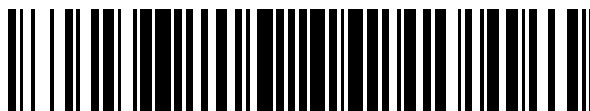


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 891 903**

51 Int. Cl.:

C12Q 1/6886 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **17.02.2016 PCT/KR2016/001630**

87 Fecha y número de publicación internacional: **25.08.2016 WO16133373**

96 Fecha de presentación y número de la solicitud europea: **17.02.2016 E 16752703 (5)**

97 Fecha y número de publicación de la concesión europea: **08.09.2021 EP 3261008**

54 Título: **Método para seleccionar agente antineoplásico basado en información de daño proteínico de individuos para evitar los efectos secundarios del agente antineoplásico**

30 Prioridad:

17.02.2015 KR 20150024443

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

31.01.2022

73 Titular/es:

**CIPHEROME, INC. (100.0%)
19925 Stevens Creek Blvd, Suite 100
Cupertino CA 95014, US**

72 Inventor/es:

KIM, JU HAN

74 Agente/Representante:

PONS ARIÑO, Ángel

ES 2 891 903 T3

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

DESCRIPCIÓN

Método para seleccionar agente antineoplásico basado en información de daño proteínico de individuos para evitar los efectos secundarios del agente antineoplásico

5

Campo de la invención

La presente divulgación se refiere a un método para seleccionar un agente antineoplásico basándose en la información de daño proteínico de un individuo para evitar los efectos secundarios de un agente antineoplásico, específicamente a un método para seleccionar un agente antineoplásico usando secuenciación genómica individual.

10

Antecedentes técnicos

Con el avance de las tecnologías de biotecnología, ahora es posible analizar toda la secuencia genómica de un ser humano y predecir una enfermedad de cada individuo y proporcionar prevención y tratamientos personalizados de la enfermedad.

15

Recientemente, como resultado de la comparación de secuencias genómicas individuales, se descubrió que pueden estar presentes diferentes bases en la misma posición en los cromosomas. Por consiguiente, dicha diferencia en las secuencias se ha usado para predecir diferencias en los individuos en la respuesta al fármaco. Por ejemplo, el metabolismo del fármaco puede ser lento o rápido dependiendo de una información de secuencia genómica específica de un individuo y, por tanto, los efectos terapéuticos o efectos secundarios de un fármaco pueden ser diferentes para cada individuo.

20

Por consiguiente, ha habido un aumento en la demanda de selección de fármacos personalizados que pueda seleccionar un fármaco y una dosis adecuados para un paciente usando la diferencia en las secuencias individuales. Además, ha surgido la farmacocinética o la farmacogenómica, que usa información genómica, por ejemplo, polimorfismo mononucleotídico (SNP), como marcador y correlación entre el marcador y la respuesta al fármaco/efectos secundarios del fármaco.

25

30

En la farmacogenética, se predicen diferencias en el metabolismo de los fármacos o productos químicos y la respuesta a los mismos en una población general o entre individuos mediante análisis genético. Algunos individuos pueden mostrar respuestas farmacéuticas inesperadas. Dichos efectos secundarios del fármaco pueden deberse a la gravedad de una enfermedad en tratamiento, la interacción de fármacos, la edad, el estado nutricional, la función hepática y renal de los pacientes y factores ambientales tales como el clima o la dieta. Sin embargo, también pueden estar provocados por diferencias genéticas relacionadas con el metabolismo del fármaco, por ejemplo, polimorfismo de enzimas metabolizantes del fármaco. Por lo tanto, se está realizando estudio a este respecto.

35

El busulfano, que es un agente antineoplásico y mielosupresor, se usa principalmente antes de trasplante de médula ósea para extirpación de médula ósea. El busulfano tiene una ventana terapéutica muy estrecha entre la mielotoxicidad letal y la toxicidad sistémica duradera. El busulfano tiene el problema de que se produce toxicidad farmacológica grave cuando la concentración del fármaco es mayor que el intervalo terapéutico y el riesgo de fracaso del injerto o recidiva aumenta cuando la concentración del fármaco es menor que el intervalo terapéutico. La toxicidad de busulfano incluye fibrosis pulmonar intersticial habitualmente denominada "pulmón de busulfano", hiperpigmentación, epilepsia, enfermedad venooclusiva hepática (VOD), emesis, trombocitopenia, etc. La probabilidad de estos problemas es elevada porque la farmacocinética del busulfano varía enormemente entre individuos.

40

45

Para reducir la toxicidad asociada con el acondicionamiento, se divulgó un método de administración de una dosis adecuada de busulfano que se determina a través de un modelo farmacocinético después de la administración de una dosis de ensayo o una dosis inicial (Kletzel M, 2006, Bartelink IH, 2008). Sin embargo, no se adopta globalmente porque el modelado farmacocinético requiere tomar muestras de sangre al menos 4-5 veces al día, así como medición y análisis farmacocinético en el día de ensayo y es necesario el caro sistema de cromatografía de líquidos/espectrometría de masas en tándem. Además, como el fármaco tiene que administrarse, es difícil predecir una dosis adecuada por anticipado y evitar complicaciones.

50

55

Puede usarse la farmacogenómica para predecir el efecto secundario de un fármaco antes de su administración. Para el busulfano también, se sabe que su nivel sanguíneo y los efectos secundarios están asociados con su farmacogenómica. En un estudio de farmacogenómica de busulfano en niños, se informó de que el gen GSTM1 afecta a la farmacocinética del busulfano (Ansari M. *et al.*, 2010). Además, se informó de que el nivel sanguíneo de busulfano puede variarse por el polimorfismo de GSTA1, que es una enzima metabolizante del fármaco importante (Kusama M. *et al.*, 2006).

60

Los fármacos basados en tiopurina con rutas metabólicas similares se usan de manera útil con diversos propósitos, incluyendo tratamiento de hematoma, tratamiento de enfermedades reumatoideas, inmunosupresión en trasplante de órganos, etc. Como uno de los fármacos basados en tiopurina, se usa azatioprina como medicación inmunosupresora en el tratamiento de colitis ulcerosa, que es una enfermedad inflamatoria del intestino, y enfermedad de Crohn. Durante

65

el tratamiento, algunos pacientes padecen supresión de la médula ósea y leucopenia. Se sabe que está estrechamente relacionado con el polimorfismo genético del gen TPMT que codifica tiopurina S-metiltransferasa. La tiopurina S-metiltransferasa es una enzima que metaboliza fármacos basados en tiopurina. Algunos genotipos conocidos para TPMT alteran el metabolismo de los fármacos basados en tiopurina inhibiendo o suprimiendo la función de TPMT. Se sabe que pueden producirse efectos secundarios graves tales como supresión de la médula ósea en pacientes que tienen dichos genotipos que suprimen la función de TPMT.

Como otro fármaco basado en tiopurina, se usa mercaptopurina junto con otros fármacos para tratamiento de mantenimiento en el tratamiento de leucemia linfocítica aguda. Como la azatioprina, la mercaptopurina también puede provocar varios efectos secundarios potencialmente mortales tales como leucopenia o agranulocitosis. Como la mercaptopurina es también un fármaco basado en tiopurina, la relación entre la variación del gen TPMT y los efectos secundarios de la mercaptopurina es conocida.

Aunque la variación del gen TPMT no se observa frecuentemente en asiáticos, muestran mayor prevalencia de leucopenia provocada por fármacos basados en tiopurina. De acuerdo con un estudio (Yang, S. K., Hong, M., Baek, J., Choi, H., Zhao, W., Joven, Y., ... y Park, S. K. (2014). A common missense variant in NUDT15 confers susceptibility to thiopurine-induced leukopenia. *Nature Genetics*, 46(9), 1017-1020.), Los pacientes con enfermedad de Crohn tratados con fármacos basados en tiopurina mostraron una alta correlación entre la variación del gen NUDT15 y la leucopenia que se producía durante los tratamientos con tiopurina.

Sacerdote *et al.*, *Int. J. Cancer*: 133, 2004-2009 (2013) divulga que polimorfismos en el gen XRCC1 modifican la supervivencia de pacientes con cáncer de vejiga tratados con quimioterapia.

Actualmente, se sabe que la variación genética de decenas a cientos de genes relacionados con el metabolismo de fármacos afecta a la función de las proteínas correspondientes en el metabolismo de fármacos en el organismo, promoviendo o suprimiendo de ese modo el metabolismo del fármaco. Aunque también es necesario dilucidar nuevas variaciones genéticas, se limita a comprender todos los diversos genes relacionados con el metabolismo de fármacos a través del análisis de una pequeña cantidad de genes. Por lo tanto, la introducción de una metodología más allá del método basado en un resultado de estudios de observación de poblaciones usando un pequeño número de genes o marcadores tal como un polimorfismo mononucleotídico es necesaria. Se necesita mucho introducir un método para proporcionar información para seleccionar un agente antineoplásico personalizado, que sea más útil y fiable, usando directamente información de variación de secuencia genómica individual y realizando deducción teórica sobre el daño a las proteínas provocado de este modo y el efecto biológico del mismo.

Descripción detallada de la invención

Problema técnico

La presente divulgación, que se concibe en vista de lo anterior, proporciona un método para proporcionar información para seleccionar un agente antineoplásico analizando información de variación de secuencia genómica individual, calculando una puntuación de daño proteínico individual a partir de la información de variación de secuencia génica implicada en la farmacodinámica y farmacocinética de un agente antineoplásico, un mielosupresor o un inmunosupresor que tenga una ruta metabólica y calculando una puntuación de fármaco individual asociando la puntuación con una relación de fármaco-proteína para predecir la probabilidad del efecto secundario del agente antineoplásico.

Solución técnica

En un aspecto, la presente invención proporciona un método implementado por ordenador para proporcionar información para seleccionar un agente antineoplásico usando variación de secuencia genómica individual, que comprende:

una etapa de obtener una o más variaciones de secuencia de un paciente en genes que se ha determinado que están implicados en la farmacodinámica o farmacocinética de un agente antineoplásico a partir de información de secuencia genómica individual;

una etapa de calcular una puntuación daño proteínico usando la [ecuación 2], en la que

la [ecuación 2] es:

$$S_g(v_1, \dots, v_n) = \left(\prod_{i=1}^n v_i^{w_i} \right)^{1/\sum_{i=1}^n w_i}$$

en la que S_g es la puntuación daño proteínico de una proteína codificada por un gen g , n es el número de la una

o más variaciones de secuencia del gen g, v_i es una puntuación de variación de secuencia génica de una variación de secuencia génica $i.^a$ y w_i es una ponderación asignada a la puntuación de variación de secuencia génica v_i de la $i.^a$ variación de secuencia génica; y una etapa de calcular una puntuación de fármaco individual usando la [ecuación 4], en la que

5

la [ecuación 4] es:

$$S_d(g_1, \dots, g_n) = \left(\prod_{i=1}^n g_i^{w_i} \right)^{1/\sum_{i=1}^n w_i}$$

10 en la que S_d es la puntuación de fármaco individual de un agente antineoplásico d, n es un número de proteínas codificadas por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d, g_i es una puntuación daño proteínico de una proteína codificada por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d, y w_i es una ponderación asignada a la puntuación daño proteínico g_i de la proteína codificada por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d; y proporcionar información para seleccionar un agente antineoplásico basándose en el cálculo de la puntuación de fármaco individual.

15

En otro aspecto, la presente invención proporciona un sistema para seleccionar un agente antineoplásico usando variación de secuencia genómica individual, que comprende:

20

- a) una base de datos de la que puede buscarse o extraerse información relevante a un gen o proteína relacionada con un agente antineoplásico aplicable a un paciente;
- b) una unidad de comunicación que es accesible para la base de datos;
- c) un primer módulo de cálculo que se configura para calcular una o más variaciones de secuencia en genes que se determina o que están implicados en la farmacodinámica o farmacocinética del agente antineoplásico basándose en la información;
- d) un segundo módulo de cálculo que se configura para calcular una puntuación daño proteínico usando la [ecuación 2], en la que

25

la [ecuación 2] es:

30

$$S_g(v_1, \dots, v_n) = \left(\prod_{i=1}^n v_i^{w_i} \right)^{1/\sum_{i=1}^n w_i}$$

en la que S_g es la puntuación daño proteínico de una proteína codificada por un gen g, n es el número de la una o más variaciones de secuencia del gen g, v_i es una puntuación de variación de secuencia génica de una variación de secuencia génica $i.^a$ y w_i es una ponderación asignada a la puntuación de variación de secuencia génica v_i de la $i.^a$ variación de secuencia génica;

35

e) un tercer módulo de cálculo que se configura para calcular una puntuación de fármaco individual usando la [ecuación 4], en la que

40

la [ecuación 4] es:

$$S_d(g_1, \dots, g_n) = \left(\prod_{i=1}^n g_i^{w_i} \right)^{1/\sum_{i=1}^n w_i}$$

en la que S_d es la puntuación de fármaco individual de un fármaco d, n es un número de proteínas codificadas por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d, g_i es una puntuación daño proteínico de una proteína codificada por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d, y w_i es una ponderación asignada a la puntuación daño proteínico g_i de la proteína codificada por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d; y

45

50

f) una unidad de presentación que se configura para presentar los valores calculados por los módulos de cálculo;

y opcionalmente comprende además:

- g) un cuarto módulo de cálculo que se configura para determinar si el agente antineoplásico se usará para el paciente usando la puntuación de fármaco individual calculada por el tercer módulo de cálculo; o
- h) una interfaz de usuario que se configura para calcular y proporcionar una puntuación de fármaco individual para el agente antineoplásico cuando el agente antineoplásico se introduce por un usuario; o
- i) una unidad de presentación que se configura para presentar además el valor calculado por cada uno de los módulos de cálculo, el procedimiento de cálculo o la información como fundamento del cálculo; y

en el que el sistema proporciona información para seleccionar un agente antineoplásico basándose en el cálculo de la puntuación de fármaco individual.

En otro aspecto, la presente invención proporciona un medio legible por ordenador que incluye un módulo de ejecución para ejecutar un procesador que realiza una operación, que comprende:

- una etapa de adquisición de una o más variaciones de secuencia de un paciente en genes que se determinan o que están implicados en la farmacodinámica o farmacocinética de un agente antineoplásico a partir de información de secuencia genómica individual;
- una etapa de calcular una puntuación daño proteínico usando la [ecuación 2], en la que

la [ecuación 2] es:

$$S_g(v_1, \dots, v_n) = \left(\prod_{i=1}^n v_i^{w_i} \right)^{1/\sum_{i=1}^n w_i}$$

en la que S_g es la puntuación daño proteínico de una proteína codificada por un gen g , n es el número de la una o más variaciones de secuencia del gen g , v_i es una puntuación de variación de secuencia génica de una variación de secuencia génica i .ª y w_i es una ponderación asignada a la puntuación de variación de secuencia génica v_i de la i .ª variación de secuencia génica; y una etapa de calcular una puntuación de fármaco individual usando la [ecuación 4], en la que la [ecuación 4] es:

la [ecuación 4] es:

$$S_d(g_1, \dots, g_n) = \left(\prod_{i=1}^n g_i^{w_i} \right)^{1/\sum_{i=1}^n w_i}$$

en la que S_d es la puntuación de fármaco individual de un fármaco d , n es un número de proteínas codificadas por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d , g_i es una puntuación daño proteínico de una proteína codificada por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d , y w_i es una ponderación asignada a la puntuación daño proteínico g_i de la proteína codificada por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d .

Efectos ventajosos

Un método y un sistema para seleccionar un agente antineoplásico personalizado basándose en la información de variación de secuencia genómica individual de la presente divulgación puede predecir la respuesta de un individuo a un fármaco específico, es decir, un agente antineoplásico, analizando la secuencia de la región exónica de un gen que codifica diversas proteínas implicadas en la farmacodinámica o farmacocinética del agente antineoplásico.

El método existente de control de la dosis de administración de busulfano basándose en el modelado farmacocinético no se adopta ampliamente porque se necesita tomar muestras varias veces al día y se necesita un equipo caro para el análisis instantáneo de la muestra de sangre. Por el contrario, como ha llegado a ser fácil usar las técnicas de genotipado (PCR, chip de SNP, secuenciación capilar, NGS, etc.) para aplicaciones clínicas con el avance en los métodos experimentales, se espera que la presente divulgación basada en farmacogenómica pueda usarse ampliamente en entornos clínicos. En particular, como el precio de la secuenciación de última generación (NGS) cae rápidamente, la estrategia genómica ha llegado a ser posible a un coste de decenas a cientos de miles de ganancia por paciente.

Además, el estudio farmacogenómico existente tiene que realizarse basándose en pares de fármaco-gen. Sin embargo, es prácticamente imposible estudiar todos los numerosos pares de fármaco-gen porque el número de pares aumenta en proporción al múltiplo del número de fármacos y el número de marcadores génicos. Por tanto, aún no se

han generado suficientes datos de soporte. Además, pueden producirse errores estadísticos grandes dependiendo de los grupos de población seleccionados. Por el contrario, de acuerdo con el método de la presente divulgación, se aplican directamente resultados de estudio y análisis a nivel molecular al tratamiento con fármacos personalizados y, por tanto, pueden establecerse fundamentos para casi todos los pares de fármaco-gen y el método puede aplicarse sin verse afectado significativamente por la diferencia entre los grupos de población.

En particular, el método y el sistema de la presente divulgación pueden usarse de manera útil para determinar si un agente antineoplásico, por ejemplo, busulfano, puede usarse para un individuo y la manera en que se usará prediciendo el efecto secundario o riesgo del fármaco por anticipado. Esto puede dar lugar a una tasa de supervivencia mejorada de los pacientes con cáncer infantil a través de la administración de busulfano a dosis óptimas. Además, puede introducirse una nueva pauta terapéutica para grupos de pacientes para los que no sea fácil el trasplante de células madre hematopoyéticas debido a la dificultad en el preacondicionamiento mejorando la seguridad del preacondicionamiento con busulfano.

La presente divulgación también puede usarse de manera útil para determinar si la mercaptopurina, que es otro agente antineoplásico, puede usarse para un individuo y la manera en que se usará prediciendo el efecto secundario o riesgo del fármaco por anticipado. Esto puede dar lugar a una tasa de supervivencia mejorada y un coste disminuido del tratamiento evitando la hospitalización prolongada o los estados físicos agravados de los pacientes provocados por los efectos secundarios tales como leucopenia, etc., durante el tratamiento.

Además, si se encuentra o se proporciona nueva información acerca de una relación de agente antineoplásico-proteína, puede añadirse fácilmente y aplicarse al método de la presente divulgación. Por tanto, es posible proporcionar un método de tratamiento mejorado de fármaco personalizado de acuerdo con mayor acumulación de información.

Breve descripción de los dibujos

La FIG. 1 es un diagrama de flujo que ilustra cada etapa de un método para proporcionar información para seleccionar un agente antineoplásico usando variación de secuencia genómica individual de acuerdo con una realización ilustrativa de la presente divulgación.

La FIG. 2 ilustra esquemáticamente la configuración de un sistema para seleccionar un agente antineoplásico usando variación de secuencia genómica individual de acuerdo con una realización ilustrativa de la presente divulgación (DB: base de datos).

Las FIG. 3A-3B muestran curvas de ROC (receiver operating characteristic - rendimiento diagnóstico) para verificar un método para seleccionar un agente antineoplásico usando variación de secuencia genómica de acuerdo con una realización ilustrativa de la presente divulgación.

Modo para realizar la invención

La presente divulgación se basa en el hallazgo de que es posible seleccionar un fármaco con alta seguridad y personalizar su dosis/uso para un individuo en un tratamiento con fármaco analizando la información de variación de secuencia genómica individual.

En un aspecto, la presente divulgación se refiere a un método para proporcionar información para seleccionar un agente antineoplásico usando variación de secuencia genómica individual, que incluye: una etapa de determinar información de una o más variaciones de secuencia génica implicadas en la farmacodinámica o farmacocinética de un agente antineoplásico a partir de información de secuencia genómica individual; una etapa de calcular una puntuación daño proteínico individual usando la información de variación de secuencia génica; y una etapa de calcular una puntuación de fármaco individual asociando la puntuación daño proteínico individual con una relación de fármaco-proteína.

En la presente divulgación, el agente antineoplásico incluye los fármacos descritos en la tabla 1 a continuación y cualquier sustancia que muestre la misma actividad farmacológica o similar, incluyendo derivados de los mismos, sales farmacéuticamente aceptables de los mismos, etc. Por ejemplo, puede ser un fármaco basado en tiopurina incluyendo tioguanina, mercaptopurina y azatioprina, un análogo desoxinucleosídico incluyendo cladribina, fludarabina, clofarabina y nelarabina, un agente alquilante incluyendo treosulfano, manosulfano y busulfano, un antimetabolito, un agente antimicrotúbulos, un inhibidor de topoisomerasa, un anticuerpo citotóxico, etc., específicamente busulfano o mercaptopurina, aunque sin limitarse a ello.

[Tabla 1]

Temozolomida	Gefitinib	Amsacrina	Etinil estradiol	Fluorouracilo	Trastuzumab
Trastuzumab (recombinación genética)	Megestrol	Celecoxib	Flutamida	Ifosfamida	Dasatinib
Procarbazina	Fulvestrant	Temsirólimus	Pemetrexed	Histamina	Tamoxifeno
Ciclosporina	Letrozol	Medroxiprogesterona	Capecitabina	Daunorrubicina	Sorafenib
Exemestano	Dactinomicina	Nilutamida	Oxaliplatino	Vinorelbina	Imatinib
Tioguanina	Topotecán	Aminoglutetimida	Erlotinib	Vindesina	Ciclosporina
Mitomicina	Hidroxiurea	Mercaptopurina	Gemcitabina	Etopósido	Vincristina
Ixabepilona	Ácido aminolevulínico	Mitoxantrona	Tenipósido	Dietilestilbestrol	Vinblastina
Busulfano	Azatioprina	Idarrubicina	Ciclofosfamida	Epirubicina	Paclitaxel
Lomustina	Bicalutamida	Trabectedina	Cisplatino	Talidomida	Bortezomib
Estreptozocina	Toremifeno	Dacarbazina	Citidina	Cladribina	Tretinoína
Tiotepa	Anastrozol	Bexaroteno	Trióxido arsénico	Irinotecán	Metotrexato
Gemtuzumab ozogamicina (recombinación genética)	Clofarabina	Estramustina	Nilotinib	Sunitinib	Nelarabina
Everólimus	Fludarabina	Lapatinib	Sirólimus	Docetaxel	Treosulfano
Afatinib	Cabazitaxel	Tacrólimus	Leflunomida	Etanol	Manosulfano

En la presente divulgación, la información de los genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico puede obtenerse de bases de datos tales como DrugBank (<http://www.drugbank.ca/>), KEGG Drug (<http://www.genome.jp/kegg/drug/>), PharmGKB (<https://www.pharmgkb.org/>), etc. Específicamente, el gen incluye ABCC1, ABCG2, APEX1, CBR3, ATP7A, CBR3, CTH, CUL9, CYP1A1, CYP2C9, CYP2D6, CYP4B1, DPEP1, DPEP2, DPYD, DRD3, EPHX1, FMO2, FMO3, GGT1, GGT5, GGT6, GNB3, GSTA2, GSTM1, MGMT, MLH1, MSH2, NAT1, OPRM1, PDE5A, PTGS1, SERPINA6, SLC15A2, SLC22A1, SLC22A2, SLC22A8, SLC22A16, SLC22A2, SLC28A2, SLC28A3, SPG7, TP53, ABCB1, ABCC2, AOX1, CHRM2, CYP1B1, CYP3A4, CYP3A5, CYP4F2, DBH, ERCC2, GGH, GGT7, GMPS, MTHFR, NUDT15, PDE3A, PDE4C, SLC22A5, SLC01B1, TLR3, UGT1A1, XDH, etc., aunque sin limitarse a ellos.

En la presente divulgación, los genes/proteínas se denominan de acuerdo con la nomenclatura HGNC (comité de nomenclatura de genes HUGO) (Gray KA, Daugherty LC, Gordon SM, Seal RL, Wright MW, Bruford EA. *genenames.org: the HGNC resources in 2013*. *Nucleic Acids Res.* enero de 2013; 41 (número de publicación de base de datos): D545-52. doi: 10.1093/nar/gks1066. Epub 17 de nov de 2012 PMID: 23161694).

En la presente divulgación, la variación de secuencia génica usada como información se refiere a variación de secuencia génica individual o polimorfismo. En la presente divulgación, la variación de secuencia génica o polimorfismo se produce en una región, particularmente una región exónica, de un gen que codifica proteínas implicadas en la farmacodinámica o farmacocinética de un agente antineoplásico, aunque sin limitarse a ello.

La expresión "información de variación de secuencia" usada en la presente divulgación significa información acerca de la sustitución, adición o eliminación de una base que constituye un exón de un gen. Dicha sustitución, adición o eliminación de la base puede ser el resultado de diversas causas, por ejemplo, diferencias estructurales incluyendo mutación, rotura, eliminación, duplicación, inversión y/o translocación de un cromosoma.

En otro aspecto, un polimorfismo de una secuencia se refiere a diferencias en una secuencia entre individuos, presentes en un genoma. El polimorfismo mononucleotídico (SNP) es la forma más frecuente de polimorfismo de una secuencia. Se refiere a variación en una base de una secuencia que consiste en A, T, C y G entre individuos. El polimorfismo de secuencia que incluye el SNP puede producirse como SNV (variación mononucleotídica), STRP (polimorfismo de repetición en tándem corta) o una variación polialélica incluyendo VNTR (diversos números de repetición en tándem) y CNV (variación del número de copias).

En el método de la presente divulgación, la información de variación de secuencia o polimorfismo encontrada en un genoma individual se recoge en asociación con una proteína implicada en la farmacodinámica o farmacocinética de un agente antineoplásico. Es decir, la información de variación de secuencia usada en el método de la presente divulgación es información de variación encontrada particularmente en una región exónica de uno o más genes implicados en la farmacodinámica o farmacocinética de un agente antineoplásico, por ejemplo, un gen que codifica una proteína diana relevante para un fármaco, una proteína enzimática implicada en el metabolismo de fármacos, una proteína transportadora o una proteína de transporte, entre la información de secuencia genómica individual obtenida, aunque sin limitarse a ello.

La información de secuencia genómica individual usada en la presente divulgación puede determinarse usando un método de secuenciación bien conocido. Además, pueden usarse los servicios disponibles en el mercado de Complete Genomics, BGI (Beijing Genome Institute), Knome, Macrogen, DNALink, etc., aunque sin limitarse a ellos.

En la presente divulgación, la información de variación de secuencia génica presente en una secuencia genómica individual puede extraerse usando diversos métodos y puede adquirirse a través de análisis comparativo de secuencias usando un programa, por ejemplo, ANNOVAR (Wang *et al.*, *Nucleic Acids Research*, 2010; 38(16): e164), SV4 (analyzer de variantes de secuencia) (Ge *et al.*, *Bioinformatics*. 2011; 27(14): 1998-2000), BreakDancer (Chen *et al.*, *Nat Methods*. septiembre de 2009; 6(9):677-81) y similares, que comparan la secuencia con un grupo de referencia, por ejemplo, la secuencia genómica de HG19.

La información de variación de secuencia génica puede recibirse/adquirirse a través de un sistema informático. En este aspecto, el método de la presente divulgación puede incluir además una etapa de recibir la información de variación de secuencia génica a través de un sistema informático. El sistema informático usado en la presente divulgación puede incluir o ser accesible para una o más bases de datos que contienen información acerca del gen implicado en la farmacodinámica o farmacocinética del agente antineoplásico, por ejemplo, un gen que codifica una proteína diana relevante para un fármaco, una proteína enzimática implicada en el metabolismo de fármacos, una proteína transportadora, una proteína de transporte, etc. Estas bases de datos pueden incluir una base de datos pública o no pública o una base de conocimientos, que proporciona información acerca del gen/proteína/interacciones de fármaco-proteína, etc., incluyendo, por ejemplo, DrugBank (<http://www.drugbank.ca/>), KEGG Drug (<http://www.genome.jp/kegg/drug/>) y PharmGKB (<http://www.pharmgkb.org/>), aunque sin limitarse a ellos.

En la presente divulgación, el agente antineoplásico puede ser información introducida por un usuario, información introducida desde una prescripción o información introducida desde una base de datos que incluye información acerca de agentes antineoplásicos. La prescripción puede incluir una prescripción electrónica, aunque sin limitarse a ella.

La expresión "farmacocinética (PK) o parámetros farmacocinéticos" usada en la presente divulgación se refiere a características de un fármaco implicadas en absorción, migración, distribución, conversión y excreción del fármaco en el organismo durante un periodo de tiempo predeterminado, e incluye el volumen de distribución (V_d), tasa de eliminación (CL), biodisponibilidad (F) y coeficiente de tasa de absorción (k_a) de un fármaco o la concentración máxima en plasma ($C_{m\acute{a}x}$), punto temporal de concentración máxima en plasma ($T_{m\acute{a}x}$), área bajo la curva (ABC) con respecto al cambio en el nivel sanguíneo de un fármaco durante un determinado periodo de tiempo, etc.

La expresión "farmacodinámica o parámetros farmacodinámicos" usada en la presente divulgación se refiere a características implicadas en las acciones fisiológicas y bioquímicas de un fármaco con respecto al organismo y mecanismos del mismo, es decir, respuestas o efectos en el organismo provocados por el fármaco.

La expresión "puntuación de variación de secuencia génica" usada en la presente divulgación se refiere a una puntuación numérica de un grado de la variación de secuencia genómica individual que provoca una variación de secuencia de aminoácidos (sustitución, adición o eliminación) de una proteína codificada por un gen o una variación de control de la transcripción y, por tanto, provoca un cambio significativo o daño en una estructura y/o función de la proteína cuando la variación de secuencia genómica se encuentra en una región exónica del gen que codifica la proteína. La puntuación de variación de secuencia génica puede calcularse considerando un grado de conservación evolutiva de aminoácidos en una secuencia genómica, un grado de un efecto de una característica física de aminoácido modificado en una estructura o función de la proteína correspondiente.

La puntuación de variación de secuencia génica usada para calcular la puntuación de daño proteínico individual y la puntuación de fármaco individual de acuerdo con la presente divulgación puede calcularse usando un método conocido en la técnica. Por ejemplo, la puntuación de variación de secuencia génica puede calcularse a partir de la información de variación de secuencia génica usando un algoritmo tal como SIFT (clasificación de intolerante frente a tolerante, Pauline C *et al.*, *Genome Res.* mayo de 2001; 11(5): 863-874; Pauline C *et al.*, *Genome Res.* marzo de 2002; 12(3): 436-446; Jing Hul *et al.*, *Genome Biol.* 2012; 13(2): R9), PolyPhen, PolyPhen-2 (fenotipado de polimorfismos, Ramensky V *et al.*, *Nucleic Acids Res.* 1 de septiembre de 2002; 30(17): 3894-3900; Adzhubei IA *et al.*, *Nat Methods* 7(4): 248-249 (2010)), MAPP (Eric A. *et al.*, análisis de múltiples variables de polimorfismo proteínico, *Genome Research* 2005; 15: 978-986), Logre (Log R de valor E Pfam, Clifford R.J *et al.*, *Bioinformatics* 2004; 20: 1006-1014), Mutation Assessor (Reva B *et al.*, *Genome Biol.* 2007; 8: R232, <http://mutationassessor.org/>), Condel (Gonzalez-Perez A *et al.*, *The American Journal of Human Genetics* 2011; 88: 440-449, <http://bg.upf.edu/fannsd/b/>), GERP (Cooper *et*

al., perfilado de tasa evolutiva genómica, *Genome Res.* 2005; 15: 901-913, <http://mendel.stanford.edu/SidowLab/downloads/gerp/>), CADD (reducción dependiente de anotación combinada, <http://cadd.gs.washington.edu/>), MutationTaster, MutationTaster2 (Schwarz *et al.*, MutationTaster2: mutation prediction for the deep-sequencing age. *Nature Methods* 2014; 11: 361-362, <http://www.mutationtaster.org/>), PROVEAN (Choi *et al.*, *PLoS One.* 2012; 7(10): e46688), PMut (Ferrer-Costa *et al.*, *Proteins* 2004; 57(4): 811-819, <http://mmb.pcb.ub.es/PMut/>), CEO (optimización de entropía combinatoria, Reva *et al.*, *Genome Biol* 2007; 8(11): R232), SNPeffct (Reumers *et al.*, *Bioinformatics.* 2006; 22(17): 2183-2185, <http://snpeffct.vib.be>), fathmm (Shihab *et al.*, *Functional Analysis through Hidden Markov Models*, *Hum Mutat* 2013; 34: 57-65, <http://fathmm.biocompute.org.uk/>), etc., aunque sin limitarse a ellos.

Los algoritmos des anteriormente se configuran para identificar cuanto efecto tiene cada variación de secuencia génica sobre una función proteínica, cuanto efecto daña la proteína o si hay o no algún otro efecto. Estos algoritmos se configuran de manera básica para considerar una secuencia de aminoácidos de una proteína codificada por un gen correspondiente y su cambio relevante provocado por una variación de secuencia génica individual y de ese modo determinar un efecto sobre una estructura y/o función de la proteína correspondiente.

En una realización ilustrativa de la presente divulgación, se usa el algoritmo SIFT (clasificación de intolerante frente a tolerante) para calcular una puntuación de variación de secuencia génica individual. En el caso del algoritmo SIFT, por ejemplo, se introduce información de variación de secuencia génica en forma de un archivo VCF (formato de identificación de variantes) y se puntúa un grado de daño provocado por cada variación de secuencia génica al gen correspondiente. En el caso del algoritmo SIFT, cuanto más cercano es un valor calculado a 0, se considera que una proteína codificada por un gen correspondiente se daña gravemente y, por tanto, se daña su función, y cuanto más cercano es la puntuación calculada a 1, se considera que la proteína codificada por el gen correspondiente mantiene su función normal.

En el caso de otro algoritmo, PolyPhen-2, cuanto mayor sea la puntuación calculada, se considera que la función de un proteína codificada por un gen correspondiente está más dañada.

Recientemente, se informó de un estudio (González-Pérez, A. y López-Bigas, N. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, *Condel. The American Journal of Human Genetics*, 2011; 88(4): 440-449), que sugiere el algoritmo Condel por comparación y combinación de SIFT, Polyphen2, MAPP, Logre y Mutation Assessor. En este estudio, los cinco algoritmos descritos anteriormente se comparan usando HumVar y HumDiv (Adzhubei, IA *et al.*, A method and server for predicting damaging missense mutations. *Nature methods*, 2010; 7(4): 248-249) como un conjunto de datos conocidos con respecto a variaciones de secuencia génica que dañan una proteína y variaciones de secuencia génica con menos efectos. Como resultado, se detectaron de manera idéntica un 97,9 % de las variaciones de secuencia génica que dañan una proteína y un 97,3 % de las variaciones de secuencia génica con menos efectos de HumVar por al menos tres de los cinco algoritmos descritos anteriormente, y un 99,7 % de las variaciones de secuencia génica que dañan una proteína y el 98,8 % de variaciones de secuencia génica con menos efectos HumDiv se detectaron de manera idéntica por al menos tres de los cinco algoritmos descritos anteriormente. Además, como resultado de trazar una curva de ROC (rendimiento diagnóstico) que muestra la precisión de los resultados de cálculo de los cinco algoritmos y una combinación de los algoritmos utilizando HumDiv y HumVar, se confirmó que la consistencia de una ABC (área bajo la curva de rendimiento diagnóstico) es considerablemente alta (de un 69 % a un 88,2 %). Es decir, los algoritmos descritos anteriormente son diferentes en los métodos de cálculo, pero las puntuaciones de variación de secuencia génica calculada están significativamente correlacionadas entre sí. Por lo tanto, se incluye en el alcance de la presente divulgación, independientemente de los algoritmos particulares que calculan las puntuaciones de variación de secuencia génica para aplicar una puntuación de variación de secuencia génica calculada aplicando los algoritmos descritos anteriormente o un método que emplea los algoritmos para las etapas de calcular una puntuación de daño proteínico individual y una puntuación de fármaco individual de acuerdo con la presente divulgación.

Cuando se produce una variación de secuencia génica en la región exónica de un gen que codifica una proteína, la variación de secuencia génica puede afectar directamente a la estructura y/o función de la proteína. Por lo tanto, la información de variación de secuencia génica puede estar asociada con el grado de daño a una función proteínica. En este aspecto, el método de la presente divulgación calcula una puntuación de daño proteínico individual basándose en la puntuación de variación de secuencia génica descrita anteriormente en la siguiente etapa.

La expresión "puntuación de daño proteínico" usada en la presente divulgación se refiere a una puntuación calculada recopilando las puntuaciones de variación de secuencia génica cuando se encuentran dos o más variaciones de secuencia significativas en una región génica que codifica una sola proteína, de modo que la proteína individual tiene dos o más puntuaciones de variación de secuencia génica. Si hay una variación de secuencia significativa individual en la región génica que codifica la proteína, una puntuación de variación de secuencia génica es idéntica a una puntuación de daño proteínico. En el presente documento, si hay dos o más variaciones de secuencia génica que codifican una proteína, una puntuación de daño proteínico se calcula como una media de las puntuaciones de variación de secuencia génica calculadas para las variaciones respectivas. Dicha media puede calcularse, por ejemplo, aunque sin limitación, midiendo una media geométrica, una media aritmética, una media armónica, una media geométrica aritmética, una media armónica aritmética, una media armónica geométrica, una media pitagórica, una media de

intercuartiles, una media cuadrática, una media truncada, una media winsorizada, una media ponderada, una media geométrica ponderada, una media aritmética ponderada, una media armónica ponderada, una media de una función, una media generalizada, una media f-generalizada, un percentil, un valor máximo, un valor mínimo, una moda, una mediana, un intervalo medio, una tendencia central, una multiplicación simple o una multiplicación ponderada, o mediante una operación funcional de los valores calculados, aunque sin limitarse a ellos.

En un aspecto ilustrativo de la presente divulgación, la puntuación de daño proteínico se calcula por la ecuación 1. La siguiente ecuación 1 puede modificarse de diversas maneras y, por tanto, la presente divulgación no se limita a ello.

[Ecuación 1]

$$S_g(v_1, \dots, v_n) = \left(\frac{1}{n} \sum_{i=1}^n v_i^p \right)^{\frac{1}{p}}$$

En la ecuación 1, S_g es una puntuación de daño proteínico de una proteína codificada por un gen g , n es el número de variaciones de secuencia diana a analizar entre variaciones de secuencia del gen g , v_i es una puntuación de variación de secuencia génica de una i .ª variación de secuencia génica y p es un número real distinto de 0. En la ecuación 1, cuando un valor de p es 1, la puntuación de daño proteínico es una media aritmética, si el valor de p es -1, la puntuación de daño proteínico es una media armónica, y si el valor de p está cercano al límite 0, la puntuación de daño proteínico es una media geométrica.

En el método de la presente invención, la puntuación de daño proteínico se calcula por la ecuación 2.

[Ecuación 2]

$$S_g(v_1, \dots, v_n) = \left(\prod_{i=1}^n v_i^{w_i} \right)^{1/\sum_{i=1}^n w_i}$$

En la ecuación 2, S_g es una puntuación de daño proteínico de una proteína codificada por un gen g , n es el número de variaciones de secuencia diana a analizar entre variaciones de secuencia del gen g , v_i es una puntuación de variación de secuencia génica de una i .ª variación de secuencia génica y w_i es una ponderación asignada a v_i . Si todas las ponderaciones w_i tienen el mismo valor, la puntuación de daño proteínico S_g es una media geométrica de las puntuaciones de variación de secuencia génica v_i . La ponderación puede asignarse considerando una clase de la correspondiente proteína, clasificación farmacodinámica o farmacocinética de la correspondiente proteína, parámetros farmacocinéticos de la proteína enzimática de un fármaco correspondiente, un grupo de población o una distribución raza.

Los "parámetros farmacocinéticos de una enzima metabolizante de fármaco" usados en la presente divulgación incluyen $V_{m\acute{a}x}$, K_m , K_{cat}/K_m , etc. $V_{m\acute{a}x}$ es la tasa de reacción enzimática máxima cuando una concentración de sustrato es muy alta, y K_m es la concentración de sustrato que provoca que la reacción alcance $1/2 V_{m\acute{a}x}$. K_m puede considerarse la afinidad entre la enzima correspondiente y el sustrato correspondiente. Según disminuye la K_m , aumenta la fuerza de unión entre la enzima correspondiente y el sustrato correspondiente. K_{cat} , también denominado número de recambio en una enzima, se refiere al número de moléculas de sustrato metabolizadas en 1 segundo en cada sitio activo enzimático cuando la enzima se activa a una tasa máxima y significa lo rápido que se produce realmente la reacción enzimática.

De acuerdo con el método de la presente divulgación, se calcula una puntuación de fármaco individual en la siguiente etapa asociando la puntuación de daño proteínico descrita anteriormente con una relación de fármaco-proteína.

La expresión "puntuación de fármaco" usada en la presente divulgación se refiere a un valor calculado con respecto a un fármaco predeterminado, por ejemplo, un agente antineoplásico, hallando una proteína diana implicada en la farmacodinámica o farmacocinética del fármaco, una proteína enzimática implicada en el metabolismo de fármacos, una proteína transportadora o una proteína de transporte, calculando las puntuaciones de daño proteínico de las proteínas y recopilando las puntuaciones.

En la presente divulgación, si se dañan dos o más proteínas implicadas en la farmacodinámica o farmacocinética de un agente antineoplásico, la puntuación del fármaco se calcula como una media de las puntuaciones de daño proteínico. Dicha media puede calcularse, por ejemplo, midiendo una media geométrica, una media aritmética, una media armónica, una media geométrica aritmética, una media armónica aritmética, una media armónica geométrica, una media pitagórica, una media de intercuartiles, una media cuadrática, una media truncada, una media winsorizada,

una media ponderada, una media geométrica ponderada, una media aritmética ponderada, una media armónica ponderada, una media de una función, una media generalizada, una media f-generalizada, un percentil, un valor máximo, un valor mínimo, una moda, una mediana, un intervalo medio, una tendencia central, una multiplicación simple o una multiplicación ponderada, o mediante una operación funcional de los valores calculados, aunque sin limitarse a ello.

La puntuación de fármaco puede calcularse ajustando las ponderaciones de una proteína diana implicada en la farmacodinámica o farmacocinética del fármaco correspondiente, es decir, el agente antineoplásico, una proteína enzimática implicada en el metabolismo de fármacos, una proteína transportadora o una proteína de transporte en consideración de las características farmacológicas. La ponderación puede asignarse considerando los parámetros farmacocinéticos de la correspondiente proteína enzimática metabolizante de fármaco, un grupo de población, una distribución de raza, etc. Además, aunque no interactúen directamente con el fármaco correspondiente, las proteínas que interactúan con precursor del fármaco correspondiente y los productos metabólicos del fármaco correspondiente, por ejemplo, proteínas implicadas en una ruta farmacológica, pueden considerarse, y las puntuaciones de daño proteínico de las mismas pueden combinarse para calcular la puntuación de fármaco. Además, las puntuaciones de daño proteínico de las proteínas que interactúan significativamente con las proteínas implicadas en la farmacodinámica o farmacocinética del fármaco correspondiente también pueden considerarse y combinarse para calcular la puntuación de fármaco. La información acerca de las proteínas implicadas en una ruta farmacológica del fármaco correspondiente, que interactúan significativamente con las proteínas en la ruta o implicadas en una ruta de transducción de señales de las mismas puede buscarse en bases de datos biológicas conocidas de manera pública tales como PharmGKB (Whirl-Carrillo *et al.*, *Clinical Pharmacology & Therapeutics* 2012; 92(4): 414-4171), The MIPS Mammalian Protein-Protein Interaction Database (Pagel *et al.*, *Bioinformatics* 2005; 21(6): 832-834), BIND (Bader *et al.*, *Biomolecular Interaction Network Database*, *Nucleic Acids Res.* 1 de enero de 2003; 31(1): 248-50), Reactome (Joshi-Tope *et al.*, *Nucleic Acids Res.* 1 de enero de 2005; 33 (número de publicación de base de datos): D428-32), etc.

En un aspecto ilustrativo de la presente divulgación, la puntuación de fármaco se calcula mediante la ecuación 3. La siguiente ecuación 3 puede modificarse de diversas maneras y, por tanto, la presente divulgación no se limita a ello.

[Ecuación 3]

$$S_d(g_1, \dots, g_n) = \left(\frac{1}{n} \sum_{i=1}^n g_i^p \right)^{\frac{1}{p}}$$

En la ecuación 3, S_d es una puntuación de fármaco de un fármaco d, n es el número de proteínas directamente implicadas en la farmacodinámica o farmacocinética del fármaco d o que interactúan con un precursor del fármaco correspondiente o productos metabólicos del fármaco correspondiente, por ejemplo, proteínas codificadas por uno más genes seleccionados de un grupo génico implicado en una ruta farmacológica, g_i es una puntuación de daño proteínico de una proteína directamente implicada en la farmacodinámica o farmacocinética del fármaco d o que interactúa con un precursor del fármaco correspondiente o productos metabólicos del fármaco correspondiente, por ejemplo, una proteína codificada por uno más genes seleccionados de un grupo génico implicado en una ruta farmacológica y p es un número real distinto de 0. En la ecuación 3, cuando el valor de p es 1, la puntuación de fármaco es una media aritmética, si el valor de p es -1, la puntuación de fármaco es una media armónica, y si el valor de p está cercano al límite 0, la puntuación de fármaco es una media geométrica.

En el método de la presente invención, la puntuación de fármaco se calcula mediante la ecuación 4.

[Ecuación 4]

$$S_d(g_1, \dots, g_n) = \left(\prod_{i=1}^n g_i^{w_i} \right)^{1/\sum_{i=1}^n w_i}$$

En la ecuación 4, S_d es una puntuación de fármaco de un fármaco d, n es el número de proteínas directamente implicadas en la farmacodinámica o farmacocinética del fármaco d o que interactúan con un precursor del fármaco correspondiente o productos metabólicos del fármaco correspondiente, por ejemplo, proteínas codificadas por uno más genes seleccionados de un grupo génico implicado en una ruta farmacológica, g_i es una puntuación de daño proteínico de una proteína directamente implicada en la farmacodinámica o farmacocinética del fármaco d o que interactúa con un precursor del fármaco correspondiente o productos metabólicos del fármaco correspondiente, por ejemplo, una proteína codificada por uno o más genes seleccionados de un grupo génico implicado en una ruta farmacológica y w_i es una ponderación asignada a g_i . Si todas las ponderaciones w_i tienen el mismo valor, la puntuación de fármaco S_d es una media geométrica de las puntuaciones de daño proteínico g_i . La ponderación puede asignarse considerando el tipo de proteína, clasificación farmacodinámica o farmacocinética de la proteína, parámetros

farmacocinéticos de la proteína enzimática de un fármaco correspondiente, un grupo de población o una distribución de raza.

5 En el caso de un método de cálculo de media geométrica usado en una realización ilustrativa de la presente divulgación, las ponderaciones se asignan equitativamente independientemente de la característica de una relación de fármaco-proteína. Sin embargo, es posible calcular una puntuación de fármaco asignando ponderaciones que consideran cada característica de una relación de fármaco-proteína como se describe en otra realización ilustrativa. Por ejemplo, pueden asignarse diferentes puntuaciones a una proteína diana de un fármaco y una proteína transportadora relacionada con el fármaco. Además, es posible calcular una puntuación de fármaco asignando
10 parámetros farmacocinéticos K_m , $V_{m\acute{a}x}$ y K_{cat}/K_m ponderaciones para la correspondiente enzima metabolizante de fármaco. Además, por ejemplo, como una proteína diana se considera más importante que una proteína transportadora en términos de acción farmacológica, puede asignarse una mayor ponderación, o una proteína transportadora o una proteína de transporte puede asignarse a ponderaciones altas con respecto a un fármaco cuya eficacia es sensible a una concentración, pero la presente divulgación no se limita a ello. La ponderación puede ajustarse minuciosamente
15 de acuerdo con las características de una relación entre un fármaco y una proteína relacionada con el fármaco y características de una interacción entre el fármaco y la proteína. Puede usarse un algoritmo sofisticado configurado para asignar una ponderación de una característica de una interacción entre un fármaco y una proteína. Por ejemplo, una proteína diana y una proteína transportadora pueden asignarse a 2 puntos y 1 punto, respectivamente.

20 En la descripción anterior, solamente la proteína que interactúa directamente con un fármaco, es decir, un agente antineoplásico, se ha ejemplificado. Sin embargo, la capacidad predictiva de las ecuaciones anteriores puede mejorarse usando información acerca de la proteína que interactúa con un precursor del fármaco correspondiente o productos metabólicos del fármaco correspondiente, interactuando la proteína significativamente con proteínas implicadas en la farmacodinámica o farmacocinética del fármaco correspondiente y la proteína implicada en una ruta
25 de transducción de señales de la misma. Es decir, utilizando información acerca de una red de interacción de proteína-proteína o ruta farmacológica, es posible usar información acerca de diversas proteínas relevantes para ello. Es decir, incluso si no se encuentra una variación significativa en la proteína que interactúa directamente con el fármaco de modo que no haya puntuación de daño proteínico calculada con respecto a la proteína o no haya daño (por ejemplo, 1,0 punto cuando se aplica el algoritmo SIFT), puede usarse una media (por ejemplo, una media geométrica) de las
30 puntuaciones de daño proteínico de las proteínas que interactúan con la proteína o implicadas en la misma ruta de transducción de señales de la proteína como puntuación de daño proteínico de la proteína para usarla para calcular una puntuación de fármaco.

35 La puntuación de fármaco individual puede calcularse con respecto a todos los fármacos a partir de los que puede adquirirse información acerca de una o más proteínas asociadas o con respecto a algunos fármacos seleccionados de los fármacos. Además, la puntuación de fármaco individual puede convertirse en una categoría.

40 El método de la presente divulgación puede incluir además una etapa de determinar si el fármaco, es decir, el agente antineoplásico, se usará individualmente usando la puntuación de fármaco individual descrita anteriormente.

45 Aunque la puntuación de fármaco individual puede aplicarse a cada uno de todos los fármacos, puede ser más útil cuando se aplica a fármacos clasificados por enfermedad, característica clínica o actividad, o fármacos médicamente comparables. El sistema de clasificación de fármacos que puede usarse en la presente divulgación puede incluir, por ejemplo, códigos ATC (sistema de clasificación química terapéutica anatómica), una lista de fármacos con marcadores farmacogenómicos conocidos que pueden influir en la información de efecto del fármaco descrita en la etiqueta del fármaco, o una lista de fármacos retirados del mercado debido a efectos secundarios de los mismos, etc.

El método de la presente divulgación puede incluir además una etapa de calcular una puntuación de prescripción.

50 La expresión "puntuación de prescripción" usada en la presente divulgación se refiere a una puntuación calculada recopilando las puntuaciones de fármaco determinadas con respecto a los fármacos, respectivamente, cuando se administran dos o más fármacos al mismo tiempo o a una corta distancia de tiempo suficiente para afectar significativamente a las acciones farmacológicas de los mismos. En la presente divulgación, cuando se determinan dos o más fármacos basándose en el orden de prioridad entre los fármacos y tienen que administrarse al mismo
55 tiempo, la puntuación de prescripción puede calcularse recopilando las puntuaciones de fármaco determinadas con respecto a los fármacos respectivos. Por ejemplo, si no hay proteína que interactúe de manera común con los fármacos, la puntuación de prescripción puede calcularse promediando simplemente, o recopilando o multiplicando las puntuaciones de fármaco de los fármacos. Si hay una proteína que interactúa comúnmente con los fármacos, la puntuación de prescripción puede calcularse asignando, por ejemplo, una doble ponderación a una puntuación de
60 daño proteínico de la correspondiente proteína que interactúa comúnmente para calcular puntuaciones de fármaco de los respectivos fármacos y después recopilando las puntuaciones de fármaco correspondientes.

La puntuación de prescripción se proporciona para determinar lo apropiado que son los fármacos o el riesgo de los fármacos incluidos en una prescripción aplicada a un individuo más allá de los efectos de los fármacos respectivos.
65 En este aspecto, el método de la presente divulgación puede incluir además una etapa de determinar lo apropiada que es una prescripción o el riesgo de una prescripción aplicada a un individuo.

El método de la presente divulgación puede realizarse con el propósito de prevenir los efectos secundarios de un agente antineoplásico, aunque sin limitarse a ello.

5 La FIG. 1 es un diagrama de flujo que ilustra cada etapa de un método para proporcionar información para seleccionar un agente antineoplásico usando variación de secuencia genómica individual de acuerdo con una realización ilustrativa de la presente divulgación. En una realización ilustrativa de la presente divulgación, el método para proporcionar información para seleccionar un agente antineoplásico se realiza secuencialmente (1) introduciendo o recibiendo información de secuencia genómica de usuario individual (S100), (2) introduciendo o recibiendo información relevante
10 para un agente antineoplásico (S110), (3) determinando la información de variación de secuencia genómica del usuario individual (S120), (4) calculando una puntuación de daño proteínico individual con respecto al agente antineoplásico (S130), (5) calculando una puntuación de fármaco individual con respecto al agente antineoplásico (S140), (6) presentando la puntuación de fármaco y clasificando los fármacos por asignación de categorías o determinando el orden de prioridad entre los fármacos de acuerdo con las clasificaciones de puntuaciones de fármaco (S150) y (7)
15 seleccionando un agente antineoplásico en consideración de la puntuación de fármaco y la prioridad y calculando una puntuación de prescripción (S160).

El método de la presente divulgación puede incluir además una etapa de ayudar a un médico a cargo de la prescripción a tomar una decisión proporcionando un procedimiento de cálculo farmacogenómico y un fundamento para calcular la puntuación de fármaco como información en forma de un diagrama, una lista, explicación, etc. cuando se selecciona una puntuación de fármaco clasificada por asignación de categorías como se describe anteriormente. Es decir, el método de acuerdo con la presente divulgación puede incluir además una etapa de proporcionar una o más
20 informaciones entre información de variación de secuencia genética, una puntuación de variación de secuencia genética, una puntuación de daño proteínico, una puntuación de fármaco e información usada para calcular los mismos.

25 En otro aspecto, la presente divulgación se refiere a un sistema para seleccionar un agente antineoplásico usando variación de secuencia genómica individual, que incluye: una base de datos a partir de la que puede buscarse o extraerse información relevante para un gen o proteína relacionada con un agente antineoplásico aplicable a un individuo; una unidad de comunicación que es accesible para la base de datos; un primer módulo de cálculo que se configura para calcular información de una o más variaciones de secuencia genética implicadas en la farmacodinámica o farmacocinética del agente antineoplásico basándose en la información; un segundo módulo de cálculo de se
30 configura para calcular una puntuación de daño proteínico individual usando la información de variación de secuencia genética; un tercer módulo de cálculo que se configura para calcular una puntuación de fármaco individual asociando la puntuación de daño proteínico individual con una relación de fármaco-proteína; y una unidad de presentación que se configura para presentar los valores calculados por los módulos de cálculo.
35

En la presente divulgación, un módulo puede significar una combinación funcional o estructural de equipo para implementar el espíritu técnico de la presente divulgación y programa informático para dirigir el equipo. Por ejemplo, el módulo puede ser un código predeterminado y una unidad lógica de un recurso de equipo por el que se ejecuta el código predeterminado. Es obvio para los expertos en la materia que el módulo no significa necesariamente códigos físicamente conectados o un tipo de equipo.
40

La expresión "módulo de cálculo" usada en la presente divulgación puede significar un código predeterminado y una unidad lógica de un recurso de equipo por el que se ejecuta el código predeterminado para calcular cada puntuación basándose en la puntuación de variación de secuencia genética, puntuación de daño proteínico, puntuación de fármaco e información como fundamento para el cálculo de los mismos con respecto a un agente antineoplásico y un gen a analizar de acuerdo con la presente divulgación, pero no significa necesariamente códigos conectados físicamente o un tipo de equipo.
45

50 El sistema de aceleradora de acuerdo con la presente divulgación puede incluir además un cuarto módulo de cálculo que se configura para determinar si el agente antineoplásico es aplicable al individuo usando la puntuación de fármaco individual calculada por el tercer módulo de cálculo.

El sistema de acuerdo con la presente divulgación puede incluir además un quinto módulo de cálculo que se configura para calcular una puntuación de prescripción recopilando las puntuaciones de fármaco determinadas con respecto a fármacos respectivos si se determinan dos o más fármacos basándose en el orden de prioridad entre los fármacos y la necesidad de administrarlos al mismo tiempo.
55

El sistema de aceleradora de acuerdo con la presente divulgación puede incluir además una interfaz de usuario que se configura para introducir una lista de agentes antineoplásicos por el usuario, o acceso a una base de datos que contiene información acerca de agentes antineoplásicos y extraer información relevante, y calcular de ese modo y proporcionar una puntuación de fármaco del fármaco.
60

El sistema de acuerdo con la presente divulgación puede incluir además una unidad de presentación que se configura para presentar los valores calculados por los respectivos módulos de cálculo o un proceso de cálculo para determinar el orden de prioridad entre los fármacos e información como fundamento para el cálculo o determinación.
65

En el sistema de acuerdo con la presente divulgación, la base de datos o un servidor que incluye información de acceso, la información calculada y la interfaz de usuario conectada al mismo pueden usarse vinculados entre sí.

5 Si se produce nueva información farmacológica/bioquímica con respecto a una relación de fármaco-proteína, el sistema de acuerdo con la presente divulgación puede actualizarse inmediatamente para usarse para selección mejorada adicional de un agente antineoplásico. En una realización ilustrativa de la presente divulgación, cuando se actualiza la base de datos o base de conocimientos, la información de variación de secuencia génica, puntuación de variación de secuencia génica, puntuación de daño proteínico, puntuación de fármaco y la información como
10 fundamento para el cálculo de los mismos almacenada en los respectivos módulos de cálculo se actualizan.

La FIG. 2 ilustra esquemáticamente la configuración de un sistema para seleccionar un agente antineoplásico usando variación de secuencia genómica individual de acuerdo con una realización ilustrativa de la presente divulgación. Un sistema 10 de la presente divulgación puede incluir una base de datos (DB) 100 a partir de la que puede buscarse o
15 extraerse información relevante para un gen o proteína relacionada con un agente antineoplásico, una unidad de comunicación 200, una interfaz de usuario o terminal 300, una unidad de cálculo 400 y una unidad de presentación 500.

En el sistema de acuerdo con la presente divulgación, la interfaz de usuario o terminal 300 puede configurarse para solicitar un procesamiento para seleccionar un agente antineoplásico usando variaciones de secuencia genómica individuales a un servidor y recibir un resultado de un servidor y/o almacenarlo. La interfaz de usuario o terminal 300 puede consistir en un terminal, tal como un teléfono móvil, un PC (ordenador personal), una tableta, un asistente digital personal (PDA) y un lector digital, que incluye una memoria y tiene una función de comunicación móvil con una capacidad de cálculo usando un microprocesador.
20

En el sistema de acuerdo con la presente divulgación, el servidor es un medio para proporcionar acceso a la base de datos 100 con respecto a un agente antineoplásico, una variación génica o una relación de fármaco-proteína y está conectado a la interfaz de usuario o terminal 300 a través de la unidad de comunicación 200 para intercambiar diversos tipos de información. En el presente documento, la unidad de comunicación 200 puede usar cualquier método de comunicación independientemente de si es con cable o inalámbrica, incluyendo no solamente comunicación en el mismo equipo, sino también una red de área local (LAN), una red de área metropolitana (MAN), una red de área ancha (WAN), Internet, redes de comunicación móvil 2G, 3G y 4G, Wi-Fi, Wibro, etc. La base de datos 100 puede instalarse directamente en el servidor y también puede conectarse a diversas bases de datos de ciencia biológica accesibles mediante Internet dependiendo de los propósitos.
25

En el sistema de acuerdo con la presente divulgación, la unidad de cálculo 400 puede incluir un primer módulo de cálculo 410 que se configura para calcular información de una o más variaciones de secuencia génica implicadas en la farmacodinámica o farmacocinética del agente antineoplásico usando la información recogida/introducida, un segundo módulo de cálculo 420 que se configura para calcular una puntuación de daño proteínico individual y un tercer módulo de cálculo 430 que se configura para calcular una puntuación de fármaco individual, como se describe anteriormente.
30

El método de acuerdo con la presente divulgación puede implementarse por equipo, programas inalterables, programa informático o combinaciones de los mismos. Cuando el método se implementa por programa informático, un medio de almacenamiento puede incluir cualquier almacenamiento o medio de transmisión legible por un dispositivo tal como un ordenador. Por ejemplo, el medio legible por ordenador puede incluir una ROM (memoria solamente de lectura); una RAM (memoria de acceso aleatorio); un medio de almacenamiento de disco magnético; un medio de almacenamiento óptico; un dispositivo de memoria flash; y otro medio de transmisión de señales eléctricas, ópticas o acústicas.
35

En este aspecto, la presente divulgación proporciona un medio legible por ordenador que incluye un medio de ejecución para ejecutar un procesador que realiza una operación que incluye: una etapa de adquirir información de una o más variaciones de secuencia génica implicadas en la farmacodinámica o farmacocinética de un agente antineoplásico a partir de información de secuencia genómica individual; una etapa de calcular una puntuación daño proteínico individual usando la información de variación de secuencia génica; y una etapa de calcular una puntuación de fármaco individual asociando la puntuación daño proteínico individual con una relación de fármaco-proteína.
40

El procesador puede incluir además una etapa de determinar si el agente antineoplásico se usará por el individuo usando el procesador de puntuación de fármaco individual.
45

En otro aspecto, la presente divulgación proporciona una composición de biomarcadores para predecir el efecto secundario de un agente antineoplásico.
50

Los genes que pueden estar contenidos en la composición de biomarcadores de acuerdo con la presente divulgación incluye ABCC1, ABCG2, APEX1, CBR3, ATP7A, CBR3, CTH, CUL9, CYP1A1, CYP2C9, CYP2D6, CYP4B1, DPEP1, DPEP2, DPYD, DRD3, EPHX1, FMO2, FMO3, GGT1, GGT5, GGT6, GNB3, GSTA2, GSTM1, MGMT, MLH1, MSH2,
55

NAT1, OPRM1, PDE5A, PTGS1, SERPINA6, SLC15A2, SLC22A1, SLC22A2, SLC22A8, SLC22A16, SLC28A2, SLC28A3, SPG7, TP53, ABCB1, ABCC2, AOX1, CHRM2, CYP1B1, CYP3A4, CYP3A5, CYP4F2, DBH, ERCC2, GGH, GGT7, GMPS, MTHFR, NUDT15, PDE3A, PDE4C, SLC22A5, SLC01B1, TLR3, UGT1A1, XDH, etc., aunque sin limitarse a ellos. Como la aparición de un efecto secundario de un agente antineoplásico puede predecirse analizando variaciones de estos genes o sus proteínas, puede usarse como marcador un agente que puede detectarlos.

A partir de ahora en el presente documento, la presente divulgación se describirá en detalle a través de ejemplos. Sin embargo, los siguientes ejemplos son solamente con propósitos ilustrativos y el alcance de la presente divulgación no está limitado por los ejemplos.

Ejemplo 1. Análisis y aplicación de información de variación de secuencia genómica individual de pacientes pediátricos con leucemia que muestran signos de aviso de efectos secundarios graves durante el tratamiento con agente antineoplásico (busulfano)

El trasplante de médula ósea es uno de los métodos de tratamiento más importantes para tratar neoplasias hemáticas tales como leucemia. Para el trasplante de médula ósea, la médula ósea de un paciente tiene que retirarse en primer lugar usando dos métodos: irradiación corporal total (TBI) y tratamiento farmacológico usando fármacos tales como busulfano. El busulfano es un agente alquilante representativo y puede reemplazar la irradiación corporal total. Sin embargo, tiene un intervalo terapéutico relativamente estrecho. Si la concentración de fármaco es mayor que el intervalo terapéutico, se produce enfermedad venooclusiva hepática (VOD) y toxicidad grave, tal como neurotoxicidad, relevante para el fármaco. Y, si la concentración de fármaco es menor que el intervalo terapéutico, la probabilidad de fracaso del injerto o recidiva aumenta. Particularmente, los niños muestran grandes diferencias en la farmacocinética de busulfano entre individuos. Por lo tanto, se usa busulfano bajo supervisión terapéutica del fármaco (TDM). La toxicidad de busulfano incluye fibrosis pulmonar intersticial habitualmente denominada "pulmón de busulfano", hiperpigmentación, epilepsia, enfermedad venooclusiva (VOD), náuseas, trombocitopenia, etc. La IARC (International Agency for Research on Cancer) clasifica el busulfano como uno de los carcinógenos del grupo 1.

Para comprobar si es posible identificar un grupo de riesgo con respecto al tratamiento con busulfano a través del método para proporcionar información para seleccionar un agente antineoplásico usando variaciones de secuencia genómica individuales de la presente divulgación, se realizó el siguiente experimento.

En primer lugar, se realizó un análisis sobre 146 pacientes pediátricos con leucemia que mostraban signos de aviso de efectos secundarios graves durante tratamiento con el agente antineoplásico busulfano (Myleran, GlaxoSmithKline, Busulfex IV, Otsuka America Pharmaceutical, Inc.) para retirar la médula ósea como precondicionamiento para trasplante de médula ósea. Entre los 146 pacientes, 124 mostraron neoplasias hemáticas malignas. El perfil de efectos secundarios de los 146 pacientes se da en la tabla 2.

[Tabla 2]

N.º	Efectos secundarios	Número de pacientes
1	Muerte relacionada con toxicidad	22
2	Toxicidad intermedia de grado 3 o mayor	11
3	Toxicidad intermedia de grado 4 o mayor	1
4	Toxicidad a corto plazo de grado 3 o mayor	31
5	Toxicidad a corto plazo de grado 4 o mayor	2
6	Enfermedad venooclusiva hepática	18
7	Muerte	36
8	ABC 170 o mayor	71

Se realizó secuenciación de exoma dirigida 280x usando Ion Proton (Life Technology) para adquirir la información de variación de secuencia genómica individual de los 146 pacientes. Como alternativa, puede realizarse secuenciación genómica completa de adquisición de la información del genoma individual completo o secuenciación de exoma completo de adquisición de la información de los genes que codifican proteínas.

Los fragmentos de secuencia analizados experimentaron limpieza de datos y comprobación de calidad y se dieron salida en forma de archivos SAM (mapa de alineación de secuencias) y BAM (mapa de alineación binaria) alineados con una secuencia de grupo de referencia humana (por ejemplo, HG19). El resultado de alineación limpiada se dio salida en forma de archivo VCF (formato de identificación de variación) mientras se detectaban variaciones tales como variaciones mononucleotídicas (SNV) e indel usando herramientas de programa informático tales como SAMTools:pileup, SAMTools:mpileup, GATK:recalibration, GATK:realignment, y similares.

Después de introducir el archivo VCF que incluye la información de variación de secuencia génica y de que se calculara la puntuación de variación de secuencia génica descrita anteriormente v_i para cada variante, se calculó una puntuación de daño proteínico individual S_g usando la ecuación 2. Después, tras comparar la puntuación de daño proteínico de los grupos de pacientes que mostraban diferentes efectos secundarios con la puntuación de daño proteínico del grupo de control, se seleccionaron 48 genes que mostraban diferencia estadísticamente significativa para los efectos secundarios respectivos como sigue. Entre ellos, 28 genes que mostraban significación estadística mayor basándose en el valor de p se clasificaron como grupo 1. Los siguientes genes son los que muestran alta relevancia para la farmacodinámica o farmacocinética de busulfano y sus productos metabólicos.

(Grupo 1)

ABCC1, ABCG2, ATP7A, CBR3, CTH, CUL9, CYP1A1, CYP2C9, DPEP1, DPEP2, DPYD, DRD3, FMO2, GGT1, GGT5, GGT6, GSTA2, GSTM1, MSH2, NAT1, PDE5A, PTGS1, SLC15A2, SLC22A1, SLC22A16, SLC22A2, SLC28A2, SPG7

(Grupo 2)

ABCB1, ABCC2, AOX1, CHRM2, CYP1B1, CYP2D6, CYP3A4, DBH, EPHX1, GGT7, GMPS, MTHFR, PDE3A, PDE4C, SERPINA6, SLC22A5, SLC22A8, TLR3, TP53, UGT1A1

Para los grupos génicos seleccionados, se calculó la puntuación de fármaco individual para los 146 pacientes pediátricos con leucemia usando el método de la presente divulgación. Más específicamente, después de calcular la puntuación de variación de secuencia génica a partir de la información de variación de secuencia génica individual usando el algoritmo SIFT, se calculó la puntuación de daño proteínico individual para los 48 genes usando la ecuación 2. Después, tras calcular la puntuación de fármaco individual para busulfano usando la ecuación 4, se realizó análisis estadístico para cada grupo. La distribución de la puntuación de daño proteínico del grupo positivo a efectos secundarios de busulfano y el grupo negativo se muestran en la tabla 3.

[Tabla 3]

	Grupo positivo a efectos secundarios de busulfano (n = 90)	Grupo negativo para efectos secundarios de busulfano (n = 56)	valor de p
Puntuación media de fármaco (fármaco de grupo 1)	0,195406408	0,468837374	0,072382486
Puntuación media de fármaco (fármacos de grupo 1 y 2)	0,273943475	0,5030009	0,034605505
Puntuación de daño proteínico de grupo 1			
SLC22A16	0,3236±0,01	0,55±0,33	2,03E-13
CBR3	0,0829±0,04	0,185±0,16	2,83E-07
FMO2	1,0E-8±0	0,1556±0,36	1,01E-06
PTGS1	0,3744±0,01	0,4543±0,21	1,07E-05
SLC15A2	0,1114±0,04	0,2232±0,32	0,000400521
DRD3	0,54±0	0,5809±0,15	0,001222641
SLC22A2	0,1±0	0,1625±0,23	0,001367927
GGT6	0,0578±0,01	0,0635±0,01	0,001375077
CUL9	0,8383±0,36	0,962±0,19	0,005778935
GSTA2	0,458±0,17	0,5249±0,15	0,006675864
SPG7	0,6653±0,34	0,8593±0,27	0,009214375
CYP1A1	0,2903±0,43	0,4599±0,48	0,013394813
GGT1	0,1±0,29	0,2655±0,43	0,014831285
ABCC1	0,432±0,26	0,5392±0,3	0,020942712
DPEP1	0,529±0,37	0,6555±0,38	0,0220818

ES 2 891 903 T3

(continuación)

Puntuación de daño proteínico de grupo 1			
ATP7A	0,5597±0,46	0,7468±0,41	0,023348076
GGT5	0,2332±0,3	0,4314±0,41	0,029370818
GSTM1	0,8037±0,35	0,9016±0,27	0,03104692
SLC22A1	0,4535±0,3	0,5463±0,29	0,031693911
DPEP2	0,8528±0,33	0,9597±0,19	0,037121203
CTH	0,48±0,44	0,6665±0,43	0,0389814
PDE5A	0,2971±0,03	0,3213±0,11	0,040084719
DPYD	0,1675±0,04	0,2002±0,16	0,043865116
ABCG2	0,4891±0,34	0,6836±0,35	0,045862428
CYP2C9	0,9679±0,16	1±0	0,046448221
SLC28A2	0,7307±0,39	0,8613±0,31	0,047310042
MSH2	0,7484±0,43	0,8897±0,3	0,048069264
NAT1	0,9433±0,13	0,9973±0,03	0,048855771
Puntuación de daño proteínico de grupo 2			
SERPINA6	0,0822±0,04	0,1017±0,09	0,050821316
EPHX1	0,1653±0,28	0,3055±0,43	0,05389215
GMPS	0,6267±0,43	0,8031±0,36	0,055682706
CYP2D6	0,1833±0,38	0,3415±0,46	0,058740693
MTHFR	0,1745±0,28	0,2669±0,37	0,066912194
ABCC2	0,6567±0,4	0,8253±0,33	0,068738095
TLR3	0,314±0,47	0,5634±0,49	0,069997413
AOX1	0,6036±0,47	0,8249±0,34	0,076843585
GGT7	0,9765±0,13	1±0	0,079490244
CHRM2	0,9763±0,14	1±0	0,07962271
SLC22A8	0,5261±0,32	0,6255±0,34	0,080067267
CYP3A4	0,9589±0,17	0,9903±0,08	0,081583332
DBH	0,7415±0,31	0,8303±0,27	0,081601181
UGT1A1	0,5425±0,47	0,6935±0,44	0,085536737
CYP1B1	0,242±0,08	0,2681±0,13	0,088335873
PDE4C	0,6012±0,48	0,7217±0,44	0,0936022
TP53	0,2161±0,2	0,2867±0,29	0,095578636
PDE3A	0,4937±0,18	0,5406±0,15	0,097645758
SLC22A5	0,9283±0,25	0,9824±0,13	0,107609865
ABCB1	0,7691±0,38	0,8404±0,35	0,210600695

Como se observa a partir de la tabla 3, se observó diferencia estadísticamente significativa entre el grupo positivo a efectos secundarios de busulfano (n = 90) y el grupo negativo a efectos secundarios de busulfano (n = 56) cuando se calculaban la puntuación de daño proteínico y la puntuación de fármaco individual a partir de la información de variación de secuencia génica usando los 28 genes seleccionados del grupo 1 (valor de $p < 0,05$).

5

Además, cuando se calculaba la puntuación de fármaco individual a partir de la variación de secuencia génica incluyendo además los genes del grupo 2, se observaba diferencia estadísticamente significativa entre los dos grupos (valor de $p < 0,05$).

5 A través de estos resultados, se demostró que es posible distinguir significativamente un grupo que muestra signos de aviso de efectos secundarios graves durante el tratamiento con busulfano (grupo de riesgo) de un grupo que no está en riesgo calculando las puntuaciones de fármaco individuales a través de análisis de la información de variación de secuencia genómica individual de acuerdo con la presente divulgación y es posible evitar los efectos secundarios indeseados.

10 Con el resultado descrito anteriormente, es posible predecir un grupo con una alta probabilidad de efectos secundarios cuando se administra busulfano a un paciente pediátrico con leucemia, de acuerdo con el método de la presente divulgación, y también es posible ajustar la concentración de fármaco o usar un método de tratamiento alternativo o método de actuación para un grupo de alto riesgo.

15 **Ejemplo 2. Análisis y aplicación de información de variación de secuencia genómica individual de pacientes con leucemia que muestran leucopenia y toxicidad hepática durante tratamiento con agente antineoplásico (mercaptopurina)**

20 Se usa mercaptopurina junto con otros fármacos para tratamiento de mantenimiento en el tratamiento de leucemia linfocítica aguda. La mercaptopurina se metaboliza en el organismo y se convierte en nucleótidos de tioguanina (TGN), que muestra efecto antineoplásico inhibiendo la síntesis del ADN. La mercaptopurina se metaboliza por la tiopurina S-metiltransferasa. Cuando el paciente tiene variación en el gen TPMT que codifica la proteína, pueden producirse efectos secundarios graves potencialmente mortales tales como leucopenia, agranulocitosis o toxicidad hepática. En particular, cuando el paciente tiene variación del gen TPMT en ambos alelos, puede producirse supresión potencialmente mortal de la médula ósea. Y, incluso cuando la variación del gen TPMT se produce solamente en un alelo, se observa mayor toxicidad y supresión de la médula ósea en comparación con cuando ambos alelos son normales.

30 Recientemente, a partir de un estudio sobre pacientes asiáticos que tienen genes TPMT normales, pero muestran la toxicidad y supresión de la médula ósea descritas anteriormente, se dilucidó la relación entre el gen NUDT15 y el fármaco basado en tiopurina. En un estudio de cohorte prospectiva (Yang, J. J., Landier, W., Yang, W., Liu, C., Hageman, L., Cheng, C., ... y Wong, F. L. (2015). Inherited NUDT15 variant is a genetic determinant of mercaptopurine intolerance in children with acute lymphoblastic leukemia. Journal of Clinical Oncology, JCO-2014.), se confirmó a través de secuenciación de todo el genoma que rs1142345, un SNP del gen TPMT, y rs116855232, un SNP del gen NUDT15, están asociados con la dosis de tratamiento de mantenimiento con mercaptopurina. A partir de la comparación de la dosis terapéutica de mercaptopurina, se descubrió que la dosis terapéutica se disminuía cuando estaban presentes los SNP. Como los médicos reducen la dosis terapéutica cuando se produce toxicidad grave durante el tratamiento, puede observarse que los SNP están muy relacionados con la aparición de toxicidad durante el tratamiento.

45 Para investigar si puede identificarse un grupo de riesgo en tratamiento con mercaptopurina usando el método para proporcionar información para seleccionar un agente antineoplásico usando variación de secuencia genómica individual de la presente divulgación, se realizó el siguiente experimento.

Se analizó un grupo de pacientes al que se había administrado tratamiento con mercaptopurina durante aproximadamente 2 años. El perfil de efectos secundarios del grupo de pacientes se da en la tabla 4.

[Tabla 4]

N.º	Efectos secundarios
1	Fiebre neutropénica
2	ANC < 500
3	Hiperbilirrubinemia de grado 3 o mayor
4	Toxicidad hepática (GOT) de grado 2 o mayor
5	Toxicidad hepática (GPT) de grado 3 o mayor
6	Toxicidad hepática de 3 o mayor en grupo con neutropenia

50 Se realizó secuenciación de exoma dirigida 280x usando Ion Proton (Life Technology) para adquirir la información de variación de secuencia genómica individual del grupo de pacientes. Como alternativa, puede realizarse secuenciación genómica completa de adquisición de la información del genoma individual completo o secuenciación de exoma completo de adquisición de la información de los genes que codifican proteínas.

Los fragmentos de secuencia analizados experimentaron limpieza de datos y comprobación de calidad y se dieron salida en forma de archivos SAM (mapa de alineación de secuencias) y BAM (mapa de alineación binaria) alineados con una secuencia de grupo de referencia humana (por ejemplo, HG19). El resultado de alineación limpiada se dio salida en forma de archivo VCF (formato de identificación de variación) mientras se detectaban variaciones tales como variaciones mononucleotídicas (SNV) e indel usando herramientas de programa informático tales como SAMTools:pileup, SAMTools:mpileup, GATK:recalibration, GATK:realignment, y similares.

Después de introducir el archivo VCF que incluye la información de variación de secuencia génica y de que se calculara la puntuación de variación de secuencia génica descrita anteriormente v_i para cada variante, se calculó una puntuación de daño proteínico individual S_g usando la ecuación 2. Después, tras comparar la puntuación de daño proteínico de los grupos de pacientes que mostraban diferentes efectos secundarios con la puntuación de daño proteínico del grupo de control, se seleccionaron 32 genes que mostraban diferencia estadísticamente significativa como sigue. Entre ellos, los genes que mostraban significación estadística mayor basándose en el valor de p se clasificaron como grupo 1. Los siguientes genes son los que muestran alta relevancia para la farmacodinámica o farmacocinética de mercaptopurina y sus productos metabólicos.

(Grupo 1)

ABCC1, APEX1, CBR3, CUL9, CYP2D6, CYP4B1, DPEP1, DPYD, EPHX1, FMO3, GNB3, MGMT, MLH1, OPRM1, PDE5A, SERPINA6, SLC15A2, SLC22A2, SLC22A8, SLC28A3, TP53

(Grupo 2)

CHRM2, CYP1B1, CYP3A5, CYP4F2, ERCC2, FMO2, GGH, NUDT15, PDE4C, SLC01B1, XDH

Para los grupos génicos seleccionados, se calculó la puntuación de fármaco individual usando el método de la presente divulgación. Más específicamente, después de calcular la puntuación de variación de secuencia génica a partir de la información de variación de secuencia génica individual usando el algoritmo SIFT, se calculó la puntuación de daño proteínico individual para los 32 genes usando la ecuación 2. Después, tras calcular la puntuación de fármaco individual para 6-mercaptopurina usando la ecuación 4, se realizó análisis estadístico para cada grupo. La distribución de la puntuación de daño proteínico del grupo positivo a efectos secundarios de 6-mercaptopurina y el grupo negativo se muestran en la tabla 5.

[Tabla 5]

		Grupo positivo a efectos secundarios de mercaptopurina (n = 6)	Grupo negativo a efectos secundarios de mercaptopurina (n = 34)	valor de p
Puntuación media de fármaco (fármaco de grupo 1)		0,1034±0,09	0,1890±0,10	0,024589877
Puntuación media de fármaco (fármacos de grupo 1 y 2)		0,0934±0,05	0,1749±0,09	0,021586521
Puntuación de daño proteínico (grupo 1)	DPEP1	0,24±0	0,7741±0,35	2,56E-11
	ABCC1	0,32±0	0,5316±0,32	0,000147391
	MGMT	0,7706±0,36	1±0	0,000382075
	SLC22A8	0,3166±0,01	0,5166±0,32	0,000416193
	CYP2D6	0,018±0	0,2319±0,41	0,00151514
	EPHX1	1,0E-8±0	0,1957±0,39	0,001937214
	APEX1	0,8209±0,36	1±0	0,003093326
	CYP4B1	0,7962±0,41	1±0	0,003093326
	CUL9	0,8047±0,39	1±0	0,003148978
	MLH1	0,8418±0,36	1±0	0,007651081
	TP53	0,17±0	0,2822±0,29	0,011587433
	OPRM1	0,3747±0,49	0,7171±0,46	0,01443295
	PDE5A	0,3±0	0,3043±0,01	0,016470184
	SERPINA6	0,07±0	0,1147±0,12	0,01668446
	FMO3	0,2853±0,35	0,6381±0,38	0,01974977
	DPYD	0,18±0	0,2765±0,27	0,021839396
	SLC15A2	0,1215±0,06	0,2519±0,34	0,029659527

(continuación)

		Grupo positivo a efectos secundarios de mercaptopurina (n = 6)	Grupo negativo a efectos secundarios de mercaptopurina (n = 34)	valor de <i>p</i>
	CBR3	0,0993±0,1	0,1811±0,1	0,039359969
	SLC28A3	0,8103±0,3	0,9569±0,2	0,039421696
	GNB3	0,9126±0,29	1±0	0,041559264
	SLC22A2	0,1±0	0,1794±0,26	0,041559264
Puntuación de daño proteínico (grupo 2)	XDH	0,9471±0,19	1±0	0,057393496
	FMO2	0,125±0,34	0,3333±0,48	0,058959933
	CYP1B1	0,1775±0,09	0,2415±0,1	0,063676645
	SLCO1B1	0,4394±0,36	0,5998±0,3	0,075169523
	CHRM2	0,9482±0,21	1±0	0,080187516
	ERCC2	0,9547±0,18	1±0	0,080187516
	GGH	0,9518±0,2	1±0	0,080187516
	CYP3A5	0,9435±0,23	1±0	0,080374542
	CYP4F2	0,175±0,4	0,4468±0,5	0,0910071
	PDE4C	0,4343±0,53	0,7454±0,44	0,092029759
	NUDT15	0,48±0,49	0,7646±0,41	0,093049258

Como se observa a partir de la tabla 5, se observó diferencia estadísticamente significativa entre el grupo positivo a efectos secundarios de mercaptopurina (n = 6) y el grupo negativo a efectos secundarios de mercaptopurina (n = 34) cuando se calculaban la puntuación de daño proteínico y la puntuación de fármaco individual a partir de la información de variación de secuencia génica usando los genes del grupo 1 (valor de $p < 0,05$).

Además, cuando se calculaba la puntuación de fármaco individual a partir de la variación de secuencia génica incluyendo además los genes del grupo 2, se observaba diferencia estadísticamente significativa entre los dos grupos.

A través de estos resultados, se demostró que es posible distinguir significativamente un grupo que muestra signos de aviso de efectos secundarios graves durante el tratamiento con mercaptopurina (grupo de riesgo) de un grupo que no está en riesgo calculando las puntuaciones de fármaco individuales a través de análisis de la información de variación de secuencia genómica individual de acuerdo con la presente divulgación y es posible evitar los efectos secundarios indeseados.

Con el resultado descrito anteriormente, es posible predecir un grupo con una alta probabilidad de efectos secundarios cuando se administra mercaptopurina a un paciente, de acuerdo con el método de la presente divulgación, y también es posible ajustar la concentración de fármaco o usar un método de tratamiento alternativo o método de actuación para un grupo de alto riesgo.

Ejemplo 3. Demostración de validez del método para personalizar la selección de fármacos basándose en información de variación de secuencia genómica individual

Los resultados de estudio fiables acerca de la información de variación de secuencia genómica individual y una diferencia individual en la respuesta farmacodinámica han sido muy limitados hasta ahora. Los estudios realizados hasta ahora han seguido un paradigma de un estudio de control de casos en que se estudia una diferencia individual en la sensibilidad comparando un grupo con una variación específica con un grupo con la variación específica para cada fármaco. En este paradigma de estudio, tiene que realizarse un estudio de control de casos costoso para cada una de todas las combinaciones de pares de numerosas variaciones de secuencia y numerosos fármacos, lo que es prácticamente imposible. Por el contrario, el método para personalizar la selección de fármacos de acuerdo con la presente divulgación es aplicable a todas las secuencias génicas sin requerir un estudio de control de casos costoso. La presente divulgación proporciona un método que puede calcular una puntuación de daño proteínico individual y una puntuación de fármaco individual simplemente calculando una variación de secuencia genómica. Por lo tanto, el método tiene una ventaja de poder hacer una deducción para personalizar la selección de fármacos con respecto a combinaciones entre todas las variaciones de secuencia genómica y todos los fármacos.

Para evaluar la validez de un resultado de selección de fármacos personalizada de acuerdo con el método de la presente divulgación, se seleccionaron 497 fármacos frecuentemente prescritos basándose en los siguientes criterios: (1) fármacos, de los que se conoce al menos un gen implicado en la farmacodinámica o farmacocinética, entre fármacos incluidos en los códigos ATC de las 15 clases de fármacos más frecuentemente prescritas durante 2005 a

2008 en Estados Unidos (Health, United States, 2011, Centers for Disease Control and Prevention), (2) fármacos con información sobre los efectos establecidos de marcadores farmacogenómicos de variación de secuencia genómica en etiquetas de fármacos de la FDA estadounidense, y (3) fármacos divulgados en la base de datos de DrugBank que se han retirado del mercado debido a los efectos secundarios del fármaco.

5 Como datos para evaluar la validez, entre el conocimiento establecido acerca de 987 pares de interacción de variación de secuencia génica-fármaco proporcionados por PharmGKB, se extrajeron 650 pares (65,9 %) que tenían al menos una vinculación con los 497 fármacos. Considerando que la diana de la presente divulgación es una variación de secuencia en una región exónica, la parte solapante entre los datos de la diana de verificación y los datos de la norma de evaluación se eliminaron para una evaluación equitativa. Para ser más específicos, se realizó una evaluación más equitativa retirando los pares con las 36 variaciones de secuencia colocadas en la región exónica entre los 650 pares y seleccionando solamente una variación de secuencia en una región no codificante. Como resultado, se seleccionaron 614 pares como norma de referencia final para la evaluación.

15 Después, se analizaron las secuencias genómicas completas de 1092 personas proporcionadas por el 1000 Genomes Project, y se aplicó de acuerdo con la presente divulgación a cada una de las 1092 personas para calcular de ese modo el riesgo farmacogenómico individual y el riesgo farmacogenómico de cada variación de secuencia génica registrada en PharmGKB.

20 Para la evaluación de la validez, se usaron la sensibilidad, especificidad y un área bajo la curva de rendimiento diagnóstico (ROC). Los 497 fármacos se clasificaron basándose en las puntuaciones de fármaco individuales y se establecieron valores de umbral para cada clasificación en 496 posiciones de segmentos entre las categorías. Después, (1) cuando una clasificación de una puntuación de fármaco de un fármaco correspondiente era mayor que un umbral y estaba presente una variación de PharmGKB en una variación genómica individual, se determinó como positivo verdadero, (2) cuando una clasificación de una puntuación de fármaco de un fármaco correspondiente era inferior que un umbral y no estaba presente una variación de PharmGKB en una variación genómica individual, se determinó como negativo verdadero, (3) cuando una clasificación de una puntuación de fármaco de un fármaco correspondiente era mayor que un umbral, pero no estaba presente una variación de PharmGKB en una variación genómica individual, se determinó como positivo falso, y (4) cuando una clasificación de una puntuación de fármaco de un fármaco correspondiente era inferior que un umbral, pero estaba presente una variación de PharmGKB en una variación genómica individual, se determinó como negativo falso. Los números de casos positivo verdadero, negativo verdadero, positivo falso y negativo falso de cada individuo con respecto a cada umbral de clasificación L se calcularon y la sensibilidad y especificidad se calcularon de acuerdo con las siguientes ecuaciones.

35
$$\text{Sensibilidad} = |D_L \cap GS| / |GS|$$

$$\text{Especificidad} = 1 - \{|D_L - GS| / |D - GS|\}$$

40 En las ecuaciones anteriores, D es un conjunto de los 497 fármacos, GS es un conjunto de fármacos PharmGKB personalizados usados como norma de referencia individual ya que una variación de secuencia génica individual en cada individuo es idéntica a un alelo de riesgo de PharmGKB, D_L es un conjunto de fármacos con altos umbrales de clasificación y el paréntesis en la barra vertical significa el número de elementos del conjunto correspondiente.

45 Como resultado del cálculo, 18 personas no tuvieron variación idéntica a la variación de PharmGKB y, por tanto, no pudo definirse un conjunto de fármacos PharmGKB personalizados usados como norma de referencia individual. Por lo tanto, las 18 personas se excluyeron de este ensayo de validez. La sensibilidad y especificidad se calcularon con respecto a todos los umbrales para trazar la curva de ROC y calcular la ABC. Para ser más específicos, las puntuaciones de variación de secuencia génica de las 1092 personas en el grupo de población total se calcularon usando el algoritmo SIFT y, después, se calcularon las puntuaciones de daño proteínico y las puntuaciones de fármaco usando la ecuación 2 y la ecuación 4, respectivamente. Además, para determinar la utilidad de la aplicación de ponderaciones de acuerdo con una distribución de raza, se calcularon la sensibilidad y especificidad específicas de raza y un valor de ABC basado en la sensibilidad y especificidad de la misma manera para cada una de las cuatro razas (africana (AFR, n = 246), americana (AMR, n = 181), asiática (ASN, n = 286), europea (EUR, n = 379)) indicadas claramente en el 1000 Genomes Project, de modo que se obtuvieron sensibilidad y especificidad específicas de raza y una ABC. Los resultados son como enumeran en la tabla 6, tabla 7 y FIG. 3.

[Tabla 6] Distribución de grupo de proteínas y puntuación media de daño proteínico

Grupo de proteínas	Número de proteínas	Número de fármacos relevantes	Número de pares de proteína-fármaco	Puntuación media de daño proteínico
Proteína diana	440	486	2357	0,798
Proteína de transporte	10	50	65	0,728

(continuación)

Grupo de proteínas	Número de proteínas	Número de fármacos relevantes	Número de pares de proteína-fármaco	Puntuación media de daño proteínico
Proteína enzimática	74	330	1347	0,733
Proteína transportadora	54	176	457	0,733
Total	545	497	4201	0,783

[Tabla 7] Validez (ABC) de puntuación de fármaco calculada para cada grupo de proteínas y cada raza usando los datos de 1000 Genomes Project

	Total	AFR	AMR	ASN	EUR
<u>Validez del cálculo de puntuación de fármaco (ABC)</u>					
Proteína diana	0,617	0,634	0,608	0,614	0,614
Proteína de transporte	0,554	0,511	0,599	0,485	0,594
Proteína enzimática	0,587	0,642	0,580	0,558	0,579
Proteína transportadora	0,497	0,492	0,488	0,489	0,512
<u>Validez de cálculo de puntuación de fármaco con y sin ponderación aplicada a cada grupo de proteínas (ABC)</u>					
Media geométrica simple	0,666	0,744	0,650	0,634	0,653
Media geométrica ponderada	0,667	0,742	0,652	0,633	0,654

5

La tabla 6 enumera la distribución de proteínas que se refieren a los 497 fármacos usados en este ejemplo para cada grupo de proteínas, e indica el número de pares de proteína-fármaco junto con una puntuación media de daño proteínico para cada grupo.

10 La tabla 7 enumera la validez del cálculo de puntuación de fármaco individual (ABC) calculada respectivamente en el caso donde no se aplican ponderaciones a cada grupo de proteínas (media geométrica simple) y el caso donde se aplican ponderaciones a cada grupo de proteínas (media geométrica ponderada) cuando se calcula la puntuación de fármaco usando la ecuación 4 con respecto a cada grupo de proteínas y cada raza.

15 Más específicamente, por ejemplo, en el grupo de población total, los valores de ABC calculados para grupos de proteínas tales como proteínas diana, proteínas de transporte, proteínas enzimáticas del metabolismo y enzimas transportadoras fueron 0,617, 0,554, 0,587 y 0,497, respectivamente. Estos valores se usaron como ponderaciones para los respectivos grupos de proteínas (cada valor se sustituyó por la ponderación w_i de la ecuación 4) para obtener de ese modo la validez del cálculo de puntuación de fármaco individual usando media geométrica ponderada (ABC = 0,667) (véase la FIG. 3B). Como resultado, se confirmó que la validez del cálculo de puntuación de fármaco individual usando la media geométrica ponderada en el caso de asignar las ponderaciones para los grupos de proteínas respectivos se aumentaba en 0,001 punto en comparación con la validez del cálculo de puntuación de fármaco individual usando media geométrica simple (ABC = 0,666) calculada aplicando una fórmula de cálculo de media geométrica simple sin asignar una ponderación (ponderación $w_i = 1$) (véase la FIG. 3A).

20 Además, como puede observarse a partir de la FIG. 3A, como otro ejemplo de aplicar ponderaciones, se asignaron ponderaciones de acuerdo con el número de personas por raza y se analizó la validez del cálculo de puntuación de fármaco individual (ABC). Como resultado, en el caso de considerar la especificidad de raza (línea gruesa), el valor de ABC del grupo de población total (Total) fue 0,666 (africana: 0,744, americana: 0,650, asiática: 0,631 y europea: 0,653), y en el caso de no considerar la especificidad de raza (línea de puntos), el valor de ABC del grupo de población total fue 0,633 (africana: 0,623, americana: 0,629, asiática: 0,64 y europea: 0,636). Por consiguiente, se confirmó que la validez del cálculo de puntuación de fármaco individual en el caso de considerar la especificidad de raza se mejoraba en comparación con el caso de no considerar la especificidad de raza.

35 Además, como se ilustra en la FIG. 3B, en el caso de asignar ponderaciones para los respectivos grupos de proteínas sin considerar la especificidad de raza (línea de puntos), la validez del cálculo de puntuación de fármaco individual (ABC) de la presente divulgación era 0,634, y en el caso de asignar ponderaciones para los respectivos grupos de proteínas mientras se considera la especificidad de raza (línea gruesa), la validez del cálculo de puntuación de fármaco individual (ABC) de la presente divulgación era 0,667. Por consiguiente, puede observarse que son útiles diferentes ponderaciones.

40

Aunque las realizaciones ilustrativas de la presente divulgación se han descrito en detalle, el alcance de la presente divulgación no se limita a las mismas. También se incluyen en el alcance de la presente divulgación diversas modificaciones y mejoras realizadas por los expertos en la materia usando el concepto básico de la presente

divulgación definida en las reivindicaciones adjuntas.

Salvo que se definan de otro modo, todos los términos técnicos usados en el presente documento tienen el mismo significado que el comprendido habitualmente por un experto en la materia a la que pertenece esta invención.

REIVINDICACIONES

1. Un método implementado por ordenador para proporcionar información para seleccionar un agente antineoplásico usando variación de secuencia genómica individual, que comprende:

- 5 una etapa de obtener una o más variaciones de secuencia de un paciente en genes que se ha determinado que están implicados en la farmacodinámica o farmacocinética de un agente antineoplásico a partir de información de secuencia genómica individual;
- 10 una etapa de calcular una puntuación daño proteínico usando la [ecuación 2], en la que

la [ecuación 2] es:

$$S_g(v_1, \dots, v_n) = \left(\prod_{i=1}^n v_i^{w_i} \right)^{1/\sum_{i=1}^n w_i}$$

- 15 en la que S_g es la puntuación daño proteínico de una proteína codificada por un gen g, n es el número de la una o más variaciones de secuencia del gen g, v_i es una puntuación de variación de secuencia génica de una variación de secuencia génica i.^a y w_i es una ponderación asignada a la puntuación de variación de secuencia génica v_i de la i.^a variación de secuencia génica; y
- 20 una etapa de calcular una puntuación de fármaco individual usando la [ecuación 4], en la que

la [ecuación 4] es:

$$S_d(g_1, \dots, g_n) = \left(\prod_{i=1}^n g_i^{w_i} \right)^{1/\sum_{i=1}^n w_i}$$

- 25 en la que S_d es la puntuación de fármaco individual de un agente antineoplásico d, n es un número de proteínas codificadas por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d, g_i es una puntuación daño proteínico de una proteína codificada por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d, y w_i es una ponderación asignada a la puntuación
- 30 daño proteínico g_i de la proteína codificada por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d; y
- proporcionar información para seleccionar un agente antineoplásico basándose en el cálculo de la puntuación de fármaco individual.

35 2. El método para proporcionar información para seleccionar un agente antineoplásico usando variación de secuencia genómica individual de acuerdo con la reivindicación 1, en el que el agente antineoplásico se selecciona de un fármaco basado en tiopurina, un análogo desoxinucleosídico y un agente alquilante.

40 3. El método para proporcionar información para seleccionar un agente antineoplásico usando variación de secuencia genómica individual de acuerdo con la reivindicación 1 o 2, en el que el agente antineoplásico es un fármaco basado en tiopurina o un análogo desoxinucleosídico.

4. El método para proporcionar información para seleccionar un agente antineoplásico usando variación de secuencia genómica individual de acuerdo con la reivindicación 1, en el que el agente antineoplásico no es busulfano.

45 5. El método para proporcionar información para seleccionar un agente antineoplásico usando variación de secuencia genómica individual de acuerdo con la reivindicación 1, en el que el gen implicado en la farmacodinámica o farmacocinética es uno o más seleccionados de un grupo que consiste en ABCC1, ABCG2, APEX1, CBR3, ATP7A, CBR3, CTH, CUL9, CYP1A1, CYP2C9, CYP2D6, CYP4B1, DPEP1, DPEP2, DPYD, DRD3, EPHX1, FMO2, FMO3, GGT1, GGT5, GGT6, GNB3, GSTA2, GSTM1, MGMT, MLH1, MSH2, NAT1, OPRM1, PDE5A, PTGS1, SERPINA6,

50 SLC15A2, SLC22A1, SLC22A2, SLC22A8, SLC22A16, SLC22A2, SLC28A2, SLC28A3, SPG7, TP53, ABCB1, ABCC2, AOX1, CHRM2, CYP1B1, CYP3A4, CYP3A5, CYP4F2, DBH, ERCC2, GGH, GGT7, GMPS, MTHFR, NUDT15, PDE3A, PDE4C, SLC22A5, SLC01B1, TLR3, UGT1A1 y XDH.

55 6. El método para proporcionar información para seleccionar un agente antineoplásico usando variación de secuencia genómica individual de acuerdo con la reivindicación 1, en el que la información de variación de secuencia génica es:

- a) una sustitución, adición o eliminación de una base que constituye un exón de un gen; o
- b) adquirida por un análisis de comparación con una secuencia genómica de un grupo de referencia.

7. El método para proporcionar información para seleccionar un agente antineoplásico usando variación de secuencia genómica individual de acuerdo con la reivindicación 1, en el que la puntuación de daño proteínico o la puntuación de fármaco se calcula:

- 5 (i) a partir de una o más puntuaciones de variación de secuencia génica calculadas usando uno o más algoritmos seleccionados de un grupo que consiste en SIFT (clasificación de intolerante frente a tolerante), PolyPhen (fenotipado de polimorfismos), PolyPhen-2, MAPP (análisis de múltiples variables de polimorfismo proteínico), Logre (Log R de valor E Pfam), MutationAssessor, MutationTaster, MutationTaster2, PROVEAN (analizador de efecto de variación proteínico), PMut, Condel, GERP (perfilado de tasa evolutiva genómica), GERP++, CEO (optimización de entropía combinatoria), SNPeffect, fathmm y CADD (reducción dependiente de anotación combinada); o
- 10 (ii) a partir de una puntuación de variación de secuencia génica, opcionalmente en el que la puntuación de variación de secuencia génica se calcula aplicando uno o más algoritmos seleccionados de un grupo que consiste en SIFT (clasificación de intolerante frente a tolerante), PolyPhen (fenotipado de polimorfismos), PolyPhen-2, MAPP (análisis de múltiples variables de polimorfismo proteínico), Logre (Log R de valor E Pfam), MutationAssessor, MutationTaster, MutationTaster2, PROVEAN (analizador de efecto de variación proteínico), PMut, Condel, GERP (perfilado de tasa evolutiva genómica), GERP++, CEO (optimización de entropía combinatoria), SNPeffect, fathmm y CADD (reducción dependiente de anotación combinada) para la variación de secuencia génica.

8. El método para proporcionar información para seleccionar un agente antineoplásico usando variación de secuencia genómica individual de acuerdo con la reivindicación 1, en el que el método comprende además una etapa de determinar si el agente antineoplásico se usará por el paciente o la manera en que se usará usando la puntuación de fármaco individual.

9. El método para proporcionar información para seleccionar un agente antineoplásico usando variación de secuencia genómica individual de acuerdo con la reivindicación 1, en el que el método comprende además una etapa de recibir la información de variación de secuencia génica a través de un sistema informático, en el que el sistema informático comprende o es accesible para una base de datos que comprende información acerca del gen implicado en la farmacodinámica o farmacocinética del agente antineoplásico.

10. El método para proporcionar información para seleccionar un agente antineoplásico usando variación de secuencia genómica individual de acuerdo con una cualquiera de las reivindicaciones 1 a 9, en el que el método se realiza para evitar los efectos secundarios del agente antineoplásico.

11. Un sistema para seleccionar un agente antineoplásico usando variación de secuencia genómica individual, que comprende:

- a) una base de datos de la que puede buscarse o extraerse información relevante a un gen o proteína relacionada con un agente antineoplásico aplicable a un paciente;
- b) una unidad de comunicación que es accesible para la base de datos;
- c) un primer módulo de cálculo que se configura para calcular una o más variaciones de secuencia en genes que se determina o que están implicados en la farmacodinámica o farmacocinética del agente antineoplásico basándose en la información;
- d) un segundo módulo de cálculo que se configura para calcular una puntuación daño proteínico usando la [ecuación 2], en la que

la [ecuación 2] es:

$$S_g(v_1, \dots, v_n) = \left(\prod_{i=1}^n v_i^{w_i} \right)^{1/\sum_{i=1}^n w_i}$$

- en la que S_g es la puntuación daño proteínico de una proteína codificada por un gen g, n es el número de la una o más variaciones de secuencia del gen g, v_i es una puntuación de variación de secuencia génica de una variación de secuencia génica i.^a y w_i es una ponderación asignada a la puntuación de variación de secuencia génica v_i de la i.^a variación de secuencia génica;
- e) un tercer módulo de cálculo que se configura para calcular una puntuación de fármaco individual usando la [ecuación 4], en la que

la [ecuación 4] es:

$$S_d(g_1, \dots, g_n) = \left(\prod_{i=1}^n g_i^{w_i} \right)^{1/\sum_{i=1}^n w_i}$$

en la que S_d es la puntuación de fármaco individual de un fármaco d , n es un número de proteínas codificadas por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d , g_i es una puntuación daño proteínico de una proteína codificada por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d , y w_i es una ponderación asignada a la puntuación daño proteínico g_i de la proteína codificada por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d ; y

f) una unidad de presentación que se configura para presentar los valores calculados por los módulos de cálculo;

y opcionalmente comprende además:

g) un cuarto módulo de cálculo que se configura para determinar si el agente antineoplásico se usará para el paciente usando la puntuación de fármaco individual calculada por el tercer módulo de cálculo; o

h) una interfaz de usuario que se configura para calcular y proporcionar una puntuación de fármaco individual para el agente antineoplásico cuando el agente antineoplásico se introduce por un usuario; o

i) una unidad de presentación que se configura para presentar además el valor calculado por cada uno de los módulos de cálculo, el procedimiento de cálculo o la información como fundamento del cálculo; y

en el que el sistema proporciona información para seleccionar un agente antineoplásico basándose en el cálculo de la puntuación de fármaco individual.

12. El sistema para seleccionar un agente antineoplásico usando variación de secuencia individual de acuerdo con la reivindicación 11, en el que el agente antineoplásico:

(a) se selecciona de un fármaco basado en tiopurina, un análogo desoxinucleosídico y un agente alquilante; o

(b) es un fármaco basado en tiopurina o un análogo desoxinucleosídico; o

(c) no es busulfano.

13. El sistema para seleccionar un agente antineoplásico usando variación de secuencia genómica individual de acuerdo con la reivindicación 11, en el que la información de variación de secuencia génica, la puntuación de daño proteínico, la puntuación de fármaco y la información como fundamento para el cálculo se almacena en cada uno de los módulos de cálculo y la información de cada módulo de cálculo se actualiza cuando se actualiza la base de datos.

14. Un medio legible por ordenador que incluye un módulo de ejecución para ejecutar un procesador que realiza una operación, que comprende:

una etapa de adquisición de una o más variaciones de secuencia de un paciente en genes que se determinan o que están implicados en la farmacodinámica o farmacocinética de un agente antineoplásico a partir de información de secuencia genómica individual;

una etapa de calcular una puntuación daño proteínico usando la [ecuación 2], en la que

la [ecuación 2] es:

$$S_g(v_1, \dots, v_n) = \left(\prod_{i=1}^n v_i^{w_i} \right)^{1/\sum_{i=1}^n w_i}$$

en la que S_g es la puntuación daño proteínico de una proteína codificada por un gen g , n es el número de la una o más variaciones de secuencia del gen g , v_i es una puntuación de variación de secuencia génica de una variación de secuencia génica i .ª y w_i es una ponderación asignada a la puntuación de variación de secuencia génica v_i de la i .ª variación de secuencia génica; y

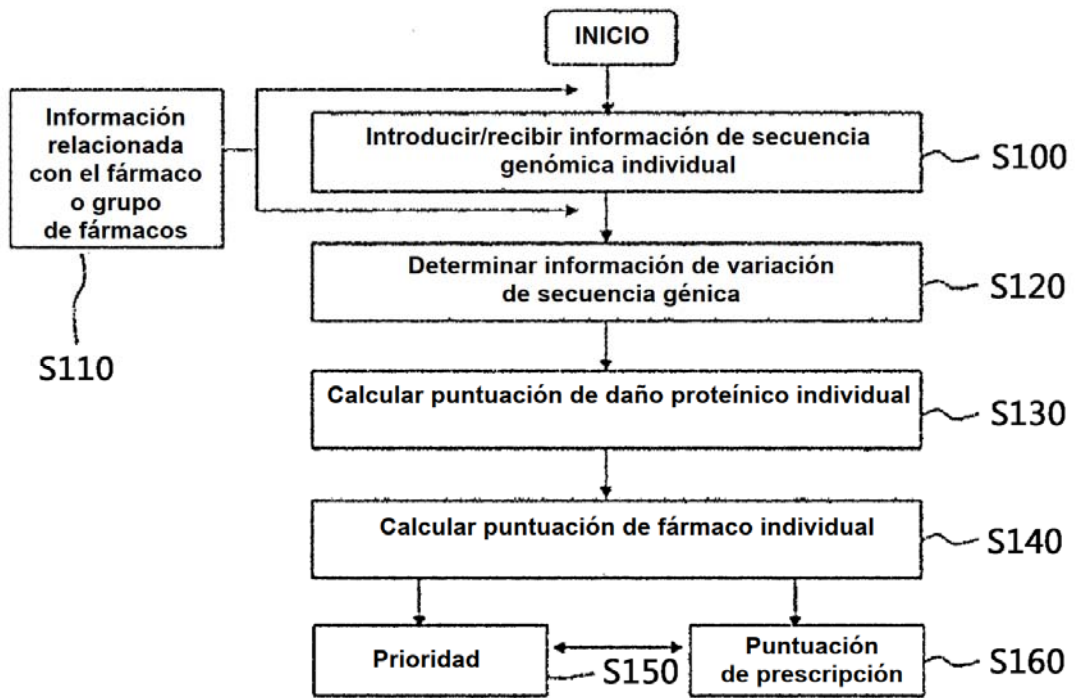
una etapa de calcular una puntuación de fármaco individual usando la [ecuación 4], en la que

la [ecuación 4] es:

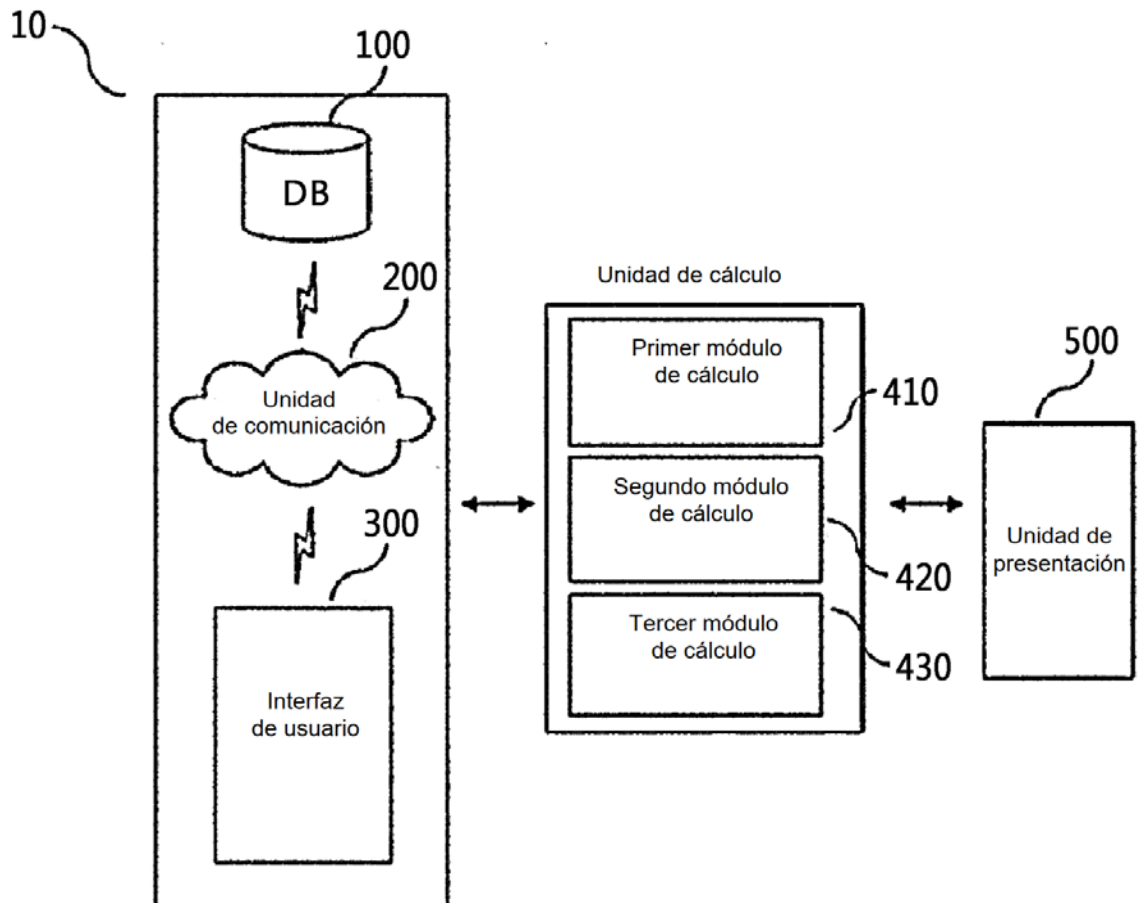
$$S_d(g_1, \dots, g_n) = \left(\prod_{i=1}^n g_i^{w_i} \right)^{1/\sum_{i=1}^n w_i}$$

- 5 en la que S_d es la puntuación de fármaco individual de un fármaco d , n es un número de proteínas codificadas por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d , g_i es una puntuación daño proteínico de una proteína codificada por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d , y w_i es una ponderación asignada a la puntuación daño proteínico g_i de la proteína codificada por uno o más genes implicados en la farmacodinámica o farmacocinética del agente antineoplásico d .
- 10 15. El medio legible por ordenador de acuerdo con la reivindicación 14, en el que el agente antineoplásico:
- (a) se selecciona de un fármaco basado en tiopurina, un análogo desoxinucleosídico y un agente alquilante; o
 - (b) es un fármaco basado en tiopurina o un análogo desoxinucleosídico; o
 - (c) no es busulfano.
- 15 16. El medio legible por ordenador de acuerdo con la reivindicación 14, en el que el procesador comprende además una etapa de determinar si el agente antineoplásico se usará por el paciente usando la puntuación de fármaco individual.

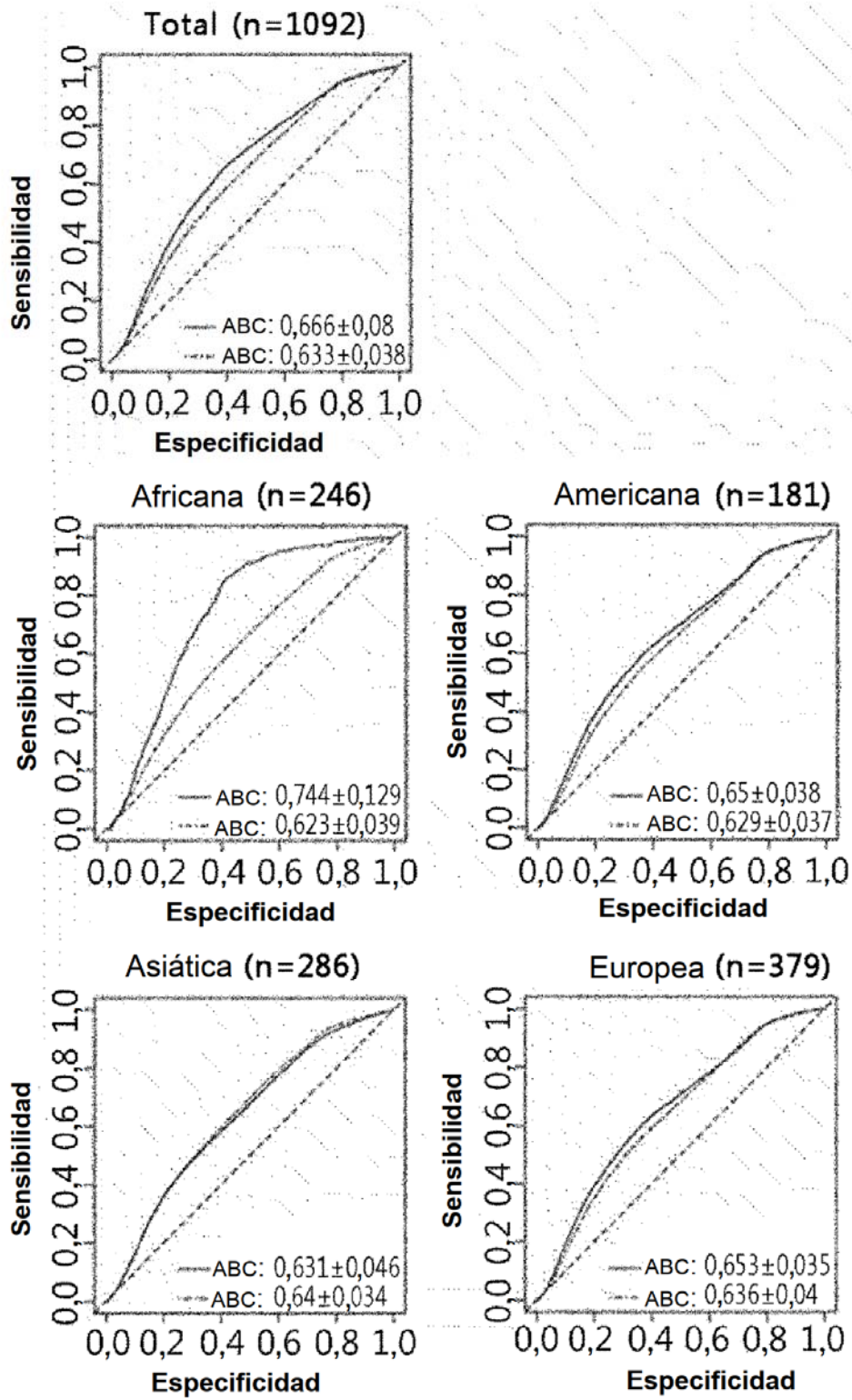
[FIG.1]



[FIG. 2]



[FIGS. 3a]



[FIGS. 3b]

