(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 728 119**

(51) Int. CI.:

***G01N 33/50*** (2006.01)
***G01N 15/14*** (2006.01)
***G06F 19/00*** (2008.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **14.01.2015 PCT/US2015/011441**

(87) Fecha y número de publicación internacional: **23.07.2015 WO15109003**

(96) Fecha de presentación y número de la solicitud europea: **14.01.2015 E 15702063 (7)**

(97) Fecha y número de publicación de la concesión europea: **27.02.2019 EP 3094974**

(54) Título: **Identificación de estados celulares funcionales**

(30) Prioridad:

**14.01.2014 US 201461927247 P**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
**22.10.2019**

(73) Titular/es:

**ASEDASCIENCES AG (100.0%)**
**Untere Paulistrasse 6A**
**8834 Schindellegi Switzer, CH**

(72) Inventor/es:

**RAJWA, BARTLOMIEJ y**
**SHANKEY, T. VINCENT**

(74) Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

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

Identificación de estados celulares funcionales

Campo de la invención

Las realizaciones se refieren a caracterizar el efecto de un agente en células.

El cribado fenotípico de compuestos es una tecnología emergente para una evaluación rápida de compuestos farmacéuticos. En los últimos años, se ha desarrollado un número de técnicas para caracterizar respuestas fenotípicas de células a agentes perturbadores como pequeñas moléculas o biologías. La gran mayoría del trabajo registrado ha utilizado ensayos bioquímicos a gran escala tradicionales, o técnicas con células individuales basadas en un cribado de alto contenido (microscopía automatizada). Por ejemplo, véase Abraham *et al*. ("High content screening applied to large-scale cell biology". Trends Biotechnol. 22,15-22,2004) y Giuliano *et al*. ("Advances in High Content Screening for Drug Discovery". ASSAY Drug Dev. Technol. 1,565-577,2003).

Más recientemente, se han implementado métodos estadísticos innovadores en el análisis de conjuntos de datos de cribado complejos. Estos métodos pueden proporcionar medios para determinar correlaciones entre conjuntos de datos. Se describen modelos basados en vía simple para cribado en Cong *et al*. ("Method for using division arrested cells in screening assays," EP 1581645).

Cong propone estudiar la transducción de señales en células con crecimiento detenido y utilizar los sistemas de este tipo para cribar agentes que modulan la actividad de receptores de superficie celular tales como los receptores adrenérgicos β2 (β2AR), un tipo de receptor acoplado a proteína G. Cong demuestra que, incluso en células con crecimiento detenido de este tipo, el tratamiento con isoproterenol (100 μM) aumenta la actividad de la fosfatasa alcalina secretada (SEAP), que a su vez establece que las células con crecimiento detenido aún tienen intactas las vías de transducción de señales hasta la respuesta transcripcional, y se pueden llevar a cabo ensayos de enzima informadora utilizando sistemas de este tipo.

Una técnica similar es proporcionada por Hytopoulos *et al*. ("Methods for analysis of biological dataset profiles". Publicación de solicitud de patente estadounidense n.º 2007-0135997).

Hytopoulos describe métodos para evaluar perfiles de conjuntos de datos biológicos. Los conjuntos de datos que comprenden información para múltiples parámetros celulares se comparan e identifican. Un conjunto de datos típico comprende lecturas a partir de múltiples parámetros celulares resultantes de exponer las células a factores biológicos en ausencia o presencia de un agente candidato. Para análisis de múltiples sistemas definidos por el contexto, se concatenan los datos de salida a partir de múltiples sistemas. Sin embargo, Hytopoulos no describe las etapas precisas del método para crear y formar los perfiles de respuesta. Adicionalmente, Hytopoulos no proporciona ninguna realización de trabajo para poner en práctica la metodología con un espécimen biológico.

Berg *et al.* ("Function homology screening". Patente estadounidense n.º 8.467.970) describe métodos para evaluar homología funcional entre fármacos. Los métodos implican exponer células a fármacos y evaluar el efecto de alterar el entorno celular monitorizando múltiples parámetros de salida. Dos entornos diferentes, tales como aquellos con diferentes compuestos presentes en el entorno, se pueden comparar directamente para determinar similitudes y diferencias. Basándose en estas comparaciones, los compuestos se pueden caracterizar a un nivel funcional, permitiendo identificar las vías y predecir los efectos secundarios de los compuestos. Berg también describe una representación de los datos medidos en forma de "biomapa", que es un mapa de calor muy simplificado que muestra gráficamente todos los parámetros celulares medidos. Berg se relaciona con la medición de las vías de señalización biológica, en lugar de las respuestas fisiológicas al estrés.

Friend *et al*. ("Methods of characterizing drug activities using consensus profiles". Patente estadounidense n.º 6.801.859) describe un método para medir patrones de respuesta biológica, tales como patrones de expresión génica, en respuesta a diferentes tratamientos con fármacos. Los perfiles (curvas) de respuesta, que se crean al exponer un sistema biológico a concentraciones variables de fármacos, pueden describir la respuesta biológica de células a un grupo o clase particular de fármacos. Las curvas de respuesta se aproximan utilizando modelos. Los vectores de datos resultantes que forman curvas o perfiles, o sus modelos paramétricos, se pueden comparar utilizando varias medidas de similitud. Esta comparación forma una matriz de distancia que se puede utilizar subsecuentemente en un algoritmo de agrupamiento jerárquico para construir un árbol que represente la similitud de los perfiles. Sin embargo, los perfiles desarrollados por Friend *et al*. se limitan a modelos matemáticos paramétricos y de tipo vectorial simples.

Además, los métodos de elaboración de perfiles de las aplicaciones mencionadas anteriormente de las publicaciones de Berg *et al*. y Friend *et al*. son limitados y, en particular, no proporcionan el uso de distribuciones para desarrollar perfiles de posibles fármacos desconocidos.

Relativamente, se ha realizado poco trabajo en esta área utilizando citometría de flujo, que permite el análisis de células individuales de estados celulares en grandes poblaciones de células. Véase, por ejemplo, Edwards *et al*. ("Flow cytometry for high-throughput, high-content screening". Curr. Opin. Chem. Biol. 8, 392...398, 2004, 2004); Oprea *et al*. ("Associating Drugs, Targets and Clinical Outcomes into an Integrated Network Affords a New Platform for Computer-Aided Drug Repurposing". Mol. Inform. 30, 100... 111,2011); Robinson *et al*. ("High-throughput secondary screening at the single-cell level". J. Lab. Autom. 18, 85-98, 2013) y Sklar *et al*. ("Flow cytometry for drug discovery, receptor pharmacology and high throughput screening." Curr. Opin. Pharmacol. 7, 527-534,2007).

Sin embargo, la disponibilidad reciente de sistemas de manipulación de líquidos de alto rendimiento para citometría ha hecho factible procesar una placa entera de 96 o 384 pocillos en pocos minutos, muestreando varios miles de células por pocillo, haciendo que la citometría sea cada vez más atractiva para ensayos celulares de alto rendimiento. Los informes que describen el uso de citometría de flujo de alto rendimiento se centran típicamente en ensayos relativamente simples que obtienen de 1 a 5 variables diferentes que describen la fisiología celular para las células analizadas. Desde una perspectiva matemática, los datos recolectados en estos ensayos se pueden describir como una matriz en la que las filas almacenan información sobre células individuales, y las columnas describen la cantidad medida (por ejemplo, características de dispersión de la luz, señales de intensidad de fluorescencia, etc.). Las características medidas se pueden resumir por una variedad de estadísticas. Más comúnmente, se utiliza la intensidad media de fluorescencia en una región de interés. Después de reducir los datos, se representan los resultados de un experimento por medio de un vector con elementos que son los valores de las estadísticas de resumen elegidas. Si un experimento implica someter a prueba en un número de concentraciones diferentes de un fármaco, el resultado final es una matriz 2-D, con columnas individuales que describen las curvas de respuesta. Se pueden utilizar parámetros adicionales (por ejemplo, diferentes periodos de tiempo de incubación del fármaco) para añadir dimensiones adicionales a la matriz.

Tradicionalmente, las curvas de respuesta de los fármacos se aproximan por medio de un modelo matemático *a priori* (como una curva log-normal sigmoidea, curva de Gompertz, Weibull, etc.) y la información de la respuesta medida del fármaco se reduce a unos pocos parámetros (o incluso a un único parámetro) que describe las curvas. El proceso completo produce un resumen de respuesta del compuesto considerablemente abreviado: típicamente una "firma" que comprende un número de valores EC50, esto es, valores que representan una concentración de un compuesto que induce una respuesta a medio camino entre la línea base y el máximo después de un tiempo de exposición específico.

Los planteamientos de este tipo se someten a importantes limitaciones inherentes que no se pueden paliar fácilmente, si acaso. Primero, estos asumen la existencia previa de un modelo matemático adecuado conocido con una parametrización apropiada que describe la respuesta de todos los compuestos sometidos a prueba. Segundo, suponen que un único parámetro ($EC_{50}$) derivado de una curva sigmoidea lleva toda la información necesaria sobre el patrón de respuesta del compuesto. Y, tercero, analizan las respuestas manifestadas por medio de los parámetros medidos de manera separada, es decir, de manera unidimensional. El análisis de los datos y la extracción de características que lleva a la formación de las curvas de respuesta también son problemáticos.

Asimismo, el procesamiento de datos citométricos tradicional y establecido depende de un proceso llamado *gating*, que implica separar manualmente las poblaciones de interés para calcular características estadísticas simples de estas poblaciones (media, mediana, coeficiente de varianza, etc.). Este *gating* puede ser sumamente subjetivo, y es difícil de reproducir en una configuración automatizada. Adicionalmente, las características calculadas no se escalan o estandarizan para reflejar el intervalo de posibles respuestas biológicas o la precisión de las mediciones de la citometría.

El documento US8467970B2 describe un método para cribar un agente biológicamente activo basado en el análisis de respuestas biológicas complejas en cultivo. Los métodos para seleccionar las células y las condiciones de cultivo para los cribados de este tipo se describen, al igual que la identificación de un conjunto de parámetros discretos que se va a medir, y el uso de análisis de biomapa para identificar y caracterizar los candidatos a fármacos y secuencias genéticas que actúan en las vías.

Las realizaciones descritas en este documento proporcionan métodos para superar los defectos significativos de los actuales métodos de cribado fenotípico, en algunas realizaciones, al emplear una nueva metodología para cuantificar las respuestas del compuesto. Las realizaciones descritas en este documento proporcionan un número de técnicas de procesamiento de datos y de adquisición de datos innovadoras, que permiten comparaciones significativas de las huellas multidimensionales del compuesto sin comprometer la calidad de la información, y sin suposiciones *a priori* sobre las respuestas.

-II-

En términos generales, la invención se refiere a un método como se describe en el punto 1 a continuación. Se describen otras realizaciones en los subsecuentes párrafos. Los párrafos numerados son autorreferenciales.

1. Un método para caracterizar el efecto de un agente en células que comprende:

a) poner en contacto muestras de una población de células con una serie de concentraciones diferentes de un agente;

b) medir cada uno de dos o más parámetros fenotípicos en células individuales de una muestra respectiva para cada concentración;

c) obtener una distribución de los valores medidos para las células para cada parámetro fenotípico y cada concentración;

d) medir cada uno de los dos o más parámetros fenotípicos en células individuales de una muestra de control no expuesta al agente;

e) obtener una distribución de los valores medidos para las células para cada parámetro fenotípico para la muestra de control;

f) calcular, para cada parámetro fenotípico, disimilitudes entre la distribución obtenida en la etapa e) y cada una de las distribuciones obtenidas en la etapa c);

g) combinar las disimilitudes calculadas para formar un tensor de respuesta multiparamétrico;

en donde el tensor de respuesta multiparamétrico es una huella característica del efecto del agente en las células que se puede comparar directamente con otras huellas generadas de la misma manera.

2. Un método según el punto 1, en donde el método se utiliza para generar y almacenar huellas.

3. Un método según el punto 1, en donde el método se utiliza para hacer comparaciones de huellas.

4. Un método según el punto 3, en donde las comparaciones se utilizan para predecir efectos *in vivo*.

5. Un método según cualquiera de los puntos precedentes, en donde los parámetros fenotípicos se miden en las células individuales por medio de citometría celular.

6. Un método según cualquiera de los puntos precedentes, en donde los parámetros fenotípicos incluyen dos o más de viabilidad celular, fase del ciclo celular, integridad de la membrana mitocondrial, toxicidad mitocondrial, concentración de glutatión, especies reactivas de oxígeno, especies reductoras, permeabilidad de la membrana citoplasmática, lesión del ADN, un constituyente de la vía de señalización de respuesta al estrés, un constituyente de la vía de respuesta inflamatoria, un constituyente de la vía de apotosis y una peroxidasa lipídica.

7. Un método según cualquiera de los puntos precedentes, en donde los parámetros fenotípicos incluyen uno o más de NFκB, SAPK, P13K, AKT, una familia de proteínas Bcl-1, GSk3B y quinasa S6 ribosomal.

8. Un método según cualquiera de los puntos precedentes, en donde uno de los parámetros fenotípicos es ciclo celular.

9. Un método según cualquiera de los puntos precedentes, en donde cada población de células se marca funcionalmente con una pluralidad de colorantes fluorescentes y los parámetros fenotípicos se detectan y cuantifican en términos de señal(es) de emisión espectral que se generan cuando dichas poblaciones de células marcadas se someten a análisis citométrico.

10. El método según cualquiera de los puntos precedentes, en donde uno de los parámetros fenotípicos es ciclo celular y se cuantifica en términos de uno o más de: HOECHST 33342, DRAQ5, yoduro YO-PRO-1, DAPI, CYTRAK ORANGE, ciclina A, ciclina B, ciclina E, o proteína histónica fosforilada.

11. Un método según cualquiera de los puntos precedentes, que comprende:

poner en contacto las muestras con n concentraciones diferentes de un agente, donde $i \in \{2,...,n\}$, y obtener al menos una muestra de control, $\kappa$, correspondiente a $m$ condiciones de control diferentes, donde m es al menos 1;

medir cada uno de *p* parámetros fenotípicos, $\psi$, en células individuales de cada muestra, donde *p* es al menos 2 y donde $\psi_p$ denota la medición de cada parámetro fenotípico particular,

obtener una distribución $C_\kappa^{(\psi_p)}$ de los valores medidos para cada condición de control $\kappa$ para cada parámetro fenotípico $\psi_p$ y obtener una distribución $S_i^{(\psi_p)}$ de los valores medidos para cada condición de concentración *i* para cada parámetro fenotípico $\psi_p$,

calcular disimilitudes calculando, para cada parámetro fenotípico $\psi_p$, distancias *d* entre pares entre las distribuciones de cada condición de control $C_\kappa^{(\psi_p)}$ y cada condición de concentración $S_i^{(\psi_p)}$,

donde

$$d_{\kappa,i}^{(\psi_p)} = D\left(C_\kappa^{(\psi_p)},\ S_i^{(\psi_p)}\right)$$

y D es una función de distancia;

en donde el tensor de respuesta multiparamétrico es un tensor de respuesta multiparamétrico A, que se representa por

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_{[\kappa_1,\psi_1]} & \cdots & \mathbf{a}_{[\kappa_1,\psi_p]} \\ \vdots & \ddots & \vdots \\ \mathbf{a}_{[\kappa_m,\psi_1]} & \cdots & \mathbf{a}_{[\kappa_m,\psi_p]} \end{bmatrix}$$

y en donde cada uno de $\mathbf{a}_{[\kappa,\psi]}$ comprende todas las distancias entre pares calculadas para un parámetro fenotípico y uno de control:

$$\mathbf{a}_{[\kappa;\psi]} = \begin{bmatrix} d_{\kappa,1}^{(\psi_p)} = D\left(C_\kappa^{(\psi_p)},\ S_1^{(\psi_p)}\right) \\ d_{\kappa,2}^{(\psi_p)} = D\left(C_\kappa^{(\psi_p)},\ S_2^{(\psi_p)}\right) \\ \vdots \\ d_{\kappa,n}^{(\psi_p)} = D\left(C_\kappa^{(\psi_p)},\ S_n^{(\psi_p)}\right) \end{bmatrix}.$$

12. Un método según el punto 11, en donde el tensor multiparamétrico A se descompone en uno de:

a) un conjunto de tensores de rango uno $\mathbf{A}^{(i)}$, en donde

$$\mathbf{A} \approx \left[\!\left[\lambda; \mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots, \mathbf{A}^{(n)}\right]\!\right] = \sum_{r=1}^{R} \lambda_r \mathbf{a}_r^{(1)} \circ \mathbf{a}_r^{(2)} \circ \cdots \circ \mathbf{a}_r^{(p)},$$

y

b) un tensor núcleo G multiplicado por una matriz $\mathbf{M}^{(j)}$ a lo largo con cada modo:

$$\mathbf{A} \approx \left[\!\left[ \mathbf{G}; \mathbf{M}^{(1)}, \mathbf{M}^{(2)}, \ldots, \mathbf{M}^{(p)} \right]\!\right] = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_p=1}^{R_p} g_{r_1 r_2 \ldots r_p} \mathbf{m}_{r_1}^{(1)} \circ \mathbf{m}_{r_2}^{(2)} \circ \cdots \circ \mathbf{m}_{r_N}^{(p)}$$

13. Un método según cualquiera de los puntos 11-12, en donde la disimilitud entre un tensor de respuesta multiparamétrico **A** (primera huella) y un tensor de respuesta multiparamétrico **B** (segunda huella) se calcula según la siguiente fórmula:

$$\mathbf{w}_{\mathbf{A},\mathbf{B}} = D(\mathbf{A},\mathbf{B}) = d\left(\mathbf{a}_{:jk}, \mathbf{b}_{:jk}\right),$$

donde D es la disimilitud entre los sensores **A** y **B**, *d* es una función de distancia que compara las fibras de modo 1 de cada uno de los tensores, y **w** es un indicativo de vector de distancia de la disimilitud entre la primera y la segunda huella.

14. Un método para predecir los efectos *in vivo* de un agente, que comprende caracterizar el efecto del agente según el método del punto 11,

comprimir el tensor de respuesta multiparamétrico **A** para reducir su complejidad por medio de un método de compresión de baja pérdida;

calcular la diferencia entre el tensor comprimido y los tensores calculados de manera comparable para un banco de compuestos de efectos *in vivo* conocidos;

a partir de los valores de las diferencias, predecir los efectos *in vivo* del agente.

15. Un método para producir un banco de huellas características de los efectos de agentes en células, que comprende generar para cada uno de una pluralidad de agentes un tensor de respuesta multiparamétrico por medio del método de cualquiera de los puntos 1-13, en donde cada tensor de respuesta multiparamétrico es una huella característica de los efectos del agente en células, y

almacenar las huellas en un medio legible por ordenador de tal manera que se puedan comparar entre sí y con las huellas dactilares de otros agentes producidos de la misma manera,

de esta manera, se forma una base de datos de huellas s características del efecto de los agentes en células.

-III-

Breve descripción de los dibujos

Varias características y ventajas de las realizaciones descritas en este documento se pueden apreciar por completo ya que estas se entenderán mejor cuando se consideren a la luz de los dibujos que acompañan:

La Figura 1A muestra perfiles citométricos de células normales como se describe en el Ejemplo 1. Las células se marcaron específicamente para medir los niveles de tiol celular/nuclear, predominantemente niveles de glutatión (GSH), como se mide con monobromobimano (MBBR). Las células se tiñeron utilizando varios colorantes fluorescentes, como se describe en el Ejemplo 1. La permeabilidad de la membrana citoplasmática se midió con CALCEÍNA AM, se midieron las especies reactivas de oxígeno (ROS; relacionadas con la función mitocondrial) con MITOSOX RED™, se midió la permeabilidad de la membrana nuclear y citoplasmática (viva/muerta) con SYTOXRED™.

La Figura 1B muestra los datos de la Figura 1A en forma de gráfico de "radar".

La Figura 2 es un conjunto de gráficos de células tratadas con 100 µM de mixotiozol (un inhibidor del complejo bcl del citocromo mitocondrial) analizado como se describe en el Ejemplo 2 para dispersión frontal (FS), glutatión, calceína AM, ROS y viabilidad celular. La comparación de la Figura 1A con la Figura 2 muestra que el tratamiento con 100 µM de mixotiozol perturba la permeabilidad (viabilidad) de la membrana nuclear y citoplasmática.

La Figura 3 es un diagrama de dispersión de células tratadas con 33 µM de carbonil cianida-4-(trifluorometoxi)fenilhidrazona "FCCP", un vehículo de iones móviles ionóforo que actúa como agente desacoplante en la mitocondria, analizado como se describe en el Ejemplo 2 para dispersión frontal, glutatión, calceína AM, ROS y viabilidad. Las flechas indican un cambio en el parámetro celular en comparación con la población de control expuesta en la Figura 1A.

La Figura 4A es un diagrama de dispersión de células tratadas con 100 µM de fluoxetina (un inhibidor selectivo de la recaptación de serotonina) analizado como se describe en el Ejemplo 2 para dispersión frontal, glutatión, calceína AM, ROS y viabilidad.

La Figura 4B es un conjunto de perfiles multidimensionales de la Figura 2, la Figura 3 y la Figura 4A, representado gráficamente a lo largo de los cuatro ejes: uno para las células tratadas con 100 µM de mixotiozol (parte superior izquierda), 33 µM de FCCP (parte superior derecha) o 100 µM de fluoxetina (parte inferior central).

La Figura 5 muestra los resultados de un análisis de control de las fases del ciclo celular en células HL-60 vivas utilizando colorante VYBRANVIOLET™ como se describe en el Ejemplo 3. El recuento de células se indica en el eje vertical. El contenido de ADN se indica en el eje horizontal. La punta morada de la izquierda indica células en fase G1. El área rosada ancha indica células en fase S. El pico morado entramado de la derecha indica células en fase G2/M.

La Figura 6A es una gráfica que muestra el potencial de la membrana mitocondrial (MMP) de células HL-60 expuestas a valinomicina, que provoca despolarización de MMP rápida, como se describe en el Ejemplo 4A. La despolarización aumenta las señales de fluorescencia verdes y rojas del colorante JC-9. Tanto la fluorescencia roja como la verde son significativamente más altas cuando las células se expusieron a la valinomicina, con el pico más alto a aproximadamente ~0,045 µM.

La Figura 6B es una gráfica que muestra JC-9 fluorescente en células HL60 tratadas con la idarubicina del fármaco antileucémico antraciclina, un análogo de daunorubicina, como se describe en el Ejemplo 4B. La idarubicina se inserta en el ADN y evita el desenrollamiento durante la replicación del ADN. La fluorescencia roja del colorante JC-9 comienza a aumentar significativamente a 1,23 µM de idarubicina y continúa aumentado por encima de eso. La fluorescencia verde no aumenta.

La Figura 7 muestra los resultados del análisis del ciclo celular de células permeabilizadas y fijadas como se describe en el Ejemplo 5A.

El panel superior izquierdo muestra una clara delineación de células en las fases G1, S y G/2M del ciclo por medio de tinción de ADN (como se describió anteriormente en la Figura 5).

El panel superior derecho muestra los resultados para histona H3 fosforilada ("H3-F") medida utilizando un anticuerpo H3-F conjugado con ALEXA FLUOR® 647 como función del ciclo celular. El contenido de ADN se midió igual que en el panel superior izquierdo. Los resultados muestran que H3-F se expresó solo en G2/M.

Los dos paneles inferiores muestran los resultados para Ciclina A2 medida utilizando un anticuerpo anti-A2 conjugado con PE como función del ciclo celular. El contenido de ADN se midió igual que en el panel superior izquierdo. La Ciclina A2 aumenta mientras las células avanzan a través de S y forman dos poblaciones de expresión de Ciclina A2 que difieren en G2/M.

El panel inferior derecho muestra los resultados para Ciclina A2 frente a aquellos para H3-F durante el progreso celular a través del final de S, hasta G2, hasta M.

Al realizar el "gating" en el análisis subsecuente solo en estas poblaciones del ciclo celular (utilizando el contenido de ADN), es evidente que la H3-F se expresa primero en células con los niveles de Ciclina A2 más altos (población G2), que la H3-F se mantiene en los niveles más altos mientras la Ciclina A2 se degrada, y que la H3-F se "pierde" entonces (por medio de desfosforilación del residuo de serina específico que se fosforiló tras la entrada en la G2) mientras el progreso celular vuelve desde la mitosis (M) al G1. En el análisis citométrico de flujo tradicional, estas secuencias de cambios en la expresión de proteínas se establecen al realizar "gating" (seleccionar) manualmente de manera cuidadosa en diferentes poblaciones celulares (basado en el contenido de ADN) y analizar subsecuentemente la expresión de proteínas (u otras dianas definidas por diferentes conjugados de anticuerpos).

La Figura 8 muestra un ejemplo de vías de transducción de señal situado más abajo del receptor 4 de tipo toll (TLR4) encontrado en monocitos sanguíneos periféricos. Los inhibidores representativos de la quinasa PI3 (PI3K) y vías de quinasa proteína asociadas a mitógenos (MAPKK) se indican con flechas.

La Figura 9 muestra la cinética (en minutos) de respuestas de transducción de señal en monocitos sanguíneos periféricos humanos a LPS en ausencia (panel izquierdo) y en presencia del inhibidor de la quinasa PI3 GDC0941, como se describe en el Ejemplo 6.

Verde - IκBα; Rojo- P-ERK; Naranja- P-Akt; Azul - P-S6.

Como se puede ver en el panel a la izquierda, el tratamiento con LPS activa la quinasa Iκ (que da lugar a la degradación proteasomal de IκBα - pérdida de verde/señal fluorescente de ALEXA FLUOR® 488), y activa (fosforila) ERK, Akt y S6 (fosfoproteínas individuales detectadas por medio de citometría de flujo utilizando anticuerpos contra P-ERK, P-Akt y P-S6 conjugados con ALEXA FLUOR® 647, PE, y azul pacífico, respectivamente).

El panel a la derecha muestra que en presencia del inhibidor de la quinasa PI3 GDC0941, P-Akt no se fosforila, y la cinética de la fosforilación de ERK y S6 se retrasa, en comparación con las células de control tratadas con LPS solo (panel a la izquierda). Se muestra además que la inhibición de PI3K no tiene impacto en la degradación proteasomal de IκBα.

La Figura 10 muestra el efecto de exponer células a GLEEVEC™ (imatinib, o STI571), los detalles de esto se proporcionan en el Ejemplo 7. Como se puede ver en la Figura, el tratamiento de células K562 durante 30 min con 2 μM da lugar a >95% de inhibición de la fosforilación de la diana STAT5 situado más abajo. STATS fosforilado actúa como activador transcripcional de varias proteínas diana, incluida Ciclina D, y la expresión constitutiva de Ciclina D mantiene a las células K562 en el ciclo celular.

Como se puede ver en la Figura, aunque la fosforilación de STATS se inhibe después de una exposición a imatinib de 30 min (como se demuestra por un cambio en la población total de células que son P-Stat5 positivas desde por encima de la línea del umbral hasta por debajo de la línea del umbral en células tratadas con GLEEVEC), no hay un cambio concomitante en el ciclo celular, como se mide por el contenido de ADN (véase el recuadro).

La Figura 11 proporciona un ejemplo en dos dimensiones y tres dimensiones de una visualización EMD multidireccional de datos de respuesta que muestran que los compuestos que parecen estar juntos en dos dimensiones, en realidad, pueden estar muy distantes entre sí cuando se mapean en tres o más dimensiones.

La Figura 12 muestra una huella de fármaco para el fármaco antidiabético troglitazona, los detalles de este se describen en el Ejemplo 8.

La Figura 13 muestra dendrogramas representativos para visualizar similitudes fisiológicas entre varios fármacos basadas en el fenotipo celular (por ejemplo, integridad de la membrana nuclear y celular o presencia de ROS) que se analizan.

La Figura 14 muestra un ejemplo de un tensor multidireccional que representa una huella de respuesta del fármaco.

La Figura 15 muestra la evolución de una nube representativa. Los cambios en el espacio definidos por disimilitudes entre nubes de puntos complejas forman una trayectoria complicada, que describe únicamente las características del compuesto que provoca esos cambios.

La Figura 16 es un diagrama de flujo que muestra las etapas generales del proceso para llevar a cabo los ensayos de fisiología celular.

La Figura 17 es un diagrama de flujo que muestra las etapas en el análisis de datos utilizando los métodos de tensor descritos en este documento.

La Figura 18 muestra los diagramas de dispersión 2-D de los resultados de un análisis de citometría de flujo simulado de células que fueron (A) expuestas y (B) no expuestas a un compuesto. (A) y (B) representan controles positivos y negativos para el agente, respectivamente. FL1 y FL2 son las señales fluorescentes 1 y 2.

La Figura 19 muestra los diagramas de dispersión 2-D de los resultados de un análisis de citometría de flujo simulado de diez muestras expuestas a concentraciones crecientes de un agente biológico, como se describe en el Ejemplo 9. Cada ejemplo contiene 5000 células. Las concentraciones se aumentaron de A a J. El número de células en el clúster tipo control positivo aumentó de 4909 en A a 29 en J. FL1 y FL2 son señales fluorescentes.

La Figura 20 proporciona dos gráficas de las curvas de respuesta para el análisis de citometría de flujo simulado descrito en el Ejemplo 9 y la Figura 18. (A) muestra los resultados para cada muestra como función del número de células en el clúster de control tipo positivo. (B) muestra los resultados para cada muestra como función del porcentaje de células en el clúster de control tipo positivo.

La Figura 21 muestra las cuatro distancias calculadas para cada muestra para proporcionar una medida de disimilitud, como se describe en el Ejemplo 9. (1) La distancia entre la muestra y el control negativo para el parámetro uno. (2) La distancia entre la muestra y el control positivo para el parámetro 1. (3) La distancia entre la muestra y el control negativo para el parámetro dos. (4) La distancia entre la muestra y el control positivo para el parámetro 2.

La Figura 22 muestra curvas de respuesta para la simulación descrita en el Ejemplo 9 derivadas por medio de descomposición de tensor poliádica. Los resultados se centraron al restar la media del vector y dividir por la desviación estándar. Los resultados se expresan como factor z, y se muestran en la gráfica de la izquierda. La gráfica de la derecha muestra los resultados normalizados a la diferencia entre los controles negativos y positivos, es decir, en los que la diferencia entre los controles negativos y positivos se define como unidad (uno) y los resultados están escalados a esta diferencia.

La Figura 23 es una gráfica de los resultados de la simulación descrita en el Ejemplo 9 expresados convencionalmente en porcentaje de células en el clúster tipo control negativo a lo largo del eje vertical y expresados en términos de distancia/disimilitud a lo largo del eje horizontal. La distancia/disimilitud se calculó para los mismos resultados utilizando controles positivos y negativos como se ilustra en la Figura 18, como se describe en el Ejemplo 9.

La Figura 24 muestra los diagramas de dispersión 2-D de los resultados de un análisis de citometría de flujo de diez muestras expuestas a concentraciones crecientes de una valinomicina, como se describe en el Ejemplo 9B. Se midió la relación de fluorescencia roja (eje vertical) y verde (eje horizontal) de JC9 para determinar el potencial de la membrana mitocondrial. La concentración de valinomicina aumentó de A (concentración menor) a J (concentración mayor).

La Figura 25 muestra la disimilitud de las curvas de respuesta para el análisis (ciclos) de valinomicina descrito en el Ejemplo 9B e ilustrado en la Figura 24, y para los datos de idarubicina (diamantes) y de acetaminofeno (triángulos) obtenidos de la misma manera. Se calcularon las disimilitudes como se describe en el ejemplo y se ilustra en la Figura 21. La gráfica de la izquierda muestra disimilitud como función de número de muestra. La gráfica de la derecha muestra disimilitud como función de concentración.

-IV-

Los aspectos ilustrativos de la descripción proporcionan métodos automatizados, independientes al observador, sólidos, reproducibles y genéricos para recolectar, compilar, representar y extraer información compleja basada en la población, particularmente, por ejemplo, información basada en citometría, como, por ejemplo, para cuantificar y comparar respuestas fisiológicas de células expuestas a compuestos químicos, tales como fármacos. Varios aspectos de la descripción proporcionan métodos para caracterizar respuestas por medio de tensores de respuesta. Los aspectos ilustrativos de la descripción proporcionan el uso de varias medidas estadísticas de distancias entre distribuciones en una o más dimensiones, y medidas de disimilitud entre vectores de respuesta agrupados en tensores multidireccionales. En varios aspectos de la descripción, las diferencias en respuestas celulares a dos (o más) compuestos químicos se caracterizan como la diferencia entre dos tensores de respuesta ("huellas ") que representan dichos compuestos. Los aspectos de la descripción proporcionan métodos para generar dichas huellas y métodos para manipularlas, procesarlas, almacenarlas, clasificarlas y utilizarlas.

En varios aspectos de la descripción descrita en este documento, por ejemplo, se analizan conjuntos de datos biológicos para determinar coincidencias entre ellos, con frecuencia, entre conjuntos de datos de prueba y de control, o entre conjuntos de datos de prueba y conjuntos de datos de perfil. Se pueden hacer comparaciones entre dos o más conjuntos de datos, donde un conjunto de datos típico comprende lecturas a partir de múltiples parámetros celulares, tales como aquellos que se originan a partir de la exposición de células a factores biológicos en ausencia o presencia de un agente candidato, donde el agente puede ser, por ejemplo, un agente genético, por ejemplo, secuencia codificante expresada, o un agente químico, por ejemplo, un posible fármaco, o una toxina ambiental. En varios aspectos de la descripción, las mediciones se realizan utilizando citometría, por ejemplo, citometría de flujo.

Citometría

Los métodos de varios aspectos de la descripción descrita en este documento son adecuados para análisis de datos multiparamétricos complejos sobre células individuales en poblaciones celulares, como se determina por citometría. Las técnicas e instrumentos citométricos, resumidos en este documento, (por ejemplo, citometría de flujo y citometría de imagen) permiten la medición simultánea de múltiples características intrínsecas (por ejemplo, dispersión de la luz, volumen celular, etc.) o características derivadas (por ejemplo, fluorescencia, absorción, etc.) de células individuales. La dispersión de la luz y la fluorescencia representan las mediciones utilizadas más comúnmente para aplicaciones citométricas actuales. Las mediciones de fluorescencia se pueden realizar tanto utilizando fluoróforos "intrínsecos" presentes de manera natural en las células (tales como, por ejemplo, porfirinas, flavinas, lipofuscinas, NADPH), como fluoróforos diseñados genéticamente para expresión específica (por ejemplo, GFP, RFP, etc.), como informadores fluorescentes que se dirigen a epítopes específicos o estructuras en o sobre varios tipos celulares (por ejemplo, anticuerpos conjugados con fluoróforo, aptámeros, presentación en fagos, o péptidos, o informadores que se convierten de estados no fluorescentes a fluorescentes por medio de enzimas específicas en o sobre células).

Las técnicas citométricas útiles en aspectos de la descripción descrita en este documento utilizan células vivas (por ejemplo, utilizando sondas que informan de la célula en aspectos de "fisiología" celular, tales como, por ejemplo, potencial de la membrana celular, ROS, contenido de glutatión, o una combinación de estos). Las técnicas citométricas útiles en algunos aspectos de la descripción emplean adicionalmente células que se fijan y se permeabilizan para permitir el transporte de fluoróforos, informadores conjugados, etc., al citoplasma y/o al núcleo.

Métodos generales para ensayos celulares utilizando citometría de flujo

Los métodos generales útiles para citometría según varios aspectos de la descripción descritos en este documento se describen a continuación y se exponen en un diagrama de flujo generalizado en la Figura 16.

Cultivo de células independientes de anclaje

Se emplean células y métodos adecuados para ensayos y análisis de actividad por medio de citometría de flujo que se conocen y se emplean habitualmente en la técnica para llevar a cabo aspectos de la descripción descrita en este documento.

Las células para ensayos se pueden obtener de fuentes comerciales o de otro tipo. Se pueden utilizar células derivadas de cáncer humano, tales como aquellas derivadas de leucemias (por ejemplo, células HL-60 utilizadas actualmente en ensayos de fisiología celular), que crecen sin fijarse al recipiente de cultivo. En términos generales, las células se pueden almacenar en nitrógeno líquido según los métodos celulares estándar. Las células congeladas se descongelan rápidamente en un baño de agua de 37 grados centígrados, y se cultivan en matraces estacionarios en medio de cultivo tisular fresco precalentado en una incubadora de cultivo tisular de 37 grados centígrados. Típicamente, el medio de cultivo tisular se reemplaza diariamente durante los primeros 2-4 días de cultivo, para diluir las células DMSO que se congelan.

Una vez que el crecimiento se establece en matraces estacionarios (el número de células y la viabilidad se monitorean utilizando un contador celular Vi-cell™), se pueden retirar partes alícuotas de células del almacenamiento de congelado (estas células de paso tempranas solo se utilizan de reserva). Además, estas células se pueden utilizar para establecer cultivos en frascos rotativos necesarios para tener un número suficiente de células para ensayos de placas. Las células que crecen en matraces se colocan en frascos rotativos a una concentración celular relativamente alta (~$10^6$ células por ml en 200 ml de medio de cultivo tisular fresco) y se cultivan en una incubadora de cultivo tisular. Inicialmente, los cultivos en frascos rotativos se alimentan típicamente al añadir medio de cultivo tisular fresco. Una vez que se establece el crecimiento, las células se retiran según se necesite para mantener las células en una concentración de 0,5-1,5 x $10^6$ de células viables/ml. Muchos tipos celulares se adaptan lentamente a los cultivos en frascos rotativos, y necesitan semanas para adaptarse con éxito a estos tipos de cultivo. La adaptación exitosa a frascos rotativos queda patente por una viabilidad alta continua (~95%) y por tasas de crecimiento consistentes (medidas utilizando tiempo de multiplicación por dos). Cuando se adaptan de manera exitosa, las reservas de células se congelan (en tubos estériles de 50 ml que contienen suficientes células para iniciar un nuevo cultivo en frascos rotativos) para mantener las células utilizadas en los ensayos en un número de paso bajo similar (se detalla más adelante). Las células mantenidas en frascos rotativos se recolectan para placas de ensayo, se centrifugan y se vuelven a suspender en medio de cultivo tisular fresco a una concentración celular adecuada para el ensayo que se va a realizar (número de células y viabilidad medidos y registrados para cada cosecha).

Como se indicó anteriormente, las células adaptadas a frascos rotativos se pueden congelar para un uso futuro, para mantener el número de células de paso bajo similar para todos los ensayos de placas. Los cultivos celulares en frascos rotativos se pueden mantener durante un mes antes de intercambiar a un nuevo lote de células congeladas de paso bajo. Durante el mes de uso habitual, un tubo de células congeladas se descongela típicamente y se vuelve a establecer el cultivo en frascos rotativos. Una vez que se adapta de manera exitosa al cultivo en frascos rotativos (como anteriormente), el lote más nuevo de células normalmente se evalúa primero para rendimiento del ensayo (véase estudios "cruzados", más adelante), antes de que este lote se utilice en ensayos de placas. Establecer células congeladas para el cultivo en frascos rotativos y someterlas a prueba habitualmente lleva de 10 a 21 días.

En términos generales, las células se someten a prueba habitualmente en múltiples etapas en el proceso de cultivo por contaminación de micoplasma. Estas incluyen cultivos iniciales en matraces, células adaptadas a frascos rotativos, y cada tubo de células congeladas (sometidas a prueba antes de cada estudio "cruzado"). Las pruebas de micoplasma se pueden proporcionar por una empresa de pruebas certificada externa, típicamente, utilizando un ensayo basado en PCR.

Almacenamiento del compuesto y preparaciones de ensayo del compuesto

Los compuestos de prueba se obtienen generalmente como reservas de 10 mM en DMSO depositadas en placas de 96 pocillos. Las placas de compuesto se almacenan selladas a temperatura ambiente en la oscuridad. Para los ensayos del compuesto, las disoluciones madre se diluyen (diluciones de compuesto de 5 a 10 etapas en DMSO) y se depositan en placas de ensayo utilizando un sistema de manipulación de líquidos. La deposición de todas las diluciones y del compuesto en placas de ensayo se realiza el mismo día que se realiza el ensayo. El volumen final del compuesto depositado en cada pocillo es 2 ul, con una concentración final de 1% de DMSO después de añadir las células.

Se debería evaluar la reproducibilidad del ensayo utilizando compuestos de prueba. Se ha utilizado un conjunto de 16 compuestos que se ha documentado que impacta en las mediciones fisiológicas celulares específicas para someter a prueba la reproducibilidad de los ensayos de fisiología celular. Estos compuestos se almacenan, como anteriormente, como disoluciones de ensayo de 10 mM en DMSO en placas de 96 pocillos. Para los estudios "cruzados", se utiliza un conjunto de 16 compuestos para comparar las respuestas fisiológicas de las células recién descongeladas y adaptadas a frascos rotativos con lotes actuales de células de producción.

Ensayos de fisiología celular

Para los ensayos de fisiología celular, puede ser conveniente utilizar 2 conjuntos de placas de 384 pocillos para medir el impacto de compuestos en diez o más parámetros de respuesta celular. Para ambos conjuntos de placas, las diluciones de compuesto se depositan primero en pocillos, y entonces se añaden 1 X $10^6$ células de ensayo a cada pocillo, como se muestra en la etapa 2602. Los compuestos se operan habitualmente con conjuntos duplicados de dilución de compuesto en la misma placa; al comienzo del estudio para un cliente individual, las placas duplicadas se ejecutan para los primeros de 16 a 32 compuestos, para medir la reproducibilidad de las respuestas. Después de mezclar de manera exhaustiva, las placas se sellan (utilizando un sello permeable de $O_2/CO_2$) y se colocan en una incubadora de cultivo tisular de 37 grados centígrados para periodos de tiempo variables (típicamente 4 horas), como se muestra en la etapa 2604. Las placas se centrifugan entonces en la etapa 2606, la mitad del líquido sobrenadante se retira, y se reemplaza por el mismo volumen de la mezcla de coloración apropiada (para la placa A, la mezcla de coloración puede incluir Monobromobimano, Calceína AM, MitoSox™ y Sytox Red™; para la placa B, la mezcla de coloración puede incluir Vybrant Violet™ (ciclo celular en vivo), JC-9 (potencial de la membrana mitocondrial)} y Sytox Red™), seguido de mezclado, como se muestra en la etapa 2608. Las placas se devuelven a la incubadora de cultivo tisular durante aproximadamente 15 (placa A) o 30 (placa B) minutos en la etapa 2610, seguido por una mezcla, y se opera inmediatamente en el citómetro de flujo en la etapa 2612.

Los datos de los pocillos de control positivos y negativos en cada fila se utilizan para calcular las respuestas como se describe con mayor detalle en este documento. Los compuestos de control positivos utilizados para la placa A y B son diferentes, y se diseñaron para proporcionar una "firma" única ("huella ") en las respuestas celulares medidas en la placa A o B, utilizando los aspectos descritos en la descripción.

Citometría de flujo de alto rendimiento

En una variedad de ensayos, el citómetro de flujo se configura utilizando un procedimiento estándar cada día que las placas se someten a ensayo. La configuración incluye garantía y control de calidad del instrumental de flujo utilizando perlas fluorescentes que se utilizan para configurar cada detector (PMT) a una diana estandarizada. Cada pocillo de una placa de 384 pocillos se muestrea secuencialmente entonces utilizando un tiempo de captación de 6 segundos, seguido de una burbuja de aire de un segundo entre muestras. El flujo de la muestra fluye a través del citómetro de flujo de una forma continua, muestreando una placa completa de 20 a 30 minutos (placas A y B, respectivamente).

El archivo de datos de la citometría de flujo (para una placa) se procesa subsecuentemente por medio de los aspectos descritos para identificar datos de pocillos individuales (utilizando alguna interacción e intervención humana), y se almacena entonces en un servidor como los datos de modo lista (LMD) para cada pocillo de una única placa.

Análisis de garantía y control de calidad de cada placa

Ambas placas (A y B) contienen controles negativos (muestras no tratadas) y controles positivos (muestras tratadas con una mezcla de 50 uM de FCCP más 50 uM de Mixatiazol para la placa A, o 25 uM de FCCP para la placa B). En este ensayo, la disimilitud entre los controles positivos y los controles negativos no define el posible intervalo de respuestas. Sin embargo, define una unidad de respuesta. Durante el periodo de tiempo del análisis de una placa entera, la disimilitud entre controles positivos y negativos puede cambiar debido al deterioro de las condiciones fisiológicas en la placa (cambio de temperatura, $O_2$, etc.). Es por esto que se espera un cierto nivel mínimo de disimilitud para cada par de controles. Para cada control positivo y negativo en una fila única, los aspectos descritos determinan la distancia QF entre las poblaciones positivas y negativas para cada respuesta de coloración individualmente. Los aspectos descritos representan gráficamente entonces el cambio en la distancia QF desde el principio (fila A) al final de la placa (fila P).

En términos generales, los experimentos descritos en este documento se han llevado a cabo según los procedimientos anteriores.

Instrumentación del citómetro

Los instrumentos de la citometría de flujo actual están equipados con múltiples láseres y múltiples detectores de fluorescencia separados que pueden cuantificar simultáneamente muchas señales de fluorescencia más características ópticas intrínsecas originadas a partir de las células individuales. Por consiguiente, las técnicas e instrumentos citométricos tales como aquellos descritos de manera ilustrativa más adelante permiten medir de miles a millones de células en una muestra. Los conjuntos de datos resultantes extremadamente grandes presentan un reto significativo para los métodos de visualización y procesamiento de datos de citometría empleados actualmente. Los métodos descritos en este documento gestionan eficazmente estos retos.

Los citómetros modernos se diseñan típicamente para detectar simultáneamente varias señales diferentes de una muestra. Una variedad de citómetros están disponibles comercialmente que se pueden utilizar según los métodos descritos en este documento. Los instrumentos típicos incluyen una célula de flujo, uno o más láseres que iluminan las células de flujo a través de una lente de enfoque, un detector o luz que pasa a través de la célula de flujo, un detector para luz dispersada hacia adelante, varias disposiciones de espejo - detector dicroicas para medir la luz de longitudes de onda específicas, típicamente para detectar fluorescencia. Una amplia variedad de otra instrumentación se incorpora con frecuencia en instrumentos comerciales.

En operaciones típicas, él láser (o láseres) ilumina la célula de flujo y las células (u otra muestra) que fluye a través de ella. El volumen iluminado por el láser se denomina como el punto de interrogación. Las células de flujo se fabrican de vidrio, cuarzo y plástico, al igual que de otros materiales. Aunque los láseres son la fuente más común de luz en los citómetros, otras fuentes de luz se pueden utilizar también. Casi todos los citómetros pueden detectar y medir una variedad de parámetros de luz dispersada hacia adelante y dispersada hacia atrás, y también varias longitudes de onda de emisión de fluorescencia. Los detectores en estos instrumentos son bastante sensibles y cuantifican fácilmente la dispersión de la luz y la fluorescencia a partir de células individuales en periodos de tiempo muy cortos. Típicamente, las señales de los detectores se digitalizan y se analizan por medio de métodos computacionales para determinar una amplia variedad de propiedades de la muestra. Existen muchos textos disponibles sobre métodos de citometría de flujo que se pueden utilizar según varios aspectos de la descripción descrita en este documento. Una referencia útil a este respecto es Practical Flow Cytometry, 4ª Edición, Howard M. Shapiro, Wiley, Nueva York (2003) ISBN: 978-0-471-41125-3.

Separación espectral de señales de citometría de flujo

Ya que las señales emitidas por las etiquetas de fluorescencia funcionales se miden por medio de una serie de detectores en un sistema de citometría (basado en flujo o imagen), los sistemas de detección son propensos a interferencia espectral. Como resultado, las intensidades de fluorocromos individuales no se pueden medir directamente a la exclusión de otros fluorocromos. Para minimizar o eliminar el ruido debido a interferencia espectral, todas las señales recolectadas se pueden modelar o procesar como mezclas lineales. La mezcla de señales para cada célula medida se descompone en aproximaciones de intensidades de señal individuales al encontrar una desviación mínima entre los resultados medidos y las composiciones aproximadas que se forman al multiplicar el estimador de la señal separada con la matriz de mezcla. La matriz de mezcla (también llamada "matriz de desbordamiento") describe la aproximación de la banda n de los espectros de fluorescencia de las etiquetas individuales (donde *n* es el número de detectores empleados en el sistema). Una aplicación de un algoritmo de minimización permite encontrar la mejor estimación de composición de la señal. Esta estimación proporciona información sobre la abundancia de diferentes etiquetas. En el caso más simple, si se asume que el error de medición es gaussiano, el proceso de separación se puede realizar utilizando minimización de mínimos cuadrados ordinarios (MCO).

Estabilización de la varianza

La estabilización de la varianza (VS) es un proceso diseñado para simplificar el análisis de datos exploratorio o para permitir el uso de técnicas de análisis de datos que hacen suposiciones sobre la homocedasticidad de los datos para conjuntos de datos heterocedásticos más complejos y, con frecuencia, ruidosos (es decir, las variables aleatorias en la secuencia tienen varianza finita diferente). Habitualmente, se ha aplicado ampliamente la VS a varios sistemas de medición biológicos basados en fluorescencia. Es una herramienta importante para el análisis de micromatrices.

En el contexto de la citometría de flujo y en el análisis de micromatrices, se ha utilizado tradicionalmente la transformación log. Sin embargo, se conocen planteamientos modernos, por ejemplo, en el contexto del análisis de micromatrices. Véase, por ejemplo, Rocke *et al*. (Approximate variance-stabilizing transformations for gene-expression microarray data." Bioinformatics, 19,966-972,2003) y Huber *et al*. ("Variance stabilization applied to microarray data calibration and to the quantification of differential expression." Bioinformatics, 18, S96-S104, 2002). Huber describe el uso de una función arcoseno hiperbólica en la estabilización de la varianza. En el contexto del análisis de datos de citometría de flujo, Moore *et al*. ("Automatic clustering of flow cytometry data with density-based merging," Adv Bioinformatics, 2009) utiliza transformación lógica. Bagwell ("Hyperlog-a flexible log-like transform for negative, zero, and positive valued data." Cytometry A. 64(1):34-42, 2005) describe el uso de transformación hyperlog en el análisis de resultados de los citómetros de flujo.

En un aspecto de la descripción, en contraste, se utiliza la técnica de arcoseno hiperbólico (logaritmo generalizado) con un parámetro encontrado empíricamente en la estabilización de la varianza.

$$\mathrm{glog}(x) = \mathrm{asinh}(x) = \log\left(x + \sqrt{x^2 + 1}\right)$$
$$\mathrm{glog}(x, a, b, c) = \mathrm{asinh}\left(a \cdot x + b\right) + c$$

Comparaciones

Ciertos aspectos de la descripción descrita en este documento proporcionan métodos que implican una etapa de comparación, en donde la distribución de las intensidades de las señales separadas se compara con la distribución de las señales separadas originadas a partir de controles u otros datos de prueba. Dependiendo del método de comparación aplicado, las distribuciones se pueden normalizar primero al dividir cada distribución por su integral.

La etapa de comparación puede implicar compilación de tensores de respuesta que contienen información sobre disimilitudes entre poblaciones de células tales como antes y después de tratamiento. Las disimilitudes se calculan como distancias entre distribuciones de señales de la población tratada de células, poblaciones no tratadas (controles "negativos" o "sin efecto"), y poblaciones tratadas con una mezcla de agentes perturbadores diseñados para maximizar la respuesta fisiológica observable (controles "positivos" o "de efecto máximo").

Para estandarizar el resultado y hacer que no se vea afectado por la variabilidad experimental, la disimilitud medida se puede expresar en unidades iguales a disimilitud media entre controles positivos y negativos.

Se pueden aplicar varias medidas de disimilitud o distancia, que incluyen (pero no se limitan a): métrica de Wasserstein, distancia de forma cuadrática (QFD), distancia de chi cuadrado, métrica de Kolmogorov, divergencia de Kullback-Leibler (simetrizada), etc. En la implementación preferida, los métodos y algoritmos de aspectos de la descripción utilizan la métrica de Wasserstein o la distancia de chi cuadrado.

En los métodos ilustrativos de más adelante, las distribuciones de la abundancia se comparan típicamente en una dimensión. Sin embargo, algunas etiquetas se codifican por medio de dos señales relacionadas (por ejemplo, JC-1, la etiqueta del potencial de la membrana mitocondrial que emite fluorescencia en dos canales separados). En este caso, se calcula una medida de disimilitud 2-D entre distribuciones. Finalmente, puede ser preferible para calcular medidas de disimilitud 2-D o 3-D utilizar distribuciones multidimensionales basadas en mediciones relacionadas con la morfología (obtenidas por medio de dispersión de la luz) y una abundancia (calculada a partir de la señal de fluorescencia). Se puede utilizar una variedad de distancias o medidas de disimilitud, asumiendo que se pueden generalizar fácilmente para múltiples dimensiones. Por ejemplo, los métodos habituales basados en la métrica de Wasserstein o la QFD se pueden utilizar en este contexto, pero no la métrica de Kolmogorov.

Una ecuación representativa para comparar poblaciones se proporciona más adelante (véase Pele *et al*. "The Quadratic-Chi Histogram Distance Family." Computer Vision -ECCV 2010, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 749-762; Daniilidis, K., Maragos, P., Paragios, N. (Eds.)):

$$QF(\mathbf{x}, \mathbf{y}) = \sqrt{\max\left[\left(\mathbf{d}\mathbf{A}\mathbf{d}^T\right), 0\right]}$$

$$\mathbf{d} = \frac{(\mathbf{x} - \mathbf{y})}{\mathbf{z}}$$

$$z = ((x + y)A)^n$$

$$A[i,j] = 1 - \sqrt{\frac{(i-j)^2}{d_{max}}}$$

en donde x e y son distribuciones de interés, y A es una matriz de disimilitud semidefinitiva positiva.

Análisis de datos de citometría utilizando tensores

Los datos multiparamétricos citométricos se pueden expresar como tensores y las comparaciones entre los controles y las muestras sometidas a prueba se pueden describir por medio de tensores de huella del compuesto. Un tensor es una matriz multidimensional y se puede considerar como una generalización de una matriz. Un tensor de primer orden (o unidireccional) es un vector, un tensor de segundo orden (bidireccional) es una matriz. Los tensores de tercer orden (tridireccionales) o superiores se llaman tensores de orden superior.

Las mediciones biológicas realizadas individualmente en un sistema de célula individual para cada célula en una población forman una distribución. Una distancia entre una distribución de mediciones realizadas en células expuestas a una presencia de un compuesto y una distribución de mediciones realizadas en células no expuestas al compuesto se pueden expresar por un único número (valor escalar). Las células se pueden exponer a un número de diferentes concentraciones de fármaco, y una medición biológica se puede realizar para cada uno de estos niveles de exposición. Un experimento de este tipo produce una serie de valores que se pueden expresar como un vector (por ejemplo, un tensor unidireccional). Si se miden múltiples parámetros biológicos, los resultados se pueden disponer en un tensor bidireccional (o una matriz), en el que cada columna contiene un parámetro medido diferente y cada fila describe una concentración diferente del compuesto.

Esta disposición de los datos se puede ampliar aún más. Si se intenta medir las distancias entre las distribuciones de las mediciones obtenidas de células tratadas y una distribución de las mediciones recolectadas de la población de células expuestas a otro compuesto, los resultados se pueden agrupar en otra matriz. Por ejemplo, puede ser beneficioso medir la disimilitud entre células tratadas con un compuesto y otro grupo de células tratadas con un compuesto bien caracterizado y diferente que crea un efecto fácil de observar que sirve como control positivo.

Las dos matrices (tensores bidireccionales) juntas producen un tensor tridireccional. La dimensionalidad de este tensor tridireccional es $I_1 \times I_2 \times I_3$, donde $I_1$ es el número de concentraciones, $I_2$ es el número de parámetros biológicos medidos, y $I_3$ es el número de disimilitudes/distancias medidas (típicamente dos: medición de distancia de control positivo y una medición de distancia de control negativo). Por lo tanto, para un tensor A que representa la medición biológica, el elemento $(i,j,k)$, denotado por $a_{i,j,k}$ describe una distancia entre mediciones del parámetro $j$ obtenido de una población de células expuestas a un compuesto a una concentración $i$, y una población de células de control $k$.

Una fibra de la columna del tensor **A**, denotada como $a_{:jk}$ contiene una serie completa de distancias entre mediciones del parámetro $j$ realizadas para una serie de muestras sometidas a prueba y un control $k$. Si se someten a prueba 10 concentraciones de un compuesto, la fibra de la columna $\mathbf{a}_{:jk}$ será un vector de 10 elementos. Un corte frontal del tensor **A**, denotado como $\mathbf{A}_{::k}$, forma una matriz que describe las distancias medidas a control $k$. Un corte lateral $\mathbf{A}_{:j:}$ es una matriz que muestra mediciones de distancias para el parámetro $j$.

El tensor **A** se puede ampliar aún más para tener en cuenta múltiples entornos, cultivos celulares o fases del ciclo celular. Por lo tanto, múltiples repeticiones de las mediciones, múltiples fases del ciclo celular en el que se realizan las mediciones, y etcétera, se pueden almacenar en un tensor de huella del compuesto. De hecho, cualquier experimento de cribado multidimensional se puede representar como tensor de orden $p^{th}$

$$\mathbf{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_p}$$

Por ejemplo, se pueden agrupar los datos de múltiples experimentos en un tensor cuadridireccional con una dimensionalidad $I_1 \times I_2 \times I_3 \times I_4$, donde $I_4$ es el número de compuestos únicos medidos. En esta configuración, un corte $\mathbf{A}_{:jkl}$ contiene disimilitudes entre muestras sometidas a prueba que contienen el compuesto $l$ y un control $k$, calculados para el parámetro $j$. Si se miden cuatro repeticiones de un compuesto $l$, el corte $A_{:jkl}$ resultante es una matriz de diez filas (concentraciones) y cuatro columnas (repeticiones).

La representación del tensor de las mediciones del compuesto se puede utilizar para definir un número de métricas y operadores relacionados con el compuesto, tales como similitud/disimilitud de un compuesto, huella del compuesto,

huella del compuesto comprimida, normalización del compuesto, etc. La disposición de los datos de cribado fenotípico en un formato de tensor posibilita la formulación de visiones únicas con respecto a las características del compuesto a través del análisis de las similitudes de respuesta del compuesto. Esto es imposible cuando la información con respecto a los compuestos se representa de manera simple como vectores de datos. Específicamente, las técnicas descritas más adelante serían imposibles de implementar si los compuestos se describieran utilizando solo valores escalares (tales como el valor $IC_{50}$ utilizado tradicionalmente).

El análisis anterior se puede exponer en términos generales en forma de la siguiente ecuación y operaciones referidas en este documento como

Análisis general del tensor de datos de la población

Método general y ecuaciones - I

En la primera etapa, un par de distancias entre los controles positivos ($C_p$) y el control negativo ($C_n$) se calcula para cada histograma marginal $\psi$. Estas distancias se mantienen como referencias para uso posterior y reescalado:

$$f_{np}^{(\psi)} = D\left(C_n^{(\psi)}, C_p^{(\psi)}\right) donde\ \psi \in \{1, \dots, p\}$$

La función de distancia D puede ser una distancia de forma cuadrática, una distancia de Wasserstein, una distancia cuadrática $\chi^2$ o cualquier otra distancia que opere en vectores que representan histogramas.

$$f_j = D\left(C_k^{(\psi)}, C_l^{(\psi)}\right)$$

Después de esta operación, una serie de distancias para las muestras biológicas se calcula de forma análoga. Las distancias se calculan para cada par hecho de un control $k$ y una muestra biológica en la serie de concentración ($S_1$, $S_2$, ..., $S_i$), donde $i$ denota la concentración de un compuesto sometido a prueba.

$$d_{k,i}^{(\psi)} = D\left(C_k^{(\psi)}, S_i^{(\psi)}\right)$$

Los valores resultantes forman una matriz multidimensional o un tensor de orden $p^{th}$.

$$\mathbf{A} \in \mathbb{R}^{I_{(\psi)1} \times I_{(\psi)2} \times \cdots \times I_{(\psi)p}}.$$

El tamaño y la dimensionalidad del tensor dependen del número de controles, del número de histogramas unidimensionales utilizado, y del número de condiciones biológicas en las que se llevó a cabo las mediciones.

Una fibra de la columna del tensor **A**, denotada como $a_{[k,\psi]}$, es un vector que contiene una serie de distancias entre mediciones de un parámetro biológico resumido por el histograma marginal $\psi$ realizadas para una serie de muestras sometidas a prueba y un control $k$.

$$a_{[k,\psi]} = \begin{bmatrix} d_{k,1}^{(\psi)} = D\left(C_k^{(\psi)}, S_1^{(\psi)}\right) \\ d_{k,2}^{(\psi)} = D\left(C_k^{(\psi)}, S_2^{(\psi)}\right) \\ \vdots \\ d_{k,1}^{(\psi)} = D\left(C_k^{(\psi)}, S_n^{(\psi)}\right) \end{bmatrix}$$

Un corte frontal del tensor **A**, denotado como $\mathbf{a}_{[k,\psi]}$, forma una matriz que describe las distancias para el parámetro biológico $\psi$, donde la medición implica todos los controles.

$$\mathbf{A}_{[\psi]} = \left[\begin{bmatrix} d^{(\psi)}_{\kappa_1,1} = D\left(C^{(\psi)}_{\kappa_1}, S^{(\psi)}_1\right) \\ d^{(\psi)}_{\kappa_1,2} = D\left(C^{(\psi)}_{\kappa_1}, S^{(\psi)}_2\right) \\ \vdots \\ d^{(\psi)}_{\kappa_1,n} = D\left(C^{(\psi)}_{\kappa_1}, S^{(\psi)}_n\right) \end{bmatrix}, \begin{bmatrix} d^{(\psi)}_{\kappa_2,1} = D\left(C^{(\psi)}_{\kappa_2}, S^{(\psi)}_1\right) \\ d^{(\psi)}_{\kappa_2,2} = D\left(C^{(\psi)}_{\kappa_2}, S^{(\psi)}_2\right) \\ \vdots \\ d^{(\psi)}_{\kappa_2,n} = D\left(C^{(\psi)}_{\kappa_2}, S^{(\psi)}_n\right) \end{bmatrix}, \cdots, \begin{bmatrix} d^{(\psi)}_{\kappa_m,1} = D\left(C^{(\psi)}_{\kappa_m}, S^{(\psi)}_1\right) \\ d^{(\psi)}_{\kappa_m,2} = D\left(C^{(\psi)}_{\kappa_m}, S^{(\psi)}_2\right) \\ \vdots \\ d^{(\psi)}_{\kappa_m,n} = D\left(C^{(\psi)}_{\kappa_m}, S^{(\psi)}_n\right) \end{bmatrix}\right]$$

Un corte lateral $\mathbf{A}_{[\psi]}$ es una matriz que muestra las mediciones de distancias para múltiples parámetros biológicos, donde la medición se llevó a cabo utilizando solo un único control *k*.

$$\mathbf{A}_{[\kappa]} = \begin{bmatrix} \left[d^{(\psi_1)}_{\kappa,1} = D\left(C^{(\psi_1)}_{\kappa}, S^{(\psi_1)}_1\right), d^{(\psi_2)}_{\kappa,2} = D\left(C^{(\psi_2)}_{\kappa}, S^{(\psi_2)}_1\right), \cdots, d^{(\psi_p)}_{\kappa,1} = D\left(C^{(\psi_p)}_{\kappa_1}, S^{(\psi_p)}_1\right)\right] \\ \left[d^{(\psi_1)}_{\kappa,1} = D\left(C^{(\psi_1)}_{\kappa}, S^{(\psi_1)}_2\right), d^{(\psi_2)}_{\kappa,1} = D\left(C^{(\psi_2)}_{\kappa}, S^{(\psi_2)}_2\right), \cdots, d^{(\psi_p)}_{\kappa,1} = D\left(C^{(\psi_p)}_{\kappa_1}, S^{(\psi_p)}_2\right)\right] \\ \vdots \\ \left[d^{(\psi_1)}_{\kappa,1} = D\left(C^{(\psi_1)}_{\kappa}, S^{(\psi_1)}_n\right), d^{(\psi_2)}_{\kappa,2} = D\left(C^{(\psi_2)}_{\kappa}, S^{(\psi_2)}_n\right), \cdots, d^{(\psi_p)}_{\kappa,n} = D\left(C^{(\psi_p)}_{\kappa_1}, S^{(\psi_p)}_n\right)\right] \end{bmatrix}$$

Por lo tanto, el tensor **A** se puede expresar como:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_{[\kappa_1,\psi_1]} & \cdots & \mathbf{a}_{[\kappa_1,\psi_p]} \\ \vdots & \ddots & \vdots \\ \mathbf{a}_{[\kappa_m,\psi_1]} & \cdots & \mathbf{a}_{[\kappa_m,\psi_p]} \end{bmatrix}$$

Las mediciones biológicas se pueden normalizar además al dividir cada $d^{(\psi)}_{\kappa,i}$ por el correspondiente $f^{(\psi)}_{n\psi}$ que da lugar a un tensor normalizado de distancias:

$$\tilde{\mathbf{A}} = \begin{bmatrix} \tilde{\mathbf{a}}_{[\kappa_1,\psi_1]} & \cdots & \tilde{\mathbf{a}}_{[\kappa_1,\psi_p]} \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{a}}_{[\kappa_m,\psi_1]} & \cdots & \tilde{\mathbf{a}}_{[\kappa_m,\psi_n]} \end{bmatrix}$$

El mismo análisis se puede llevar a cabo para una variedad de condiciones, lo que da lugar a una serie de tensores normalizados de distancias.

Compresión

Un tensor **A** obtenido de una serie de mediciones forma una huella del compuesto única, ya que contiene todas las características fenotípicas de un compuesto sometido a prueba. Este tensor A se puede "comprimir" utilizando técnicas de aproximación de tensor de rango bajo tales como descomposición de tensor poliádica u otros métodos. Véase Kolda y Bader, "Tensor decompositions and applications," SIAM Rev. 51,455,2009.

La descomposición del tensor factoriza un tensor en una suma de tensores de componentes de rango uno. Un tensor

$$\mathbf{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$$

por definición es de rango uno si se puede expresar como el producto exterior de los vectores *p*:

$$\mathbf{A} = \mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \cdots \circ \mathbf{a}^{(p)}$$

El objetivo de una descomposición poliádica canónica (CP) es encontrar una aproximación del tensor A, denotada como **Â**, que satisfaga los siguientes criterios:

$$\min_{\hat{\mathbf{A}}} \left\| \mathbf{A} - \hat{\mathbf{A}} \right\|, \quad \text{donde} \quad \hat{\mathbf{A}} = \left[ \lambda; \mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots, \mathbf{A}^{(n)} \right] = \sum_{r=1}^{R} \lambda_r \mathbf{a}_r^{(1)} \circ \mathbf{a}_r^{(2)} \circ \cdots \circ \mathbf{a}_r^{(p)}$$

Los tensores se pueden descomponer también utilizando otras técnicas. Por ejemplo, la descomposición de Tucker descompone un tensor en un tensor núcleo multiplicado por una matriz a lo largo de cada modo:

$$\hat{\mathbf{A}} = \left[ \mathbf{G}; \mathbf{M}^{(1)}, \mathbf{M}^{(2)}, \ldots, \mathbf{M}^{(p)} \right] = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_p=1}^{R_p} g_{r_1 r_2 \ldots r_p} \mathbf{a}_{r_1}^{(1)} \circ \mathbf{a}_{r_2}^{(2)} \circ \cdots \circ \mathbf{a}_{r_N}^{(p)}$$

Los diversos métodos para descomponer un tensor permiten aproximar una huella de compuesto multidimensional complicada a una versión abreviada, "comprimida". Este objetivo se puede conseguir al retirar la parte de una descomposición que no contribuye significativamente a la precisión de reconstrucción de un tensor.

El ejemplo de más adelante ilustra este punto. Se asume que **A** es un tensor tridireccional que contiene la información de un compuesto. Los cortes frontales del tensor **A**, denotados como $\mathbf{A}_{::1}$ y $\mathbf{A}_{::2}$, son:

| | [,1] | [,2] | [,3] |
|---|---|---|---|
| [1,] | 9.02 | 17.69 | 27.87 |
| [2,] | 18.23 | 36.13 | 53.95 |
| [3,] | 27.16 | 53.81 | 81.07 |
| [4,] | 35.82 | 72.05 | 108.15 |
| [5,] | 44.85 | 89.85 | 134.69 |

Corte $\mathbf{A}_{::1}$

| | [,1] | [,2] | [,3] |
|---|---|---|---|
| [1,] | 9.98 | 20.25 | 29.92 |
| [2,] | 19.92 | 40.12 | 59.78 |
| [3,] | 29.73 | 59.67 | 90.24 |
| [4,] | 39.61 | 80.16 | 119.79 |
| [5,] | 49.58 | 99.94 | 149.88 |

Corte $\mathbf{A}_{::2}$

Después de realizar la descomposición CP, se obtienen los siguientes vectores:

$$\mathbf{a}_1^{(1)} = \left[-0.1348, -0.2698, -0.4048, -0.5398, -0.6737\right]$$

$$\mathbf{a}_2^{(1)} = \left[0.2366, 0.4133, -0.7798, 0.4011, -0.0657\right]$$

$$\mathbf{a}_1^{(2)} = \left[-0.2663, -0.5347, -0.802\right]$$

$$\mathbf{a}_2^{(2)} = \left[0.1966, -0.8447, 0.4979\right]$$

$$\mathbf{a}_1^{(3)} = \left[0.6691, 0.7431\right]$$

$$\mathbf{a}_2^{(3)} = \left[-0.2009, -0.9796\right]$$

$$\lambda_1 = 372.9803, \lambda_2 = 0.5591139$$

El tensor original se puede reconstruir a partir de este resultado siguiendo el modelo CP:

$$\hat{\mathbf{A}} = \lambda_1 \mathbf{a}_1^{(1)} \circ \mathbf{a}_1^{(2)} \mathbf{a}_1^{(3)} + \lambda_2 \mathbf{a}_2^{(1)} \circ \mathbf{a}_2^{(2)} \mathbf{a}_2^{(3)}$$

Después de la reconstrucción, el tensor aproximado **Â** contiene los siguientes valores:

| | [,1] | [,2] | [,3] |
|---|---|---|---|
| [1,] | 8.954503 | 18.01323 | 26.98964 |
| [2,] | 17.924563 | 36.04534 | 53.98549 |
| [3,] | 26.920336 | 53.94805 | 81.06428 |
| [4,] | 35.865694 | 72.06479 | 108.01836 |
| [5,] | 44.777858 | 89.69133 | 134.84858 |

Corte $\hat{\mathbf{A}}_{::1}$

| | [,1] | [,2] | [,3] |
|---|---|---|---|
| [1,] | 9.924882 | 20.08719 | 29.90171 |
| [2,] | 19.872031 | 40.17834 | 59.86737 |
| [3,] | 29.961619 | 59.62578 | 90.19134 |
| [4,] | 39.797839 | 80.17582 | 119.87475 |
| [5,] | 49.733276 | 99.80678 | 149.77195 |

Corte $\hat{\mathbf{A}}_{::2}$

Sin embargo, el tensor **A** se puede reconstruir utilizando solo vectores $a_1$ en vez de $a_1$ y $a_2$. Aunque la aproximación resultante tiene un error mayor, aún captura la mayor parte de la información representada por A:

$$\hat{\mathbf{A}} = \lambda_1 \mathbf{a}_1^{(1)} \circ \mathbf{a}_1^{(2)} \circ \mathbf{a}_1^{(3)}$$

| | [,1] | [,2] | [,3] |
|---|---|---|---|
| [1,] | 8.95973 | 17.98877 | 26.98267 |
| [2,] | 17.93369 | 35.88512 | 54.00360 |
| [3,] | 26.90311 | 54.01435 | 81.02666 |
| [4,] | 35.87455 | 72.02604 | 108.03888 |
| [5,] | 44.77559 | 89.89757 | 134.84490 |

Corte $\hat{\mathbf{A}}_{::1}$

| | [,1] | [,2] | [,3] |
|---|---|---|---|
| [1,] | 9.998363 | 19.97770 | 29.96624 |
| [2,] | 19.916531 | 39.98714 | 59.98007 |
| [3,] | 29.877697 | 59.98645 | 89.97872 |
| [4,] | 39.841823 | 79.96027 | 119.94411 |
| [5,] | 49.726282 | 99.81718 | 149.75464 |

Corte $\hat{\mathbf{A}}_{::2}$

Como se demostró anteriormente, el tensor **A** se puede aproximar utilizando solo vectores $a_1$. Si la precisión resultante es suficiente, la descomposición comprime eficazmente el tensor **A** aproximadamente 6 veces —se necesitan 60 números para codificar el tensor original pero solo 11 números para codificar el tensor comprimido $\hat{\mathbf{A}}$ (vectores $\mathbf{a}^{(1)}$, $\mathbf{a}^{(2)}$, $\mathbf{a}^{(3)}$ y valor de λ). La forma descompuesta del tensor se puede utilizar además para la comparación tensor frente a tensor, al igual que una entrada para métodos de aprendizaje automático supervisado, tales como, por ejemplo, para respaldar el aprendizaje automático de vectores.

Se debería tener en cuenta que la descomposición CP simple y la descomposición de Tucker más compleja, aunque la utilizada más comúnmente para expresar el contenido de una información dispuesta en un tensor de una forma más simple y fácil de comprender, no son los únicos métodos útiles para descomponer tensores. Se pueden utilizar también otros métodos de descomposición de tensores con los datos de cribado, mientras la organización de los conjuntos de datos siga una convención aceptable.

Además, como se mencionó anteriormente, la descomposición de tensores se puede aplicar a tensores de orden superior. En este caso, las múltiples mediciones realizadas para múltiples compuestos se descompondrán a lo largo de los modos de un tensor, revelando las diferencias cruciales entre compuestos. También, se pueden utilizar múltiples repeticiones del mismo compuesto sometido a prueba, lo que lleva eficazmente al cálculo de una huella de compuesto "promediada" sintética que se puede comparar entonces con otras huellas.

Gating automatizado basado en la morfología

Un aspecto de la descripción proporciona el uso de un *gating* automático guiado por modelos (aunque el uso de algoritmos de *gating* es opcional). En este documento, las técnicas del estado de la técnica de modelado de mezclas con o sin adiciones patentadas se pueden añadir al algoritmo. El sistema puede depender de un planteamiento iterativo para mejorar la eficacia del ensayo.

En un aspecto de la descripción, la técnica de *gating* comprende 3 distribuciones de probabilidad de asimetría normal que representan "células vivas", "células moribundas" y "células muertas" (detritus). Dependiendo de los datos, se puede utilizar un modelo existente (por ejemplo, validado antiguo) o uno nuevo generado basado en los controles. Por ejemplo, es posible proceder a calcular la verosimilitud log (LL) para cada modelo de mezcla. Los modelos específicos para los que LL es superior se retienen entonces para uso futuro.

Disimilitud del compuesto

Las mediciones de disimilitud del compuesto se pueden realizar después de la descomposición de una huella de compuesto individual, después de una descomposición de un tensor de resultado en el que todos los compuestos (y todas las repeticiones de mediciones) se almacenan, o se utiliza directamente la versión completa, "no comprimida" de las huellas. Si la comparación implica tensores descompuestos, el problema se reduce esencialmente a comparar vectores.

Sin embargo, como se mencionó anteriormente, los tensores de huella de compuesto se pueden comparar también directamente. En este caso, la disimilitud entre tensores de compuesto se expresa como un vector que almacena las disimilitudes (distancias) entre las fibras de modo 1 $a_{:jk}$ del tensor **A** de compuesto. Ya que las fibras $a_{:jk}$ almacenan información sobre las respuestas de cultivos celulares expuestos a concentraciones crecientes de los compuestos sometidos a prueba, los valores forman una serie, que se puede ver como una curva poligonal. En consecuencia, la disposición orientada al tensor de datos permite utilizar distancias entre curvas discretas (vectores).

Una de las distancias de este tipo es la distancia de Fréchet discreta, que se define como la longitud mínima de una correa necesaria para que el perro y el guía se muevan desde los puntos de partida de las dos curvas que describen sus

caminos hasta los puntos finales, proporcionada de modo que tanto el perro como el guía solo puedan moverse hacia adelante, aunque pueden controlar su velocidad. Ya que las fibras se pueden ver como curvas poligonales, en esta configuración tanto el perro como el guía solo se pueden parar en los vértices de las curvas y cada uno de ellos puede tanto permanecer en su vértice actual como saltar al siguiente mientras camina. Véase Aronov *et al*., "Fréchet Distance for Curves," Revisited, in: Azar, Y., Erlebach, T. (Eds.), Algorithms - ESA 2006, Lecture Notes in Computer Science. Springer Berlín Heidelberg, pp. 52-63, 2006.

Una medida de distancia conceptualmente similar es una distancia de alineamiento temporal no lineal dinámica, que mide la distancia entre dos curvas poligonales al deformarlas de una manera no lineal para encontrar el mejor alineamiento. El coste de realizar el alineamiento se define como el valor de distancia. Véase Berndt y Clifford, "Using dynamic time warping to find patterns in time series." AAAI94 Workshop on Knowledge Discovery in Databases, pp. 359-370,1994.

Por lo tanto, en un caso general, se puede expresar la disimilitud entre dos huellas de compuesto (tensores) **A** y **B** como:

$$w_{A,B} = D(\mathbf{A},\mathbf{B}) = d\left(\mathbf{a}_{:jk}, \mathbf{b}_{:jk}\right),$$

donde *D* es la disimilitud entre los tensores de compuesto **A** y **B**, *d* es la función de distancia que compara las fibras de modo 1 de cada uno de los tensores (tales como distancia de Fréchet o distancia de alineamiento temporal no lineal dinámica), y w es el vector de distancia. Para un tensor tridireccional producido por medio de experimentos que implican una curva de respuesta de 10 dosis, cuatro parámetros y controles positivos y negativos, el vector w tendría una longitud de 8. Cada elemento del vector se calcularía como una curva-distancia entre las correspondientes fibras de los tensores **A** y **B**.

Representaciones gráficas de las disimilitudes de compuestos

Las similitudes/disimilitudes entre compuestos se pueden demostrar gráficamente utilizando varios métodos que dependen de la función utilizada para evaluar la disimilitud. Un ejemplo se ilustra en la Figura 11. Si una huella abreviada de un compuesto se puede expresar como un vector después de una descomposición (formada por múltiples vectores concatenados), entonces se puede aplicar una distancia elegida *a priori* entre dos vectores, o enseñar una función de distancia utilizando métodos métricos de aprendizaje. Véase Weinberger y Saul, "Distance metric learning for large margin nearest neighbor classification," J. Mach Learn. Res. 10, 207-244, 2009. El resultado final se puede visualizar utilizando escalamiento multidimensional (EMD), agrupamiento jerárquico, o métodos de aprendizaje de variaciones.

Si no se utiliza descomposición/compresión, la visualización de similitudes entre compuestos requiere un planteamiento que utilice vectores de disimilitudes, en vez de medidas escalares de disimilitud. Un método de este tipo es el algoritmo EMD multimodal que utiliza mayorización del estrés. Véase Leeuw y Mair, "Multidimensional Scaling Using Majorization: SMACOF," R. J. Stat. Softw. 31, 1-30,2009.

Utilizar las matrices de disimilitud formadas por técnicas de escalamiento multidimensional multidireccional coloca las firmas medidas del compuesto en un espacio que define la similitud entre compuestos. Véase Groenen *et al.* ("The majorization approach to multidimensional scaling for Minkowski distances." J. Classif. 12, 3-19, 1995). Los compuestos localizados cerca en el espacio resultante se considerarán "similares" en el sentido del proceso descrito, y el compuesto ubicado lejos se considerará "disimilar".

Comparación de distribuciones utilizando múltiples subconjuntos celulares

La disimilitud de una muestra de control frente a una sometida a prueba se puede calcular empleando todos los casos (células) medidos o calculados de manera separada para cada subconjunto definido funcionalmente de la población celular. Por ejemplo, se pueden definir dos subconjuntos de células morfológicamente distintos con base en la dispersión de la luz medida. En esta configuración, los dos subconjuntos celulares se pueden comparar con los controles de manera separada. Lo que sigue es que el número de comparaciones es dependiente del número de subconjuntos formados. Por ejemplo, si cinco etiquetas funcionales se designan para demostrar cinco aspectos diferentes de la fisiología celular, y se elige para comparar el resultado tratado frente al resultado no tratado para cada una de las poblaciones de células morfológicamente diferentes, se pueden producir hasta 20 distancias diferentes (5 etiquetas x 2 controles x 2 subconjuntos celulares).

Ya que la concentración del atente perturbador resultante en el efecto fisiológico deseado no se puede determinar *a priori*, una función que describe el aumento gradual del efecto con la concentración perturbadora se puede utilizar para indicar propiedades del agente perturbador, la medición se realiza con varias concentraciones de un fármaco. Por lo tanto, para el escenario descrito anteriormente asumiendo el uso de 10 concentraciones, el conjunto final de datos comprendería 200 valores (5 etiquetas x 2 controles x 2 subpoblaciones x 10 concentraciones). Los valores medidos no son independientes; por lo tanto, se tratan para procesamiento posterior como un tensor multidireccional. Por ejemplo, si solo se utiliza un tipo de control (por ejemplo, un control "positivo"), el conjunto de datos formaría un tensor tridireccional con las dimensiones de 5 x 10 x 2. La Figura 14 muestra un ejemplo de un tensor multidireccional que representa una huella de respuesta de un fármaco.

Uso de un espacio de características multidimensionales

En un aspecto de la descripción, la evolución de una nube de puntos en un espacio de características multidimensionales se puede calcular al calcular directamente las distancias entre distribuciones. Los ejemplos de nubes de puntos que evolucionan a un espacio de características multidimensionales incluyen, pero no se limitan a, por ejemplo, variables tales como concentración de fármaco, tiempo de incubación, etc. Los cambios en la evolución de la nube en el espacio definidos por disimilitudes entre nubes de puntos complejas forman una trayectoria complicada, que describe únicamente las características del compuesto que provoca esos cambios.

Una evolución de nube representativa se muestra en la Figura 15.

Comparación de distribuciones utilizando información del ciclo celular

En aspectos de la descripción, los tensores de huella se crean para cada uno de los clústeres o subconjuntos celulares definidos por etiquetas dependientes del ciclo celular. En esta configuración, la evaluación de las fases del ciclo celular se realiza simultáneamente con otras mediciones tales como cuantificando la intensidad de etiquetas de fluorescencia que se intercalan en ADN (tales como HOECHST 33342, DRAQ5, yoduro YO-PRO-1, IODIDE, DAPI, CYTRAK ORANGE), cuantificando la presencia de proteínas inmunomarcadas asociadas al ciclo celular (por ejemplo, ciclinas, proteínas histónicas fosforiladas, etc.), o midiendo la señal de fluorocromos codificados unidos genéticamente a fases del ciclo celular. En el caso más simple, estas técnicas permiten determinar tres compartimientos del ciclo celular, lo que lleva a la formación de un tensor de resultado 5 x 10 x 3 (o tensores más complejos, dependiendo del número de etiquetas adicionales). Esto se ilustra en el Ejemplo 5A.

Clasificación supervisada y no supervisada

Los aspectos de la descripción proporcionan métodos de clasificación, en donde se realizaron análisis subsecuentes utilizando múltiples planteamientos. Estas técnicas analíticas se pueden implementar dependiendo de la representación o visualización final preferida de los resultados. Por ejemplo, en la configuración no supervisada, el tensor de distancia se puede procesar por medio de un algoritmo SMACOF (escalamiento vía mayorización de una función complicada) multidimensional y multimodal u otra técnica matemática adecuada para escalado. Después de la aplicación del método EMD, las huellas del fármaco se colocarán en un plano 2-D o en cualquier otra superficie n-dimensional (por ejemplo, esfera) definida por medio de los investigadores. Esto permite una interpretación y visualización fácil de los conjuntos de datos, que forman un "mapa" tras proyección EMD.

En el marco supervisado, se utilizan las huellas que representan compuestos que pertenecen a varios grupos definidos por función, tipo de respuesta, estructura química, etc., para crear una biblioteca de formación con múltiples clases. La biblioteca de formación se puede utilizar subsecuentemente para enseñar a un clasificador (tal como un clasificador de red neuronal, una máquina de vectores de soporte, u otro tipo de sistema de aprendizaje automático) que categorizará huellas desconocidas previamente en clases definidas por compuestos conocidos.

Por extensión, los aspectos relacionados de la descripción proporcionan métodos para evaluar los efectos de agentes perturbadores desconocidos (no conocidos *a priori* para impactar en el MMP) para determinar si tiene un impacto diferencial en el MMP (y otras mediciones de fisiología celular, por ejemplo, células bajas frente a altas de SYTOX™ RED) en diferentes fases del ciclo celular. Por lo tanto, para cada compuesto sometido a prueba, se proporciona un tensor de medición multidireccional (firma del compuesto) que comprende un tensor de estado relacionado con el ciclo celular como una de las dimensiones en las que el estado celular se evalúa.

Lo siguiente es un método representativo para comparar diferentes agentes perturbadores utilizando los sistemas y los algoritmos de aspectos de la descripción.

Primero, para cada canal de fluorescencia (por ejemplo, de manera separada para células vivas y células moribundas), se crea una matriz de distancia entre todos los compuestos utilizando la técnica descrita anteriormente. Por ejemplo, la distancia puede ser la distancia de alineamiento temporal no lineal dinámica (dtw) entre curvas de respuesta. Subsecuentemente, un escalamiento multidimensional y multidireccional (EMD) se utiliza para combinar varias matrices en un espacio euclídeo. Entonces, se puede el EMD para un grupo de matrices que contiene solo datos del MMP (CALCEÍNA, SYTOX™) y otro grupo con solo datos de permeabilidad (MBBR, MITOSOX™). Las diversas similitudes fisiológicas entre los compuestos se pueden visualizar con la ayuda de técnicas de mapeo y agrupamiento conocidas, por ejemplo, visualización de dendogramas o gráficas (por ejemplo, creando nodos intermediarios y exportando los datos a cualquier variedad de paquetes de *software* de visualización externa). Asimismo, se puede emplear también un planteamiento supervisado utilizando uno o más clasificadores.

Ciclo celular

Los aspectos de la descripción descritos en este documento permiten mediciones de expresión coordinada de proteína (u otro marcador) en poblaciones de células como función del ciclo celular (por ejemplo, G1, S, G2M), y para determinar los efectos dependientes del ciclo celular de los compuestos de prueba. El análisis multiparamétrico se puede llevar a cabo así analizando el efecto de cada agente perturbador en diferentes concentraciones y/o puntos temporales para investigar el efecto de dichos compuestos en los diversos parámetros celulares (por ejemplo, potencial de la membrana mitocondrial, permeabilidad de la membrana nuclear o citoplasmática, ROS, muerte celular o apoptosis).

Un ejemplo de análisis dependiente del ciclo celular se basa en la medición de la expresión de Ciclina A2 en células normales (no perturbadas). En este documento, los posibles "estados" incluyen Ciclina A2 negativa, Ciclina A2 baja y Ciclina A2 alta. De manera similar, para fosfo-histona 3 (H3-F), que es un segundo marcador en el análisis del ciclo celular, los posibles "estados" incluyen "negativo" y "positivo". Estos dos marcadores del ciclo celular se pueden analizar también en combinación, produciendo así nueve combinaciones ("estados") diferentes posibles. No siempre es necesario investigar todos los posibles "estados" porque todos los estados pueden no existir en un espacio biológico (matriz dispersa) normal.

Por consiguiente, dependiendo del estado del ciclo celular en el que se encuentre una célula particular, las perturbaciones diferenciales causadas por fármacos o compuestos de interés se pueden investigar al poblar células en elementos de matriz discretos (normales). Como ejemplo, los fármacos que bloquean la progresión normal de la mitosis de vuelta a la G1, que causa cambios cuantitativos en poblaciones de matrices "normales" (es decir, acumulación de células en los compartimientos del ciclo celular "tardíos" (normales) (por ejemplo, G2 y M)) y/o depletan células en la fase G1, se pueden analizar junto con tinción de Ciclina A2 y/o de H3-F. De manera similar, un fármaco que evita la separación de los núcleos hija se esperaría que mostrara un patrón de huella cuantitativo diferente en comparación con un fármaco que detiene a las células en la fase S (por ejemplo, un fármaco que inhibe nueva síntesis de ADN). Por consiguiente, los compuestos que causan que las células aparezcan en diferentes elementos de una matriz no solo crean una firma única, sino que también el elemento específico de matriz que está ocupado podría proporcionar información con respecto al mecanismo de acción del fármaco. Por ejemplo, la expresión de Ciclina A2 en la G1 y/o M puede ser el resultado de un inhibidor de proteasoma que evita la degradación de Ciclina A2 normal.

Sistema de ensayo de tipo celular múltiple

En un aspecto, la descripción proporciona métodos para someter a ensayo a estados celulares utilizando una pluralidad de tipos celulares, por ejemplo, dos o más líneas celulares (a partir de cultivo tisular) en un único ensayo. Una ventaja de este planteamiento es que permite analizar la lesión/respuestas del ADN. Una ventaja adicional es que permite estudiar las vías de señalización tanto constitutivas como inducibles en el mismo ensayo (utilizando una línea celular con expresión constitutiva y otra que puede activar la misma vía utilizando un antagonista apropiado). Utilizando dos (o más) líneas celulares simultáneamente, sería posible abarcar múltiples vías de señalización en un ensayo.

Por ejemplo, utilizando líneas celulares mieloides humanas (derivadas de pacientes con leucemia mieloide), una línea celular que responde a LPS activará las vías de NF-κB y PI3 quinasa, mientras que otra que responde a TNF-α activará múltiples vías de MAP quinasa; en ambos casos, se pueden evaluar en sentido ascendente (quinasa Iκ para NF-κB) y en sentido descendente (P-S6 para ERK y mTOR para PI3K). Además, estos ensayos pueden incluir marcadores de lesión/respuesta del ADN, como se indicó anteriormente. La línea celular que responde en mezclas de células se puede identificar utilizando tanto el contenido de ADN (algunas líneas celulares son diploides, otras son aneuploides con diferente contenido anormal de ADN), o características biológicas (marcadores de superficie celular), o células que se pueden "marcar con código de barras" (G. Nolan et al.). Finalmente, los ensayos de señalización incluyen análisis del ciclo celular (por ejemplo, contenido de ADN) para permitir la correlación de respuestas de la vía de transducción de señal con la fisiología celular en respuesta a los mismos fármacos.

**Ejemplos**

Los siguientes ejemplos se proporcionan solo a modo de ilustración por medio de varios aspectos particulares de la descripción y no son de ningún modo exhaustivos ni exclusivos.

Metodología general de los siguientes ejemplos

Cultivos tisulares

Las células de cultivo tisular humanas se obtuvieron de Sigma-Aldrich (San Luis, Misuri, EE. UU.) como existencias congeladas. Las líneas celulares se descongelaron y se colocaron inicialmente en un cultivo estático, seguido de cultivo continuo en frascos rotativos de 1 l, en RPMI 1640/10%FBS/Glut (+/- 0,5-2 mM)/Pen/Strep. Las células para ensayos de cribado *in vivo* se mantuvieron en crecimiento de fase log en condiciones altamente estandarizadas (densidad de las células de entrada y de recolecta, viabilidad, etc.).

Preparación de las placas de ensayo de fisiología celular 1a y 1b

Las células se depositaron en placas de 384 pocillos (1 X $10^6$ células en 40 µl de medio de cultivo tisular) y se depositaron diluciones en serie de compuestos de prueba (o vehículo en el caso de los controles) en pocillos individuales utilizando técnicas de robótica estándar. Se realizaron la deposición y dilución del compuesto por medio de BIOMEK FX (diluciones de 5 o 10 etapas de 100 a 0,005 µM). Las diluciones se prepararon a partir de disoluciones madre de compuestos disueltos inicialmente en DMSO al 100%. El volumen final depositado en cada pocillo es 2 µl de cada dilución de compuesto (los compuestos tienen una concentración final de DMSO al 1%). Después de incubación en una incubadora de cultivo tisular ($CO_2$ al 5% a 37 °C) para periodos de tiempo variables (típicamente 4 horas), las placas se centrifugaron a temperatura ambiente, seguido de retirada de 20 µl de líquido sobrenadante. Después de vibración de la placa para suspender de nuevo las células, una mezcla de coloración (Placa 1a Monobromobimano, CALCEÍNA AM, MITOSOX™ y SYTOSOX™ RED; placa 1b colorante de ciclo celular en vivo VYBRANVIOLET™, JC-9 y SYTOSOX™ RED) se añadió en un volumen total de 20 µl, seguido de una etapa de mezclado adicional. Las placas se devolvieron a la incubadora de cultivo tisular ($CO_2$ al 5% a 37 °C) durante 15 min (placa 1a) o 30 min (placa 1b) adicionales, seguido por análisis inmediato utilizando HYPERCYT-CYAN™. Preparación de placas de ensayo de ciclo celular o señalización celular 2 y 3

Las células se depositaron en diluciones de compuestos de prueba depositadas en pocillos individuales de placas de 96 pocillos, como se indicó anteriormente. Después de la incubación en una incubadora de cultivo tisular ($CO_2$ al 5% a 37 °C) durante de 30 min a 12 h, las células se sedimentaron por centrifugación, el sobrenadante se retiró, las células se volvieron a suspender por vibración en el líquido sobrenadante remanente, y se añadieron 50 µl de fijador (formaldehído al 1,2% en PBS). Después de la incubación de 15 min (a 37 °C?), las células se sedimentaron por centrifugación, el sobrenadante ser retiró, y las células se volvieron a suspender por vibración de la placa.

Se permeabilizó la célula al añadir 150 µl de metanol absoluto frío, seguido de incubación a 4 °C durante 15 minutos. Las células permeabilizadas se centrifugaron, el sobrenadante se retiró, y las células se lavaron tres veces con PBS frío (4 °C) que contenía FBS al 4%. Después del lavado final, las células se centrifugaron, el sobrenadante se retiró, y las células se volvieron a suspender en un cóctel de anticuerpos (volumen total 20 µl) apropiado para el experimento.

Para análisis del contenido de ADN/ciclo celular, las células se volvieron a suspender (después de lavados para retirar los conjugados de anticuerpo no unidos) en PBS/FBS que contenía DAPI (1 µg/ml), seguido de análisis utilizando HYPERCYT-CYAN™.

Citometría de flujo de alto rendimiento con HYPERCYT-CYAN

Cada pocillo de la placa de 384 pocillos se muestreó con un tiempo de captación de 6 segundos, más 1 segundo de retraso entre pocillos, lo que dio lugar a un tiempo de lectura total para cada placa de aproximadamente 45 minutos. Los datos se obtuvieron utilizando un citómetro de flujo CYAN™, con el *software* HYPERCYT™ utilizado para compilar archivos para cada pocillo en una placa individual utilizando un formato FCS. Los datos para cada placa se analizaron subsecuentemente para varios parámetros de garantía de calidad (número total de sucesos por pocillo, número de sucesos asociados con sucesos morfológicamente normales, número de células muertas), como se describe más adelante.

**Ejemplo 1 células normales**

Se llevaron a cabo análisis de citometría de flujo de propiedades celulares intrínsecas (dispersión de la luz hacia adelante y luz dispersada ortogonalmente al haz láser) y señales de fluorescencia derivadas de los colorantes enumerados en la Tabla 1 para las placas 1a y 1b.

Tabla 1. Lista de colorantes utilizados en las Placas 1a y 1b con dianas fisiológicas

| Colorante | Placa | Respuesta fisiológica |
|---|---|---|
| Monobromobimano (MBBR) | 1a | SH celular/nuclear (predominantemente glutatión (GSH)) |
| CALCEÍNA AM | 1a | Permeabilidad de la membrana citoplasmática |
| MITOSOX RED™ | 1a | Especies de oxígeno reactivo (relacionadas con función mitocondrial) |
| SYTOXRED™ | 1a y 1b | Permeabilidad de la membrana nuclear y citoplasmática (viva/muerta) |
| VYBRANVIOLET™ | 1b | Medición del ciclo celular que incluye células viables |
| JC-9 | 1b | Potencial de membrana (MMP) |
| ™Sondas moleculares/Invitrogen | | |

Los resultados se muestran en el diagrama de dispersión en la Figura 1A. Como se ve en el diagrama, la dispersión de la luz sola permite la identificación de células que tienen morfología "normal". En esta figura, las células no se tratan con ningún fármaco, y la mayoría de las células (aquí, superior al 95% en todos los sucesos) muestra morfología "normal".

Las células mostradas en esta figura se trataron también con una mezcla de colorantes de la Tabla 1 como a continuación:

(1) Monobromobimano (MBBR) para medir el contenido de glutatión intracelular (GSH), como se describe previamente en Cossarizza *et al*. ("Simultaneous analysis of reactive oxygen species and reduced glutathione content in living cells by polychromatic flow cytometry," Nat. Protoc. 4, 1790-1797, 2009); Kosower *et al*. ("Bimane fluorescent labels: labeling of normal human red cells under physiological conditions." Proc. Natl. Acad. Sci. EE. UU. 76, 3382-3386, 1979); Radkowsky *et al*. ("Bimanes. 17. (Haloalkyl)-1,5-diazabiciclo[3.3.O]octadienediones (halo-9,10-dioxabimanes): reactivity toward the tripeptide thiol, glutathione." J. Am. Chem. Soc. 108, 4527-4531, 1986);

(2) CALCEÍNA AM para medir la integridad de la membrana citoplasmática, como se describe Ivnitski-Steele *et al*. ("High-throughput flow cytometry to detect selective inhibitors of ABCB1, ABCC1, and ABCG2 transporters." Assay Drug Dev. Technol. 6, 263-276, 2008). La CALCEÍNA AM es un compuesto no fluorescente que pasa a través de la membrana citoplasmática y, en células vivas, el compuesto se convierte en un compuesto fluorescente muy verde que no puede pasar a través de la membrana citoplasmática; si subsecuentemente las células pierden integridad de la membrana citoplasmática, el compuesto fluorescente abandona la célula por difusión pasiva;

(3) MITOSOX™ RED para medir especies de oxígeno reactivo intracelulares, como se describe en Zielonka *et al*. ("Detection of 2-hydroxyethidium in cellular systems: a unique marker product of superoxide and hydroethidine." Nat. Protoc. 3, 8-21, 2008). El MITOSOX™ RED es un compuesto que se puede difundir en células viables y reacciona con especies de aniones superóxidos, particularmente productos de la respiración mitocondrial; las condiciones tales como estrés celular pueden causar que la mitocondria responda (a veces rápidamente) aumentando la respiración y/o liberando aniones superóxidos; y

(3) SYTOXRED™ para medir la viabilidad celular (pérdida de integridad de la membrana nuclear). El SYTOXRED™ es un colorante de ADN de gran afinidad que solo puede pasar a través de las membranas citoplasmáticas y nucleares de las células que han perdido la integridad de ambas membranas; por lo tanto, las células fluorescentes brillantes están muertas.

Basado en la perturbación específica, las células pueden perder (o, en algunos casos, ganar) GSH intracelular, lo que da lugar a una pérdida (o ganancia) de señal de fluorescencia.

Los tensores de parámetros para la población de la gráfica de nubes de células normales se calcularon y se representaron o describieron gráficamente en la Figura 1B.

**Ejemplo 2: Células tratadas con mixotiazol, FCCP y fluoxetina**

Las células se trataron con un agente perturbador (por ejemplo, mixotiazol, FCCP y fluoxetina) y se analizaron como se describe en el Ejemplo 1.

Los efectos de exponer las células a 100 µM de mixotiazol se muestran en la Figura 2. El tratamiento de las células con mixotiazol dio lugar a un cambio significativo en la morfología celular (panel superior izquierdo), disminución en GSH y CALCEÍNA AM (paneles superiores centrales y derechos), y un aumento en ROS y SYTOXRED™ (paneles inferiores) en la mayoría de las células (azul oscuro en todos los paneles). Cada una de estas respuestas es visualmente diferente del patrón para cada colorante en células "normales", como se muestra en la Figura 1A.

La exposición de células a 33 µM de FCCP provocó una curva de respuesta diferente, que se muestra en la Figura 3. Los resultados demuestran otro patrón adicional, distinto al patrón "normal" y al patrón para 100 µM de mixotiazol. Una población distinta de células con "morfología normal" está presente después de tratamiento con FCCP (células verdes en el panel superior izquierdo); estas mismas células están coloreadas en verde en todos los paneles y, como se muestra, con niveles normales de GSH y CALCEÍNA AM (paneles superiores centrales y derechos), y membranas nucleares y citoplasmáticas intactas, como se determina por la falta de absorción de SYTOX™ Red (panel inferior derecho). Sin embargo, las células tratadas con FCCP tienen ROS aumentadas (un indicador de estrés mitocondrial; panel inferior izquierdo). Además, hay una población distinta de células que muestra morfología alterada, GSH reducido, CALCEÍNA aumentada, ROS aumentadas y SYTOXRED™ mayor (mostradas en todos los paneles en la Figura 3 como células coloreadas en azul claro) en comparación con células con "morfología normal".

Otro patrón de respuesta adicional se muestra en la Figura 4A para células expuestas a 100 µM de fluoxetina. Aproximadamente un 30% de las células ha perdido la "morfología normal", muestran GSH reducido, CALCEÍNA AM reducida, ROS aumentadas y niveles significativamente mayores de SYTOXRED™ (células azul oscuro en todos los paneles en la Figura 4A). Un "firma" distintiva de esta respuesta es el patrón visto en el histograma de SYTOXRED™ (parte inferior derecha) —todas las células con contenido alto de SYTOXRED™ muestran una absorción de colorante que está leyendo el contenido de ADN (ciclo celular) de la línea celular HL-60 parental, que indica que, en este periodo de tiempo del tratamiento, las células se mueren durante el ciclo celular, y su ADN no se ha degradado todavía (un distintivo normal de células apoptóticas).

Un conjunto de perfiles multidimensionales de la Figura 2, la Figura 3 y la Figura 4A se representa gráficamente a lo largo de cuatro ejes en la Figura 4B: uno para las células tratadas con 100 µM de mixotiozol (parte superior izquierda), 33 µM de FCCP (parte superior derecha) o 100 µM de fluoxetina (parte inferior central).

**Ejemplo 3: Ciclo celular**

Las mediciones idénticas a aquellas descritas anteriormente en el Ejemplo 2 se hicieron utilizando una segunda combinación de colorantes:

(1) VYBRANVIOLET™ para medir el ciclo celular en células vivas;

(2) JC-9 para medir el potencial de la membrana mitocondrial (MMP); y

(3) SYTOXRED™ (como se describió anteriormente en este documento).

La medición del ciclo celular de células no tratadas, que muestra poblaciones distintas de células en las fases G1, S y G2M del ciclo celular, se representan gráficamente en la Figura 5. Una primera población que constituye el pico afilado y alto a la izquierda de la gráfica son células en la G1. La porción plana central de la gráfica surge de las células en la fase S. El pico a la derecha de la gráfica se produce por células en la G2/M.

Los compartimientos del ciclo celular mostrados aquí son idénticos a aquellos medidos en células HL-60 fijadas, lo que demuestra equivalencia de la tinción de células vivas con planteamientos "convencionales" utilizando fijación de células para permitir que los colorantes de unión al ADN accedan al ADN nuclear, como se describe más adelante.

**Ejemplo 4A: Efecto de valinomicina**

La medición del potencial de la membrana mitocondrial (MMP) con JC-9 se realizó con dos canales de fluorescencia independientes (señales de fluorescencia verde y roja), ya que los cambios en el MMP causan agregación del colorante, con diferente emisión de fluorescencia por las moléculas de colorante agregadas frente a las no agregadas. Como se muestra en la Figura 6A, tanto la fluorescencia roja como la verde aumentaron inicialmente en respuesta a una concentración de valinomicina creciente, con una florescencia verde pico a ~0,05 µM de valinomicina. En concentraciones mayores de valinomicina, la fluorescencia roja aumenta, con la relación de fluorescencia roja/verde en aumento. Un aspecto único de este ensayo es que correlaciona cambios en el MMP (JC-9) y en la muerte celular (SYTOXRED™) con las fases del ciclo celular (G1, S y G2M).

**Ejemplo 4B: Efecto de idarubicina**

Se llevó a cabo un estudio idéntico a uno descrito anteriormente con exposición de células HL60 expuestas a 1,23 µM de idarubicina, un análogo de daunorubicina que se inserta en el ADN y evita el desenrollamiento durante la replicación del ADN y detiene la división celular. Como se muestra en la Figura 6B, la fluorescencia roja del colorante JC-9 es significativamente mayor en células tratadas con idarubicina que en las células de control no tratadas (los primeros dos puntos de datos en la derecha indican células de control).

**Ejemplo 5: Análisis de células fijadas**

Los estudios han establecido que la mitocondria experimenta cambios estructurales significativos durante las diferentes fases del ciclo celular normal. Véase Bami *et al*. ("Static cytofluorometry and fluorescence morphology of mitochondria and DNA in proliferating fibroblasts." Biotech. Histochem. Off. Publ. Biol. Stain Comm. 71, 66-70, 1996); Margineantu *et al*. ("Cell cycle dependent morphology changes and associated mitochondrial DNA." Mitochondrion 1, 425-435, 2002); Schieke *et al*. ("Coordination of mitochondrial bioenergetics with G1 phase cell cycle progression." Cell Cycle 7, 1782-1787, 2008) y Sweet *et al*. ("Changes in mitochondrial mass, membrane potential, and cellular adenosine triphosphate content during the cell cycle of human leukemic (HL-60) cells." J. Cell. Physiol. 180, 91-96, 1999). Los datos limitados sugieren diferencias en el potencial mitocondrial ($\Delta\Psi m$) en las poblaciones de G1 frente a las de S+G2M (Schieke *et al*., Cell Cycle 7, 1782-1787, 2008); sin embargo, no se han publicado informes que correlacionen los cambios en el MMP y en las fases del ciclo celular a nivel de células individuales. Además, no hay estudios que investiguen los efectos diferenciales de diferentes compuestos químicos, por ejemplo, compuestos conocidos por cambiar el MMP, en diferentes fases del ciclo celular. En los dos aspectos ilustrativos de la descripción descritos más adelante, el análisis del ciclo celular (Ejemplo 5A) y de la transducción de señal (Ejemplo 5B) se llevaron a cabo utilizando células fijadas.

**Ejemplo 5A: Análisis del ciclo celular utilizando células fijadas**

Las fases del ciclo celular se pueden medir en células vivas utilizando colorantes de unión al ADN que pueden pasar a través de las membranas citoplasmáticas y nucleares (por ejemplo, VYBRANVIOLET™, como se muestra en la Figura 5). Sin embargo, este planteamiento es problemático a veces. Algunos colorantes de unión al ADN se transportan fuera de las células viables por "bombas" de la membrana citoplasmática que retiran compuestos potencialmente tóxicos (por ejemplo, transportadores de P-glicoproteína encontrados comúnmente en células cancerosas y células tipo madre). Un problema adicional es que las membranas citoplasmáticas y nucleares intactas bloquean el acceso a proteínas internas por medio de conjugados de anticuerpos para citometría de flujo de fluorescencia. Para superar estas limitaciones, las células se fijaron y permeabilizaron, utilizando una combinación de fijadores químicos, que evitan la degradación de proteínas celulares, lípidos, carbohidratos al igual que de la estructura celular. Después de la fijación, las membranas citoplasmáticas y nucleares se permeabilizaron utilizando detergentes habituales, dicho proceso permite rastrear moléculas (por ejemplo, conjugados de anticuerpos) para acceder a los compartimientos intracelulares.

Las células K-562 (derivadas de una línea CML) de leucemia mieloide crónica humana se fijaron y permeabilizaron utilizando técnicas conocidas (por ejemplo, fijación con formaldehído seguida por permeabilización utilizando metanol a menos 20 °C). Estas células fijadas son adecuadas para medición simultánea del contenido de ADN (ciclo celular) y proteínas intracelulares e intranucleares. Véase, por ejemplo, Pollice *et al*. ("Sequential paraformaldehyde and methanol fixation for simultaneous flow cytometric analysis of DNA, cell surface proteins, and intracellular proteins." Cytometry 13, 432-444, 1992).

Después de lavado, las células fijadas y permeabilizadas se incubaron con uno o más conjugados fluoróforo-anticuerpo, entonces se lavaron para retirar los anticuerpos no unidos, y se analizaron por medio de citometría de flujo.

Como se muestra en la Figura 7, el análisis del ciclo celular, como se indica por el contenido de ADN, mostró un patrón similar de poblaciones de G1, S y G2M (panel superior izquierdo) como se mostró para la tinción del ciclo celular viable en la Figura 5.

Las mismas células se tiñeron también con un anticuerpo contra Ciclina A2 (conjugada con PE) y con un anticuerpo contra la forma fosforilada de la proteína histónica H3 (H3-F), conjugada con ALEXA FLUOR® 647. El análisis del ciclo celular frente a la H3-F (histograma superior derecho) muestra que la forma fosforilada de la proteína histónica H3 solo se expresa en la población de G2M; el análisis del ciclo celular frente a la expresión de la proteína Ciclina A2 (histograma inferior izquierdo) muestra un patrón más complejo de niveles crecientes de Ciclina A2 durante la progresión a través de la S, y dos poblaciones de expresión de Ciclina A2 en la G2M.

La biología de la expresión de Ciclina A2 en células no perturbadas está bien caracterizada, y después de alcanzar niveles pico de expresión durante la G2, la Ciclina A2 se degrada rápidamente al entrar en la mitosis (M). Como se muestra en el panel inferior derecho, el análisis de la expresión de Ciclina A2 frente a H3-F revela un patrón más complejo de la expresión de estas dos proteínas durante la progresión a través de la S tardía, a G2, a M. Al realizar "gating" en el análisis subsecuente solo en estas poblaciones del ciclo celular (utilizando el contenido de ADN), es evidente que la H3-F se expresa primero en células con los niveles de Ciclina A2 más altos (población G2), que la H3-F se mantiene en los niveles más altos mientras la Ciclina A2 se degrada, y que la H3-F se "pierde" entonces (por medio de desfosforilación del residuo de serina específico que se fosforiló tras la entrada en la G2) mientras el progreso celular vuelve desde la mitosis (M) a la G1. En el análisis citométrico de flujo tradicional, estas secuencias de cambios en la expresión de proteínas se establecen al realizar "gating" (seleccionar) manualmente de manera cuidadosa en diferentes poblaciones celulares (basado en el contenido de ADN) y analizar subsecuentemente la expresión de proteínas (u otras dianas definidas por diferentes conjugados de anticuerpos).

**Ejemplo 6: Transducción de señal**

Las vías de transducción de señal regulan múltiples vías del destino celular, que incluyen división celular (proliferación), muerte celular (apoptosis, anoikis) y diferenciación celular. En general, la señalización actúa a través de receptores citoplasmáticos que activan las vías descendentes después de unirse a un ligando específico reconocido por el receptor. El distintivo de las vías de señalización es una cascada de modificaciones de proteínas (frecuentemente, fosforilación de residuos de serina, o treonina o tirosina, aunque la señalización puede incluir otras modificaciones de proteínas, que incluyen metilación, acetilación, ubiquitinación y SUMOlización.

En la fosforilación, los sucesos de transducción de señal estudiados de manera más común en la actualidad, la activación del receptor de superficie celular por medio de su ligando cognado genera una actividad de quinasa en el receptor. Esta quinasa activa fosforila entonces los residuos específicos de aminoácidos en su compañero situado más abajo, normalmente, generando una actividad de quinasa. En términos generales, esta cascada continua hasta que la quinasa terminal se transloca al núcleo, donde se une a sitios de activación de la transcripción específicos en el ADN. Además, muchas vías de señalización actúan "lateralmente" para activar o inhibir otras vías de señalización (por ejemplo, la PI3 quinasa puede activar vías de MAP quinasa, además de la vía de síntesis de proteína Akt>mTOR>quinasa S6 ribosomal>S6 situadas más abajo).

Un ejemplo de vías de transducción de señal situadas más abajo del receptor 4 de tipo toll (TLR4) encontrado en monocitos sanguíneos periféricos se muestra en la Figura 8. En contraste con la mayoría de las interacciones receptor-ligando de superficie, en esta respuesta sumamente inusual, la unión del ligando cognado, lipopolisacárido (LPS), a TLR4 activa múltiples vías de transducción de señal, que incluyen todas las tres de MAP (proteína quinasa activada por mitógenos) quinasas (ERK, p38 y SAP/JNK MAP quinasas), PI3 quinasa y la vía NF-κB.

Todas estas tres vías se han monitorizado utilizando citometría de flujo en monocitos sanguíneos periféricos. Como anteriormente (para monitorizar proteínas asociadas con el ciclo celular), las células se pueden fijar y permeabilizar para permitir acceso intracelular (y nuclear) para conjugados fluoróforo-anticuerpo.

En este documento, se pueden utilizar los colorantes específicos que monitorean la permeabilidad de la membrana celular (marcador para muerte celular), permeabilidad de la membrana nuclear (marcador para muerte celular), vías de transducción de señal implicadas en la lesión del ADN (fosfo-histona H2A, fosfo-ATM/ATR, fosfo-p53), vías de transducción de señal implicadas en la inflamación (por ejemplo, NF-kB), transducción de señal por medio de vías de MAP quinasa (ERK, SAP/JNK, p38), transducción de señal por medio de la vía de PI3 quinasa (Akt, GSK3B, RS6K, S6), vías apoptóticas tempranas (por ejemplo, anexina V), vías apoptóticas medias (por ejemplo, caspasa escindida), vías apoptóticas tardías (por ejemplo, fragmentación de PARP/ADN), etc. Los resultados de un experimento representativo que mide la cinética de cuatro vías de señalización simultáneamente utilizando citometría de flujo se presentan en la

Figura 9. Aquí, se trató sangre entera humana con LPS en presencia (muestra experimental) o ausencia (muestra de control) del inhibidor GDC0941 de PI3 quinasa. Las células se fijaron y permeabilizaron según los métodos publicados previamente en Shankey *et al*. ("Whole blood fixation and penneabilization protocol with red blood cell lysis for flow cytometry of intracellular phosphorylated epitopes in leukocyte subpopulations." Cytometry 67A;4-17, 2005) y se tiñeron con CD14-PECy7 para identificar monocitos en la mezcla celular en células sanguíneas blancas normales), más anticuerpos para IκBα (conjugados con ALEXAFLUOR™488), P-ERK (conjugados con ALEXAFLUOR™647), P-Akt (conjugados con PE), y P-S6 (conjugados con azul pacífico).

Como se muestra en la Figura 9, el tratamiento con LPS activó la Ik quinasa, lo que da lugar a degradación proteasomal de IκBα y a pérdida de señal de fluorescencia verde/ALEXA FLUOR® 488. El tratamiento con LPS también activó ERK, Akt y S6 (como se sometió a ensayo por medio de fosforilacion de efectores situados más abajo). En presencia del inhibidor GDC0941 de la PI3 quinasa (panel derecho), P-Akt no se fosforiló, y la cinética de la fosforilación de ERK y S6 se retrasó (en comparación con las células tratadas con LPS solo). La inhibición de PI3K no tuvo impacto tampoco en la degradación proteasomal de IκBα. Los resultados mostrados en este documento demuestran:

(1) Cinética reproducible de respuestas en una población celular normal

(2) Interacciones en vías (horizontales) (P-ERK activa P-S6 en monocitos)

(3) S6 no se activa por Akt, como se ha visto en otras células no hematopoyéticas

(4) El uso de inhibidores bien "mapeados" para estudiar vías.

**Ejemplo 7: Muestras de cultivo tisular**

Para proporcionar un sistema celular disponible inmediatamente y reproducible (y para evitar los problemas vistos en otros métodos existentes), los sistemas experimentales basados en líneas celulares de cultivo tisular se pueden utilizar para monitorizar el impacto de fármacos en vías de señalización.

Los métodos de citometría de flujo que utilizan células de cultivo tisular se han utilizado habitualmente para investigar los efectos de fármacos, por ejemplo, inhibidores de Bcr/Abl quinasa que son útiles en la terapia de leucemia mieloide crónica (LMC). La LMC se asocia con el cromosoma Filadelfia, una translocación genética que fusiona el gen Abl1 en el cromosoma 9 con parte del gen BCR en el cromosoma 22. La proteína de fusión resultante contiene un receptor tirosina quinasa que activa de manera constitutiva varias vías de señalización situadas más abajo, que incluyen P-STAT5, P-Crkl, P-mTOR y P-HSF. La Abl quinasa es la diana de varias terapéuticas utilizadas clínicamente en la actualidad, que incluyen imatinib (GLEEVEC™), nilotinib y dasatinib. Estos compuestos actúan inhibiendo la actividad de la tirosina quinasa a nivel de receptor, y también inhiben concomitantemente todas las vías de señalización descendentes.

Como modelo representativo de la LMC, la línea celular K562 humana, que expresa la proteína de fusión Bcr/Abl y fosforila de manera constitutiva la diana STAT5 en sentido descendente (Cytometry 54A; 75-88, 2003), se utilizó en el siguiente experimento. Como se muestra en la Figura 10, el tratamiento de células K562 durante 30 min con 2 µM de GLEEVEC™ (imatinib, o STI571) da lugar a >95% de inhibición de la fosforilación de la diana STAT5 en sentido descendente. También, como se muestra en la Figura 10, aunque la fosforilación de STAT5 se inhibe después de exposición a imanitib de 30 min, no hay un cambio en el ciclo celular, como se mide por el contenido de ADN.

STAT5 fosforilado (P-STAT5) actúa como activador transcripcional de varias proteínas diana, incluida la Ciclina D. La expresión constitutiva de Ciclina D (inducida por P-STAT5) mantiene a las células K562 en el ciclo celular. Se ha descubierto que la exposición a imatinib durante 24 h disminuye la fase S (como marcador de proliferación celular) un ~50%, y la exposición adicional a imatinib durante 24 h adicionales disminuye la fase S en un 50-70% adicional (datos no mostrados).

**Ejemplo 8: Efecto de troglitazona**

Las células se trataron con troglitazona, seguido por incubación con los diversos colorantes de fluorescencia (CALCEÍNA AM, MBBR, MITOSOX™ y CYTOSOX™). Para cada pocillo de no control, la distancia de chi cuadrado a las plantillas de control positivo y negativo (para cada canal de fluorescencia) se calculó y normalizó con el factor de escala. Los datos normalizados se reorganizaron en tensores, y se calculó la distancia desde las plantillas de control positivo y negativo. Las disimilitudes del tensor de troglitazona se pueden calcular utilizando técnicas conocidas por medio de comparación de las columnas de fibra de los tensores utilizando distancia de alineamiento temporal no lineal dinámica (véase Giorgino *et al*. "Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package." Journal of Statistical

Software, 31(7), 1-24, 2009). Las disimilitudes o distancias a todos los otros compuestos se pueden calcular de esta manera.

Los resultados se muestran en la Figura 12.

**Ejemplo 9: Derivación automatizada de respuestas de MMP celular sin *gating***

Las mediciones del potencial de la membrana mitocondrial se realizan con frecuencia en citometría de flujo utilizando sensores mitocondriales tales como JC-1 y JC-9. Estas especies químicas son colorantes que exhiben acumulación dependiente de potencial en la mitocondria, indicada por medio de un cambio en la emisión de fluorescencia a partir de la parte verde del espectro de luz visible (~525 nm para JC-1) a la roja (~590 nm para JC-1). Por lo tanto, los cambios fisiológicos en la mitocondria se indican por medio de un cambio en la relación de intensidad entre las dos bandas observadas de fluorescencia.

En citometría, la detección de un cambio general (a nivel de población) en MMP se mide típicamente por medio de *gating* (selección) de las subpoblaciones de células caracterizadas por niveles mayoritariamente bajos y altos de MMP. Esto requiere procesado manual de datos, o uso de varios algoritmos de agrupamiento que intentan identificar las subpoblaciones relevantes de células. Cuando se identifican las subpoblaciones, es posible crear curvas de respuesta del MMP, que ilustran funciones que unen concentraciones crecientes de compuestos y cambio correspondiente en el MMP.

Estas funciones (curvas) se crean típicamente utilizando el porcentaje de células en una ventana de análisis específica (por ejemplo, una ventana de análisis que delinea población con "MMP alto") como una variable sustitutiva que proporciona información con respecto a la respuesta de una población celular a exposición a un compuesto. Por lo tanto, las curvas de respuesta caracterizan los compuestos bioquímicos de interés y se pueden utilizar para agrupar además los compuestos en clústeres, para extraer las características del compuesto a partir de bases de datos de compuestos químicos, o para predecir la funcionalidad de los compuestos.

Aunque el *gating* se utiliza típicamente en las aplicaciones descritas, Bernas *et al*. demostró que es factible determinar las diferencias entre poblaciones de células en muestras de citometría de flujo sin él, al calcular una métrica (distancia) de disimilitud sólida entre las distribuciones que describen estas poblaciones (Bernas *et al*. (2008): Quadratic form: A robust metric for quantitative comparison of flow cytometric histograms; Cytometry A 73A, 715-726. doi:10.1002/cyto.a.20586). La métrica propuesta por Bernas (distancia de forma cuadrática) se puede operar en una dimensión o en un número mayor. Rajwa *et al*. enseña un método para calcular estas distancias y utilizarlas en el contexto de cuantificación de diferencias entre muestras biológicas en citometría de flujo (Rajwa *et al*. (2011): Quantification of differences between measured values and statistical validation based on the differences, publicación de solicitud de patente estadounidense n.º US20110066385 A1). Robinson *et al*. enseña un método para calcular curvas de respuesta y derivar los parámetros acompañantes de estas curvas, tales como IC50, etc. utilizando funciones de distancia (Robinson *et al*. (2013): Gate-free flow cytometry data analysis, publicación de solicitud de patente estadounidense n.º US20130226469 A1). En esta descripción, Robinson calcula las distancias entre una serie de muestras y un control. El control utilizado podría ser un control positivo, o un control negativo. La descripción demuestra calcular disimilitudes para canales individuales de fluorescencia.

Una aplicación de los métodos descritos por las referencias anteriores produciría cuatro curvas de respuesta complementarias: una que codifica la disimilitud entre la serie de muestras y un control positivo para la primera banda medida de fluorescencia, una segunda que codifica la disimilitud entre las muestras y el control negativo en la misma banda, y dos curvas más construidas análogamente para la segunda banda medida. Aunque es evidente que las cuatro curvas de respuesta tomadas conjuntamente contienen información similar a la información derivada utilizando el planteamiento tradicional que utiliza los porcentajes de células en el clúster identificado, no queda claro cómo las curvas derivadas se podrían utilizar en lugar de una única curva de respuesta, que tiene una utilidad bien establecida en el análisis de compuestos químicos y química medicinal.

Por lo tanto, las descripciones citadas anteriormente no se aplican al caso del MMP, aunque proporcionan una solución práctica a muchas otras mediciones de citometría de flujo.

En general, se puede decir que la técnica anterior no incluye un método para calcular una única curva de respuesta para muestras en las que la información sobre la respuesta biológica se codifica en más de una dimensión, y se requiere el uso de control tanto positivo como negativo. Bernaset *et al*. menciona el hecho de que la función de distancia propuesta es escalable a citometría multiparamétrica (multidimensional), los autores no describen una implementación aplicable de manera práctica o una demostración de un caso multidimensional. Robinson *et al*. enseña que las múltiples

comparaciones entre muestras en una serie y un control se pueden disponer en una serie que lleva a una curva de respuesta similar a una curva de respuesta extraída a partir de *gating* de datos manual. Sin embargo, Robinson *et al*. guarda silencio sobre el uso del método descrito en un caso bidimensional como las mediciones de JC-, JC-9, JC-10 y otros colorantes radiométricos. Esta descripción enseña la derivación de curvas de respuesta que utilizan funciones de distancia para las mediciones del MMP y otras mediciones en las que se utilizan dos o más tipos de controles, y las muestras biológicas se caracterizan por dos o más mediciones.

El método descrito más adelante es capaz de derivar una única curva de respuesta que contiene toda la información necesaria para determinar las características de la respuesta exhibida por el compuesto biológico de interés cuando actúa en la población de células.

El método implica la aplicación de funciones de distancia a histogramas marginales (histogramas unidimensionales) formados por mediciones de citometría de flujo.

Simulación *in silico*

Un método que utiliza una aplicación de una función de distancia a histogramas marginales (histogramas unidimensionales) formados por mediciones de citometría de flujo según aspectos de la descripción descritos en este documento se ilustra en el siguiente ejemplo *in silico*.

La Figura 19 ilustra una simulación por ordenador de diez muestras de citometría de flujo expuestas a una concentración creciente de un agente que influye en el potencial de la membrana mitocondrial. La Figura 18 muestra controles negativos y positivos. El cambio en el número y en el porcentaje de células que pertenecen al clúster de células definido por el control negativo se ilustra en la Figura 20. El panel A muestra curvas de respuesta como una función del número de células. El B muestra curvas de respuesta como una función del porcentaje de células.

Primero, se calcula un par de distancias entre los controles positivos ($C_p$) y el control negativo ($C_n$). Las dos distancias se definen por histogramas marginales ($C^I$) de FL1 e histogramas marginales ($C^{II}$) de FL2. Estas distancias son referencias para uso adicional y reescalado:

$$f_1 = D(C^I_p, C^I_i)$$

$$f_2 = D(C^{II}_p, C^{II}_i)$$

donde *i* es el número que indica la muestra en una serie de diluciones, y los números romanos (I, II) indican el número de canal. La ecuación muestra un caso bidimensional (dos histogramas marginales); por lo tanto, solo se utilizan dos números romanos. La función de distancia *D* puede ser una distancia de forma cuadrática, una distancia de Wasserstein, una distancia cuadrática $\chi^2$ o cualquier otra distancia que funcione en vectores que representan histogramas.

Después de esta operación, una serie de distancias para las muestras biológicas se calcula de forma análoga. Las distancias se calculan para cada par del control positivo y una muestra biológica en la serie ($\mathbf{S}_1$, $\mathbf{S}_2$, ..., $\mathbf{S}_1$), al igual que las muestras del control negativo (véase la Figura 21):

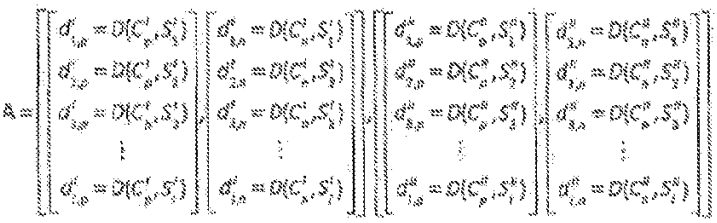
$$d^I_{n,i} = D(C^I_n, S^I_i)$$

$$d^I_{p,i} = D(C^I_p, S^I_i)$$

$$d^{II}_{n,i} = D(C^{II}_n, S^{II}_i)$$

$$d^{II}_{p,i} = D(C^{II}_p, S^{II}_i)$$

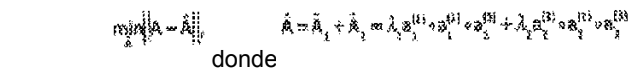
Los valores resultantes forman una matriz. El tamaño y la dimensionalidad de la matriz dependen del número de controles, del número de histogramas unidimensionales utilizados como se explica más adelante. En el caso del MMP, se utilizan los controles, y se miden los dos canales de fluorescencia.



La matriz de mediciones (dos matrices mostradas anteriormente) se pueden ver como un tensor tridireccional. La dimensionalidad de este tensor tridireccional es $I_1 \times I_2 \times I_3$, donde $I_1$ es el número de concentraciones (típicamente 10), $I_2$ es el número de respuestas medidas (en el caso de etiquetas tipo JC - dos), y $I_3$ es el número de disimilitudes/distancias medidas (típicamente dos: medición de distancia de control positivo, y una medición de distancia de control negativo). Por lo tanto, para un tensor **A** que representa la medición biológica, el elemento (*i,j,k*), denotado por $\alpha_{i,j,k}$, describe una distancia entre mediciones del parámetro *j* obtenido de una población de células expuestas a un compuesto a una concentración *i*, y una población de células de control *k*.

En la tercera etapa, después de la disposición de los datos en el tensor **A**, se realiza la descomposición del tensor. El objetivo de la descomposición del tensor es formar tensores de rango bajo, que contienen toda o la mayor parte de la información del tensor de rango superior. En el ejemplo demostrado, la descomposición utilizada es la descomposición de tensor poliádica (CP). El tensor A se descompone en dos tensores descompuestos.

El objetivo de una descomposición (CP) es encontrar una aproximación del tensor **A**, denotada como $\hat{\mathbf{A}}$, que satisfaga los siguientes criterios:



donde

donde o denota el producto exterior.

El primer vector que construye el primer tensor $\hat{\mathbf{A}}_1$ (denotado de aquí en adelante como $A_1$ para simplificar) $a_1^{(1)}$ (denotado de aquí en adelante como $a_1$) describe el cambio total en respuesta exhibido por la serie de las muestras biológicas sobre la concentración en reposo de los compuestos (Figura 22A). El vector $a_1$ se puede caracterizar además al ajustar un modelo preconcebido de respuesta (tal como log-logístico, Gompertz, Weibull, Cedergreen-Ritz-Streibig, Brain-Cousens, etc) (Cedergreen *et al*., 2005; Meddings et al., 1989). El ejemplo de un ajuste de este tipo se muestra en la Figura 22B.

En todas las muestras simuladas *in silico* mostradas en la Figura 19, el número total de células es 5000. El número de células en el clúster tipo "control positivo" mostrado en de las Figuras 19A a las Figuras 19J es 4909, 4753, 4409, 3650, 2478, 1370, 581, 234, 87 y 29, respectivamente. Como se mencionó anteriormente, la información sobre el cambio que se puede visualizar como un porcentaje de células en el clúster, o como una diferencia entre el porcentaje de células en el clúster y el porcentaje de células que pertenecen al clúster en el control negativo. En el ejemplo demostrado, estos porcentajes son 98,18, 95,06, 88,18, 73, 49,56, 27,4, 11,62, 4,68, 1,74, y 0,58 para las muestras de A a J (Figura 19), respectivamente. En todos los casos una forma sigmoidea caracteriza la función representada gráficamente.

La curva de respuesta calculada por descomposición de tensor poliádica (curva descrita por el vector $a_1$) se demuestra en la Figura 22A. La curva retiene las características sigmoideas que proporcionan la misma información sobre dependencia funcional entre concentración (codificada por número de muestra) y un nivel relativo de respuesta. La curva se puede reescalar utilizando distancias entre controles. Después del reescalado, la disimilitud de 1 es igual a la disimilitud entre control negativo y positivo. La gráfica que presenta la correspondencia entre las curvas de respuesta derivadas utilizando la metodología tradicional, y el método descrito se proporciona en la Figura 23.

(B) Análisis de respuestas celulares a valinomicina y a otros dos agentes

El procedimiento descrito anteriormente se realizó en datos de citometría de flujo reales. Los resultados se muestran en las Figuras 24 y 25. Los datos se obtuvieron utilizando los métodos descritos anteriormente en este documento.

La Figura 24 muestra una serie de gráficas de citometría de flujo que demuestran el cambio en poblaciones de células expuestas a valinomicina. La Figura 25A proporciona curvas de respuesta derivadas de los datos de citometría de flujo.

Además de la valinomicina en la Figura 24, adicionalmente, se utilizaron muestras expuestas a idarubicina y acetomifeno. La unidad de disimilitud en la Figura 25A es la diferencia entre control positivo y negativo, donde el control positivo es una muestra tratada con FCCP en una concentración de 25 µm.

Los datos de respuesta calculados se pueden utilizar también para ajustar un modelo de curva sigmoidea definido. Por ejemplo, un modelo log-logístico de cuatro parámetros ajustado para proporcionar ejemplos representa la dependencia funcional aproximada entre concentración y respuesta de la muestra en la Figura 25B. Aspectos adicionales de la descripción se exponen en las siguientes cláusulas numeradas como ejemplos comparativos. Los párrafos numerados son autorreferenciales. En particular, la expresión "según cualquiera de lo anterior o de lo siguiente" utilizada en estos párrafos se refiere a otros párrafos. La expresión significa, en los siguientes párrafos, aspectos de la descripción descritos en este documento que incluyen tanto la materia descrita en los párrafos individuales por sí solos como la materia descrita por los párrafos tomados en combinación.

Métodos y análisis

A1. Un método para caracterizar una o más respuestas celulares a un agente, que comprende:

exponer las poblaciones de células a una pluralidad de concentraciones, c, de un agente;

medir por medio de citometría una pluralidad de parámetros fisiológicos, p, de células en la población en cada concentración;

calcular un conjunto de distancias entre poblaciones y controles para cada parámetro para la población de células en cada concentración; y

compilar uno o más tensores para cada compuesto a partir de distancias calculadas; y

comprimir los tensores por descomposición de tensores para producir una huella abreviada de compuesto en forma de un vector.

A2. Un método para comparar una o más respuestas celulares a un agente, que comprende:

(A) exponer primeras poblaciones de células a una pluralidad de concentraciones de un primer agente;

medir por medio de citometría una pluralidad de parámetros fisiológicos, p, de células en la población en cada concentración de dicho primer agente;

compilar uno o más tensores a partir de mediciones, describiendo así dicho primer agente;

comprimir el tensor(es) por descomposición de tensores para obtener una huella abreviada de compuesto en forma de un vector;

(B) exponer segundas poblaciones de las células a una segunda pluralidad de concentraciones de un segundo agente; medir por medio de citometría la pluralidad de parámetros fisiológicos de células en la población en cada concentración de dicho segundo agente;

compilar uno o más tensores a partir de mediciones, describiendo así dicho segundo agente;

comprimir el tensor(es) por descomposición de tensores para obtener una huella(s) abreviada de compuesto en forma de un vector, y

(C) calcular una disimilitud entre la primera y la segunda huella abreviada para determinar la diferencia entre la respuesta de las células al primero y al segundo agente.

A3. Un método para determinar una o más respuestas de células a un agente, que comprende:

medir dos o más respuestas de fisiología celular para uno o más controles negativos, o uno o más positivos y para una o más concentraciones de un compuesto;

seleccionar subpoblaciones de células para los controles y la serie de concentraciones por medio de restricción matemática realizando así *gating* de las células en un compartimiento particular del ciclo celular y una clase morfológica particular;

calcular una disimilitud entre las distribuciones de mediciones celulares para cada control positivo y negativo y cada una de las concentraciones;

caracterizar la respuesta de las células al compuesto por medio de la disimilitud calculada.

A4. Un método según cualquiera de lo anterior o de lo siguiente, en donde las respuestas se calculan utilizando una métrica matemática que opera en distribuciones.

A5. Un método según cualquiera de lo anterior o de lo siguiente, en donde las mediciones forman una nuble de puntos de datos multidimensionales.

A6. Un método según cualquiera de lo anterior o de lo siguiente, en donde los cambios en las nubes de puntos de datos multidimensionales se calculan como distancias por medio de uno o más de lo siguiente una métrica de Wasserstein, una métrica definida como solución al problema de transporte de Kantorovich-Rubinstein, una distancia de forma cuadrática, una distancia de chi cuadrado, divergencia de Kullback-Leibler, una divergencia de Jensen-Shannon, métrica de Kolmogorov, ϕ-divergencia de Csiszár, una divergencia de Burbea y Rao y divergencia de Bregman.

A7. Un método según cualquiera de lo anterior o de lo siguiente, en donde el valor en el percentil 95 de las distancias entre pares en un grupo de control se elige como el límite de precisión de la medición (límite de significancia estadística).

A8. Un método según cualquiera de lo anterior o de lo siguiente, en donde una medida sólida de disimilitud entre un grupo de controles positivos y un grupo de controles negativos define una unidad de disimilitud entre las respuestas.

A9. Un método según cualquiera de lo anterior o de lo siguiente, que comprende cuantificar los cambios en distribución de las nubes de puntos de datos multidimensionales medidos en una pluralidad de concentraciones diferentes del compuesto.

A10. Un método según cualquiera de lo anterior o de lo siguiente, que emplea además reducción de dimensionalidad por medio de características de resumen a partir de una pluralidad de tablas de respuestas en donde cada vector de respuesta se resume por medio de un número de cantidades derivadas más pequeño que el número de concentraciones de dicho compuesto.

A11. Un método según cualquiera de lo anterior o de lo siguiente, que representa además una respuesta a un agente por medio de un tensor multidireccional que comprende características de resumen.

A12. Un método según cualquiera de lo anterior o de lo siguiente, en donde los tensores calculados para diferentes agentes se comparan entre sí calculando una medida de disimilitud.

A13. Un método según cualquiera de lo anterior o de lo siguiente, en donde los tensores se descomponen utilizando uno o más de lo siguiente descomposición de Tucker, descomposiciones de CANDECOMP, PARAFAC, PARAFAC2, INDSCAL, CANDELINC, DEDICOM o PARATUCK2.

Tensores y tensores de estado

TSt1. Un tensor de respuesta celular según cualquiera de lo anterior o de lo siguiente, en donde el tensor de respuesta se genera a partir de un conjunto de datos de valores medidos de una pluralidad de dos o más parámetros fisiológicos, en donde el tensor de respuesta cuantifica un fenotipo celular multiparamétrico y multifactorial.

TSt2. Un tensor de respuesta celular según cualquiera de lo anterior o de lo siguiente, codificado de forma tangible de modo que se pueda acceder a él, copiarlo, utilizarlo y/o recuperarlo por completo o en parte por un usuario.

TSt3. Un tensor de respuesta celular según cualquiera de lo anterior o de lo siguiente, en donde un tensor de respuesta para un conjunto de datos de control proporciona la base para determinar si un conjunto de datos de prueba es diferente de un conjunto de datos de control o de perfil.

Controles

Ctrl. Un método según cualquiera de lo anterior o de lo siguiente, en donde se tratan células de control positivo con uno o más compuestos conocidos que desencadenan un efecto máximo medible en una o más de las respuestas de fisiología celular medidas.

Ctr2. Un método según cualquiera de lo anterior o de lo siguiente, en donde los controles negativos son células no tratadas, células tratadas con tampón, células tratadas con medio, o células tratadas con un compuesto simulado.

Ciclo celular

Ccy1. Un método según cualquiera de lo anterior o de lo siguiente, en donde el estado celular es una medición de la fase de crecimiento de las células, preferiblemente, una medición de la división celular.

Ccy2. El método según cualquiera de lo anterior o de lo siguiente, en donde se detecta el estado celular o la fase del ciclo celular por medio de citometría de flujo a nivel de célula individual.

Ccy3. Un método según cualquiera de lo anterior o de lo siguiente, en donde uno de los parámetros fisiológicos es el ciclo celular.

Ccy4. Un método según cualquiera de lo anterior o de lo siguiente, en donde uno de los parámetros fisiológicos es el compartimiento del ciclo celular G1, S, G2 y M.

Ccy5. Un método según cualquiera de lo anterior o de lo siguiente, en donde uno de los compartimientos del ciclo celular G1, S, G2/M.

Ccy6. Un método según cualquiera de lo anterior o de lo siguiente, en donde todas las respuestas fisiológicas se miden como una función del compartimiento del ciclo celular.

Ccy7. Un método según cualquiera de lo anterior o de lo siguiente, en donde las fases del ciclo celular se miden utilizando etiquetas de fluorescencia.

Ccy8. Un método según cualquiera de lo anterior o de lo siguiente, en donde las fases del ciclo celular se miden utilizando uno o más colorantes intercalantes fluorescentes de ADN.

Ccy9. Un método según cualquiera de lo anterior o de lo siguiente, en donde las fases del ciclo celular se miden utilizando uno o más de los colorantes intercalantes fluorescentes HOECHST 33342(2'-(4-etoxifonil)-6-(4-metil-1-piperazinil)-1H,3'H-2,5'-bibenzanidazol), DRAQ5™ (1,5-bis {[2-(di-metilamino) etil]amino)-4, 8-dihidroxiantraceno-9,10-diona), YODURO YO-PRO-1 (Quinolinium, 4-((3-metil-2(3H)-benzoxazolilideno)metil)-1-(3-(trimetilammonio)propil)-,diYODURO), DAPI (4', 6-diamidino-2-fenilindol) y CYTRAK ORANGE (derivado de 1,5-bis{[2-(di-metilamino) etil]amino}-4, 8- dihidroxiantraceno-9,10-diona).

Ccy10. Un método según cualquiera de lo anterior o de lo siguiente, en donde las fases del ciclo celular se miden por medio de inmunomarcaje de proteínas dependientes del ciclo celular.

Ccy11. Un método según cualquiera de lo anterior o de lo siguiente, en donde las fases del ciclo celular se miden utilizando inmunomarcaje de una o más de ciclinas A, ciclina B y ciclina E.

Ccy12. Un método según cualquiera de lo anterior o de lo siguiente, en donde las fases del ciclo celular se miden por medio de inmunomarcaje de una o más proteínas histónicas fosforiladas.

Ccyl3. Un método según cualquiera de lo anterior o de lo siguiente, en donde las fases del ciclo celular utilizan fluorocromos dependientes del ciclo celular marcados genéticamente, por ejemplo, proteína Rb hiperfosforilada y el ciclo celular se puede medir por medio de citometría de flujo (véase Juan *et al.* "Phosphorylation of retinoblastoma susceptibility gene protein assayed in individual lymphocytes during their mitogenic stimulation," Experimented Cell Res 239: 104-110, 1998) y la expresión de la proteína ciclina (o su estado de fosforilación) se puede monitorear utilizando citometría de flujo (véase Darzynkiewicz *et al*. "Cytometry of cell cycle regulatory proteins." Chapter in: Progress in Cell Cycle Research 5;533-542, 2003.

Ccy14. Un método según cualquiera de lo anterior o de lo siguiente, en donde las fases del ciclo celular se miden por medio de la expresión de un resto de proteína de fusión codificada genéticamente que comprende una proteína oscilante expresada de manera natural unida a un resto de proteína fluorescente, por ejemplo, un ciclo celular detenido en la G2/M (Cheng *et al*., "Cell-cycle arrest at G2/M and proliferation inhibition by adenovirus-expressed mitofusin-2 gene in human colorectal cancer cell lines," Neoplasma 60; 620-626, 2013); regulación de la entrada a la fase S (McGowan *et al*., "Platelet-derived growth factor-A regulates lung fibroblast S-phase entry through p27kipl and FoxO3a," Respiratory Research, 14;68-81, 2013); o identificación de células proliferantes vivas utilizando un informador de fusión ciclina B1-GFP (véase Klochendler *et al*., "A transgenic mouse marking live replicating cells reveals in vivo transcriptional program of proliferation," Developmental Cell, 16;681-690,2012).

Ccy15. Un método según cualquiera de lo anterior o de lo siguiente, en donde se altera el ciclo celular por medio de un agente.

Ccy16. Un método según cualquiera de lo anterior o de lo siguiente, en donde se altera el ciclo celular por medio de una variación en el método de cultivo celular.

Ccy17. Un método según cualquiera de lo anterior o de lo siguiente, en donde se altera el ciclo celular por medio de cambios en los niveles de uno o más de lo siguiente en el medio de cultivo: glucosa, aminoácidos esenciales y no esenciales, concentración de $O_2$, pH, galastosa y/o glutamina/glutamato.

Ccy18. El método según cualquiera de lo anterior o de lo siguiente, que comprende además detectar el estado celular o la fase del ciclo celular en una población de control de células expuestas a una pluralidad de químicos o agentes que son conocidos por perturbar el estado del ciclo celular.

Células

Cls1. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células cultivadas *in vitro.*

Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células de biopsia.

Cls2. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células vivas.

Cls3. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células fijadas.

Cls4. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son una línea celular.

Cls5. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son características de un tipo celular sano que existe de manera natural.

Cls6. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son características de una enfermedad.

Cls7. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son características de un trastorno genético congénito.

Cls8. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son características de un cáncer.

Cls9. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son características de un trastorno metabólico.

Cls10. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células animales.

Cls11. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células de mamífero.

Cls12. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células humanas.

Cls13. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células germinales o células madre, que incluyen, células madre pluripotentes.

Cls14. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células somáticas.

Cls15. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células madre.

Cls16. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células madre embrionarias.

Cls17. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células madre pluripotentes.

Cls18. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células madre pluripotentes inducidas.

Cls19. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células blásticas.

Cls20. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células diferenciadas.

Cls21. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son células somáticas diferenciadas terminalmente.

Cls22. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son cardiomiocitos, hepatocitos, neuronas o una combinación de estos.

Cls23. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células son una o más de las siguientes células: células primarias, células transformadas, células madre, células de insectos, células de levadura, preferiblemente células independientes de anclaje, tales como, por ejemplo, líneas celulares hematopoyéticas humanas (que incluyen, pero no se limitan a, HL-60, K562, CCRF-CEM, Jurkat, THP-1, etc.); o líneas celulares dependientes de anclaje (que incluyen, pero no se limitan a, HT-29 (colon), T-24 (vejiga), SKBR (pecho), PC-3 (próstata), etc.).

Duración

Durl. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células se exponen a un agente durante una pluralidad de duraciones o de varios periodos de tiempo, por ejemplo, midiendo el curso del tiempo (cinética) para la activación de vías de señalización en células (véase, por ejemplo, Woost *et al*., "High-resolution kinetics of cytokine signaling in human CD34/CD117-positive cells in unfractionated bone marrow," Blood, 117; 131-141, 2011). Se prefieren los aspectos de la descripción que implican análisis de la cinética por encima de los aspectos de la descripción que no requieren un análisis de este tipo (véase Komblau *et al*. "Dynamic single-cell network profiles in acute myelogenous leukemia are associated with patient response to standard induction therapy," Clin Cancer Res, 16;3721-3733,2010, which does not teach kinetic analysis).

Dur2. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células se exponen a un agente durante 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 15, 16, 18, 20, 22, 24, 26, 28, 30, 35, 40, 44, 48, 52, 56, 60, 66, 72, 78 o más horas o cualquier combinación de estas.

Concentración

Cnc1. Un método según cualquiera de lo anterior o de lo siguiente, en donde una pluralidad de dos o más series de concentraciones de un agente.

Pluralidad (número) de muestras

Plr1. Un método según cualquiera de lo anterior o de lo siguiente, que comprende en donde se mide una pluralidad de muestras.

Plr2. Un método según cualquiera de lo anterior o de lo siguiente, que comprende medir una pluralidad de muestras dispuestas en pocillos de una placa multipocillo.

Plr3. Un método según cualquiera de lo anterior o de lo siguiente, que comprende medir una pluralidad de muestras dispuestas en pocillos de una placa de 96, 384 o 1536 pocillos.

Métodos de instrumentación / básicos

Ins1. Un método según cualquiera de lo anterior o de lo siguiente, en donde las respuestas se miden por medio de citometría.

Ins2. Un método según cualquiera de lo anterior o de lo siguiente, en donde las respuestas se miden por medio de citometría de flujo.

Ins3. Un método según cualquiera de lo anterior o de lo siguiente, en donde las respuestas se miden por medio de citometría de flujo de células vivas.

Ins4. Un método según cualquiera de lo anterior o de lo siguiente, en donde las respuestas se miden por medio de citometría de flujo de células fijadas.

Ins5. Un método según cualquiera de lo anterior o de lo siguiente, en donde las respuestas se miden por medio de diagnóstico por imágenes de células inmovilizadas.

Ins6. Un método según cualquiera de lo anterior o de lo siguiente, en donde las respuestas se miden por medio de fluorimetría.

Ins7. Un método según cualquiera de lo anterior o de lo siguiente, en donde una pluralidad de dos o más parámetros de respuesta se mide por medio de una matriz de sensores multicanal.

Procesamiento de la señal

Sig1. Un método según cualquiera de lo anterior o de lo siguiente, que comprende correlacionar señales de fluorescencia por medio de separación lineal de las señales obtenidas al multiplicar el vector de los valores medidos por una matriz inversa de una matriz que contiene en sus columnas los espectros de las especies fluorescentes utilizadas; dicha matriz se normaliza por columna a 1.

Sig2. Un método según cualquiera de lo anterior o de lo siguiente, que comprende correlacionar señales de fluorescencia por medio de separación lineal de las señales obtenidas al multiplicar el vector de los valores medidos por una matriz inversa de una matriz que contiene en sus columnas los espectros de las especies fluorescentes utilizadas; dicha matriz se normaliza por diagonal a 1.

Agentes

Agt1. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células se exponen a un único compuesto.

Agt2. Un método según cualquiera de lo anterior o de lo siguiente, en donde las células se exponen a dos o más compuestos.

Agt3. Un método según cualquiera de lo anterior o de lo siguiente, en donde uno o más de los compuestos estimulan una respuesta fisiológica.

Agt4. Un método según cualquiera de lo anterior o de lo siguiente, en donde el agente puede ser un agente genético, por ejemplo, una secuencia codificante expresada; o un agente químico, por ejemplo, un posible fármaco.

Parámetros fisiológicos

MMP

MMP1. Un método según cualquiera de lo anterior o de lo siguiente, en donde se mide toxicidad mitocondrial.

MMP2. Un método según cualquiera de lo anterior o de lo siguiente, en donde se mide la pérdida de potencial o de integridad de la membrana mitocondrial.

MMP3. Un método según cualquiera de lo anterior o de lo siguiente, en donde se mide la pérdida de potencial o de integridad de la membrana mitocondrial utilizando colorante fluorescente.

MMP4. Un método según cualquiera de lo anterior o de lo siguiente, en donde se mide la pérdida de potencial o de integridad de la membrana mitocondrial utilizando uno o más de lo siguiente JC-1 (YODURO de 5,5',6,6'-tetracloro-1,1',3,3'-tetraetilbenzimi-dazolilcarbocianina), JC-9 ((YODURO de 3,3'-dimetil-β-naftoxazolio, MTTOPROBE™, Molecular Probes), JC-10 (por ejemplo, derivado of JC-1), DiOC2(3) ((YODURO de 3,3'-diemiloxacarbocianina; MITOPROBE™ Molecular Probes), DiICI(5)

((YODURO de 1,1',3,3,3',3'-hexametilindodicarbo - cianina; MITOPROBE™, Molecular Probes), MITOTRACKER™ (Molecular Probes), ORANGE CMTMROS (diacetato de clorometil-diclorodihidrofluoresceína, MTTOTRACKER™ ORANGE, Molecular Probes) y CMXROS (1H,5H,11H,15H-Xanteno[2,3,4-ij:5,6,7-i'j']diquinolizin-18-io, 9-[4-(clorometil)fenil]-2,3,6,7,12,13,16,17-octahidro-, cloruro, MITOTRACKER™ RED, Molecular Probes).

Dispersión de la luz

LSg1. Un método según cualquiera de lo anterior o de lo siguiente, en donde un parámetro fisiológico del estado celular se mide por medio de dispersión de la luz.

LSg2. Un método según cualquiera de lo anterior o de lo siguiente, en donde un parámetro fisiológico del estado celular se mide por medio de dispersión de la luz láser.

LSg3. Un método según cualquiera de lo anterior o de lo siguiente, en donde un parámetro fisiológico del estado celular se mide al cuantificar la cantidad de luz láser dispersada a partir de una célula individual en dos o más ángulos.

LSg4. Un método según cualquiera de lo anterior o de lo siguiente, en donde un parámetro fisiológico del estado celular se mide por medio de dispersión de la luz láser.

LSg5. Un método según cualquiera de lo anterior o de lo siguiente, en donde un parámetro fisiológico del estado celular se mide por medio de dispersión de la luz láser, en donde la longitud de onda de la luz emitida por el láser se encuentra en el intervalo de uno o más de lo siguiente 403-408 nm, 483-493 nm,525-535 nm, 635-635 nm y 640-650 nm.

Viabilidad celular

Via1. Un método según cualquiera de lo anterior o de lo siguiente, en donde se mide la viabilidad celular.

Via2. Un método según cualquiera de lo anterior o de lo siguiente, en donde se mide la integridad de la membrana celular.

Via3. Un método según cualquiera de lo anterior o de lo siguiente, en donde la viabilidad celular se determina al medir la integridad de la membrana.

Via4. Un método según cualquiera de lo anterior o de lo siguiente, en donde la pérdida de integridad de la membrana se detecta utilizando un colorante.

Via5. Un método según cualquiera de lo anterior o de lo siguiente, en donde la pérdida de integridad de la membrana se detecta utilizando un colorante que entra en las células con membranas dañadas características de células moribundas o muertas, pero que no entra en células con membranas intactas características de células vivas.

Via6. Un método según cualquiera de lo anterior o de lo siguiente, en donde la pérdida de integridad de la membrana se detecta utilizando un colorante que entra en las células con membranas dañadas características de células moribundas o muertas, pero que no entra en células con membranas intactas características de células vivas, en donde los colorantes fluororescen al unirse al ADN.

Via7. Un método según cualquiera de lo anterior o de lo siguiente, en donde la pérdida de integridad de la membrana se detecta utilizando uno o más de los siguientes colorantes: YODURO DE PROPIDIO, DAPI, y 7-aminoactinomicina D.

Via8. Un método según cualquiera de lo anterior o de lo siguiente, en donde la pérdida de integridad de la membrana se mide utilizando uno o más colorantes que cruzan membranas celulares intactas y fluorescen tras interactuar con enzimas intracelulares y permanecen en el citoplasma de células vivas pero se difunden fuera en ausencia de una membrana citoplasmática intacta.

Via9. Un método según cualquiera de lo anterior o de lo siguiente, en donde la pérdida de integridad de la membrana se mide utilizando uno o más colorantes que cruzan membranas celulares intactas y fluorescen tras interactuar con enzima intracelulares y permanecen en el citoplasma de células vivas pero se difunden fuera en ausencia de una membrana citoplasmática intacta, en donde los colorantes son uno o más de los siguientes diacetato de fluoresceína, CALCEÍNA AM, BCECF AM, diacetato de carboxieosina, CELLTRACKER™ GREEN CMFDA, acetato de clorometilo SNARF-1 y diacetato de ácido carboxílico OREGON GREEEN 488.

Via10. Un método según cualquiera de lo anterior o de lo siguiente, en donde la vialibilidad se mide por medio de uno o más de lo siguiente Anexina V, caspasas escindidas / activación de caspasa, que incluye fosforilación y/o degradación de la lámina nuclear.

GLU, ROS, MMP, CMP y viabilidad

GRC1. Un método según cualquiera de lo anterior o de lo siguiente, en donde se miden uno o más de los siguientes parámetros fisiológicos: concentración de glutatión ("GLU"), radicales libres y/o especies de oxígeno reactivo ("ROS"), potencial/permeabilidad de la membrana mitocondrial ("MMP"), permeabilidad de la membrana citoplasmática y viabilidad celular.

Lesión del ADN; estrés, inflamación, metabolismo, apoptosis

DSI1. Un método según cualquiera de lo anterior o de lo siguiente, en donde se miden uno o más de los siguientes parámetros fisiológicos: lesión del ADN; un constituyente de la vía de señalización de respuesta al estrés; un constituyente de la vía de la respuesta inflamatoria; un constituyente regulador de la ruta metabólica o un constituyente de la vía apoptótica.

DSI2. Un método según cualquiera de lo anterior o de lo siguiente, en donde se mide el constituyente de la vía de señalización de respuesta al estrés SAPK.

DSI3. Un método según cualquiera de lo anterior o de lo siguiente, en donde se mide el constituyente de la vía de señalización de respuestas inflamatorias NF-kB.

DS14. Un método según cualquiera de lo anterior o de lo siguiente, en donde el constituyente regulador de la ruta metabólica medido es una peroxidasa lipídica, GSk3B y/o quinasa S6 ribosomal.

DSI5. Un método según cualquiera de lo anterior o de lo siguiente, en donde el constituyente de la vía apoptótica medido es PI3K, AKT y/o una proteína de la familia Bcl.

Bancos de referencia

Rbk1. Un método según cualquiera de lo anterior o de lo siguiente, en donde los químicos perturbadores conocidos o agentes moleculares exógenos se subdividen además en función de sus efectos conocidos.

Rbk2. Un método según cualquiera de lo anterior o de lo siguiente, que comprende además crear tablas de respuestas que comprenden información sobre los cambios en viabilidad celular, toxicidad mitocondrial y al menos un descriptor fisiológico o fenotípico adicional en cada concentración empleada de dicho compuesto calculada para cada fase del ciclo celular definida por marcadores dependientes del ciclo celular.

Rbk3. Un método según cualquiera de lo anterior o de lo siguiente, en donde los tensores que describen compuestos conocidos utilizados para tratar una enfermedad particular se agrupan en una única clase definida o en una pluralidad de clases definidas, y los tensores de compuestos se utilizan como conjunto de formación para un aprendizaje supervisado que clasifica compuestos desconocidos y no caracterizados previamente en dichas clases definidas.

Rbk4. Un método según cualquiera de lo anterior o de lo siguiente, en donde los tensores que describen compuestos conocidos se agrupan en clases con base en sus respuestas no específicas, tales como, efectos secundarios.

Rbk5. El método según cualquiera de lo anterior o de lo siguiente, en donde se utilizan tensores de características para descubrir clústeres de compuestos similares utilizando aprendizaje no supervisado.

Rbk6. El método según cualquiera de lo anterior o de lo siguiente, en donde los tensores de características se vectorizan.

Clasificación de la acción del agente

Cls1. Un método para clasificar compuestos biológicamente activos según cualquiera de lo siguiente que comprende una pluralidad de características celulares de una población de células expuestas a dichos compuestos, en donde dichas características se correlacionan con propiedades morfológicas cuantificadas simultáneamente por medio de proporciones de intensidad de la dispersión de la luz medida en dos o más ángulos.

Cls2. Un método según cualquiera de lo anterior o de lo siguiente, que comprende exponer un cultivo de dicha población de células a una pluralidad de compuestos y detectar la respuesta fisiológica de dicha población de células en presencia y ausencia de dicho compuesto.

Cls3. Un método según cualquiera de lo anterior o de lo siguiente, que comprende detectar la respuesta fisiológica de las células individuales muestreadas a partir de dicho cultivo.

Cls4. Un método según cualquiera de lo anterior o de lo siguiente, en donde la respuesta fisiológica es toxicidad mitocondrial, que se cuantifica en términos de pérdida del potencial de la membrana mitocondrial o una pérdida de la integridad de la membrana mitocondrial utilizando una o más etiquetas de fluorescencia seleccionadas de entre el grupo que consiste en JC-1, JC-9, JC-10, DiOC2(3), DiIC1(5), MITO TRACKER® ORANGE CMTMROS, MITO TRACKER® RED CMXROS.

Cls5. Un método según cualquiera de lo anterior o de lo siguiente, en donde la respuesta fisiológica es viabilidad celular total, que se cuantifica en términos de pérdida de la integridad de la membrana celular utilizando una o más etiquetas de fluorescencia.

Cls6. Un método según cualquiera de lo anterior o de lo siguiente, en donde las etiquetas de fluorescencia se seleccionan de entre los grupos que consisten en colorantes que entran en el interior celular lo que da lugar a una fluorescencia muy brillante (por ejemplo, YODURO de propidio y 7-aminoactinomicina D); colorantes que cruzan membranas de membranas celulares intactas y producen moléculas fluorescentes tras interactuar con enzimas intracelulares (por ejemplo, acetato de fluorosceína, CALCEÍNA AM, BCECF AM, diacetato de carboxieosina, CELLTRACKER™ GREEN CMFDA, acetato de clorometilo SNARF-1 y diacetato de ácido carboxílico OREGON GREEEN 488).

Cls7. Un método según cualquiera de lo anterior o de lo siguiente, que comprende además detectar al menos un descriptor fisiológico o fenotípico adicional de entre el grupo que consiste en concentración de glutatión, presencia de especies de oxígeno reactivo o de radicales libres.

Sistemas

Sys1. Un sistema para evaluar / comparar conjuntos de datos biológicos, que comprende un medio de almacenamiento legible por ordenador no transitorio que almacena un programa informático que, cuando se ejecuta en un ordenador, provoca que el ordenador realice cualquiera de los métodos anteriores o siguientes.

Sys2. Un sistema para evaluar / comparar conjuntos de datos biológicos, que comprende un medio de almacenamiento legible por ordenador no transitorio que almacena un programa informático que, cuando se ejecuta en un ordenador, provoca que el ordenador realice cualquiera de los métodos anteriores o siguientes para caracterizar una o más respuestas celulares a un agente, dicho método comprende:

medir por medio de citometría una pluralidad de parámetros fisiológicos, p, de células en la población que se exponen a una concentración de dicho agente;

calcular un conjunto de distancias entre poblaciones y controles para cada parámetro para la población de células en cada concentración; y

compilar un tensor o un conjunto de tensores para cada compuesto (donde los tensores contienen huellas de compuestos); y

comprimir los tensores por medio de descomposición de tensores para producir una huella abreviada de compuesto en forma de un vector.

Sys3. Un sistema informático para evaluar / comparar conjuntos de datos biológicos, que comprende un medio de almacenamiento legible por ordenador no transitorio que almacena un programa informático que, cuando se ejecuta en un ordenador, provoca que el ordenador realice un método para caracterizar una o más respuestas celulares a un agente, dicho método comprende:

(A) exponer primeras poblaciones de células a una pluralidad de concentraciones, c, de un primer agente; medir por medio de citometría una pluralidad de parámetros fisiológicos, p, de células en la población en cada concentración de dicho primer agente;

compilar uno o más tensores a partir de la descripción de dicho primer agente;

comprimir el tensor(es) de huella por medio de descomposición para obtener una huella abreviada de compuesto en forma de un vector;

(B) exponer segundas poblaciones de células a una segunda pluralidad de concentraciones, c2, de un segundo agente;

medir por medio de citometría una pluralidad de parámetros fisiológicos, p, de células en la población en cada concentración de dicho segundo agente;

compilar uno o más tensores a partir de la descripción de dicho segundo agente;

comprimir el tensor(es) de huella por medio de descomposición para obtener una huella abreviada de compuesto en forma de un vector y

(C) calcular una disimilitud entre la primera y la segunda huella abreviada para determinar la diferencia entre la respuesta de las células al primero y al segundo agente.

Sys4. Un sistema informático para evaluar / comparar conjuntos de datos biológicos, que comprende un medio de almacenamiento legible por ordenador no transitorio que almacena un programa informático que, cuando se ejecuta en un ordenador, provoca que el ordenador realice un método para caracterizar una o más respuestas celulares a un agente, dicho método comprende:

medir dos o más respuestas de fisiología celular para uno o más controles negativos, o uno o más positivos y para una o más concentraciones de un compuesto;

seleccionar subpoblaciones de células para los controles y la serie de concentraciones por medio de *gating* de las células en compartimientos particulares del ciclo celular y una clase morfológica particular;

calcular una disimilitud entre las distribuciones de mediciones celulares para cada control positivo y negativo y cada una de las concentraciones;

para determinar así la respuesta de las células al compuesto.

Conjuntos de datos y bases de datos

Dbs1. Un conjunto de datos que comprende valores para uno o más parámetros celulares.

Dbs2. Un conjunto de datos que comprende valores medidos para múltiples parámetros celulares para células expuestas a factores biológicos en ausencia o presencia de un agente candidato.

Dbs3. Una base de datos que comprende conjuntos de datos de huellas de compuestos en forma de tensores de respuesta de compuestos.

Dbs4. Una base de datos de perfiles de confianza para la identificación de perfiles de prueba, donde el perfil de confianza es un tensor de respuesta de un compuesto conocido *a priori* y bien caracterizado.

Dbs5. Los conjuntos de datos pueden ser conjuntos de datos de control, o conjuntos de datos de perfil que reflejan los cambios de parámetros de agentes conocidos. Para análisis de múltiples sistemas definidos por el contexto, se pueden concatenar los datos de salida a partir de múltiples sistemas.

Huella

Fpt1. Una huella de fármaco que comprende un valor de múltiples parámetros de respuesta celular.

Fpt2. Una huella de fármaco de un género de compuestos, que comprende un promedio de mediciones repetidas de tensores de estado.

Fpt2. Una huella de fármaco de un género de compuestos, que comprende un vector o una matriz producidos por medio de descomposición de tensores donde dicho tensor contenía mediciones de múltiples compuestos.

**REIVINDICACIONES**

1. Un método para caracterizar el efecto de un agente en células que comprende:

a) poner en contacto muestras de una población de células con una serie de concentraciones diferentes de un agente;

b) medir cada uno de dos o más parámetros fenotípicos en células individuales de una muestra respectiva para cada concentración;

c) obtener una distribución de los valores medidos para las células para cada parámetro fenotípico y cada concentración;

d) medir cada uno de los dos o más parámetros fenotípicos en células individuales de una muestra de control no expuesta al agente;

e) obtener una distribución de los valores medidos para las células para cada parámetro fenotípico para la muestra de control;

f) calcular, para cada parámetro fenotípico, disimilitudes entre la distribución obtenida en la etapa e) y cada una de las distribuciones obtenidas en la etapa c);

g) combinar las disimilitudes para formar un tensor de respuesta multiparamétrico; en donde el tensor de respuesta multiparamétrico es una huella característica del efecto del agente en las células que se puede comparar directamente con otras huellas generadas de la misma manera.

2. Un método según la reivindicación 1, en donde el método se utiliza para generar y almacenar huellas.

3. Un método según la reivindicación 1, en donde el método se utiliza para hacer comparaciones de huellas.

4. Un método según la reivindicación 3, en donde las comparaciones se utilizan para predecir efectos *in vivo*.

5. Un método según cualquiera de las reivindicaciones precedentes, en donde los parámetros fenotípicos se miden en las células individuales por medio de citometría celular.

6. Un método según cualquiera de las reivindicaciones precedentes, en donde los parámetros fenotípicos incluyen dos o más de lo siguiente viabilidad celular, fase del ciclo celular, integridad de la membrana mitocondrial, toxicidad mitocondrial, concentración de glutatión, especies reactivas de oxígeno, especies reductoras, permeabilidad de la membrana citoplasmática, lesión del ADN, un constituyente de la vía de señalización de respuesta al estrés, un constituyente de la vía de la respuesta inflamatoria, un constituyente de la vía de apotosis y una peroxidasa lipídica.

7. Un método según cualquiera de las reivindicaciones precedentes, en donde los parámetros fenotípicos incluyen uno o más de NFκB, SAPK, P13K, AKT, una familia de proteínas Bcl-1, GSk3B y quinasa S6 ribosomal.

8. Un método según cualquiera de las reivindicaciones precedentes, en donde uno de los parámetros fenotípicos es ciclo celular.

9. Un método según cualquiera de las reivindicaciones precedentes, en donde cada población de células se marca funcionalmente con una pluralidad de colorantes fluorescentes y los parámetros fenotípicos se detectan y cuantifican en términos de señal(es) de emisión espectral que se genera cuando dichas poblaciones de células marcadas se someten a análisis citométrico.

10. El método según cualquiera de las reivindicaciones precedentes, en donde uno de los parámetros fenotípicos es ciclo celular y se cuantifica en términos de uno o más de: HOECHST 33342, DRAQ5, yoduro YO-PRO-1, DAPI, CYTRAK ORANGE, ciclina A, ciclina B, ciclina E, o proteína histónica fosforilada.

11. Un método según cualquiera de las reivindicaciones precedentes, que comprende:

poner en contacto las muestras con $n$ concentraciones diferentes de un agente, donde $i \in \{2,...,n\}$, y obtener al menos una muestra de control, $\kappa$, correspondiente a $m$ condiciones de control diferentes, donde $m$ es al menos 1;

medir cada uno de $p$ parámetros fenotípicos, $\psi$, en células individuales de cada muestra, donde $p$ es al menos 2 y donde $\psi_p$ denota la medición de cada parámetro fenotípico particular,

obtener una distribución $C_{\kappa}^{(\psi_p)}$ de los valores medidos para cada condición de control *κ* para cada parámetro fenotípico ψp y obtener una distribución $S_i^{(\psi_p)}$ de los valores medidos para cada condición de concentración *i* para cada parámetro fenotípico ψp,

calcular disimilitudes calculando, para cada parámetro fenotípico $\psi_p$, distancias *d* entre pares entre las distribuciones de cada condición de control $C_{\kappa}^{(\psi_p)}$ y cada condición de concentración $S_i^{(\psi_p)}$,

donde

$$d_{\kappa,i}^{(\psi_p)} = D\left(C_{\kappa}^{(\psi_p)},\ S_i^{(\psi_p)}\right)$$

y *D* es una función de distancia;

en donde el tensor de respuesta multiparamétrico es un tensor de respuesta multiparamétrico A, que se representa por

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_{[\kappa_1,\psi_1]} & \cdots & \mathbf{a}_{[\kappa_1,\psi_p]} \\ \vdots & \ddots & \vdots \\ \mathbf{a}_{[\kappa_m,\psi_1]} & \cdots & \mathbf{a}_{[\kappa_m,\psi_p]} \end{bmatrix}$$

y en donde cada uno de $\mathbf{a}_{[\kappa,\psi]}$ comprende todas las distancias entre pares calculadas para un parámetro fenotípico y uno de control:

$$\mathbf{a}_{[\kappa,\psi]} = \begin{bmatrix} d_{\kappa,1}^{(\psi_p)} = D\left(C_{\kappa}^{(\psi_p)},\ S_1^{(\psi_p)}\right) \\ d_{\kappa,2}^{(\psi_p)} = D\left(C_{\kappa}^{(\psi_p)},\ S_2^{(\psi_p)}\right) \\ \vdots \\ d_{\kappa,n}^{(\psi_p)} = D\left(C_{\kappa}^{(\psi_p)},\ S_n^{(\psi_p)}\right) \end{bmatrix}.$$

12. Un método según la reivindicación 11, en donde el tensor multiparamétrico **A** se descompone en uno de:

(a) un conjunto de tensores de rango uno $\mathbf{A}^{(i)}$, en donde

$$\mathbf{A} \approx \left[\!\left[\lambda; \mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots, \mathbf{A}^{(n)}\right]\!\right] = \sum_{r=1}^{R} \lambda_r \mathbf{a}_r^{(1)} \circ \mathbf{a}_r^{(2)} \circ \cdots \circ \mathbf{a}_r^{(p)},$$

y

(b) un tensor núcleo **G** multiplicado por una matriz $\mathbf{M}^{(j)}$ a lo largo con cada modo:

$$\mathbf{A} \approx \left[\!\left[\mathbf{G}; \mathbf{M}^{(1)}, \mathbf{M}^{(2)}, \ldots, \mathbf{M}^{(p)}\right]\!\right] = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_p=1}^{R_p} g_{r_1 r_2 \ldots r_p} \mathbf{m}_{r_1}^{(1)} \circ \mathbf{m}_{r_2}^{(2)} \circ \cdots \circ \mathbf{m}_{r_N}^{(p)}$$

13. Un método según cualquiera de las reivindicaciones 11-12, en donde la disimilitud entre un tensor de respuesta multiparamétrico **A** (primera huella) y un tensor de respuesta multiparamétrico **B** (segunda huella) se calcula según la siguiente fórmula:

$$\mathbf{w}_{\mathbf{A},\mathbf{B}} = D(\mathbf{A},\mathbf{B}) = d\left(\mathbf{a}_{:jk},\mathbf{b}_{:jk}\right),$$

donde D es la disimilitud entre los sensores **A** y **B**, *d* es una función de distancia que compara las fibras de modo 1 de cada uno de los tensores, y **w** es un indicativo de vector de distancia de la disimilitud entre la primera y la segunda huella.

14. Un método para predecir los efectos *in vivo* de un agente, que comprende caracterizar el efecto del agente según el método de la reivindicación 11,

comprimir el tensor de respuesta multiparamétrico **A** para reducir su complejidad por medio de un método de compresión de baja pérdida;

calcular la diferencia entre el tensor comprimido y los tensores calculados de manera comparable para un banco de compuestos de efectos *in vivo* conocidos;

a partir de los valores de las diferencias, predecir los efectos *in vivo* del agente.

15. Un método para producir un banco de huellas características de los efectos de agentes en células, que comprende generar para cada uno de una pluralidad de agentes un tensor de respuesta multiparamétrico por medio del método de cualquiera de las reivindicaciones 1-13, en donde cada tensor de respuesta multiparamétrico es una huella característica de los efectos del agente en células, y

almacenar las huellas en un medio legible por ordenador de tal manera que se puedan comparar entre sí y con las huellas de otros agentes producidos de la misma manera,

de esta manera, se forma una base de datos de huellas características de los efectos de los agentes en células.

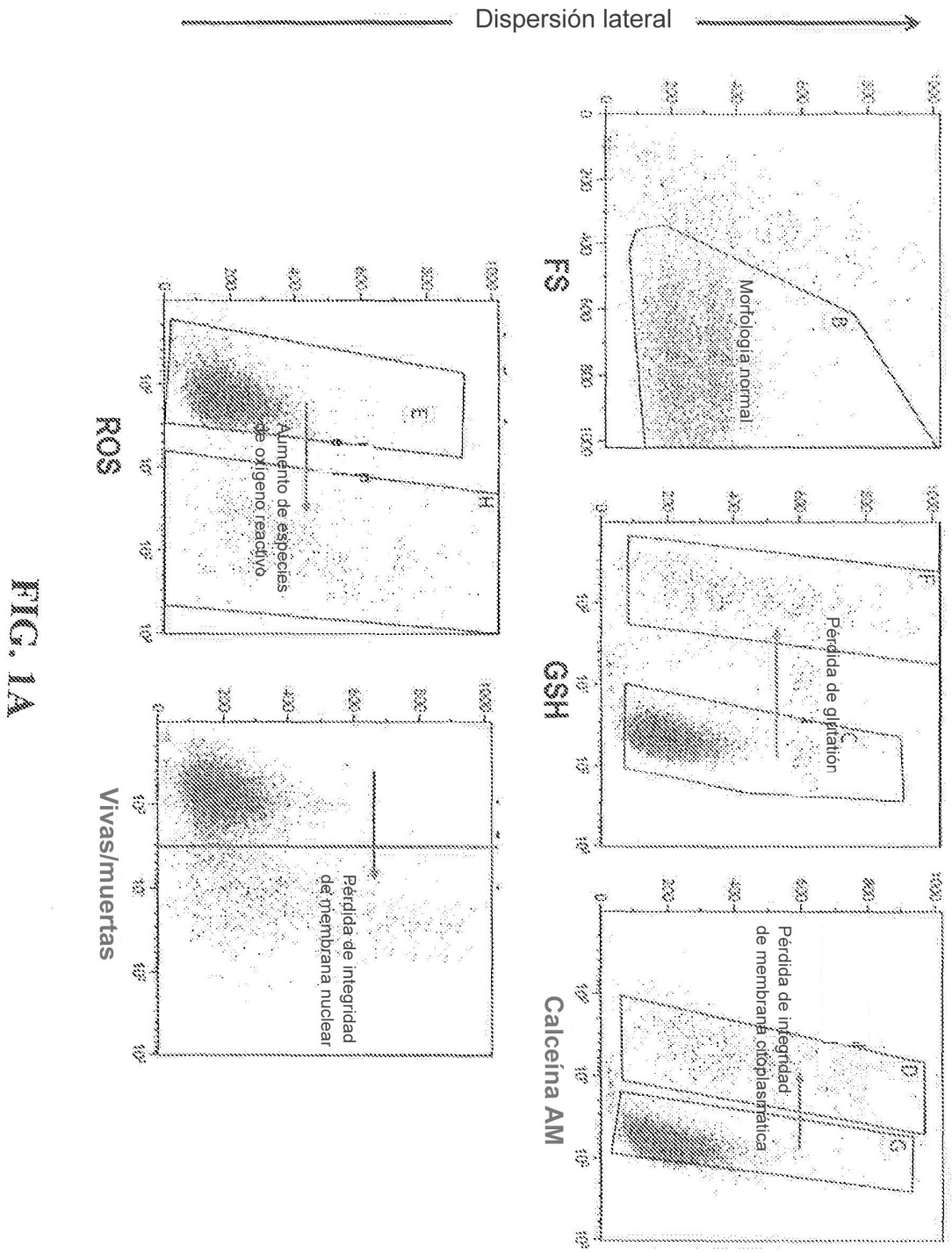

Dispersión lateral
FS
Morfología normal
B
GSH
Pérdida de glutation
F
C
Calceína AM
Pérdida de integridad
de membrana citoplasmática
D
G
ROS
Aumento de especies
de oxígeno reactivo
E
H
Vivas/muertas
Pérdida de integridad
de membrana nuclear

FIG. 1A

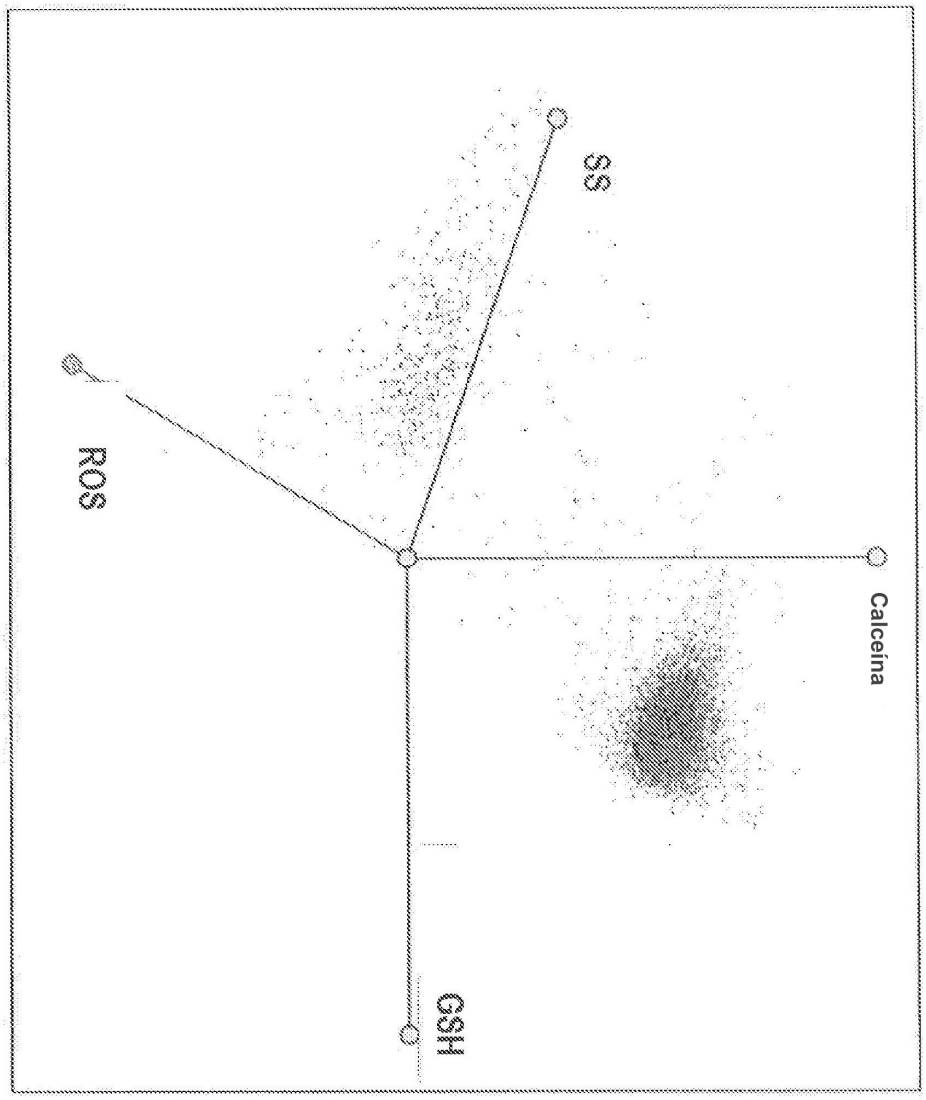



FIG. 1B

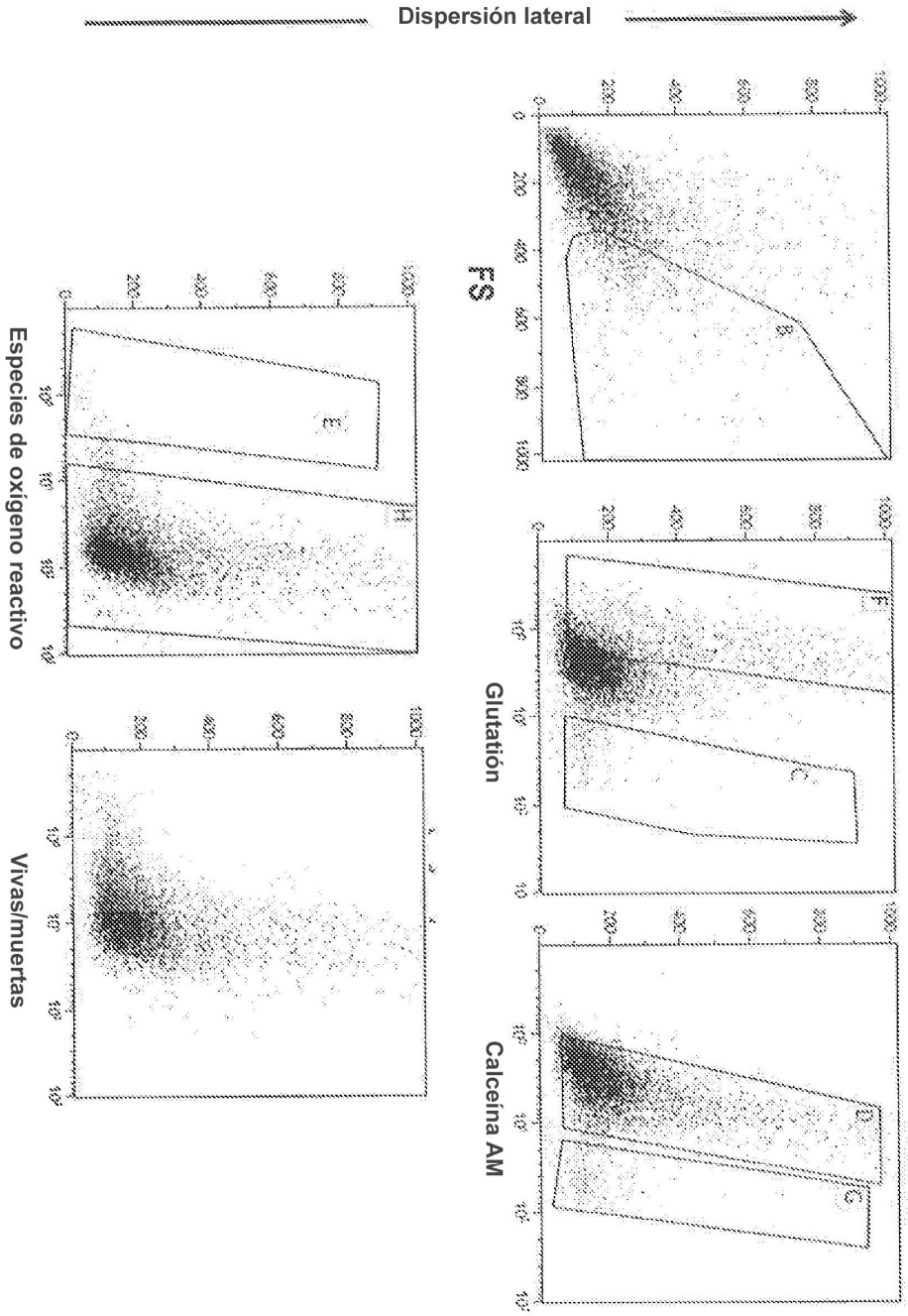

Dispersión lateral
FS
Especies de oxígeno reactivo
Glutatión
Vivas/muertas
Calceína AM

FIG. 2

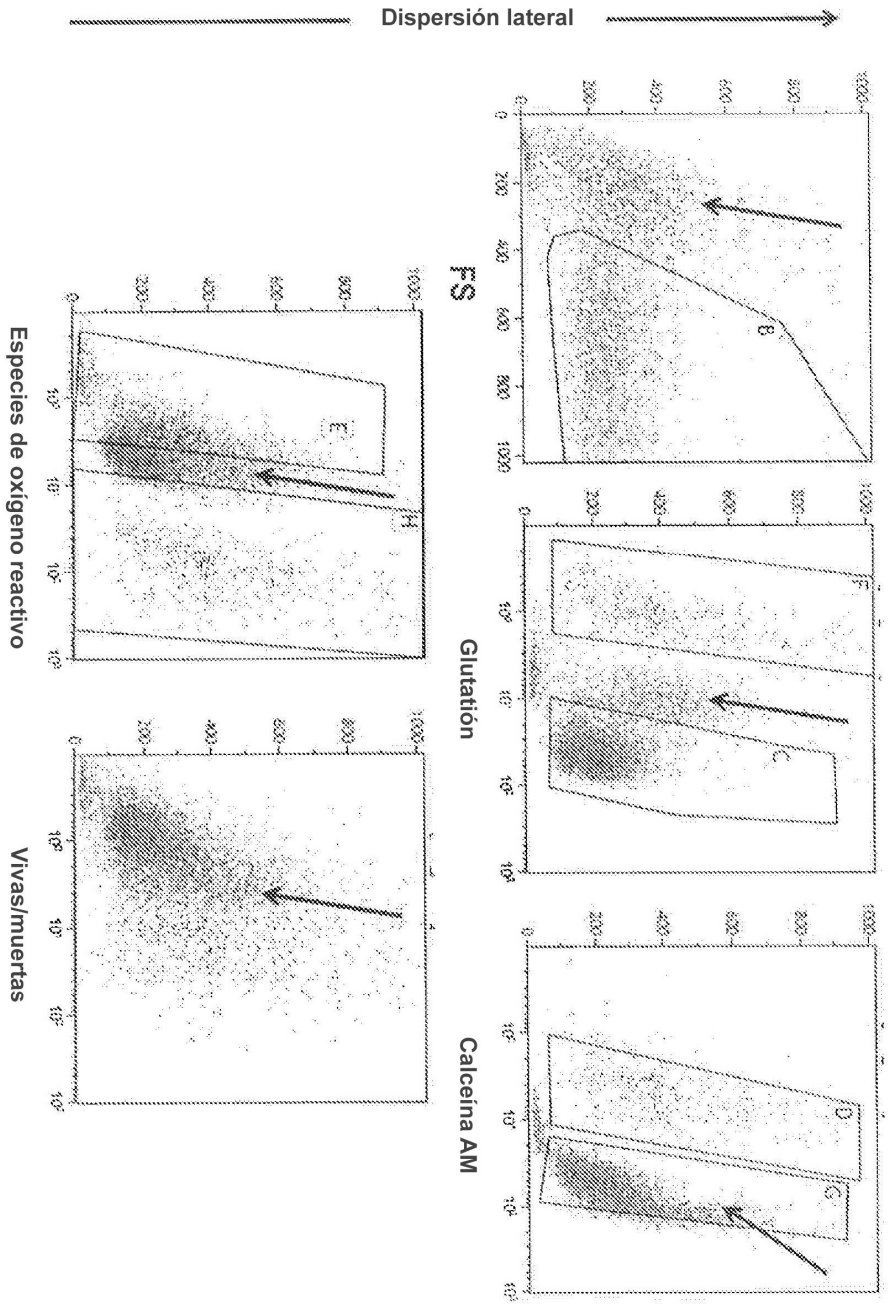

Dispersión lateral
FS
Especies de oxígeno reactivo
Glutatión
Vivas/muertas
Calceína AM
A
B
C
D
E
F
G
H

FIG. 3

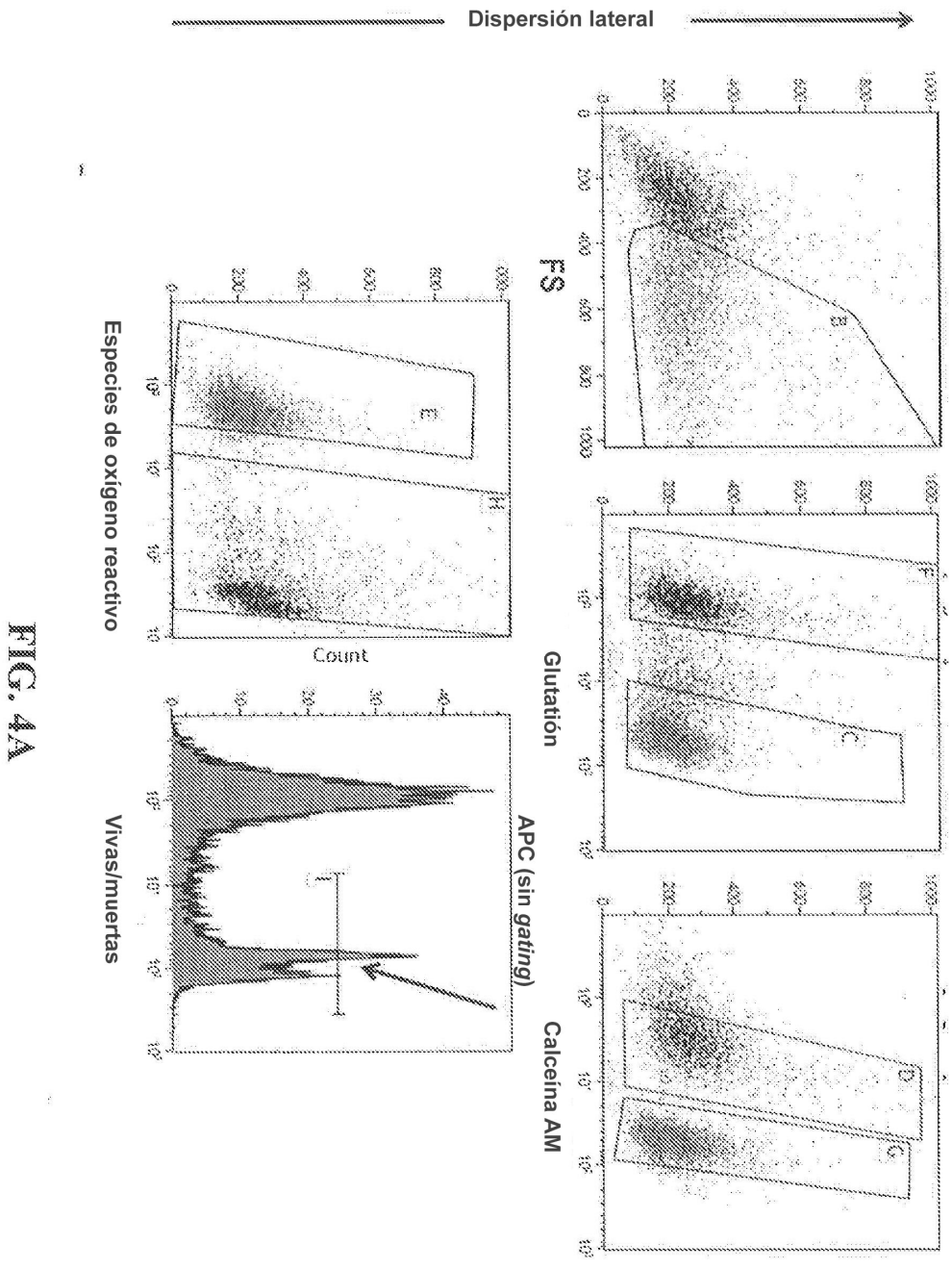

Dispersión lateral
FS
Especies de oxígeno reactivo
Count
Glutatión
APC (sin gating)
Vivas/muertas
Calceína AM

FIG. 4A

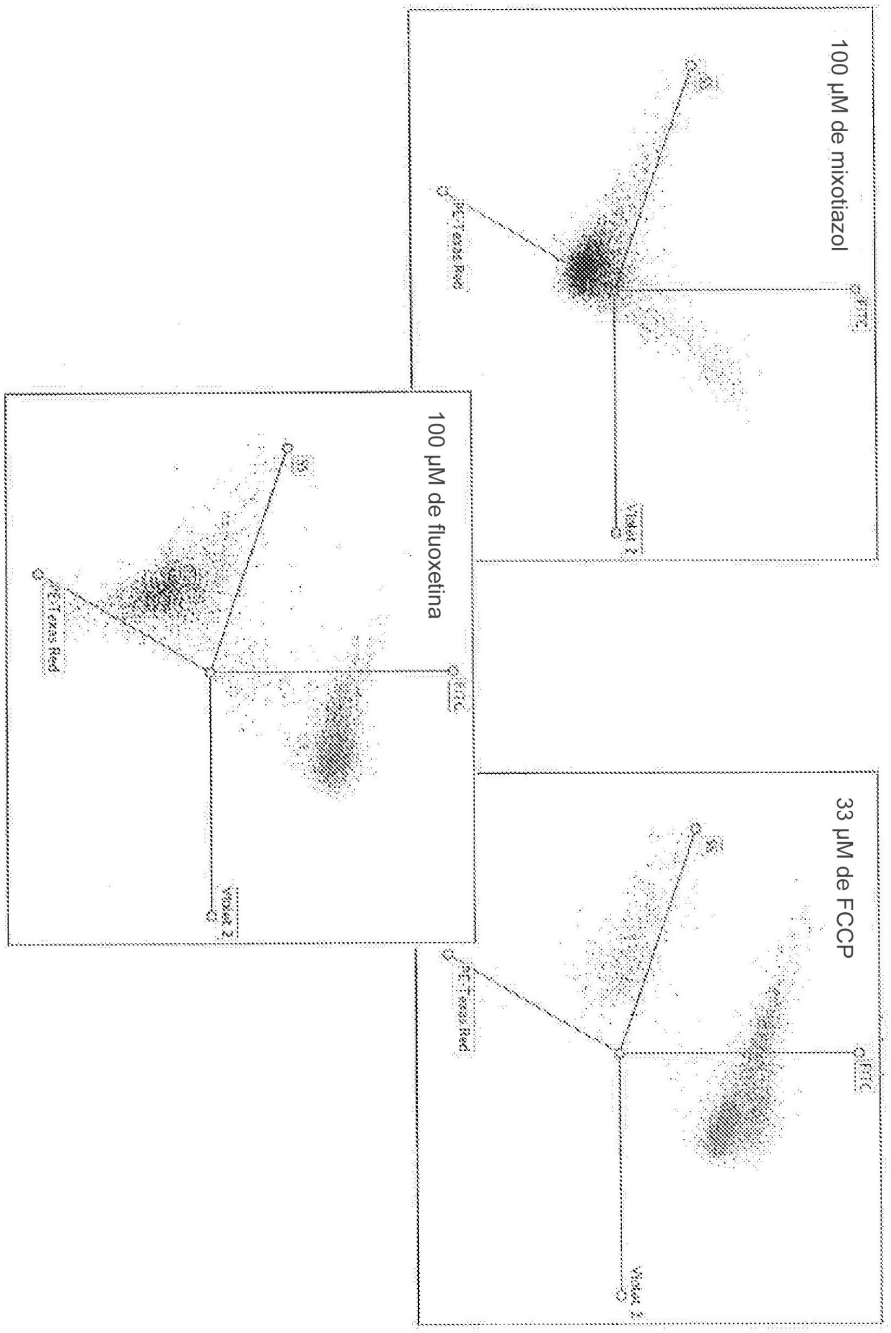
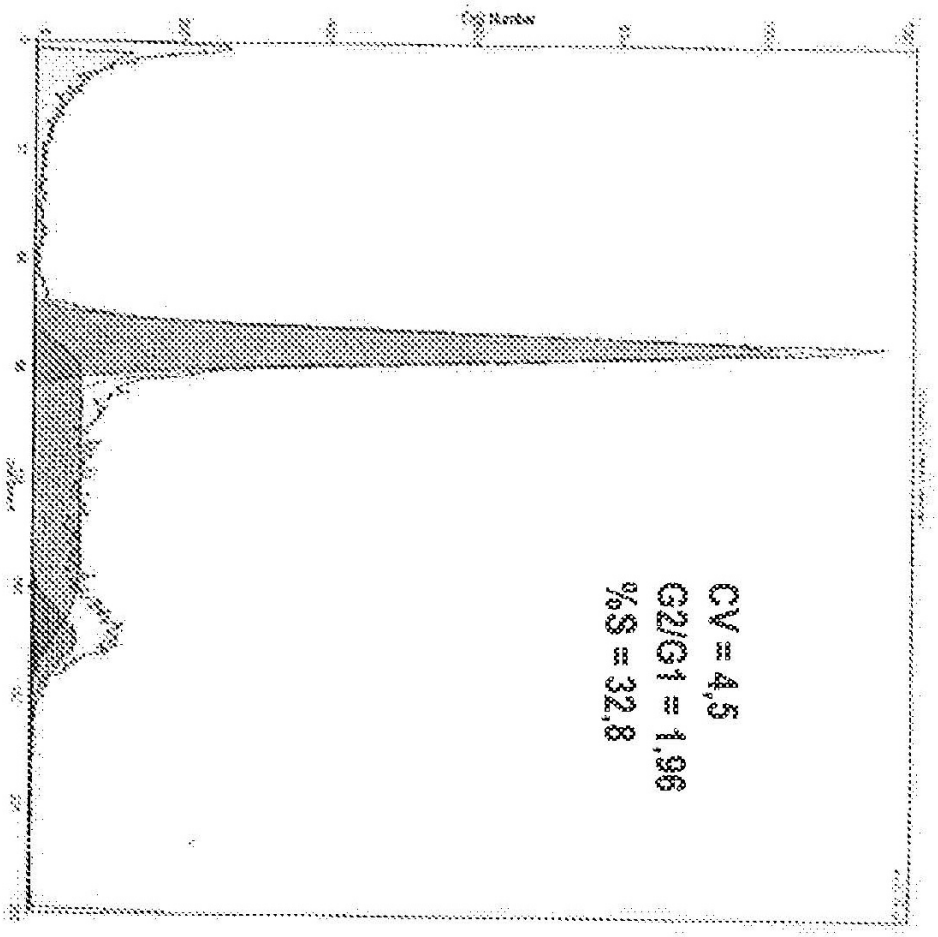



FIG. 4D

FIG. 5



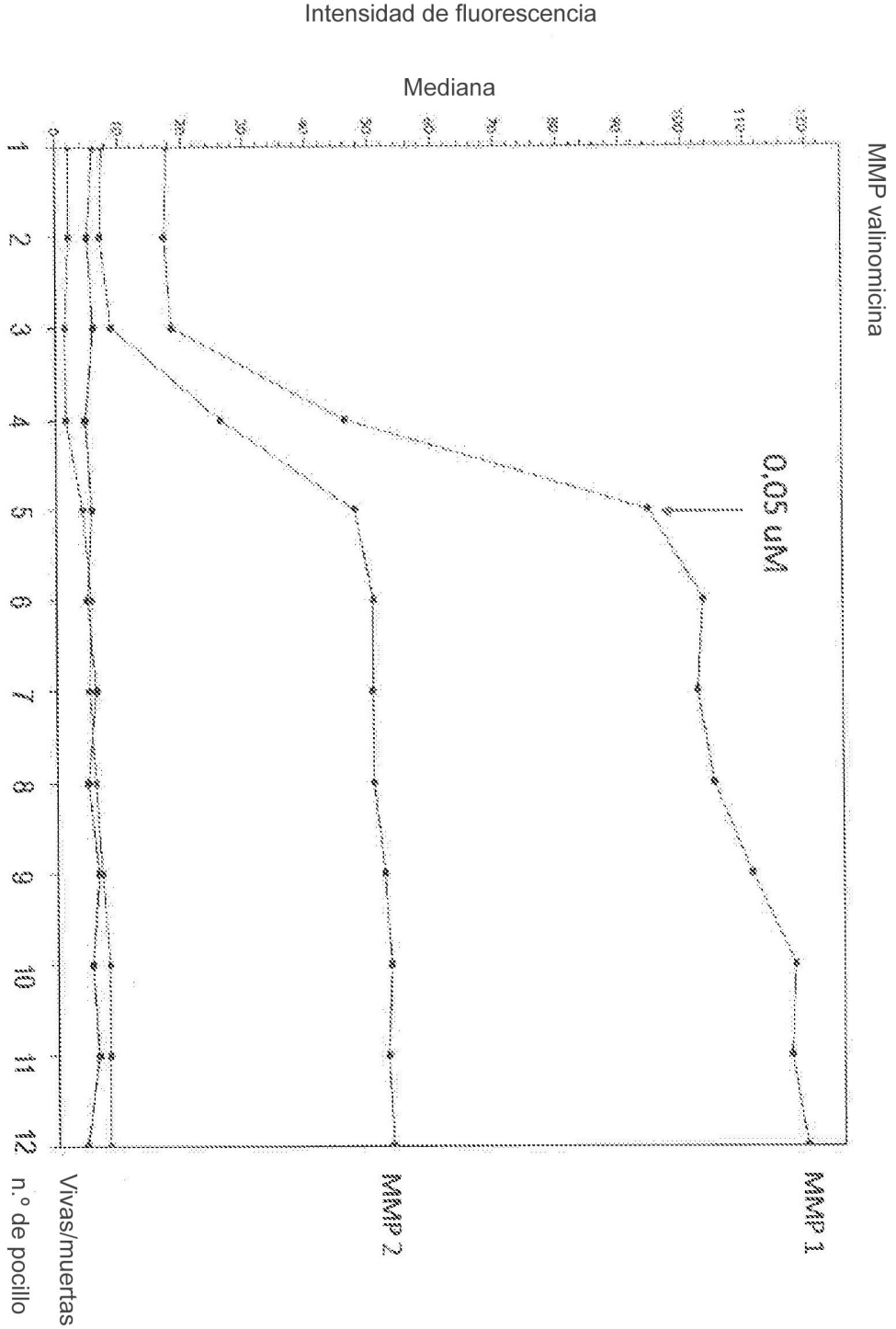

Intensidad de fluorescencia
Mediana
MMP valinomicina
0,05 uM
MMP 1
MMP 2
Vivas/muertas
n.º de pocillo
1
2
3
4
5
6
7
8
9
10
11
12

FIG. 6A

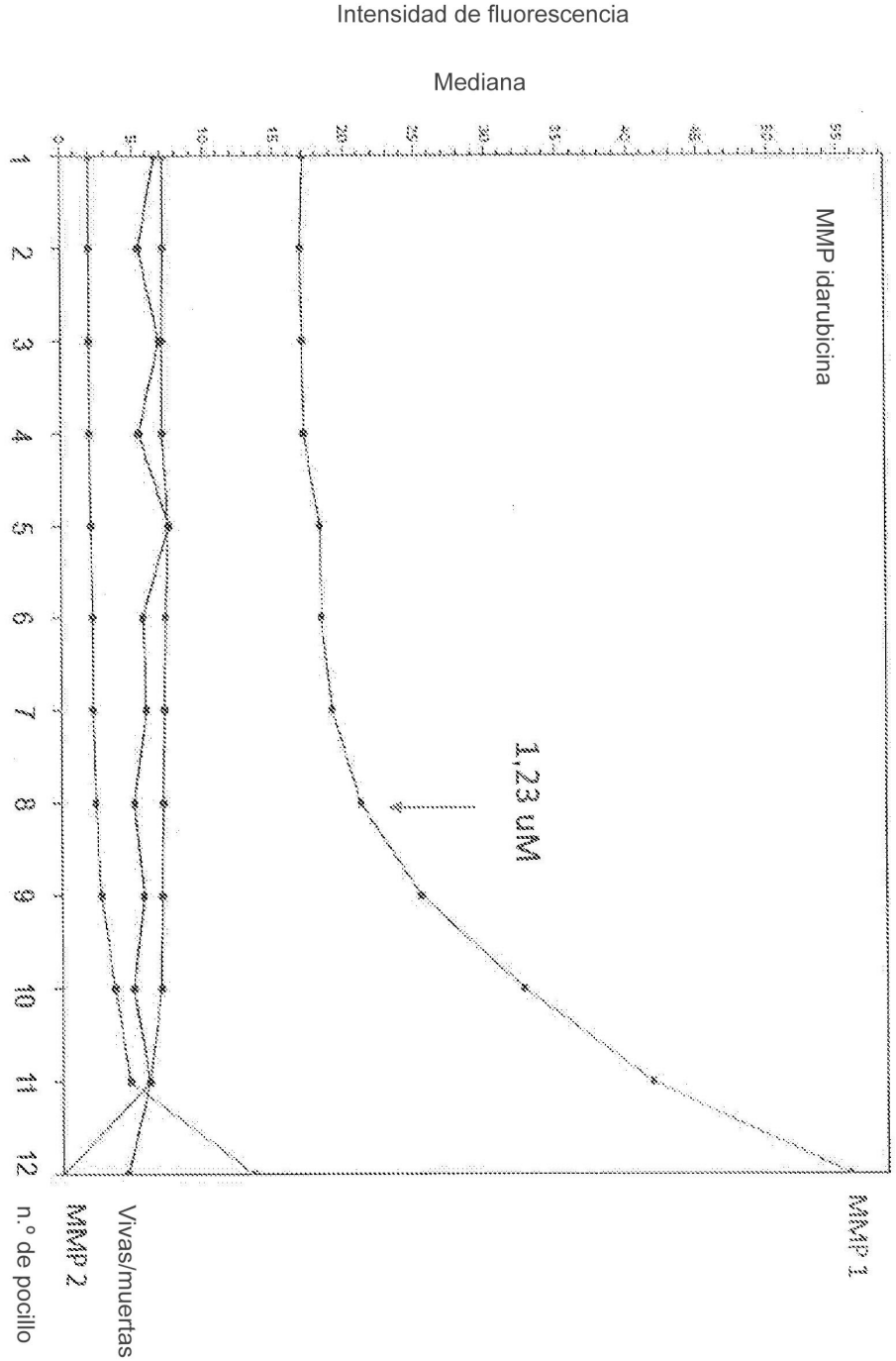
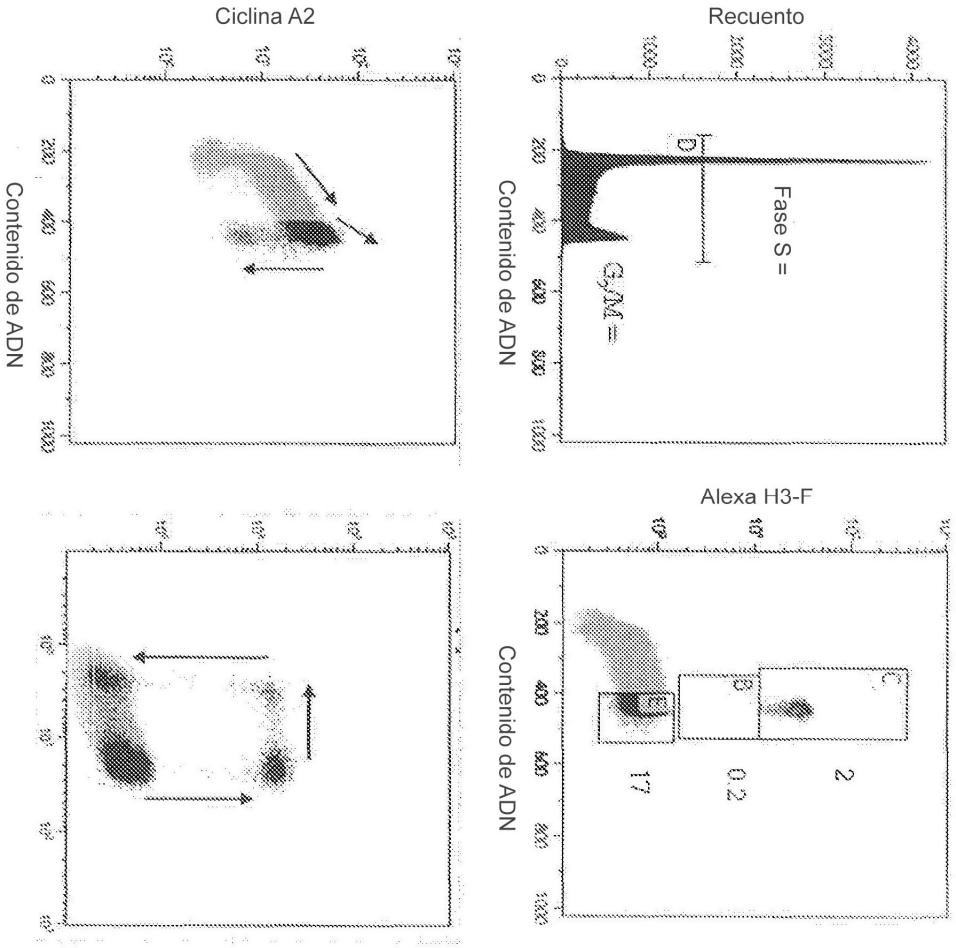
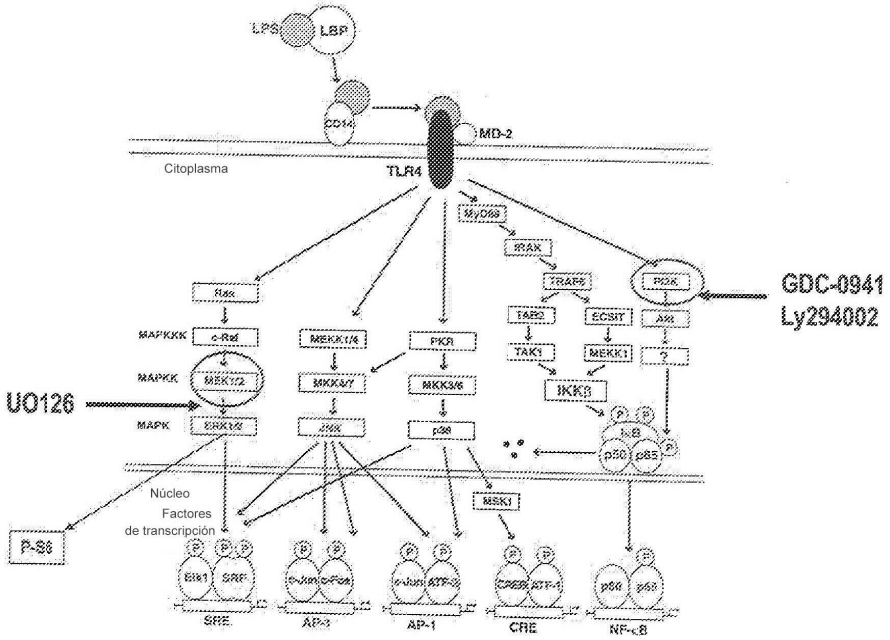

Intensidad de fluorescencia
Mediana
MMP idarubicina
1,23 uM
MMP 1
Vivas/muertas
MMP 2
n.º de pocillo

FIG. 6B

FIG. 7


Recuento
Contenido de ADN
D
Fase S =
G1/M =
Alexa H3-F
Contenido de ADN
A
B
C
17
0.2
2
Ciclina A2
Contenido de ADN

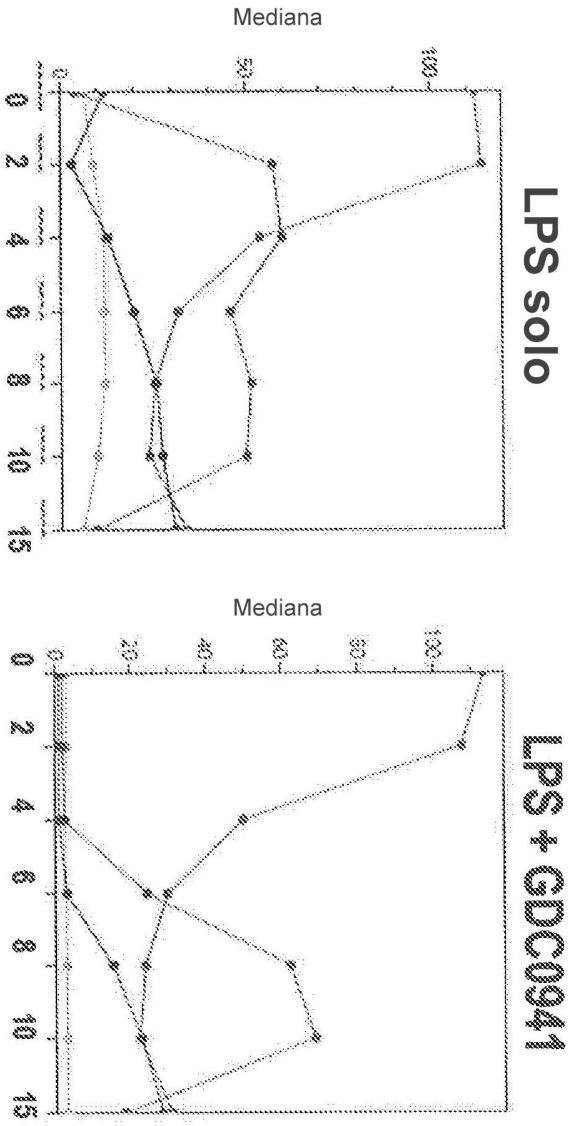
## Inhibidores y vías de transducción de señal inducida por LPS




De Guha y Mackman, Cellular Signaling 13;85 2001

FIG. 8

FIG. 9



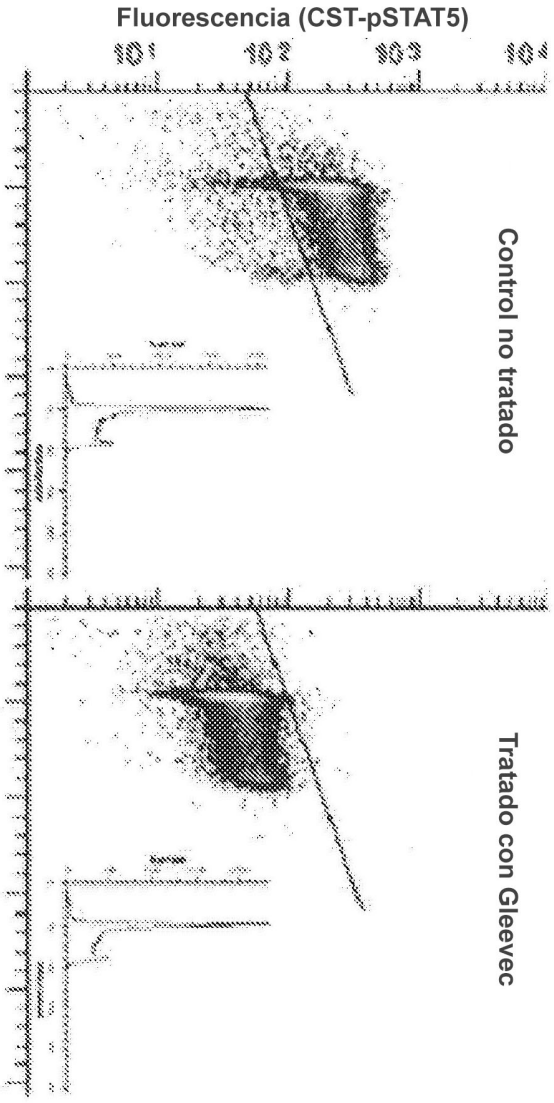

Fluorescencia (CST-pSTAT5)
10 1
10 2
10 3
10 4
Control no tratado
Tratado con Gleevec

FIG. 10

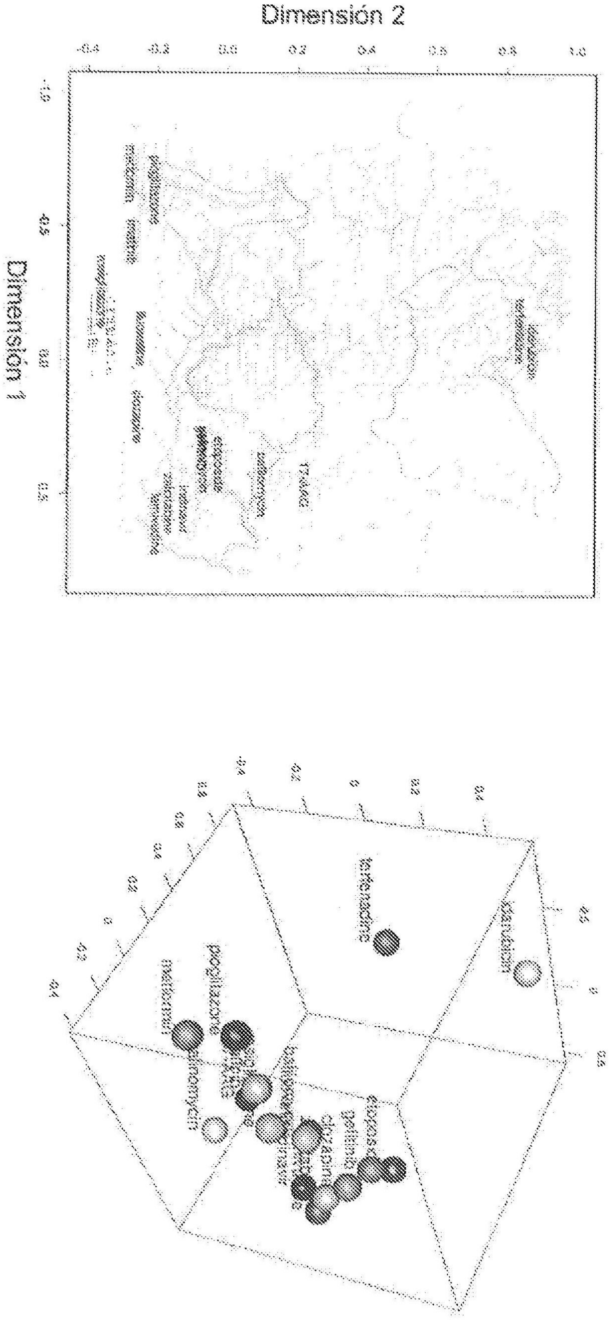



FIG. 11

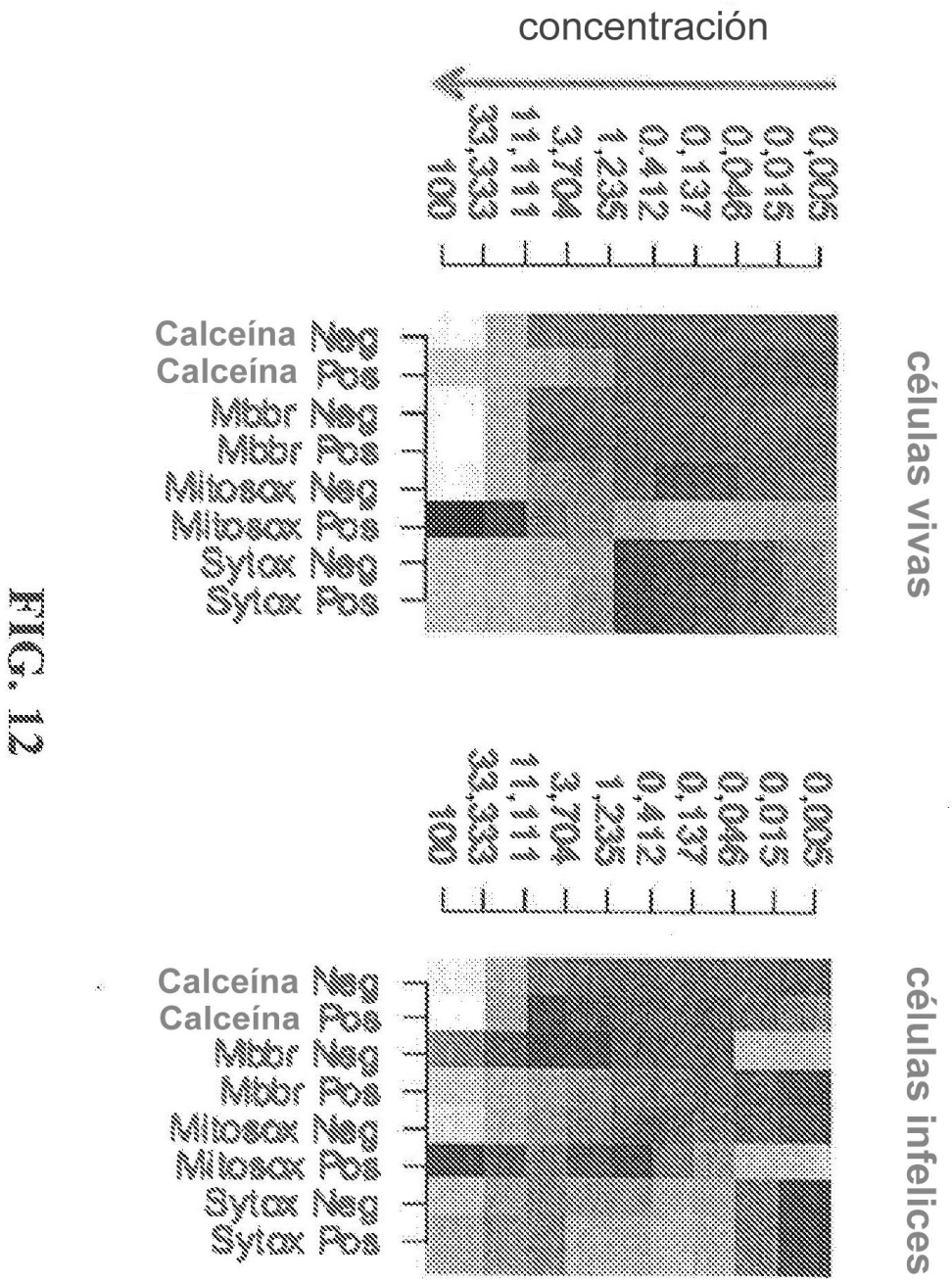
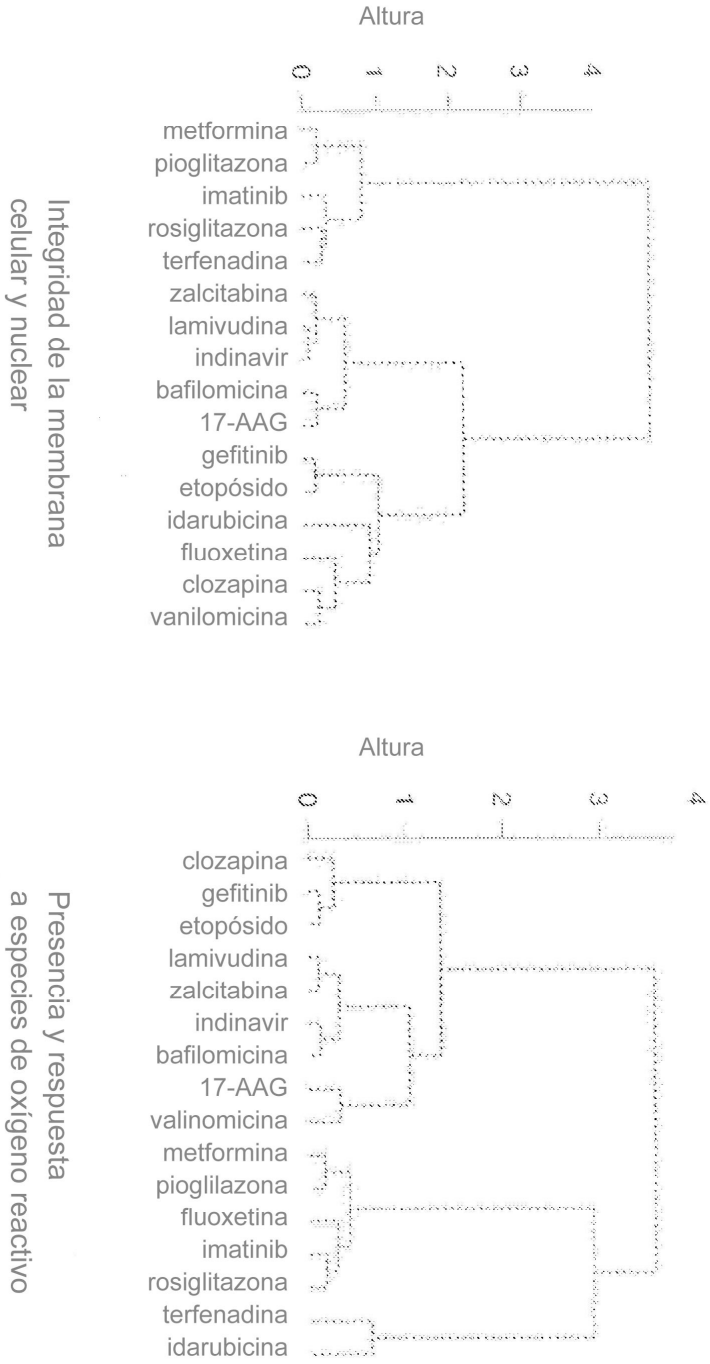
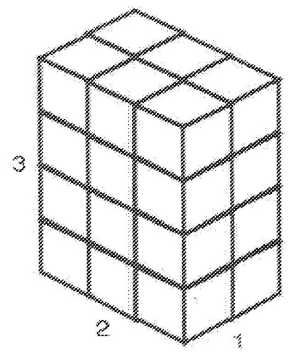
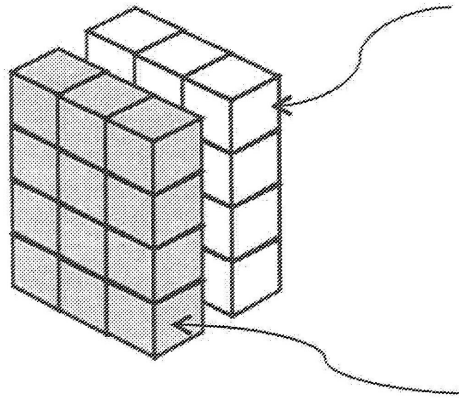

concentración
100
33,333
11,111
3,704
1,235
0,412
0,137
0,046
0,015
0,005
0,0
células vivas
Calceína Neg
Calceína Pos
Mtbr Neg
Mtbr Pos
Mtbsax Neg
Mtbsax Pos
Sytox Neg
Sytox Pos
células infelices

FIG. 12

FIG. 13


Integridad de la membrana celular y nuclear
Altura
0
1
2
3
4
metformina
pioglitazona
imatinib
rosiglitazona
terfenadina
zalcitabina
lamivudina
indinavir
bafilomicina
17-AAG
gefitinib
etopósido
idarubicina
fluoxetina
clozapina
vanilomicina
Presencia y respuesta a especies de oxígeno reactivo
Altura
0
1
2
3
4
clozapina
gefitinib
etopósido
lamivudina
zalcitabina
indinavir
bafilomicina
17-AAG
valinomicina
metformina
pioglilazona
fluoxetina
imatinib
rosiglitazona
terfenadina
idarubicina



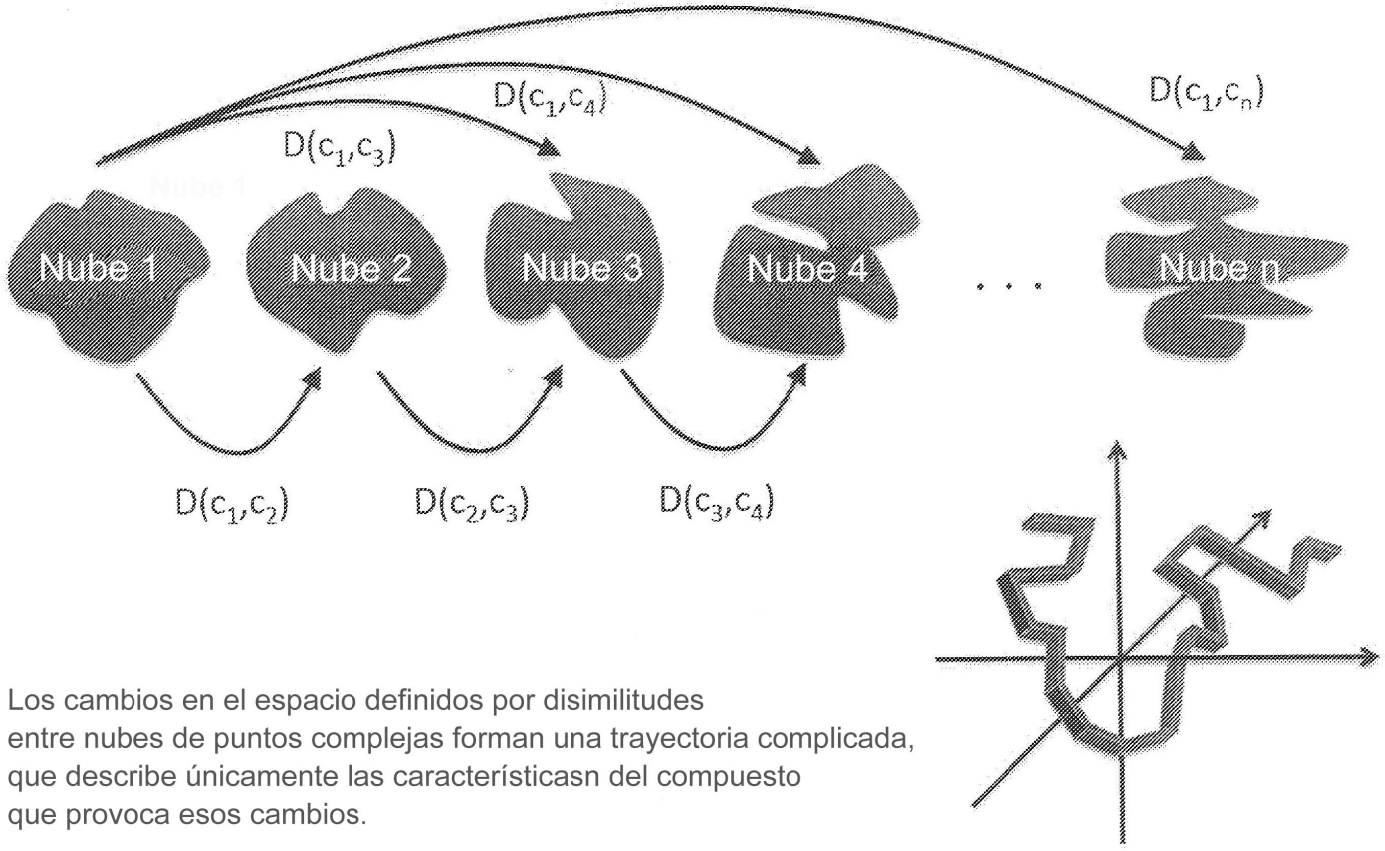



Disimilitud entre distribución de señales de fluorescencia que indican respuesta fisiológica $X_3$ a partir de células de tipo de morfología $Y_2$ a un nivel de concentración $Z_1$ y un control relevante.

Disimilitud entre distribución de señales de fluorescencia que indican respuesta fisiológica X3 a partir de células de tipo de morfología Y1 a un nivel de concentración Z4 y un control relevante.

Eje 1: Morfología celular
Eje 2: Etiqueta de fluorescencia funcional
Eje 3: Concentración de agente perturbador

FIG. 14




Los cambios en el espacio definidos por disimilitudes entre nubes de puntos complejas forman una trayectoria complicada, que describe únicamente las característicasn del compuesto que provoca esos cambios.

FIG. 15

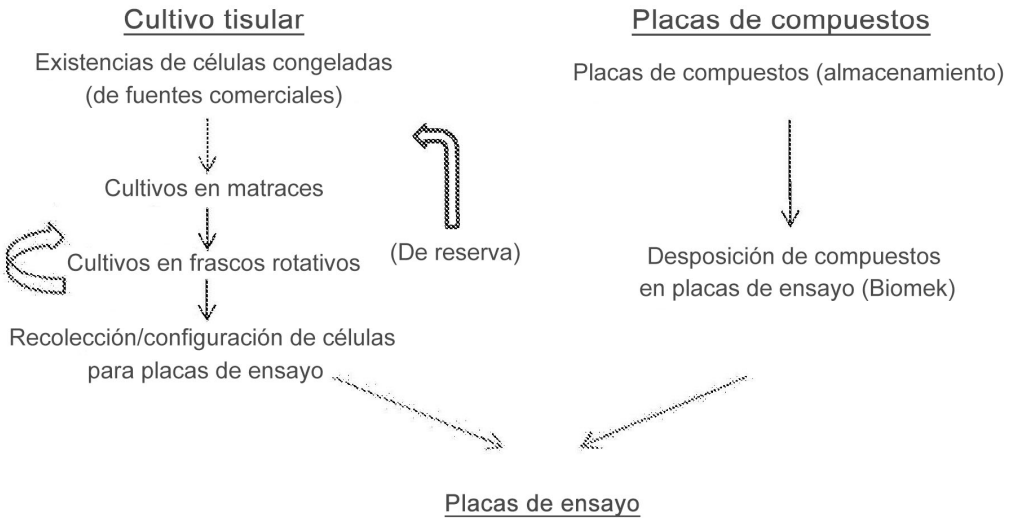

Ensayo de fisiología celular
Cultivo tisular
Existencias de células congeladas
(de fuentes comerciales)
Cultivos en matraces
Cultivos en frascos rotativos
Recolección/configuración de células
para placas de ensayo
(De reserva)
Placas de compuestos
Placas de compuestos (almacenamiento)
Desposición de compuestos
en placas de ensayo (Biomek)
Placas de ensayo
1. Depositar células en las placas de ensayo de compuestos formateados
2. Incubar (de 4 a ¿? horas) a 37 ºC en incubadora de cultivo tisular
3. Centrifugar las células, retirar sobrenadante
4. Añadir mezcla de colorantes en todos los pocillos (Biomek)
5. Incubar durante 15-30 min a 37 ºC en incubadora de cultivo tisular
Análisis celulares
1. Citometría de flujo (HyperCyt/Cian)
2. Obtención/almacenamiento de datos
3. Análisis
a. Análisis de datos manual de los pocillos (Kaluza)
b. Análisis aumotatizado/semiautomatizado
I. Análisis de garantía y control de calidad
II. Curvas de respuesta
III. Agrupamiento (no supervisado)

FIG. 16

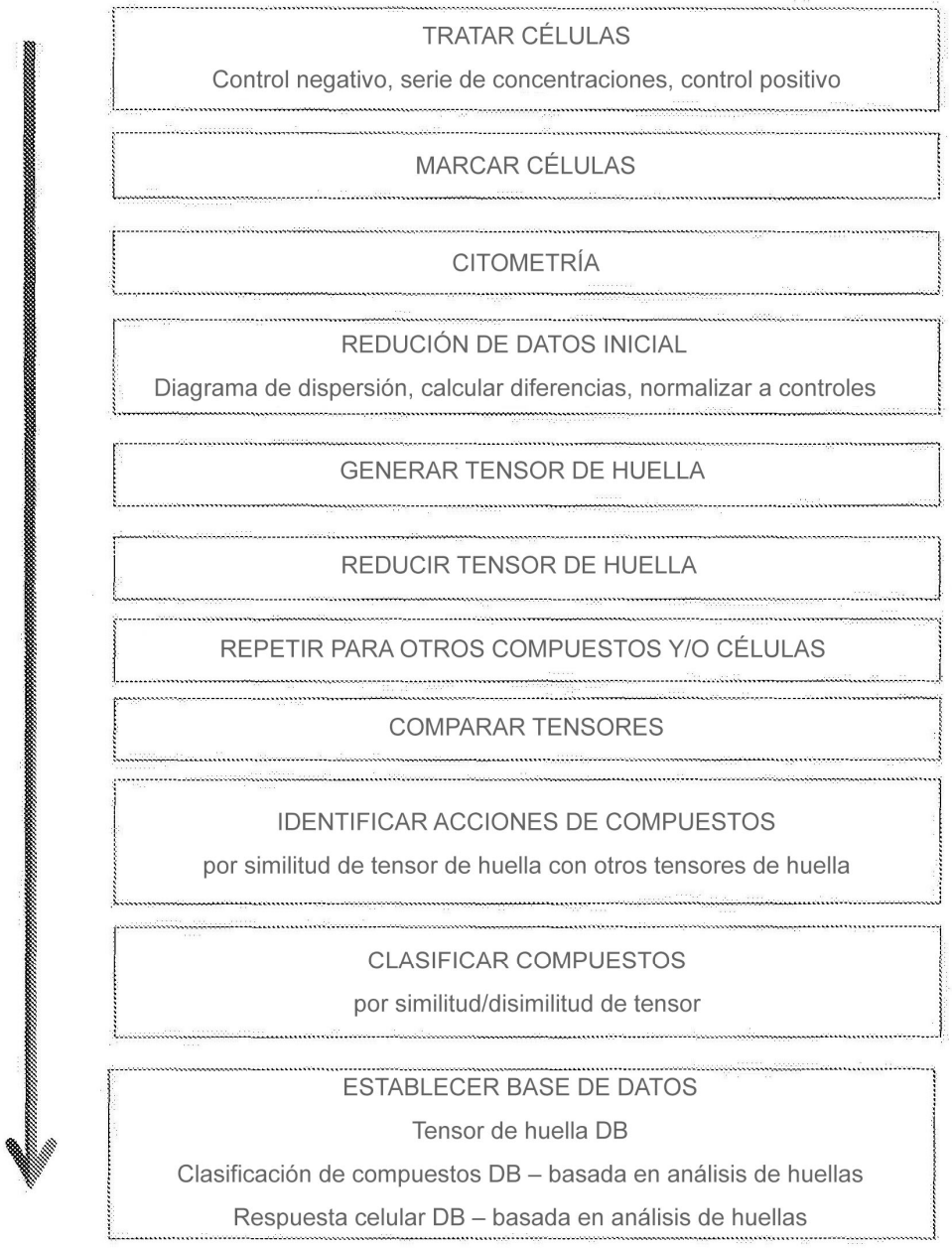
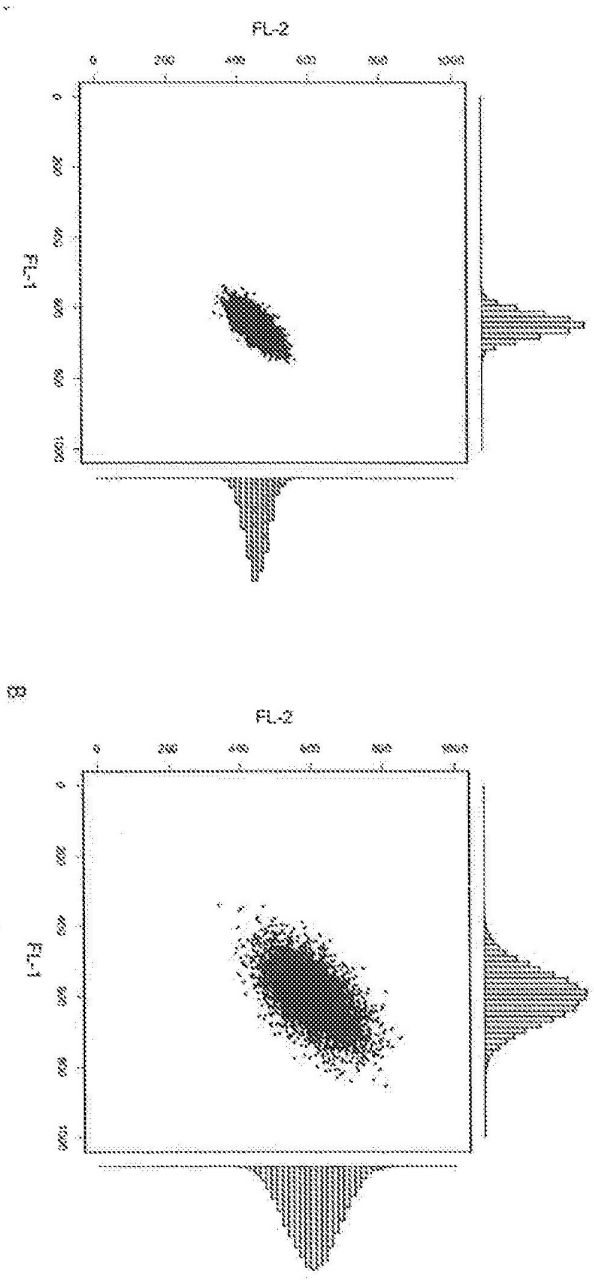

TRATAR CÉLULAS
Control negativo, serie de concentraciones, control positivo
MARCAR CÉLULAS
CITOMETRÍA
REDUCIÓN DE DATOS INICIAL
Diagrama de dispersión, calcular diferencias, normalizar a controles
GENERAR TENSOR DE HUELLA
REDUCIR TENSOR DE HUELLA
REPETIR PARA OTROS COMPUESTOS Y/O CÉLULAS
COMPARAR TENSORES
IDENTIFICAR ACCIONES DE COMPUESTOS
por similitud de tensor de huella con otros tensores de huella
CLASIFICAR COMPUESTOS
por similitud/disimilitud de tensor
ESTABLECER BASE DE DATOS
Tensor de huella DB
Clasificación de compuestos DB – basada en análisis de huellas
Respuesta celular DB – basada en análisis de huellas

FIG. 17

FIG. 18



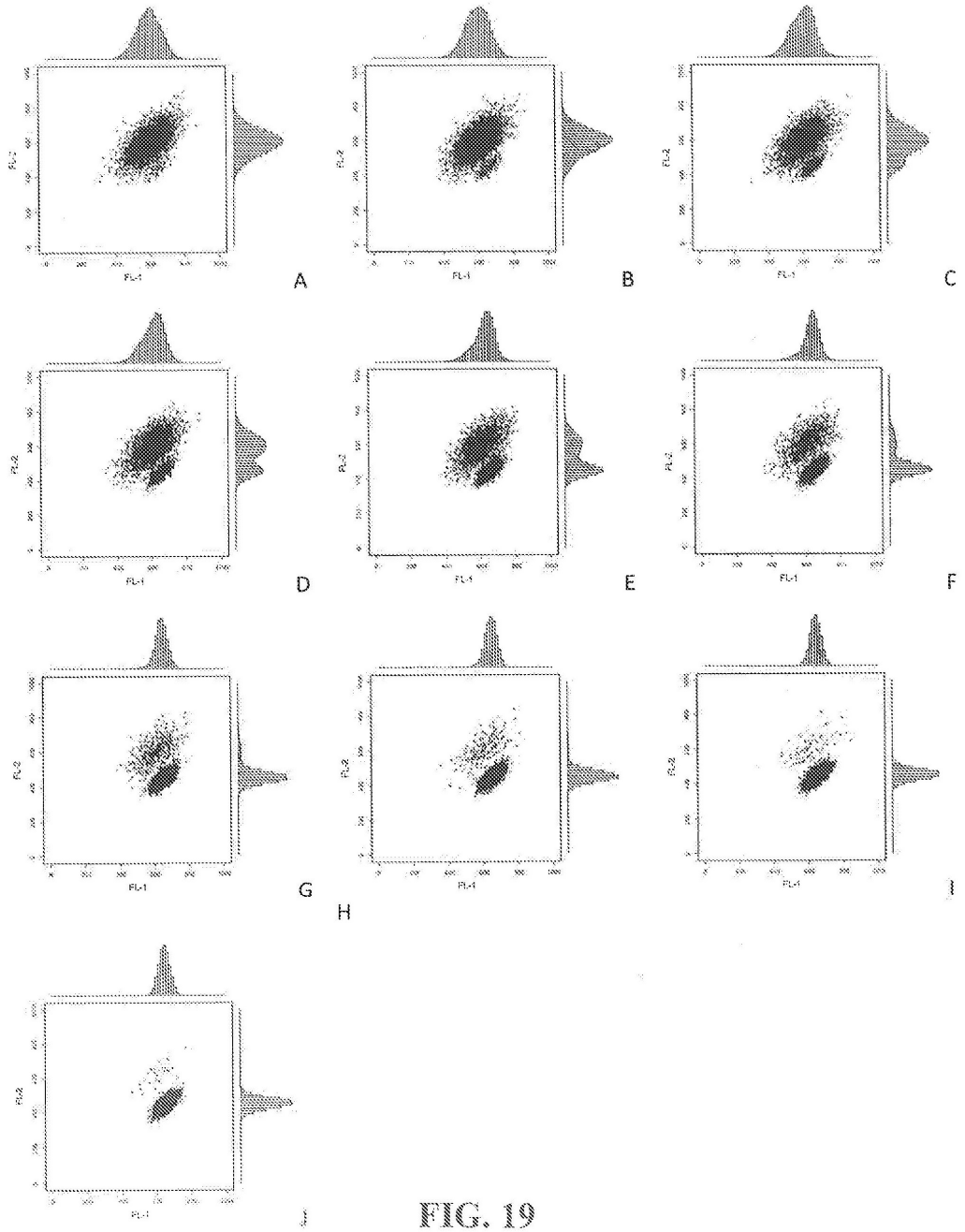



FIG. 19

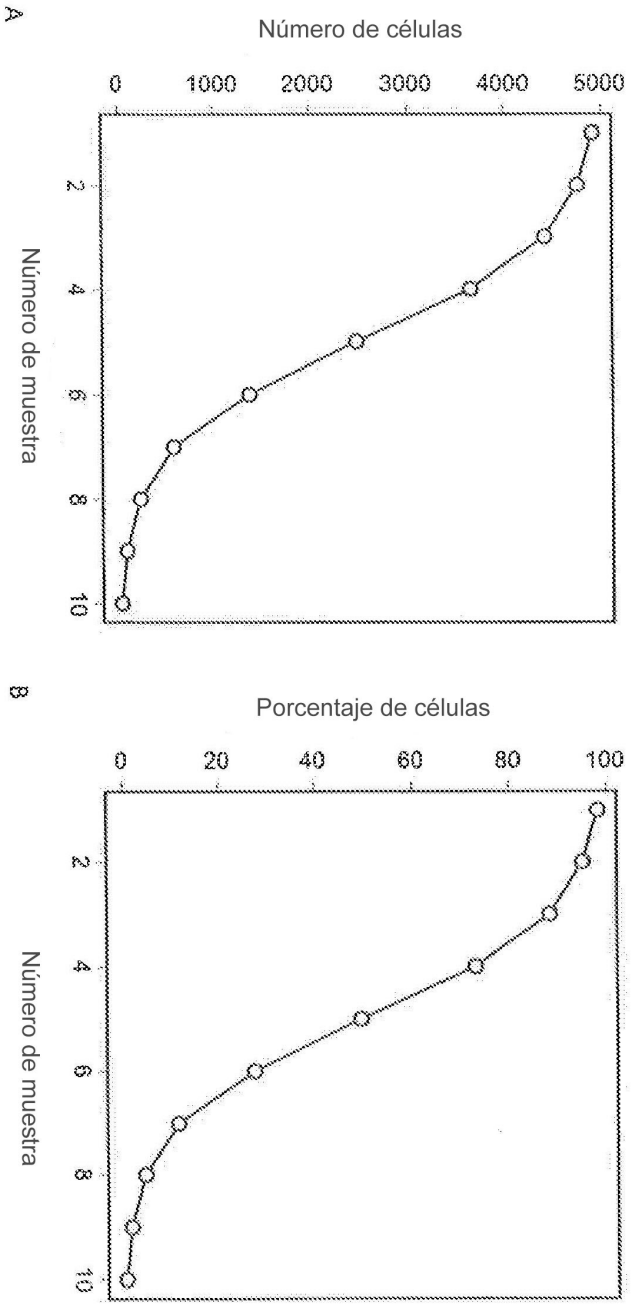
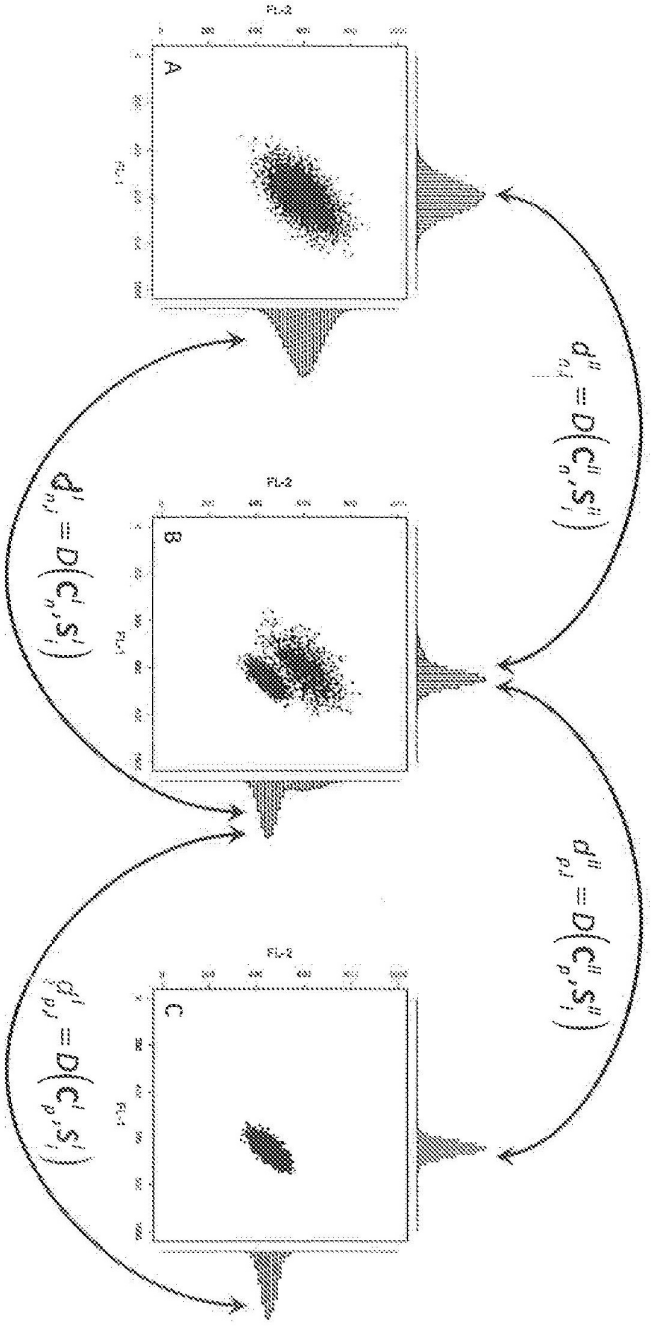

A
Número de células
0
1000
2000
3000
4000
5000
Número de muestra
2
4
6
8
10
B
Porcentaje de células
0
20
40
60
80
100
Número de muestra
2
4
6
8
10

FIG. 20

FIG. 21



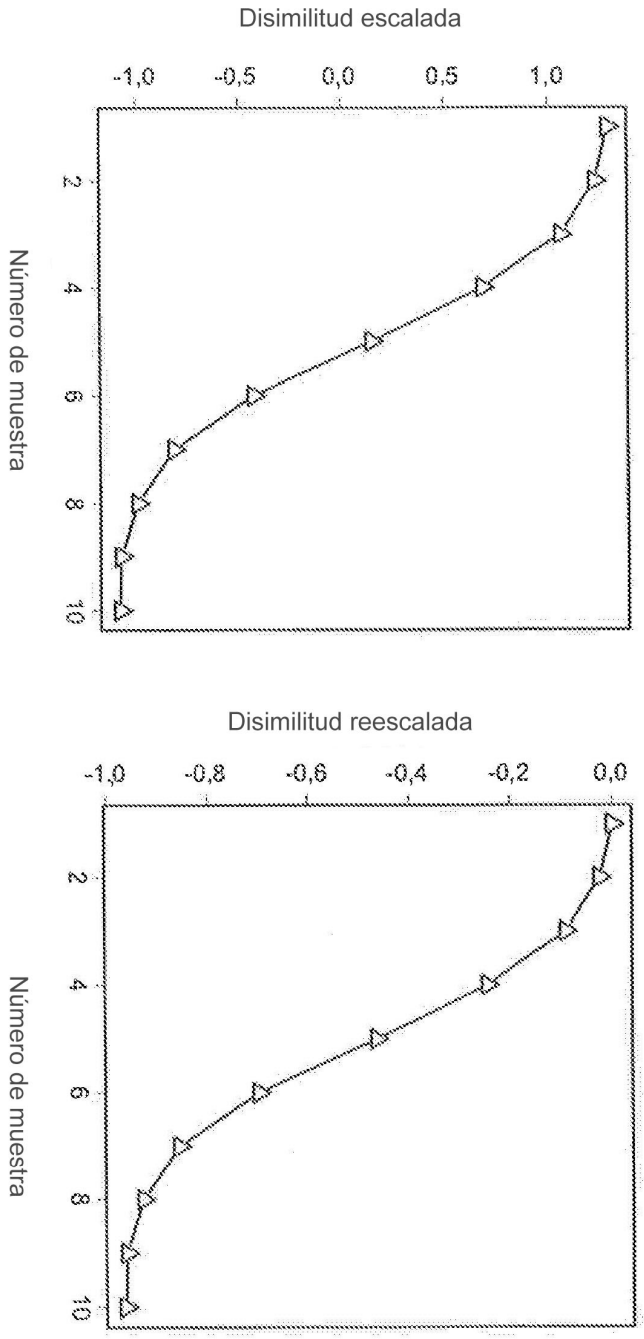
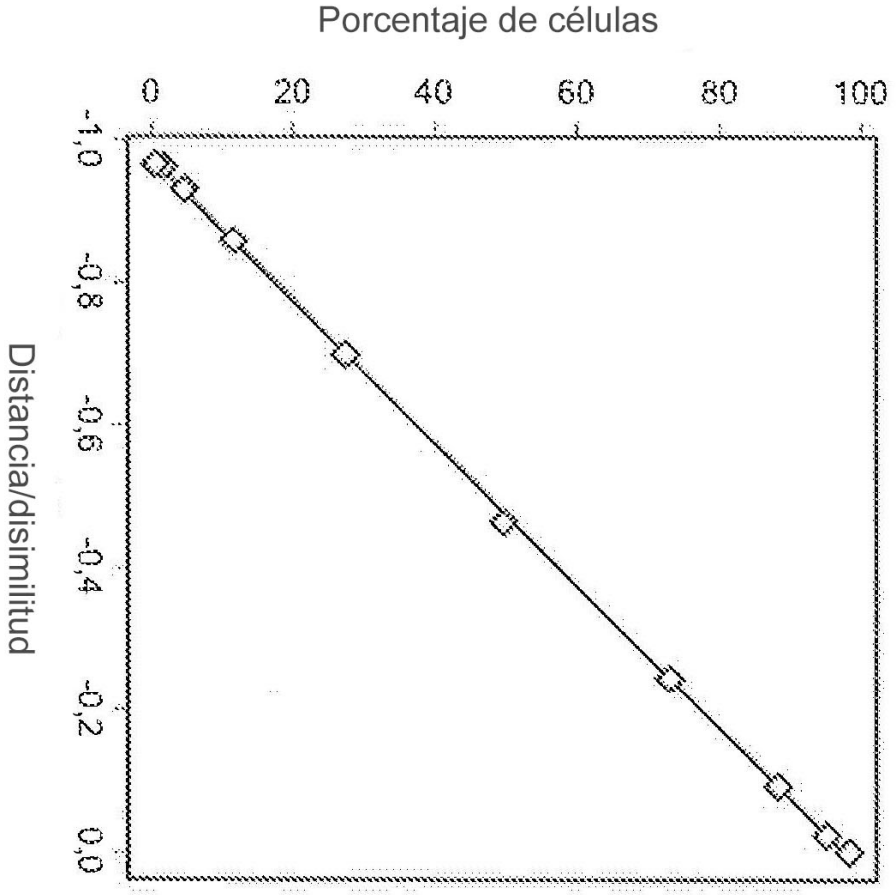

Disimilitud escalada
-1,0
-0,5
0,0
0,5
1,0
Número de muestra
2
4
6
8
10
Disimilitud reescalada
-1,0
-0,8
-0,6
-0,4
-0,2
0,0
Número de muestra
2
4
6
8
10

FIG. 22

FIG. 23


Porcentaje de células
0
20
40
60
80
100
Distancia/disimilitud
-1,0
-0,8
-0,6
-0,4
-0,2
0,0

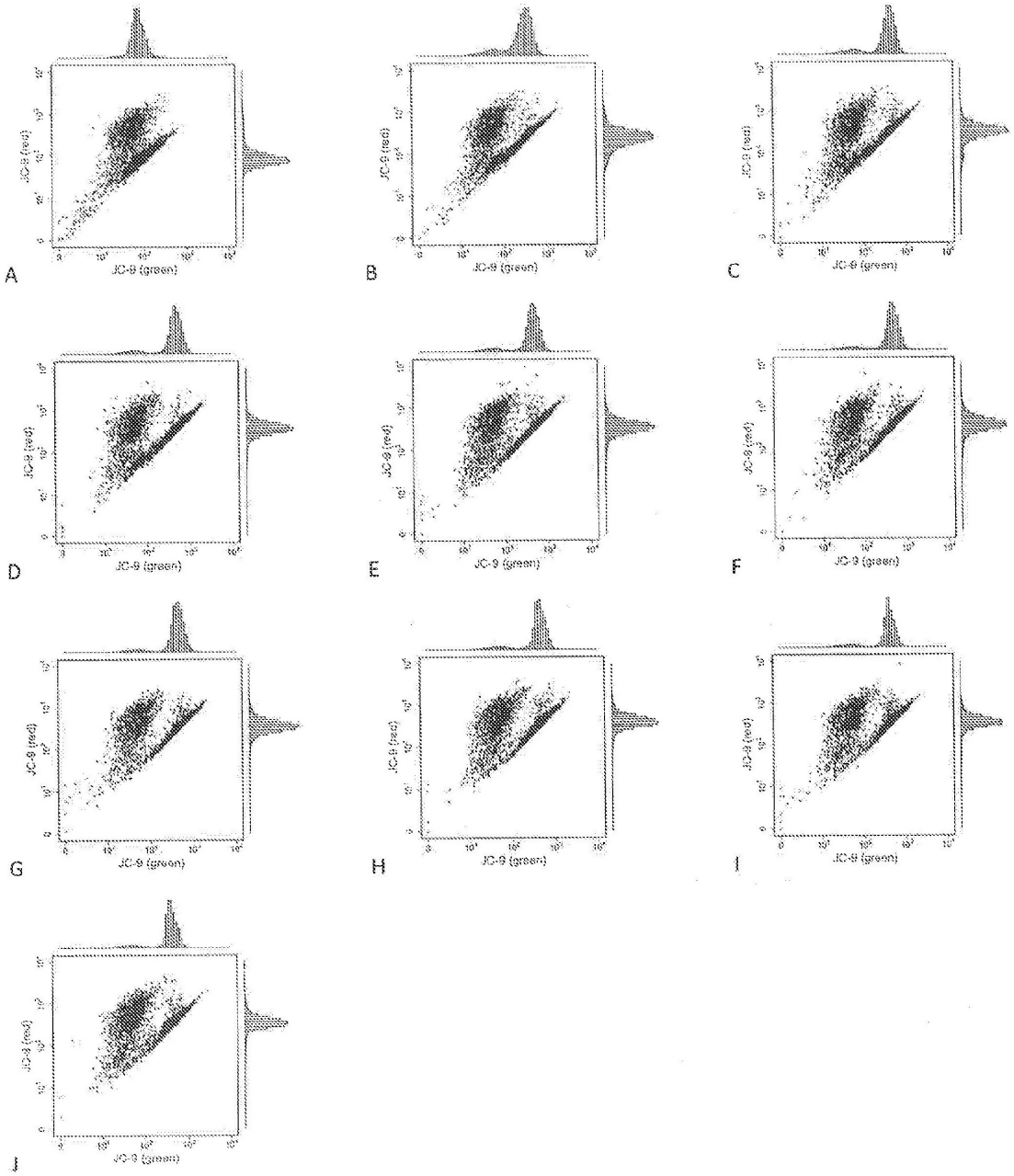



FIG. 24

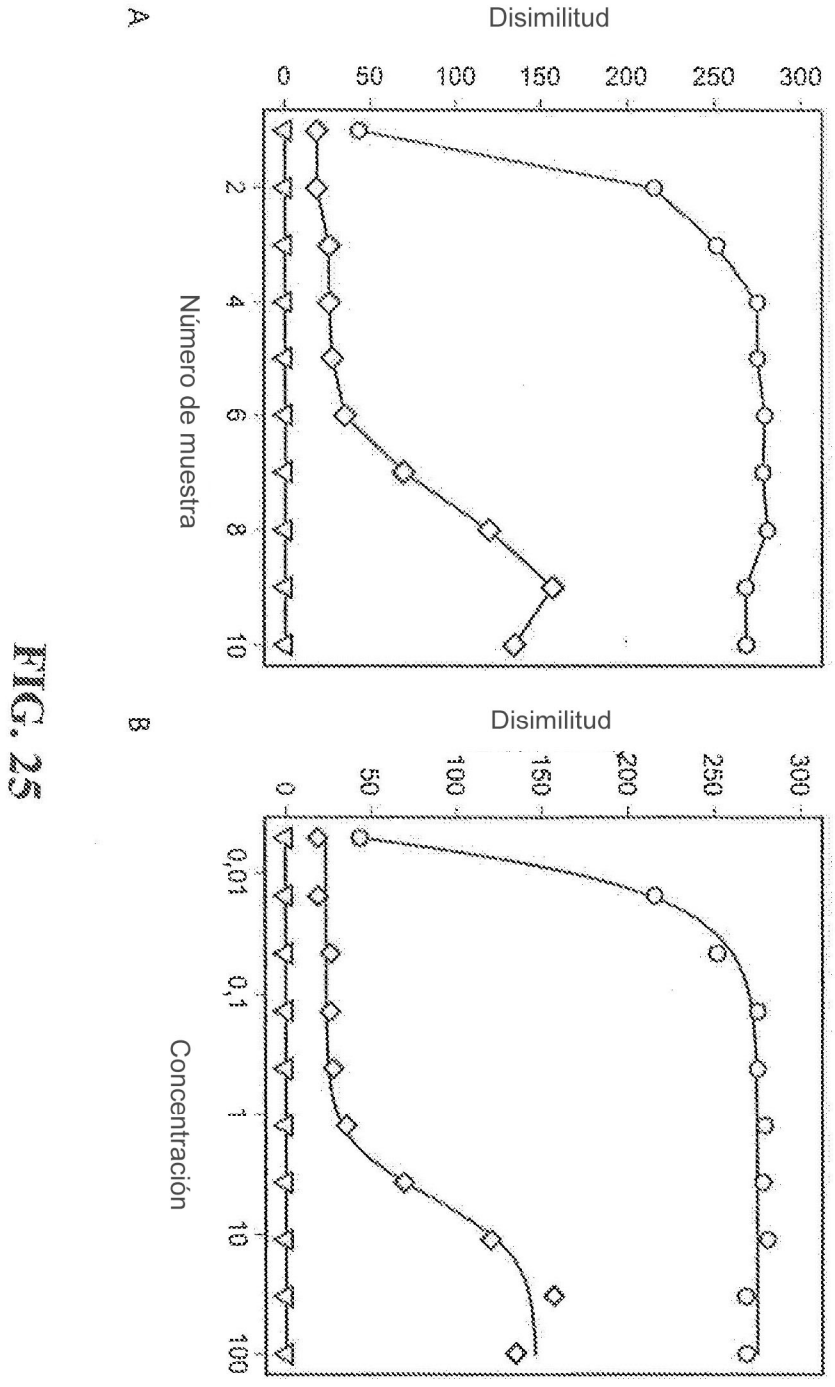

A
Disimilitud
0
50
100
150
200
250
300
2
4
6
8
10
Número de muestra
B
Disimilitud
0
50
100
150
200
250
300
0,01
0,1
1
10
100
Concentración

FIG. 2S