Industria y
SUPERINTENDENCIA

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
No. 13-233174- -00000-0000
Fecha: 2013-10-01 16:38:27 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 429

DELEGATURA DE PROPIEDAD INDUSTRIAL

SOLICITUD

PATENTE DE INVENCIÓN

21. **EXPEDIENTE No.** _____

54. **TÍTULO** ENSAYOS INTERROGATORIOS BASADOS EN CELULAS Y USOS DE LOS MISMOS

51. **CLASIFICACIÓN INTERNACIONAL** _____

71. **SOLICITANTE** BERG , LLC

DOMICILIO NASHVILLE, TENNESSEE, ESTADOS UNIDOS DE AMERICA

74. **APODERADO** LUZ CLEMENCIA DE PAEZ
C.C. 35.456.344 de Usaquén

22. **BOGOTÁ, D. C.,** _____

(FORMA P-10)

117144-Imprenta Nacional de Colombia-2012

Industria y Comercio
SUPERINTENDENCIA

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

No. 13-233174- -00000-0000

Fecha: 2013-10-01 16:38:27 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 429

/o de radicación

DIRECCIÓN DE NUEVAS CREACIONES
SOLICITUD FASE NACIONAL -PCT

1	TIPO DE SOLICITUD	☒ Patente de invención	☐ Patente de Modelo de Utilidad
		☒ Capítulo I	☐ Capitulo II

2 DATOS SOLICITUD INTERNACIONAL PCT (86)

Solicitud Internacional No.	PCT/US2012/027615	Fecha	2012/03/02
Publicación Internacional No.	WO 2012/119129 A1	Fecha	2012/09/07

3 TÍTULO DE LA INVENCIÓN

ENSAYOS INTERROGATORIOS BASADOS EN CÉLULAS Y USOS DE LOS MISMOS

4 CLASIFICACIÓN INTERNACIONAL (CIP)

A61P 035/000

5 SOLICITANTE (S) ☐ Esta persona también es inventor. Para datos adicionales utilizar hoja de información complementaria

APELLIDOS O RAZÓN SOCIAL	NOMBRE	IDENTIFICACIÓN	TIPO
BERG LLC			

6 DATOS DEL SOLICITANTE

DIRECCIÓN	1845 Elm Hill Pike, Nashville, Tennessee 37210, Estados Unidos de América	No. TELÉFONO	3 47 36 11
CIUDAD	Nashville	CORREO ELECTRÓNICO	clientes@cavelier.com
DEPARTAMENTO/ESTADO	Tennessee	NACIONALIDAD O LUGAR DE CONSTITUCIÓN	Estados Unidos de América
PAÍS DE RESIDENCIA	Estados Unidos de América		

7 INVENTOR (ES) Para datos adicionales utilizar hoja de información complementaria

APELLIDOS	NOMBRES	NACIONALIDAD
1. NARAIN	Niven Rajin	Guyanés
2. SARANGARAJAN	Rangaprasad	Estadounidense
3. VISHNUDAS	Vivek K.	Hindú

DIRECCIÓN DE CORREO ELECTRÓNICO:

8 DATOS INVENTOR (ES)

PAÍS RESIDENCIA	DEPARTAMENTO/ESTADO	CIUDAD	DIRECCIÓN
1. Estados Unidos de América	Massachusetts	Cambridge	73 Fresh Pond Parkway, Cambridge, Massachusetts 02138, Estados Unidos de América
2. Estados Unidos de América	Massachusetts	Boylston	454 Central Srteet, Boylston, Massachusetts 01505, Estados Unidos de América
3. Estados Unidos de América	Massachusetts	Waltham	8 Garden Circle, Apt.#4, Waltham, Massachusetts 02452, Estados Unidos de América

OTRO(S) SOLICITANTE(S) Y/O (OTRO(S) INVENTOR(ES)

☐ Los demás solicitantes y/o (demás) inventores se indican en una hoja a continuación.

V

9	☐ REPRESENTANTE LEGAL ☒ APODERADO		
APELLIDOS	NOMBRES	IDENTIFICACIÓN	
De Páez	Luz Clemencia	C.C . 35.456.344 T.P . 23.555	

DIRECCIÓN	Carrera 4ª No. 72-35	No. TELÉFONO	3 47 36 11
CIUDAD	BOGOTA	CORREO ELECTRÓNICO	cavelier@cavelier.com
PAÍS	COLOMBIA	No. RADICACIÓN DE PROTOCOLO DE PODER GENERAL	

10 DECLARACIONES DE PRIORIDAD ☒ SI ☐ NO

(33) PAÍS DE ORIGEN	CÓDIGO PAÍS	(31) NÚMERO	(32) FECHA (AAAA/MM/DD)
1. U.S.A.	US	61/448,587	2011/03/02
2. U.S.A	US	61/593,848	2012/02/01

11 DECLARACIÓN SOBRE USO DE RECURSOS GENÉTICOS O BIOLÓGICOS

Declaro que el objeto de la presente solicitud de patente fue obtenido a partir de recursos genéticos o biológicos de los que cualquiera de los países miembros de la Comunidad Andina es país de origen.

☐ SI ☐ NO

Nota: En caso afirmativo deberá anexar copia del contrato de acceso de recursos genéticos o productos derivados, o certificado o número de registro, expedido por la Autoridad competente.

12 DECLARACIÓN SOBRE USO DE CONOCIMIENTOS TRADICIONALES

Declaro que el objeto de la presente solicitud de patente fue obtenido a partir de conocimientos tradicionales de comunidades indígenas, afroamericanas o locales de países miembros de la Comunidad Andina.

☐ SI ☐ NO

Nota: En caso afirmativo deberá anexar la licencia o autorización de uso de conocimiento tradicional, o certificado, o número de registro expedido por la Autoridad competente.

13 REDUCCIÓN DE TASAS

Declaro que carezco de medios económicos para presentar la solicitud de patente.

☐ SI ☐ NO

Nota: En caso de ser persona natural y carecer de medios económicos, y por lo tanto, aplique la reducción de tasas a que se refiere la resolución vigente en tarifas, debe firmar la presente solicitud bajo la gravedad de juramento.

Micro, pequeñas y medianas empresas ☐

Universidades públicas o privadas ☐

Entidades sin ánimo de lucro ☐

Debe aportar los documentos que se indican en el numeral 17 de anexos

14 AUTORIZACIÓN DE NOTIFICACIÓN EN LÍNEA ☐ SI ☐ NO

Manifiesto que he leído y entendido perfectamente los términos y condiciones de uso de medios electrónicos para las notificaciones en línea a través de Internet de los actos administrativos proferidos por la Superintendencia de Industria y Comercio que deben ser notificados personalmente y, en consecuencia, autorizo el servicio de notificación a través de internet.

15	COMPROBANTE DE PAGO O PAGO ELECTRÓNICO	**No.** 13-0098033	**Fecha:** 2013/09/30

16 FIRMA DEL SOLICITANTE, DEL APODERADO O DEL REPRESENTANTE LEGAL

Junto a cada firma, indicar el nombre del firmante y su calidad (si tal calidad no es obvia al leer el petitorio)

Nombre del Firmante	Firma
LUZ CLEMENCIA DE PAEZ	Luz Clemencia de Páez T.P. 23555 CSJ
C.C 35.456.344 de Usaquén	Tarjeta Profesional 23.555

de6 JGG\SDC\ID-AF252426\WPC20488

17	ANEXOS

Documentación Técnica
Solicitud Internacional en castellano

1. ☒ Descripción N° de folios: 339
2. ☒ Reivindicaciones N° Reivindicaciones: 44
3. ☒ Dibujos y/o figuras N° folios: 40
4. ☒ Resumen.
5. ☐ Certificado de depósito de material biológico si fuera el caso.
6. ☐ Listado de secuencias de nucleótidos y/o aminoácidos en forma digital si fuera el caso.
7. ☒ Arte final 12 x 12.
8. ☐ Anexo formato digital.

☒ **Forma PCT/IB/306.** Notificación correspondiente a la inscripción del cambio de nombre del solicitante **BERG BIOSYSTEMS, LLC** por **BERG PHARMA LLC.**

☒ **Forma PCT/IB/306.** Notificación correspondiente a la inscripción del cambio de nombre del solicitante **BERG PHARMA LLC** por **BERG LLC.**

☒ **Forma PCT/IB/306.** Notificación correspondiente a la inscripción del cambio de dirección del inventor **Niven Rajin NARAIN.**

Documentación Jurídica

9. ☐ Poderes, si fuera el caso.
10. ☒ Tres (3) documentos que legalmente prueban la cesión de los tres (3) inventores al solicitante o a su causante.
11. ☐ Copia del contrato de acceso de recursos genéticos o productos derivados, certificado o número de registro, si fuera el caso.
12. ☐ Copia de la licencia o autorización de Conocimientos Tradicionales, certificado o número de registro, si fuera el caso.
13. **Reducción de tasas**

Micro, pequeñas o medianas empresas

☐ Copia simple de la declaración de renta del año inmediatamente anterior, o en su defecto prueba documental idónea.

☐ Documento de constancia de cumplimiento con lo establecido en la ley 905 de 2004.

Universidades públicas o privadas

☐ Copia acto de reconocimiento institucional emitido por el Ministerio de Educación

Entidades sin ánimo de lucro

☐ Copia de registro vigente en Cámara de comercio.
☐ Hoja de información complementaria.
☐ Otros, especificar

14. ☒ Comprobante de pago de la tasa de presentación de la solicitud por $500.000.
15. ☐ Comprobante de pago por reivindicación de prioridad.
16. ☐ Comprobante de pago de la tasa por concepto de excedente de palabras en la publicación.
17. ☐ Comprobante de pago por reivindicación adicional a 10.

4

FIGURA CARACTERISTICA

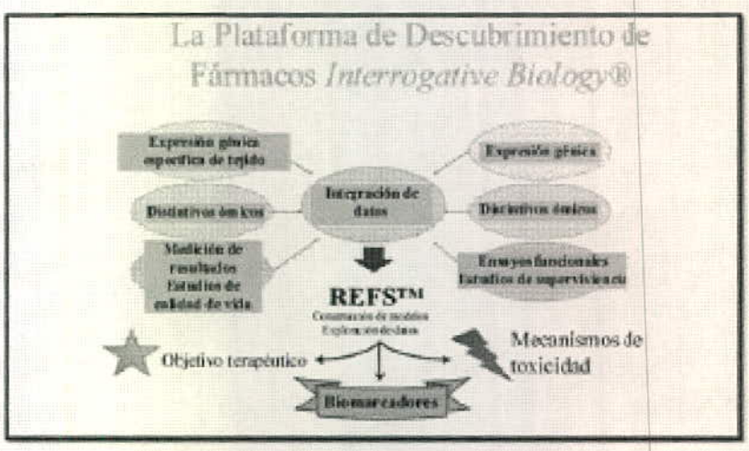


Figura 7

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 September 2012 (07.09.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/119129 A1

(51) International Patent Classification:
A61P 35/00 (2006.01)

(21) International Application Number:
PCT/US2012/027615

(22) International Filing Date:
2 March 2012 (02.03.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/448,587 2 March 2011 (02.03.2011) US
61/593,848 1 February 2012 (01.02.2012) US

(71) Applicant *(for all designated States except US)*: **BERG BIOSYSTEMS, LLC** [US/US]; 1845 Elm Hill Pike, Nashville, TN 37210 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **NARAIN, Niven, Rajin** [GY/US]; 350 Third Street, Suite 2005, Cambridge, MA 02142 (US). **SARANGARAJAN, Rangaprasad** [US/US]; 454 Central Street, Boylston, MA 01505 (US). **VISHNUDAS, Vivek, K.** [IN/US]; 8 Garden Circle, Apt. #4, Waltham, MA 02452 (US).

(74) Agents: **HANLEY, Elizabeth, A.** et al.; McCarter & English, LLP, 265 Franklin Street, Boston, MA 02110 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

— *with international search report (Art. 21(3))*

(54) **Title**: INTERROGATORY CELL-BASED ASSAYS AND USES THEREOF

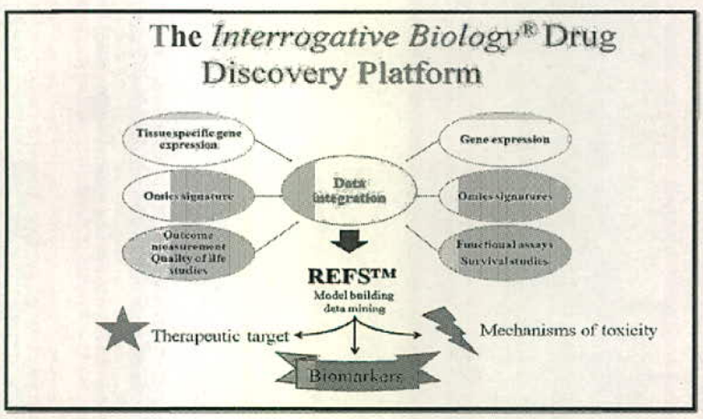


Figure 7

(57) **Abstract**: Described herein is a discovery Platform Technology for analyzing a biological system or process (e.g. a disease condition, such as cancer) via model building. In particular, described herein is a method for identifying a modulator of a biological system by establishing a model for the biological system, obtaining a first data set from the model for the biological system, obtaining a second data set from the model for the biological system, and generating a consensus causal relationship network among the expression levels of the plurality of genes and the functional activity or cellular response based solely on the first data set and the second data set using a programmed computing device.

WO 2012/119129 A1

ENSAYOS INTERROGATORIOS BASADOS EN CÉLULAS Y USOS DE LOS MISMOS

Referencia cruzada a solicitudes relacionadas

La presente solicitud reivindica la prioridad de la Solicitud Provisional de los Estados Unidos No. 61/448.587, presentada el 2 de marzo de 2011 y la Solicitud Provisional de los Estados Unidos No. 61/593.848, presentada el 1° de febrero de 2012, que se incorporan a la presente a modo de referencia en su totalidad.

Antecedentes de la invención

El desarrollo de nuevos fármacos ha mejorado considerablemente desde el descubrimiento del ADN en 1964 por James Watson y Francis Crick, pioneros de los que actualmente denominamos biología molecular. Las herramientas y productos de la biología molecular posibilitan la medición rápida, detallada y precisa de la regulación génica a nivel de ADN y ARN. Las siguientes tres décadas luego del descubrimiento que cambió el paradigma verían la génesis de modelos animales inactivados, reacciones relacionadas a enzimas clave y nuevos entendimientos de mecanismos de enfermedades y patofisiología de las plataformas mencionados anteriormente. En la primavera del 2000, cuando Craig

Ventor y Francis Collins anunciaron la secuenciación inicial del genoma humano, el mundo científico ingresó en una nueva ola de la medicina.

El mapeo del genoma inmediatamente despertó esperanzas, por ejemplo, de poder controlar enfermedades incluso antes de su aparición, de usar la terapia génica para revertir los procesos degenerativos cerebrales que provocan la enfermedad de Alzheimer o Parkinson y de una construcción que podría introducirse en un sitio tumoral y provocar la erradicación de la enfermedad mientras se restauraba la arquitectura y fisiología del tejido normal. Otros tomaron giros controvertidos y propusieron la noción de crear progenies elegidas con respecto al color de ojos o cabellos, altura, etc. Diez años más tarde, sin embargo, todavía estamos esperando, sin ningún camino específico en vista, el éxito sostenido de la terapia génica o incluso el control elemental de los procesos genéticos.

Por consiguiente, una realidad evidente es que la genética, al menos independiente de construcciones de soporte, no impulsa los objetivos clínicos de la fisiología. De hecho, muchos procesos tales como modificaciones postranscripcionales, mutaciones, polimorfismos de nucleótido simple (SNP, por sus siglas en inglés), y modificaciones translacionales podrían alterar el

origen de un gen y/o su proteína complementaria codificada y contribuir así al proceso de la enfermedad.

Compendio de la invención

La era de la información y la creación de internet posibilitó una sobrecarga de información, mientras que también facilitó la colaboración y la crítica internacionales. Irónicamente, las realidades mencionadas anteriormente también pueden ser la causa de que la comunidad científica esté pasando por alto algunos puntos simples, incluida que la comunicación de cascadas de señal y la interferencia dentro y entre células y/o tejidos permite la homeostasis y la mensajería para que se produzcan mecanismos correctivos cuando algo sale mal.

Un punto se refiere a las enfermedades cardiovasculares (CVD, por sus siglas en inglés), que sigue siendo la principal causa de muertes en Estados Unidos y gran parte del mundo desarrollado, siendo la causa de 1 de cada 2,8 muertes solo en Estados Unidos. Adicionalmente, las CVD sirven como una patología subyacente que contribuye a complicaciones asociadas tales como Enfermedad Renal Crónica (~19 millones de casos en EE. UU.), síndrome de cansancio crónico y un factor clave en el síndrome metabólico. Los avances significativos en la tecnología relacionada con diagnósticos, técnicas de cirugía mínimamente invasivas, stents de elución de

fármacos y control clínico eficaz han contribuido a un período de crecimiento sin comparación en el campo de la cardiología intervencionista y han posibilitado una gestión más eficaz de las CVD. Sin embargo, la etiología de la enfermedad relacionada con las CVD y comorbilidades asociadas tales como diabetes y enfermedad vascular periférica todavía deben elucidarse completamente.

Todavía faltan nuevos enfoques para explorar los mecanismos y vías implicados en un proceso biológico, tales como la etiología de enfermedades (por ejemplo, CVD) y para identificar vías de regulación claves y/o moléculas objetivo *(por ejemplo,* "objetivos tratables con fármacos") y/o marcadores para hacer un diagnóstico mejor de las enfermedades y/o tratamientos.

La invención que se describe en la presente se basa, al menos en parte, en un nueva utilización en colaboración de herramientas de biología, genómicas, proteómicas, metabolómicas, transcriptómicas y de bioinformática en línea, que, cuando se combinan, pueden usarse para estudiar cualquier sistema biológico de interés, tales como enfermedades seleccionadas incluidos el cáncer, diabetes, obesidad y enfermedades cardiovasculares, usando un abordaje de biología de sistemas. En una primera etapa, se desarrollan sistemas de modelado celular para explorar varios sistemas biológicos, tal como un proceso de enfermedad, que

comprende células relacionadas con la enfermedad sometidas a varios estímulos ambientales relevantes para enfermedad (por ejemplo, hiperglucemia, hipoxia, inmunoestrés y peroxidación lipídica). En algunas realizaciones, el sistema de modelado celular implica mecanismos de interferencia entre varios tipos de células que interactúan (tales como células de músculo liso de la aorta (HASMC, por sus siglas en inglés), células tubulares proximales renales (HK-2), células endoteliales de la aorta (HAEC, por sus siglas en inglés) y fibroblastos dérmicos (HDFa)). Las lecturas biológicas de alta productividad del sistema de modelado celular se obtienen usando una combinación de técnicas, incluidas, por ejemplo espectrometría de masas de punta (LC/MSMS), citometría de flujo, ensayos basados en células y ensayos funcionales. Las lecturas biológicas de alta productividad a continuación se someten a un análisis bioinformático para estudiar tendencias de datos congruentes mediante modelado *in vitro, in vivo,* y *in silico* . Las matrices resultantes posibilitan la extracción de datos conexos donde se desarrollaron análisis de regresión lineales y no lineales para alcanzar puntos de presión concluyentes (o "concentradores"). Estos, "concentradores", tales como se presentan en la presente, son candidatos para el descubrimiento de fármacos. Específicamente, estos concentradores representan objetivos de fármacos y/o marcadores de enfermedades potenciales.

Los distintivos moleculares de los diferenciales posibilitan comprender los mecanismos que dictan las alteraciones en el microambiente tisular que conducen al inicio y la evolución de la enfermedad. Tomados juntos, la combinación de las plataformas tecnológicas mencionadas anteriormente y con el modelado celular estratégico posibilitan una inteligencia sólida que puede emplearse para establecer una comprensión adicional de las enfermedades mientras se crean bibliotecas de biomarcadores y fármacos candidatos que pueden aumentar clínicamente el estándar de cuidado en la cardiología intervencionista.

Adicionalmente, este abordaje no es solo útil para el diagnóstico de enfermedades o la intervención, sino que también tiene una aplicabilidad general en prácticamente todas las afecciones patológicas o no patológicas en sistemas biológicos, tales como sistemas biológicos donde dos o más sistemas celulares interactúan. Por ejemplo, este abordaje es útil para obtener conocimiento sobre los mecanismos asociados o causales de la toxicidad farmacológica. Por lo tanto, la invención proporciona un marco para una evaluación biológica interrogativa que puede aplicarse en general en un amplio espectro de escenarios.

Una característica significativa de la plataforma de la invención es que el sistema basado en AI se basa en los conjuntos de datos obtenidos del sistema de modelo celular, sin recurrir ni tomar en

cuenta ningún conocimiento existente en la técnica, tal como las relaciones biológicas conocidas (es decir, ningún punto de datos es artificial), en relación con el proceso biológico. Por consiguiente, los modelos estadísticos resultantes generados a partir de la plataforma son imparciales. Otra característica significativa de la plataforma de la invención y sus componentes, por ejemplo, los sistemas de modelo celular y los conjuntos de datos obtenidos de los mismos, es que permite la construcción continua sobre los modelos celulares con el tiempo (por ejemplo, mediante la introducción de nuevas células y/o condiciones), de forma que una red de relaciones causales de consenso inicial, "de primera generación", generada a partir de un modelo celular para un sistema o proceso biológico puede evolucionar junto con la evolución del modelo celular en sí hasta una red de relaciones de múltiples generaciones (y las redes delta o delta-delta obtenidas de las mismas). De esta forma, los modelos celulares, los conjuntos de datos y las redes de relaciones causales generadas a partir de los modelos celulares mediante el uso de los métodos de tecnología de plataforma pueden evolucionar constantemente y ampliarse a partir de conocimiento previo obtenido a partir de la tecnología de plataforma.

Por consiguiente, en un aspecto, la invención se refiere a un método para identificar un modulador de un sistema biológico, comprendiendo dicho método: (1) establecer un modelo para el

sistema biológico, usando células asociadas con el sistema biológico, para representar un aspecto característico del sistema biológico; (2) obtener un primer conjunto de datos a partir del modelo para el sistema biológico, en donde el primer conjunto de datos representa niveles de expresión de una pluralidad de genes en las células asociadas con el sistema biológico; (3) obtener un segundo conjunto de datos a partir del modelo para el sistema biológico, en donde el segundo conjunto de datos representa una actividad funcional o una respuesta celular de las células asociadas con el sistema biológico; (4) generar una red de relaciones causales de consenso entre los niveles de expresión de la pluralidad de genes y la actividad funcional o respuesta celular basada solamente en el primer conjunto de datos y el segundo conjunto de datos usando un dispositivo informático programado, en donde la generación de la red de relaciones causales de consenso no se basa en ninguna relación biológica conocida distinta del primer conjunto de datos y el segundo conjunto de datos; (5) identificar, a partir de la red de relaciones causales de consenso, una relación causal única en el sistema biológico, en donde un gen asociado con la relación causal única se identifica como un modulador del sistema biológico.

En ciertas realizaciones, el modulador estimula o activa el sistema biológico.

En ciertas realizaciones, el modulador inhibe el sistema biológico.

En ciertas realizaciones, el modelo del sistema biológico comprende un cultivo in vitro de células asociadas con el sistema biológico, comprendiendo además opcionalmente un cultivo in vitro correspondiente de células testigo.

En ciertas realizaciones, el cultivo in vitro de las células se somete a una perturbación ambiental, y el cultivo in vitro de las células testigo correspondiente es de células idénticas que no se sometieron a la perturbación ambiental.

En ciertas realizaciones, la perturbación ambiental comprende uno o más de un contacto con un agente, un cambio en la condición de cultivo, una modificación / mutación genética introducida y un vehículo *(por ejemplo,* vector) que provoca una modificación / mutación genética.

En ciertas realizaciones, el primer conjunto de datos comprende niveles de expresión de proteína y/o ARNm de la pluralidad de genes.

En ciertas realizaciones, el primer conjunto de datos comprende además uno o más datos lipidómicos, datos metabolómicos, datos

transcriptómicos y datos de polimorfismos de nucleótidos simples (SNP).

En ciertas realizaciones, el segundo conjunto de datos comprende uno o más de perfil bioenergético, proliferación celular, apoptosis, función organelar y una asociación genotipo-fenotipo actualizados mediante modelos funcionales seleccionados de ensayos ATP, ROS, OXPHOS y Seahorse.

En ciertas realizaciones, la etapa (4) se lleva a cabo mediante una plataforma informática basada en inteligencia artificial (AI).

En ciertas realizaciones, la plataforma informática basada en AI comprende REFS(TM).

En ciertas realizaciones, la plataforma informática basada en AI recibe todos los datos de entrada del primer conjunto de datos y el segundo conjunto de datos sin aplicar un punto de corte estadístico.

En ciertas realizaciones, la red de relaciones causales de consenso establecida en la etapa (4) se refina adicionalmente hasta una red de relaciones causales de estimulación, antes de la etapa (5), mediante estimulación *in silico* basada en los datos de

entrada, para proporcionar un nivel de confianza de predicción para una o más relaciones causales dentro de la red de relaciones causales de consenso.

En ciertas realizaciones, la relación causal única se identifica como parte de una red de relaciones causales diferenciales que está presente exclusivamente en células y está ausente en las células testigo correspondientes.

En ciertas realizaciones, el método comprende además validar la relación causal única identificada en un sistema biológico.

En otro aspecto, la invención se refiere a un método para identificar un modulador de un proceso de enfermedad, comprendiendo dicho método: (1) establecer un modelo de enfermedad para el proceso de enfermedad, usando células relacionadas, para representar un aspecto característico del proceso de enfermedad; (2) obtener un primer conjunto de datos a partir del modelo de enfermedad, en donde el primer conjunto de datos representa niveles de expresión de una pluralidad de genes en las células relacionadas; (3) obtener un segundo conjunto de datos a partir del modelo de enfermedad, en donde el segundo conjunto de datos representa una actividad funcional o una respuesta celular de las células relacionadas con la enfermedad; (4) generar una red de relaciones causales de consenso entre los niveles de expresión de la pluralidad de genes y la actividad

funcional o respuesta celular basada solamente en el primer conjunto de datos y el segundo conjunto de datos usando un dispositivo informático programado, en donde la generación de la red de relaciones causales de consenso no se basa en ninguna relación biológica conocida distinta del primer conjunto de datos y el segundo conjunto de datos; (5) identificar, a partir de la red de relaciones causales de consenso, una relación causal única en el proceso de enfermedad, en donde un gen asociado con la relación causal única se identifica como un modulador de un proceso de enfermedad.

En ciertas realizaciones, el proceso de enfermedad es cáncer, diabetes, obesidad o enfermedad cardiovascular.

En ciertas realizaciones, el cáncer es cáncer pulmonar, cáncer de mama, cáncer de próstata, melanoma, carcinoma escamoso, cáncer colorrectal, cáncer pancreático, cáncer tiroideo, cáncer endometrial, cáncer de vejiga, cáncer de riñones, un tumor sólido, leucemia, linfoma no hodgkiniano o un cáncer resistente a fármacos.

En ciertas realizaciones, el modulador estimula o activa el proceso de enfermedad.

En ciertas realizaciones, el modulador inhibe proceso de enfermedad.

En ciertas realizaciones, el modulador cambia la vía metabólica energética específicamente en las células de la enfermedad de una vía glucolítica hacia una vía de fosforilación oxidativa.

En ciertas realizaciones, el modelo de enfermedad comprende un cultivo in vitro de células de enfermedad, que comprende además opcionalmente un cultivo in vitro correspondiente de células testigo o normales.

En ciertas realizaciones, el cultivo in vitro de las células de enfermedad se somete a una perturbación ambiental, y el cultivo in vitro de las células testigo correspondiente es de células de enfermedad idénticas que no se sometieron a la perturbación ambiental.

En ciertas realizaciones, la perturbación ambiental comprende uno o más de un contacto con un agente, un cambio en la condición del cultivo, un modificación / mutación genética introducida y un vehículo *(por ejemplo,* vector) que provoca una modificación / mutación genética.

En ciertas realizaciones, el aspecto característico del proceso de enfermedad comprende una condición de hipoxia, una condición hiperglucémica, una condición de cultivo rico en ácido láctico o combinaciones de los mismos.

En ciertas realizaciones, el primer conjunto de datos comprende niveles de expresión de proteína y/o ARNm de la pluralidad de genes.

En ciertas realizaciones, el primer conjunto de datos comprende además uno o más datos lipidómicos, datos metabolómicos, datos transcriptómicos y datos de polimorfismos de nucleótidos simples (SNP).

En ciertas realizaciones, el segundo conjunto de datos comprende uno o más de perfil bioenergético, proliferación celular, apoptosis, función organelar y una asociación genotipo-fenotipo actualizados mediante modelos funcionales seleccionados de ensayos ATP, ROS, OXPHOS y Seahorse.

En ciertas realizaciones, la etapa (4) se lleva a cabo mediante una plataforma informática basada en inteligencia artificial (AI).

En ciertas realizaciones, la plataforma informática basada en AI comprende REFS(TM).

En ciertas realizaciones, la plataforma informática basada en AI recibe todos los datos de entrada del primer conjunto de datos y el segundo conjunto de datos sin aplicar un punto de corte estadístico.

En ciertas realizaciones, la red de relaciones causales de consenso establecida en la etapa (4) se refina adicionalmente hasta una red de relaciones causales de estimulación, antes de la etapa (5), mediante estimulación *in silico* basada en los datos de entrada, para proporcionar un nivel de confianza de predicción para una o más relaciones causales dentro de la red de relaciones causales de consenso.

En ciertas realizaciones, la relación causal única se identifica como parte de una red de relaciones causales diferenciales que está presente exclusivamente en células de enfermedad y está ausente en las células testigo correspondientes.

En ciertas realizaciones, el método comprende además validar la relación causal única identificada en un sistema biológico.

En otro aspecto, la invención se refiere a un método para proporcionar un modelo para un sistema biológico para uso en un

método de plataforma, que comprende: establecer un modelo para un sistema biológico, usando células asociadas con el sistema biológico, para representar un aspecto característico del sistema biológico, en donde el modelo para el sistema biológico es útil para generar conjuntos de datos usados en el método de plataforma; proporcionando así un modelo para un sistema biológico para uso en un método de plataforma.

En otro aspecto, la invención se refiere a un método para obtener un primera conjunto de datos y un segundo conjunto de datos a partir de un modelo para un sistema biológico para uso en un método de plataforma, que comprende: (1) obtener un primer conjunto de datos a partir del modelo para un sistema biológico para uso en un método de plataforma, en donde el modelo para el sistema biológico comprende células asociadas con el sistema biológico y en donde el primer conjunto de datos representa niveles de expresión de una pluralidad de genes en las células asociadas con el sistema biológico; (2) obtener un segundo conjunto de datos a partir del modelo para un sistema biológico para uso en el método de plataforma, en donde el segundo conjunto de datos respuesta una actividad funcional o una respuesta celular de las células asociadas con el sistema biológico; obteniendo así un primer conjunto de datos y un segundo conjunto de datos a partir del modelo para el sistema biológico para uso en un método de plataforma.

En otro aspecto, la invención se refiere a un método para identificar un modulador de un sistema biológico, comprendiendo dicho método: (1) generar una red de relaciones causales de consenso entre un primer conjunto de datos y un segundo conjunto de datos obtenidos a partir de un modelo para el sistema biológico, en donde el modelo comprende células asociadas con el sistema biológico y en donde el primer conjunto de datos representa niveles de expresión de una pluralidad de genes en las células y el segundo conjunto de datos representa una actividad funcional o una respuesta celular de las células, usando un dispositivo informático programado, en donde la generación de la red de relaciones causales de consenso no se basa en ninguna relación biológica conocida distinta del primer conjunto de datos y el segundo conjunto de datos; (2) identificar, a partir de la red de relaciones causales de consenso, una relación causal única en el sistema biológico, en donde se identifica un gen asociado con la relación causal única como un modulador del sistema biológico; identificando así un modulador de un sistema biológico.

En otro aspecto, la invención se refiere a un método para identificar un modulador de un sistema biológico, comprendiendo dicho método: 1) proporcionar un red de relaciones causales de consenso generada a partir de un modelo para el sistema biológico; 2) identificar, a partir de la red de relaciones

causal de consenso, una relación causal única en el sistema biológico, en donde se identifica un gen asociado con la relación causal única como un modulador del sistema biológico; identificando así un modulador de un sistema biológico.

En ciertas realizaciones de los diversos métodos, la red de relaciones causales de consenso se genera entre un primer conjunto de datos y un segundo conjunto de datos obtenidos a partir del modelo para el sistema biológico, en donde el modelo comprende células asociadas con el sistema biológico y en donde el primer conjunto de datos representa niveles de expresión de una pluralidad de genes en las células y el segundo conjunto de datos representa una actividad funcional o una respuesta celular de las células, usando un dispositivo informático programado, en donde la generación de la red de relaciones causales de consenso no se basa en ninguna relación biológica conocida distinta del primer conjunto de datos y el segundo conjunto de datos.

En otro aspecto, la invención se refiere a un método para proporcionar un modelo de enfermedad para uso en un método de plataforma, que comprende: establecer un modelo de enfermedad para un proceso de enfermedad, usando células relacionadas con la enfermedad, para representar un aspecto característico del proceso de enfermedad, en donde el modelo de establece es útil para generar conjuntos de datos de modelo de enfermedad usados en

el método de plataforma; proporcionando así un modelo de enfermedad para uso en un método de plataforma.

En otro aspecto, la invención se refiere a un método para obtener un primera conjunto de datos y un segundo conjunto de datos a partir de un modelo de enfermedad para uso en un método de plataforma, que comprende: (1) obtener un primer conjunto de datos a partir de un modelo de enfermedad para uso en un método de plataforma, en donde el modelo de enfermedad comprende células relacionadas con la enfermedad, y en donde el primer conjunto de datos representa niveles de expresión de una pluralidad de genes en las células relacionadas con la enfermedad; (2) obtener un segundo conjunto de datos a partir del modelo de enfermedad para uso en un método de plataforma, en donde el segundo conjunto de datos representa una actividad funcional o una respuesta celular de las células relacionadas con la enfermedad; obteniendo así un primer conjunto de datos y un segundo conjunto de datos a partir del modelo de la enfermedad; obteniendo así una primer conjunto de datos y un segundo conjunto de datos a partir de un modelo de enfermedad para uso en un método de plataforma.

En otro aspecto, la invención se refiere a un método para identificar un modulador de un proceso de enfermedad, comprendiendo dicho método: (1) generar una red de relaciones causales de consenso entre un primer conjunto de datos y un

segundo conjunto de datos obtenidos a partir de un modelo de enfermedad, en donde el modelo de enfermedad comprende células de la enfermedad y en donde el primer conjunto de datos representa niveles de expresión de una pluralidad de genes en las células y el segundo conjunto de datos representa una actividad funcional o una respuesta celular de las células, usando un dispositivo informático programado, en donde la generación de la red de relaciones causales de consenso no se basa en ninguna relación biológica conocida distinta del primer conjunto de datos y el segundo conjunto de datos; (2) identificar, a partir de una red de relaciones causales de consenso, una relación causal única en el proceso de la enfermedad, en donde se identifica un gen asociado con la relación causal única como un modulador de un proceso de enfermedad; identificando así un modulador de un proceso de enfermedad.

En otro aspecto, la invención se refiere a un método para identificar un modulador de un proceso de enfermedad, comprendiendo dicho método: 1) proporcionar un red de relaciones causales de consenso generada a partir de un modelo de enfermedad para el proceso de enfermedad; 2) identificar, a partir de la red de relaciones causal de consenso, una relación causal única en el proceso de enfermedad, en donde se identifica un gen asociado con la relación causal única como un modulador de un proceso de

enfermedad; identificando así un modulador de un proceso de enfermedad.

En ciertas realizaciones, la red de relaciones causales de consenso se genera entre un primer conjunto de datos y un segundo conjunto de datos obtenidos a partir de un modelo de enfermedad para el proceso de enfermedad, en donde el modelo de enfermedad comprende células de la enfermedad y en donde el primer conjunto de datos representa niveles de expresión de una pluralidad de genes en las células relacionadas con la enfermedad y el segundo conjunto de datos representa una actividad funcional o una respuesta celular de las células, usando un dispositivo informático programado, en donde la generación de la red de relaciones causales de consenso no se basa en ninguna relación biológica conocida distinta del primer conjunto de datos y el segundo conjunto de datos.

En ciertas realizaciones, la "perturbación ambiental", también denominada en la presente "componente de estímulo externo", es un agente terapéutico. En ciertas realizaciones, el componente de estímulo externo es una molécula pequeña *(por ejemplo,* una molécula pequeña de no más de 5 kDa, 4 kDa, 3 kDa, 2 kDa, 1 kDa, 500 Dalton o 250 Dalton). En ciertas realizaciones, el componente de estímulo externo es biológico. En ciertas realizaciones, el componente de estímulo externo es un químico. En ciertas

realizaciones, el componente de estímulo externo es endógeno o exógeno a las células. En ciertas realizaciones, el componente de estímulo externo es un MIM o desplazador epi. En ciertas realizaciones, el componente de estímulo externo es un factor de estrés para el sistema celular, tal como hipoxia, hiperglucemia, hiperlipidemia, hiperinsulinemia, y/o condiciones ricas en ácido láctico.

En ciertas realizaciones, el componente de estímulo externo puede incluir un agente terapéutico o un agente terapéutico candidato para tratar una enfermedad, incluido un agente quimioterapéutico, fármacos biológicos basados en proteínas, anticuerpos, proteínas de fusión, fármacos de molécula pequeña, lípidos, polisacáridos, ácidos nucleicos, *etc.*

En ciertas realizaciones, el componente de estímulo externo puede ser uno o más factores de estrés, tales como los encontrados típicamente *in vivo* en las diversas enfermedades, incluidos hipoxia, condiciones hiperglicémicas, ambiente ácido (que puede imitarse mediante tratamiento con ácido láctico), *etc.*

En otras realizaciones, el componente de estímulo externo puede incluir uno o más MIM y/o desplazadores epi, tales como se definen más adelante. Los MIM ejemplares incluyen la Coenzima Q10

(también denominada en la presente CoQ10) y compuestos en la familia de la Vitamina B, o nucleósidos, mononucleótidos o dinucleótidos que comprenden un compuesto en la familia de la Vitamina B.

Al llevar a cabo mediciones de rendimiento celular (tales como expresión proteica), se pueden usar cantidades absolutas (por ejemplo, cantidad de expresión) o niveles relativos (por ejemplo, nivel de expresión relativo). En una realización, se usan cantidades absolutas (por ejemplo, cantidades de expresión). En una realización, se usan cantidades o niveles relativos (por ejemplo, niveles de expresión relativos). Por ejemplo, para determinar el nivel de expresión proteica relativo de un sistema celular, la cantidad de cualquier proteína dada en el sistema celular, con o sin el estímulo externo para el sistema celular, puede comprarse con una línea celular o mezcla de líneas celulares testigo adecuadas (tales como todas las células usadas en el mismo experimento) y proporcionar un valor de aumento o disminución. El experto en la técnica apreciará que se pueden emplear cantidades absolutas o cantidades relativas en cualquier medición de rendimiento celular, tal como nivel de transcripción de gen y/o ARN, nivel lipídico o cualquier rendimiento funcional, *por ejemplo,* nivel de apoptosis, nivel de toxicidad o ECAR o OCR, tales como se describen en la presente. Se puede usar un nivel umbral predeterminado para un aumento *(por ejemplo,* al menos 1,2,

1,3, 1,4, 1,5, 1,6, 1,7, 1,8, 1,9, 2, 2,5, 3, 3,5, 4, 4,5, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75 o 100 o un aumento mayor) o disminución *(por ejemplo,* al menos una disminución de 0,9, 0,8, 0,75, 0,7, 0,6, 0,5, 0,45, 0,4, 0,35, 0,3, 0,25, 0,2, 0,15, 0,1 o 0,05 veces o una disminución del 90%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10% o 5% o menor) para seleccionar diferenciales considerables, y a continuación, los datos del rendimiento celular para los diferenciales considerables pueden incluirse en los conjuntos de datos (por ejemplo, primero y segundo conjuntos de datos) utilizados en los métodos de tecnología de plataforma de la invención. Todos los valores presentados en la lista precedente también pueden ser el límite superior o inferior de los rangos, *por ejemplo,* entre 1,5 y 5 veces, 5 y 10 veces, 2 y 5 veces o entre 0,9 y 0,7, 0,9 y 0,5 o 0,7 y 0,3 veces, y se pretende que sean parte de la presente invención.

A lo largo de la presente solicitud, todos los valores presentados en una lista, por ejemplo, tales como los mencionados anteriormente, también puede ser el límite superior o inferior de rangos que se pretende que sean parte de la presente invención.

En una realización de los métodos de la invención, no toda relación causal observada en una red de relaciones causales puede ser de importancia biológica. En relación con cualquier sistema biológico dado para el cual se aplica la evaluación biológica

interrogativa objeto, algunas (o tal vez todas) las relaciones causales (y los genes asociados con las mismas) pueden ser "determinativas" en relación con el problema biológico específico en cuestión, *por ejemplo,* responsable por provocar una enfermedad (un objetivo potencial para la intervención terapéutica) o es un biomarcador para la enfermedad (un potencial factor de diagnóstico o pronóstico). En una realización, una relación causal única observada en el sistema biológico es determinativa en relación con el problema biológico específico en cuestión. En una realización, una relación causal única observada en el sistema biológico es determinativa en relación con el problema específico en cuestión.

Dichas relaciones causales determinativas pueden ser seleccionados por un usuario final del método objeto, o pueden ser seleccionadas por un programa informático de bioinformática, tal como REFS, el programa de análisis comparativo de vías con capacidad DAVID o el programa de análisis de vías KEGG. En ciertas realizaciones, más de una programa informático de bioinformática se usa y se prefieren resultados por consenso de dos o más programas informáticos de bioinformática.

Tal como se usa en la presente, "diferenciales" de rendimientos celulares incluyen diferencias *(por ejemplo,* niveles aumentados o disminuidos) en cualquier uno o más parámetros de los

rendimientos celulares. En ciertas realizaciones, los diferenciales se seleccionan cada uno independientemente del grupo que consiste en diferenciales en la transcripción de ARNm, expresión proteica, actividad proteica, nivel de metabolitos / intermedio y/o interacción ligando-objetivo. Por ejemplo, en términos de nivel de expresión proteica, los diferencias entre dos rendimientos celulares, tales como los rendimientos asociados con un sistema celular antes y después del tratamiento mediante un componente de estímulo externo, puede medirse y cuantificarse usando tecnologías conocidas en la técnica, tales como ensayos basados en espectrometría de masas *(por ejemplo,* iTRAQ, 2D-LC-MSMS, *etc.)*.

En un aspecto, el modelo celular para un sistema biológico comprende un sistema de interacción celular, en donde un primer sistema celular que tiene un primer ambiente celular con un componente de estímulo externo genera un primer ambiente celular modificado; de forma que se establece un sistema celular de interacción exponiendo un segundo sistema celular que tiene un segundo ambiente celular al primer ambiente celular modificado. En una realización, se genera al menos un diferencial de interacción celular considerable del sistema celular de interacción; y se identifica al menos un diferencial de interacción celular determinativo de forma que se produce una evaluación biológico interrogativa. En ciertas realizaciones, el

al menos un diferencial de interacción celular considerable es una pluralidad de diferenciales.

En ciertas realizaciones, el al menos un diferencial de interacción celular determinativo es seleccionado por el usuario final. Alternativamente, en otra realización, el al menos un diferencial de interacción celular determinativo se selecciona mediante un programa informático de bioinformática (tal como, por ejemplo, REFS, análisis de vías KEGG o análisis comparativo de vías con capacidad DAVID) basándose en los datos proteómicos cuantitativos.

En ciertas realizaciones, el método además comprende generar un diferencial de rendimiento celular considerable para el primer sistema celular.

En ciertas realizaciones, los diferenciales se seleccionan cada uno independientemente del grupo que consiste en diferenciales en la transcripción de ARNm, expresión proteica, actividad proteica, nivel de metabolitos / intermedio y/o interacción ligando-objetivo.

En ciertas realizaciones, el primer sistema celular y el segundo sistema celular se seleccionan independientemente de: un

población homogénea de células primarias, una línea celular de cáncer o una línea celular normal.

En ciertas realizaciones, el primer ambiente celular modificado comprende factores secretados por el primer sistema celular en el primer ambiente celular, como resultado de la puesta en contacto del primer sistema celular con el componente de estímulo externo. Los factores pueden comprender proteínas secretadas u otras moléculas de señalización. En ciertas realizaciones, el primer ambiente celular modificado es sustancialmente libre del componente de estímulo externo original.

En ciertas realizaciones, el sistema celular de interacción comprende una placa Transwell que tiene un compartimento de inserto y un compartimento de pocillo separados por una membrana. Por ejemplo, el primer sistema celular puede cultivarse en el compartimento de inserto (o el compartimento de pocillo) y el segundo sistema celular puede cultivarse en el compartimento de pocillo (o el compartimento de inserto).

En ciertas realizaciones, el sistema celular de interacción comprende un primer cultivo para cultivar el primer sistema celular y un segundo cultivo para cultivar el segundo sistema celular.

En este caso, el primer ambiente celular modificado puede ser un medio acondicionado del primer sistema celular.

En ciertas realizaciones, el primer ambiente celular y segundo el ambiente celular pueden ser idénticos. En ciertas realizaciones, el primer ambiente celular y segundo el ambiente celular pueden ser diferentes.

En ciertas realizaciones, el sistema celular de interacción comprende un cocultivo del primer sistema celular y el segundo sistema celular.

Los métodos de la invención pueden usarse, o aplicarse, en cualquier cantidad de "evaluaciones biológicas interrogativas". La aplicación de los métodos de la invención en una evaluación biológica interrogativa posibilitan la identificación de uno o más moduladores de un sistema biológico o "activadores" de proceso celular determinativo de un sistema o proceso biológico. Los métodos de la invención pueden usarse para llevar a cabo un amplio rango de evaluaciones biológicas interrogativas. En ciertas realizaciones, la evaluación biológica interrogativa es el diagnóstico de un estado de enfermedad. En ciertas realizaciones, la evaluación biológica interrogativa es la determinación de la eficacia de un fármaco. En ciertas realizaciones, la evaluación biológica interrogativa es la determinación de la toxicidad de un fármaco. En ciertas

realizaciones, la evaluación biológica interrogativa es la estadificación de un estado de enfermedad. En ciertas realizaciones, la evaluación biológica interrogativa identifica objetivos para cosméticos contra el envejecimiento.

Tal como se usa en la presente, una "evaluación biológica interrogativa" puede incluir la identificación de uno o más moduladores de un sistema biológico, por ejemplo, "activadores" de proceso celular determinativo *(por ejemplo,* un aumento o disminución en la actividad de una vía biológica, o miembros clave de la vía o reguladores clave de miembros de la vía) asociados con la perturbación ambiental o componente de estímulo externo, o una relación causal única en un sistema o proceso biológico. Puede incluir además etapas adicionales diseñadas para evaluar o verificar si los activadores de proceso celular determinativos identificados son necesarios y/o suficientes para eventos descendentes asociados con la perturbación ambiental o componente de estímulo externo, incluidos experimentos con modelos animales *in vivo* y/o con cultivo de tejidos *in vitro* . En ciertas realizaciones, la evaluación biológica interrogativa es el diagnóstico o estadificación de un estado de enfermedad, en donde los moduladores identificados de un sistema biológico, por ejemplo, activadores de proceso celular determinativos (por ejemplo, diferenciales de interacción o relaciones causales únicas en un sistema o proceso biológico) representan marcadores

de enfermedad u objetivos terapéuticos que pueden someterse a intervención terapéutica. La evaluación biológica interrogativa es adecuada para cualquier enfermedad en teoría, pero puede ser particularmente útil en áreas tales como la oncología / biología del cáncer, diabetes, obesidad, enfermedades cardiovasculares y afecciones neurológicas (especialmente enfermedades neurodegenerativas, tales como, a modo no taxativo, enfermedad de Alzheimer, enfermedad de Parkinson, enfermedad de Huntington, esclerosis lateral amiotrófica (ELA) y neurodegeneración relacionada con el envejecimiento).

En ciertas realizaciones, la evaluación biológica interrogativa es la determinación de la eficacia de un fármaco, en donde los moduladores identificados de un sistema biológico, por ejemplo, activador de proceso celular determinativo (por ejemplo, diferenciales de interacción o relaciones causales únicas en un proceso o sistema biológico) pueden ser los signos característicos de un fármaco exitoso y a su vez pueden usarse para identificar agentes adicionales, tales como MIM o desplazadores epi, para tratar la misma enfermedad.

En ciertas realizaciones, la evaluación biológica interrogativa es la identificación de objetivos de fármacos para prevenir o tratar infecciones, en donde el activador de proceso celular determinativo identificado (por ejemplo, diferenciales de interacción o relaciones causales únicas en un proceso o sistema

biológico) puede ser marcadores/indicadores o moléculas biológicas clave causantes del estado infeccioso, y a su vez pueden usarse para identificar agentes antiinfecciosos.

En ciertas realizaciones, la evaluación biológica interrogativa es la evaluación de un efecto molecular de un agente, *por ejemplo, un* fármaco, en un perfil de enfermedad dado, en donde los moduladores identificados de un sistema biológico, por ejemplo, activador de proceso celular determinativo (por ejemplo, diferenciales de interacción celular o relaciones causales únicas en un sistema o proceso biológico) pueden ser un aumento o disminución en la actividad de una o más vías biológicas, o miembros clave de la vía(s), o reguladores clave de miembros de la vía(s), y a su vez pueden usarse, *por ejemplo,* para predecir la eficacia terapéutica del agente para la enfermedad dada.

En ciertas realizaciones, la evaluación biológica interrogativa es la evaluación del perfil toxicológico de un agente, *por ejemplo,* un fármaco, en una célula, tejido, órgano u organismo, en donde los moduladores identificados de un sistema biológico, por ejemplo, activador de proceso celular determinativo (por ejemplo, diferenciales de interacción celular o relaciones causales únicas en un sistema o proceso biológico) pueden ser indicadores de toxicidad, *por ejemplo,* citotoxicidad y a su vez puede usarse para predecir o identificar el perfil toxicológico del agente. En una realización, los moduladores identificados de

un sistema biológico, por ejemplo, activador de proceso celular determinativo (por ejemplo, diferenciales de interacción celular o relaciones causales únicas en un sistema o proceso biológico) es un indicador de la cardiotoxicidad de un fármaco o fármaco candidato, y a su vez, puede usarse para predecir o identificar el perfil cardiotoxicológico del fármaco o fármaco candidato.

En ciertas realizaciones, la evaluación biológica interrogativa es la identificación de objetivos de fármacos para prevenir o tratar una enfermedad o trastorno provocado por armas biológicas, tales como protozoarios, hongos, bacterias, protestos, virus o toxinas, en donde los moduladores identificados de un sistema biológico, por ejemplo, el activador de proceso celular determinativo (por ejemplo, diferenciales de interacción o relaciones causales únicas en un proceso o sistema biológico) pueden ser marcadores/indicadores o moléculas biológicas clave causantes de dicha enfermedad o trastorno, y a su vez pueden usarse para identificar agentes de biodefensa.

En ciertas realizaciones, la evaluación biológica interrogativa es la identificación de objetivos para agentes contra el envejecimiento, tales como cosméticos contra el envejecimiento, en donde los moduladores identificados de un sistema biológico, por ejemplo, el activador de proceso celular determinativo (por ejemplo, diferenciales de interacción o relaciones causales únicas en un proceso o sistema biológico) pueden ser marcadores o

indicadores del proceso de envejecimiento, específicamente el proceso de envejecimiento en la piel, y a su vez pueden usarse para identificar agentes contra el envejecimiento.

En un modelo celular ejemplar para el envejecimiento que se usa en los métodos de la invención para identificar objetivos para cosméticos contra el envejecimiento, el modelo celular comprende una célula epitelial envejecida que, por ejemplo, se trata con luz UV (una perturbación ambiental o componente de estímulo externo), y/o células neonatales, que también se tratan opcionalmente con luz UV. En una realización, un modelo celular del envejecimiento comprende un sistema de interacción celular. En un sistema de interacción entre dos células ejemplar establecido para identificar objetivos para cosméticos contra el envejecimiento, una célula epitelial envejecida (primer sistema celular) puede tratarse con luz UV (un componente de estímulo externo), y se pueden medir cambios, por ejemplo, cambios proteómicos y/o cambios funcionales, en una célula neonatal (segundo sistema celular) resultante del contacto de las células neonatales con medio acondicionado de la célula epitelial envejecida tratada, por ejemplo, pueden medirse cambios proteómicos usando espectrometría de masas cuantitativa convencional, o puede identificarse una relación causal única de una red de relaciones causales generada a partir de los datos.

En otro aspecto, la invención proporciona un kit para llevar a cabo una evaluación biológica interrogativa usando una tecnología de plataforma descubierta, que comprende uno o más reactivos para detectar la presencia de, y/o para cuantificar la cantidad de, un analito que es el objeto de una red de relaciones causales generada a partir de los métodos de la invención. En una realización, dicho analito es el objeto de una relación causal única en el sistema biológico, por ejemplo, un gen asociado con una relación causal única en el sistema biológico. En ciertas realizaciones, el analito es una proteína, y los reactivos comprenden un anticuerpo contra la proteína, una etiqueta para la proteína, y/o uno o más agentes para preparar la proteína para análisis de alta productividad *(por ejemplo,* secuenciación basada en espectrometría de masas).

Se entenderá que todas las realizaciones descritas en la presente, incluidas aquellas descritas solo en los ejemplos, son partes de la descripción general de la invención, y pueden combinarse con cualquier otra realización de la invención a menos que se desestimen o consideren inaplicables explícitamente.

Breve descripción de los dibujos

A continuación, en la presente se describirán varias realizaciones de la presente divulgación en referencia a las figuras, en donde:

Figura 1: Ilustración del abordaje para identificar agentes terapéuticos.

Figura 2: Ilustración de sistemas de biología del cáncer y la consecuencia de la regulación de rendimiento interactivo multifisiológico integrado.

Figura 3: Ilustración de interrogación sistémica de importancia biológica usando
MIMS.

Figura 4: Ilustración de red de modelado del cáncer para permitir una consulta biológica interrogativa.

Figura 5: Ilustración de la tecnología de plataforma de biología interrogativa.

Figura 6: Ilustración de tecnologías empleadas en la tecnología de plataforma.

Figura 7: Representación esquemática de los componentes de la plataforma incluida la recolección de datos, la integración de datos y la exploración de datos.

Figura 8: Representación esquemática de la interrogación sistemática usando MIMS y recolección de datos de respuesta de la cascada "ómica".

Figura 9: Dibujo de los componentes empleados para construir los modelos *in vitro* que representan los estados normal y diabético.

Figura 10: Representación esquemática de la plataforma informática REFS ™usada para generar redes causales de la proteína así como su relación con la patofisiología de la enfermedad.

Figura 11: Representación esquemática del abordaje para la generación de una red diferencial en estados diabéticos con respecto a normales y nodos diabéticos que se restauran a estados normales mediante el tratamiento con MIMS.

Figura 12: Una red diferencial representativa en estados diabéticos con respecto a estados normales.

Figura 13: Una representación esquemática de un nodo y bordes de interés asociados (Nodelin en el centro). Se representa la funcionalidad celular asociada con cada borde.

Figura 14: Diagrama de flujo de alto nivel de un método ejemplar, de acuerdo con algunas realizaciones.

Figura 15A-15D: Ilustración esquemática de alto nivel de los componentes y proceso para un sistema informático basado en AI que puede usarse con realizaciones ejemplares.

Figura 16: Diagrama de flujo de proceso en sistema informático basado en AI que puede usarse con algunas realizaciones ejemplares.

Figura 17: Representa esquemáticamente un ambiente informático ejemplar adecuado para poner en práctica realizaciones ejemplares descritas en la presente.

Figura 18: Ilustración de un diseño de estudio de caso descrito en el Ejemplo 1.

Figura 19: Efecto de tratamientos con CoQ10 en nodos descendentes.

Figura 20: El tratamiento con CoQ10 disminuye la expresión de LDHA en la línea celular de cáncer HepG2.

Figura 21: Red de consenso de interacción proteica ejemplar a 70% de frecuencia de fragmento basada en datos de las líneas celulares Paca2, HepG2 y THLE2.

Figura 22: Se identificaron proteínas sensibles a la estimulación de la expresión de LDHA en dos líneas celulares de cáncer usando la tecnología de plataforma.

Figura 23: Análisis con Ingenuity Pathway Assist® de la red LDHA - PARK7 identifica a TP53 como concentrador ascendente.

Figura 24: Efecto del tratamiento con CoQ10 en los niveles de expresión de TP53 en la línea celular de cáncer SKMEL28.

Figura 25: Activación de TP53 asociada con la expresión alterada de proteínas BCL-2 que producen la apoptosis en la línea celular de cáncer SKMEL28 y efecto del trata con CoQ10 en los niveles de expresión de Bcl-2, Bax y Caspase3 en SKMEL28.

Figura 26: Ilustración del abordaje matemático para la generación de redes delta-delta.

Figura 27: Red de diferencial cáncer-saludable (delta-delta) que activa ECAR y OCR. Cada activador tiene efectos diferenciales en el punto extremo tal como se representa por el espesor del borde. El espesor del borde en cytoscape representa la fuerza del cambio.

Figura 28: Mapeo de PARK7 nodos asociados a partir de los rendimientos de la tecnología de plataforma interrogativa usando IPA: Las formas grises incluyen todos los nodos asociados con PARK7 a partir de los rendimientos de biología interrogativa que se importaron a IPA. La formas sin rellenar (con nombres) son nuevas conexiones incorporadas por IPA para crear un mapa completo.

Figura 29: La tecnología de plataforma interrogativa de la invención, que demuestra nuevas asociaciones de nodos asociados con PARK7. Los bordes que se muestran con líneas con rayas son conexiones entre dos nodos en las simulaciones que tienen nodos intermedios, pero no tienen nodos intermedios en IPA. Los bordes que se muestran con líneas punteadas son conexiones entre dos nodos en las simulaciones que tienen nodos intermedios, pero tienen nodos intermedios diferentes en IPA.

Figura 30: Ilustración del abordaje matemático para la generación de redes delta-delta. Comparación de bordes únicos de NG en la red delta NG∩HG con bordes únicos de HGT1 en la red delta HG∩HGT1. Los bordes en la intersección de NG y HGT1 son bordes HG que s restauran a NG con TI.

Figura 31: Red delta-delta de bordes diabéticos restaurados a normales con tratamiento con la Coenzima Q10 superimpuesta en la red delta NG ∩ HG.

Figura 32: Red delta-delta de bordes hiperlipidémicos restaurados a normales con tratamiento con la Coenzima Q10 superimpuesta en la red delta lipidemia normal ∩ Hiperlipidemia.

Figura 33: Esquema que representa el destino alterado del ácido grasos en la enfermedad y el tratamiento con fármaco. Un equilibrio entre la utilización de ácido graso libre (FFA) para la generación de ATP y remodelación de membrana en respuesta a la alteración de la biología de la membrana ha sido implicado en la cardiotoxicidad inducida por fármaco.

Figura 34: Esquema que representa parámetros de diseño experimental y modelado usados para estudiar la toxicidad inducida por fármaco en cardiomiocitos diabéticos.

Figura 35: Desregulación de red transcripcional y expresión de genes del metabolismo energético mitocondrial humano en cardiomiocitos diabéticos mediante tratamiento con fármaco (T): molécula de rescate (R) normaliza la expresión génica.

Figura 36: A. Tratamiento con fármaco (T) inducido por la expresión de GPAT1 y TAZ en mitocondrias de cardiomiocitos acondicionados en hiperglucemia. En combinación con la molécula de rescate (T+R) los niveles de GPAT1 y TAZ se normalizaron. B. Síntesis de TAG a G3P.

Figura 37: A. Tratamiento con fármaco (T) disminuye la OCR (siglas en inglés para tasa de consumo de oxígeno) mitocondrial en cardiomiocitos acondicionados en hiperglucemia. La molécula de rescate (T+R) normaliza la OCR. B. Tratamiento con fármaco (T) reprime la síntesis mitocondrial de ATP en cardiomiocitos acondicionados en hiperglucemia.

Figura 38: Anotación GO de proteínas reguladas a la baja mediante tratamiento con fármaco. Proteínas que participan en el metabolismo energético mitocondrial se regularon a la baja con tratamiento con fármaco.

Figura 39: Ilustración del abordaje matemático para la generación de redes delta. Compare bordes únicos de T con respecto a UT, estando ambos modelos en ambiente diabético.

Figura 40: Esquema que representa concentradores y redes de proteína potenciales que activan la patofisiología de toxicidad inducida por fármaco.

Descripción detallada de la invención

I. Descripción general

Las realizaciones ejemplares de la presente invención incorporan métodos que pueden llevarse a cabo usando una plataforma de biología interrogativa ("la Plataforma") que es una herramienta para comprender una amplia variedad de procesos biológicos, tales

como patofisiología de la enfermedad, y los activadores moleculares clave subyacentes tales como procesos biológicos, incluidos los factores que posibilitan un proceso de enfermedad. Algunas realizaciones ejemplares incluyen sistemas que pueden incorporar al menos una porción, o toda, la Plataforma. Algunos métodos ejemplares emplean al menos parte de, o toda, la Plataforma. Las metas y objetivos de algunas realizaciones ejemplares en las que participa la plataforma se resumen en general a continuación a efectos ilustrativos:

i) crear distintivos moleculares específicos como activadores de componentes fundamentales del proceso biológico (por ejemplo, proceso de enfermedad) ya que se relacionan con la patofisiología general del proceso biológico;

ii) generar distintivos moleculares o mapas diferenciales que pertenecen al proceso biológico, que pueden ayudar a identificar distintivos moleculares diferenciales que distinguen un estado biológico (por ejemplo, un estado de enfermedad) con respecto a un estado biológico diferente (por ejemplo, un estado normal) y desarrollar una comprensión sobre los distintivos o entidades moleculares ya que arbitran mecanismos de cambio entre los dos estados biológicos (por ejemplo, de un estado normal al de enfermedad); y

iii) investigar el rol de "concentradores" de la actividad molecular como objetivos de intervención potenciales para el control externo del proceso biológico (por ejemplo, para usar el

concentrados como un objetivo terapéutico potencial), o como biomarcadores potenciales para el proceso biológico en cuestión (por ejemplo, biomarcadores específicos de la enfermedad, en usos de pronóstico y/o teranóstico).

Algunos métodos ejemplares en los que participa la Plataforma pueden incluir una o más de las siguientes características:

1) modelar el proceso biológico (por ejemplo, proceso de enfermedad) y/o componentes del proceso biológico (por ejemplo, fisiología y patofisiología de la enfermedad) en uno o más modelos, preferiblemente modelos in vitro, usando células asociadas con el proceso biológico. Por ejemplo, las células pueden ser células derivadas de humanos que participan normalmente en el proceso biológico en cuestión. El modelo puede incluir varias pistas / condiciones / perturbaciones celulares que son específicas del proceso biológico (por ejemplo, enfermedad). De forma ideal, el modelo representa varios estados (de enfermedad) y componentes de flujo, en lugar de una evaluación estática de la condición (enfermedad) biológica.

2) creación de perfiles de distintivos de ARNm y/o proteína usando cualquier medio conocido en la técnica. Por ejemplo, reacción en cadena de la polimerasa cuantitativa (qPCR) y herramientas de análisis proteómico tales como Espectrometría de

masas (MS). Dichos conjuntos de datos de ARNm y proteína representan la reacción biológica al ambiente / perturbación. Cuando sea aplicable y posible, también pueden integrarse datos lipidómicos, metabolómicos y transcriptómicos como medidas complementarias o alternativas para el proceso biológico en cuestión. El análisis SNP es otro componente que puede usarse algunas veces en el proceso. Puede ser útil para la investigación, por ejemplo, si el SNP o una mutación específica tiene algún efecto sobre el proceso biológico. Estas variables pueden usarse para describir el proceso biológico, como una "instantánea" estática, o como una representación de un proceso dinámico.

3) someter a ensayo para determinar una o más respuestas celulares a pistas y perturbaciones, incluidos, a modo no taxativo, creación de perfiles bioenergéticos, proliferación celular, apoptosis y función organelar. La asociación verdadera genotipo-fenotipo se actualiza mediante el empleo de modelos funcionales, tales como ATP, ROS, OXPHOS, ensayos Seahorse, etc. Dichas respuestas celulares representan la reacción de las células en el proceso biológico (o modelos del mismo) en respuesta al estado(s) correspondiente de expresión de ARNm / proteína, y cualesquiera otras estados relacionados en el 2) mencionado anteriormente.

4) integrar los datos de ensayos funcionales obtenidos en 3) con datos proteómicos y otros datos obtenidos en 2), y determinar las asociaciones proteicas impulsadas por la causalidad, empleando un sistema o plataforma informática basada en inteligencia artificial (basada en AI). Dicho sistema basado en AI se basa, y preferiblemente solo se basa en, los conjuntos de datos obtenidos en 2) y/o 3), sin recurrir al conocimiento existente relacionado con el proceso biológico. Preferiblemente, ningún punto de datos se corta estadísticamente ni artificialmente. En cambio, todos los datos obtenidos se introducen en el sistema basado en AI para determinar las asociaciones proteicas. Una meta o rendimiento del proceso de integración es una o más redes diferenciales (pueden llamarse de otra forma en la presente "redes delta" o, en algunos casos, "redes delta-delta" según sea el caso) entre los diferentes estados biológicos (por ejemplo, estado de enfermedad con respecto a estado normal).

5) crear perfiles de rendimientos a partir de la plataforma informática basada en AI para explorar cada concentrador de actividad como un objetivo y/o biomarcador terapéutico potencial. Dicha creación de perfiles puede llevarse a cabo totalmente *in silico* basada en los conjuntos de datos obtenidos, sin recurrir a ningún experimento real en laboratorio húmedo.

6) validar el concentrador de actividad empleando técnicas moleculares y celulares. Dicha validación posinformática de

rendimiento con experimentos basados en células en laboratorio húmedo puede ser opcional, pero ayudan a crear una círculo completo de interrogación.

Todos y cualquiera de los abordajes resumidos anteriormente pueden usarse en cualquier aplicación específica relacionada con cualquier proceso biológico, dependiendo, al menos en parte, de la naturaleza de la aplicación específica. Es decir, uno o más de los abordajes resumidos anteriormente pueden omitirse o modificarse, y uno o más abordajes adicionales pueden emplearse, dependiendo de la aplicación específica.

Se proporcionan varios esquemas que ilustran la plataforma. Específicamente, en la Figura 1 se representa una ilustración de un abordaje ejemplar para identificar agentes terapéuticos usando la plataforma. En la Figura 2 se representa una ilustración de sistemas de biología del cáncer y la consecuencia de la regulación de rendimiento interactivo multifisiológico integrado. En la Figura 3 se representa una ilustración de una interrogación sistemática de importancia biológica usando MIMS. En la Figura 4 se representa una ilustración de un modelado de una red de cáncer para posibilitar una consulta biológica interrogativa.

Las Figuras 5 y 6 representan ilustraciones de la plataforma de biología interrogativa y las tecnologías empleadas en la

plataforma. La Figura 7 es una representación esquemática de los componentes de la plataforma incluida la recolección de datos, la integración de datos y la exploración de datos. La Figura 8 es una representación esquemática de la interrogación sistemática usando MIMS y recolección de datos de respuesta de la cascada "ómica".

La Figura 14 es un diagrama de flujo de alto nivel de un método ejemplar 10, en el que se indican los componentes de un sistema ejemplar que pueden usarse para llevar a cabo el método ejemplar. Inicialmente, se establece un modelo (por ejemplo, un modelo in vitro) para un proceso biológico (por ejemplo, un proceso de enfermedad) y/o componentes del proceso biológico (por ejemplo, fisiología y patofisiología de enfermedad) usando células normalmente asociadas con el proceso biológico (etapa 12). Por ejemplo, las células pueden ser células derivadas de humanos que participan normalmente en el proceso biológico (por ejemplo, enfermedad). El modelo celular puede incluir varias pistas / condiciones / perturbaciones celulares que son específicas para el proceso biológico (por ejemplo, enfermedad). De forma ideal, el modelo celular representa varios estados (de enfermedad) y componentes de flujo del proceso biológico (por ejemplo, enfermedad) en lugar de una evaluación estática del proceso biológico. El modelo celular de comparación puede incluir células testigo o células normales (por ejemplo, sin enfermedad). Se

proporciona una descripción adicional de los modelos celulares en las secciones III.A y IV, a continuación.

Un primer conjunto de datos se obtiene a partir del modelo celular para el proceso biológico, que incluye información que representa niveles de expresión de una pluralidad de genes (por ejemplo, distintivos de ARNm y/o proteína) (etapa 16) usando cualquier proceso o sistema conocido (por ejemplo, reacción en cadena de la polimerasa cuantitativa (qPCR) y herramientas de análisis proteómico tales como Espectrometría de masas (MS, por sus siglas en inglés)).

Un tercer conjunto de datos se obtiene a partir del modelo celular de comparación para el proceso biológico (etapa 18). El tercer conjunto de datos incluye información que representa niveles de expresión de una pluralidad de genes en las células de comparación del modelo celular de comparación.

En ciertas realizaciones de los métodos de la invención, estos primer y tercer conjuntos de datos se denominan colectivamente en la presente "primero conjunto de datos" que representa niveles de expresión de una pluralidad de genes en las células (todas las células incluidas las células de comparación) asociadas con el sistema biológico.

El primer conjunto de datos y el tercer conjunto de datos pueden obtenerse de uno o más sistema(s) de análisis de distintivos de ARNm y/o proteína. Los datos de ARNm y proteína en el primer y tercer conjuntos de datos pueden representar reacciones biológicas al ambiente y/o perturbación. Cuando sea aplicable y posible, también pueden integrarse datos lipidómicos, metabolómicos y transcriptómicos como medidas complementarias o alternativas para el proceso biológico. El análisis de SNP es otro componente que puede usarse algunas veces en el proceso. Puede ser útil para la investigación, por ejemplo, si el polimorfismo de nucleótido simple (SNP) o una mutación específica tiene algún efecto sobre el proceso biológico. Las variables de datos pueden usarse para describir el proceso biológico, como una "instantánea" estática, o como una representación de un proceso dinámico. A continuación, en la sección III.B., figura una descripción adicional relacionada con obtener información que representa niveles de expresión de una pluralidad de genes en células.

Un segundo conjunto de datos se obtiene a partir del modelo celular para el proceso biológico, que incluye información que representa una actividad funcional o respuesta de células (etapa 20). De forma similar, un cuarto conjunto de datos se obtiene a partir del modelo celular para el proceso biológico, que incluye

información que representa una actividad funcional o respuesta de las células de comparación (etapa 22).

En ciertas realizaciones de los métodos de la invención, estos segundo y cuarto conjuntos de datos se denominan colectivamente en la presente "segundo conjunto de datos" que representa una actividad funcional o una respuesta celular de las células (todas las células incluidas las células de comparación) asociadas con el sistema biológico.

Se pueden usar uno o más sistemas de ensayo funcional para obtener información en relación con la actividad funcional o respuesta de células o de células de comparación. La información sobe las respuestas celulares funcionales a pistas y perturbaciones pueden incluir, a modo no taxativo, creación de perfiles bioenergéticos, proliferación celular, apoptosis y función organelar. Se pueden emplear modelos funcionales para procesos y vías (por ejemplo, trifosfato de adenosina (ATP), especies reactiva de oxígeno (ROS), fosforilación oxidativa (OXPHOS), ensayos Seahorse, etc.,) para obtener la verdadera asociación genotipo-fenotipo. La actividad funcional o respuestas celulares representan la reacción de las células en el proceso biológico (o modelos del mismo) en respuesta al estado(s) correspondiente de expresión de ARNm / proteína, y cualesquiera otras condiciones o perturbaciones aplicadas relacionadas. A

continuación, en la sección III.B., se proporciona información en relación con obtener información que representa la actividad funcional o respuesta de las células.

El método también incluye generar modelos implementados por computadora de los procesos biológicos en las células y en las células testigo. Por ejemplo, una o más (por ejemplo, un grupo de) redes Bayesian de relaciones causales entre el nivel de expresión de la pluralidad de genes y la actividad funcional o respuesta celular pueden generarse para el modelo celular (las "redes del modelo celular generadas") a partir del primer conjunto de datos y el segundo conjunto de datos (etapa 24). Las redes de modelo celular generadas, individualmente o colectivamente, incluyen información direccional probabilística cuantitativa en conexión con las relaciones. Las redes de modelo celular generadas no se basan en relaciones biológicas conocidas entre la expresión génica y/o actividad funcional o respuesta celular, distintas del primer conjunto de datos y el segundo conjunto de datos. La uno o más redes de modelo celular generadas pueden denominarse colectivamente red de modelo celular de consenso.

Una o más (por ejemplo, un grupo de) redes Bayesian de relaciones causales entre el nivel de expresión de la pluralidad de genes y la actividad funcional o respuesta celular pueden generarse para

el modelo celular de comparación (las "redes del modelo celular de comparación generadas") a partir del primer conjunto de datos y el segundo conjunto de datos (etapa 26). Las redes de modelo celular de comparación generadas, individualmente o colectivamente, incluyen información direccional probabilística cuantitativa en conexión con las relaciones. Las redes celulares generadas no se basan en relaciones biológicas conocidas entre la expresión génica y/o actividad funcional o respuesta celular, distintas del primer conjunto de datos y el segundo conjunto de datos. La uno o más redes de modelo de comparación generadas pueden denominarse colectivamente red de modelo celular de consenso.

Las redes de modelo celular generadas y las redes de modelo celular de comparación generadas pueden crearse usando una plataforma informática basada en inteligencia artificial (basada en AI). Detalles adicionales relacionados con la creación de las redes de modelo celular generadas,
la creación de las redes de modelo celular de comparación generadas y el sistema informático basado en AI figuran, a continuación, en la sección III. C y en la descripción de las Figuras 2A-3.

Cabe señalar que muchas plataformas o sistemas diferentes basados en AI pueden emplearse para generar las redes Bayesian de

relaciones causales que incluyen información direccional probabilística cuantitativa. Aunque ciertos ejemplos descritos en la presente emplean un sistema específico disponible en el mercado, es decir, REFS ™(Reverse Engineering/Forward Simulation) de GNS (Cambridge, MA), las realizaciones no están limitadas. Las plataformas o sistemas basados en AI adecuados para implementar algunas realizaciones emplean algoritmos matemáticos para establecer relaciones causales entre las variables de entrada (por ejemplo, el primer y segundo conjuntos de datos), basados solo en los datos de entrada sin tomar en cuenta conocimiento anterior existente sobre cualquier relación biológica potencial, establecida y/o verificada.

Por ejemplo, la plataforma informática basada en AI REFS ™utiliza datos biológicos de entrada mínimamente procesados o en bruto derivados experimentalmente (original) (por ejemplo, datos genéticos, genómicos, epigenéticos, proteómicos, metabolómicos y clínicos) y rápidamente lleva a cabo billones de cálculos para determinar cómo interactúan las moléculas entre sí en un sistema completo. La plataforma informática basada en AI REFS ™lleva a cabo un proceso de ingeniería inversa con el fin de crear un modelo celular *in silico* implementado por computadora (por ejemplo, redes de modelo celular generadas), basado en los datos de entrada, que representa cuantitativamente el sistema biológico subyacente. Adicionalmente, se pueden desarrollar hipótesis sobre

el sistema biológico subyacente y simulares rápidamente en base al modelo celular implementado por computadora, a efectos de obtener predicciones, acompañado por niveles de confianza asociados, en relación con las hipótesis.

Con este abordaje, los sistemas biológicos se representan mediante modelos celulares implementados por computadora cuantitativos en los que se simulan "intervenciones" para aprender sobre mecanismos detallados del sistema biológico (por ejemplo, enfermedad), estrategias de intervención eficaces y/o biomarcadores clínicos que determinan qué pacientes responderán a un régimen de tratamiento dado. Los abordajes estadísticos y de bioinformática convencionales, así como abordajes basados en el modelado de la biología conocida, típicamente son incapaces de proporcionar estos tipos de perspectivas.

Después que se crean las redes de modelo celular generadas y las redes de modelo celular de comparación generadas, se comparan. Se identifican una o más relaciones causales presentes en al menos algunas de las redes de modelo celular generadas, y ausentes de, o que tienen al menos un parámetro considerablemente diferente en las redes de modelo celular de comparación generadas (etapa 28). Dicha comparación puede resultar en la creación de un red diferencial. La comparación, identificación y/o creación de la red diferencial (delta) pueden llevarse a cabo usando un módulo

de creación de red diferencial, que se describe en más detalla a continuación en la sección III.D y con respecto a la descripción de la Figura 26.

En algunas realizaciones, los conjuntos de datos de entrada son de un tipo celular y un tipo celular de comparación, que crea un grupo de redes de modelo celular basado en el tipo celular y otro grupo de redes de modelo celular de comparación basado en el tipo celular testigo de comparación. Se puede llevar a cabo un diferencial entre el grupo de redes del tipo celular y el grupo de redes del tipo(s) celular de comparación.

En otras realizaciones, los conjuntos de datos de entrada son de múltiples tipos celulares (por ejemplo, dos o más tipos de células de cáncer) y múltiples tipos celulares de comparación (por ejemplo, dos o más tipos celulares no cancerosos, normales). Se puede generar un grupo de redes de modelo celular para cada tipo celular y cada tipo celular de comparación individualmente, y/o los datos de los múltiples tipos celulares y de los múltiples tipos celulares de comparación pueden combinarse en conjuntos de datos compuestos respectivos. Los conjuntos de datos compuestos producen un grupo de redes correspondiente a los múltiples tipos celulares (datos compuestos) y otro grupo de redes correspondiente a las múltiples tipos celulares de comparación (datos compuestos de comparación). Se puede llevar a cabo un

diferencial entre el grupo de redes para los datos compuestos en comparación con el grupo de redes para los datos compuestos de comparación.

En algunas realizaciones, se puede llevar a cabo un diferencial entre dos redes diferenciales diferentes. Este rendimiento puede denominarse red delta-delta y se describe más adelante en relación con la Figura 26.

Se puede identificar información de relación cuantitativa para cada relación en las redes de modelo celular generadas (etapa 30). De forma similar, se puede identificar la información de relación cuantitativa para cada relación en las redes de modelo celular de comparación generadas (etapa 32). La información cuantitativa en referencia a la relación puede incluir una dirección que indica causalidad, un medición de la incertidumbre estadística con respecto a la relación (por ejemplo, una medición estadística del Área debajo de la curva (AUC)), y/o una expresión de la magnitud cuantitativa de la potencia de la relación (por ejemplo, una vez). Pueden crearse perfiles con las diversas relaciones en las redes de modelo celular generadas usando la información de relación cuantitativa para explorar cada concentrador de actividad en las redes como un objetivo terapéutico y/o biomarcador potencial. Dicha creación de perfiles puede llevarse a cabo totalmente *in silico* basada en los

resultados de las redes de modelo celular generadas, sin recurrir a ningún experimento real en laboratorio húmedo.

En algunas realizaciones, un concentrador de actividad en las redes puede validarse empleando técnicas moleculares y celulares. No es necesario llevar cabo dicha validación posinformática de rendimiento con experimentos basados en células en laboratorio húmedo, pero puede ayudar a crear un círculo completo de interrogación. La Figura 15 representa esquemáticamente una representación de alto nivel simplificada de la funcionalidad de un sistema informático basado en AI ejemplar (por ejemplo, el sistema informático basado en AI REFS™) e interacciones entre el sistema basado en AI y otros elementos o porciones de una plataforma de biología interrogativa ("la Plataforma"). En la Figura 15A, se introducen en un sistema basado en AI varios conjuntos de datos obtenidos a partir de un modelo para un proceso biológico (por ejemplo, un modelo de enfermedad), tal como una dosificación de fármaco, dosificación tratamiento, expresión de proteína, expresión de ARNm y cualquiera de muchas mediciones funcionales asociadas (tales como OCR, ECAR). Tal como se muestra en la Figura 15B, a partir de los conjuntos de datos de entrada, el sistema de AI crea una biblioteca de "fragmentos de red" que incluye variables (proteínas, lípidos y metabolitos) que activan mecanismos moleculares en el proceso biológico (por

ejemplo, enfermedad), en un proceso denominado Bayesian Fragment Enumeration (enumeración de fragmentos de Bayesian) (Figura 15B). En la Figura 15C, el sistema basado en AI selecciona un subconjunto de los fragmentos de red en la biblioteca y construye una red de prueba inicial a partir de los fragmentos. El sistema basado en AI también selecciona un subconjunto diferente de los fragmentos de red en la biblioteca para construir otra red de prueba inicial. Finalmente, se crea un grupo de redes de prueba iniciales (por ejemplo, 1000 redes) a partir de diferentes subconjuntos de fragmentos de red en la biblioteca. Este proceso puede denominarse muestreo de grupos paralelo. Cada red de prueba en el grupo se evoluciona u optimiza agregando, restando y/o sustituyendo fragmentos de red adicionales de la biblioteca. Si se obtienen datos adicionales, los datos adicionales pueden incorporarse a los fragmentos de red en la biblioteca y pueden incorporarse en el grupo de redes de prueba a través de la evolución de cada red de prueba. Después de completar el proceso de optimización/evolución, el grupo de redes de prueba puede describirse como las redes de modelo celar generadas.

Tal como se muestra en la Figura 15D, el grupo de redes de modelo celular generadas puede usarse para estimular el comportamiento del sistema biológico. La simulación puede usarse para predecir el comportamiento del sistema biológico frente a cambios en las condiciones, que pueden verificarse experimentalmente usando

experimentos basados en células o basados en animales en laboratorio húmedo. Adicionalmente, se pueden extraer los parámetros cuantitativos de relaciones en las redes de modelo celular generadas usando la funcionalidad de simulación mediante la aplicación perturbaciones simuladas a cada nodo individualmente mientras se observan los efectos en los otros nodos en las redes de modelo celular generadas. Se proporcionan más detalles más adelante en la sección III.C.

El proceso de ingeniería inversa automatizado del sistema informático basado en AI, que se representa en las Figuras 2A-2D, crea un grupo de redes de modelo celular generadas que es un modelo basado en computadora sistemático e imparcial de las células.

La ingeniería inversa determina las conexiones de red direccional probabilística entre las mediciones moleculares en los datos y los resultados fenotípicos de interés. La variación en las mediciones moleculares posibilita conocer las relaciones de causa y efecto probabilísticas entre estas entidades y cambios en los valores de referencia. La naturaleza de conocimiento de la máquina de la plataforma también posibilita la capacitaciones cruzada y predicciones basadas en un conjunto de datos que está evolucionando constantemente.

Las conexiones de red entre las mediciones moleculares en los datos son "probabilísticas", parcialmente porque la conexión puede basarse en correlaciones entre los conjuntos de datos observados "conocidos" por el algoritmo informático. Por ejemplo, si el nivel de expresión de la proteína X y el de la proteína Y están relacionados positivamente o negativamente, en base al análisis estadístico del conjunto de datos, se puede atribuir una relación causal para establecer una conexión de red entre las proteínas X e Y. La confiabilidad de dicha relación causal putativa puede definirse adicionalmente por una probabilidad de la conexión, que puede medirse por mediante un valor p (por ejemplo, $p < 0,1$, 0,05, 0,01, etc.).

Las conexiones de red entre las mediciones moleculares en los datos son "direccionales", parcialmente porque las conexiones de red entre las mediciones moleculares, tales como se determinan mediante el proceso de ingeniería inversa, refleje la causa y el efecto de la relación entre el gen / proteína conectados, de forma que elevar el nivel de expresión de una proteína puede provocar que el nivel de expresión de la otra se eleve o caiga, dependiendo de si la conexión es estimuladora o inhibidora.

Las conexiones de red entre las mediciones moleculares en los datos son "cuantitativas", parcialmente porque las conexiones de red entre las mediciones moleculares, tales como se determinan mediante el proceso, pueden simularse *in silico,* en base al

conjunto de datos existente y las mediciones probabilísticas asociadas con el mismo. Por ejemplo, en las conexiones de red establecidas entre las mediciones moleculares, puede ser posible aumentar o disminuir teóricamente (por ejemplo, en 1, 2, 3, 5, 10, 20, 30, 50, 100 veces o más) el nivel de expresión de una proteína dada (o un "nodo" en la red) y simular cuantitativamente sus efectos en otras proteicas conectadas en la red.

Las conexiones de red entre las mediciones moleculares en los datos son "imparciales", al menos parcialmente porque ningún punto de datos se corta estadísticamente o artificialmente, y parcialmente porque las conexiones de red se basan solo en datos de entrada, sin referirse a conocimiento preexistente sobre el proceso biológico en cuestión.

Las conexiones de red entre las mediciones moleculares en los datos son "sistémicas" e (imparciales), parcialmente porque todas las conexiones potenciales entre todas las variables de entrada han sido exploradas sistémicamente, por ejemplo, por pares. La confianza en el poder informático para ejecutar dicha exploración sistémica aumenta exponencialmente a medida que la cantidad de variables de entrada aumenta.

En general, un grupo de ~1000 redes es habitualmente suficiente para predecir relaciones cuantitativas causales probabilísticas

entre todas las entidades medidas. El grupo de redes captura la incertidumbre en los datos y permite el cálculo de las métricas de confianza para cada predicción de modelo. Las predicciones generadas usando el grupo de redes juntas, cuando hay diferencias en las predicciones de redes individuales en el grupo representan el grado de incertidumbre en la predicción. Esta característica permite la atribución de métricas de confianza para predicciones de respuesta clínica generadas a partir del modelo.

Una vez que los modelos se someten a ingeniería inversa, se pueden llevar a cabo consultas de estimulación adicionales en el grupo de modelos para determinar activadores moleculares clave para el proceso biológico en cuestión, tal como una enfermedad. En la Figura 9 se muestra un dibujo de componentes empleados para construir modelos *in vitro* ejemplares que representan estados normales y diabéticos. En la Figura 10 se muestra una representación esquemática de la plataforma informática REFS ™ ejemplar usada para generar redes causales de la proteína así como su relación con la patofisiología de la enfermedad. La Figura 11 muestra una representación esquemática del abordaje ejemplar para la generación de una red diferencial en estados diabéticos con respecto a normales y nodos diabéticos que se restauran a estados normales mediante el tratamiento con MIMS. En la Figura 12 se muestra una red diferencial representativa en estados diabéticos con respecto a normales. La Figura 13 se

muestra una representación esquemática de un nodo y bordes asociados de interés (Nodel en el centro) y la funcionalidad celular asociada con cada borde.

Luego de describir en general la invención más atrás, las secciones a continuación proporcionan una descripción más detallada de varios aspectos o elementos de la invención general, en conjunción con uno o más sistemas biológicos específicos que pueden analizarse usando los métodos de la presente. Cabe señalar, sin embargo, que los sistemas biológicos específicos usados a efectos de ilustración más adelante no son limitantes. Por el contrario, se pretende que otros sistemas biológicos diferentes, incluidos cualquier alternativa, modificación y equivalente de los mismos, puedan analizarse de forma similar usando la tecnología de Plataforma objeto.

II. Definiciones

Tales como se usan en la presente, ciertos términos que se pretende definir específicamente, pero que aún no han sido definidos en otras secciones de la memoria descriptiva, se definen en la presente.

Los artículos "el" y "la" se usan en la presente para referirse a uno o más de uno (*es decir,* a al menos uno) del sujeto

gramatical del artículo. A modo de ejemplo, "un elemento " significa un elemento o más de un elemento.

El término "incluye" se usa en la presente para significar, y se usa de forma intercambiable con, la frase "incluidos a modo no taxativo".

El término "o" se usa en la presente para significar, y se usa de forma intercambiable con, el término "y/o", a menos que el contexto indique claramente lo contrario.

La expresión "tal como" se usa en la presente para significar, y se usa de forma intercambiable con, la frase "tal como pero a modo no taxativo".

"Vía metabólica" se refiere a una secuencia de reacciones mediadas por enzimas que transforman un compuesto en otro y proporcionan intermediarios y energía para funciones celulares. La vía metabólica puede ser lineal o cíclica o ramificada.

"Estado metabólico" se refiere al contenido molecular de un ambiente específico celular, multicelular o tejido en un punto dado de tiempo tal como se mide mediante varios indicadores químicos y biológicos, a medida que se relacionan con un estado de salud o enfermedad.

El término "micromatriz" se refiere a una matriz de distintos polinucleótidos, oligonucleótidos, polipéptidos (por ejemplo, anticuerpos) o péptidos sintetizados en una sustrato, tal como papel, nailon u otro tipo de membrana, filtro, chip, portaobjetos o cualquier otro soporte sólido adecuado.

Los términos "trastornos" y "enfermedades" se usan inclusivamente y se refieren a cualquier desviación de la estructura o función normal de cualquier parte, órgano o sistema del cuerpo (o cualquier combinación de los mismos). Una enfermedad específica se manifiesta mediante síntomas y signos característicos, incluidos cambios biológicos, químicos y físicos, y a menudo se asocia con una variedad de otros factores incluidos, a modo no taxativo, factores demográficos, ambientales, de empleo, genéticos y médicamente históricos. Ciertos signos, síntomas y factores característicos relacionados pueden cuantificarse a través de una variedad de métodos para proporcionar importante información diagnóstica.

El término "expresión" incluye el proceso mediante el cual se produce un polipéptido a partir de polinucleótidos, tales como ADN. El proceso puede implicar la transcripción de un gen en ARNm y la traducción de este ARNm en un polipéptido. Dependiendo del contexto en que se usa, "expresión" puede referirse a la producción de ARN, proteína o ambos.

Los términos "nivel de expresión de un gen" o "nivel de expresión génica" se refiere al nivel de ARNm, así como al de transcripto(s) nuevo(s) preARNm, intermediarios del procesamiento del transcripto, ARNm(s) maduros y productos de degradación, o el nivel de proteína, codificada por el gen en la célula.

El término "modulación" se refiere a la regulación ascendente (es decir, activación o estimulación), a la regulación descendente

(es decir, inhibición o supresión) de una respuesta, o las dos en combinación o separadas. Un "modulador" es un compuesto o molécula que modula, y puede ser, por ejemplo, un agonista, antagonista, activador, estimulador, supresor o inhibidor.

El término "Trolamina", tal como se usa en la presente, se refiere a Trolamina NF, Trietanolamina, TEALAN®, TEAlan 99%, Trietanolamina, 99%, Trietanolamina, NF o Trietanolamina, 99%, NF. Estos términos puede usarse de forma intercambiable en la presente.

El término "genoma" se refiere a la totalidad de la información genética de una entidad biológica (célula, tejido, órgano, sistema, organismo). Está codificado en ADN o ARN (en ciertas virus, por ejemplo). El genoma incluye los genes y las secuencias no codificantes del ADN.

El término "proteoma" se refiere al conjunto total de proteínas expresadas por un genoma, un célula, un tejido o un organismo en un momento dado. Más específicamente, puede referirse al conjunto total de proteínas expresadas en tipo de células u organismo dado en un momento dado en condiciones definidas. El proteoma puede incluir variantes proteicas debidas a, por ejemplo,empalme alternativo de genes y/o modificaciones postranslacionales (tales como glucosilación o fosforilación).

El término "transcriptoma" se refiere al conjunto total de moléculas de ARN transcritas, incluidas ARNm, ARNr, ARNt y otros ARN no codificantes producidos en una o una población de conjunto

en un momento dado. El término puede aplicarse al conjunto total de transcriptos en un organismo dado o al subconjunto específico de transcriptos presentes en un tipo de célula específico. A diferencia del genoma, que es aproximadamente fijo para una línea celular dada (excluidas las mutaciones), el transcriptoma puede variar con las condiciones ambientales externas. Debido a que incluye todos los transcritos de ARNm en la célula, el transcriptoma refleja los genes que se están expresando activamente en cualquier momento dado, con la excepción de los fenómenos de degradación del ARNm tales como la atenuación transcripcional.

El estudio de la transcriptómica, también denominadas como creación de perfiles de expresión, examina el nivel de expresión de ARNm en una población celular dada, a menudo usando técnicas de alta productividad basadas en tecnología de micromatrices de ADN.

El término "metaboloma" se refiere al conjunto completo de metabolitos de molécula pequeña (tales como intermediarios metabólicos, hormonas y otras moléculas de señalización y metabolitos secundarios) que se encuentran dentro de una muestra biológica, tal como un organismo simple, en un momento dado en condiciones dadas. El metaboloma es dinámico y puede cambiar de segundo en segundo.

El término "lipidoma" se refiere al conjunto completo de lípidos que se encuentran dentro de una muestra biológica, tal como un

organismo simple, en un momento dado en condiciones dadas. El lipidoma es dinámico y puede cambiar de segundo en segundo.

El término "interactoma" se refiere al conjunto total de interacciones moleculares en un sistema biológico en estudio *(por ejemplo,* células). Se puede exhibir como una gráfica dirigida. Las interacciones moleculares pueden producirse entre moléculas que pertenecen a diferentes familias bioquímicas (proteínas, ácidos nucleicos, lípidos, carbohidratos, *etc.)* y también dentro de una familia dada. Cuando se habla en términos de proteómica, el interactoma se refiere a red de interacción proteína-proteína(PPI) o red de interacción proteica (PIN). Otro tipo ampliamente estudiado de interactoma es el interactoma proteína-ADN (red formada por factores de transcripción (y ADN o proteínas reguladoras de cromatina) y sus genes objetivo.

El término "rendimiento celular" incluye una colección de parámetros, preferiblemente parámetros medibles, relacionados con el estados celular, incluidos (a modo no taxativo): nivel de transcripción de uno o más genes *(por ejemplo,* medible por RT-PCR, qPCR, micromatriz,

etc.), nivel de expresión de una o más proteínas *(por ejemplo,* medible por espectrometría de masas o transferencia Western), actividad absoluta *(por ejemplo,* medible como tasas de conversión de sustrato) o actividad relativa *(por ejemplo,* medible como un valor % comparado con la actividad máxima) de una o más enzimas o proteínas, nivel de uno o más metabolitos o intermediarios, nivel

de fosforilación oxidativa *(por ejemplo,* medible por la Tasa de consumo de oxígeno u OCR), nivel de glucolisis *(por ejemplo,* medible por tasa de acidificación extracelular o ECAR), nivel de unión o interacción ligando-objetivo, actividad de moléculas secretadas extracelularmente, *etc.* El rendimiento celular puede incluir datos para una cantidad predeterminada de genes o proteínas objetivo, *etc.*, o puede incluir una evaluación global para todos los genes o proteínas detectables. Por ejemplo, se puede usar espectrometría de masas para identificar y/o cuantificar todas las proteínas detectables expresadas en una muestra o población celular dada, sin conocimiento previo sobre si alguna proteína específica puede expresarse en la muestra o población celular.

Tal como se usa en la presente, un "sistema celular" incluye una población de células homogéneas o heterogéneas. El células dentro del sistema puede crecer *in vivo,* en el ambiente natural o fisiológico, o pueden crecer *in vitro* en, por ejemplo, ambientes de cultivo tisular controlados. Las células dentro del sistema pueden ser relativamente homogéneas *(por ejemplo,* no menos de 70%, 80%, 90%, 95%, 99%, 99,5%, 99,9% homogéneas), o pueden contener dos o más tipos celulares, tales como los tipos celulares que generalmente se encuentran creciendo en gran proximidad *in vivo,* o tipos celulares que pueden interactuar entre sí *in vivo* a través de, *por ejemplo,* paracrina u otra comunicación intercelular de larga distancia. Las células dentro

del sistema celular pueden derivar de líneas celulares establecidas, incluidas líneas celulares de cáncer, líneas celulares inmortales o líneas celulares normales, o pueden ser células primarias o células recientemente aisladas de tejidos u órganos vivos.

Las células en el sistema celular típicamente están en contacto con un "ambiente celular" que puede proporcionar nutrientes, gases (oxígeno o CO_2, *etc.)*, químicos o estimulantes proteináceos / no proteináceos que pueden definir las condiciones que afectan el comportamiento celular. El ambiente celular puede ser medios químicos con componentes químicos definidos y/o extractos de tejido o componentes de suero menos definidos, y pueden incluir un pH, CONTENIDO DE CO_2, presión y temperatura definidos en los que crecen las células. Alternativamente, el ambiente celular puede ser el ambiente natural o fisiológico que se encuentra *in vivo* para el sistema celular específico.

En ciertas realizaciones, un ambiente celular comprende condiciones que simulan un aspecto de un sistema o proceso biológico, por ejemplo, simular un estado, proceso o ambiente de enfermedad. Dichas condiciones de cultivo incluyen, por ejemplo, hiperglucemia, hipoxia o condiciones ricas en ácido láctico. Varias condiciones diferentes se describen en la presente.

En ciertas realizaciones, un ambiente celular para un sistema celular específico también incluye ciertas características de superficie celular del sistema celular, tal como los tipos de

receptores o ligandos en la superficie celular y sus respectivas actividades, la estructura de carbohidrato o moléculas lipídicas, polaridad de o fluidez de membrana, estado de aglomeración de ciertas proteínas de membrana, *etc.* Estas características de superficie celular pueden afectar la función de las células cercanas, tales como células que pertenecen a un sistema celular diferente. En ciertas realizaciones diferentes, sin embargo, el ambiente celular de un sistema celular no incluye características de superficie celular del sistema celular.

El ambiente celular puede alterarse para que se vuelva un "ambiente celular modificado". Las alteraciones pueden incluir cambios *(por ejemplo,* aumento o disminución) en cualquier uno o más componentes encontrados en el ambiente celular, incluida la adición de uno o más "componentes de estímulo externo" al ambiente celular. La perturbación o componente de estímulo externo puede ser endógeno al ambiente celular *(por ejemplo,* el ambiente celular contiene algunos niveles del estimulante, y más del mismo se agrega para aumentar su nivel), o puede ser exógeno al ambiente celular *(por ejemplo,* el estimulante está ausente en gran medida del ambiente celular antes de la alteración). El ambiente celular puede alterarse adicionalmente mediante cambios secundarios resultantes de la adición del componente de estímulo externo, ya que el componente de estímulo externo puede cambiar el rendimiento celular del sistema celular, incluidas moléculas secretadas en el ambiente celular por el sistema celular.

Tal como se usa en la presente, "componente de estímulo externo", también denominado en la presente "perturbación ambiental", incluye cualquier estímulo físico y/o químico externo que puede afectar la función celular. Este puede incluir cualquier molécula orgánica o inorgánica, grande o pequeña, químicos naturales o sintéticos, cambio de temperatura, cambio de pH, radiación, luz (UVA, UVB *etc.)*, microondas, ondas de sonido, corriente eléctrica, campos magnéticos modulados o no modulados, *etc.*

La expresión "**M**olécula **I**ntracelular **M**ultidimensional (**MIM**)", es una versión aislada o una versión producida sintéticamente de una molécula endógena que se produce naturalmente en el cuerpo y/o que está presente en al menos una célula de un humano. Una MIM eses capaz de ingresar a un célula y la entrada en la célula incluye una entrada completa o parcial en la célula siempre que la porción biológicamente activa de la molécula ingrese completamente en la célula. Las MIM son capaces de inducir una transducción de señal y/o mecanismo de expresión génica dentro de una célula. Las MIM son multidimensionales porque las moléculas tiene un agente terapéutico y un portador, por ejemplo, administración de fármaco, efecto. Las MIM también son multidimensionales debido a que las moléculas actúan de una forma en un estado de enfermedad y de una forma diferente en un estado normal. Por ejemplo, en el caso de la CoQ-10, la administración de la CoQ-10 a una célula de melanoma en presencia de VEGF conduce a un nivel disminuido de Bcl2 que, a su vez, conduce a un

potencial encógenos disminuido para la célula de melanoma. En cambio, en un fibroblasto normal, la coadministración de CoQ-10 y VEFG no tiene ningún efecto en los niveles de Bcl2.

En una realización, una MIM también es un desplazador epi. En otra realización, una MIM no es un desplazador epi. En otra realización, una MIM se caracteriza por una o más de las funciones precedentes. En otra realización, una MIM se caracteriza por dos o más de las funciones precedentes. En una realización adicional, una MIM se caracteriza por tres o más de las funciones precedentes. En otra realización, una MIM se caracteriza por todas las funciones precedentes. El experto en la técnica apreciará que una MIM de la invención también pretende abarcar una mezcla de dos o más moléculas endógenas, en donde la mezcla se caracteriza por una o más de las funciones precedentes. Las moléculas endógenas en la mezcla están presentes en una relación tal que la mezcla funciona como una MIM.

Las MIM pueden ser moléculas basadas en lípidos o no basadas en lípidos. Ejemplos de MIM incluyen, a modo no taxativo, CoQ10, acetil Co-A, palmitil Co-A, L-carnitina, aminoácidos tales como, por ejemplo, tirosina, fenilalanina y cisteína. En una realización, la MIM es una molécula pequeña. En una realización de la invención, la MIM no es CoQ10. Un experto en la técnica puede identificar rutinariamente las MIM usando cualquiera de los ensayos descritos en detalle en la presente. Las MIM se describen en mayor detalle en US 12/777.902 (US 2011-0110914), cuyo

contenido total se incorpora expresamente a la presente a modo de referencia.

Tal como se usa en la presente, un "desplazador epimetabólico" (desplazador epi) es una molécula que modula el desplazamiento metabólico de un estado saludable (o normal) a un estado de enfermedad y viceversa, manteniendo o restableciendo así la salud celular, de tejido, órgano, sistema y/o huésped en un humano. Los desplazadores epi son capaces de efectuar la normalización en un microambiente tisular. Por ejemplo, un desplazador epi incluye cualquier molécula que es capaz,cuando se agrega o reduce en una célula, de afectar el microambiente *(por ejemplo,* el estado metabólico) de una célula. El experto en la técnica apreciará que un desplazador epi de la invención también pretende abarcar una mezcla de dos o más moléculas, en donde la mezcla se caracteriza por una o más de las funciones precedentes. Las moléculas en la mezcla están presentes en una relación tal que la mezcla funciona como un desplazador epi. Ejemplos de desplazadores epi incluyen, a modo no taxativo, CoQ-10; vitamina D3; componentes de ECM tales como fibronectina; inmunomoduladores, tales como TNFa o cualquier de las interleucinas, por ejemplo, IL-5, IL-12, IL-23; factores angiogénicos; y factores apoptóticos.

En una realización, el desplazador epi también es una MIM. En una realización, el desplazador epi no es CoQ10. Un experto en la técnica puede identificar rutinariamente los desplazadores epi usando cualquiera de los ensayos descritos en detalle en la

presente. Los desplazadores epi se describen en mayor detalle en US 12/777.902 (US 2011-0110914), cuyo contenido total se incorpora expresamente a la presente a modo de referencia.

Otros términos no definidos explícitamente en la presente solicitud tiene el significado que entendería un experto en la técnica.

III. Etapas y componentes ejemplares de la tecnología de plataforma

Solamente a efectos de ilustración, las siguientes etapas de la tecnología de plataforma objeto pueden describirse a continuación en la presente como utilidad ejemplar para integrar datos obtenidos de un modelo de cáncer construido a medida, y para identificar nuevas proteínas / vías que activan la patogénesis del cáncer. Los mapas relacionales resultantes de este análisis proporcionan objetivos de tratamiento del cáncer, así como marcadores de diagnóstico / pronóstico asociados con el cáncer. Sin embargo, la tecnología de plataforma objeto tiene aplicabilidad general para cualquier sistema o proceso biológico, y no se limita a ningún cáncer específico ni otros modelos de enfermedad específicos.

Adicionalmente, aunque la descripción a continuación se presenta en algunas porciones como etapas diferenciadas, es a efectos ilustrativos y de simplificación, y por consiguiente, en realidad, no implica dicho orden y/o demarcación rígida de las etapas. Además, las etapas de la invención pueden llevarse a cabo

por separado, y la invención proporcionada en la presente pretende abarcar cada una de las etapas individuales por separado, así como también combinaciones de una o más (por ejemplo, cualquier una, dos, tres, cuatro, cinco, seis o todas las siete etapas) etapas de la tecnología de plataforma objeto, que pueden llevarse a cabo independientemente de las etapas restantes.

También se pretende que la invención incluya todos los aspectos de la tecnología de plataforma como componentes y realizaciones separados de la invención. Por ejemplo, se pretende que los conjuntos de datos generados sean realizaciones de la invención. Como ejemplos adicionales, las redes de relación causal generadas, redes de relaciones causales de consenso generadas y/o redes de relación causal simuladas generadas, también se pretende que sean realizaciones de la invención. Las relaciones causales identificadas como únicas en el sistema biológico se pretende que sean realizaciones de la invención. Adicionalmente, los modelos construidos a medida para un sistema biológico específico también se pretende que sean realizaciones de la invención. Por ejemplo, los modelos construidos a medida para un estados o proceso de enfermedad, tal como, por ejemplo, modelos celulares de cáncer, obesidad /diabetes/enfermedad cardiovascular o un modelo construido a medida para toxicidad (por ejemplo, cardiotoxicidad) de un fármaco, también se pretende que sean realizaciones de la invención.

82

A. Construcción de modelos a medida

La primera etapa en la tecnología de plataforma es el establecimiento de un modelo para un sistema o proceso biológico. Un ejemplo de un sistema o proceso biológico es el cáncer. Como cualquier otro proceso o sistema biológico complicado, el cáncer es una afección patológica complicada caracterizada por múltiples aspectos únicos. Por ejemplo, debido a su alta tasa de crecimiento, muchas células de cáncer están adaptadas a crecer en condiciones de hipoxia, tiene glucólisis regulada de forma ascendente y vías metabólicas de fosforilación oxidativa reducidas. Como resultado, las células de cáncer pueden reaccionar de forma diferente a una perturbación ambiental, tal como tratamiento mediante un fármaco potencial, en comparación con la reacción de una célula normal en respuesta al mismo tratamiento. Por consiguiente, sería de interés descifrar las respuestas únicas del cáncer al tratamiento con el fármaco en comparación con las respuestas de las células normales. Con este fin, se puede establecer un modelo de cáncer a medida para simular el ambiente de una célula cancerosa, por ejemplo, dentro de un tumor *in vivo,* creando condiciones de cultivo celular muy próximas a las condiciones de una célula cancerosa en un tumor in vivo, o imitar varios aspectos del crecimiento del cáncer, aislando diferentes condiciones de crecimiento de las células cancerosas.

Uno de dicho "ambiente" de cáncer, o condición de crecimiento de estrés, es hipoxia, una condición que se encuentra típicamente dentro de un tumor sólido. La hipoxia puede inducirse en las células usando métodos conocidos en la técnica. Por ejemplo, la hipoxia puede inducirse colocando sistemas celulares en una Modular Incubator Chamber (cámara incubadora modular) (MIC-101, Billups-Rothenberg Inc. Del Mar,CA), que puede inundarse con una mezcla de gas industrial que contiene 5% de CO_2, 2% de O_2 y 93% de nitrógeno. Los efectos pueden medirse después de un período predeterminado, *por ejemplo,* 24 horas después del tratamiento de hipoxia, con o sin componentes de estímulo externo adicionales *(por ejemplo,* CoQ10 a 0, 50 o 100 μM).

Asimismo, el tratamiento con ácido láctico de las células imita un ambiente celular donde la actividad de glucólisis es alta, tal como la que existe en el ambiente del tumor in vivo. El estrés inducido por ácido láctico puede investigarse a una concentración final de ácido láctico de aproximadamente 12,5 mM en un momento predeterminado, *por ejemplo,* a las 24 horas, con o sin componentes de estímulo externo adicionales *(por ejemplo,* CoQ10 a 0, 50 o 100 μM).

La hiperglucemia es normalmente una condición encontrada en la diabetes; sin embargo, en cierta medida, la hiperglucemia también imita un aspecto del crecimiento del cáncer debido a que muchas células cancerosas dependen de la glucosa como su fuente principal de energía. La exposición de células objeto a un

condición típica de hiperglucemia puede incluir agregar 10% de glucosa de grado de cultivo a medios adecuados, de forma que la concentración final de la glucosa en los medios sea de aproximadamente 22 mM.

Las condiciones individuales que reflejan aspectos del crecimiento canceroso pueden investigarse por separado en el modelo de cáncer construido a medida y/o pueden combinarse entre sí. En una realización, las combinaciones de al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50 o más condiciones que reflejan o simulan diferentes aspectos del crecimiento / condiciones del cáncer se investigan en el modelo de cáncer construido a medida. En una realización, las condiciones individuales y, además, las combinaciones de al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50 o más condiciones que reflejan o simulan diferentes aspectos del crecimiento / condiciones del cáncer se investigan en el modelo de cáncer construido a medida. Todos los valores presentados en la lista precedente también pueden ser los límites superiores o inferiores de rangos, que se pretenden que sean parte de esta invención, *por ejemplo,* entre 1 y 5, 1 y 10, 1 y 20, 1 y 30, 2 y 5, 2 y 10, 5 y 10, 1 y 20, 5 y 20, 10 y 20, 10 y 25, 10 y 30 o 10 y 50 condiciones diferentes.

A continuación se indican algunas combinaciones ejemplares de condiciones que pueden usarse para tratar las células. Otras combinaciones pueden formularse fácilmente dependiendo de la

evaluación biológica interrogativa específica que se está llevando a cabo.

1. Solo medios
2. 50 µM de CTL Coenzima Q10 (CoQ10)
3. 100 µM de CTL Coenzima Q10
4. 12,5 mM de ácido láctico
5. 12,5 mM de ácido láctico + 50 µM de CTL Coenzima Q10
6. 12,5 mM de ácido láctico + 100 µM de CTL Coenzima Q10
7. Hipoxia
8. Hipoxia + 50 µM de CTL Coenzima Q10
9. Hipoxia + 100 µM de CTL Coenzima Q10
10. Hipoxia + 12,5 mM de ácido láctico
11. Hipoxia + 12,5 mM de ácido láctico + 50 µM de CTL Coenzima Q10
12. Hipoxia + 12,5 mM de ácido láctico + 100 µM de CTL Coenzima Q10
13. Medios + 22 mM de Glucosa
14. 50 µM de CTL Coenzima Q10 + 22 mM de Glucosa
15. 100 µM de CTL Coenzima Q10 + 22 mM de Glucosa
16. 12,5 mM de ácido láctico + 22 mM de Glucosa
17. 12,5 mM de ácido láctico + 22 mM de Glucosa + 50 µM de CTL Coenzima Q10
18. 12,5 mM de ácido láctico + 22 mM de Glucosa + 100 µM de CTL Coenzima Q10
19. Hipoxia + 22 mM de Glucosa

20. Hipoxia + 22 mM de Glucosa + 50 µM de CTL Coenzima Q10

21. Hipoxia + 22 mM de Glucosa + 100 µM de CTL Coenzima Q10

22. Hipoxia + 12,5 mM de ácido láctico + 22 mM de Glucosa

23. Hipoxia + 12,5 mM de ácido láctico + 22 mM de Glucosa + 50 µM de CTL Coenzima Q10

24. Hipoxia + 12,5 mM de ácido láctico + 22 mM de Glucosa + 100 µM de CTL Coenzima Q10

Como testigo se cultivan una o más líneas celulares normales (por ejemplo, THLE2 y HDFa) en condiciones similares a efectos de identificar proteínas o vías únicas del cáncer (ver más adelante). El testigo puede ser el modelo celular de comparación descrito anteriormente.

Se pueden incluir múltiples células cancerosas de igual o diferente origen (por ejemplo las líneas cancerosas PaCa2, HepG2, PC3 y MCF7), en oposición a un único tipo de célula cancerosa, en el modelo de cáncer. En ciertas situaciones, se pueden llevar a cabo experimentos de interacción o ECS entre diferentes células cancerosas (por ejemplo, HepG2 y PaCa2) para varios fines interrelacionados.

En algunas realizaciones que implican interacción, los experimentos que se llevan a cabo en los modelos celulares se diseñan para determinar la modulación del estado o función celular de un sistema o población celular *(por ejemplo,* célula de hepatocarcinoma, HepG2) mediante otro sistema o población celular *(por ejemplo,* cáncer pancreático, PaCa2) en condiciones de

tratamiento definidas *(por ejemplo,* hiperglucemia, hipoxia (isquemia)). De acuerdo con un escenario típico, se pone en contacto un primer sistema / población celular con componentes de estímulo externo, tales como una molécula candidata *(por ejemplo,* una molécula de fármaco pequeño, una proteína) o una condición candidata *(por ejemplo,* hipoxia, ambiente con alto contenido de glucosa). En respuesta, el primer sistema/población celular cambia su transcriptoma, proteoma, metaboloma y/o interactoma, conduciendo a cambios que pueden detectarse fácilmente dentro y fuera de la célula. Por ejemplo, se pueden medir cambios en el transcriptoma mediante el nivel de transcripción de una pluralidad de ARNm objetivo; los cambios en el proteoma pueden medirse mediante el nivel de expresión de una pluralidad de proteínas objetivo; y los cambios en el metaboloma pueden medirse mediante el nivel de una pluralidad de metabolitos objetivo mediante ensayos diseñados específicamente para metabolitos dados. Alternativamente, los cambios mencionados anteriormente en el metaboloma y/o proteoma, al menos con respecto a ciertos metabolitos o proteínas secretados, también pueden medirse por sus efectos en el segundo sistema / población celular, incluida la modulación del transcriptoma, proteoma, metaboloma e interactoma del segundo sistema / población celular. Por lo tanto, los experimentos pueden usarse para identificar los efectos de la molécula(s) de interés secretada por el primer sistema / población celular en un segundo sistema / población

celular en diferentes condiciones de tratamiento. Los experimentos también pueden usarse para identificar cualquier proteína modulada como resultado de la señalización del primer sistema celular (en respuesta al tratamiento con el componente de estímulo externo) en otro sistema celular, mediante, por ejemplo, detección diferencial de datos proteómicos. El mismo escenario experimental también puede adaptarse para un escenario inverso, de forma que los efectos recíprocos entre los dos sistemas celulares también pueden evaluarse. En general, para este tipo de experimento, la elección de los pares de líneas celulares se basa en gran medida en los factores tales como el origen, estado de enfermedad y función celular.

Aunque típicamente participan sistemas de dos células en este tipo de escenario experimental, también se pueden diseñar experimentos similares para más de dos sistemas celulares, por ejemplo, inmovilizando cada sistema celular diferenciado en un soporte sólido separado.

Una vez construido el modelo a medida, se pueden aplicar una o más "perturbaciones" al sistema, tal como una variación genética de paciente en paciente, o con / sin tratamiento mediante ciertos fármacos o profármacos. *Ver* Figura 15D. Los efectos de dichas perturbaciones en elsistema, incluido el efecto en células cancerosas relacionadas con la enfermedad y células testigo normales relacionadas con la enfermedad, pueden medirse usando

varios medios conocidos en la técnica o patentados, tales como se describen en la sección III.B, más adelante.

En un experimento ejemplar, las líneas de cáncer PaCa2, HepG2, PC3 y MCF7 y las líneas celulares normales THLE2 y HDFa, se acondicionan en cada una de hiperglucemia, hipoxia y condiciones ricas en ácido láctico, así como en todas las combinaciones de dos o tres de estas condiciones, y además, con o sin una perturbación ambiental, específicamente tratamiento con Coenzima Q10.

El modelo celular construido a medida puede establecerse y usarse en las etapas de la tecnología de plataforma de la invención para identificar finalmente una relación causal única en el sistema biológico, llevando a cabo las etapas descritas en la presente. El experto en la técnica entenderá, sin embargo, que un modelo celular construido a medida que se usa para generar una red de relaciones causales de consenso de "primera generación" inicial para un proceso biológico puede evolucionar continuamente p expandirse con el tiempo, por ejemplo, mediante la introducción de líneas celulares adicionales de cáncer o normales y/o condiciones de cáncer adicionales. Se pueden recoger datos adicionales del modelo celular evolucionado, es decir, datos de la nueva porción(es) agregada del modelo celular. Los nuevos datos recogidos de un modelo celular expandido o evolucionado, es decir, de la nueva porción(es) agregada del modelo celular, a continuación pueden introducirse a los conjuntos de datos usados

previamente para generar la red de relaciones causales de consenso de "primera generación" a efectos de generar una red de relaciones causales de consenso de "segunda generación" más sólida. Las nuevas relaciones causales únicas para el sistema biológico a continuación pueden identificarse a partir de la red de relaciones causales de consenso de "segunda generación". De esta forma, la evolución del modelo celular proporciona una evolución de las redes de relaciones causales de consenso, proporcionando así perspectivas nuevas y/o más confiables con respecto a los moduladores del sistema biológico.

Se describen en mayor detalle ejemplos adicionales de modelos celulares construidos a medida.

B. Recolección de datos

En general, se pueden recoger dos tipos de datos de cualquier sistema de modelo construido a medida. Un tipo de datos (por ejemplo, el primer conjunto de datos, el tercer conjunto de datos) generalmente se refiere al nivel de ciertas macromoléculas, tales como ADN, ARN, proteína, lípido, etc. Un conjunto de datos ejemplar en esta categoría son los datos proteómicos (por ejemplo, datos cualitativos y cuantitativosrelacionados con la expresión de todas o sustancialmente todas las proteínas medibles de una muestra). El otro tipo de datos generalmente son datos funcionales (por ejemplo, el segundo conjunto de datos, el cuarto conjunto de

datos) que reflejan los cambios fenotípicos resultantes de los cambios en el primer tipo de datos.

Con respecto al primer tipo de datos, en algunas realizaciones ejemplares, se llevan a cabo reacción en cadena de la polimerasa cuantitativa (qPCR) y análisis proteómico para crear perfiles de cambio en ARNm celular y expresión proteica mediante reacción en cadena de la polimerasa cuantitativa (qPCR) y análisis proteómico. El ARN total puede aislarse usando un kit de aislamiento de ARN disponible en el mercado. Luego de la síntesis del ADNc, se pueden emplear matrices de qPCR disponibles en el mercado (por ejemplo, las de SA Biosciences) para área de enfermedad o procesos celulares tales como angiogénesis, apoptosis y diabetes, para crear un perfil de un conjunto predeterminado de genes siguiendo las instrucciones del fabricante. Por ejemplo, puede usarse el sistema de amplificación Biorad cfx-384 para todos los experimentos de creación de perfil transcripcional. Luego de la recolección de datos (Ct), el cambio final con respecto al testigo puede determinarse usando el método δCt tal como se indica en el protocolo del fabricante. El análisis de la muestra proteómica puede llevarse a cabo tal como se describe en las secciones posteriores.

El método objeto puede emplear análisis proteómico de alta productividad a gran escala de cientos de muestras de carácter similar, y proporciona los datos necesarios para identificar los diferenciales de rendimiento celular.

Existen numerosas tecnologías conocidas en la técnica adecuadas a estos efectos. Un técnica ejemplar es el análisis iTRAQ en combinación con espectrometría de masas y se describe brevemente a continuación.

El abordaje proteómico cuantitativo se basa en el etiquetado del isótopo estable con el reactivo 8-plex iTRAQ y 2D-LC MALDI MS/MS para la identificación y cuantificación del péptido. La cuantificación con esta técnica es relativa: se atribuyen a los péptidos u proteínas relaciones de abundancia con respecto a la muestra de referencia. Las muestras de referencia comunes en múltiples experimentos iTRAQ facilitan la comparación de muestras a lo largo de múltiples experimentos iTRAQ.

Por ejemplo, para implementar este esquema de análisis, se pueden combinar seis muestras primarias y dos muestras agrupadas testigo en una mezcla 8-plex iTRAQ de acuerdo con las sugerencias del fabricante. Esta mezcla de ocho muestras a continuación puede fraccionarse mediante cromatografía líquida bidimensional; intercambio catiónico fuerte (SCX) en la primera dimensión y HPLC de fase inversa en la segunda dimensión, a continuación puede someterse a análisis de espectrometría de masas.

En la presente se proporciona una descripción general breve de los procedimientos de laboratorio que pueden emplearse.

Extracción proteica: Las células pueden lisarse con solución amortiguadora de lisis de urea 8 M con inhibidores de proteasa

(Thermo Scientific Halt Protease inhibitor EDTA-free) e incubarse sobre hielo durante 30 minutos con agitación vorticial durante 5 segundos cada 10 minutos. Las lisis puede completarse sometiéndose a ondas ultrasónicas en impulsos de 5 segundos. Los lisados celulares pueden centrifugarse a 14000 x g durante 15 minutos (4 oC) para eliminar los restos celulares. Se puede llevar a cabo el ensayo de Bradford para determinar la concentración proteica. Se pueden reducir 100ug de proteína de cada muestra (10mM Ditiotreitol (DTT), 55 °C, 1 h), alquilarse (25 mM de yodoacetamida, temperatura ambiente, 30 minutos) y digerirse con Tripsina (1:25 p/p, 200 mM de bicarbonato de trietilamonio (TEAB), 37 oC, 16 h).

Preparación de muestra de secretoma: 1) En una realización, las células pueden cultivarse en medio libre de suero: Los medios acondicionados pueden concentrarse mediante un secador por congelación, reducirse (10mM de Ditiotreitol (DTT), 55 °C, 1 h), alquilarse (25 mM de yodoacetamida, a temperatura ambiente, incubarse durante 30 minutos) y luego desalarse mediante precipitación en acetona. Se puede digerir una cantidad igual de proteínas a partir de los medios acondicionados concentrados con Tripsina (1:25 p/p, 200 mM de bicarbonato de trietilamonio (TEAB), 37 oC, 16 h).

En una realización, las células pueden cultivarse en medio que contiene suero: El volumen del medio puede reducirse usando columnas 3k MWCO Vivaspin (GE Healthcare Life Sciences), y luego

reconstituirse con 1xPBS (Invitrogen). La seroalbúmina puede retirarse de todas las muestras usando columna AlbuVoid (Biotech Support Group, LLC) siguiendo las instrucciones del fabricante con las modificaciones de intercambio de solución amortiguadora para optimizar la aplicación del medio acondicionado.

Etiquetado con iTRAQ 8 Plex: Se pueden agrupar alícuotas de cada digesto tríptico en cada conjunto experimental para crear la muestra testigo agrupada. Las alícuotas iguales de cada muestra y la muestra testigo agrupada puede etiquetarse con reactivos iTRAQ 8 Plex, de acuerdo con los protocolos del fabricante (AB Sciex). Las reacciones pueden combinarse, someterse al vacío hasta secarse, volver a suspenderse agregando ácido fórmico al 0,1% y analizarse mediante LC-MS/MS.

2D-Nano LC-MS/MS: Todas las mezclas peptídicas etiquetadas pueden separarse mediante 2D-nanoLC en línea y analizarse mediante espectrometría de masas en tándem con electropulverización. Los experimentospueden llevarse a cabo en un sistema Eksigent 2D NanoLC Ultra conectado a un espectrómetro de masas LTQ Orbitrap Velos equipado con una fuente de iones por nanoelectropulverización (Thermo Electron, Bremen, Alemania).

Las mezclas peptídicas pueden inyectarse en una columna SCX de 5 cm (300µm ID, 5µm, columna PolySULFOETHYL Aspartamide de PolyLC, Columbia, MD) con un flujo de 4 µL / min y eluírse en 10 segmentos de elución de intercambio iónico en una columna C18 trap (2,5 cm, 100µm ID, 5µm, 300 Å ProteoPep II de New Objective,

Woburn, MA) y lavarse durante 5 min con H2O/0,1% FA. La separación a continuación puede llevarse a cabo adicionalmente a 300 nL/min usando un gradiente de 2-45% B (H20 /0,1%FA (disolvente A) y ACN /0,1%FA (disolvente B)) durante 120 minutos en una columna de sílice fusionada de 15 cm (75µm ID, 5µm, 300 Å ProteoPep II de New Objective, Woburn, MA).

Se pueden adquirir espectros MS de escaneo completo (m/z 300-2000) en el Orbitrap con resolución de 30000. Los iones más intensos (hasta 10) pueden aislarse secuencialmente para fragmentación usando Disociación C-trap de gran energía (HCD) y excluirse dinámicamente durante 30 segundos. La HCD puede llevarse a cabo con un ancho de aislamiento de 1,2 Da. Los iones de fragmento resultantes pueden escanearse en el orbitrap con la resolución de 7500. El LTQ Orbitrap Velos puede controlarse mediante Xcalibur 2.1 con base 1.0.1.

Identificación y cuantificación de péptidos/proteínas: Los péptidos y proteínas pueden identificarse mediante búsqueda de base de datos automatizada usando el programa informático Proteome Discoverer (Thermo Electron) con el buscador Mascot en la base de datos SwissProt. Los parámetros de búsqueda pueden incluir 10 ppm para tolerancia de MS, 0,02 Da para tolerancia MS2 y digestión total de tripsina permitiendo hasta 2 escisiones perdidas. La carbamidometilación (C) puede fijarse como la modificación fija. La oxidación (M), TMT6 y desamidación (NQ) pueden fijarse como modificaciones dinámicas. Las

identificaciones de péptidos y proteína pueden filtrarse con Mascot Significant Threshold ($p<0,05$). Se puede permitir a los filtros un nivel de confianza de 99% de identificación proteica (1% FDA).

El programa informático Proteome Discoverer puede aplicar factores de corrección en los iones reporteros y pueden rechazar todos los valores de análisis cuantitativo si no están presentes todos los canales de análisis cuantitativo. El análisis cuantitativo de proteína relativa puede lograrse mediante normalización en la intensidad media.

Con respecto al segundo tipo de datos, en algunas realizaciones ejemplares, la creación de perfiles bioenergéticos de modelos de cáncer y normales pueden emplear el analizador Seahorse ™XF24 para permitir la comprensión de los componentes de glucólisis y fosforilación oxidativa.

Específicamente, las células pueden colocarse en placas de cultivo Seahorse a densidades óptimas. Estas células pueden colocarse en placas en 100 µl de medios o tratamiento y dejarse a 37°C en una incubadora con 5% de CO_2. Dos horas después, cuando las células están adheridas a las placas de 24 pocillos, puede agregarse 150 µl adicionales de medios o solución de tratamiento y las placas pueden dejarse en la incubadora de cultivo durante toda la noche. Este procedimiento de sembrado de dos etapas permite una distribución uniforme de células en la placa de cultivo. Los cartuchos Seahorse que contienen el sensor de

oxígeno y pH pueden hidratarse durante toda la noche en el fluido de calibración en una incubadora sin CO_2 a 37°C. Típicamente se cargan tres fármacos mitocondriales en los tres puertos en el cartucho. Se pueden cargar oligomicina, un inhibidor de complejo III, FCCP, un desacoplador y Rotenone, un inhibidor de complejo I en los puertos A, B y C respectivamente del cartucho. Todos los fármacos concentrados pueden prepararse a una concentración de 10x en medios de DMEM sin solución amortiguadora. Los cartuchos pueden incubarse primero con los compuestos mitocondriales en una incubadora sin CO_2 durante aproximadamente 15 minutos antes del ensayo. Las placas de cultivo Seahorse pueden lavarse en medios basados en DMEM sin solución amortiguadora que contienen glucosa a una concentración que se encuentra en los medios de crecimiento normales. Las células pueden colocarse en capas con 630 ul de los medios sin solución amortiguadora y pueden equilibrarse en una incubadora sin CO_2 antes de colocarse en el instrumento Seahorse con un cartucho precalibrado. El instrumento puede pasarse durante tres-cuatro bucles con una mezcla, esperar y medir el ciclo para obtener una cifra inicial, antes de iniciar la inyección de fármacos a través del puerto. Pueden hacerse dos bucles antes que se introduzca el siguiente fármaco.

La OCR (tasa de consumo de oxígeno) y la ECAR (tasa de acidificación extracelular) pueden registrarse mediante los electrodos en una cámara de 7 **µl** y pueden crearse con el cartucho empujando contra la placa de cultivo Seahorse.

C. Integración de datos y generación de modelos *in silico*

Una vez que se hayan obtenido los conjuntos de datos relevantes, se puede llevar a cabo la integración de los conjuntos de datos y la generación de modelos estadísticos implementados por computadora usando un sistema o plataforma informático basado en AI (por ejemplo, la plataforma REFS™). Por ejemplo, un sistema basado a AI ejemplar pueden producir redes basadas en simulación de asociaciones proteicas como activadores clave de puntos extremos metabólicos (ECAR/OCR). *Ver* Figura 15. Se pueden encontrar algunos antecedentes relacionados con el sistema REFS ™ en Xing et al., "Causal Modeling Using Network Ensemble Simulations of Genetic and Gene Expression Data Predicts Genes Involved in Rheumatoid Arthritis," *PLoS Computational Biology,* tomo 7,

número 3, 1-19 (March 2011) (e100105) y la Patente de los Estados Unidos 7.512.497 de Periwal, cuyo contenido total de cada una se incorpora expresamente a la presente a modo de referencia en su totalidad. Esencialmente, tal como se describió anteriormente, el sistema REFS ™es un sistema basado en AI que emplea algoritmos matemáticos para establecer relaciones causales entre las variables de entrada (por ejemplo, niveles de expresión proteica, niveles de expresión de ARNm y los datos funcionales correspondientes, tales como los valores OCR / ECAR medidos en las placas de cultivo Seahorse). Este proceso se basa solamente

en los datos de entrada, sin tomar en cuenta el conocimiento anterior existente sobre cualquier relación biológica potencial, establecida y/o verificada.

Específicamente, una ventaja significativa de la plataforma de la invención es que el sistema basado en AI se basa en los conjuntos de datos obtenidos del modelo celular, sin recurrir ni tomar en cuenta ningún conocimiento existente en relación con el proceso biológico. Además, preferiblemente, ningún punto de datos se corta estadísticamente o artificialmente y, en cambio, todos los datos obtenidos se ingresan en el sistema AI para determinar las asociaciones proteicas. Por consiguiente, los modelos estadísticos resultantes generados a partir de la plataforma son imparciales, ya que no toman en cuenta ninguna relación biológica conocida.

Específicamente, los datos proteómicos y de ECAR/OCR pueden ingresarse en el sistema de información en base a AI, que construye modelos estadísticos basados en asociaciones de datos, tal como se describió anteriormente. Las redes basadas en simulación de asociaciones proteicas a continuación se derivan para cada enfermedad con respecto al escenario normal, incluidos los tratamientos y condiciones usando los siguientes métodos.

Una descripción detallada de un proceso ejemplar para construir las redes generadas (por ejemplo, optimizadas o evolucionadas) se muestra más adelante con respecto a la Figura 16. Tal como se describió anteriormente, los datos proteómicos y datos de célula

funcionales se ingresan en el sistema basado en AI (etapa 210). Los datos de ingreso, que pueden ser datos en bruto o datos mínimamente procesados, se preprocesan, que pueden incluir la normalización (por ejemplo, usando una función cuantil o estándares internos) (etapa 212). El preprocesamiento también puede incluir ingresar valores de datos faltantes (por ejemplo, usando el algoritmo de los K vecinos más cercanos (K-NN)) (etapa 212).

Los datos preprocesados se usan para construir una biblioteca de fragmentos de red (etapa 214). Los fragmentos de red definen relaciones cuantitativas, continuas entre todos los pequeños conjuntos posibles (por ejemplo, conjuntos de 2-3 miembros o conjuntos de 2-4 miembros) de variables medidas (datos de ingreso). Las relaciones entre las variables en un fragmento pueden ser lineales,

logísticas, multinomiales, dominantes o recesivas homocigóticas, etc. A la relación en cada fragmento se atribuye un puntaje probabilísitico de Bayesian que refleja cuan probable es la relación candidata dados los datos de entrada y también penaliza la relación por su complejidad matemática. Al puntear todos las relaciones posibles en pares o en tríos (y en algunas realizaciones también relaciones de cuarteto) inferidas a partir de los datos de entrada, se pueden identificar los fragmentos más probables en la biblioteca (los fragmentos probable). Los parámetros cuantitativos de la relación también se computan en

base a los datos de entrada y se almacenan para cada fragmento. Se pueden usar varios tipos de modelos en la enumeración de fragmentos incluidos, a modo no taxativo, regresión lineal, regresión logística, (Análisis de Varianza) modelos ANOVA, (Análisis de Covarianza) modelos ANCOVA, modelos de regresión no lineal/polinomial e incluso regresión no paramétrica. Las presunciones anteriores en los parámetros de modelo pueden tomar distribuciones Gull o penalidades de Bayesian Information Criterion (BIC) relacionadas con la cantidad de parámetros usados en el modelo. En un proceso de inferencia de red, cada red en un grupo de redes de prueba iniciales se construye a partir de un subconjunto de fragmentos en la biblioteca de fragmentos. Cada red de prueba inicial en el grupo de redes de prueba iniciales se construye con un subconjunto diferente de los fragmentos de la biblioteca de fragmentos (etapa 216).

Una descripción global de las representaciones matemáticas subyacentes de las redes de Bayesian y los fragmentos de red, que se basan en Xing et al., "Causal Modeling Using Network Ensemble Simulations of Genetic and Gene Expression Data Predicts Genes Involved in Rheumatoid Arthritis," *PLoS Computational Biology,* tomo 7, número 3, 1-19 (Marzo, 2011) (e100105), se presenta más adelante.

Un sistema multivariante con variables aleatorias $X = X_1,\ldots,X_n$ puede caracterizare mediante una función de

distribución de probabilidad multivariante $P(X_1,\ldots,X_n;\Theta)$, que incluye una gran cantidad de parámetros Θ. La función de distribución de probabilidad multivariante puede factorizarse y representarse mediante un producto de distribuciones de probabilidad condicional local:

$$P(X_1,\ldots,X_n;\Theta)=\prod_{i-1}^{n}P_i\left(X_i\middle|Y_{j1},\ldots,Y_{jK_i};\Theta_i\right),$$

en las que cada variable X_i es independiente de sus variables no descendentes dadas sus variables genitores K_i, que son $Y_{j1},\ldots,Y_{jK_i}$. Después de la factorización, cada distribución de probabilidad local tiene sus propios parámetros Θ_i.

La función de distribución de probabilidad multivariante puede factorizarse de diferentes formas siendo cada factorización específica y parámetros correspondientes un modelo probabilístico diferenciado. Cada factorización específica (modela) puede representarse mediante una Directed Acrylic Graph (gráfica de acrílico dirigida) (DAC) que tiene un vértice para cada variable X_i y los bordes dirigidos entre vértices representan dependencias entre variables en las distribuciones condicionales locales $P_i\left(X_i\middle|Y_{j1},\ldots,Y_{jK_i}\right)$. Las subgráficas de una DAG,

incluyendo cada una un vértice y bordes dirigidos asociados son fragmentos de red.

Un modelo se evoluciona u optimiza determinando la factorización más probable y los parámetros más probables dados los datos de entrada. Esto puede describirse como "conociendo una red de Bayesian" o, en otras palabras, dado un conjunto de entrenamiento de datos de entrada, encontrar una red que coincide mejor con los datos de entrada. Esto se logra usando una función de puntaje que evalúa cada red con respecto a los datos de entrada.

Una red de Bayesian se usa para determinar la probabilidad de una factorización dados los datos de entrada. La Ley de Bayes establece que la probabilidad posterior, $P(D|\mathrm{M})$, de un modelo M, dados los datos D es proporcional al producto del producto de la probabilidad posterior de los datos dadas las presunciones del modelo, $P(D|\mathrm{M})$, multiplicados por la probabilidad previa del modelo, $P(M)$, suponiendo que las probabilidad de los datos, $P(D)$, es constante en los modelos. Esto se expresa en la siguiente ecuación:

$$P(M|\mathrm{D}) = \frac{P(D|\mathrm{M}) * P(M)}{P(D)}$$

La probabilidad posterior de los datos suponiendo el modelo es la integral de la probabilidad de datos con respecto a la distribución previa de los parámetros:

$$P(D|\mathrm{M}) = \int P(D|\mathrm{M}(\Theta))P(\Theta|M)d\Theta$$

Suponiendo que todos los modelos son igualmente probables (es decir, que $P(M)$ es una constante), la probabilidad posterior del modelo M dados los datos D puede factorizarse en el producto de las integrales con respecto a los parámetros para cada fragmento de red local M_i de la siguiente forma:

$$P(M|D) = \prod_{i=1}^{n} \int P_i\left(X_i \middle| Y_{j1},\ldots,Y_{jK_i};\Theta_i\right)$$

Cabe señalar que en la ecuación anterior, un término constante líder ha sido omitido. En algunas realizaciones, el Bayesian Information Criterion (BIC), que toma un logaritmo negativo de la probabilidad posterior del modelo $P(D|\mathrm{M})$ puede usarse para "Puntuar" cada modelo de la siguiente manera:

$$S_{tot}(M) = -\log P(M|D) = \sum_{i=1}^{n} S(M_i),$$

donde el puntaje total S_{tot} para un modelo M es una suma de los puntajes locales S_i para cada fragmento de red local. El BIC además proporciona una expresión para determinar un puntaje para cada fragmento de red individual:

$$S(M_i) \approx S_{BIC}(M_i) = S_{MLE}(M_i) + \frac{\kappa(M_i)}{2} \log N$$

donde $\kappa(M_i)$ es la cantidad de parámetros de ajuste en el modelo M_i y N es la cantidad de muestras (puntos de datos). $S_{MLE}(M_i)$ es el logaritmo negativo de la función de probabilidad para un fragmento de red, que puede calcularse a partir de las relaciones funcional usadas para cada fragmento de red. Para un puntaje BIC, cuanto más bajo es el puntaje, más probable es que un modelo se ajuste a los datos de entrada.

El grupo de redes de prueba se optimiza globalmente, que puede describirse como optimizar o evolucionar las redes (etapa 218). Por ejemplo, las redes de prueba puede evolucionarse y optimizarse de acuerdo con un algoritmo de muestreo Metropolis Monte Carlo. El apareamiento simulado puede usarse para optimizar o evolucionar cada red de prueba en el conjunto a través de transformaciones locales. En un ejemplo de procesos de apareamientos simulados, cada red de prueba se cambia mediante la adición de un fragmento de red de la biblioteca, mediante la supresión de un fragmento de red de la red de prueba, mediante la sustitución de un fragmento de red o mediante un cambio por otro medio de la topología de la red, y a continuación se calcula un nuevo puntaje para la red. En términos generales, si el puntaje mejora, el cambio se mantiene y si el puntaje empeora el cambio

se rechaza. Una parámetro de "temperatura" permite un cambio local parcial que empeora el puntaje a mantener, lo cual ayuda al proceso de optimización para evitar algunas mínimas locales. El parámetro de "temperatura" se disminuye con el tiempo para permitir que el proceso de optimización/evolución converja.

Todo o parte del proceso de inferencia de la red puede llevarse a cabo en paralelo para las diferentes redes de prueba. Cada red puede optimizarse en paralelo en un procesador separado y/o en un dispositivo informático separado. En algunas realizaciones, el proceso de optimización puede llevarse a cabo en una supercomputadora que incorpora cientos de miles de procesadores que operan en paralelo. La información puede compartirse entre los procesos de optimización llevados a cabo en procesadores paralelos.

El proceso de optimización puede incluir un filtro de red que baja cualquier red del grupo que no alcanza un umbral estándar para el puntaje general. La red bajada puede reemplazarse por una nueva red inicial. Adicionalmente, cualquier red que no está "libre de escala" puede bajarse del grupo. Después que el grupo de redes ha sido optimizado o evolucionado, el resultado puede denominarse un conjunto de redes de modelo celular generadas, que pueden denominarse colectivamente como la red de consenso generada.

D. <u>Simulación para extraer información de relación cuantitativa y para predicción</u>

Se puede usar simulación para extraer información de parámetro cuantitativa con respecto a cada relación en las redes de modelo celular generadas (etapa 220). Por ejemplo, la simulación para la extracción de información cuantitativa puede implicar perturbar (aumentando o disminuyendo) cada nodo en la red 10 veces y calcular las distribuciones posteriores para los otros nodos (por ejemplo, proteínas) en los modelos. Los puntos extremos se comparan mediante una prueba de t con el supuesto de 100 muestras por grupo y el corte de significación de 0,01. La estadística de la prueba de t es la media de 100 pruebas de t. A través del uso de esta técnica de simulación, se genera un AUC (área debajo de la curva) que representa la intensidad de la predicción y el cambio en veces que representa la magnitud *in silico* de un nodo que activa un punto extremo para cada relación en el grupo de redes.

Se puede emplear un módulo de cuantificación de relación de un sistema informático local para dirigir el sistema basado en AI para llevar a cabo las perturbaciones y para extraer la información de AUC y la información de las veces. La información cuantitativa extraída puede incluir el cambio en veces y la AUC para cada borde que conecta un nodo genitor a un nodo descendiente. En algunas realizaciones, se puede usar un programa R construido a medida para extraer la información cuantitativa.

En algunas realizaciones, el grupo de redes de modelo celular generadas puede usarse a través de la simulación para predecir

respuestas a cambios en las condiciones, que luego pueden verificarse a través de experimentos basados en animales en laboratorios húmedos.

El rendimiento del sistema basado en AI puede ser parámetros de relación cuantitativa y/u otras predicciones de simulación (222).

E. Generación de redes diferenciales (delta)

Un módulo de creación de red diferencial puede usarse para generar redes diferenciales (delta) entre las redes de modelo celular generadas y las redes de modelo celular de comparación generadas. Tal como se describió anteriormente, la red diferencial compara todos los parámetros cuantitativos de las relaciones en las redes de modelo celular generadas y la redes de modelo celular de comparación generadas. Los parámetros cuantitativos para cada relación en la red diferencial se basan en la comparación. En algunas realizaciones, un diferencial puede llevarse a cabo entre varias redes diferenciales, que pueden denominarse red delta-delta. Un ejemplo de una red delta-delta se describe más adelante y con respecto a la Figura 26 en la sección de Ejemplos. El módulo de creación de red diferencial puede ser un programa o guión escrito in PERL.

F. Visualización de redes

Los valores de relación para el grupo de redes y para las redes diferenciales pueden visualizarse usando un programa de visualización de red (por ejemplo, plataforma de fuente abierta para análisis de redes complejo y visualización Cytoscape del

Consorcio Cytoscape). En las representaciones visuales de las redes, el espesor de cada borde (por ejemplo, cada línea que conecta a las proteínas) representa la intensidad del cambio en veces. Los bordes también son direccionales indicando la causalidad, y cada borde tiene un nivel de confianza de predicción asociado.

G. **Sistema informático ejemplar**

La Figura 17 representa esquemáticamente un sistema/ambiente informático ejemplar que puede emplearse en algunas realizaciones para comunicarse con el sistema informático basado en AI, para generar redes diferencial, para visualizar redes, para guardar y almacenar datos y/o para interactuar con el usuario. Tal como se explicó anteriormente, los cálculos para un sistema informático basado en AI pueden llevarse a cabo en unasupercomputadora aparte con cientos o miles de procesadores paralelos que interactúan, directa o indirectamente, con el sistema informático ejemplar. El ambiente incluye un dispositivo informático 100 con dispositivos periféricos asociados. El dispositivo informático 100 puede programarse para implementar un código ejecutable 150 para llevar a cabo varios métodos o porciones de métodos descritos en la presente. El dispositivo informático 100 incluye un dispositivo de almacenamiento 116, tal como un disco duro, CD-ROM u otros medios legibles por computadora no transitorios. El dispositivo de almacenamiento 116 puede almacenar un sistema operativo 118 y otros programas informáticos relacionados. El dispositivo

informático 100 adicionalmente puede incluir una memoria 106. La memoria 106 puede comprender una memoria de sistema informático o memoria de acceso aleatorio, tal como DRAM, SRAM, EDO RAM, etc. La memoria 106 puede comprender otros tipos de memoria también, o combinaciones de las mismas. El dispositivo informático 100 puede almacenar, en el dispositivo de almacenamiento 116 y/o en la memoria 106, instrucciones para implementar y procesar cada porción del código ejecutable 150.

El código ejecutable 150 puede incluir un código para la comunicación con el sistema informático basado en AI 190, para generar redes diferenciales (por ejemplo, un módulo de creación de red diferencial), para extraer información de relación cuantitativa del sistema informático basado en AI (por ejemplo, un módulo de cuantificación de relación) y para visualizar redes (por ejemplo, Cytoscape).

En algunas realizaciones, el dispositivo informático 100 puede comunicarse directa o indirectamente con el sistema informático basado en AI 190 (por ejemplo, un sistema para ejecutar REFS). Por ejemplo, el dispositivo informático 100 puede comunicarse con el sistema informático basado en AI 190 transfiriendo archivos de datos (por ejemplo, secuencias de datos) al sistema informático basado en AI 190 a través de una red. Adicionalmente, el dispositivo informático 100 puede tener un código ejecutable 150 que proporciona una interfaz e instrucciones al sistema informático basado en AI 190.

En algunas realizaciones, el dispositivo informático 100 puede comunicarse directa o indirectamente con uno o más sistemas experimentales 180 que proporcionan datos para el conjunto de datos de entrada. Los sistemas experimentales 180 para generar datos pueden incluir sistemas para datos proteómicos basados en espectrometría de masas, expresión génica en micromatriz, expresión génica en qPCR, datos metabolómicos basados en espectrometría de masas y datos lipidómicos basados en espectrometría de masas, micromatrices de SNP, un panel de ensayos funcionales y otros plataformas y tecnologías de biología in-vitro.

El dispositivo informático 100 también incluye un procesador 102 y puede incluir uno o más procesador(es) adicional(es) 102', para ejecutar el programa informático almacenado en la memoria 106 y otrosprogramas para controlar el equipo del sistema, dispositivos periféricos y/o equipos periféricos. El procesador 102 y procesador(es) 102' pueden ser cada uno un procesador de un único núcleo o un procesador de múltiples núcleos (104 y 104'). Se puede emplear virtualización en el dispositivo informático 100 de forma que la infraestructura y los recursos en el dispositivo informático puedan compartirse dinámicamente. También se pueden usar procesadores virtuales con código ejecutable 150 y otros programas informáticos en el dispositivo de almacenamiento 116. Se puede proporciona una máquina virtual 114 para manejar un proceso que está activo en múltiples procesadores de forma que

el proceso parezca estar usando solo un recurso informático en lugar de múltiples. También se pueden usar múltiples máquinas virtuales con un procesador.

Un usuario puede interactuar con un dispositivo informático 100 a través de un dispositivo de presentación visual 122, tal como un monitor de computadora, que puede exhibir una interfaz de usuario 124 o cualquier otra interfaz. La interfaz de usuario 124 del dispositivo de presentación 122 puede usarse para exhibir los datos en bruto, representaciones visuales de redes, etc. El dispositivo de presentación visual 122 también puede exhibir otros aspectos o elementos de realizaciones ejemplares (por ejemplo, un ícono para el dispositivo de almacenamiento 116). El dispositivo informático 100 puede incluir otros dispositivo I/O tales como un teclado o una interfaz táctil multipuntos (por ejemplo, una pantalla táctil) 108 y un dispositivo de puntero 110, (por ejemplo, un ratón, ratón de bola y/o panel táctil) para recibir la entrada de un usuario. El teclado 108 y el dispositivo puntero 110 pueden estar conectados al dispositivo de presentación visual 122 y/o el dispositivo informático 100 a través de una conexión cableada y/o inalámbrica.

El dispositivo informático 100 puede incluir una interfaz de red 112 para interactuar con un dispositivo de red 126 a través de una Red de Área Local (LAN, por sus siglas en inglés), una Red de Área Extensa (WAN, por sus siglas en inglés) o Internet a través de una variedad de conexiones que incluyen, a modo no taxativo,

líneas telefónicas estándares, enlaces LAN o WAN (por ejemplo, 802.11, TI, T3, 56kb, X.25), conexiones de banda ancha (por ejemplo, ISDN, Frame Relay, ATM), conexiones inalámbricas, red de controlador de área (CAN) o alguna combinación de cualquiera o todas las mencionadas anteriormente. La interfaz de red 112 puede comprender un adaptador de red incorporado, tarjeta de interfaz de red, tarjeta de red PCMCIA, adaptador de red de bus de tarjeta, adaptador de red inalámbrica, adaptador de red USB, módem o cualquier otro dispositivo adecuado para posibilitar que el dispositivo informático 100 interactúe con cualquier tipo de red capaz de comunicación y que lleve a cabo las operaciones descritas en la presente.

Asimismo, el dispositivo informático 100 puede ser cualquier sistema informático tal como una estación de trabajo, computadora de escritorio, servidor, computadora portátil, computadora de bolsillo u otra forma de dispositivo informático o de telecomunicaciones que es capaz de comunicación y que tiene suficiente poder de procesamiento y capacidad de memoria para llevar a cabo las operaciones descritas en la presente.

El dispositivo informático 100 puede ejecutarse en cualquier sistema operativo 118 tal como cualquiera de las versiones de los sistemas operativos MICROSOFT WINDOWS, las diferentes versiones de los sistemas operativos de Unix y Linux, cualquier versión de MACOS para las computadoras de Macintosh, cualquier sistema operativo incorporado, cualquier sistema operativo en tiempo

real, cualquier sistema operativo de fuente abierta, cualquier sistema operativo patentado, cualquier sistema operativo para dispositivos informáticos móviles o cualquier otro sistema operativo capaz de ejecutarse en el dispositivo informático y llevar a cabo las operaciones descrito en la presente. El sistema operativo puede ejecutarse en modo nativo o modo emulado.

IV. Modelos para un sistema biológico y usos de los mismos

A. Establecer un modelo para un sistema biológico

Prácticamente todos los sistemas o procesos biológicos implican interacciones complicadas entre diferentes tipos de células y/o sistemas de órganos. La perturbación de funciones fundamentales en un tipo de célula u órgano puede conducir a efectos secundarios en otros tipos de células u órganos interactuantes, y dichos cambios posteriores, a su vez, pueden respaldar los cambios iniciales y provocar complicaciones adicionales. Por lo tanto, es beneficioso diseccionar un sistema o proceso biológico dado hasta sus componentes, tal como la interacción entre pares de tipos celulares u órganos, y explorar sistémicamente las interacciones entre estos componentes a efectos de obtener una visión global más completa del sistema o proceso biológico.

Por consiguiente, la presente invención proporciona modelos celulares para sistemas biológicos. Con este fin, los inventores construyeron modelos celulares para varios sistemas biológicos ejemplares que han sido empleados en la tecnología de plataforma

objeto de descubrimiento. Los inventores llevaron a cabo experimentos con los modelos celulares usando la tecnología de plataforma objeto de descubrimiento para generar redes de relaciones causales de consenso, incluidas relaciones causales únicas en el sistema biológico y así identificaron "moduladores" o "activadores" moleculares fundamentales importantes para los sistemas o procesos biológicos específicos.

Una ventaja significativa de la tecnología de plataforma y sus componentes, por ejemplo, los modelos celulares construidos a medida y los conjuntos de datos obtenidos a partir de los modelos celulares, es que la red de relaciones causales de consenso de "primera generación" inicial generada para unsistema o proceso biológico puede evolucionar o expandirse continuamente con el tiempo, por ejemplo, mediante la introducción de líneas/tipos celulares adicionales y/o condiciones adicionales. Se pueden recoger datos adicionales del modelo celular evolucionado, es decir, datos de la nueva porción(es) agregada del modelo celular. Los nuevos datos recogidos de un modelo celular expandido o evolucionado, es decir, de la nueva porción(es) agregada del modelo celular, a continuación pueden introducirse a los conjuntos de datos usados previamente para generar la red de relaciones causales de consenso de "primera generación" a efectos de generar una red de relaciones causales de consenso de "segunda generación" más sólida. Las nuevas relaciones causales únicas para el sistema biológico a continuación pueden identificarse a

partir de la red de relaciones causales de consenso de "segunda generación". De esta forma, la evolución del modelo celular proporciona una evolución de las redes de relaciones causales de consenso, proporcionando así perspectivas nuevas y/o más confiables con respecto a los moduladores del sistema biológico. De esta forma, los modelos celulares, los conjuntos de datos y las redes de relaciones causales generadas a partir de los modelos celulares mediante el uso de los métodos de la tecnología de plataforma pueden evolucionar constantemente y ampliarse a partir de conocimiento previo obtenido a partir de la tecnología de plataforma.

Por consiguiente, la invención proporciona redes de relaciones causales de consenso generadas a partir de los modelos celulares empleados en la tecnología de plataforma. Estas redes de relaciones causales de consenso pueden ser redes de relaciones causales de consenso de primera generación o pueden ser redes de relaciones causales de consenso de generaciones múltiples, por ejemplo, redes de relaciones causales de consenso de 2a, 3a, 4a, 5a, 6a, 7a, 8a, 9a, 10a, 11a, 12a, 13a, 14a, 15a, 16a, 17a, 18a, 19a, 20a o mayor generación. Además, la invención proporciona redes de relaciones causales de consenso generadas a partir de los modelos celulares empleados en la tecnología de plataforma. Estas redes de relaciones causales de consenso simuladas pueden ser redes de relaciones causales de consenso simuladas de primer generación o pueden ser redes de relaciones

causales de consenso simuladas de múltiples generaciones, por ejemplo redes de relaciones causales de consenso simuladas de 2a, 3a, 4a, 5a, 6a, 7a, 8a, 9a, 10a, 11a, 12a, 13a, 14a, 15a, 16a, 17a, 18a, 19a, 20a o mayor generación. La invención además proporciona redes delta y redes delta-delta generadas a partir de cuales red de relaciones causales de consenso de la invención.

Un modelo celular construido a medida para un sistema o proceso biológico comprende una o más células asociadas con el sistema biológico. El modelo para un sistema/proceso biológico puede establecerse para simular un ambiente de sistema biológico, por ejemplo,ambiente de una célula cancerosa *in vivo,* creando condiciones (por ejemplo, condiciones de cultivo celular) que imitan un aspecto característico del sistema o proceso biológico. Se pueden incluir múltiples célula de igual o diferente origen, en oposición a un único tipo de célula, en el modelo celular. En una realización, se incluyen al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 50 o más líneas celulares o tipos celulares diferentes en el modelo celular. En una realización, las células son todas del mismo tipo, por ejemplo, todas células de cáncer de mama o células vegetales, pero son líneas celulares diferentes establecidas, por ejemplo, líneas celulares diferentes establecidas de células de cáncer de mama o células vegetales. Todos los valores presentados en la lista precedente también pueden ser el límite superior o inferior de rangos, que se pretende que sean parte de la presente

invención, *por ejemplo,* entre 1 y 5, 1 y 10, 2 y 5 o 5 y 15 líneas celulares o tipos celulares diferentes.

Ejemplos de tipos celulares que pueden incluirse en los modelos celulares de la invención incluyen, a modo no taxativo, células humanas, células de animales, células de mamíferos, células vegetales, levaduras, bacterias u hongos. En una realización, las células del modelo celular pueden incluir células enfermas, tales como células de cáncer o células infectadas bacterianamente o viralmente. En una realización, las células del modelo celular pueden incluir células asociadas con la enfermedad, tales como células implicadas en la diabetes, obesidad o estado de enfermedad cardiovascular, por ejemplo, células de músculo liso de la aorta o hepatocitos. El experto en la técnica reconocería aquellas células que están implicadas o asociadas con un estado/proceso biológico específico, por ejemplo, estado/proceso de enfermedad y cualquiera de dichas células puede incluirse en un modelo celular de la invención.

Los modelos celulares de la invención pueden incluir una o más "células testigo". En una realización, una célula testigo puede ser una célula no tratada o no perturbada. En otra realización, una "célula testigo" puede ser una célula normal, por ejemplo, no enferma. En una realización, se incluyen al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 50 o más células testigo diferentes en el modelo celular. Todos los valores presentados en la lista precedente también

pueden ser el límite superior o inferior de rangos, que se pretende que sean parte de la presente invención, *por ejemplo,* entre 1 y 5, 1 y 10, 2 y 5 o 5 y 15 líneas celulares o tipos celulares testigo diferentes. En una realización, las células testigo son todos del mismo tipo pero son líneas celulares diferentes establecidas para dicho tipo celular. En una realización, como testigo, una o más líneas celulares normales, por ejemplo, no enfermas, pueden cultivarse en condiciones similares y/o exponerse a la misma perturbación, que las células primarias del modelo celular a efectos de identificar proteínas o vías únicas para el estado o proceso biológico.

Un modelo celular a medida de la invención también puede comprender condiciones que imitan un aspecto característico del estado o proceso biológico. Por ejemplo, las condiciones de cultivo celular pueden seleccionarse para aproximarse a las condiciones de una célula cancerosa en un ambiente de tumor in vivo, o de una célula de músculo liso de la aorta de un paciente que padece una enfermedad cardiovascular. En algunos casos, las condiciones son condiciones de estrés. Se pueden emplear varias condiciones / estresantes en los modelos celulares de la invención. En una realización, estos estresantes / condiciones pueden constituir la "perturbación", por ejemplo, estímulo externo, para los sistemas celulares. Una condición de estrés ejemplar es la hipoxia, una condición que se encuentra típicamente, por ejemplo, en los tumores sólidos. La hipoxia

puede inducirse usando métodos conocidos en la técnica. Por ejemplo, la hipoxia puede inducirse colocando sistemas celulares en una Modular Incubator Chamber (Cámara Incubadora Modular) (MIC-101, Billups-Rothenberg Inc. Del Mar, CA), que pueden inundarse con una mezcla gaseosa industrial que contenga 5% de CO_2, 2% de O_2 y 93% de nitrógeno. Los efectos pueden medirse después de un período predeterminado, *por ejemplo,* 24 horas después del tratamiento de hipoxia, con o sin componentes de estímulo externo adicionales *(por ejemplo,* CoQ10 a 0, 50 o 100 µM). Asimismo, el tratamiento con ácido láctico imita un ambiente celular donde la actividad de glucólisis es alta. El estrés inducido por ácido láctico puede investigarse a una concentración final de ácido láctico de aproximadamente 12,5 mM en un momento predeterminado, *por ejemplo,* a las 24 horas, con o sin componentes de estímulo externo adicionales *(por ejemplo,* CoQ10 a 0, 50 o 100 µM). La hiperglucemia es una condición encontrada en la diabetes así como también en el cáncer. Una condición típica de hiperglucemia que puede usarse para tratar células objeto incluye 10% de glucosa de grado de cultivo agregados a medios adecuados para proporcionar una concentración final de la glucosa en los medios de aproximadamente 22 mM. La hiperlipidemia es una condición encontrada, por ejemplo, en la obesidad y en la enfermedad cardiovascular. Las condiciones hiperlipidémicas pueden proporcionarse cultivando células en medios que contengan 0,15 mM de palmitato de sodio. La hiperinsulinemia es una

condición encontrada, por ejemplo, en la diabetes. Las condiciones hiperinsulinémicas pueden inducirse mediante el cultivo de células en medios que contengan 1000 nM de insulina.

Las condiciones individuales pueden investigarse por separado en modelos celulares construidos a medida de la invención y/o pueden combinarse entre sí. En una realización, una combinación de al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50 o más condiciones que reflejan o simulan diferentes aspectos característicos del sistema biológico se investigan en el modelo celular construido a medida. En una realización, las condiciones individuales y, además, combinaciones de al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20,

25, 30, 35, 40, 45, 50 o más condiciones que reflejan o simulan diferentes aspectos característicos del sistema biológico se investigan en el modelo celular construido a medida. Todos los valores presentados en la lista precedente también pueden ser los límites superiores o inferiores de rangos, que se pretenden que sean parte de esta invención, *por ejemplo,* entre 1 y 5, 1 y 10, 1 y 20, 1 y 30, 2 y 5, 2 y 10, 5 y 10, 1 y 20, 5 y 20, 10 y 20, 10 y 25, 10 y 30 o 10 y 50 condiciones diferentes.

Una vez construido el modelo celular a medida, se pueden aplicar una o más "perturbaciones" al sistema, tal como una variación genética de paciente en paciente, o con / sin tratamiento mediante ciertos fármacos o profármacos. *Ver* Figura 15D. Los efectos de dichas perturbaciones al sistema del modelo celular

pueden medirse usando varios medios conocidos en la técnica o patentados, tales como se describen en la sección III.B, más adelante.

El modelo celular construido a medida puede exponerse a una perturbación, por ejemplo, una "perturbación ambiental" o "componente de estímulo externo". La "perturbación ambiental" o "componente de estímulo externo" puede ser endógeno al ambiente celular *(por ejemplo,* el ambiente celular contiene algunos niveles del estimulante, y más del mismo se agrega para aumentar su nivel), o puede ser exógeno al ambiente celular *(por ejemplo,* el estimulante/perturbación está ausente en gran medida del ambiente celular antes de la alteración). El ambiente celular puede alterarse adicionalmente mediante cambios secundarios resultantes de la adición de la perturbación ambiental o componente de estímulo externo, ya que el componente de estímulo externo puede cambiar el rendimiento celular del sistema celular, incluidas moléculas secretadas en el ambiente celular por el sistema celular. La perturbación ambiental o componente de estímulo externo pueden incluir cualquier estímulo físico y/o químico externo que pueda afectar la función celular. Este puede incluir cualquier molécula orgánica o inorgánica, grande o pequeña, químicos naturales o sintéticos, cambio de temperatura, cambio de pH, radiación, luz (UVA, UVB *etc.)*, microondas, ondas de sonido, corriente eléctrica, campos magnéticos modulados o no modulados, *etc.* La perturbación ambiental o componente de

estímulo externo también puede incluir una modificación o mutación genética introducida o un vehículo (por ejemplo, vector) que provoque una modificación/mutación genética.

(i) Sistemas celulares de interacción

En ciertas situaciones, cuando se desea investigar una interacción entre dos o más sistemas celulares, se puede formar un "sistema celular de interacción", por ejemplo, poniendo en contacto el ambiente celular modificado de un primer sistema celular en contacto con un segundo sistema celular para afectar el rendimiento celular del segundo sistema celular.

Tal como se usa en la presente, un "sistema celular de interacción" comprende dos o más sistemas celulares, en los que el ambiente celular de al menos un sistema celular entra en contacto con un segundo sistema celular, de forma que al menos un rendimiento celular en el segundo sistema celular sea cambiado o afectado. En ciertas realizaciones, los sistemas celulares dentro del sistema celular de interacción pueden estar en contacto directo entre sí. En otras realizaciones, ninguno de los sistemas celulares está en contacto directo entre sí.

Por ejemplo, en ciertas realizaciones, el sistema celular de interacción puede estar en formar de una placa Transwell, en la que se cultiva un primer sistema celular en un inserto y un segundo sistema celular se cultiva en un compartimento de pocillo correspondiente. Los dos sistemas celulares pueden estar en contacto con los mismos medios o medios diferentes, y pueden

intercambiar parte o todos los componentes de los medios. El componente de estímulo externo agregado a un sistema celular puede absorberse sustancialmente por un sistema celular y/o degradado antes de poder dispersarse hasta el otro sistema celular. Alternativamente, el componente de estímulo externo puede aproximarse o alcanzar finalmente un equilibrio entre los dos sistemas celulares.

En ciertas realizaciones, el sistema celular de interacción puede adoptar la forma de sistema celulares cultivados separadamente, en donde cada sistema celular tiene su propio medio y/o condiciones de cultivo (temperatura, contenido de CO_2 , pH, *etc.)*, o condiciones de cultivo similares o idénticas. Los dos sistemas celulares pueden entrar en contacto, por ejemplo, tomando el medio acondicionado de un sistema celular y poniéndolo en contacto con el otro sistema celular. También se pueden llevar a cabo contactos directos célula-célula entre los dos sistemas celulares si se desea. Por ejemplo, las células de los dos sistemas celulares pueden cocultivarse en cualquier momento si se desea, y luego los sistemas celulares cocultivados pueden separar, por ejemplo, mediante clasificación FACS cuando las células en al menos un sistema celular tienen un marcador o etiqueta clasificable (tal como una proteína marcadora fluorescente expresada establemente GFP).

De forma similar, en ciertas realizaciones, el sistema celular de interacción puede ser simplemente un cocultivo. El tratamiento

selectivo de células en un sistema celular puede llevarse a cabo tratando primero las células en dicho sistema celular, antes de cultivar las células tratadas en cocultivo con células en otro sistema celular. El escenario de sistema celular de interacción de cocultivo puede ser útil cuando se desea estudiar, por ejemplo, los efectos en un segundo sistema celular provocados por cambios en la superficie celular en un primer sistema celular, después de la estimulación del primer sistema celular mediante un componente de estímulo externo.

El sistema celular de interacción de la invención es particularmente adecuado para explorar el efecto de ciertos componentes de estímulo externo predeterminados en el rendimiento celular de uno o ambos sistemas celulares. El efecto primario de dicho estímulo en el primer sistema celular (con el que el estímulo hace contacto directo) puede determinarse comparando los rendimientos celulares *(por ejemplo,* nivel de expresión proteica) antes y después del contacto del primer sistema celular con el estímulo externo, que, tal como se usa en la presente, puede denominarse como "diferenciales de rendimiento celular (significativos)". El efecto secundario de dicho estímulo sobre el segundo sistema celular, que es mediado a través del ambiente celular modificado del primer sistema celular (tal como su secretoma), también puede medirse de forma similar. Allí, por ejemplo, puede hacerse una comparación en el proteoma del segundo sistema celular entre el proteoma del segundo sistema celular *con*

el tratamiento con el estímulo externo en el primer sistema celular y el proteoma del segundo sistema celular *sin* el tratamiento con el estímulo externo en el primer sistema celular. Cualquier cambio significativo observado(en el proteoma o cualquier otro rendimiento celular de interés) puede denominarse "diferencial de interacción celular significativo".

Al llevar a cabo mediciones de rendimiento celular (tales como expresión proteica), se pueden usar cantidades absolutas o niveles relativos. Por ejemplo, para determinar el nivel de expresión proteica relativo de un segundo sistema celular, la cantidad de cualquier proteína dada en el segundo sistema celular, con o sin el estímulo externo para el primer sistema celular, puede comprarse con una línea celular o mezcla de líneas celulares testigo adecuadas y proporcionar un valor de aumento o disminución en veces. Se puede usar un nivel umbral predeterminado para dicho aumento en veces *(por ejemplo,* al menos 1,2, 1,3, 1,4, 1,5, 1,6, 1,7, 1,8, 1,9, 2, 2,5, 3, 3,5, 4, 4,5, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75 o 100 o un aumento mayor en veces) o disminución en veces *(por ejemplo,* al menos una disminución de 0,95, 0,9, 0,8, 0,75, 0,7, 0,6, 0,5, 0,45, 0,4, 0,35, 0,3, 0,25, 0,2, 0,15, 0,1 o 0,05 veces o una disminución del 90%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10% o 5% o menor) para seleccionar diferenciales de interacción celular significativos. Todos los valores presentados en la lista precedente también pueden ser el

límite superior o inferior de los rangos, *por ejemplo,* entre 1,5 y 5 veces, entre 2 y 10 veces, entre 1 y 2 veces o entre 0,9 y 0,7 veces, y se pretende que sean parte de la presente invención.

A lo largo de la presente solicitud, todos los valores presentados en una lista, por ejemplo, tales como los mencionados anteriormente, también puede ser el límite superior o inferior de rangos que se pretende que sean parte de la presente invención.

A efectos de ilustración, en un sistema de dos células ejemplar establecido para imitar aspectos de un modelo de enfermedad cardiovascular, se puede tratar una línea celular de músculo liso del corazón (primer sistema celular) con una condición de hipoxia (un componente de estímulo externo) y se pueden medir los cambios en el proteoma en una línea celular de riñón (segundo sistema celular) que resulten de poner en contacto las células de riñón con el medio acondicionado del músculo liso del corazón usando espectrometría de masas cuantitativa convencional. Los diferenciales de interacción celular significativos en estas células de riñón pueden determinarse, en base a comparación con células de riñón testigo apropiadas *(por ejemplo,* células de riñón cultivadas de forma similar puestas en contacto con medio acondicionado de células de músculo liso del corazón cultivadas de forma similar no tratadas con condiciones de hipoxia).

No todos los diferenciales de interacción celular significativos puede ser de importancia biológica. En relación con cualquier

sistema biológico dado para el cual se aplica la evaluación biológica interrogativa objeto, algunos (o tal vez todos) los diferenciales de interacción celular significativos pueden ser "determinativos" en relación con el problema biológico específico en cuestión, *por ejemplo,* responsable por provocar una enfermedad (un objetivo potencial para la intervención terapéutica) o es un biomarcador para la enfermedad (un potencial factor de diagnóstico o pronóstico).

Dichos diferenciales de interacción determinativos pueden ser seleccionados por un usuario final del método objeto, o pueden ser seleccionadas por un programa informático de bioinformática, tal como el programa de análisis comparativo de vías con capacidad DAVID o el programa de análisis de vías KEGG. En ciertas realizaciones, más de una programa informático de bioinformática se usa y se prefieren resultados por consenso de dos o más programas informáticos de bioinformática.

Tal como se usa en la presente, "diferenciales" de rendimientos celulares incluyen diferencias *(por ejemplo,* niveles aumentados o disminuidos) en cualquier uno o más parámetros de los rendimientos celulares. Por ejemplo, en términos de nivel de expresión proteica, los diferencias entre dos rendimientos celulares, tales como los rendimientos asociados con un sistema celular antes y después del tratamiento mediante un componente de estímulo externo, puede medirse y cuantificarse usando tecnologías conocidas en la técnica, tales como ensayos basados

en espectrometría de masas *(por ejemplo,* iTRAQ, 2D-LC-MSMS, *etc.)*.

(ii) Modelos específicos de cáncer

Un ejemplo de un sistema o proceso biológico es el cáncer. Como cualquier otro proceso o sistema biológico complicado, el cáncer es una afección patológica complicada caracterizada por múltiples aspectos únicos. Por ejemplo, debido a su alta tasa de crecimiento, muchas células de cáncer están adaptadas a crecer en condiciones de hipoxia, tiene glucólisis regulada de forma ascendente y vías metabólicas de fosforilación oxidativa reducidas. Como resultado, las células de cáncer pueden reaccionar de forma diferente a una perturbación ambiental, tal como tratamiento mediante un fármaco potencial, en comparación con la reacción de una célula normal en respuesta al mismo tratamiento. Por consiguiente, sería de interés descifrar las respuestas únicas del cáncer al tratamiento con el fármaco en comparación con las respuestas de las células normales. Con este fin, se puede establecer un modelo de cáncer a medida para simular el ambiente de una célula cancerosa, por ejemplo, dentro de un tumor *in vivo,* seleccionando las líneas celulares de cáncer apropiadas y creando las condiciones de cultivo celular que imitan un aspecto característico del estado o proceso de enfermedad. Por ejemplo, las condiciones de cultivo celular pueden seleccionarse para aproximarse a las condiciones de una célula cancerosa en un tumor in vivo, o para imitar varios

aspectos del crecimiento del cáncer, aislando diferentes condiciones de crecimiento de las células cancerosas.

Se pueden incluir múltiples células cancerosas de igual o diferente origen (por ejemplo las líneas cancerosas PaCa2, HepG2, PC3 y MCF7), en oposición a un único tipo de célula cancerosa, en el modelo de cáncer. En una realización, se incluyen al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50 o más líneas celulares de cáncer o tipos celulares de cáncer diferentes en el modelo de cáncer. Todos los valores presentados en la lista precedente también pueden ser el límite superior o inferior de rangos, que se pretende que sean parte de la presente invención, *por ejemplo,* entre 1 y 5, 1 y 10, 2 y 5 o 5 y 15 líneas celulares o tipos celulares de cáncer diferentes.

En una realización, las células cancerosas son todas del mismo tipo, por ejemplo, todas células de cáncer de mama, pero son líneas celulares diferentes establecidas, por ejemplo, líneas celulares diferentes establecidas de células de cáncer de mama.

Ejemplos de tipos de células cancerosas que pueden incluirse en el modelo de cáncer incluyen a modo no taxativo, cáncer pulmonar, cáncer de mama, cáncer de próstata, melanoma, carcinoma escamoso, cáncer colorrectal, cáncer pancreático, cáncer tiroideo, cáncer endometrial, cáncer de vejiga, cáncer de riñones, un tumor sólido, leucemia, linfoma no hodgkiniano. En una realización, se puede incluir una célula cancerosa resistente a fármacos en el

modelo de cáncer. Ejemplos específicos de líneas celulares que pueden incluirse en un modelo de cáncer incluyen, a modo no taxativo, células PaCa2, HepG2, PC3 y MCF7. En la técnica se conocen numerosas líneas celulares de cáncer y cualquiera de dichas líneas celulares de cáncer puede incluirse en un modelo de cáncer de la invención.

Los modelos celulares de la invención pueden incluir una o más "células testigo". En una realización, una célula testigo puede ser una célula cancerosa no tratada o no perturbada. En otra realización, una "célula testigo" puede ser una célula normal no cancerosa. Se puede incluir cualquier de numerosas líneas celulares normales, no cancerosas en el modelo celular. En una realización, las células normales son una o más células THLE2 y HDFa. En una realización, se incluyen al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50 o más tipos celulares normales diferentes en el modelo de cáncer. Todos los valores presentados en la lista precedente también pueden ser el límite superior o inferior de rangos, que se pretende que sean parte de la presente invención, *por ejemplo,* entre 1 y 5, 1 y 10, 2 y 5 o 5 y 15 líneas celulares o tipos celulares de cáncer normales diferentes. En una realización, las células normales son todas del mismo tipo, por ejemplo, todas células epiteliales o de mama sanas, pero son líneas celulares diferentes establecidas, por ejemplo, líneas celulares diferentes establecidas de células epiteliales o de mama. En una

realización, como testigo, una o más líneas celulares normales no cancerosas (por ejemplo, THLE2 y HDFa) se cultivan en condiciones similares y/o se exponen a la misma perturbación, que las células cancerosas del modelo celular a efectos de identificar proteínas o vías únicas del cáncer.

Un modelo celular de cáncer a medida también puede comprender condiciones de cultivo celular que imitan un aspecto característico del estado o proceso canceroso. Por ejemplo, las condiciones de cultivo celular pueden seleccionarse para aproximarse a las condiciones de una célula cancerosa en un ambiente de tumor in vivo, o para imitar varios aspectos del crecimiento del cáncer, aislando diferentes condiciones de crecimiento de las células cancerosas. En algunos casos, las condiciones de cultivo celular son condiciones de estrés.

Uno de dicho "ambiente" de cáncer, o condición de estrés, es hipoxia, una condición que se encuentra típicamente dentro de un tumor sólido. La hipoxia puede inducirse en las células usando métodos conocidos en la técnica. Por ejemplo, la hipoxia puede inducirse colocando sistemas celulares en una Modular Incubator Chamber (Cámara Incubadora Modular) (MIC-101, Billups-Rothenberg Inc. Del Mar, CA), que pueden inundarse con una mezcla gaseosa industrial que contenga 5% de CO_2, 2% de O_2 y 93% de nitrógeno. Los efectos pueden medirse después de un período predeterminado, *por ejemplo,* 24 horas después del tratamiento de hipoxia, con o sin

componentes de estímulo externo adicionales *(por ejemplo,* CoQ10 a 0, 50 o 100 µM).

Asimismo, el tratamiento con ácido láctico de las células imita un ambiente celular donde la actividad de glucólisis es alta, tal como la que existe en el ambiente del tumor in vivo. El estrés inducido por el ácido láctico puede investigarse a una concentración final de ácido láctico de aproximadamente 12,5 mM en un momento predeterminado, *por ejemplo,* a las 24 horas, con o sin componentes de estímulo externo adicionales *(por ejemplo,* CoQ10 a 0, 50 o 100 µM).

La hiperglucemia es normalmente una condición encontrada en la diabetes; sin embargo, en cierta medida, la hiperglucemia también imita un aspecto del crecimiento del cáncer debido a que muchas células cancerosas dependen de la glucosa como su fuente principal de energía. La exposición de células objeto a un condición típica de hiperglucemia puede incluir agregar 10% de glucosa de grado de cultivo a medios adecuados, de forma que la concentración final de la glucosa en los medios sea de aproximadamente 22 mM.

Las condiciones individuales que reflejan aspectos del crecimiento canceroso pueden investigarse por separado en el modelo de cáncer construido a medida y/o pueden combinarse entre sí. En una realización, las combinaciones de al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50 o más condiciones que reflejan o simulan diferentes aspectos del crecimiento /

condiciones del cáncer se investigan en el modelo de cáncer construido a medida. En una realización, las condiciones individuales y, además, las combinaciones de al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50 o más condiciones que reflejan o simulan diferentes aspectos del crecimiento / condiciones del cáncer se investigan en el modelo de cáncer construido a medida. Todos los valores presentados en la lista precedente también pueden ser los límites superiores o inferiores de rangos, que se pretenden que sean parte de esta invención, *por ejemplo,* entre 1 y 5, 1 y 10, 1 y 20, 1 y 30, 2 y 5, 2 y 10, 5 y 10, 1 y 20, 5 y 20, 10 y 20, 10 y 25, 10 y 30 o 10 y 50 condiciones diferentes.

Una vez construido el modelo celular a medida, se pueden aplicar una o más "perturbaciones" al sistema, tal como una variación genética de paciente en paciente, o con / sin tratamiento mediante ciertos fármacos o profármacos. *Ver* Figura 15D. Los efectos de dichas perturbaciones al sistema, incluido el efecto en células cancerosas relacionadas con la enfermedad y células testigo normales relacionadas con la enfermedad, pueden medirse usando varios medios conocidos en la técnica o patentados, tales como se describen en la sección III.B, más adelante.

En un experimento ejemplar, las líneas de cáncer PaCa2, HepG2, PC3 y MCF7 y las líneas celulares normales THLE2 y HDFa, se acondicionan en cada una de hiperglucemia, hipoxia y condiciones ricas en ácido láctico, así como en todas las combinaciones de

dos o tres de estas condiciones, y además, con o sin una perturbación ambiental, específicamente tratamiento con Coenzima Q10. A continuación se indican dichas combinación ejemplares de condiciones, con o sin una perturbación, tratamiento con Coenzima Q10, que pueden usarse para tratar células cancerosas y/o células testigo (por ejemplo, normales) del modelo celular de cáncer. Otras combinaciones pueden formularse fácilmente dependiendo de la evaluación biológica interrogativa específica que se está llevando a cabo.

1. Solo medios
2. 50 µM de CTL Coenzima Q10
3. 100 µM de CTL Coenzima Q10 4. 12,5 mM de ácido láctico
5. 12,5 mM de ácido láctico + 50 µM de CTL Coenzima Q10
6. 12,5 mM de ácido láctico + 100 µM de CTL Coenzima Q10
7. Hipoxia
8. Hipoxia + 50 µM de CTL Coenzima Q10
9. Hipoxia + 100 µM de CTL Coenzima Q10
10. Hipoxia + 12,5 mM de ácido láctico
11. Hipoxia + 12,5 mM de ácido láctico + 50 µM de CTL Coenzima Q10
12. Hipoxia + 12,5 mM de ácido láctico + 100 µM de CTL Coenzima Q10
13. Medios + 22 mM de Glucosa
14. 50 µM de CTL Coenzima Q10 + 22 mM de Glucosa

15. 100 µM de CTL Coenzima Q10 + 22 mM de Glucosa

16. 12,5 mM de ácido láctico + 22 mM de Glucosa

17. 12,5 mM de ácido láctico + 22 mM de Glucosa + 50 µM de CTL Coenzima Q10

18. 12,5 mM de ácido láctico + 22 mM de Glucosa + 100 µM de CTL Coenzima Q10

19. Hipoxia + 22 mM de Glucosa

20. Hipoxia + 22 mM de Glucosa + 50 µM de CTL Coenzima Q10

21. Hipoxia + 22 mM de Glucosa + 100 µM de CTL Coenzima Q10

22. Hipoxia + 12,5 mM de ácido láctico + 22 mM de Glucosa

23. Hipoxia + 12,5 mM de ácido láctico + 22 mM de Glucosa + 50 µM de CTL Coenzima Q10

24. Hipoxia + 12,5 mM de ácido láctico + 22 mM de Glucosa + 100 µM de CTL Coenzima Q10

En ciertas situaciones, se pueden llevar a cabo experimentos de interacción o ECS entre diferentes células cancerosas (por ejemplo, HepG2 y PaCa2) para varios fines interrelacionados. En algunas realizaciones que implican interacción, los experimentos que se llevan a cabo en los modelos celulares se diseñan para determinar la modulación del estado o función celular de un sistema o población celular *(por ejemplo,* célula de hepatocarcinoma, HepG2) mediante otro sistema o población celular *(por ejemplo,* cáncer pancreático, PaCa2) en condiciones de tratamiento definidas *(por ejemplo,* hiperglucemia, hipoxia (isquemia)). De acuerdo con un escenario típico, se pone en

contacto un primer sistema / población celular con componentes de estímulo externo, tales como una molécula candidata *(por ejemplo,* una molécula de fármaco pequeño, una proteína) o una condición candidata *(por ejemplo,* hipoxia, ambiente con alto contenido de glucosa). En respuesta, el primer sistema/población celular cambia su transcriptoma, proteoma, metaboloma y/o interactoma, conduciendo a cambios que pueden detectarse fácilmente dentro y fuera de la célula. Por ejemplo, se pueden medir cambios en el transcriptoma mediante el nivel de transcripción de una pluralidad de ARNm objetivo; los cambios en el proteoma pueden medirse mediante el nivel de expresión de una pluralidad de proteínas objetivo; y los cambios en el metaboloma pueden medirse mediante el nivel de una pluralidad de metabolitos objetivo mediante ensayos diseñados específicamente para metabolitos dados. Alternativamente, los cambios mencionados anteriormente en el metaboloma y/o proteoma, al menos con respecto a ciertos metabolitos o proteínas secretados, también pueden medirse por sus efectos en el segundo sistema / población celular, incluida la modulación del transcriptoma, proteoma, metaboloma e interactoma del segundo sistema / población celular. Por lo tanto, los experimentos pueden usarse para identificar los efectos de la molécula(s) de interés secretada por el primer sistema / población celular en un segundo sistema / población celular en diferentes condiciones de tratamiento. Los experimentos también pueden usarse para identificar cualquier

proteína modulada como resultado de la señalización del primer sistema celular (en respuesta al tratamiento con el componente de estímulo externo) en otro sistema celular, mediante, por ejemplo, detección diferencial de datos proteómicos. El mismo escenario experimental también puede adaptarse para un escenario inverso, de forma que los efectos recíprocos entre los dos sistemas celulares también pueden evaluarse. En general, para este tipo de experimento, la elección de los pares de líneas celulares se basa en gran medida en los factores tales como el origen, estado de enfermedad y función celular.

Aunque típicamente participan sistemas de dos células en este tipo de escenario experimental, también se pueden diseñar experimentos similares para más de dos sistemas celulares, por ejemplo, inmovilizando cada sistema celular diferenciado en un soporte sólido separado.

El modelo celular de cáncer construido a medida puede establecerse y usarse en las etapas de la tecnología de plataforma de la invención para identificar finalmente una relación causal única en el sistema biológico, llevando a cabo las etapas descritas en la presente. El experto en la técnica entenderá, sin embargo, que un modelo celular de cáncer construido a medida que se usa para generar una red de relaciones causales de consenso de "primera generación" inicial puede evolucionar continuamente o expandirse con el tiempo, por ejemplo, mediante la introducción de líneas celulares adicionales

de cáncer o normales y/o condiciones de cáncer adicionales. Se pueden recoger datos adicionales del modelo de cáncer evolucionado, es decir, datos de la nueva porción(es) agregada del modelo de cáncer. Los nuevos datos recogidos de un modelo de cáncer expandido o evolucionado, es decir, de la nueva porción(es) agregada del modelo de cáncer, a continuación pueden introducirse a los conjuntos de datos usados previamente para generar la red de relaciones causales de consenso de "primera generación" a efectos de generar una red de relaciones causales de consenso de "segunda generación" más sólida. A continuación pueden identificarse nuevas relaciones causales únicas para el estado de cáncer (o únicas para la respuesta del estado de cáncer a una perturbación) a partir de red de relaciones causales de consenso de "segunda generación". De esta forma, la evolución del modelo de cáncer proporciona una evolución de las redes de relaciones causales de consenso, proporcionando así perspectivas nuevas y/o más confiables con respecto a los activadores determinativos (o moduladores) del estado de cáncer.

(iii) Modelos de diabetes/obesidad/enfermedad cardiovascular

Otros ejemplos de un sistema o proceso biológico son la diabetes, la obesidad y la enfermedad cardiovascular. Tal como con el cáncer, los estados de enfermedad relacionados de diabetes, obesidad y enfermedad cardiovascular con afecciones patológicas complicadas caracterizadas por múltiple aspectos únicos. Sería de interés identificar las proteínas/vías que activan la patogénesis

de la diabetes/obesidad/ enfermedad cardiovascular. Además sería de interés descifrar las respuesta única de células asociadas con la diabetes/obesidad/enfermedad cardiovascular al tratamiento con fármacos en comparación con las respuestas de células normales. Con este fin, se puede establecer un modelo de diabetes/obesidad/enfermedad cardiovascular a medida para simular un ambiente experimentado por células relevantes para la enfermedad, seleccionando las líneas celulares apropiadas y creando las condiciones de cultivo celular que imitan un aspecto característico del estado o proceso de enfermedad. Por ejemplo, las condiciones de cultivo celular pueden seleccionarse para aproximarse a la hiperglucemia, hiperlipidemia, hiperinsulinemia, hipoxia o condiciones ricas en ácido láctico.

Se pueden incluir cualesquiera células relevantes para la diabetes/obesidad/enfermedad cardiovascular en el modelo de diabetes/obesidad/enfermedad cardiovascular. Ejemplos de células relevantes para la diabetes/obesidad/enfermedad cardiovascular incluyen, por ejemplo, adipocitos, miotubos, hepatocitos, células de músculo liso de la aorta (HASMC) y células tubulares proximales (por ejemplo,HK2). Se pueden incluir múltiples tipos celulares de igual o diferente origen, en oposición a un único tipo de célula, en el modelo de diabetes/obesidad/enfermedad cardiovascular. En una realización, se incluyen al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25,

30, 35, 40, 45, 50 o más tipos celulares diferentes en el modelo de

diabetes/obesidad/enfermedad cardiovascular. Todos los valores presentados en la lista precedente también pueden ser el límite superior o inferior de rangos, que se pretende que sean parte de la presente invención, *por ejemplo,* entre 1 y 5, 1 y 10, 2 y 5 o 5 y 15 tipos celulares diferentes. En una realización, las células son todas del mismo tipo, por ejemplo, todas adipocitos, pero son líneas celulares diferentes establecidas, por ejemplo, líneas celulares diferentes establecidas de células de adipocitos. Se conocen en la técnica varios otros tipos celulares que participan en el estado de diabetes/obesidad/enfermedad cardiovascular y dichas células pueden incluirse en un modelo de diabetes/obesidad/enfermedad cardiovascular de la invención.

Los modelos celulares de diabetes/obesidad/enfermedad cardiovascular de la invención pueden incluir una o más "células testigo". En una realización, una célula testigo puede ser una célula relevante para la enfermedad no tratada o no perturbada, por ejemplo, una célula que no se expone a una condición hiperlipidémica o hiperinsulinémica. En otra realización, un "célula testigo" puede ser una célula no relevante para la enfermedad, tal como una célula epitelial. Se puede incluir cualquiera de numerosas células no relevantes para la enfermedad en el modelo celular. En una realización, se incluyen al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20,

25, 30, 35, 40, 45, 50 o más tipos celulares no relevantes para la enfermedad diferentes en el modelo celular. Todos los valores presentados en la lista precedente también pueden ser el límite superior o inferior de rangos, que se pretende que sean parte de la presente invención, *por ejemplo,* entre 1 y 5, 1 y 10, 2 y 5 o 5 y 15 líneas celulares o tipos celulares no relevantes para la enfermedad diferentes. En una realización, las células no relevantes para la enfermedad son todas del mismo tipo, por ejemplo, todas células epiteliales o de mama sanas, pero son líneas celulares diferentes establecidas, por ejemplo, líneas celulares diferentes establecidas de células epiteliales o de mama. En una realización, como testigo, una o más líneas celulares no relevantes para la enfermedad se cultivan en condiciones similares y/o se exponen a la misma perturbación, que las células relevantes para la enfermedad del modelo celular a efectos de identificar proteínas o vías únicas para la diabetes/obesidad/enfermedad cardiovascular.

Un modelo de diabetes/obesidad/enfermedad cardiovascular a medida también puede comprender condiciones de cultivo celular que imitan un aspecto característico del (representan la patofisiología del) estado o proceso de la diabetes/obesidad/enfermedad cardiovascular. Por ejemplo, las condiciones de cultivo celular pueden seleccionarse para aproximarse a las condiciones de una célula

relevante para la diabetes/obesidad/enfermedad cardiovascular en su ambiente in vivo, o para imitar varios aspectos de la diabetes/obesidad/enfermedad cardiovascular. En algunos casos, las condiciones de cultivo celular son condiciones de estrés.

Las condiciones ejemplares que representan la patofisiología de la diabetes/obesidad/enfermedad cardiovascular incluyen, por ejemplo, una o más de hipoxia, condiciones ricas en ácido láctico, hiperglucemia, hiperlipidemia e hiperinsulinemia. La hipoxia puede inducirse en las células usando métodos conocidos en la técnica. Por ejemplo, la hipoxia puede inducirse colocando sistemas celulares en una Modular Incubator Chamber (Cámara Incubadora Modular) (MIC-101, Billups-Rothenberg Inc. Del Mar, CA), que pueden inundarse con una mezcla gaseosa industrial que contenga 5% de CO_2, 2% de O_2 y 93% de nitrógeno. Los efectos pueden medirse después de un período predeterminado, *por ejemplo,* 24 horas después del tratamiento de hipoxia, con o sin componentes de estímulo externo adicionales *(por ejemplo,* CoQ10 a 0, 50 o 100 µM).

Asimismo, el tratamiento de células con ácido láctico imita un ambiente celular donde la actividad de glucólisis es alta. El estrés inducido por ácido láctico puede investigarse a una concentración final de ácido láctico de aproximadamente 12,5 mM en un momento predeterminado, *por ejemplo,* a las 24 horas, con o sin componentes de estímulo externo adicionales *(por ejemplo,* CoQ10 a 0, 50 o 100 µM). La hiperglucemia es una condición

encontrada en la diabetes. La exposición de células objeto a un condición típica de hiperglucemia puede incluir agregar 10% de glucosa de grado de cultivo a medios adecuados, de forma que la concentración final de la glucosa en los medios sea de aproximadamente 22 mM. La hiperlipidemia es una condición encontrada en la obesidad y en la enfermedad cardiovascular. Las condiciones hiperlipidémicas pueden proporcionarse cultivando células en medios que contengan 0,15 mM de palmitato de sodio. La hiperinsulinemia es una condición encontrada en la diabetes. Las condiciones hiperinsulinémicas pueden inducirse mediante el cultivo de células en medios que contengan 1000 nM de insulina. Las condiciones adicionales que representan la patofisiología de la diabetes/obesidad/enfermedad cardiovascular incluyen, por ejemplo, cualquier una o más de inflamación, estrés del retículo endoplasmático, estrés mitrocondrial y estrés peroxisómico. Los métodos para crear una condición similar a una inflamatoria en células se conocen en la técnica. Por ejemplo, una condición inflamatoria puede simularse cultivando células en presencia de TNFalfa y/o IL-6. Los métodos para crear una condiciones que simulan el estrés del retículo endoplasmático también se conocen en la técnica. Por ejemplo, se pueden crear condiciones para simular el estrés del retículo endoplasmático cultivando células en presencia de tapsigargina y/o tunicamicina. Los métodos para crear una condiciones que simulan el estrés mitocondrial también

se conocen en la técnica. Por ejemplo, se pueden crear condiciones para simular el estrés mitocondrial cultivando células en presencia de rapamicina y/o galactosa. Los métodos para crear una condiciones que simulan el estrés peroxisómico también se conocen en la técnica. Por ejemplo, se pueden crear condiciones para simular el estrés peroxisómico cultivando células en presencia ácido abscísico.

Las condiciones individuales que reflejan diferentes aspectos de la diabetes/obesidad/enfermedad cardiovascular pueden investigarse por separado en el modelo de diabetes/obesidad/enfermedad cardiovascular construido a medida y/o pueden combinarse entre sí. En una realización, las combinaciones de al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50 o más condiciones que reflejan o simulan diferentes aspectos de la diabetes/obesidad/enfermedad cardiovascular se investigan en el modelo de diabetes/obesidad/enfermedad cardiovascular construido a medida. En una realización, las condiciones individuales y, además, las combinaciones de al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50 o más condiciones que reflejan o simulan diferentes aspectos de la diabetes/obesidad/enfermedad cardiovascular se investigan en el modelo de diabetes/obesidad/enfermedad cardiovascular construido a medida. Todos los valores presentados en la lista precedente también pueden ser los límites superiores o inferiores de rangos, que se pretenden que sean parte de esta invención, *por ejemplo,*

entre 1 y 5, 1 y 10, 1 y 20, 1 y 30, 2 y 5, 2 y 10, 5 y 10, 1 y 20, 5 y 20, 10 y 20, 10 y 25, 10 y 30 o 10 y 50 condiciones diferentes.

Una vez construido el modelo celular a medida, se pueden aplicar una o más "perturbaciones" al sistema, tal como una variación genética de paciente en paciente, o con / sin tratamiento mediante ciertos fármacos o profármacos. *Ver* Figura 15D. Los efectos de dichas perturbaciones al sistema, incluido el efecto sobre las células relacionadas con la diabetes/obesidad/enfermedad cardiovascular pueden medirse usando varios medios conocidos en la técnica o patentados, tales como se describen en la sección III .B, más adelante.

En un experimento ejemplar, cada uno de adipocitos, miotubos, hepatocitos, células de músculo liso de la aorta(HASMC) y células tubulares proximales (HK2), se acondicionan en cada una de hiperglucemia, hipoxia, hiperlipidemia, hiperinsulinemia y condiciones ricas en ácido láctico, así como en todas las combinaciones de dos, tres, cuatro o todas las cinco condiciones, y además, con o sin una perturbación ambiental, específicamente tratamiento con Coenzima Q10. Además de las combinaciones ejemplares de condiciones descritas anteriormente en el contexto del modelo de cáncer, a continuación se indican en la presente algunas combinaciones de condiciones

ejemplares adicionales, con o sin una perturbación, por ejemplo, tratamiento con Coenzima Q10, que pueden usarse para tratar las

células relevantes para la diabetes/obesidad/enfermedad cardiovascular (y/o células testigo) del modelo celular de diabetes/obesidad/enfermedad cardiovascular. Se pretende que sean meramente ejemplares y el experto en la técnica apreciará que cualquier condición individual y/o combinación mencionada anteriormente que representa la patofisiología de la diabetes/obesidad/enfermedad cardiovascular puede emplearse en el modelo celular para producir conjunto de datos de rendimiento. Otras combinaciones pueden formularse fácilmente dependiendo de la evaluación biológica interrogativa específica que se está llevando a cabo.

1. Solo medios
2. 50 µM de CTL Coenzima Q10
3. 100 µM de CTL Coenzima Q10
4. 0,15 mM de palmitato de sodio
5. 0,15 mM de palmitato de sodio + 50 µM de CTL Coenzima Q10
6. 0,15 mM de palmitato de sodio + 100 µM de CTL Coenzima Q10
7. 1000 nM de insulina
8. 1000 mM de insulina + 50 µM de CTL Coenzima Q10
9. 1000 mM de insulina + 100 µM de CTL Coenzima Q10
10. 1000 nM de insulina + 0,15 mM de palmitato de sodio
11. 1000 nM de insulina + 0,15 mM de palmitato de sodio + 50 µM de CTL Coenzima Q10
12. 1000 nM de insulina + 0,15 mM de palmitato de sodio + 100 µM de CTL Coenzima Q10

En ciertas situaciones, se pueden llevar a cabo experimentos de interacción o ECS entre diferentes células relevantes para la enfermedad (por ejemplo, células HASMC y HK2 o células hepáticas y adipocitos) para varios fines interrelacionados. En algunas realizaciones que implican interacción, los experimentos que se llevan a cabo en los modelos celulares se diseñan para determinar la modulación del estado o función celular de un sistema o población celular *(por ejemplo,* células hepáticas) mediante otro sistema o población celular *(por ejemplo,* adipocitos) en condiciones de tratamiento definidas *(por ejemplo,* hiperglucemia, hipoxia, hiperlipidemia, hiperinsulinemia). De acuerdo con un escenario típico, un primer sistema/población celular se pone en contacto con componentes de estímulo externo, tales como una molécula candidata *(por ejemplo,* una molécula de fármaco pequeña, una proteína) o una condición candidata *(por ejemplo,* hipoxia, ambiente con alto contenido de glucosa). En respuesta, el primer sistema/población celular cambia su transcriptoma, proteoma, metaboloma y/o interactoma, conduciendo a cambios que pueden detectarse fácilmente dentro y fuera de la célula. Por ejemplo, se pueden medir cambios en el transcriptoma mediante el nivel de transcripción de una pluralidad de ARNm objetivo; los cambios en el proteoma pueden medirse mediante el nivel de expresión de una pluralidad de proteínas objetivo; y los cambios en el metaboloma pueden medirse mediante el nivel de una pluralidad de metabolitos objetivo mediante ensayos diseñados específicamente para

metabolitos dados. Alternativamente, los cambios mencionados anteriormente en el metaboloma y/o proteoma, al menos con respecto a ciertos metabolitos o proteínas secretados, también pueden medirse por sus efectos en el segundo sistema / población celular, incluida la modulación del transcriptoma, proteoma, metaboloma e interactoma del segundo sistema / población celular. Por lo tanto, los experimentos pueden usarse para identificar los efectos de la molécula(s) de interés secretada por el primer sistema / población celular en un segundo sistema / población celular en diferentes condiciones de tratamiento. Los experimentos también pueden usarse para identificar cualquier proteína modulada como resultado de la señalización del primer sistema celular (en respuesta al tratamiento con el componente de estímulo externo) en otro sistema celular, mediante, por ejemplo, detección diferencial de datos proteómicos. El mismo escenario experimental también puede adaptarse para un escenario inverso, de forma que los efectos recíprocos entre los dos sistemas celulares también pueden evaluarse. En general, para este tipo de experimento, la elección de los pares de líneas celulares se basa en gran medida en los factores tales como el origen, estado de enfermedad y función celular.

Aunque típicamente participan sistemas de dos células en este tipo de escenario experimental, también se pueden diseñar experimentos similares para más de dos sistemas celulares, por

ejemplo, inmovilizando cada sistema celular diferenciado en un soporte sólido separado.

El modelo celular de diabetes/obesidad/enfermedad cardiovascular construido a medida puede establecerse y usarse en las etapas de la tecnología de plataforma de la invención para identificar finalmente una relación causal única en el estado de diabetes/obesidad/enfermedad cardiovascular, llevando a cabo las etapas descritas en la presente. El experto en la técnica entenderá, sin embargo, que tal como en el modelo de cáncer, un modelo de diabetes/obesidad/enfermedad cardiovascular construido a medida que se usa para generar una red de relaciones causales de consenso de "primera generación" inicial para un proceso biológico puede evolucionar continuamente p expandirse con el tiempo, por ejemplo, mediante la introducción de líneas celulares adicionales relevantes para la enfermedad y/ocondiciones adicionales relevantes para la enfermedad. Se pueden recoger datos adicionales del modelo de diabetes/obesidad/enfermedad cardiovascular evolucionado, es decir, datos de la nueva porción(es) agregada del modelo de diabetes/obesidad/enfermedad cardiovascular. Los nuevos datos recogidos de un modelo expandido o evolucionado, es decir, de la nueva porción(es) agregada del modelo, a continuación pueden introducirse en los conjuntos de datos usados previamente para generar la red de relaciones causales de consenso de "primera generación" a efectos de generar una red de relaciones causales de consenso de "segunda

generación" más sólida. A continuación pueden identificarse nuevas relaciones causales únicas para el estado de diabetes/obesidad/enfermedad cardiovascular (o únicas para la respuesta del estado de diabetes/obesidad/enfermedad cardiovascular a una perturbación) a partir de red de relaciones causales de consenso de "segunda generación". De esta forma, la evolución del modelo de diabetes/obesidad/enfermedad cardiovascular proporciona una evolución de las redes de relaciones causales de consenso, proporcionando así perspectivas nuevas y/o más confiables con respecto a los activadores determinativos (o moduladores) del estado de diabetes/obesidad/enfermedad cardiovascular.

B. Uso de modelos celulares para evaluaciones biológicas interrogativas

Los métodos y modelos celulares proporcionados en la presente la invención pueden usarse, o aplicarse, en cualquier cantidad de "evaluaciones biológicas interrogativas". El uso de los métodos de la invención en una evaluación biológica interrogativa facilitan la identificación de "moduladores" o "activadores" de proceso celular determinativos de un sistema biológico.

Tal como se usa en la presente, una "evaluación biológica interrogativa" puede incluir la identificación de uno o más moduladores de un sistema biológico, por ejemplo, "activadores" de proceso celular determinativos (por ejemplo, un aumento o disminución en la actividad de una vía biológica, o miembros

clave de la vía o reguladores clave de miembros de la vía) asociados con la perturbación ambiental o componente de estímulo externo, o una relación causal única en un sistema o proceso biológico. Puede incluir además etapas adicionales diseñadas para evaluar o verificar si los activadores de proceso celular determinativos identificados son necesarios y/o suficientes para eventos descendentes asociados con la perturbación ambiental o componente de estímulo externo, incluidos experimentos con modelos animales in vivo y/o con cultivo de tejidos in vitro.

En ciertas realizaciones, la evaluación biológica interrogativa es el diagnóstico o estadificación de un estado de enfermedad, en donde los moduladores identificados de un sistema biológico, por ejemplo, activadores de proceso celular determinativos (por ejemplo, diferenciales de interacción o relacionescausales únicas en un sistema o proceso biológico) representan marcadores de enfermedad u objetivos terapéuticos que pueden someterse a intervención terapéutica. La evaluación biológica interrogativa es adecuada para cualquier enfermedad en teoría, pero puede ser particularmente útil en áreas tales como la oncología / biología del cáncer, diabetes, obesidad, enfermedades cardiovasculares y afecciones neurológicas (especialmente enfermedades neurodegenerativas, tales como, a modo no taxativo, enfermedad de Alzheimer, enfermedad de Parkinson, enfermedad de Huntington, esclerosis lateral amiotrófica (ELA) y neurodegeneración relacionada con el envejecimiento).

En ciertas realizaciones, la evaluación biológica interrogativa es la determinación de la eficacia de un fármaco, en donde los moduladores identificados de un sistema biológico, por ejemplo, activador de proceso celular determinativo (por ejemplo, diferenciales de interacción o relaciones causales únicas en un proceso o sistema biológico) pueden ser los signos característicos de un fármaco exitoso y a su vez pueden usarse para identificar agentes adicionales, tales como MIM o desplazadores epi, para tratar la misma enfermedad.

En ciertas realizaciones, la evaluación biológica interrogativa es la identificación de objetivos de fármacos para prevenir o tratar infecciones (por ejemplo, infección bacteriana o viral), en donde el activador de proceso celular determinativo identificado (por ejemplo, diferenciales de interacción o relaciones causales únicas en un proceso o sistema biológico) puede ser marcadores/indicadores o moléculas biológicas clave causantes del estado infeccioso, y a su vez pueden usarse para identificar agentes antiinfecciosos.

En ciertas realizaciones, la evaluación biológica interrogativa es la evaluación de un efecto molecular de un agente, por ejemplo, un fármaco, en un perfil de enfermedad dado, en donde los moduladores identificados de un sistema biológico, por ejemplo, activador de proceso celular determinativo (por ejemplo, diferenciales de interacción celular o relaciones causales únicas en un sistema o proceso biológico) pueden ser un aumento o

disminución en la actividad de una o más vías biológicas, o miembros clave de la vía(s), o reguladores clave de miembros de la vía(s), y a su vez pueden usarse, por ejemplo, para predecir la eficacia terapéutica del agente para la enfermedad dada.

En ciertas realizaciones, la evaluación biológica interrogativa es la evaluación del perfil toxicológico de un agente, por ejemplo, un fármaco, en una célula, tejido, órgano u organismo, en donde los moduladores identificados de un sistema biológico, por ejemplo, activador de proceso celular determinativo (por ejemplo, diferenciales de interacción celular o relaciones causales únicas en una sistema o proceso biológico) pueden ser indicadores de toxicidad, por ejemplo, citotoxicidad, y pueden a su vez usarse para predecir o identificar el perfil toxicológico de un agente. En una realización, los moduladores identificados de un sistema biológico, por ejemplo, activador de proceso celular determinativo (por ejemplo, diferenciales de interacción celular o relaciones causales únicas en un sistema o proceso biológico) es un indicador de la cardiotoxicidad de un fármaco o fármaco candidato, y a su vez, puede usarse para predecir o identificar el perfil cardiotoxicológico del fármaco o fármaco candidato.

En ciertas realizaciones, la evaluación biológica interrogativa es la identificación de objetivos de fármacos para prevenir o tratar una enfermedad o trastorno provocado por armas biológicas, tales como protozoarios, hongos, bacterias, protestos, virus o

toxinas, en donde los moduladores identificados de un sistema biológico, por ejemplo, el activador de proceso celular determinativo (por ejemplo, diferenciales de interacción o relaciones causales únicas en un proceso o sistema biológico) pueden ser marcadores/indicadores o moléculas biológicas clave causantes de dicha enfermedad o trastorno, y a su vez pueden usarse para identificar agentes de biodefensa.

En ciertas realizaciones, la evaluación biológica interrogativa es la identificación de objetivos para agentes contra el envejecimiento, tales como cosméticos contra el envejecimiento, en donde los moduladores identificados de un sistema biológico, por ejemplo, el activador de proceso celular determinativo (por ejemplo, diferenciales de interacción o relaciones causales únicas en un proceso o sistema biológico) pueden ser marcadores o indicadores del proceso de envejecimiento, específicamente el proceso de envejecimiento en la piel, y a su vez pueden usarse para identificar agentes contra el envejecimiento.

En un modelo celular ejemplar para el envejecimiento que se usa en los métodos de la invención para identificar objetivos para cosméticos contra el envejecimiento, el modelo celular comprende una célula epitelial envejecida que, por ejemplo, se trata con luz UV (una perturbación ambiental o componente de estímulo externo), y/o células neonatales, que también se tratan opcionalmente con luz UV. En una realización, un modelo celular del envejecimiento comprende un sistema de interacción celular.

En un sistema de interacción entre dos células ejemplar establecido para identificar objetivos para cosméticos contra el envejecimiento, una célula epitelial envejecida (primer sistema celular) puede tratarse con luz UV (un componente de estímulo externo), y se pueden medir cambios, por ejemplo, cambios proteómicos y/o cambios funcionales, en una célula neonatal (segundo sistema celular) resultante del contacto de las células neonatales con medio acondicionado de la célula epitelial envejecida tratada, por ejemplo, pueden medirse cambios proteómicos usando espectrometría de masas, o se puede identificar una relación causal única a partir de una red de relaciones causales generada a partir de los datos.

V. Análisis proteómico de muestra

En ciertas realizaciones, el método objeto puede emplear análisis proteómico cuantitativo de alta productividad a gran escala de cientos de muestras de carácter similar, y proporciona los datos necesarios para identificar los diferenciales de rendimiento celular.

Existen numerosas tecnologías conocidas en la técnica adecuadas a estos efectos. Un técnica ejemplar es el análisis iTRAQ en combinación con espectrometría de masas y se describe brevemente a continuación.

Para proporcionar muestras de referencia para la cuantificación relativa con la técnica iTRAQ, se crean múltiples agrupaciones QC. Se generaron dos acumulaciones QC separadas, que consisten en

alícuotas de cada muestra, de las muestras de Célula No. 1 y Célula No. 2 - estas muestras se denominan QCS1 y QCS2, y QCP1 y QCP2 para los sobrenadantes y gránulos, respectivamente. A efectos posibilitar la comparación de concentración proteica en las dos líneas celulares, se combinan alícuotas de *granulado* celular de las agrupaciones QC descritas anteriormente en volúmenes iguales para generar muestras de referencia (QCP).

El abordaje proteómico cuantitativo se basa en el etiquetado del isótopo estable con el reactivo 8-plex iTRAQ y 2D-LC MALDI MS/MS para la identificación y cuantificación del péptido. La cuantificación con esta técnica es relativa: se atribuyen a los péptidos u proteínas relaciones de abundancia con respecto a la muestra de referencia. Las muestras de referencia comunes en múltiples experimentos iTRAQ facilitan la comparación de muestras a lo largo de múltiples experimentos iTRAQ.

Para implementar este esquema de análisis, se combinan seis muestras primarias y dos muestras de agrupación testigo en una mezcla 8-plex iTRAQ, estando las muestras de agrupación testigo etiquetadas con los reactivos 113 y 117 de acuerdo con las sugerencias del fabricante. A continuación está mezcla de ocho muestras se fracciona mediante cromatografía líquida bidimensional; intercambio catiónico intenso (SCX) en la primera dimensión y HPLC de fase inversa en la segunda dimensión. El eluyen de HPLC se fracciona directamente en placas MALDI y las

placas se analizan en un espectrómetro de masas MDS SCIEX/AB 4800 MALDI TOF/TOF.

En ausencia de información adicional, se supone que los cambios más importantes en la expresión proteica son aquellos dentro de los mismos tipos celulares en diferentes condiciones de tratamiento. Por esta razón, las muestras primarias de Célula No. 1 y Célula No. 2 se analizan en mezclas de iTRAQ separadas. Para facilitar la comparación de la expresión proteica en las muestras de Célula No. 1 con respecto a la Célula No. 2, se analizan muestras QCP universales en las "ranuras de iTRAQ" disponibles no ocupadas por muestras QC primarias o de línea celular específica (QC1 y QC2). En la presente se proporciona una descripción general breve de los procedimientos de laboratorio empleados.

A. <u>Extracción de proteínas a partir de muestras de sobrenadante celular</u>

Para muestras de sobrenadante celular (CSN), las proteínas del medio de cultivo están presentes en gran exceso con respecto a las proteínas secretadas por las células cultivadas. En un intento por reducir este antecedentes, se implementó una reducción de proteína abundante por anticipado. Dado que no existen disponibles columnas de afinidad específicas para proteínas séricas de bovino ni caballo, se usó una columna de anti-IgY14 humana. Aunque estos anticuerpos se dirigen contra proteínas humanas, la amplia especificidad proporcionada por la

naturaleza policlonal de los anticuerpos se anticipó para lograr la reducción de las proteínas bovinas y equinas presentes en los medios de cultivo celular que se usaron.

Se cargó una alícuota de 200µl del material de CSN QC en una columna de reducción de IgY14 de 10 mL antes del inicio del estudio para determinar la concentración proteica total (ensayo con ácido bicinconínico (BCA)) en el material de flujo. A continuación se seleccionó el volumen de carga para lograr una fracción reducida que contuviese aproximadamente 40 µg de proteína total.

B. Extracción de proteínas a partir de granulados celulares

Se lisó una alícuota de Célula No. 1 y Célula No. 2 en la solución amortiguadora de lisis "estándar" usada para el análisis de las muestras de tejido en BGM, y se determinó el contenido total de proteína mediante el ensayo BCA. Luego de establecer el contenido proteico de estos lisados celulares representativos, todas las muestras de granulado celular (incluidas las muestras QC descritas en la Sección 1.1) se procesaron hasta obtener lisados celulares. Cantidades de lisado de aproximadamente 40 µg de proteína total se hicieron avanzar en el flujo de proceso.

C. Preparación de muestras para espectrometría de masas

La preparación de muestras sigue los procedimientos de operación estándares y constituyen los siguientes:

- Reducción y alquilación de proteínas

•Limpieza de proteínas en columna de fase inversa (solo granulados celulares)

•Digestión con tripsina

•Etiquetado con iTRAQ

•Cromatografía de intercambio catiónico intenso - recolección de seis fracciones (sistema Agilent 1200)

•Fraccionamiento mediante HPLC y manchado en placas MALDI (sistema Dionex Ultimate3000/Probot)

D. MALDI MS y MS/MS

La HPLC-MS generalmente emplea estrategias de ESI MS/MS en línea. BG Medicine usa una plataforma LC-MALDI MS/MS fuera de línea que resulta en una mejor concordancia de conjuntos de proteína observados en las muestras primarias sin la necesidad de inyectar la misma muestra múltiples veces. Luego de la primera recolección de datos en todas las mezclas iTRAQ, dado que las fracciones de péptidos se retuvieron en las placas MALDI objetivo, las muestras pueden analizarse una segunda vez usando un patrón de adquisición MS/MS dirigido derivado del conocimiento obtenido durante la primera adquisición. De esta forma, se logra la frecuencia de observación máxima para todas las proteínas identificadas (idealmente, cada proteína debería medirse en cada mezcla iTRAQ).

E. Procesamiento de datos

Los proceso de procesamiento de datos dentro del flujo de trabajo de BGM Proteomics pueden separarse en los procedimientos tales como identificación y cuantificación preliminar de péptidos que

se completan para cada mezcla iTRAQ individualmente (Sección 1.5.1) y aquellos procesos (Sección 1.5.2) tales como atribución final de péptidos a proteínas y cuantificación final de proteínas, que no se completan hasta que no se completa la adquisición de datos para el proyecto.

Las principales etapas de procesamiento de datos dentro del flujo de trabajo de BGM Proteomics son:

•Identificación de péptidos usando el buscador de base de datos Mascot (Matrix Sciences)

•Validación interna automatizada de Mascot IDs

•Cuantificación de péptidos y cuantificación preliminar de proteínas

•Depuración por experto de conjunto de datos final

•Atribución final de péptidos de cada mezcla en un conjunto común de proteínas usando la herramienta PVT automatizada

•Eliminación de valores atípicos y cuantificación final de proteínas

(i) Procesamiento de datos de mezclas iTRAQ individuales

A medida que cada mezcla iTRAQ se procesa a través del flujo de trabajo los espectros MS/MS se analizan usando herramientas informáticas BGM patentadas para identificaciones de péptidos y proteínas, así como una evaluación inicial de la información de cuantificación. En base a los resultados de este análisis preliminar, la calidad del flujo de trabajo para cada muestra primaria en la mezcla se juzga con respecto a un conjunto de

métricas de rendimiento de BGM. Si una muestra dada (o mezcla) no aprueba las métricas de rendimiento mínimas especificadas, y existe material adicional disponible, dicha muestra se repite en su totalidad y son los datos de esta segunda implementación del flujo de trabajo los que se incorporan en el conjunto de datos final.

(ii) Identificación de péptidos

Los espectros de MS/MS se sometieron a búsqueda con respecto a la base de datos de secuencias proteicas Uniprot que contenía secuencia humanas, bovinas y equinas incrementadas mediante secuencia contaminantes comunes tales como tripsina porcina. Los detalles de los parámetros de búsqueda Mascot, incluida la lista completa de modificaciones, se muestran en la Tabla 3.

Tabla 3: Parámetros de búsqueda Mascot

Tolerancia de la masa del precursor	100 ppm
Tolerancia de la masa del fragmento	0,4 Da
Modificaciones variables	ITRAQ8 extremo N ITRAQ8 lisina Cys carbamidometil Pyro-Glu (extremo N) Pyro-Carbamidometil Cys (extremo N) Desamidación (solo N) Oxidación (M)

Especificidad enzimática	Completamente tríptica
Cantidad de sitios trípticos perdidos permitidos	2
Rango peptídico considerado	1

Después de finalizada la búsqueda Mascot, se usa un procedimiento de autovalidación para promover *(es decir,* validar) las coincidencias peptídicas específicas de Mascot. La diferenciación entre coincidencias válidas y no válidas se basa en el puntaje Mascot alcanzado con respecto al puntaje Mascot esperado y la diferencia entre los puntajes de Mascot de péptidos de Rango 1 y péptidos de Rango 2. Los criterios exigidos para la validación se suavizan un tanto si el péptido es uno de los varios coincidentes con una única proteína en el mezcla iTRAQ o si el péptido está presente en un catálogo de péptidos validados anteriormente.

(iii) Cuantificación de péptidos y proteínas

El conjunto de péptidos validados para cada mezcla se utiliza para calcular las métricas de cuantificación proteica preliminar para cada mezcla. Las relaciones de los péptidos se calculan dividiendo el área pico de la etiqueta iTRAQ *(es decir, m/z* 114, 115, 116, 118, 119 o 121) para cada péptido validado entre la

mejor representación del área pico de la agrupación de referencia (QC1 o QC2). Esta área de pico es el promedio de los picos 113 y 117 siempre que ambas muestras aprueben los criterios de aceptación QC. Las relaciones de proteína preliminares se determinan calculando la relación media de todos los péptidos validados "útiles" que coinciden con dicha proteína. Los péptidos "útiles" se etiquetan completamente con iTRAQ (todos los extremos N se etiquetan con Lisina o PiroGlu) y se etiquetan completamente con Cisteína *(por ejemplo.,* todos los residuos Cys se alquilan con Carbamidometil o Pyro-cmc en el extremo N).

(iv) Procesamiento de posadquisición

Una vez que todas las pasadas de adquisición de datos MS/MS se completan para cada mezcla en el proyecto, los datos se recogen usando las tres etapas mencionadas a continuación que tienen como fin facilitar que los resultados para cada muestra primaria se comparen de forma simple y significativa con los de otra.

(v) Atribución global de secuencias peptídicas a proteínas

La distribución final de secuencias peptídicas a números de acceso a proteínas se lleva a cabo a través de la Protein Validation Tool (PVT) (Herramienta de validación proteica) patentada. El procedimiento PVT determina el mejor conjunto de proteína no redundante mínimo para describir la colección entera de péptidos identificados en el proyecto. Este es un procedimiento automatizado que ha sido optimizado para manipular datos de una taxonomía homogénea.

Las atribuciones de proteínas para los experimentos con sobrenadante se depuraron manualmente para tratar las complejidades de taxonomías mixtas en la base de datos. Dado que el paradigma automatizado no es válido para cultivos celulares cultivados en medios complementados con suero bovino y equino, se necesita depuración manual extensiva para minimizar la ambigüedad de la fuente de cualquier proteína dada.

(vi) Normalización de relaciones peptídicas

Las relaciones peptídicas para cada muestra se normalizan en base al método de Vandesompele *et al. Genome Biology,* **2002,** *3(7),* research 0034.1-11. Este procedimiento se aplica solamente a las mediciones de granulado celular. Para las muestras de sobrenadante,

los datos cuantitativos no se normalizan teniendo en cuenta que la mayor contribución a las identificación de péptidos viene de los medios.

(vii) Cálculo final de relaciones proteicas

Se utiliza un procedimiento de eliminación de valores atípicos estadística estándar para eliminar los valores atípicos alrededor de cada relación media de proteínas, más allá del nivel 1,96 σ en el conjunto de datos transformado por logaritmo. Luego del proceso de eliminación, el conjunto final de relaciones proteicas se (vuelve a) calcula.

VI. Marcadores de la invención y usos de los mismos

La presente invención se basa, al menos en parte, en la identificación de nuevos biomarcadores que se asocian con un sistema biológico, tal como un proceso de enfermedad o respuesta de una sistema biológico a una perturbación, tal como una agente terapéutico.

En especial0, la invención se refiere a marcadores (de aquí en adelante "marcadores" o "marcadores de la invención"), que se describen en los ejemplos. La invención proporciona ácidos nucleicos y proteínas que se codifican o corresponden a los marcadores (de aquí en adelante "ácidos nucleicos marcadores" y "proteínas marcadoras" respectivamente). Estos marcadores son particularmente útiles para el diagnósticos de enfermedades; pronóstico de enfermedades; desarrollo de objetivos de fármacos para varias enfermedades; detección de la presencia de toxicidad, preferiblemente toxicidad inducida por fármacos, por ejemplo, cardiotoxicidad; identificar un agente que provoca o está en riesgo de provocar toxicidad; identificar un agente que puede reducir o prevenir la toxicidad inducida por fármacos; aliviar, reducir o prevenir la cardiotoxicidad inducida por fármaco; e identificar marcadores predictores de cardiotoxicidad inducida por fármacos.

Un "marcador" es un gen cuyo nivel de expresión alterado en un tejido o célula con respecto a su nivel de expresión en tejidos o células normales o sanos se asocia con un estado de enfermedad tal como cáncer, diabetes, obesidad, enfermedad cardiovascular o

un estado de toxicidad, tal como toxicidad inducida por fármaco, por ejemplo, cardiotoxicidad. Un "ácido nucleico marcadores" es un ácido nucleico *(por ejemplo,* ARNm, ADNc) codificado o correspondiente a un marcador de la invención. Dichos ácidos nucleicos marcadores incluyen ADN *(por ejemplo,* ADNc) que comprende la secuencia total o parcial de cualquiera de los genes que son marcadores de la invención o el complemento de dicha secuencia. Los expertos en la técnica conocen dichas secuencias y se pueden encontrar, por ejemplo, en el sitio web pubmed gubernamental de NIH. Los ácidos nucleicos marcadores también incluyen ARNque comprende la secuencia total o parcial de cualquier a de los marcadores génicos de la invención o el complemento de dicha secuencia, en todos los residuos de timidina se reemplazan por residuos de uridina. Una "proteína marcadora" es una proteína codificada o correspondiente a un marcador de la invención. Una proteína marcadora comprende la secuencia total o parcial de cualquiera de las proteínas marcadoras de la invención. Los expertos en la técnica conocen dichas secuencias y se pueden encontrar, por ejemplo, en el sitio web pubmed gubernamental de NIH. Los términos "proteína" y "polipéptido" se usan de forma intercambiable.

Un fluido corporal "asociado con un estado de enfermedad o estado tóxico" es un fluido que, cuando está en el cuerpo del paciente, hace contacto o pasa a través de las células de sarcoma o hacia las células o proteínas cubiertas por células de sarcoma que es

capaz de pasar. Los fluidos corporales asociados con un estado de enfermedad o estado tóxico ejemplares incluyen fluidos corporales *(por ejemplo* sangre completa, suero sanguíneo, sangre de la que se retiraron las plaquetas) y se describen en más detalle a continuación. Los fluidos corporales asociados con un estado de enfermedad o estado tóxico no se limitan a sangre completa, sangre de la que se retiraron las plaquetas, linfa, fluido prostático, orina y semen.

El nivel de expresión "normal" de un marcador es el nivel de expresión del marcador en células de un sujeto humano o paciente que no padece un estado de enfermedad o un estado de toxicidad.

Una "sobreexpresión" o "nivel de expresión más alto" de un marcador se refiere a un nivel de expresión en una muestra de prueba que es mayor que el error estándar del ensayo empleado para evaluar la expresión, y es preferiblemente al menos el doble y más preferiblemente tres, cuatro, cinco, seis, siete, ocho, nueve o diez veces mayor que el nivel de expresión del marcador en una muestra testigo *(por ejemplo,* muestra de un sujetos sano que no tiene el marcador asociado con un estado de enfermedad o un estado de toxicidad, *por ejemplo,* cáncer, diabetes, obesidad, enfermedad cardiovascular y cardiotoxicidad) y preferiblemente, el nivel de expresión promedio del marcador en varias muestras testigo.

Un "nivel de expresión más bajo" de un marcador se refiere a un nivel de expresión en una muestra de prueba que al menos el doble

y más preferiblemente tres, cuatro, cinco, seis, siete, ocho, nueve o diez veces más bajo que el nivel de expresión del marcador en una muestra testigo *(por ejemplo,* muestra de un sujetos sano que no tiene el marcador asociado con un estado de enfermedad o un estado de toxicidad, *por ejemplo,* cáncer, diabetes, obesidad, enfermedad cardiovascular y cardiotoxicidad) y preferiblemente, el nivel de expresión promedio del marcador en varias muestras testigo.

Un "polinucleótido transcrito" o "transcrito nucleotídico" es un polinucleótido *(por ejemplo,* un ARNm, ARNnh, un ADNc o un análogo de dicho ARN o ADNc) que escomplementario u homólogo con todo o una porción de un ARNm maduro hecho mediante transcripción de un marcador de la invención y procesamiento postranscripcional normal *(por ejemplo,* empalme), si lo hay, del transcrito de ARN y transcripción inversa del transcrito de ARN.

"Complementario" se refiere al concepto amplio de complementariedad secuencial entre regiones de dos hebras de ácido nucleico o entre dos regiones de la misma hebra de ácido nucleico. Se sabe que un residuo de adenina de una primera región de ácido nucleico es capaz de formar enlaces de hidrógeno específicos ("pares de bases") con un residuo de una segunda región de ácido nucleico que es antiparalela a la primera región si el residuo es timina o uracilo. De forma similar, se sabe que un residuo de citosina de una primera hebra de ácido nucleico es capaz de formar pares de bases con un residuo de una segunda

hebra de ácido nucleico que es antiparalela a la primera hebra si el residuo es guanina. Una primera región de un ácido nucleico es complementaria a una segunda región del mismo o un ácido nucleico diferente si, cuando las dos regiones se disponen de forma antiparalela, al menos un residuo nucleotídico de la primera región es capaz de formar un par de bases con un residuo de la segunda región. Preferiblemente, la primera región comprende una primera porción y la segunda región comprende una segunda porción, por las cuales, cuando las primera y segunda porciones se disponen de forma antiparalela, al menos aproximadamente 50%, y preferiblemente al menos aproximadamente 75%, al menos aproximadamente 90%, o al menos aproximadamente 95% de los residuos nucleotídicos de la primera porción son capaces de formar pares de bases con los residuos nucleotídicos en la segunda porción. Más preferiblemente, todos los residuos nucleotídica de la primera porción son capaces de formar pares de bases con residuos nucleotídicos en la segunda porción.

"Homólogo" tal como se usa en la presente, se refiere a la similitud de secuencia de nucelótidos entre dos regiones de la misma hebra de ácido nucleico o entre regiones de dos hebras de ácido nucleico diferentes. Cuando una posición de residuo nucleotídico en ambas regiones está ocupada por el mismo residuo nucleotídico, entonces las regiones son homólogas en dicha posición. Una primera región es homóloga con una segunda región si al menos una posición de residuo nucleotídico de cada región

está ocupada por el mismo residuo. La homología entre dos regiones se expresa en términos de la proporción de posiciones de residuos nucleotídicos de las dos regiones que están ocupadas por el mismo residuo nucleotídico. A modo de ejemplo, una región que tiene la secuencia de nucleótidos 5'-ATTGCC-3' y una región que tiene la secuencia de nucleótidos 5'-TATGGC-3' comparten un 50% de homología. Preferiblemente, la primera región comprende una primera porción y la segunda región comprende una segunda porción, por las cuales, al menos aproximadamente 50%, y preferiblemente al menos aproximadamente 75%, al menos aproximadamente 90%, o al menos aproximadamente 95% de las posicionesde residuos nucleotídicos de cada una de las porciones están ocupadas por el mismo residuo nucleotídico. Más preferiblemente, todas las posiciones de residuos nucleotídicos de cada una de las porciones están ocupadas por el mismo residuo nucleotídico.

"Proteínas de la invención" abarca proteínas marcadoras y fragmentos de las mismas; variantes de proteínas marcadoras y sus fragmentos; péptidos y polipéptidos que comprenden al menos un segmento de 15 aminoácidos de una proteína marcadora o variante de proteína marcadora; y proteínas de fusión que comprenden un una proteína marcadora o variante de proteína marcadora, o un segmento de al menos 15 aminoácidos de una proteína marcadora o variante de proteína marcadora.

La invención además proporciona anticuerpos, derivados de anticuerpos y fragmentos de anticuerpos que se unen específicamente con las proteínas marcadores y fragmentos de las proteínas marcadoras de la presente invención. A menos que se especifique los contrario en la presente, los términos "anticuerpo" y "anticuerpos" abarcan ampliamente formar de anticuerpos de origen natural *(por ejemplo,* IgG, IgA, IgM, IgE) y anticuerpos recombinantes tales como anticuerpos monocatenarios, anticuerpos quiméricos y humanizados y anticuerpos multiespecíficos, así como fragmentos y derivados de todos los precedentes, cuyos fragmentos y derivados tienen al menos un sitio de unión antigénico. Los derivados de anticuerpos puede comprender una proteína o resto químico conjugado con un anticuerpo.

En ciertas realizaciones, los marcadores de la invención incluyen uno o más genes (o proteínas) seleccionados del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1, CANX, GRP78, GRP75, TIMP1, PTX3, HSP76, PDIA4, PDIA1, CA2D1, GPAT1 y TAZ. En algunas realizaciones, los marcadores son una combinación de al menos dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez, once, doce, trece, catorce, quince, dieciséis, diecisiete, dieciocho, diecinueve, veinte, veinticinco, treinta o más de los genes (o proteínas) precedentes. Todos los valores presentados en la lista

precedente también pueden ser los límites superiores o inferiores de rangos, que se pretenden que sean parte de esta invención, *por ejemplo,* entre 1 y 5, 1 y 10, 1 y 20, 1 y 30, 2 y 5, 2 y 10, 5 y 10, 1 y 20, 5 y 20, 10 y 20, 10 y 25, 10 y 30 de los genes (o proteínas) precedentes.

En una realización, los marcadores de la invención son genes o proteínas asociados o que participan en el cáncer. Dichos genes o proteínas que participan en el cáncer incluyen, por ejemplo, HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9,PARP1, HADHA, PHB2, ATP5A1, y/o CANX. En algunas realizaciones, los marcadores de la invención son una combinación de al menos dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez, once, doce, trece, catorce, quince, dieciséis, diecisiete, dieciocho, diecinueve, veinte o más de los genes (o proteínas) precedentes. Todos los valores presentados en la lista precedente también pueden ser los límites superiores o inferiores de rangos, que se pretenden que sean parte de esta invención, *por ejemplo,* entre 1 y 5, 1 y 10, 1 y 20, 1 y 30, 2 y 5, 2 y 10, 5 y 10, 1 y 20, 5 y 20, 10 y 20, 10 y 25, 10 y 30 de los genes (o proteínas) precedentes.

En una realización, los marcadores de la invención son genes o proteínas asociados o que participan en la toxicidad inducida por fármacos. Dichos genes o proteínas que participan en la toxicidad inducida por fármacos incluyen, por ejemplo, GRP78, GRP75, TIMP1,

PTX3, HSP76, PDIA4, PDIA1, CA2D1, GPAT1 y/o TAZ. En algunas realizaciones, los marcadores de la invención son una combinación de al menos dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez de los genes (o proteínas) precedentes. Todos los valores presentados en la lista precedente también pueden ser los límites superiores o inferiores de rangos, que se pretenden que sean parte de esta invención, *por ejemplo,* entre 1 y 5, 1 y 10, 1 y 20, 1 y 30, 2 y 5, 2 y 10, 5 y 10, 1 y 20, 5 y 20, 10 y 20, 10 y 25, 10 y 30 de los genes (o proteínas) precedentes.

A. Marcadores asociados con cardiotoxicidad

La presente invención se basa, al menos en parte, en la identificación de biomarcadores nuevos que se asocian con la cardiotoxicidad inducida por fármacos. La invención se basa adicionalmente, al menos en parte, en el descubrimiento de que la Coenzima Q10 es capaz de reducir o prevenir la cardiotoxicidad inducida por fármacos.

Por consiguiente, la invención proporciona métodos para identificar un agente que provoca o está en riesgo de provocar toxicidad. En una realización, el agente es una fármaco o fármaco candidato. En una realización, la toxicidad es toxicidad inducida por fármacos, por ejemplo, cardiotoxicidad. En una realización, el agente es un fármaco o fármaco candidato para tratar la diabetes, obesidad o un trastorno cardiovascular. En estos métodos, se evalúa la cantidad de uno o más biomarcadores/proteínas en un par de muestras (una primera

muestra no sometida al tratamiento con fármaco y una segunda muestra sometida al tratamiento con fármaco). Una modulación en el nivel de expresión de los uno o más biomarcadores en la segunda muestra en comparación con la primera muestra es una indicación de que el fármaco provoca o está en riesgo de provocar toxicidad inducida por fármaco, por ejemplo, cardiotoxicidad. En una realización, los uno o más biomarcadores se seleccionan del grupo que consiste en GRP78, GRP75, TIMP1, PTX3, HSP76, PDIA4, PDIA1, CA2D1, GPAT1 y TAZ. Los métodos de la presente invención pueden ponerse en práctica conjuntamente con cualquier otro método usado por un médico experto para identificar un fármaco en riesgo de provocar cardiotoxocidad inducida por fármaco.

Por consiguiente, en un aspecto, la invención proporciona un método para identificar un fármaco que provoca o está en riesgo de provocar toxicidad inducida por fármaco (por ejemplo, cardiotoxicidad), que comprende: comparar (i) el nivel de expresión de uno o más biomarcadores presentes en una primera muestra celular obtenida antes del tratamiento con el fármaco; con (ii) el nivel de expresión de uno o más biomarcadores presentes en una segunda muestra celular obtenida después del tratamiento con el fármaco; en donde los uno o más biomarcadores se seleccionan del grupo que consiste en GRP78, GRP75, TIMP1, PTX3, HSP76, PDIA4, PDIA1, CA2D1, GPAT1 y TAZ; en donde la modulación en el nivel de expresión de uno o más biomarcadores en

la segunda muestra en comparación con la primera muestra es una indicación de que el fármaco provoca o está en riesgo de provocar toxicidad inducida por fármaco (por ejemplo, cardiotoxicidad).

En una realización, la toxicidad inducida por fármaco es cardiotoxicidad inducida por fármaco. En una realización, las células son células del sistema cardiovascular, por ejemplo, cardiomiocitos. En una realización, las células son cardiomiocitos diabéticos. En una realización, el fármaco es un fármaco o fármaco candidato para tratar la diabetes, obesidad o una enfermedad cardiovascular.

En una realización, la modulación (por ejemplo, un aumento o disminución) en el nivel de expresión de uno, dos, tres, cuatro, cinco, seis, siete, ocho, nueve o todos los diez biomarcadores seleccionados del grupo que consiste en GRP78, GRP75, TIMP1, PTX3, HSP76, PDIA4, PDIA1, CA2D1, GPAT1 y TAZ en la segunda muestra en comparación con la primera muestra es una indicación de que el fármaco provoca o está en riesgo de provocar toxicidad inducida por fármaco.

Los métodos para identificar un agente que puede reducir p prevenir la toxicidad inducida por fármaco también se proporcionan en la invención. En una realización, la toxicidad inducida por fármaco es cardiotoxicidad. En una realización, el fármaco es un fármaco o fármaco candidato para tratar la diabetes, obesidad o un trastorno cardiovascular. En estos métodos, se evalúa la cantidad de uno o más biomarcadores en tres

muestras (una primera muestra no sometida al tratamiento con fármaco, una segunda muestra sometida al tratamiento con el fármaco y una tercera muestra sometida al tratamiento con el fármaco y al agente). Aproximadamente el mismo nivel de expresión de los uno o más biomarcadores en la tercera muestra en comparación con la primera muestra es una indicación de que el agente puede reducir o prevenir la toxicidad inducida por fármaco, por ejemplo,cardiotoxicidad inducida por fármaco. En una realización, los uno o más biomarcadores se seleccionan del grupo que consiste en GRP78, GRP75, TIMP1, PTX3, HSP76, PDIA4, PDIA1, CA2D1, GPAT1 y TAZ.

Usando los métodos descritos en la presente, se puede someter a detección una variedad de moléculas, particularmente incluidas moléculas suficientemente pequeñas para ser capaces de cruzar la membrana celular, a efecto de identificar moléculas que modulan, *por ejemplo,* aumentan o disminuyen la expresión y/o actividad de un marcador de la invención. Los compuestos identificados de esta forma pueden proporcionarse a un sujeto a efectos de reducir, aliviar o prevenir la toxicidad inducida por fármaco en el sujeto.

Por consiguiente, en otro aspecto, la invención proporciona un método para identificar un agente que puede reducir o prevenir la toxicidad inducida por fármaco, que comprende: (i) determinar el nivel de expresión de uno o más biomarcadores presentes en una primera muestra celular obtenida antes del tratamiento con un

fármaco que induce la toxicidad; (ii) determinar el nivel de expresión de uno o más biomarcadores presentes en una segunda muestra celular obtenida luego del tratamiento con el fármaco que induce la toxicidad; (iii) determinar el nivel de expresión de uno o más biomarcadores presentes en una tercera muestra celular obtenida después del tratamiento con el fármaco que induce la toxicidad y el agente; y (iv) comparar el nivel de expresión de los uno o más biomarcadores presentes en la tercera muestra con la primera muestra; en donde el uno más biomarcadores se seleccionan del grupo que consiste en GRP78, GRP75, TIMP1, PTX3, HSP76, PDIA4, PDIA1, CA2D1, GPAT1 y TAZ; y en donde aproximadamente el mismo nivel de expresión de los uno o más marcadores en la tercera muestra en comparación con la primera muestra es una indicación de que el agente puede reducir o prevenir la toxicidad inducida por fármaco.

En una realización, la toxicidad inducida por fármaco es cardiotoxicidad inducida por fármaco. En una realización, las células son células del sistema cardiovascular, por ejemplo, cardiomiocitos. En una realización, las células son cardiomiocitos diabéticos. En una realización, el fármaco es un fármaco o fármaco candidato para tratar la diabetes, obesidad o una enfermedad cardiovascular.

En una realización, aproximadamente el mismo nivel de expresión de uno, dos, tres, cuatro, cinco, seis, siete, ocho, nueve o todos los diez biomarcadores seleccionados del grupo que consiste

en GRP78, GRP75, TIMP1, PTX3, HSP76, PDIA4, PDIA1, CA2D1, GPAT1 y TAZ en la tercera muestra en comparación con la primera muestra es una indicación de que el agente puede reducir o provocar la toxicidad inducida por fármaco.

La invención además proporciona métodos para aliviar, reducir o prevenir la cardiotoxicidad inducida por fármaco en un sujeto que los necesita, que comprenden administrar a un sujeto (por ejemplo, un mamífero, un humano o un animal no humano) un agente identificado mediante los métodos de detección proporcionados en la presente, reduciendo o previniendo así la cardiotoxicidad inducida por fármaco en el sujeto. En una realización, el agente se administra a un sujeto que ya ha sido tratado con un fármaco que induce cardiotoxicidad. En una realización, el agente se administra a un sujeto al mismo tiempo que el tratamiento de un sujeto con un fármaco que induce cardiotoxicidad. En una realización, el agente se administra a un sujeto antes del tratamiento del sujeto con un fármaco que induce cardiotoxicidad.

La invención además proporciona métodos para aliviar, reducir o prevenir la cardiotoxicidad inducida por fármaco en un sujeto que los necesita, que comprenden administrar Coenzima Q10 a un sujeto (por ejemplo, un mamífero, un humano o un animal no humano), reduciendo o previniendo así la cardiotoxicidad inducida por fármaco en el sujeto. En una realización, la Coenzima Q10 se administra a un sujeto que ya ha sido tratado con un fármaco que induce cardiotoxicidad. En una realización, la Coenzima Q10 se

administra a un sujeto al mismo tiempo que el tratamiento del sujeto con un fármaco que induce la cardiotoxicidad. En una realización, la Coenzima Q10 se administra a un sujeto antes del tratamiento del sujeto con un fármaco que induce la cardiotoxicidad. En una realización, la cardiotoxicidad se asocia con la modulación de la expresión de uno, dos, tres, cuatro, cinco, seis, siete, ocho, nueve o todos los diez biomarcadores seleccionados del grupo que consiste en GRP78, GRP75, TIMP1, PTX3, HSP76, PDIA4, PDIA1, CA2D1, GPAT1 y TAZ. Todos los valores presentados en la lista precedente también pueden ser el límite superior o inferior de rangos, que se pretende que sean parte de la presente invención, *por ejemplo,* entre 1 y 5, 1 y 10, 2 y 5, 2 y 10 o 5 y 10 de los genes (o proteínas) precedentes.

La invención además proporciona biomarcadores (por ejemplo, genes y/o proteínas) que son útiles como marcadores predictores para la cardiotoxicidad, por ejemplo, cardiotoxicidad inducida por fármaco. Estos biomarcadores incluyen GRP78, GRP75, TIMP1, PTX3, HSP76, PDIA4, PDIA1, CA2D1, GPAT1 y TAZ. Sin embargo, el experto en la técnica será capaz de identificar biomarcadores predictivos adicionales de cardiotoxicidad inducida por fármaco empleado los métodos descritos en la presente, por ejemplo, llevar a cabo los métodos descritos en el Ejemplo 3, pero usando un fármaco diferente conocido por inducir cardiotoxicidad. Los biomarcadores de cardiotoxicidad inducida por fármaco de la invención se describen en más detalle a continuación.

GRP78 y GRP75 también se denominan proteínas de respuesta a glucosa. Estas proteínas se asocian con el estrés del retículo endo/sarcoplasmático (estrés de ER) de cardiomiocitos. SERCA, o ATPasa de calcio del retículo sarcoendoplasmático, regula la homeostasis de Ca2+ en las células cardíacos. Cualquier interrupción de estas ATPasas puede conducir a una disfunción cardíaca y a insuficiencia cardíaca. En base a los datos proporcionados en la presente, GRP75 y GRP78 los bordes alrededor de los mismos son nuevos predictores de cardiotoxicidad inducida por fármaco.

TIMP1, también denominada inhibidor de metaloproteasa TIMP 1, participa en la remodelación de la matriz celular extra en asociación con MMP. La expresión de TIMP1 se correlaciona con la fibrosis del corazón y la hipoxia de células endoteliales vasculares también induce la expresión de TIMP1. En base a los datos proporcionados en la presente, TIMP1 es un nuevo predictor de cardiotoxicidad inducida por fármaco.

PTX3, también denominado Pentraxina 3, pertenece a las familia de las proteínas C reactivas (CRP) y es un buen marcador de una afección inflamatoria del corazón. Sin embargo, el PTX3 en plasma también podría ser representativo de respuesta inflamatoria sistémica debido a sepsis u otras afecciones médicas. En base a los datos proporcionados en la presente, PTX3 puede ser un nuevo marcador de la función cardíaca o cardiotoxicidad.

Adicionalmente, los bordes asociados con PTX 3 en la red podrían formar un nuevo panel de biomarcadores.

HSP76, también denominado HSPA6, solo se conoce como expresado en células endoteliales y linfocitos B. No existe un papel conocido para esta proteína en la función cardíaca. En base a los datos proporcionados en la presente, HSP76 es un nuevo predictor de cardiotoxicidad inducida por fármaco.

PDIA4, PDIA1, también denominados proteínas de la familia de proteínas disulfuro isomerasa A, se asocian con la respuesta de estrés de ER, al igual que GRP. No existe un papel conocido para estas proteínas en la función cardíaca. En base a los datos proporcionados en la presente, estas proteínas puede ser nuevos predictores de cardiotoxicidad inducida por fármaco.

CA2D1 también denominada subunidad de canal de calcio dependiente del voltaje alfa 2/delta. La subunidad de canal de calcio dependiente del voltaje alfa 2/delta regula la densidad de la corriente de calcio y la activación/desactivación de la cinética del canal de calcio. CA2D1 tiene un papel importante en el acoplamiento de excitación-contracción en el corazón. No existe un papel conocido para esta proteína en la función cardíaca. En base a los datos proporcionados en la presente, CA2D1 es un nuevo predictor de cardiotoxicidad inducida por fármaco.

GPAT1 es una de cuatro isoformas de glicerol-3-fosfato aciltransferasa conocidas, y se ubica en la membrana externa mitocondrial, permitiendo la regulación recíproca concarnitina

palmitoiltransferasa-1. GPAT1 se regula de forma ascendente transcripcionalmente mediante insulina y SREBP-1c y se regula de forma descendente intensamente mediante la proteína cinasa activada por AMP, de forma consistente con su rol en la síntesis de triacilglicerol. En base a los datos proporcionados en la presente, GPAT1 es un nuevo predictor de cardiotoxicidad inducida por fármaco.

TAZ, también denominado Tafazzina, se expresa altamente en el músculo cardíaco y esquelético. TAZ participa en el metabolismo de cardiolipina y funciona como una fosfolípido-lisofosfolípido transacilasa. La tafazzina es responsable por remodelan la cardiolipina fosfolipídica (CL), el lípido distintivo de la membrana interna mitocondrial. En base a los datos proporcionados en la presente, TAZ es un nuevo predictor de cardiotoxicidad inducida por fármaco.

B. <u>Marcadores asociados con el cáncer</u>

La presente invención se basa, al menos en parte, en la identificación de biomarcadores novedosos que están asociados con el cáncer. Dichos marcadores asociados con el cáncer incluyen, por ejemplo, HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1, y/o CANX. En algunas realizaciones, los marcadores de la invención son una combinación de al menos dos, tres, cuatro, cinco, seis, siete, ocho, nueve,

diez, once, doce, trece, catorce, quince, dieciséis, diecisiete, dieciocho, diecinueve, veinte y más de los antemencionados marcadores.

Por consiguiente, la invención proporciona métodos para identificar un agente que causa o está en riesgo de causar cáncer. En una realización, el agente es un fármaco o candidato a fármaco. En estos métodos, la cantidad de uno o más biomarcadores/proteínas en un par de muestras (una primera muestra sometida al tratamiento de fármaco y una segunda muestra sometida al tratamiento de fármaco) se somete a análisis. Una modulación en el nivel de expresión del uno o más biomarcadores en la segunda muestra, comparado con la primera muestra, es un indicador de que el fármaco causa o está en riesgo de causar cáncer. En una realización, el uno o más biomarcadores se seleccionan del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1, y CANX. Los métodos de la presente invención pueden llevarse a cabo en conjunto con cualquier otro método usado por el experto en la técnica para identificar un fármaco en riesgo de causar cáncer.

En un aspecto, la invención proporciona métodos para analizar le eficacia de una terapia para tratar un cáncer en un

sujeto, el método comprende: comparar el nivel de expresión de uno o más marcadores presentes en una primera muestra obtenida de un sujeto antes de administrarle al menos una parte del régimen de tratamiento al sujeto, donde el uno o más marcadores se selecciona del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1 y CANX; y el nivel de expresión del uno o más marcadores presentes en una segunda muestra obtenida de un sujeto después de la administración de al menos una parte del régimen de tratamiento, donde una modulación en el nivel de expresión del uno o más marcadores en la segunda muestra, comparado con la primera muestra es un indicador de que la terapia es eficaz para el tratamiento del cáncer en el sujeto.

En una realización, la muestra comprende un fluido obtenido del sujeto. En una realización, el fluido se selecciona del grupo que consiste en fluidos sanguíneos, vómito, saliva, linfa, fluido quístico, orina, fluidos recogidos de un lavado bronquial, fluidos recogidos del enjuague peritoneal y fluidos ginecológicos. En una realización, la muestra es una muestra de sangre o un componente de esta.

En otra realización, la muestra comprende un tejido o un componente de este, obtenido del sujeto. En una realización, el tejido se selecciona del grupo que consiste en óseo, tejido

conectivo, cartílago, pulmonar, hepático, renal, tejido muscular, cardíaco, pancreático y cutáneo.

En una realización el sujeto es un ser humano.

En una realización, el nivel de expresión del uno o más marcadores en la muestra biológica se determina mediante el análisis de un polinucleótido transcripto o una porción de este en la muestra. En una realización, el análisis del polinucleótido transcripto comprende amplificar el polinucleótido transcripto.

En una realización, el nivel de expresión del marcador en el sujeto se determina mediante el análisis de una proteína o una porción de esta en la muestra. En una realización, la proteína se analiza usando un reactivo que se une específicamente con la proteína.

En una realización, el nivel de expresión del uno o más marcadores en la muestra se determina usando una técnica que se selecciona del grupo que consiste en reacción en cadena de la polimerasa (PCR), reacción de amplificación, análisis PCR de transriptasa inversa, análisis de polimorfismo de conformación de cadena simple (SSCP), detección de la escisión de malapareamientos, análisis de heterodúplex, análisis de transferencia Southern, análisis de transferencia Northern,

análisis de transferencia Western, hibridación in situ, análisis de matrices, secuenciación de ácido desoxirribonucleico, análisis de polimorfismo de la restricción de la longitud del fragmento y combinaciones o subcombinaciones de estos, de dicha muestra.

En una realización, el nivel de expresión del marcador en la muestra se determina usando una técnica que se selecciona del grupo que consiste en inmunohistoquímica, inmunocitoquímica, citometría de flujo, ELISA y espectrometría de masa.

En una realización, se determina el nivel de expresión de una pluralidad de marcadores.

En una realización, el sujeto se está tratando con una terapia que se selecciona del grupo que consiste en un compuesto influenciador ambiental, cirugía, radiación, terapia hormonal, terapia de anticuerpos, terapia con factores de crecimiento, citocinas, quimioterapia, terapia de células madre alogénicas. En una realización, el compuesto influenciador ambiental es una molécula de la Coenzima Q10.

La invención proporciona adicionalmente métodos para analizar si un sujeto que padece de cáncer, el método comprende: determinar el nivel de expresión de uno o más marcadores presentes en una muestra biológica obtenida del sujeto, donde el

uno o más marcadores se selecciona del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1 y CANX; y comparar el nivel de expresión del uno o más marcadores presentes en la muestra biológica obtenida del sujeto con el nivel de expresión del uno o más marcadores presentes en una muestra de control, donde una modulación en el nivel de expresión del uno o más marcadores en la muestra biológica obtenida del sujeto en relación con el nivel de expresión del uno o más marcadores en la muestra de control es un indicador de que el sujeto padece de cáncer, y se analiza de esta manera si el sujeto padece de cáncer.

En una realización, la muestra comprende un fluido obtenido del sujeto. En una realización, el fluido se selecciona del grupo que consiste en fluidos sanguíneos, vómito, saliva, linfa, fluido quístico, orina, fluidos recogidos de un lavado bronquial, fluidos recogidos del enjuague peritoneal y fluidos ginecológicos. En una realización, la muestra es una muestra de sangre o un componente de esta.

En otra realización, la muestra comprende un tejido o un componente de este, obtenido del sujeto. En una realización, el tejido se selecciona del grupo que consiste en óseo, tejido

conectivo, cartílago, pulmonar, hepático, renal, tejido muscular, cardíaco, pancreático y cutáneo.

En una realización el sujeto es un ser humano.

En una realización, el nivel de expresión del uno o más marcadores en la muestra biológica se determina mediante el análisis de un polinucleótido transcripto o una porción de este en la muestra. En una realización, el análisis del polinucleótido transcripto comprende amplificar el polinucleótido transcripto.

En una realización, el nivel de expresión del marcador en el sujeto se determina mediante el análisis de una proteína o una porción de esta en la muestra. En una realización, la proteína se analiza usando un reactivo que se une específicamente con la proteína.

En una realización, el nivel de expresión del uno o más marcadores en la muestra se determina usando una técnica que se selecciona del grupo que consiste en reacción en cadena de la polimerasa (PCR), reacción de amplificación, análisis PCR de transriptasa inversa, análisis de polimorfismo de conformación de cadena simple (SSCP), detección de la escisión de malapareamientos, análisis de heterodúplex, análisis de transferencia Southern, análisis de transferencia Northern,

análisis de transferencia Western, hibridación in situ, análisis de matrices, secuenciación de ácido desoxirribonucleico, análisis de polimorfismo de la restricción de la longitud del fragmento y combinaciones o subcombinaciones de estos, de dicha muestra.

En una realización, el nivel de expresión del marcador en la muestra se determina usando una técnica que se selecciona del grupo que consiste en inmunohistoquímica, inmunocitoquímica, citometría de flujo, ELISA y espectrometría de masa.

En una realización, se determina el nivel de expresión de una pluralidad de marcadores.

En una realización, el sujeto se está tratando con una terapia que se selecciona del grupo que consiste en un compuesto influenciador ambiental, cirugía, radiación, terapia hormonal, terapia de anticuerpos, terapia con factores de crecimiento, citocinas, quimioterapia, terapia de células madre alogénicas. En una realización, el compuesto influenciador ambiental es una molécula de la Coenzima Q10.

La invención proporciona adicionalmente métodos para pronosticar si un sujeto está predispuesto a desarrollar un cáncer, el método comprende: determinar el nivel de expresión de uno o más marcadores presentes en una muestra biológica obtenida

del sujeto, donde el uno o más marcadores se selecciona del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1 y CANX; y comparar el nivel de expresión del uno o más marcadores presentes en la muestra biológica obtenida del sujeto con el nivel de expresión del uno o más marcadores presentes en una muestra de control, donde una modulación en el nivel de expresión del uno o más marcadores en la muestra biológica obtenida del sujeto en relación con el nivel de expresión del uno o más marcadores en la muestra de control es un indicador de que el sujeto está predispuesto a desarrollar cáncer, y se realiza de esta manera el pronóstico de si el sujeto está predispuesto a desarrollar el cáncer.

En una realización, la muestra comprende un fluido obtenido del sujeto. En una realización, el fluido se selecciona del grupo que consiste en fluidos sanguíneos, vómito, saliva, linfa, fluido quístico, orina, fluidos recogidos de un lavado bronquial, fluidos recogidos del enjuague peritoneal y fluidos ginecológicos. En una realización, la muestra es una muestra de sangre o un componente de esta.

En otra realización, la muestra comprende un tejido o un componente de este, obtenido del sujeto. En una realización, el

tejido se selecciona del grupo que consiste en óseo, tejido conectivo, cartílago, pulmonar, hepático, renal, tejido muscular, cardíaco, pancreático y cutáneo.

En una realización el sujeto es un ser humano.

En una realización, el nivel de expresión del uno o más marcadores en la muestra biológica se determina mediante el análisis de un polinucleótido transcripto o una porción de este en la muestra. En una realización, el análisis del polinucleótido transcripto comprende amplificar el polinucleótido transcripto.

En una realización, el nivel de expresión del marcador en el sujeto se determina mediante el análisis de una proteína o una porción de esta en la muestra. En una realización, la proteína se analiza usando un reactivo que se une específicamente con la proteína.

En una realización, el nivel de expresión del uno o más marcadores en la muestra se determina usando una técnica que se selecciona del grupo que consiste en reacción en cadena de la polimerasa (PCR), reacción de amplificación, análisis PCR de transriptasa inversa, análisis de polimorfismo de conformación de cadena simple (SSCP), detección de la escisión de malapareamientos, análisis de heterodúplex, análisis de

transferencia Southern, análisis de transferencia Northern, análisis de transferencia Western, hibridación in situ, análisis de matrices, secuenciación de ácido desoxirribonucleico, análisis de polimorfismo de la restricción de la longitud del fragmento y combinaciones o subcombinaciones de estos, de dicha muestra.

En una realización, el nivel de expresión del marcador en la muestra se determina usando una técnica que se selecciona del grupo que consiste en inmunohistoquímica, inmunocitoquímica, citometría de flujo, ELISA y espectrometría de masa.

En una realización, se determina el nivel de expresión de una pluralidad de marcadores.

En una realización, el sujeto se está tratando con una terapia que se selecciona del grupo que consiste en un compuesto influenciador ambiental, cirugía, radiación, terapia hormonal, terapia de anticuerpos, terapia con factores de crecimiento, citocinas, quimioterapia, terapia de células madre alogénicas. En una realización, el compuesto influenciador ambiental es una molécula de la Coenzima Q10.

La invención proporciona adicionalmente métodos para pronosticar la recurrencia de un cáncer en un sujeto, el método comprende: determinar el nivel de expresión de uno o más

marcadores presentes en una muestra biológica obtenida del sujeto, donde el uno o más marcadores se selecciona del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1 y CANX; y comparar el nivel de expresión del uno o más marcadores presentes en la muestra biológica obtenida del sujeto con el nivel de expresión del uno o más marcadores presentes en una muestra de control, donde una modulación en el nivel de expresión del uno o más marcadores en la muestra biológica obtenida del sujeto en relación con el nivel de expresión del uno o más marcadores en la muestra de control es un indicador de la recurrencia del cáncer, y se realiza de esta manera el pronóstico de la recurrencia del cáncer en el sujeto.

En una realización, la muestra comprende un fluido obtenido del sujeto. En una realización, el fluido se selecciona del grupo que consiste en fluidos sanguíneos, vómito, saliva, linfa, fluido quístico, orina, fluidos recogidos de un lavado bronquial, fluidos recogidos del enjuague peritoneal y fluidos ginecológicos. En una realización, la muestra es una muestra de sangre o un componente de esta.

En otra realización, la muestra comprende un tejido o un componente de este, obtenido del sujeto. En una realización, el tejido se selecciona del grupo que consiste en óseo, tejido

conectivo, cartílago, pulmonar, hepático, renal, tejido muscular, cardíaco, pancreático y cutáneo.

En una realización el sujeto es un ser humano.

En una realización, el nivel de expresión del uno o más marcadores en la muestra biológica se determina mediante el análisis de un polinucleótido transcripto o una porción de este en la muestra. En una realización, el análisis del polinucleótido transcripto comprende amplificar el polinucleótido transcripto.

En una realización, el nivel de expresión del marcador en el sujeto se determina mediante el análisis de una proteína o una porción de esta en la muestra. En una realización, la proteína se analiza usando un reactivo que se une específicamente con la proteína.

En una realización, el nivel de expresión del uno o más marcadores en la muestra se determina usando una técnica que se selecciona del grupo que consiste en reacción en cadena de la polimerasa (PCR), reacción de amplificación, análisis PCR de transriptasa inversa, análisis de polimorfismo de conformación de cadena simple (SSCP), detección de la escisión de malapareamientos, análisis de heterodúplex, análisis de transferencia Southern, análisis de transferencia Northern,

análisis de transferencia Western, hibridación in situ, análisis de matrices, secuenciación de ácido desoxirribonucleico, análisis de polimorfismo de la restricción de la longitud del fragmento y combinaciones o subcombinaciones de estos, de dicha muestra.

En una realización, el nivel de expresión del marcador en la muestra se determina usando una técnica que se selecciona del grupo que consiste en inmunohistoquímica, inmunocitoquímica, citometría de flujo, ELISA y espectrometría de masa.

En una realización, se determina el nivel de expresión de una pluralidad de marcadores.

En una realización, el sujeto se está tratando con una terapia que se selecciona del grupo que consiste en un compuesto influenciador ambiental, cirugía, radiación, terapia hormonal, terapia de anticuerpos, terapia con factores de crecimiento, citocinas, quimioterapia, terapia de células madre alogénicas. En una realización, el compuesto influenciador ambiental es una molécula de la Coenzima Q10.

La invención proporciona adicionalmente métodos para realizar el pronóstico de supervivencia de un sujeto con cáncer, el método comprende: determinar el nivel de expresión de uno o másmarcadores presentes en una muestra biológica obtenida del

sujeto, donde el uno o más marcadores se selecciona del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1 y CANX; y comparar el nivel de expresión del uno o más marcadores presentes en la muestra biológica obtenida del sujeto con el nivel de expresión del uno o más marcadores presentes en una muestra de control, donde una modulación en el nivel de expresión del uno o más marcadores en la muestra biológica obtenida del sujeto en relación con el nivel de expresión del uno o más marcadores en la muestra de control es un indicador de la supervivencia del sujeto, y se realiza de esta manera el pronóstico de supervivencia del sujeto con cáncer.

En una realización, la muestra comprende un fluido obtenido del sujeto. En una realización, el fluido se selecciona del grupo que consiste en fluidos sanguíneos, vómito, saliva, linfa, fluido quístico, orina, fluidos recogidos de un lavado bronquial, fluidos recogidos del enjuague peritoneal y fluidos ginecológicos. En una realización, la muestra es una muestra de sangre o un componente de esta.

En otra realización, la muestra comprende un tejido o un componente de este, obtenido del sujeto. En una realización, el tejido se selecciona del grupo que consiste en óseo, tejido

conectivo, cartílago, pulmonar, hepático, renal, tejido muscular, cardíaco, pancreático y cutáneo.

En una realización el sujeto es un ser humano.

En una realización, el nivel de expresión del uno o más marcadores en la muestra biológica se determina mediante el análisis de un polinucleótido transcripto o una porción de este en la muestra. En una realización, el análisis del polinucleótido transcripto comprende amplificar el polinucleótido transcripto.

En una realización, el nivel de expresión del marcador en el sujeto se determina mediante el análisis de una proteína o una porción de esta en la muestra. En una realización, la proteína se analiza usando un reactivo que se une específicamente con la proteína.

En una realización, el nivel de expresión del uno o más marcadores en la muestra se determina usando una técnica que se selecciona del grupo que consiste en reacción en cadena de la polimerasa (PCR), reacción de amplificación, análisis PCR de transriptasa inversa, análisis de polimorfismo de conformación de cadena simple (SSCP), detección de la escisión de malapareamientos, análisis de heterodúplex, análisis de transferencia Southern, análisis de transferencia Northern,

análisis de transferencia Western, hibridación in situ, análisis de matrices, secuenciación de ácido desoxirribonucleico, análisis de polimorfismo de la restricción de la longitud del fragmento y combinaciones o subcombinaciones de estos, de dicha muestra.

En una realización, el nivel de expresión del marcador en la muestra se determina usando una técnica que se selecciona del grupo que consiste en inmunohistoquímica, inmunocitoquímica, citometría de flujo, ELISA y espectrometría de masa.

En una realización, se determina el nivel de expresión de una pluralidad de marcadores.

En una realización, el sujeto se está tratando con una terapia que se selecciona del grupo que consiste en un compuesto influenciador ambiental, cirugía, radiación, terapia hormonal, terapia de anticuerpos, terapia con factores de crecimiento, citocinas, quimioterapia, terapia de células madre alogénicas. En una realización, el compuesto influenciador ambiental es una molécula de la Coenzima Q10.

La invención proporciona adicionalmente métodos para monitorear la evolución de un cáncer en un sujeto, el método comprende: comparar el nivel de expresión de uno o más marcadores presentes en una primera muestra obtenida del sujeto antes de

administrarle al menos una parte de un régimen de tratamiento al sujeto y el nivel de expresión del uno o más marcadores presentes en una segunda muestra obtenida del sujeto después de la administración de al menos una parte del régimen de tratamiento, donde el uno o más marcadores se selecciona del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1 y CANX, y se monitorea de esta manera la evolución del cáncer en el sujeto.

En una realización, la muestra comprende un fluido obtenido del sujeto. En una realización, el fluido se selecciona del grupo que consiste en fluidos sanguíneos, vómito, saliva, linfa, fluido quístico, orina, fluidos recogidos de un lavado bronquial, fluidos recogidos del enjuague peritoneal y fluidos ginecológicos. En una realización, la muestra es una muestra de sangre o un componente de esta.

En otra realización, la muestra comprende un tejido o un componente de este, obtenido del sujeto. En una realización, el tejido se selecciona del grupo que consiste en óseo, tejido conectivo, cartílago, pulmonar, hepático, renal, tejido muscular, cardíaco, pancreático y cutáneo.

En una realización el sujeto es un ser humano.

En una realización, el nivel de expresión del uno o más marcadores en la muestra biológica se determina mediante el análisis de un polinucleótido transcripto o una porción de este en la muestra. En una realización, el análisis del polinucleótido transcripto comprende amplificar el polinucleótido transcripto.

En una realización, el nivel de expresión del marcador en el sujeto se determina mediante el análisis de una proteína o una porción de esta en la muestra. En una realización, la proteína se analiza usando un reactivo que se une específicamente con la proteína.

En una realización, el nivel de expresión del uno o más marcadores en la muestra se determina usando una técnica que se selecciona del grupo que consiste en reacción en cadena de la polimerasa (PCR), reacción de amplificación, análisis PCR de transriptasa inversa, análisis de polimorfismo de conformación de cadena simple (SSCP), detección de la escisión de malapareamientos, análisis de heterodúplex, análisis de transferencia Southern, análisis de transferencia Northern, análisis de transferencia Western, hibridación in situ, análisis de matrices, secuenciación de ácido desoxirribonucleico, análisis de polimorfismo de la restricción de la longitud del fragmento y combinaciones o subcombinaciones de estos, de dicha muestra.

En una realización, el nivel de expresión del marcador en la muestra se determina usando una técnica que se selecciona del grupo que consiste en inmunohistoquímica, inmunocitoquímica, citometría de flujo, ELISA y espectrometría de masa.

En una realización, se determina el nivel de expresión de una pluralidad de marcadores.

En una realización, el sujeto se está tratando con una terapia que se selecciona del grupo que consiste en un compuesto influenciador ambiental, cirugía, radiación, terapia hormonal, terapia de anticuerpos, terapia con factores de crecimiento, citocinas, quimioterapia, terapia de células madre alogénicas. En una realización, el compuesto influenciador ambiental es una molécula de la Coenzima Q10.

La invención proporciona adicionalmente métodos para identificar un compuesto para el tratamiento de un cáncer en un sujeto, el método comprende: obtener una muestra biológica del sujeto; poner en contacto la muestra biológica con un compuesto de prueba; determinar el nivel de expresión de uno o más marcadores presentes en la muestra biológica obtenida del sujeto, donde el uno o más marcadores se selecciona del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF,

KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1 y CANX con un cambio de múltiplo positivo y/o con un cambio de múltiplo negativo; comparar el nivel de expresión del uno o más marcadores en la muestra biológica con un control adecuado; y seleccionar un compuesto de prueba que disminuya el nivel de expresión del uno o más marcadores con un cambio de múltiplo negativo presentes en la muestra biológica y/o aumente el nivel de expresión del uno o más marcadores con un cambio de múltiplo positivo presentes en la muestra biológica, y se identifica de esta manera un compuesto para tratar el cáncer en un sujeto.

En una realización, la muestra comprende un fluido obtenido del sujeto. En una realización, el fluido se selecciona del grupo que consiste en fluidos sanguíneos, vómito, saliva, linfa, fluido quístico, orina, fluidos recogidos de un lavado bronquial, fluidos recogidos del enjuague peritoneal y fluidos ginecológicos. En una realización, la muestra es una muestra de sangre o un componente de esta.

En otra realización, la muestra comprende un tejido o un componente de este, obtenido del sujeto. En una realización, el tejido se selecciona del grupo que consiste en óseo, tejido conectivo, cartílago, pulmonar, hepático, renal, tejido muscular, cardíaco, pancreático y cutáneo.

En una realización el sujeto es un ser humano.

En una realización, el nivel de expresión del uno o más marcadores en la muestra biológica se determina mediante el análisis de un polinucleótido transcripto o una porción de este en la muestra. En una realización, el análisis del polinucleótido transcripto comprende amplificar el polinucleótido transcripto.

En una realización, el nivel de expresión del marcador en el sujeto se determina mediante el análisis de una proteína o una porción de esta en la muestra. En una realización, la proteína se analiza usando un reactivo que se une específicamente con la proteína.

En una realización, el nivel de expresión del uno o más marcadores en la muestra se determina usando una técnica que se selecciona del grupo que consiste en reacción en cadena de la polimerasa (PCR), reacción de amplificación, análisis PCR de transriptasa inversa, análisis de polimorfismo de conformación de cadena simple (SSCP), detección de la escisión de malapareamientos, análisis de heterodúplex, análisis de transferencia Southern, análisis de transferencia Northern, análisis de transferencia Western, hibridación in situ, análisis de matrices, secuenciación de ácido desoxirribonucleico, análisis

de polimorfismo de la restricción de la longitud del fragmento y combinaciones o subcombinaciones de estos, de dicha muestra.

En una realización, el nivel de expresión del marcador en la muestra se determina usando una técnica que se selecciona del grupo que consiste en inmunohistoquímica, inmunocitoquímica, citometría de flujo, ELISA y espectrometría de masa.

En una realización, se determina el nivel de expresión de una pluralidad de marcadores.

En una realización, el sujeto se está tratando con una terapia que se selecciona del grupo que consiste en un compuesto influenciador ambiental, cirugía, radiación, terapia hormonal, terapia de anticuerpos, terapia con factores de crecimiento, citocinas, quimioterapia, terapia de células madre alogénicas. En una realización, el compuesto influenciador ambiental es una molécula de la Coenzima Q10.

La invención proporciona adicionalmente un kit para analizar la eficacia de una terapia para tratar un cáncer, el kit comprende reactivos para determinar el nivel de expresión de al menos un marcador seleccionado del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA,

PHB2, ATP5A1 y CANX, e instrucciones para usar el kit para analizar la eficacia de la terapia para tratar el cáncer.

La invención proporciona adicionalmente un kit para analizar si un sujeto padece de cáncer o no, el kit comprende reactivos para determinar el nivel de expresión de al menos un marcador seleccionado del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1 y CANX, e instrucciones para usar el kit para analizar si el sujeto padece de cáncer o no.

La invención proporciona adicionalmente un kit para realizar el pronóstico de si un sujeto está predispuesto a desarrollar un cáncer, el kit comprende reactivos para determinar el nivel de expresión de al menos un marcador seleccionado del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1 y CANX, e instrucciones para usar el kit para realizar el pronóstico de si el sujeto está predispuesto a desarrollar el cáncer.

La invención proporciona adicionalmente un kit para realizar el pronóstico de la recurrencia de un cáncer en un sujeto, el kit comprende reactivos para analizar el nivel de

expresión de al menos un marcador seleccionado del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1 y CANX, e instrucciones para usar el kit para realizar el pronóstico de recurrencia del cáncer.

La invención proporciona adicionalmente un kit para realizar el pronóstico de la recurrencia de un cáncer, el kit comprende reactivos para determinar el nivel de expresión de al menos un marcador seleccionado del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1 y CANX, e instrucciones para usar el kit para realizar el pronóstico de recurrencia del cáncer.

La invención proporciona adicionalmente un kit para realizar el pronóstico de la supervivencia de un sujeto con cáncer, el kit comprende reactivos para determinar el nivel de expresión de al menos un marcador seleccionado del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1 y CANX, e instrucciones para usar el kit para realizar el pronóstico de supervivencia del sujeto con cáncer.

La invención proporciona adicionalmente un kit para monitorear la evolución de un cáncer en un sujeto, el kit comprende reactivos para determinar el nivel de expresión de al menos un marcador seleccionado del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1 y CANX, e instrucciones para usar el kit para realizar el pronóstico de la evolución del cáncer en un sujeto.

Los kits de la invención pueden comprender adicionalmente medios para obtener una muestra biológica de un sujeto, una muestra de control y/o un compuesto influenciador ambiental.

El medio para determinar el nivel de expresión del al menos un marcador puede comprender medios para analizar un polinucleótido transcripto o una parte de este en la muestra y/o medios para analizar una proteína o una parte de esta en la muestra.

En una realización, los kits comprenden reactivos para determinar el nivel de expresión de una pluralidad de marcadores.

Varios aspectos de la invención se describen en más detalle en las siguientes subsecciones.

C. Moléculas de ácido nucleico aisladas

Un aspecto de la invención se refiere a moléculas de ácido nucleico aisladas, incluyendo ácidos nucleicos que codifican una proteína marcadora o una parte de esta. Los ácidos nucleicos de la invención también incluyen moléculas de ácido nucleico suficientes para su uso como sondas de hibridación para identificar moléculas de ácido nucleico marcadoras, por ejemplo, aquellas aptas para su uso como cebadores de PCR para la amplificación o mutación de moléculas de ácido nucleico marcadoras. Tal como se usa en la presente, el término "molécula de ácido nucleico" se pretende que incluya moléculas de ADN (por ejemplo, ADNc o ADN genómico) y moléculas de *ARN* (por ejemplo, ARNm) y a análogos de ADN y ARN generados usando análogos de nucleótidos. La molécula de ácido nucleico puede ser de cadena simple o de cadena doble, pero preferentemente es un ADN de cadena doble.

Una molécula de ácido nucleico "aislada" es la cual se separa de otras moléculas de ácido nucleico que están presentes en la fuente natural de la molécula de ácido nucleico. En una realización, una molécula de ácido nucleico "aislada" está libre de secuencias (preferentemente secuencias que codifican proteínas) que flanquean naturalmente el ácido nucleico *(es*

decir, secuencias ubicadas en los extremos 5' y 3' del ácido nucleico) en el ADN genómico del orgánico del cual se deriva el ácido nucleico. Por ejemplo, en varias realizaciones, la molécula de ácido nucleico puede contener menos de aproximadamente 5kB, 4kB, 3kB, 2kB, 1kB, 0,5 kB o 0,1 kB de secuencias de nucleótidos que flanquean naturalmente la molécula de ácido nucleico en ADN genómico de la célula de la cual deriva el ácido nucleico. En otra realización, una molécula de ácido nucleico "aislada", tal como una molécula de ADNc, puede estar sustancialmente libre de otro material celular o medio de cultivo cuando se produce mediante técnicas recombinantes, o sustancialmente de precursores químicos u otros químicos cuando se sintetiza químicamente. Una molécula de ácido nucleico que está sustancialmente libre de material celular incluye preparaciones que tienen menos del 30%, 20%, 10% o 5% de ácido nucleico heterólogo (también mencionado en la presente como "ácido nucleico contaminante").

Una molécula de ácido nucleico de la presente invención puede aislarse usando técnicas de biología molecular estándar y la secuencia de información en los registros de la base de datos descritos en la presente. Usando todas o una parte de dichas secuencias de ácido nucleico, las moléculas de ácido nucleico de la invención pueden aislarse usando técnicas de hibridación y clonación estándar *(por ejemplo,* como las descritas en Sambrook

et ál., ed., *Molecular Cloning: A Laboratory Manual, 2ª ed.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Una molécula de ácido nucleico de la invención puede amplificarse usando ADNc, ARNm o ADN genómico como plantilla y cebadores de oligonucleótidos adecuados para las técnicas de amplificación por PCR estándar. El ácido nucleico amplificado de esta manera puede clonarse en un vector adecuado y caracterizarse mediante el análisis de secuencia de ADN. Además, los nucleótidos correspondientes a toda o parte de la molécula de ácido nucleico de la invención pueden prepararse mediante técnicas de síntesis estándar, *por ejemplo,* usando un sintetizador de ADN automatizado.

En otra realización preferida, una molécula de ácido nucleico aislada de la invención comprende una molécula de ácido nucleico que tiene una secuencia de nucleótidos complementaria a la secuencia de nucleótidos de un ácido nucleico marcador o de la secuencia de nucleótidos de un ácido nucleico que codifica una proteína marcadora. Una molécula de ácido nucleico complementaria a una secuencia de nucleótidos dada es la cual es suficientemente complementaria a la secuencia de nucleótidos dada, de manera que esta se hibrida a la secuencia de nucleótidos dada, y forma de esta manera un dúplex estable.

Además, una molécula de ácido nucleico de la invención puede comprender solamente una parte de una secuencia de ácido nucleico, donde la secuencia de ácido nucleico de longitud completa comprende un ácido nucleico marcador o que codifica una proteína marcadora. Dichos ácidos nucleicos pueden usarse, por ejemplo, como sonda o cebador. La sonda/el cebador se usa comúnmente como uno o más oligonucleótidos sustancialmente purificados. El oligonucleótido comprende comúnmente una región de la secuencia de nucleótidos que se hibrida bajo condiciones estrictas a al menos aproximadamente 7, preferentemente aproximadamente 15, más preferentemente aproximadamente 25, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350 o 400 o más nucleótidos consecutivos de un ácido nucleico de la invención.

Las sondas basadas en la secuencia de una molécula de ácido nucleico de la invención pueden usarse para detectar secuencias transcriptas o genómicas correspondientes a uno o más marcadores de la invención. La sonda comprende un grupo etiqueta unido a esta, *por ejemplo, un* radioisótopo, un compuesto fluorescente, una enzima o un co-factor de enzima. Dichas sondas pueden usarse como parte de un kit de prueba diagnóstica para identificar células o tejidos que pueden expresar incorrectamente la proteína, tal como mediante niveles de una molécula de ácido nucleico que codifica la proteína en una muestra de células de un sujeto, *por ejemplo,* mediante la detección de niveles de ARNm o

la determinación de si un gen codifica la proteína que fue mutada o eliminada.

La invención comprende adicionalmente moléculas de ácido nucleico que difieren, debido a la degeneración del código genético, de la secuencia de nucleótidos de los ácidos nucleicos que codifican una proteína marcadora, y por lo tanto codifican la misma proteína.

Los expertos en la técnica apreciarán que los polimorfismos de la secuencia de ADN que llevan a cambios en la secuencia de aminoácidos pueden existir en una población *(por ejemplo,* la población humana). Dichos polimorfismos genéticos pueden existir entre individuos de una población debido a variantes alélicas naturales. Un alelo es uno de un grupo de genes que se presentan de manera alternativa en un locus génico dado. Además, se apreciará que los polimorfismos de ADN que afectan los niveles de expresión del ARN pueden también existir de manera que afecten la totalidad del nivel de expresión de ese gen (por ejemplo, mediante la afección de la regulación o la degradación).

Tal como se usa en la presente, la frase "variante alélica" se refiere a una secuencia de nucleótidos que se

presenta en un locus dado o a un polipéptido que es codificado por la secuencia de nucleótidos.

Tal como se usa en la presente, los términos "gen" y "gen recombinante" se refieren a moléculas de ácido nucleico que comprenden un marco de lectura abierto que codifica un polipéptido correspondiente a un marcador de la invención. Dichas variantes alélicas naturales pueden dar como resultado comúnmente una variación del 1-5% en la secuencia de nucleótidos de un gen dado. Pueden identificarse alelos alternativos mediante la secuenciación del gen de interés en una cantidad de individuos diferentes. Esto se puede llevar a cabo fácilmente usando sondas de hibridación para identificar el mismo locus génico en una variedad de individuos. Cualquiera y todas dichas variantes de nucleótidos y polimorfismos de aminoácido resultantes, o variantes que son el resultado de una variación alélica natural, y que no alteran la actividad funcional, se pretende que estén incluidos dentro del alcance de la invención.

En otra realización, una molécula de ácido nucleico aislada de la invención es de al menos 7, 15, 20, 25, 30, 40, 60, 80, 100, 150, 200, 250, 300, 350, 400, 450, 550, 650, 700, 800, 900, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2400, 2600, 2800, 3000, 3500, 4000, 4500 o más nucleótidos de longitud, y se hibrida bajo condiciones estrictas a un ácido nucleico marcador o

a un ácido nucleico que codifica una proteína marcadora. Tal como se usa en la presente, el término "se hibrida bajo condiciones estrictas" se pretende que describa las condiciones de hibridación y de lavado bajo las cuales las secuencias de nucleótidos al menos 60% (65%, 70% y preferentemente 75%) idénticas unas a otras permanecen típicamente hibridadas entre sí. Dichas condiciones estrictas son conocidas para los expertos en la técnica y pueden encontrarse en las secciones 6.3.1-6.3.6 de *Current Protocols in Molecular Biology,* John Wiley & Sons, N.Y. (1989). Un ejemplo no taxativo preferido de condiciones de hibridación estrictas es la hibridación en 6X de cloruro de sodio/citrato de sodio (SSC) a aproximadamente 45°C, seguida de uno o más lavados en 0,2X SSC, 0,1% de SDS a 50-65°C.

Además de las variantes alélicas de origen natural de la molécula de ácido nucleico de la invención que puede existir en la población, el experto en la técnica apreciará además que pueden introducirse cambios en la secuencia mediante mutación, y de esta manera se producen cambios en la secuencia de aminoácidos de la proteína codificada, sin alterar la actividad biológica de la proteína codificada de esta manera. Por ejemplo, se pueden realizar sustituciones de nucleótidos, lo cual lleva a sustituciones de aminoácidos en residuos aminoácidos "no esenciales". Un residuo aminoácido "no esencial" es un residuo que puede alterarse de la secuencia de tipo salvaje sin alterar

la actividad biológica, mientras que un residuo aminoácido "esencial" es necesario para la actividad biológica. Por ejemplo, los residuos aminoácidos que no se conservan o solo se conservan parcialmente entre homólogos de varias especias pueden ser no esenciales para la actividad y, por lo tanto, serían objetivos probables para su alteración. De manera alternativa, los residuos aminoácidos que se conservan entre los homólogos de varias especies (por ejemplo, murina y humana) pueden ser esenciales para la actividad y por lo tanto no serían objetivos probables para su alteración.

Por consiguiente, otro aspecto de la invención se refiere a moléculas de ácido nucleico que codifican una variante de una proteína marcadora que contiene cambios en los residuos aminoácidos que no son esenciales para la actividad. Dichas variantes de proteínas marcadoras en secuencias de aminoácidos de proteínas marcadoras de origen natural aún conservan la actividad biológica. En una realización, dicha variante de proteína marcadora tiene una secuencia de aminoácidos que es al menos aproximadamente 40% idéntica, 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% idéntica a la secuencia de aminoácidos de una proteína marcadora.

Una molécula de ácido nucleico aislada que codifica una variante de una proteína marcadora puede crearse mediante la

introducción de una o más sustituciones adiciones o deleciones de nucleótidos dentro de la secuencia de nucleótidos de los ácidos nucleicos marcadores, de manera que una o más sustituciones, adiciones o deleciones de residuos aminoácidos sean introducidas en la proteína codificada. Las mutaciones pueden introducirse mediante técnicas estándar, tal como mutagénesis dirigida al sitio y mutagénesis mediada por PCR. Preferentemente, las sustituciones de aminoácidos conservadoras se hacen en uno o más residuos aminoácidos no esenciales que fueron predichos. Una "sustitución de aminoácidos conservadora" es una en la que el residuo de aminoácido se reemplaza por un residuo de aminoácido que tiene una cadena lateral similar. Las familias de residuos aminoácidos que tienen cadenas laterales similares se han definido en la técnica. Estas familias incluyen aminoácidos con cadenas laterales básicas *(por ejemplo,* lisina, arginina, histidina), cadenas laterales ácidas *(por ejemplo,* ácido aspártico, ácido glutámico), cadenas laterales polares no cargadas *(por ejemplo,* glicina, asparagina, glutamina, serina, treonina, tirosina, cisteina), cadenas laterales no polares *(por ejemplo,* alanina, valina, leucina, isoleucina, prolina, fenulalanina, metionina, triptófano), cadenas laterales beta ramificadas (por ejemplo, *treonina, valina, isoleucina) y cadenas laterales aromáticas*(por ejemplo, *tirosina,* fenilalanina, triptófano, histidina). De manera alternativa, las mutaciones pueden introducirse al azar a lo largo de toda o parte de la

secuencia codificante, tal como por medio de mutagénesis de saturación, y los mutantes resultantes pueden cribarse para verificar la actividad biológica para identificar mutantes que conservan la actividad. Después del a mutagénesis, la proteína codificada puede expresarse de manera recombinante y puede determinarse la actividad de la proteína.

La presente invención abarca moléculas de ácido nucleico antisentido, *es decir,* moléculas que son complementarias al ácido nucleico sentido de la invención, por ejemplo, complementarias a la cadena codificante de una molécula de ADNc marcadora de cadena doble o complementarias a una secuencia de ARNm marcadora. Por consiguiente, un ácido nucleico antisentido de la invención puede ser un enlace hidrógeno unido a (es decir apareado con) un ácido nucleico de la invención. El ácido nucleico antisentido puede ser complementario a una cadena codificante completa, o solamente a una parte de esta, por ejemplo, la totalidad o parte de una región codificante de la proteína (o marco de lectura abierto). Una molécula de ácido nucleico antisentido también puede ser antisentido a la totalidad o parte de una región no codificante de la cadena codificante de una secuencia de nucleótidos que codifican una proteína marcadora. Las regiones no codificantes ("regiones 5' y 3' no traducidas") son las secuencias 5' y 3' que flanquean la región codificante y no se traducen en aminoácidos.

Un oligonucleótido antisentido puede ser de, por ejemplo, aproximadamente 5, 10, 15, 20, 25, 30, 35, 40, 45, o 50 o más nucleótidos de longitud. Un ácido nucleico antisentido de la invención puede construirse usando síntesis química y reacciones de ligación enzimáticas usando procedimientos conocidos en la técnica. Por ejemplo, un ácido nucleico antisentido *(por ejemplo,* un oligonucleótido antisentido) puede sintetizarse químicamente usando nucleótidos de origen natural o nucleótidos modificados de varias maneras diseñados para aumentar la estabilidad biológica de las moléculas o para aumentar la estabilidad física del dúplex formado entre los ácidos nucleicos sentido y antisentido, por ejemplo, pueden usarse derivados de fosforotioato y nucleótidos sustituidos de acridina. Los ejemplos de nucleótidos modificados que pueden usarse para generar ácido nucleico antisentido incluyen 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-yodouracilo, hipoxantina, xantina, 4-acetilcitosina, 5-(carboxihidroxilmetil)uracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidrouracilo, beta-D-galactosilqueosina, inosina, N6-isopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-adenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta-D-mannosilqueosina, 5'-metoxicarboximetiluracilo, 5-metoxiuracilo, 2-metiltio-N6-

isopenteniladenina, ácido uracilo-5-oxiacetico (v), , pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, éster metílico del ácido uracilo-5-oxiacético, ácido uracilo-5-oxiacético (v), 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil) uracilo, (acp3)w y 2,6-diaminopurina. De manera alternativa, el ácido nucleico antisentido puede producirse biológicamente usando un vector de expresión en el cual se ha clonado el ácido nucleico en una orientación antisentido *(es decir,* el ARN transcripto de un ácido nucleico insertado será de una orientación antisentido a un ácido nucleico objetivo de interés, descrito adicionalmente en la siguiente subsección).

Las moléculas de ácido nucleico antisentido de la invención comúnmente se administran a un sujeto o se generan *in situ* de manera que se hibriden o se unan con el ARNm celular y/o el ADN genómico que codifica una proteína marcadora para inhibir de esta manera la expresión del marcador, *por ejemplo*, mediante la inhibición de la transcripción y/o la traducción. La hibridación puede ser mediante complementariedad de nucleótido convencional para formar un dúplex estable o, por ejemplo, en el caso de una molécula de ácido nucleico antisentido que se une a dúplex de ADN, a través de interacciones en la ranura mayor de la doble hélice. Los ejemplos de una vía de administración de moléculas de ácido nucleico antisentido de la invención incluyen

inyección directa al sitio de tejido o infusión del ácido nucleico antisentido en el fluido corporal asociado con la enfermedad o estado de toxicidad. De manera alternativa, las moléculas de ácido nucleico antisentido pueden modificarse para dirigirse a células seleccionadas y entonces administrarse de manera sistémica. Por ejemplo, para la administración sistémica, las moléculas antisentido pueden modificarse de manera que se unan específicamente a receptores o antígenos expresados en una superficie celular seleccionada, *por ejemplo,* mediante la unión de las moléculas de ácido nucleico antisentido a péptidos o anticuerpos que se unen a receptores o antígenos de la superficie celular. Las moléculas de ácido nucleico antisentido también pueden administrarse a células usando los vectores descritos en la presente. Para lograr concentraciones intracelulares suficientes de las moléculas de antisentido, se prefieren las construcciones de vectores en las cuales se coloca la molécula de ácido nucleico antisentido bajo el control de un promotor pol II o pol III fuerte.

Una molécula de ácido nucleico antisentido de la invención puede ser una molécula de ácido nucleico α-anomérica. Una molécula de ácido nucleico α-anomérica forma híbridos de cadena doble específicos con ARN complementario en los cuales, al contrario de las unidades α usuales, las cadenas corren paralelas una a la otra (Gaultier *et ál.,* 1987, *Nucleic Acias Res.* 15:6625-

6641). La molécula de ácido nucleico antisentido puede comprender también un 2'-o-metilribonucleotidó (Inoue *et ál.*, 1987, *Nucleic Acias Res.* 15:6131-6148) o un análogo de ADN-ARN quimérico (Inoue *et ál.*, 1987, *FEBS Lett.* 215:327-330).

La invención también abarca las ribozimas. Las ribozimas son molécula de ARN catalítico con actividad de ribonucleasa que son capaces de escindir un ácido nucleico de cadena simple, tal como un ARNm, respecto al cual tienen una región de complementariedad. Por lo tanto, las ribozimas (por ejemplo, las ribozimas de cabeza de martillo descritas en Haselhoff y Gerlach, 1988, *Nature* 334:585-591) pueden usarse para escindir catalíticamente transcriptos de ARNm para de esta manera inhibir la traducción de la proteína codificada por el ARNm. Una ribozima que tiene una especificidad por una molécula de ácido nucleico que codifica una proteína marcadora puede diseñarse basándose en la secuencia de nucleótidos de un ADNc correspondiente al marcador. Por ejemplo, puede construirse un derivado de un ARN de IVS L-19 de *Tetrahymena* en el cual la secuencia de nucleótidos del sitio activo es complementaria a la secuencia de nucleótidos que será escindida (véase Cech *et ál.* patente estadounidense N.° 4.987071; y Cech *et ál.* patente estadounidense N.° 5,116,742). De manera alternativa, un ARNm que codifica un polipéptido de la invención puede usarse para seleccionar un ARN catalítico que tiene una actividad de ribonucleasa específica de un conjunto de

moléculas de ARN (véase, por ejemplo, Bartel y Szostak, 1993, *Science* 261:1411-1418).

La invención también abarca moléculas de ácido nucleico que forman estructuras de triple hélice. Por ejemplo, la expresión de un marcador de la invención puede inhibirse mediante el direccionamiento hacia secuencias de nucleótidos complementarias a la región reguladora del gen que codifica el ácido o proteína marcadores *(por ejemplo,* el promotor y/o mejorador) para formar estructuras de triple hélice que evitan la transcripción del gen en células objetivo. Véase en general Helene (1991) *Anticancer Drug Des.* 6(6):569-84; Helene (1992) *Ann. N.Y. Acad. Sel* 660:27-36; y Maher (1992) *Bioassays* 14(12):807-15.

En varias realizaciones, las moléculas de ácido nucleico de la invención pueden modificarse en el resto base, en el resto azúcar o en la estructura fosfato para mejorar, *por ejemplo,* la estabilidad, hibridación o las solubilidad de la molécula. Por ejemplo, la estructura de fosfato desoxirribosa de los ácidos nucleicos puede modificarse para generar ácidos péptidonucleicos (véase Hyrup *et ál.,* 1996, *Bioorganic & Medicinal Chemistry* 4(1): 5-23). Tal como se usa en la presente, los términos "ácido péptidonucleicos" o "PNA" se refieren a miméticos de ácido nucleico, *por ejemplo,* miméticos de ADN, en los cuales la cadena

principal de fosfato desoxirribosa se reemplaza por una estructura de pseudopéptido y solamente se conservan las cuatro nucleobases naturales. Se ha demostrado que la estructura neutra de los PNA permite la hibridación específica al ADN y al ARN bajo condiciones de baja fuerza iónica. La síntesis de los oligómeros de PNA puede llevarse a cabo usando los protocolos de síntesis de péptidos de fase sólida estándar como se describen en Hyrup *et ál.* (1996), *supra;* Perry-O'Keefe *et ál.* (1996) *Proc. Nati. Acad. Sci. USA* 93:14670-675.

Los PNA pueden usarse en aplicaciones terapéuticas y diagnósticas. Por ejemplo, los PNA pueden usarse como agentes antisentido o sentido para la modulación específica de secuencia de la expresión génica mediante, *por ejemplo,* la inducción de la interrupción de la transcripción o traducción, o de la inhibición de la replicación. Los PNA también pueden usarse, *por ejemplo,* en el análisis de mutaciones de un par de base sencillo en un gen mediante, *por ejemplo,* fijación por PCR dirigida a PNA; como enzimas de restricción artificiales cuando se usa en combinación con otras enzimas, *por ejemplo,* nucleasas S1 (Hyrup (1996), *supra;* o como sondas o cebadores para la hibridación y secuencia de ADN (Hyrup, 1996, *supra;* Perry-O'Keefe *et ál.*, 1996, *Proc. Nati. Acad. Sci. USA* 93:14670-675).

En otra realización, los PNA se pueden modificar, *por ejemplo,* para mejorar su estabilidad o absorción celular, mediante la unión de grupos lipofílicos u otros grupos colaboradores al PNA, mediante la formación de quimeras PNA-ADN, o mediante el uso de liposomas u otras técnicas de administración de fármacos conocidas en la técnica. Por ejemplo, pueden generarse quimeras de PNA-ADN que pueden combinar las propiedades ventajosas del PNA y del ADN. Dichas quimeras permiten que las enzimas de reconocimiento de ADN, *por ejemplo,* RNasa H y polimerasas de ADN, interactúen con la porción ADN mientras que la porción PNA proporcionaría una afinidad y especificidad de unión altas. Las quimeras de PNA-ADN pueden estar unidas usando conectores de longitudes apropiadas seleccionados en término de apilamiento de bases, la cantidad de enlaces entre las nucleobases y la orientación (Hyrup, 1996, *supra).* La síntesis de las quimeras de PNA-ADN puede llevarse a cabo como se describe en Hyrup (1996), *supra,* y Finn *et ál.* (1996) *Nucleic Acids Res.* 24(17):3357-63. Por ejemplo, una cadena de ADN puede sintetizarse en un soporte sólido usando química de acoplamiento de fosforamidita estándar y análogos de nucleósidos modificados. Pueden usarse compuestos tal como 5'-(4-metoxitritil)amino-5'-desoxi-timidina fosforamidita como el enlace entre el PNA y el extremo 5' del ADN (Mag *et ál.,* 1989, *Nucleic Acids Res.* 17:5973-88). Los monómeros entonces se acoplan de un modo en etapas para producir una molécula quimérica con un segmento de PNA 5' y un

segmento de ADN 3' (Finn *et ál.*, 1996, *Nucleic Acids Res.* 24(17):3357-63). De manera alternativa, las moléculas quiméricas pueden sintetizarse con un segmento de ADN 5' y un segmento de PNA 3' (Peterser *et ál.*, 1975, *B100rganic Med. Chem. Lett.* 5:1119-11124).

En otras realizaciones, el oligonucleótido puede incluir otros grupos adjuntos tales como péptidos *(por ejemplo,* para dirigirse a receptores de células hospedadoras *in vivo),* o agentes que faciliten el transporte a través de la membrana celular (véase, *por ejemplo,* Letsinger *et ál.*, 1989, *Proc. Nati Acad. Sel USA* 86:6553-6556; Lemaitre *et ál.*, 1987, *Proc. Nati Acad. Sel USA* 84:648-652; publicación PCT N.° WO 88/09810) o la barrera hematoencefálica (véase, *por ejemplo,* la publicación PCT N.° WO 89/10134). Además, los oligonucleótidos pueden modificarse con agentes de escisión accionada por hibridación (véase, *por ejemplo,* Krol *et ál.*, 1988, *Bio/Techniques* 6:958-976) o agentes de intercalación (véase, *por ejemplo,* Zon, 1988, *Pharm. Res.* 5:539-549). Para este fin, el oligonucleótido puede conjugarse con otra molécula, *por ejemplo,* un péptido, un agente de reticulación accionada por hibridación, un agente de transporte, un agente de escisión accionada por hibridación, etc.

La invención también incluye ácidos nucleicos de balizas moleculares que tienen al menos una región que es complementaria

al ácido nucleico de la invención, de manera que la baliza molecular es útil para cuantificar la presencia del ácido nucleico de la invención en una muestra. Un ácido nucleico de "baliza molecular" es un ácido nucleico que comprende un par de regiones complementarias, y que tiene un fluoróforo y un inactivador fluorescente asociado a este. El fluoróforo y el inactivador están asociados con diferentes partes del ácido nucleico, en una orientación tal que cuando las regiones complementarias están apareadas entre sí, la fluorescencia del fluoróforo es inactivada por el incactivador. Cuando las regiones complementarias del ácido nucleico no están apareadas entre sí, la fluorescencia del fluoróforo es inactivada en un grado menor. Los ácidos nucleicos de baliza molecular se describen, por ejemplo, en la patente estadounidense 5.876.930.

D. <u>Anticuerpos y proteínas aislados</u>

Un aspecto de la invención se refiere a proteínas marcadoras aisladas y a porciones biológicamente activas de estas, así como a fragmentos de polipéptidos adecuados para su uso como inmunógenos para generar anticuerpos dirigidos contra una proteína marcadora o a un fragmento de esta. En una realización, la proteína marcadora natural puede aislarse de fuentes de células o tejido mediante un esquema de purificación adecuado usando técnicas de purificación de proteínas estándar.

En otra realización, una proteína o péptido que comprende la totalidad o un segmento de la proteína marcadora se produce mediante técnicas de ADN recombinante. Como alternativa a la expresión recombinante, dicha proteína o péptido puede sintetizarse químicamente usando técnicas de síntesis de péptidos estándar.

Una proteína "aislada" o "purificada" o una porción biológicamente activa de esta está sustancialmente libre de material celular o de otras proteínas contaminantes de la fuente de células o tejido de la cual se deriva la proteína, o sustancialmente libre de precursores químicos u otros químicos cuando se sintetiza químicamente. El término "sustancialmente libre de material celular" incluye preparaciones de proteína en las cuales la proteína se separa de los componentes celulares de las células de las cuales se aísla y se produce de forma recombinante. Por lo tanto, una proteína que está sustancialmente libre de material celular incluye preparaciones de proteína que tienen menos del 30%, 20%, 10% o 5% (en peso en seco) de proteína heteróloga (también mencionada en la presente como "proteína contaminante"). Cuando la proteína o porción biológicamente activa de esta se produce de manera recombinante, también está, de manera preferente, sustancialmente libre de medio de cultivo, es decir , el medio de cultivo representa menos de aproximadamente 20%, 10% o 5% del volumen de la preparación de

proteína. Cuando la proteína s produce mediante sustancialmente química, está, de manera preferente, sustancialmente libre de precursores químicos o de otros químicos, *es decir*, está separada de los precursores químicos u otros químicos que están involucrados en la síntesis de la proteína. Por consiguiente, dichas preparaciones de la proteína tienen menos de aproximadamente 30%, 20%, 10%, 5% (en peso en seco) de compuestos o precursores químicos distintos al polipéptido de interés.

Las porciones biológicamente activas de una proteína marcadora incluyen polipéptidos que comprenden secuencias de aminoácidos suficientemente idénticas o derivadas de secuencias de aminoácidos de la proteína marcadora, que incluyen menos aminoácidos que la proteína de longitud completa, y exhiben al menos una actividad de la proteína de longitud completa correspondiente. Comúnmente, las porciones biológicamente activas comprenden un dominio o motivo con al menos una actividad de la proteína de longitud completa correspondiente. Una porción biológicamente activa de una proteína marcadora de la invención puede ser un polipéptido que tiene, por ejemplo, 10, 25, 50, 100 o más aminoácidos de longitud. Además, pueden prepararse otras porciones biológicamente activas, en las cuales otras regiones de la proteína marcadora se eliminan, mediante técnicas recombinantes y pueden evaluarse para verificar una o más de las

actividades funcionales de la forma natural de la proteína marcadora.

Las proteínas marcadoras preferidas son codificadas por secuencias de nucleótidos que comprenden las secuencias que codifican cualquiera de los genes descritos en los ejemplos. Otras proteínas útiles son sustancialmente idénticas *(por ejemplo,* al menos aproximadamente 40%, preferentemente 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99%) a una de estas secuencias y conservan la actividad funcional de la proteína marcadora de origen natural correspondiente, pero se diferencian en la secuencia de aminoácidos, debido a la mutagénesis o variación alélica natural.

Para determinar el porcentaje de identidad de dos secuencias de aminoácidos o de dos ácidos nucleicos, las secuencias se alinean con fines de una comparación óptima *(por ejemplo,* se pueden introducir huecos en la secuencia de un primer aminoácido o secuencia de ácido nucleico para una alineación óptima con una segunda secuencia de aminoácido o ácido nucleico). Entonces se comparan los residuos de aminoácidos o nucleótidos en las posiciones de aminoácidos o posiciones de nucleótidos correspondientes. Cuando una posición en la primera secuencia está ocupada por el mismo residuo aminoácido o nucleótido que en la posición correspondiente en la segunda secuencia, entonces las

moléculas son idénticas en esa posición. Preferentemente, el porcentaje de identidad entre dos secuencias se calcula usando una alineación global. De manera alternativa, el porcentaje de identidad entre dos secuencias se calcula usando una alineación local. El porcentaje de identidad entre las dos secuencias está en función de la cantidad de posiciones idénticas compartidas por las secuencias *(es decir,* % de identidad = N.° de posiciones idénticas/N.° total de posiciones *(por ejemplo,* posiciones superpuestas) x 100). En una realización, las dos secuencias son de la misma longitud. En otra realización, las dos secuencias no son de la misma longitud.

La determinación del porcentaje de identidad entre las dos secuencias puede lograrse usando un algoritmo matemático. Un ejemplo preferido, no taxativo, de un algoritmo matemático utilizado para la comparación de dos secuencias es el algoritmo de Karlin y Altschul (1990) *Proc. Nati. Acad. Sci. USA* 87:2264-2268, modificado como en Karlin y Altschul (1993) *Proc. Nati. Acad. Sci. USA* 90:5873-5877. Dicho algoritmo se incorporó en los programas BLASTN y BLASTX de Altschul, *et ál.* (1990) J. *Mol. Biol.* 215:403-410. Las búsquedas de nucleótidos de BLAST pueden llevarse a cabo con el programa BLASTN, puntaje = 100, longitud de palabra = 12, para obtener una secuencia de nucleótidos homóloga a las moléculas de ácido nucleico de la invención. Las búsquedas de proteínas de BLAST pueden llevarse a cabo con el

programa BLASTP, puntaje = 50, longitud de palabra = 3, para obtener una secuencia de aminoácidos homóloga a las moléculas de proteína de la invención. Para obtener la alineación de huevos con fines comparativos, puede usarse una nueva versión del algoritmo BLAST llamada Gapped BLAST como se describe en Altschul *et ál.* (1997) *Nucleic Acids Res.*25:3389-3402, que es capaz de llevar a cabo alineaciones de huecos locales para los programas BLASTN, BLASTP y BLASTX. De manera alternativa, puede usarse PSI-Blast para llevar a cabo búsquedas reiteradas que detecten relaciones distantes entre moléculas. Cuando se usan los programas BLAST, Gapped BLAST y PSI-Blast, pueden usarse los parámetros por defecto de cada programa respectivo *(por ejemplo,* BLASTX y BLASTN). Véase http://www.ncbi.nlm.nih.gov. Otro ejemplo preferido, no taxativo, de un algoritmo matemático utilizado para la comparación de secuencias es el algoritmo de Myers y Miller, (1988) *CABIOS* 4:11-17. Dicho algoritmo se incorpora en el programa ALIGN (versión 2.0) que es parte del paquete de software de alineación de secuencias GCG. Cuando se utiliza el programa ALIGN para comparar secuencias de aminoácidos, pueden usarse una tabla de pesos de residuos PAM120, una penalización por longitud de hueco de 12 y una penalización por hueco de 4. Otro algoritmo útil para identificar regiones de similitud y alineación de secuencia local es el algoritmo de FASTA como se describe en Pearson y Lipman (1988) *Proc. Nati. Acad. Sel USA* 85:2444-2448. Cuando se usa el algoritmo FASTA para comparar secuencias de

nucleótidos y aminoácidos, puede usarse una tabla de peso de residuo PAM120, por ejemplo, con un valor de fc-tuple de 2.

El porcentaje de identidad entre dos secuencias puede determinarse usando técnicas similares a las descritas anteriormente, con y sin huecos. Al calcular el porcentaje de identidad solo se cuentan las coincidencias exactas.

La invención también proporciona proteínas quiméricas o de fusión que comprenden una proteína marcadora o un segmento de esta. Tal como se usa en la presente, una "proteína quimérica" o "proteína de fusión" comprende la totalidad o parte (preferentemente una parte biológicamente activa) de una proteína marcadora unida operativamente a un polipéptido heterólogo (es decir, un polipéptido distinto a la proteína marcadora). Dentro de la proteína de fusión el término "unido operativamente" pretende indicar que la proteína marcadora o segmento de esta y el polipéptido heterólogo están fusionados sin cambio de marco de lectura entre sí. El polipéptido heterólogo puede estar fusionado al extremo amino o al extremo carboxilo de la proteína marcadora o al segmento.

Una proteína de fusión útil es la proteína de fusión GST en la cual una proteína marcadora o segmento está fusionado al extremo carboxilo de las secuencias GST. Dichas proteínas de

fusión facilitan la purificación de un polipéptido recombinante de la invención.

En otra realización, la proteína de fusión contiene una secuencia de señalización heteróloga en su extremo amino. Por ejemplo, la secuencia de señalización natural de una proteína marcadora puede eliminarse y reemplazarse con una secuencia de señalización de otra proteína. Por ejemplo, la secuencia secretora gp67 de la proteína de la cubierta del baculovirus puede usarse como la secuencia de señalización heteróloga (Ausubel *et ál.,* ed., *Current Protocols in Molecular Biology,* John Wiley & Sons, NY, 1992). Otros ejemplos de secuencias de señalización heterólogas eucariotas incluyen las secuencias secretoras de melitina y fosfatasa alcalina placentaria (Stratagene; La Jolla, California). En todavía otro ejemplo, la secuencia de señalización procriota útil incluye una señal secretora phoA (Sambrook *et ál., supra)* y la señal secretora de proteína A (Pharmacia Biotech; Piscataway, Nueva Jersey).

En todavía otra realización, la proteína de fusión es una proteína de fusión de inmunoglobulina en la cual la totalidad o parte de la proteína marcadora está fusionada a secuencias que derivan de un miembro de la familia de proteínas inmunoglobulinas. Las proteínas de fusión inmunoglobulinas de la invención pueden incorporarse a composiciones farmacéuticas y

administrarse al sujeto para inhibir una interacción entre un ligando (soluble o unido a la membrana) y una proteína en la superficie de una célula (receptor), para suprimir de esta manera la transducción de señalización *in vivo*. La proteína de fusión inmunoglobulina puede usarse para afectar la biodisponibilidad de un ligando cognado de una proteína marcadora. La inhibición de la interacción ligando/cognado puede ser útil terapéuticamente, tanto para el tratamiento de trastornos proliferativos como diferenciativos y para la modulación (por ejemplo promover o inhibir) de la supervivencia celular. Además, las proteínas de fusión inmunoglobulinas de la invención pueden usarse como inmunógenos para producir anticuerpos dirigidos contra una proteína marcadora en un sujeto, para purificar los ligandos y en ensayos de cribado para identificar moléculas que inhiben la interacción de la proteína marcadora con los ligandos.

Las proteínas quiméricas y de fusión de la invención pueden producirse mediante técnicas de ADN recombinante estándar. En otra realización, el gen de fusión puede sintetizarse mediante técnicas convencionales incluyendo sintetizadores de ADN automatizados. De manera alternativa, la amplificación por PCR de fragmentos de gen puede llevarse a cabo usando cebadores de anclaje que generan extremos sin complementos unidos entre dos fragmentos génicos consecutivos que pueden aparearse subsiguientemente y volver a amplificarse para generar una secuencia génica quimérica (véase, por ejemplo, Ausubel *et ál.*,

supra). Además, hay muchos vectores de expresión disponibles comercialmente que ya codifican un resto de fusión (por ejemplo, un polipéptido GST). Un ácido nucleico que codifica un polipéptido de la invención puede clonarse en dicho vector de expresión de manera que el resto de fusión esté unido sin cambio de marco de lectura al polipéptido de la invención.

Una secuencia de señalización puede usarse para facilitar la secreción y aislamiento de las proteínas marcadoras. Las secuencias de señalización se caracterizan típicamente por un núcleo de aminoácidos hidrofóbos que generalmente se escinden de la proteína madura durante la secreción en uno o más eventos de escisión. Dichos péptidos de señalización contienen sitios de procesamiento que permiten la escisión de la secuencia de señalización de las proteínas maduras mientras pasan a través de la vía secretora. Por lo tanto, la invención se refiere a proteínas marcadoras, proteínas de fusión y segmentos de estas que tienen una secuencia de señalización, así como a aquellas proteínas de las cuales se escindió proteolíticamente la secuencia de señalización *(es decir,* los productos de escisión). En una realización, una secuencia de ácido nucleico que codifica una secuencia de señalización puede estar unida operativamente en un vector de expresión a una proteína de interés, tal como una proteína marcadora o un segmento de esta. La secuencia de señalización dirige la secreción de la proteína, de manera de

formar un hospedador eucariota en el cual el vector de expresión se transforma y la secuencia de señalización se escinde de manera subsecuente o concurrente. La proteína puede purificarse fácilmente del medio extracelular mediante métodos reconocidos en la técnica. De manera alternativa, la secuencia de señalización puede unirse a la proteína de interés usando una secuencia que facilita la purificación, tal como con un dominio GST.

La presente invención también se refiere a variantes de las proteínas marcadoras. Dichas variantes tienen una secuencia de aminoácidos alterada que pueden funcionar ya sea como agonistas (miméticos) o como antagonistas. Las variantes pueden generarse mediante mutagénesis, *por ejemplo*, truncamiento o mutación puntual. Un agonista puede conservar sustancialmente la misma o un subconjunto de las actividades biológicas de la forma de origen natural de la proteína. Un antagonista de una proteína puede inhibir una o más de las actividades de la forma de origen natural de la proteína mediante, por ejemplo, unión competitiva a un miembro en dirección a 3 o en dirección a 5 de una cascada de señalización celular que incluye la proteína de interés. Por lo tanto, los efectos biológicos específicos pueden provocarse mediante el tratamiento con una variante de función limitada. El tratamiento de un sujeto con una variante que tiene un subconjunto de las actividades biológicas de la forma de origen natural de la proteína puede tener menos efectos secundarios en

un sujeto, en comparación con el tratamiento con la forma de origen natural de la proteína.

Las variantes de una proteína marcadora que funcionan ya sea como agonistas (miméticos) o como antagonistas, pueden identificarse mediante el cribado de bibliotecas combinatorias de mutantes, *por ejemplo,* mutantes truncados, de una proteína de la invención para verificar la actividad agonista o antagonista. En una realización, se genera una biblioteca variegada de variantes mediante mutagénesis combinatoria a nivel del ácido nucleico y se codifica mediante una biblioteca génica variegada. Una biblioteca génica variegada de variantes se puede producir mediante, por ejemplo, ligación enzimática de una mezcla de oligonucleótidos sintéticos en secuencias génicas, de manera que un conjunto degenerado de secuencias de proteínas potenciales es expresable como polipéptidos individuales o, de manera alternativa, como un conjunto de proteínas de fusión más grandes *(por ejemplo,* para la visualización de fagos). Hay una variedad de métodos que pueden usarse para producir bibliotecas de variantes potenciales de las proteínas marcadoras a partir de una secuencia degenerada de oligonucleótidos. Los métodos para sintetizar oligonucleótidos degenerados son conocidos en la técnica(véase, por ejemplo, Narang, 1983, *Tetrahedron* 39:3; Itakura *et ál.*, 1984, *Annu. Rev. Biochem.* 53:323; Itakura *et ál.*, 1984, *Science* 198:1056; Ike *et ál.*, 1983 *Nucleic Acid Res.* 11:477).

Además, las bibliotecas de segmentos de una proteína marcadora pueden usarse para generar una población variegada de péptidos para el cribado y subsiguiente selección de variantes de proteínas marcadoras y segmentos de estas. Por ejemplo, una biblioteca de fragmentos de secuencia codificante puede generarse mediante el tratamiento de un fragmento de PCR de cadena doble de la secuencia codificante de interés con una nucleasa bajo condiciones donde ocurra la actividad de corte, solo aproximadamente una vez por molécula, desnaturalizando el ADN de doble cadena, renaturalizando el ADN para formar un ADN de doble cadena que puede incluir pares sentido/antisentido de diferentes productos de corte, eliminar las partes de cadena simple de los dúplex reformados mediante el tratamiento con nucleasa S1, y ligando la biblioteca de fragmentos resultante en un vector de expresión. Mediante este método, una biblioteca de expresión puede derivarse para que codifique fragmentos internos y del extremo amino de varios tamaños de la proteína de interés.

Se conocen en la técnica varias técnicas para cribar los productos génicos de bibliotecas combinatorias hechas a partir de mutaciones puntuales o truncamiento, y para cribar las bibliotecas de ADNc para verificar los productos génicos que tienen una propiedad seleccionada. Las técnicas usadas más ampliamente, que son adaptables al análisis de alto rendimiento,

para cribar bibliotecas génicas grandes típicamente incluyen la clonación de la biblioteca génica en vectores de expresión replicables, transformar las células adecuadas con los vectores de la biblioteca resultante, y expresar los genes combinatorios bajo condiciones en las cuales la detección de una actividad deseada facilite el aislamiento del vector que codifica el gen cuyo producto fue detectado. La mutagénesis recursiva de conjunto (REM), una técnica que mejora la frecuencia de mutantes funcionales en las bibliotecas, puede usarse en combinación con los ensayos de cribado para identificar variantes de una proteína de la invención (Arkin y Yourvan, 1992, *Proc. Nati. Acad. Sci. USA* 89:7811-7815; Delgrave *et ál,* 1993, *Protein Engineering* 6(3):327- 331).

Otro aspecto de la invención se refiere a anticuerpos dirigidos contra una proteína la invención. En realizaciones preferidas, los anticuerpos se unen específicamente a una proteína marcadora o a un fragmento de esta. Los términos "anticuerpo" y "anticuerpos" tal como se usan en la presente se refieren a moléculas de inmunoglobulina, así como a fragmentos y derivados de esta, que comprenden una porción inmunológicamente activa de una molécula de inmunoglobulina, *(es decir,* dicha porción contiene un sitio de unión al antígeno que se une específicamente a un antígeno, de manera tal como una proteína marcadora, *por ejemplo,* un epítopo de una proteína marcadora). Un

anticuerpo que se une específicamente a una proteína de la invención es un anticuerpo que se une a la proteína, pero sustancialmente no se une a otras moléculas en una muestra, *por ejemplo,* una muestra biológica que contiene naturalmente la proteína.. Los ejemplos de una porción inmunológicamente activa de una molécula de inmunoglobulina incluyen, de modo no taxativo, anticuerpos de cadena simple (scAb), fragmentos F(ab) y $F(ab')_2$.

Una proteína aislada de la invención o un fragmento de esta puede usarse como inmunógeno para generar anticuerpos. Puede usarse la proteína de longitud completa o, de manera alternativa, la invención proporciona fragmentos de péptido antigénico para su uso como inmunógenos. El péptido antigénico de una proteína de la invención comprende al menos 8 (preferentemente 10, 15, 20, o 30 o más) residuos aminoácido de la secuencia de aminoácidos de una de las proteínas de la invención, y comprende al menos un epítopo de la proteína de manera que un anticuerpo generado contra el péptido forma un complejo inmunitario específico con la proteína. Los epítopos preferidos comprendidos por el péptido antigénico son regiones que están ubicadas en la superficie de la proteína, por ejemplo, regiones hidrofilias. Pueden usarse análisis de secuencia de hidrofobicidad, análisis de secuencia de hidrofilicidad, o análisis similares para identificar regiones hidrofilias. En realizaciones preferidas, una proteína marcadora aislada o fragmento de esta se usa como inmunógeno.

Se usa comúnmente un inmunógeno para preparar anticuerpos mediante la inmunización de un sujeto adecuado *(por ejemplo* inmunocompetente) tal como un conejo, cabra, ratón, u otro mamífero o vertebrado. Una preparación inmunogénica adecuada puede contener, por ejemplo, proteínas o péptidos expresados de manera recombinante o sintetizados químicamente. La preparación puede incluir adicionalmente un adyuvante, tal como un adyuvante de Freund completo o incompleto, o un agente inmunoestimulador similar. Las composiciones inmunógenas preferidas son aquellas que no contienen otras proteínas humanas tal como, por ejemplo, composiciones inmunógenas preparadas usando una célula hospedadora no humana para la expresión recombinante de una proteína de la invención. De esta manera, las composiciones de anticuerpo resultantes tienen bajo contenido o no tienen proteínas humanas que no sean la proteína de la invención.

La invención proporcionar anticuerpos policlonales o monoclonales. El término "anticuerpo monoclonal" o "composición de anticuerpos monoclonales", tal como se usa en la presente, se refiere a una población de moléculas de anticuerpo que contienen solamente una especia de un sitio de unión al antígeno capaz de inmunorreaccionar con un epítopo particular. Las composiciones de anticuerpos policlonales y monoclonales preferidas son aquellas que se seleccionaron para anticuerpos dirigidos contra una

proteína de la invención. Las preparaciones de anticuerpos policlonales y monoclonales particularmente preferidas son aquellas que contienen solamente anticuerpos dirigidos contra una proteína marcadora o un fragmento de esta.

Los anticuerpos policlonales pueden prepararse mediante la inmunización de un sujeto adecuado con una proteína de la invención como inmunógeno. La titulación del anticuerpo en el sujeto inmunizado puede monitorearse en el tiempo mediante técnicas estándar, tal como ensayo inmunoabsorbente ligado a enzimas (ELISA) usando polipéptidos inmovilizados. En un momento adecuado luego de la inmunización, por ejemplo, cuando los títulos de anticuerpo específico están más elevados, las células que producen anticuerpos pueden obtenerse del sujeto y usarse para preparar anticuerpos monoclonales (mAb) mediante técnicas estándar, tal como la técnica de hibridoma descrita originalmente por Kohler y Milstein (1975) *Nature* 256:495-497, la técnica del hibridoma de células B (véase Kozbor *et ál.*, 1983, *Immunol. Today* 4:72), la técnica de EBV-hibridoma (véase Cole *et ál.*, pp. 77-96 In *Monoclonal Antibodies and Cancer Therapy,* Alan R. Liss, Inc., 1985) o técnicas de trioma. La tecnología para producir hibridomas (véase en general *Current Protocols in Immunology,* Coligan *et ál.* ed., John Wiley & Sons, Nueva York, 1994). Las células de hibridoma que producen un anticuerpo monoclonal de la invención se detectan mediante el cribado de sobrenadantes del

cultivo de hibridoma para detectar anticuerpos que se unen al polipéptido de interés, por ejemplo, usando un ensayo ELISA estándar.

De manera alternativa para preparar hibridomas que secretan anticuerpos monoclonales, un anticuerpo monoclonal dirigido contra una proteína de la invención puede identificarse y aislarse mediante cribado de una biblioteca combinatoria recombinante *(por ejemplo,* una biblioteca de visualización de fagos de anticuerpos) con el polipéptido de interés. Los kits para generar y cribar bibliotecas de visualización de fagos se encuentran disponibles comercialmente *(por ejemplo,* el Sistema de Anticuerpos de Fagos Recombinantes *de Pharmacia,* N.° de catálogo 27-9400-01; y el Kit de Visualización de Fagos SurfZAP *de Stratagene,* N.° de catálogo 240612). Adicionalmente, los ejemplos de métodos y reactivos particularmente adaptables para su uso en la generación y el cribado de la biblioteca de visualización pueden encontrarse en, por ejemplo, patente estadounidense N.° 5.223.409; publicación PCT N.° WO 92/18619; publicación PCT N.° WO 91/17271; publicación PCT N.° WO 92/20791; publicación PCT N.° WO 92/15679; publicación PCT N.° WO 93/01288; publicación PCT N.° WO 92/01047; publicación PCT N.° WO 92/09690; publicación PCT N.° WO 90/02809; Fuchs *et ál.* (1991) *Bio/Technology* 9:1370-1372; Hay *et ál.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et ál.*

(1989) *Science* 246:1275- 1281; Griffiths *et ál.* (1993) *EMBO J.* 12:725-734.

La invención también proporciona anticuerpos recombinantes que se unen específicamente a una proteína de la invención. En realizaciones preferidas, los anticuerpos recombinantes se unen específicamente a una proteína marcadora o a un fragmento de esta. Los anticuerpos recombinantes incluyen, de modo no taxativo, anticuerpos monoclonales quiméricos y humanizados, que comprenden tanto partes humanas como no humanas, anticuerpos de cadena simple y anticuerpos multiespecíficos. Un anticuerpo quimérico es una molécula en la cual diferentes partes se derivan de diferentes especies animales, tal como aquellos que tienen una región variable derivada de un mAb murino y una región constante de inmunoglobulina humana. (Véase, *por ejemplo,* Cabilly et ál., patente estadounidense N.° 4.816.567; y Boss et ál., patente estadounidense N.° 4.816.397, que se incorporan a la presente en su totalidad mediante esta referencia.) Los anticuerpos de cadena simple tienen un sitio de unión al antígeno que consiste en un único polipéptido. Pueden producirse mediante técnicas conocidas en la técnica, por ejemplo, usando los métodos descritos en Ladner *et. ál* patente estadounidense N.° 4.946.778 (que se incorpora a la presente mediante esta referencia en su totalidad); Bird *et ál.,* (1988) *Science* 242:423-426; Whitlow *et ál.,* (1991) *Methods in Enzymology* 2:1-9; Whitlow *et ál.,* (1991)

Methods in Enzymology 2:97-105; y Huston *et ál.*, (1991) *Methods in Enzymology Molecular Design and Modeling: Concepts and Applications* 203:46-88. Los anticuerpos multiespecíficos son moléculas de anticuerpo que tienen al menos dos sitios de unión al antígeno que se unen específicamente a antígenos diferentes. Dichas moléculas pueden producirse mediante técnicas conocidas en la técnica, por ejemplo usando los métodos descritos en Segal, patente estadounidense N.° 4.676.980 (cuya divulgación se incorpora a la presente mediante esta referencia en su totalidad); Holliger et ál., (1993) *Proc. Nati. Acad. Sel USA* 90:6444-6448; *Whitlow et ál,* (1994) *Protein Eng.* 7:1017-1026 y la patente estadounidense N.° 6.121.424.

Los anticuerpos humanizados son moléculas de anticuerpo de especias no humanas que tienen una o más regiones determinantes de la complementariedad (CDR) de la especia no humana y una región de marco de una molécula de inmunoglobulina humana. (Véase, por ejemplo, Queen, patente estadounidense N.° 5.585.089, que se incorpora a la presente mediante esta referencia en su totalidad.) Los anticuerpos monoclonales humanizados pueden producirse mediante técnicas de ADN recombinante conocidas en la técnica, por ejemplo usando los métodos descritos en la publicación PCT N.° WO 87/02671; solicitud de patente europea 184.187; solicitud de patente europea 171.496; solicitud de patente europea 173.494; publicación PCT N.° WO 86/01533; patente

estadounidense N.° 4.816.567; solicitud de patente europea 125.023; Better *et ál.* (1988) *Science* 240:1041-1043; Liu *et ál.* (1987) *Proc. Nati. Acad. Sci. USA* 84:3439-3443; Liu *et ál.* (1987) J. *Immunol.* 139:3521- 3526; Sun *et ál.* (1987) *Proc. Nati. Acad. Sci. USA* 84:214-218; Nishimura *et ál.* (1987) *Cancer Res.* 47:999-1005; Wood *et ál.* (1985) *Nature* 314:446-449; y Shaw *et ál.* (1988) J. *Nati. CancerInst.* 80:1553-1559); Morrison (1985) *Science* 229:1202-1207; Oi *et ál.* (1986) *Bio/Techniques* 4:214; patente estadounidense 5.225.539; Jones *et ál.* (1986) *Nature* 321:552-525; Verhoeyan *et ál.* (1988) *Science* 239:1534; y Beidler *et ál.* (1988) J. *Immunol.* 141:4053-4060.

Más particularmente, los anticuerpos humanizados pueden producirse, por ejemplo, usando ratones transgénicos que son incapaces de expresar genes de cadena pesada y ligera de inmunoglobulina endógena, pero que pueden expresar genes de cadena pesada y ligera humanos. Los ratones transgénicos están inmunizados de manera normal con un antígeno seleccionado, *por ejemplo,* la totalidad o una parte de un polipéptido correspondiente a un marcador de la invención. Pueden obtenerse anticuerpos monoclonales dirigidos contra el antígeno usando tecnología de hibridoma convencional. Los transgenes de inmunoglobulina humana albergados por los ratones transgénicos se redistribuyen durante la diferenciación de células B, y subsecuentemente experimentan un cambio de clase y mutación

somática. Por lo tanto, usando dicha técnica, es posible producir anticuerpos de IgG, IgA e IgE terapéuticamente útiles. Por un resumen de esta tecnología para producir anticuerpos humanos, véase Lonberg y Huszar (1995) *Int. Rev. Immunol.* 13:65-93). Por una discusión detallada de esta tecnología para producir anticuerpos humanos y anticuerpos monoclonales humanos y protocolos para producir dichos anticuerpos, véase, *por ejemplo,* la patente estadounidense 5.625.126; patente estadounidense 5.633.425; patente estadounidense 5.569.825; patente estadounidense 5.661.016; y la patente estadounidense 5.545.806. Además, las empresas como Abgenix, Inc. (Freemont, CA), pueden proporcionar anticuerpos humanos dirigidos contra un antígeno seleccionado usando tecnología similar a la descrita anteriormente.

Los anticuerpos completamente humanos que reconocen un epítopo seleccionado pueden generarse usando una técnica mencionada como "selección guiada". En este abordaje un anticuerpo monoclonal no humano seleccionado, por ejemplo, un anticuerpo murino, se usa para guiar la selección de un anticuerpo completamente humano que reconoce el mismo epítopo (Jespers *et ál,* 1994, *Bio/technology* 12:899-903).

Los anticuerpos de la invención pueden aislarse después de la producción (por ejemplo, de la sangre o del suero del sujeto)

o de la síntesis, y pueden purificarse adicionalmente mediante técnicas conocidas. Por ejemplo, los anticuerpos de IgG pueden purificarse usando cromatografía de proteína A. Los anticuerpos específicos para una proteína de la invención pueden seleccionarse o *(por ejemplo,* purificarse parcialmente) o purificarse mediante, *por ejemplo,* cromatografía de afinidad. Por ejemplo, una proteína expresada de manera recombinante y purificada (o parcialmente purificada) de la invención se produce como se describió en la presente, y se acopla de manera covalente o no covalente a un soporte sólido tal como, por ejemplo, una columna de cromatografía. La columna puede entonces usarse para purificar por afinidad anticuerpos específicos para las proteínas de la invención de una muestra que contiene anticuerpos dirigidos contra una cantidad grade de diferentes epítopos, generando de esta manera una composición de anticuerpos sustancialmente purificada, es decir, la cual está sustancialmente libre de anticuerpos contaminantes. Composición de anticuerpos sustancialmente purificada significa, en este contexto, que la muestra de anticuerpos contienen como máximo únicamente 30% (en peso en seco) de anticuerpos contaminantes dirigidos contra epítopos distintos a aquellos de las proteínas deseadas de la invención, y preferentemente como máximo un 20%, más preferentemente como máximo un 10% y todavía más preferentemente como máximo un 5% (en peso en seco) de la muestra está compuesto de anticuerpos contaminantes. Una composición de anticuerpos

purificados significa que al menos 99% de los anticuerpos de la composición están dirigidos contra la proteína deseada de la invención.

En una realización preferida, los anticuerpos sustancialmente purificados de la invención pueden unirse específicamente a un péptido de señalización, a una secuencia secretada, a un dominio extracelular, a un domino transmembrana o citoplasmático, o a la membrana citoplasmática de una proteína de la invención. En una realización particularmente preferida, los anticuerpos sustancialmente purificados de la invención se unen específicamente a una secuencia secretada o a un dominio extracelular de las secuencias de aminoácidos de una proteína de la invención. En una realización más preferida, los anticuerpos sustancialmente purificados de la invención se unen específicamente a una secuencia secretada o a un dominio extracelular de las secuencias de aminoácidos de una proteína marcadora.

Un anticuerpo dirigido contra una proteína de la invención puede usarse para aislar la proteína por medio de técnicas estándar, tal como cromatografía de afinidad o inmunoprecipitación. Además, dicho anticuerpo puede usarse para detectar la proteína marcadora o un fragmento de esta (por ejemplo, en un lisado celular o en un sobrenadante celular) para

poder evaluar el nivel y el patrón de expresión del marcador. Los anticuerpos también pueden usarse como diagnóstico para monitorear los niveles de proteína en tejidos o fluidos corporales (por ejemplo en un fluido asociado con una enfermedad o estado de toxicidad) como parte de un procedimiento te análisis clínico, por ejemplo, para, por ejemplo, determinar la eficacia de un régimen de tratamiento dado. La detección puede facilitarse mediante el uso de un derivado de anticuerpo, que comprende un anticuerpo de la invención acoplado a una sustancia detectable. Los ejemplos de sustancias detectables incluyen varias enzimas, grupos prostéticos, materiales fluorescentes, materiales luminiscentes, materiales bioluminiscentes y materiales radiactivos. Los ejemplos de enzimas adecuadas incluyen peroxidasa de rábano picante, fosfatasa alcalina, α-galactosidasa o acetilcolinesterasa; los ejemplos de complejos de grupos prostéticos adecuados incluyen estreptavidina/biotina y avidina/biotina; los ejemplos de materiales fluorescentes adecuados incluyen umbeliferona, fluoresceina, isotiocianato de fluoresceina, rodamina, diclorotriazinilamina fluoresceina, cloruro de dansilo o ficoeritrina; un ejemplo de un material luminiscente adecuado incluye luminol; los ejemplos de materiales bioluminiscentes adecuados incluyen luciferasa, luciferina y aequorina, y los ejemplos de materiales radiactivos adecuados incluyen ^{125}I, ^{131}I, ^{35}S o ^{3}H.

Los anticuerpos de la invención también pueden usarse como agentes terapéuticos en el tratamiento del cáncer. En una realización preferida, los anticuerpos completamente humanos de la invención se usan para el tratamiento terapéutico de pacientes de cáncer humanos, particularmente aquellos que tienen cáncer. En otra realización preferida, los anticuerpos que se unen específicamente a una proteína marcadora o a un fragmento de esta se usan para el tratamiento terapéutico. Además, dicho anticuerpos terapéutico puede ser un derivado de anticuerpo o inmunotoxina que comprende un anticuerpo conjugado a un resto terapéutico, tal como una citotoxina, un agente terapéutico o un ion metálico radiactivo. Una citotoxina o un agente citotóxico incluye cualquier agente que es dañino para las células. Los ejemplos incluyen taxol, citocalasina B, gramicidina D, bromuro de etidio, emetina, mitomicina, etopósido, tenopósido, vincristina, vinblastina, colquicina, doxorrubicina, daunorrubicina, dihidroxi antracin diona, mitoxantrona, mitramicina, actinomicina D, 1-dehidrotestosterona, glucocorticoides, procaína, tetracaína, lidocaína, propranolol, y puromicina, y análogos y homólogos de estos. Los agentes terapéuticos incluyen, de modo no taxativo, antimetabolitos (por ejemplo, metotrexato, 6-mercaptopurina, 6-tioguanina, citarabina, 5-fluorouracilo decarbazina), agentes alquilantes (por ejemplo,mecloretamina, tioepa clorambucilo, melfalán, carmustina (BSNU) y lomustina (CCNU), ciclotosfamida, busulfán,

dibromomanitol, estreptozotocina, mitomicina C y cis-diclorodiamina platino (II) (DDP) cisplatino), antraciclinas (por ejemplo, daunorrubicina (anteriormente daunomicina) y doxorrubicina), antibióticos *(por ejemplo,* dactinomicina (anteriormente actinomicina), bleomicina, mitramicina y antramicina (AMC)), y agentes anti-mitóticos *(por ejemplo,* vincristina y vinblastina).

Los anticuerpos conjugados de la invención pueden ser usados para modificar una respuesta biológica dada, dado que el resto de fármaco no se entiende como limitado a agentes terapéuticos químicos clásicos. Por ejemplo, el resto de fármaco puede ser una proteína o por ejemplo que posee una actividad biológica dada. Dichas proteínas pueden incluir, por ejemplo, una toxina tal como una proteína que inhibe ribosomas (véase Better et ál., patente estadounidense N.° 6.146.631, cuya divulgación se incorpora a la presente en su totalidad), abrina, ricina A, exotoxinas pseudomonas, o toxina difteria; una proteína como un factor de necrosis tumoral, .alfa.-interferón, β-interferón, factor de crecimiento nervioso, factor de crecimiento derivado de plaquetas, activador de tejido plasminógeno; o modificadores de la respuesta biológica, tales como, por ejemplo, linfocinas, interleucina-1 ("IL-1"), interleucina-2 ("IL-2"), interleucina-6 ("IL-6"), factor estimulante de colonias de macrofase [sic] de

granulocitos ("GM-CSF"), factor estimulante de colonias de granulocitos ("G-CSF") u otros factores de crecimiento.

Las técnicas para conjugar dichos restos terapéuticos con anticuerpos son conocidas, véase, por ejemplo, Arnon et ál., "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", en Monoclonal Antibodies And Cancer Therapy, Reisfeld et ál. (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom et ál., "Antibodies For Drug Delivery", en Controlled Drug Delivery (2ª Ed.), Robinson et ál. (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", en Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera et ál. (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", en Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin et ál. (eds.), pp. 303-16 (Academic Press 1985), y Thorpe et ál., "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", Immunol. Rev., 62:119-58 (1982).

Por consiguiente, en un aspecto, la invención proporciona anticuerpos sustancialmente purificados, fragmentos y derivados de anticuerpos, todos los cuales se unen específicamente a una proteína de la invención y, preferentemente, a una proteína marcadora. En varias realizaciones, los anticuerpos

sustancialmente purificados de la invención, o fragmentos o derivados de estos, pueden ser anticuerpos humanos, no humanos, quiméricos y/o humanizados. En otro aspecto, la invención proporciona anticuerpos no humanos, fragmentos de anticuerpos y derivados, todos los cuales se unen específicamente a una proteína de la invención y, preferentemente, a una proteína marcadora. Dichos anticuerpos no humanos pueden ser anticuerpos de cabra, ratón, oveja, caballo, pollo, conejo o rata. De manera alternativa, los anticuerpos no humanos de la invención pueden ser anticuerpos quiméricos y/o humanizados. Además, los anticuerpos no humanos de la invención pueden ser anticuerpos policlonales y/o anticuerpos monoclonales. En un aspecto adicional, la invención proporciona anticuerpos monoclonales, fragmentos y derivados de anticuerpos, todos los cuales se unen específicamente a una proteína de la invención y, preferentemente, a una proteína marcadora. Los anticuerpos monoclonales pueden ser anticuerpos humanos, quiméricos y/o no humanos.

La invención también proporciona un kit que contiene un anticuerpo de la invención conjugado con una sustancia detectable, e instrucciones de uso. Todavía otro aspecto de la invención es un una composición farmacéutica que comprende un anticuerpo de la invención. En una realización, la composición

farmacéutica comprende un anticuerpo de la invención y un portador farmacéuticamente aceptable.

E. Medicina predictiva

La presente invención pertenece al campo de la medicina predictiva en el cual los ensayos de diagnóstico, ensayos de pronóstico, farmacogenómica y los ensayos clínicos de monitoreo se usan para fines de pronóstico (predictivos) para tratar de esta manera a un individuo profilácticamente. Por consiguiente, un aspecto de la presente invención se refiere a ensayos de diagnóstico para determinar el nivel de expresión de una o más proteínas o ácidos nucleicos marcadores, para poder determinar si un individuo está en riesgo de desarrollar determinada enfermedad o toxicidad inducida por fármacos. Dichos ensayos pueden usarse para fines de pronóstico o predictivos para tratar de esta manera a un individuo profilácticamente antes de la aparición del trastorno.

Todavía otro aspecto de la invención se refiere al monitoreo de la influencia de agentes (por ejemplo, fármacos u otros compuestos administrados ya sea para inhibir o tratar o prevenir un trastorno o toxicidad inducida por fármacos {es *decir* para poder entender cualquier efecto tóxico inducido por fármacos que dicho tratamiento pueda tener}) en la expresión o

actividad de un marcador de la invención en ensayos clínicos. Este y otros agentes de la invención se describen en más detalle en las siguientes secciones.

F. Ensayos de diagnóstico

Un método ejemplar para detectar la presencia o ausencia de una proteína o ácido nucleico marcadores en una muestra biológica involucra la obtención de una muestra biológica (por ejemplo, muestra de tejido o fluido corporal asociado con la toxicidad) de un sujeto de prueba y poner en contacto la muestra biológica con un compuesto o agente capaz de detectar el polipéptido o ácido nucleico (por ejemplo, ARNm, ADN genómico o ADNc). Los métodos de detección de la invención pueden, por lo tanto, usarse para detectar ARNm, proteínas, ADNc o ADN genómico *in vitro* así como *in vivo*. Por ejemplo, las técnicas *in vitro* para la detección de ARNm incluyen hibridaciones Northern e hibridaciones *in situ*. Las técnicas *in vitro* para detectar una proteína marcadora incluyen ensayos inmunoabsorbentes ligados a enzimas (ELISA), transferencias Western, inmunoprecipitaciones e inmunofluorescencia. Las técnicas *in vitro* para la detección de ADN genómico incluyen hibridaciones Southern. *Las técnicas in vivo* para la detección de ARNm incluyen reacción en cadena de la polimerasa (PCR), hibridaciones Northerne hibridaciones *in situ*. Además, las técnicas *in vivo* para la detección de una proteína

marcadora incluyen introducir en un sujeto un anticuerpo marcado dirigido contra la proteína o fragmento de esta. Por ejemplo, el anticuerpo puede estar marcado con un marcador radiactivo cuya presencia y ubicación en el sujeto pueda detectarse mediante técnicas de imagenología estándar.

Un principio general de dichos ensayos diagnósticos y de pronóstico involucra preparar una muestra o mezcla de reacción que puede contener un marcador y una sonda, bajo condiciones apropiadas y durante un tiempo suficiente para permitir que el marcador y la sonda interactúen y se unan, formando así un complejo que puede ser retirado y/o detectado en la mezcla de reacción. Estos ensayos pueden llevarse a cabo en una variedad de formas.

Por ejemplo, un método para llevar a cabo dicho ensayo involucraría anclar el marcador o sonda a un soporte de fase sólida, también mencionado como sustrato, y detectar los complejos de marcador/sonda objetivos anclados en la fase sólida al final de la reacción. En una realización de dicho método, una muestra de un sujeto, que va a analizarse para detectar la presencia y/o concentración de marcador, puede anclarse a un soporte de fase sólida o portador. En otra realización, la situación inversa es posible, en la cual la sonda puede anclarse

a una fase sólida y una muestra del sujeto puede dejarse reaccionar como un componente sin anclar del ensayo.

Hay muchos métodos establecidos para anclar componentes de ensayo a una fase sólida. Estos incluyen, de modo no taxativo, molécula de marcador o sonda que están inmovilizadas a través de conjugación de biotina o estreptavidina. Dichos componentes de ensayo biotinilados pueden prepararse a partid de biotina-NHS (N-hidroxi-succinimida) usando técnicas conocidas en la técnica *(por ejemplo,* kit de biotinilado, Pierce Chemicals, Rockford, IL), e inmovilizarse en pocillos de placas de 96 pocillos recubiertos de estreptavidina (Pierce Chemical). En ciertas realizaciones, las superficies con componentes de ensayo inmovilizados pueden prepararse por adelantado y almacenarse.

Otros portadores o soportes de fase sólida adecuados para dichos ensayos incluyen cualquier material capaz de unirse a la clase de molécula a la cual pertenece el marcador o sonda. Los soportes y portadores conocidos incluyen, de modo no taxativo, vidrio, poliestireno, nylon, polipropileno, nylon, polietileno, dextrano, amilasas, celulosas naturales y modificadas, poliacrilamidas, gabros y magnetita.

Para poder llevar a cabo los ensayos con los abordajes antemencionados, el componente no inmovilizado se agrega a la

fase sólida a la cual está anclado el segundo componente. Después de que la reacción está completa, los componentes que no están en forma de complejo pueden retirarse (por ejemplo, mediante lavado) bajo condiciones de manera que cualquier complejo formado permanecerá inmovilizado en la fase sólida. La detección de los complejos de marcador/sonda anclados a la fase sólida puede llevarse a cabo en una cantidad de métodos descritos en la presente.

En una realización preferida, la sonda, cuando es el componente no anclado del ensayo, puede marcarse con fines de detección y lectura del ensayo, ya sea directa o indirectamente, con marcadores detectables discutidos en la presente y que son conocidos para los expertos en la técnica.

También es posible detectar directamente la formación del complejo de marcador/sonda sin manipulación o marcado adicionales de cualquiera de los componentes (marcador o sonda), por ejemplo, mediante la utilización de la técnica de transferencia de energía fluorescente, Lakowicz *et ál.*, patente estadounidense N.° 5.631.169; Stavrianopoulos, *et ál.*, patente estadounidense N.° 4.868.103). Un marcador fluoróforo en la primera molécula "donante" se selecciona de manera que al excitarla con luz incidente de una longitud de onda adecuada, su energía fluorescente emitida será absorbida por una etiqueta fluorescente

en una segunda molécula "aceptora", que, a su vez, es capaz de generar fluorescencia debido a la energía absorbida. De manera alternativa, la molécula de proteína "donante" puede usar simplemente la energía fluorescente natural de los residuos de triptófano. Los marcadores se eligen de manera que emitan luz de diferentes longitudes de onda, de manera que la etiqueta de la molécula "aceptora" puede diferenciarse del de la "donante". Dado que la eficiencia de le transferencia de energía entre los marcadores se relaciona con la distancia que separa las moléculas, las relaciones espaciales entre las moléculas puede analizarse. En una situación en la cual la unión ocurre entre las moléculas, la emisión fluorescente de la etiqueta de la molécula "aceptora" en el ensayo debe ser máxima. Puede medirse de manera conveniente un evento de unión FET a través de medios de detección fluorométrica conocidos en la técnica *(por ejemplo,* usando un fluorímetro).

En otra realización, la determinación de la capacidad de una sonda de reconocer un marcador puede lograrse sin marcar ninguno de los componentes del ensayo (sonda o marcador) utilizando una tecnología tal como el Análisis de Interacción Biomolecular (BIA) de tiempo real (véase, *por ejemplo,* Sjolander, S. yUrbaniczky, C, 1991, *Anal. Chem.* 63:2338-2345 y Szabo *et ál.*, 1995, *Curr. Opin. Struct. Biol.* 5:699-705). Tal como se usa en la presente, "BIA" o "resonancia de plasmón de superficie" es una

tecnología para estudiar interacciones bioespecíficas en tiempo real, sin marcar ninguno de los elementos que interactúan (por ejemplo, BIAcore). Los cambios en la masa en la superficie de unión (indicador de un evento de unión) da como resultado alteraciones del índice de refracción de la luz cerca de la superficie (el fenómeno óptico de la resonancia de plasmón de superficie (SPR)), lo cual da como resultado una señal detectable que puede usarse como un indicador de reacciones en tiempo real entre las moléculas biológicas.

De manera alternativa, en otra realización, se pueden llevar a cabo análisis de diagnóstico y pronóstico análogos con el marcador y la sonda como solutos en una fase líquida. En dicho ensayo, el marcador y sonda en complejo se separan de los componentes que no están en complejo mediante una cantidad de técnicas estándar, incluyendo, de modo no taxativo: centrifugación diferencial, cromatografía, electroforesis e inmunoprecipitación. En la centrifugación diferencial, los complejos de marcador/sonda pueden separarse de los componentes de ensayo que no están en complejo a través de una serie de etapas de centrifugación, debido a los diferentes equilibrios de sedimentación de los complejos basados en sus diferentes tamaños y densidades (véase, por ejemplo, Rivas, G., y Minton, A.P., 1993, *Trenas Biochem Sci.* 18(8):284-7). Las técnicas de cromatografía estándar pueden también utilizarse para separar

moléculas en complejo de las que no están en complejo. Por ejemplo, la cromatografía de filtración de gel separa las moléculas basándose en el tamaño y a través de la utilización de una resina de filtración de gel adecuada en un formato de columna, por ejemplo, los componentes relativamente más grandes del complejo pueden separarse de los relativamente más pequeños. De manera similar, las propiedades de carga relativamente diferentes del complejo de marcador/sonda cuando se compara con los componentes que no están en complejo pueden explotarse para diferenciar los componentes del complejo de los que no están en complejo, por ejemplo a través de la utilización de resinas de cromatografía de intercambio de iones. Dichas resinas y técnicas cromatográficas son conocidas por los expertos en la técnica (véase, *por ejemplo,* Heegaard, N.H., 1998, J. *Mol. Recognit.* Winter 11(16): 141-8; Hage, D.S., y Tweed, S.A. *J Chromatogr B Biomed Sci Appl* 1997 10 de oct;699(1-2):499-525). También puede emplearse la electroforesis en gel para separar los componentes de ensayo en complejo de los componentes no unidos (véase, *por ejemplo,* Ausubel *et ál.,* ed., *Current Protocols in Molecular Biology,* John Wiley & Sons, Nueva York, 1987-1999). En esta técnica, los complejos de proteína o ácido nucleico se separan basándose en tamaño o carga, por ejemplo. Para poder mantener la interacción de unión durante el proceso electroforético, se prefieren típicamente los materiales y condiciones de matriz de gel no desnaturalizantes en la ausencia de agente reductor. Las

condiciones adecuadas para el ensayo particular y componentes de este serán conocidas para un experto en la técnica.

En una realización particular, el nivel de ARNm marcador puede determinarse tanto mediante formatos *in situ* como *in vitro* en una muestra biológica usando métodos conocidos en la técnica. El término "muestra biológica" se pretende que incluya tejidos, células, fluidos biológicos y aislados de estos, aislados de un sujeto, así como tejidos, células y fluidos presentes dentro de un sujeto. Muchos métodos de detección de la expresión usan ARN aislado. Para los métodos *in vitro*, puede usarse cualquier técnica de aislamiento que no evite el aislamiento del ARNm para la purificación del ARN de las células (véase, *por ejemplo*, Ausubel *et ál.*, ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, Nueva York 1987-1999). Adicionalmente, grandes cantidades de muestras de tejido pueden procesarse fácilmente usando técnicas conocidas por los expertos en la técnica, tal como, por ejemplo, el proceso de aislamiento de ARN de un solo paso de Chomczynski (1989, patente estadounidense N.° 4.843.155).

El ARNm aislado puede usarse en los ensayos de hibridación o amplificación que incluyen, de modo no taxativo, análisis Southern o Northern, análisis de reacción en cadena de la polimerasa y matrices de sondas. Un método de diagnóstico preferido para la detección de los niveles de ARNm involucra

poner en contacto el ARNm aislado con una molécula de ácido nucleico (sonda) que pueda hibridarse con el ARNm codificado por el gen que se está detectando. La sonda de ácido nucleico puede ser, por ejemplo, un ADNc de longitud completa o una parte de este, tal como un oligonucleótido de al menos 7, 15, 30, 50, 100, 250 o 500 nucleótidos de longitud y suficiente para hibridarse de manera específica bajo condiciones estrictas a un ARNm o ADN genómico que codifica un marcador de la presente invención. Otras sondas adecuadas para su uso en ensayos diagnósticos de la invención se describen en la presente. La hibridación de un ARNm con la sonda indica que el marcador en cuestión está siendo expresado.

En un formato, el ARNm se inmoviliza en una superficie sólida y se pone en contacto con una sonda, por ejemplo, mediante el pase del ARNm aislado en un gel de agarosa y transfiriendo el ARNm desde el gel a una membrana, tal como nitrocelulosa. En un formato alternativo, la o las sondas se inmovilizan en una superficie sólida y el ARNm se pone en contacto con la o las sondas, por ejemplo, en un chip de ADN Affymetrix. Un experto en la técnica puede adaptar fácilmente métodos de detección de ARNm conocidos para su uso en la detección del nivel de ARNm codificado por los marcadores de la presente invención.

Un método alternativo para determinar el nivel de marcador de ARNm en una muestra involucra el proceso de amplificación de ácido nucleico, por ejemplo, mediante RT-PCR (la realización experimental establecida en Mullis, 1987, patente estadounidense N.° 4.683.202), reacción en cadena de la ligasa (Barany, 1991, *Proc. Nati. Acad. Sci. USA,* 88:189-193), replicación de secuencia autosostenida (Guatelli *et ál.,* 1990, *Proc. Nati. Acad. Sel USA* 87:1874-1878), sistema de amplificación transcriptional (Kwoh *et ál.,* 1989, *Proc. Nati. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicasa (Lizardi *et ál.,* 1988, *Bio/Technology* 6:1197), replicación de círculo rodante (Lizardi *et ál.,* patente estadounidense N.° 5.854.033) o cualquier otro método de amplificación del ácido nucleico, seguido de la detección de moléculas amplificadas usando técnicas conocidas para los expertos en la técnica. Estos esquemas de detección son especialmente útiles para la detección de moléculas de ácido nucleico si dichas moléculas están presentes en cantidades muy bajas. Tal como se usa en la presente, los cebadores se definen como un par de moléculas de ácido nucleico que pueden aparear las regiones 5' y 3' de un gen (cadenas positiva y negativa, respectivamente, y viceversa) y contener una región corta entre estas. En general, los cebadores de amplificación son de aproximadamente 10 a 30 nucleótidos de longitud y flanquean una región de aproximadamente 50 a 200 nucleótidos de longitud. Bajo las condiciones adecuadas y con los reactivos adecuados, dichos

cebadores permiten la amplificación de una molécula de ácido nucleico que comprende la secuencia de nucleótidos flanqueada por los cebadores.

Para los métodos *in situ*, no es necesario aislar el ARNm antes de la detección. En dichos métodos, una muestra de células o tejido se prepara/procesa usando métodos histológicos conocidos. La muestra entonces se inmoviliza en un soporte, comúnmente un portaobjetos de vidrio, y luego se pone en contacto con la sonda que puede hibridarse al ARNm que codifica el marcador.

Como una forma alternativa a realizar las determinaciones basadas en el nivel de expresión absoluto del marcador, las determinaciones pueden basarse en el nivel de expresión normalizado del marcador. Los niveles de expresión se normalizan mediante la corrección del nivel de expresión absoluto mediante la comparación de su expresión con la expresión de un gen que no es un marcador, *por ejemplo,* un gen de mantenimiento que se expresa de manera constitutiva. Los genes adecuados para la normalización incluyen los genes de mantenimiento tales como genes de actina, o genes específicos de células epiteliales. Esta normalización permite la comparación del nivel de expresión en una muestra, *por ejemplo,* una muestra de un paciente, con otra muestra, por ejemplo, una muestra que no pertenece a la

enfermedad o que no es tóxica, o entre muestras de diferentes fuentes.

De manera alternativa, el nivel de expresión puede proporcionarse como un nivel de expresión relativo. Para determinar un nivel de expresión relativo de un marcador, el nivel de expresión del marcador se determina para 10 o más muestras de aislados de células normales en comparación con aislados de células de la enfermedad o tóxicas, preferentemente 50 o más muestras, antes de la determinación del nivel de expresión de la muestra en cuestión. EL nivel de expresión medio de cada uno de los genes analizados en el número mayor de muestras se determinó y este se usa como valor de referencia del nivel de expresión para el marcador. El nivel de expresión del marcador determinado para la mejor muestra (nivel de expresión absoluto) entonces se dividió entre el valor de expresión medio obtenido para el marcador. Esto proporciona un nivel de expresión relativo.

Preferentemente, las muestras usadas en la determinación del valor de referencia estarán en forma de células que no son de la enfermedad o que no son tóxicas. La elección del a fuente de células depende del uso del nivel de expresión relativo. Usar la expresión encontrada en los tejidos normales como puntaje de expresión medio ayuda a la validación de si el marcador analizado

es específico de la enfermedad o toxicidad (en comparación con células normales). Además, cuando se acumulan más datos, el valor de expresión medio se puede revisar, siempre que mejore los niveles de expresión relativos basados en los datos acumulados. Los datos de expresión de células de la enfermedad o células tóxicas proporcionan un medio para evaluar la gravedad de esta enfermedad o estado tóxico.

En otra realización de la presente invención, se detecta una proteína marcadora. Un agente preferido para la detección de la proteína marcadora de la invención es un anticuerpo capaz de unirse a dicha proteína o a un fragmento de esta, preferentemente un anticuerpo con una etiqueta detectable. Los anticuerpos pueden ser policlonales o, más preferentemente, monoclonales. Puede usarse un anticuerpo intacto o un fragmento o derivado de este (por ejemplo, Fab o $F(ab')_2$) El término "marcado", con respecto a la sonda o anticuerpo, se pretende que abarque el marcado directo de la sonda o anticuerpo mediante acoplamiento (es *decir,* enlace físico) de una sustancia detectable con la sonda o anticuerpo, así como al marcado indirecto de la sonda o anticuerpo mediante la reactividad con otro reactivo que está marcado directamente. Los ejemplos de marcado indirecto incluyen la detección de un anticuerpo primario usando un anticuerpo secundario marcado de manera fluorescente y al marcado en el extremo de una sonda de

ADN con biotina, de manera que pueda ser detectado con estreptavidina marcada mediante fluorescencia.

Las proteínas de las células pueden aislarse usando técnicas que son conocidas para los expertos en la técnica. Los métodos de aislamiento de proteínas empleados pueden ser, por ejemplo, los descritos en Harlow y Lane (Harlow y Lane, 1988, *Antibodies: A Laboratory Manual,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York).

Puede emplearse una variedad de formatos para determinar si una muestra contiene una proteína que se une a un anticuerpo dado. Los ejemplos de dichos formatos incluyen, de modo no taxativo, inmunoensayo de enzimas (EIA), radioinmunoensayo (RIA), análisis de transferencia Western y ensayo inmunoabsorbente ligado a enzimas (ELISA). Un experto en la técnica puede adaptar fácilmente métodos de detección de proteína/anticuerpo conocidos para su uso en la determinación de si las células expresan un marcador de la presente invención.

En un formato, los anticuerpos, o fragmentos o derivados de anticuerpo, pueden usarse en los métodos tal como transferencias Western o técnicas de inmunofluorescencia para detectar las proteínas expresadas. En dichos usos, es generalmente preferible inmovilizar ya sea el anticuerpo o las

proteínas en un soporte sólido. Los portadores o soportes de fase sólida adecuados incluyen cualquier soporte capaz de unir un antígeno o un anticuerpo. Los soportes y portadores conocidos incluyen vidrio, poliestireno, polipropileno, polietileno, dextrano, nylon, amilasas, celulosas naturales y modificadas, poliacrilamidas, gabros y magnetita.

Un experto en la técnica conocerá muchos otros portadores adecuados para unir el anticuerpo o antígeno, y será capaz de adaptar dicho soporte para su uso con la presente invención. Por ejemplo, la proteína aislada de células de una enfermedad o células tóxicas puede someterse a electroforesis en gel de poliacrilamida e inmovilizarse en un soporte de fase sólida tal como nitrocelulosa. Entonces puede lavarse el soporte con amortiguadores adecuados seguido de tratamiento con el anticuerpo marcado de manera detectable. El soporte de fase sólida puede entonces lavarse con un amortiguador una segunda vez para retirar el anticuerpo no unido. La cantidad de etiqueta unida en el soporte sólido puede detectarse por medios convencionales.

La invención también abarca kits para detectar la presencia de una proteína o ácido nucleico marcadores en una muestra biológica. Dichos kits pueden usarse para determinar si un sujeto padece o está en riesgo de desarrollar cierta enfermedad o toxicidad inducida por fármacos. Por ejemplo, el kit

puede comprender un compuesto o agente marcado capaz de detectar una proteína o ácido nucleico marcadores en una muestra biológica y medios para determinar la cantidad de la proteína o ARNm en la muestra *(por ejemplo,* un anticuerpo que se une a la proteína o a un fragmento de esta, o una sonda de oligonucleótido que se une al ADN o ARNm que codifica la proteína). Los kits también pueden incluir instrucciones para interpretar los resultados obtenidos usando el kit.

Para los kits basados en anticuerpos, el kit puede comprender, por ejemplo: (1) un primer anticuerpo *(por ejemplo,* unido a un soporte sólido) que se une a una proteína marcadora; y, opcionalmente, (2) un segundo anticuerpo diferente que se une ya sea a la proteína o al primer anticuerpo y está conjugado con una etiqueta detectable.

Para los kits basados en oligonucleótidos, el kit puede comprender, por ejemplo: (1) un oligonucleótido, por ejemplo, *un* oligonucleótido marcado de manera detectable, que se hibrida con una secuencia de ácido nucleico que codifica una proteína marcadora o (2) un par de cebadores útiles para amplificar una molécula de ácido nucleico marcador. El kit también puede comprender, *por ejemplo,* un agente amortiguador, un conservante o un agente estabilizador de proteínas. El kit puede comprender adicionalmente componentes necesarios para detectar una etiqueta

detectable (por ejemplo, una enzima o un sustrato). EL kit también puede contener una muestra de control o una serie de muestras de control que pueden analizarse y compararse con la muestra de prueba. Cada componente del kit puede estar contenido dentro de un recipiente individual y todos los varios recipientes pueden estar dentro de un único empaque, junto con instrucciones para interpretar los resultados de los ensayos llevados a cabo usando el kit.

G. Farmacogenómica

Los marcadores de la invención también son útiles como marcadores farmacogenómicos. Tal como se usa en la presente, un "marcador farmacogenómico" es un marcador bioquímico objetivo cuyo nivel de expresión se correlacione con una respuesta o susceptibilidad específica a un fármaco clínico en un paciente (véase, por ejemplo, McLeod *et ál.* (1999) *Eur. J. Cancer* 35(12): 1650-1652). La presencia o cantidad de la expresión del marcador farmacogenómico está relacionada con la respuesta predicha del paciente y más particularmente de las células enfermas o tóxicas del paciente a la terapia con un fármaco o clase de fármacos específicos. Mediante le evaluación de la presencia o cantidad de la expresión de uno o más marcadores farmacogenómicos en un paciente, puede seleccionarse una terapia de fármacos que es más adecuada para el paciente, o que se predice que tendrá un mayor

grado de éxito. Por ejemplo, basándose en la presencia o cantidad de un ARN o proteína codificada por marcadores específicos de tumor en un paciente, puede seleccionarse un fármaco o un tratamiento que está optimizado para el tratamiento del tumor específico que es más probable que esté presente en el paciente. El uso de marcadores farmacogenómicos, por lo tanto, permite seleccionar o diseñar el tratamiento más adecuado para cada paciente de cáncer sin probar diferentes fármacos o regímenes.

Otro aspecto de los farmacogenómicos se refiere a las afecciones genéticas que alteran la forma en que el cuerpo reacciona a los fármacos. Estas condiciones farmacogenómicas pueden ocurrir ya sea como defectos raros o como polimorfismos. Por ejemplo, la deficiencia de glucosa-6-fosfato deshidrogenasa (G6PD) es una enzimopatía hereditaria común en la cual la complicación clínica principal es la hemolisis después de la ingestión de fármacos oxidantes (fármacos anti-malaria, sulfonamidas, analgésicos, nitrofuranos) y del consumo de habas.

Como realización ilustrativa, la actividad de enzimas metabolizantes de fármacos es un determinante principal tanto de la intensidad como de la duración de la acción de un fármaco. El descubrimiento de polimorfismos genéticos de enzimas metabolizantes de fármacos *(por ejemplo,* N-acetiltransferasa 2 (NAT 2) y citocromo P450, enzimas CYP2D6 y CYP2C19) ha

proporcionado una explicación de por qué algunos pacientes no obtienen los efectos esperados de los fármacos o muestran una respuesta al fármaco exagerada, y una toxicidad grave después de tomar una dosis estándar y segura de un fármaco. Estos polimorfismos se expresan en dos fenotipos en la población, el metabolizador extensivo (EM) y el metabolizador pobre (PM). La prevalencia de PM es diferente entre diferentes poblaciones. Por ejemplo, el gen que codifica CYP2D6 es altamente polimórfico y se han identificado varias mutaciones en PM, lo cual lleva a la ausencia de CYP2D6 funcional. Los metabolizadores pobres de CYP2D6 y CYP2C19 experimentan bastante frecuentemente una respuesta exagerada a los fármacos y efectos secundarios al recibir dosis estándar. Si un metabolito es el resto terapéutico activo, un PM no mostrará una respuesta terapéutica, como se demostró para el efecto analgésico de la codeína mediada por su metabolito de morfina formado por CYP2D6. El otro extremo son los llamados metabolizadores ultrarrápidos que no responden a las dosis estándar. Recientemente, se identificó que la base molecular del metabolismo ultrarrápido se debe a la amplificación génica de CYP2D6.

Por lo tanto, el nivel de expresión de un marcador de la invención en un individuo puede determinarse para seleccionar de esta manera el o los agentes adecuados para el tratamiento terapéutico o profiláctico del individuo. Además, los estudios

farmacogenéticos pueden usarse para aplicar la genotipificación de alelos polimórficos que codifican enzimas que metabolizan fármacos a la identificación del fenotipo de la respuesta a un fármaco de un individuo. Este conocimiento, cuando se aplica a la dosificación o selección de un fármaco, puede evitar reacciones adversas o el fallo terapéutico y por lo tanto mejorar la eficiencia terapéutica y profiláctica cuando se trata a un sujeto con un modulador de la expresión de un marcador de la invención.

H. Ensayos clínicos de monitoreo

El monitoreo de la influencia de los agentes (por ejemplo, compuestos de fármaco) en el nivel de expresión de un marcador de la invención puede aplicarse no solo al cribado básico de fármacos, sino también a ensayos clínicos. Por ejemplo, la eficacia de un agente para afectar la expresión del marcador puede monitorearse en ensayos clínicos de sujetos que reciben el tratamiento para ciertas enfermedades, tal como cáncer, diabetes, obesidad, enfermedad cardiovascular y cardiotoxicidad, o toxicidad inducida por fármacos. En una realización preferida, la presente invención proporciona un método para monitorear la efectividad del tratamiento de un sujeto con un agente (por ejemplo, un agonista, antagonista, peptidomimético, proteína, péptido, ácido nucleico, molécula pequeña y otro candidato a fármaco) que comprende las etapas de (i) obtener una muestra

antes de la administración de un sujeto, antes de la administración del agente; (ii) detectar el nivel de expresión de uno o más marcadores seleccionados de la invención en la muestra antes de la administración; (iii) obtener una o más muestras después de la administración del sujeto; (iv) detectar el nivel de expresión del o los marcadores en las muestras después de la administración; (v) comparara el nivel de expresión del o los marcadores en la muestra antes de la administración con el nivel de expresión del o los marcadores en la o las muestras después de la administración; y (vi) alterar la administración del agente al sujeto en consecuencia. Por ejemplo, la expresión aumentada del o los genes marcadores durante el curso del tratamiento puede indicar una dosificación no efectiva y la conveniencia de aumentar la dosificación. Por el contrario, la expresión disminuida del o los genes marcadores puede indicar la eficacia del tratamiento, y la no necesidad de cambiar la dosificación.

H. Matrices

La invención también incluye una matriz que comprende un marcador de la presente invención. La matriz puede usarse para evaluar la expresión de uno o más genes en la matriz. En una realización, la matriz puede usarse para evaluar la expresión génica en un tejido para establecer la especificidad de tejido de los genes en la matriz. De esta manera, pueden evaluarse hasta

7600 genes de manera simultánea para verificar su expresión. Esto permite un perfil que puede desarrollarse para mostrar una batería de genes expresados específicamente en uno o más tejidos.

Además, dicha determinación cualitativa de la invención permite la cuantificación de la expresión génica. Por lo tanto, puede determinarse no solo la especificidad de tejido, sino también el nivel de expresión de la batería de genes en el tejido. Por lo tanto, los genes pueden agruparse en base a su expresión en el tejido *per se* y al nivel de expresión en ese tejido. Esto es útil, por ejemplo, en la determinación de la relación entre la expresión génica entre tejidos. Por lo tanto, un tejido puede perturbarse y puede determinarse el efecto de la expresión génica en un segundo tejido. En este contexto, puede determinarse el efecto de un tipo de célula en otro tipo de célula en respuesta al estímulo biológico. Dicha determinación es útil, por ejemplo, para saber el efecto de la interacción entre células a nivel de la expresión génica. Si el agente se administra terapéuticamente para tratar un tipo celular pero tiene un efecto no deseable en otro tipo celular, la invención proporciona un ensayo para determinar la base molecular del efecto no deseado y por lo tanto proporciona la oportunidad de co-administrar un agente que contrarresta o trata de otra manera el efecto no deseado. De manera similar, incluso dentro de un tipo celular, pueden determinarse los efectos biológicos

indeseables a nivel molecular. Por lo tanto, los efectos de un agente en la expresión de un gen distinto al gen objetivo pueden determinarse y contrarrestarse.

En otra realización, la matriz puede usarse para monitorear el transcurso de tiempo de la expresión de uno o más genes en la matriz. Esto puede ocurrir en varios contextos biológicos, según se describe en la presente, por ejemplo, en el desarrollo de toxicidad inducida por fármacos, la evolución de toxicidad inducida por fármacos, y procesos, tales como la transformación asociada con la toxicidad inducida por fármacos.

La matriz también es útil para determinar el efecto de la expresión de un gen en la expresión de otros genes en la misma célula o en diferentes células. Esto proporciona, por ejemplo, una selección de objetivos moleculares alternativos para la intervención terapéutica si el objetivo final o en dirección a 3 no puede ser regulado.

La matriz también es útil para determinar los patrones de expresión diferencial de uno o más genes en células normales y anormales. Esto proporciona una batería de genes que pueden servir como objetivo molecular para el diagnóstico o la intervención terapéutica.

VII. Métodos para obtener las muestras

Las muestras útiles en los métodos de la invención incluyen muestras de tejido, células, biopsia o fluidos corporales que expresan un marcador de la invención. En una realización, una muestra puede ser un tejido, una célula, sangre entera, suero, plasma, raspado bucal, saliva, fluido cefalorraquídeo, orina, heces o lavado bronquioalveolar. En realizaciones preferidas, la muestra de tejido es una muestra de una enfermedad o estado de toxicidad. En realizaciones más preferidas, la muestra de tejido es una muestra de un cáncer, una muestra de diabetes, una muestra de obesidad, una muestra cardiovascular o una muestra de toxicidad inducida por fármacos.

Las muestras corporales pueden obtenerse de un sujeto mediante una variedad de técnicas conocidas en la técnica, incluyendo, por ejemplo, mediante el uso de biopsia o mediante el raspado o frotado de un área, o usando una aguja para aspirar fluidos corporales. Los métodos para recoger varias muestras corporales son conocidos en la técnica.

Las muestras de tejido adecuadas para detectar y cuantificar un marcador de la invención puede ser fresca, congelada o fijada, de acuerdo con los métodos conocidos por los expertos en la técnica. Las muestras de tejido adecuadas se

seleccionan y se colocan preferentemente en un portaobjetos de un microscopio para su análisis adicional. De manera alternativa, las muestras sólidas, *es decir,* las muestras de tejido, pueden ser solubilizadas y/u homogeneizadas, y subsiguientemente analizadas como extractos solubles.

En una realización, una muestra de biopsia recién obtenida se congela usando, por ejemplo, nitrógeno líquido o difluorodiclorometano. La muestra congelada se monta para ser seccionada usando, por ejemplo, OCT, y se secciona en serie en un criostato. Las secciones en serie se recogen en un portaobjetos de vidrio de un microscopio. Para la tinción inmunohistoquímica los portaobjetos pueden recubrirse con, por ejemplo, alumbre de cromo, gelatina o poli-L-lisina para asegurarse de que las secciones se adhieran a los portaobjetos. En otra realización, las muestras se fijan y se incorporan antes de ser seccionadas. Por ejemplo, una muestra de tejido puede fijarse en, por ejemplo, formalina, deshidratarse en serie e incorporarse en, por ejemplo, parafina.

Una vez que se obtiene una muestra, puede usarse cualquier método conocido en la técnica adecuado para detectar y cuantificar un marcador de la invención (ya sea a nivel del ácido nucleico o de la proteína). Dichos métodos son conocidos en la técnica e incluyen, de modo no taxativo, transferencia western,

transferencia northern, transferencia southern, inmunohisstoquímica, ELISA, *por ejemplo,* ELISA amplificado, inmunoprecipitación, inmunofluorescencia, citometría de flujo, inmunocitoquímica, análisis espectrométricos de masas, *por ejemplo,* MALDI-TOF y SELDI-TOF, técnicas de hibridación de ácido nucleico, métodos de transcripción inversa de ácido nucleico y métodos de amplificación de ácido nucleico. En realizaciones particulares, la expresión de un marcador de la invención es detectada a nivel de proteínas usando, por ejemplo, anticuerpos que se unen específicamente a estas proteínas.

Las muestras pueden necesitar ser modificadas para hacer que el marcador de la invención sea accesible para la unión del anticuerpo. En un aspecto particular de los métodos de inmunocitoquímica o inmunohistoquímica, los portaobjetos pueden transferirse a un amortiguador de pretratamiento y calentarse opcionalmente para aumentar la accesibilidad del antígeno. El calentamiento de la muestra en el amortiguador de pretratamiento altera rápidamente la bi-capa lipídica de las células y hace que los antígenos sean más accesibles para la unión del anticuerpo (puede ser el caso de muestras frescas, pero no es lo que ocurre comúnmente en muestras fijadas). Los términos "amortiguador de pretratamiento" y "amortiguador de preparación" se usan de manera intercambiable en la presente para referirse a un amortiguador que se usa para preparar muestras de citología o histología para

inmunotinción, particularmente mediante el aumento de la accesibilidad de un marcador de la invención para la unión del anticuerpo. El amortiguador de pretratamiento puede comprender una solución salina específica del pH, un polímero, un detergente, o un tensioactivo no iónico o aniónico tal como, por ejemplo, un tensioactivo aniónico o no iónico etiloxilado, un alcanoato o un alcoxilato o incluso mezclas de estos tensioactivos o incluso el uso de una sal biliar. El amortiguador de pretratamiento puede, por ejemplo, ser una solución de 0,1% a 1% de ácido desoxicólico, sal sódica o una solución de lauret-13-carboxilato de sodio (por ejemplo, Sandopan LS) o un complejo aniónico etoxilado. En algunas realizaciones, el amortiguador de pretratamiento puede también usarse como un amortiguador de almacenamiento deslizante.

Puede usarse cualquier método para hacer que proteínas marcadoras de la invención sean más accesibles para la unión con el anticuerpo para llevar a la práctica la invención, incluyendo los métodos de recuperación antigénica conocidos en la técnica. Véase, por ejemplo, Bibbo, *et ál.* (2002) *Acta. Cytol.* 46:25-29; Saqi, *et ál.* (2003) *Diagn. Cytopathol.* 27:365-370; Bibbo, *et ál.* (2003) *Anal. Quant. Cytol. Histol.* 25:8-11, cuyos contenidos se incorporan a la presente mediante esta referencia.

Después del pretratamiento para aumentar la accesibilidad de la proteína marcadora, las muestras pueden bloquearse usando un agente de bloqueo adecuado, *por ejemplo, un reactivo de bloqueo de* peroxidasa, tal como peróxido de hidrógeno. En algunas realizaciones, las muestras pueden bloquearse usando un reactivo de bloqueo de proteínas para evitar la unión no específica del anticuerpo. El reactivo de bloqueo de proteínas puede comprender, por ejemplo, caseína purificada. Un anticuerpo, particularmente un anticuerpo monoclonal o policlonal que se une específicamente al marcador de la invención es incubado entonces con la muestra. Un experto en la técnica apreciará que se puede obtener un pronóstico o un diagnóstico más acertados en algunos casos mediante la detección de múltiples epítopos en una proteína marcadora de la invención en una muestra de un paciente. Por lo tanto, en realizaciones particulares, se usan al menos dos anticuerpos dirigidos a diferentes epítopos de un marcador de la invención. Cuando se usa más de un anticuerpo, estos anticuerpos pueden agregarse a una única muestra de manera secuencial como reactivos de anticuerpos individuales o simultáneamente como un cóctel de anticuerpos. De manera alternativa, cada anticuerpo individual puede agregarse a una muestra separada del mismo paciente, y agrupar los datos resultantes.

Las técnicas para detectar la unión de anticuerpos se conocen en la técnica. La unión de los anticuerpos a un marcador

de la invención puede detectarse a través del uso de reactivos químicos que generan una señal detectable que corresponde al nivel de unión de anticuerpo y, por consiguiente, al nivel de expresión de la proteína marcadora. En uno de los métodos de inmunohistoquímica o inmunocitoquímica de la invención, la unión del anticuerpo se detecta a través del uso de un anticuerpo secundario que está conjugado con un polímero marcado. Los ejemplos de polímeros marcados incluyen, de modo no taxativo, conjugados de polímero-enzima. Las enzimas en estos complejos se usan típicamente para catalizar la deposición de un cromógeno en el sitio de unión del antígeno-anticuerpo, dando como resultado de esa manera la tinción de la célula que corresponde al nivel de expresión del biomarcador de interés. La enzimas de particular interés incluyen, de modo no taxativo, peroxidasa de rábano picante (HRP) y fosfatasa alcalina (AP).

En uno método particular de inmunohistoquímica o inmunocitoquímica de la invención, la unión del anticuerpo a un marcador de la invención se detecta a través del uso de un polímero marcado con HRP que está conjugado con un anticuerpo secundario. La unión del anticuerpo también se pude detectar a través del uso de un reactivo de sonda específico de una especie, que se une a anticuerpos monoclonales o policlonales, y un polímero conjugado con HRP, que se una a la especie específica del reactivo de sonda. Los portaobjetos se tiñen para detectar la

unión del anticuerpo usando cualquier chromagen, *por ejemplo,* el chromagen 3,3-diaminobenzidina (DAB), y luego se contra tiñen con hematoxilina y, opcionalmente, un colorante azul, tal como hidróxido de amonio o TBS/Tween-20. Otros chromagen adecuados incluyen, por ejemplo, 3-amino-9-etilcarbazol (AEC). En algunos aspectos de la invención, los portaobjetos son revisados microscópicamente por un cito tecnólogo y/o un patólogo para evaluar la tinción de las células, *por ejemplo,* tinción fluorescente *(es decir,* la expresión del marcador). De manera alternativa, las muestras pueden ser revisadas por medio *de* microscopía automatizada o por personal con la asistentica de software informático que facilita la identificación de células teñidas positivas.

La detección de la unión del anticuerpo puede facilitarse mediante el acoplamiento de anticuerpos anti-marcador a una sustancia detectable. Los ejemplos de sustancias detectables incluyen varias enzimas, grupos prostéticos, materiales fluorescentes, materiales luminiscentes, materiales bioluminiscentes y materiales radiactivos. Los ejemplos de enzimas adecuadas incluyen la peroxidasa de rábano picante, fosfatasa alcalina, □-galactosidasa o acetilcolinesterasa; los ejemplos de complejos de grupos prostéticos adecuados incluyen estreptavidina/biotina y avidina/biotina; los ejemplos de materiales fluorescentes adecuados incluyen umbeliferona,

fluoresceina, isotiocianato de fluoresceina, rodamina, diclorotriazinilamina fluoresceina, cloruro de dansilo o ficoeritrina; un ejemplo de un material luminiscente adecuado incluye luminol; los ejemplos de materiales bioluminiscentes adecuados incluyen luciferasa, luciferina y aequorina, y los ejemplos de materiales radiactivos adecuados incluyen ^{125}I, ^{131}I, ^{35}S, ^{14}C o ^{3}H.

En una realización de la invención las muestras congeladas se preparan como se describió anteriormente y se tiñen subsiguientemente con anticuerpos contra un marcador de la invención diluidos a una concentración adecuada, usando, por ejemplo, solución salina amortiguada con Tris (TBS). Los anticuerpos primarios pueden detectarse mediante la incubación de los portaobjetos en anti-inmunoglobulina biotinilada. Esta señal puede amplificarse y visualizarse opcionalmente usando precipitación dediaminobencidina del antígeno. Además, los portaobjetos pueden contrateñirse opcionalmente con, por ejemplo, hematoxilina, para visualizar las células.

En otra realización, las muestras fijadas e incorporadas se tiñen con anticuerpos contra un marcador de la invención y se contratiñen como se describe para las secciones congeladas. Además, las muestras pueden tratarse opcionalmente con agentes para amplificar la señal, para poder visualizar la tinción del

anticuerpo. Por ejemplo, puede utilizarse una deposición catalizada por peroxidasa de biotinil-tiramida, que a su vez se hace reaccionar con estreptavidina conjugada con peroxidasa (Sistema de Amplificación de Señal Catalizada (CSA), DAKO, Carpintería, CA).

Los ensayos basados en tejido *(es decir,* inmunohistoquímica) son los métodos preferidos para detectar y cuantificar un marcador de la invención. En una realización, la presencia o ausencia de un marcador de la invención puede determinarse mediante inmunohistoquímica. En una realización, el análisis inmunohistoquímico usa concentraciones bajas de un anticuerpo anti-marcador de manera que las células que carecen del marcador no se tiñen. En otra realización, la presencia o ausencia de un marcador de la invención se determina usando un método inmunohistoquímico que usa concentraciones altas de anticuerpo anti-marcador, de manera que las células que carecen de la proteína marcadora se tiñen de manera potente. Las células que no se tiñen contienen o bien el marcador mutado y no producen la proteína marcadora antigénicamente reconocible, o son células en las cuales las vía que regulan los niveles de marcador están desreguladas, lo cual da como resultado un estado constante de expresión de proteína marcadora insignificante.

Un experto en la técnica reconocerá que la concentración de un anticuerpo particular usado para llevar a la práctica los métodos de la invención variará dependiendo de factores tales como el tiempo de unión, el nivel de especificidad del anticuerpo por el marcador de la invención y el método de preparación de la muestra. Además, cuando se usan múltiples anticuerpos, la concentración necesaria puede verse afectada por el orden en el cual se aplican los anticuerpos a la muestra, *por ejemplo,* de manera simultánea como un cóctel o secuencial como reactivos de anticuerpos individuales. Además, la química de detección usada para visualizar la unión del anticuerpo al marcador de la invención también debe optimizarse para producir la relación deseada entre señal y ruido.

En una realización de la invención, se usan métodos proteómicos, por ejemplo, espectrometría de masas, para detectar y cuantificar las proteínas marcadoras de la invención. Por ejemplo, espectrometría de masa de tiempo de vuelo de desorción/ionización láser asistida por matriz (MALDI-TOF MS) o espectrometría de masa de tiempo de vuelo de desorción/ionización láser en superficie (SELDI-TOF MS) que involucra la aplicación a una muestra biológica, tal como suero, a un chip de unión a proteínas (Wright, G.L., Jr., *et ál.* (2002) *Expert Rev Mol Diagn* 2:549; Li, J., *et ál.* (2002) *Clin Chem* 48:1296; Laronga, C, *et ál.* (2003) *DisMarkers* 19:229; Petricoin, E.F., *et ál.* (2002)

359:572; Adam, B.L., *et ál.* (2002) *Cancer Res* 62:3609; Toison, J., *et ál.* (2004) *Lab Invest* 84:845; Xiao, Z., *et ál.* (2001) *Cancer Res* 61:6029) pueden usarse para detectar y cuantificar las proteínas PY-Shc y/o p66-Shc. Se describen métodos espectrométricos de masas en, por ejemplo, las patentes estadounidenses N.° 5.622.824, 5.605.798 y 5.547.835, cuyos contenidos completos de incorporan a la presente mediante esta referencia.

En otras realizaciones, la expresión de un marcador de la invención se detecta a nivel del ácido nucleico. Las técnicas basadas en ácidos nucleicos para evaluar la expresión son conocidas en la técnica e incluyen, por ejemplo, determinación del nivel de ARNm marcador en una muestra de un sujeto. Muchos métodos de detección de la expresión usan ARN aislado. Puede usarse cualquier técnica de aislamiento que no evite el aislamiento del ARNm para la purificación del ARN de las células que expresan un marcador de la invención (véase, por ejemplo, Ausubel et ál., ed., (1987-1999) Current Protocols in Molecular Biology (John Wiley & Sons, Nueva York). Adicionalmente, grandes cantidades de muestras de tejido pueden procesarse fácilmente usando técnicas conocidas por los expertos en la técnica, tal como, por ejemplo, el proceso de aislamiento de ARN de un solo paso de Chomczynski (1989, patente estadounidense N.° 4.843.155).

El término "sonda" se refiere a cualquier molécula que es capaz de unirse selectivamente a un marcador de la invención, por ejemplo, un transcripto de nucleótido y/o una proteína. Las sondas pueden ser sintetizadas por un experto en la técnica, o derivadas de preparaciones biológicas adecuadas. Las sondas pueden diseñarse específicamente para ser marcadas. Los ejemplos de moléculas que pueden utilizarse como sondas incluyen, de modo no taxativo, ARN, ADN, proteínas, anticuerpos y moléculas orgánicas.

El ARNm aislado puede usarse en los ensayos de hibridación o amplificación que incluyen, de modo no taxativo, análisis Southern o Northern, análisis de reacción en cadena de la polimerasa y matrices de sondas. Un método para la detección de los niveles de ARNm involucra poner en contacto el ARNm aislado con una molécula de ácido nucleico (sonda) que pueda hibridarse con el ARNm marcador. La sonda de ácido nucleico puede ser, por ejemplo, un ADNc de longitud completa o una parte de este, tal como un oligonucleótido de al menos 7, 15, 30, 50, 100, 250 o 500 nucleótidos de longitud y suficiente para hibridarse de manera específica bajo condiciones estrictas a un ADN genómico marcador.

En una realización, el ARNm se inmoviliza en una superficie sólida y se pone en contacto con una sonda, por ejemplo, mediante el pase del ARNm aislado en un gel de agarosa y

transfiriendo el ARNm desde el gel a una membrana, tal como nitrocelulosa. En una realización alternativa, la o las sondas se inmovilizan en una superficie sólida y el ARNm se pone en contacto con la o las sondas, por ejemplo, en un chip de ADN Affymetrix. Un experto en la técnica puede adaptar fácilmente métodos de detección de ARNm conocidos para su uso en la detección del nivel de ARNm marcador.

Un método alternativo para determinar el nivel de ARNm marcador en una muestra involucra el proceso de amplificación de ácido nucleico, *por ejemplo,* mediante RT-PCR (la realización experimental establecida en Mullis, 1987, patente estadounidense N.° 4,683,202), reacción en cadena de la ligasa (Barany (1991) *Proc. Nati. Acad. Sel USA* 88:189-193), replicación de secuencia autosostenida (Guatelli et ál. (1990) *Proc. Nati Acad. Sel USA* 87:1874-1878), sistema de amplificación transcriptional (Kwoh *et al.* (1989) *Proc. Nati Acad. Sel USA* 86:1173-1177), Q-Beta Replicasa (Lizardi *et ál.* (1988) *Bio/Technology* 6:1197), replicación de círculo rodante (Lizardi et ál., patente estadounidense N.° 5.854.033) o cualquier otro método de amplificación del ácido nucleico, seguido de la detección de moléculas amplificadas usando técnicas conocidas para los expertos en la técnica. Estos esquemas de detección son especialmente útiles para la detección de moléculas de ácido nucleico si dichas moléculas están presentes en cantidades muy

bajas. En aspectos particulares de la invención, la expresión del marcador se determina mediante RT-PCR fluorogénica cuantitativa *(es decir,* el sistema TaqMan™). Dichos métodos comúnmente utilizan pares de cebadores de oligonucleótidos que son específicos para un marcador de la invención. Los métodos para diseñar los cebadores de oligonucleótidos específicos para una secuencia conocida son conocidos en la técnica.

Los niveles de expresión de un marcador de la invención pueden monitorearse usando un ensayo de transferencia de membrana (tal como el usado en un análisis de hibridación, tal como Northern, Southern, dot, y similares), o micropocillos, tubos de ensayo, geles, perlas o fibras (o cualquier soporte sólido que comprende el ácido nucleico unido). Véanse las patentes estadounidenses N.° 5.770.722, 5.874.219, 5.744.305, 5.677.195 y 5.445.934, las cuales se incorporan a la presente mediante esta referencia. La detección de la expresión del marcador puede también comprender el uso de sondas de ácido nucleico en solución.

En una realización de la invención, se usan micromatrices para detectar la expresión de un marcador de la invención. Las micromatrices son especialmente útiles para este fin debido a la reproducibilidad entre diferentes experimentos. Las micromatrices de ADN proporcionan un método para la medición simultánea de los

niveles de expresión de una gran cantidad de genes. Cada matriz consiste en un patrón reproducible de sondas de captura acopladas a un soporte sólido. El ARN o ADN marcado se hibrida a sondas complementarias en la matriz y entonces se detectan mediante escaneo con láser. Las intensidades de hibridación para cada sonda en la matriz se determinan y se convierten a un valor cuantitativo representando los niveles de expresión génica relativos. Véanse las patentes estadounidenses N.° 6.040.138, 5.800.992, 6.020.135, 6.033.860 y 6.344.316, las cuales se incorporan a la presente mediante esta referencia. Las matrices de oligonucleótidos de alta densidad son particularmente útiles para determinar el perfil de expresión génica para una gran cantidad de ARN en la muestra.

Las cantidades de un marcador y/o una relación matemática de las cantidades de un marcador de la invención pueden usarse para calcular el riesgo de recurrencia de una enfermedad, por ejemplo, cáncer, diabetes, obesidad, enfermedad cardiovascular, o un estado de toxicidad, por ejemplo, una toxicidad o cardiotoxicidad inducidas por fármacos, en un sujeto que se está tratando contra la enfermedad o estado de toxicidad, la supervivencia de un sujeto que se está tratando contra una enfermedad o estado de toxicidad, si una enfermedad o estado de toxicidad es agresivo, la eficacia de un régimen de tratamiento para tratar una enfermedad o estado de toxicidad, y similares,

usando los métodos de la invención, que pueden incluir métodos de análisis de regresión conocidos por los expertos en la técnica. Por ejemplo, los modelos de regresión adecuados incluyen, de modo no taxativo, CART (por ejemplo, Hill, T, y Lewicki, P. (2006) "STATISTICS Methods and Applications" StatSoft, Tulsa, OK), Cox (por ejemplo,www.evidence-based-medicine.co.uk), exponencial, normal y normal logarítmico (por ejemplo, www.obgyn.cam.ac.uk/mrg/statsbook/stsurvan.html), logístico (por ejemplo,www.en.wikipedia.org/wiki/Logistic_regression), paramétrico, no paramétrico, semiparamétrico (por ejemplo,www.socserv.mcmaster.ca/jfox/Books/Companion), lineal (por ejemplo,www.en.wikipedia.org/wiki/Linear_regression), o aditivo (por ejemplo,www.en.wikipedia.org/wiki/Generalized_additive_model).

En una realización, un análisis de regresión incluye las cantidades de marcador. En otra realización, un análisis de regresión incluye una relación matemática del marcador. En todavía otra realización, un análisis de regresión de las cantidades de marcador, y/o una relación matemática del marcador puede incluir co-variantes adicionales clínicas y/o moleculares. Dichas co-variantes clínicas incluyen, de modo no taxativo, estado nodal, estadificación del tumor, grado del tumor, tamaño del tumor, régimen de tratamiento, *por ejemplo,* quimioterapia y/o radioterapia, resultado clínico *(por ejemplo,* recaída,

supervivencia específica de la enfermedad, falla de la terapia), y/o resultados clínicos en función del tiempo después del diagnóstico, el tiempo después del inicio de la terapia y/o el tiempo después de completar el tratamiento.

VIII. Kits

La invención también proporciona composiciones y kits para realizar el pronóstico de una enfermedad, por ejemplo, cáncer, diabetes, obesidad, enfermedad cardiovascular, o un estado de toxicidad, por ejemplo, toxicidad inducida por fármacos o cardiotoxicidad, recurrencia de una enfermedad o de un estado de toxicidad, o la supervivencia de un sujeto que se está tratando contra una enfermedad o estado de toxicidad. Estos kits incluyen uno o más de los siguientes: un anticuerpo detectable que se une específicamente a un marcador de la invención, un anticuerpo detectable que se une específicamente a un marcador de la invención, reactivos para obtener y/o preparar muestras de tejido del sujeto para su tinción, e instrucciones de uso.

Los kits de la invención pueden comprender opcionalmente componentes adicionales útiles para llevar a cabo los métodos de la invención. A modo de ejemplo, los kits de la invención pueden comprender fluidos *(por ejemplo,* amortiguador SSC) adecuados para aparear ácidos nucleicos complementarios o para unir un

anticuerpo con una proteína con la cual se une específicamente, uno o más compartimientos de muestra, un material instructivo que describe la puesta en práctica de un método de la invención y controles/estándares específicos de tejido.

IX. Ensayos de cribado

Los objetivos de la invención incluyen, de modo no taxativo, genes y/o proteína listados en la presente. Basándose en los resultados de los experimentos descritos en la presente por los solicitantes, las proteínas clave moduladas en una enfermedad o estado de toxicidad están asociadas con o pueden ser clasificadas en diferentes vías o grupos de moléculas, incluyendo componentes citoesqueléticos, factores de transcripción respuesta apoptótica, vía fosfato pentosa, vía biosintética, tensión oxidativa (pro-oxidante), alteraciones de la membrana y metabolismo de fosforilación oxidativa. Por consiguiente, en una realización de la invención, un marcador puede incluir uno o más genes (o proteínas) seleccionadas del grupo que consiste en HSPA8, FLNB, PARK7, HSPA1A/HSPA1B, ST13, TUBB3, MIF, KARS, NARS, LGALS1, DDX17, EIF5A, HSPA5, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1, CANX, GRP78, GRP75, TIMP1, PTX3, HSP76, PDIA4, PDIA1, CA2D1, GPAT1 y TAZ. En una realización, un marcador puede incluir uno o más genes (o proteínas) seleccionadas del grupo que consiste en GRP78, GRP75, TIMP1, PTX3, HSP76, PDIA4,

PDIA1, CA2D1, GPAT1 y TAZ. En algunas realizaciones, los marcadores son una combinación de al menos dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez, once, doce, trece, catorce, quince, dieciséis, diecisiete, dieciocho, diecinueve, veinte, veinticinco, treinta y más de los antemencionados gens (o proteínas).

Los ensayos de cribado útiles para identificar los moduladores de marcadores identificados se describen más adelante.

La invención también proporciona métodos (también mencionados en la presente como "ensayos de cribado") para identificar moduladores, *es decir,* compuesto o agentes candidatos o de prueba o agentes *(por ejemplo,* proteínas, péptidos, peptidomiméticos, peptoides, moléculas pequeñas u otros fármacos), que son útiles para tratar o prevenir una enfermedad o un estado de toxicidad mediante la modulación de la expresión y/o actividad de un marcador de la invención. Dichos ensayos comúnmente comprenden una reacción entre un marcador de la invención y uno o más componentes del ensayo. Los otros componentes del ensayo pueden ser el compuesto de prueba en sí, o una combinación de compuestos de prueba y un compañero de unión natural de un marcador de la invención. Los compuesto identificados por medio de los ensayos descritos en la presente

pueden ser útiles, por ejemplo, para modular, *por ejemplo,* inhibir, mejorar, tratar o prevenir la agresividad de una enfermedad o estado de toxicidad.

Los compuestos de prueba usados en los ensayos de cribado de la presente invención pueden obtenerse de varias fuentes disponibles, incluyendo bibliotecas sistémicas de compuestos naturales y/o sintéticos. Los compuestos de prueba pueden obtenerse mediante cualquiera de los varios abordajes en métodos de biblioteca combinatoria conocidos en la técnica, incluyendo: bibliotecas biológicas; bibliotecas de peptoides (bibliotecas de moléculas que tienen las funcionalidades de los péptidos, pero con una estructura no peptídica novedosa que es resistente a la degradación enzimática, pero que permanece activa de todos modos; véase, *por ejemplo,* Zuckermann *et él.,* 1994, J. *Med. Chem.*37:2678-85); bibliotecas de fase de solución o fase sólida paralela dirigible espacialmente; métodos de bibliotecas sintéticas que requieren desconvulción; el método de bibliotecas de "una perla, un compuesto"; y métodos de bibliotecas sintéticas usando selección de cromatografía de afinidad. Los abordajes de biblioteca biológica y biblioteca de peptoides se limitan a bibliotecas de péptidos, mientras que los otros cuatro abordajes son aplicables a bibliotecas de compuestos de péptidos, oligómeros no peptídicos o moléculas pequeñas (Lam, 1997, *Anticancer Drug Des.* 12:145).

Los ejemplos de métodos para la síntesis de bibliotecas moleculares pueden encontrarse en la técnica, por ejemplo en: DeWitt *et ál.* (1993) *Proc. Nati. Acad. Sci. U.S.A.* 90:6909; Erb *et ál.* (1994) *Proc. Nati. Acad. Sci. USA* 91:11422; Zuckermann *et ál.* (1994). J. *Med. Chem.* 37:2678; Cho *et ál.* (1993) *Science* 261:1303; Carrell *et ál.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et ál. (1994) Angew. Chem. Int. Ed. Engl.* 33:2061; y en Gallop *et ál.* (1994) J. *Med. Chem.* 37:1233.

Las bibliotecas de compuestos pueden presentarse en solución (por ejemplo, Houghten, 1992, *Biotechniques* 13:412-421), o en perlas (Lam, 1991, *Nature* 354:82-84), chips (Fodor, 1993, *Nature* 364:555-556), bacterias y/o esporas, (Ladner, USP 5,223,409), plásmidos (Culi *et ál.*, 1992, *Proc Nati Acad Sci USA* 89:1865-1869) o en fagos (Scott y Smith, 1990, *Science* 249:386-390; Devlin, 1990, *Science* 249:404-406; Cwirla *et ál.*, 1990, *Proc. Nati. Acad. Sci.* 87:6378-6382; Felici, 1991, J. *Mol. Biol.* 222:301-310; Ladner, *supra.)*.

Los métodos de cribado de la invención comprenden poner en contacto una célula enferma o una célula en estado de toxicidad con un compuesto de prueba y determinar la capacidad del compuesto de prueba de modular la expresión y/o actividad de un marcador de la invención en la célula. La expresión y/o actividad

de un marcador de la invención puede determinarse como se describe en la presente.

En otra realización, la invención proporciona ensayos para analizar compuestos candidatos o de prueba que son sustratos de un marcador de la invención o porciones biológicamente activas de estos. En otra realización, la invención proporciona ensayos para analizar compuestos candidatos o de prueba que se unen a un marcador de la invención o porciones biológicamente activas de estos. La determinación de la capacidad del compuesto de prueba para unirse directamente a un marcador puede llevarse a cabo, por ejemplo, mediante el acoplamiento del compuesto con un radioisótopo o etiqueta enzimática, de manera que la unión del compuesto al marcador se pueda determinar mediante la detección del compuesto marcador etiquetado en un complejo. Por ejemplo, los compuestos (por ejemplo, sustratos marcadores) pueden etiquetarse con ^{131}I, ^{125}I, ^{35}S, ^{14}C o ^{3}H, ya sea directa o indirectamente, y el radioisótopo se detecta mediante conteo directo de la radioemisión o por conteo de centelleo. De manera alternativa, los componentes de ensayo pueden etiquetarse enzimáticamente con, por ejemplo, peroxidasa de rábano picante, fosfatasa alcalina o luciferasa, y la etiqueta enzimática puede detectarse mediante la determinación del a conversión de un sustrato apropiado en un producto.

Esta invención se refiere adicionalmente a agentes novedosos identificados en los ensayos de cribado antemencionados. Por consiguiente, está dentro del alcance de la invención el uso adicional de un agente identificado como se describe en la presente en un modelo animal adecuado. Por ejemplo, un agente capaz de modular la expresión y/o actividad de un marcador de la invención identificado como se describe en la presente puede usarse en un modelo animal para determinar la eficacia, toxicidad o efectos secundarios del tratamiento con dicho agente. De manera alternativa, un agente identificado como se describe en la presente puede ser usado en un modelo animal para determinar el mecanismo de acción de dicho agente. Adicionalmente, la presente invención se refiera a los usos de agentes novedosos identificados mediante los ensayos de cribado antemencionados para el tratamiento que se describió anteriormente.

Ejemplificación de la invención

EJEMPLO 1: Empleo de Tecnología de Plataforma para construir un consenso de cáncer y redes de simulación

En este ejemplo, la tecnología de plataforma descrito en detalle anteriormente se empleó para integrar los datos obtenidos

de un modelo de cáncer construido a medida *in vitro*, e identificar de esta manera las proteínas/vías novedosas que dirigen la patogénesis del cáncer. Los mapas relacionales que resultan de estos análisis proporcionaron objetivos de tratamiento del cáncer, así como marcadores de diagnóstico/pronóstico asociados con el cáncer.

El diseño del estudio se muestra en la Figura 18. Brevemente, se sometieron dos líneas celulares de cáncer (PaCa2, HepG2) y una línea celular normal (THLE2) a una de siete condiciones simulando un ambiente que experimentan las células cancerosas *in vivo*. Específicamente, las células se expusieron a una condición hiperglucémica, condición de hipoxia, condición de ácido láctico, combinación de condición hiperglucémica + hipoxia, combinación de condición hiperglucémica + ácido láctico, combinación de condición de hipoxia + ácido láctico, o combinación de condición hiperglucémica + hipoxia + ácido láctico. Las diferentes condiciones se crearon como sigue:

-La condición hiperglucémica se creó mediante el cultivo de las células en medio que contenía glucosa 22 mM.

-La condición de hipoxia se indujo mediante la colocación de las células en una cámara incubadora modular (MIC-101, Billups-Rothenberg Inc. Del Mar, CA), que se llenó con una mezcla de gas industrial que contenía 5% de CO_2, 2% de O_2 y 93% de nitrógeno.

-La condición ácido láctico se creó mediante el cultivo de las células en medio que contenía ácido láctico 12,5 mM.

—La combinación de la condición hiperglucémica + hipoxia se creó mediante el cultivo de las células en medio que contenía glucosa 22 mM y las células se colocaron en una cámara incubadora modular que se llenó con una mezcla de gas industrial que contenía 5% de CO_2, 2% de O_2y 93% de nitrógeno.

-La combinación de condición hiperglucémica + ácido láctico se creó mediante el cultivo de las células en medio que contenía glucosa 22 nM y ácido láctico 12,5 mM.

—La combinación de la condición hipoxia + ácido láctico se creó mediante el cultivo de las células en medio que contenía ácido láctico 12,5 mM y las células se colocaron en una cámara incubadora modular que se llenó con una mezcla de gas industrial que contenía 5% de CO_2, 2% de O_2y 93% de nitrógeno.

—La combinación de la condición hiperglucémica + hipoxia + ácido láctico se creó mediante el cultivo de las células en medio que contenía glucosa 22 mM, ácido láctico 12,5 mM y las células se colocaron en una cámara incubadora modular que se llenó con una mezcla de gas industrial que contenía 5% de CO_2, 2% de O_2y 93% de nitrógeno.

El modelo celular que comprende las células antemencionadas, donde las células se expusieron a cada condición descrita anteriormente, se interrogaron adicionalmente mediante

la exposición de las células a una perturbación ambiental mediante el tratamiento con la Coenzima Q10. Específicamente, las células se trataron con la Coenzima Q10 a 0, 50 µM o 100 µM.

Las muestras celulares así como las muestras de medio para cada línea celular con cada condición y cada tratamiento con Coenzima Q10 se recogieron en varios momentos después del tratamiento, incluyendo después de 24 horas y 48 horas de tratamiento.

Además, se llevaron a cabo experimentos de interacción entre dos células cancerosas diferentes, células PaCa2 y HepG2, en los cuales las células de PaCa2 y HepG2 se cocultivaron. Este abordaje de cocultivo se menciona como experimento de secretoma extracelular (ECS). El primer sistema celular (PaCa2) se cultivó primero en los insertos de los pocillos de una cámara de cultivo del tipo transwell. Se usaron placas de seis pocillos para permitir un mejor análisis estadístico. En el momento de sembrar el primer sistema celular en los insertos, se colocaron los insertos en una placa de 6 pocillos separada. El segundo sistema celular (HepG2) se sembró en la bandeja primaria. La bandeja de inserto que contenía el primer sistema celular y la bandeja primaria que contenía el segundo sistema celular se incubaron a 37 °C durante la noche. Cada uno de los sistemas celulares se cultivaron en el medio específico de células específicas (donde,

de manera alternativa, cada uno de los sistemas celulares puede cultivarse en un medio adaptado para sustentar el crecimiento de ambos tipos celulares). En el segundo día, se dio el tratamiento predeterminado mediante intercambio de medios. Específicamente, los insertos que contenían el primer sistema celular se colocaron en la bandeja primaria que contenía el segundo sistema celular. La bandeja se incubó durante un período de tiempo predeterminado, por ejemplo,24 horas o 48 horas. Se establecieron pocillos duplicados con las mismas condiciones, y las células se agruparon para proporcionar suficiente material para el análisis 2D. El medio (alícuota de 1 mL), las células de los insertos y las células de los pocillos de la bandeja primaria se cultivaron como muestras separadas. Los experimentos se llevaron a cabo por triplicado para proporcionar un mejor poder de análisis estadístico.

Los experimentos de interacción también se llevaron a cabo mediante experimentos de "cambio de medio". Específicamente, un medio cultivado o "secretoma" del primer sistema celular (PaCa2) se recogió después de 24 horas o 48 horas luego de la perturbación o condicionamiento descrito anteriormente y luego se agregó al segundo sistema celular (HepG2) durante 24-48 horas. Entonces se recogió medio cultivado final o "secretoma" del segundo sistema celular. Todos los secretomas finales se sometieron a análisis proteómico.

El método de iProfiling de los cambios en la expresión de proteína celular total mediante proteómica cuantitativa se llevó a cabo para las muestras celulares y de medio recogidas de cada línea celular en cada condición y con cada "perturbación ambiental", es decir, tratamiento de Coenzima Q10, usando las técnicas descritas anteriormente en la descripción detallada. El método de iProfiling de cambios en la expresión de proteína celular total mediante proteómica cuantitativa se llevó a cabo de manera similar para las muestras celulares y de medio recogidas para cada línea celular cocultivada en cada condición con cada tratamiento.

Además, el perfilado bioenergético del cáncer, células normales y células en los experimentos de interacción expuestas a cada condición, y con o sin la perturbación por Coenzima Q10 se generó mediante el empleo del analizador Seahorse, esencialmente como lo recomienda el fabricante. Se registraron la OCR (tasa de consumo de oxígeno) y la ECAR (tasa de acidificación extracelular) mediante los electrodos en una cámara de 7 µl creada con el cartucho apoyado contra la placa de cultivo seahorse.

Los datos proteómicos recogidos para cada línea celular (incluyendo células en los experimentos de interacción) en cada

condición y con cada perturbación, y los datos del perfilado bioenergético recogidos para cada línea celular en cada condición y con cada perturbación, se ingresaron y procesaron mediante el sistema REFS™. Los datos en bruto de los experimentos de aca2, HepG2, THLE2 y interacción se combinaron usando una nomenclatura estandarizada. Los genes con más del 15% de ausencia de datos proteómicos se retiraron. Se desarrolló una estrategia de imputación de datos. Por ejemplo, se usó un modelo de error de repetición interna para imputar los datos de las condiciones experimentales con duplicados. Un algoritmo K-NN basado en 10 vecinos se usó para imputar los datos sin repetición. Se construyeron diferentes modelos de REFS™ para tres sistemas biológico juntos, solo para el sistema de Paca2, o solo para el sistema HepG2, vinculados a los datos fenotípicos.

El área bajo la curva y los cambios de múltiplo para cada borde que conecta un nodo padre a un nodo hijo en las redes de simulación se extrajeron mediante un programa hecho a medida usando el lenguaje de programación R, donde el lenguaje de programación R es un entorno de software de código abierto para cómputos y gráficos estadísticos.

Los datos de salida del programa R se ingresaron en Cytoscape, un programa de código abierto, para generar una representación visual de una red de consenso.

Entre todos los modelos construidos, un ejemplo de red de consenso REFS de interacción de proteínas a una frecuencia de 70% de fragmento se muestra en la figura 21.

Cada nodo en la red de consenso mostrada en la figura 21 se simuló mediante el aumento o la disminución de la expresión de LDHA de 4 veces, para generar una red de simulación usando REFS™, como se describe en más detalle anteriormente en la descripción detallada.

Se investigaron el efecto del cambio en la expresión LDHA simulado en PARK7 y las proteínas en notas asociadas con PARK7 a altos niveles en la red de consenso de ejemplo que se muestran en la figura 21. Las proteínas que responden a la simulación de LDHA en dos líneas celulares de cáncer, es decir, Paca2 y HepG2, se identificaron usando REFS™ (véase la figura 22). Las cantidades representan el cambio de múltiplo en el nivel de expresión de proteína particular.

Para validar las conexiones de proteína identificadas usando los métodos antemencionados, los marcadores que se identificaron en la proximidad inmediata de LDHA en la red de simulación se ingresaron en IPA, un programa informático que utiliza redes neurales para determinar el enlace molecular entre

datos de salida experimentales y redes basado en bibliografía previamente publicada. Los datos de salida del programa IPA se muestran en la figura 23, donde los marcadores en formas grises se identificaron como que estaban en la proximidad inmediata de LDHA en la red de simulación generada por la plataforma, y los marcadores en formas sin rellenar son conexiones identificadas por IPA basándose en conocimiento de bibliografía publicada anteriormente.

Se observó que los marcadores identificados en los datos de salida de la tecnología de plataforma de Biología Interrogativa (mostrados en la Figura 21), es decir, DHX9, HNRNPC, CKAP4, HSPA9, PARP1, HADHA, PHB2, ATP5A1 y CANX, estaban conectados con los marcadores de cáncer conocidos, tal como TP53 y PARK7 dentro de la red general por IPA (mostrada en la Figura 23). El hecho de que los factores identificados mediante el uso de la plataforma de Biología Interrogativa compartan la conectividad con otros factores conocidos publicados en la bibliografía científica, valida la exactitud de la red creada mediante el uso de la Plataforma de Biología Interrogativa. Además, la asociación de red dentro de la sub-red de LDHA creada mediante el uso de datos de salida de la plataforma de Biología Interrogativa demostró la presencia de la influencia direccional de cada factor, al contrario de la red IPA donde el vínculo entre las entidades moleculares no proporciona direccionalidad

funcional entre los nodos que interactúan. Por lo tanto, el empleo de un abordaje imparcial para la generación de datos, integración e ingeniería inversa para crear un modelo computacional seguido de simulación y análisis de red diferencial, el descubrimiento de la plataforma de Biología Interrogativa permite el entendimiento de los mecanismos desconocidos hasta ahora en la patofisiología del cáncer que son congruentes con conocimientos científicos establecidos de la patofisiología de enfermedades.

La Figura 19 muestra el efecto del tratamiento con CoQ10 en los nodos descendentes (los números de acceso de proteína pubmed se listan en la Figura 19) basado en los datos de expresión de proteína del método de iProfiling. El número de acceso de proteína P00338 es LDHA. La validación de laboratorio húmedo de los datos proteómicos se llevó a cabo para la expresión de LDHA en células HepG2 (véase la Figura 20). Como se muestra en la Figura 20, los niveles de expresión de LDHA disminuyeron cuando se trataron las HepG2 con CoQ10 50 uM o CoQ10 100 uM durante 24 o 48 horas.

Para los marcadores de cáncer conocidos TP53, Bcl-2, Bax y Caspasa3, se llevó a cabo la validación de laboratorio húmedo de los efectos del tratamiento con CoQ10 de los niveles de expresión

de estos marcadores en células SKMEL 28 (véase la Figura 24 y la Figura 25).

EJEMPLO 2: Empleo de Tecnología de Plataforma para construir una red delta-delta de cáncer

En este ejemplo, la tecnología de plataforma descrito en detalle anteriormente se empleó para integrar los datos obtenidos de un modelo de cáncer construido a medida *in vitro*, e identificar de esta manera las proteínas/vías novedosas que dirigen la patogénesis del cáncer Los mapas relacionales que resultan de estos análisis proporcionaron objetivos de tratamiento del cáncer, así como marcadores de diagnóstico/pronóstico asociados con el cáncer.

Brevemente, se sometieron cuatro líneas celulares de cáncer (PaCa2, HepG2, PC3 y MCF7) y dos líneas celulares normales (THLE2 y HDFa) a varias condiciones simulando un ambiente que experimentan las células cancerosas *in vivo*. Específicamente, las células se expusieron por separado a cada una de condiciones hiperglucémicas, condiciones de hipoxia y tratamiento con ácido láctico. Por ejemplo, una condición hiperglucémica se creó mediante el cultivo de las células en medio que contenía glucosa 22 mM. -Una condición de hipoxia se indujo mediante la colocación de las células en una cámara incubadora modular (MIC-101,

Billups-Rothenberg Inc. Del Mar, CA), que se llenó con una mezcla de gas industrial que contenía 5% de CO_2, 2% de O_2y 93% de nitrógeno. Para el tratamiento con ácido láctico, cada línea celular se trató con ácido láctico 0 o 12,5 mM. Además de exponer las células a cada una de las tres condiciones antemencionadas por separado, las células también se expusieron a combinaciones de dos o las tres condiciones (es decir, condiciones hiperglucémica e hipóxica; condición hiperglucémica y ácido láctico; condición hipóxica y ácido láctico; y condiciones hiperglucémica y hepóxica, y ácido láctico).

El modelo celular que comprende las células antemencionadas, donde cada tipo de célula se expuso a cada condición descrita anteriormente, se interrogaron adicionalmente mediante la exposición de las células a una perturbación ambiental mediante el tratamiento con la Coenzima Q10. Específicamente, las células se trataron con Coenzima Q10 a 0, 50 µM o 100 µM.

Las muestras celulares así como las muestras de medio que contenían el secretoma de las células, para cada línea celular expuesta a cada condición (o combinación de condiciones), con y sin tratamiento con Coenzima Q10 se recogieron en varios momentos después del tratamiento, incluyendo después de 24 horas y 48 horas de tratamiento.

Además, se llevaron a cabo experimentos de interacción entre dos células cancerosas diferentes, células PaCa2 y HepG2, en los cuales las células de PaCa2 y HepG2 se cocultivaron. Este abordaje de cocultivo se menciona como experimento de secretoma extracelular (ECS). El primer sistema celular (PaCa2) se cultivó en los insertos de los pocillos de una cámara de cultivo del tipo transwell. Generalmente se usaron placas de seis pocillos para permitir un mejor análisis estadístico. En el momento de sembrar el primer sistema celular en los insertos, se colocaron los insertos en una placa de 6 pocillos separada. El segundo sistema celular (HepG2) se sembró en la bandeja primaria. La placa de 6 pocillos que contenía los insertos, que contenía el primer sistema celular y la bandeja primaria que contenía el segundo sistema celular se incubaron a 37°C durante la noche. Cada uno de los sistemas celulares se cultivaron en su medio respectivo de células específicas (donde, de manera alternativa, cada uno de los sistemas celulares puede cultivarse en un medio adaptado para sustentar el crecimiento de ambos tipos celulares). En el segundo día, se dio el tratamiento predeterminado mediante intercambio de medios. Específicamente, los insertos que contienen el primer sistema celular y el medio respectivo del primer sistema celular se colocan en la bandeja primaria que contiene el segundo sistema celular y el medio respectivo del segundo sistema celular. En todos los casos de cocultivo, sin embargo, las células

cocultivadas que han sido expuestas a la misma "condición de cáncer"(por ejemplo, hiperglucemia, hipoxia, ácido láctico o combinaciones de estos), aunque separados, el primer día antes del cocultivo. Es decir, el primer sistema celular en los insertos y el segundo sistema celular en las bandejas se expusieron a la misma condición antes de ser movidos a la distribución de "cocultivo". La bandeja se incubó durante un período de tiempo predeterminado, *por ejemplo,* 24 horas o 48 horas. Se establecieron pocillos duplicados con las mismas condiciones, y las células se agruparon para proporcionar suficiente material para el subsiguiente análisis proteómico. El medio que contenía el secretoma (alícuota de 1 mL), las células de los insertos y las células de los pocillos de la bandeja primaria se cultivaron como muestras separadas. Los experimentos se llevaron a cabo por triplicado para proporcionar un mejor poder estadístico.

Los experimentos de interacción también se llevaron a cabo mediante experimentos de "cambio de medio". Específicamente, un medio cultivado o "secretoma" del primer sistema celular (PaCa2) se recogió después de 24 horas o 48 horas luego de la perturbación y/o condicionamiento y luego se agregó al segundo sistema celular durante 24-48 horas. Entonces se recogió medio cultivado final o "secretoma" del segundo sistema celular. Todos los secretomas finales se sometieron a análisis proteómico.

Después de la exposición del sistema celular a las "condiciones de cáncer" descritas anteriormente, la perturbación (es decir, tratamiento con Coenzima Q10), y/o las afecciones producidas por el secretoma de una célula apareada de un experimento de cocultivo, la respuesta de las células fue entonces analizada mediante un análisis de varias lecturas del sistema celular. Las lecturas incluyen datos proteómicos, específicamente la expresión de proteína intracelular así como de proteínas secretadas en el medio cultivo celular, y datos funcionales, específicamente la bioenergía celular.

El método de iProfiling de los cambios en la expresión de proteína celular total mediante proteómica cuantitativa se llevó a cabo para las muestras celulares y de medio recogidas de cada línea celular (línea celular normal y de cáncer) expuesta a cada condición (o combinación de condiciones), con o sin la "perturbación ambiental", es decir, tratamiento de Coenzima Q10, usando las técnicas descritas anteriormente en la descripción detallada.

Además, el perfilado bioenergético de cada línea celular (línea celular normal y de cáncer) expuesta a cada condición (o combinación de condiciones), con o sin la "perturbación ambiental", es decir, tratamiento de Coenzima Q10, se generó

mediante el empleo del analizador Seahorse, esencialmente como lo recomienda el fabricante. Se registraron la tasa de consumo de oxígeno (OCR) y la tasa de acidificación extracelular (ECAR) mediante los electrodos en una cámara de 7 µl creada con el cartucho apoyado contra la placa de cultivo seahorse.

Los datos proteómicos recogidos para cada línea celular en cada condición o condiciones, y con/sin cada perturbación, y los datos del perfilado bioenergético recogidos para cada línea celular en cada condición o condiciones, y con/sin cada perturbación, se procesaron entonces mediante el sistema REFS™. Una "red compuesta perturbada de cáncer" se generó a partir de los datos combinados de todas las líneas celulares de cáncer, cada una de las cuales fue expuesta a cada condición específica (o combinación de condiciones), y expuestas adicionalmente a perturbación (CoQ10). Una "red compuesta no perturbada de cáncer" se generó a partir de los datos combinados de todas las líneas celulares de cáncer, cada una de las cuales fue expuesta a cada condición específica (o combinación de condiciones), y sin perturbación (sin CoQ10). De manera similar, una "red compuesta perturbada normal" se generó a partir de los datos combinados de todas las líneas celulares normales, cada una de las cuales fue expuesta a cada condición específica (o combinación de condiciones), y expuestas adicionalmente a perturbación (CoQ10). Una "red compuesta no perturbada normal" se generó a partir de

los datos combinados de todas las líneas celulares normales, cada una de las cuales fue expuesta a cada condición específica (o combinación de condiciones), y sin perturbación (sin CoQ10).

A continuación, se generaron "redes compuestas de simulación" (también mencionadas en la presente como "redes de simulación") para cada una de las cuatro redes compuestas descritas anteriormente usando REFS™. Para lograr esto, cada nodo en la red de consenso compuesta dada se simuló (mediante el aumento o la disminución 10 veces) para generar una redes de simulación usando REFS™, como se describe en más detalle anteriormente en la descripción detallada.

El área bajo la curva y los cambios de múltiplo para cada borde que conecta un nodo padre a un nodo hijo en las redes de simulación se extrajeron mediante un programa hecho a medida usando el lenguaje de programación R, donde el lenguaje de programación R es un entorno de software de código abierto para cómputos y gráficos estadísticos.

Finalmente, se generaron las redes delta, donde las redes delta representan el diferencial entre dos redes compuestas de simulación. Las redes delta se generaron a partir de las redes compuestas de simulación. Para generar una red diferencial de cáncer en relación con un estado normal en respuesta a la

Coenzima Q10 (red delta-delta), se llevaron a cabo etapas consecutivas de comparación, como se ilustra en la Figura 26, mediante un programa construido a medida usando el lenguaje de programación PERL.

Primero, las redes de cáncer sin tratar (T0) y cáncer tratado (T1) se compararon usando el programa R, y las redes únicas T1 de cáncer tratado se separaron (véase la forma creciente en gris oscuro de la Figura 26). Esta representa la red "delta" de cáncer T1 ∩ (intersección) cáncer T0. La interacción/asociaciones de proteínas dentro de esta red delta pueden visualizarse como una representación de la respuesta única del cáncer al tratamiento con Coenzima Q10.

De manera similar, las redes normales sin tratar (T0) y normal tratada (T1) se compararon usando el programa R, y las redes únicas T1 normales tratadas se separaron (véase la forma creciente en gris claro de la Figura 26). Esta representa la red "delta de Normal T1 ∩ Normal T0. Las interacciones/asociaciones de proteínas dentro de esta red delta pueden visualizarse como una representación de la respuesta única de las células normales al tratamiento con Coenzima Q10.

Finalmente, las redes únicas de cáncer T1 (véase la forma creciente en gris oscuro en la Figura 26) y las redes únicas

normales T1 (véase la forma creciente en gris claro en la Figura 26) se compararon usando el programa R, y se generaron redes que son únicas para el cáncer únicamente, y no están presentes en células normales, en respuesta a la Coenzima Q10 (véase la Figura 26). Esta recolección de interacciones/asociaciones de las proteínas representan vías únicas dentro de las células cancerosas que no están presentes en células normales luego del tratamiento con Coenzima Q10. Esta recolección de interacciones/asociaciones de las proteínas se llama una "red delta-delta", dado que es un mapa diferencial producido a partir de una comparación de un mapa diferencial de células cancerosas y un mapa diferencial de células de control normales.

Los datos de salida de los programas PERL y R se ingresaron en Cytoscape, un programa de código abierto, para generar una representación visual de una red Delta-Delta.

Las redes delta-delta identificadas usando el método descrito en la presente son altamente útiles para identificar objetivos para un tratamiento de cáncer. Por ejemplo, de acuerdo con la red delta-delta presentada en la Figura 27, la Proteína A inhibe OCR3 (una medida de fosforilación oxidativa) y mejora ECAR3 (una medida para la glicólisis). Dado que la interacción es única en células cancerosas (dado que la red delta-delta sustrajo cualquier interacción que está comúnmente presente en células

normales después del tratamiento con Coenzima Q10), la inhibición de la expresión de proteína A se espera que reduzca el metabolismo de energía basado en glicólisis, que es una marca distintiva de la vía metabólica del cáncer, y el cambio de células hacia un metabolismo de energía basado en fosforilación oxidativa, que es un fenotipo más estrechamente asociado con células normales. Por lo tanto, se espera que una terapia de combinación que usa Coenzima Q10 e inhibidor de proteína A sea eficaz para tratar el cáncer, al menos en parte mediante un cambio en el perfil de metabolismo de energía de la célula cancerosa a uno que se asemeja al de una célula normal.

La ventaja de la plataforma de Biología Interrogativa de la invención se ilustra adicionalmente mediante el uso de un ejemplo sustantivo donde una sub-red derivada de redes causales se compara con una red molecular usando IPA, un programa informático que utiliza una red neural para determinar el vínculo molecular entre datos de salida experimentales y redes basado en *bibliografía previamente publicada*. La sub-red causal que contiene PARK7 generada usando la plataforma de Biología Interrogativa (mostrada en la Figura 29) se usa como un ejemplo sustantivo. Todas las firmas moleculares de la red PARK7 de la plataforma de Biología Interrogativa se incorporaron a IPA para generar una red basada en pruebas de la bibliografía conocida/existente. Los datos de salida entre los datos de salida

de la Biología Interrogativa y los generados mediante el uso de IPA se compararon.

Se observó que seis marcadores identificados mediante los datos de salida de la tecnología de plataforma de Biología Interrogativa (mostrados en la Figura 29), es decir, A, B, C, X, Y y Z en las Figuras 27-29, están conectados a TP53 dentro de la red generada por IPA (Figura 28). Entre los seis marcadores, A, B y C han sido reportados en la bibliografía como asociados con el cáncer, así como HSPA1A/HSPA1B. X, Y y X se identificaron como "concentradores" o activadores clave del estado canceroso, y por lo tanto se identificaron como marcadores de cáncer novedosos. Además, MIF1 y KARS también se identificaron como "concentradores" o activadores clave del estado canceroso, y por lo tanto se identificaron como marcadores de cáncer novedosos. El hecho de que los factores identificados mediante el uso de la plataforma de Biología Interrogativa compartan la conectividad con otros factores conocidos publicados en la bibliografía científica, valida la exactitud de la red creada mediante el uso de la Plataforma de Biología Interrogativa Además, la asociación de la red dentro de la sub-red de PARK7 creada mediante el uso de datos de salida de la plataforma de Biología Interrogativa (mostrados en la Figura 29) demostró la presencia de la influencia direccional de cada factor, al contrario de la red IPA (mostrada en la Figura 28) donde el vínculo entre las entidades

moleculares no proporciona direccionalidad funcional entre los nodos que interactúan. Además, los datos de salida de la plataforma de Biología Interrogativa (mostrados como líneas punteadas en la Figura 29) demostraron la asociación de estos componentes que lleva a un mecanismo potencial a través de PARK7. Se observó que la Proteína C, la Proteína A y otros nodos de PARK7 son activadores clave del metabolismo del cáncer (Figura 27).

Como lo prueba el presente ejemplo, el empleo de un abordaje imparcial para la generación de datos, integración e ingeniería inversa para crear un modelo computacional seguido de simulación y análisis de red diferencial, el descubrimiento de la plataforma de Biología Interrogativa permite el entendimiento de los mecanismos desconocidos hasta ahora en la patofisiología del cáncer que son congruentes con conocimientos científicos establecidos de la patofisiología de enfermedades.

EJEMPLO 3: Empleo de Tecnología de Plataforma para construir una red delta-delta de diabetes/obesidad/enfermedad cardiovascular

En este ejemplo, la tecnología de plataforma descrita en detalle anteriormente en la descripción detallada se empleó para integrar los datos obtenidos de un modelo de

diabetes/obesidad/enfermedad cardiovascular (CVD) construido a medida, y para identificar proteínas/vías novedosas que dirigen la patogénesis de la diabetes/obesidad/CVD. Los mapas relacionales que resultan de estos análisis proporcionaron objetivos de tratamiento de la diabetes/obesidad/CVD, así como marcadores de diagnóstico/pronóstico asociados con la diabetes/obesidad/CVD.

Se sometieron cinco líneas celulares humanas primarias, a saber, adipocitos, miotubos, hepatocitos, células de músculo liso aórtico (HASMC), y células tubulares proximales (HK2) a una de cinco condiciones simulando un ambiente experimentado por estas células relevantes de enfermedades *in vivo*. Específicamente, cada una de las cinco líneas celulares se expusieron por separado a cada una de las siguientes condiciones: condiciones hiperglucémicas, condiciones hiperlipidímicas, condiciones hiperinsulinémicas, condiciones de hipoxia y exposición al ácido láctico. La condición hiperglucémica se indujo mediante el cultivo de las células en medio que contenía glucosa 22 mM. La condición hiperlipidémica se indujo mediante el cultivo de las células en medio que contenía palmitato de sodio 0,15 mM. La condición hiperinsulinémica se indujo mediante el cultivo de las células en medio que contenía insulina 1000 nM. La condición hipóxica se indujo mediante la colocación de las células en una cámara incubadora modular (MIC-101, Billups-Rothenberg Inc. Del

Mar, CA), que se llenó con una mezcla de gas industrial que contenía 5% de CO_2, 2% de O_2y 93% de nitrógeno. Cada línea celular también se trató con ácido láctico 0 o 12,5 mM.

Además, los experimentos de interacción entre dos pares diferentes de células, HASMC (sistema celular 1) y células HK2 (sistema celular 2) o células hepáticas (sistema celular 1) y adipocitos (sistema celular 2), se llevaron a cabo de manera que los pares de células se cocultivaran. Este abordaje de cocultivo se menciona como experimento de secretoma extracelular (ECS). El primer sistema celular (por ejemplo, HASMC) se cultivó primero en los insertos de los pocillos de una cámara de cultivo del tipo transwell. Se usaron placas de seis pocillos para permitir un mejor análisis estadístico. En el momento de sembrar el primer sistema celular en los insertos, se colocaron los insertos en una placa de 6 pocillos separada. El segundo sistema celular (por ejemplo, HK2) se sembró en la bandeja primaria. La bandeja de inserto que contenía el primer sistema celular y la bandeja primaria que contenía el segundo sistema celular se incubaron a 37 °C durante la noche. Cada uno de los sistemas celulares se cultivaron en el medio específico de células específicas (donde, de manera alternativa, cada uno de los sistemas celulares puede cultivarse en un medio adaptado para sustentar el crecimiento de ambos tipos celulares). En el segundo día, se dio el tratamiento predeterminado mediante intercambio de medios. Específicamente,

los insertos que contenían el primer sistema celular se colocaron en la bandeja primaria que contenía el segundo sistema celular. La bandeja se incubó durante un período de tiempo predeterminado, *por ejemplo,* 24 horas o 48 horas. Se establecieron pocillos duplicados con las mismas condiciones, y las células se agruparon para proporcionar suficiente material para el análisis 2D. El medio (alícuota de 1 mL), las células de los insertos y las células de los pocillos de la bandeja primaria se cultivaron como muestras separadas. Los experimentos se llevaron a cabo por triplicado para proporcionar un mejor poder de análisis estadístico.

Los experimentos de interacción también se llevaron a cabo mediante experimentos de "cambio de medio". Específicamente, un medio cultivado o "secretoma" del primer sistema celular, HASMC, se recogió después de 24 horas o 48 horas luego de la perturbación o condicionamiento y luego se agregó al segundo sistema celular, adipocitos, durante 24-48 horas. Entonces se recogió medio cultivado final o "secretoma" del segundo sistema celular. Todos los secretomas finales se sometieron a análisis proteómico.

El modelo celular que comprende las células antemencionadas, donde las células se expusieron a cada condición descrita anteriormente, se "interrogaron" adicionalmente mediante

la exposición de las células a una "perturbación ambiental" mediante el tratamiento con la Coenzima Q10. Específicamente, las células se trataron con la Coenzima Q10 A 0,50 µM o 100 µM.

Las muestras para cada línea celular con cada condición y cada tratamiento con Coenzima Q10 se recogieron en varios momentos después del tratamiento, incluyendo después de 24 horas y 48 horas de tratamiento. Para ciertas células y bajo ciertas condiciones, también se recogieron muestras de medio y se analizaron.

El método de iProfiling de los cambios en la expresión de proteína celular total mediante proteómica cuantitativa se llevó a cabo para las muestras celulares y de medio recogidas de cada línea celular en cada condición y con cada "perturbación ambiental", es decir, tratamiento de Coenzima Q10, usando las técnicas descritas anteriormente en la descripción detallada.

Los datos proteómicos recogidos para cada línea celular listados anteriormente en cada condición, y con cada perturbación, y los datos del perfilado bioenergético recogidos para cada línea celular en cada condición y con cada perturbación, se procesaron entonces mediante el sistema REFS™. Una red compuesta perturbada se generó a partir de datos

combinados obtenidos de todas las líneas celulares para una condición específica (por ejemplo, hiperglucemia) expuesta a la perturbación (CoQ10). Una red compuesta no perturbada se generó a partir de datos combinados obtenidos de todas las líneas celulares para la misma condición específica (por ejemplo, hiperglucemia) sin la perturbación (sin CoQ10). De manera similar, una red compuesta perturbada se generó a partir de datos combinados obtenidos de todas las líneas celulares para una segunda condición de control (por ejemplo, glucemia normal) expuesta a la perturbación (CoQ10). Una red compuesta no perturbada se generó a partir de datos combinados obtenidos de todas las líneas celulares para la misma segunda condición de control (por ejemplo, glucemia normal) sin la perturbación (sin CoQ10).

Cada nodo en las redes compuestas de consenso descritas anteriormente se simuló (mediante el aumento o la disminución 10 veces) para generar una redes de simulación usando REFS™, como se describe en más detalle anteriormente en la descripción detallada.

El área bajo la curva y los cambios de múltiplo para cada borde que conecta un nodo padre a un nodo hijo en las redes de simulación se extrajeron mediante un programa hecho a medida usando el lenguaje de programación R, donde el lenguaje de

programación R es un entorno de software de código abierto para cómputos y gráficos estadísticos.

Las redes delta se generaron a partir de las redes compuestas simuladas. Para generar una red diferencial de una condición diabetes/obesidad/enfermedad cardiovascular en relación con un estado normal en respuesta a la Coenzima Q10 (red delta-delta), se llevaron a cabo etapas de comparación, como se ilustra en la Figura 30, mediante un programa construido a medida usando el lenguaje de programación PERL.

Específicamente, como se muestra en la Figura 30, Tratamiento T1 se refiere al tratamiento con Coenzima Q1 y NG y HG se refiere a una estado normal o una condición de hiperglucemia. Se compararon los bordes únicos de NG en la red delta NG∩HG con bordes únicos de HGT1 en la red delta HG∩HGT1. Los bordes en la intersección de NG y HGT1 son bordes HG que se restauran a NG con T1. Los bordes HG restaurados a NG con T1 se superpusieron con la red delta NG∩HG (mostrada en círculos de color oscuro en la Figura 31).

Específicamente, un mapa compuesto simulado de un estado de glucemia normal (NG) y un mapa compuesto simulado de una condición de hiperglucemia (HG) se compararon usando un programa Perl hecho a medida para generar bordes únicos de la condición de

glicemia normal. Un mapa compuesto simulado de una condición de hiperglucemia sin tratamiento de Coenzima Q10 (HG) y un mapa simulado de una condición de hiperglucemia con tratamiento de Coenzima Q10 (HGT1) se compararon usando un programa Perl hecho a medida para generar bordes únicos de la condición de hiperglucemia con tratamiento con Coenzima Q10 (HGT1). Los bordes en la intersección de los bordes únicos del estado normal de glucemia (NG) y los bordes únicos de la condición de hiperglucemia con tratamiento con Coenzima Q10 (HGT1) se identificaron usando el programa Perl. Estos bordes representan factores/redes que se restauran al estado normal de glucemia a partir de la condición de hiperglucemia mediante el tratamiento con Coenzima Q10. La red delta-delta de bordes hiperglicémicos restaurados a normales con tratamiento con la Coenzima Q10 se superpuso en la red delta de glucemia normal ∩ Hiperglucemia. Se muestra una muestra de las redes superpuestas en la Figura 31. La Figura 31 es un ejemplo de red diferencial de condición de diabetes/obesidad/enfermedad contra el estado normal en respuesta a la Coenzima Q10 (red delta-delta). Los círculos de colores más oscuros en la Figura 31 son bordes identificados que se restauraron al estado de glucemia normal a partir de la condición de hiperglucemia mediante el tratamiento de Coenzima Q10. Los círculos de colores más claros en la Figura 31 son bordes de hipercemia normal única identificada.

Los datos de salida de los programas PERL y R se ingresaron en Cytoscape, un programa de código abierto, para generar una representación visual de una red Delta-Delta.

De manera similar a los experimentos descritos anteriormente para la condición de hiperglucemia contra estado normal glucémico, se compararon una red compuesta simulada de condición de hiperlipidemia (combinando datos de todas las células relacionadas con diabetes/obesidad/enfermedad cardiovascular antemencionadas) sin tratamiento de Coenzima Q10 y una red compuesta simulada de condición de hiperlipidemia (combinando datos de todas las células relacionadas con diabetes/obesidad/enfermedad cardiovascular antemencionada) con tratamiento de Coenzima Q10, usando el programa Perl para generar bordes únicos de la condición de hiperlipidemia con tratamiento de Coenzima Q10. Los bordes en la intersección de los bordes únicos del estado normal de lipidemia y los bordes únicos de la condición de hiperlipidémica con tratamiento con Coenzima Q10 se identificaron usando el programa Perl. Estos bordes representan factores/redes que se restauraron al estado de lipidemia normal a partir de la condición de hiperlipidemia mediante el tratamiento de Coenzima Q10. Una red delta-delta de bordes hiperlipidémicos restaurados a normales con tratamiento con la Coenzima Q10 se superpuso en la red delta de lipidemia normal ∩ Hiperlipidemia. Se muestra una muestra de las redes superpuestas en la Figura 32.

Los círculos de colores más oscuros en la Figura 32 son bordes identificados que se restauraron al estado de lipidemia normal a partir de la condición de hiperlipidemia mediante el tratamiento de Coenzima Q10. Los círculos de colores más claros en la Figura 32 son bordes de lipidemia normal única identificada. Se identificó FASN como un factor importante de una vía de señalización que modula el efecto de la Coenzima Q10 para restaurar la hiperlipidemia a un estado normal de lipidemia.

La sintasa de ácido graso - enzimas de síntesis de ácido graso tal como FASN se han implicado en casi todos los aspectos de las alteraciones metabólicas humanas tal como obesidad, resistencia a la insulina o dislipidemia. Se han propuesto los inhibidores de FASN como moléculas líder para el tratamiento de la obesidad, aunque sus mecanismos moleculares son desconocidos (Mobbs et ál. 2002). Se ha demostrado que la cerulenina y el compuesto sintético C75, que son inhibidores de FASN, tienen un efecto en la reducción de la ingesta de alimentos y producen la disminución del peso (Loftus *et ál.* 2000).

El hecho de que el FASN haya sido identificado mediante la plataforma de tecnología descrita en la presente como un factor importante en la vía de señalización que modula el efecto de la Coenzima Q10 para restaurar a un diabético al estado normal, como se muestra en la Figura 32, valida la exactitud de esta red

delta-delta. Por lo tanto, otros factores novedosos identificados en esta red delta-delta serán factores terapéuticos potenciales o fármacos objetivo para su investigación adicional.

EJEMPLO 4: Empleo de la tecnología de plataforma para construir modelos de cardiotoxicidad inducida por fármacos

En este ejemplo, la tecnología de plataforma descrito en detalle anteriormente en la descripción detallada se empleó para integrar los datos obtenidos de un modelo cardiotoxicidad construido a medida, y para identificar proteínas/vías novedosas que dirigen la patogénesis/toxicidad de los fármacos. Los mapas relacionales resultantes de este análisis proporcionaron biomarcadores de toxicidad.

En un corazón sano, la función de contracción depende del balance de la oxidación de carbohidratos y ácidos grasos. Se piensa que el desbalance crónico en la absorción, utilización, biogénesis organelar y secreción en tejido no adiposo (corazón e hígado) se encuentra en el centro del daño y disfunción mitocondrial y es un punto clave en la cardiotoxicidad inducida por fármacos.

Los solicitantes describen en la presente un abordaje de sistema que combina firmas de proteína y lípidos con ensayos de punto extremo funcionales que apuntan específicamente a la bioenergía celular y la función de la membrana mitocondrial. Se trataron modelos in vitro que comprenden cardiomiocitos diabéticos y normales complementados con un exceso de ácidos grasos e hiperglucemia, con un panel de fármacos para crear firmas y mecanismos potenciales para la toxicidad. Los solicitantes demostraron los efectos variados de los fármacos en la desestabilización de la mitocondria mediante la interrupción del componente de metabolismo de energía en varios niveles, incluyendo (i) Desregulación de las redes transcripcionales que controlan la expresión de los genes de metabolismo de energía mitocondriales; (ii) Inducción de GPAT1 y tafazzina en cardiomiocitos diabéticos, iniciando así de novo la síntesis de fosfolípidos y remodelando la membrana mitocondrial; y (iii) Destino alterado del ácido graso en cardiomiocitos diabéticos, influenciando la absorción, oxidación de ácido graso y síntesis de ATP. Además, los solicitantes combinaron el poder de la biología de laboratorio húmedo y una plataforma de recolección de datos basada en AI para generar una red causal basada en modelos bayesianos. Las redes de proteínas y lípidos que son causal de la pérdida de la función celular normal se usaron para discernir los mecanismos de la toxicidad inducida por fármacos a partir de mecanismos protectores celulares. Este

abordaje novedoso servirá como una poderosa nueva herramienta para entender el mecanismo de la toxicidad, mientras se permite el desarrollo de terapias más seguras para corregir un fenotipo alterado.

Los cardiomiocitos se sometieron a condiciones que simulaban un ambiente diabético que es experimentado por las células relevantes de la enfermedad *in vivo*. Específicamente, las células se expusieron a condiciones hiperglucémicas y condiciones de hiperlipidemia. La condición hiperglucémica se indujo mediante el cultivo de las células en medio que contenía glucosa 22 mM. La condición de hiperlipidemia se indujo mediante el cultivo de las células en medio que contenía L-carnitina 1 mM, ácido oleico 0,7 mM y ácido linoleico 0,7 mM.

El modelo celular que comprende las células antemencionadas, donde las células se expusieron a cada condición descrita anteriormente, se "interrogó" adicionalmente mediante la exposición de las células a una "perturbación ambiental" mediante el tratamiento con un fármaco contra la diabetes (T) que se sabe que causa cardiotoxicidad, una molécula de rescate (R) o ambos, el fármaco contra la diabetes y la molécula de rescate (T+R). Específicamente las células se trataron con fármaco contra la diabetes; o se trataron con molécula de rescate Coenzima Q10 a 0,

50 µM o 100 µM; o se trataron con ambos, el fármaco contra la diabetes y la molécula de rescate Coenzima Q10.

Las muestras de cada condición con cada tratamiento de perturbación se recogieron en varios momentos después del tratamiento, incluyendo después de 6 horas de tratamiento. Para ciertas condiciones, también se recogieron muestras de medio y se analizaron.

El método de iProfiling de los cambios en la expresión de proteína celular total mediante proteómica cuantitativa se llevó a cabo para las muestras celulares y de medio recogidas de cada condición y con cada "perturbación ambiental", es decir, tratamiento con fármaco contra la diabetes, tratamiento con Coenzima Q10 o ambos, usando las técnicas descritas anteriormente en la descripción detallada. Los experimentos de perfilado transcripcional se llevaron a cabo usando un sistema de amplificación cfx-384 de Biorad. Luego de la recolección de datos (Ct), el cambio final con respecto al testigo se determinó usando el método δCt tal como se indica en el protocolo del fabricante. Los experimentos lipidómicos se llevaron a cabo usando espectrometría de masa. Los ensayos funcionales, tal como tasa de consumo de oxígeno OCR se midieron mediante el uso del analizador Seahorse, esencialmente del modo recomendado por el fabricante. La OCR) se registró mediante los electrodos en una cámara de 7 µl

creada con el cartucho apoyado contra la placa de cultivo seahorse.

Como se muestra en la Figura 35, la red transcripcional y la expresión de los genes de metabolismo de energía mitocondrial humanos en cardiomiocitos diabéticos (cardiomiocitos condicionados en hiperglucemia e hiperlipidemia) se compararon entre tratamientos perturbados y no perturbados. Específicamente, los datos de la red transcripcional y los genes de metabolismo de energía mitocondrial humanos se compararon entre cardiomiocitos diabéticos tratados con un fármaco contra la diabetes (T) y muestras de cardiomiocitos diabéticos no tratados (UT). Los datos de la red Transcripcional y los genes de metabolismo de energía mitocondrial humanos se compararon entre cardiomiocitos diabéticos tratados con un fármaco contra la diabetes y una molécula de rescate Coenzima Q10 (T+R) y muestras de cardiomiocitos diabéticos no tratados (UT). Al comparar los datos de cardiomiocitos diabéticos no tratados, ciertos genes de expresión y transcripción se alteraron cuando los cardiomiocitos diabéticos fueron tratados con el fármaco contra la diabetes. Se demostró que la molécula de rescate Coenzima Q10 revierte el efecto tóxico del fármaco contra la diabetes y normaliza la expresión génica y la transcripción.

Como se muestra en la Figura 36A, los cardiomiocitos se cultivaron ya sea en normoglucemia (NG) o condición de hiperglucemia (HG) y se trataron con un fármaco contra la diabetes solo (T) o con ambos, un fármaco contra la diabetes y una molécula de rescate Coenzima Q10 (T+R). Los niveles de expresión de proteínas de GPAT1 y TAZ para cada condición y cada tratamiento se analizaron con transferencia western. Tanto GPAT1 como TAZ se regularon por aumento en los cardiomiocitos condicionados por hiperglucemia y tratados con fármaco contra la diabetes. Cuando los cardiomiocitos condicionados por hiperglucemia se trataron tanto con el fármaco contra la diabetes como con la molécula de rescate Coenzima Q10, el nivel de expresión de la proteína regulada por aumento de GPAT1 y TAZ se normalizó.

Como se muestra en la Figura 37A, los experimentos de la tasa de consumo de oxígeno mitocondrial (%) se llevaron a cabo para las muestras de cardiomiocitos condicionados por hiperglucemia. Los cardiomiocitos condicionados por hiperglucemia fueron no tratados (UT), tratados con un fármaco contra la diabetes T1 que se sabe que causa cardiotoxicidad, tratados con un fármaco contra la diabetes T2 que se sabe que causa cardiotoxicidad, tratados con un fármaco contra la diabetes T1 y una molécula de rescate Coenzima Q10 (T1+R), o tratados con un fármaco contra la diabetes T2 y una molécula de rescate Coenzima

Q10 (T2+R). Al comparar de las muestras de control sin tratar, la OCR mitocondrial disminuyó cuando los cardiomiocitos condicionados por hiperglucemia se trataron con el fármaco contra la diabetes T1 o T2. Sin embargo, la OCR mitocondrial se normalizó cuando los cardiomiocitos condicionados por hiperglucemia se trataron con el fármaco contra la diabetes y la molécula de rescate Coenzima Q10 (T1+R o T2+R).

Como se muestra en la Figura 37B, los experimentos de síntesis de ATP mitocondrial se llevaron a cabo para las muestras de cardiomiocitos condicionados por hiperglucemia. Los cardiomiocitos condicionados por hiperglucemia fueron no tratados (UT), tratados con un fármaco contra la diabetes (T), o tratados con un fármaco contra la diabetes y una molécula de rescate Coenzima Q10 (T+R). Al comparar de las muestras de control sin tratar, la síntesis de ATP mitocondrial se reprimió cuando los cardiomiocitos condicionados por hiperglucemia se trataron con el fármaco contra la diabetes (T).

Como se muestra en la Figura 38, basándose en los datos proteómicos recogidos, las proteínas reguladas por disminución mediante el tratamiento con fármacos se anotaron con términos GO. Las proteínas involucradas en el metabolismo de energía mitocondrial se regularon por disminución cuando los

cardiomiocitos condicionados por hiperglucemia se trataron con un fármaco contra la diabetes que se sabe causa cardiotoxicidad.

Los datos de proteómica, lipidómica, perfilado transcripcional, ensayos funcionales y transferencia western recogidos para cada condición y cada perturbación se procesaron entonces mediante el sistema REFS™. Las redes compuestas perturbadas se generaron a partir de los datos combinados de una condición específica (por ejemplo, hiperglucemia o hiperlipidemia) expuesta a cada perturbación (por ejemplo, fármaco contra la diabetes, CoQ10 o ambos). Las redes compuestas no perturbadas se generaron a partir de los datos combinados obtenidos de la misma condición específica (por ejemplo, hiperglucemia o hiperlipidemia) sin perturbación (sin tratar). De manera similar, las redes compuestas perturbadas se generaron a partir de los datos combinados de una segunda condición de control (por ejemplo, glucemia normal) expuesta a cada perturbación (por ejemplo, fármaco contra la diabetes, CoQ10 o ambos). Las redes compuestas no perturbadas se generaron a partir de los datos combinados obtenidos de la misma segunda condición de control (por ejemplo, glucemia normal) sin perturbación (sin tratar).

Cada nodo en las redes compuestas de consenso descritas anteriormente se simuló (mediante el aumento o la disminución 10

veces) para generar una redes de simulación usando REFS™, como se describe en más detalle anteriormente en la descripción detallada.

El área bajo la curva y los cambios de múltiplo para cada borde que conecta un nodo padre a un nodo hijo en las redes de simulación se extrajeron mediante un programa hecho a medida usando el lenguaje de programación R, donde el lenguaje de programación R es un entorno de software de código abierto para cómputos y gráficos estadísticos.

Las redes delta se generaron a partir de las redes compuestas simuladas. Para generar una red diferencial de una condición de toxicidad inducida por fármacos en relación con un estado normal en respuesta al fármaco contra la diabetes (red delta), se llevaron a cabo etapas de comparación, como se ilustra en la Figura 39, mediante un programa construido a medida usando el lenguaje de programación PERL.

Específicamente, como se muestra en la Figura 39, UT se refiere a las redes de expresión de proteínas de los cardiomiocitos de control sin tratar en una condición de hiperglucemia. Tratamiento T se refiere a redes de expresión de proteínas de cardiomiocitos tratados con fármacos contra la

diabetes en una condición de hiperglucemia. Los bordes únicos de T en la red delta UT∩T se presentan en la Figura 40.

Específicamente, un mapa compuesto simulado de una condición de hiperglucemia en cardiomiocito no tratados y un mapa compuesto simulado de una condición de hiperglucemia en cardiomiocitos no tratados con un fármaco contra la diabetes se compararon usando un programa Perl hecho a medida para generar bordes únicos de los cardiomiocitos tratados con un fármaco contra la diabetes en una condición de hiperglucemia. Los datos de salida de los programas PERL y R se ingresaron en Cytoscape, un programa de código abierto, para generar una representación visual de una red delta. Como se muestra en la Figura 40, la red representa las redes delta que son dirigidas por el fármaco contra la diabetes con respecto a los cardiomiocitos no tratados/modelos de cardiotoxicidad en una condición de hiperglucemia.

A partir de la red diferencial de la condición de toxicidad inducida por fármacos con respecto a la condición normal mostrada en la Figura 40, se identificaron proteínas que impulsan la patofisiología de la cardiotoxicidad inducida por fármacos, tal como GRP78, GRP75, TIMP1, PTX3, HSP76, PDIA4, PDIA1, CA2D1. Estas proteínas pueden funcionar como biomarcadores para la identificación de otros fármacos que inducen la

cardiotoxicidad. Estas proteínas también pueden funcionar como biomarcadores para la identificación de agentes que pueden aliviar la cardiotoxicidad.

Los experimentos descritos en este Ejemplo demuestran que la biología de membrana perturbada y el destino alterado del ácido graso libre en cardiomiocitos diabéticos expuestos al tratamiento con fármacos representa una pieza central en la toxicidad inducida por fármacos. La integración de datos y la biología de red permitieron una comprensión mejorada de la cardiotoxicidad, y la identificación de biomarcadores predictivos novedosos para la cardiotoxicidad.

Incorporación mediante referencia

El contenido de todas las referencias citadas (incluyendo las referencias bibliográficas, patentes, solicitudes de patente y sitios web) que pueden citarse a lo largo de la presente, se incorporan expresamente mediante esta referencia en su totalidad, como lo están las referencias citadas en estas. La puesta en práctica de la presente invención empleará, a menos que se indique lo contrario, técnicas convencionales de formulación de proteínas, que son conocidas en la técnica.

Equivalentes

La invención puede realizarse en otras formas específicas sin apartarse del espíritu o características esenciales de esta. Las realizaciones antemencionadas deben, por lo tanto, considerarse en todo aspecto como ilustrativas y no como limitantes de la invención descrita en la presente. El alcance de la invención está indicado, por lo tanto, por las reivindicaciones adjuntas en vez de por la descripción que antecede, y todos los cambios que surjan del significado y rango de equivalencia de las reivindicaciones se pretende que estén incorporadas a estas.

REIVINDICACIONES

Reivindicamos:

1. Un método para identificar un modulador de un sistema biológico, dicho método comprende:

(1) establecer el modelo para el sistema biológico, usar células asociadas con el sistema biológico, para representar un aspecto característico del sistema biológico;

(2) obtener un primer conjunto de datos del modelo para el sistema biológico, donde el primer conjunto de datos representa los niveles de expresión de una pluralidad de genes en las células asociadas con el sistema biológico;

(3) obtener un segundo conjunto de datos del modelo para el sistema biológico, donde el segundo conjunto de datos representa una actividad funcional o una respuesta celular de las células asociadas con el sistema biológico;

(4) generar una red de relaciones causales de consenso entre los niveles de expresión de la pluralidad de genes y la actividad funcional o la respuesta celular basada únicamente en el primer conjunto de datos y el segundo conjunto de datos, usando un dispositivo de computación programado, donde la generación de la red de relaciones causales de consenso no está basada en ninguna relación biológica conocida que no sea el primer conjunto de datos o el segundo conjunto de datos;

(5) identificar, a partir de la red de relaciones causales de consenso, una relación causal única en el sistema biológico, donde un gen asociado con la relación causal única se identifica como un modulador del sistema biológico.

2. El método de la reivindicación 1, donde el modulador estimula o promueve el sistema biológico.

3. El método de la reivindicación 1, donde el modulador inhibe el sistema biológico.

4. El método de la reivindicación 1, en donde el modelo del sistema biológico comprende un cultivo in vitro de células asociadas con el sistema biológico, comprendiendo además opcionalmente un cultivo in vitro correspondiente de células testigo.

5. El método de la reivindicación 4, en donde el cultivo in vitro de las células se somete a una perturbación ambiental, y el cultivo in vitro de las células testigo correspondiente es de células idénticas que no se sometieron a la perturbación ambiental.

6. El método de la reivindicación 4, en donde la perturbación ambiental comprende uno o más de un contacto con un

agente, un cambio en la condición de cultivo, una modificación / mutación genética introducida y un vehículo *(por ejemplo,* vector) que provoca una modificación / mutación genética.

7. El método de la reivindicación 1, en donde el primer conjunto de datos comprende niveles de expresión de proteína y/o ARNm de la pluralidad de genes.

8. El método de la reivindicación 1, en donde el primer conjunto de datos comprende además uno o más datos lipidómicos, datos metabolómicos, datos transcriptómicos y datos de polimorfismos de nucleótidos simples (SNP).

9. El método de la reivindicación 1, en donde el segundo conjunto de datos comprende uno o más de perfil bioenergético, proliferación celular, apoptosis, función organelar y una asociación genotipo-fenotipo actualizados mediante modelos funcionales seleccionados de ensayos ATP, ROS, OXPHOS y Seahorse.

10. El método de la reivindicación 1, en donde la etapa (4) se lleva a cabo mediante una plataforma informática basada en inteligencia artificial (AI).

11. El método de la reivindicación 10, en donde la plataforma informática basada en AI comprende REFS (TM).

12. El método de la reivindicación 1, en donde la plataforma informática basada en AI recibe todos los datos de entrada del primer conjunto de datos y el segundo conjunto de datos sin aplicar un punto de corte estadístico.

13. El método de la reivindicación 1, en donde la red de relaciones causales de consenso establecida en la etapa (4) se refina adicionalmente hasta una red de relaciones causales de simulación, antes de la etapa (5), mediante simulación *in silico* basada en los datos de entrada, para proporcionar un nivel de confianza de predicción para una o más relaciones causales dentro de la red de relaciones causales de consenso.

14. El método de la reivindicación 1, en donde la relación causal única se identifica como parte de una red de relaciones causales diferenciales que está presente exclusivamente en células y está ausente en las células testigo correspondientes.

15. El método de la reivindicación 1, que comprende además validar la relación causal única identificada en un sistema biológico.

16. Un método para identificar un modulador de un proceso de enfermedad, dicho método comprende:

(1) establecer un modelo de enfermedad para el proceso de enfermedad, usando células relacionadas con la enfermedad para representar un aspecto característico del proceso de enfermedad;

(2) obtener un primer conjunto de datos del modelo de enfermedad, donde el primer conjunto de datos representa los niveles de expresión de una pluralidad de genes en las células relacionadas con la enfermedad;

(3) obtener un segundo conjunto de datos del modelo de enfermedad, donde el segundo conjunto de datos representa una actividad funcional o una respuesta celular de las células relacionadas con la enfermedad;

(4) generar una red de relaciones causales de consenso entre los niveles de expresión de la pluralidad de genes y la actividad funcional o la respuesta celular basada únicamente en el primer conjunto de datos y el segundo conjunto de datos, usando un dispositivo de computación programado, donde la generación de la red de relaciones causales de consenso no está basada en ninguna relación biológica conocida que no sea el primer conjunto de datos o el segundo conjunto de datos;

(5) identificar, a partir de la red de relaciones causales de consenso, una relación causal única en el proceso de enfermedad, donde un gen asociado con la relación causal única se identifica como un modulador de un proceso de enfermedad.

17. El método de la reivindicación 16, en donde el proceso de enfermedad es cáncer, diabetes, obesidad o enfermedad cardiovascular.

18. El método de la reivindicación 17, en donde el cáncer es cáncer pulmonar, cáncer de mama, cáncer de próstata, melanoma, carcinoma escamoso, cáncer colorrectal, cáncer pancreático, cáncer tiroideo, cáncer endometrial, cáncer de vejiga, cáncer de riñones, un tumor sólido, leucemia, linfoma no hodgkiniano o un cáncer resistente a fármacos.

19. El método de la reivindicación 16, donde el modulador estimula o promueve el proceso de enfermedad.

20. El método de la reivindicación 16, donde el modulador inhibe el proceso de enfermedad.

21. El método de la reivindicación 20, en donde el modulador cambia la vía metabólica energética específicamente en las células de la enfermedad de una vía glucolítica hacia una vía de fosforilación oxidativa.

22. El método de la reivindicación 16, en donde el modelo de enfermedad comprende un cultivo in vitro de células de

enfermedad, que comprende además opcionalmente un cultivo in vitro correspondiente de células testigo o normales.

23. El método de la reivindicación 22, en donde el cultivo in vitro de las células de la enfermedad se somete a una perturbación ambiental, y el cultivo in vitro de las células testigo correspondiente es de células de la enfermedad idénticas que no se sometieron a la perturbación ambiental.

24. El método de la reivindicación 22, en donde la perturbación ambiental comprende uno o más de un contacto con un agente, un cambio en la condición de cultivo, una modificación / mutación genética introducida y un vehículo (por ejemplo, vector) que provoca una modificación / mutación genética.

25. El método de la reivindicación 16, en donde el aspecto característico del proceso de enfermedad comprende una condición de hipoxia, una condición hiperglucémica, una condición de cultivo rico en ácido láctico o combinaciones de los mismos.

26. El método de la reivindicación 16, en donde el primer conjunto de datos comprende niveles de expresión de proteína y/o ARNm de la pluralidad de genes.

27. El método de la reivindicación 16, en donde el primer conjunto de datos comprende además uno o más datos lipidómicos, datos metabolómicos, datos transcriptómicos y datos de polimorfismos de nucleótidos simples (SNP).

28. El método de la reivindicación 16, en donde el segundo conjunto de datos comprende uno o más de perfil bioenergético, proliferación celular, apoptosis, función organelar y una asociación genotipo-fenotipo actualizados mediante modelos funcionales seleccionados de ensayos ATP, ROS, OXPHOS y Seahorse.

29. El método de la reivindicación 16, en donde la etapa (4) se lleva a cabo mediante una plataforma informática basada en inteligencia artificial (AI).

30. El método de la reivindicación 29, en donde la plataforma informática basada en AI comprende REFS (TM).

31. El método de la reivindicación 16, en donde la plataforma informática basada en AI recibe todos los datos de entrada del primer conjunto de datos y el segundo conjunto de datos sin aplicar un punto de corte estadístico.

32. El método de la reivindicación 16, en donde la red de relaciones causales de consenso establecida en la etapa (4) se

refina adicionalmente hasta una red de relaciones causales de simulación, antes de la etapa (5), mediante simulación *in silico* basada en los datos de entrada, para proporcionar un nivel de confianza de predicción para una o más relaciones causales dentro de la red de relaciones causales de consenso.

33. El método de la reivindicación 16, en donde la relación causal única se identifica como parte de una red de relaciones causales diferenciales que está presente exclusivamente en células de enfermedad y está ausente en las células testigo correspondientes.

34. El método de la reivindicación 16, que comprende además validar la relación causal única identificada en un sistema biológico.

35. Un método para proporcionar un modelo para un sistema biológico para su uso en un método de plataforma, que comprende:

establecer un modelo para un sistema biológico, usando células asociadas con el sistema biológico, para representar un aspecto característico del sistema biológico, en donde el modelo para el sistema, donde el modelo para el sistema biológico es útil para generar conjuntos de datos usados en el método de plataforma; proporcionando así un modelo para un sistema biológico para uso en un método de plataforma.

36. Un método para obtener un primer conjunto de datos y un segundo conjunto de datos a partir de un modelo para un sistema biológico para uso en un método de plataforma, que comprende:

(1) obtener un primer conjunto de datos a partir del modelo para un sistema biológico para uso en un método de plataforma, en donde el modelo para el sistema biológico comprende células asociadas con el sistema biológico y en donde el primer conjunto de datos representa niveles de expresión de una pluralidad de genes en las células asociadas con el sistema biológico;

(2) obtener un segundo conjunto de datos a partir del modelo para un sistema biológico para uso en el método de plataforma, en donde el segundo conjunto de datos representa una actividad funcional o una respuesta celular de las células asociadas con el sistema biológico;

obteniendo así un primer conjunto de datos y un segundo conjunto de datos a partir del modelo para el sistema biológico para uso en un método de plataforma.

37. Un método para identificar un modulador de un sistema biológico, dicho método comprende:

(1) generar una red de relaciones causales de consenso entre un primer conjunto de datos y un segundo conjunto de datos

obtenidos a partir de un modelo para el sistema biológico, en donde el modelo comprende células asociadas con el sistema biológico y en donde el primer conjunto de datos representa niveles de expresión de una pluralidad de genes en las células y el segundo conjunto de datos representa una actividad funcional o una respuesta celular de las células, usando un dispositivo informático programado, en donde la generación de la red de relaciones causales de consenso no se base en ninguna relación biológica conocida distinta del primer conjunto de datos y el segundo conjunto de datos;

(2) identificar, a partir de la red de relaciones causales de consenso, una relación causal única en el sistema biológico, donde un gen asociado con la relación causal única se identifica como un modulador del sistema biológico;

identificando así un modulador de un sistema biológico.

38. Un método para identificar un modulador de un sistema biológico, dicho método comprende:

1) proporcionar una red de relaciones causales de consenso a partir de un modelo para el sistema biológico;

2) identificar, a partir de la red de relaciones causales de consenso, una relación causal única en el sistema biológico, donde un gen asociado con la relación causal única se identifica como un modulador del sistema biológico;

identificando así un modulador de un sistema biológico.

39. El método de la reivindicación 38, en donde la red de relaciones causales de consenso se genera entre un primer conjunto de datos y un segundo conjunto de datos obtenidos a partir del modelo para el sistema biológico, en donde el modelo comprende células asociadas con el sistema biológico y en donde el primer conjunto de datos representa niveles de expresión de una pluralidad de genes en las células y el segundo conjunto de datos representa una actividad funcional o una respuesta celular de las células, usando un dispositivo informático programado, en donde la generación de la red de relaciones causales de consenso no se basa en ninguna relación biológica conocida distinta del primer conjunto de datos y el segundo conjunto de datos.

40. Un método para proporcionar un modelo de enfermedad para su uso en un método de plataforma, que comprende:

establecer un modelo de enfermedad para un proceso de enfermedad, usando células relacionadas con la enfermedad, para representar un aspecto característico del proceso de enfermedad, en donde el modelo de enfermedad es útil para generar conjuntos de datos de modelo de enfermedad usados en el método de plataforma; proporcionando así un modelo de enfermedad para uso en un método de plataforma.

41. Un método para obtener un primer conjunto de datos y un segundo conjunto de datos a partir de un modelo de enfermedad para uso en un método de plataforma, que comprende:

(1) obtener un primer conjunto de datos de un modelo de enfermedad para su uso en un método de plataforma, en donde el modelo de enfermedad comprende célula relacionadas con la enfermedad y en donde el primer conjunto de datos representa los niveles de expresión de una pluralidad de genes en las células relacionadas con la enfermedad;

(2) obtener un segundo conjunto de datos del modelo de enfermedad para su uso en un método de plataforma, en donde el segundo conjunto de datos representa una actividad funcional o una respuesta celular de las células relacionadas con la enfermedad; obteniendo así un primer conjunto de datos y un segundo conjunto de datos del modelo de enfermedad;

obteniendo así un primer conjunto de datos y un segundo conjunto de datos a partir de un modelo de enfermedad para uso en un método de plataforma.

42. Un método para identificar un modulador de un proceso de enfermedad, dicho método comprende:

(1) generar una red de relaciones causales de consenso entre un primer conjunto de datos y un segundo conjunto de datos obtenidos a partir de un modelo de enfermedad, en donde el modelo de enfermedad comprende células de la enfermedad y en donde el

primer conjunto de datos representa niveles de expresión de una pluralidad de genes en las células y el segundo conjunto de datos representa una actividad funcional o una respuesta celular de las células, usando un dispositivo informático programado, en donde la generación de la red de relaciones causales de consenso no se basa en ninguna relación biológica conocida distinta del primer conjunto de datos y el segundo conjunto de datos;

(2) identificar, a partir de la red de relaciones causales de consenso, una relación causal única en el proceso de enfermedad, donde un gen asociado con la relación causal única se identifica como un modulador de un proceso de enfermedad;

identificando así un modulador de un proceso de enfermedad.

43. Un método para identificar un modulador de un proceso de enfermedad, dicho método comprende:

1) proporcionar una red de relaciones causales de consenso a partir de un modelo de enfermedad para el proceso de enfermedad;

2) identificar, a partir de la red de relaciones causales de consenso, una relación causal única en el proceso de enfermedad, donde un gen asociado con la relación causal única se identifica como un modulador de un proceso de enfermedad;

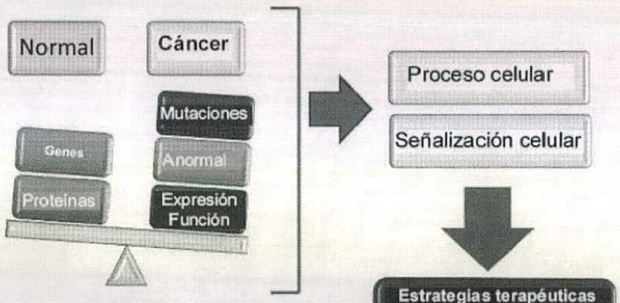
identificando así un modulador de un proceso de enfermedad.

44. El método de la reivindicación 43, en donde la red de relaciones causales de consenso se genera entre un primer conjunto de datos y un segundo conjunto de datos obtenidos a partir de un modelo de enfermedad para el proceso de enfermedad, en donde el modelo de enfermedad comprende células de la enfermedad y en donde el primer conjunto de datos representa niveles de expresión de una pluralidad de genes en las células relacionadas con la enfermedad y el segundo conjunto de datos representa una actividad funcional o una respuesta celular de las células, usando un dispositivo informático programado, en donde la generación de la red de relaciones causales de consenso no se basa en ninguna relación biológica conocida distinta del primer conjunto de datos y el segundo conjunto de datos.

ENSAYOS INTERROGATORIOS BASADOS EN CÉLULAS Y USOS DE LOS MISMOS

RESUMEN

En la presente se describe un Tecnología de Plataforma de descubrimiento para analizar un sistema o proceso biológico (por ejemplo, una enfermedad, tal como cáncer) por medio de un modelo de construcción. En particular, se describe en la presente un método para identificar un modulador de un sistema biológico mediante el establecimiento de un modelo para el sistema biológico, obtener un primer conjunto de datos del modelo para el sistema biológico, obtener un segundo conjunto de datos del modelo para el sistema biológico, y generar una red de relaciones causales de consenso entre los niveles de expresión de la pluralidad de genes y la actividad funcional o la respuesta celular basada únicamente en el primer conjunto de datos y el segundo conjunto de datos, usando un dispositivo informático programado.



- Desregulación de interacciones intra e intercelulares con ambiente extracelular
- Mutación génica/proteína aberrante - influye en el proceso/señalización celular

Figura 1

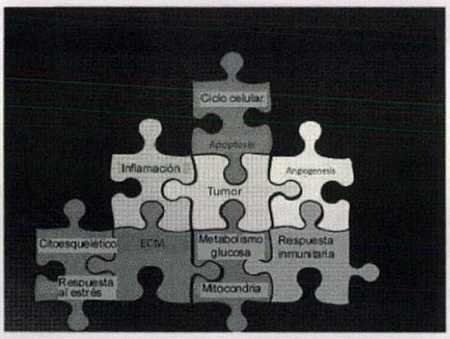

Ciclo celular
Apoptosis
Inflamación
Angiogenesis
Tumor
Citoesquelético
ECM
Metabolismo glucosa
Respuesta inmunitaria
Respuesta al estrés
Mtocondria

Figura 2

362

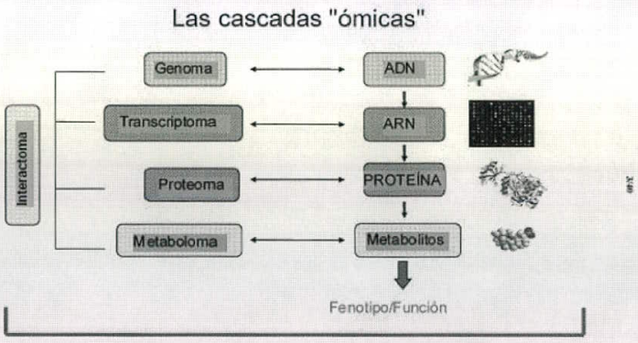

Las cascadas "ómicas"
Genoma
ADN
Transcriptoma
ARN
Proteoma
PROTEÍNA
Metaboloma
Metabolitos
Interactoma
Fenotipo/Función

Figura 3

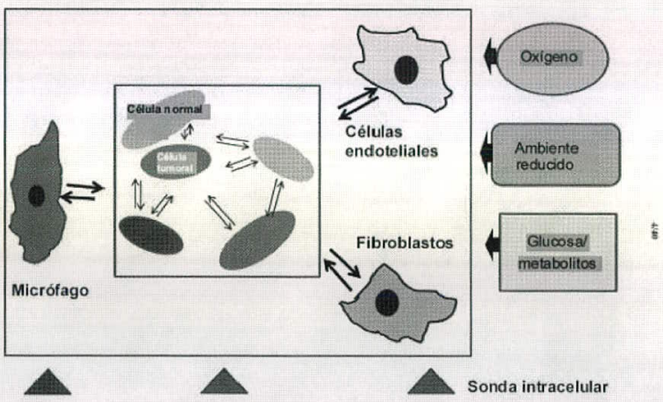

Célula normal
Célula tumoral
Micrófago
Células endoteliales
Fibroblastos
Oxígeno
Ambiente reducido
Glucosa/ metabolitos
Sonda intracelular

Figura 4

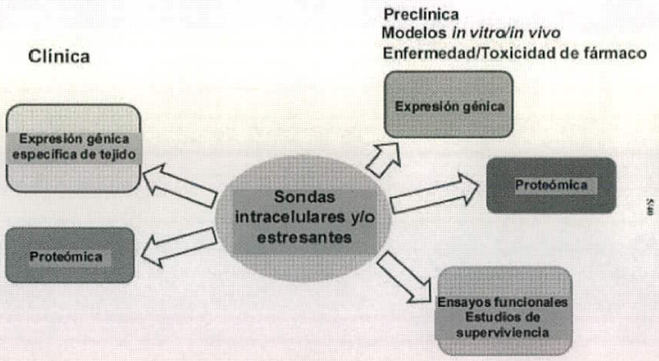

Clínica
Preclínica
Modelos *in vitro/in vivo*
Enfermedad/Toxicidad de fármaco
Expresión génica específica de tejido
Proteómica
Sondas intracelulares y/o estresantes
Expresión génica
Proteómica
Ensayos funcionales
Estudios de supervivencia

Figura 5

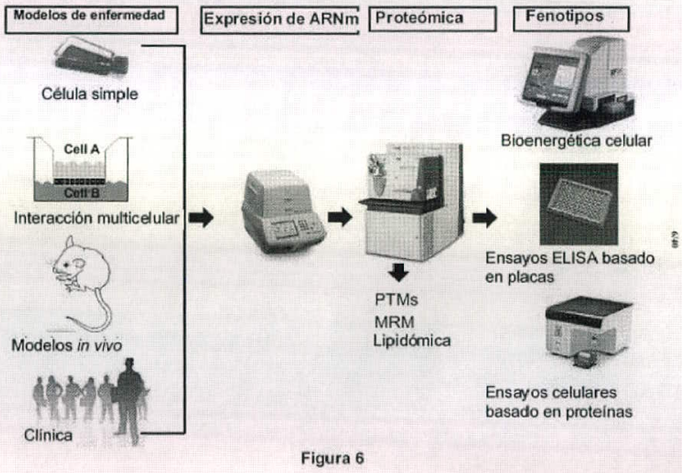

Modelos de enfermedad
Expresión de ARNm
Proteómica
Fenotipos
Célula simple
Cell A
Cell B
Interacción multicelular
Modelos in vivo
Clínica
PTMs
MRM
Lipidómica
Bioenergética celular
Ensayos ELISA basado en placas
Ensayos celulares basado en proteínas

Figura 6

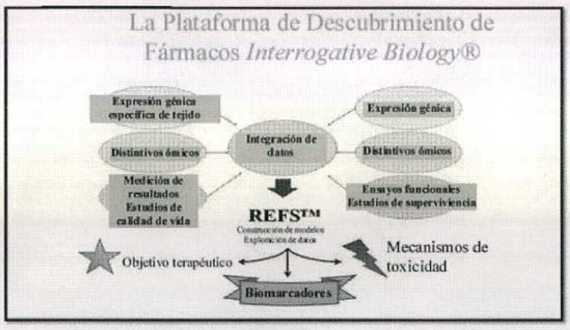

La Plataforma de Descubrimiento de Fármacos *Interrogative Biology*®
Expresión génica específica de tejido
Distintivos ómicos
Medición de resultados Estudios de calidad de vida
Integración de datos
Expresión génica
Distintivos ómicos
Ensayos funcionales Estudios de supervivencia
REFS™
Construcción de modelos
Explotación de datos
Objetivo terapéutico
Mecanismos de toxicidad
Biomarcadores

Figura 7

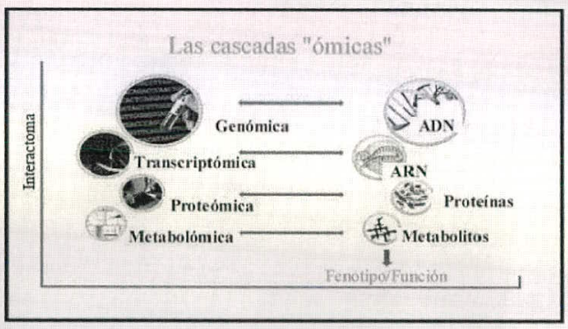

Las cascadas "ómicas"
Interactoma
Genómica
ADN
Transcriptómica
ARN
Proteómica
Proteínas
Metabolómica
Metabolitos
Fenotipo/Función

Figura 8

368

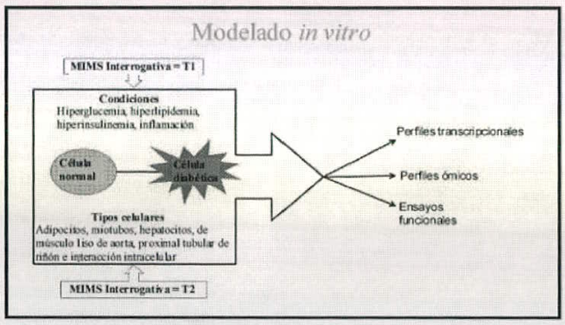

Modelado *in vitro*
MIMS Interrogativa = T1
Condiciones
Hiperglucemia, hiperlipidemia, hiperinsulinemia, inflamación
Célula normal
Célula diabética
Tipos celulares
Adipocitos, miotubos, hepatocitos, de músculo liso de aorta, proximal tubular de riñón e interacción intracelular
MIMS Interrogativa = T2
Perfiles transcripcionales
Perfiles ómicos
Ensayos funcionales

Figura 9

565

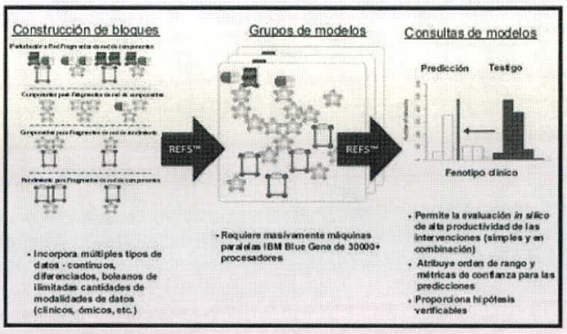

Construcción de bloques
Grupos de modelos
Consultas de modelos
REFS™
REFS™
Predicción
Testigo
Fenotipo clínico
• Incorpora múltiples tipos de datos - continuos, diferenciados, booleanos de ilimitadas cantidades de modalidades de datos (clínicos, ómicos, etc.)
• Requiere masivamente máquinas paralelas IBM Blue Gene de 30000+ procesadores
• Permite la evaluación *in silico* de alta productividad de las intervenciones (simples y en combinación)
• Atribuye orden de rango y métricas de confianza para las predicciones
• Proporciona hipótesis verificables

Figura 10

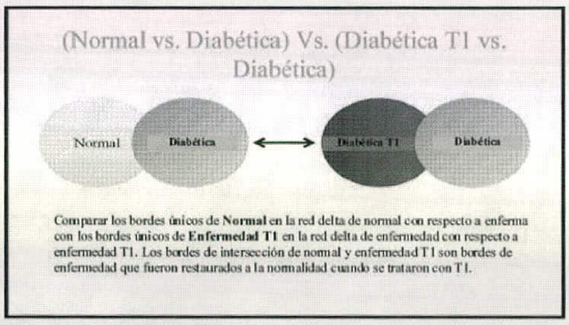

(Normal vs. Diabética) Vs. (Diabética T1 vs. Diabética)
Normal
Diabética
Diabética T1
Diabética
Comparar los bordes únicos de **Normal** en la red delta de normal con respecto a enferma con los bordes únicos de **Enfermedad T1** en la red delta de enfermedad con respecto a enfermedad T1. Los bordes de intersección de normal y enfermedad T1 son bordes de enfermedad que fueron restaurados a la normalidad cuando se trataron con T1.

Figura 11

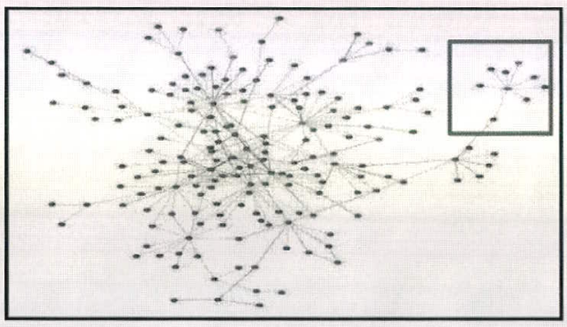


Figura 12

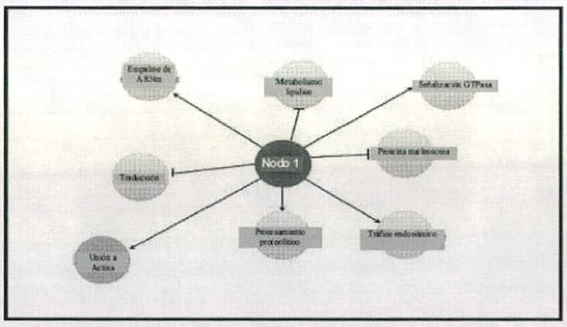

Empalme de ARNm
Metabolismo lipídico
Señalización GTPasa
Nodo 1
Traducción
Proteína mal plegada
Procesamiento proteolítico
Tráfico endosomático
Unión a Actina

Figura 13

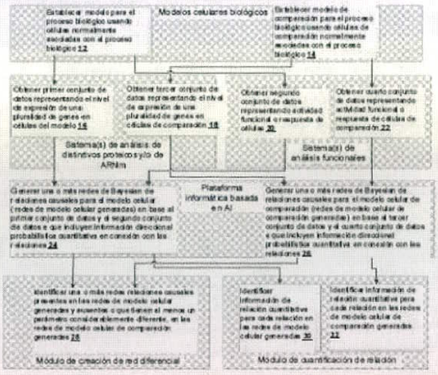


Figura 14

08781

334

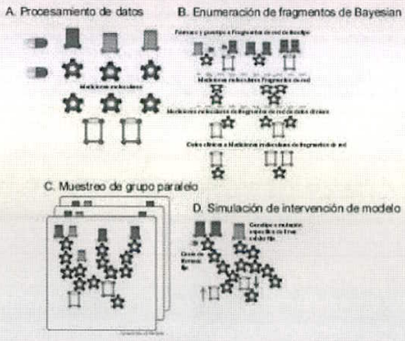

A. Procesamiento de datos
B. Enumeración de fragmentos de Bayesian
C. Muestreo de grupo paralelo
D. Simulación de intervención de modelo

Figura 15

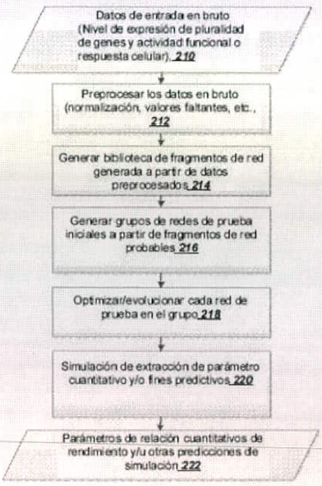


Figura 16

16/49

342

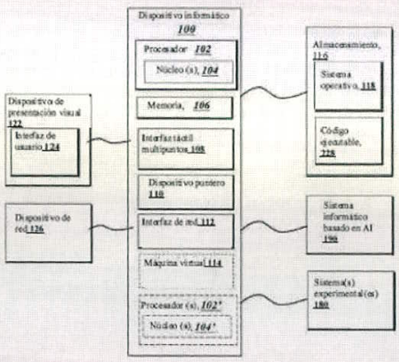


Figura 17

17289

5224

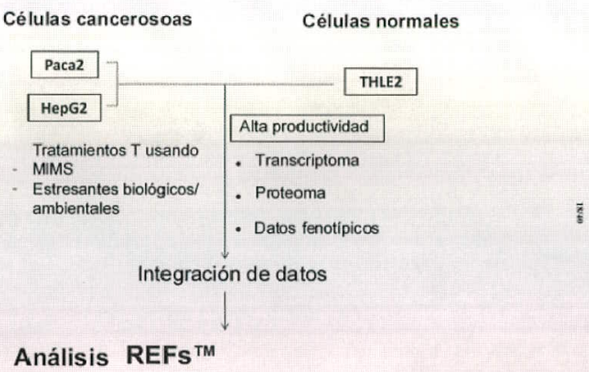

Células cancerosoas
Células normales
Paca2
HepG2
THLE2
Alta productividad
Tratamientos T usando
- MIMS
- Estresantes biológicos/ ambientales
• Transcriptoma
• Proteoma
• Datos fenotípicos
Integración de datos
Análisis REFs™

Figura 18

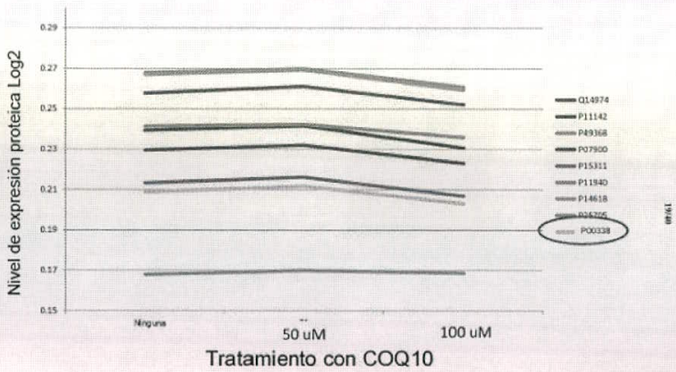

Nivel de expresión proteica Log2
0.29
0.27
0.25
0.23
0.21
0.19
0.17
0.15
Ninguna
50 uM
100 uM
Tratamiento con COQ10
Q14974
P11142
P49368
P07900
P15311
P11940
P14618
P26705
P00338

Figura 19

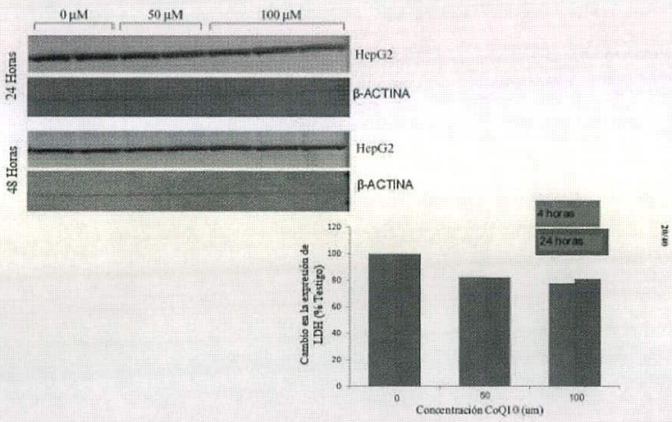

0 µM
50 µM
100 µM
24 Horas
HepG2
β-ACTINA
48 Horas
HepG2
β-ACTINA
4 horas
24 horas
Cambio en la expresión de LDH (% Testigo)
120
100
80
60
40
20
0
0
50
100
Concentración CoQ10 (um)

Figura 20

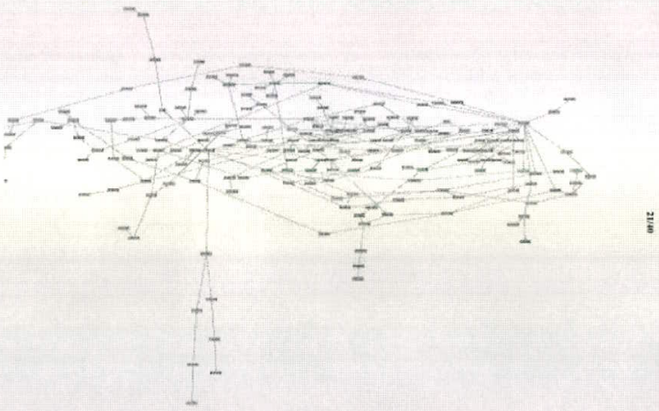


Figura 21

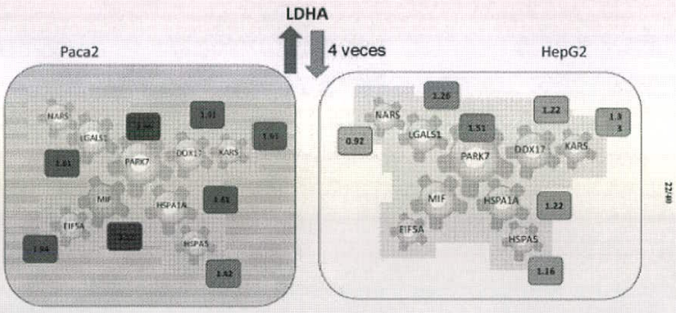

LDHA
4 veces
Paca2
HepG2
NARS
LGALS1
PARK7
DDX17
KARS
MIF
HSPA1A
EIF5A
HSPA5
1.61
1.91
1.85
1.94
1.42
3.26
1.22
1.3
1
0.92
1.51
1.22
1.16

Figura 22

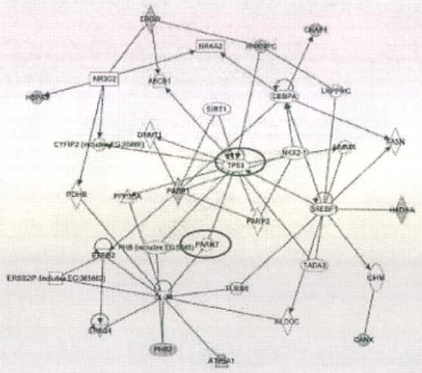

EGOR
CKAP4
NR3C2
ABCB1
CEBPA
SIRT1
DRAM2
TP53
NCOA1
ASN
PDHB
PARP1
SREBF1
PARK2
TADA3
CHN
ATP6AP1

Figura 23

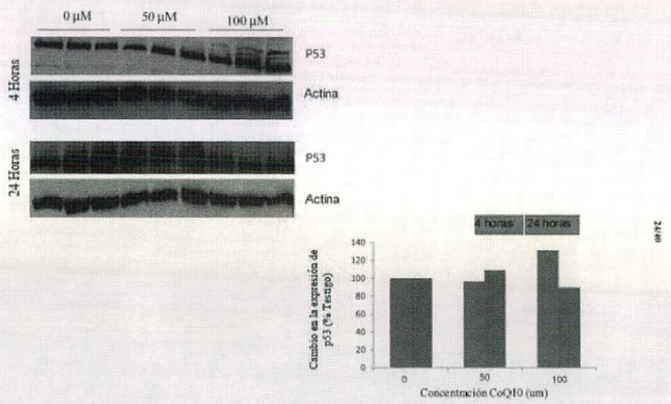

0 µM
50 µM
100 µM
4 Horas
P53
Actina
24 Horas
P53
Actina
4 horas
24 horas
Cambio en la expresión de p53 (% Testigo)
140
120
100
80
60
40
20
0
0
50
100
Concentración CoQ10 (um)

Figura 24

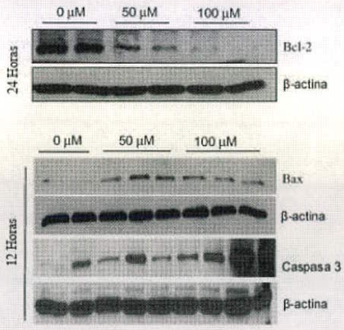

0 µM
50 µM
100 µM
24 Horas
Bcl-2
β-actina
0 µM
50 µM
100 µM
12 Horas
Bax
β-actina
Caspasa 3
β-actina

Figura 25

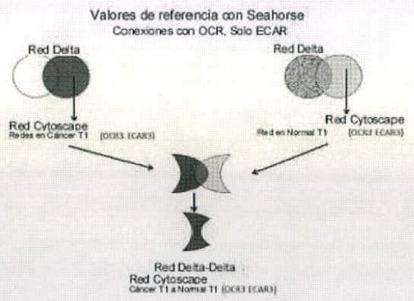

Valores de referencia con Seahorse
Conexiones con OCR, Solo ECAR
Red Delta
Red Cytoscape
Redes en Cáncer T1 (OCR1 ECAR1)
Red Delta
Red Cytoscape
Red en Normal T1 (OCR2 ECAR2)
Red Delta-Delta
Red Cytoscape
Cáncer T1 a Normal T1 (OCR3 ECAR3)

Figura 26

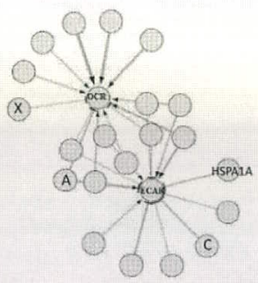


Figura 27

27249

585

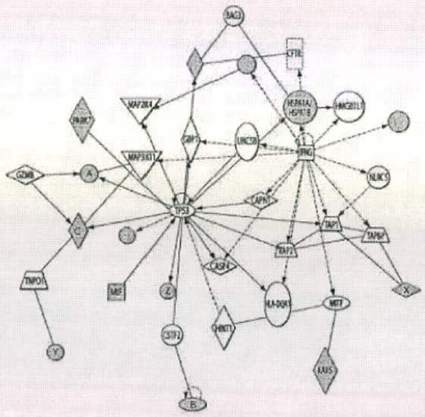


Figura 28

2KNB

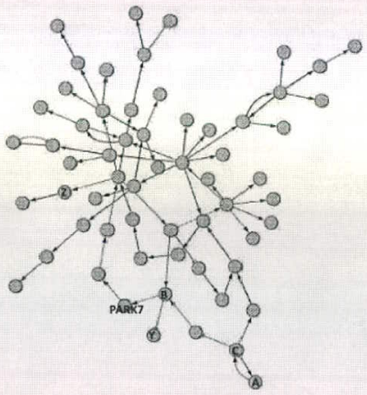


Figura 29

Redes Delta-Delta

DD_NG.vs.HG.vs.HGT1.cys

[NG vs. HG] VS. [HG vs. HGT1]

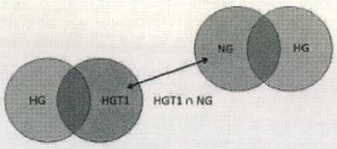


Figura 30

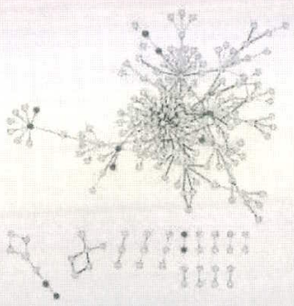


Figura 31

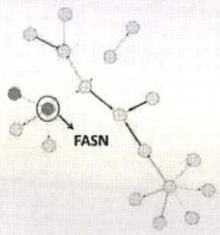


Figura 32


Ácido graso libre
Ingesta
Mitocondria
β-oxidación
Acyl CoA
β-oxidación
Acyl CoA
Tratamiento con fármaco de Diabetes
Pérdida de función de membrana
Remodelación de membrana
Utilización de síntesis ATP

Figura 33


Normo/Hiperglucemia
36 Hrs
X hrs
Creación de perfil transcripcional
Proteómica/lipidómica
Ensayos funcionales / Transferencia Western
Cardiomiocitos humanos
Ácido linoleico
Ácido oleico
L Carnitina
Tratamiento con fármaco

Figura 34

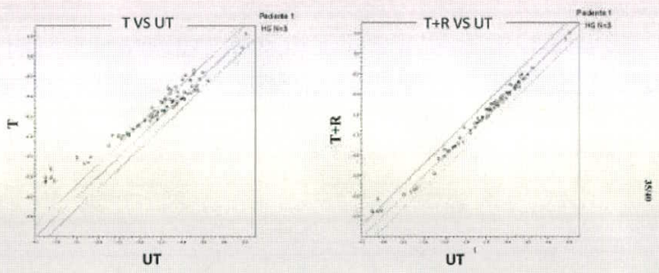


Figura 35

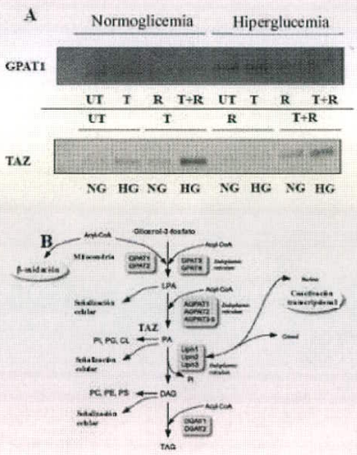

A
Normoglicemia
Hiperglucemia
GPAT1
UT T R T+R UT T R T+R
UT T R T+R
TAZ
NG HG NG HG NG HG NG HG
B
Acyl-CoA
Glicerol-3-fosfato
Mitocondria
β-oxidación
GPAT1
GPAT2
GPAT3
GPAT4
LPA
Acyl-CoA
Señalización celular
AGPAT1
AGPAT2
AGPAT3-5
TAZ
PI, PG, CL
PA
Señalización celular
Lipin1
Lipin2
Lipin3
PI
Retina
Coactivador transcripcional
Grasa
PC, PE, PS
DAG
Señalización celular
Acyl-CoA
DGAT1
DGAT2
TAG

Figura 36

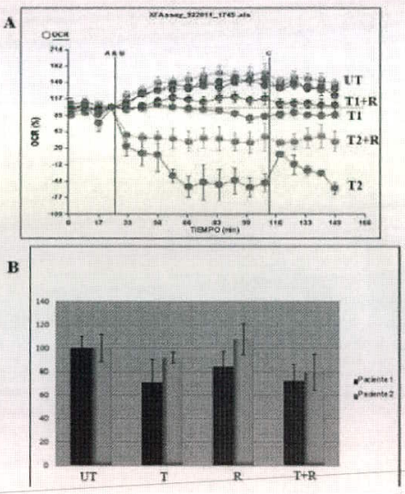

A
OCR
OCR (%)
TIEMPO (min)
UT
T1+R
T1
T2+R
T2
B
Paciente 1
Paciente 2
UT
T
R
T+R

Figura 37

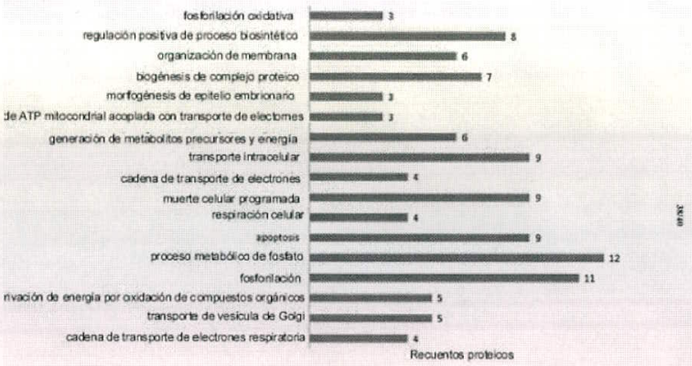

fosforilación oxidativa
3
regulación positiva de proceso biosintético
8
organización de membrana
6
biogénesis de complejo proteico
7
morfogénesis de epitelio embrionario
3
de ATP mitocondrial acoplada con transporte de electrones
3
generación de metabolitos precursores y energía
6
transporte intracelular
9
cadena de transporte de electrones
4
muerte celular programada
9
respiración celular
4
apoptosis
9
proceso metabólico de fosfato
12
fosforilación
11
ivación de energía por oxidación de compuestos orgánicos
5
transporte de vesícula de Golgi
5
cadena de transporte de electrones respiratoria
4
Recuentos proteicos

Figura 38

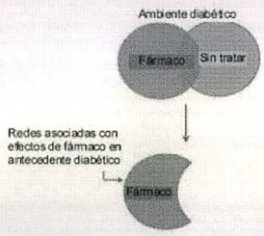

Ambiente diabético
Fármaco
Sin tratar
Redes asociadas con efectos de fármaco en antecedente diabético
Fármaco

Figura 39

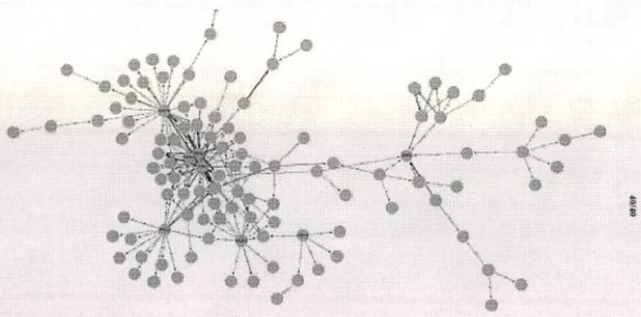


Figura 40

401

PATENT COOPERATION TREATY PCT/US2012/027615

ADVANCE E-MAIL

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF THE RECORDING OF A CHANGE

(PCT Rule 92*bis*.1 and Administrative Instructions, Section 422)

To:

HANLEY, Elizabeth, A.
McCarter & English, LLP
265 Franklin Street
Boston, MA 02110
ETATS-UNIS D'AMERIQUE

Date of mailing *(day/month/year)*
02 October 2012 (02.10.2012)

Applicant's or agent's file reference
119992-04420

IMPORTANT NOTIFICATION

International application No.
PCT/US2012/027615

International filing date *(day/month/year)*
02 March 2012 (02.03.2012)

1. The following indications appeared on record concerning:

☒ the applicant ☐ the inventor ☐ the agent ☐ the common representative

Name and Address
BERG BIOSYSTEMS, LLC
1845 Elm Hill Pike
Nashville, TN 37210
United States of America

State of Nationality: US
State of Residence: US
Telephone No.:
Facsimile No.:
E-mail address:

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

☐ the person ☒ the name ☐ the address ☐ the nationality ☐ the residence

Name and Address
BERG PHARMA LLC
1845 Elm Hill Pike
Nashville, TN 37210
United States of America

State of Nationality: US
State of Residence: US
Telephone No.:
Facsimile No.:
E-mail address:
☐ Notifications by e-mail authorized

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

☒ the receiving Office
☐ the International Searching Authority
☐ the Authority(ies) specified for supplementary search
☐ the International Preliminary Examining Authority
☒ the designated Offices concerned
☐ the elected Offices concerned
☐ other:

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No. +41 22 338 70 80

Authorized officer
Safi Horiya
e-mail pt03.pct@wipo.int
Telephone No. +41 22 338 74 03

402

PATENT COOPERATION TREATY PCT/US2012/027615

ADVANCE E-MAIL

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF THE RECORDING OF A CHANGE

(PCT Rule 92*bis*.1 and Administrative Instructions, Section 422)

To:

HANLEY, Elizabeth, A.
McCarter & English, LLP
265 Franklin Street
Boston, MA 02110
ETATS-UNIS D'AMERIQUE

Date of mailing *(day/month/year)*
27 August 2013 (27.08.2013)

Applicant's or agent's file reference
119992-04420

IMPORTANT NOTIFICATION

International application No.
PCT/US2012/027615

International filing date *(day/month/year)*
02 March 2012 (02.03.2012)

1. The following indications appeared on record concerning:

☒ the applicant ☐ the inventor ☐ the agent ☐ the common representative

Name and Address
BERG PHARMA LLC
1845 Elm Hill Pike
Nashville, TN 37210
United States of America

State of Nationality: US
State of Residence: US
Telephone No.:
Facsimile No.:
E-mail address:

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

☐ the person ☒ the name ☐ the address ☐ the nationality ☐ the residence

Name and Address
BERG LLC
1845 Elm Hill Pike
Nashville, TN 37210
United States of America

State of Nationality: US
State of Residence: US
Telephone No.:
Facsimile No.:
E-mail address:
☐ Notifications by e-mail authorized

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

☒ the receiving Office
☐ the International Searching Authority
☐ the Authority(ies) specified for supplementary search
☐ the International Preliminary Examining Authority
☒ the designated Offices concerned
☐ the elected Offices concerned
☐ other:

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No. +41 22 338 70 80

Authorized officer
Campin Helene
e-mail pt03.pct@wipo.int
Telephone No. +41 22 338 74 03

403

PATENT COOPERATION TREATY PCT/US2012/027615

ADVANCE E-MAIL

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF THE RECORDING OF A CHANGE

(PCT Rule 92*bis*.1 and Administrative Instructions, Section 422)

To:

HANLEY, Elizabeth, A.
McCarter & English, LLP
265 Franklin Street
Boston, MA 02110
ETATS-UNIS D'AMERIQUE

Date of mailing *(day/month/year)*
27 August 2013 (27.08.2013)

Applicant's or agent's file reference
119992-04420

IMPORTANT NOTIFICATION

International application No.
PCT/US2012/027615

International filing date *(day/month/year)*
02 March 2012 (02.03.2012)

1. The following indications appeared on record concerning:

☒ the applicant ☒ the inventor ☐ the agent ☐ the common representative

Name and Address
NARAIN, Niven, Rajin
350 Third Street
Suite 2005
Cambridge, MA 02142
United States of America

State of Nationality: GY
State of Residence: US
Telephone No.:
Facsimile No.:
E-mail address:

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

☐ the person ☐ the name ☒ the address ☒ the nationality ☐ the residence

Name and Address
NARAIN, Niven, Rajin
73 Fresh Pond Parkway
Cambridge, MA 02138
United States of America

State of Nationality: US
State of Residence: US
Telephone No.:
Facsimile No.:
E-mail address:
☐ Notifications by e-mail authorized

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

☒ the receiving Office
☐ the International Searching Authority
☐ the Authority(ies) specified for supplementary search
☐ the International Preliminary Examining Authority
☒ the designated Offices concerned
☐ the elected Offices concerned
☐ other:

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No. +41 22 338 70 80

Authorized officer
Campin Helene
e-mail pt03.pct@wipo.int
Telephone No. +41 22 338 74 03

(119992-04429)

IN THE PATENT OFFICE OF COLOMBIA

Assignee(s): **Berg Pharma LLC**

Assignor(s): **Berg Biosystems, LLC**

Serial Nos. &
Filing Dates: **USSN 61/448587 filed 02 March 2011**
USSN 61/593848 filed 01 February 2012

I, Jill A. Mello, Ph.D., J.D., attorney for the applicant, hereby attest that the attached document is a true and exact copy of the Assignment executed by **Berg Biosystems, LLC**, a company organized and existing under the laws of Delaware. The Assignment is from Berg Biosystems, LLC to **Berg Pharma LLC**, a company organized and existing under the laws of Delaware.

Jill A. Mello, Ph.D., J.D.
Registration No. 55,450
Attorney for Applicant

United States of America)
Commonwealth of Massachusetts) ss.:
County of Suffolk)

On the 20th day of August, 2013 before me, the undersigned notary public, personally appeared Jill A. Mello (name of document signer), proved to me through satisfactory evidence of identification, which was that she is known to me personally, to be the person whose name is signed on the preceding or attached document in my presence.

(official signature and seal of notary)
Jessica Dale Bell
My commission expires March 16, 2018.

JESSICA DALE BELL
Notary Public
Massachusetts
Commission Expires Mar 16, 2018

405



UNITED STATES PATENT AND TRADEMARK OFFICE

UNDER SECRETARY OF COMMERCE FOR INTELLECTUAL PROPERTY AND
DIRECTOR OF THE UNITED STATES PATENT AND TRADEMARK OFFICE

JULY 20, 2012

PTAS

JILL A. MELLO, PH.D.
265 FRANKLIN STREET
MCCARTER & ENGLISH, LLP
BOSTON, MA 02110

501995195

UNITED STATES PATENT AND TRADEMARK OFFICE
NOTICE OF RECORDATION OF ASSIGNMENT DOCUMENT

THE ENCLOSED DOCUMENT HAS BEEN RECORDED BY THE ASSIGNMENT RECORDATION BRANCH OF THE U.S. PATENT AND TRADEMARK OFFICE. A COMPLETE COPY IS AVAILABLE AT THE ASSIGNMENT SEARCH ROOM ON THE REEL AND FRAME NUMBER REFERENCED BELOW.

PLEASE REVIEW ALL INFORMATION CONTAINED ON THIS NOTICE. THE INFORMATION CONTAINED ON THIS RECORDATION NOTICE REFLECTS THE DATA PRESENT IN THE PATENT AND TRADEMARK ASSIGNMENT SYSTEM. IF YOU SHOULD FIND ANY ERRORS OR HAVE QUESTIONS CONCERNING THIS NOTICE, YOU MAY CONTACT THE ASSIGNMENT RECORDATION BRANCH AT 571-272-3350. PLEASE SEND REQUEST FOR CORRECTION TO: U.S. PATENT AND TRADEMARK OFFICE, MAIL STOP: ASSIGNMENT RECORDATION BRANCH, P.O. BOX 1450, ALEXANDRIA, VA 22313.

RECORDATION DATE: 07/19/2012 REEL/FRAME: 028590/0308
NUMBER OF PAGES: 5

BRIEF: ASSIGNMENT OF ASSIGNORS INTEREST (SEE DOCUMENT FOR DETAILS).

DOCKET NUMBER: 119992-00001

ASSIGNOR:
BERG BIOSYSTEMS, LLC DOC DATE: 06/11/2012

ASSIGNEE:
BERG PHARMA LLC
1845 ELM HILL PIKE
NASHVILLE, TENNESSEE 37210

APPLICATION NUMBER: 12695113 FILING DATE: 01/27/2010
PATENT NUMBER: ISSUE DATE:
TITLE: VITAMIN D3 AND ANALOGS THEREOF FOR ALLEVIATING SIDE EFFECTS ASSOCIATED WITH CHEMOTHERAPY

APPLICATION NUMBER: 12777902 FILING DATE: 05/11/2010
PATENT NUMBER: ISSUE DATE:
TITLE: METHODS FOR TREATMENT OF ONCOLOGICAL DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS

APPLICATION NUMBER: 12778010 FILING DATE: 05/11/2010
PATENT NUMBER: ISSUE DATE:
TITLE: METHODS FOR THE DIAGNOSIS OF METABOLIC DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS

APPLICATION NUMBER: 12778029 FILING DATE: 05/11/2010
PATENT NUMBER: ISSUE DATE:
TITLE: METHODS FOR THE DIAGNOSIS OF ONCOLOGICAL DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS

APPLICATION NUMBER: 12778054 FILING DATE: 05/11/2010
PATENT NUMBER: ISSUE DATE:
TITLE: METHODS FOR TREATMENT OF METABOLIC DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS

APPLICATION NUMBER: 12778094 FILING DATE: 05/11/2010
PATENT NUMBER: ISSUE DATE:
TITLE: METHODS FOR TREATMENT OF ONCOLOGICAL DISORDERS USING AN EPIMETABOLIC SHIFTER (COENZYME Q10)

APPLICATION NUMBER: 12853431 FILING DATE: 08/10/2010
PATENT NUMBER: ISSUE DATE:
TITLE: VITAMIN D3 AND ANALOGS THEREOF FOR TREATING ALOPECIA

APPLICATION NUMBER: 12864185 FILING DATE: 10/18/2010
PATENT NUMBER: ISSUE DATE:
TITLE: ASSAY SYSTEM FOR THE ASSESSMENT OF ONCOGENICITY, TUMOR PROGRESSION, AND TREATMENT EFFICACY

APPLICATION NUMBER: 61023570 FILING DATE: 01/25/2008
PATENT NUMBER: ISSUE DATE:
TITLE: ASSAY SYSTEM FOR THE ASSESSMENT OF ONCOGENICITY, TUMOR PROGRESSION, AND TREATMENT EFFICACY

APPLICATION NUMBER: 61147549 FILING DATE: 01/27/2009
PATENT NUMBER: ISSUE DATE:
TITLE: VITAMIN D3 AND ANALOGS THEREOF FOR ALLEVIATING SIDE EFFECTS ASSOCIATED WITH CHEMOTHERAPY

APPLICATION NUMBER: 61177241 FILING DATE: 05/11/2009
PATENT NUMBER: ISSUE DATE:
TITLE: METHODS FOR TREATMENT OF ONCOLOGICAL DISORDERS USING AN EPIMETABOLIC SHIFTER (COENZYME Q10)

APPLICATION NUMBER: 61177243 FILING DATE: 05/11/2009
PATENT NUMBER: ISSUE DATE:
TITLE: METHODS FOR TREATMENT OF ONCOLOGICAL DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS

APPLICATION NUMBER: 61177244 FILING DATE: 05/11/2009
PATENT NUMBER: ISSUE DATE:
TITLE: METHODS FOR THE DIAGNOSIS OF ONCOLOGICAL DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS

APPLICATION NUMBER: 61177245 FILING DATE: 05/11/2009
PATENT NUMBER: ISSUE DATE:
TITLE: METHODS FOR TREATMENT OF METABOLIC DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS

APPLICATION NUMBER: 61177246 FILING DATE: 05/11/2009
PATENT NUMBER: ISSUE DATE:
TITLE: METHODS FOR THE DIAGNOSIS OF METABOLIC DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS

APPLICATION NUMBER: 61234178 FILING DATE: 08/14/2009
PATENT NUMBER: ISSUE DATE:
TITLE: VITAMIN D3 AND ANALOGS THEREOF FOR TREATING ALOPECIA

APPLICATION NUMBER: 61239003 FILING DATE: 09/01/2009
PATENT NUMBER: ISSUE DATE:
TITLE: VITAMIN D3 AND ANALOGS THEREOF FOR ALLEVIATING SIDE EFFECTS ASSOCIATED WITH CHEMOTHERAPY

APPLICATION NUMBER: 61363562 FILING DATE: 07/12/2010
PATENT NUMBER: ISSUE DATE:
TITLE: METHODS FOR TREATMENT OF ENDOMETRIOSIS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS

APPLICATION NUMBER: 61448587 FILING DATE: 03/02/2011
PATENT NUMBER: ISSUE DATE:
TITLE: INTERROGATORY CELL-BASED ASSAYS AND USES THEREOF

APPLICATION NUMBER: 61474708 FILING DATE: 04/12/2011
PATENT NUMBER: ISSUE DATE:
TITLE: COMPOSITIONS AND METHODS FOR TREATMENT OF PERVASIVE DEVELOPMENTAL DISORDER

APPLICATION NUMBER: 61491132 FILING DATE: 05/27/2011
PATENT NUMBER: ISSUE DATE:
TITLE: MITOCHONDRIAL METABOLIC MODULATORS AND USES THEREOF

APPLICATION NUMBER: 61506221 FILING DATE: 07/11/2011
PATENT NUMBER: ISSUE DATE:
TITLE: METHODS FOR TREATMENT OF ENDOMETRIOSIS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS

APPLICATION NUMBER: 61511487 FILING DATE: 07/25/2011
PATENT NUMBER: ISSUE DATE:
TITLE: USE OF MALATE DEHYDROGENASE 2 (MDH2) IN THE DIAGNOSIS AND TREATMENT OF DISEASE STATES

APPLICATION NUMBER: 61514388 FILING DATE: 08/02/2011
PATENT NUMBER: ISSUE DATE:
TITLE: METHODS AND KITS FOR DETECTION OF COENZYME Q10

APPLICATION NUMBER: 61593848 FILING DATE: 02/01/2012
PATENT NUMBER: ISSUE DATE:
TITLE: INTERROGATORY CELL-BASED ASSAYS AND USES THEREOF

ASSIGNMENT RECORDATION BRANCH
PUBLIC RECORDS DIVISION

ASSIGNMENT

This Assignment (this "***Assignment***") is executed as of June 11, 2012 between **Berg Biosystems, LLC**, a corporation duly organized under and pursuant to the laws of the State of Delaware, and having its principal place of business at 1845 Elm Hill Pike, Nashville, Tennessee 37210 ("***Assignor***"), and **Berg Pharma LLC**, a corporation duly organized under and pursuant to the laws of the State of Delaware, and having its principal place of business at 1845 Elm Hill Pike, Nashville, Tennessee 37210 ("***Assignee***").

Background

Assignor owns those certain patents and patent applications listed in Schedule A, and the inventions and new and useful improvements described and claimed therein and all know-how associated therewith existing as of the date set forth above (all of the foregoing collectively, the "***Patents***"). Assignee is desirous of acquiring the entire right, title and interest of Assignor in and to the Patents.

Terms

In consideration of **One Dollar ($1.00)**, and for other good and valuable consideration, the receipt, adequacy and sufficiency of which is hereby acknowledged, Assignor does hereby assign, transfer and convey, unto Assignee, and Assignee's successors and assigns, Assignor's entire right, title, and interest in and to the Patents, and in and to all Letters Patent that may or shall be granted on patent applications or inventions within the Patents, or any parts thereof, or any divisional, continuation, continuation-in-part, substitution, renewal, reissue or other applications based in whole or in part thereon, and all reissues, re-examinations and extensions thereof, in each case in any country and including any foreign counterpart thereof or thereto, together with the right to file applications and the right to claim for the same the priority rights derived from the patent applications within the Patents; and further together with all claims for damages by reason of past, present or future infringement, with the right to sue for, and collect the same for the use of Assignee, Assignee's successors and assigns and all proceeds of the foregoing, including, without limitation, any claim by Assignor against third parties for, or any profits arising from, past, present, or future infringement of the Patents, the same to be held and enjoyed by the Assignee and Assignee's successors and assigns as fully as the same would have been held and enjoyed by the Assignor had this Assignment not been made.

Assignor agrees, for itself and Assignor's executors, administrators, and successors in interest, at Assignee's expense, to execute all applications, amended specifications, deeds or other instruments, and to do all acts necessary or proper to secure the grant to Assignee of Letters Patent on the patent applications within the Patents, with specifications and claims in such form as shall be approved by the counsel of Assignee, and to vest and confirm in Assignee and Assignee's successors and assigns the legal title to all of the Patents.

Assignor hereby authorizes and requests the Commissioner of Patents and Trademarks of the United States, and any official of any country or countries foreign to the United States whose duty it is to issue patents or other evidence or forms of industrial property protection on applications as aforesaid, to issue any such patents on any such applications within the Patents to the Assignee, and Assignee's successors, legal representatives and assigns, as fully and entirely as the same would have been held by Assignor had this Assignment not been made.

In Witness Whereof, this Assignment is duly executed as of the date first set forth above.

BERG BIOSYSTEMS, LLC
ASSIGNOR:

By: RANGAPRASAD SARANGARAJAN
Title: V.P - R&D.

Witness:

Date: 06/11/2012

BERG PHARMA LLC
ASSIGNEE:

By: NIVEN R. NARAIN
Title: PRESIDENT

Witness:

Date: 06/11/2012

SCHEDULE A

Confidential – Attorney-Client Privileged

Serial No.	Filing Date	Atty. Docket No.	Title
61/177241	11-May-2009	119992-00601	METHODS FOR TREATMENT OF ONCOLOGICAL DISORDERS USING AN EPIMETABOLIC SHIFTER (COENZYME Q10)
12/778094	11-May-2010	119992-00602	METHODS FOR TREATMENT OF ONCOLOGICAL DISORDERS USING AN EPIMETABOLIC SHIFTER (COENZYME Q10)
61/177243	11-May-2009	119992-00701	METHODS FOR TREATMENT OF ONCOLOGICAL DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS
12/777902	11-May-2010	119992-00702	METHODS FOR TREATMENT OF ONCOLOGICAL DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS
61/177244	11-May-2009	119992-00801	METHODS FOR THE DIAGNOSIS OF ONCOLOGICAL DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS
12/778029	11-May-2010	119992-00802	METHODS FOR THE DIAGNOSIS OF ONCOLOGICAL DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS
61/177245	11-May-2009	119992-00901	METHODS FOR TREATMENT OF METABOLIC DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS
12/778054	11-May-2010	119992-00902	METHODS FOR TREATMENT OF METABOLIC DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS
61/491132	27-May-2011	119992-00903	MITOCHONDRIAL METABOLIC MODULATORS AND USES THEREOF
61/177246	11-May-2009	119992-01001	METHODS FOR THE DIAGNOSIS OF METABOLIC DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS
12/778010	11-May-2010	119992-01002	METHODS FOR THE DIAGNOSIS OF METABOLIC DISORDERS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS
61/147549	27-Jan-2009	119992-01501	VITAMIN D3 AND ANALOGS THEREOF FOR ALLEVIATING SIDE EFFECTS ASSOCIATED WITH CHEMOTHERAPY
61/239003	01-Sep-2009	119992-01502	VITAMIN D3 AND ANALOGS THEREOF FOR ALLEVIATING SIDE EFFECTS ASSOCIATED WITH CHEMOTHERAPY
12/695113	27-Jan-2010	119992-01503	VITAMIN D3 AND ANALOGS THEREOF FOR ALLEVIATING SIDE EFFECTS ASSOCIATED WITH CHEMOTHERAPY
61/234178	14-Aug-2009	119992-01701	VITAMIN D3 AND ANALOGS THEREOF FOR TREATING ALOPECIA
12/853431	10-Aug-2010	119992-01703	VITAMIN D3 AND ANALOGS THEREOF FOR TREATING ALOPECIA
61/023570	25-Jan-2008	119992-03201	ASSAY SYSTEM FOR THE ASSESSMENT OF ONCOGENICITY, TUMOR PROGRESSION, AND TREATMENT EFFICACY
12/864185	18-Oct-2010	119992-03202	ASSAY SYSTEM FOR THE ASSESSMENT OF ONCOGENICITY, TUMOR PROGRESSION, AND TREATMENT EFFICACY
61/363562	12-Jul-2010	119992-04101	METHODS FOR TREATMENT OF ENDOMETRIOSIS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS
61/506221	11-Jul-2011	119992-04102	METHODS FOR TREATMENT OF ENDOMETRIOSIS USING EPIMETABOLIC SHIFTERS, MULTIDIMENSIONAL INTRACELLULAR MOLECULES, OR ENVIRONMENTAL INFLUENCERS
61/514388	02-Aug-2011	119992-04201	METHODS AND KITS FOR DETECTION OF COENZYME Q10
61/474708	12-Apr-2011	119992-04301	COMPOSITIONS AND METHODS FOR TREATMENT OF PERVASIVE DEVELOPMENTAL DISORDER
61/448587	02-Mar-2011	119992-04403	INTERROGATORY CELL-BASED ASSAYS AND USES THEREOF
61/593848	01-Feb-2012	119992-04404	INTERROGATORY CELL-BASED ASSAYS AND USES THEREOF
61/511487	25-Jul-2011	119992-05101	USE OF MALATE DEHYDROGENASE 2 (MDH2) IN THE DIAGNOSIS AND TREATMENT OF DISEASE STATES

412

(119992-04429)

EN LA OFICINA DE PATENTES DE COLOMBIA

Cesionario(s): **Berg Pharma, LLC**

Cedente(s): **Berg Biosystems, LLC**

Nos. de serie y
Fechas de presentación: **USSN 61/448587 presentada 02 de marzo 2011**
USSN 61/593848 presentada 01 de febrero 2012

Yo, Jill A. Mello, Ph.D., J.D., abogada de los solicitantes, por la presente declaro que el documento adjunto es una copia fiel y exacta de la asignación ejecutada por **Berg Biosystems, LLC**, una compañía constituida y regida bajo las leyes del estado de Delaware. La asignación es de Berg Biosystems, a **Berg Pharma, LLC**, sociedad constituida y existente bajo las leyes de Delaware.

/ Firmado /
Jill A. Mello, Ph.D., J.D.
No. de Registro Fiscal 55.450
Abogada por los Solicitantes

Estados Unidos de América)
Estado de Massachusetts) ss.:
Condado de Suffolk)

El día 20 de agosto de 2013 ante mí, el notario que suscribe, compareció personalmente Jill A. Mello (nombre del firmante de documentos), que demostró con evidencia satisfactoria de identificación, en que ella me es conocida personalmente, ser la persona cuyo nombre se halla firmado en el documento anterior o adjunto en mi presencia.

/Firmado/ (firma y sello oficiales de notario)
Jessica Dale Bell
Mi comisión expira el 16 de marzo 2018.

[Sello que lee:]
JESSICA DALE BELL
Notario público
Massachusetts
Comisión Expira 16 de marzo 2018

OFICINA DE PATENTES Y MARCAS DE LOS ESTADOS UNIDOS
SUBSECRETARIO DE COMERCIO PARA LA PROPIEDAD INTELECTUAL Y
DIRECTOR DE LA OFICINA DE PATENTES Y MARCAS DE LOS ESTADOS UNIDOS

20 DE JULIO DE 2012

PTAS

JILL A. MELLO, PH.D.
265 FRANKLIN STREET
MCCARTER & ENGLISH, LLP
BOSTON, MA 02110

501995195

OFICINA DE PATENTES Y MARCAS DE LOS ESTADOS UNIDOS
AVISO DE INSCRIPCIÓN DE DOCUMENTO DE ASIGNACIÓN

EL DOCUMENTO ADJUNTO HA SIDO REGISTRADO POR LA SUCUURSAL DE INSCRIPCIÓN DE ASIGNACIONES DE LA OFICINA DE PATENTES Y MARCAS DE LOS ESTADOS UNIDOS. UNA COPIA COMPLETA ESTÁ DISPONIBLE EN LA SALA DE ASIGNACIÓN BAJO EL CARRETE Y NÚMERO DE MARCO LISTADO A CONTINUACION.

POR FAVOR LEA TODA LA INFORMACIÓN CONTENIDA EN ESTE AVISO. LA INFORMACIÓN CONTENIDA EN ESTE AVISO DE INSCRIPCIÓN REFLEJA LA PRESENCIA DE DATOS EN EL SISTEMA DE ASIGNACIONES DE PATENTES Y MARCAS. SI ENCUENTRA ERRORES O TIENE PREGUNTAS SOBRE ESTE AVISO, PUEDE CONTACTAR A LA SUCURSAL DE INSCRIPCIÓN DE ASIGNACIONES, 571-272-3350. POR FAVOR ENVIAR SOLICITUD DE CORRECCIÓN A: OFICINA DE PATENTES Y MARCAS DE LOS ESTADOS UNIDOS, PARADA DE CORREO: SUCURSAL DE INSCRIPCIÓN DE ASIGNACIONES, APARTADO POSTAL 1450, ALEXANDRIA, VIRGINIA 22313.

FECHA DE INSCRIPCIÓN: 19/07/2012 CARRETE/MARCO: 028590/0308
NÚMERO DE PÁGINAS: 5

MEMORIAL: ASIGNACIÓN INTERÉS DE CEDENTES (VÉASE EL DOCUMENTO PARA MÁS DETALLES).

NÚMERO DE EXPEDIENTE: 119992-00001

CEDENTE:
BERG BIOSYSTEMS, LLC DOC FECHA: 11/06/2012

CESIONARIO:
BERG PHARMA, LLC
1845 ELM HILL PIKE
NASHVILLE, TENNESSEE 37210

NÚMERO DE SOLICITUD: 12695113 FECHA DE RADICACION: 27/01/2010
NÚMERO DE PATENTE: FECHA DE EMISION:
TÍTULO: LA VITAMINA D3 Y SUS ANÁLOGOS PARA ALIVIAR LOS EFECTOS SECUNDARIOS ASOCIADOS CON LA QUIMIOTERAPIA

NÚMERO DE SOLICITUD: 12777902 FECHA DE PRESENTACIÓN: 11/05/2010
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: MÉTODOS PARA EL TRATAMIENTO DE TRASTORNOS ONCOLÓGICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES

NÚMERO DE SOLICITUD: 12778010 FECHA DE PRESENTACIÓN: 11/05/2010
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: MÉTODO PARA EL DIAGNÓSTICO DE TRASTORNOS METABÓLICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES

NÚMERO DE SOLICITUD: 12778029 FECHA DE PRESENTACIÓN: 11/05/2010
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: MÉTODOS PARA EL DIAGNÓSTICO DE TRASTORNOS ONCOLÓGICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES

NÚMERO DE SOLICITUD: 12778054 FECHA DE PRESENTACIÓN: 11/05/2010
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: MÉTODOS PARA EL TRATAMIENTO DE TRASTORNOS METABÓLICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES

NÚMERO DE SOLICITUD: 12778094 FECHA DE PRESENTACIÓN: 11/05/2010
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: MÉTODOS PARA EL TRATAMIENTO DE TRASTORNOS ONCOLÓGICOS UTILIZANDO UN MODIFICADOR EPIMETABÓLICO (COENZIMA Q10)

NÚMERO DE SOLICITUD: 12853431 FECHA DE PRESENTACIÓN: 10/08/2010
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: LA VITAMINA D3 Y SUS ANÁLOGOS PARA EL TRATAMIENTO DE LA ALOPECIA

NÚMERO DE SOLICITUD: 12864185 FECHA DE PRESENTACIÓN: 18/10/2010
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: SISTEMA DE ENSAYO PARA LA EVALUACIÓN DE ONCOGENICIDAD, PROGRESIÓN DE UN TUMOR, Y EFICACIA DE UN TRATAMIENTO

NÚMERO DE SOLICITUD: 61023570 FECHA DE PRESENTACIÓN: 25/01/2008
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: SISTEMA DE ENSAYO PARA LA EVALUACIÓN DE ONCOGENICIDAD, PROGRESIÓN DE UN TUMOR, Y EFICACIA DE UN TRATAMIENTO

NÚMERO DE SOLICITUD: 61147549 FECHA DE PRESENTACIÓN: 27/01/2009
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: LA VITAMINA D3 Y SUS ANÁLOGOS PARA ALIVIAR LOS EFECTOS SECUNDARIOS ASOCIADOS CON LA QUIMIOTERAPIA

NÚMERO DE SOLICITUD: 61177241 FECHA DE PRESENTACIÓN: 11/05/2009
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: MÉTODOS PARA EL TRATAMIENTO DE TRASTORNOS ONCOLÓGICOS UTILIZANDO UN MODIFICADOR EPIMETABÓLICO (COENZIMA Q10)

NÚMERO DE SOLICITUD: 61177243 FECHA DE PRESENTACIÓN: 11/05/2009
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: MÉTODOS PARA EL TRATAMIENTO DE TRASTORNOS ONCOLÓGICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES

NÚMERO DE SOLICITUD: 61177244 FECHA DE PRESENTACIÓN: 11/05/2009
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: MÉTODOS PARA EL DIAGNÓSTICO DE TRASTORNOS ONCOLÓGICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES

NÚMERO DE SOLICITUD: 61177245 FECHA DE PRESENTACIÓN: 11/05/2009
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: MÉTODOS PARA EL TRATAMIENTO DE TRASTORNOS METABÓLICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES

NÚMERO DE SOLICITUD: 61177246 FECHA DE PRESENTACIÓN: 11/05/2009
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: MÉTODO PARA EL DIAGNÓSTICO DE TRASTORNOS METABÓLICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES

NÚMERO DE SOLICITUD: 61234178 FECHA DE PRESENTACIÓN: 14/08/2009
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: LA VITAMINA D3 Y SUS ANÁLOGOS PARA EL TRATAMIENTO DE LA ALOPECIA

NÚMERO DE SOLICITUD: 61239003 FECHA DE PRESENTACIÓN: 01/09/2009
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: LA VITAMINA D3 Y SUS ANÁLOGOS PARA ALIVIAR LOS EFECTOS SECUNDARIOS ASOCIADOS CON LA QUIMIOTERAPIA

NÚMERO DE SOLICITUD: 61363562 FECHA DE PRESENTACIÓN: 12/07/2010
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: MÉTODOS PARA EL TRATAMIENTO DE LA ENDOMETRIOSIS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES

NÚMERO DE SOLICITUD: 61448587 FECHA DE PRESENTACIÓN: 02/03/2011
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: ENSAYOS INTERROGATORIOS BASADOS EN CÉLULAS Y USOS DE LOS MISMOS

NÚMERO DE SOLICITUD: 61474708 FECHA DE PRESENTACIÓN: 12/04/2011
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: COMPOSICIONES Y MÉTODOS PARA EL TRATAMIENTO DEL TRASTORNO GENERALIZADO DEL DESARROLLO

NÚMERO DE SOLICITUD: 61491132 FECHA DE PRESENTACIÓN: 27/05/2011
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: MODULADORES DEL METABOLISMO MITOCONDRIAL Y USOS DE LOS MISMOS

NÚMERO DE SOLICITUD: 61506221 FECHA DE PRESENTACIÓN: 11/07/2011
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: MÉTODOS PARA EL TRATAMIENTO DE LA ENDOMETRIOSIS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES

NÚMERO DE SOLICITUD: 61511487 FECHA DE PRESENTACIÓN: 25/07/2011
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: EL USO DE MALATO-DESHIDROGENASA 2 (MDH2) EN EL DIAGNÓSTICO Y TRATAMIENTO DE ESTADOS DE ENFERMEDAD

NÚMERO DE SOLICITUD: 61514388 FECHA DE PRESENTACIÓN: 02/08/2011
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: MÉTODOS Y KITS PARA LA DETECCIÓN DE LA COENZIMA Q10

NÚMERO DE SOLICITUD: 61593848 FECHA DE PRESENTACIÓN: 01/02/2012
NÚMERO DE PATENTE: FECHA DE EMISIÓN:
TÍTULO: ENSAYOS DE INTERROGATORIO BASADOS EN CÉLULAS Y USOS DE LOS MISMOS

SUCURSAL DE ASIGNACIÓN DE INSCRIPCIONES
DIVISIÓN DE REGISTROS PUBLICOS

CESIÓN

Esta Cesión (la "***Cesión***") se ejecuta a partir del 11 de junio de 2012 entre **Berg Biosystems, LLC**, una sociedad debidamente constituida conforme y de acuerdo con las leyes del Estado de Delaware, y que tiene su centro de actividad principal en 1845 Elm Hill Pike, Nashville, Tennessee 37210 ("Cedente"), y **Berg Pharma LLC**, una sociedad debidamente constituida conforme y de acuerdo con las leyes del Estado de Delaware, y que tiene su centro de actividad principal en 1845 Elm Hill Pike, Nashville, Tennessee 37210 ("***Cesionario***").

Trasfondo

El Cedente posee esas ciertas patentes y solicitudes de patentes que figuran en el Listado A y los inventos y mejoras nuevas y útiles que se describen y reivindican en él y todos los conocimientos relacionados con el mismo que existen a partir de la fecha indicada anteriormente (todo ello conjuntamente, las "***Patentes***"). El Cesionario está deseoso de adquirir la totalidad de los derechos, títulos e intereses del Cedente en y de los patentes.

Condiciones

Por consideración de **un dólar ($1,00)**, y por otra consideración buena y valiosa, el recibo, la adecuación y suficiencia de la cual se reconoce por medio de la presente, el Cedente por medio de este instrumento asigna, transfiere y transmite al Cesionario, y los sucesores y cesionarios del Cesionario, todos los derechos, títulos e intereses del Cedente sobre y para las patentes, y a todas las Cartas patentes que puedan ser o se concederán referentes a las solicitudes de patentes o invenciones dentro de las patentes, o cualquier parte de las mismas, o cualquier solicitud divisional, de continuación, de continuación en parte, sustitución, renovación, nueva publicación o sobre otras solicitudes basadas en su totalidad o en parte en ellas, y todas las reediciones, la repetición del examen y la prórroga de las mismas, en cada caso, en cualquier país y que incluye cualquier contraparte extranjera del mismo o la misma, junto con el derecho de presentar las solicitudes y el derecho a reclamar los mismos los derechos de prioridad derivados de las solicitudes de patentes en las patentes, y aún más, junto con todas las reclamaciones por daños y perjuicios a causa de una infracción pasada, presente o futura, con el derecho de demandar por, y recoger las mismas para el uso del Cesionario, los sucesores y cesionarios del Cesionario y todas las ganancias de los anteriores, incluyendo, sin limitación, cualquier reclamo por el Cedente frente a terceros, o los beneficios derivados de una infracción pasada, presente o futura de las patentes, tales beneficios a ser mantenidos y disfrutados por el Cesionario y los cesionarios y sucesores del Cesionario tan plenamente como los mismos se hubieran celebrado y disfrutado por el Cedente si no se hubiera llevado a cabo esta cesión.

El Cedente se compromete, por sí mismo y los ejecutores, administradores y sucesores en interés del Cedente, a ejecutar, por cuenta del Cesionario, todas las aplicaciones, especificaciones otrosíes, escrituras u otros instrumentos, y realizar todos los actos necesarios o apropiados para garantizar la concesión del Cesionario de la patente de letras en las solicitudes en las patentes, con las especificaciones y reivindicaciones en forma que deberá ser aprobada por la Junta del Cesionario, para investir y confirmar en el Cesionario y los cesionarios y sucesores del Cesionario, la titularidad legal de todas las patentes.

El Cedente autoriza y solicita al Comisionado de Patentes y Marcas de los Estados Unidos, y a cualquier funcionario de cualquier país o países extranjeros a los Estados Unidos cuya misión es la emisión de patentes o de otras pruebas o formas de protección de la propiedad industrial en las solicitudes, como queda dicho, a emitir dichas patentes sobre cualquiera de estas solicitudes dentro de las patentes al Cesionario y los sucesores, representantes legales y cesionarios del Cesionario, tan plena y completamente como las mismas se hubieran celebrado y disfrutado por el Cedente si no se hubiera llevado a cabo esta cesión.

En constancia de lo cual, esta Cesión queda debidamente firmada a la fecha indicada a continuación.

BERG BIOSYSTEMS, LLC
CEDENTE:
/Firmado/
Por: Rangaprasad Sarangarajan
Cargo: Vicepresidente de Investigacion y Desarrollo

Testigo:

Fecha: 06/11/2012 /Firmado/

BERG PHARMA, LLC
CESIONARIO:

/Firmado/
Por: Niven R. Narain
Cargo: Presidente

Testigo:

Fecha: 06/11/2012 /Firmado/

LISTADO A

Confidencial - Privilegio Cliente - Abogado

N° de Serie	Fecha presentación	Expediente Legal N°	Título
61/177241	11 -May- 2009	119992-00601	MÉTODOS PARA EL TRATAMIENTO DE TRASTORNOS ONCOLÓGICOS UTILIZANDO UN MODIFICADOR EPIMETABÓLICO (COENZIMA Q10)
12/778094	11 -May- 2010	119992-00602	MÉTODOS PARA EL TRATAMIENTO DE TRASTORNOS ONCOLÓGICOS UTILIZANDO UN MODIFICADOR EPIMETABÓLICO (COENZIMA Q10)
61/177243	11 -May- 2009	119992-00701	MÉTODOS PARA EL TRATAMIENTO DE TRASTORNOS ONCOLÓGICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES
12/777902	11 -May- 2010	119992-00702	MÉTODOS PARA EL TRATAMIENTO DE TRASTORNOS ONCOLÓGICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES

61/177244	11 -May- 2009	119992-00801	MÉTODOS PARA EL DIAGNÓSTICO DE TRASTORNOS ONCOLÓGICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES
12/778029	11 -May- 2010	119992-00802	MÉTODOS PARA EL DIAGNÓSTICO DE TRASTORNOS ONCOLÓGICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES
61/177245	11 -May- 2009	119992-00901	MÉTODOS PARA EL TRATAMIENTO DE TRASTORNOS METABÓLICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES
12/778054	11 -May- 2010	119992-00902	MÉTODOS PARA EL TRATAMIENTO DE TRASTORNOS METABÓLICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES
61/491132	27 -May- 2011	119992-00903	MODULADORES DEL METABOLISMO MITOCONDRIAL Y USOS DE LOS MISMOS
61/177246	11 -May- 2009	119992-01001	MÉTODO PARA EL DIAGNÓSTICO DE TRASTORNOS METABÓLICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES
12/778010	11 -May- 2010	119992-01002	MÉTODO PARA EL DIAGNÓSTICO DE TRASTORNOS METABÓLICOS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES
61/147549	27 -May- 2009	119992-01501	LA VITAMINA D3 Y SUS ANÁLOGOS PARA ALIVIAR LOS EFECTOS SECUNDARIOS ASOCIADOS CON LA QUIMIOTERAPIA
61/239003	01 - Sep- 2009	119992-01502	LA VITAMINA D3 Y SUS ANÁLOGOS PARA ALIVIAR LOS EFECTOS SECUNDARIOS ASOCIADOS CON LA QUIMIOTERAPIA
12/695113	27 -Ene- 2010	119992-01503	LA VITAMINA D3 Y SUS ANÁLOGOS PARA ALIVIAR LOS EFECTOS SECUNDARIOS ASOCIADOS CON LA QUIMIOTERAPIA
61/234178	14 -Ago- 2009	119992-01701	LA VITAMINA D3 Y SUS ANÁLOGOS PARA EL TRATAMIENTO DE LA ALOPECIA
12/853431	10 -Ago- 2010	119992-01703	LA VITAMINA D3 Y SUS ANÁLOGOS PARA EL TRATAMIENTO DE LA ALOPECIA
61/023570	25 - Ene- 2008	119992-03201	SISTEMA DE ENSAYO PARA LA EVALUACIÓN DE ONCOGENICIDAD, PROGRESIÓN DE UN TUMOR, Y EFICACIA DE UN TRATAMIENTO
12/864185	18 -Oct- 2010	119992-03202	SISTEMA DE ENSAYO PARA LA EVALUACIÓN DE ONCOGENICIDAD, PROGRESIÓN DE UN TUMOR, Y EFICACIA DE UN TRATAMIENTO
61/363562	12 - Jul- 2010	119992-04101	MÉTODOS PARA EL TRATAMIENTO DE LA ENDOMETRIOSIS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES
61/506221	11 - Jul- 2011	119992-04102	MÉTODOS PARA EL TRATAMIENTO DE LA ENDOMETRIOSIS USANDO MODIFICADORES EPIMETABÓLICOS, MOLÉCULAS MULTIDIMENSIONALES INTRACELULARES O INFLUYENTES AMBIENTALES
61/514388	02 -Ago- 2011	119992-04201	MÉTODOS Y KITS PARA LA DETECCIÓN DE LA COENZIMA Q10
61/474708	12 -Abr- 2011	119992-04301	COMPOSICIONES Y MÉTODOS PARA EL TRATAMIENTO DEL TRASTORNO GENERALIZADO DEL DESARROLLO
61/448587	02 -Mar- 2011	119992-04403	ENSAYOS INTERROGATORIOS BASADOS EN CÉLULAS Y USOS DE LOS MISMOS
61/593848	01-Feb- 2012	119992-04404	ENSAYOS DE INTERROGATORIO BASADOS EN CÉLULAS Y USOS DE LOS MISMOS
61/511487	25 - Jul- 2011	119992-05101	EL USO DE MALATO-DESHIDROGENASA 2 (MDH2) EN EL DIAGNÓSTICO Y TRATAMIENTO DE ESTADOS DE ENFERMEDAD

Traducción efectuada por Alvaro Becerra
Id.250799-WPC20488

ME13317390v1

(119992-04429)

IN THE PATENT OFFICE OF COLOMBIA

Assignee(s): **Berg Biosystems, LLC**

Inventor(s): **Niven Rajin Narain et al.**

Serial Nos. &
Filing Dates: **USSN 61/448587 filed 02 March 2011**
USSN 61/593848 filed 01 February 2012

I, Jill A. Mello, Ph.D., J.D., attorney for the applicant, hereby attest that the attached document is a true and exact copy of the Assignment executed by inventors, **Niven Rajin Narain**, **Rangaprasad Sarangarajan**, and **Vivek Kannoth Vishnudas**. The Assignment is from said inventor(s) to **Berg Biosystems, LLC**, a company organized and existing under the laws of Delaware.

Jill A. Mello, Ph.D., J.D.
Registration No. 55,450
Attorney for Applicant

United States of America)
Commonwealth of Massachusetts) ss.:
County of Suffolk)

On the 20th day of August, 2013 before me, the undersigned notary public, personally appeared Jill A. Mello (name of document signer), proved to me through satisfactory evidence of identification, which was that she is known to me personally, to be the person whose name is signed on the preceding or attached document in my presence.

(official signature and seal of notary)
Jessica Dale Bell
My commission expires March 16, 2018.

JESSICA DALE BELL
Notary Public
Massachusetts
Commission Expires Mar 16, 2018

417

EAH/JAM/JW/Pat

119992-04403
119992-04404
ENTERED

UNITED STATES PATENT AND TRADEMARK OFFICE

UNDER SECRETARY OF COMMERCE FOR INTELLECTUAL PROPERTY AND
DIRECTOR OF THE UNITED STATES PATENT AND TRADEMARK OFFICE

MARCH 8, 2012

PTAS

MCCARTER & ENGLISH, LLP
265 FRANKLIN STREET
BOSTON, MA 02110

501842776

UNITED STATES PATENT AND TRADEMARK OFFICE
NOTICE OF RECORDATION OF ASSIGNMENT DOCUMENT

THE ENCLOSED DOCUMENT HAS BEEN RECORDED BY THE ASSIGNMENT RECORDATION BRANCH OF THE U.S. PATENT AND TRADEMARK OFFICE. A COMPLETE COPY IS AVAILABLE AT THE ASSIGNMENT SEARCH ROOM ON THE REEL AND FRAME NUMBER REFERENCED BELOW.

PLEASE REVIEW ALL INFORMATION CONTAINED ON THIS NOTICE. THE INFORMATION CONTAINED ON THIS RECORDATION NOTICE REFLECTS THE DATA PRESENT IN THE PATENT AND TRADEMARK ASSIGNMENT SYSTEM. IF YOU SHOULD FIND ANY ERRORS OR HAVE QUESTIONS CONCERNING THIS NOTICE, YOU MAY CONTACT THE ASSIGNMENT RECORDATION BRANCH AT 571-272-3350. PLEASE SEND REQUEST FOR CORRECTION TO: U.S. PATENT AND TRADEMARK OFFICE, MAIL STOP: ASSIGNMENT RECORDATION BRANCH, P.O. BOX 1450, ALEXANDRIA, VA 22313.

RECORDATION DATE: 03/06/2012

REEL/FRAME: 027817/0890
NUMBER OF PAGES: 3

BRIEF: ASSIGNMENT OF ASSIGNORS INTEREST (SEE DOCUMENT FOR DETAILS).

DOCKET NUMBER: 119992-04403 & 04404

ASSIGNOR:
NARAIN, NIVEN RAJIN

DOC DATE: 03/02/2012

ASSIGNOR:
SARANGARAJAN, RANGAPRASAD

DOC DATE: 03/02/2012

ASSIGNOR:
VISHNUDAS, VIVEK KANNOTH

DOC DATE: 03/02/2012

ASSIGNEE:
BERG BIOSYSTEMS, LLC
1845 ELM HILL PIKE
NASHVILLE, TENNESSEE 37210

APPLICATION NUMBER: 61448587
PATENT NUMBER:
TITLE: INTERROGATORY CELL-BASED ASSAYS AND USES THEREOF

FILING DATE: 03/02/2011
ISSUE DATE:

APPLICATION NUMBER: 61593848
PATENT NUMBER:
TITLE: INTERROGATORY CELL-BASED ASSAYS AND USES THEREOF

FILING DATE: 02/01/2012
ISSUE DATE:

RECEIVED
MAR 09 2012
McCARTER & ENGLISH

ASSIGNMENT RECORDATION BRANCH
PUBLIC RECORDS DIVISION

ASSIGNMENT BY INVENTOR

WE, Niven Rajin NARAIN, Rangaprasad SARANGARAJAN and Vivek Kannoth VISHNUDAS of Cambridge, Massachusetts; Boylston, Massachusetts and Waltham, Massachusetts, respectively, in consideration of One Dollar and other valuable consideration paid to me by

Berg Biosystems, LLC

a Corporation of Delaware, having its principal place of business at 1845 Elm Hill Pike, Nashville, Tennessee 37210, the receipt of which is hereby acknowledged, do hereby sell, assign and transfer unto said

Berg Biosystems, LLC

its successors and assigns, the entire interest for the United States of America and all foreign countries including all rights of priority under the International Convention for the Protection of Industrial Property in a certain invention or improvement in

INTERROGATORY CELL-BASED ASSAYS AND USES THEREOF

described in an application

☐ Executed by us of even date herewith and about to be filed

or

☒ Application No. 61/448587, filed on March 2, 2011
☒ Application No. 61/593848, filed on February 1, 2012

in the United States Patent and Trademark Office, and in all Letters Patent of the United States and all foreign countries which may or shall be granted on said invention, or any parts thereof, or on said application, or any divisional, continuing, reissue or other applications based in whole or in part thereon. And we agree, for ourselves and our executors and administrators, with said Corporation and its successors and assigns but at its or their expense and charges, hereafter to execute all applications, amended specifications, deed or other instrument, and to do all acts necessary or proper to secure the grant of Letters Patent in the United States and in all other countries to said Corporation, with specifications and claims in such form as shall be approved by the counsel of said Corporation and to vest and confirm in said Corporation, its successors and assigns, the legal title to all such patents.

U.S.S.N. 61/448587
U.S.S.N. 61/593848

Attorney Docket No. 119992-04403
Attorney Docket No. 119992-04404

And we do hereby authorize and request the Commissioner of Patents and Trademarks of the United States to issue such Letters Patent as shall be granted upon said application or applications based thereon to said Corporation, its successors and assigns.

03/02/2012
Date

Niven Rajin NARAIN

Witness:
March 2, 2012

03/02/2012
Date

Rangaprasad SARANGARAJAN

Witness:
March 2, 2012

03/02/2012
Date

Vivek Kannoth VISHNUDAS

Witness:
March 2, 2012

421

(119992-04429)

EN LA OFICINA DE PATENTES DE COLOMBIA

Cesionario(s): **Berg Biosystems, LLC**

Inventor(es): **Niven Rajin Narain et al.**

Nos. de serie y
Fechas de presentación: **USSN 61/448587 presentada 02 de marzo 2011**
USSN 61/593848 presentada 01 de febrero 2012

Yo, Jill A. Mello, Ph.D., J.D., abogada de los solicitantes, por la presente declaro que el documento adjunto es una copia fiel y exacta de la asignación ejecutada por los inventores, **Niven Rajin Narain, Rangaprasad Sarangarajan** y **Vivek Kannoth Vishnudas**. La asignación es de dicho(s) inventor(es) a **Berg Biosystems, LLC**, sociedad constituida y existente bajo las leyes de Delaware.

/ Firmado /
Jill A. Mello, Ph.D., J.D.
No. de Registro Fiscal 55.450
Abogada por los Solicitantes

Estados Unidos de América)
Estado de Massachusetts) ss.:
Condado de Suffolk)

El día 20 de agosto de 2013 ante mí, el notario que suscribe, compareció personalmente Jill A. Mello (nombre del firmante de documentos), quien demostró con evidencia satisfactoria para su identificación, la cual era que ella me es conocida personalmente, ser la persona cuyo nombre se halla firmado en el documento anterior o adjunto en mi presencia.

/Firmado/ (firma y sello oficiales de notario)
Jessica Dale Bell
Mi comisión expira el 16 de marzo 2018.

[Sello que lee:]
JESSICA DALE BELL
Notario público
Massachusetts
Comisión Expira 16 de marzo 2018

EAH/JAM/JW/Pat

119992-04403
119992-04404
ARCHIVADO

OFICINA DE PATENTES Y MARCAS DE LOS ESTADOS UNIDOS

SUBSECRETARIO DE COMERCIO PARA LA PROPIEDAD INTELECTUAL Y
DIRECTOR DE LA OFICINA DE PATENTES Y MARCAS DE LOS ESTADOS UNIDOS

08 DE MARZO 2012

PTAS

MCCARTER & ENGLISH, LLP
265 FRANKLIN STREET
BOSTON, MA 02110

501842776

OFICINA DE PATENTES Y MARCAS DE LOS ESTADOS UNIDOS
AVISO DE INSCRIPCIÓN DE DOCUMENTO DE ASIGNACIÓN

EL DOCUMENTO ADJUNTO HA SIDO REGISTRADO POR LA RAMA DE INSCRIPCIÓN DE ASIGNACIONES DE LA OFICINA DE PATENTES Y MARCAS DE LOS ESTADOS UNIDOS. UNA COPIA COMPLETA ESTÁ DISPONIBLE EN LA SALA DE ASIGNACIÓN BAJO EL CARRETE Y NÚMERO DE MARCO LISTADO A CONTINUACION.

POR FAVOR LEA TODA LA INFORMACIÓN CONTENIDA EN ESTE AVISO. LA INFORMACIÓN CONTENIDA EN ESTE AVISO DE INSCRIPCIÓN REFLEJA LA PRESENCIA DE DATOS EN EL SISTEMA DE ASIGNACIONES DE PATENTES Y MARCAS. SI ENCUENTRA ERRORES O TIENE PREGUNTAS SOBRE ESTE AVISO, PUEDE CONTACTAR A LA RAMA DE INSCRIPCIÓN DE ASIGNACIONES, 571-272-3350. POR FAVOR ENVIAR SOLICITUD DE CORRECCIÓN A: OFICINA DE PATENTES Y MARCAS DE LOS ESTADOS UNIDOS, PARADA DE CORREO: SUCURSAL DE INSCRIPCIÓN DE ASIGNACIONES, APARTADO POSTAL 1450, ALEXANDRIA, VIRGINIA 22313.

FECHA DE INSCRIPCIÓN: 03/06/2012 CARRETE/MARCO: 027817 / 0890
NÚMERO DE PÁGINAS: 3

MEMORIAL: ASIGNACIÓN INTERÉS DE CEDENTES (VÉASE DOCUMENTO PARA DETALLES).

NÚMERO DE EXPEDIENTE: 119992-04403 & 04404

CEDENTE:
NARAIN, NIVEN RAJIN DOC FECHA: 02/03/2012

CEDENTE:
SARANGARAJAN, RANGAPRASAD DOC FECHA: 02/03/2012

CEDENTE:
VISHNUDAS, VIVEK KANNOTH DOC FECHA: 02/03/2012

CESIONARIO:
BERG BIOSYSTEMS, LLC
1845 ELM HILL PIKE
NASHVILLE, TENNESSEE 37210

NÚMERO DE SOLICITUD: 61448587 FECHA DE PRESENTACION: 02/03/2011
NÚMERO DE PATENTE: FECHA DE EMISION:
TÍTULO: ENSAYOS INTERROGATORIOS BASADOS EN CÉLULAS, Y USOS DE LOS MISMOS

NÚMERO DE SOLICITUD: 61593848
NÚMERO DE PATENTE: FECHA DE EMISION:
TÍTULO: ENSAYOS INTERROGATORIOS BASADOS EN CÉLULAS, Y USOS DE LOS MISMOS

[Sello que lee:]
RECIBIDO
9 DE MARZO DE 2012
McCARTER & ENGLISH

SUCURSAL DE INSCRIPCIÓN DE CESIONES
DIVISION DE REGISTROS PUBLICOS

CESION POR INVENTOR

NOSOTROS, Niven Rajin NARAIN, Rangaprasad SARANGARAJAN y Vivek Kannoth VISHNUDAS de Cambridge, Massachusetts, Boylston, Massachusetts y Waltham, Massachusetts, respectivamente, por la consideración de un dólar y otras consideraciones de valor pagadas por

Berg Biosystems, LLC,

una corporación de Delaware, que tiene su centro de actividad principal en 1845 Elm Hill Pike, Nashville, Tennessee 37210, cuyo recibo queda reconocido, por la presente vendemos, cedemos y transferimos a dicha

Berg Biosystems, LLC,

sus sucesores y cesionarios, todo el interés en los Estados Unidos y todos los países extranjeros, incluyendo todos los derechos de prioridad en virtud del Convenio Internacional para la Protección de la Propiedad Industrial en la invención o mejora

ENSAYOS DE INTERROGATORIO BASADOS EN CÉLULAS Y USOS DE LOS MISMOS,

descrita en una solicitud

☐ Ejecutada por nuestra parte en esta misma fecha, a punto de ser presentada

o

☒ Solicitud N° 61/448587, presentada el 02 de marzo de 2011
☒ Solicitud N° 61/593848, presentada el 01 de febrero 2012

en la Oficina de Marcas y Patentes de Estados Unidos, y en todas las Cartas Patente de los Estados Unidos y todos los países extranjeros en los que se puedan conceder o se concederán a dicha invención, o cualquier parte del mismo, o en dicha solicitud, o en cualquier aplicación divisional, de continuación, reedición u otras basadas en su totalidad o en parte al respecto. Y estamos de acuerdo, a nombre propio y de nuestros ejecutores y administradores, con que dicha Corporación y sus sucesores y cesionarios, pero a su o sus expensas y gastos, en lo sucesivo ejecute(n) todas las aplicaciones, especificaciones modificadas, escrituras u otros instrumentos, y realicen todos los actos necesarios o adecuados para garantizar la concesión de Cartas patentes en los Estados Unidos y en todos los demás países a dicha Corporación, con las especificaciones y reivindicaciones en forma que deberá ser aprobada por la Junta de dicha Corporación y para investir y confirmar en dicha Corporación, sus sucesores y cesionarios, el título de propiedad legal de todas tales patentes.

U.S.S.N. 61/448587
U.S.S.N. 61/593848

Expediente Legal N ° 119992-04403
Expediente Legal N ° 119992-04404

Y por la presente autorizamos y solicitamos al Comisionado de Marcas y Patentes de los Estados Unidos que emita tales Cartas Patente como sean otorgadas sobre tal o tales solicitud o solicitudes basadas sobre las mismas a dicha Corporación, sus sucesores y cesionarios.

02/03/2012 / Firmado /

Fecha Niven Rajin NARAIN

Testigo:

02 de marzo 2012 / Firmado /

02/03/2012 / Firmado /

Fecha Rangaprasad SARANGARAJAN

Testigo:

02 de marzo 2012 / Firmado /

03/02/2012 / Firmado /

Fecha Vivek Kannoth VISHNUDAS

Testigo:

02 de marzo 2012 / Firmado /

Traducción efectuada por Alvaro Becerra
Id.250799-WPC20488

(119992-04429)

IN THE PATENT OFFICE OF COLOMBIA

CERTIFICATION

I, Jill A. Mello, Ph.D., J.D., attorney for the applicant, hereby attest that the attached document is a true and exact copy of the Certificate of Amendment, filed with the Secretary of State of the State of Delaware on the 5th day of April 2013, changing the name of **Berg Pharma LLC** to **Berg LLC**.

Jill A. Mello, Ph.D., J.D.
Registration No. 55,450
Attorney for Applicant

United States of America)
Commonwealth of Massachusetts) ss.:
County of Suffolk)

On the 20th day of August, 2013 before me, the undersigned notary public, personally appeared Jill A. Mello (name of document signer), proved to me through satisfactory evidence of identification, which was that she is known to me personally, to be the person whose name is signed on the preceding or attached document in my presence.

(official signature and seal of notary)
Jessica Dale Bell
My commission expires March 16, 2018.

JESSICA DALE BELL
Notary Public
Massachusetts
Commission Expires Mar 16, 2018

Delaware

The First State

I, JEFFREY W. BULLOCK, SECRETARY OF STATE OF THE STATE OF DELAWARE, DO HEREBY CERTIFY THE ATTACHED IS A TRUE AND CORRECT COPY OF THE CERTIFICATE OF AMENDMENT OF "BERG PHARMA LLC", CHANGING ITS NAME FROM "BERG PHARMA LLC" TO "BERG LLC", FILED IN THIS OFFICE ON THE FIFTH DAY OF APRIL, A.D. 2013, AT 1:35 O'CLOCK P.M.

Jeffrey W. Bullock, Secretary of State

4242094 8100

130403872

AUTHENTICATION: 0350177

DATE: 04-11-13

You may verify this certificate online at corp.delaware.gov/authver.shtml

4025

State of Delaware
Secretary of State
Division of Corporations
Delivered 01:35 PM 04/05/2013
FILED 01:35 PM 04/05/2013
SRV 130403872 - 4242094 FILE

STATE OF DELAWARE
CERTIFICATE OF AMENDMENT

1. Name of Limited Liability Company: BERG PHARMA LLC

2. The Certificate of Formation of the limited liability company is hereby amended as follows:

THE NAME OF THE LIMITED LIABILITY COMPANY IS HEREBY CHANGED FROM 'BERG PHARMA LLC' TO 'BERG LLC'.

IN WITNESS WHEREOF, the undersigned have executed this Certificate on the 5th day of APRIL, A.D. 2013.

By: _____
Authorized Person(s)

Name: ERNEST R. MARTINEZ
Print or Type

(119992-04429)

EN LA OFICINA DE PATENTES DE COLOMBIA

CERTIFICACION

Yo, Jill A. Mello, Ph.D., J.D., abogada de los solicitantes, por la presente declaro que el documento adjunto es una copia fiel y exacta del Certificado de Enmienda, radicado ante el Secretario de Estado del estado de Delaware el día 5 de abril del 2013, en el que se cambia el nombre de **Berg Pharma LLC** a **Berg LLC**.

/ Firmado /
Jill A. Mello, Ph.D., J.D.
No. de Registro Fiscal 55.450
Abogada por los Solicitantes

Estados Unidos de América)
Estado de Massachusetts) ss.:
Condado de Suffolk)

El día 20 de agosto de 2013 ante mí, el notario que suscribe, compareció personalmente Jill A. Mello (nombre del firmante de documentos), quien demostró con evidencia satisfactoria de identificación, en que ella me es conocida personalmente, ser la persona cuyo nombre se halla firmado en el documento anterior o adjunto en mi presencia.

/Firmado/ (firma y sello oficiales de notario)
Jessica Dale Bell
Mi comisión expira el 16 de marzo 2018.

[Sello que lee:]
JESSICA DALE BELL
Notario público
Massachusetts
Comisión Expira 16 de marzo 2018

Delaware

El Primer Estado

YO, JEFFREY W. BULLOCK, SECRETARIO DE ESTADO DEL ESTADO DE DELAWARE , POR LA PRESENTE CERTIFICO QUE LOS DOCUMENTOS ADJUNTOS SON COPIA FIEL Y CORRECTA DEL CERTIFICADO DE ENMIENDA DE "BERG PHARMA LLC", CAMBIANDO SU NOMBRE DE "BERG PHARMA LLC" A "BERG LLC", PRESENTADO EN ESTA OFICINA EN EL QUINTO DÍA DE ABRIL , AD 2013 , A LA 1:35 PM EN PUNTO.

/Firmado/
Jeffrey W.

Bullock, Secretario De Estado

FECHA: ***4242094 8100***

AUTENTICACIÓN: 0350177

04-04-11-13

Usted ***130403872*** *puede verificar este certificado*

En línea en corp.delaware.gov/authver.shtml

Estado de Delaware
Secretaria de Estado
División de Corporaciones
Entregado 13:35 04/05/2013
PRESENTADO 13:35 04/05/2013
SRV 130403872-4242090 ARCHIVO

ESTADO DE DELAWARE
CERTIFICADO DE ENMIENDA

1. Nombre de la Sociedad de Responsabilidad Limitada: BERG PHARMA LLC
2. El Certificado de Formación de la sociedad de responsabilidad limitada se modifica como sigue:

EL NOMBRE DE LA SOCIEDAD DE RESPONSABILIDAD LIMITADA POR LA PRESENTE SE CAMBIÓ DE 'BERG PHARMA LLC' A 'BERG LLC'.

EN FE DE LO CUAL, los abajo firmantes han suscrito este certificado el quinto día de abril, AD 2013.

Por: /Firmado/
Persona(a) Autorizada(s)

Nombre: ERNEST R MARTINEZ
Imprenta o a máquina

Traducción efectuada por Alvaro Becerra
Id.250799-WPC20488

428

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
NIT : 800.176.089-2
- / -

Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 13 - 0098033

Bogotá D.C., Septiembre 30 de 2013 - 11:59:00

RECIBIDO DE : CAVELIER ABOGADOS NI 860.041.367

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	599246756	30/09/2013	10.237.000.00

*** Conceptos Pagados ***

CANT. RENTISTICO	CONCEPTO	Vr.UNDITARIO	Vr.CONCEPTO
1 50005-01-01 SOLICITUDES	1 TRAMITES DE SOL. DE PATENTE DE INVENCION	500.000.00	500.000.00
			$500.000.00

SON: **QUINIENTOS MIL PESOS MONEDA CORRIENTE***

Responsable: _____
-***-
Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
No. 13-233174- -00000-0000
Fecha: 2013-10-01 16:38:27 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 429

COJO: 20488-841-001-4

Industria y Comercio
SUPERINTENDENCIA

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
DIVISION DE NUEVAS CREACIONES
EXTRACTO PARA PUBLICACIÓN
SOLICITUD FASE NACIONAL PCT

(21) N° de solicitud:

22) Fecha de solicitud:

(30) Prioridad

(31) *Prioridad*	(32) *Fecha*	(33) *País*
61/448/587	**2011/03/02**	**US**
61/593,848	**2012/02/01**	**US**

(51) Int Cl: **A61P 035/000**

(71)Solicitante(s) **BERG LLC**

(72) Inventor(es) **NIVEN RAJIN NARAIN**
Y OTROS

(74) Apoderado: **LUZ CLEMENCIA DE PAEZ**

(85) Fecha limite inicio fase nacional: 2013/10/02
(86) Datos relativos a la presentación de la solicitud PCT
Fecha de presentación de la solicitud: 2012/03/02
No. de solicitud PCT/US2012/027615

(87) Publicación internacional
Fecha: 2012/09/07
N° publicación: ***WO 2012/119129***

(54) Titulo: **ENSAYOS INTERROGATORIOS BASADOS EN CÉLULAS Y USOS DE LOS MISMOS**

Resumen:

En la presente se describe un Tecnología de Plataforma de descubrimiento para analizar un sistema o proceso biológico (por ejemplo, una enfermedad, tal como cáncer) por medio de un modelo de construcción. En particular, se describe en la presente un método para identificar un modulador de un sistema biológico mediante el establecimiento de un modelo para el sistema biológico, obtener un primer conjunto de datos del modelo para el sistema biológico, obtener un segundo conjunto de datos del modelo para el sistema biológico, y generar una red de relaciones causales de consenso entre los niveles de expresión de la pluralidad de genes y la actividad funcional o la respuesta celular basada únicamente en el primer conjunto de datos y el segundo conjunto de datos, usando un dispositivo informático programado.

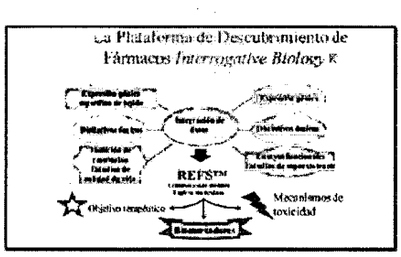


Figura 7

Industria y Comercio
SUPERINTENDENCIA

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

No. 13-233174- -00000-0000

Fecha: 2013-10-01 16:38:27 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 429

IES

PATENTE DE INVENCIÓN ☐ MODE ;/00

- ☐ Indicación que se solicita una patente.
- ☐ Datos de identificación del solicitante o de la persona que presenta la solicitud
- ☐ Descripción de la invención
- ☐ Dibujos de ser estos pertinentes
- ☐ Comprobante de pago de las tasas establecidas (De ser el caso formato de descuento)

Completa ☐ Incompleta ☐

PATENTE DE INVENCIÓN PCT ☐ MODELO DE UTILIDAD PCT ☐ Art.33 Decisión 486/00, Circular Única

- ☐ Indicación que se solicita una PCT
- ☐ Copia de la solicitud en español, tal como fue presentada inicialmente (capítulo descriptivo, reivindicatorio, resumen)
- ☐ Dibujos de ser estos pertinentes
- ☐ Comprobante de pago de las tasas establecidas (de ser el caso formato de descuento)

Completa ☐ Incompleta ☐

DISEÑO INDUSTRIAL ☐ (Art. 119 Decisión 486/00)

- ☐ Indicación que se solicita Diseño industrial
- ☐ Datos de identificación del solicitante o de la persona que presenta la solicitud
- ☐ Representación gráfica y fotográfica del Diseño industrial o muestra del material que incorpora el diseño
- ☐ Comprobante de pago de las tasas establecidas

Completa ☐ Incompleta ☐

ESQUEMA DE TRAZADO ☐ (Art. 92 Decisión 486/00)

- ☐ Indicación que se solicita un esquema de trazado
- ☐ Datos de identificación del solicitante o de la persona que presenta la solicitud
- ☐ Representación gráfica de un esquema de trazado
- ☐ Comprobante de pago de las tasas establecidas

Completa ☐ Incompleta ☐