo general, las enfermedades que pueden detectarse en el desarrollo previo a la enfermedad o en los primeros estados son más fáciles de tratar y brindan un resultado más positivo para el paciente. Por ejemplo, para el cáncer, las primeras etapas de una enfermedad pueden incluir cáncer en etapa 0 y etapa 1. El cáncer en etapa 0 describe el cáncer *in situ*, lo que significa "en su lugar", lo que significa que el cáncer aún se encuentra en el lugar donde comenzó y no se ha propagado a los tejidos cercanos. Esta etapa del cáncer a menudo es altamente curable, generalmente mediante la extirpación de todo el tumor con cirugía. El cáncer en etapa 1 suele ser un cáncer o tumor pequeño que no ha crecido profundamente en el tejido cercano y no se ha propagado a los ganglios linfáticos u otras partes del cuerpo. Además, la etapa temprana de una enfermedad puede ser una etapa en la que no hay signos o síntomas externos. Por ejemplo, en la enfermedad de Alzheimer, una etapa temprana puede ser una etapa previa a la enfermedad de Alzheimer en la que no se detectan síntomas y, sin embargo, el paciente desarrollará la enfermedad de Alzheimer meses o años después.

- 50 55 60 En algunas realizaciones, las matrices de sensores y los métodos pueden detectar etapas intermedias de la enfermedad. Los estados intermedios de la enfermedad describen etapas de la enfermedad que han superado los primeros signos y síntomas y el paciente experimenta uno o más síntomas de la enfermedad. Por ejemplo, para el cáncer, los cánceres en etapa II o III se consideran etapas intermedias, lo que indica cánceres más grandes o tumores que han crecido más profundamente en el tejido cercano. En algunos casos, los cánceres en etapa II o III también pueden haberse propagado a los ganglios linfáticos, pero no a otras partes del cuerpo.

- Además, las matrices de sensores y los métodos pueden detectar etapas tardías o avanzadas de la enfermedad. Las etapas tardías o avanzadas de la enfermedad también pueden denominarse "graves" o "avanzadas" y generalmente indican que el sujeto sufre múltiples síntomas y efectos de la enfermedad. Por ejemplo, el cáncer en etapa grave incluye la etapa IV, en la que el cáncer se diseminó a otros órganos o partes del cuerpo y, a veces, se lo denomina cáncer avanzado o metastásico.

En algunas realizaciones, los métodos de la presente tecnología incluyen comparar la huella de proteína de la muestra con un panel de huellas de proteína asociadas con una pluralidad de enfermedades y/o una pluralidad de estados de enfermedad para determinar si la muestra indica una enfermedad y/o estado de enfermedad.. Por ejemplo, se pueden recolectar muestras de una población de sujetos a lo largo del tiempo. Una vez que los sujetos desarrollan una enfermedad o trastorno, la presente invención permite la capacidad de caracterizar y detectar los cambios en las huellas de biomoléculas a lo largo del tiempo en el sujeto comparando la huella de biomoléculas de la muestra del mismo sujeto antes de que hayan desarrollado una enfermedad con la huella de biomolécula del sujeto después de haber desarrollado la enfermedad. En algunas realizaciones, las muestras se pueden tomar de cohortes de pacientes que desarrollan la misma enfermedad, lo que permite el análisis y la caracterización de las huellas de biomoléculas que están asociadas con las diferentes etapas de la enfermedad para estos pacientes (por ejemplo, desde estados previos a la enfermedad hasta estados de enfermedad).

Solo con fines ilustrativos, los ejemplos han demostrado que los métodos y conjuntos de sensores de la presente invención pueden distinguir no solo entre diferentes tipos de cáncer, sino también entre las diferentes etapas del cáncer (por ejemplo, etapas tempranas del cáncer).

Se contemplan métodos para determinar la huella de una biomolécula asociada con al menos una enfermedad o trastorno y/o un estado de enfermedad. Los métodos comprenden los pasos de obtener una muestra de al menos dos sujetos diagnosticados con al menos una enfermedad o trastorno o que tienen el mismo estado de enfermedad; poner en contacto cada muestra con una matriz de sensores descrita en el presente documento para determinar una huella de biomolécula para cada matriz sensora, y analizar la huella de al menos dos muestras para determinar una huella de biomolécula asociada con al menos una enfermedad o trastorno y/o estado de enfermedad.

Clasificación de corona de biomolécula

El método para determinar la huella de la biomolécula asociada con la enfermedad o trastorno y/o el estado de enfermedad incluye el análisis de las huellas de la biomolécula de al menos dos muestras. Esta determinación, análisis o clasificación estadística se realiza mediante métodos conocidos en la técnica, que incluyen, pero no se limitan a, una amplia variedad de enfoques de análisis de datos supervisados y no supervisados, aprendizaje automático, aprendizaje profundo y agrupación, incluido el análisis de agrupación jerárquica (HCA), análisis de componentes principales (PCA), análisis discriminante de mínimos cuadrados parciales (PLS-DA), bosque aleatorio, regresión logística, árboles de decisión, máquina de vectores de soporte (SVM), k-vecinos más cercanos, bayes ingenuo, regresión lineal, regresión polinomial, SVM para regresión, agrupamiento de K-medias y modelos ocultos de Markov, entre otros. En otras palabras, las huellas de biomoléculas de cada muestra se comparan/analizan entre sí para determinar con significancia estadística qué patrones son comunes entre las huellas individuales para determinar una huella de biomoléculas que está asociada con la enfermedad o trastorno o estado de enfermedad.

En general, los algoritmos de aprendizaje automático se utilizan para construir modelos que asignan con precisión etiquetas de clase a ejemplos con base en las características de entrada que describen el ejemplo. En algunos casos, puede ser ventajoso emplear enfoques de aprendizaje automático y/o aprendizaje profundo para los métodos descritos en este documento. Por ejemplo, el aprendizaje automático se puede utilizar para asociar la huella de la biomolécula con varios estados de enfermedad (por ejemplo, sin enfermedad, precursor de una enfermedad, etapa temprana o tardía de la enfermedad, etc.). Por ejemplo, en algunos casos, se emplean uno o más algoritmos de aprendizaje automático en relación con un método de la invención para analizar los datos detectados y obtenidos por la corona de biomoléculas y las huellas de biomoléculas derivadas de la misma. Por ejemplo, en una realización, el aprendizaje automático se puede acoplar con la matriz de sensores descrita en el presente documento para determinar no solo si un sujeto tiene una etapa previa de cáncer, cáncer o no tiene o desarrolla cáncer, sino también para distinguir el tipo de cáncer.

Los siguientes Ejemplos han mostrado la capacidad de la matriz de sensores descrita aquí para determinar el estado de enfermedad de una serie de enfermedades diferentes, incluyendo cáncer, enfermedad cardiovascular y enfermedad neurológica (por ejemplo, enfermedad de Alzheimer) con significación estadística.

Este ensayo no se limita a estas realizaciones específicas, ya que la matriz de sensores se puede aplicar a una variedad de enfermedades y estados de enfermedad como se describe en este documento.

En algunas realizaciones, el método incluye la obtención de muestras de sujetos de control que se ponen en contacto con la matriz de sensores para producir una huella de biomolécula de control. Estas huellas de biomoléculas de control se pueden usar luego para compararlas con las huellas de biomoléculas de los sujetos con una enfermedad o trastorno y/o estado de enfermedad específico para determinar una huella de biomolécula específica para esa enfermedad o trastorno y/o estado de enfermedad específico.

El método puede incluir, por ejemplo, obtener una muestra de control de al menos un sujeto de control, poner en contacto la muestra de control con la matriz de sensores para producir una pluralidad de coronas de biomoléculas de control y analizar la pluralidad de biomoléculas de cada corona de biomoléculas de control. El

método puede comprender además comparar la pluralidad de biomoléculas de la pluralidad de xorona de biomoléculas de control con la pluralidad de biomoléculas de la pluralidad de corona de biomoléculas del sujeto con la enfermedad o trastorno para determinar una huella de biomolécula asociada con al menos una enfermedad o trastorno.

- 5 También se contemplan métodos para diagnosticar o pronosticar una enfermedad o trastorno. Los métodos comprenden obtener una muestra de un sujeto; poner en contacto la muestra con una matriz de sensores para producir una huella de biomolécula, y comparar la huella de biomolécula con un panel de huellas de biomolécula asociada con una pluralidad de enfermedades o trastornos; y diagnosticar o pronosticar la enfermedad o trastorno.
- 10 En algunas realizaciones, se contemplan métodos para identificar patrones de biomarcadores o biomarcadores específicos asociados con una enfermedad o trastorno. Los métodos adecuados incluyen, por ejemplo, realizar los métodos descritos anteriormente (por ejemplo, obtener muestras de al menos dos sujetos diagnosticados con la enfermedad o trastorno y al menos dos sujetos de control; poner en contacto cada muestra con la matriz de sensores para producir una huella de biomolécula, y comparar la huella de la biomolécula de los sujetos con la enfermedad o el trastorno con la huella de la biomolécula de los sujetos de control para determinar al menos un patrón y/o biomarcador asociado con la enfermedad o el trastorno. De manera adecuada, el método puede comprender al menos 2 sujetos con enfermedad y al menos dos sujetos de control, alternativamente al menos 5 sujetos con enfermedad y al menos 5 sujetos de control, alternativamente al menos 10 sujetos con enfermedad y al menos 10 sujetos de control, alternativamente al menos 15 sujetos con enfermedad y al menos 15 sujetos de control, alternativamente al menos 20 sujetos con enfermedad y al menos 20 sujetos de control, e incluye cualquier variación intermedia (por ejemplo, sujetos con enfermedad de al menos 2-100 y sujetos de control de al menos 2-100).

25 En algunas realizaciones, las matrices y métodos permiten la determinación de un patrón de biomarcadores asociados con el estado de enfermedad o enfermedad o trastorno o, en algunas realizaciones, biomarcadores específicos que están asociados con la enfermedad o trastorno. No solo se podrán identificar los biomarcadores que pueden estar asociados con un estado de enfermedad, por ejemplo, los biomarcadores enumerados en este documento, sino que también se pueden determinar nuevos biomarcadores o patrones de biomarcadores que pueden estar asociados con un estado de enfermedad o una enfermedad o trastorno. Como se discutió anteriormente, algunos biomarcadores o patrones de biomarcadores para una enfermedad o trastorno específico pueden ser un cambio en una biomolécula asociada con la matriz de sensores de la presente invención y diferir de lo que generalmente se denomina biomarcadores en la técnica, por ejemplo, y aumentar expresión de una biomolécula específica asociada a una enfermedad. Como se discutió anteriormente, puede ser la interacción de una biomolécula, por ejemplo biomolécula X, con otras biomoléculas, por ejemplo biomolécula Y y Z, que da como resultado la capacidad de asociarse con un estado de enfermedad específico y puede no correlacionarse con ningún cambio en la concentración absoluta del biomarcador X en la muestra a lo largo del tiempo o estado de enfermedad. Por lo tanto, una molécula que en el sentido convencional no se consideraría un biomarcador ya que no cambia en concentración absoluta en una muestra desde el estado previo a la enfermedad hasta el estado de enfermedad, puede, en vista de la presente descripción, considerarse una biomolécula ya que sus cambios relativos que son medidos por la matriz de la presente invención están asociados con un estado de enfermedad. En otras palabras, puede ser un aumento o disminución en la interacción de la biomolécula X (debido a las interacciones de X con los elementos sensores y otras biomoléculas en la muestra) con la matriz que proporciona una señal de que un biomarcador está asociado con un estado de enfermedad.

45 Los biomarcadores de cáncer adecuados incluyen, pero no se limitan a, AHSG (α2-HS-Glicoproteína), AKR7A2 (Aflatoxina B1 aldehído reductasa), AKT3 (PKB γ), ASGR1 (ASGPR1), BDNF, BMP1 (BMP-1), BMPER, C9, CA6 (anhidrasa carbónica VI), CAPG (CapG), CDH1 (cadherina-1), CHRDL1 (Tipo Cordina 1), CKB-CKM- (CK-MB), CLIC1 (canal intracelular de cloruro 1), CMA1 (quimasa), CNTN1 (contactina-1), COL18A1 (endostatina), CRP, CTSL2 (catepsina V), DDC (dopa descarboxilasa), EGFR (ERBB1), FGA-FGB-FGG (dímero D), FN1 (fibronectina FN1.4), GHR (receptor de la hormona del crecimiento), GPI (glucosa fosfato isomerasa), HMGB1 (HMG-1), HNRNPAB (hnRNP A/B), HP (haptoglobina, tipo mixto), HSP90AA1 (HSP 90α), HSPA1A (HSP 70), IGFBP2 (IGFBP-2), IGFBP4 (IGFBP-4), IL12B-IL23A (IL-23), ITIH4 (cadena pesada H4 del inhibidor de inter-α-tripsina), KIT (SCF sR), KLK3-SERPINA3 (PSA-ACT), L1CAM (NCAM-L1), LRIG3, MMP12 (MMP-12), MMP7 (MMP-7), NME2 (NDP quinasa B), PA2G4 (proteína de unión a ErbB3 Ebp1), PLA2G7 (LpPLA2/PAFAH), PLAUR (suPAR), PRKACA (PRKA C -α), PRKCB (PKC-β-II), PROK1 (EG-VEGF), PRSS2 (tripsina-2), PTN (pleiotrofina), SERPINA1 (α1-antitripsina), STC1 (estaniocalcina-1), STX1A (sintaxis 1A), TACSTD2 (proteína GA733-1), TFF3 (factor de trébol 3), TGFB1 (βIGH3), TPI1 (triosafosfato isomerasa), TPT1 (fortilina), YWHAG (14-3-3 proteína γ), YWHAH (14-3-3 proteína eta), biomarcadores de cáncer de próstata, por ejemplo, PSA, Pro-PSA, PHI, PCA3, TMPRSS3:ERG, PCMT, MTEN, marcadores de cáncer de mama, por ejemplo, oncogén del receptor 2 del factor de crecimiento epidérmico (HER2), biomarcador de melanoma BRAF, biomarcador de cáncer de pulmón EML4-ALK, A2ML1, BAX, C10orf47, Clorf62, CSDA, EIFC3, EFB, GABARAPL2, GUKI, GZMH, HIST1H3B, HLA-A, HSP90AA1, NRGN, PRDX5, PTMA, RABAC1, RABAGAP1L, RPL22, SAP 18, SEP1, SOX1, EGFR, EGFRVIII, apolipoproteína AI, apolipoproteína CIII, mioglobina, tenascina C, MSH6, claudina-3, claudina-4, caveolina-1, factor de coagulación III, CD9, CD36, CD37, CD53, CD63, CD81, CD136,

CD147, Hsp70, Hsp90, Rab13, Desmocollin-1, EMP-2, CK7, CK20, GCDF15, CD82, Rab-5b, Anexo n V, MFG-E8, HLA-DR, un microARN miR200, MDC, NME-2, KGF, PIGF, Flt-3L, HGF, MCP1, SAT-1, MIP-1-b, GCLM, OPG, TNF RII, VEGF-D, ITAC, MMP-10, GPI, PPP2R4, AKR1B1, Amy1A, MIP-1b, P-cadherina, EPO y similares. Por ejemplo, los biomarcadores para el cáncer de mama incluyen, pero no se limitan a, ER/PR, HER-2/neu y similares. Los biomarcadores para el cáncer colorrectal incluyen, pero no se limitan a, por ejemplo, EGFR, KRAS, UGT1A1 y similares. Los biomarcadores asociados con la leucemia/linfoma incluyen, pero no se limitan a, por ejemplo, antígeno CD20, CD30, FIP1L1-PDGFRalpha, PDGFR, Philadelphia Chromosome (BCR/ABL), PML/RAR alfa, TPMT, UGT1A1 y similares. Los biomarcadores asociados con el cáncer de pulmón incluyen, pero no se limitan a, por ejemplo, ALK, EGFR, KRAS y similares. Los biomarcadores son conocidos en la técnica y se pueden encontrar, por ejemplo, en Bigbee W, Herberman RB. *Tumor markers and immunodiagnosis*. In: Bast RC Jr, Kufe DW, Pollock RE, et al., editors. *Cancer Medicine*. 6th ed. Hamilton, Ontario, Canada: BC Decker Inc., 2003; Andriole G, Crawford E, Grubb R, et al. Mortality results from a randomized prostate-cancer screening trial. *New England Journal of Medicine* 2009; 360(13):1310-1319; Schröder FH, Hugosson J, Roobol MJ, et al. Screening and prostate-cancer mortality in a randomized European study. *New England Journal of Medicine* 2009; 360(13):1320-1328; Bys SS, Partridge E, Black A, et al. Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. *JAMA* 2011; 305(22):2295-2303; Cramer DW, Bast RC Jr, Berg CD, et al. Ovarian cancer biomarker performance in prostate, lung, colorectal, and ovarian cancer screening trial specimens. *Cancer Prevention Research* 2011; 4(3):365-374; Sparano JA, Gray RJ, Makower DF, et al. Prospective validation of a 21-gene expression assay in breast cancer. *New England Journal of Medicine* 2015; Publicado por primera vez en línea el 28 de septiembre de 2015. doi: 10.1056/NEJMoa1510764.

De manera adecuada, estos métodos pueden usarse para determinar biomarcadores asociados con el cáncer. Por ejemplo, en una realización el cáncer es glioblastoma, y en el que el biomarcador se selecciona del grupo que consiste en HABP1, VTNC, CO3, ITIH2, ITIH1, CO7, FHR5, CBPN, ALBU, PLMN, CO4A, PRDX2, VWF, C4BPA, APOB, HBB, CNDP1, CRP, SAA4, APOE, CSCL7 y combinaciones de los mismos. En otra realización, el cáncer es meningioma, y en donde el biomarcador se selecciona del grupo que consiste en FCN3, RET4, HABP2, CBPN y combinaciones de los mismos. En otra realización, el cáncer es cáncer de páncreas y en el que el biomarcador se selecciona del grupo que consiste en KNG1, IC1, CBPB2, TRFE, GELS, CXCL7, HPTR, PGK1, AACT, LUM, APOE, FIBB, APOA2, A1BG, A1AT, LBP, APOA1, H4, FIBG y combinaciones de los mismos. En otra realización, el cáncer es cáncer de pulmón y el biomarcador se selecciona del grupo que consiste en CO0, CRP, SAA4, APOA1, A1AT, GELS y combinaciones de los mismos. En otra realización, el cáncer es mieloma y el biomarcador es ALBU.

Los biomarcadores también pueden estar asociados con la enfermedad cardiovascular que se conocen en la técnica e incluyen, pero no se limitan a, perfil de lípidos, nivel de glucosa y hormonas y biomarcadores fisiológicos basados en la medición de niveles de biomoléculas importantes como ferritina sérica, triglicéridos para Relación HDLp (lipoproteínas de alta densidad), relación lipoforina-colesterol, relación lípido-lipoforina, nivel de colesterol LDL, niveles de HDLp y apolipoproteínas, relación de lipoforinas y LTP, esfingolípidos, índice Omega-3 y nivel de ST2, entre otros. Los biomarcadores adecuados para enfermedades cardiovasculares se pueden encontrar en la técnica, por ejemplo, pero sin limitarse a, en van Holten et al. "Circulating Biomarkers for Predicting Cardiovascular Disease Risk; a Systemic Review and Comprehensive Overview of Meta-Analyses" *PLoS One*, 2013 8(4): e62080.

Los biomarcadores también pueden estar asociados con una enfermedad neurológica. Los biomarcadores adecuados son conocidos en la técnica e incluyen, pero no se limitan a, por ejemplo, A β 1-42, t-tau y p-tau 181, α -sinucleína, entre otros. Véase, por ejemplo, Chintamaneni and Bhaskar "Biomarkers in Alzheimer's Disease: A Review" *ISRN Pharmacol*. 2012. 2012: 984786. Publicado en línea el 28 de junio de 2012.

Los biomarcadores para enfermedades inflamatorias son conocidos en la técnica e incluyen, pero no se limitan a, por ejemplo, citoquinas/quimioquinas, efectores relacionados con el sistema inmunitario, proteínas de fase aguda [proteína C reactiva (PCR) y amiloide A sérico (SAA)], especies reactivas de oxígeno (ROS) y especies reactivas de nitrógeno (RNS), factores relacionados con prostaglandinas y ciclooxygenasa (COX), y mediadores como factores de transcripción y factores de crecimiento, que pueden incluir, por ejemplo, proteína C reactiva (CRP), S100, LIF, CXCL1, CXCL2, CXCL4, CXCL5, CXCL8, CXCL9, CXCL10, CCL2, CCL23, IL- β , IL-IRa, TNF, IL-6, IL-10, IL-17A, IL-17F, IL-21, IL-22, IFN γ , CXCR1, CXCR4, CXCR5, GM-CSF, GM-CSFR, G-CSF, G-CSFR, EGF, VEGFA, LEP, SAA1, VCAM1, CRP, MMP1, MMP3, TNFRSF1A, RETN, CHI3L1, anticuerpos antinucleares (ANA), factor reumatoide (FR), anticuerpos contra el péptido cíclico citrulinado (anti-CCP)] y para la EII crónica (calprotectina fecal), entre otros. Los biomarcadores adecuados para la enfermedad inflamatoria intestinal, por ejemplo, incluyen CRP, ESR, pANCA, ASCA y calprotectina fecal. Véase, por ejemplo, Yi Fengming and Wu Jianbing, "Biomarkers of Inflammatory Bowel Disease," *Disease Markers*, vol. 2014, Article ID 710915, 11 páginas, 2014. doi:10.1155/2014/710915.

Los términos "individuo", "sujeto" y "paciente" se usan indistintamente en el presente documento independientemente de si el sujeto tiene o está experimentando actualmente alguna forma de tratamiento. Como se usa en el presente documento, el término "sujeto" generalmente se refiere a cualquier vertebrado, incluidos, pero no limitados a, los mamíferos. Ejemplos de mamíferos que incluyen primates, incluidos simios

y humanos, equinos (por ejemplo, caballos), caninos (por ejemplo, perros), felinos, varios tipos de ganado domesticado (por ejemplo, ungulados, como cerdos, cerdos, cabras, ovejas y similares), así como mascotas domesticadas (por ejemplo, gatos, hámsteres, ratones y conejillos de indias). Preferiblemente, el sujeto es un ser humano.

5 Las matrices y los métodos descritos en el presente documento se pueden usar en varias condiciones diferentes para proporcionar la huella de la biomolécula deseada. Por ejemplo, el tamaño de los elementos del sensor, la tasa de flujo de la muestra a través del sensor, el tiempo de incubación de la matriz de sensores con la muestra y la temperatura a la que se incuba la matriz de sensores se pueden cambiar para proporcionar una huella de biomolécula reproducible. Los tamaños adecuados del elemento sensor incluyen elementos sensores a nanoescala que tienen menos de una micra en al menos una dirección.

10 El tiempo adecuado para incubar la matriz o la pluralidad de elementos sensores incluye, al menos unos pocos segundos, por ejemplo, al menos 10 segundos a aproximadamente 24 horas, por ejemplo, al menos aproximadamente 10 segundos, al menos aproximadamente 15 segundos, al menos aproximadamente 20 segundos, al menos al menos unos 25 segundos, al menos unos 30 segundos, al menos unos 40 segundos, 15 al menos unos 50 segundos, al menos unos 60 segundos, al menos unos 90 segundos, al menos unos 2 minutos, al menos unos 3 minutos, al menos unos 4 minutos, al menos unos 5 minutos, al menos unos 6 minutos, al menos unos 7 minutos, al menos unos 8 minutos, al menos unos 9 minutos, al menos unos 10 minutos, al menos unos 15 minutos, al menos unos 20 minutos, al menos unos 25 minutos, al menos unos 30 minutos, al menos unos 45 minutos, al menos unos 50 minutos, al menos unos 60 minutos, al menos unos 90 minutos, 20 al menos unas 2 horas, al menos unas 3 horas, al menos al menos unas 4 horas, al menos unas 5 horas, al menos unas 6 horas, al menos unas 7 horas, a 1 al este unas 8 horas, al menos unas 9 horas, al menos unas 10 horas, al menos unas 12 horas, al menos unas 14 horas, al menos unas 15 horas, al menos unas 16 horas, al menos unas 17 horas, al menos unas 18 horas, al menos unas 19 horas, al menos unas 20 horas, e incluir cualquier tiempo e incremento intermedio (por ejemplo, 10 segundos, 11, 12, 13, 14, 15, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60 segundos, etc.; 1 minuto, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, etc.; 1 hora, 2 horas, 3 horas, 4 horas, 5 horas, 6 horas, 7 horas, 8 horas, etc.)

30 Además, un experto en la materia puede determinar la temperatura a la que se realiza el ensayo, e incluye temperaturas entre aproximadamente 4 °C y aproximadamente 40 °C, alternativamente desde aproximadamente 4 °C hasta aproximadamente 20 °C, alternativamente desde aproximadamente 10 °C. °C a unos 15 °C, alternativamente de unos 10 °C a unos 40 °C, por ejemplo, a unos 4 °C, a unos 5 °C, a unos 6 °C, a unos 7 °C, a unos 8 °C, a unos 9 °C, aproximadamente 10 °C, aproximadamente 11 °C, aproximadamente 12 °C, aproximadamente 13 °C, aproximadamente 14 °C, aproximadamente 15 °C, aproximadamente 16 °C, 35 aproximadamente 17 °C, aproximadamente 18 °C, aproximadamente 19 °C, aproximadamente 20 °C, aproximadamente 21 °C, aproximadamente 22 °C, aproximadamente 25 °C, aproximadamente 30 °C, aproximadamente 35 °C, aproximadamente 37 °C, etc. Adecuado, el ensayo puede realizarse a temperatura ambiente temperatura (por ejemplo, aproximadamente aproximadamente 37 °C, por ejemplo, de 40 °C a aproximadamente 35 °C a aproximadamente 40 °C).

40 Los métodos de la presente invención pueden comprender poner en contacto la muestra con la matriz de sensores. El contacto de la muestra con la matriz de sensores se puede realizar utilizando cualquier caudal adecuado en el que la muestra pueda fluir sobre la matriz de sensores. En algunos aspectos, la velocidad de flujo de las corrientes, el número de Reynolds o las áreas de sección transversal relativas de las corrientes de flujo pueden modificarse para proporcionar un contacto adecuado entre la muestra y la matriz de sensores.

45 Por ejemplo, en realizaciones que usan un nanocanal o un microcanal, el área de la sección transversal de la primera corriente puede ser más del 1 % del área de la sección transversal del canal. En otro ejemplo, el área de la sección transversal de la primera corriente puede ser inferior al 90 % del área de la sección transversal del canal. La relación del área de la sección transversal puede ser de 10:1 a 1:10, de 1:5 a 5:1, de 1:3 a 3:1, 50 de 1:2 a 2:1 o de 1:1.

En ciertas otras circunstancias, el flujo de la muestra sobre la matriz tiene un número de Reynolds en el lugar de la introducción de la muestra de entre 300 y 1,000,000. En algunos casos, la ubicación de la introducción de la muestra es un nanocanal o microcanal.

Kits (no reivindicado *per se* pero puede ser útil para poner en práctica la invención)

55 Los aspectos de la presente divulgación que se describen con respecto a los métodos se pueden utilizar en el contexto de la matriz de sensores o kits discutidos en esta divulgación. De manera similar, los aspectos de la presente divulgación que se describen con respecto a la matriz de sensores y los métodos se pueden utilizar en el contexto de los kits, y los aspectos de la presente divulgación que se describen con respecto a los kits se pueden utilizar en el contexto de los métodos y la matriz de sensores.

Esta divulgación proporciona kits. Los kits pueden ser adecuados para su uso en los métodos descritos en este documento. Los kits adecuados incluyen un kit para determinar la huella de una biomolécula para una muestra que comprende una matriz de sensores como se describe en el presente documento. En un aspecto, el kit proporciona una matriz de sensores que comprende al menos dos elementos sensores que tienen propiedades fisicoquímicas diferentes entre sí. En algunos aspectos, los kits proporcionan un panel comparativo de huellas de biomoléculas para usar la huella de biomoléculas para determinar un estado de enfermedad para el sujeto. En algunos aspectos, se incluyen instrucciones sobre cómo determinar la huella de la biomolécula. En algunas realizaciones adecuadas, las matrices de sensores se proporcionan como matrices de chips en el kit.

En otros aspectos, se proporcionan kits para determinar un estado de enfermedad de un sujeto o diagnosticar o pronosticar una enfermedad en un sujeto. Los kits adecuados incluyen una matriz de sensores que comprende al menos dos elementos sensores que tienen diferentes propiedades fisicoquímicas entre sí para determinar la huella de una biomolécula. Además, el kit puede incluir además un panel comparativo de huellas de biomoléculas de diferentes estados patológicos o diferentes enfermedades o trastornos. Se proporcionan instrucciones para determinar la huella y el análisis de la biomolécula.

Debería ser evidente para los expertos en la técnica que son posibles muchas modificaciones adicionales además de las ya descritas sin apartarse de los conceptos inventivos. Al interpretar esta divulgación, todos los términos deben interpretarse de la manera más amplia posible de acuerdo con el contexto. Las variaciones del término "que comprende" deben interpretarse en el sentido de referirse a elementos, componentes o pasos de manera no excluyente, por lo que los elementos, componentes o pasos a los que se hace referencia pueden combinarse con otros elementos, componentes o pasos que no están expresamente referenciado. Las realizaciones a las que se hace referencia como "que comprenden" ciertos elementos también se contemplan como "que consisten esencialmente en" y "que consisten en" esos elementos.

Los siguientes ejemplos no limitantes se incluyen solo con fines ilustrativos y no pretenden limitar el alcance del rango de técnicas y protocolos en los que las composiciones y los métodos de la presente invención pueden encontrar utilidad, como apreciará una persona con experiencia en la técnica y puede implementarse fácilmente.

Ejemplos

Ejemplo 1A. matriz de sensores sin etiquetas para la detección temprana del cáncer

El presente ejemplo proporciona una matriz de sensores sin etiquetas para la detección temprana de varios tipos de cáncer. La matriz de sensores consiste en tres liposomas de reacción cruzada diferentes con varias cargas superficiales (es decir, catiónico (DOPG (1,2-dioleoil-sn-glicero-3-fosfo-(1'-rac-glicerol)), aniónico (DOTAP(1,2-Dioleoil-3-trimetilamonio-propano)-DOPE(dioleoilfosfatidiletanolamina)), y neutro (CHOL (DOPC-Colesterol)), cuya composición de corona proteica cambia en respuesta a sus interacciones con el plasma de pacientes que tienen diferentes tipos de varios cánceres, es decir, pulmón, páncreas, mieloma, meningioma y glioblastoma. Aunque ninguna composición de corona de proteínas es específica para ningún tipo de cáncer, los cambios en el patrón de composición de corona proporcionan una "huella" única para cada tipo de cáncer.

Los perfiles de corona duros de los elementos de la matriz de sensores utilizan plasma de pacientes con cánceres en etapas tempranas, intermedias y avanzadas. La composición de la corona de proteínas que se forma en la superficie de los elementos de la matriz de sensores (nanopartículas) depende en gran medida de las propiedades fisicoquímicas de esas nanopartículas y, al mismo tiempo, puede verse fuertemente afectada por el tipo de enfermedad presente en el donante de la plasma humano utilizado para la incubación. El tamaño y la carga de las nanopartículas recubiertas de corona, después de la incubación con plasma de pacientes con cinco tipos diferentes de cáncer (es decir, glioblastoma multiforme, cáncer de pulmón, meningioma, mieloma múltiple y cáncer de páncreas) e individuos sanos (véase tabla 1), se probaron mediante dispersión de luz dinámica (DLS/Nanosight) y microscopía electrónica de transmisión (TEM), y los resultados demostraron que las propiedades fisicoquímicas de las nanopartículas recubiertas de corona dependían sustancialmente del tipo de cáncer (Figura 2A,B).

Tabla 1. Información general sobre los pacientes (y sus tipos de cáncer) cuyo plasma se utilizó en este estudio.

| ETIQUETA DEL PACIENTE | EDAD | GÉNERO | ETAPA DEL CÁNCER | CATEGORÍA DE ETAPA DEL CÁNCER |
|-----------------------|------|--------|------------------|-------------------------------|
| SANO 1 | 43 | F | ----- - | ----- |
| SANO 2 | 54 | M | ----- - | ----- |
| SANO 3 | 67 | F | ----- - | ----- |

| ETIQUETA DEL PACIENTE | EDAD | GÉNERO | ETAPA DEL CÁNCER | CATEGORÍA DE ETAPA DEL CÁNCER |
|-----------------------|------|--------|---|-------------------------------|
| SANO 4 | 69 | M | ----- - | ----- |
| SANO 5 | 61 | F | ----- - | ----- |
| PÁNCREAS 1 | 81 | M | TNM: cT4 N+M+ | Avanzado |
| PÁNCREAS 2 | 76 | M | TNM: cT2 N+ | Moderado |
| PÁNCREAS 3 | 60 | F | TNM: cT4 N+ | Avanzado |
| PÁNCREAS 4 | 75 | M | TNM: cT3 N+M+ | Avanzado |
| PÁNCREAS 5 | 61 | F | TNM: cT3 N+M+ | Avanzado |
| PÁNCREAS 6 | 71 | F | TNM: cT3 N+ | Avanzado |
| PÁNCREAS 7 | 61 | M | TNM: cT3 N+ | Avanzado |
| PÁNCREAS 8 | 67 | M | TNM: cT3 N+ | Avanzado |
| PULMÓN 1 | 45 | F | TNM: TIA N0 M0 | Temprano |
| PULMÓN 2 | 24 | M | TNM: TIA N0 M0 | Temprano |
| PULMÓN 3 | 44 | F | TNM: TIA N0 M0 | Temprano |
| PULMÓN 4 | 48 | F | TNM: TIA N0 M0 | Temprano |
| PULMÓN 5 | 40 | M | TNM: TIA N0 M0 | Temprano |
| PULMÓN 6 | 47 | F | TNM: TIA N0 M0 | Temprano |
| PULMÓN 7 | 51 | M | TNM: TIA N0 M0 | Temprano |
| PULMÓN 8 | 52 | F | TNM: TIA N0 M0 | Temprano |
| MIELOMA 1 | 41 | F | Inicio II (52% de células plasmáticas en médula ósea; IgG-k monoclonal) | Moderado |
| MIELOMA 2 | 60 | F | Inicio I (28% de células plasmáticas en médula ósea; IgG-k monoclonal) | Temprano |
| MIELOMA 3 | 57 | F | Inicio I (8% de células plasmáticas en médula ósea; IgG-k monoclonal) | Temprano |
| MIELOMA 4 | 63 | M | Inicio I (9% de células plasmáticas en médula ósea; IgG-k monoclonal) | Temprano |
| MIELOMA 5 | 75 | M | Inicio II (42% de células plasmáticas en médula ósea; IgG-k monoclonal) | Moderado |
| MIELOMA 6 | 46 | F | Inicio I (15% de células plasmáticas en médula ósea; IgG-L monoclonal) | Temprano |
| MIELOMA 7 | 56 | F | Inicio I (44% de células plasmáticas en médula ósea; IgG-L monoclonal) | Temprano |
| MIELOMA 8 | 70 | F | Inicio I (27% de células plasmáticas en médula ósea; IgG-L monoclonal) | Temprano |
| GLIOBLASTOMA 1 | 58 | M | WHO 4 | Avanzado |

| ETIQUETA DEL PACIENTE | EDAD | GÉNERO | ETAPA DEL CÁNCER | CATEGORÍA DE ETAPA DEL CÁNCER |
|-----------------------|------|--------|------------------|-------------------------------|
| GLIOBLASTOMA 2 | 75 | M | WHO 4 | Avanzado |
| GLIOBLASTOMA 3 | 76 | M | WHO 4 | Avanzado |
| GLIOBLASTOMA 4 | 73 | M | WHO 4 | Avanzado |
| GLIOBLASTOMA 5 | 76 | F | WHO 4 | Avanzado |
| GLIOBLASTOMA 6 | 62 | M | WHO 4 | Avanzado |
| GLIOBLASTOMA 7 | 48 | M | WHO 4 | Avanzado |
| GLIOBLASTOMA 8 | 56 | M | WHO 4 | Avanzado |
| MENINGIOMA 1 | 82 | F | WHO 2 | Moderado |
| MENINGIOMA 2 | 50 | F | WHO 2 | Moderado |
| MENINGIOMA 3 | 80 | M | WHO 2 | Moderado |
| MENINGIOMA 4 | 67 | M | WHO 2 | Moderado |
| MENINGIOMA 5 | 64 | M | WHO 2 | Moderado |
| MENINGIOMA 6 | 84 | M | WHO 2 | Temprano |
| MENINGIOMA 7 | 58 | F | WHO 2 | Temprano |
| MENINGIOMA 8 | 70 | M | WHO 2 | Temprano |

Se realizó una evaluación cuantitativa de la proteína total adsorbida en las nanopartículas a través del ensayo BCA o NanoOrange, y los resultados mostraron diferencias significativas en las cantidades de proteínas adsorbidas después de la incubación en plasma de pacientes con varios tipos de cáncer (Figura 2B). La evaluación cuantitativa de la proteína total adsorbida en la superficie de los liposomas mostró una fuerte dependencia de la cantidad de proteína del tipo de cáncer (Figura 2B). La composición de la corona de proteínas en la superficie de tres liposomas se evaluó mediante cromatografía líquida-espectrometría de masas (LC-MS/MS) en la que se definió la abundancia de 1,800 proteínas conocidas. La contribución de las proteínas individuales y sus categorías (es decir, complemento, coagulación, fuga tisular, lipoproteínas, fase aguda, inmunoglobulinas y otras proteínas plasmáticas) a la composición de la corona se definieron (Figura 2C; Figura 3A-F). Estos resultados demostraron asociaciones significativas entre la composición proteica y no solo el tipo de cáncer sino también el tipo de elemento sensor (es decir, tipo de nanopartícula).

De acuerdo con un extenso cuerpo de literatura, existen relaciones considerables entre el desarrollo del cáncer y las variaciones en el complemento, la coagulación, la pérdida de tejidos, las lipoproteínas, la fase aguda y las inmunoglobulinas. Por lo tanto, las interacciones de reacción cruzada de estas categorías de proteínas con nanopartículas pueden proporcionar "huellas" únicas para cada tipo de cáncer, lo que puede facilitar la identificación y discriminación del cáncer. En consecuencia, uno esperaría que la matriz de sensores de corona de proteínas adsorbiera de forma cruzada un amplio rango de proteínas involucradas en la inducción y el desarrollo del cáncer que podrían usarse para la identificación y discriminación del cáncer.

Desarrollo de análisis de clasificación supervisado para identificar y discriminar entre los tipos de cáncer utilizando los resultados de la matriz de sensores de corona de proteínas. Con el fin de investigar si las huellas de la corona de proteínas (PCF) de varios elementos sensores podrían utilizarse como biosensores y formar patrones únicos para diferentes enfermedades, hemos aplicado enfoques de clasificación enfocados a los datos proteómicos sobre la composición de la corona de proteínas de tres liposomas (catiónico, aniónico, y neutro).

Los detalles de los métodos se describen en la sección Métodos. Se introduce una puntuación de importancia de variable ponderada en la proyección (VIP) y se aplica para clasificar las variables según el análisis discriminante de mínimos cuadrados parciales (PLS-DA) como un método de proyección lineal. La selección de las variables más relevantes (proteínas) en la construcción del modelo de clasificación puede guiarse por un conjunto de variables clasificadas obtenidas. En este sentido, las variables mejor clasificadas se agregaron al modelo una por una y se monitoreó el error de clasificación del modelo PLS-DA. Se observó que el modelo

de clasificación tiene el error mínimo al usar solo las 69 características principales (Figura 4A). El nuevo espacio de características de 69 dimensiones se usó con éxito para discriminar 30 muestras pertenecientes a seis clases mediante PLS-DA con una alta precisión de clasificación (0.97) usando una validación cruzada de 10 veces y dejar uno fuera (Figura 4A,B). Los parámetros de clasificación se dan en la Tabla 2. La contribución de cada proteína individual seleccionada a la separación de cada grupo de cáncer (VIP) se representa en el eje Y y X, respectivamente, para proporcionar una representación visual de la especificidad relativa de los hallazgos (Figura 4B-F). Las proteínas con puntajes VIP más altos podrían considerarse el mejor conjunto informativo o de diagnóstico para discriminar cada enfermedad de los controles y de entre todas las categorías de cáncer.

5 Tabla 2. Resultados de clasificación obtenidos a partir de dos modelos desarrollados y no lineales para seis grupos de muestras.

| Modelos/parámetros de clasificación | PLS-DA, 10 veces CV, LV=5 20 iteraciones | CPANN (8x8) CV de 10 veces |
|-------------------------------------|--|----------------------------|
| Especificidad (CV) | 1 | 1.00 |
| | 2 | 0.96 |
| | 3 | 1.00 |
| | 4 | 1.00 |
| | 5 | 1.00 |
| | 6 | 1.00 |
| Sensibilidad (CV) | 1 | 1.00 |
| | 2 | 1.00 |
| | 3 | 1.00 |
| | 4 | 1.00 |
| | 5 | 1.00 |
| | 6 | 1.00 |
| error de clase (CV) | 1 | 0.00 |
| | 2 | 0.03 |
| | 3 | 0.00 |
| | 4 | 0.00 |
| | 5 | 0.00 |
| | 6 | 0.00 |

10 1) Control 2) Glioblastoma 3) Meningioma 4) Mieloma 5) Páncreas 6) Cánceres de pulmón

15 Luego se aplicaron PLS-DA y la red neuronal artificial de contrapropagación (CPANN) a las variables seleccionadas y muestras completas como clasificación supervisada, enfoques lineales y no lineales, respectivamente. Cabe señalar que el conjunto de datos primarios con todas las variables antes de la selección de variables tiene una discriminación deficiente y no se pudo separar en seis grupos.

20 Luego, para verificar y analizar adicionalmente los datos, se decidió aprovechar un método de clasificación no lineal. Visualizar el espacio de características puede ayudarnos a comprender las estructuras ocultas y las relaciones topológicas entre los patrones. Para reducir la dimensionalidad del espacio de características mientras se preserva la relación topológica en la estructura de datos, se usó CPANN (una variante supervisada de mapas autoorganizados/SOM) para aprender y predecir la pertenencia a clases de los patrones, produciendo simultáneamente un mapa bidimensional de las neuronas y proporcionar información valiosa (desde un enfoque no lineal) sobre la estructura de datos. Los detalles del CPANN se proporcionan en la sección Métodos. Se verificaron diferentes tamaños para el mapa CPANN utilizando una validación cruzada de 10 veces; Se eligió un mapa que incluye 64 (8 x 8) neuronas debido al error mínimo de clasificación (Figura 9C). Además, la estructura topológica de los datos en el espacio de alta dimensión se refleja en el mapa de asignación producido por CPANN (Figura 5C,D). Teniendo en cuenta la similitud de las neuronas con los vectores de entrada, el mapa se puede dividir en seis zonas distintas relacionadas con diferentes tipos de cáncer y muestras de control. Las muestras con la misma etiqueta de clase se asignan a neuronas cercanas o iguales, lo que significa que las variables seleccionadas brindan información valiosa para discriminar las muestras en el espacio de características. La posición relativa y la orientación de seis zonas en el mapa pueden

aportar información cualitativa sobre las similitudes entre los tipos de cáncer. Para representar el efecto de la selección de variables en la calidad del mapeo, otro CPANN fue entrenado con las 1823 variables, y el mapa resultante muestra que los biomarcadores (variables) seleccionados juegan un papel importante en la discriminación entre tipos de cáncer y su clasificación adecuada (Figura 5C,D).

5 Sobre la base de los resultados obtenidos, tanto los modelos lineales como los no lineales mostraron una alta precisión, deducida de sus valores aceptables de especificidad, sensibilidad y error de clase. De acuerdo con estos hallazgos, la agrupación no supervisada (HCA) basada en 69 marcadores pudo separar fuertemente varios tipos de muestras cancerosas y de control (Figura 5E-F). Como puede verse en la Figura 5, existe una gran similitud entre los grupos de muestras de glioblastoma y meningioma, lo que implica una dificultad en la 10 discriminación, muy probablemente relacionada con patrones de proteómica plasmática similares en estos dos cánceres cerebrales. Estos resultados reflejan el hecho de que las concentraciones plasmáticas de muchas proteínas en la corona difieren considerablemente, no solo entre sujetos con diferentes tipos de cáncer, sino también entre individuos sanos.

15 Para ilustrar la capacidad de los biosensores para el reconocimiento de patrones, se realizó una serie de análisis de los datos obtenidos de nanopartículas individuales. Como era de esperar, el patrón de las huellas específicas del cáncer no se pudo extraer únicamente del PCF de cada nanopartícula. Se descubrió que las técnicas de reconocimiento de patrones aplicadas a la abundancia de proteínas en la corona de proteínas formada en tres liposomas diferentes (catiónicos, aniónicos y neutros) distinguieron correctamente no solo el cáncer de las muestras de control, sino también cada tipo de cáncer bajo consideración de los demás.

20 *Identificación de proteínas con funciones cruciales en la detección y discriminación del cáncer como biomarcadores prometedores para tipos específicos de cáncer.* El uso de biomarcadores tanto antes del diagnóstico de cáncer (en la evaluación del riesgo y el cribado/detección temprana) como después del diagnóstico (en el monitoreo de la terapia, la selección de la terapia adicional y la detección de la recurrencia) produciría importantes beneficios terapéuticos y económicos para la salud. Para comprender la relevancia 25 biológica potencial de las 69 proteínas seleccionadas que discriminan muestras cancerosas, buscamos manualmente en informes publicados previamente en PubMed sobre biomarcadores de proteínas de tipos específicos de cánceres que están sobreregulados o subregulados de acuerdo con las diferentes etapas de la enfermedad. Los datos resultantes se compararon con las proteínas seleccionadas en el modelo para identificar 30 marcadores coincidentes y determinar la relevancia biológica del modelo propuesto. Curiosamente, se notaron cantidades significativas de biomarcadores específicos para cinco grupos de cánceres investigados entre los predictores seleccionados que habían sido informados como biomarcadores de cáncer específicos (Figura 7B).

35 La alta especificidad de los marcadores seleccionados para discriminar entre los cinco grupos de cánceres, que se deriva del enfoque de matriz de sensores de corona de proteínas introducido, demuestra una correlación significativa con el trabajo que se está realizando actualmente en el complejo espacio de la proteómica del cáncer; por lo tanto, esta estrategia no solo proporciona una base para la predicción del cáncer, sino que también convierte esa promesa en realidad. Es de destacar que la discriminación entre diferentes grupos ocurre como resultado de varios predictores (y no biomarcadores individuales) que cambian simultáneamente de manera sistemática, formando patrones únicos para cada tipo específico de cáncer. Sobre la base de esta evidencia, los predictores más informativos seleccionados por el modelo propuesto que aún no se han 40 informado como biomarcadores específicos del cáncer pueden tener un gran potencial como nuevos candidatos a biomarcadores de diagnóstico. Para definir su papel en el desarrollo del cáncer, la variación y la funcionalidad de estos candidatos prometedores en pacientes con cáncer deben monitorearse cuidadosamente. Al centrarse en los patrones únicos derivados de un gran número de sujetos a través de un conjunto de predictores informativos, los investigadores deberían poder predecir los cánceres en diferentes etapas con mayor precisión 45 de lo que es posible con los métodos actuales.

50 *Análisis de datos de cohorte.* Para probar la capacidad de esta innovadora matriz de sensores para la detección muy temprana de cánceres, se utilizó plasma de cohortes de personas sanas a las que se les diagnosticó uno de los cinco tipos de cánceres varios años después de la recolección de plasma. Utilizando las muestras de la cohorte, evaluamos si nuestros modelos propuestos, tanto lineales como no lineales, con 69 predictores seleccionados podrían utilizarse para la predicción del cáncer.

55 Con este fin, se pusieron en el modelo los valores relacionados con 69 variables y se predijo la pertenencia a la clase de cada objeto de la cohorte dados los valores de los parámetros óptimos fijos. Es de destacar que 19 variables (proteínas), de las 69 variables, estaban ausentes en el perfil proteómico de la matriz de sensores de corona de proteínas de muestras de cohortes; por lo que su monto en la matriz de datos de validación fue cero. Curiosamente, tanto los modelos lineales como los no lineales proporcionaron buenas predicciones para las cinco muestras. La distancia observada entre las muestras de entrenamiento y de cohorte (Figura 6A,B), muy probablemente influenciado por los valores cero que se agregaron a la matriz de datos para esas 19 variables ausentes. Como se muestra en la Figura 6C,D, las muestras de la cohorte se colocaron en la neurona correcta relacionada con el glioblastoma en el mapa CPANN. MÉTODOS

5 **Liposomas.** El colesterol (Chol) se adquirió de Sigma Aldrich (St. Louis, MO, EE. UU.). DOPC (dioleoilfosfatidilcolina), DOPE (dioleoilfosfatidiletanolamina), DOPG (1,2-dioleoil-sn-glicero-3-fosfo-(1'-rac-glicerol)) y DOTAP (1,2-dioleoil-3-trimetilamonio-propano) se adquirieron de Avanti Polar Lipids (Alabaster, AL, EE. UU.). Se prepararon liposomas DOPG, DOTAP-DOPE (relación molar 1:1) y DOPC-Chol (relación molar 1:1) disolviendo cantidades apropiadas de lípidos 9:1 (v/v) en cloroformo:metanol. La mezcla de cloroformo:metanol se evaporó mediante evaporación rotatoria. Las películas de lípidos se mantuvieron al vacío durante la noche y se hidrataron con regulador de solución salina de fosfato (PBS) 10 mmol/l (pH 7.4) hasta una concentración final de lípidos de 1 mg/ml. Las suspensiones de liposomas obtenidas se clasificaron por extrusión a través de un filtro de carbonato de policarbonato de 50 nm mediante el Avanti Mini-Extruder (Avanti Polar Lipids, Alabaster, AL).

10 **Recolección, preparación y almacenamiento de plasma humano.** Se recolectó plasma humano (HP) de pacientes sanos y con cáncer diagnosticados con glioblastoma multiforme, cáncer de pulmón, meningioma, mieloma múltiple o cáncer de páncreas. El presente estudio fue aprobado por los Ethical Committees of the Sapienza University of Rome (glioblastoma multiforme, meningioma, mieloma múltiple), la University of Napoli Federico II (cáncer de pulmón) y el University Campus Bio-Medico di Roma (cáncer de páncreas). En resumen, 15 se recolectó sangre por venopunción de sujetos sanos y pacientes con cáncer por medio de un Sistema de Recolección de Sangre BD P100 (Franklin Lakes, NJ, EUA) con tecnología de pulsador que reduce el desperdicio de sangre y minimiza el riesgo de contaminación. Después de la formación del coágulo, las muestras se centrifugaron a 1000 × g durante 5 min para sedimentar las células sanguíneas y se eliminó el sobrenadante. 20 Después de confirmar la ausencia de hemólisis, el plasma recogido de cada donante (1 ml) se dividió en alícuotas de 200 microlitros y se almacenó a -80 °C en tubos Protein LoBind etiquetados hasta su uso. Para el análisis, las alícuotas se descongelaron a 4 °C y luego se dejaron calentar a temperatura ambiente (RT).

25 **Muestras de plasma de cohorte.** Utilizamos plasma humano de personas sanas diagnosticadas con cáncer de cerebro, pulmón y páncreas dentro de los ocho años posteriores a la recolección de plasma. Las muestras de plasma se recolectaron a través del NIH-funded Golestan Cohort Study, financiado por NIH, realizado por el National Cancer Institute (NCI) en los EE. UU., la International Agency for Research on Cancer (IARC) en Francia y la Tehran University of Medical Sciences (TUMS) en Iran. Este estudio involucró la recolección y 30 almacenamiento de plasma de 50,000 sujetos sanos, más de 1,000 de los cuales desarrollaron varios tipos de cáncer en los años siguientes. En este estudio se utilizaron muestras de cinco individuos por cáncer. Estas importantes muestras de plasma nos brindan la oportunidad única de probar la capacidad de nuestra innovadora matriz de sensores de corona de proteínas para la detección temprana de cánceres.

35 **Tamaño y potencial zeta.** Los liposomas desnudos se incubaron con HP (1:1 v/v) durante 1 hora a 37 °C. Posteriormente, las muestras se centrifugaron a 14000 rpm durante 15 minutos a 4 °C para sedimentar los complejos liposoma-HP. El sedimento resultante se lavó tres veces con solución salina regulada con fosfato (PBS) y se resuspendió en agua ultrapura. Para las medidas de tamaño y potencial zeta, se diluyeron 10 µL de 40 cada muestra con 990 µL de agua destilada. Todas las mediciones de tamaño y potencial zeta se realizaron a temperatura ambiente utilizando un Zetasizer Nano ZS90 (Malvern, Reino Unido) equipado con un láser HeNe de 5 mW (longitud de onda λ = 632.8 nm) y un correlador logarítmico digital. Las funciones de autocorrelación de intensidad normalizada se analizaron por el método CONTIN para obtener la distribución del coeficiente de difusión D de las partículas. D se convierte en un radio hidrodinámico efectivo R_H por la ecuación de Stokes-Einstein ($R_H = k_B T / 6\pi\eta D$), donde $k_B T$ es la energía térmica y η es la viscosidad del solvente. La movilidad electroforética de las muestras, u , se midió mediante electroforesis láser Doppler. El potencial zeta se calculó 45 mediante la relación de Smoluchowski (potencial zeta = $u\eta/\epsilon$) donde η y ϵ son la viscosidad y la permitividad de la fase solvente, respectivamente). El tamaño y el potencial zeta de los complejos liposoma-HP se dan como media ± desviación estándar (S.D.) de cinco mediciones independientes.

50 **Ensayo de proteínas.** Las formulaciones de liposomas se incubaron con HP (1:1 v/v) durante 1 hora a 37 °C. Posteriormente, los complejos liposoma-HP se sedimentaron a 15000 × g durante 15 minutos a 4 °C y se lavaron tres veces con PBS. El sedimento lavado se resuspendió en urea 8 mol/l, NH_4CO_3 50mmol/l. Se añadieron 10 microlitros de cada muestra a cinco pocillos de una placa de 96 pocillos. La cuantificación de 55 proteínas se realizó agregando 150 microlitros/pocillo de reactivo Protein Assay (Pierce, Thermo Scientific, Waltham, MA, EE. UU.). El multipocillo se agitó y se incubó a temperatura ambiente durante 5 minutos. La absorbancia se midió con GloMax Discover System (Promega, Madison, WI, EE. UU.) a 660 nm. Los efectos de fondo se corrigieron adecuadamente y la concentración de proteína se calculó usando la curva estándar. Los resultados se dan como media ± S.D. de cinco repeticiones independientes.

60 **Identificación y cuantificación de proteínas.** El procedimiento de incubación se realizó como se describe en otra parte (Label-free quantitative analysis for studying the interactions between nanoparticles and plasma proteins. Analytical and Bioanalytical Chemistry, 2013, 405, 2-3, 635-645, incorporado como referencia en su totalidad). Las formulaciones de liposomas se incubaron con HP (1:1 v/v) durante 1 hora a 37 °C. Las muestras se centrifugaron a 15000 × g durante 15 min para sedimentar los complejos liposoma-HP. El sedimento se lavó tres veces con Tris HCl 10 mmol/l (pH 7.4), NaCl 150 mmol/l y EDTA 1 mmol/l. Después del lavado, el sedimento se secó al aire y se resuspendió en el regulador de digestión. La digestión y la desalinización de péptidos se

llevaron a cabo como se describió anteriormente (Shotgun proteomic analytical approach for studying proteins adsorbed onto liposome surface. Analytical and Bioanalytical Chemistry, 2013, 401, 4, 1195-1202). En resumen, el sedimento se resuspendió en 40 microlitros de urea 8 mol/l, NH4CO3 50 mmol/l y se digirió añadiendo 2 microgramos de tripsina. Los péptidos digeridos se desalinizaron utilizando la columna SPE C18, se reconstituyeron con un volumen adecuado de una solución de ácido fórmico al 0.1 % y se almacenaron a -80 °C hasta el análisis. Los péptidos digeridos se analizaron mediante nanocromatografía líquida de alta resolución (HPLC) acoplada a espectrometría de masas en tandem (MS/MS). El análisis NanoHPLC MS/MS se llevó a cabo utilizando un Dionex Ultimate 3000 (Dionex Corporation Sunnyvale, CA, EE. UU.) directamente conectado a un espectrómetro de masas híbrido de trampa de iones lineal-Orbitrap (Orbitrap LTQ-XL, Thermo Scientific, Bremen, Alemania) mediante un nanoelectroaspersión. fuente de iones Las mezclas de péptidos se enriquecieron en una precolumna Acclaim PepMap 100 C18 de 300 µm DI × 5 mm (Dionex Corporation Sunnyvale, CA, EE. UU.), empleando una fase móvil premezclada compuesta por ddH2O/ACN, 98/2 (v/v) que contenía (v/v) HCOOH, a un caudal de 10 microlitros/min. A continuación, las mezclas de péptidos se separaron mediante cromatografía de fase inversa (RP). El mayor conjunto de péptidos se detectó utilizando un gradiente de LC optimizado de 3 horas compuesto por una fase móvil A de ddH2O/HCOOH (99.9/0.1, v/v) y una fase móvil B de ACN/HCOOH (99.9/0.1, v/v). Los espectros de MS de los péptidos eluidos se recogieron en un rango m/z de 350-1700 usando un ajuste de resolución de 60 000 (ancho completo a la mitad del máximo en m/z 400), operando en el modo dependiente de datos. Se recogieron espectros de MS/MS para los cinco iones más abundantes en cada cribado de MS. Se pueden encontrar detalles adicionales en otro lugar ((Shotgun proteomic analytical approach for studying proteins adsorbed onto liposome surface. Analytical and Bioanalytical Chemistry, 2013, 401, 4, 1195-1202). Para cada condición experimental se prepararon tres muestras independientes (réplicas biológicas), cada una de las cuales se midió por triplicado (réplicas técnicas), lo que arrojó nueve mediciones para cada condición experimental. Los archivos de datos RAW se enviaron a Mascot (v2.3, Matrix Science, Londres, Reino Unido) utilizando el filtro de importación de datos de archivos RAW LCQ/DECA de Thermo-Finnigan para realizar búsquedas contra la base de datos Swiss-Prot no redundante (09-2014, 546000 secuencias, restricción taxonómica de Homo Sapiens). Para la búsqueda en la base de datos, la tripsina se especificó como la enzima proteolítica con un máximo de dos escisiones perdidas. La carbamidometilación se fijó como modificación fija de la cisteína, mientras que la oxidación de la metionina se eligió como modificación variable. La tolerancia a la masa monoisotópica para los iones precursores y los iones de fragmentación se estableció en 10 ppm y 0.8 Da, respectivamente. Se seleccionaron estados de carga de +2 o +3 como iones precursores. Los archivos de salida de proteoma se enviaron al software comercial Scaffold (v3.6, Proteome Software, Portland, Oregon, EE. UU.). Las identificaciones de péptidos se validaron si superaban un umbral de probabilidad del 95 % establecido por el algoritmo PeptideProphet. Se aceptaron las identificaciones de proteínas si podían establecerse con una probabilidad superior al 99.0 % y contenían al menos dos péptidos únicos. Se agruparon las proteínas que contenían péptidos compartidos y que no se pudieron diferenciar sobre la base del análisis MS/MS solo para satisfacer los principios de parsimonia. Se utilizaron recuentos de espectro no ponderados (USC) para evaluar la coherencia de las réplicas biológicas en el análisis cuantitativo, y se utilizaron recuentos de espectro normalizados (NCS) para recuperar la abundancia de proteínas.

40 **Análisis estadístico.** Todos los análisis estadísticos se realizaron con las cajas de herramientas PLS, Kohonen y CPANN, y los gráficos se crearon con Microsoft Excel, XLSTAT y MATLAB.

45 **Matriz de datos.** Cada materia prima de la matriz predictora (X) relacionada con cada individuo se deriva de la abundancia de todas las proteínas obtenidas de la matriz de sensores de corona de tres proteínas (Figura 4A). En el paso de preprocessamiento, los datos normalizados en la matriz X, la abundancia relativa de proteínas (RPA), se escalaron automáticamente.

Clasificación y agrupamiento.

50 **Análisis discriminante por mínimos cuadrados parciales (PLS-DA).** El análisis discriminante de mínimos cuadrados parciales es un enfoque multivariante bien conocido considerado como un método de clasificación lineal y reducción de dimensiones que consiste en dos partes principales, una parte estructural, que busca variables latentes como combinaciones lineales de variables independientes originales (es decir, matriz de datos X), que tienen la covariación máxima con las variables dependientes correspondientes (es decir, pertenencia a una clase, Y). Los componentes medidos incluyen las variables latentes como puntajes y cargas, que muestran cómo se relacionan las variables latentes y sus originales. Basado en la capacidad de PLSDA para reducir la dimensionalidad de los datos, permite un mapeo lineal y una visualización gráfica de los diferentes patrones de datos. PLS-DA es particularmente adecuado para tratar con patrones altamente colineales y ruidosos. Los principales problemas asociados con el gran conjunto de datos en proteómica son la gran cantidad de variables monitoreadas (es decir, proteínas) y la cantidad relativamente pequeña de muestras. Por lo tanto, puede haber una gran redundancia entre las variables, lo que hace que muchas de ellas sean poco informativas e irrelevantes para la clasificación. De esta forma, eliminar variables no informativas o encontrar nuevas no correlacionadas puede mejorar el rendimiento predictivo de la clasificación. Dado que en aplicaciones biomédicas como el presente trabajo, no solo debemos tomar decisiones sobre si una muestra pertenece a uno de varios grupos conocidos, sino también determinar qué variables son más relevantes para la mejor discriminación entre clases, un método como PLS- DA es un buen enfoque candidato para encontrar

nuevas variables latentes no correlacionadas mientras se preserva la variación de los datos. El poder de impotencia de las variables originales para producir proyecciones latentes también se puede calcular mediante la importancia de la variable en el análisis de proyección (VIP) y puede jugar un papel significativo en la decisión sobre las variables.

- 5 *Identificación de las variables más relevantes cpn base en el VIP ponderado.* Se utilizó el análisis discriminante de mínimos cuadrados parciales (PLS-DA) para explorar los valores de VIP asociados con las variables. VIP es una medida combinada de cuánto contribuye una variable a una descripción de los dos conjuntos de datos: la variable dependiente (Y) y la variable independiente (X). Los pesos en un modelo PLS reflejan la covarianza entre las variables independientes y dependientes, y la inclusión de los pesos permite que VIP refleje no solo 10 qué tan bien se describe la variable dependiente, sino también qué tan importante es esa información para el modelo de las variables independientes.

Se desarrolló un enfoque basado en la puntuación VIP para identificar el mejor subconjunto de variables. Las puntuaciones VIP se pueden calcular realizando PLS-DA en el conjunto de datos. En ese enfoque, las puntuaciones VIP de las variables se calculan 50 veces, cada vez usando una permutación aleatoria de 15 conjuntos de entrenamiento y validación (los conjuntos de entrenamiento aleatorios se seleccionaron iterativamente considerando una cobertura del 80 por ciento de cada clase de objetos). Teniendo en cuenta las variables más importantes, los valores grandes de puntuación VIP (>2), las 200 variables principales se pueden seleccionar en cada repetición y agregar al grupo de variables principales. Posteriormente, una frecuencia de 20 ocurrencia ($Freq_i$) y un puntaje VIP promedio (VIP_i) para cada variable se puede obtener de acuerdo con el grupo de variables principales. Así, la selección de variables i (que tiene un alto valor VIP_i y bajo valor VIP_i) es menos recomendable debido a la dependencia de los conjuntos de entrenamiento y validación. Por lo tanto, el valor VIP_i de cada variable puede ser ponderado por $Freq_i$, y la clasificación de las variables más relevantes se 25 puede hacer utilizando el VIP_i ponderado. Fig_appr es un diagrama esquemático del enfoque propuesto. La selección de las variables más relevantes para construir el modelo de clasificación puede guiarse por el ranking obtenido de la siguiente manera: Las variables altamente clasificadas se agregaron una por una al conjunto de 30 datos y se calculó el error de clasificación de PLS-DA para encontrar el número mínimo de predictores relevantes (Figura 4A).

Red Neuronal Artificial de Contrapropagación (CPANN). La red neuronal artificial de contrapropagación (CPANN) es una variante supervisada del mapa de autoorganización que consiste en dos capas de neuronas 35 dispuestas en una cuadrícula NxN predefinida. CPANN se puede usar para mapear datos de un espacio de características de alta dimensión a un espacio discreto de neuronas de baja dimensión (típicamente 2), así como para predecir la pertenencia a la clase de las muestras desconocidas. Los vectores de entrada (vector de características de muestra) y los vectores de membresía de clase correspondientes (un vector binario) se presentan a la capa de entrada y salida de CPANN, respectivamente. La corrección del peso de las neuronas 40 en ambas capas se realiza con base en el aprendizaje competitivo y la cooperación de las neuronas (véase Tabla 3). Por lo tanto, se pueden mapear vectores de entrada similares en las mismas neuronas o en neuronas adyacentes y viceversa. El mapa de asignación final revela adecuadamente la estructura de los datos en el espacio de características y preserva la distancia de los patrones en la red de neuronas de baja dimensión. La Figura 9C muestra un mapa de asignación de alta calidad de CPANN utilizando biomarcadores de primera 45 categoría. De acuerdo con las distintas regiones para cada clase, se minimiza el riesgo de error de clasificación. El tamaño adecuado del mapa se puede decidir realizando una validación cruzada de 10 veces en diferentes tamaños de mapa. El CPANN capacitado se puede utilizar para asignar una pertenencia a una clase a una muestra sin etiquetar. La presencia de variables redundantes y no informativas en los datos de entrenamiento afectará la calidad del mapa y aumentará el riesgo de un error de clasificación (Figura 5C). El proceso es un mapeo no lineal, que ayuda a visualizar un objeto de entrada de alta dimensión en una cuadrícula de neuronas bidimensional. Es un procedimiento autogestionado y resuelve el tema de la clasificación de forma transparente. Se pueden encontrar más detalles sobre el método CPANN en las siguientes referencias.

Análisis de agrupamiento jerárquico (HCA). El análisis de agrupamiento jerárquico es un método no supervisado ampliamente utilizado para explorar y visualizar grandes conjuntos de datos heterogéneos como 50 los que se usan a menudo en proteómica en grupos distintos y homogéneos. Las estrategias para el agrupamiento jerárquico se pueden dividir en dos categorías: métodos aglomerativos y divisivos. El procedimiento de aglomeración primero separa cada objeto en su propio grupo individual y luego combina los grupos secuencialmente; los objetos o grupos similares se fusionan hasta que cada objeto pertenece a un solo grupo. El procedimiento de división, por el contrario, comienza con todos los objetos en un grupo grande y los 55 divide gradualmente en grupos más pequeños hasta que cada objeto está en un grupo individual. Finalmente, los objetos se organizan en un dendograma cuyas ramificaciones son las agrupaciones definidas. En el análisis de agrupamientos, para identificar subgrupos homogéneos, se deben definir los dos conceptos importantes, similitud (determinar un valor numérico para la similitud entre objetos y construir una matriz de similitud) y vinculación (conexión de un objeto a un grupo o no). Aquí, aplicamos el agrupamiento jerárquico aglomerativo 60 con el algoritmo de vinculación del vecino más lejano para el análisis no supervisado basado en las variables seleccionadas.

5 *Predicción de muestras de cohortes.* La capacidad predictiva de variables lineales y no lineales, con 69 predictores seleccionados, se evaluó mediante análisis de muestras de cohortes. Con este fin, se pusieron en el modelo los valores relacionados con 69 variables, y se predijo la pertenencia a la clase de cada objeto de la cohorte en los valores de parámetros óptimos fijos. Las 19 variables (proteínas) en el perfil proteómico de

10 nuestra matriz de sensores de corona de proteínas de muestras de cohortes no se detectaron, y se consideraron cero para estas variables en la matriz de datos de validación. Curiosamente, tanto los modelos lineales como los no lineales proporcionan buenas predicciones para las cinco muestras. La distancia entre las muestras de entrenamiento y de cohorte que se muestra en la Figura 6A probablemente esté relacionado con los valores cero que se agregaron a la matriz de datos para esas 19 variables ausentes. Examinamos el efecto de los valores cero reemplazados para las variables ausentes al eliminar estas 19 proteínas de ambas matrices de datos de los conjuntos de entrenamiento y predicción; luego se construyó el modelo PLS-DA y se predijeron muestras de cohortes. Como era de esperar, no se observó una gran distancia entre la cohorte y las muestras de entrenamiento.

Tabla 3. Coeficiente de correlación del mapa de pesos de CPANN para cada variable en las seis clases.

| ID de biomarcador | CC(1) Control | CC(2) Glioblastoma | CC(3) meningioma | CC(4) mieloma | CC(5) Páncreas | CC(6) Pulmón |
|-------------------|---------------|--------------------|------------------|---------------|----------------|--------------|
| 1323 | -0.16 | 0.04 | 0.80 | -0.41 | 0.16 | -0.51 |
| 1282 | -0.47 | -0.04 | 0.11 | 0.90 | -0.03 | -0.37 |
| 1518 | -0.22 | 0.76 | -0.12 | -0.12 | -0.15 | 0.05 |
| 742 | 0.38 | -0.06 | -0.21 | -0.72 | -0.03 | 0.58 |
| 1462 | -0.14 | -0.25 | -0.17 | -0.20 | -0.12 | 0.89 |
| 1373 | -0.30 | -0.13 | -0.10 | 0.98 | -0.21 | -0.16 |
| 1466 | -0.37 | 0.03 | 0.84 | -0.31 | 0.14 | -0.37 |
| 1313 | -0.28 | 0.03 | 0.59 | -0.05 | 0.37 | -0.68 |
| 1630 | 0.97 | -0.23 | -0.31 | -0.26 | -0.19 | -0.19 |
| 815 | -0.33 | -0.15 | -0.24 | 0.65 | -0.45 | 0.61 |
| 882 | -0.16 | -0.16 | -0.25 | -0.13 | -0.16 | 0.91 |
| 1284 | -0.47 | -0.01 | 0.42 | 0.74 | -0.34 | -0.28 |
| 891 | -0.37 | -0.05 | -0.33 | 0.55 | -0.34 | 0.68 |
| 954 | -0.19 | -0.10 | -0.25 | -0.20 | 0.99 | -0.20 |
| 782 | -0.37 | -0.32 | -0.39 | 0.66 | 0.02 | 0.50 |
| 820 | -0.34 | -0.24 | -0.43 | 0.36 | 0.02 | 0.72 |
| 1556 | -0.18 | -0.09 | -0.25 | -0.19 | 0.98 | -0.20 |
| 1286 | 0.41 | -0.35 | 0.18 | -0.31 | -0.51 | 0.40 |
| 868 | 0.15 | -0.28 | -0.44 | 0.16 | -0.32 | 0.72 |
| 1308 | 0.42 | 0.26 | -0.38 | -0.45 | -0.31 | 0.46 |
| 881 | -0.12 | -0.09 | -0.30 | -0.18 | 0.98 | -0.24 |
| 848 | -0.47 | -0.04 | -0.07 | 0.72 | -0.49 | 0.47 |
| 1267 | 0.56 | -0.01 | -0.01 | -0.61 | -0.44 | 0.39 |
| 950 | -0.18 | -0.13 | -0.26 | -0.14 | 0.96 | -0.20 |
| 1561 | 0.95 | -0.24 | -0.35 | -0.28 | -0.10 | -0.17 |
| 856 | -0.33 | -0.18 | -0.21 | 0.64 | -0.43 | 0.60 |
| 895 | -0.03 | -0.16 | -0.32 | -0.20 | -0.13 | 0.87 |
| 1270 | -0.60 | 0.65 | 0.57 | -0.08 | -0.02 | -0.37 |
| 1476 | 0.42 | -0.11 | -0.39 | -0.32 | -0.32 | 0.65 |
| 1416 | 0.59 | -0.05 | -0.10 | -0.61 | -0.36 | 0.40 |

ES 2 993 503 T3

| ID de biomarcador | CC(1) Control | CC(2) Glioblastoma | CC(3) meningioma | CC(4) mieloma | CC(5) Páncreas | CC(6) Pulmón |
|-------------------|---------------|--------------------|------------------|---------------|----------------|--------------|
| 786 | -0.15 | -0.19 | -0.32 | 0.02 | 0.94 | -0.25 |
| 751 | -0.38 | -0.23 | -0.36 | 0.33 | 0.05 | 0.70 |
| 1316 | 0.81 | -0.04 | -0.43 | -0.35 | -0.22 | 0.11 |
| 1351 | 0.87 | -0.35 | -0.21 | -0.05 | -0.36 | -0.12 |
| 992 | 0.90 | -0.24 | -0.29 | -0.21 | -0.17 | -0.18 |
| 1272 | -0.45 | 0.37 | 0.74 | -0.17 | 0.07 | -0.51 |
| 1363 | -0.63 | 0.12 | 0.81 | -0.12 | 0.05 | -0.19 |
| 1486 | -0.33 | 0.66 | 0.07 | -0.18 | -0.26 | 0.22 |
| 1496 | -0.18 | -0.08 | -0.24 | -0.18 | 0.93 | -0.19 |
| 96 | -0.23 | -0.13 | -0.22 | -0.18 | 0.92 | -0.11 |
| 1294 | -0.23 | -0.08 | 0.01 | 0.78 | -0.08 | -0.37 |
| 1027 | -0.18 | 0.69 | -0.02 | -0.26 | -0.05 | -0.02 |
| 1007 | -0.29 | -0.11 | 0.86 | -0.08 | -0.23 | -0.23 |
| 1362 | -0.21 | -0.11 | -0.05 | -0.23 | 0.87 | -0.25 |
| 1302 | 0.29 | -0.32 | -0.37 | 0.75 | -0.20 | -0.19 |
| 802 | -0.51 | -0.22 | -0.15 | 0.55 | 0.57 | -0.13 |
| 1265 | -0.37 | 0.33 | 0.67 | -0.17 | 0.16 | -0.59 |
| 1489 | -0.45 | 0.10 | 0.78 | -0.20 | 0.03 | -0.26 |
| 878 | -0.09 | -0.26 | -0.32 | 0.85 | -0.06 | -0.08 |
| 1338 | -0.45 | 0.82 | 0.12 | -0.07 | -0.05 | -0.16 |
| 826 | 0.18 | 0.40 | 0.38 | -0.55 | 0.02 | -0.47 |
| 1424 | -0.19 | -0.21 | -0.32 | -0.09 | 0.09 | 0.77 |
| 355 | -0.20 | -0.10 | -0.27 | -0.15 | 0.93 | -0.16 |
| 879 | -0.20 | -0.03 | -0.19 | -0.21 | 0.89 | -0.20 |
| 1305 | -0.57 | 0.14 | 0.78 | 0.14 | -0.03 | -0.43 |
| 978 | -0.12 | -0.15 | -0.25 | -0.09 | -0.17 | 0.82 |
| 769 | -0.01 | -0.36 | -0.58 | 0.42 | 0.01 | 0.56 |
| 1567 | -0.16 | 0.58 | -0.15 | -0.16 | 0.18 | -0.14 |
| 35 | 0.60 | -0.31 | -0.39 | -0.37 | -0.13 | 0.46 |
| 1632 | -0.15 | -0.12 | 0.64 | -0.10 | -0.16 | -0.19 |
| 1431 | 0.53 | -0.35 | 0.42 | -0.44 | -0.24 | -0.16 |
| 788 | -0.45 | 0.11 | -0.13 | 0.87 | -0.22 | -0.03 |
| 1333 | -0.40 | -0.10 | -0.32 | 0.03 | 0.69 | 0.22 |
| 1357 | 0.47 | 0.01 | -0.56 | -0.14 | -0.30 | 0.50 |
| 997 | -0.16 | -0.13 | -0.22 | -0.12 | -0.16 | 0.83 |
| 1325 | 0.68 | -0.19 | 0.04 | -0.41 | -0.44 | 0.13 |
| 1275 | -0.45 | 0.50 | -0.08 | 0.52 | 0.05 | -0.34 |
| 1315 | -0.15 | 0.34 | 0.66 | -0.62 | 0.08 | -0.33 |
| 1545 | -0.12 | -0.05 | -0.26 | -0.15 | -0.21 | 0.85 |

Para cada variable, se puede calcular el coeficiente de correlación del mapa de peso correspondiente con el patrón del mapa de asignación. $CC = 0$: indica que no hay correlación entre el biomarcador y la clase relacionada; $1 > CC(i) > 0$: concordancia entre la intensidad del biomarcador y la clase relacionada con el cáncer. $0 < CC < -1$: una correlación inversa entre el valor del biomarcador y la clase relacionada con el cáncer. Los valores de $CC > 0.5$ o < -0.5 están coloreados. Por ejemplo, el mapa de peso del biomarcador 1282 está altamente correlacionado con el patrón de cáncer de clase 4 en el mapa de asignación y puede ser un biomarcador importante para las muestras de pacientes con mieloma.

Ejemplo 1B. El análisis en profundidad del proteoma humano utilizando la caracterización de la corona de proteínas de múltiples nanopartículas y el aprendizaje automático permiten una identificación y discriminación precisas de los cánceres en etapas tempranas

En una segunda realización, los datos colectivos de corona de proteínas para una muestra de plasma dada, derivados de los perfiles de corona de proteínas de nanopartículas individuales, pueden identificar y discriminar diferentes tipos de cánceres utilizando el aprendizaje automático (por ejemplo, enfoque de bosque aleatorio) para analizar los datos. Se usó la matriz de sensores como se describe en el Ejemplo 1A y los datos recopilados se analizaron más detalladamente mediante aprendizaje automático. Esta matriz de sensores (es decir, el nanosistema de corona de proteínas) identificó el cáncer de manera inequívoca y sólida y permite la discriminación entre diferentes tipos de cáncer. Este sistema se puede utilizar para predecir tipos de cáncer utilizando muestras de plasma ciegas. La capacidad de detección muy temprana mediante el uso de plasma de personas sanas en cohortes existentes que fueron diagnosticadas con cánceres varios años después de la recolección de plasma para determinar una huella de biomolécula precancerosa.

La Figura 1 presenta una descripción general esquemática del uso de la matriz de sensores (sistema de nanopartículas) para usar la proteómica multiproteína-corona para la detección del cáncer en pacientes con cáncer conocido y para la predicción del cáncer usando plasma ciego y muestras de cohortes.

Resultados

Perfiles de corona de proteínas del nanosistema utilizando plasma de pacientes con cánceres en estadios tempranos, intermedios y avanzados

Nuestro nanosistema de corona de proteína consiste en tres liposomas de reacción cruzada diferentes con diversas cargas superficiales [aniónico (DOPG (1,2-dioleoil-sn-glicero-3-fosfo-(1'-rac-glicerol))), catiónico (DOTAP (1,2-dioleoil-3-trimetilamonio-propano)-DOPE (dioleoilfosfatidiletanolamina)), y neutro (CHOL (DOPC-colesterol))], cuyos perfiles de corona de proteínas se midieron después de la exposición al plasma de sujetos sanos o pacientes con uno de cinco tipos de cáncer: pulmón, páncreas, mieloma, meningioma o glioblastoma.

Se realizaron análisis proteómicos en cada paciente (5 pacientes/grupo) por triplicado para cada uno de los tres liposomas (véase Figura 2A,B para obtener detalles sobre el tamaño y la carga de los liposomas) en nuestro nanosistema de corona de proteína (es decir, ensayos totales: $3 \times (29) \times 3 = 261$). Aunque ninguna composición de corona de proteína única es específica para ningún tipo de cáncer como también se describe en el Ejemplo 1A, la composición de corona de proteína colectiva para una muestra de plasma dada derivada de diferentes liposomas proporcionó una "huella" única para cada tipo de cáncer.

Aunque la composición de la corona de proteínas que se forma en la superficie de los liposomas depende en gran medida de sus propiedades fisicoquímicas, también puede verse muy afectada por el tipo, la concentración y la conformación únicos de las proteínas y otras biomoléculas presentes en el plasma del paciente determinado. El tamaño y la carga de los liposomas recubiertos de corona se probaron mediante dispersión de luz dinámica (DLS/Nanosight), después de la incubación con plasma de pacientes con cinco tipos diferentes de cáncer (véase Tabla 1) e individuos sanos. Los resultados confirmaron que las propiedades fisicoquímicas de las nanopartículas recubiertas de corona variaban según los diferentes tipos de cáncer (Figura 2A,B).

Se realizó una evaluación cuantitativa de la proteína total adsorbida en los liposomas. a través de los ensayos BCA (ácido bicinconílico) o NanoOrange, y los resultados confirmaron diferencias significativas en las cantidades de proteínas adsorbidas después de la incubación en plasma de pacientes con diversos tipos de cáncer (Figura 3A-F). La evaluación cuantitativa de la proteína total adsorbida en la superficie de los liposomas mostró la dependencia de la cantidad de proteína del tipo de cáncer (Figura 3A-F). La composición de la corona de proteínas en la superficie de tres liposomas se evaluó mediante cromatografía líquida-espectrometría de masas en tandem (LC-MS/MS) en la que se definieron 876 proteínas conocidas. Se definieron las contribuciones de las proteínas individuales y sus categorías (es decir, complemento, coagulación, pérdida tisular, lipoproteínas, fase aguda, inmunoglobulinas y otras proteínas plasmáticas) a la composición de la corona (Figura 3A-F). Los resultados demostraron asociaciones entre la composición proteica y no solo el tipo de cáncer sino también el tipo de liposomas. El mecanismo detrás de esta variación es que la alta relación superficie-volumen de las nanopartículas brinda una oportunidad única para que un amplio rango de proteínas plasmáticas humanas (proteínas muy abundantes, proteínas plasmáticas menos abundantes y proteínas muy raras) participen en la corona, sin necesidad de agotamiento de proteínas muy abundantes y sin una correlación

directa con las concentraciones de proteínas plasmáticas. Además, en nuestro sistema, los cambios conformacionales en las proteínas plasmáticas también pueden cambiar la composición de la corona de proteínas, alterando significativamente el sitio de interacción de las proteínas con las nanopartículas. Estas características hacen que el perfil de corona de proteínas sea único entre otros enfoques desarrollados para el análisis de proteínas en plasma humano.

Las composiciones de la corona de proteínas en la superficie de los liposomas contenían un amplio rango de proteínas plasmáticas humanas, incluidas proteínas muy abundantes (por ejemplo, albúmina, transferrina, proteínas del complemento, apolipoproteínas y alfa-2-macroglobulina) y proteínas muy raras (definidas como <100 ng /ml) como la proteína de la transcripción 1 inducida por el factor de crecimiento transformante beta-1 (~10 ng/ml), fructosa-bifosfato aldolasa A (~20 ng/ml), tiorredoxina (~18 ng/ml) y L-selectina (~92ng/ml). También se identificaron 388 proteínas sin concentraciones plasmáticas previas conocidas (<http://plasmaproteomedatabase.org/>). Luego, la información de proteínas obtenida se analizó mediante un enfoque de aprendizaje automático para probar la capacidad de nuestro nanosistema de corona de proteínas para la detección sólida y precisa del cáncer.

Desarrollo de clasificador para detectar y discriminar entre cánceres utilizando resultados de nanosistemas de corona de proteínas.

Para evaluar la capacidad del nanosistema de corona de proteína para detectar diversos tipos de cáncer, se utilizaron datos proteómicos recopilados de los 3 liposomas distintos en 29 muestras de plasma (5 pacientes de cada uno de los 5 tipos de cáncer y 4 muestras sanas) para entrenar un clasificador, específicamente un algoritmo, que recibe mediciones de matriz de un paciente y genera una de seis etiquetas (ya sea uno de los cinco tipos de cáncer o saludable). Antes de entrenar al clasificador, los datos sin procesar de las 3 nanopartículas primero se 'eliminaron el ruido' a través de una factorización de tensor de bajo rango que se analiza en profundidad en una sección posterior. Esta eliminación de ruido mitiga implícitamente la importante variabilidad observada en elementos de corona individuales. Luego entrenamos un clasificador de bosque aleatorio, un popular algoritmo de clasificación no lineal, en los datos sin ruido resultantes.

Se probó la precisión de este clasificador en 16 muestras ciegas (3 pacientes de cada uno de 5 tipos de cáncer y 1 muestra sana;-). Se midió la precisión general de la clasificación para la tarea de asignar correctamente estas muestras ciegas a una de las seis etiquetas, junto con la sensibilidad y la especificidad para cada uno de los cinco tipos de cáncer por separado (Tabla 4). Debido al número relativamente pequeño de muestras de plasma (45 en total) y para garantizar la solidez de estos resultados, se realizó este procedimiento (entrenando un clasificador de bosque aleatorio en 29 muestras y midiendo la precisión, sensibilidad y especificidad en las 16 muestras restantes) un total de 1000 veces. Los conjuntos de entrenamiento y pruebas en cada una de las 1000 réplicas del experimento se eligieron al azar de entre todas las particiones de datos estratificadas por clases. Este enfoque nos permite calcular estimaciones no sesgadas del valor *p* para las precisiones de clasificación que se informaron (véase Tabla 4). Específicamente, la última fila de la Tabla 4 muestra que la precisión general promedio en las 1000 repeticiones fue del 96.2 %, o lo que es lo mismo, un error general del 3.8 %. También se observó un valor de *p* de 0.04 para la hipótesis nula de que el error de clasificación general es inferior al 87.5 % y, por lo tanto, se puede rechazar esta hipótesis nula con un 95 % de confianza. Las sensibilidades para los cinco tipos de cáncer individuales oscilan entre el 87.4 % y el 100.0 %, y las especificidades individuales oscilan entre el 97.0 % y el 100.0 %.

Tabla 4. Precisión, sensibilidad y especificidad de la clasificación general. Precisión de clasificación general para el nanosistema de corona de proteína con uno, dos y tres liposomas (Columna 2). Tanto la precisión de la clasificación como los valores *p* asociados mejoran con liposomas adicionales. La sensibilidad y la especificidad individuales para los tipos de glioblastoma, pulmón, meningioma, mieloma y páncreas (índices 3 a 12) también muestran que la sensibilidad y la especificidad mejoran con liposomas adicionales. Los resultados experimentales se promediaron sobre 1000 extracciones independientes de un conjunto de entrenamiento que comprende 29 plasmas, con evaluación de los 16 plasmas restantes. Los valores de *p* son para la hipótesis nula de un error de clasificación inferior al 87.5 %.

| Tamaño de matriz | Precisión | Glioblastoma | | Pulmón | | Meningioma | | Mieloma | | Páncreático | |
|------------------|-------------|--------------|----------|--------|----------|------------|----------|---------|----------|-------------|----------|
| | | Sens. | Especif. | Sens. | Especif. | Sens. | Especif. | Sens. | Especif. | Sens. | Especif. |
| Uno | 86.0 (0.43) | 80.1 | 96.5 | 85.9 | 98.4 | 83.9 | 96.3 | 88.1 | 97.1 | 89.6 | 97.7 |
| Dos | 92.4 (0.18) | 89.6 | 97.5 | 94.0 | 99.5 | 91.6 | 98.2 | 99.9 | 99.6 | 85.8 | 99.9 |
| Tres | 96.2 (0.34) | 100.0 | 96.9 | 94.0 | 100.0 | 98.5 | 97.0 | 100.0 | 100.0 | 87.4 | 99.5 |

El valor de los multiliposomas en el nanosistema de corona de proteína

Para evaluar la importancia de incluir múltiples liposomas en el nanosistema de corona de proteína, se repitió todo el procedimiento de clasificación usando 1000 divisiones diferentes de los datos en muestras de

entrenamiento y prueba, pero esta vez usando datos medidos de un solo liposoma. Esto se hizo para cada uno de los tres liposomas, y los resultados se informan en la primera fila de la Tabla 4. En relación con la matriz completa, incluidos los tres liposomas, las matrices de un solo liposoma mostraron una precisión, sensibilidad y especificidad significativamente menores. También se volvió a realizar el procedimiento para conjuntos de dos liposomas (hay tres conjuntos únicos); estos resultados se encuentran en la segunda fila de la Tabla 4. Los sistemas de dos liposomas son más precisos que un sistema de un liposoma, pero aún más débiles que la serie completa de tres liposomas. En general, esto indica el valor de los cambios en el patrón de composición de la corona entre diferentes liposomas y la necesidad de incluir tales observaciones múltiples.

Importancia y estabilidad de variables y proteínas

10 El modelo de bosque aleatorio también produce una puntuación de importancia para cada variable (es decir, cada par liposoma-proteína). Esta puntuación mide esencialmente qué tan importante fue esa variable para discriminar pacientes de diferentes tipos de cáncer. En un 'árbol' individual del bosque aleatorio, la puntuación de importancia de cualquier variable utilizada en la construcción del árbol se define como la proporción del conjunto de entrenamiento que se encuentra en las 'hojas' de los nodos que utilizan esa variable (a las variables 15 no utilizadas en la construcción del árbol se les asigna una puntuación de cero); entonces el puntaje de importancia general para una variable es el promedio de sus puntajes de importancia en cada árbol.

20 Para diferentes familias biológicas de proteínas (las mismas que se utilizan en la Figura 3A-F), se calculó la importancia general en la discriminación de diferentes tipos de cáncer (Figura 10A, (a)-(c)). Los resultados sugieren que diferentes familias de proteínas son importantes para detectar diferentes tipos de cáncer. Por ejemplo, las proteínas de fase aguda fueron relativamente importantes para detectar meningiomas y las 25 lipoproteínas fueron relativamente importantes para detectar glioblastomas. En particular, estas variaciones fueron distintas de las variaciones en el porcentaje de cada categoría adsorbida en los liposomas (Figura 10A, (d)-(f)). Curiosamente, estos resultados están en buen acuerdo con la función biológica de estas categorías de proteínas. Por ejemplo, está bien aceptado que el metabolismo de los lípidos está sustancialmente alterado en el glioblastoma en comparación con el tejido sano, lo que puede ser la razón principal de los cambios sustanciales observados en la interacción de las lipoproteínas con los liposomas. FIGURAS. 2C y 10A). Se proporcionan detalles sobre el número de proteínas identificadas y proteínas únicas en la composición corona de diferentes liposomas y sus combinaciones, Figura 10B.

30 También se calcularon las proteínas globales más importantes (Figura 10C y Tabla 5). Estas proteínas se detectaron en combinaciones de los tres liposomas, lo que nuevamente muestra el papel crítico de los múltiples liposomas en el nanosistema. También se evaluó la solidez de este conjunto de proteínas "importantes" en los clasificadores estimados en las 1000 divisiones de datos de entrenamiento discutidos anteriormente. Específicamente, la Figura 10B muestra el 25° a 75° percentiles de las puntuaciones de importancia para las 30 35 proteínas más importantes en promedio. Estos muestran que el conjunto de proteínas importantes es robusto a la división de datos utilizados para el entrenamiento del modelo. Entre estas proteínas más importantes (véase Tabla 5), algunos han sido reconocidas por desempeñar un papel fundamental en el desarrollo del cáncer. Por ejemplo, las ficolinas (tanto la Ficolina 2 como la Ficolina 3) son moléculas de reconocimiento de 40 patrones séricos con propiedades opsónicas con una capacidad sustancial para regular la activación del complemento. Se ha demostrado que las concentraciones séricas de Ficolin 3 son más altas en pacientes con cáncer de ovario que en sujetos sanos. Además, la Ficolina 3 se identificó en un análisis proteómico diferencial del suero de cáncer de próstata, lo que sugiere un papel para esta proteína también en el cáncer de próstata. Por otro lado, está bien establecido que la apolipoproteína A2 y sus isoformas se sobreexpresan en el suero 45 del cáncer de próstata y que la concentración de proteínas de fase aguda (por ejemplo, proteínas del complemento) puede cambiar en ≥ 25 % en presencia de trastornos inflamatorios como cáncer. Durante mucho tiempo se ha documentado una clara asociación entre el cáncer y el sistema hemostático. La hemostasia modula el flujo sanguíneo al regular la adhesión de plaquetas y el depósito de fibrina. Varias proteínas implicadas en la hemostasia se han relacionado con la regulación de la angiogénesis. Entre estos, el fibrinógeno es la proteína principal en el proceso de hemostasia y se ha encontrado en muchos tumores; modula la angiogénesis y el crecimiento tumoral y se ha implicado en la formación de metástasis. De hecho, los niveles 50 plasmáticos de fibrinógeno se han utilizado para pronosticar resultados clínicos en pacientes con carcinoma de células renales no metastásico y para predecir metástasis a distancia en cáncer de páncreas.

Tabla 5. Información sobre las proteínas generales más importantes, detectadas en combinaciones de los tres liposomas. * Las concentraciones son valores obtenidos mediante recuento espectral de la base de datos de proteoma de plasma (www.plasmaproteomedatabase.org).

| Nombre de entrada Uniprot | Descripción de la proteína | Función | Cantidad en plasma* |
|---------------------------|------------------------------|---|---------------------|
| FCN3 | Ficolina -3 | activación del complemento, vía de la lectina | 1 µg/ml |
| SAA4 | Proteína sérica amiloide A-4 | actividad quimioatrayente | 30 µg/ml |

| | | | |
|-------|--|--|-------------|
| CBPN | Cadena catalítica de carboxipeptidasa N | protege al cuerpo de péptidos vasoactivos e inflamatorios que contienen Arg o Lys C-terminal | 720 ng/ml |
| APOA2 | Apolipoproteína A2 | respuesta inflamatoria aguda, transporte de lípidos | 750 µg/ml |
| CO7 | Complemento componente C7 | regulador de la respuesta inmune innata y adaptativa | 2.6 µg/ml |
| FHR5 | Proteína 5 relacionada con el factor H del complemento | activación del complemento, vía alternativa | 11 ng/ml |
| COF1 | Cofilina-1 | organización del citoesqueleto de actina | 140 ng/ml |
| HABP2 | Proteína de unión a hialuronano 2 | actividad de endopeptidasa de tipo serina | 1.1 µg/ml |
| IGHG1 | Inmunoglobulina pesada constante gamma 1 | unión al antígeno | N.A.. |
| IGHG3 | Inmunoglobulina pesada constante gamma 3 | unión al antígeno | N.A.. |
| IGHG2 | Inmunoglobulina pesada constante gamma 2 | unión al antígeno | N.A.. |
| RET4 | Proteína de unión a retinol 4 | transportador de retinol | 580 µg/ml |
| VTNC | vitronectina | Adhesión celular | 35 µg/ml |
| GRP78 | Proteína regulada por glucosa de 78 kDa | Actividad ATPasa | 100 ng/ml |
| KV118 | Cadena Ig kappa región V-I WEA | Unión a antígeno | N.A.. |
| CPN2 | Carboxipeptidasa subunidad 2 N | regulación de la activación del complemento | 2 µg/ml |
| COL11 | Colectina-11 | unión a manosa, activación del complemento | N.A.. |
| MASP1 | Serina proteasa 1 de lectina de unión a manano | activación del complemento, vía de la lectina | 240 ng/ml |
| FIBB | Cadena beta de fibrinógeno | hemostasia | 706 µg/ml |
| FIBA | Cadena alfa de fibrinógeno | hemostasia | 2.5 mg/ml |
| C1S | Subcomponente del complemento C1s | regulación de la activación del complemento | 50 µg/ml |
| FGL1 | Proteína1 similar al fibrinógeno | hemostasia | 2.3 ng/ml |
| VWF | factor de von Willebrand | hemostasia | 110 µg/ml |
| CO6 | Complemento componente C6 | regulación de la activación del complemento | 40 µg/ml |
| IGHA1 | Inmunoglobulina pesada constante alfa 1 | unión al antígeno | N.A.. |
| IGLL5 | Polipéptido 5 similar a inmunoglobulina lambda | unión al antígeno | N.A.. |
| CO8A | Componente del complemento cadena alfa C8 | regulación de la activación del complemento | 70-90 µg/ml |
| K1C14 | Queratina, tipo citoesquelético 14 | I constituyente estructural del citoesqueleto | 210 ng/ml |
| CRP | Proteína C-reactiva | activación del complemento, vía clásica | 2 µg/ml |

| | | |
|------|--|-----------|
| CO8G | Componente del regulación de la activación del complemento cadena gamma C8 | 1.1 µg/ml |
|------|--|-----------|

Para investigar más a fondo el papel de las proteínas importantes identificadas por nuestro enfoque de aprendizaje automático, las buscamos en la base de datos Open Targets, una plataforma para la identificación y validación de objetivos terapéuticos. Esa base de datos calcula una puntuación de asociación de enfermedades para cada proteína en función de la evidencia de varias otras bases de datos (incluidos GWAS Catalog, UniProt, Gene2Phenotype, Cancer Gene Census, IntOGen, Europe PMC y Reactome) para derivar una puntuación en una escala de 0 (más bajo) a 1.0 (más alto) de asociación de enfermedades. De las proteínas enumeradas, tres tienen fuertes asociaciones con el cáncer. La proteína de unión a hialuronano tiene una asociación general muy fuerte (1.0) con los cánceres; La cadena beta del fibrinógeno tiene una asociación moderadamente fuerte (0.4) con el cáncer de pulmón; mientras que la Queratina, Tipo 1 citoesquelético 14 está fuertemente asociada (0.72) con el cáncer de próstata. Casi todas las demás proteínas de la Tabla 5 tienen algún grado de asociación débil (0.05 - 0.4) con varios tipos de cánceres. Por lo tanto, en general, las proteínas tienen un vínculo con asociaciones de cáncer conocidas.

Superación de la variabilidad en mediciones individuales con factorización de tensor

Debido a la variabilidad significativa en las poblaciones de pacientes, junto con el ruido introducido por el error de medición, encontramos que cualquier proteína individual era insuficiente para la clasificación. Para cada una de las 100 proteínas más abundantes (que producen $3 \times 100 = 300$ variables), calculamos el puntaje z absoluto promedio de esa concentración de proteínas en la corona observada en pacientes de cada tipo de cáncer. Una puntuación z más alta para un tipo de cáncer en una proteína dada indica que esta proteína puede ser útil para detectar este tipo de cáncer. Figura 11 (barras azules) muestra un histograma de estos puntajes z absolutos promedio para cada tipo de cáncer y el grupo saludable. También medimos puntajes z absolutos promedio para proteínas previamente vinculadas a estos tipos de cáncer específicos, que se muestran en el mismo histograma en la Figura 10 (barras de color gris claro). En todos estos cientos de variables, solo una proteína tiene una puntuación z absoluta superior a 2.0, y en un solo tipo de cáncer (mieloma), lo que sugiere que ninguna proteína es suficiente para una clasificación precisa.

La descomposición de tensores que precede a nuestra construcción de un bosque aleatorio cumple efectivamente el papel de calcular un pequeño número de *promedios ponderados* de composición proteica en la corona observada en una muestra dada. Este promedio es para todas las proteínas y nanopartículas en una muestra de plasma dada, y sirve para mitigar la variabilidad en la concentración observada de cualquier proteína dada. Por ejemplo, en la Figura 11 (barras negras largas), la puntuación z absoluta promedio para uno de estos promedios ponderados se representa para cada tipo de cáncer. Los puntajes z absolutos de estos promedios ponderados son significativamente más altos que los de cualquier variable individual. Además de la intuición que esto proporciona sobre cómo nuestro enfoque supera las variaciones en una sola proteína en muestras con la misma indicación, esta eliminación de ruido también afecta materialmente nuestra capacidad de clasificación. En ausencia de eliminación de ruido, el simple entrenamiento de un clasificador de bosque aleatorio en los datos sin procesar de la corona produciría una precisión de clasificación del 94 % (a un valor p de 0.07) en comparación con el 96.2 % (a un valor p de 0.04) informado en Tabla 4. Para los datos de cohorte informados en la Tabla 7 en ausencia de eliminación de ruido nos permitió aumentar la precisión de la clasificación del 92 % (a un valor p de 0.04) al 94.1 % (a un valor p de 0.01).

Dependencia de la precisión de la clasificación en los datos

Nuestra evaluación final sobre los 45 pacientes que no pertenecen a la cohorte fue sobre el valor de los datos. Nuestro clasificador original fue entrenado en 5 muestras por tipo de cáncer. Para medir el valor de incluir más o menos datos de entrenamiento, repetimos el procedimiento de clasificación (dividir los datos en muestras de entrenamiento y prueba, entrenar un clasificador en las muestras de entrenamiento y medir el rendimiento en las muestras de prueba, todo hecho 1000 veces) para un número variable de muestras de entrenamiento por clase, que van de cuatro a seis (Tabla 6). La precisión, la sensibilidad y la especificidad de la clasificación aumentaron con el aumento del número de muestras de entrenamiento. Con seis muestras de entrenamiento por clase, la precisión general alcanza el 96.8 % (*es decir* una tasa de error global del 3.2 %). Esto sugiere fuertemente que la precisión, la sensibilidad y la especificidad seguirán aumentando a medida que se incluyan más datos en el entrenamiento del clasificador.

Tabla 6. La precisión de la clasificación mejora con más datos. (Columna 2) Error de clasificación general cuando el conjunto de entrenamiento consiste en cuatro, cinco y seis muestras de cada cáncer y los valores p asociados. Tanto el error de clasificación como los valores p asociados mejoran con muestras adicionales de cada cáncer. (Columnas 3-12) Sensibilidad y especificidad para cánceres de glioblastoma, pulmón, meningioma, mieloma y pancreas cuando el conjunto de entrenamiento consiste en cuatro a seis muestras. Los resultados experimentales se presentan en 1000 extracciones independientes de un conjunto de entrenamiento con cuatro pacientes sanos y de cuatro a seis pacientes con cada tipo de cáncer. Los valores n son para la hipótesis nula de al menos dos errores de clasificación.

| | | Glioblastoma | | Pulmón | | Meningioma | | Mieloma | | Pancreático | |
|---------------|-------------|--------------|----------|--------|----------|------------|----------|---------|----------|-------------|----------|
| # de Muestras | Predicción | Sens. | Especif. | Sens. | Especif. | Sens. | Especif. | Sens. | Especif. | Sens. | Especif. |
| cuatro | 94.4 (0.14) | 98.8 | 98.1 | 91.2 | 100.0 | 93.0 | 96.7 | 100.0 | 99.9 | 88.9 | 98.4 |
| Cinco | 96.2 (0.04) | 100.0 | 98.9 | 94.0 | 100.0 | 98.5 | 97.9 | 100.0 | 100.0 | 97.4 | 99.5 |
| Seis | 96.8 (0.01) | 100.0 | 99.4 | 96.6 | 100.0 | 98.9 | 97.9 | 100.0 | 100.0 | 97.8 | 99.5 |

Muestreo de proteínas plasmáticas de baja y alta abundancia con liposomas sin necesidad de agotamiento de proteínas

5 El nanosistema de corona de proteína multiliposoma, utilizando técnicas de aprendizaje automático, produce un patrón de proteína de "huella" único para cada tipo de cáncer y para individuos sanos. En la búsqueda del mecanismo que subyace a la capacidad única de la corona de proteína en la identificación y discriminación del cáncer, hemos analizado minuciosamente la composición de la corona para determinar la contribución de proteínas de alta y baja abundancia, y hemos comparado esos resultados con la concentración de corona de proteínas en plasma humano. En este sentido, la contribución de cada proteína (por ejemplo, proteína X) a la composición de la corona y el plasma se normalizó con respecto a la albúmina utilizando (péptidos totales de albúmina en corona de proteína)/(péptidos totales de proteína X en corona de proteína) y (concentración de albúmina en plasma)/(concentración de proteína X en plasma), respectivamente. Es de destacar que buscamos manualmente las concentraciones de corona de proteínas identificadas en plasma utilizando una base de datos de proteoma en línea (<http://plasmaproteomedatabase.org/>). Como los péptidos totales de la albúmina en la corona de proteína son proporcionales al peso total de la albúmina en la corona de proteína, dividir estos péptidos en los péptidos de la proteína X es comparable con la relación entre la concentración de albúmina y la concentración de proteína X en el plasma. Como se muestra en la Figura 12, encontramos que las concentraciones plasmáticas varían en 10 órdenes de magnitud (en escala logarítmica), mientras que los liposomas detectan estas mismas proteínas en 4-5 órdenes de magnitud. En otras palabras, revelamos que la composición de la corona de proteínas tiene una gran capacidad para concentrar un amplio rango de proteínas de baja abundancia (≤ 100 ng/mg) y proteínas muy raras (≤ 10 ng/mg). Esto significa que la obtención de datos sobre proteínas muy abundantes en la corona de proteínas no interfiere con la detección de péptidos derivados de proteínas raras y menos abundantes. Vale la pena señalar que la mayoría de las proteínas detectadas son proteínas poco abundantes, raras o desconocidas/no declaradas. La participación de las proteínas de baja abundancia en la composición corona se debe principalmente al intercambio de corona de proteínas con baja afinidad por proteínas con mayor afinidad y cinética de adsorción más lenta.

30 La detección de proteínas de baja abundancia en plasma humano requiere agotamiento de proteínas muy abundantes y estrategias de fraccionamiento de plasma posteriores al agotamiento; sin embargo, incluso utilizando estas estrategias de agotamiento, la proteómica del plasma no ha tenido un gran éxito en la detección temprana de cánceres. Los resultados obtenidos con nuestro sistema indican que, a diferencia de las proteínas plasmáticas humanas, la composición de la corona de proteínas contiene un amplio rango de proteínas plasmáticas menos abundantes y muy raras, sin necesidad de agotamiento (que puede provocar la eliminación no intencionada de proteínas de baja abundancia/ biomarcadores). Además, encontramos varias proteínas en la corona de proteína cuyas concentraciones en el plasma humano son desconocidas/no informadas, posiblemente debido a su muy baja concentración en el plasma humano. Más específicamente, los liposomas aniónicos, neutros y catiónicos representan el 323, 189 y 155 de las proteínas con concentración plasmática desconocida/no informada, respectivamente). La contribución de estas proteínas a la corona de proteínas se indica mediante el cuadro rectangular a la derecha en la Figura 12.

40 El uso de multiliposomas (con distintas propiedades superficiales) brinda una oportunidad única para aumentar la profundidad de detección de proteínas poco abundantes y proteínas muy raras (véase Figura 12), que mejoró sustancialmente la sensibilidad, la especificidad y la precisión predictiva del nanosistema de corona de proteína. Cada liposoma proporcionó diferentes patrones de proteínas aportadas con una fuerte dependencia del tipo de cáncer.

Detección de cáncer y discriminación entre muestras de cohortes

45 Finalmente, para investigar la capacidad del nanosistema de corona de proteína para detectar cánceres en etapas muy tempranas, utilizamos plasma de cohorte (obtenido del Estudio de Cohorte de Golestan financiado por

NIH; los detalles se proporcionan en la sección Métodos) de personas sanas que fueron diagnosticadas hace varios años después de la recolección de plasma con cánceres de páncreas, pulmón y cerebro. Seguimos el mismo procedimiento, realizando 1000 experimentos. En cada experimento, dividimos los datos en 12 (4 de cada 3 tipos de cáncer) muestras de entrenamiento y 3 muestras de prueba (uno de cada 3 tipos de cáncer).

5 Se eliminó el ruido de las observaciones a través del procedimiento descrito anteriormente, y luego se entrenó un clasificador de bosque aleatorio en las muestras de entrenamiento. Finalmente, se midieron la precisión general, la sensibilidad y la especificidad en las muestras de prueba (Tabla 7). La precisión general fue del 94.1 % (una tasa de error del 5.9 %), con un valor de p de 0.01 para la hipótesis nula de una precisión general <66.7 % (por lo tanto, rechazada con un 95 % de confianza). Las sensibilidades en los tres tipos de cáncer oscilaron 10 entre el 83.2 % y el 100.0 %, y las especificidades entre el 91.6 % y el 100.0 %. Estos valores relativamente altos sugieren que la matriz de proteínas en corona puede detectar con éxito los cánceres en sus etapas más tempranas.

15 Tabla 7. Precisión, sensibilidad y especificidad de la clasificación general para muestras de cohortes. (Columna 2) Precisión de clasificación general para uno, dos y tres liposomas. Tanto el error de clasificación como los valores p asociados mejoran con la adición de liposomas al nanosistema. (Columnas 3-8) Sensibilidad y especificidad para los cánceres de cerebro, pulmón y páncreas, respectivamente. De nuevo, la sensibilidad y la especificidad mejoran con la adición de liposomas. Los resultados experimentales se promedian sobre 1000 sorteos independientes de un conjunto de entrenamiento que comprende 12 pacientes, con evaluación de los 3 pacientes restantes. Los valores de p son para la hipótesis nula de una precisión de clasificación <66.7 %.

| Tamaño de matriz | Precisión | Cerebro | | Pulmón | | Pancreático | |
|------------------|-------------|---------|----------|--------|----------|-------------|----------|
| | | Sens. | Especif. | Sens. | Especif. | Sens. | Especif. |
| Uno | 75.4 (0.23) | 74.7 | 86.1 | 90.7 | 90.9 | 60.8 | 86.9 |
| Dos | 80.5 (0.11) | 92.3 | 89.7 | 76.3 | 92.1 | 73.1 | 88.9 |
| Tres | 94.1 (0.01) | 100.0 | 100.0 | 83.2 | 99.6 | 99.2 | 91.6 |

20

También se realizaron los mismos experimentos con subconjuntos más pequeños de liposomas, incluyendo cada liposoma individualmente y los tres conjuntos únicos de dos liposomas. Nuevamente, se demostró que la precisión general, las sensibilidades y las especificidades aumentaron drásticamente a medida que aumentaba el número de liposomas, lo que nuevamente indica el valor del sistema de matriz.

25

Discusión

Ningún estudio anterior que haya utilizado la corona de proteína o cualquier otro enfoque para el fraccionamiento del proteoma plasmático ha permitido la detección simultánea de múltiples cánceres con una especificidad, sensibilidad y precisión predictiva aceptables. Esta es la primera vez que se desarrolla una matriz de sensores con especificidad, sensibilidad y precisión predictiva no solo para el cáncer, sino también para subtipos específicos de cáncer.

30 Se ha demostrado que el fraccionamiento del proteoma plasmático por la corona de proteínas no está relacionado con la abundancia de proteínas plasmáticas específicas y, en cambio, ocurre a través de muchos factores, incluida la afinidad de la proteína por la superficie de la nanopartícula a través de un amplio rango de fuerzas que incluyen fuerzas de Coulomb, dispersión de London, acidez y basicidad del enlace de hidrógeno, 35 polarizabilidad, pares de electrones solitarios e interacciones proteína-proteína entre las proteínas participantes en la estructura de la corona. Se ha demostrado que la composición de la corona de proteínas es dinámica e inicialmente dominada por abundantes proteínas, incluidas la albúmina, la inmunoglobulina y el fibrinógeno, que junto con otras 19 proteínas comprenden más del 99 % de la masa proteína en el proteoma plasmático. El 1 % restante del proteoma plasmático está compuesto por más de 10 000 proteínas; un subconjunto de estas 40 proteínas con afinidades más altas y/o cinéticas de absorción más bajas por la superficie de las nanopartículas compiten con las proteínas de gran abundancia por su inclusión en la composición de la corona. Además del intercambio de las proteínas con mayor afinidad de unión por aquellas con menor afinidad de unión en la superficie de las nanopartículas, existe la posibilidad de que otras proteínas de baja afinidad contribuyan en la 45 capa corona externa de proteínas debido a su favorable relación proteína-proteína. Interacciones con la capa corona de proteína ya formada. Esto significa que las proteínas poco abundantes intercambiadas pueden dirigir la formación de la corona de proteínas hacia la adsorción de más proteínas de baja abundancia.

50 Aquí analizamos por primera vez la capacidad de concentración relativa de la corona proteica para enriquecer proteínas de baja abundancia (definidas como <100 ng/ml), demostrando que muchas de estas proteínas se encuentran en concentraciones <10 ng/ml y acercándose a 1 pg/ millilitro. Más importante aún, mostramos que muchas de estas proteínas juegan un papel crucial en la identificación y discriminación del cáncer a través de enfoques de aprendizaje automático. El papel de la corona de proteínas en la concentración de un amplio rango de proteínas muy raras y de baja abundancia puede contribuir en gran medida a superar las razones

principales detrás del éxito limitado de las técnicas actuales de espectrometría de masas (incluida LC-MS/MS) en la detección temprana de cánceres, ya que estas técnicas pueden detectar proteínas en un rango dinámico de 4-6 órdenes de magnitud. En otras palabras, la composición de la corona de proteínas permite el muestreo en un amplio rango dinámico del proteoma plasmático, lo que puede mejorar sustancialmente la profundidad de la cobertura de proteínas sin agotamiento de proteínas. Dado que la corona de proteína obtenida de una sola nanopartícula constituye como máximo varios cientos de proteínas distintas (un pequeño subconjunto del proteoma total), postulamos que el uso de múltiples nanopartículas con distintas propiedades fisicoquímicas podría crear dimensiones adicionales de información proteómica: 1) cada nanopartícula adicional con distintas propiedades fisicoquímicas permite potencialmente el reclutamiento de proteínas únicas adicionales de baja abundancia (Figura 12); 2) las coronas de proteína que se superponen a la superficie de más de una nanopartícula participan en diferentes porcentajes de contribución corona y, por lo tanto, alteran la concentración y la identidad de otras corona de proteínas participantes; y 3rd) tanto la información de corona de proteína única como la superpuesta de cada nanopartícula sirven como variables únicas y, por lo tanto, proporcionan más datos a nuestro enfoque de aprendizaje automático. La combinación de más nanopartículas para el fraccionamiento de plasma y el análisis proteómico proporciona significativamente más información para la detección y discriminación del cáncer con una sensibilidad, especificidad y precisión de predicción superiores en comparación con menos nanopartículas (Tabla 4). Conceptualmente, nuestro enfoque de nanopartículas múltiples es similar al sistema olfativo, en el que la especificidad del reconocimiento de olores se origina en el patrón de respuestas de varios cientos de receptores olfativos altamente reactivos cruzados, donde cualquier receptor proporciona información incompleta pero la combinación es altamente específica en la identificación de un odorante determinado (es decir, en los seres humanos ~400 receptores activos pueden detectar y diferenciar ~10 000 odorantes, y en perros ~1200 receptores activos pueden detectar ~1 000 000 de odorantes). Otros investigadores también han utilizado enfoques similares para detectar y diferenciar entre diversas familias de analitos, diversos alimentos y bebidas, bacterias y hongos patógenos, biomoléculas e incluso las propias nanopartículas.

Se utilizaron tres liposomas de reacción cruzada diferentes (con cargas superficiales negativa, neutra y positiva) cuyos perfiles de corona de proteína se midieron después de la exposición al plasma de pacientes individuales que tenían uno de cinco tipos de cáncer: pulmón, páncreas, mieloma, meningioma o glioblastoma. Para identificar y discriminar entre los tipos de cáncer utilizando los resultados del nanosistema de corona de proteína, utilizamos un enfoque de aprendizaje automático de bosque aleatorio bien definido. Hemos designado 1000 conjuntos diferentes de muestras como muestras de entrenamiento (es decir, plasmas con cánceres conocidos y condiciones saludables) o muestras de prueba (es decir, plasmas ciegos) para garantizar que nuestro nanosistema de corona de proteína sea robusto y preciso para la detección de cáncer con excelente precisión predictiva. Aunque ninguna composición de corona de proteína de una sola nanopartícula fue específica para ningún tipo de cáncer en particular con una precisión predictiva aceptable (es decir, 86.0 % con un valor de *p* de 0.43 para el error de clasificación inferior al 87.5 %), encontramos que el patrón de composición de corona derivado de los multiliposomas proporciona una "huella" única para cada tipo de cáncer con una precisión predictiva excelente (es decir, 96.2 % con un valor de *p* de 0.04 para el error de clasificación inferior al 87.5 %). Estos resultados, basados en el análisis profundo del proteoma humano utilizando la caracterización de la corona de proteínas multiliposomáticas y el aprendizaje automático, confirmaron la promesa de este sistema para la identificación y discriminación inequívocas de los cánceres y la discriminación sin errores contra sujetos sanos con una especificidad excelente (de 97.0 % al 100.0 %).

Para probar la capacidad del nanosistema de corona de proteína para la detección muy temprana de cánceres, se utilizaron muestras de plasma de cohortes. Estas muestras se recolectaron de personas sanas a las que se les diagnosticó cáncer de pulmón, páncreas o cerebro ocho años después de la recolección de plasma. Los resultados del nanosistema de corona de proteína, utilizando plasma de 15 pacientes en el estudio de cohorte, revelaron que nuestro enfoque identificó y discriminó con precisión los cánceres incluso en estas muestras de cohortes previas al diagnóstico, para las cuales la detección de cáncer no era posible con las alternativas actuales. De acuerdo con nuestros hallazgos en las muestras de plasma fresco, el muestreo de plasma con múltiples liposomas proporcionó una precisión de clasificación superior (94.1 % frente a 75.4 %) y especificidad [es decir, cerebro (100.0 % frente a 86.1 %) pulmón (96.6 % frente a 90.9 %) y pancreático (91.6 % frente a 86.9 %)] en comparación con el muestreo de un solo liposoma. Cabe señalar que los perfiles de corona de proteínas de las muestras de la cohorte fueron diferentes en comparación con las muestras de cáncer frescas anteriores. Esto se debe principalmente al almacenamiento prolongado en congelación (aproximadamente 10 años) de las muestras de la cohorte, ya que las muestras se recolectaron en el momento de la selección de individuos sanos. Cada vez se acepta más que el almacenamiento a largo plazo de muestras de plasma puede afectar fuertemente la concentración y la integridad de un subconjunto de proteínas, alterando a su vez la composición de la corona de proteínas y disminuyendo la sensibilidad de nuestro nanosistema de corona de proteínas. Por lo tanto, la sensibilidad máxima del nanosistema se puede lograr cuando se usa plasma fresco, ya sea como parte de una prueba de detección de cáncer o como parte de un estudio de cáncer después del diagnóstico. El valor en este último escenario puede ser que los cambios en el patrón de la corona de proteínas a lo largo del tiempo proporcionen información valiosa relevante para la recurrencia temprana del tumor o la presencia de enfermedad residual después de la resección del tumor.

En resumen, presentamos una prueba de concepto de que el nanosistema de corona de proteínas de múltiples nanopartículas tiene un potencial considerable para aumentar la profundidad de detección de proteínas y, por lo tanto, una capacidad única para detectar y discriminar con precisión los cánceres incluso en sus etapas más tempranas, cuando las tecnologías existentes no logran detectar enfermedad. El patrón de corona de proteínas de nanopartículas múltiples derivado de nuestro nanosistema de corona de proteína proporciona una "huella" multivariada única para la detección del cáncer, lo que no es posible cuando se analiza la corona de proteína de una sola nanopartícula. Además, el patrón de corona de proteínas representa el enriquecimiento colectivo de un amplio rango de proteínas plasmáticas (incluidas proteínas abundantes y raras) utilizando nanopartículas distintas, lo que lo distingue claramente de otros enfoques proteómicos multivariados de plasma completo que no han logrado producir resultados adecuados para la detección de cáncer temprano. La profundidad de detección de proteínas (la principal limitación del análisis de proteínas en plasma humano) y la precisión predictiva del nanosistema se pueden mejorar aún más mediante el uso de nanopartículas adicionales con diferentes propiedades fisicoquímicas. El resultado predictivo exitoso de nuestro enfoque de caracterización con corona de proteínas y aprendizaje automático para plasmas frescos ciegos y plasmas de cohortes retrospectivos proporciona una base adecuada para estudios prospectivos posteriores en cáncer. Además, la utilidad del nanosistema de corona de proteína multinanopartícula se puede aplicar a otras enfermedades humanas importantes para las cuales la detección temprana puede mejorar significativamente tanto la longevidad como la calidad de vida.

liposomas se prepararon como se describe en el Ejemplo 1A.

20 *Recolección, preparación y almacenamiento de plasma humano* se realizó como se describe en el Ejemplo 1A.

Muestras de plasma de cohorte se prepararon como se describe en el Ejemplo 1A.

25 *Microscopía electrónica de transmisión (TEM)*. Las formulaciones de liposomas se han caracterizado por TEM como se informó anteriormente. En resumen, se depositaron 10 μl de cada muestra en rejillas recubiertas con Formvar, se tiñeron negativamente con acetato de uranilo al 1 %, se lavaron con agua ultrapura y se secaron al aire. Las mediciones se realizaron con un Zeiss Libra 120 y el análisis de imágenes se realizó con el software ImageJ.

Tamaño y potencial zeta se determinó como se detalla en el Ejemplo 1A.

Ensayo de proteínas se llevó a cabo como se describe en el Ejemplo 1A.

30 *Identificación y cuantificación de proteínas* se realizó como en el Ejemplo 1A. Se usaron recuentos de espectro no ponderados (USC) para evaluar la consistencia de las réplicas biológicas en el análisis cuantitativo, y se usaron recuentos de espectro normalizados (NCS) para recuperar la abundancia de proteínas.

Análisis estadístico. Todos los análisis estadísticos del texto principal se realizaron en Python con los paquetes scikit-learn, numpy y scipy, y las figuras y gráficos se crearon con el paquete bokeh en Python, junto con Microsoft Excel, XLSTAT y MATLAB.

35 *Matrices de datos*. Para las 60 muestras de plasma (45 de no cohorte y 15 de cohorte), etiquetadas como $i = 1, \dots, 60$, una matriz de datos X_i (con 3 filas y ~ 900 columnas) se generó de manera que cada fila de la matriz corresponde a las abundancias de proteínas de una sola nanopartícula, obtenidas del nanosistema de corona de proteínas. Como paso de preprocessamiento, convertimos las abundancias de proteínas en abundancias relativas de proteínas (RPA) mediante la normalización de las filas de todas las matrices.

40 *Clasificación y agrupamiento*

Factorización tensorial. Tratamos los datos como un tensor de tres modos, los dos primeros modos correspondientes a nanopartículas y proteínas, y el tercer modo correspondiente a muestras de plasma; esto es esencialmente equivalente a apilar las matrices de observación correspondientes a cada muestra, X_i , encima del otro. Los datos se eliminaron del ruido a través de una factorización de tensor de rango bajo de Tucker (135,136,137) usando código implementado en Python para este proyecto (disponible para uso académico previa solicitud). Cada matriz X_i se aproxima mediante una descomposición tensorial que toma la forma

$$X_i \sim US_i V^\top,$$

50 donde U es una matriz cuyas filas pueden verse como características latentes correspondientes a cada una de las nanopartículas, y de manera similar V es una matriz cuyas filas pueden verse como características latentes correspondientes a cada una de las proteínas; estas características latentes se comparten en todas las matrices de datos. Finalmente, cada S_i es una matriz que codifica las interacciones entre las características de las nanopartículas y las proteínas, y se permite que estas sean únicas entre las muestras. Estimamos esta descomposición en dos pasos: (a) estimamos U y V a través de una descomposición de valor singular truncada en los desarrollos de modo 1 y modo 2 del tensor, y luego dadas estas estimaciones, (b) ajustamos cada S_i matriz por separado a través de un cálculo de mínimos cuadrados.

Clasificación aleatoria de bosques. El modelo de bosque aleatorio es un conocido algoritmo de aprendizaje automático para la clasificación. Un bosque aleatorio se compone de múltiples árboles de decisión, cada uno de los cuales toma decisiones de clasificación simples basadas en relativamente pocas variables. Estos árboles se crean (o "entrenan") con diferentes subconjuntos de variables extraídos al azar, de modo que es probable que no haya dos árboles idénticos. Dada una nueva muestra, cada árbol se recorre de arriba hacia abajo hasta que se alcanza un conjunto de muestras de entrenamiento en la parte inferior. Usar el bosque como un todo para la clasificación equivale a hacer que los árboles de decisiones múltiples "voten" en una etiqueta (en este caso, uno de los cinco tipos de cáncer o saludable), donde el voto de cada árbol se realiza a partir de las etiquetas del conjunto inferior de muestras de entrenamiento. Para nuestro propio algoritmo, cada bosque aleatorio constaba de 1000 árboles de decisión y se entrenó con el paquete scikit-learn. Las puntuaciones de importancia también se calcularon utilizando el mismo paquete.

Ejemplo 2. Construcción de matriz de sensores adicional para la detección de enfermedades.

Se fabrica una matriz de sensores que consiste en 12 nanopartículas de reacción cruzada diferentes, incluidos tres liposomas, tres nanopartículas de óxido de hierro superparamagnético y seis nanopartículas de oro. Estos tipos de liposomas (DOPG (1,2-dioleoil-sn-glicero-3-fosfo-(1'-rac-glicerol)), DOTAP(1,2-dioleoil-3-trimetilamonio-propano)-DOPE(dioleoilfosfatidiletanolamina), y CHOL (DOPC-colesterol)) con cargas superficiales negativas, neutras y positivas se sintetizan de acuerdo con nuestros informes anteriores. Nanopartículas de óxido de hierro superparamagnético recubiertas con PEG ultra uniformes de 20 nm de tamaño y de varios pesos moleculares de PEG (es decir, 300, 3000 y 6000) se obtienen de Micromod®. Nanopartículas de oro, con tamaño de núcleo de ~ 2 nm, y se sintetizan diferentes funcionalidades de superficie.

Todas las nanopartículas tienen el mismo tamaño pero tienen diferentes propiedades de superficie, que se espera que formen composiciones de corona de proteínas significativamente diferentes en respuesta al plasma de pacientes que desarrollan varios tipos de cáncer. La matriz de sensores prueba la capacidad de la matriz de sensores de corona de proteínas para identificar y discriminar entre los tipos de cáncer a través de las proteínas plasmáticas del paciente, como se describe en el Ejemplo 1.

Se analizan muestras de un amplio rango de cánceres, incluidos pulmón, páncreas, mieloma, meningioma, glioblastoma, carcinoma de células escamosas de esófago y adenocarcinoma gástrico. Las matrices de corona de proteínas basadas en diferentes nanopartículas (liposomas, óxido de hierro y oro) están diseñadas para investigar la variación en la huella de biomoléculas entre varios tipos de cáncer. El ensayo del sensor también permitirá la identificación de proteínas que son marcadores para tipos específicos de cáncer, incluidos biomarcadores nuevos y desconocidos para diferentes tipos de células cancerosas. El reconocimiento de patrones que se puede determinar en función de los diferentes tipos de cáncer o la etapa del cáncer.

Se analizará un amplio rango de plasmas humanos de personas sanas que luego fueron diagnosticadas con varios tipos de cánceres varios años después de la recolección de plasma. Las muestras de plasma se recolectaron a través de un estudio de cohortes financiado por los NIH, denominado Estudio de cohortes de Golestan, realizado por el Instituto Nacional del Cáncer (NCI) en los EE. UU., la Agencia Internacional para la Investigación del Cáncer (IARC) en Francia y la Universidad de Teherán de Ciencias Médicas (TUMS) en Irán. Este estudio involucró la recolección y almacenamiento de plasma de 50,000 sujetos sanos. Más de 1000 de estos sujetos desarrollaron varios tipos de cáncer en los años siguientes. Las muestras se almacenan en IARC y nuestro equipo las utiliza para su análisis. Estas importantes muestras de plasma nos brindan una oportunidad única para probar la capacidad de nuestra innovadora matriz de sensores de corona de proteínas para la detección temprana de cánceres. Además, también evaluaremos e identificaremos proteínas útiles para la identificación y discriminación de estos cánceres.

Creemos que los resultados de la aplicación de este innovador matriz de sensores a muestras de plasma de cohorte no solo serán fundamentales en la detección y detección de cánceres en etapas tempranas, sino que también ayudarán a identificar nuevos marcadores de proteínas involucrados en el desarrollo del cáncer. Nuestras opciones de componentes de matriz de sensores tienen una capacidad más específica en comparación con otros métodos desarrollados para proporcionar huellas para una amplia variedad de proteínas de una manera no específica y de reacción cruzada para la identificación y discriminación de cánceres.

Los perfiles de corona duros de los elementos de la matriz de sensores se prueban utilizando plasma de pacientes con cánceres en estadios intermedios y avanzados:

La composición de la corona de proteínas que se forma en la superficie de los elementos de la matriz de sensores depende en gran medida de las propiedades fisicoquímicas de esas nanopartículas y, al mismo tiempo, puede verse fuertemente afectada por el tipo de enfermedad presente en el donante del plasma humano utilizado. para la incubación. Para preparar las nanopartículas recubiertas de corona, las 12 nanopartículas preparadas se incuban con plasma humano (por separado) de nueve tipos de cáncer (pulmón, páncreas, mieloma, leucemia mieloide, meningioma, glioblastoma, mama, carcinoma de células escamosas de esófago y adenocarcinoma gástrico) y aislado de proteínas libres a través de un enfoque de centrifugación bien

definido. La centrifugación se suele realizar a 13000 g durante 30 min a 15 °C. Se eliminará el sobrenadante y las partículas recogidas se redispersarán en 500 microlitros de PBS. El procedimiento se repetirá para eliminar las proteínas sueltas. Para eliminar las proteínas sueltas de la superficie de las nanopartículas, las nanopartículas recolectadas se redispersarán en PBS frío (15 °C) y se recolectarán mediante centrifugación. A continuación, se determinará el tamaño y la carga de las nanopartículas recubiertas de corona mediante DLS/Nanosight y se compararán con sus valores iniciales obtenidos en el regulador. Se realizará una evaluación cuantitativa de la proteína total adsorbida en las nanopartículas, a través de el ensayo BCA o NanoOrange, mientras que el análisis de proteómica de escopeta cualitativa identifica las proteínas adsorbidas en la superficie de las 12 nanopartículas. En resumen, después de la separación de las proteínas de la superficie de las nanopartículas (de acuerdo con el protocolo (Saha, K.; Rahimi, M.; Yazdani, M.; Kim, ST; Moyano, DF; Hou, S.; Das, R.; Mout, R.; Rezaee, F.; Mahmoudi, M. ACS nano 2016, 10, (4), 4421-4430), las proteínas se inyectan en un aparato de cromatografía líquida-espectrometría de masas (LC-MS/MS). Las proteínas se identifican a partir de los datos resultantes mediante la selección de bases de datos relevantes. Para obtener el número total de espectros LC-MS/MS para todos los péptidos atribuidos a una proteína emparejada, el índice de semi-^aPolidispersidad de la evaluación cuantitativa del ajuste cumulante de la cantidad de proteína se realizará a través del conteo espectral (SpC). Las cantidades normalizadas de SpC (NpSpC) de cada proteína identificada en los espectros de LC-MS/MS se calcularán mediante la siguiente ecuación:

$$NpSpCk = \left(\frac{\frac{SpC}{(M_w)_k}}{\sum_{t=1}^n \left(\frac{SpC}{(M_w)_t} \right)} \right) \times 100$$

donde NpSpC es el porcentaje normalizado del recuento espectral (es decir, el recuento crudo de iones) para la proteína k, SpC es el recuento espectral y Mw es el peso molecular (en kDa) de la proteína k.

Desarrollo de análisis de agrupamiento supervisado y no supervisado para identificar y discriminar entre los tipos de cáncer utilizando los resultados de la matriz de sensores:

Con el fin de investigar si las huellas de corona de proteína (PCF) de varios elementos sensores podrían utilizarse como biosensores y formar patrones únicos para diferentes enfermedades (firma de corona de biomolécula), hemos aplicado enfoques de clasificación enfocados a los datos proteómicos de la composición de corona de proteína de tres liposomas. (catiónico, aniónico y neutro) como se describe en el Ejemplo 1.

Ejemplo 3. Conjugación de nanopartículas al sustrato para hacer matrices de sensores.

Pueden usarse diferentes tipos de partículas como elementos sensores a nanoscala en la práctica de esta invención. Además, se proporcionan diferentes métodos de conjugación de los elementos sensores a nanoscala con el sustrato. Además, los elementos sensores a nanoscala se pueden unir al sustrato en diferentes patrones.

Los ejemplos específicos de diferentes configuraciones de elementos sensores en diferentes sustratos se ejemplifican en la Figura 15-43.

Ejemplo 4. Conjunto de sensores que comprende nanopartículas de sílice y poliestireno para la detección del cáncer

Este ejemplo demuestra que una matriz de sensores de la presente invención con nanopartículas diferentes a las utilizadas en el ejemplo 1 todavía es capaz de detectar muestras de cáncer a partir de muestras de pacientes sanos.

Utilizando el protocolo experimental del Ejemplo 1, se diseñó una matriz de sensores utilizando partículas de sílice y poliestireno funcionalizadas. En este ejemplo, un total de seis nanopartículas: dos tipos de nanopartículas, es decir poliestireno (P) y sílice (S), con tres diferentes funcionalizaciones superficiales, es decir ninguno, modificación de amina y modificación de carboxilo (P-NH₂, P-COOH, S-NH₂ y S-COOH).

La caracterización de nanopartículas desnudas de poliestireno y sílice con diferente funcionalización (ninguna, modificación de amina (NK2) y modificación de carboxilo (COOH) se muestra en las FIGURAS 44A-44D demostrando sus tamaños, DLS y potencial zeta de las partículas desnudas e imágenes TEM. La caracterización de nanopartículas de sílice y poliestireno recubiertas de proteína con diferente funcionalización se demuestra en las FIGURAS 45A-45D con sus tamaños, DLS, potencial zeta y TEM de las nanopartículas de poliestireno y sílice cargadas con corona de proteína.

La matriz de sensores de 6 nanopartículas se puso en contacto con el plasma de individuos sanos o pacientes con cáncer de recto, cáncer de mama, cáncer de vejiga, cáncer de tiroides, cáncer de útero, cáncer de ovario, cáncer de riñón (5 pacientes por cáncer) como se muestra en la Figura 46 por el método descrito en el ejemplo 1. Los perfiles de corona de proteínas de poliestireno y nanopartículas de sílice (100 nm) se analizaron

mediante SDS PAGE. La comparación de la corona de proteína de partículas simples, modificadas con amina y modificadas con carboxilo por SDS PAGE se muestra en la Figura 47.

Los perfiles de corona de proteínas para plasma sano para poliestireno y nanopartículas de sílice (100 nm) analizados por SDS-PAGE se muestran en la Figura 48. Comparación de la corona proteica de partículas simples, modificadas con amina y modificadas con carboxilo.

Los datos se analizaron como se describe en el Ejemplo 1. Los resultados del análisis estadístico y del agrupamiento se representan en la Figura 49. Como se muestra, los individuos sanos pueden identificarse y clasificarse en comparación con los pacientes con cáncer (los controles sanos están en el espacio ortogonal frente a las muestras de pacientes con cáncer como se ve en la Figura 49).

Este ejemplo demuestra que esta matriz de sensores que utiliza una matriz de seis nanopartículas puede discriminar y detectar un paciente con cáncer de un individuo sano.

Materiales: Tres partículas de sílice funcionalizadas de manera diferente se adquirieron de Kisker-Products (<https://www.kisker-biotech.com/>); tres partículas de poliestireno funcionalizadas de manera diferente se adquirieron de Polyscience, Inc. (<http://www.polysciences.com/>). Todas las partículas tenían el mismo tamaño (100 nm). Su morfología, tamaño promedio, índice de polidispersidad (PDI) y potencial zeta se caracterizaron mediante mediciones de TEM, DLS y potencial zeta.

Información experimental: 1h de incubación en 50 % de plasma humano sano. SDS PAGE: 4-20 % acrilamida 45 min 40 mA/gel. tinción: Comassie azul coloidal durante la noche. nanopartículas utilizadas: 0.5 mg).

Ejemplo 5. El nanosistema de matriz de sensores de corona de proteínas identifica la enfermedad de las arterias coronarias

La enfermedad de las arterias coronarias (CAD) es el tipo más común de enfermedad cardíaca y representa la principal causa de muerte tanto en hombres como en mujeres. La detección temprana de CAD es crucial para prevenir la muerte, prolongar la supervivencia y mejorar la calidad de vida de los pacientes. Este ejemplo describe el nanosistema de matriz de sensores no invasivo que contiene seis nanopartículas para la detección ultraprecisa de CAD mediante el reconocimiento de patrones de PC específicos. Si bien la PC de una sola nanopartícula no brinda la especificidad requerida, las PC multivariadas en seis nanopartículas distintas con diferentes químicas de superficie brindan la información deseable para discriminar selectivamente cada afección cardiovascular bajo investigación.

La CAD es una condición crónica que comienza durante la adolescencia y progresó gradualmente a lo largo de toda la vida de la persona afectada. Se caracteriza por la presencia de placas ateroscleróticas en las arterias coronarias. La génesis de la aterosclerosis radica en la disfunción del endotelio: cuando se someten a estímulos de estrés y factores inflamatorios (por ejemplo, estrés oxidativo y fuerzas hemodinámicas), las células endoteliales expresan moléculas de adhesión a la superficie que inducen el reclutamiento de leucocitos circulantes y lipoproteínas de baja densidad (LDL) que contienen colesterol.⁵ Estos eventos inducen la formación de placas ateroscleróticas, que estrechan la arteria coronaria y, por lo tanto, perjudican el flujo sanguíneo. Dependiendo de la velocidad de desarrollo de la placa y de la gravedad de la obstrucción arterial, los síntomas pueden culminar en un infarto de miocardio.⁵

Un diagnóstico preciso y a tiempo de CAD en sujetos en riesgo es muy importante para iniciar de inmediato una terapia *ad hoc* y evitar complicaciones posteriores. La angiografía coronaria es hasta la fecha el método más preciso y confiable para el diagnóstico de CAD. Sin embargo, insertar un catéter en una arteria del brazo (o cuello o parte superior apretada) hasta el corazón es invasivo, costoso y causa muchos efectos secundarios, incluidas infecciones, lesiones en la arteria cateterizada, alergia y sangrado excesivo. Por lo tanto, existe una necesidad urgente de desarrollar nuevas pruebas para la detección de CAD. Si bien se ha informado que varios biomarcadores inflamatorios son útiles para el diagnóstico, desafortunadamente, ninguno de ellos se usa en la práctica clínica, lo que destaca la necesidad aún muy frecuente de nuevas pruebas de diagnóstico. Los inventores han desarrollado un análisis de sangre basado en nano como una nueva herramienta para el diagnóstico de CAD.

Como se demostró en los Ejemplos anteriores para el cáncer, las PC personalizadas actúan como huella de una condición de plasma determinada. Este Ejemplo usa el mismo enfoque para definir con precisión la formación de la placa aterosclerótica, a través de sus cambios inducidos en la composición del plasma. De hecho, las células asociadas a la placa (por ejemplo, células espumosas, macrófagos, mastocitos, monocitos y células T) arrojan un amplio rango de biomoléculas (por ejemplo, citoquinas, proteasas y biomoléculas vasoactivas) a la sangre,²² y por lo tanto puede inducir cambios en el patrón de composición de la PC en la superficie de varias nanopartículas con respecto a la PC de pacientes sin formación de placa.

Este ejemplo analizó la PC formada alrededor de nanopartículas utilizando plasma derivado de i) pacientes a los que se les diagnosticó CAD después de una angiografía coronaria (CAD), ii) pacientes con síntomas que se sometieron a una angiografía coronaria y sus vasos coronarios se encontraron sanos (NO CAD), iii)

reestenosis (recurrencia de CAD después del tratamiento) y iv) voluntarios sanos sin factores de riesgo (por ejemplo, antecedentes familiares, tabaquismo, obesidad, hipertensión) de enfermedad cardiovascular (CONTROL).

5 Con el fin de tener un espectro más amplio de proteínas adsorbidas, el Ejemplo usó seis nanopartículas disponibles comercialmente como elementos de la matriz de sensores, con diferente composición y química de superficie y/o funcionalización, creando una matriz de sensores basada en PC de 6 nanopartículas capaz de usarse como un prueba diagnóstica fácil y no invasiva. Esta novedosa prueba diagnóstica de CAD se utilizará como detección pre-no invasiva para pacientes en riesgo. En particular, estos resultados demostraron que los 10 patrones de PC permitieron la discriminación ultraprecisa entre CAD, NO CAD, reestenosis y pacientes de control, proporcionando así una herramienta novedosa, precisa y no invasiva nunca antes desarrollada para el diagnóstico de CAD basado en sangre.

Resultados

15 En este ejemplo, un total de seis nanopartículas: dos tipos de nanopartículas, *es decir* poliestireno (P) y sílice (S), con tres diferentes funcionalizaciones superficiales, *es decir* ninguno, modificación de amina y modificación de carboxilo (P-NH₂, P-COOH, S-NH₂ y S-COOH), como se describe en la Figura 50 fueron usados. Se ha medido el tamaño, el potencial zeta y la morfología de las nanopartículas antes y después de la incubación en plasma para comparar las diferencias en los resultados entre la identidad sintética de las nanopartículas 20 desnudas y su identidad biológica correspondiente (nanopartículas recubiertas de PC). El análisis de dispersión de luz dinámica mostró que todas las nanopartículas desnudas estaban altamente monodispersas, como lo demuestra el índice de polidispersidad ≤ 0.02 , y de tamaño homogéneo de aproximadamente 100 nm, en el rango de 93 nm a 120 nm (Figura 50A). Después de 1 h de incubación con plasma de pacientes, el tamaño de 25 todas las nanopartículas aumentó debido a la presencia de una capa de proteínas adsorbidas (PC), cuyo grosor y composición han demostrado depender de la concentración de proteínas, las propiedades superficiales y el tamaño de las nanopartículas. Todas las nanopartículas desnudas tenían carga superficial negativa (Figura 50B), siendo las funcionalizadas con amina ligeramente menos negativas que otras debido a la contribución de los grupos amino positivos. Estos resultados estaban en línea con las especificaciones proporcionadas por el proveedor y otros estudios. La modificación con grupos amina no fue suficiente para cambiar la carga superficial de las nanopartículas de sílice y poliestireno caracterizadas por todos los residuos cargados negativamente en sus superficies a pH fisiológico.

30 Una vez expuesto al plasma, todas las cargas superficiales se volvieron menos negativas (de -5 mV a -25 mV) debido a la carga de la mayoría de las proteínas plasmáticas a pH fisiológico. En general, las propiedades fisicoquímicas de las nanopartículas recubiertas con PC mostraron tendencias similares, siendo siempre más grandes y menos negativas que sus contrapartes desnudas, independientemente del plasma utilizado para la incubación. Sin embargo, cuando se incubaron con plasma NO CAD, las nanopartículas de P y S exhibieron 35 un aumento de tamaño de ≈ 85 nm, mayor que el mostrado usando plasma de otras condiciones (40-50 nm). Por otro lado, el espesor de PC de las nanopartículas de S-NH₂ incubadas con plasma CAD fue mayor (≈ 40 nm) que las derivadas de la incubación de las mismas nanopartículas con otros tipos de plasma (≈ 30 nm). La microscopía electrónica de transmisión mostró que las nanopartículas no cambiaron su morfología y estructura 40 después de la incubación en plasma (Figura 50C). Además, se observó un aumento en el tamaño después del recubrimiento, lo que confirma los resultados obtenidos por dispersión de luz dinámica. Las concentraciones de proteínas de diferentes PC se evaluaron mediante el ensayo de Bradford, lo que muestra que, en general, todas las nanopartículas de sílice adsorberon menos proteínas en la PC que las nanopartículas de poliestireno (Figura 45 51A). Esta observación se confirmó mediante el análisis de las PC mediante 1D-SDS PAGE. Se utilizaron cinco pacientes con CAD, SIN CAD y sin riesgo de CAD (CONTROL) para recolectar plasma y las PC obtenidas para todas las nanopartículas se resolvieron en geles SDS PAGE teñidos con Comassie (Figura 51B). Se han analizado geles por densitometría y se han detectado diferencias en la cantidad de proteínas en los PCs CAD, NO CAD y CONTROL de una misma nanopartícula (Figura 51C, flechas). En algunos casos, 50 también notamos la presencia y/o ausencia de algunas proteínas en PC específicas (Figura 51C, flechas azules). Estos resultados confirman que la formación de placa de aterosclerosis induce cambios en la composición del plasma y, en consecuencia, diferencias en la PC. La corona de proteína se analizó mediante análisis LC-MS/MS. Se han identificado más de 150 proteínas en cada muestra de nanopartículas recubiertas con PC. Además, los valores de conteo espectral, que representan el número total de espectros de fragmentación para todos los péptidos atribuidos a una proteína específica, se han utilizado para obtener 55 información sobre la abundancia de las proteínas y, en consecuencia, el porcentaje de contribución de cada proteína identificada en las PC. Las diferencias en la contribución porcentual de las 20 proteínas más abundantes en las PC se informan en la Figura 52. Los resultados demostraron correlación entre la composición de PC, la condición del plasma, el tipo de nanopartícula y la funcionalización de la superficie.

60 Los datos obtenidos de todas las muestras se recopilaron, analizaron y clasificaron: se creó una clave para cada medición mediante la concatenación de nanopartículas, modificación de la superficie, tipo de plasma y etiqueta. Las proteínas identificadas con menos de 2 péptidos se eliminaron de la consideración y se usó TIBCO Spotfire Analyst 7.6.1 para pivotar los datos de modo que las filas se identificaran por entrada de proteína, las columnas por la clave y los valores que contenían el porcentaje de contribución.

Además del análisis matemático, analizamos la presencia de proteínas exclusivas en la corona de una determinada condición. Para hacer esto, creamos Diagramas de Venn que facilitan la investigación de proteínas comunes y únicas en grandes conjuntos de datos de proteínas o genes. Las PC de cada nanopartícula se han analizado por separado, para un total de 6 clasificaciones. Buscamos proteínas exclusivas comparando 5 proteínas comunes a los 5 PC derivados de pacientes para CAD, NO CAD y control. Nuestros resultados muestran que para cada nanopartícula utilizada en este estudio, varias proteínas específicas se identifican exclusivamente en patrones de PC específicos. Entre estas, varias proteínas implicadas en la regulación de la activación del complemento (proteína 3 relacionada con el factor H del complemento, componente C8 gamma del complemento) se identificaron exclusivamente en el patrón de PC de pacientes con CAD. Este resultado 10 está en consonancia con el papel recientemente descrito de la activación del complemento en la patogenia de las enfermedades cardiovasculares (desarrollo de aterosclerosis, rotura de placa y trombosis). Otro ejemplo está representado por la apolipoproteína (a), considerada un biomarcador atractivo candidato para su uso en la práctica clínica para CAD, que hemos detectado exclusivamente en el patrón de PC de pacientes con CAD. 15 La apolipoproteína (a) es el componente principal de la lipoproteína (a), se sabe que está sujeta a escisión proteolítica y sus fragmentos se acumulan en placas ateroscleróticas.

20 Confirmamos la precisión del enfoque en la discriminación de pacientes con CAD y NO CAD mediante el análisis de 9 muestras de plasma ciegas (3 por cada condición y 3 de control). La PC se formó aproximadamente las 6 nanopartículas de nuestro nanosistema de matriz de sensores de PC. Luego, las proteínas en las PC se identificaron mediante LC-MS/MS (Información de apoyo) y los resultados se analizaron utilizando los mismos enfoques de clasificación y agrupación.

25 El análisis estadístico y de datos se representa gráficamente en la FIG. 54 mostrando el discreto aislamiento de la clasificación de la CAD, NO CAD y CONTROL (sin riesgo de CAD).

30 En conclusión, hemos demostrado la precisión de una matriz de sensores de PC de 6 nanopartículas para la detección de CAD. El nanosistema de matriz de sensores de PC desarrollado en este trabajo demostró ser también sensible y preciso en la discriminación de CAD, reestenosis, NO CAD e individuos sanos, lo que demuestra aún más su gran valor potencial como plataforma tecnológica para ser utilizada en el entorno clínico. De hecho, a pesar de la presencia de muchos síntomas típicamente asociados con la placa aterosclerótica, el 35 grupo de pacientes NO CAD que se investiga en este trabajo no tenía ninguna obstrucción en sus arterias.

35 Los patrones de PC resultantes de esta plataforma representan una huella multivariada única, que es más precisa y de amplio espectro que las obtenidas usando las PC de una sola nanopartícula. El enfoque presentado aquí puede ser de gran valor para la detección no solo de CAD, sino también de otras diversas enfermedades humanas, mejorando la calidad de vida de muchos pacientes. Especialmente debido a la no invasividad de la prueba, prevemos que los pacientes utilizarían dicha prueba con más gusto y con mayor frecuencia que la angiografía. Y a diferencia de la angiografía, esta prueba es fácil de administrar y no tiene efectos secundarios, lo que permite a los pacientes verificar el estado de las arterias desde los primeros 40 síntomas y reducir significativamente las complicaciones de la CAD debido a la detección temprana.

Métodos

40 Nanopartículas. Se adquirieron tres partículas de sílice funcionalizadas de manera diferente de Kisker-Products (<https://www.kisker-biotech.com/>); se adquirieron tres partículas de poliestireno funcionalizadas de manera diferente de Polyscience, Inc. (<http://www.polysciences.com/>). Todas las partículas tenían el mismo tamaño (100 nm). Su morfología, tamaño promedio, índice de polidispersidad (PDI) y potencial zeta se caracterizaron mediante mediciones de TEM, DLS y potencial zeta.

45 Formación de corona proteica. Las PC se crearon incubando 0.5 mg de nanopartículas en H₂O desionizado con el mismo volumen de plasma humano. La incubación se realizó a 37 °C con agitación durante 1 hora. Inmediatamente después de la incubación, se centrifugó a 14 000 rpm y 10 °C durante 30 minutos para formar un sedimento. A continuación, el sedimento se lavó y se suspendió en 200 µl de solución salina regulada con fosfato (PBS) a 4 °C. Las medidas de centrifugación se repitieron tres veces en las mismas condiciones anteriores. El sedimento de las nanopartículas recubiertas de PC se resuspendió en urea 8 M, bicarbonato de amonio 50 mM para luego ejecutar geles SDS-PAGE y análisis LC/MSMS o en H₂O desionizado para luego 50 analizar el tamaño y el potencial ζ .

Caracterización de nanopartículas. El tamaño y el potencial ζ de nanopartículas desnudas y recubiertas con corona de proteína se han caracterizado diluyendo 10 µl de cada muestra en 1 ml total de agua destilada. Las mediciones se han realizado con un Zetasizer Nano ZS90 (Malvern, Reino Unido). Los valores de tamaño y carga superficial se dan como media \pm S.D. de tres mediciones independientes.

55 Ensayo de concentración de proteínas. La cantidad de proteínas dentro de la corona se determinó mediante el ensayo de Bradford (Bio-rad) usando albúmina de suero bovino a una concentración conocida como estándar para construir una curva estándar de 5 puntos ($R^2=0.99$). Las concentraciones de proteína se registran como un promedio de tres experimentos \pm S.D.

Geles 1D-SDS PAGE. Las proteínas de la corona se disolvieron en urea 8 M, bicarbonato de amonio 50 mM. Se añadió una cantidad igual de regulador Laemmli 2X al sedimento y se calentó durante 5 min a 90 °C antes de cargarlo y resolverlo en un 4-20% Mini-PROTEAN® TGX™ Precast Gels (Bio-Rad Laboratories, Hercules, CA) durante 1 h a 120 V. Las proteínas se tiñeron con azul brillante de Coomassie (Fisher Scientific, Fair Lawn, NJ, EE. UU.) durante la noche, seguido de un lavado extenso en agua ultrapura.

5 Espectrometría de masas/ Estadística/PLS-DA. La espectrometría de masas, el análisis estadístico y PLS-DA se realizaron como se describe en el Ejemplo 1.

Ejemplo 6. Una prueba de corona de proteína de nanopartículas múltiples para la detección de la enfermedad de Alzheimer en una etapa temprana

10 La patología de la enfermedad de Alzheimer comienza décadas antes de la detección de los síntomas clínicos. La necesidad de un diagnóstico temprano preciso y no invasivo para la enfermedad de Alzheimer está creciendo rápidamente. Para abordar el problema, se desarrolló un nanosistema de matriz de sensores de 15 última generación, que detecta con éxito cambios minúsculos en los patrones de proteínas plasmáticas y utiliza técnicas de agrupación para determinar la presencia de la enfermedad de Alzheimer. La tecnología desarrollada también se puede aplicar en el futuro para diagnosticar otras enfermedades, ya que la matriz de sensores crea huellas únicas para cada cambio del proteoma plasmático debido a enfermedades y es lo suficientemente sensible para capturar dichos cambios a diferencia de las tecnologías actuales.

20 Como se discutió en los Ejemplos anteriores, la composición de las PC cambia en función del proteoma del plasma que puede alterarse como consecuencia de una enfermedad, lo que da como resultado "corona de proteínas personalizadas" (PPC). Sin embargo, las PPC no pueden proporcionar una estrategia robusta y precisa para la detección temprana de enfermedades, principalmente debido a la gran superposición de 25 proteínas similares en la composición de la corona de proteínas. En este estudio, utilizando los PPC formados aproximadamente una plataforma multi-NP que contiene seis nanopartículas, proporcionamos patrones de huellas para una detección robusta y precisa de AD con precisión y especificidad de predicción sin precedentes. 25 La prueba de matriz de sensores desarrollada distinguió con éxito entre pacientes con y sin AD, así como pacientes que desarrollaron AD varios años después (usando plasmas de cohortes). Estos ejemplos proporcionan una alternativa factible y no invasiva a la detección actual de AD, lo que permite la capacidad de proporcionar una detección y un tratamiento temprano sin precedentes.

Resultados y discusión

30 La prueba de matriz de sensores consiste en incubar muestras de plasma con 6 nanopartículas diferentes. Estas últimas son nanopartículas de poliestireno y sílice de 100 nm, cada una con químicas de superficie ajustables (simple, -amino y -carboxilo conjugado, denominados en lo sucesivo P, P-NH₂, P-COOH, S, S-NH₂, 35 S-COOH) y distribución de tamaño estrecho como se describe en el Ejemplo 5 para análisis CAD. Como prueba de concepto inicial, para buscar diferencias entre AD y plasmas de control, se ha analizado el tamaño, la carga superficial y la morfología de las nanopartículas antes y después de la formación de PC. Se ha aplicado el análisis de seguimiento de nanopartículas (Nanosight) para caracterizar el tamaño de las nanopartículas. Durante tal análisis, un haz láser ilumina las nanopartículas, de las cuales se visualiza la luz dispersada a través de un microscopio óptico. Mientras tanto, una cámara alineada con el haz graba un video que muestra el movimiento de las nanopartículas (30-60 fotogramas/seg). Antes de la incubación en plasma, todas las 40 nanopartículas eran de tamaño homogéneo (Figura 58, nanopartículas de poliestireno 90-100 nm; nanopartículas de sílice 80-100 nm) con una carga superficial negativa consistente con las proporcionadas por el fabricante (Tabla 8). Después de 1 h de incubación a 37 °C bajo agitación en plasma, las nanopartículas recubiertas de PC se recuperaron mediante centrifugación seguida de lavados extensos para eliminar las 45 proteínas no unidas y poco unidas. En todos los casos, revelamos un aumento en el tamaño de las nanopartículas recubiertas de PC y una distribución de tamaño más amplia, lo que indica una población menos homogénea (Figura 54, gráfico de dispersión). El aumento promedio de tamaño fue de 30 nm, lo que indica un espesor de 15 nm de la capa de PC, lo cual es consistente con los datos reportados en la literatura ¹ y con lo 50 observado usando plasma de voluntarios sanos (datos no mostrados). La presencia de proteínas plasmáticas en la superficie de las nanopartículas indujo un cambio en su carga, que se hizo menos negativa alcanzando los valores típicos de las proteínas plasmáticas (-20 mV a -0 mV a Tabla 8).

Tabla 8: potenciales zeta de nanopartículas desnudas o después de entrar en contacto con plasma

| Potenciales Zeta (mV) | Nanopartículas desnudas | Plasma sano | Plasma AD |
|-----------------------|-------------------------|---------------|---------------|
| P | -42.27 ± 0.54 | -20.36 ± 1.54 | -26.56 ± 2.14 |
| P-NH ₂ | -27.57 ± 1.63 | -29.81 ± 1.55 | -30.62 ± 0.74 |
| P-COOH | -49.38 ± 1.1 | -34.87 ± 0.78 | -27.02 ± 0.71 |
| S | -53.72 ± 1.05 | -23.22 ± 1.55 | -21.23 ± 0.54 |

| | | | |
|-------------------|---------------|---------------|---------------|
| S-NH ₂ | -46.44 ± 0.86 | -24.26 ± 0.95 | -21.21 ± 0.38 |
| S-COOH | -56.77 ± 0.14 | -28.37 ± 0.87 | -25.69 ± 1.31 |

Se observaron ligeras diferencias en la carga superficial de la misma nanopartícula incubada con plasma AD y plasma sano (Tabla 8). Sin embargo, esas diferencias no eran sustanciales y, por lo tanto, no eran útiles en la discriminación entre diferentes condiciones de plasma. También se utilizó microscopía electrónica de transmisión (TEM) para evaluar la morfología de las nanopartículas, que permaneció sin cambios después de la formación de PC. De hecho, tanto antes como después de la incubación con plasma, todas las nanopartículas tenían una forma homogénea redonda (Figura 55). Se observó la presencia de una fina capa de proteínas, asociada a un aumento del tamaño de las nanopartículas recubiertas de PC, lo que confirma los resultados obtenidos por Nanosight (Figura 55).

- 10 Las PC asociadas con las partículas se resolvieron mediante electroforesis en gel (SDS-PAGE) y luego se visualizaron mediante tinción con azul brillante de Coomassie (Figura 56). En general, al evaluar visualmente los carriles, las 6 nanopartículas tenían diferentes patrones de PC, que era lo que se esperaba y deseaba para obtener los mejores resultados de nuestro enfoque. Las nanopartículas de sílice parecían adsorber menos proteínas en su superficie que las nanopartículas de poliestireno. En la mayoría de las muestras, la PC de poliestireno simple estaba más enriquecida que otras en proteínas con un peso molecular de ≈60 kDa, que muy probablemente se atribuyen a la albúmina. Esto fue cierto tanto para el plasma de pacientes con AD como para el plasma de control (Figura 56, flechas). El análisis densitométrico de bandas asociadas a cada PC confirmó que las nanopartículas de sílice generalmente adsorbieron menor cantidad de proteínas, pero también reveló que un mayor número de proteínas constituyen sus PC: como por ejemplo, más de una banda está presente a nivel de albúmina para nanopartículas de PC de poliestireno (Figura 56, flechas). Por otro lado, se registraron varias diferencias en los perfiles de PPC de pacientes con AD e individuos sanos, particularmente en el caso de las nanopartículas de sílice (Figura 56, flechas). Sin embargo, esas diferencias no fueron estadísticamente significativas y no fueron suficientes para discriminar con precisión entre los dos grupos de pacientes bajo investigación. Para profundizar en los PCs y conocer la identidad exacta y la cantidad de proteínas que componen los diferentes PCs, todas las muestras han sido analizadas por espectrometría de masas. Se utilizó el análisis sin etiquetas de conteo espectral, ampliamente utilizado para el perfilado cuantitativo de las PC alrededor de las nanopartículas 2-4, para determinar la contribución porcentual de cada proteína en las PC. Este cálculo se realizó seis veces para cada paciente de plasma (incubado por separado con cada nanopartícula), lo que permitió la formación de un perfil de PC específico de AD derivado de la combinación de la contribución de las PC relacionadas con seis nanopartículas. La Tabla 9 describe la población de pacientes utilizada.

Tabla 9: Información de los pacientes en los que se utilizaron sus plasmas en este ejemplo.

| Enquista | Paciente n. | ID del caso | Edad en la escisión | Sexo | Altura (cm) | Peso (kg) | IRaC | Año de escisión | Diagnóstico Clínico (Pacientes) |
|----------|-------------|-------------|---------------------|------|-------------|-----------|-------|-----------------|--|
| A | 1 | 104688 | 80 | F | 167 | 70 | 28.4 | 2016 | Histerectomía (1980); Enfermedad de Alzheimer (2015) |
| B | 2 | 104683 | 87 | M | 165 | 48 | 18 | 2016 | Enfermedad de Alzheimer (2016) |
| C | 3 | 104684 | 79 | F | 160 | 60 | 23.44 | 2016 | enfermedad de alzheimer (2016) |
| D | 4 | 104682 | 87 | F | | | | 2016 | Enfermedad de Alzheimer (2014) |
| E | 5 | 104686 | 83 | M | 166 | 82 | 28.4 | 2016 | Enfermedad de Alzheimer (2016) |
| F | 6 | 104689 | 86 | F | 160 | 72 | 28.13 | 2016 | Enfermedad de Alzheimer (2015) |
| G | 7 | 104687 | 77 | M | 173 | 72 | 28.98 | 2016 | Enfermedad de Alzheimer (2016) |
| H | 8 | 104658 | 70 | M | 173 | 70 | 28.38 | 2016 | Enfermedad de Alzheimer (2016) |
| I | 9 | 104680 | 85 | M | 168 | 72 | 29.37 | 2016 | Enfermedad de Alzheimer, demencia de Alzheimer (2016) |
| K | 10 | 104657 | 73 | M | 162 | 68 | 28.36 | 2016 | Enfermedad de Alzheimer (2008); Radiculopatía, Estenosis de la arteria carótida |
| L | 11 | 66393 | 88 | F | 170 | 68 | 23.88 | 2008 | Fibrilación auricular (1999); Hipodrópsico, Encefalitis viral (1984); Fractura del pie izquierdo (1981); Fractura del anillo pérvico (1978); Histerectomía (1989); Catarata (1999); Enfermedad de Alzheimer (2009) |

Conclusión

- Este trabajo representa un estudio de prueba de concepto para el desarrollo de un análisis de sangre MNPC para el diagnóstico de la AD. La prueba MNPC mostró una precisión y especificidad de predicción sin precedentes. Si bien las pruebas de sangre de biomarcadores individuales a menudo se asocian con falsos

positivos y, por lo tanto, requieren más análisis para confirmar el diagnóstico, nuestro enfoque registra las interacciones de todas las proteínas de alta afinidad con el conjunto de 6 nanopartículas, lo que permite la creación de una huella de PC de alta especificidad de una determinada enfermedad. Las nanopartículas actúan como nano-concentrador de proteínas plasmáticas en su superficie (cada nanopartícula concentra las proteínas plasmáticas con mayor afinidad hacia su superficie): esto ayuda en la revelación de proteínas cuyos niveles solo cambian levemente en condiciones patológicas respecto a un estado saludable. Estos pequeños cambios pueden pertenecer tanto a proteínas plasmáticas de alta abundancia como de baja abundancia. En este estudio, informamos un número relativamente pequeño de pacientes porque el plasma de cada paciente se analizó 6 veces con diferentes nanopartículas, por lo que los resultados obtenidos para cada paciente son 6 veces más específicos que los obtenidos cuando se realiza un solo análisis.

Materiales y métodos

Nanopartículas. Partículas de sílice (simple, amino y carboxilo conjugado) fueron adquiridas de Kisker-Products (<https://www.kisker-biotech.com/>). Partículas de poliestireno (simple, conjugado con amino y carboxilo) fueron adquiridas de Polyscience, Inc. (<http://www.polysciences.com/>). De acuerdo con el fabricante, todas las partículas tenían el mismo tamaño (90-100 nm). Su morfología, tamaño promedio y potencial zeta se caracterizaron como se describe más adelante en esta sección.

Formación de corona de proteína personalizada.

Las PC se crearon incubando 0.5 mg de nanopartículas en H₂O desionizado con el mismo volumen de plasma humano. La incubación se realizó a 37 °C con agitación durante 1 hora. Inmediatamente después de la incubación, las nanopartículas recubiertas de PC se recuperaron mediante centrifugación (14,000 rpm y 10 °C durante 30 minutos) y un lavado exhaustivo en solución salina regulada con fosfato frío (PBS) para eliminar las proteínas no unidas o débilmente unidas. El sedimento de nanopartículas recubiertas de PC se resuspendió en urea 8 M, bicarbonato de amonio 50 mM para geles SDS-PAGE y análisis LC/MSMS o en H₂O desionizado para dispersión de luz dinámica y análisis de potencial ζ .

Caracterización fisicoquímica de nanopartículas. El tamaño se midió mediante análisis de seguimiento de nanopartículas (Nanosight, Malvern, Reino Unido). El software calcula el tamaño de acuerdo con el potencial ζ de nanopartículas desnudas y recubiertas de PC que se ha determinado utilizando un Zetasizer Nano ZS90 (Malvern). las nanopartículas se diluyeron en agua antes del análisis a una concentración de 50 μ g/ml. Los valores de tamaño y carga superficial se dan como media \pm S.D. de tres mediciones independientes. Para el análisis por microscopía electrónica de transmisión (TEM), las muestras y las rejillas se han etiquetado con acetato de uranilo al 1 %. Se utilizó el microscopio electrónico de transmisión Tecnai G2 Spirit BiotWIN equipado con una cámara CCD AMT 2k.

Electroforesis en gel unidimensional. Las coronas de proteínas personalizadas se disolvieron en urea 8 M, bicarbonato de amonio 50 mM. Se añadió una cantidad igual de regulador Laemmli 2X al sedimento y se calentó durante 5 min a 90 °C antes de cargarlo y resolverlo en un 4-20% Mini-PROTEAN® TGX™ Precast Gels (Bio-Rad Laboratories, Hercules, CA) durante 1 h a 120 V. Las proteínas se tiñeron con azul brillante de Coomassie (Fisher Scientific, Fair Lawn, NJ, EE. UU.) durante la noche, seguido de un lavado extenso en agua ultrapura. El análisis densitométrico de las intensidades de las bandas ha sido realizado por ImageJ (sitio web: imagej.nih.gov/ij/).

Identificación y cuantificación de proteínas por espectrometría de masas. Las proteínas se redujeron con ditioltreitol 10 mM (Sigma) durante 1 hora a 56 °C y luego se alquilaron con yodoacetamida 55 mM (Sigma-Aldrich, St Louis, MO, EE. UU.) durante 1 hora a 25 °C en la oscuridad. A continuación, las proteínas se digirieron con tripsina modificada (Promega, Madison, WI, EE. UU.) a una relación enzima/sustrato de 1:50 en acetato de amonio 100 mM, pH 8.9 a 25 °C durante la noche. La actividad de la tripsina se detuvo mediante la adición de ácido acético (99.9 %, Sigma-Aldrich) hasta una concentración final del 5 %. Los péptidos se desalinizaron utilizando C18 SpinTips (Protea, Morgantown, WV), luego se centrifugaron al vacío y se almacenaron a -80 °C hasta el día del análisis. Los péptidos se separaron mediante HPLC de fase inversa (Thermo Fisher, Waltham, MA Easy nLC1000) utilizando una precolumna (fabricada internamente, 6 cm de 10 μ m C18) y una columna analítica de punta de 5 μ m autoenvasada (12 cm de 5 μ m C18, New Objective) durante un gradiente de 140 minutos antes de la nanolectroaspersión con un espectrómetro de masas QExactive (Thermo Fisher). El disolvente A era ácido fórmico al 0.1 % y el disolvente B era MeCN al 80 %/ácido fórmico al 0.1 %. Las condiciones de gradiente fueron 2-10 % B (0-3 min), 10-30 % B (3-107 min), 30-40 % B (107-121 min), 40-60 % B (121-126 min.), 60-100 % B (126-127 min), 100 % B (127-137 min), 100-0 % B (137-138 min), 0 % B (138-140 min) y el espectrómetro de masas fue operado en un modo dependiente de datos. Los parámetros para el MS de escaneo completo fueron: resolución de 70,000 a través de 350-2000 m/z, AGC 3e6 e IT máximo de 50 ms. El escaneo completo de MS fue seguido por MS/MS para los 10 iones precursores principales en cada ciclo con un NCE de 28 y una exclusión dinámica de 30 s. Los archivos de datos espectrales de masas sin procesar (.raw) se buscaron utilizando Proteome Discoverer (Thermo Fisher) y Mascot versión 2.4.1 (Matrix Science). Los parámetros de la búsqueda Mascot fueron: Tolerancia a la masa de 10 ppm para iones precursores; unidades de 15 milimasa (mmu) para tolerancia a la masa de iones en fragmentos; 2

escisiones perdidas de tripsina; la modificación fija fue la carbamidometilación de la cisteína; modificaciones variables fueron la oxidación de metionina. Solo se incluyeron en el análisis de datos los péptidos con una puntuación de Mascot ≥ 25 . El recuento espectral se realizó sumando el número total de péptidos seleccionados para la fragmentación de cada proteína.

5 **Ejemplo 7. El tamaño de las partículas afecta la cantidad de proteína unida**

Este ejemplo demuestra que el uso de nanopartículas de diferentes tamaños hechas del mismo material proporciona una huella de biomolécula diferente para cada tamaño. Se incubaron nanopartículas de sílice de tres diámetros diferentes (100 nm (0.1 μm), 3 μm y 4 μm) con la misma muestra de plasma. Las perlas se analizaron por SDS-PAGE. Como se demuestra en la Figura 58, cuanto más grandes son las perlas, más 10 proteínas se incluyen en la firma biomolécula. Además, cada tamaño de perla tiene una firma de corona de biomolécula distinta y, por lo tanto, la combinación de diferentes tamaños de nanopartículas solo permite el desarrollo de una huella de biomolécula distinta. Como se muestra, existen claras diferencias en el patrón o las proteínas entre las tres perlas de diferentes tamaños.

15 Se puede hacer una matriz de sensores utilizando nanopartículas de diferentes tamaños, cada una de las cuales dará una huella de biomolécula diferente.

Ejemplo 8. La matriz de sensores puede proporcionar una huella de biomolécula que comprenda ácidos nucleicos

La matriz de sensores y la firma de corona de biomolécula asociada junto con las proteínas contienen ácidos nucleicos que forman la corona de biomolécula. Este ejemplo demuestra la composición de los ácidos nucleicos que se unen a nanopartículas de sílice con nanopartículas de superficie neutra. Las nanopartículas se 20 incubaron con plasma y se analizaron los ácidos nucleicos asociados (FIGURA 59). Las FIGURAS 60 y 61 representan el análisis de ácido nucleico en todas las muestras y su contenido en plasma (la cantidad de ácido nucleico en plasma fue de 33.8 pg/ μl). Se analizó el contenido de ácido nucleico de las coronas de biomoléculas 25 asociadas a la nanopartícula. Las proteínas se disociaron de nanopartículas utilizando urea y los ácidos nucleicos se analizaron posteriormente (Figura 62, cantidad de ácido nucleico 14.0) o las proteínas no se disociaron de la corona y los ácidos nucleicos se analizaron (Figura 63, 14.4 pg/ μl). Alternativamente, las nanopartículas se pueden incubar con ácidos nucleicos que se han purificado del plasma con un kit de plasma y luego se incuban con las partículas desnudas (Figura 64, cantidad de ácido nucleico 13.2 pg/ μl).

30 Los resultados revelaron la capacidad de las nanopartículas en la adsorción de ácido nucleico en la composición de la corona biomolecular.

REIVINDICACIONES

1. Un método para detectar un estado de enfermedad en un sujeto, que comprende:
 - (a) proporcionar una muestra biológica obtenida del sujeto;
 - (b) poner en contacto la muestra con una matriz de sensores, en el que la matriz de sensores comprende una pluralidad de elementos sensores que difieren entre sí en al menos una propiedad fisicoquímica; y en el que cada elemento sensor se une a una pluralidad de biomoléculas en la muestra para producir una pluralidad de firmas de corona de biomoléculas que forman firmas de corona de biomoléculas distintas entre sí con base en la propiedad fisicoquímica de cada elemento sensor; y
 - (c) determinar una composición combinada de una primera firma de corona de biomoléculas de la pluralidad de firmas de corona de biomoléculas y una segunda firma de corona de biomoléculas de la pluralidad de firmas de corona de biomoléculas para producir una huella de biomoléculas asociada con la muestra, en el que la huella de biomoléculas diferencia el estado de enfermedad del sujeto.
2. El método de la reivindicación 1, en el que el estado de enfermedad se selecciona entre no tener enfermedad o trastorno, tener un precursor de una enfermedad o trastorno y tener enfermedad o trastorno.
3. El método de la reivindicación 1 o 2, en el que el paso (c) comprende ensayar la primera firma de corona de biomolécula para identificar una abundancia de una primera biomolécula de la primera firma de corona de biomolécula y ensayar la segunda firma de corona de biomolécula para identificar una abundancia de una segunda biomolécula de la segunda firma de corona de biomolécula.
4. El método de la reivindicación 1 o 2, en el que el paso (c) comprende ensayar una firma de corona de biomolécula de cada elemento sensor para identificar un monto, nivel, cantidad o estado conformacional de una biomolécula en cada firma de corona de biomolécula.
5. El método de una cualquiera de las reivindicaciones 1-4, en el que el paso (c) comprende separar la pluralidad de biomoléculas de cada corona de biomolécula de cada elemento sensor y ensayar la pluralidad de biomoléculas para determinar la composición de biomolécula de la corona de biomolécula para determinar la huella de biomolécula.
6. El método de una cualquiera de las reivindicaciones 1-5, en el que el método comprende además comparar la huella de biomolécula de la muestra con un panel de huellas de biomolécula asociada con una pluralidad de estados de enfermedad para determinar qué estado de enfermedad está asociado con la muestra.
7. El método de una cualquiera de las reivindicaciones 1-6, en el que la pluralidad de elementos sensores comprende una pluralidad de partículas.
8. El método de la reivindicación 7, en el que la pluralidad de partículas comprende nanopartículas o micropartículas.
9. El método de una cualquiera de las reivindicaciones 1-8, en el que la pluralidad de elementos sensores comprende una pluralidad de cargas diferentes.
10. El método de una cualquiera de las reivindicaciones 1-9, en el que la pluralidad de elementos sensores se fabrican mediante impresión 3D o 4D.
11. El método de una cualquiera de las reivindicaciones 1-10, en el que la al menos una propiedad fisicoquímica se selecciona del grupo que consiste en composición, tamaño, carga superficial, hidrofobicidad, hidrofilicidad, funcionalidad superficial, topografía superficial y forma.
12. El método de la reivindicación 11, en el que la al menos una propiedad fisicoquímica comprende funcionalidad superficial.
13. El método de la reivindicación 12, en el que las funcionalidades superficiales se seleccionan de aminopropilo, amina, ácido borónico, ácido carboxílico, metilo, éster N-succinimidilo, polietilenglicol, estreptavidina, éter metílico, trietoxipropilaminosilano, tiol, citrato, ácido lipoico, polietilenimina, hidroxilo y cualquier combinación de los mismos.
14. El método de la reivindicación 13, en el que las funcionalidades superficiales son amina y ácido carboxílico.
15. El método de una cualquiera de las reivindicaciones 1-14, en el que la pluralidad de elementos sensores están inmovilizados sobre un sustrato.

FIG. 1

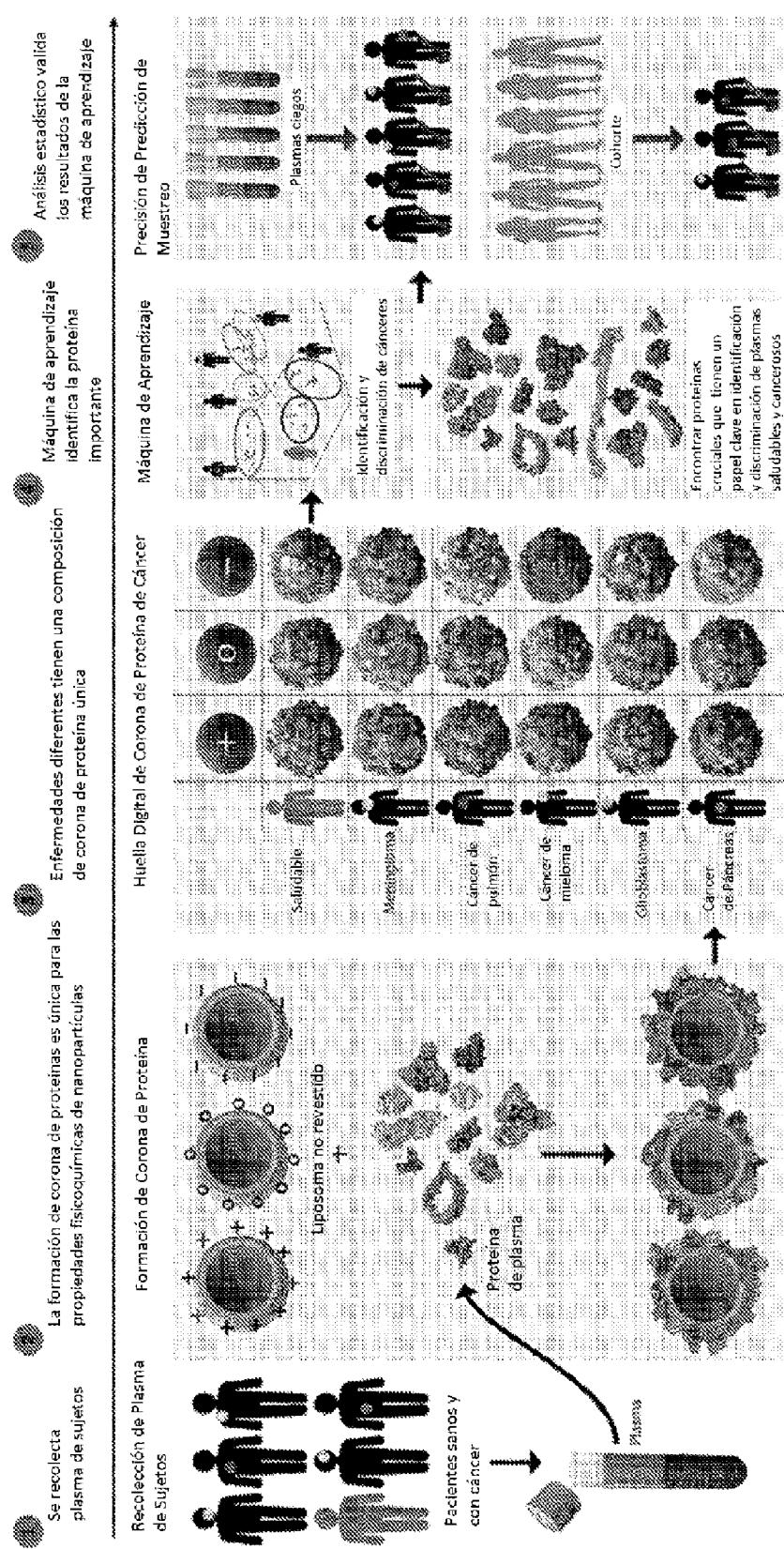


FIG. 2A

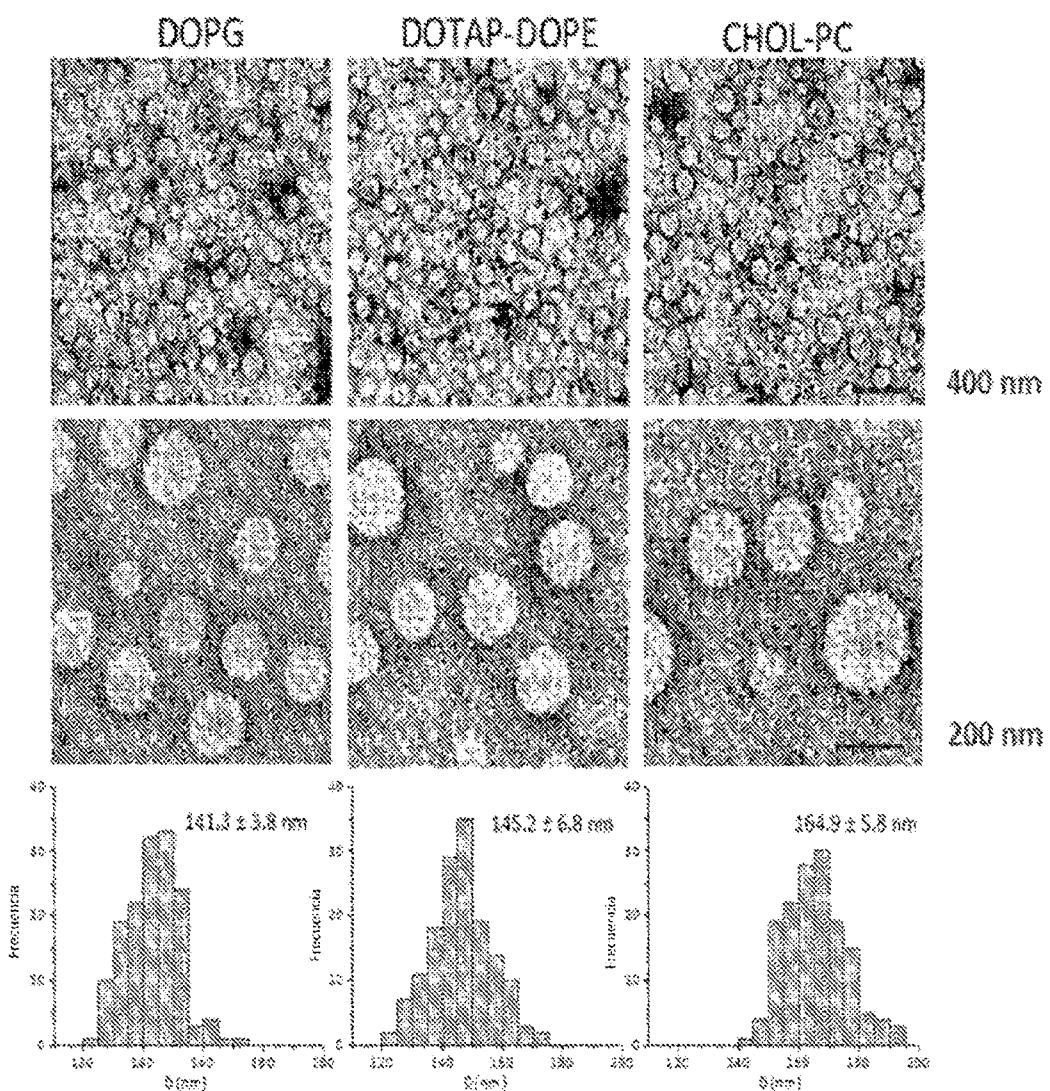
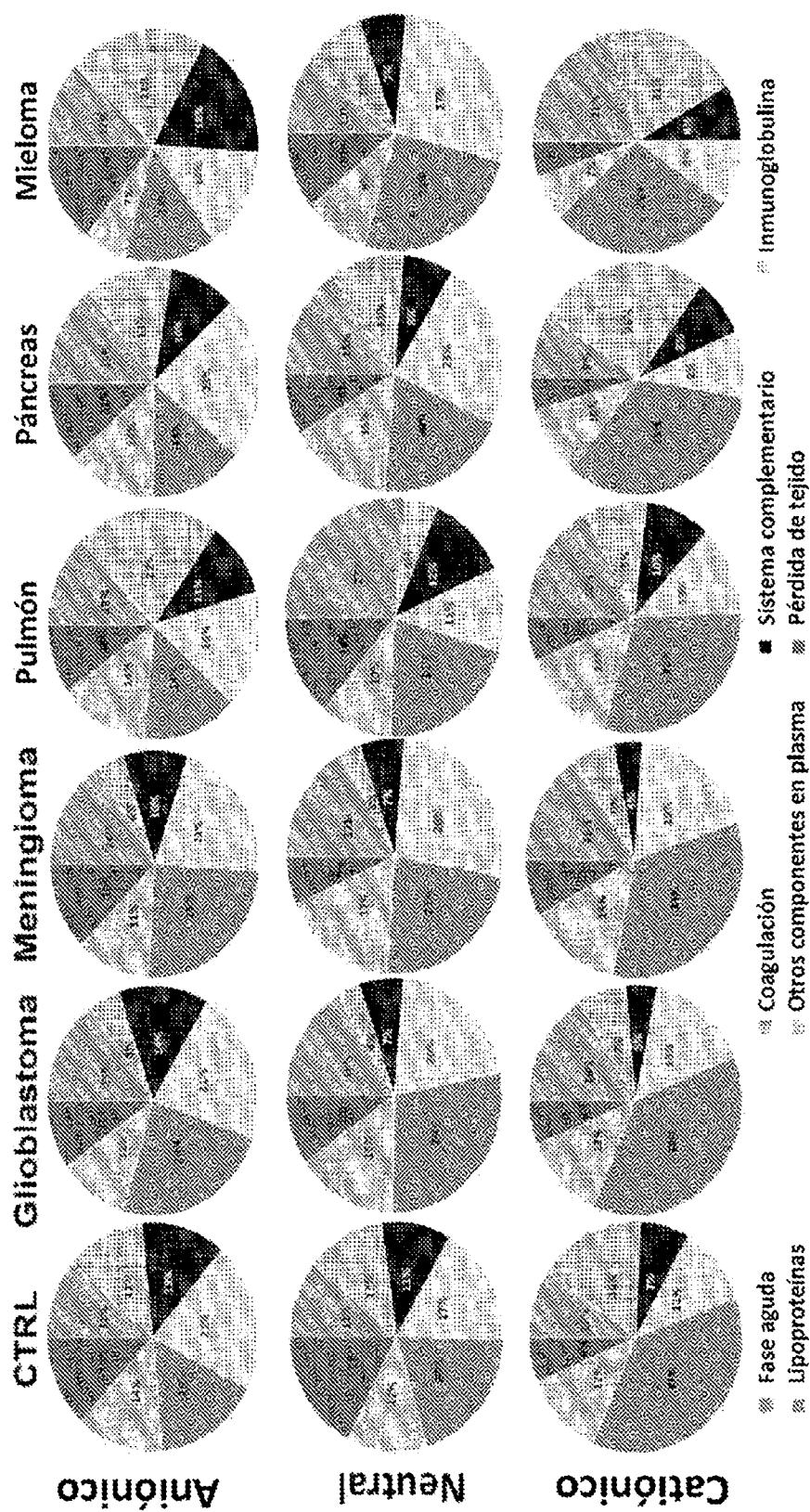


FIG. 2B

| Liposomas desnudos | | | | |
|----------------------------------|---------------------------|-------------|---------------------|-------------------------|
| | Intensidad de tamaño (nm) | PDI | Potencial Zeta (mV) | Diámetro TFM (nm) |
| Catiónico | 153.59211.40 | 0.173934 | 46.50±1.30 | 149.21±5.8 |
| Aniónico | 146.43±15 | 0.11310.076 | -48.13±4.54 | 141.31±3.8 |
| Neutral | 172.91±8 | 0.0510.001 | -3.51±1.2 | 164.9±5.8 |
| Liposomas Recubiertos con Corona | | | | |
| Catiónico | | | | |
| | Intensidad de tamaño (nm) | PDI | Zeta-Potencial (mV) | Ensayo de proteína (ug) |
| Saludable | 232.97±9.44 | 0.3910.00 | -48.13±1.33 | 3.88±0.33 |
| Páncreas | 240.57±4.20 | 0.3510.02 | -48.33±1.32 | 6.84±0.69 |
| Pulmón | 200.47±5.52 | 0.3610.03 | -33.27±1.03 | 6.02±0.70 |
| Glioblastoma | 233.30±7.87 | 0.3810.05 | -40.70±1.44 | 6.81±0.68 |
| Meningioma | 233.27±4.22 | 0.3410.06 | -47.27±0.45 | 6.42±0.78 |
| Mieloma | 238.42±8.50 | 0.4310.09 | -9.38±0.37 | 8.02±0.77 |
| Aniónico | | | | |
| Saludable | 172.70±1.82 | 0.4000.002 | -42.70±1.31 | 7.42±0.86 |
| Páncreas | 235.23±4.37 | 0.36210.066 | -35.63±0.72 | 6.65±0.45 |
| Pulmón | 192.37±4.57 | 0.2900.028 | -34.43±0.75 | 6.87±0.31 |
| Glioblastoma | 203.50±8.57 | 0.47510.047 | -38.50±0.82 | 6.48±0.95 |
| Meningioma | 198.60±4.30 | 0.33710.031 | -42.83±0.89 | 6.29±0.69 |
| Mieloma | 182.50±4.20 | 0.26510.010 | -22.88±3.29 | 8.32±0.36 |
| Neutral | | | | |
| Saludable | 251.40±7.50 | 0.37810.003 | -28.45±0.34 | 3.38±0.18 |
| Páncreas | 268.50±3.87 | 0.37010.003 | -31.36±0.09 | 4.13±0.23 |
| Pulmón | 238.63±8.15 | 0.37610.074 | -30.81±1.36 | 4.38±0.40 |
| Glioblastoma | 233.87±3.87 | 0.29710.040 | -38.00±0.34 | 3.07±0.28 |
| Meningioma | 228.97±9.69 | 0.41310.074 | -32.55±3.03 | 2.99±0.32 |
| Mieloma | 233.27±3.82 | 0.32710.024 | -6.81±0.41 | 3.38±1.37 |

FIG. 2C



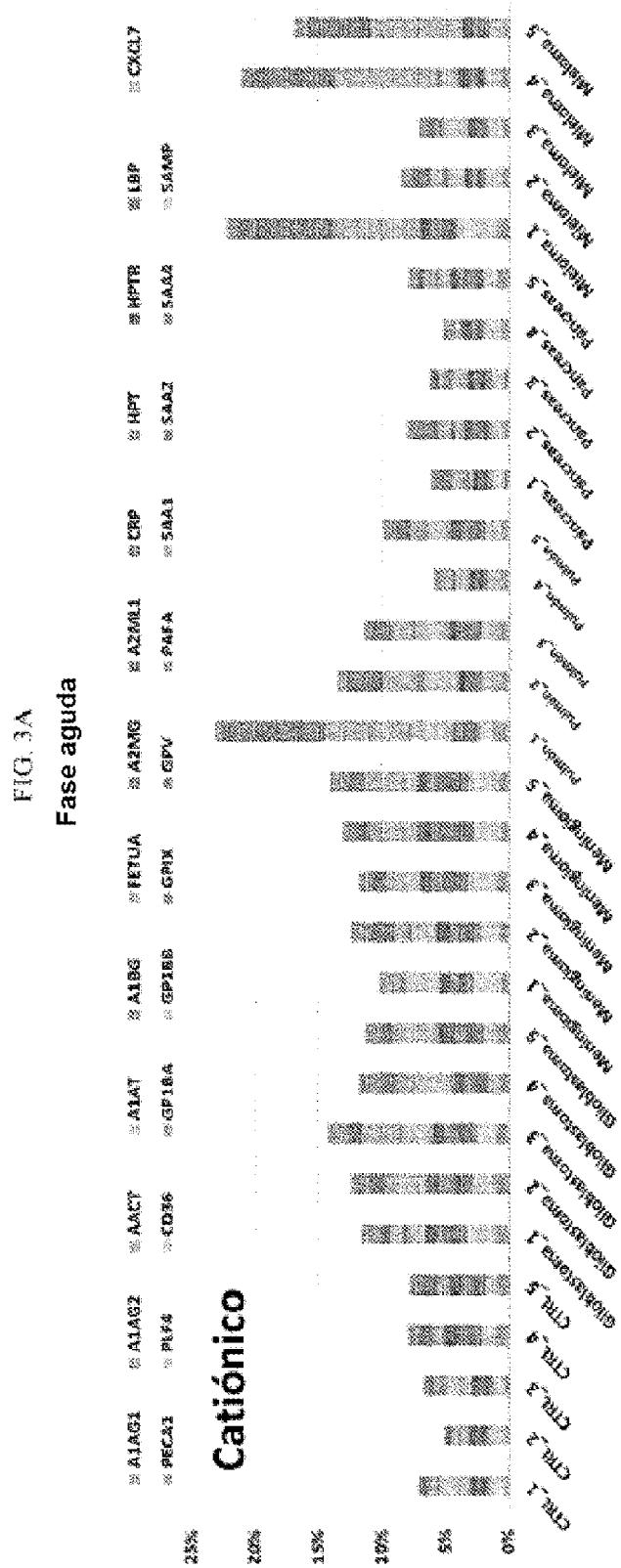


FIG. 3A CONTINUACIÓN



FIG. 3A CONTINUACIÓN

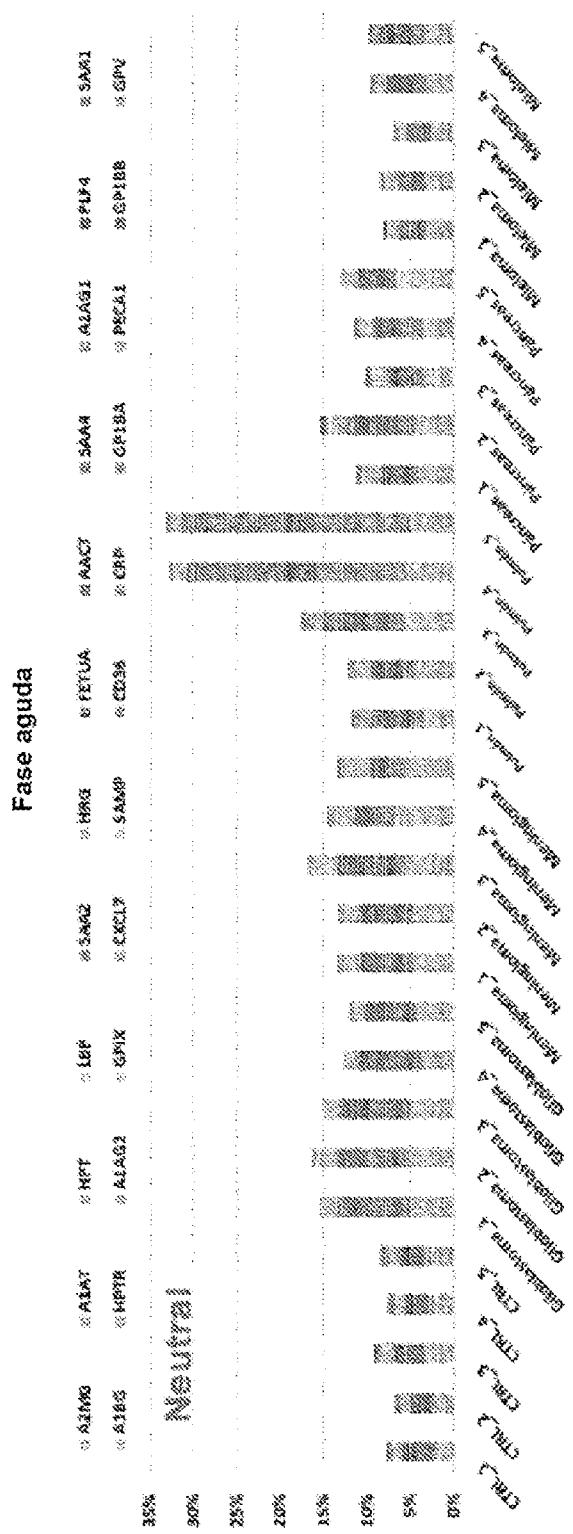
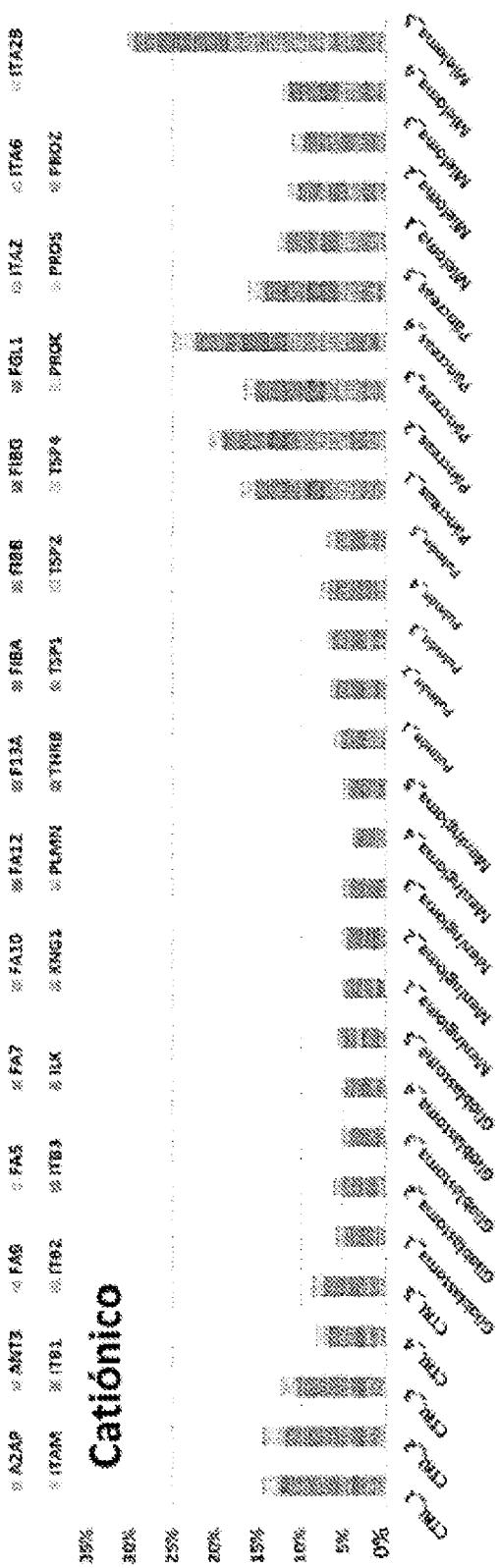


FIG. 3B
COAGULACIÓ



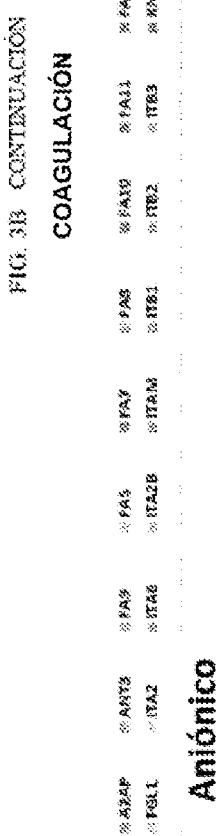


FIG. 3B - CONTINUACIÓN
COAGULACIÓN

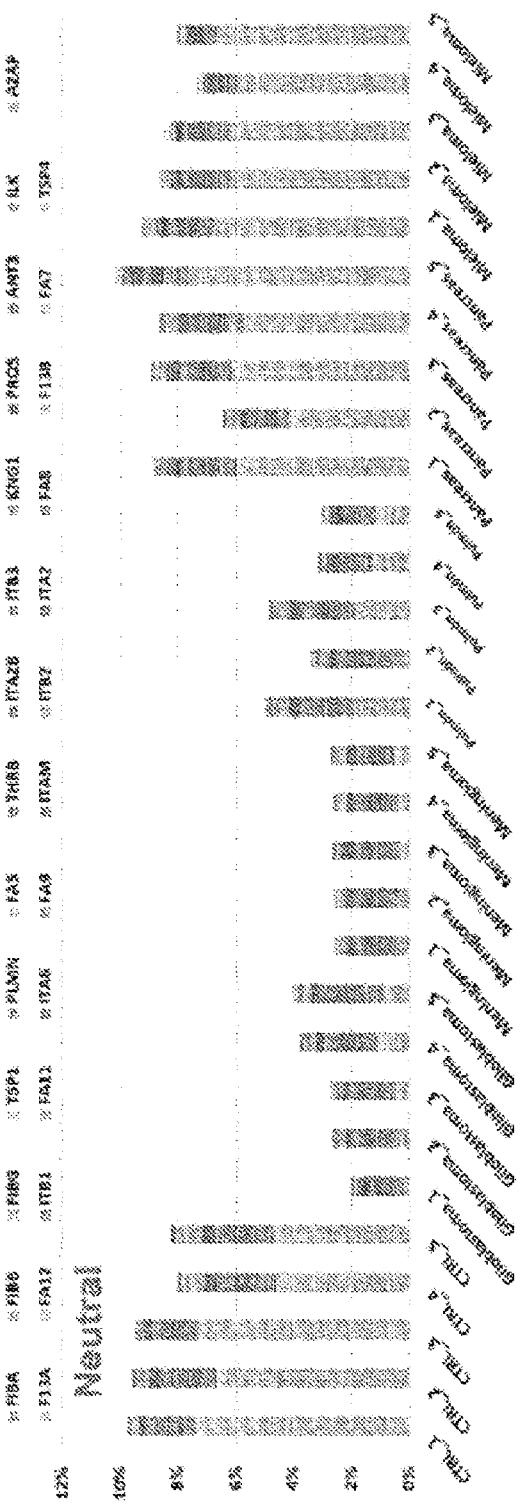


FIG. 3C
IMMUNOGLOBULINA

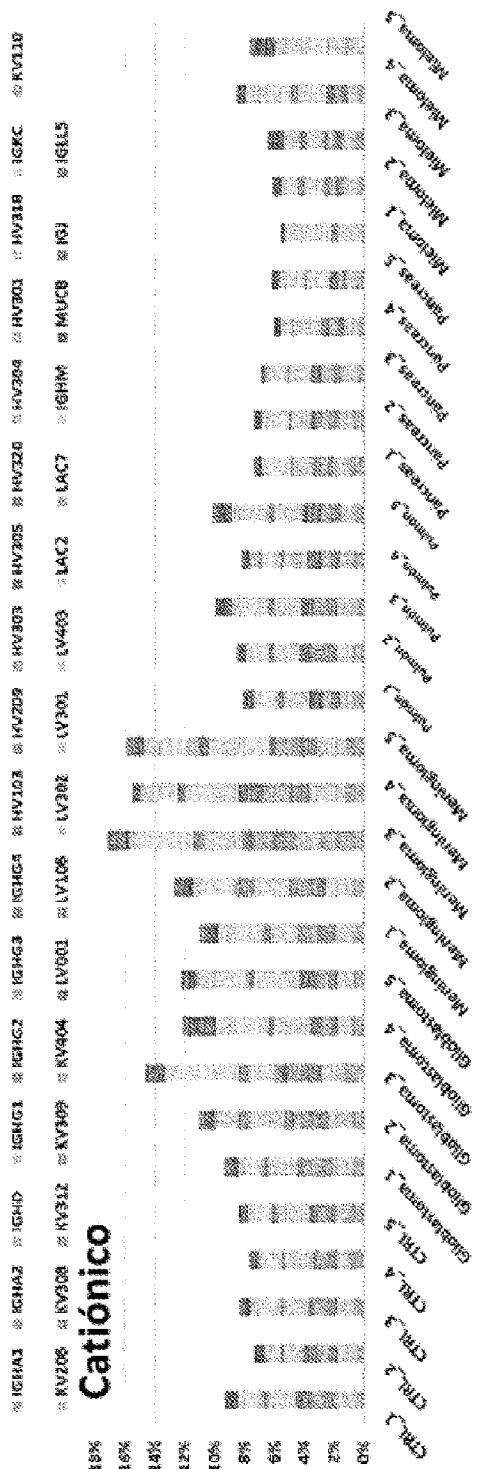


FIG. 3C. CONTENCIÓN

IMMUNOGLOBULINAS

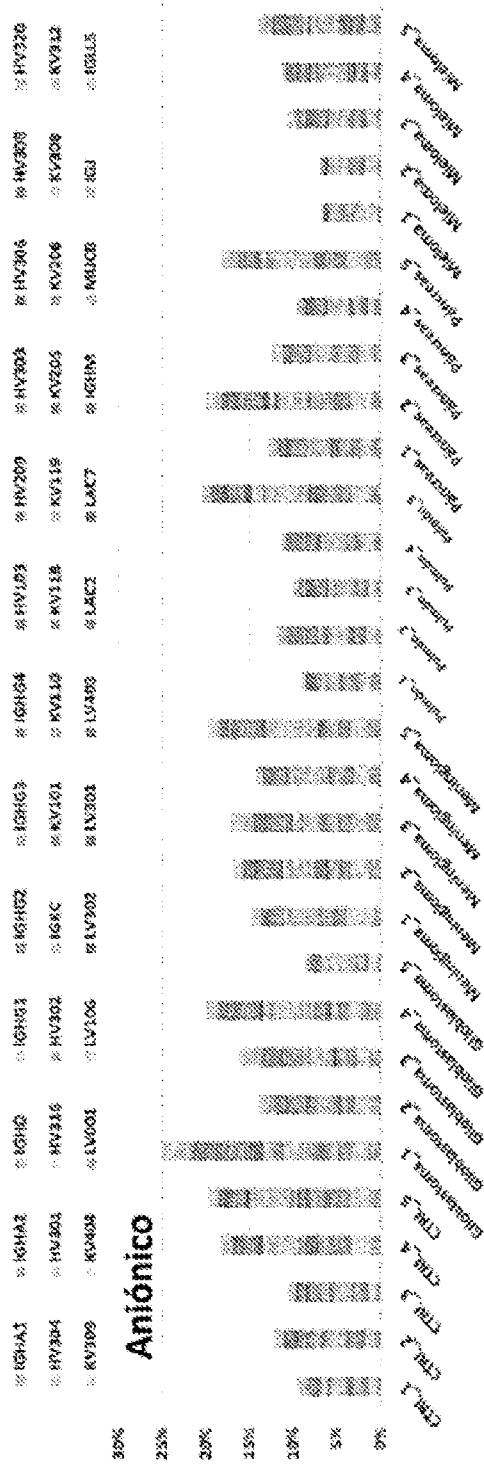
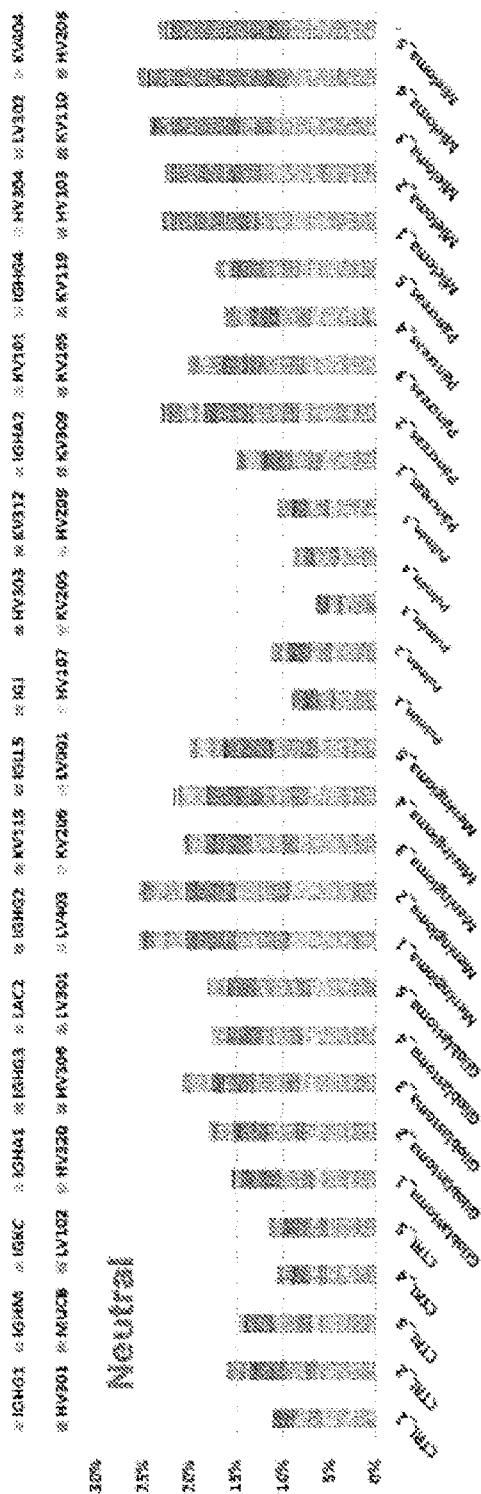


FIG 3C CONTINUACIÓN

INMUNOGLOBULINAS



三

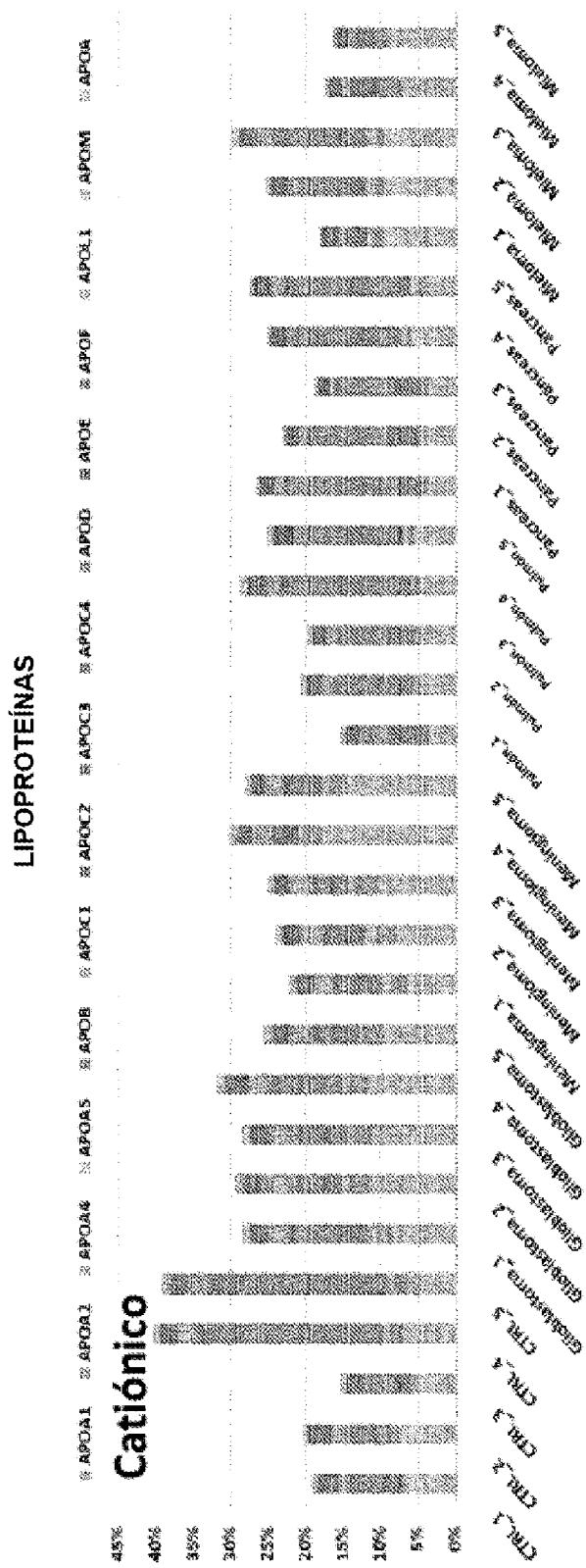


FIG. 3D CONTINUACIÓN
LIPOPROTEÍNAS

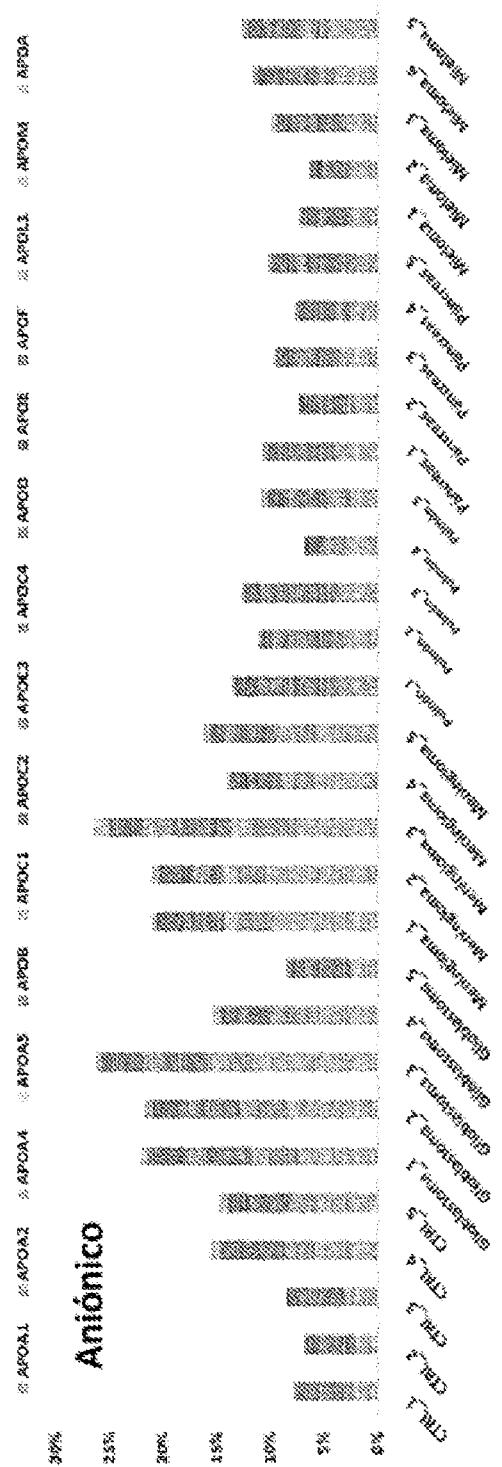
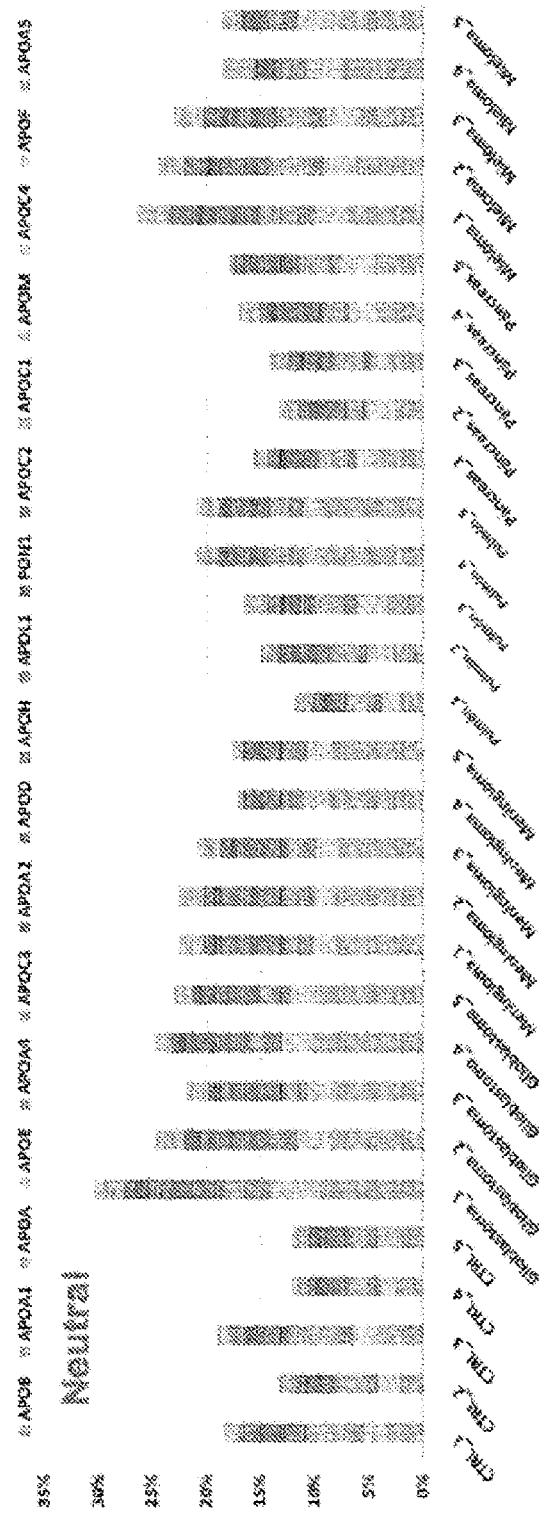


FIG. 3D CONTENCIACIÓN
LIPOPROTEÍNAS



610

ENLACE TISULAR

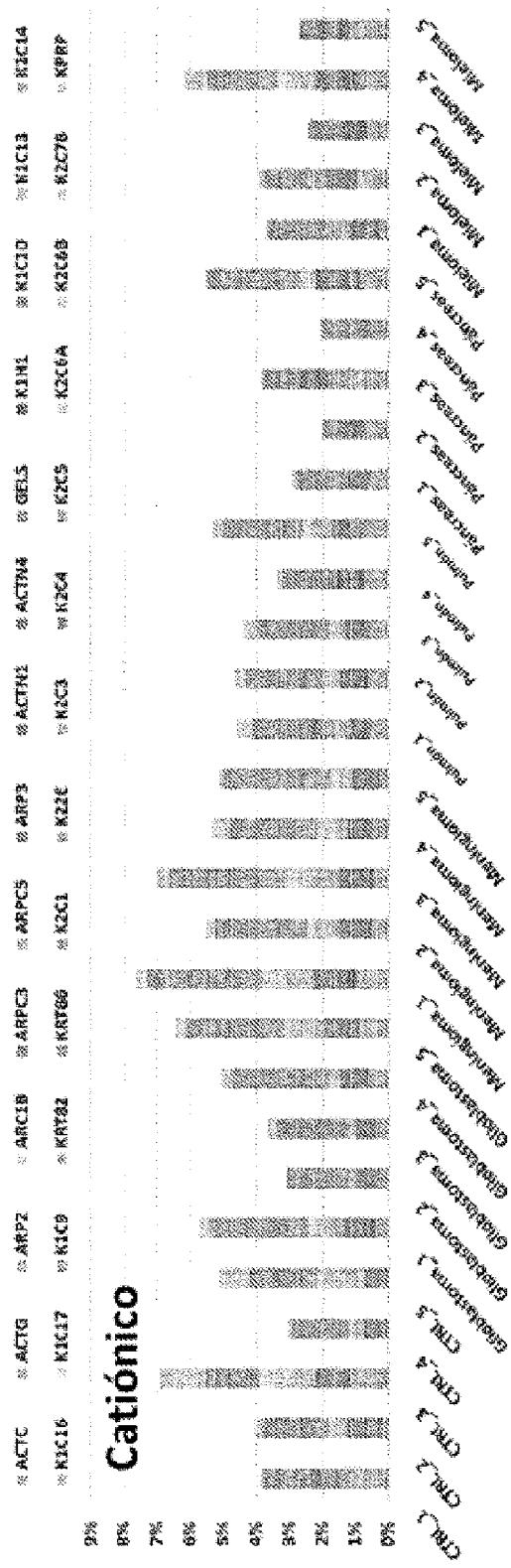
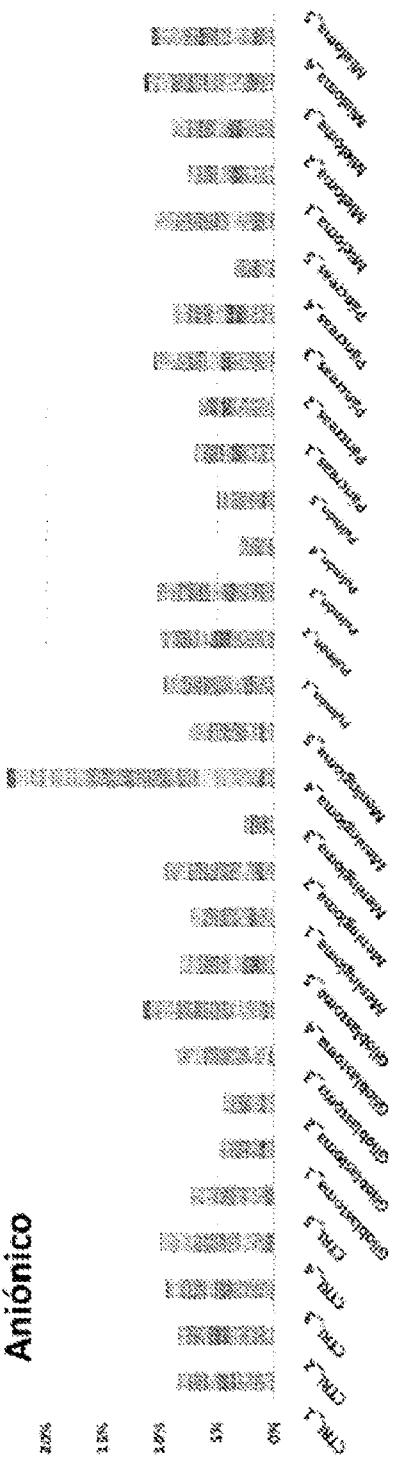


FIG. 3E CONTINUACIÓN

ENLACE TISULAR



Aniónico

FIG. 3E CONTINUACIÓN ENLACE TISULAR

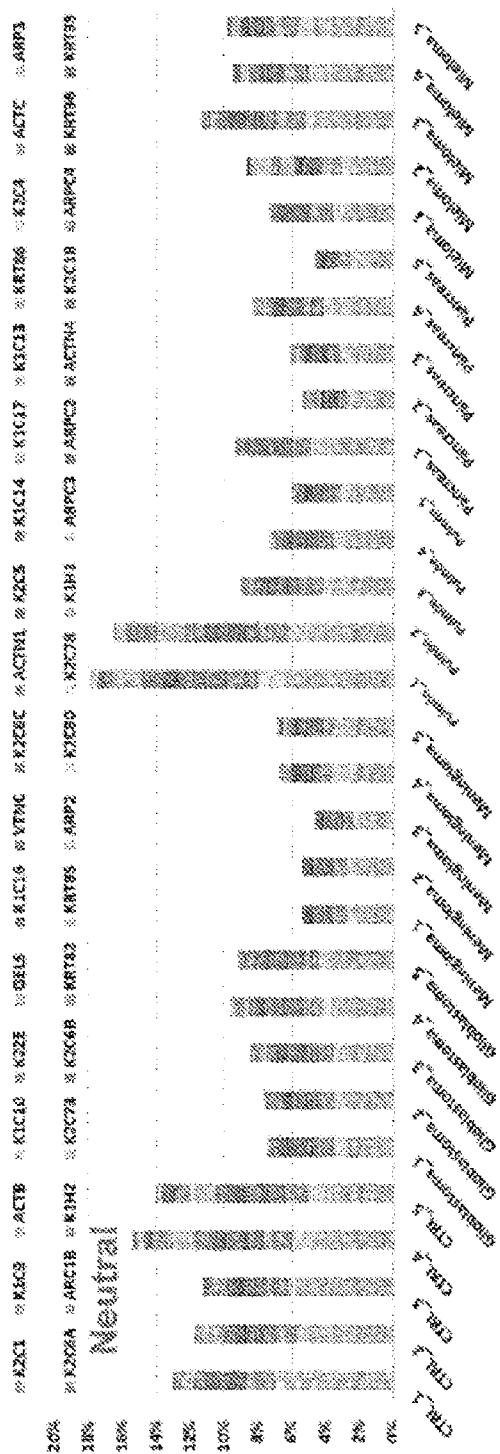
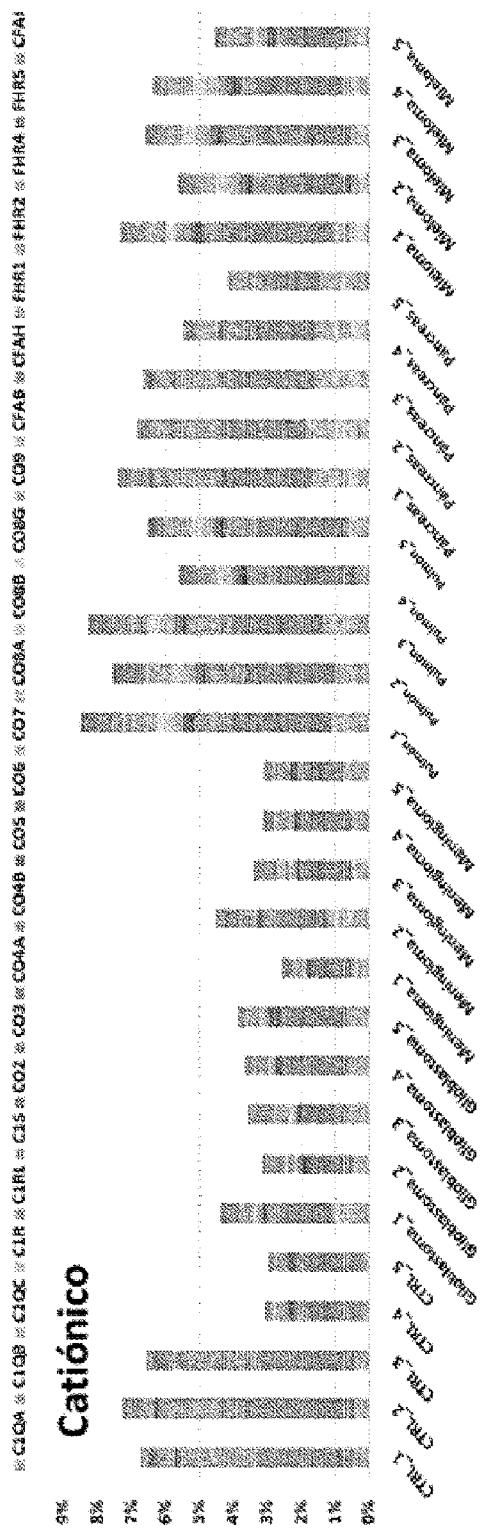


FIG. 31 COMPLEMENTO



COMPLEMENTO

FIG. 3. *CONTINUACIÓN*

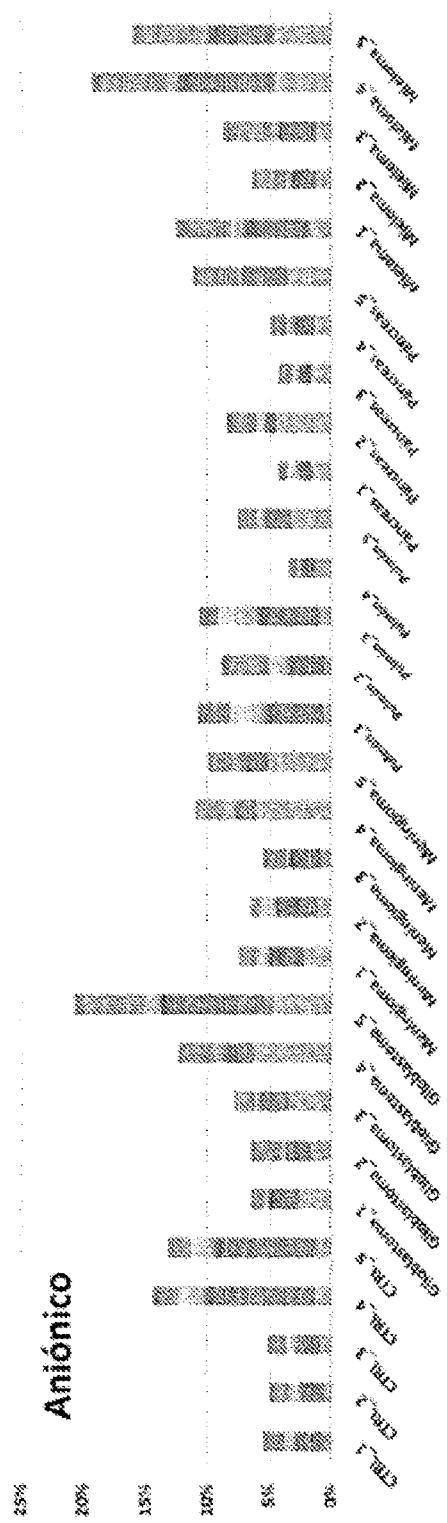


FIG. 37. CONTINUACIÓN.

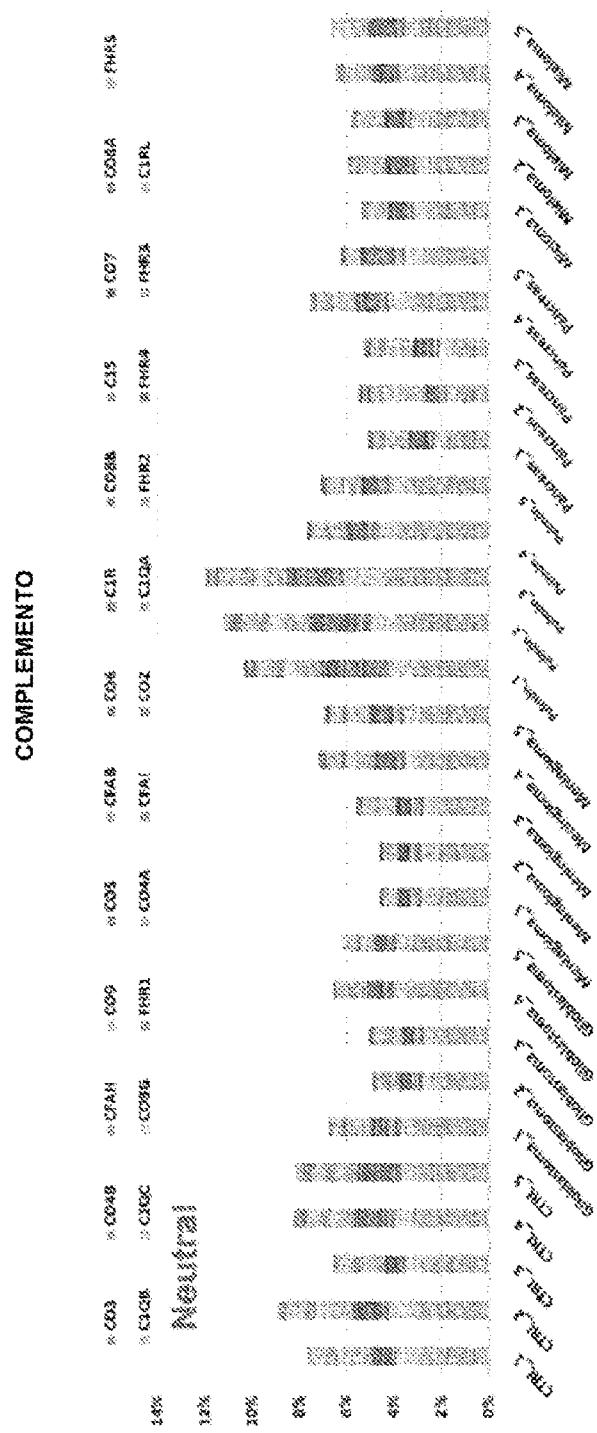


FIG. 3G
OTRAS

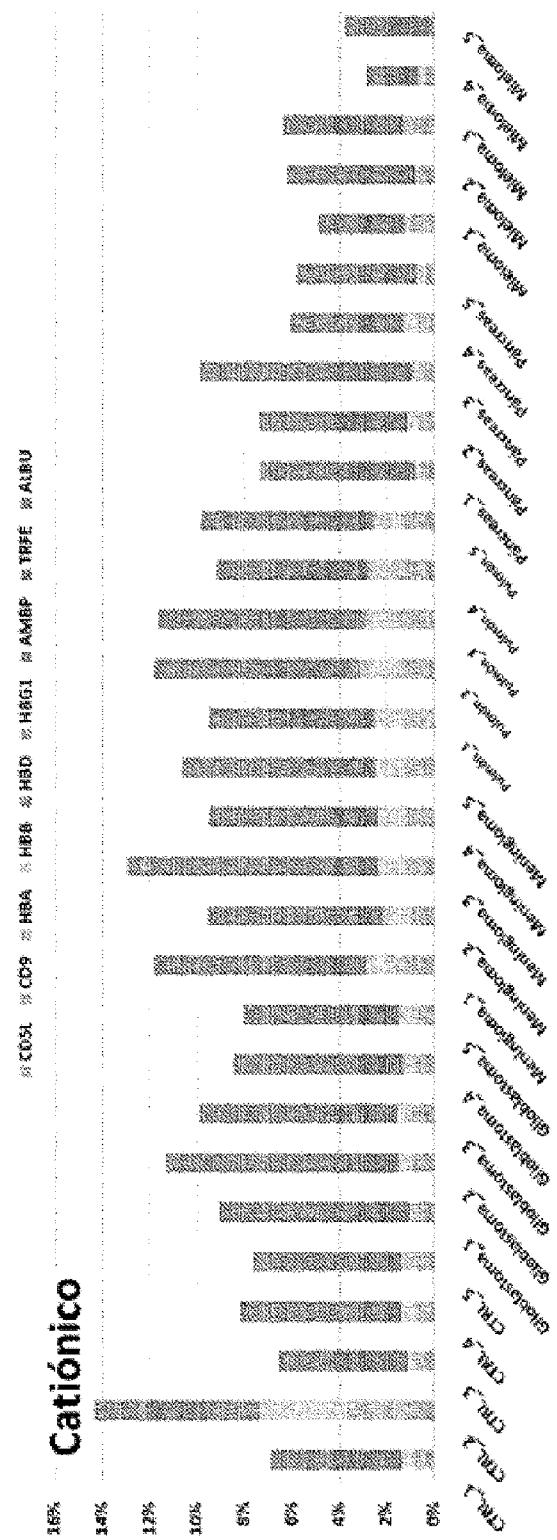


FIG. 3G CONTINUACIÓN

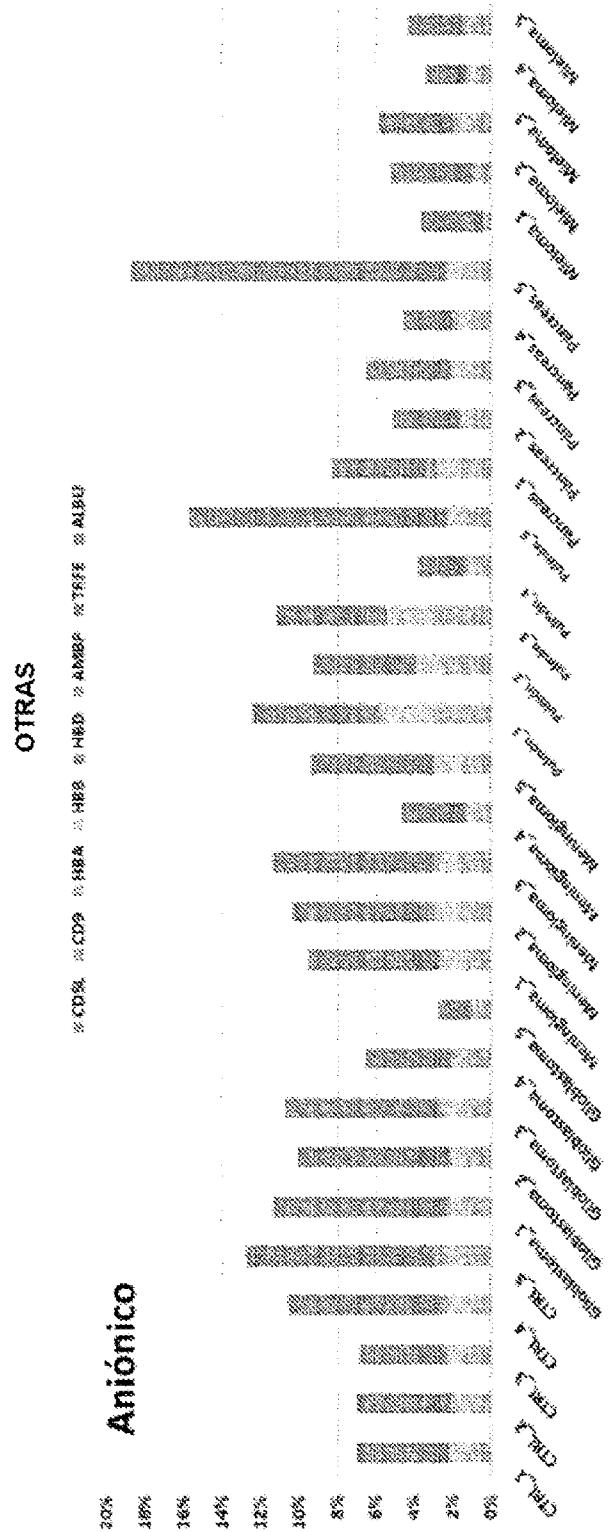


FIG. 30. CONTINUACIÓN

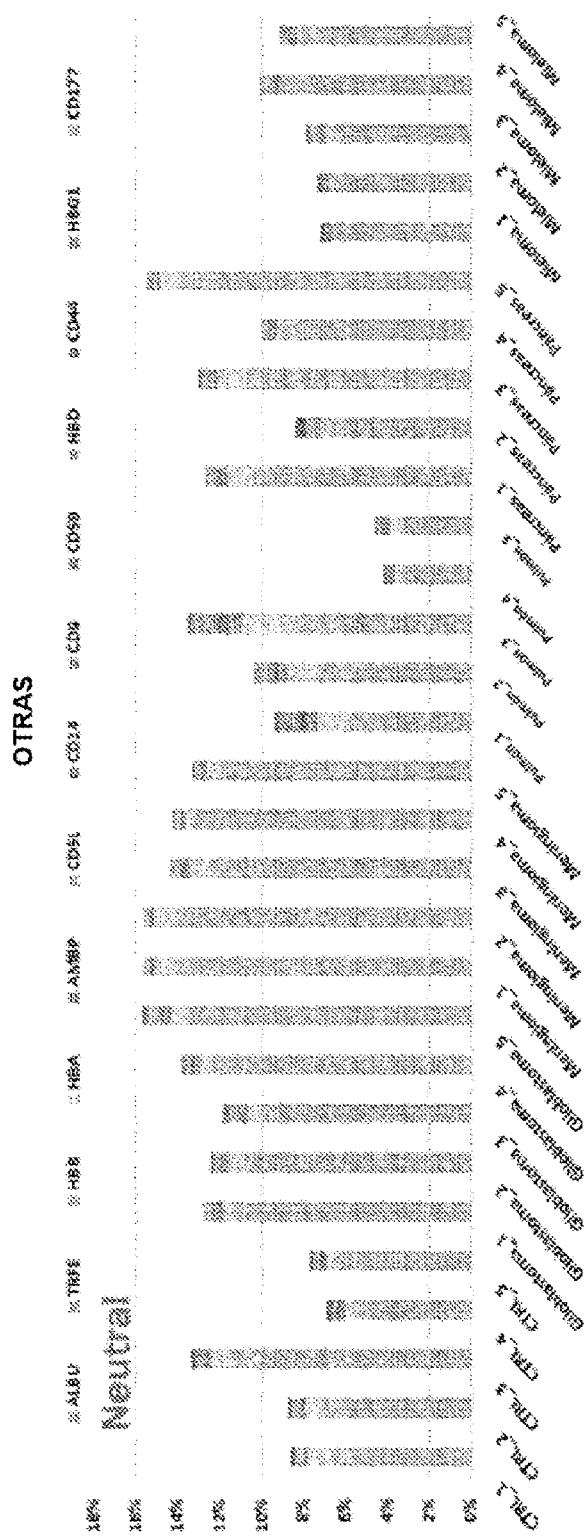


FIG. 4A

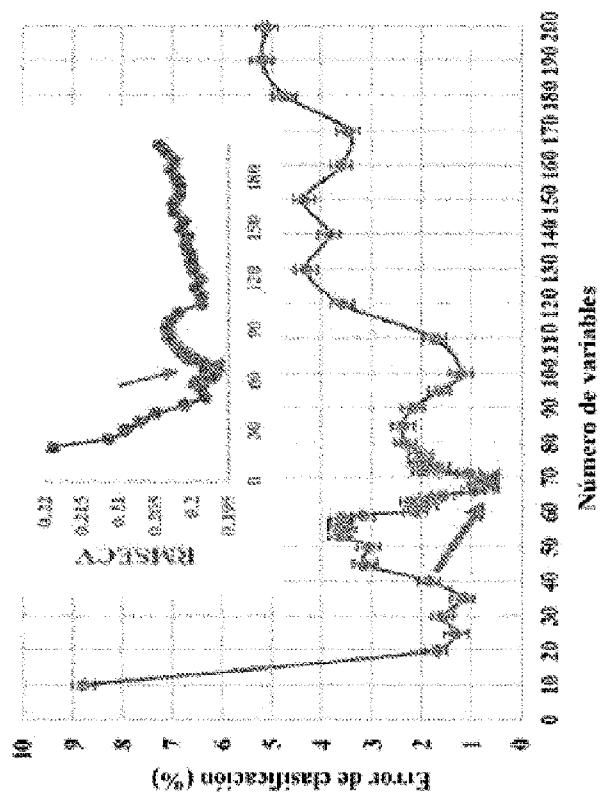


FIG. 4B

GLIOBLASTOMA

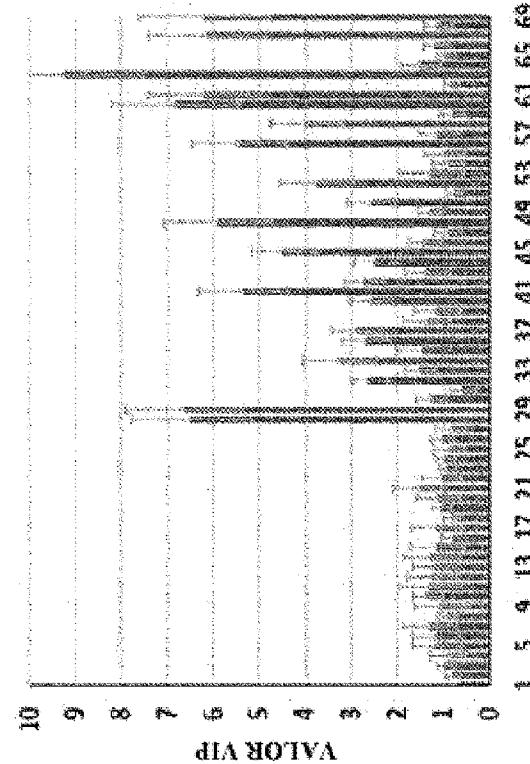


FIG. 4C

MENINGIOMA

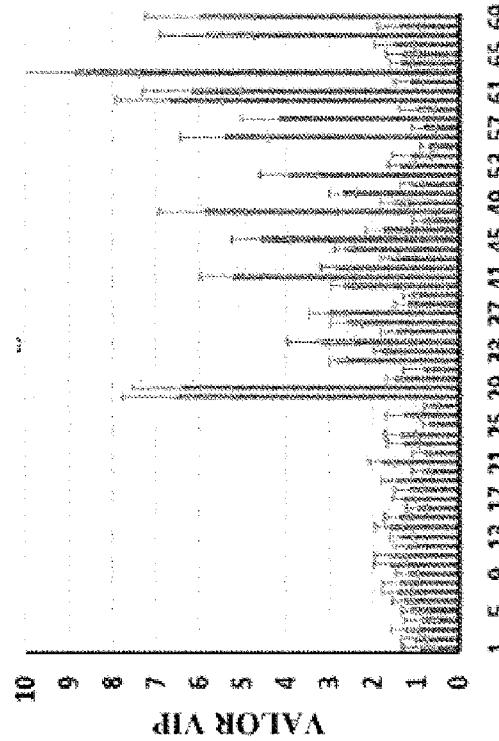


FIG. 4D

PULMÓN

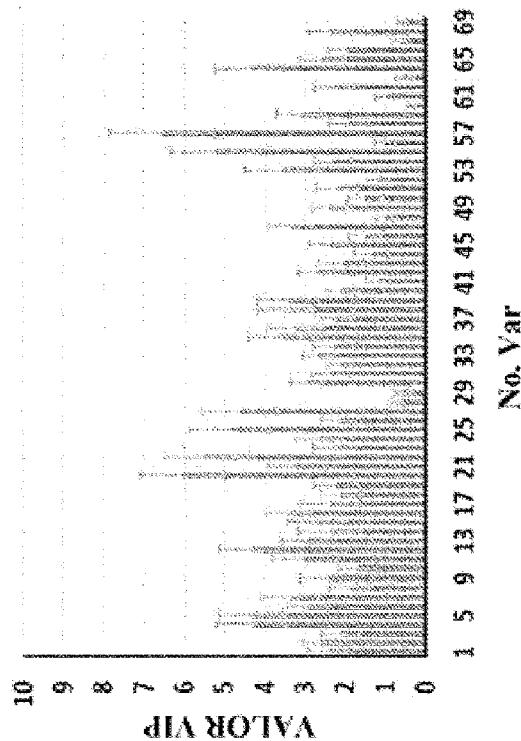


FIG. 4E

PÁNCREAS

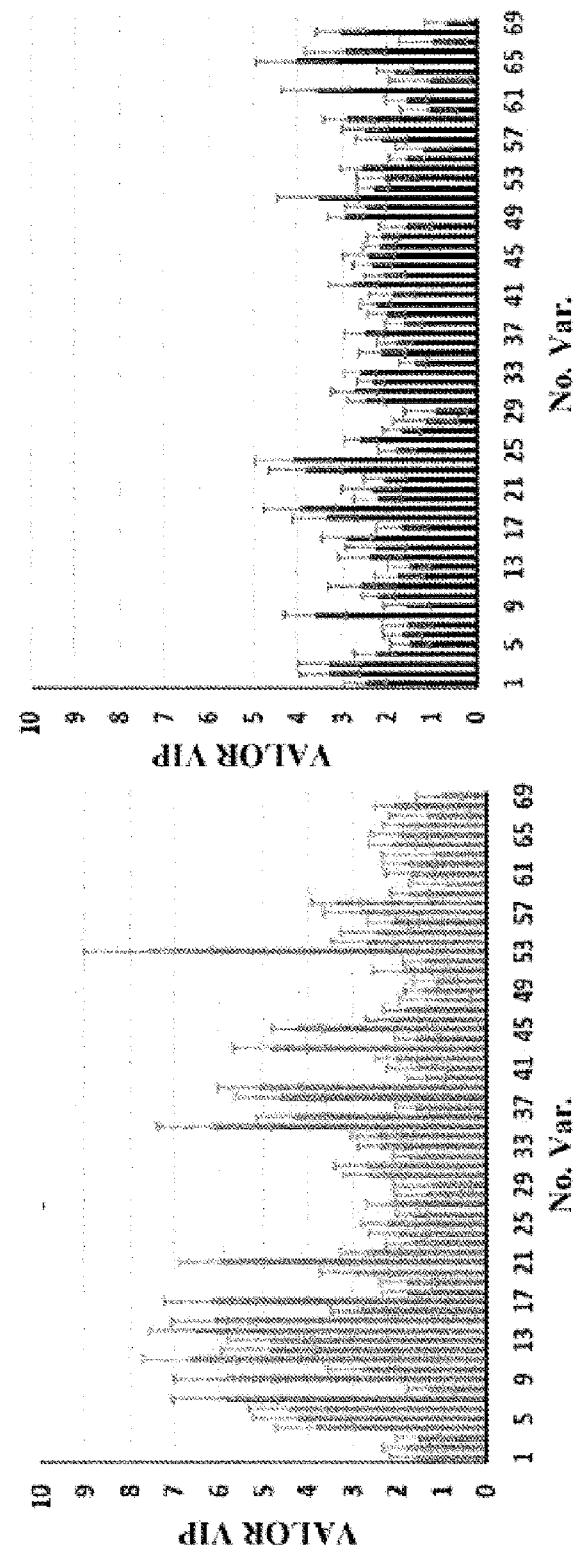
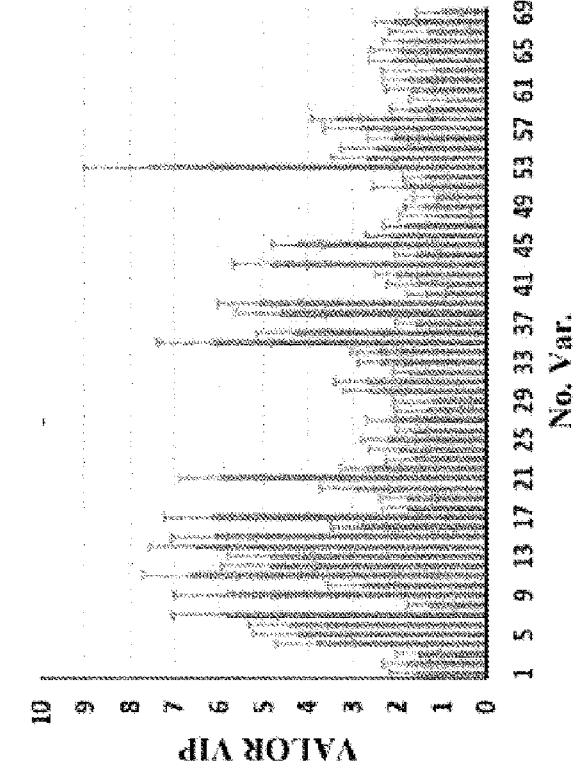
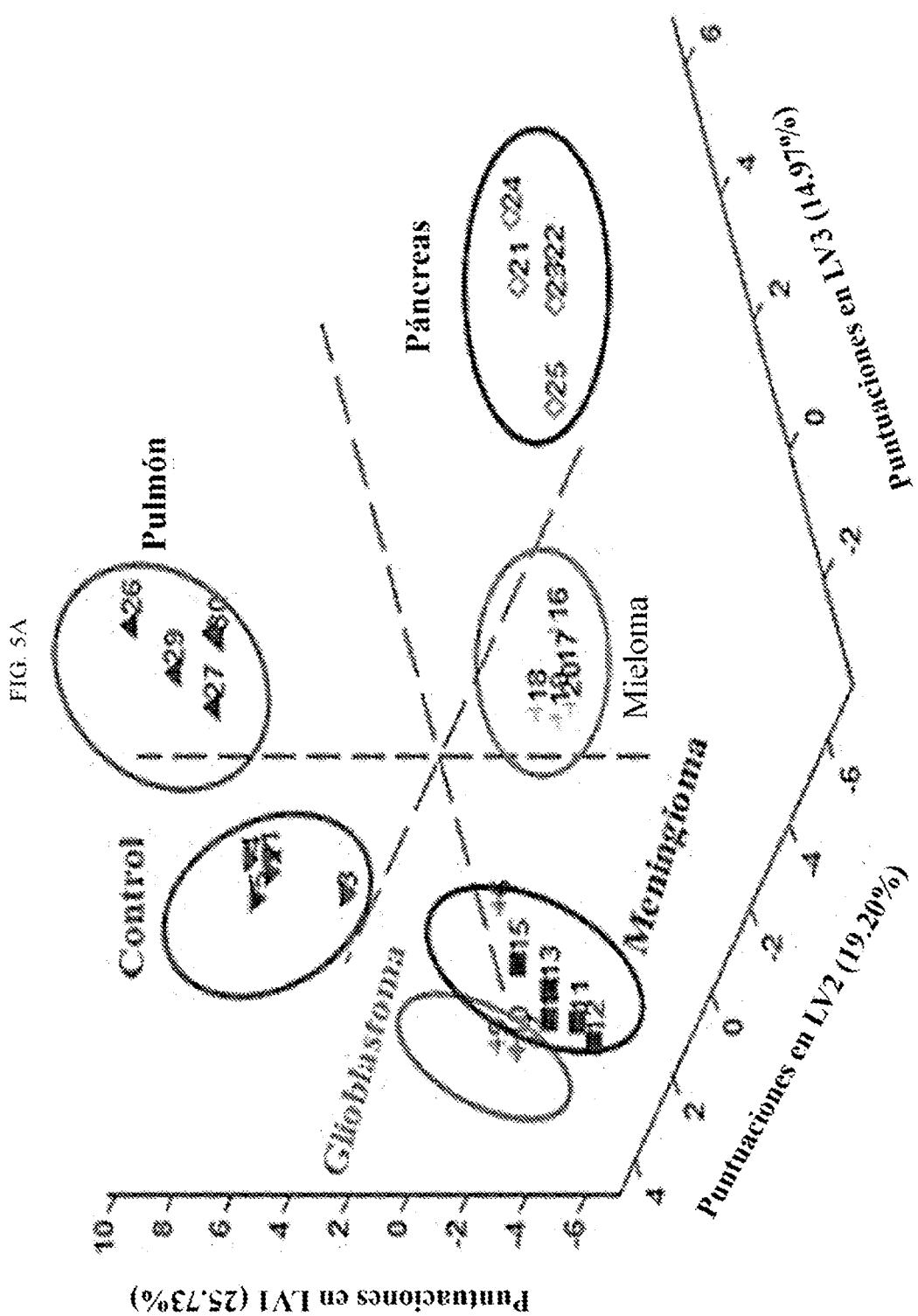
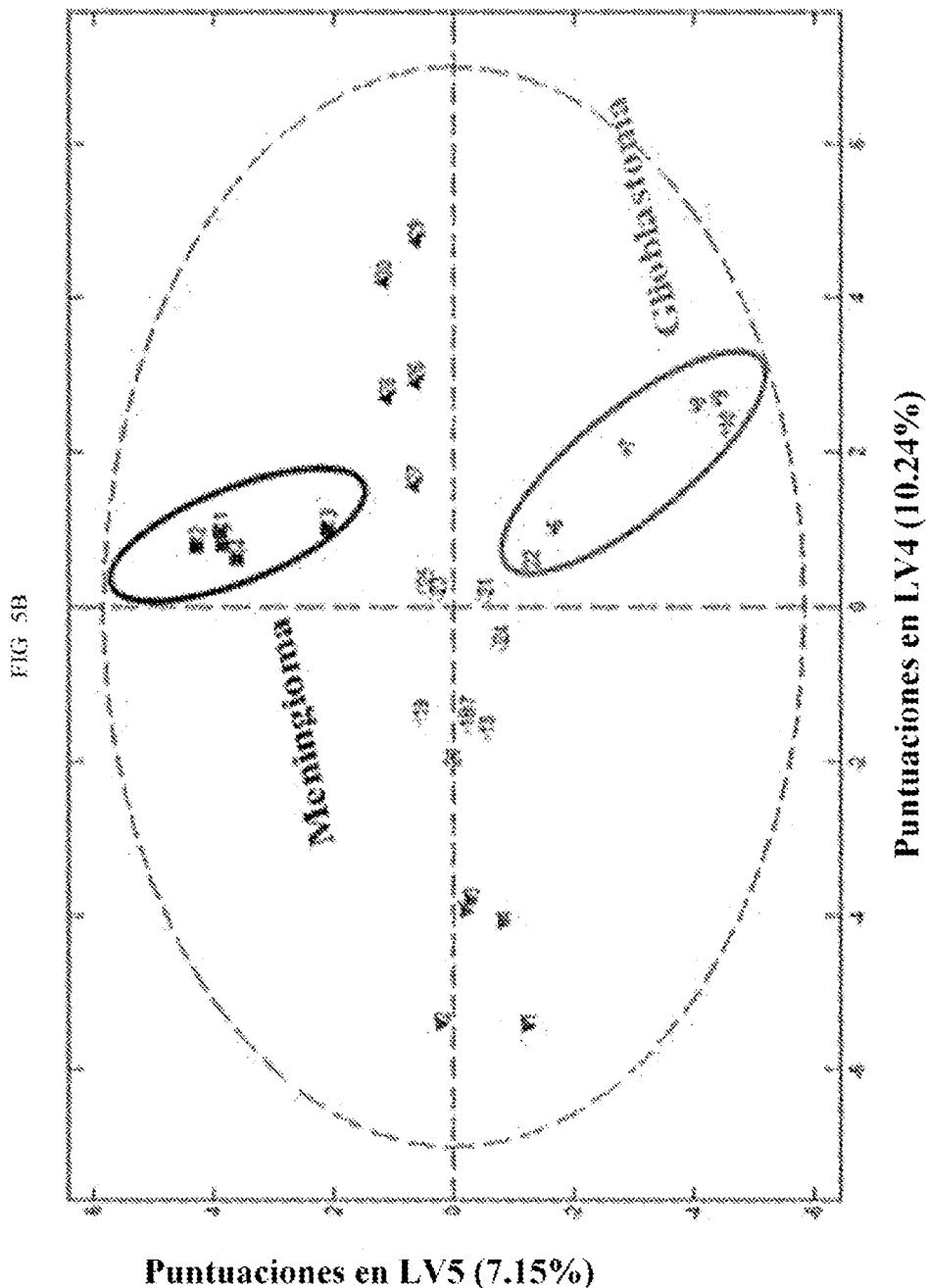


FIG. 4E

MIÉLOMA

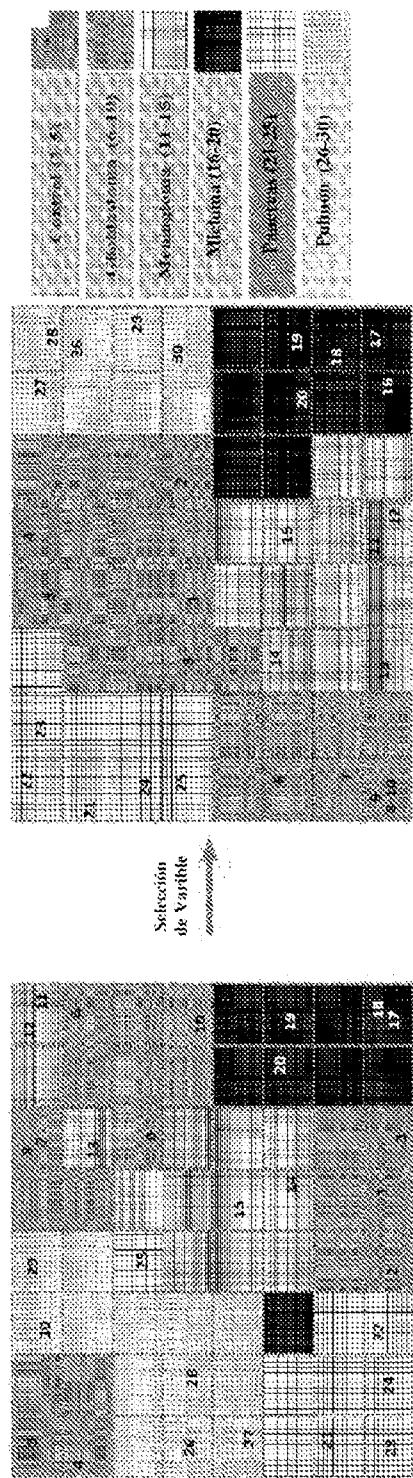




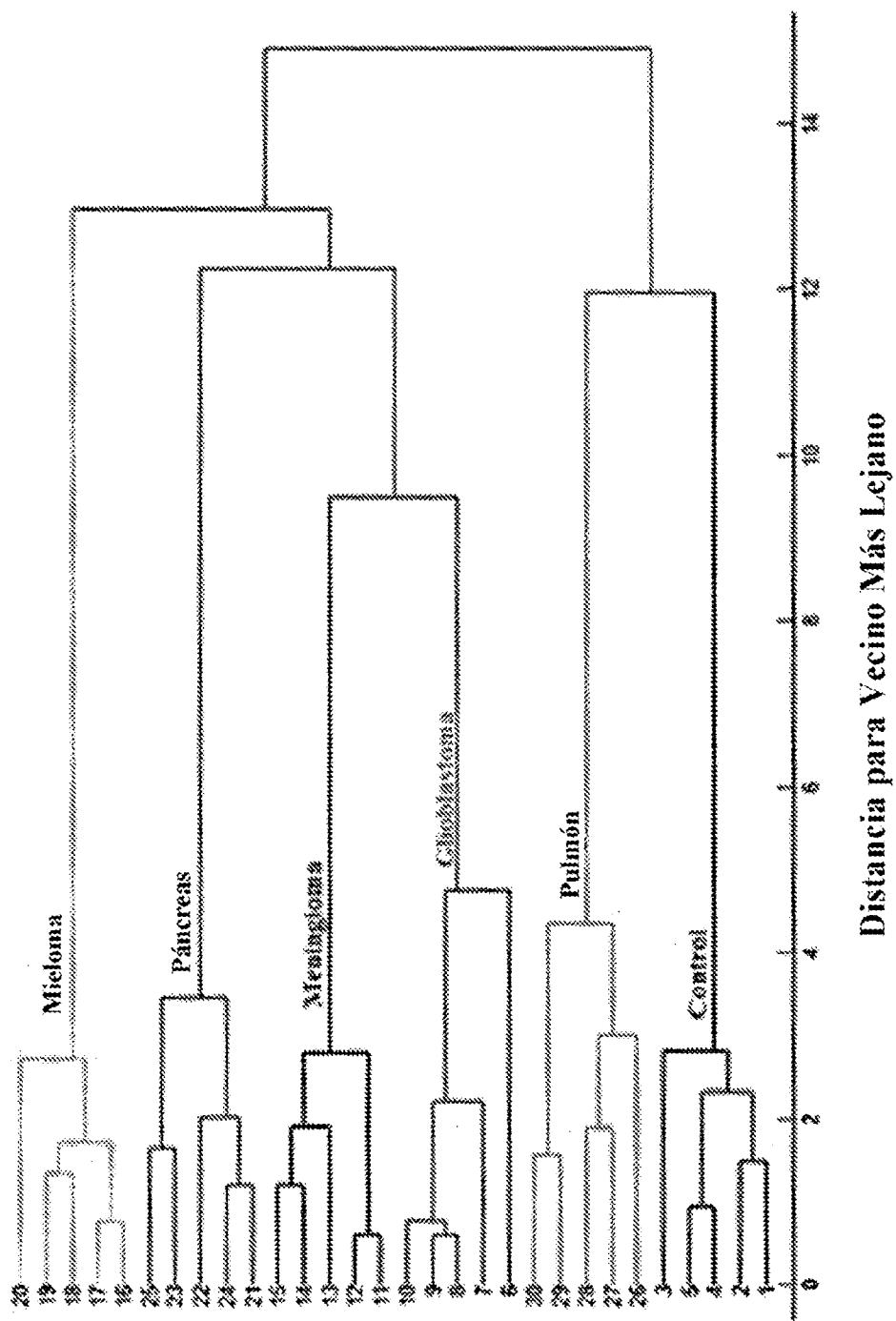


ES 2 993 503 T3

FIG. 5D



53



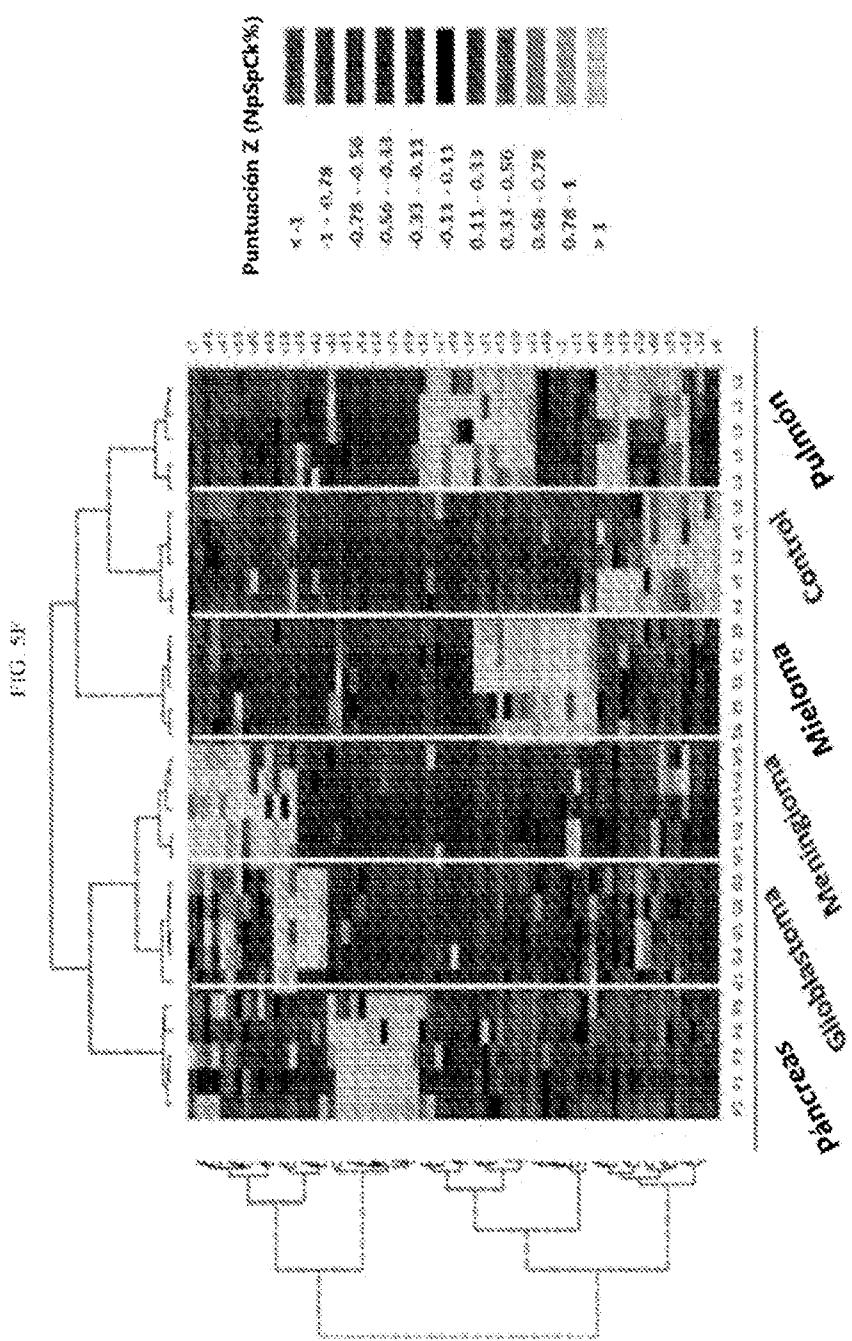


FIG. 6A

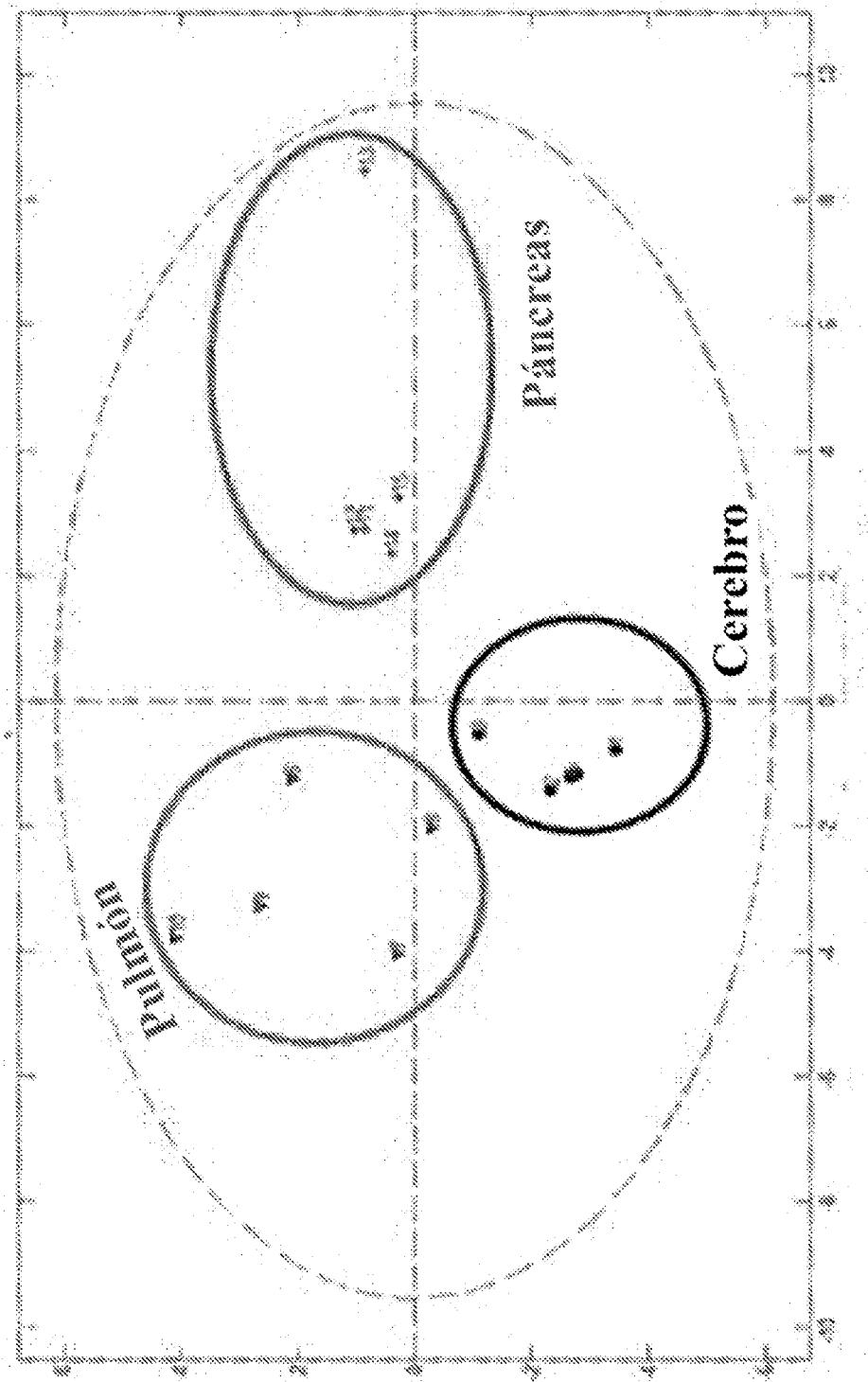
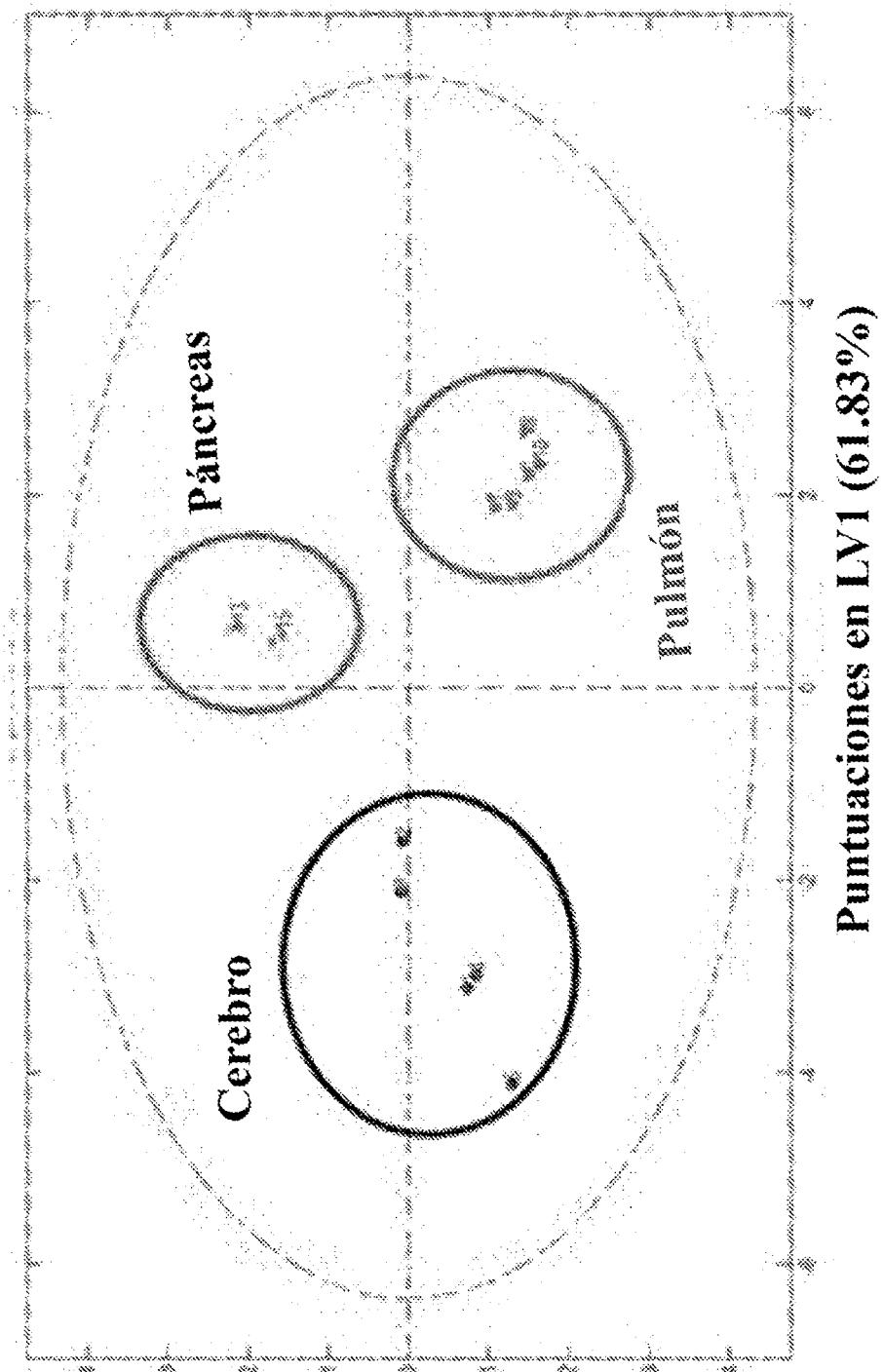


FIG. 6B



Puntuaciones en LV2 (28.62%)

Puntuaciones en LV1 (61.83%)

FIG. 6D

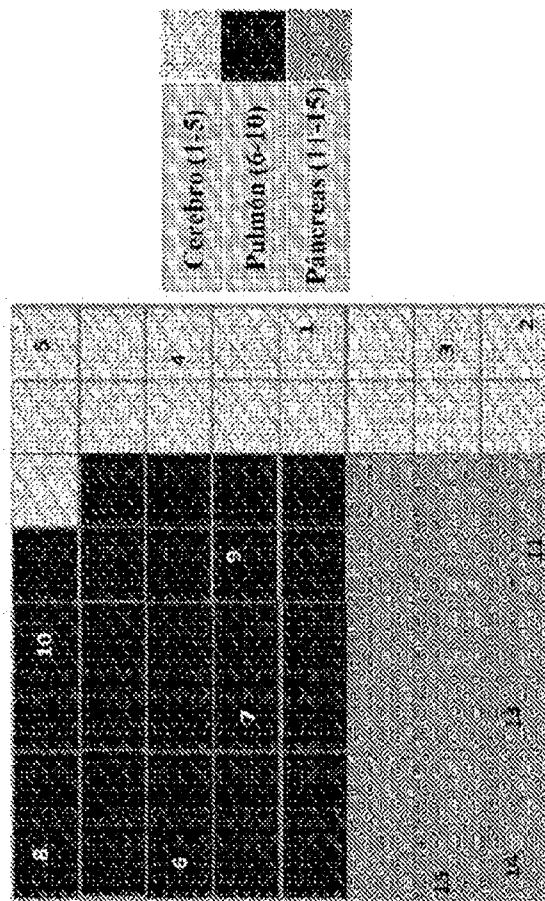


FIG. 6C

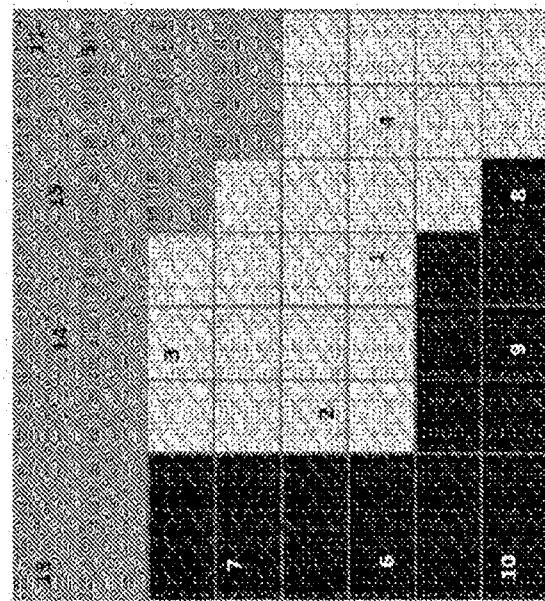


FIG. 7B

| ID de Variable (modelo PLS) | Nombre de Proteína | ID de Variable (modelo PLS) | Nombre de Proteína |
|-----------------------------|--------------------|-----------------------------|--------------------|
| 1 | PLMN | 33 | TRFE |
| 2, 19, 51 | FCN3 | 34 | APOE |
| 3 | BIN2 | 35 | IGHG1 |
| 4 | VTNC | 36 | FINC |
| 5 | ITIH1 | 37 | C4BPA |
| 6 | KNG1 | 38 | IGKC |
| 7 | CO9 | 39 | GELS |
| 8 | C1S | 40 | VTDB |
| 9 | RET4 | 41 | K1C16 |
| 10, 47 | C1R | 42 | IGHG3 |
| 11 | CO6 | 43 | LAC2 |
| 12 | IC1 | 44 | HORN |
| 13 | APOH | 45 | IGHG2 |
| 14 | CO7 | 46 | HRG |
| 15 | C08A | 48 | SAA4 |
| 16, 49 | PROP | 50 | K1C14 |
| 17 | GRP78 | 52 | HPTR |
| 18, 61 | PHLD | 53 | KV118 |
| 20 | FHR5 | 54 | CO4A |
| 21 | CO8G | 55 | ZPI |
| 22 | COF1 | 56 | CXCL7 |
| 23, 62 | CBPN | 57 | CRP |
| 24 | FGL1 | 58 | IGHG4 |
| 25 | CBPB2 | 59 | SBSN |
| 26 | COL10 | 60 | ITAM |
| 27 | KRT86 | 63 | ITB2 |
| 28, 67 | CDS9 | 64 | PGK1 |
| 29 | KAIN | 65 | COL11 |
| 30 | ALBU | 66 | EMIL1 |
| 31 | CO4B | 68 | QCR2 |
| 32 | A2MG | 69 | HV269 |

FIG. 7A

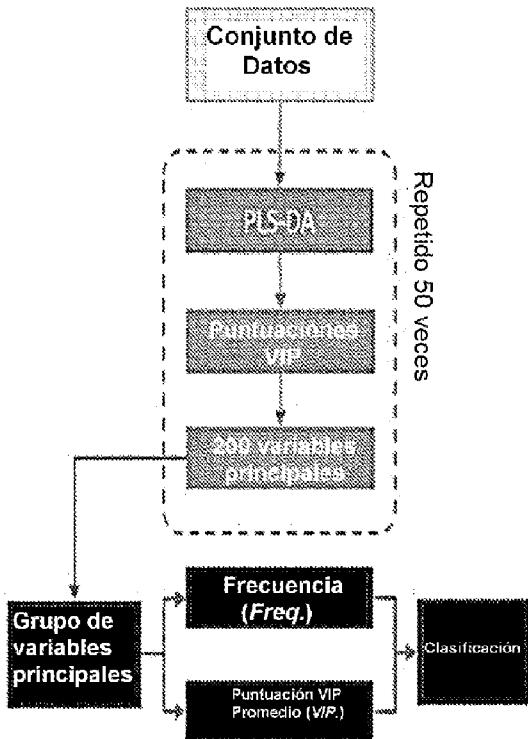


FIG. 7C

| Tipo de Cáncer | Nombre de Proteína |
|----------------|---|
| Glioblastoma | PLMN, VTNC, ITIH1, CO7, FHR5, CBPN, ALBU, C4BPA, CO4A, CRP, APOE, CXCL7 |
| Meningioma | FCN3, RET4, CBPN |
| Páncreas | KNG1, IC1, CBPB2, TRFE, APOE, GELS, HPTR, CXCL7, PGK1 |
| Pulmón | CO9, SAA4, CRP, GELS |
| Mieloma | ALBU |

FIG. 8A

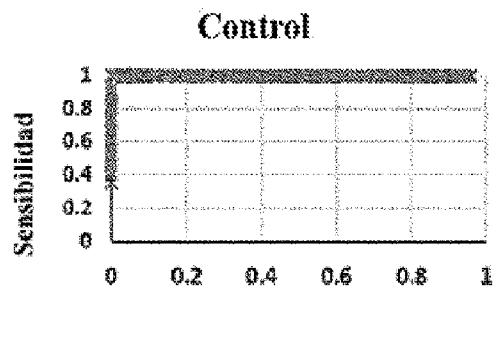


FIG. 8B

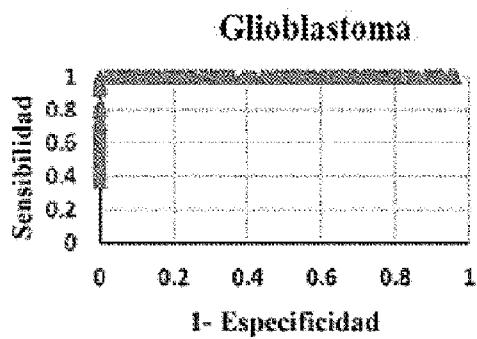


FIG. 8C

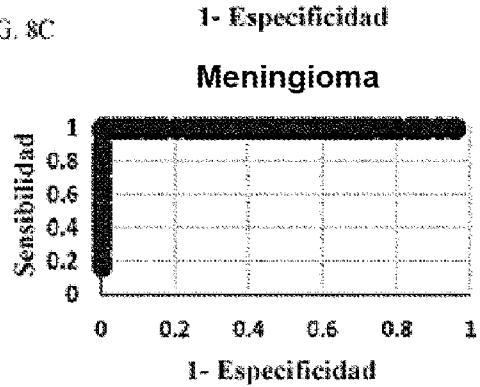


FIG. 8D

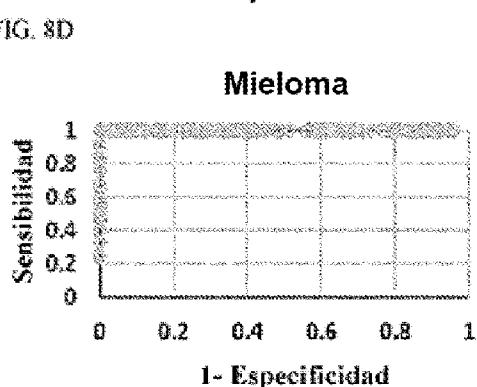


FIG. 8E

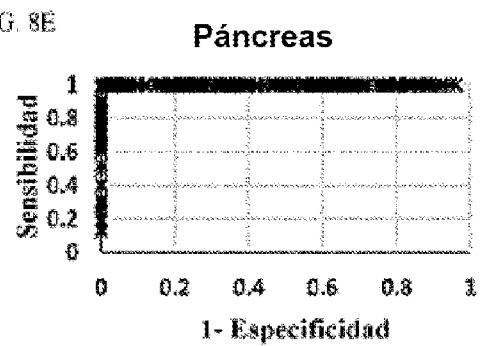


FIG. 8F

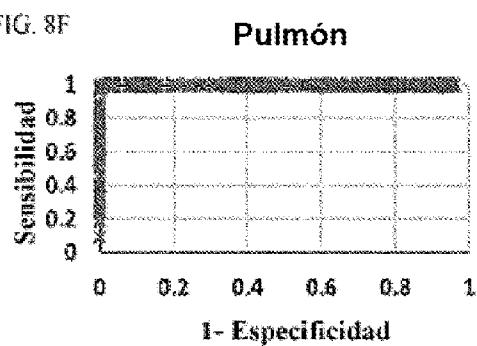


FIG. 9A

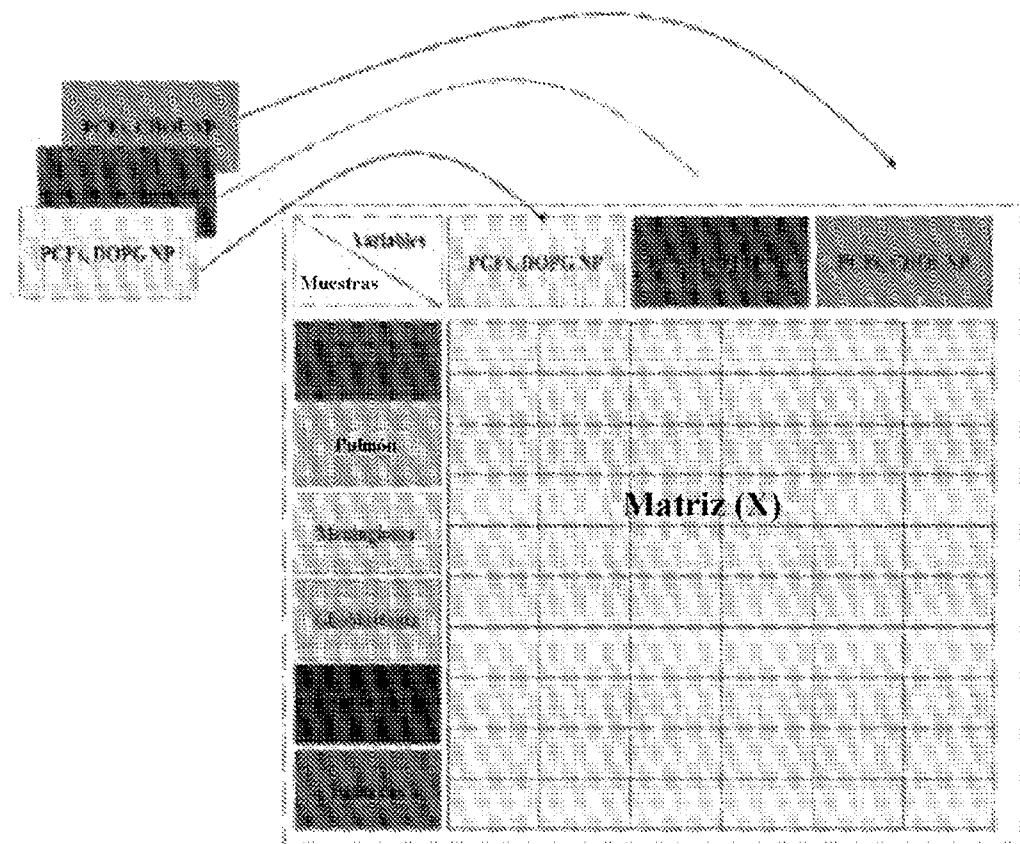


FIG. 9B

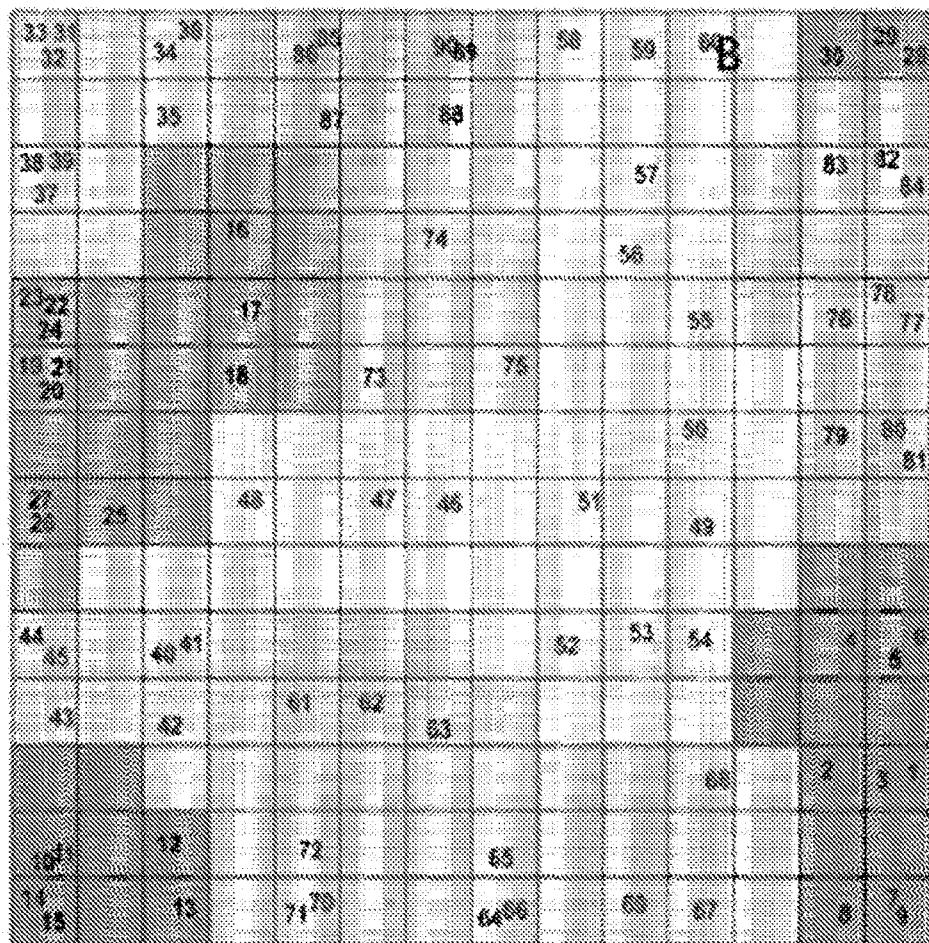


FIG. 9C

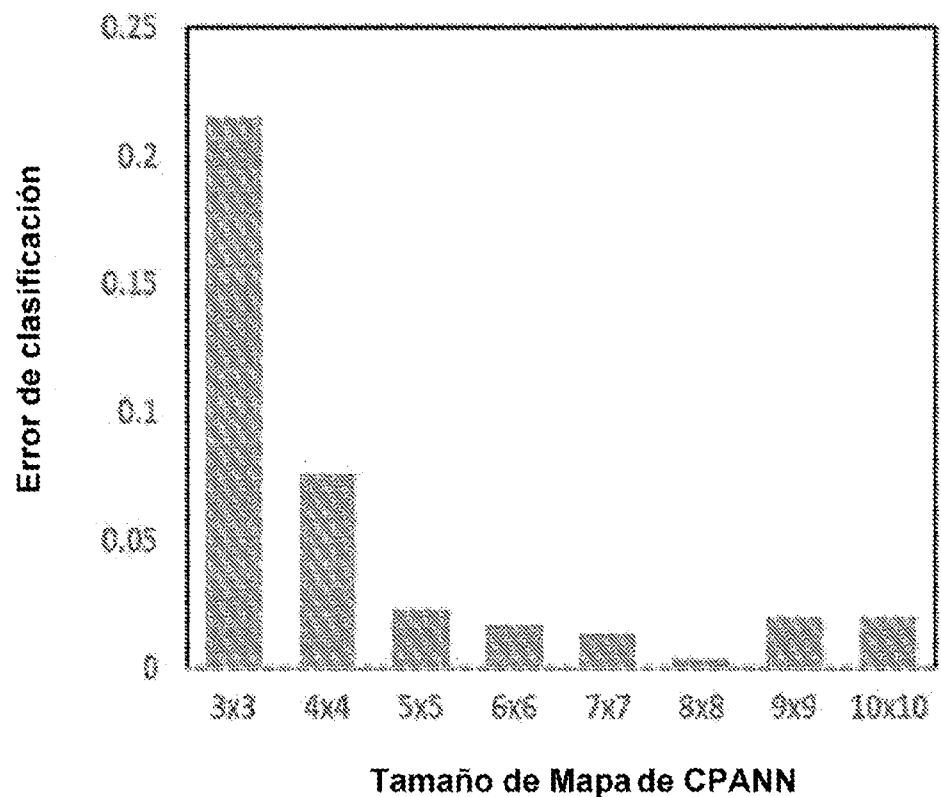


FIG. 9D

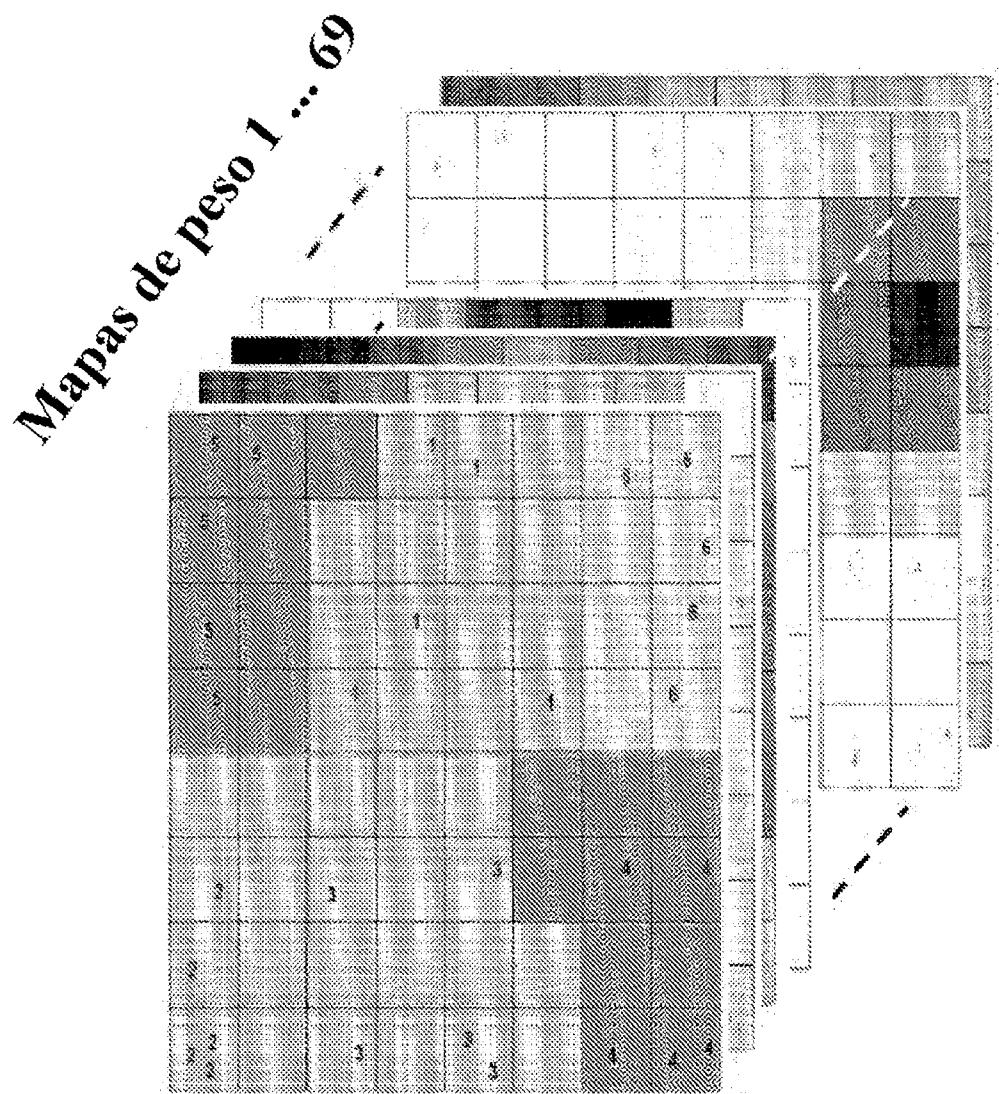


FIG. 9E

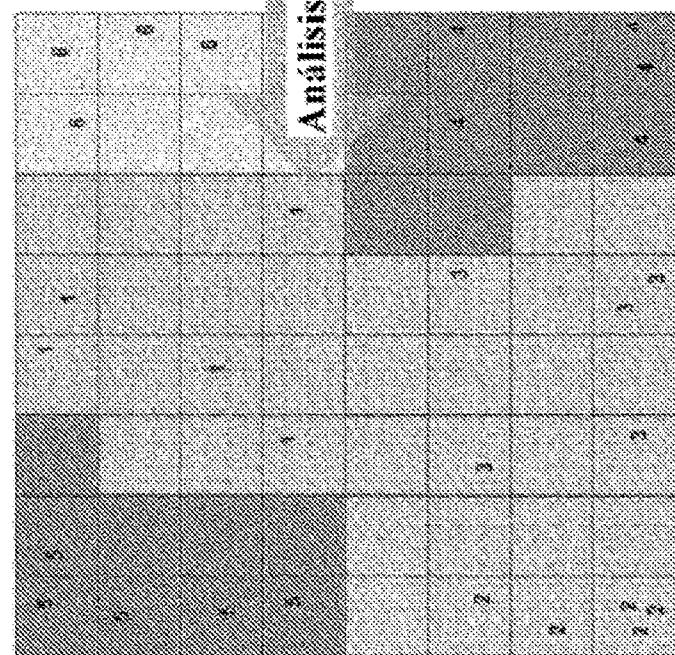


FIG. 9F

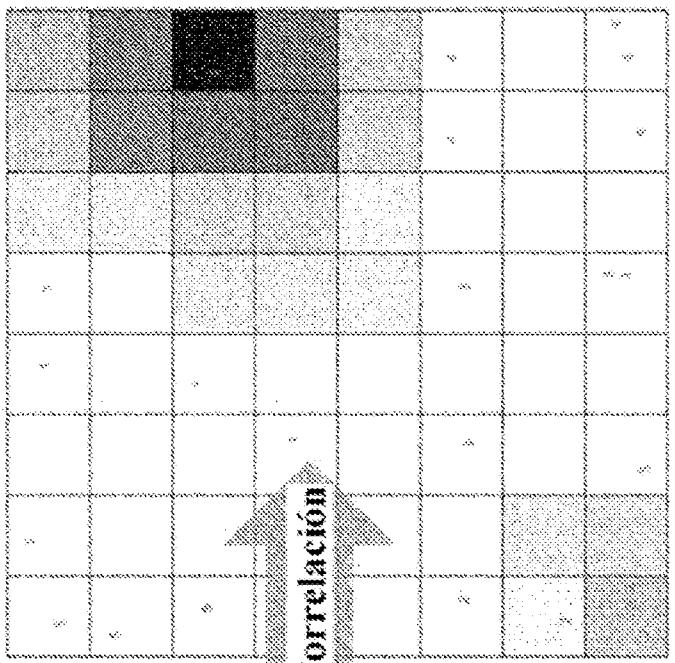


FIG. 10A

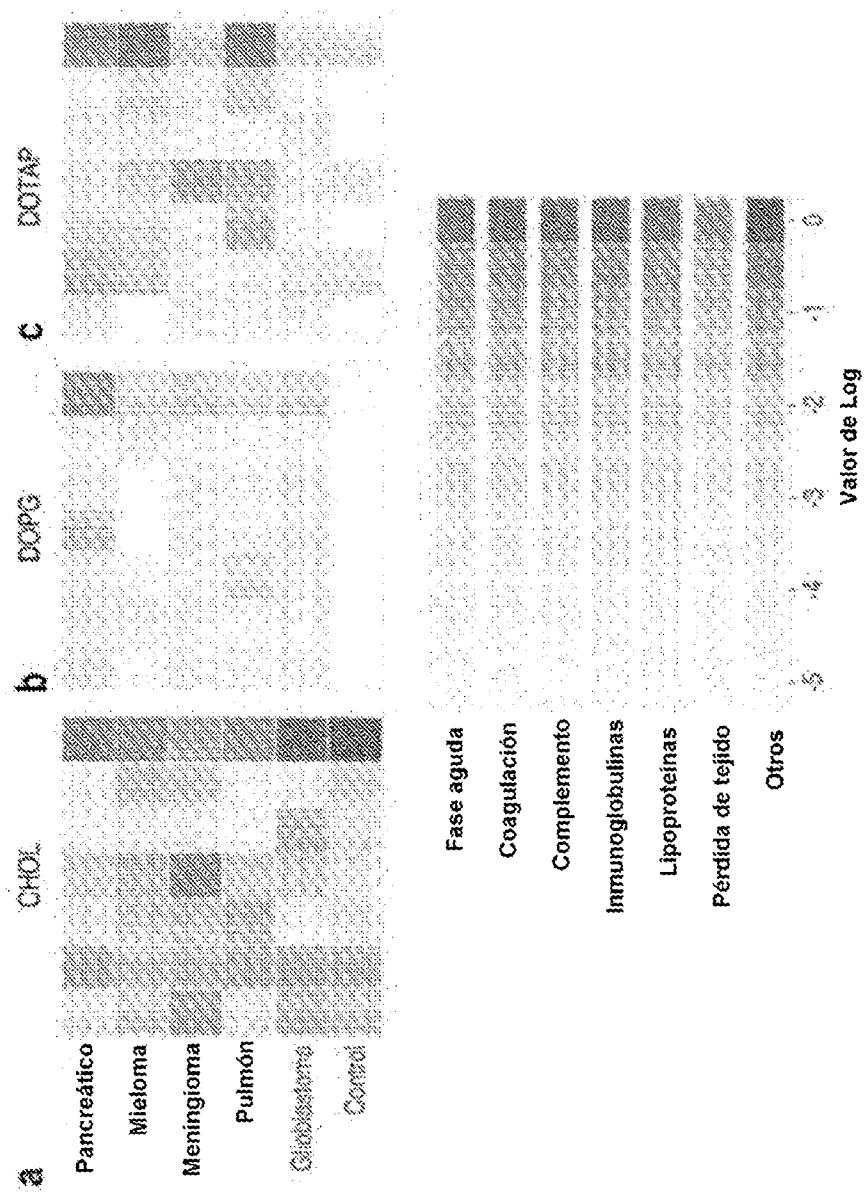


FIG. 10B

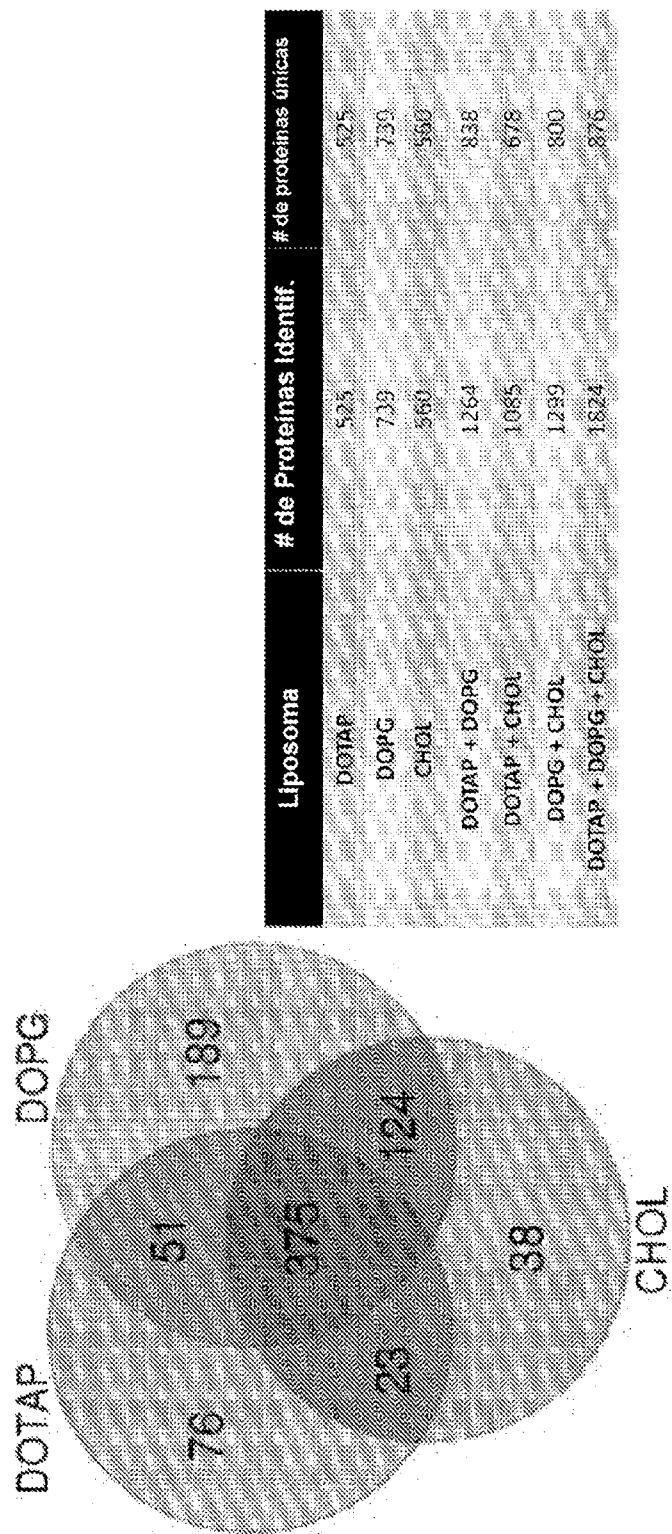
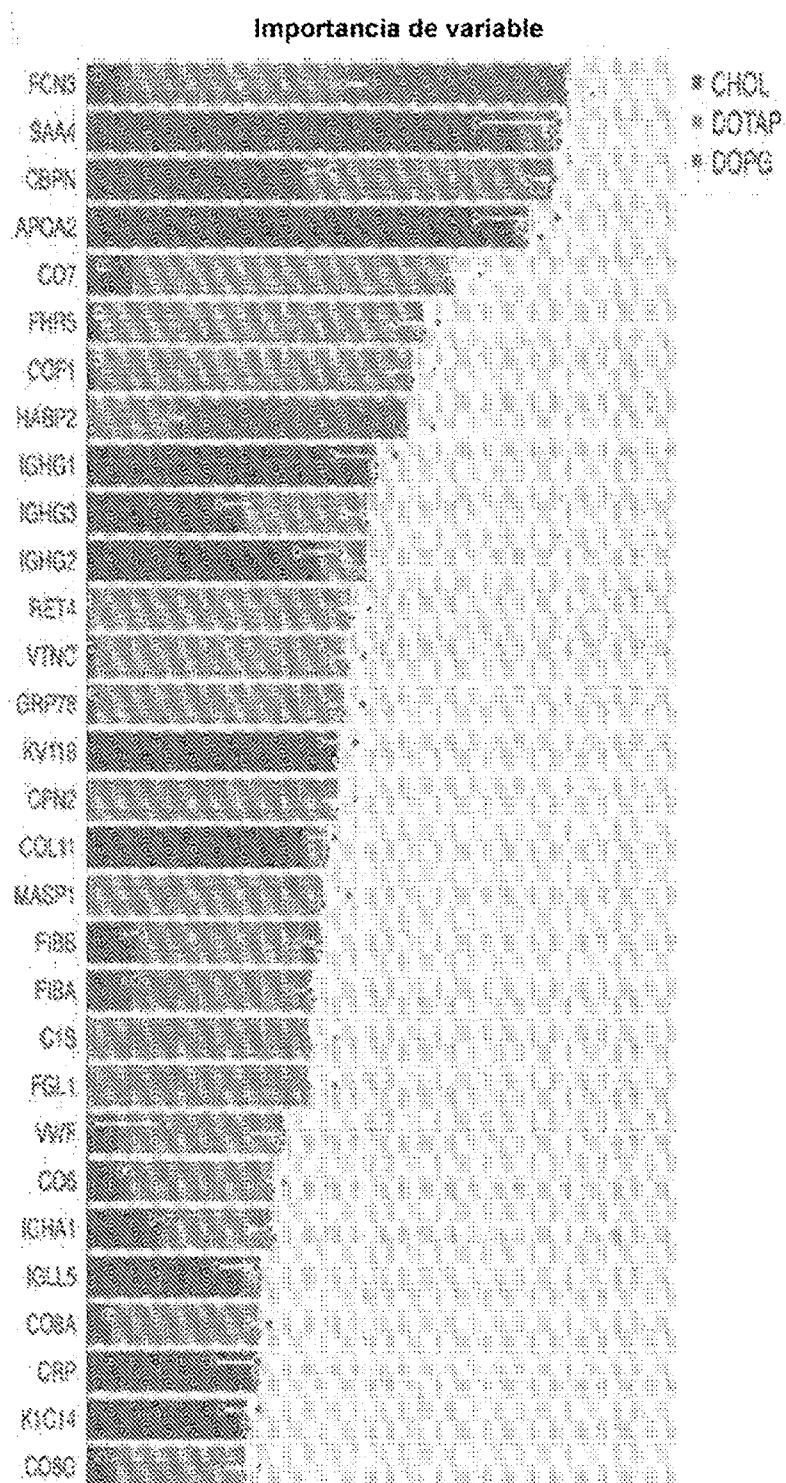


FIG. 10C



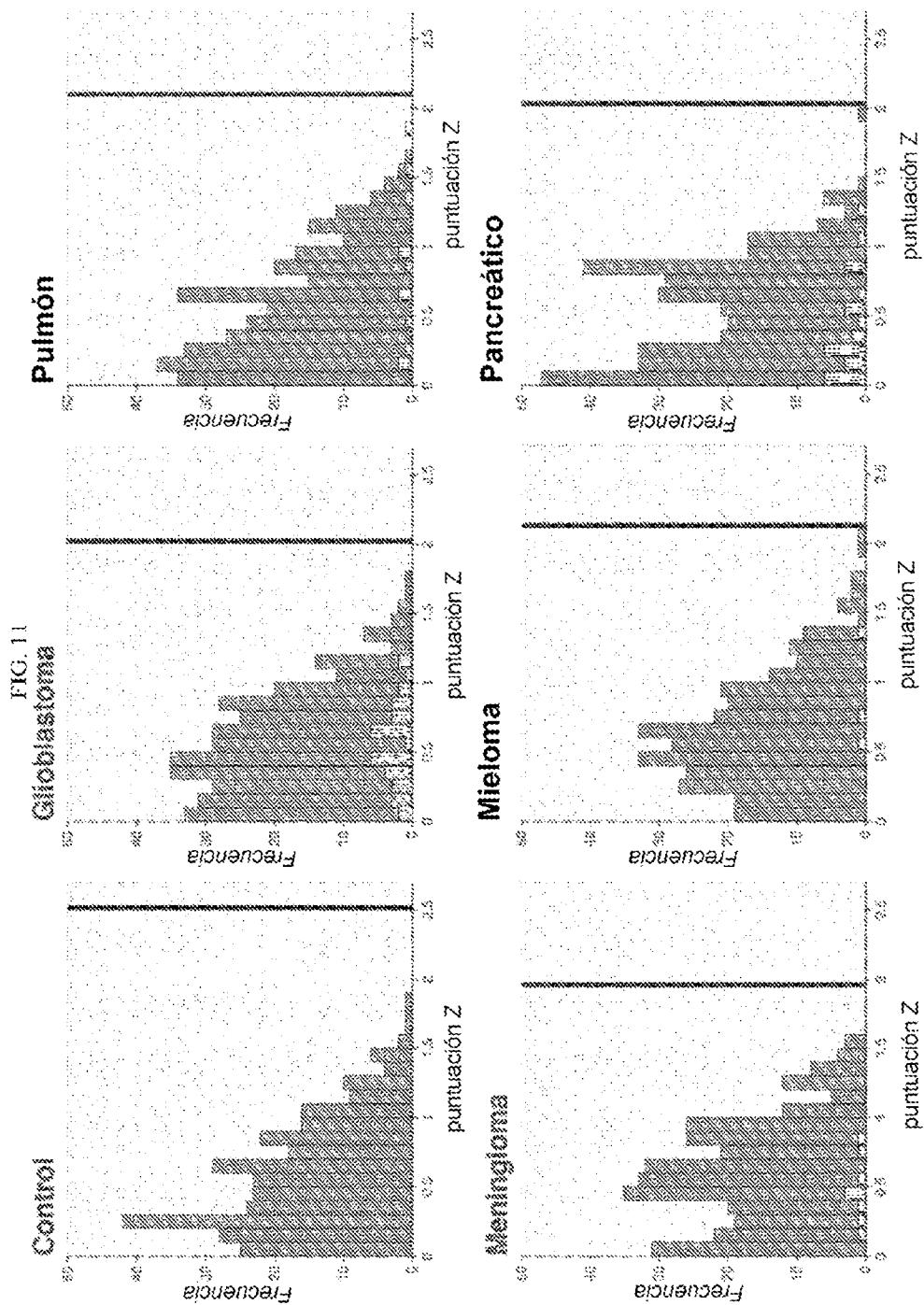
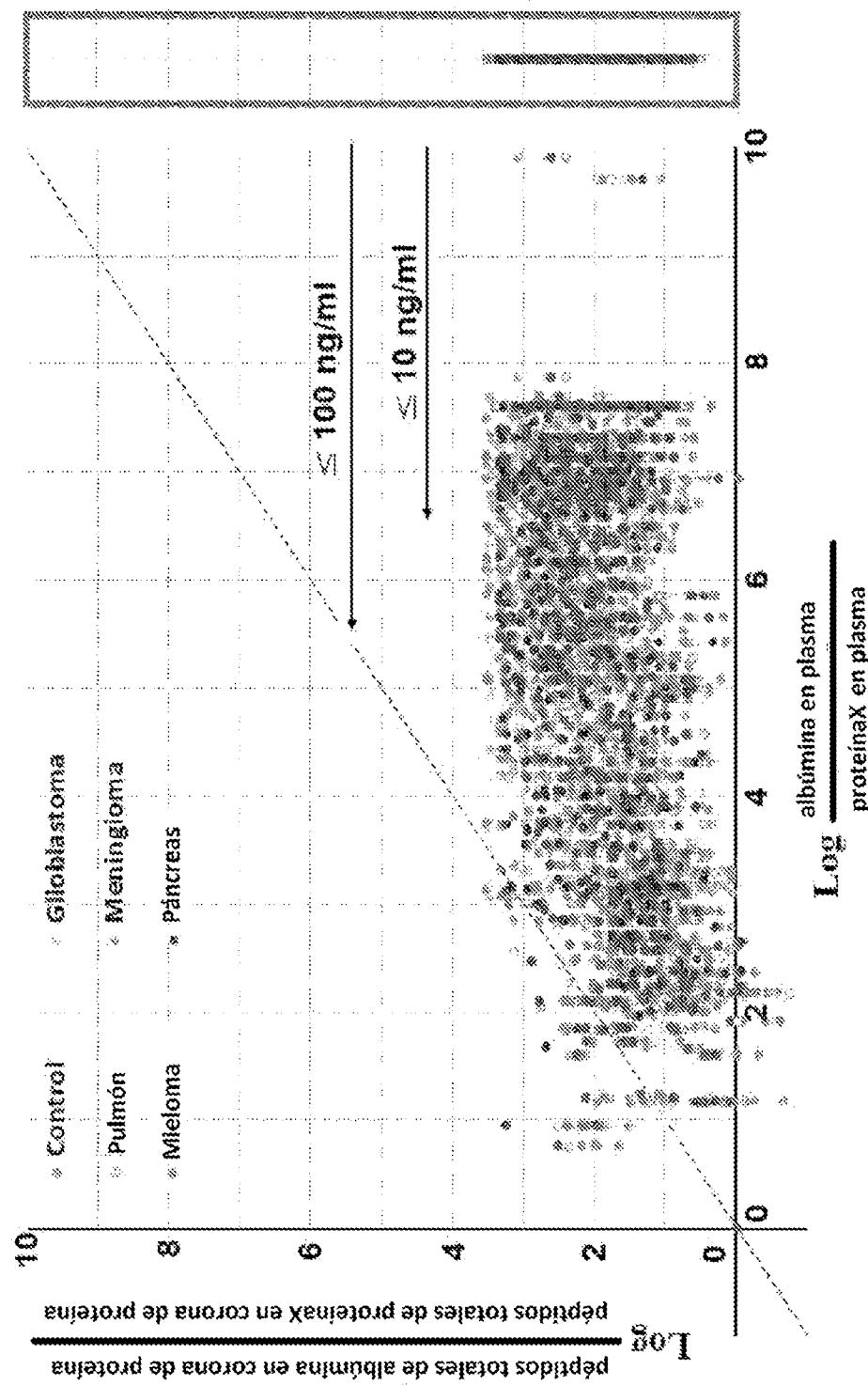


FIG. 12

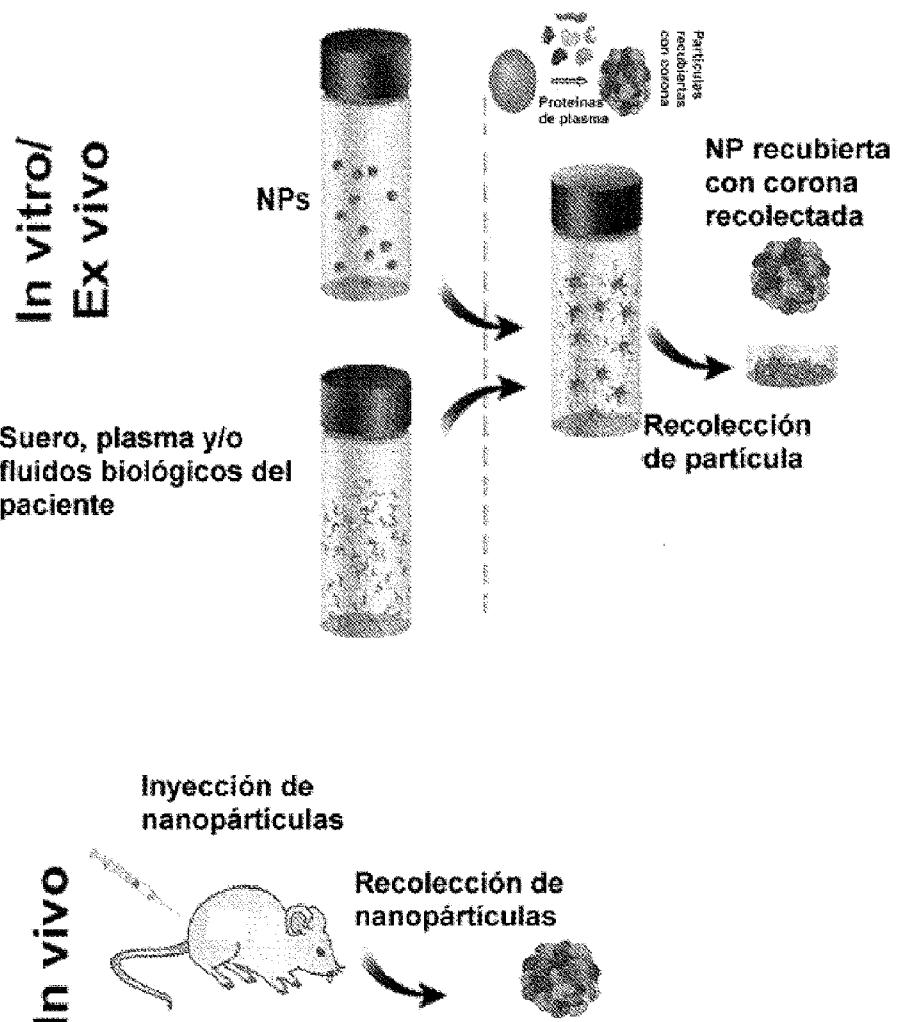


**Elementos de la
matriz de sensores**
**Diversos tipos de NPs (orgánicos,
inorgánicos, y o poliméricos)
con diferentes propiedades
fisicoquímicas**

Ejemplos

| Liposomas Funcionalidades/Cargas de Superficie Diferentes | Diferentes Tamaños | Composiciones y conjugados de Lipidos Diferentes p.e., colesterol |
|--|--------------------|--|
| | | |
| Nanopartículas superparamagnéticas de óxido de hierro Funcionalidades/Cargas de Superficie Diferentes | Diferentes Tamaños | Diferentes Formas |
| | | |
| Nanopartículas de oro Funcionalidades/Cargas de Superficie Diferentes | Diferentes Tamaños | Diferentes Formas |
| | | |

FIG. 14

Formación de la corona de proteína

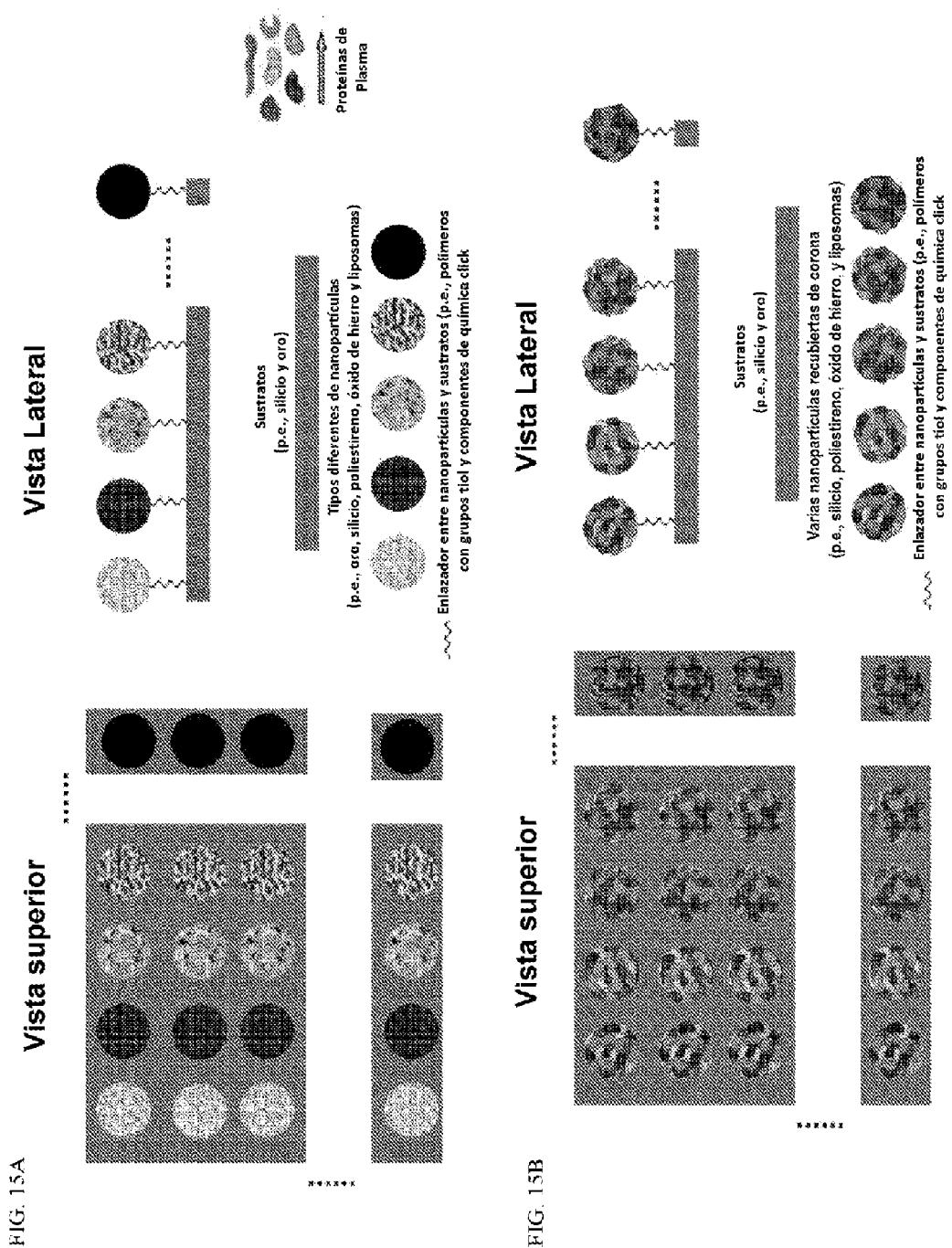


FIG. 16A Vista superior Vista Lateral

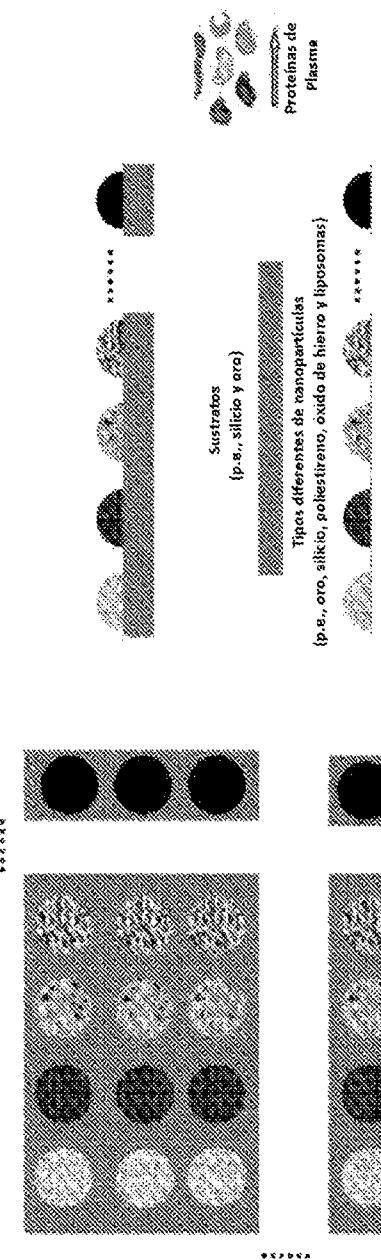
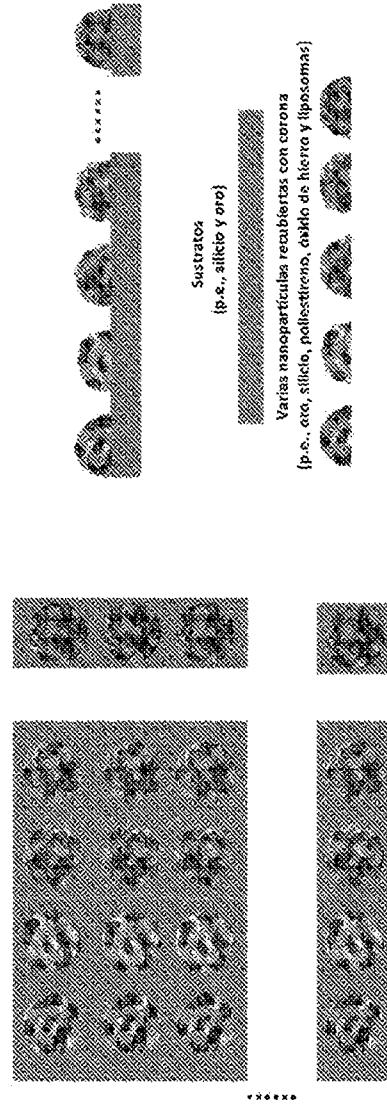
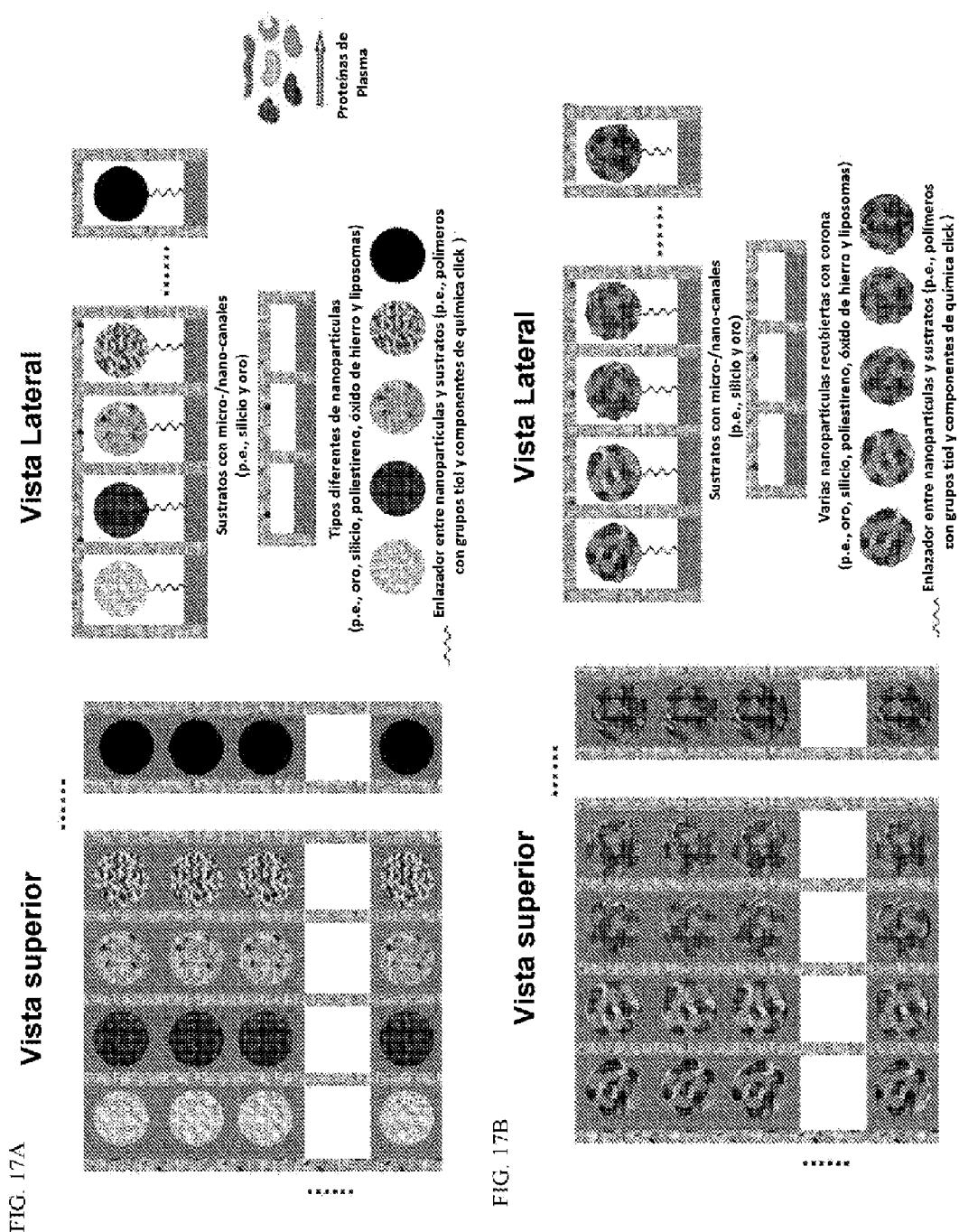
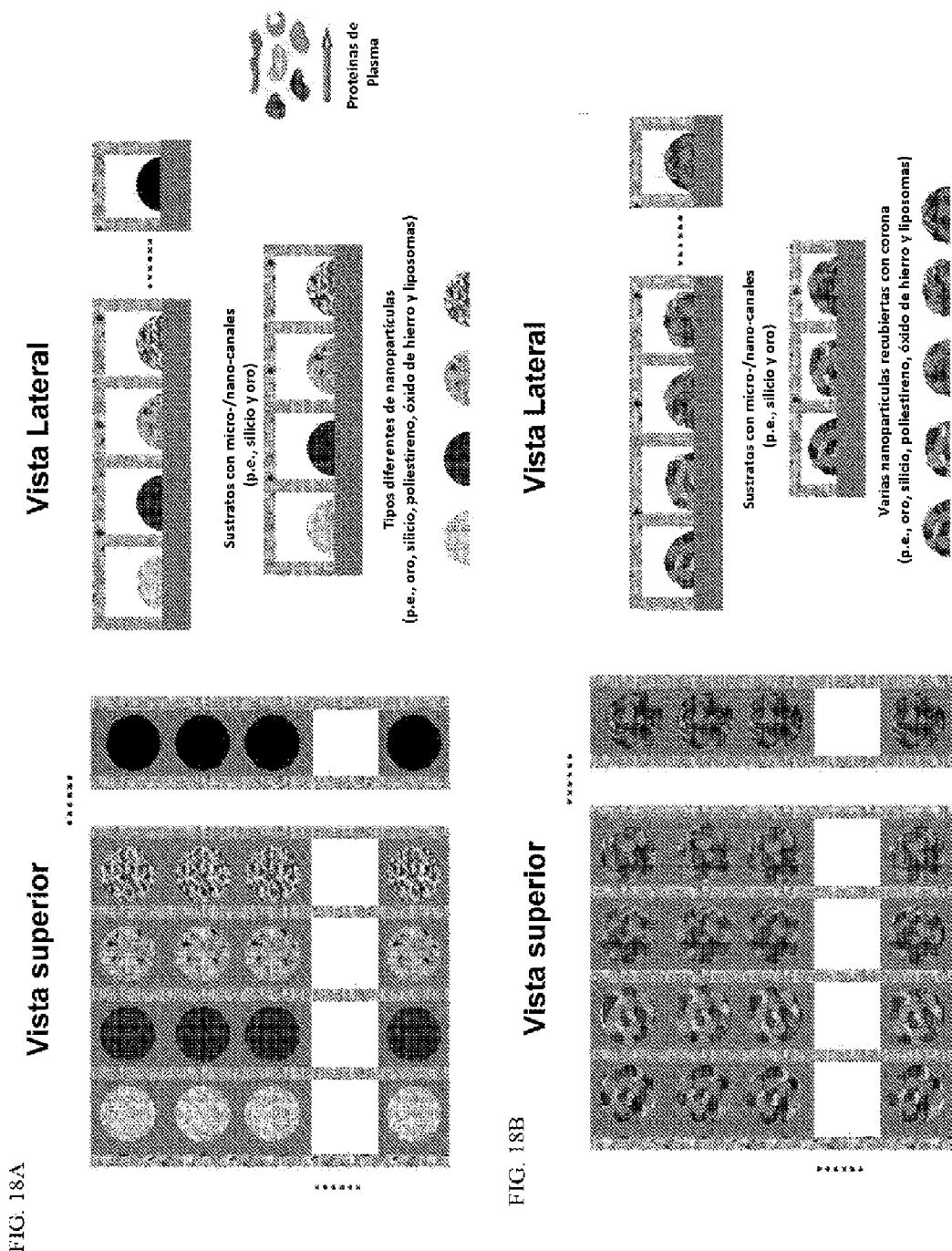


FIG. 16B Vista superior Vista Lateral







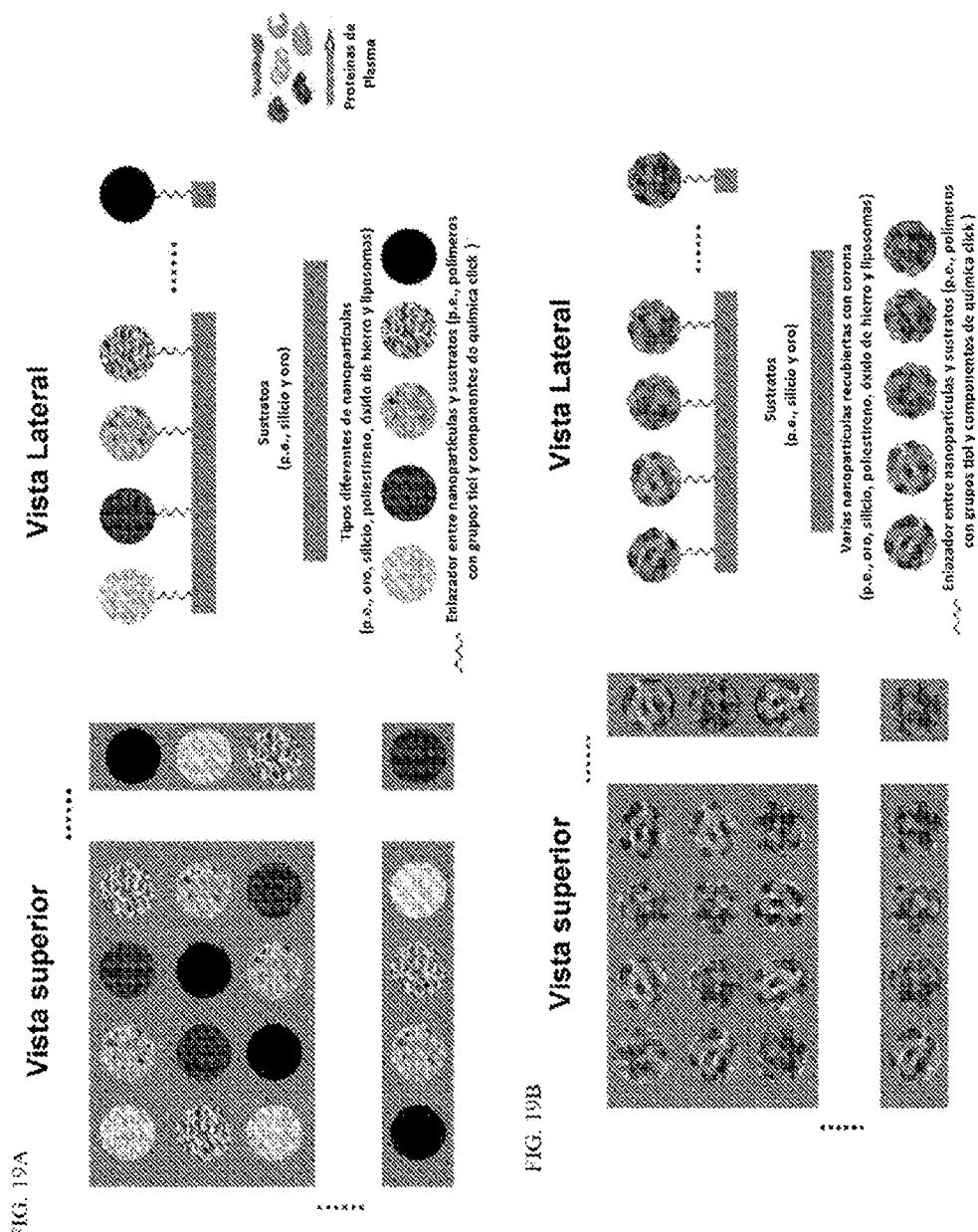


FIG. 20A Vista superior Vista Lateral

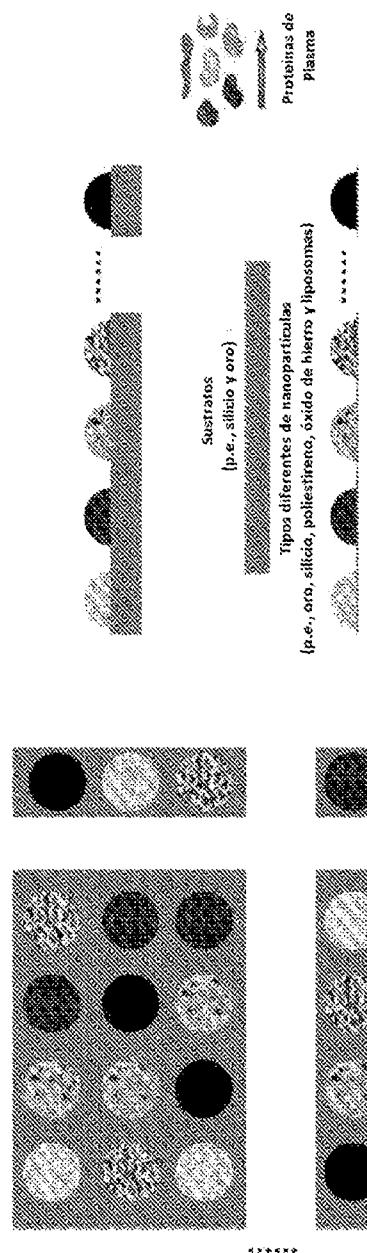


FIG. 20B Vista superior Vista Lateral

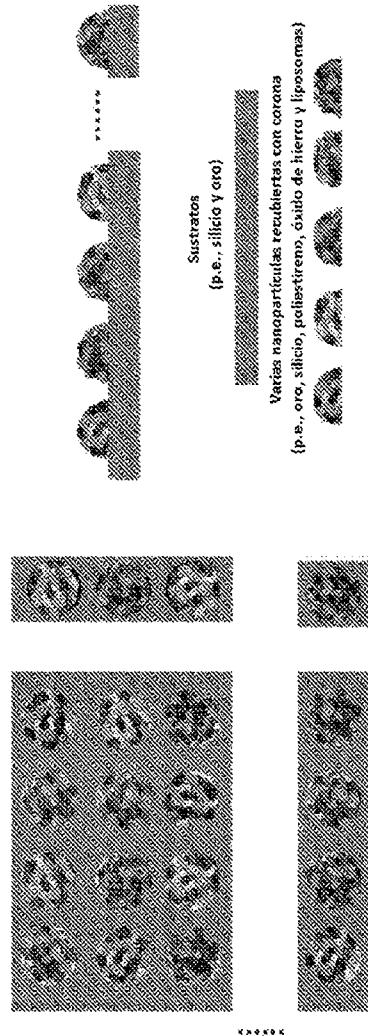


FIG. 21A Vista superior

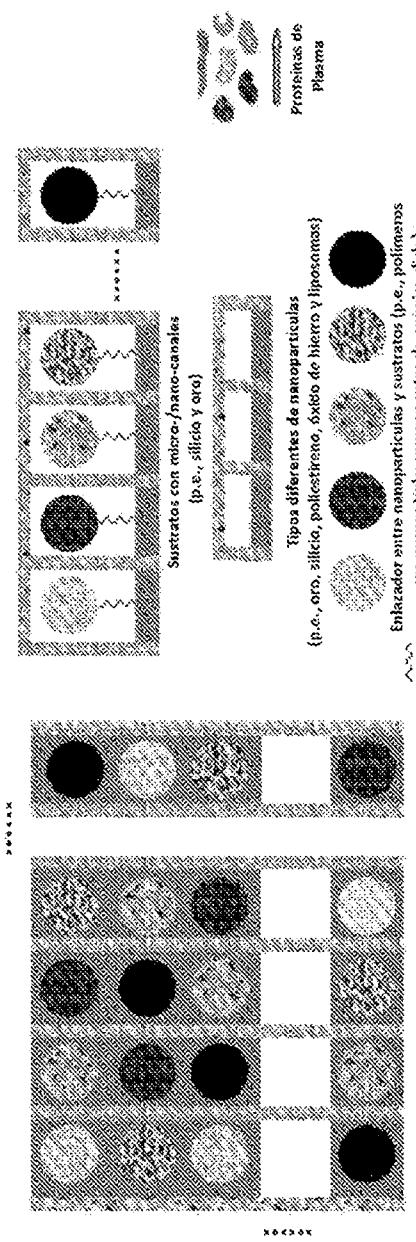


FIG. 21B Vista superior

Vista lateral

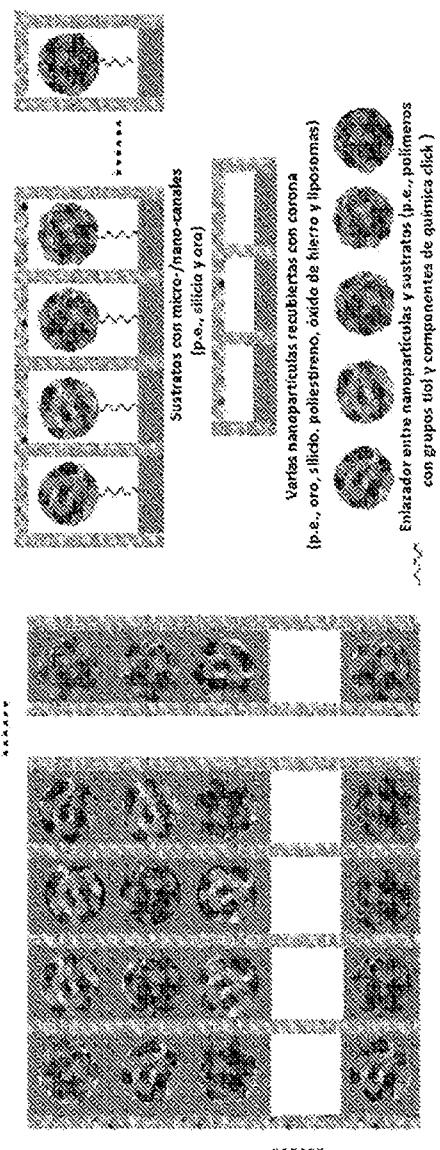


FIG. 22A Vista superior

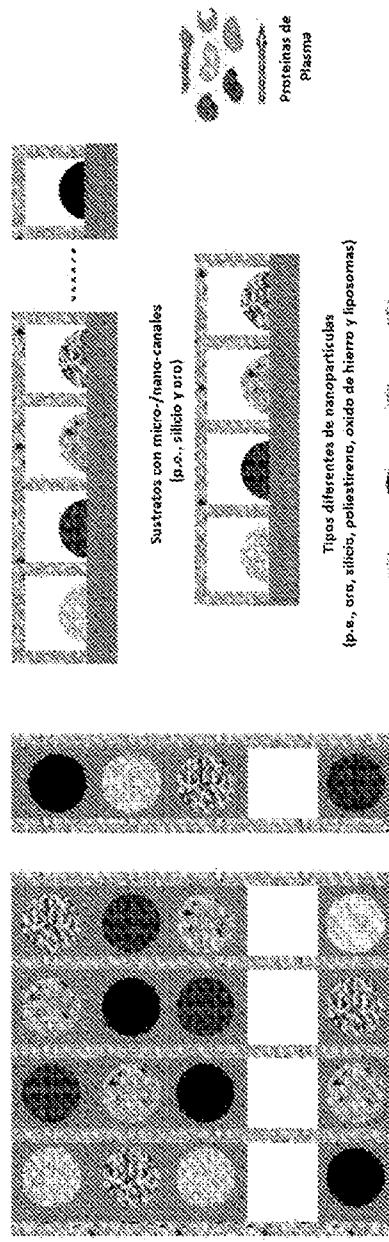
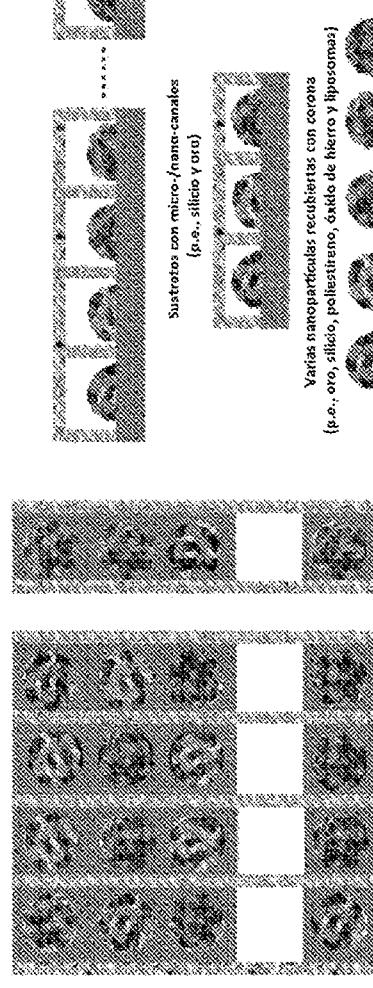


FIG. 22B Vista superior



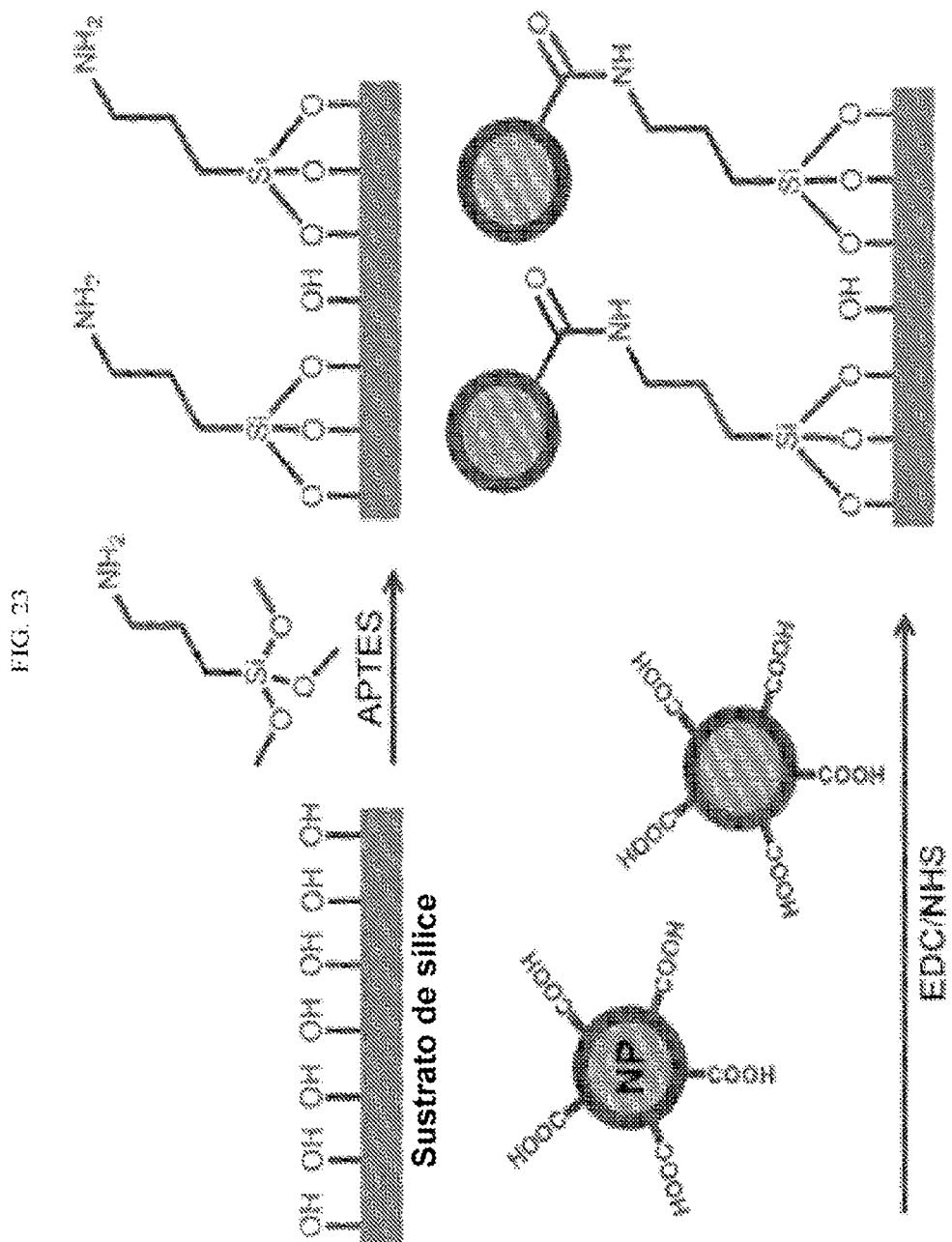
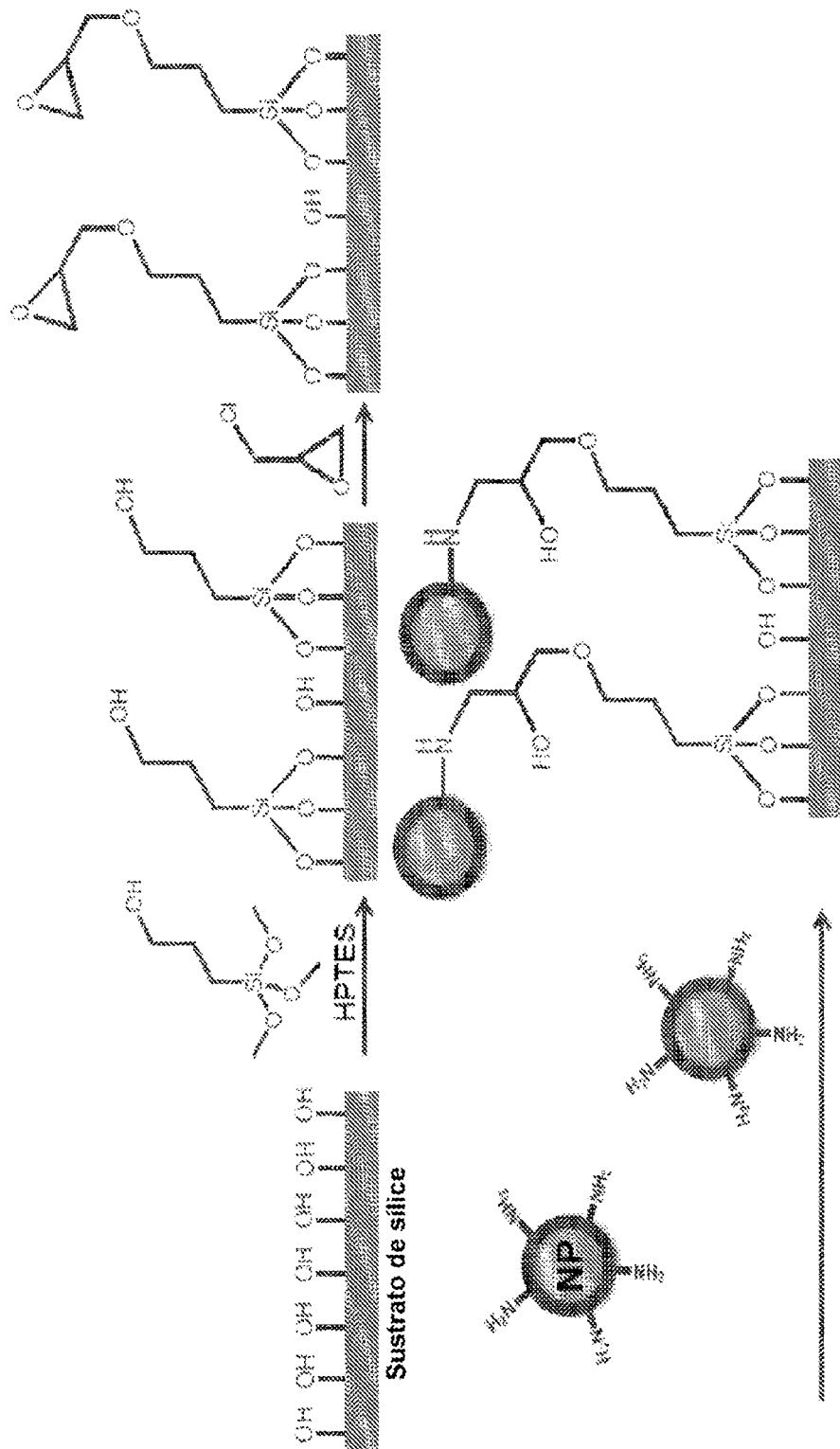
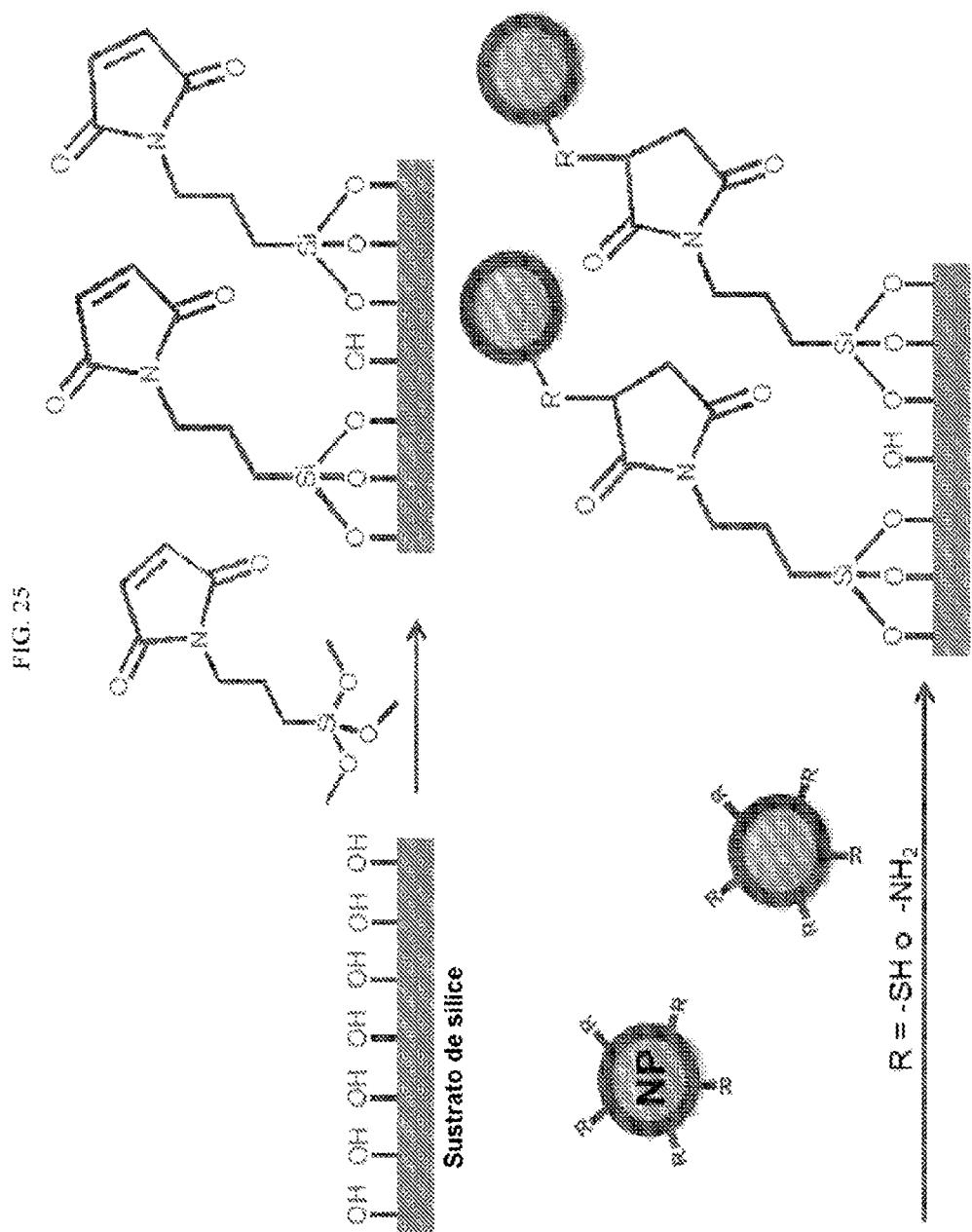


FIG. 24





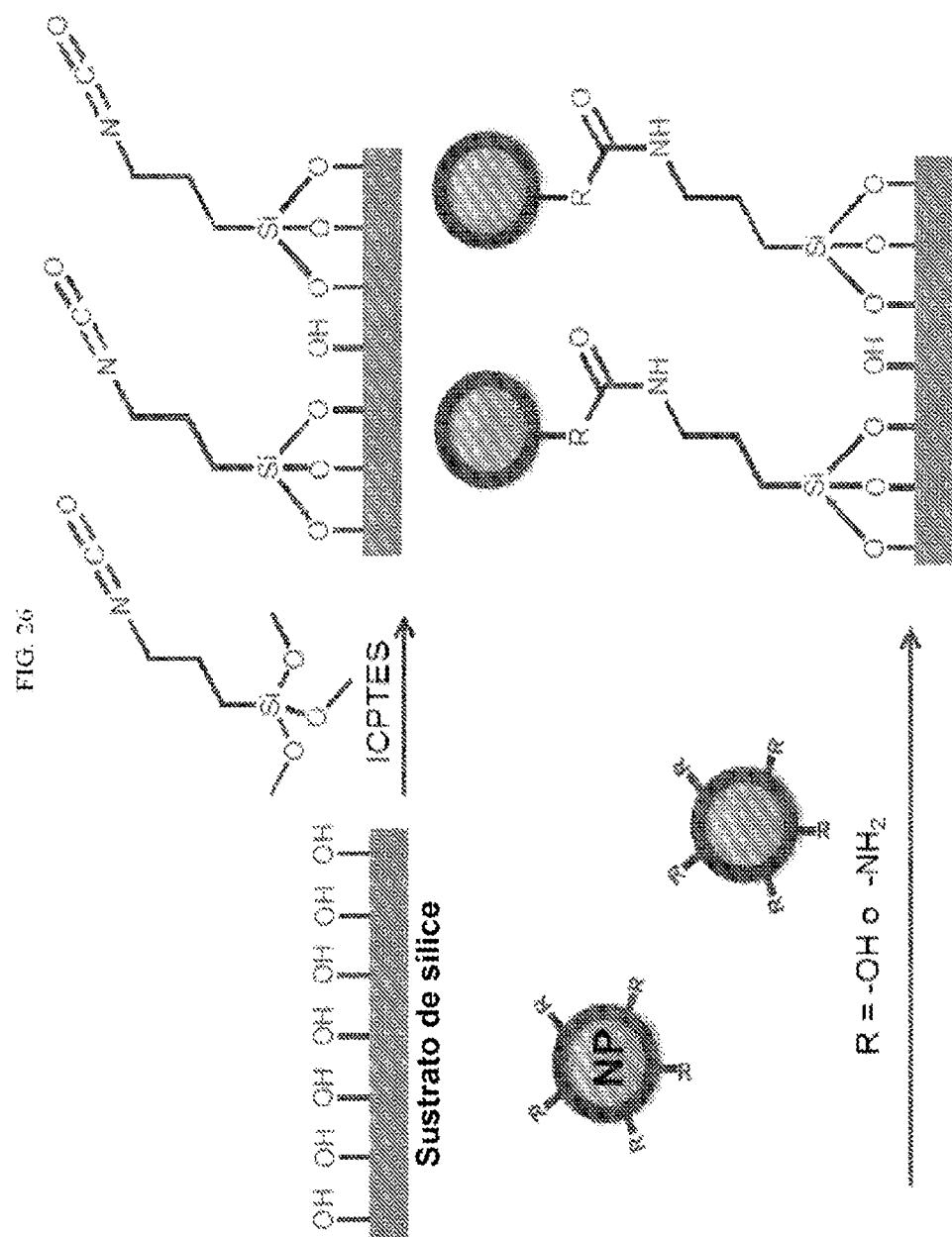


FIG. 27

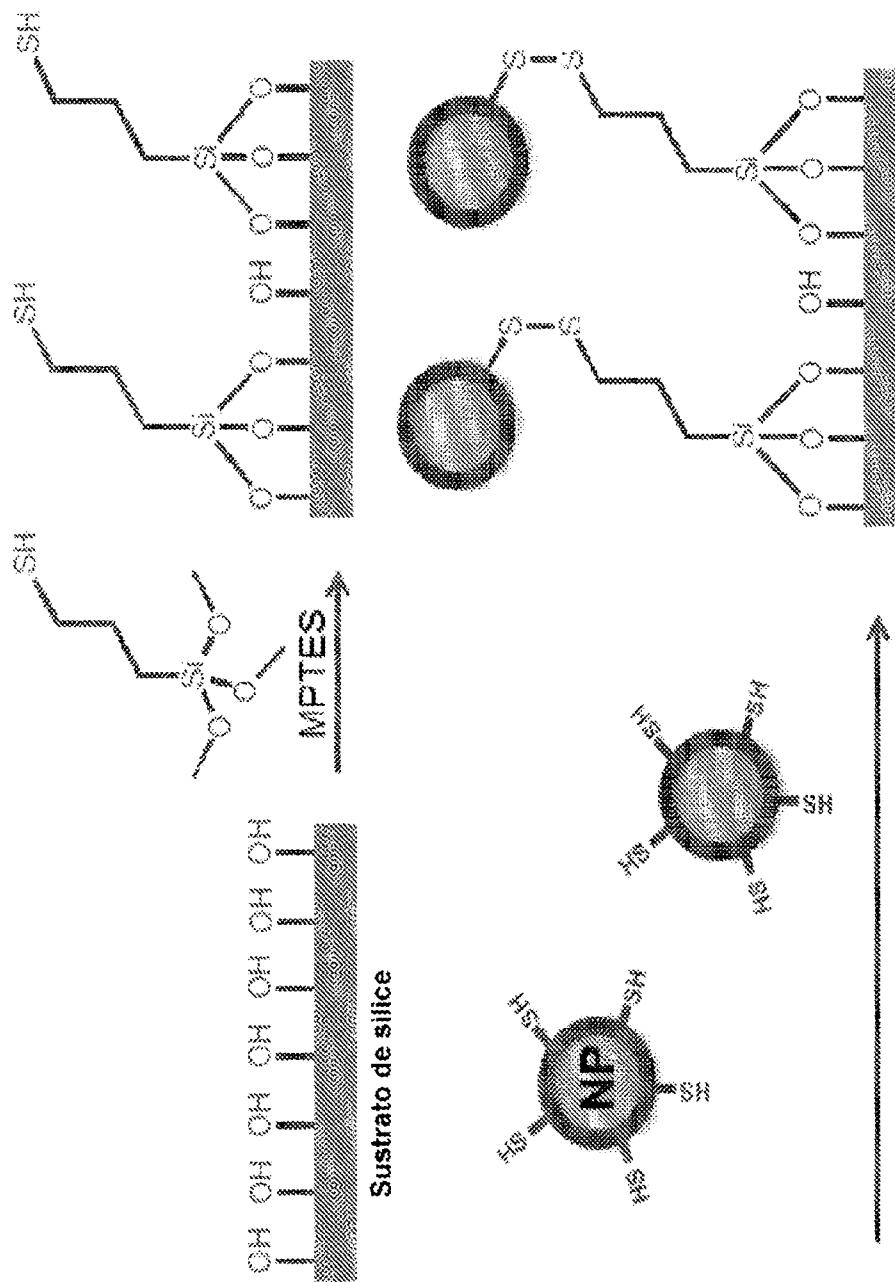


FIG. 23

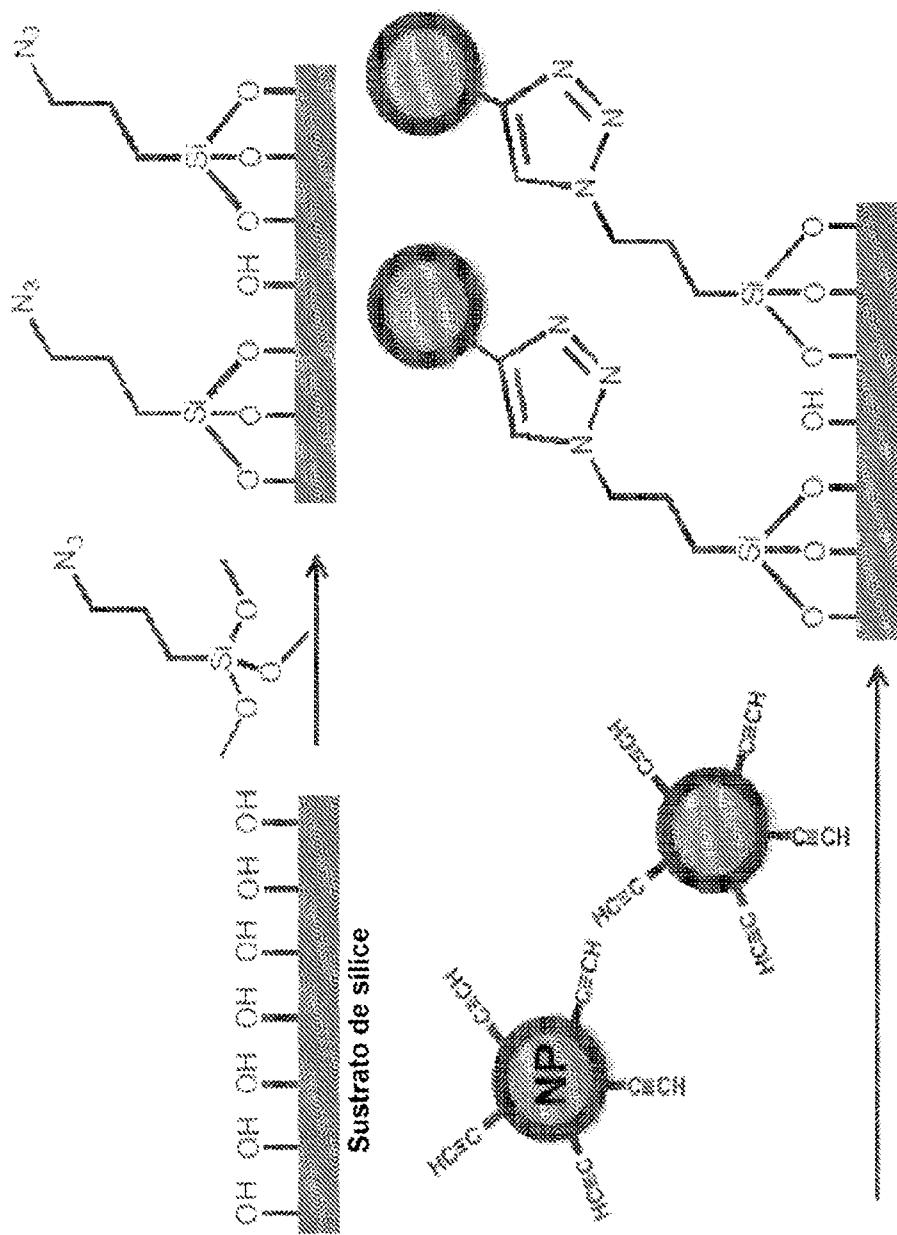


FIG. 29

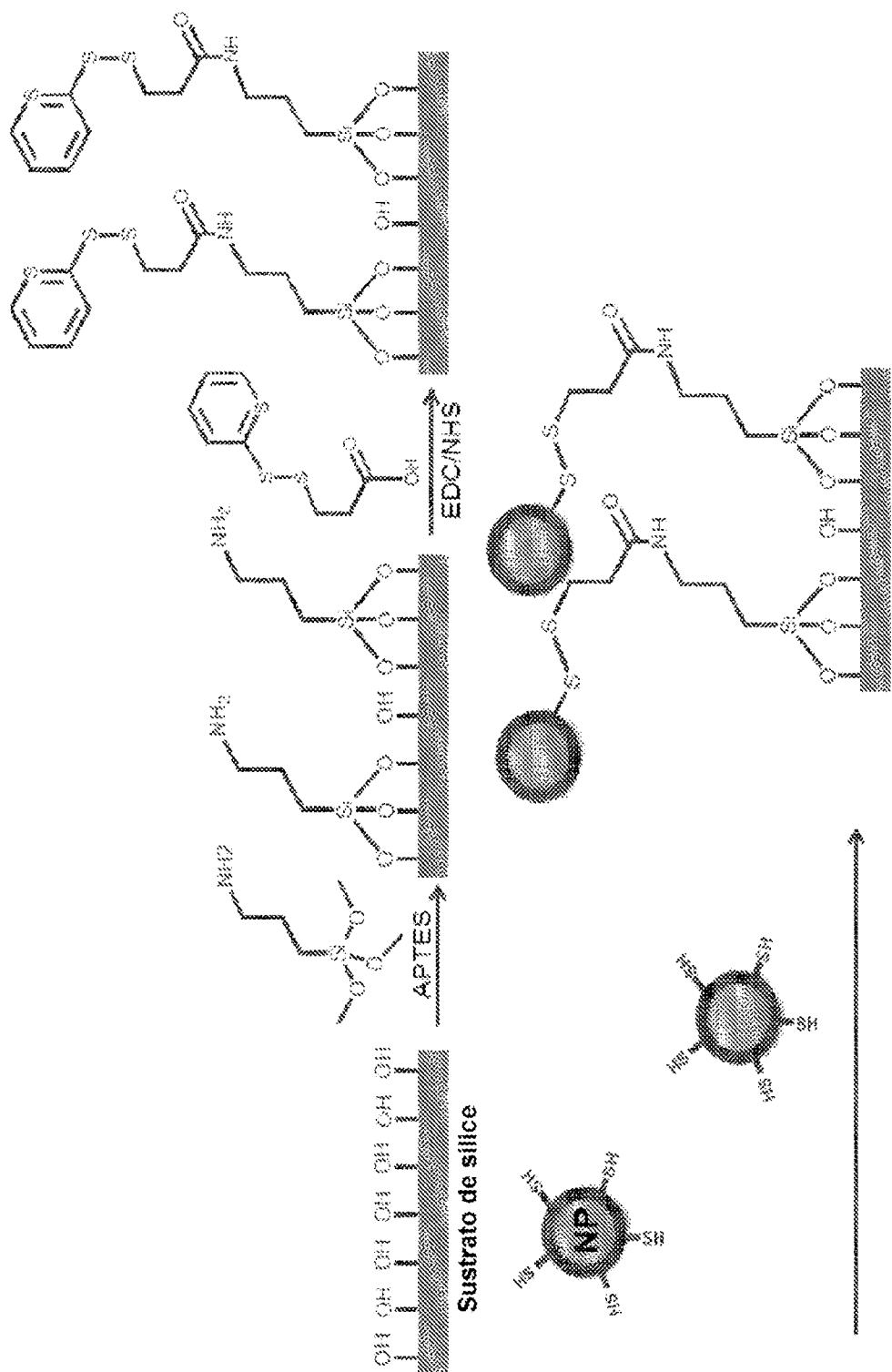


FIG. 30

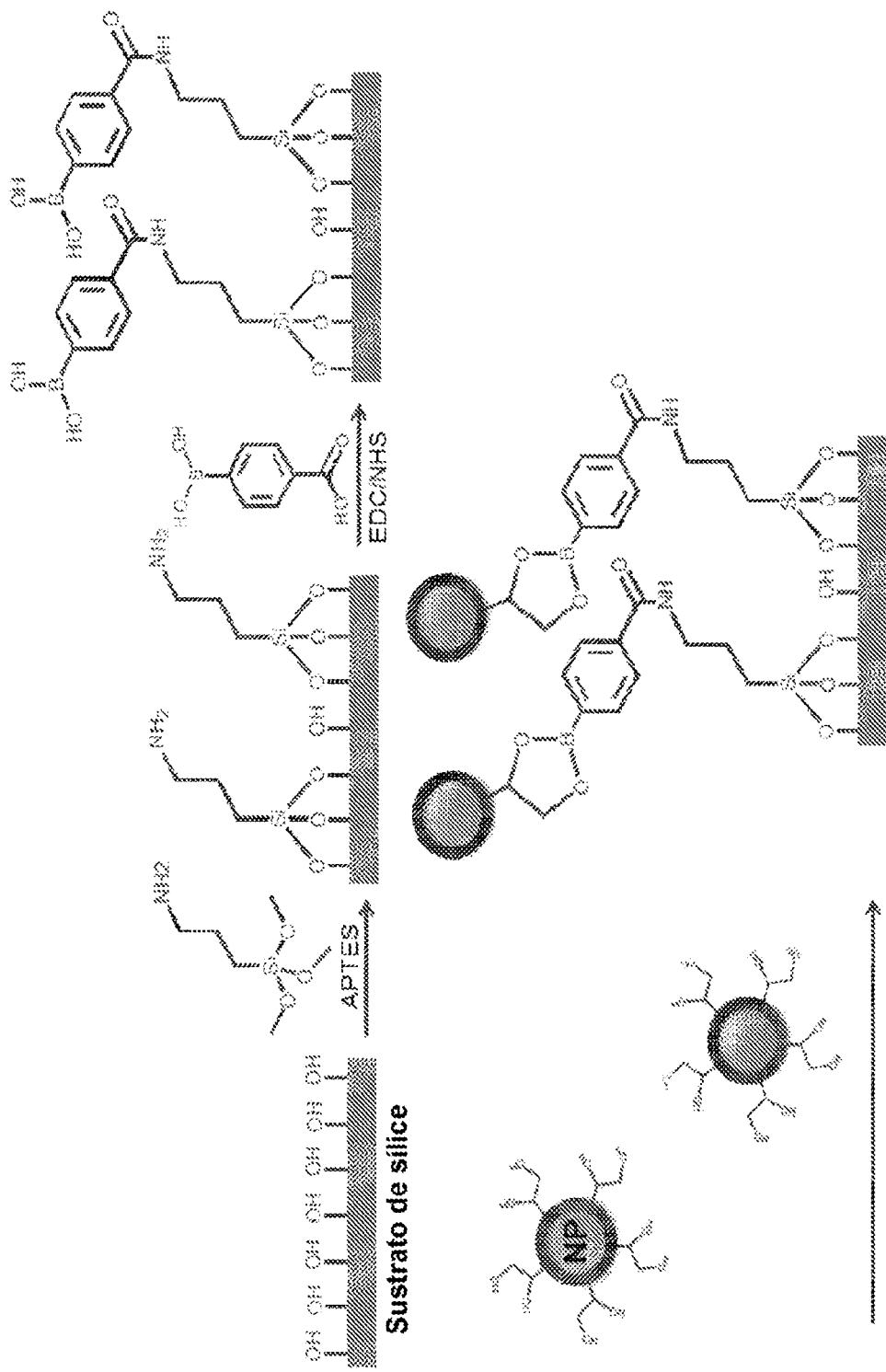
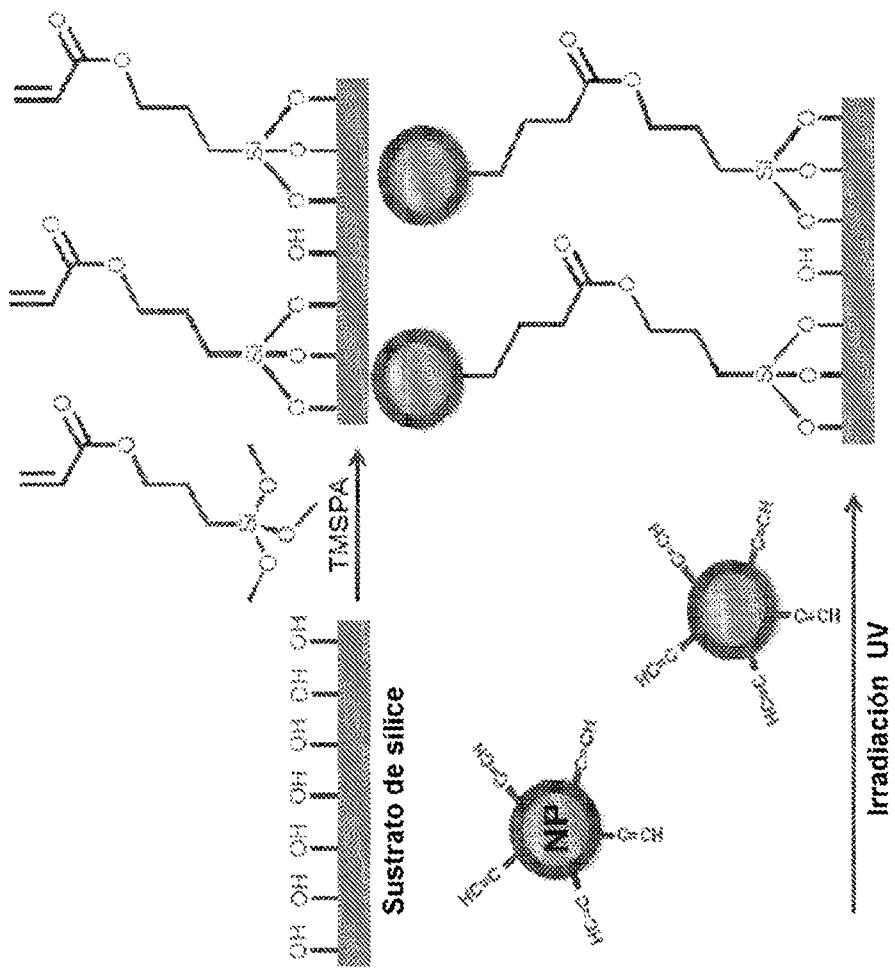


FIG. 31



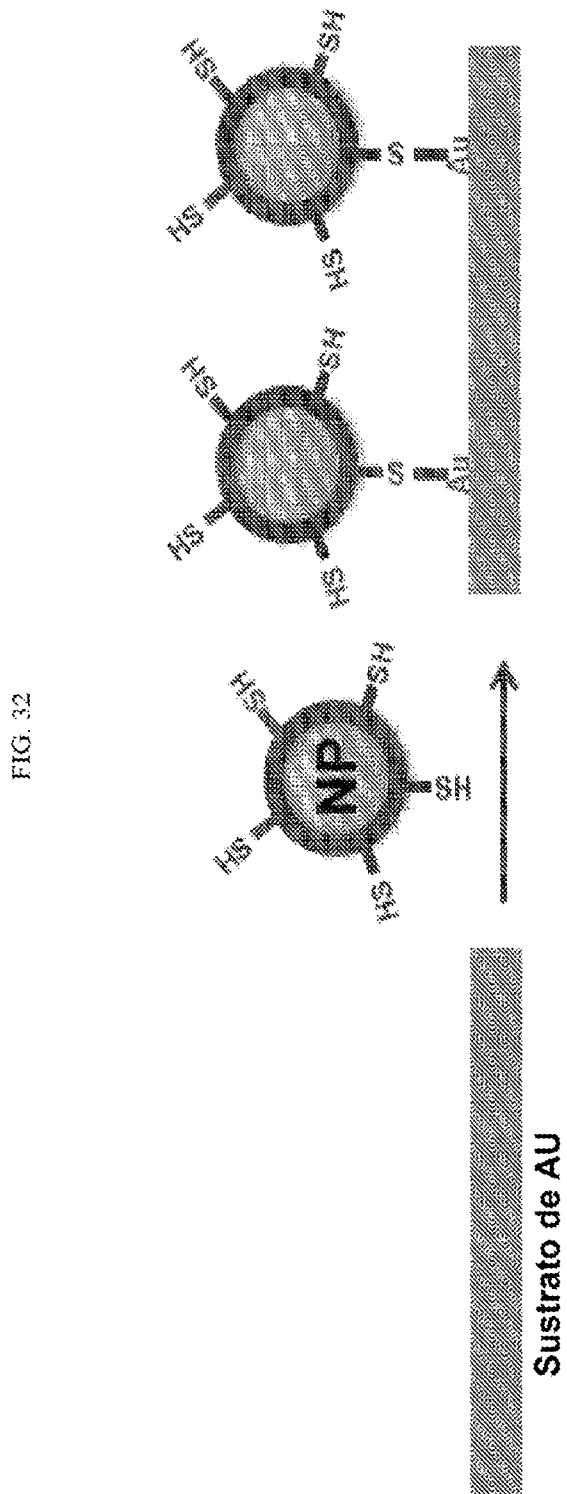


FIG. 33

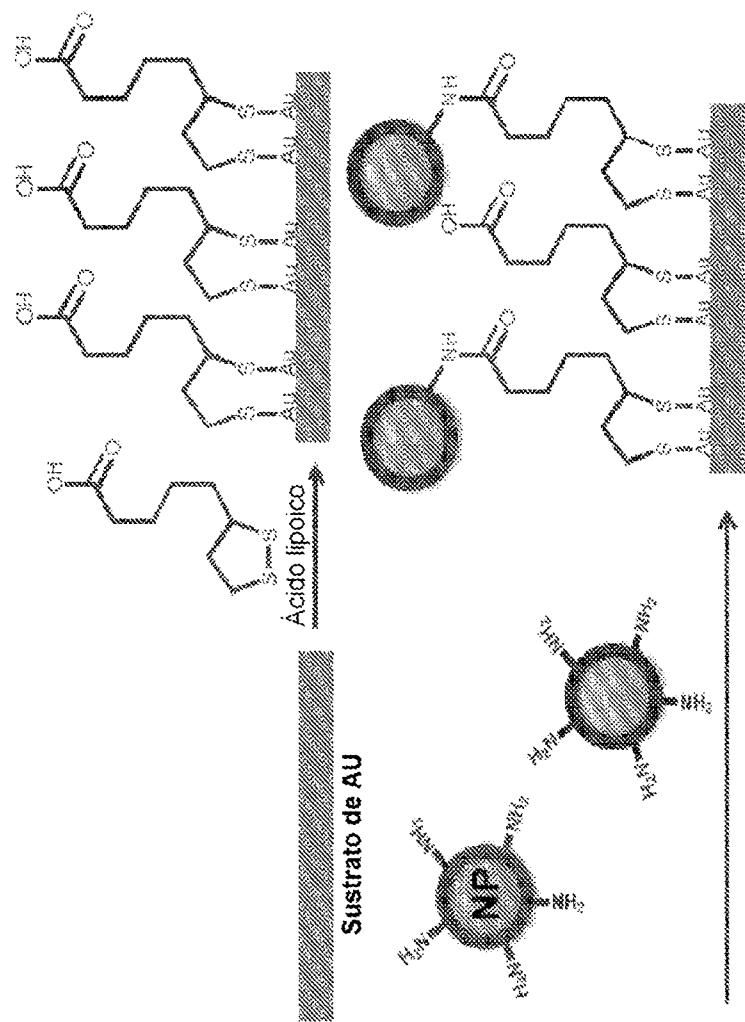


FIG. 34

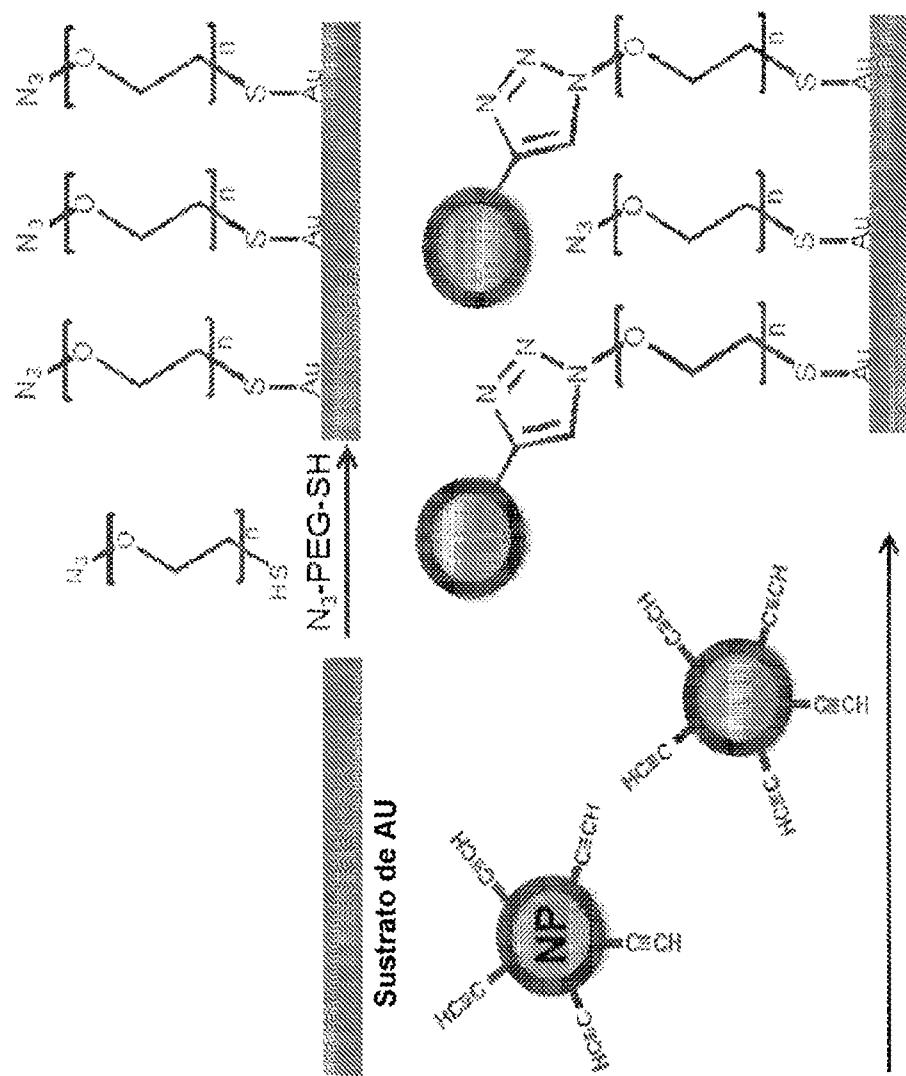


FIG. 35

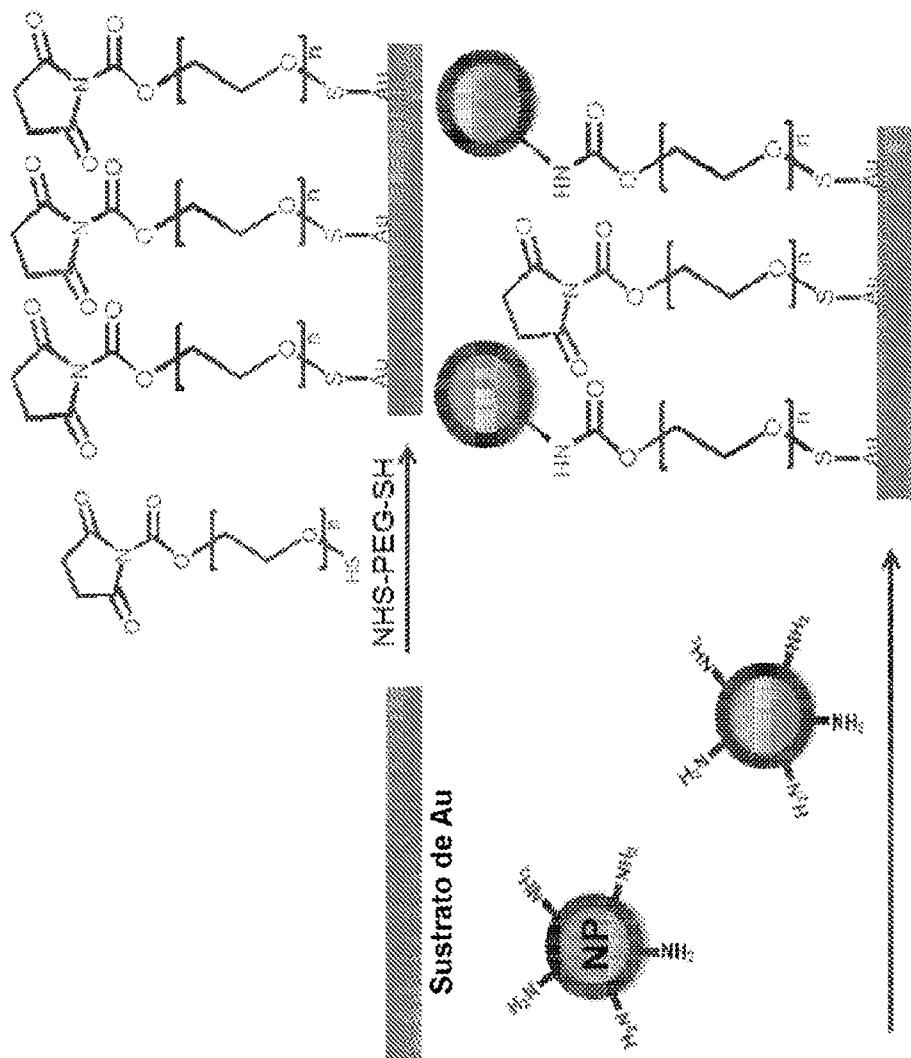


FIG. 36

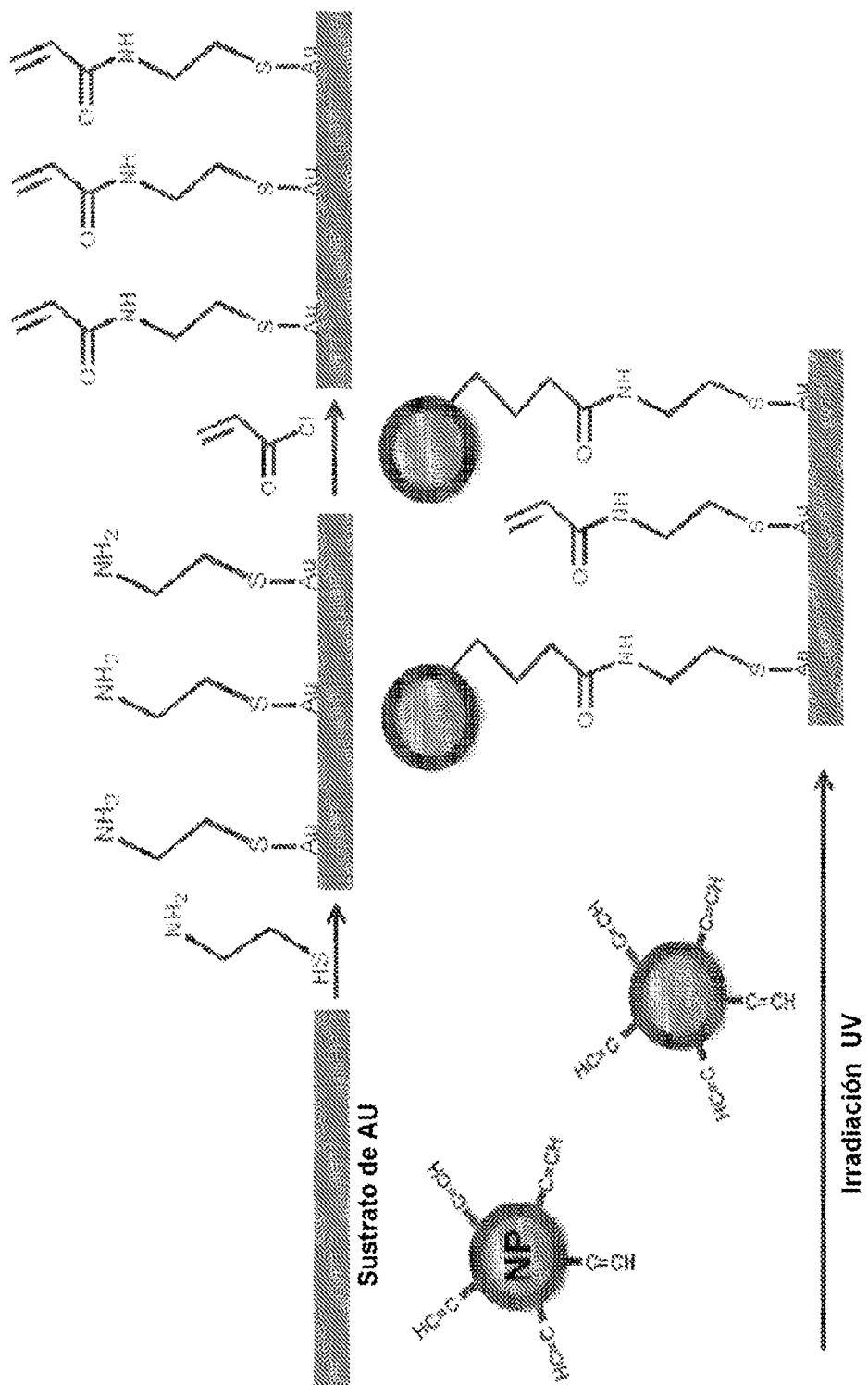


FIG. 37

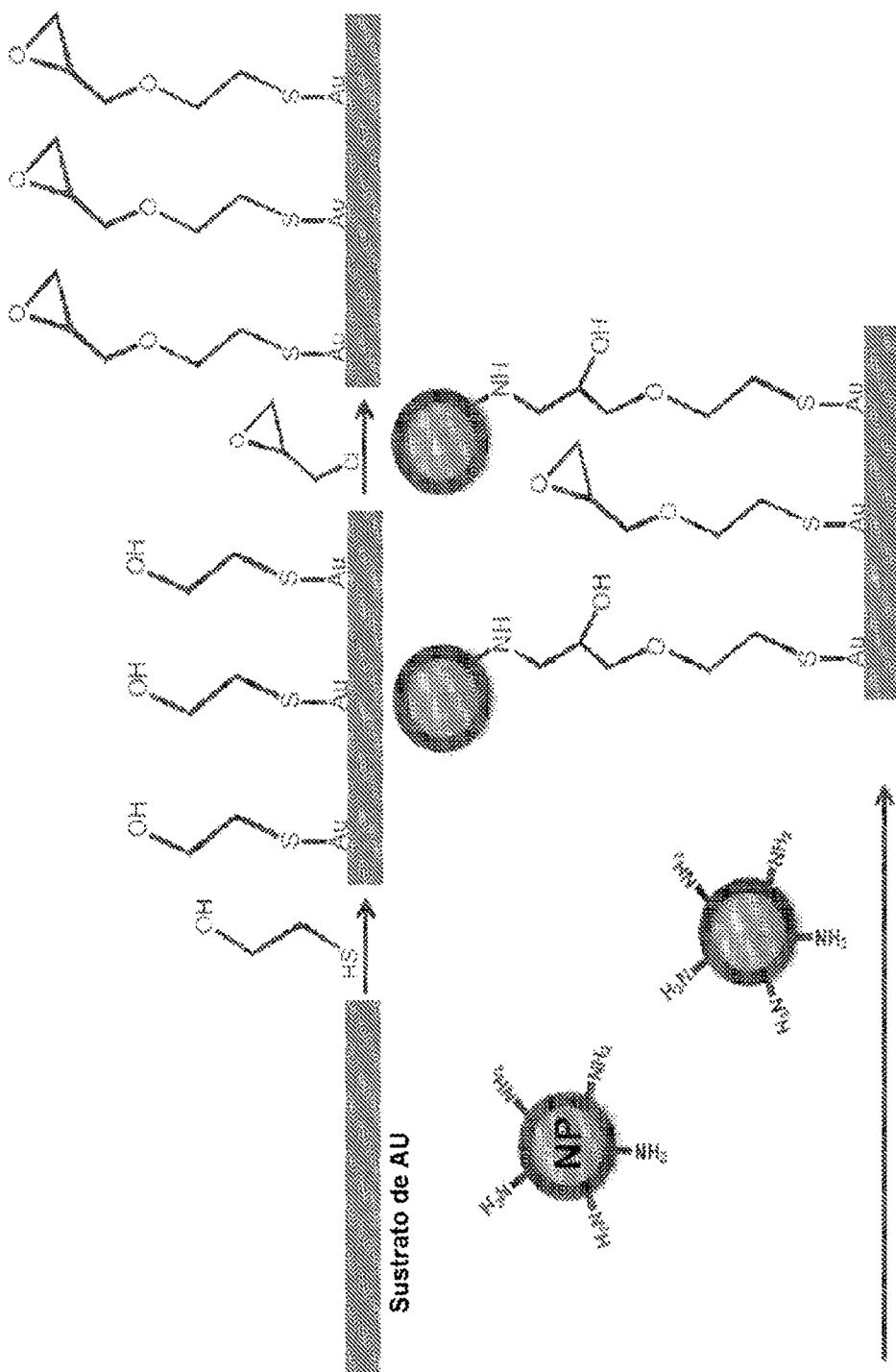


FIG. 38

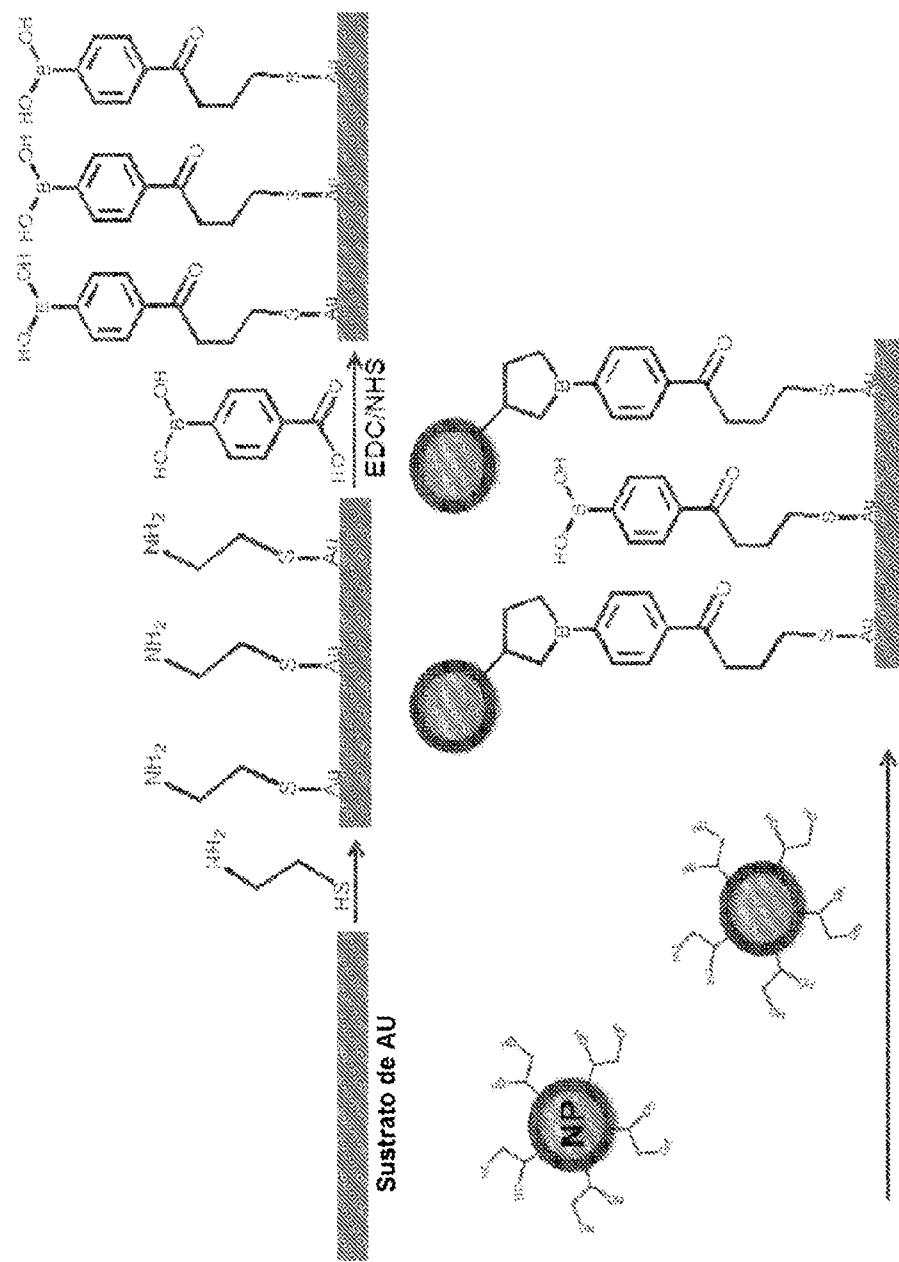


FIG. 39

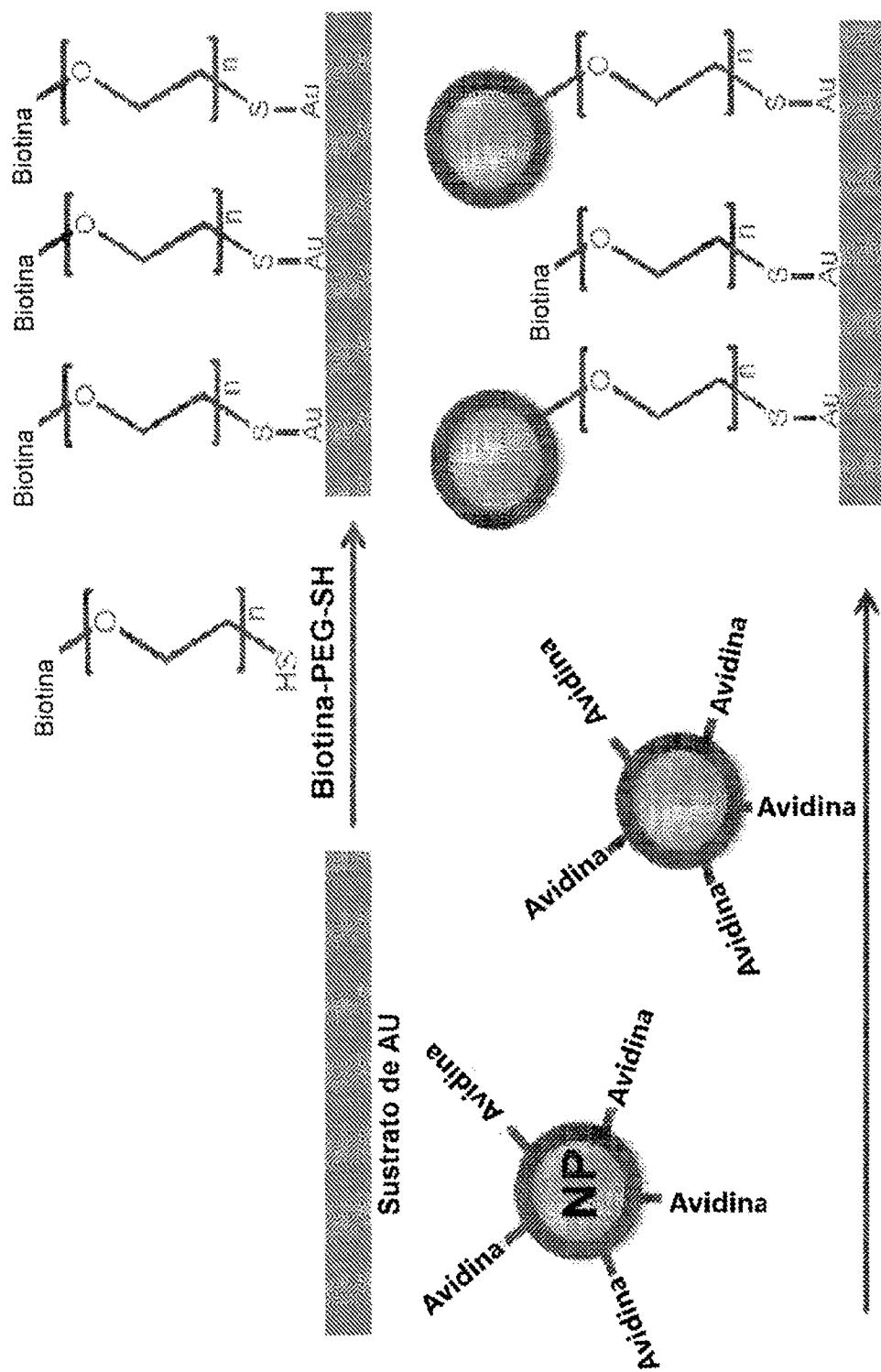
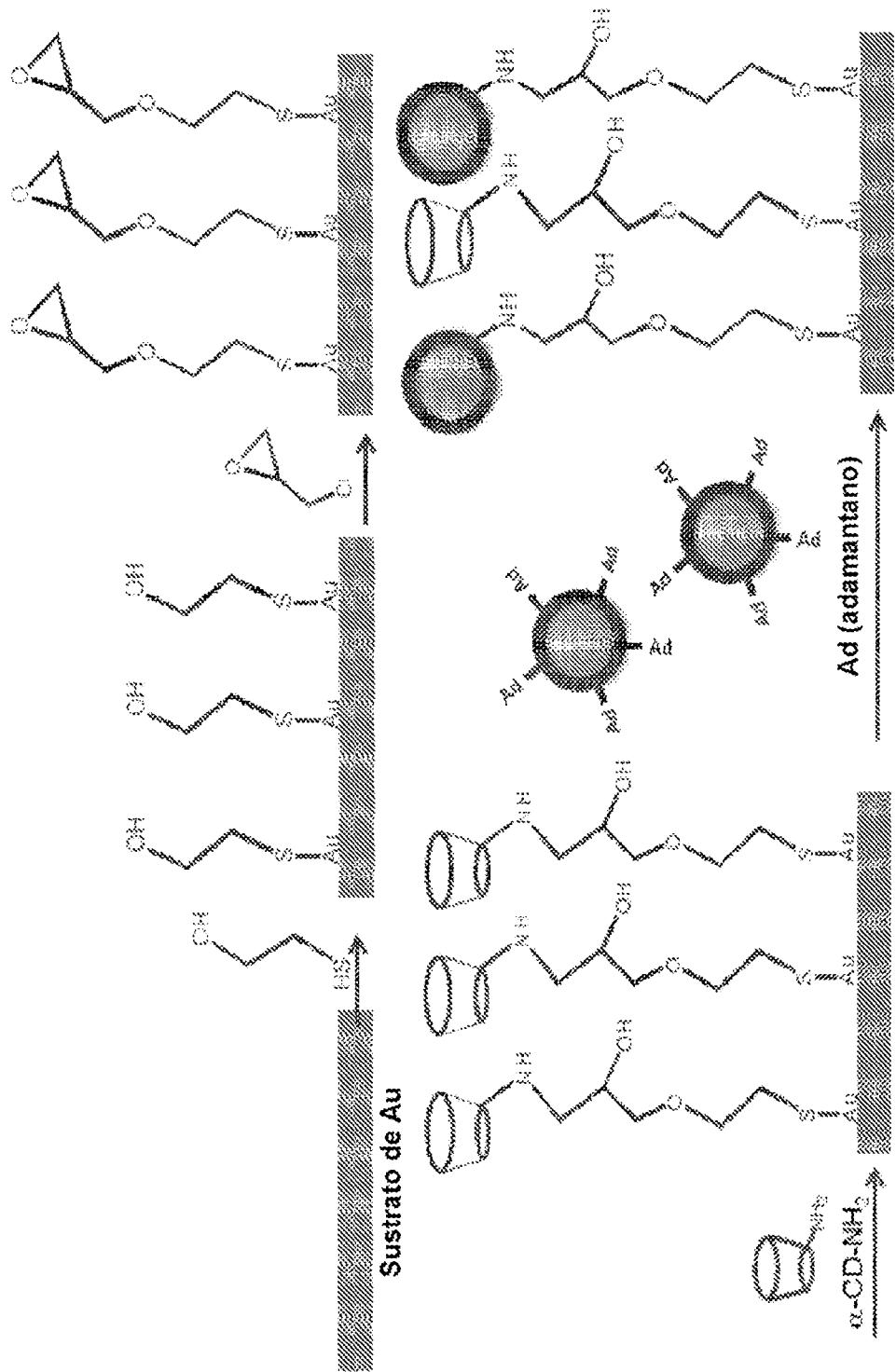
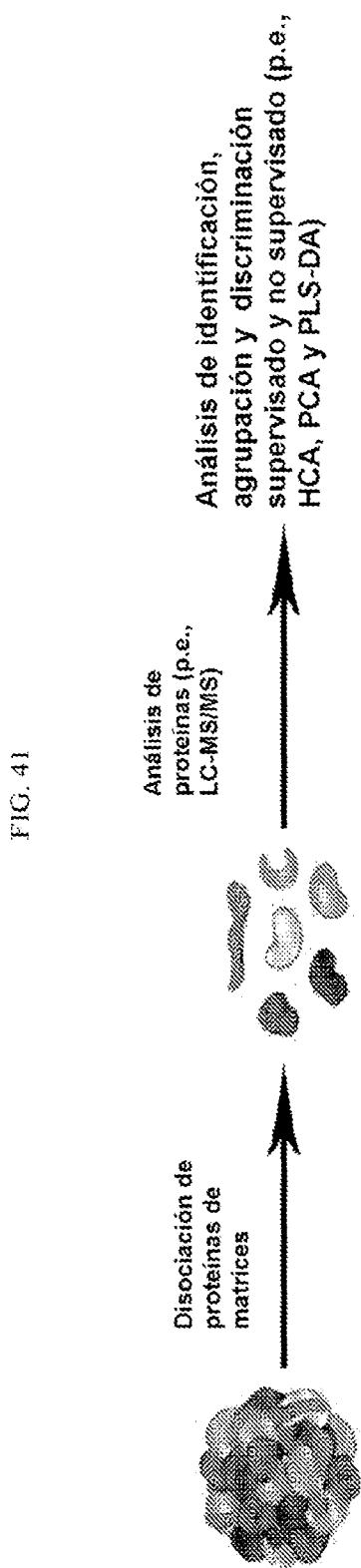
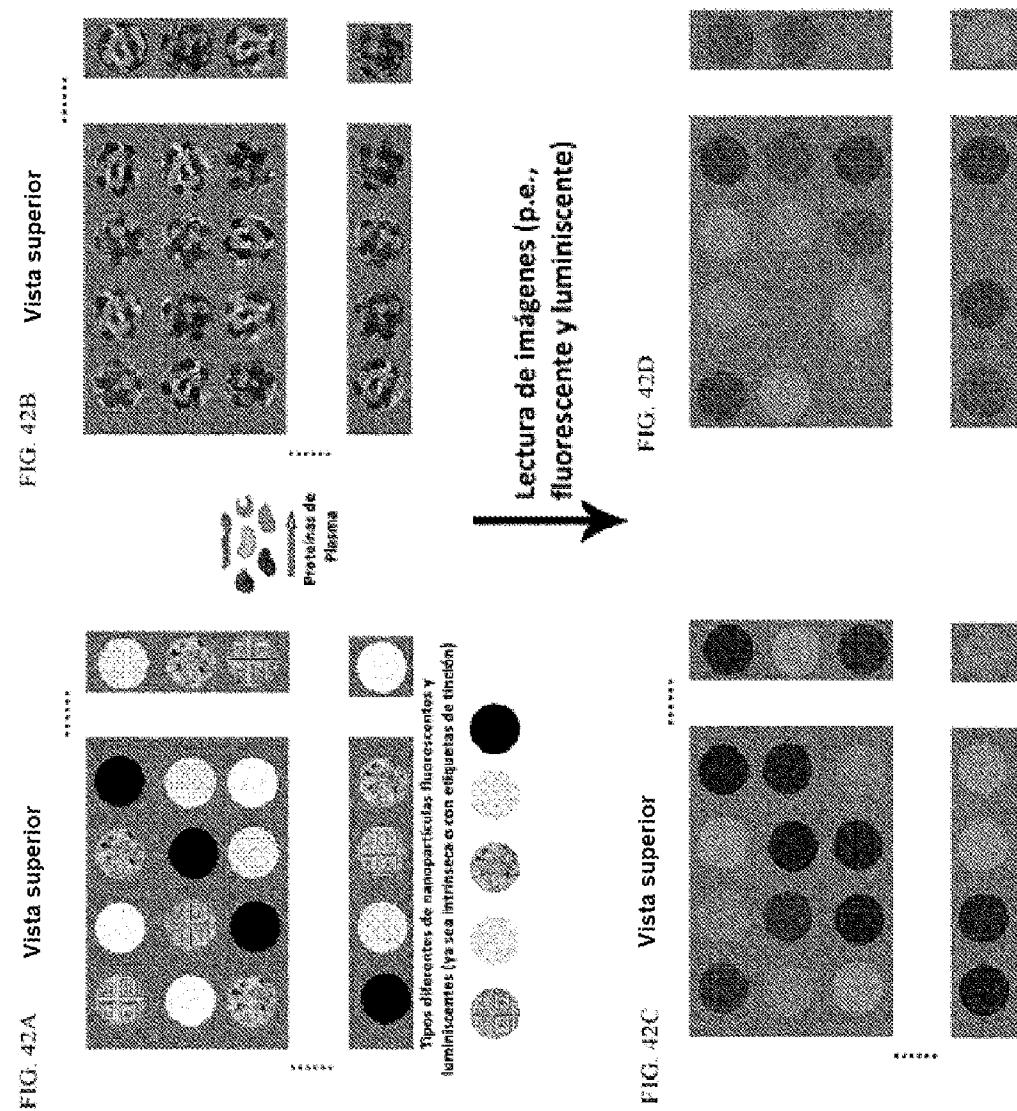


FIG. 40







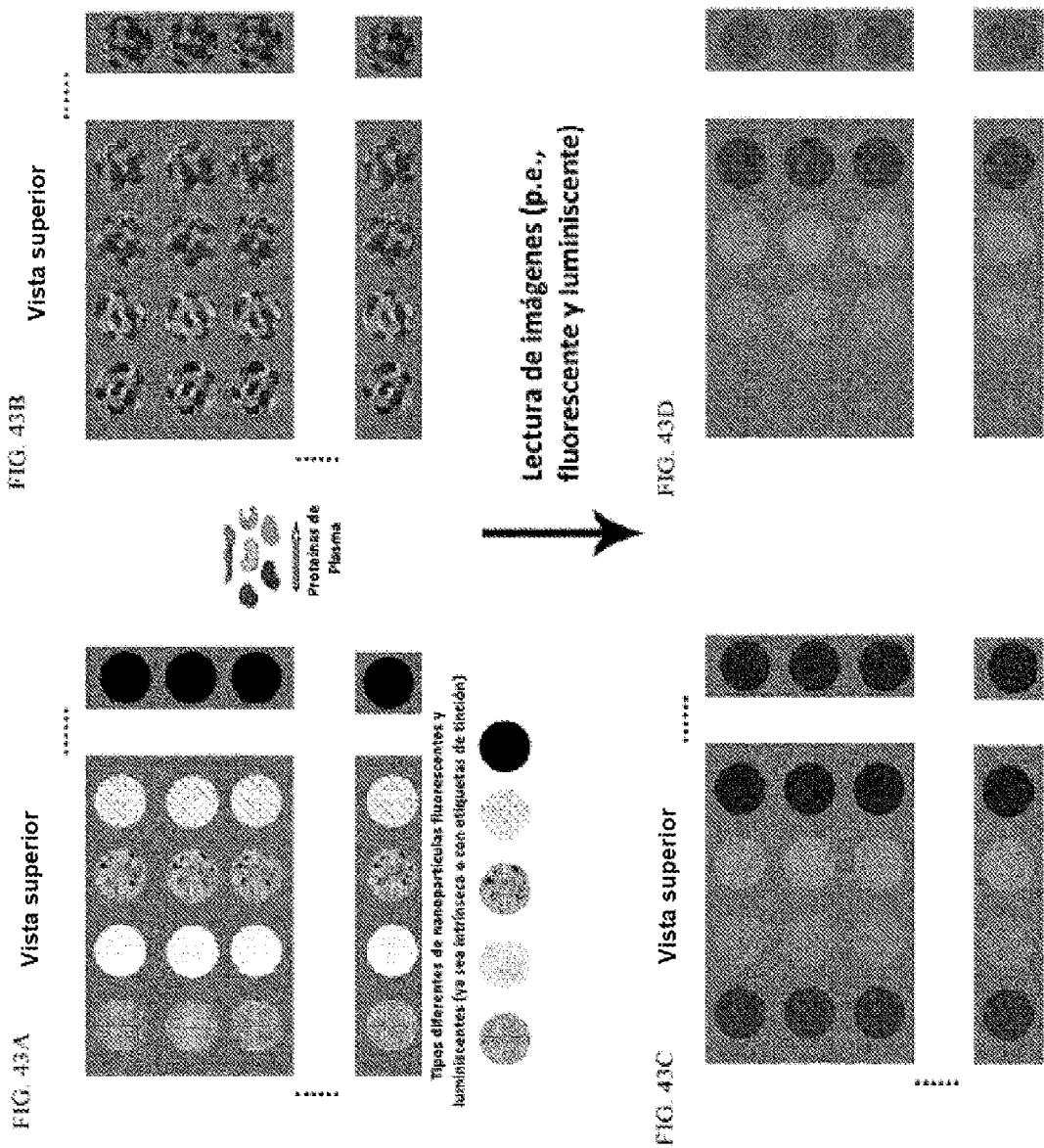


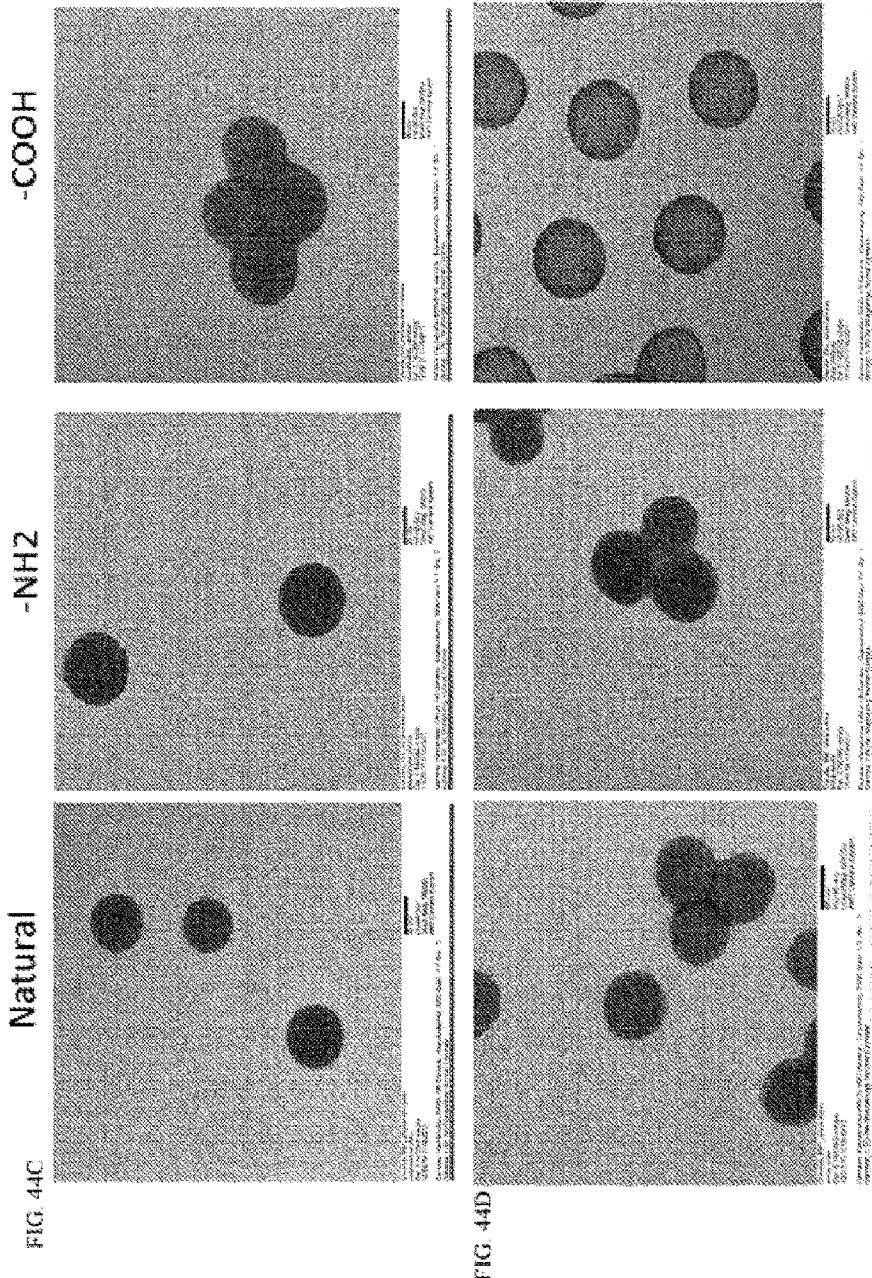
FIG. 44A

NPs de poliestireno

| | Tamaño (nm) | Potencial zeta (mV) | PDI |
|-----------------|-----------------|---------------------|-------|
| ○ | 98.6 \pm 1.1 | -42.27 \pm 1.35 | 0.09 |
| NH ₂ | 110.5 \pm 0.9 | -27.57 \pm 0.89 | 0.05 |
| COOH | 106.6 \pm 0.5 | -49.38 \pm 1.27 | 0.067 |

FIG. 44B

| | Tamaño (nm) | Potencial zeta (mV) | PDI |
|-----------------|-----------------|---------------------|--------|
| ○ | 100.5 \pm 1.6 | -53.72 \pm 1.35 | 0.0035 |
| NH ₂ | 93.9 \pm 0.9 | -46.44 \pm 0.92 | 0.055 |
| COOH | 120 \pm 0.7 | -55.77 \pm 1.4 | 0.02 |



Imágenes TEM. Caracterización morfológica y estructural de NPs desnudas usanfo teñidas usando acetato de uracilo. En este ejemplo, se empleó plasma saludable humano.

FIG. 45A

NPs de poliestireno

| Tamaño (nm) | Potencial zeta (mV) | pDl |
|-------------|---------------------|-------|
| 148.4 ± 1.4 | -25.9 ± 0.32 | 0.049 |
| 148.9 ± 4 | -26.25 ± 1.59 | 0.035 |
| 125.7 ± 1.8 | -21.6 ± 1.19 | 0.064 |

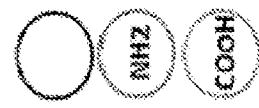


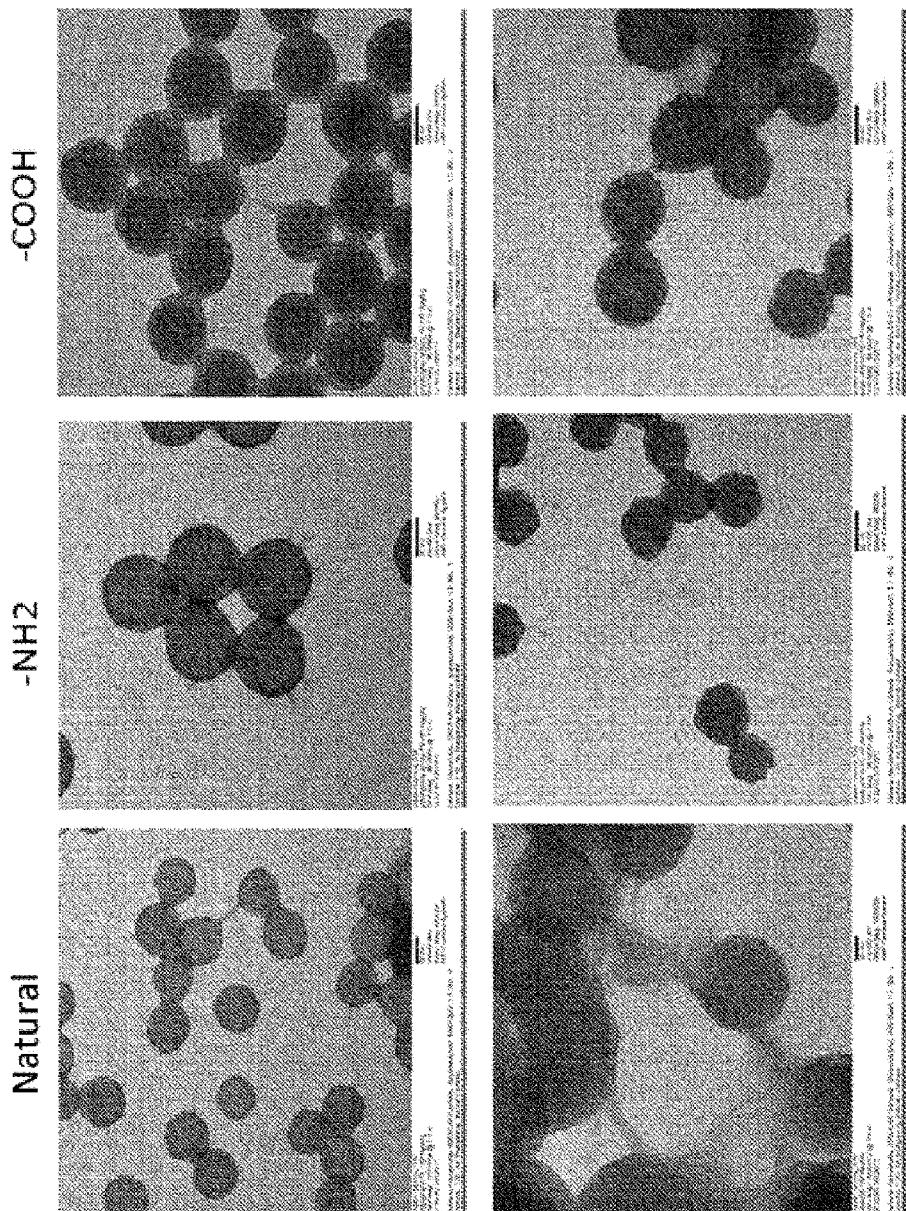
FIG. 45B

NPs de silice

| Tamaño (nm) | Potencial zeta (mV) | pDl |
|-------------|---------------------|-------|
| 225.6 ± 2.3 | -23.4 ± 1.68 | 0.283 |
| 198.2 ± 1.7 | -17.04 ± 0.71 | 0.179 |
| 175.1 ± 2.4 | -10.02 ± 0.71 | 0.173 |



Natural FIG. 45C



450

Imágenes TEM. Caracterización morfológica y estructural de NPs desnudas usanfó teñidas usando acetato de uracilo. En este ejemplo, se empleó plasma saludable humano.

FIG 46

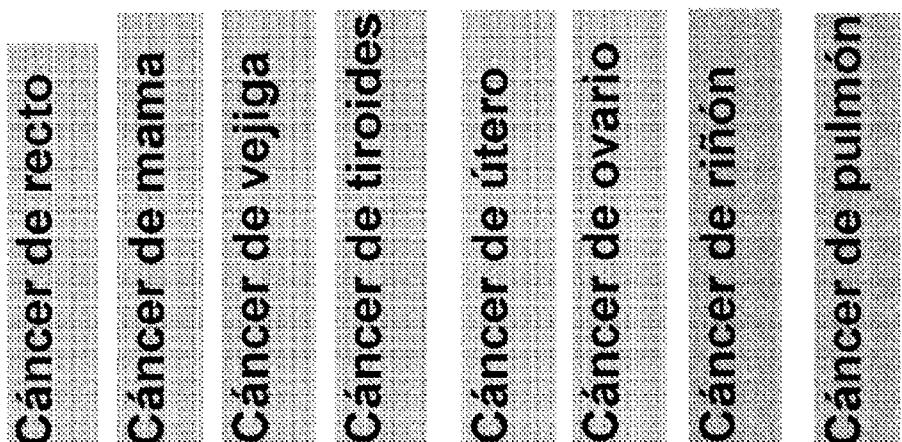


FIG. 47

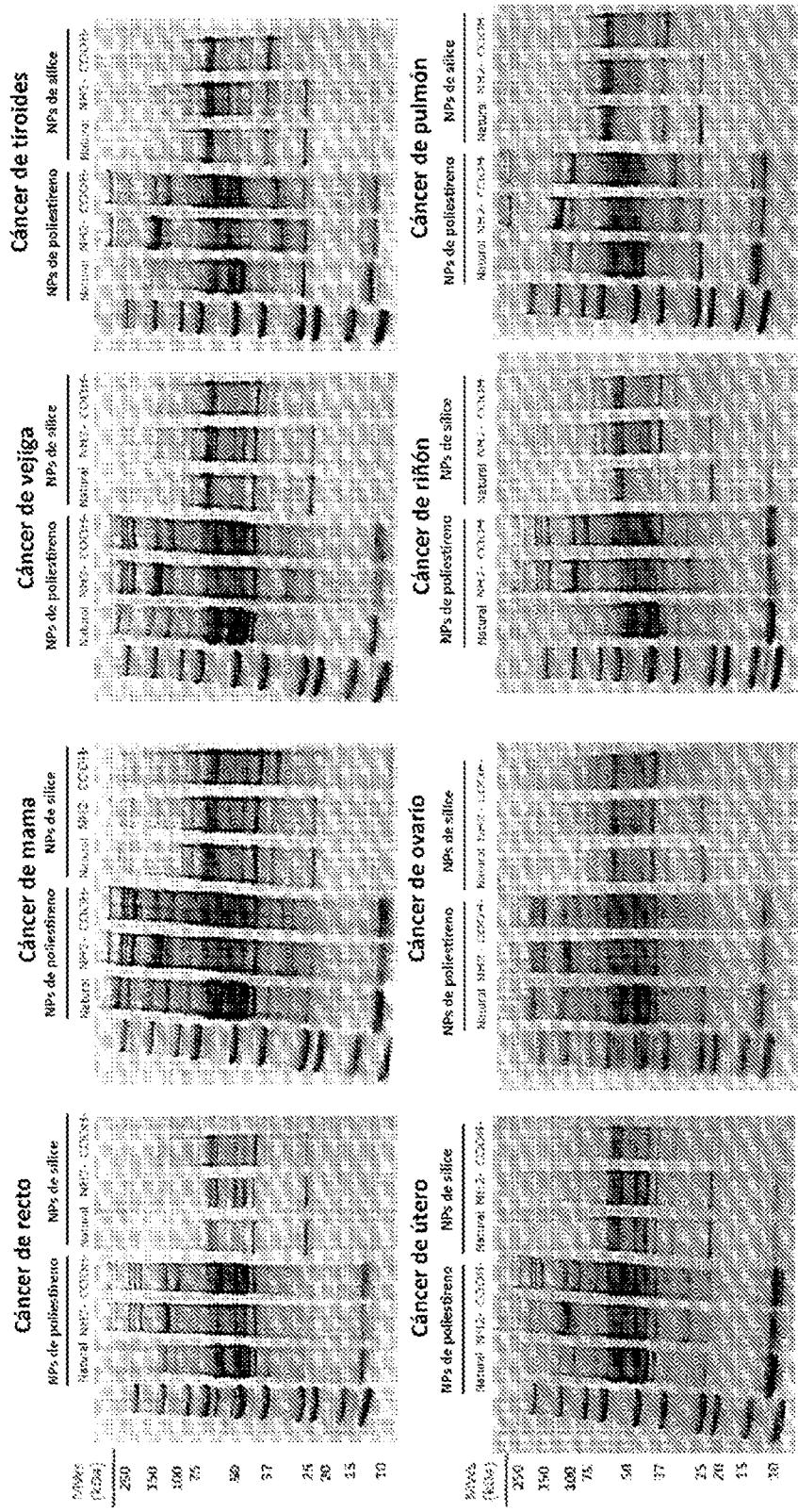
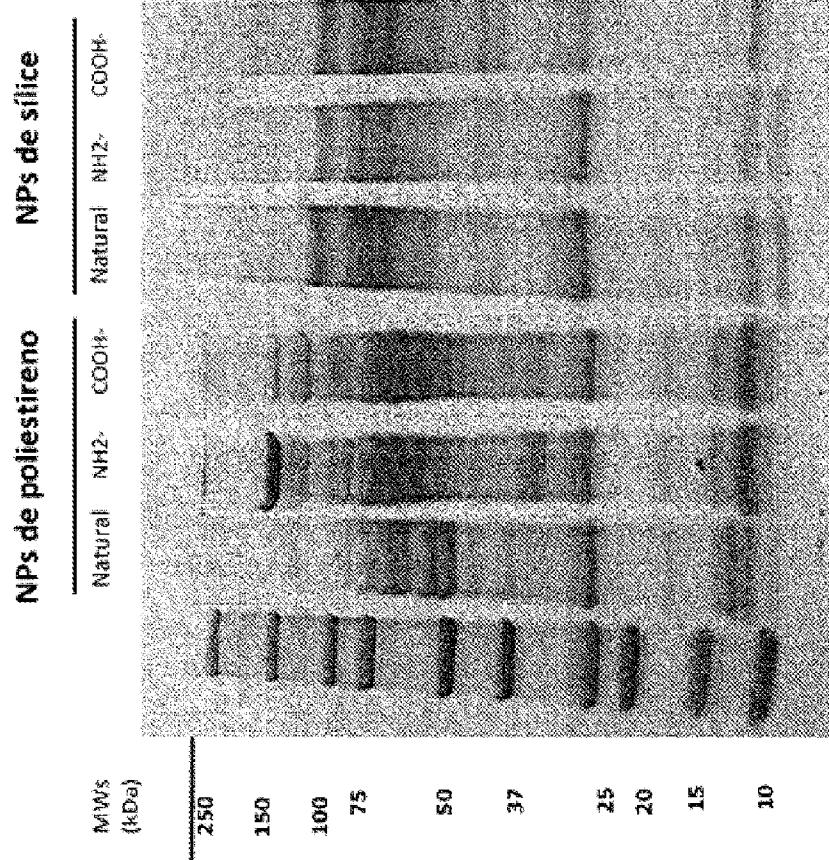


FIG. 48

Plasma saludable

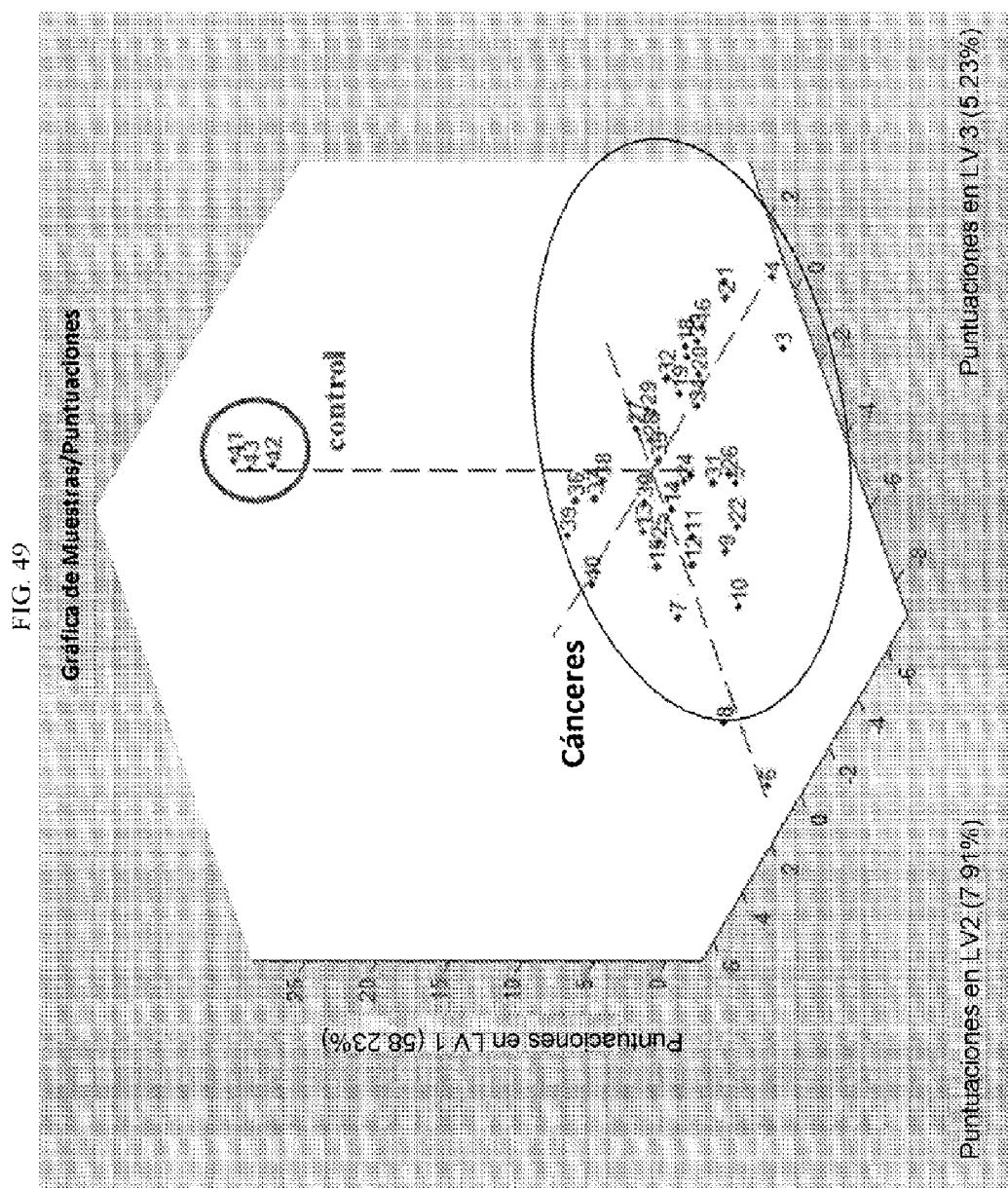
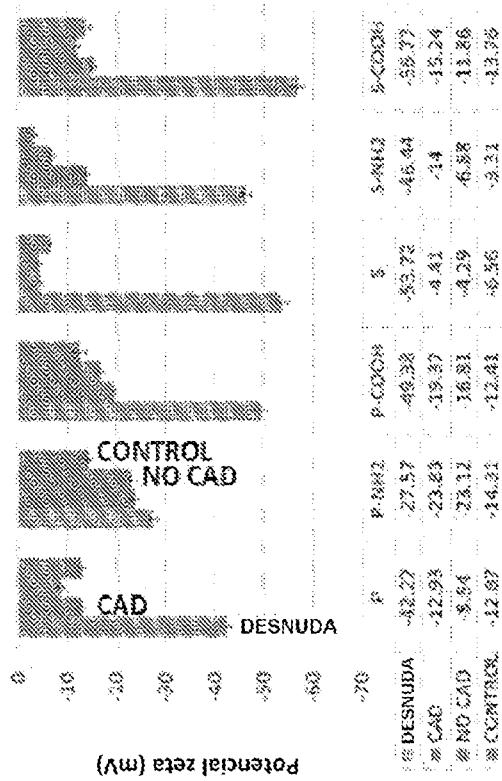
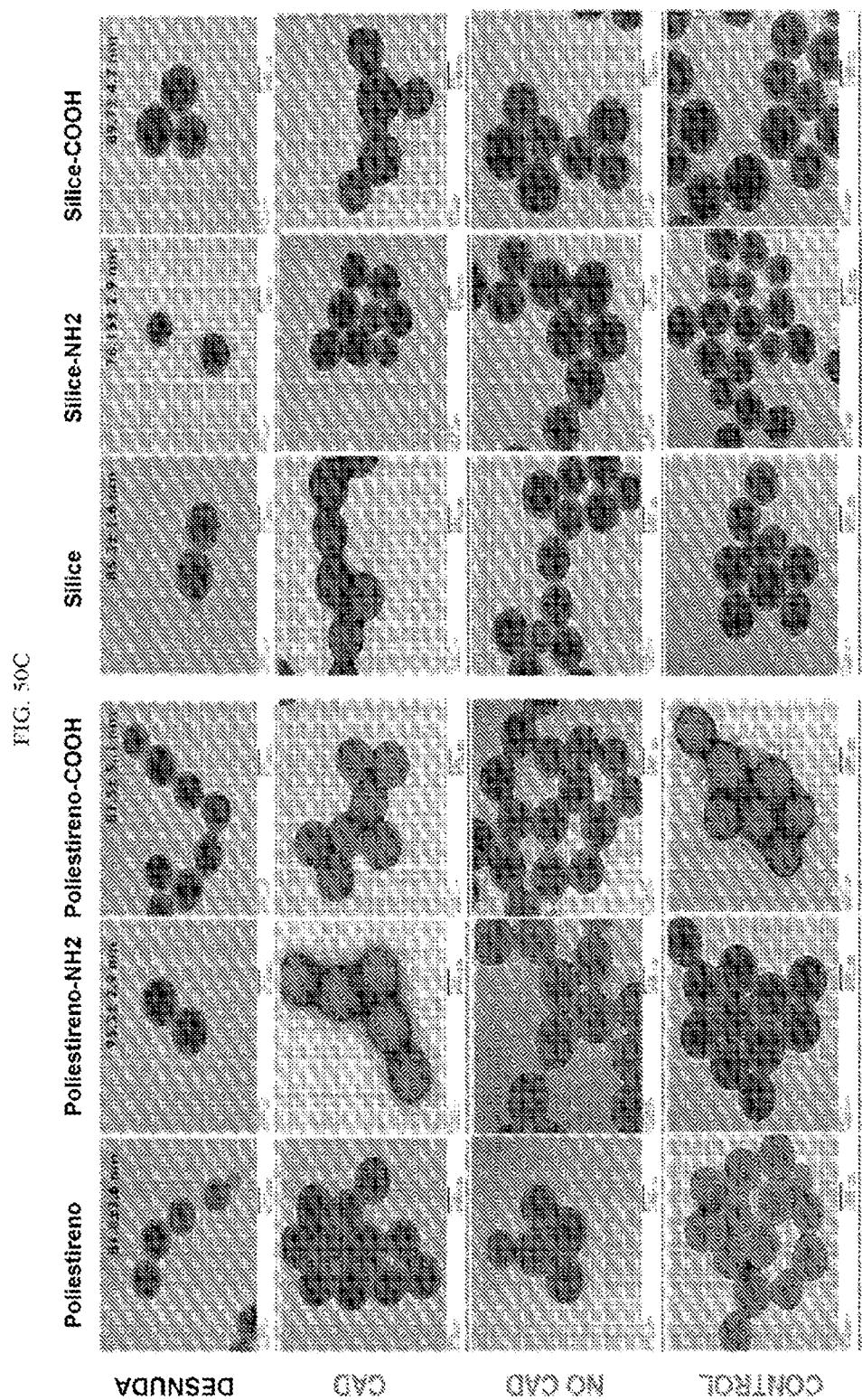


FIG. 50A

| NPs | DESNUDA | | | CAD | | | NO CAD | | | CONTROL | | |
|--------|-------------|-----|-------|-------------|-----|-------|-------------|------|-------|-------------|-----|-------|
| | Tamaño (nm) | SD | PD1 | Tamaño (nm) | SD | PD1 | Tamaño (nm) | SD | PD1 | Tamaño (nm) | SD | PD1 |
| P | 98.6 | 1.1 | 0.09 | 153.6 | 3.2 | 0.24 | 187.5 | 10.7 | 0.16 | 146.4 | 1.4 | 0.288 |
| P-NH2 | 110.5 | 0.9 | 0.05 | 149.5 | 1.8 | 0.302 | 145.3 | 6.3 | 0.086 | 148.9 | 4 | 0.129 |
| P-COOH | 108.6 | 0.5 | 0.067 | 140.8 | 1.2 | 0.216 | 138.1 | 1.3 | 0.327 | 125.7 | 1.8 | 0.268 |
| S | 160.5 | 1.6 | 0.035 | 148.6 | 3.9 | 0.137 | 183.7 | 1.3 | 0.236 | 138.6 | 2.1 | 0.005 |
| S-NH2 | 93.8 | 0.9 | 0.035 | 175.7 | 3.3 | 0.238 | 114.7 | 14.5 | 0.444 | 140.9 | 6.8 | 0.007 |
| S-COOH | 120 | 0.7 | 0.02 | 178.4 | 2.1 | 0.294 | 171.3 | 13.7 | 0.196 | 175.1 | 3.4 | 0.17 |

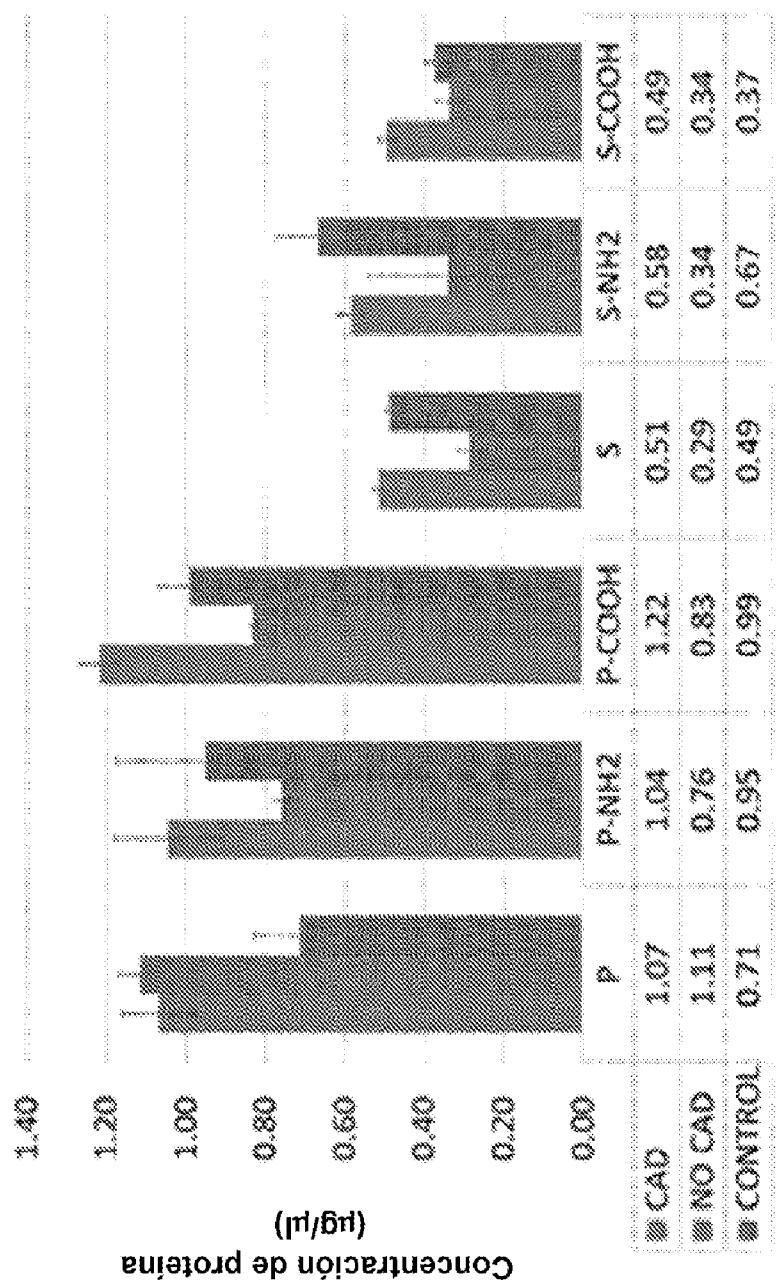
FIG. 50B





Ensayo de Bradford

FIG. 51A



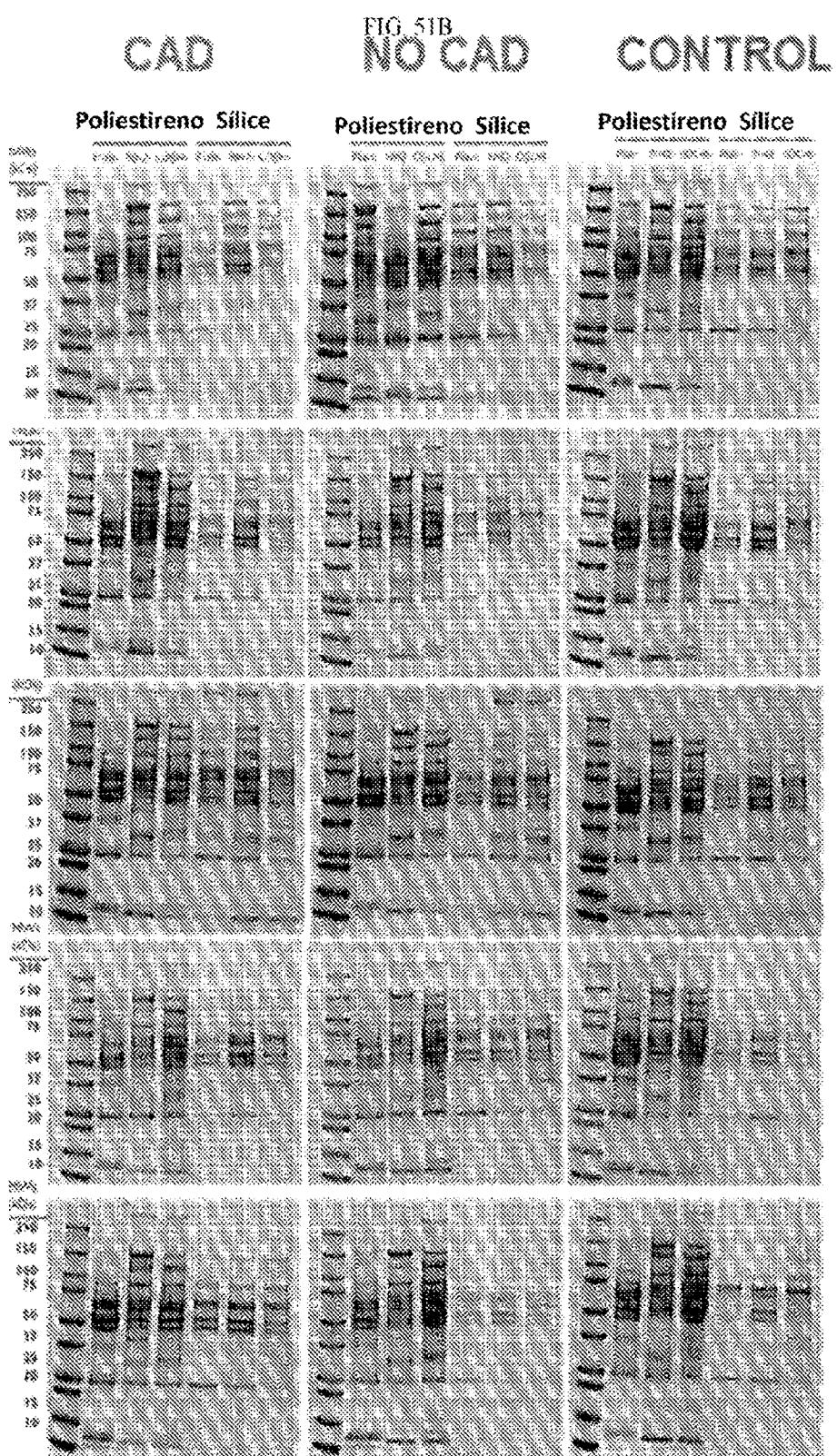
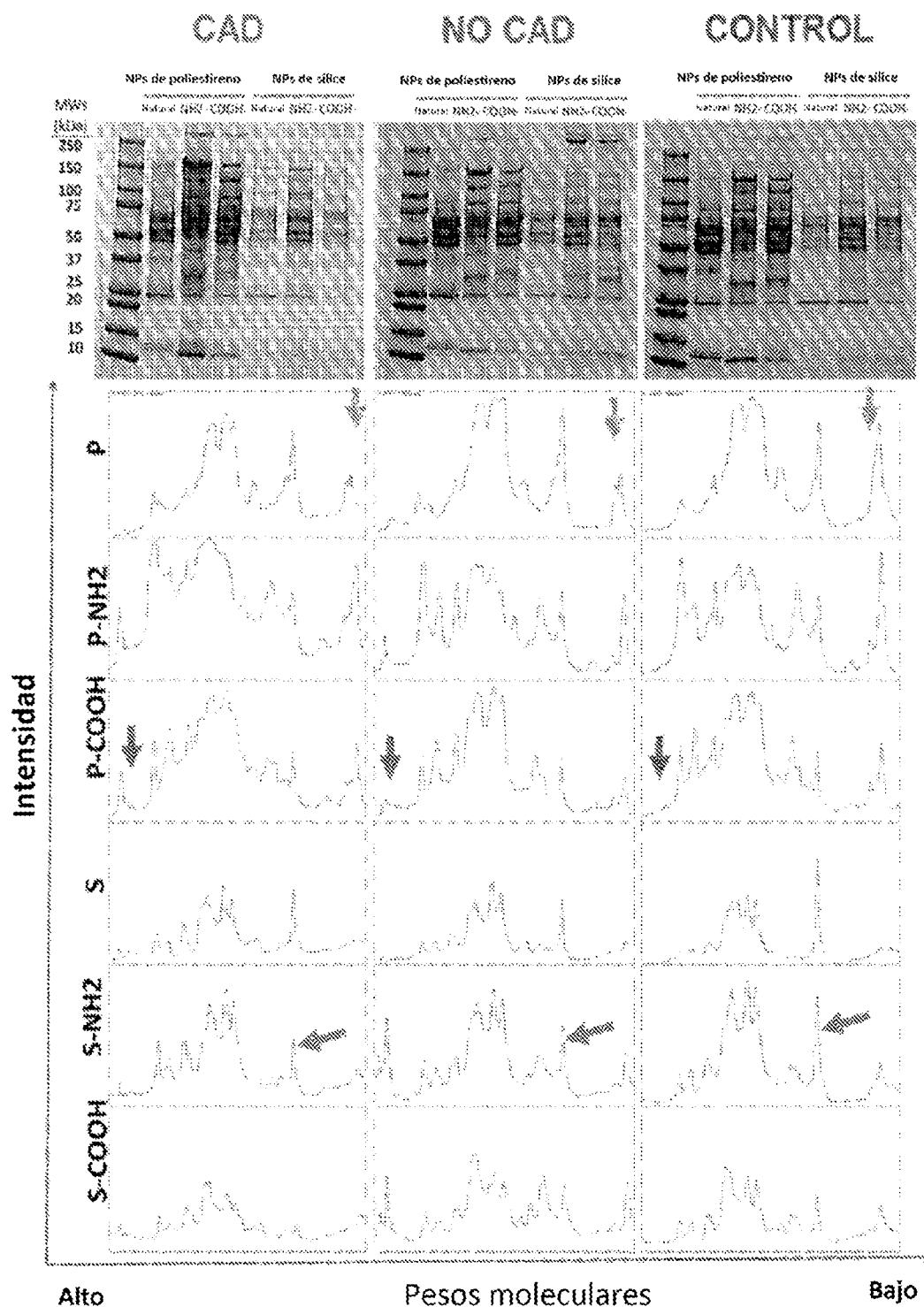


FIG. 51C



Poliesstirene

FIG. 52.

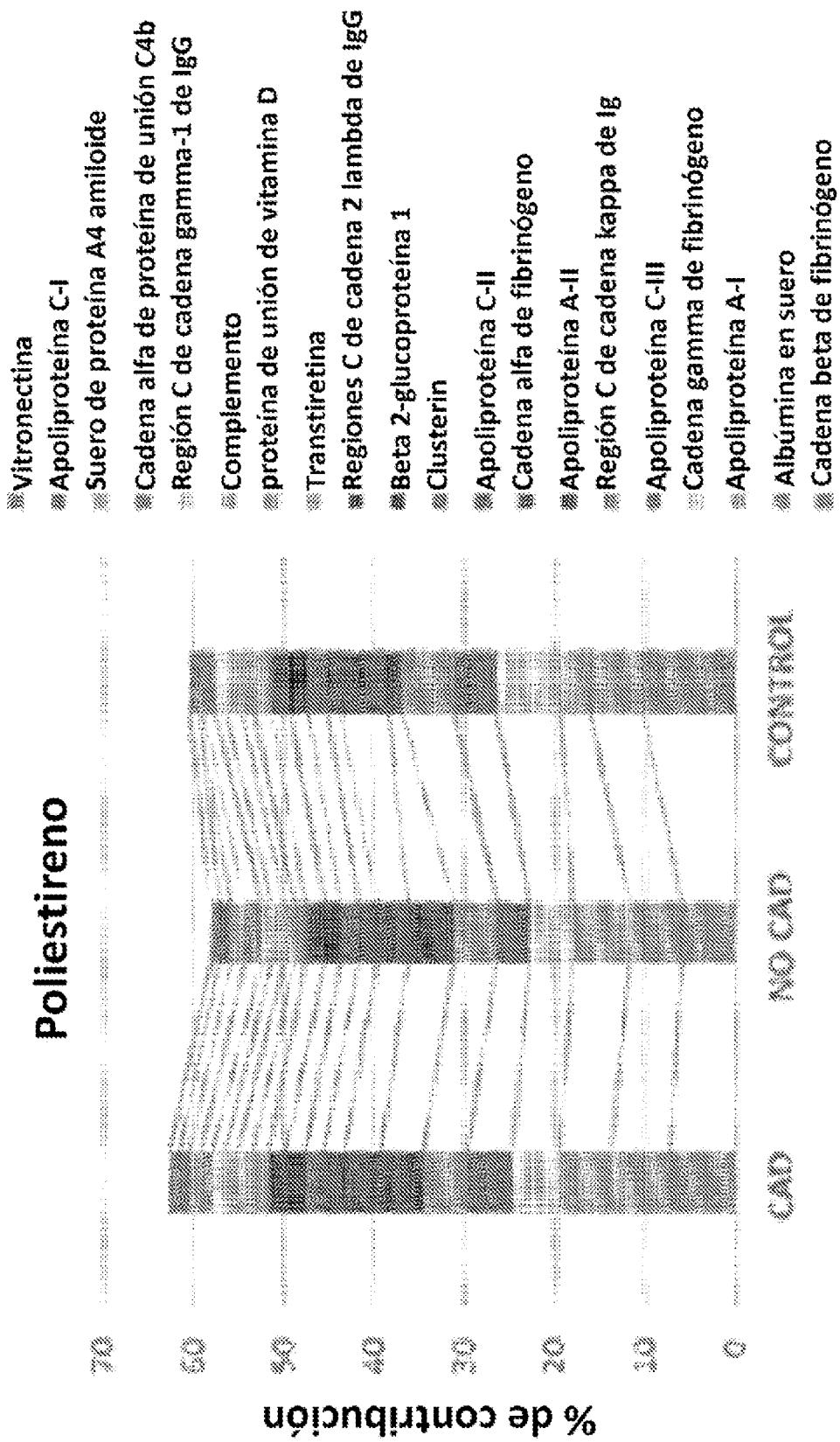


FIG. 52 (CONTINUACIÓN)

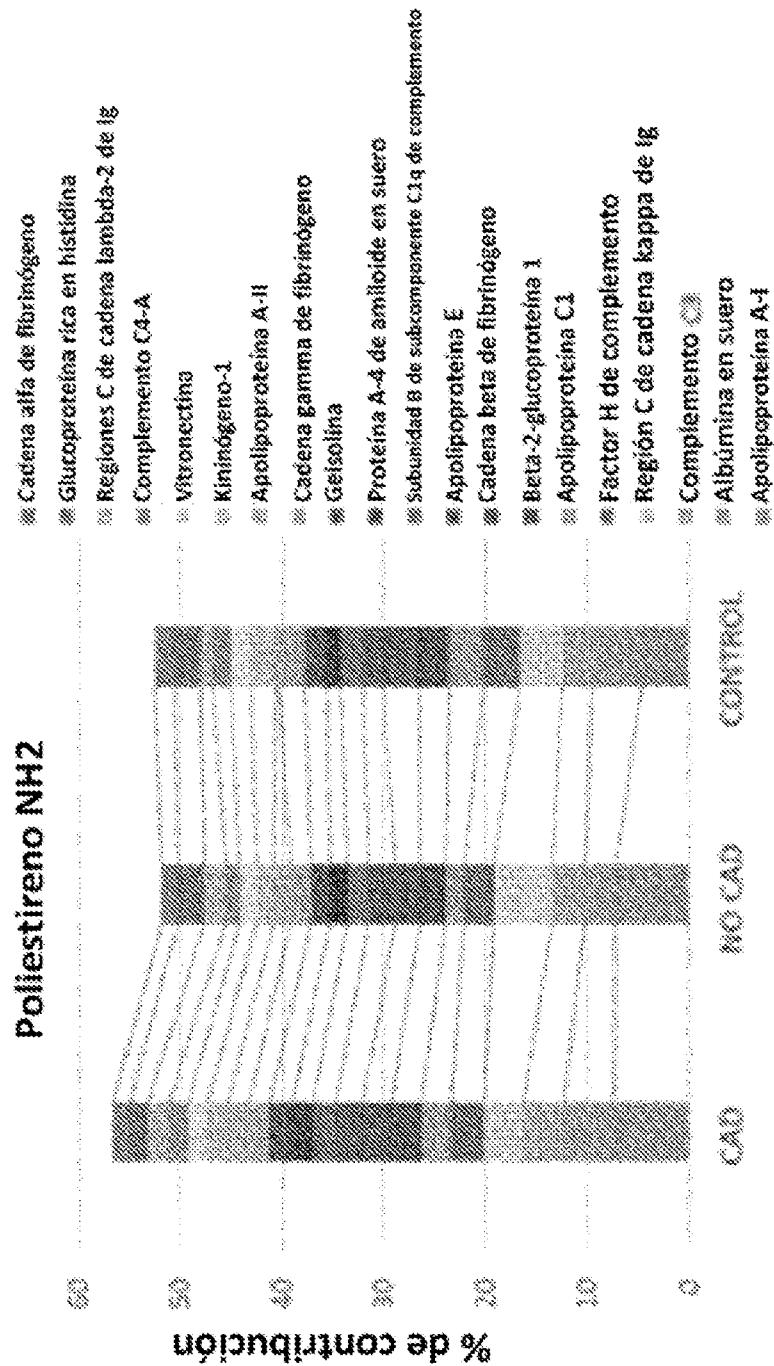


FIG. 82 (CONTINUACIÓN)

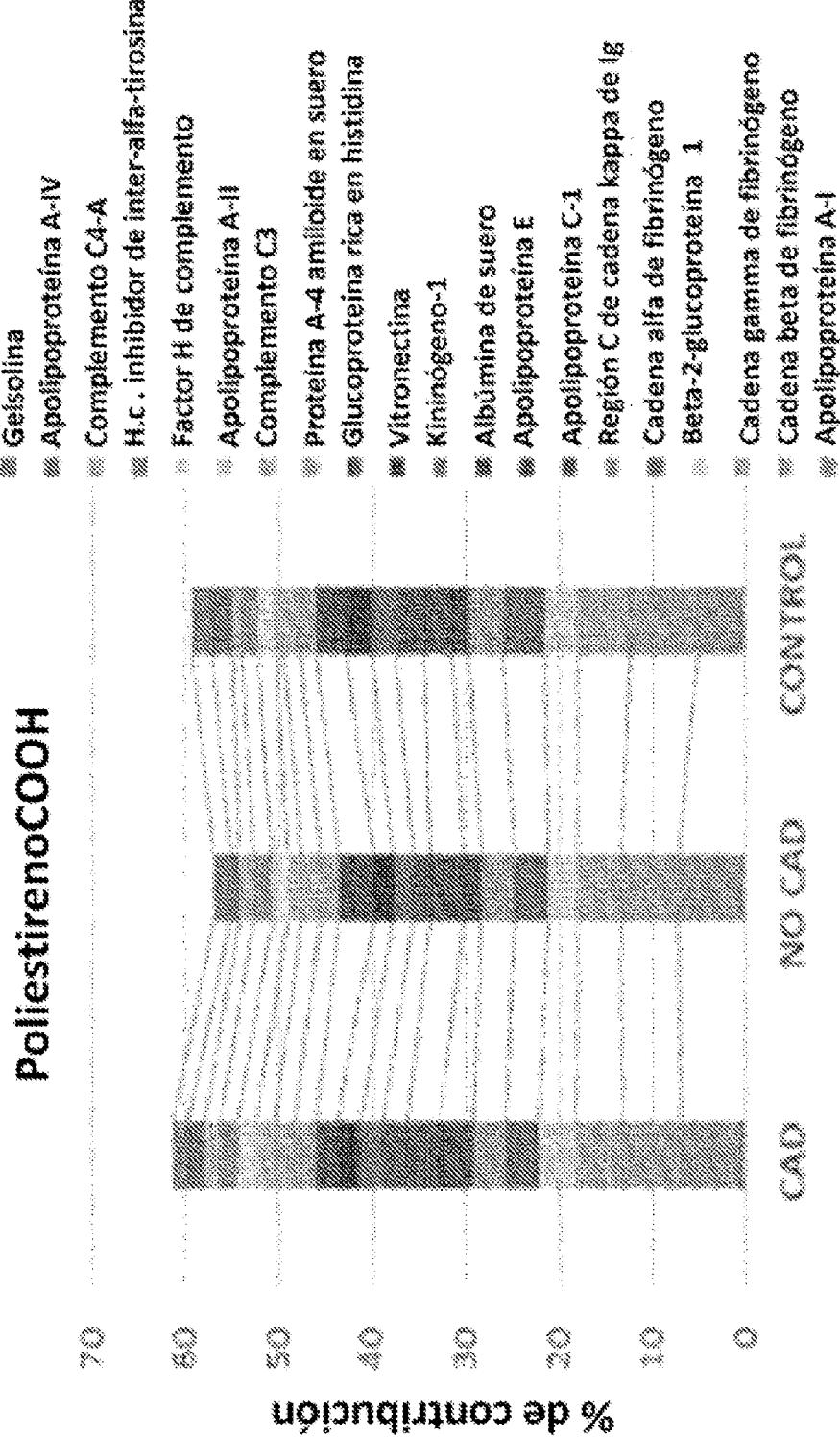


FIG. 52 (CONTINUACIÓN)

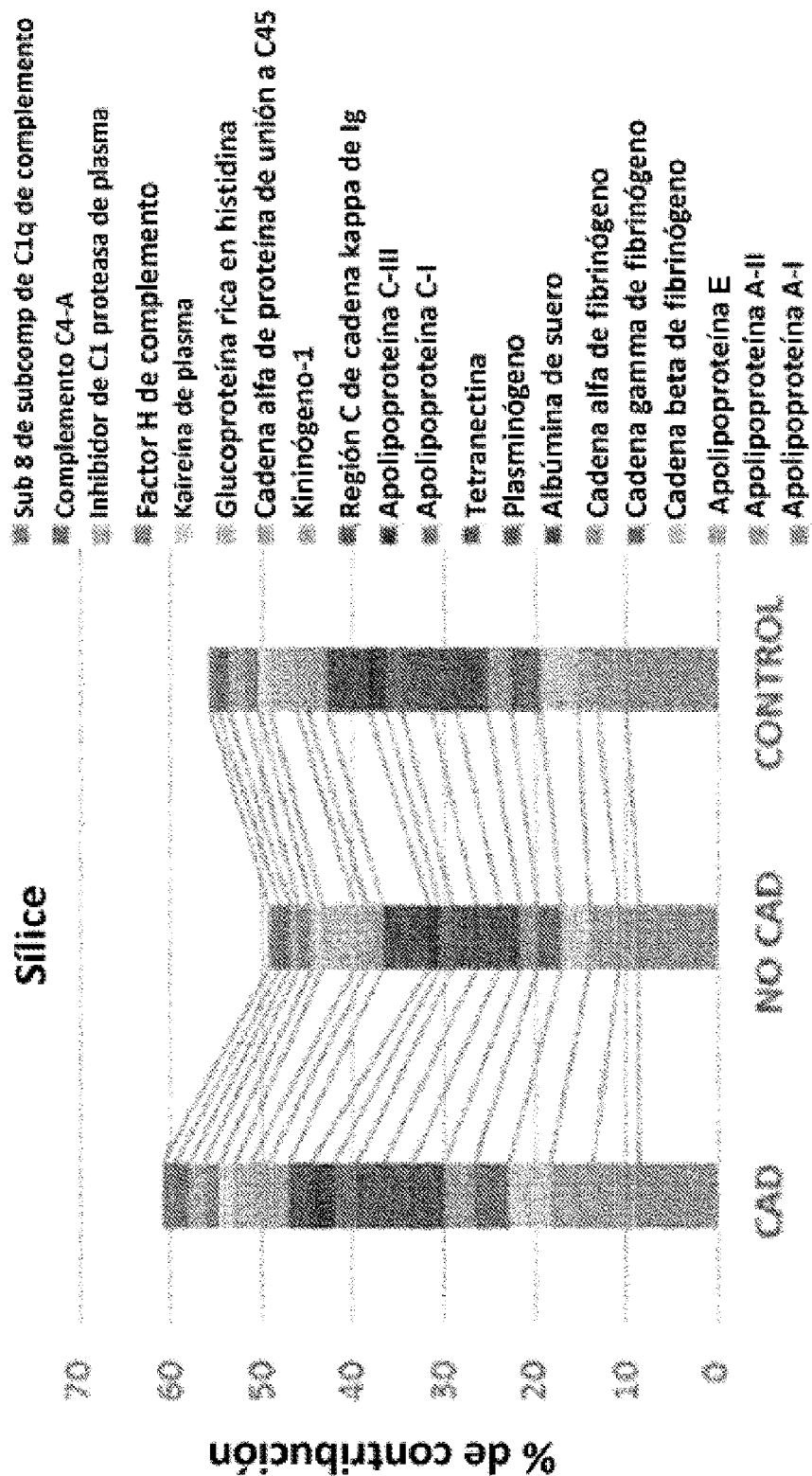


FIG. 52 (CONTINUACION)

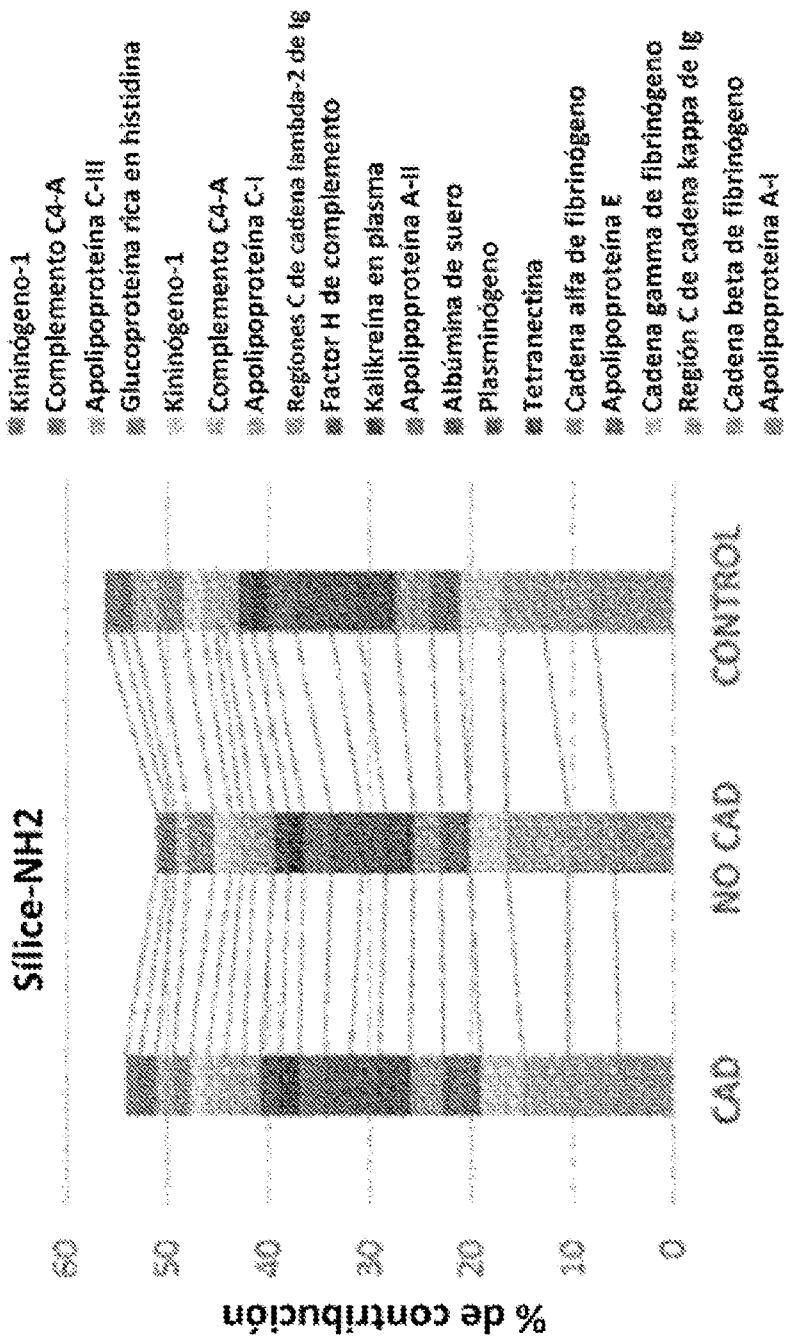


FIG. 52 (CONTINUACIÓN)

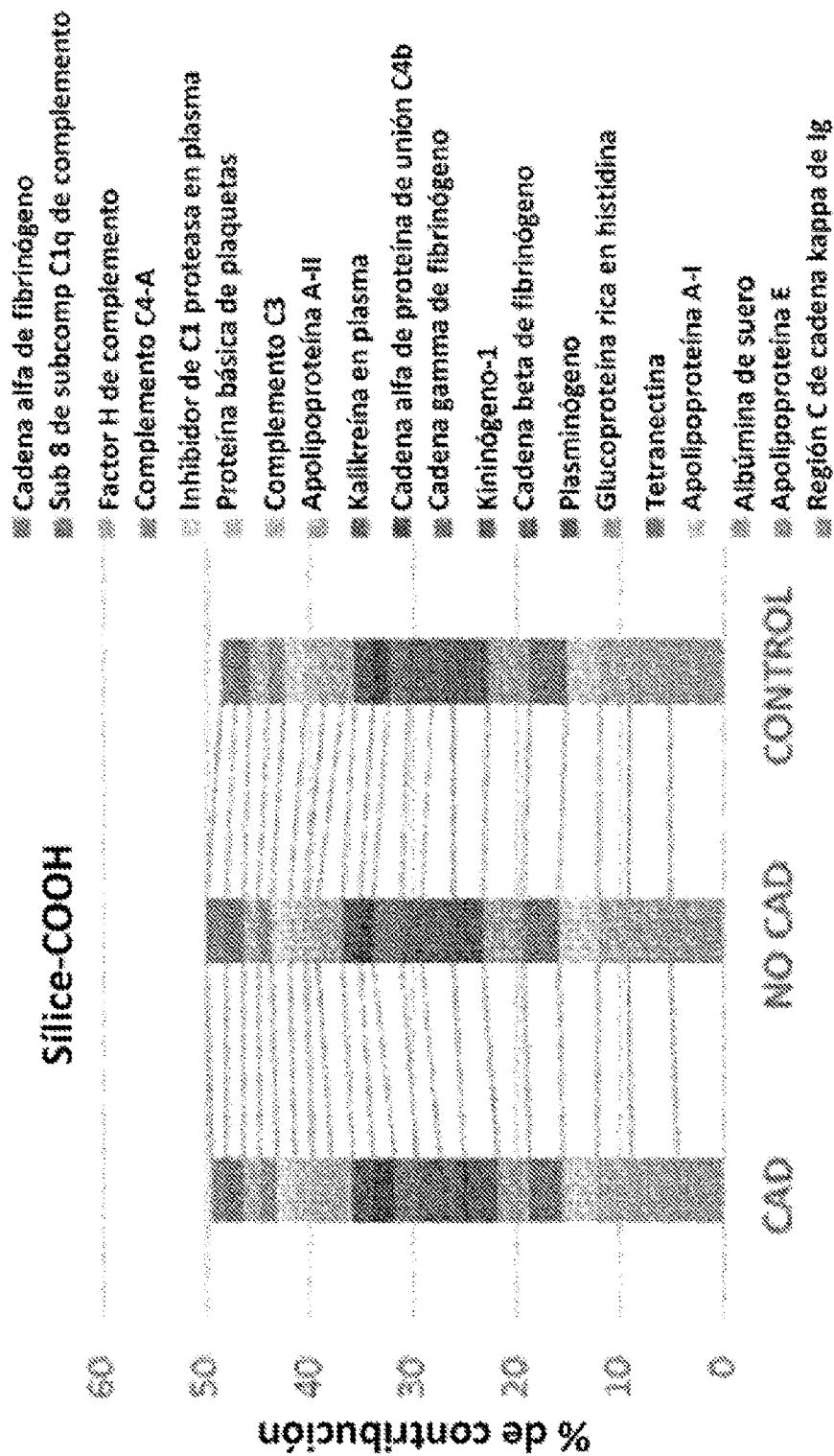
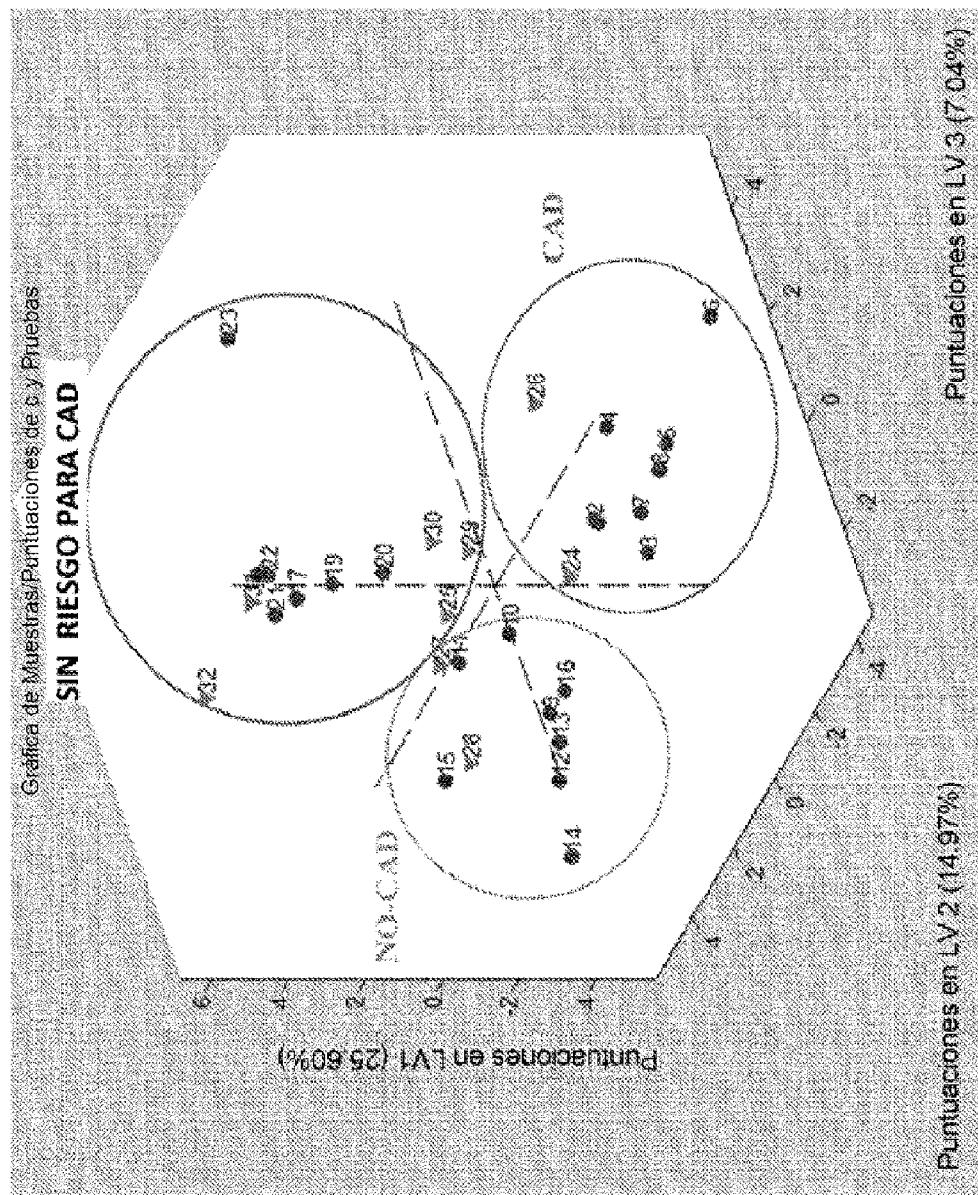
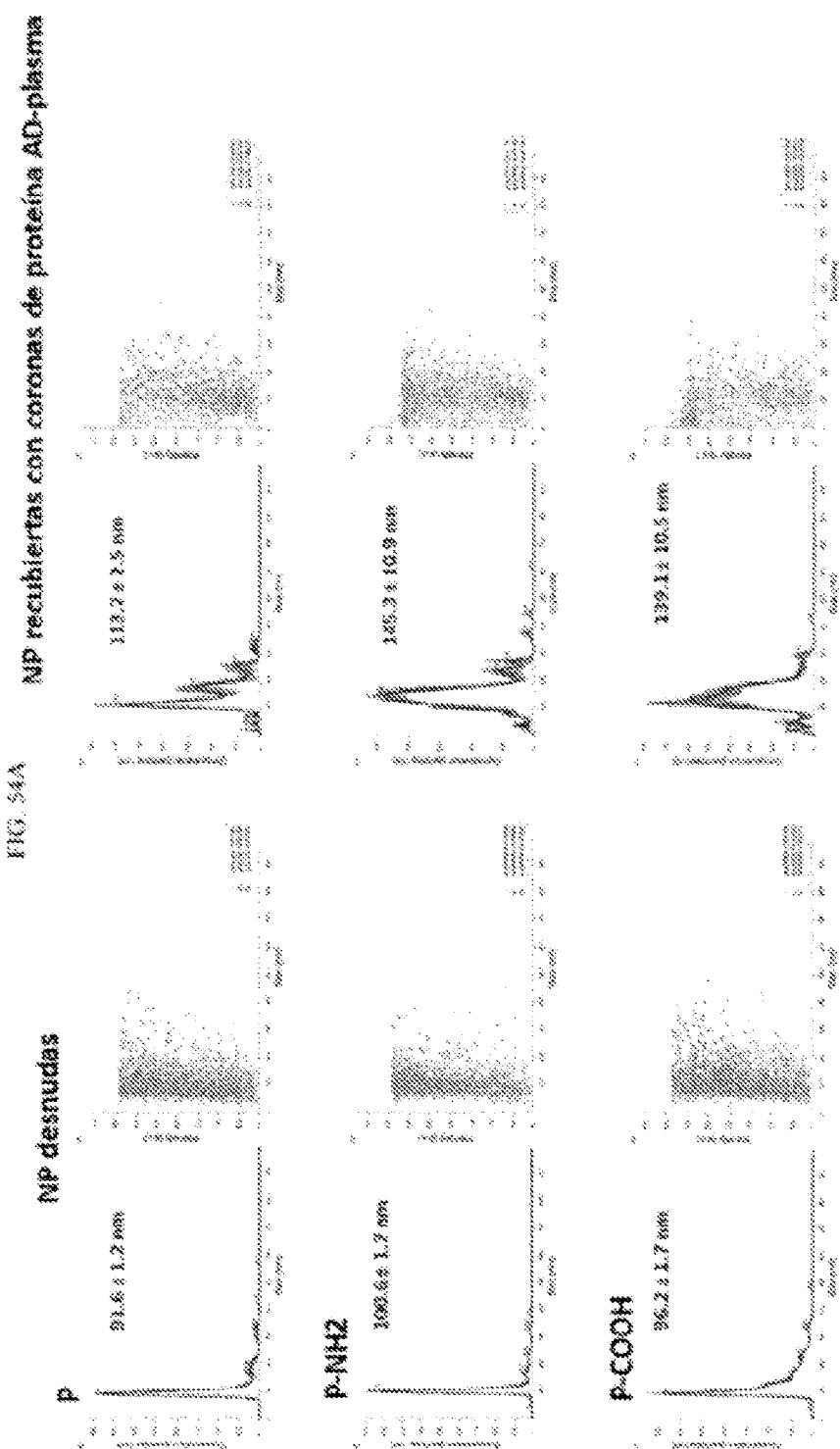
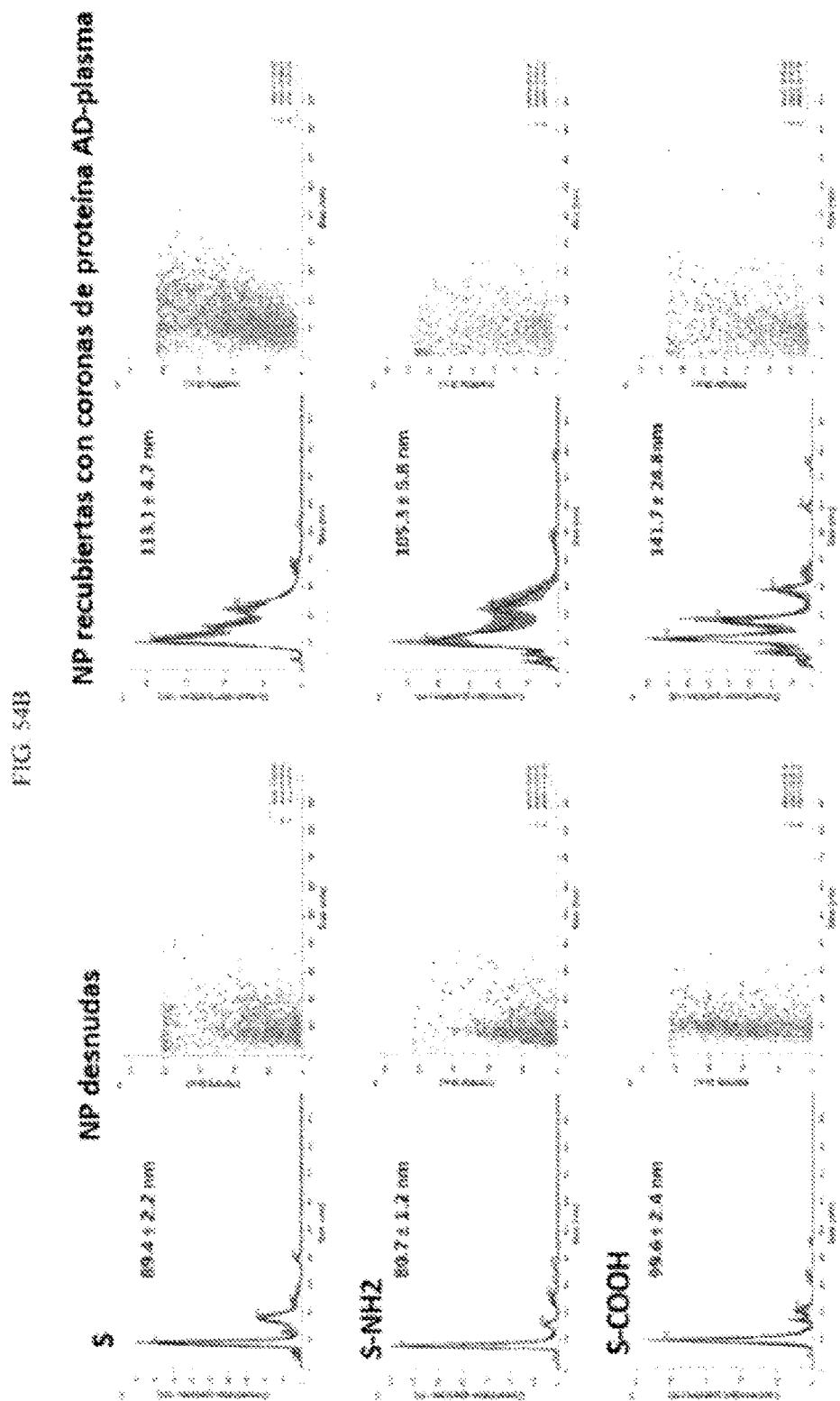


FIG. 53

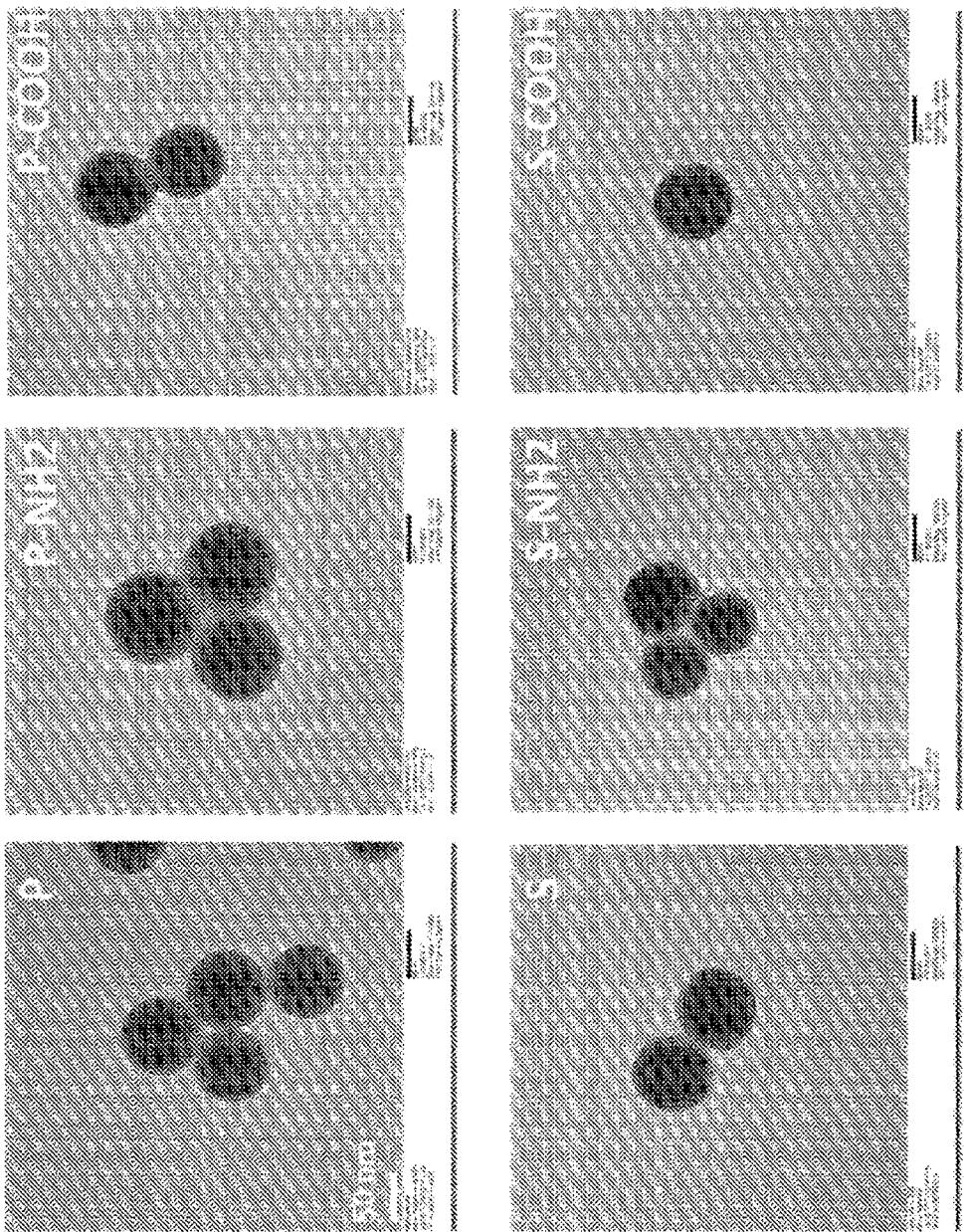


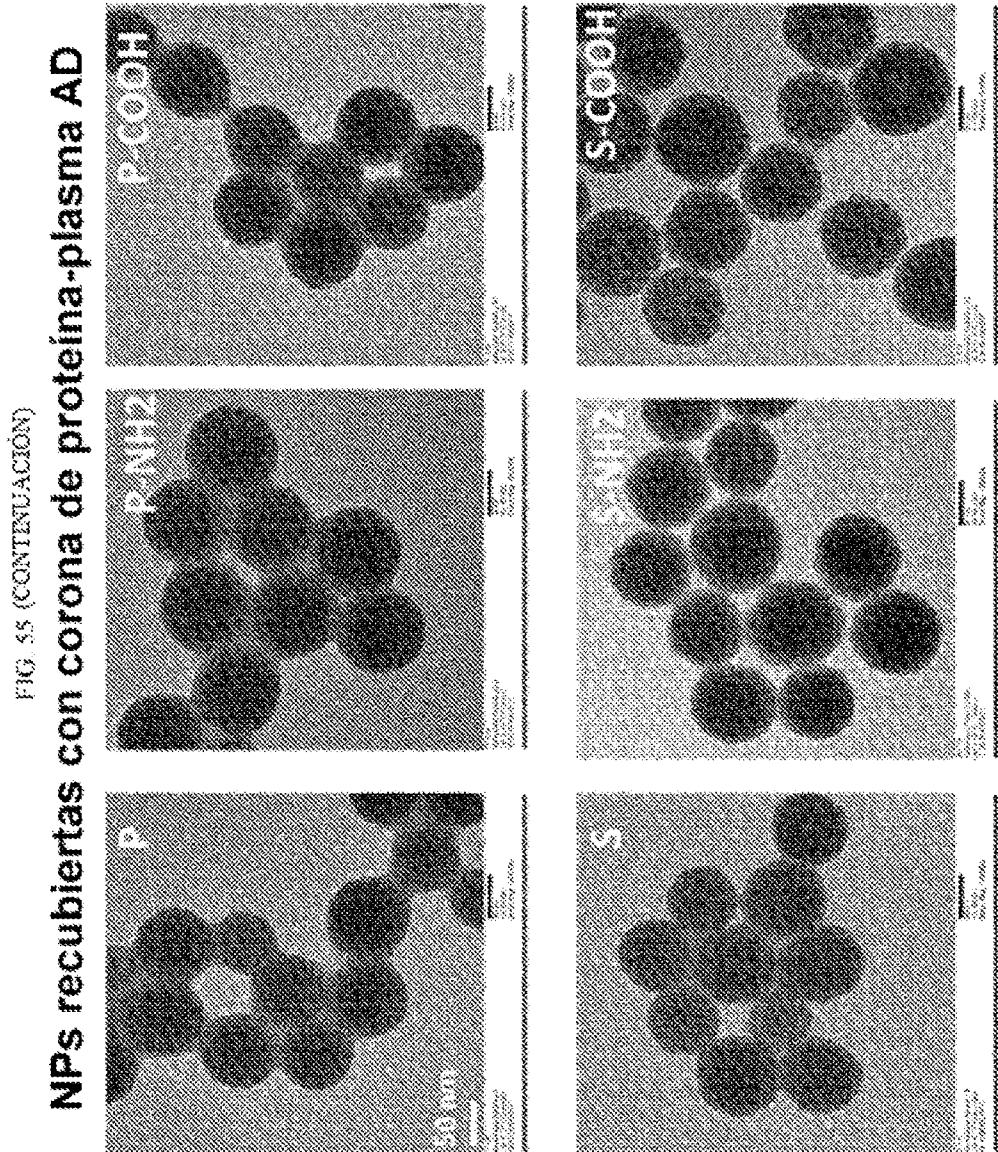


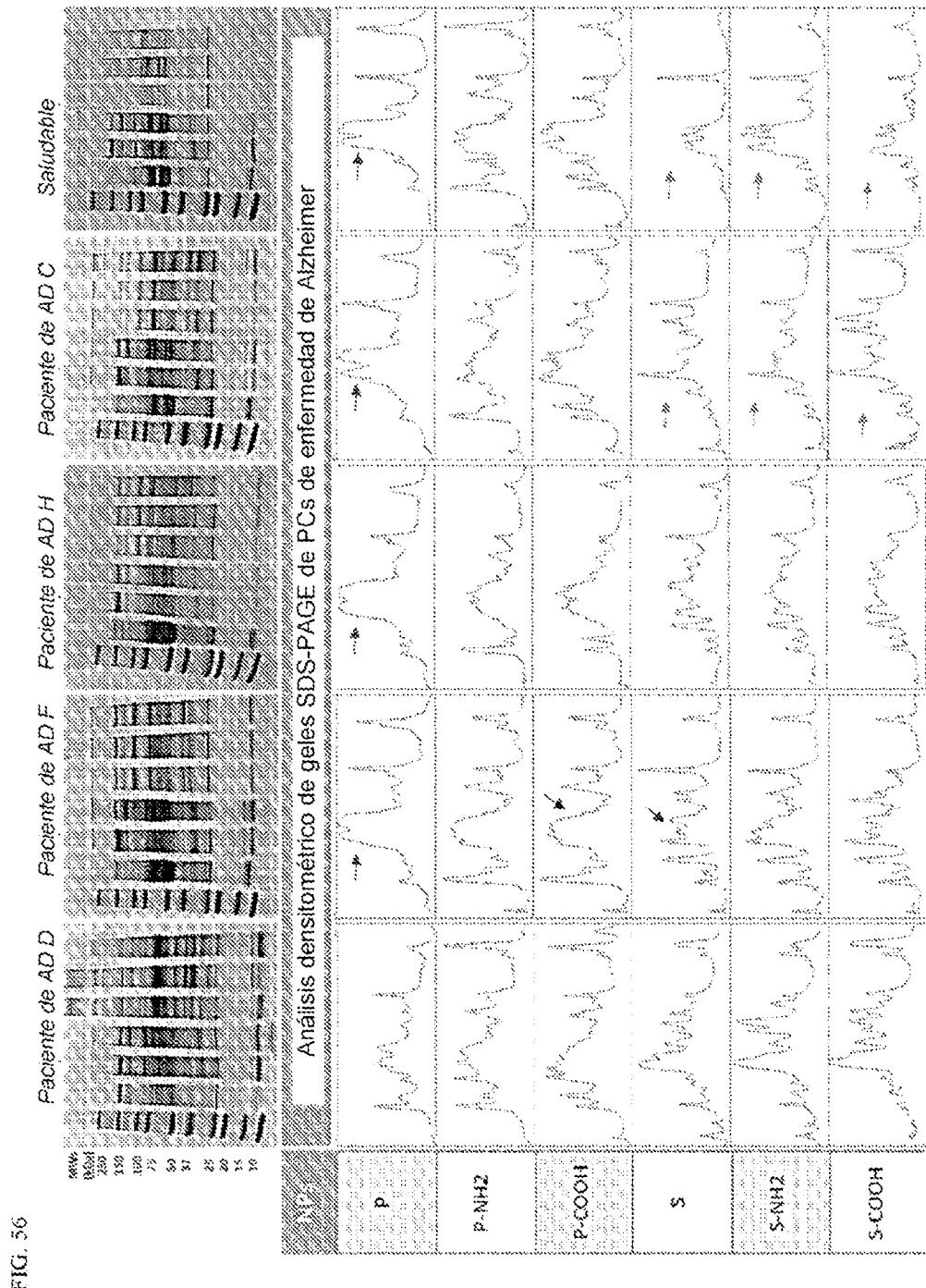


NPs Desnudas

FIG. 55







ES 2 993 503 T3

FIG. S7

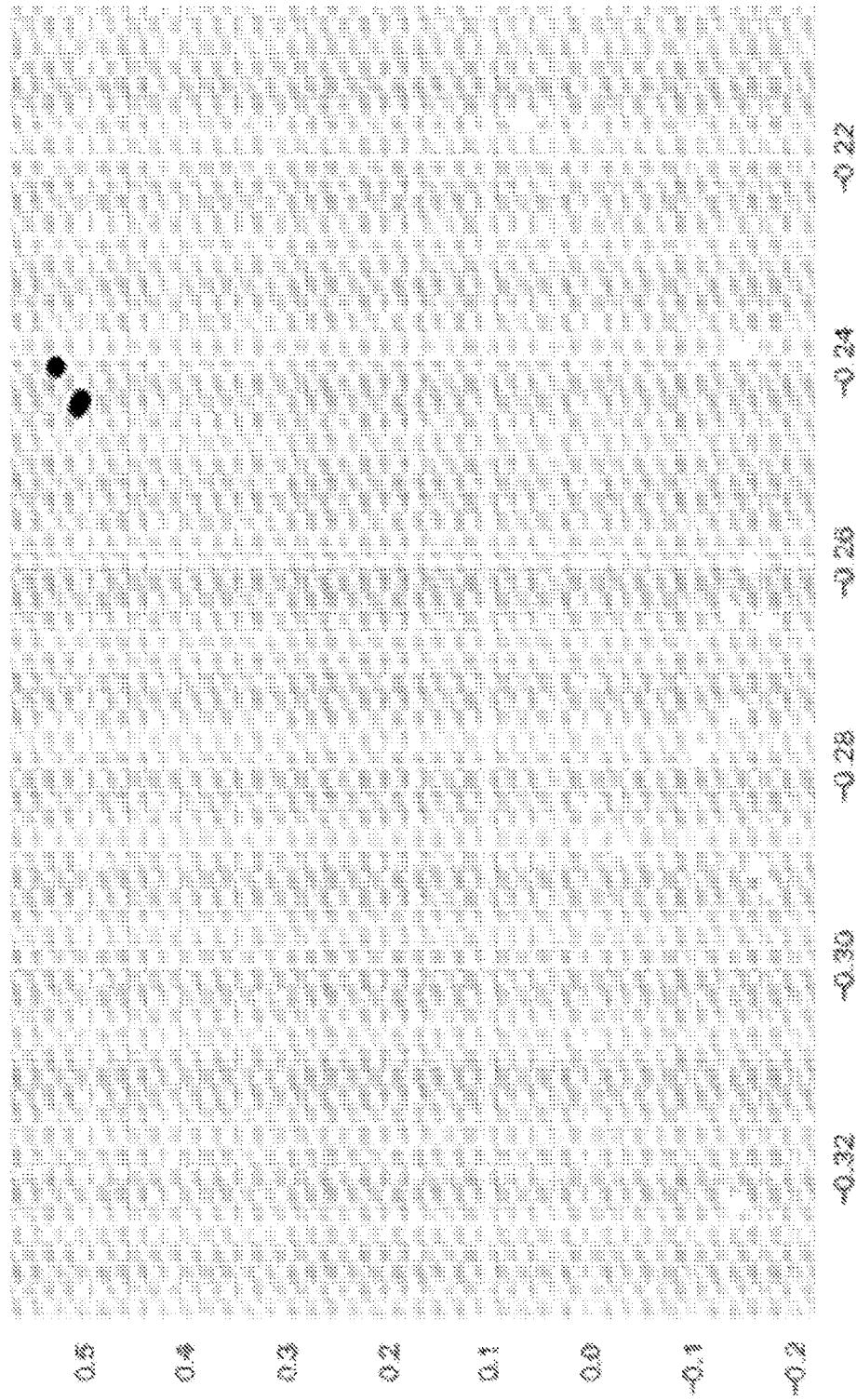
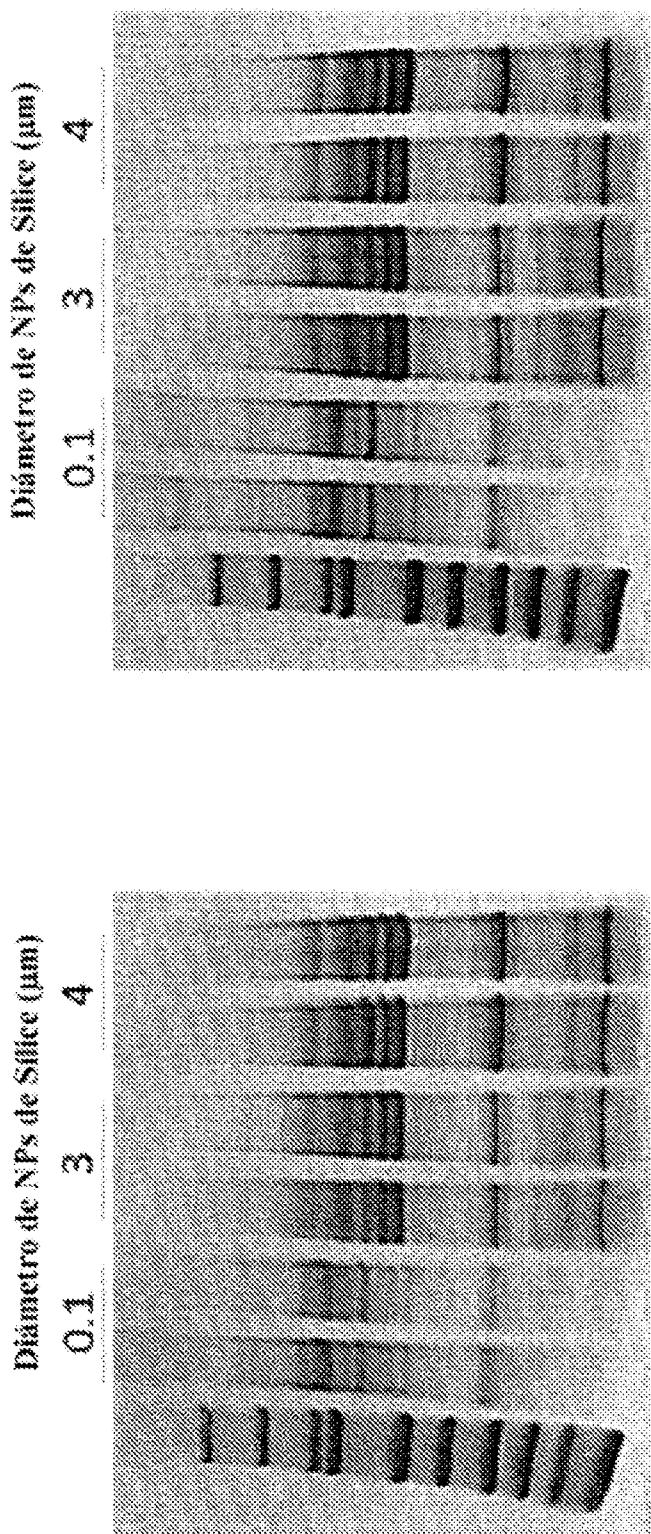


FIG. 58



Misma volumen cargado (10 μ g)
 Dos diferentes protocolos de centrifugación se han
 usado para nano y micro partículas

Misma cantidad (10 μ g) cargada después de que se
 determinó concentración de proteína total mediante
 ensayo de Bradford

FIG. 59

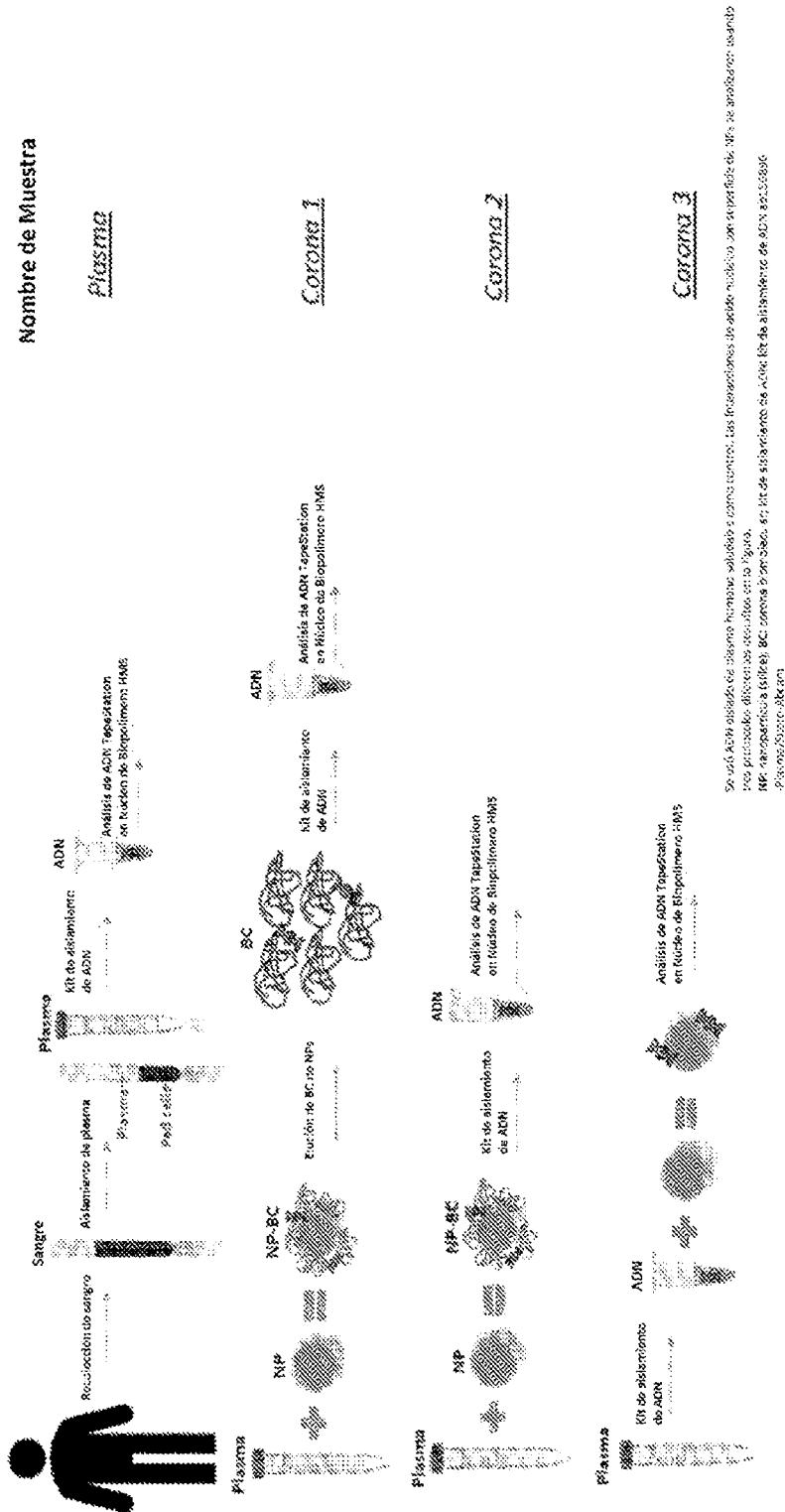


FIG. 60
Imagen de gel

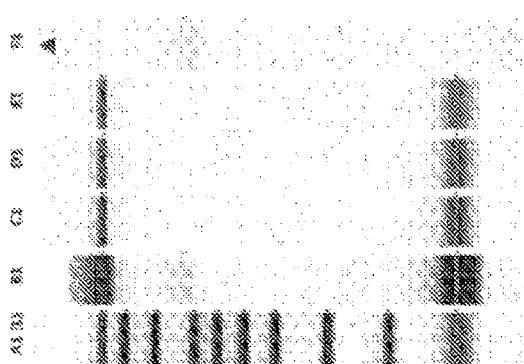


Imagen predeterminada [Contraste 100%], se escala la imagen para ver rango de peso molecular más grande

Información de muestra

| Nombre | Descripción de muestra | Aldea | Observaciones |
|--------|------------------------|----------|--------------------------|
| A1 | 22% | Escobela | |
| B1 | 13% | Escobela | |
| C1 | 23% | Escobela | |
| D1 | 23% | Escobela | |
| E1 | 23% | Escobela | |
| F1 | | | Marcadores no detectados |

FIG. 61

FIG. 61

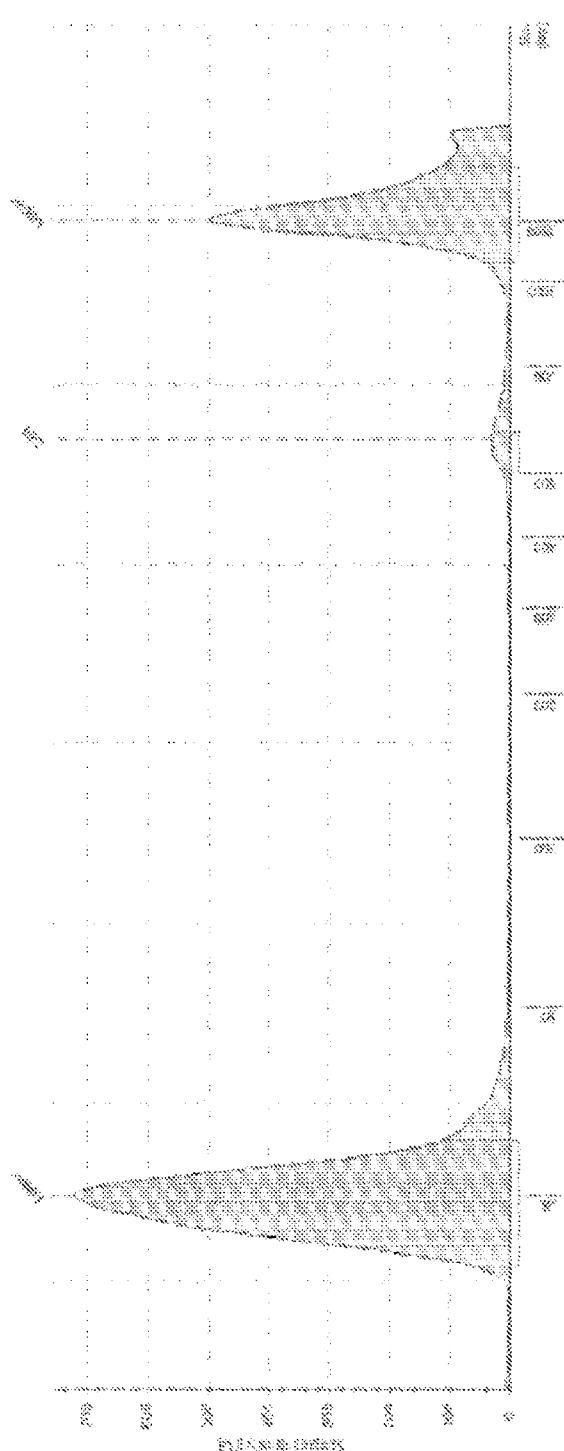


Tabla de muestra

| Perfil | Línea | Presión de muestra | Área | Observaciones |
|--------|-------|--------------------|------|---------------|
| 133 | 100 | 100 | 100 | |

Tabla de Pico

FIG. 61 (CONTINUACIÓN)

| Yardas [yd] | Cien. Cárbaras | Cien. Asignada | Mortalidad de Pico | % de Árboles infectados | Concentración de Pico | Observaciones |
|-------------|----------------|----------------|--------------------|-------------------------|-----------------------|----------------------|
| 25 | 55.1 | — | 1.000.0 | — | — | Mortalidad intelect. |
| 50.4 | 10.8 | — | 27.6 | 100.00 | — | Mortalidad intelect. |
| 130.0 | 23.9 | 23.9 | 25.6 | — | — | Mortalidad intelect. |

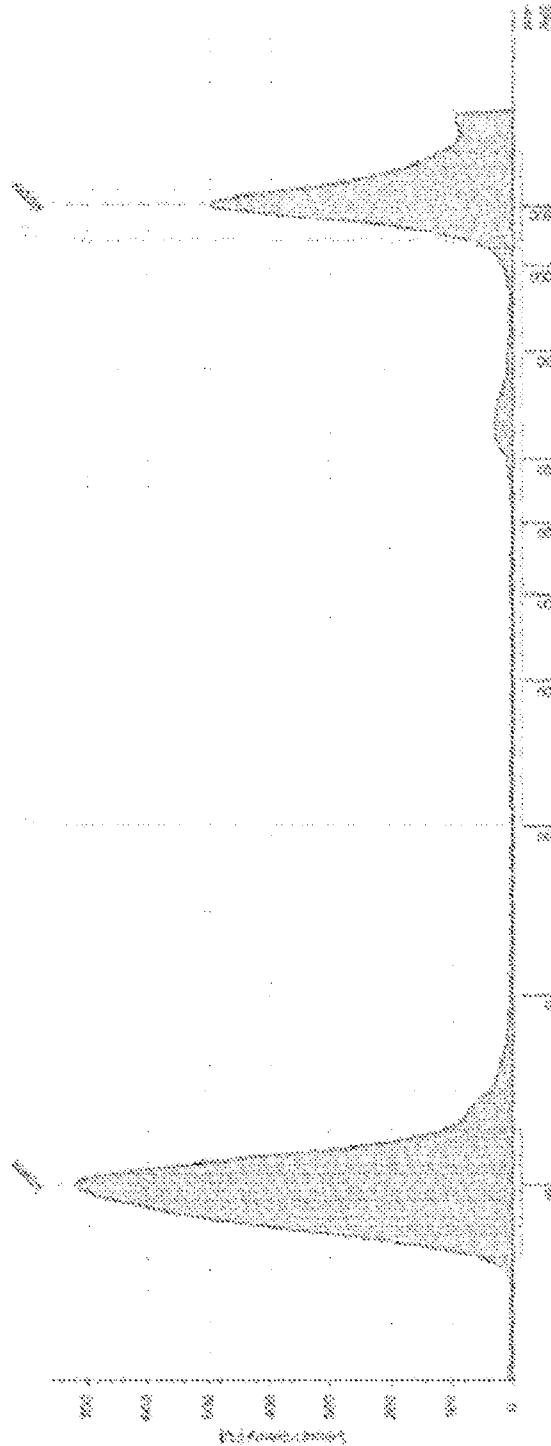
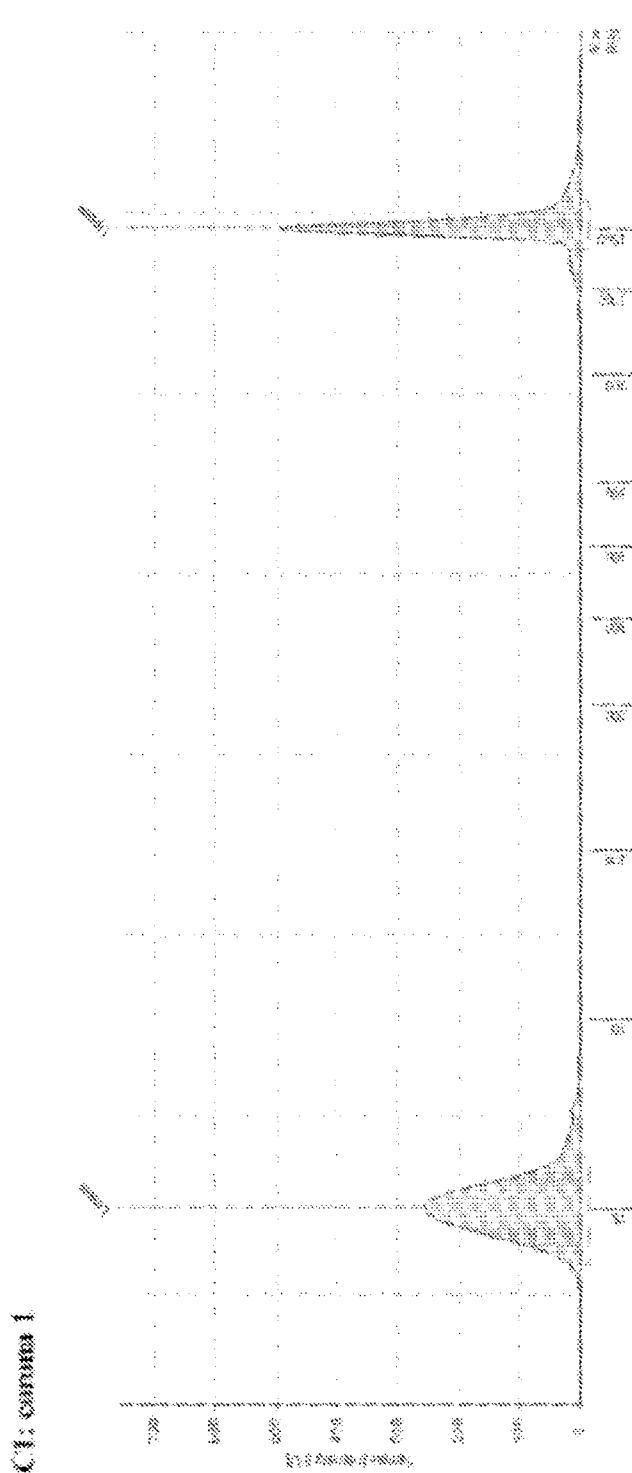


Tabla de Región

| De [yd] | A [yd] | Variedad poseída | Variedad deseada | Mortalidad de Ríos | % de Piel | Concentración de Pico | % de Piel |
|---------|--------|------------------|------------------|--------------------|-----------|-----------------------|-----------|
| 1303 | 1333 | 662 | 33.8 | 117 | 23.67 | — | — |

ES 2 993 503 T3

FIG. 62



卷之三

Tabla de nuestros

| Species | Common Name | Description | Observations |
|---------|-------------|---------------------|--------------|
| Proba | Proba | Described by author | |
| C1 | C1 | Described by author | |

FIG. 62 (CONTINUACIÓN)

Tabla de Pie

| Área de Pie | Cuadros Cuadrados | Área Asignada | Máx. de Pico | % de Área Superada | Concentración de Pico | Observaciones |
|-------------|-------------------|---------------|--------------|--------------------|-----------------------|---------------|
| 25 | 43% | 25% | 26480 | — | — | Máx. de Pico |
| 25 | 43% | 25% | 276 | — | — | Máx. de Pico |
| 3300 | — | — | — | — | — | Máx. de Pico |

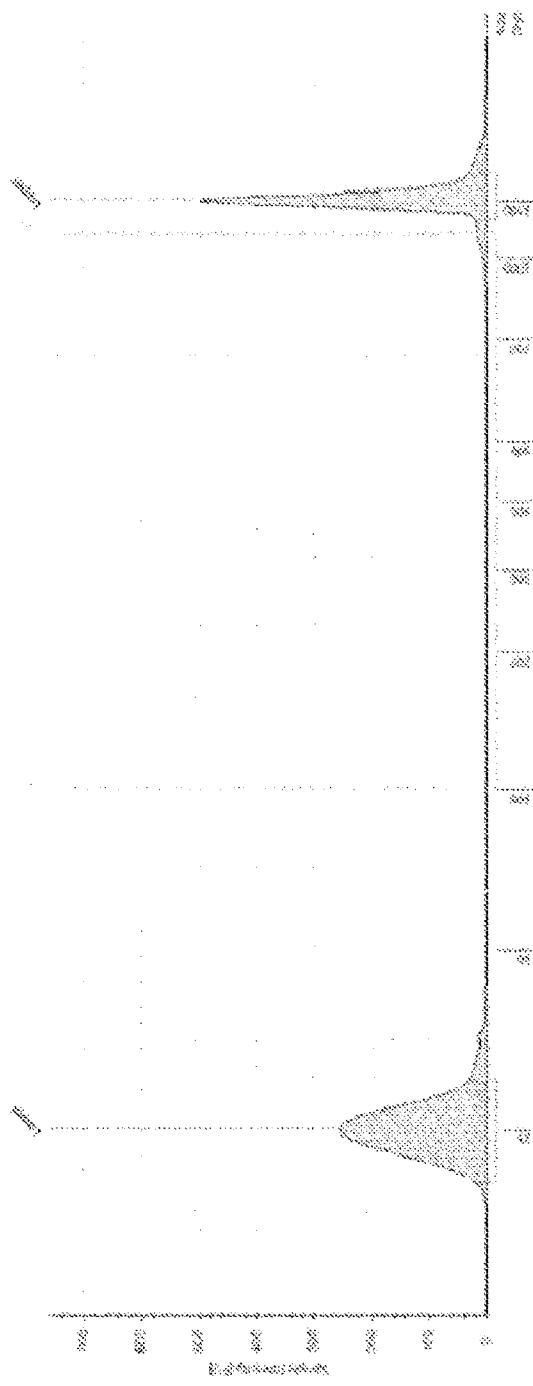


Tabla de Región

| Región | A. [100] | Transito promedio [hrs] | Área de Pico | Máx. de Región | % de Total | Concentración de region | C. [hrs] |
|--------|----------|-------------------------|--------------|----------------|------------|-------------------------|----------|
| 100 | 1.200 | 846 | 14.0 | 46.4 | 17.84 | — | — |

FIGURA 2

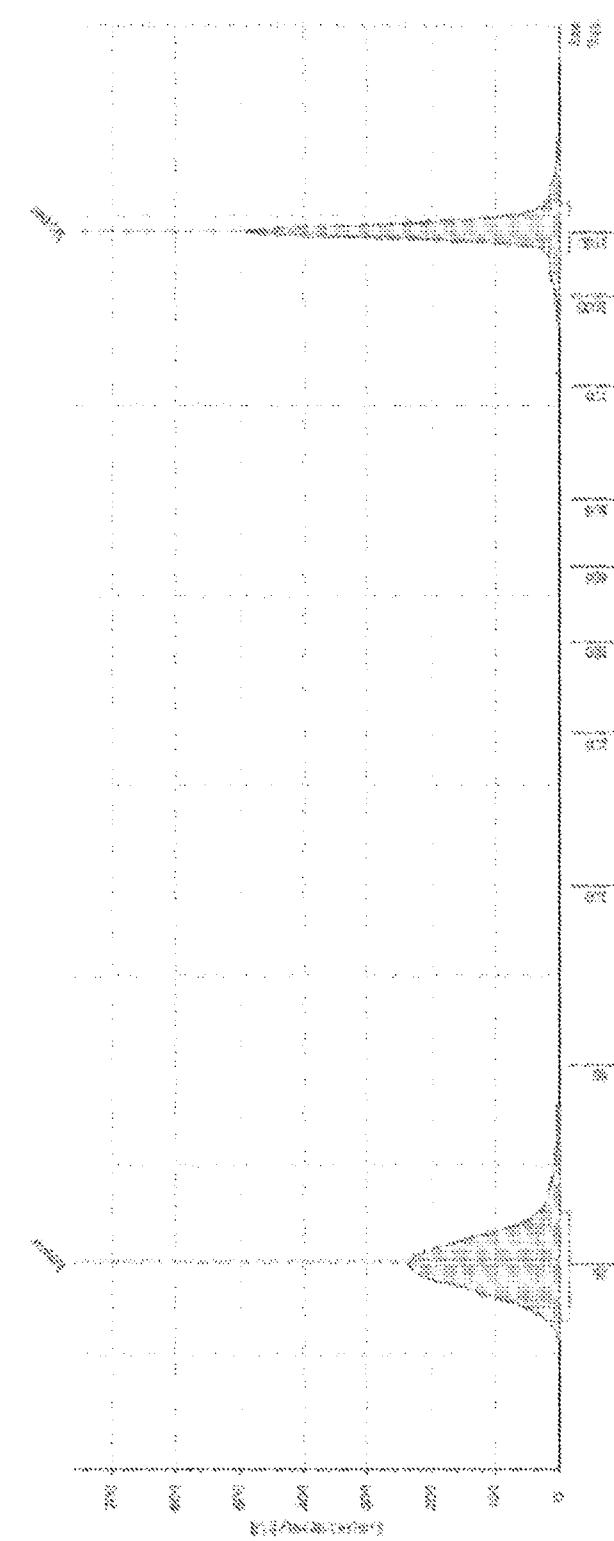


TABLA DE MUESTRA

| Periodo | Distribución de muestra | | | | | | | | | | Afecto | Observaciones | |
|---------|-------------------------|-----|-----|-----|-----|-----------|-----|-----|-----|-----|--------|---------------|-----|
| | Control 2 | | | | | Control 3 | | | | | | | |
| 13 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |

Tabla de Pico

FIG. 63 (CONTINUACION)

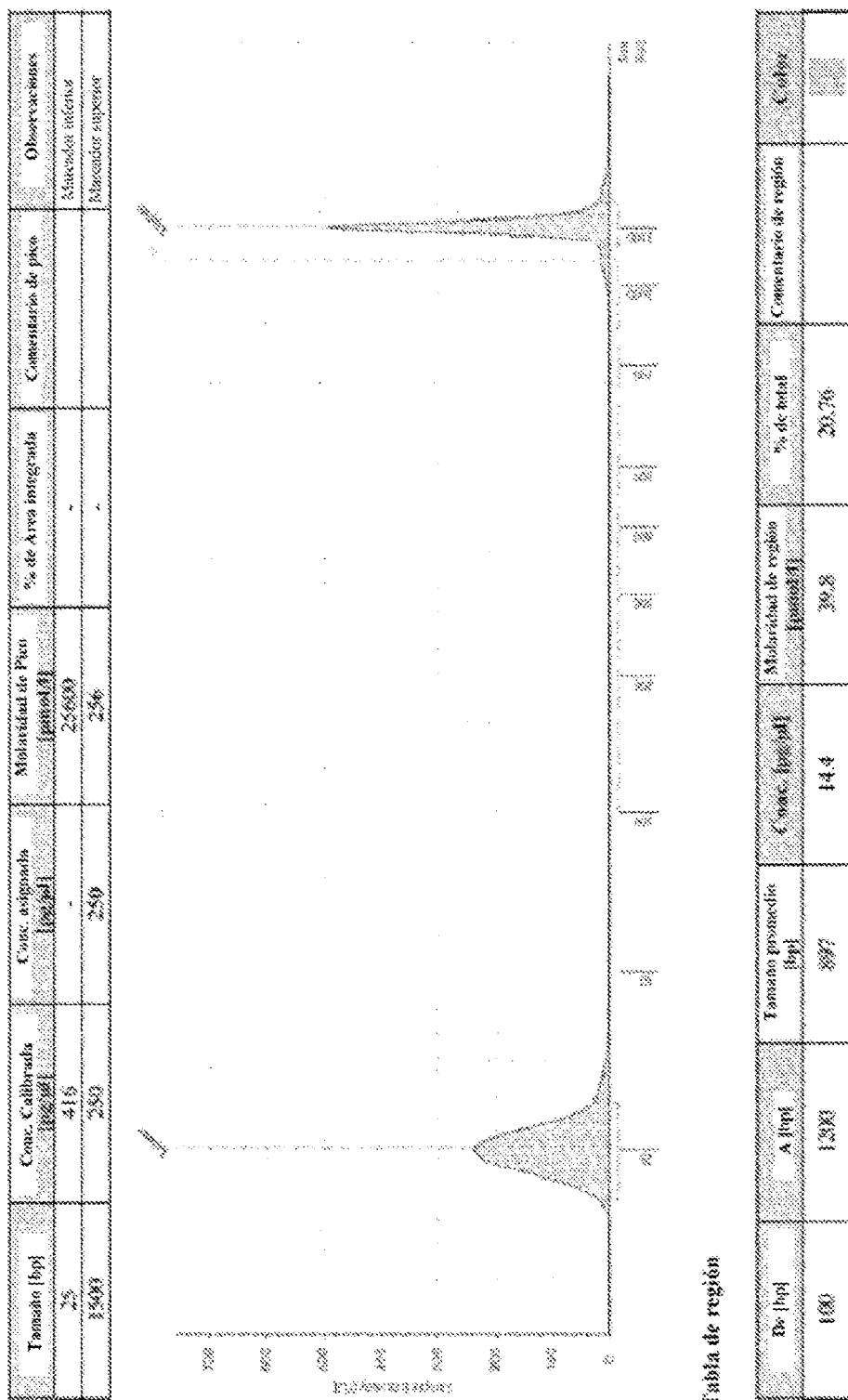


Tabla de region

| Región | Porcentaje de region | Concentración de region |
|-----------------|----------------------|-------------------------|
| Región de fondo | 20.36 | 39.8 |

FIG. 64

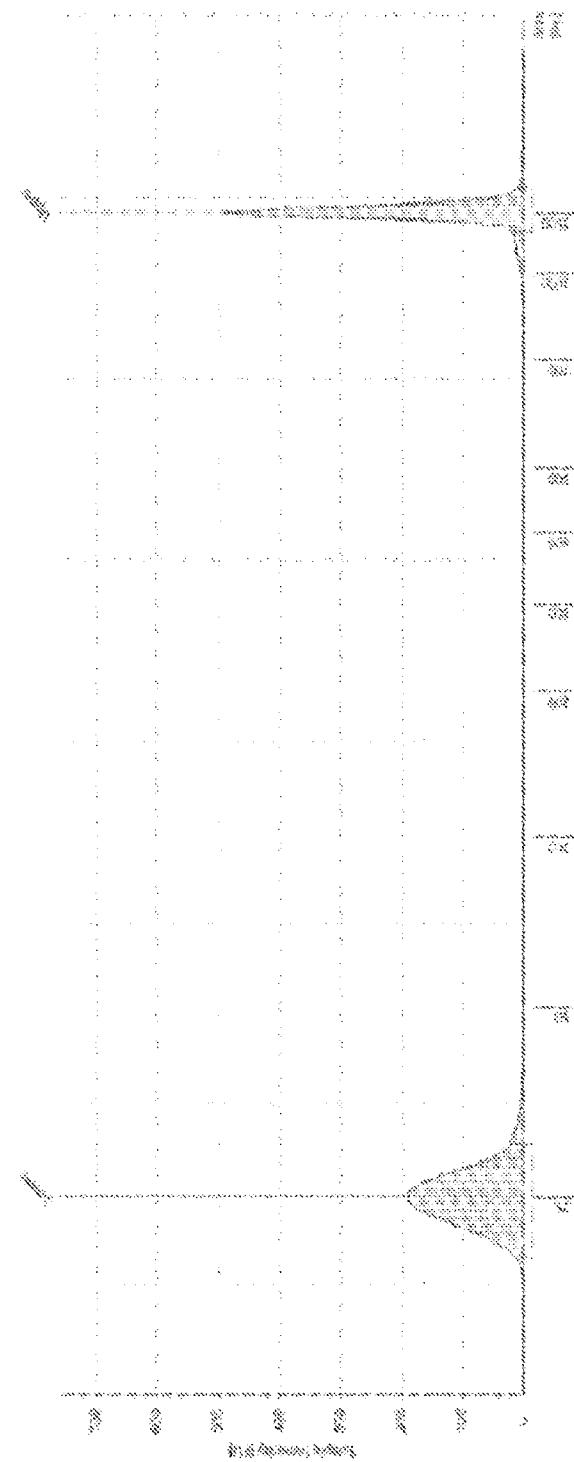


FIG. 64: CORRELATION

| Objetivo | Distribución de los datos | | Número de observaciones |
|----------|---------------------------|-------------|-------------------------|
| | Exponencial | Lineal | |
| ES 2 993 | Exponencial | Lineal | 39 |
| ES 2 993 | Lineal | Exponencial | 39 |

| Objetivo | Distribución de los datos | | Número de observaciones |
|----------|---------------------------|-------------|-------------------------|
| | Exponencial | Lineal | |
| ES 2 993 | Exponencial | Lineal | 39 |
| ES 2 993 | Lineal | Exponencial | 39 |

FIG. 64 (CONTINUACIÓN)

Tabla de Pico

| Frecuente (p/p) | Cone. Calibrada | % de superficie | Altitud media de Pico | % de Áreas Interiores | Contaminación de pico | Observaciones |
|-----------------|-----------------|-----------------|-----------------------|-----------------------|-----------------------|---------------|
| 25 | 305 | 100% | 2000 | — | — | Manzanares |
| 1500 | 263 | 25% | 2300 | 2.9% | — | Manzanares |

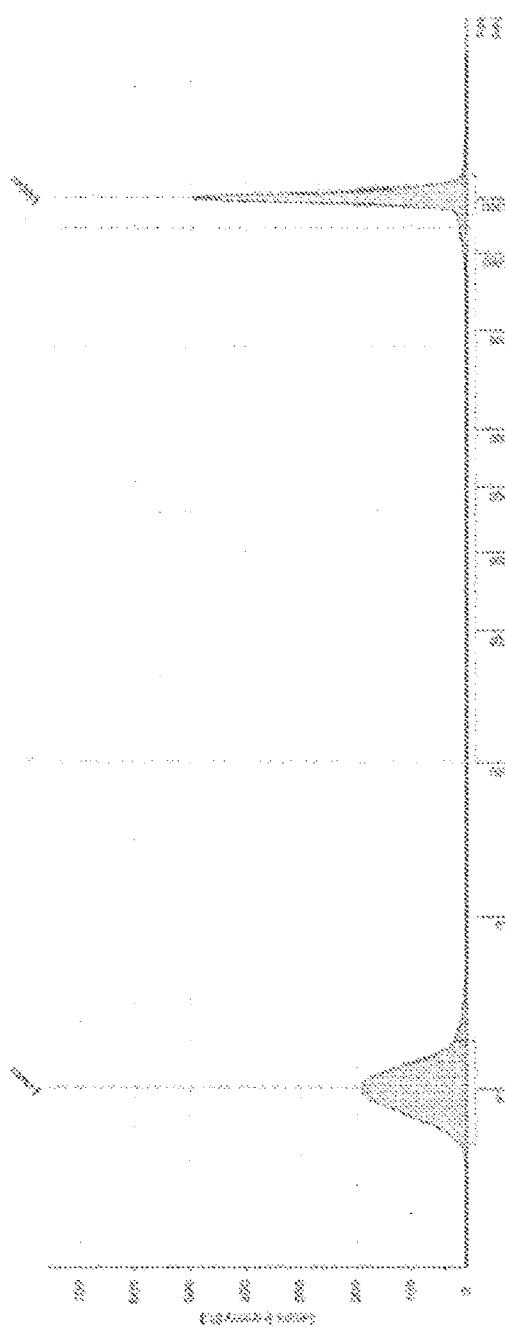


Tabla de región

| De (p/p) | A (p/p) | Tamaño promedio (p/p) | % de superficie | Número de casas | % de total | Contaminación de región | % de total |
|----------|---------|-----------------------|-----------------|-----------------|------------|-------------------------|------------|
| 300 | 1200 | 318 | 15.2 | 35.6 | 25.7% | 548 | 50.5% |

FIG. 65

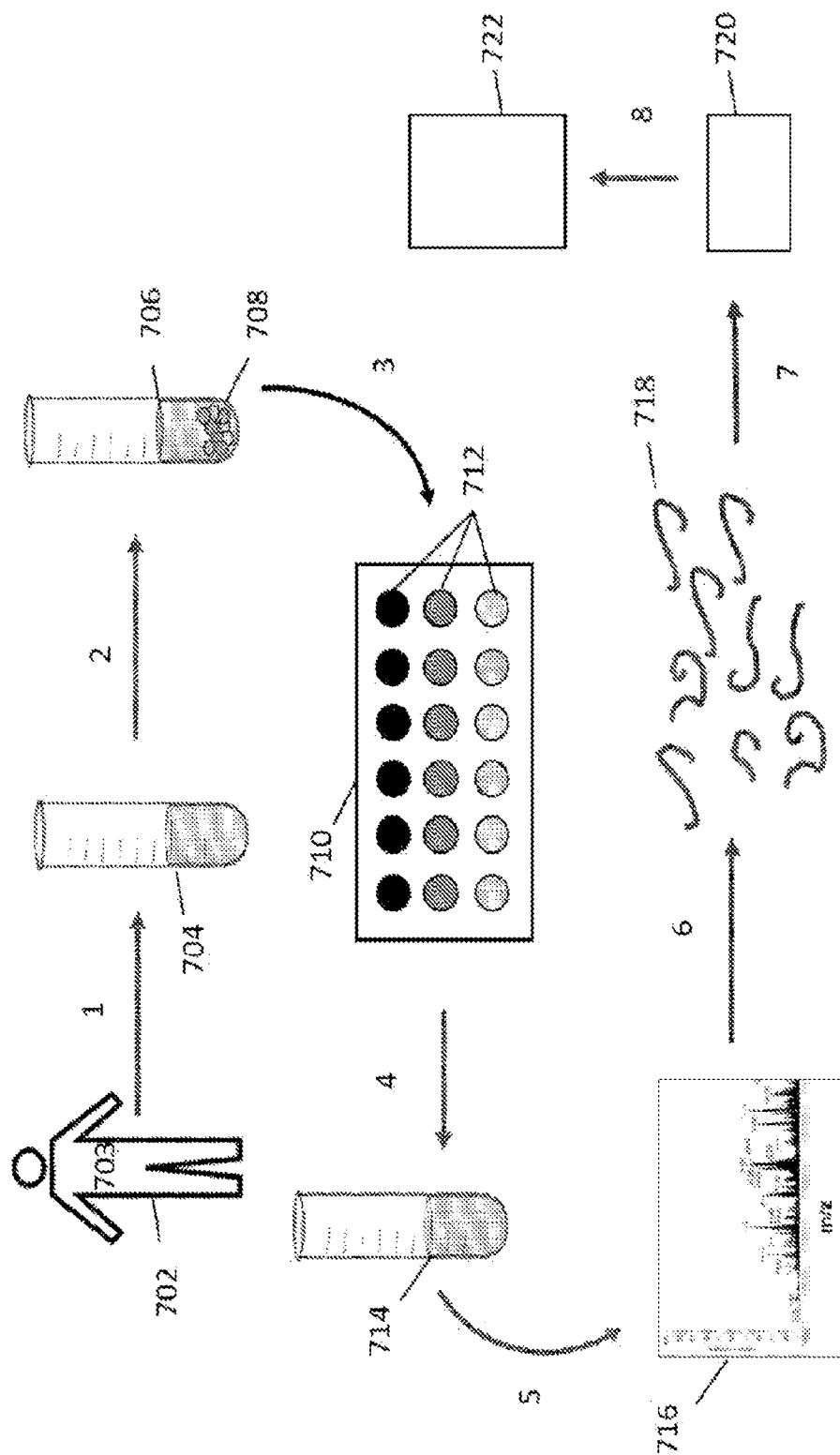


FIG. 66

