(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 776 144**

(51) Int. CI.:

C12Q 1/68 (2008.01)
G01N 33/50 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **26.10.2012 PCT/US2012/062263**

(87) Fecha y número de publicación internacional: **02.05.2013 WO13063496**

(96) Fecha de presentación y número de la solicitud europea: **26.10.2012 E 12844605 (1)**

(97) Fecha y número de publicación de la concesión europea: **11.12.2019 EP 2775844**

(54) Título: **Biomarcadores de respuesta a inhibidores de NAE**

(30) Prioridad:

28.10.2011 US 201161552686 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
29.07.2020

(73) Titular/es:

MILLENNIUM PHARMACEUTICALS, INC. (50.0%)
40 Landsdowne Street
Cambridge, MA 02139, US y
MASSACHUSETTS GENERAL HOSPITAL (50.0%)

(72) Inventor/es:

BENES, CYRIL, H.;
BLAKEMORE, STEPHEN, J.;
BLANK, JONATHAN, L.;
LIGHTCAP, ERIC, S.;
MULLIGAN, GEORGE, J.;
SCHU, MATTHEW, C.;
SMITH, PETER, G. y
SETTLEMAN, JEFFREY E.

(74) Agente/Representante:

ISERN JARA, Jorge

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

DESCRIPCIÓN

Biomarcadores de respuesta a inhibidores de NAE

Solicitudes relacionadas

Esta solicitud reivindica la prioridad a la Solicitud Provisional de Estados Unidos número 61/552.686 presentada el 28 de octubre de 2011.

Listado de secuencias

Esta solicitud contiene un Listado de Secuencias que se presenta aquí en formato de lectura electrónica. El archivo del Listado de Secuencias se creó el 26 de octubre de 2012, se denomina "sequencelisting.txt", y su tamaño es de 149 kb (153.088 bytes).

Antecedentes

Las células se vuelven cancerosas cuando su genotipo o fenotipo se altera de modo que existe un crecimiento descontrolado que no está sujeto a los límites del entorno normal de los tejidos. Uno o más genes están mutados, amplificados, suprimidos, sobreexpresados o subexpresados. Las partes de cromosomas se pueden perder o mover de un lugar a otro. Algunos cánceres tienen patrones característicos por los cuales se alteran los genotipos o fenotipos.

Muchos genes tienen mutaciones asociadas con cáncer. Algunos genes tienen múltiples sitios donde pueden ocurrir mutaciones. Muchos cánceres tienen mutaciones y/o expresión errónea de más de un gen. Las mutaciones genéticas pueden facilitar la progresión tumoral,, tasa de crecimiento tumoral o si un tumor hará metástasis. Algunas mutaciones pueden influir en si una célula tumoral responderá a terapia.

Una variedad de agentes tratan el cáncer. Los cánceres de sangre y médula ósea a menudo se tratan con esteroides/glucocorticoides, imids, inhibidores de proteasoma y agentes alquilantes. Los cánceres de otros tejidos a menudo se tratan con agentes alquilantes, inhibidores de topoisomerasa, inhibidores de quinasa, inhibidores de los microtúbulos, inhibidores de angiogénesis u otros agentes. Algunos pacientes responden a una terapia mejor que a otra, presentando el potencial para que un paciente siga múltiples rutas terapéuticas para una terapia eficaz. Se puede perder un tiempo valioso temprano en el programa de tratamiento de un paciente siguiendo una terapia que finalmente se demuestra ineficaz para ese paciente. Muchos pacientes no pueden permitirse el tiempo de elecciones de ensayo y error de los regímenes terapéuticos. Las decisiones de tratamiento rápidas y precisas conducen a una gestión eficaz de la enfermedad.

BASE DE DATOS GeO [en línea] ("Affymetrix GeneChip Human Genome U133 Array Set HGU133A", GeO, NCBI, (20020311), Registro en la base de datos n.º gl96, Páginas 61, 243, 418) describe una matriz con sondas para genes marcadores NF2, SMAD4 y FBXW7.

Bignell, G. R. *et al.* (2010) (Signatures of mutation and selection in the cancer genome. Nature, 463(7283), 893) describe genes involucrados en cáncer y describe el uso de cebadores para detectar mutaciones en dichos genes.

La solicitud internacional (PCT) publicada como WO 2011/068863 (A1) describe una combinación de agentes terapéuticos para uso en el tratamiento de un paciente que padece cáncer, comprendiendo la combinación al menos un agente inductor de poliploidía y al menos un inhibidor de proteínas de la familia Bel-2. También se describen ensayos de diagnóstico útiles en la clasificación de pacientes para tratamiento con uno o más agentes terapéuticos.

Sumario

La materia objeto para la que se busca protección es la definida en las reivindicaciones. La presente divulgación se refiere a pronóstico y planificación para tratamiento de tumores midiendo la cantidad, presencia o cambios de marcadores proporcionados en este documento. Los marcadores predicen si habrá un resultado favorable (por ejemplo, buena respuesta, tiempo de progresión prolongado y/o supervivencia a largo plazo) después del tratamiento con un inhibidor de la enzima activadora de NEDD8 (NAE), como un sulfamato de metilo 1-sustituido. El análisis de muestras que comprenden células tumorales, por ejemplo, *in vitro,* para determinar la presencia, cantidades o cambios de marcadores genéticos, por ejemplo, el estado mutacional de al menos un gen marcador, identifica pacientes particulares que se espera que tengan un resultado favorable con el tratamiento, por ejemplo, con un inhibidor de NAE, como un sulfamato de metilo 1-sustituido, y cuya enfermedad puede tratarse con un tratamiento estándar o menos agresivo, así como los pacientes que se espera que tengan un resultado desfavorable con el tratamiento y pueden requerir un tratamiento alternativo para, una combinación de tratamientos y/o tratamientos más agresivos con un inhibidor de NAE para asegurar un resultado favorable y/o una gestión satisfactoria de la enfermedad.

En un aspecto, la invención proporciona kits útiles para determinar características, por ejemplo, cantidades, presencia o cambios, de los marcadores. En otro aspecto, la invención proporciona métodos para determinar el pronóstico y el tratamiento o estrategias de gestión de enfermedades. En estos aspectos, se mide la característica, por ejemplo, tamaño, secuencia, composición o cantidad de marcador en una muestra que comprende células tumorales. En una realización, el tumor es un tumor líquido, por ejemplo, tumor hematológico, por ejemplo, leucemia mielógena aguda, síndrome mielodisplásico o mieloma múltiple. En otra realización, el tumor es un tumor sólido, por ejemplo, melanoma, cáncer de pulmón no microcítico, cáncer de esófago, cáncer de vejiga, cáncer neuroblastoma, mesotelioma, cáncer de páncreas.

En diversas realizaciones, se mide la característica, por ejemplo, tamaño, secuencia, composición o cantidad de ADN, el tamaño, secuencia, composición o cantidad de ARN y/o el tamaño, secuencia, composición o cantidad de proteína correspondiente a un gen marcador con una o más mutaciones, por ejemplo, mutación somática, descrito en este documento. La información útil que conduce al pronóstico o tratamiento o estrategias de gestión de la enfermedad se obtiene cuando los ensayos revelan información sobre un gen marcador, por ejemplo, si el gen está mutado o no, la identidad de la mutación, y/o si la cantidad de ARN o proteína de un gen o genes mutados indica sobreexpresión o subexpresión. En una realización, la estrategia se determina para inhibición de la enzima E1, por ejemplo, inhibición de NAE, por ejemplo, MLN4924, terapia.

Un gen marcador útil para probar la determinación del pronóstico o el tratamiento o la estrategia de gestión de la enfermedad según los métodos de la divulgación se selecciona del grupo que consiste en neurofibromina 2 (NF2), madres contra el homólogo 4 decapentapléjico (SMAD4), desmetilasa 6A específica de lisina (KDM6A), proteína tumoral p53 (TP53), inhibidor de quinasa dependiente de ciclina 2A (CDKN2A), variante p14 del inhibidor de quinasa dependiente de ciclina 2A (CDKN2A_p14), en algunos casos, dominio de repetición de F-box y WD que contiene 7 (FBXW7) y, en algunos casos, poliposis coli adenomatosa (APC). Cada gen marcador incluye mutaciones o alteraciones cuya presencia en ADN o cuyos efectos, por ejemplo, en las características del ARN marcador y/o proteínas, por ejemplo, cantidades, tamaño, secuencia o composición, puede proporcionar información para determinar el pronóstico o el tratamiento o la gestión de la enfermedad. En alguna descripción, un gen o una forma mutante o modificada del mismo útil como marcador, tiene una característica de ADN, ARN y/o proteína, por ejemplo, tamaño, secuencia, composición o cantidad, por ejemplo, en una muestra que comprende células tumorales, que es diferente a un ADN, ARN y/o proteína normal. En este documento se describen ejemplos de modificaciones de estos genes, denominados "genes marcadores" cuya mutación o cantidades pueden proporcionar dicha información.

La mutación de los marcadores de la presente invención, proporciona información sobre el resultado después del tratamiento, por ejemplo, con un inhibidor de NAE, como un sulfamato de metilo 1-sustituido. Al examinar la característica, por ejemplo, tamaño, secuencia, composición o cantidad de uno o más de los marcadores identificados en un tumor, es posible determinar qué agente terapéutico, combinación de agentes, régimen de dosificación y/o administración se espera que proporcione un resultado favorable después del tratamiento. Al examinar la característica, por ejemplo, tamaño, secuencia, composición o cantidad de uno o más de los marcadores o conjuntos de marcadores identificados en un cáncer, también es posible determinar qué agente terapéutico, combinación de agentes, régimen de dosificación y/o administración es menos probable que proporcione un resultado favorable después del tratamiento. Al examinar la característica, por ejemplo, tamaño, secuencia, composición o cantidad de uno o más de los marcadores identificados, por tanto es posible eliminar agentes o regímenes terapéuticos ineficaces o inapropiados. Importante, estas determinaciones se pueden hacer paciente por paciente. Así, se puede determinar si es probable que un régimen terapéutico en particular beneficie a un paciente o tipo de paciente en particular, y/o si se debe iniciar o evitar un régimen en particular, continuado, descontinuado o alterado.

La presente invención se refiere a métodos de identificación y/o selección de un paciente con cáncer que se espera que demuestre un resultado favorable tras la administración de un régimen terapéutico, por ejemplo, un régimen terapéutico que comprende un inhibidor de NAE, como un tratamiento con sulfamato de metilo 1-sustituido. Además se proporcionan métodos para identificar a un paciente que se espera que tenga un resultado desfavorable tras la administración de dicho régimen terapéutico. Estos métodos generalmente incluyen medir, determinar, recibir, almacenar o transmitir información sobre la característica, por ejemplo, tamaño, secuencia, composición o cantidad de uno o más marcadores o mutación de gen(es) marcador en un tumor del paciente *(por ejemplo,* células cancerosas de un paciente, por ejemplo, células cancerosas hematológicas o células tumorales sólidas), opcionalmente comparando eso con la característica, por ejemplo, tamaño, secuencia, composición o cantidad de un marcador de referencia, y en una realización adicional, identificar o aconsejar si el resultado de la muestra corresponde a un resultado favorable de un régimen de tratamiento, por ejemplo, un inhibidor de NAE, como un régimen de tratamiento con sulfamato de metilo 1-sustituido.

Además los métodos proporcionados incluyen métodos terapéuticos que incluyen además la etapa de comenzar, continuar, o comenzar una terapia en consecuencia donde la presencia de una mutación en un gen marcador o la característica, por ejemplo, tamaño, secuencia, la composición o cantidad de marcador o marcadores de un paciente indica que se espera que el paciente demuestre un resultado favorable con la terapia, *por ejemplo,* el inhibidor de NAE, como un régimen terapéutico de sulfamato de metilo 1-sustituido. Además, los métodos incluyen métodos

terapéuticos que incluyen además la etapa de parar, interrumpir, alterar o detener una terapia en consecuencia donde la presencia de una mutación en un gen marcador o la característica, por ejemplo, tamaño, secuencia, composición o cantidad del marcador de un paciente indica que se espera que el paciente demuestre un resultado desfavorable con el tratamiento, por ejemplo, con el inhibidor de NAE, como un régimen de sulfamato de metilo 1-sustituido, por ejemplo, en comparación con un paciente identificado con un resultado favorable que recibe el mismo régimen terapéutico. En otro aspecto, se proporcionan métodos para análisis de un paciente que aún no recibe tratamiento, por ejemplo, un inhibidor de NAE, como una terapia con sulfamato de metilo 1-sustituido e identificación y predicción del resultado del tratamiento basándose en la presencia de una mutación en un gen marcador o característica, por ejemplo, tamaño, secuencia, composición o cantidad de uno o más marcadores de un paciente descritos en este documento. Dichos métodos pueden incluir no ser tratado con la terapia, por ejemplo, inhibidor de NAE, como una terapia de sulfamato de metilo 1-sustituido, ser tratado con terapia, por ejemplo, inhibidor de NAE, o ser tratado con una terapia de sulfamato de metilo 1-sustituido en combinación con una terapia adicional más, ser tratado con una terapia alternativa a un inhibidor de NAE, como una terapia con sulfamato de metilo 1-sustituido, o ser tratado con un régimen de dosificación y/o administración más agresivo de una terapia, por ejemplo, Inhibidor de la enzima E1, como un inhibidor de NAE, por ejemplo, en comparación con el régimen de dosificación y/o administración de un paciente identificado con un resultado favorable al inhibidor de NAE convencional, como una terapia de sulfamato de metilo 1-sustituido. Así, los métodos proporcionados de la invención pueden eliminar el uso ineficaz o inapropiado de la terapia, por ejemplo, inhibidor de NAE, como regímenes de terapia con sulfamato de metilo 1-sustituido.

Los métodos adicionales de la divulgación incluyen métodos para determinar la actividad de un agente, la eficacia de un agente, o identificar nuevos agentes terapéuticos o combinaciones. Dichos métodos incluyen métodos para identificar un agente como útil, por ejemplo, como un inhibidor de NAE, como un sulfamato de metilo 1-sustituido, para tratar un cáncer, por ejemplo, un cáncer hematológico (por ejemplo, mieloma múltiple, leucemias, linfoma, etc.) o cáncer de tumor sólido (por ejemplo, melanoma, cáncer de esófago o cáncer de vejiga), basándose en su capacidad de influir en la presencia de una mutación en un gen marcador o característica, por ejemplo, tamaño, secuencia, composición o cantidad de un marcador o marcadores de la invención. Por ejemplo, un inhibidor que disminuye o aumenta la presencia de una mutación en un gen marcador o característica, por ejemplo, tamaño, secuencia, composición o cantidad de un marcador o marcadores proporcionados de modo que indique un resultado favorable de un paciente con cáncer sería un agente candidato para el cáncer. Alternativamente, un agente que sea capaz de disminuir la viabilidad de una célula tumoral que comprende un marcador indicativo de un resultado desfavorable sería un agente candidato para el cáncer.

La presente invención también se refiere a un inhibidor de NAE para uso en métodos de tratamiento de un paciente con cáncer, con un régimen terapéutico, por ejemplo, un inhibidor de NAE, como un régimen de terapia con sulfamato de metilo 1-sustituido (por ejemplo, solo, o en combinación con un agente adicional como un agente quimioterapéutico, por ejemplo, un agente glucocorticoide, un inhibidor del proteasoma, un agente alquilante, un inhibidor de quinasa o un inhibidor de topoisomerasa), que incluye la etapa de seleccionar para el tratamiento a un paciente cuya característica de marcador, por ejemplo, tamaño, secuencia, composición o cantidad indica que se espera que el paciente tenga un resultado favorable con el régimen terapéutico y tratar al paciente con la terapia, por ejemplo, inhibición de NAE, como una terapia de sulfamato de metilo 1-sustituido. En algunas realizaciones, el método puede incluir la etapa de seleccionar un paciente cuya característica de marcador, por ejemplo, tamaño, secuencia, composición o cantidad o cantidades indica que se espera que el paciente tenga un resultado favorable y que administre una terapia que no sea una terapia con inhibidor de NAE que demuestre tiempos de supervivencia esperados similares a los del inhibidor de NAE, como una terapia de sulfamato de metilo 1-sustituido.

Los métodos adicionales para tratar a un paciente con cáncer incluyen seleccionar pacientes que sea poco probable que experimenten un resultado favorable después del tratamiento con una terapia contra el cáncer (por ejemplo, inhibidor de NAE, como una terapia de sulfamato de metilo 1-sustituido). Dichos métodos pueden incluir además uno o más de: administrar una dosis más alta o un programa de dosificación incrementada de una terapia, por ejemplo, inhibidor de NAE, como un sulfamato de metilo 1-sustituido en comparación con la dosis o el programa de dosificación de un paciente identificado como que tiene un resultado favorable con la terapia convencional; administrar una terapia contra el cáncer que no sea un inhibidor de NAE, como una terapia de sulfamato de metilo 1-sustituido; administrar un inhibidor de NAE, como un agente de sulfamato de metilo 1-sustituido en combinación con un agente adicional. Además se proporcionan métodos para seleccionar un paciente con enfermedad agresiva que se espera que demuestre un tiempo más rápido para progresión y muerte.

Los métodos adicionales de la divulgación incluyen un método para evaluar si tratar o pagar el tratamiento del cáncer, por ejemplo, cáncer hematológico (por ejemplo, mieloma múltiple, leucemias, linfoma, *etc.)* o cáncer de tumor sólido (por ejemplo, melanoma, cáncer de esófago o de vejiga) revisando la cantidad de marcador o marcadores de un paciente para indicar el resultado de una terapia contra el cáncer, por ejemplo, un inhibidor de NAE, como un régimen de terapia con sulfamato de metilo 1-sustituido, y tomar una decisión o asesorar sobre si se debería realizar el pago.

Se hace referencia al contenido completo de todas las publicaciones, solicitudes de patente, patentes y otras referencias mencionadas en este documento.

Otras características y ventajas de la invención serán evidentes a partir de la siguiente descripción detallada, dibujos y reivindicaciones.

Dibujos

Figura 1. Estructura general de sulfamato de metilo 1-sustituido. G^1 es -O- o $-CH_2-$; G^2 es -H u -OH; G^3 es -H u -OH; G^4 es -NH-, -O- o un enlace covalente; y G^5 es heteroarilo sustituido.

Figura 2. Rutas generales para ubiquitinación de sustratos proteicos con ligasa culina-RING (CRL) y para nedilación. En las CRL, la subunidad de culina debe modificarse en una lisina conservada por la proteína NEDD8 similar a ubiquitina para activar la actividad de holoenzima. La activación y conjugación de NEDD8 con proteínas de culina se cataliza mediante una cascada enzimática que es homóloga a la ubiquitinación que involucra E1 (NAE) y E2 (Ubc12) de NEDD8. La eliminación de NEDD8 de la culina es catalizada por el señalosoma COP9. La desnedilación facilita la disociación de los componentes de CRL. El núcleo de culina-RING se secuestra en un estado inactivo uniéndose a CAND1 hasta que se recluta para formar una nueva CRL.

Figura 3. Respuesta de un panel 2 de línea celular a MLN4924. Cada punto representa una línea celular.

Figuras 4A-B. Comparación de respuestas de paneles de líneas celulares a MLN4924. A. Ordenación del panel 2 de línea celular por CE50. Las líneas oscuras representan líneas celulares que están presentes en el panel 1. En ambos paneles hay 114 líneas celulares con nombres idénticos. B. Comparación de la viabilidad del porcentaje de control (POC) para las líneas celulares que están presentes en ambos paneles. Los resultados de las líneas celulares superpuestas tienen una correlación de orden de rango de Spearman de 0,72, valor $p < 2,2e-16$.

Figura 5. Asociación tisular de resistencia con mutaciones TP53. Las líneas celulares de cáncer de colon mutante TP53 son más resistentes (mayor porcentaje de viabilidad de control) al tratamiento MLN4924 que las líneas celulares wt TP53.

Figuras 6A-D. Efecto de la pérdida de TP53 en la viabilidad de líneas celulares de cáncer después del tratamiento con diferentes dosis de MLN4924 en múltiples puntos temporales. El efecto de la supresión de TP53 en la sensibilidad de líneas celulares de cáncer de colon HCT-116 emparejado a MLN4924 se midió con ATPlite, en un intervalo de concentraciones MLN4924 y puntos temporales. Los datos se representan como media ± ETM, N = 3. Línea discontinua, supresión de p53; línea continua, p53 tipo silvestre.

Descripción detallada

Uno de los problemas continuos con la terapia en pacientes con cáncer son las diferencias individuales de la respuesta a las terapias. Aunque los avances en el desarrollo de terapias exitosas contra el cáncer avanzan, solo un subconjunto de pacientes responde a una terapia particular. Con el índice terapéutico estrecho y el potencial tóxico de muchas terapias contra el cáncer disponibles, tales respuestas diferenciales contribuyen potencialmente a que los pacientes se sometan innecesariamente, a regímenes de terapia ineficaces e incluso potencialmente dañinos. Si una terapia diseñada pudiera optimizarse para tratar pacientes individuales, tales situaciones podrían reducirse o incluso eliminarse. Además, la terapia diseñada dirigida puede proporcionar terapia satisfactoria general para el paciente. En consecuencia, existe la necesidad de identificar pacientes con cáncer particulares que se espera que tengan un resultado favorable cuando se administran terapias contra el cáncer particular, así como pacientes con cáncer particular que pueden tener un resultado favorable utilizando terapias contra el cáncer más agresivas y/o alternativas, por ejemplo, alternativa a terapias previas contra el cáncer administradas al paciente. Por tanto, sería beneficioso proporcionar el diagnóstico, estadio, pronóstico y monitorización de pacientes con cáncer, incluyendo, por ejemplo, pacientes con cáncer hematológico (por ejemplo, mieloma múltiple, leucemias, linfoma, *etc.)* o cáncer de tumor sólido (por ejemplo, melanoma, cáncer de esófago o cáncer de vejiga) que se beneficiarían de terapias de inhibición del cáncer particulares, así como aquellos que se beneficiarían de una terapia de inhibición del cáncer más agresiva y/o alternativa, por ejemplo, alternativa a una terapia o terapias contra el cáncer que el paciente ha recibido, dando como resultado medidas preventivas apropiadas.

La presente invención se basa, en parte, en el reconocimiento de que la mutación de un gen marcador puede estar asociada con la sensibilidad de una célula que comprende el gen mutado a un inhibidor de NAE, como un sulfamato de metilo 1-sustituido. En algunas realizaciones, el gen marcador está involucrado en la ruta de la ligasa culina-RING (CRL), por ejemplo, un gen cuya proteína codificada interactúa con una CRL o una proteína asociada a CRL, o es un sustrato de CRL. Una proteína codificada por un gen marcador puede tener una función de tipo silvestre como supresor tumoral. Los ejemplos de genes marcadores incluyen NF2, SMAD4 y/o KDM6A. Otros ejemplos de genes marcadores incluyen TP53, APC, CDKN2A y/o CDKN2A_p14. Los genes marcadores pueden exhibir mutaciones, por ejemplo, mutaciones somáticas, cuya presencia puede afectar la expresión o actividad del producto génico codificado. En algunas realizaciones, puede haber más de una mutación en un gen marcador o más de un gen marcador con una mutación en una célula tumoral o tumor. En realizaciones adicionales, puede haber mutaciones del gen marcador en células que tienen mutaciones en genes adicionales, incluyendo mutaciones que pueden conducir a tumorigénesis, pero los genes mutados adicionales pueden no ser genes marcadores como se considera en este documento. En algunas realizaciones, la mutación es una mutación inactivadora. En otras realizaciones, la mutación afecta la expresión del gen marcador. En otras realizaciones, una mutación puede provocar una interacción alterada del producto génico codificado con un compañero de unión celular. La identificación y/o medición de la mutación en el gen marcador se puede usar para determinar si se puede esperar un resultado

favorable por el tratamiento de un tumor, por ejemplo, con un inhibidor de NAE, como una terapia con sulfamato de metilo 1-sustituido o si se trata de una terapia alternativa y/o una terapia más agresiva con, por ejemplo, un inhibidor de NAE, como un inhibidor de sulfamato de metilo 1-sustituido puede mejorar el tiempo de supervivencia esperado. Por ejemplo, las composiciones y métodos proporcionados en este documento pueden usarse para determinar si se espera que un paciente tenga un resultado favorable a un inhibidor de NAE, como un agente terapéutico de sulfamato de metilo 1-sustituido o un inhibidor de NAE, como un régimen de administración o dosificación de sulfamato de metilo 1-sustituido. En general, la mutación en los genes marcadores supresores de tumores descritos en este documento está asociada con la sensibilidad o el resultado favorable del tratamiento con un inhibidor de NAE. Los ejemplos de genes marcadores que pueden funcionar como un supresor tumoral en brutas relacionadas con la ligasa culina-RING y cuya mutación está asociada con la sensibilidad a la inhibición de NAE incluyen NF2, SMAD4, KDM6A, FBXW7, CDKN2A y/o CDKN2A_p14. Sin embargo, TP53 y APC también son genes marcadores supresores tumorales. En particular, los genes de la ruta TP53 están asociados con efectos inhibidores de NAE. Como se describe en este documento, en algunas realizaciones para muchos tipos de tumores, la mutación en TP53, y en algunos casos, APC, conduce a resistencia a la inhibición de NAE. En consecuencia, un gen marcador de tipo silvestre del grupo que consiste en TP53 y APC puede asociarse con sensibilidad a NAE. En algunas realizaciones, la mutación de un gen marcador seleccionado entre el grupo que consiste en TP53 y APC está asociada con la resistencia a un inhibidor de NAE.

Basándose en estas identificaciones, la presente divulgación proporciona, sin limitación: 1) métodos y composiciones para determinar si un inhibidor de NAE, como un régimen de terapia con sulfamato de metilo 1-sustituido será o no eficaz para lograr un resultado favorable y/o controlar el cáncer; 2) métodos y composiciones para monitorizar la eficacia de un inhibidor de NAE, como una terapia de sulfamato de metilo 1-sustituido (solo o en una combinación de agentes) y dosificación y administraciones usadas para el tratamiento de tumores; 3) métodos y composiciones para tratamientos de tumores que comprenden, por ejemplo, inhibidor de NAE, como un régimen de terapia de inhibición de sulfamato de metilo 1-sustituido; 4) métodos y composiciones para identificar agentes terapéuticos específicos y combinaciones de agentes terapéuticos, así como regímenes de dosificación y administración que son eficaces para el tratamiento de tumores en pacientes específicos; y 5) métodos y composiciones para identificar estrategias de gestión de enfermedades.

La ubiquitina y otras moléculas similares a la ubiquitina (ubls) se activan mediante una enzima específica (una enzima E1) que cataliza la formación de un intermedio de acil-adenilato con la glicina C-terminal de la ubl. La ubl activada se transfiere luego a un resto de cisteína catalítico dentro de la enzima E1 mediante la formación de un enlace intermedio de tioéster. El intermedio E1-ubl y un E2 asociado, dan como resultado un intercambio de tioéster en donde la ubl se transfiere a la cisteína de sitio activo del E2. La ubl se conjuga con la proteína diana, directamente o en combinación con una ligasa E3, mediante la formación de enlaces isopeptídicos con el grupo amino de una cadena lateral de lisina en la proteína diana. La ubl denominada precursor neuronal expresado por células de desarrollo desregulado 8 (NEDD8) se activada con la enzima heterodimérica activadora de NEDD8 (NAE, también conocida como APPBP1-UBA3, UBE1C (enzima activadora de ubiquitina E1C)) y se transfiere a una de las dos enzimas conjugadoras de E2 (proteína transportadora de ubiquitina 12 (UBC12) y UBC17), que finalmente da como resultado el ligamiento de NEDD8 a proteínas culina por el subtipo de ubiquitina ligasas Culina-RING (véase Figura 2). Una función de la nedilación es la activación de ubiquitina ligasas basadas en culina involucradas en la renovación de muchas proteínas del ciclo celular y de señalización celular, incluyendo p27 e I-κB. Véase Pan *et al.,* Oncogene 23:1985-97 (2004). La inhibición de NAE puede alterar el recambio proteico mediado por ligasa culina-RING y puede conducir a la muerte apoptótica en células, por ejemplo, células tumorales o células de un organismo patógeno, por ejemplo, un parásito. Véase Soucy *et al*. (2010) Genes & Cancer 1:708-716.

Como se usa en este documento, el término "E1", "Enzima E1", o "enzima activadora E1" se refiere a cualquiera de una familia de enzimas activadoras dependientes de ATP relacionadas involucradas en la activación o promoción de la conjugación a ubiquitina o similar a ubiquitina (colectivamente "ubl") a moléculas diana. Las enzimas activadoras de E1 funcionan mediante una formación intermedia de adenilación/tioéster para transferir la ubl apropiada a la enzima conjugadora de E2 respectiva mediante una reacción de transtiolación. La ubl-E2 activada resultante favorece la conjugación final de la ubl a una proteína diana. Una variedad de proteínas celulares que desempeñan un papel en señalización celular, ciclo celular y recambio proteico son sustratos para conjugación de ubl que se regula mediante enzimas activadoras E1 (por ejemplo, NAE, EAU, SAE). A menos que el contexto indique lo contrario, el término "enzima E1" se refiere a cualquier proteína enzimática activadora E1, incluyendo, sin limitación, enzima activadora de NEDD8 (NAE (APPBP1/Uba3)), enzima activadora de ubiquitina (UAE (Uba1)), enzima activadora de sumo (SAE (Aos1/Uba2)), UBA4, UBA5, UBA6, ATG7 o enzima activadora de ISG15 (Ube1L).

La expresión "inhibidor enzimático E1" o "inhibidor de la enzima E1" se usa para significar un compuesto que tiene una estructura como se define en este documento, que es capaz de interactuar con una enzima E1 e inhibir su actividad enzimática. Inhibir la actividad enzimática E1 significa reducir la capacidad de una enzima E1 para activar la conjugación como ubiquitina (ubl) a un péptido o proteína sustrato (por ejemplo, ubiquitinación, nedilación, sumoilación). En algunas realizaciones, un inhibidor de la enzima E1 puede inhibir más de una enzima E1. En otras realizaciones, un inhibidor de la enzima E1 es específico para una enzima E1 particular. En diversas realizaciones, tal reducción de la actividad enzimática E1 es al menos aproximadamente 50 %, al menos aproximadamente 75 %, al menos aproximadamente 90 %, al menos aproximadamente 95 %, o al menos aproximadamente 99 %. En

diversas realizaciones, la concentración de inhibidor de enzima E1 requerida para reducir una actividad enzimática E1 es inferior a aproximadamente 1 µM, inferior a aproximadamente 500 nM, inferior a aproximadamente 100 nM, inferior a aproximadamente 50 nM, o inferior a aproximadamente 10 nM.

Como se usa en este documento, la expresión "inhibidor de NAE" se refiere a un inhibidor del heterodímero de NAE. Los ejemplos de inhibidores de NAE incluyen sulfamatos de metilo 1-sustituidos (véase Figura 1), incluyendo MLN4924. Langston S. *et al.* Solicitud de patente de Estados Unidos n.º serie 11/700.614, cuya solicitud PCT se publicó como WO07/092213, WO06084281 y WO2008/019124, desvelan compuestos que son inhibidores eficaces de enzimas activadoras de E1, por ejemplo, NAE. En algunas realizaciones, los inhibidores de NAE no inhiben, o no son buenos inhibiendo, otras enzimas E1 (no NAE). Los compuestos son útiles para inhibir la actividad E1 *in vitro* e *in vivo* y son útiles para el tratamiento de trastornos de proliferación celular, por ejemplo, cáncer y otros trastornos asociados con la actividad E1, como infecciones patógenas y trastornos neurodegenerativos. Una clase de compuestos descritos en Langston *et al*., son sulfamatos de ((IS, 2S, 4*R*)-2-hidroxi-4-{7H-pirrolo[2,3-d]pirimidin-7-il}ciclopentil)metilo 4-sustituidos. MLN4924 (sulfamato de ((1S,2S,4R)-4-{4-[(1S)-2,3-dihidro-1H-inden-1-ilamino]-7H-pirrolo[2,3-d]pirimidin-7-il}-2-hidroxiciclopentil)metilo) es un inhibidor E1 específico de NAE que interrumpe el recambio proteico mediado por ligasa culina-RING que conduce a la muerte apoptótica en células tumorales humanas por la perturbación de la homeostasis de la proteína celular (Soucy *et al*. (2009) Nature 458:732-736).

La evaluación de MLN4924 en estudios de xenoinjerto celular y tumoral ha revelado dos mecanismos de acción distintos. El primero es la inducción de la re-replicación del ADN, daño en el ADN y muerte celular mediante la desregulación mediada por MLN4924 del sustrato Cdt-1 de $CRL1^{SKP2}$ y $CRL4^{DDB1}$ (Milhollen *et al.* (2011) Cancer Res. 71:3042-3051). Se ha demostrado que el estado de p53 no influye en la inducción de re-replicación del ADN, pero puede hacer que las células sean más propensas a sufrir apoptosis o senescencia, dependiendo del contexto genético apropiado (Milhollen *et al.* (2011) *supra,* Lin *et al*. (2010) Nature 464:374-379 y Lin *et al*. (2010) Cancer Res.70:10310-20). El segundo mecanismo es la inhibición de la actividad de la ruta NF-κB en linfomas difusos de linfocitos B grandes dependientes de NF-κB principalmente mediante la desregulación de la renovación mediada por $CRL1^{\beta TRCP}$ de IκBα fosforilado (Milhollen *et al*. (2010) Blood 116:1515-1523). Además, los modelos preclínicos de leucemia mielógena aguda (AML) son sensibles a la inhibición de MLN4924 tanto en líneas celulares como en blastos primarios de pacientes mediante mecanismos relacionados con la desregulación de Cdt-1, inhibición e inducción de NF-κB de especies reactivas de oxígeno (Swords *et al*. (2010) Blood 115:3796-3800).

Genes como NF2 (revisado por Ahronowitz *et al*. (2007) Human Mutation 28:1-2), KDM6A (revisado por van Haaften *et al*. (2009) Nat. Genet. 41:521-523), FBXW7, TP53, CDKN2A y CDKN2A_p14 están mutados en muchos tipos de cáncer. SMAD4 está mutado en varios tipos de cáncer, pero muchas de las mutaciones SMAD4 se encuentran en cánceres de intestino, páncreas (revisado por Miyaki y Kuroki (2003) Biochem. Biophys. Res. Commun. 306:799-804) o glándula tiroides.

Como se usa en este documento, "NF2" se refiere a la isoforma más larga del gen asociado con el número de acceso GenBank NM_000268, SEQ ID NO: 1 (marco de lectura abierto es SEQ ID NO: 2, nucleótidos 444 a 2231 de SEQ ID NO: 1), codificando el número de acceso GenPept NP_000259, SEQ ID NO: 3). Otros nombres para NF2 incluyen ACN, BANF, SCH y merlina (proteína similar a moesinezrina-radixina). NF2 funciona como gen supresor tumoral y se puede encontrar en el cromosoma 22. NF2 interactúa con el citoesqueleto, proteínas de superficie celular y puede estar involucrada en dinámica citoesquelética y regulación del transporte iónico. Las funciones de NF2 que pueden relacionarlo con sensibilidad a la inhibición de NAE, por ejemplo, MLN4924 incluyen su capacidad para inhibir la ubiquitina ligasa E3 $CRL4^{DCAF1}$ (Li *et al*. (2010) Cell 140:477-490). Las mutaciones en NF2 pueden alterar su actividad inhibitoria y conducir a ubiquitinación incontrolada de sustratos de $CRL4^{DCAF1}$ y proliferación de células que albergan el gen mutado.

Como se usa en este documento, "SMAD4" se refiere al gen asociado con el número de acceso GenBank NM_005359, SEQ ID NO: 4 (marco de lectura abierto es SEQ ID NO: 5, nucleótidos 539 a 2197 de SEQ ID NO: 4), codificando el número de acceso GenPept NP_005350, SEQ ID NO: 6. Otros nombres para SMAD4 incluyen eliminado en el locus 4 de carcinoma pancreático (DPC4), JIP, o madres contra decapentapléjico, Drosophila, homólogo de, 4 (MAD4). SMAD4 es una proteína de transducción de señales implicada en la señalización del factor de crecimiento transformante (TGF)-beta. SMAD4 puede actuar como supresor tumoral y puede ser bien de degradación por ubiquitinación con el complejo proteico de caja Skp-Culina-F (SCF).

Como se usa en este documento, "KDM6A" se refiere al gen asociado con el número de acceso GenBank NM_021140, SEQ ID NO: 7 (marco de lectura abierto es SEQ ID NO: 8, nucleótidos 376 a 4581 de SEQ ID NO: 7, o SEQ ID NO: 9), codificando el número de acceso GenPept NP_066963, SEQ ID NO: 10 o SEQ ID NO: 11 (SEQ ID NO: 10 con una V en lugar de una L en la posición 173 y una R en lugar de L en 584, una N en lugar de S en la posición 601 y/o K en lugar de E en la posición 629). Otros nombres para KDM6A incluyen gen TPR ligado a proteína X de repetición tetratricopeptídica transcrito ubicuamente o transcrito ubicuamente en el cromosoma X (UTX), o bA286N14.2. KDM6A es una histona desmetilasa y puede funcionar como supresor tumoral.

Como se usa en este documento, "FBXW7" se refiere al gen asociado con el número de acceso GenBank NM_033632, SEQ ID NO: 12 (marco de lectura abierto es SEQ ID NO: 13, nucleótidos 150 a 2273 de SEQ ID NO:

12), que codifica el número de acceso GenPept NP_361014, SEQ ID NO: 14. Otros nombres para FBXW7 incluyen homólogo de *C elegans* sel-10 (SEL10), homólogo de archipiélago (AGO), proteína FBX30 de caja F (FBXO30), o proteína 4 de control de división celular (CDC4). FBXW7 puede asociarse en un complejo de ubiquitina proteína ligasa para participar en la ubiquitinación de proteínas dependiente de fosforilación, incluyendo proteínas involucradas en ciclo celular y supervivencia. FBXW7 puede actuar como supresor tumoral. El uso de FBXW7 como gen marcador puede ser específico de órgano, es decir, puede ser un marcador de sensibilidad en tumores que surgen en algunos tejidos pero no en otros. Por ejemplo, FBXW7 puede ser un marcador de sensibilidad en tumores del útero, cuello uterino o hígado, pero no es un marcador de sensibilidad en tumores del tracto digestivo, donde las mutaciones en otros genes pueden dominar para dar lugar a insensibilidad o resistencia de células de esos tumores a MLN4924.

Como se usa en este documento, "TP53" se refiere al gen asociado con el número de acceso GenBank NM_000546, SEQ ID NO: 15 (marco de lectura abierto es SEQ ID NO: 16, nucleótidos 203 a 1384 de SEQ ID NO: 15, o una variante en donde el nucleótido en la posición 417 es una guanina en lugar de una citosina), codificando el número de acceso GenPept NP_000537, SEQ ID NO: 17 o una variante en donde el resto de aminoácido en la posición 72 es una arginina, R en lugar de una prolina, P). Otros nombres para TP53 incluyen BCC7, LFS1 y p53. TP53 se une al ADN y activa factores de transcripción y puede funcionar como supresor tumoral.

Como se usa en este documento, "CDKN2A" se refiere al gen asociado con el número de acceso GenBank NM_000077, SEQ ID NO: 18 (marco de lectura abierto es SEQ ID NO: 19, nucleótidos 307 a 777 de SEQ ID NO: 18), codificando el número de acceso GenPept NP_000068, SEQ ID NO: 20. Otros nombres para CDKN2A incluyen marco de lectura abierto alternativo (ARF), p16, p16ARF, inhibidor de la quinasa 4 dependiente de ciclina (INK4) y gen-1 supresor tumoral múltiple (MTS1). Las variantes de CDKN2A difieren en el primer exón. Una variante es "CDKN2A_p14" o "CDKN2A.p14", también conocido como p14ARF, está asociado con el número de acceso GenBank NM_058195, SEQ ID NO: 21 (marco de lectura abierto es SEQ ID NO: 22, nucleótidos 38 a 559 de SEQ ID NO: 21); GenPept NP_478102, SEQ ID NO: 23 o una variante que comienza en el resto de aminoácido 42 de SEQ ID NO: 23. CDKN2A_p14 resulta de la traducción en un marco de lectura diferente que p16ARF (p16INK4a, CDKN2A). CDKN2A y CDKN2A_p14 inhiben la quinasa 4 dependiente de ciclina, puede estabilizar p53 y puede regular la progresión del ciclo celular G1. CDKN2A y CDKN2A_p14 pueden actuar como supresor tumoral.

Como se usa en este documento, "APC" se refiere a poliposis adenomatosa coli, el gen asociado con el número de acceso GenBank NM_000038, SEQ ID NO: 24 (marco de lectura abierto SEQ ID NO: 25, o una variante con una timina en lugar de una citosina en el nucleótido 1458), que codifica el número de acceso GenPept NP_000029, SEQ ID NO: 26. Otros nombres para APC incluyen BTSP2 y DP2. APC se une a los microtúbulos e inhibe la ruta de señalización de Wnt y puede funcionar como supresor tumoral.

Ha habido interés en la catalogación pública de mutaciones asociadas con cánceres. Los ejemplos de bases de datos públicas que incluyen información sobre mutaciones asociadas con cánceres son la Base de Datos de Genotipos y Fenotipos (dbGaP) mantenida por el Centro Nacional de Información Biotecnológica (Bethesda, MD) y la base de datos del Catálogo de Mutaciones Somáticas en Cáncer (COSMIC) mantenida por Wellcome Trust Sanger Institute (Cambridge, Reino Unido).

Se proporcionan composiciones y métodos para determinar el estado mutacional, por ejemplo, para identificar mutaciones en genes marcadores en tumores hematológicos (por ejemplo, mieloma múltiple, leucemias, linfoma, *etc.)* o sólido (por ejemplo, melanoma, cáncer de esófago, cáncer de pulmón o de vejiga) para predecir la respuesta al tratamiento, tiempo hasta progresión y supervivencia después del tratamiento. Las composiciones y los métodos proporcionados en este documento también pueden identificar mutaciones en genes marcadores en tumores sólidos como cáncer de colon, cáncer de mama, cáncer de cabeza y cuello, o cáncer del sistema nervioso central.

Los marcadores se identificaron en función de los perfiles genéticos de células tumorales que muestran sensibilidad al tratamiento para MLN4924. El marcador TP53 también se identificó en función del comportamiento de líneas celulares isogénicas que difieren en la deleción del gen TP53. La sensibilidad observada puede ser coherente entre células tumorales probadas con más de un método.

A menos que se defina lo contrario, todos los términos técnicos y científicos utilizados en este documento tienen los significados que comúnmente entiende un experto en la materia a la que pertenece esta invención. Generalmente, la nomenclatura utilizada en relación con, y técnicas de cultivo celular y tisular, biología molecular y la química e hibridación de proteínas y oligo- o polinucleótidos descritas en este documento son las conocidas en la técnica. Los números de acceso GenBank o GenPept y las secuencias útiles de ácidos nucleicos y péptidos se pueden encontrar en el sitio web mantenido por el Centro Nacional de Información Biotecnológica, Bethesda, MD. Se menciona el contenido de todos los registros de acceso a la base de datos (por ejemplo, de los archivos de anotaciones Affymetrix HG133, Entrez, GenBank, RefSeq, COSMIC) mencionado en esta solicitud (incluyendo las Tablas). Se usan técnicas convencionales para ADN recombinante, síntesis de oligonucleótidos, purificación de proteínas, cultivo tisular y transformación y transfección (por ejemplo, electroporación, lipofección, etc.). Las reacciones enzimáticas se realizan según las especificaciones del fabricante o como se logra comúnmente en la técnica o como se describe aquí. Las técnicas y procedimientos anteriores generalmente se realizan según métodos conocidos en la técnica, por

ejemplo, como se describe en varias referencias generales y más específicas que se citan y discuten a lo largo de la presente memoria descriptiva. Véanse, por ejemplo, Sambrook *et al*. (2000) Molecular Cloning: A Laboratory Manual (3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) o Harlow, E. y Lane, D. (1988) Antibodies: A Laboratory Manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY). Las nomenclaturas utilizadas en relación con, y los procedimientos y técnicas de laboratorio de, química analítica, química orgánica sintética y química medicinal y farmacéutica descritas en este documento se conocen en la técnica. Para síntesis químicas, análisis químicos, preparación farmacéutica, formulación y administración, y tratamiento de pacientes se usan técnicas convencionales. Además, a no ser que el contexto requiera otra cosa, los términos en singular incluirán las pluralidades, y los términos en plural incluirán el singular. En caso de conflicto, la presente memoria descriptiva, incluyendo las definiciones, prevalecerá.

Los artículos "un", "uno" y "al menos uno" se usan en el presente documento para referirse a uno o más de uno de los objetos gramaticales del artículo. A modo de ejemplo, "un elemento" significa uno o más de un elemento, al menos un elemento. En caso de conflicto, la presente memoria descriptiva, incluyendo las definiciones, prevalecerá.

Como se usa en este documento, un resultado o pronóstico "favorable" se refiere a supervivencia a largo plazo, largo tiempo hasta progresión (TTP) y/o buena respuesta. Por el contrario, un pronóstico "desfavorable" se refiere a la supervivencia a corto plazo, corto tiempo de progresión (TTP) y/o mala respuesta.

Un "marcador" como se usa en este documento, incluye un material asociado con un gen marcador que se ha identificado que tiene una mutación en las células tumorales de un paciente y, además, esa mutación es característica de un paciente cuyo resultado es favorable o desfavorable con el tratamiento, por ejemplo, con un inhibidor de NAE, como un sulfamato de metilo 1-sustituido. Los ejemplos de un marcador incluyen un material, por ejemplo, un locus cromosómico, ADN para un gen, ARN para un gen o proteína para un gen. Por ejemplo, un marcador incluye un material génico marcador, por ejemplo, un locus cromosómico, ADN, ARN o proteína que demuestra una característica, por ejemplo, tamaño, secuencia, composición o cantidad indicativa de un paciente de supervivencia a corto plazo; alternativamente, un marcador incluye un material genético marcador, por ejemplo, un locus cromosómico, ADN, ARN o proteína que demuestra una mutación o característica, por ejemplo, tamaño, secuencia, composición o cantidad indicativa de un paciente de supervivencia a largo plazo. En otro ejemplo, un marcador incluye un material génico marcador, por ejemplo, un locus cromosómico, ADN, ARN o proteína cuya mutación o característica, por ejemplo, tamaño, secuencia, composición o cantidad es indicativa de un paciente con una mala respuesta al tratamiento; alternativamente, un marcador incluye un material genético marcador, por ejemplo, un locus cromosómico, ADN, ARN o proteína cuya mutación o característica, por ejemplo, tamaño, secuencia, composición o cantidad es indicativo de un paciente con una buena respuesta. En un ejemplo adicional, un marcador incluye un material génico marcador, por ejemplo, un locus cromosómico, ADN, ARN o proteína cuya mutación o característica, por ejemplo, tamaño, secuencia, composición o cantidad es indicativa de un paciente cuya enfermedad tiene un corto tiempo hasta progresión (TTP) después del tratamiento; alternativamente, un marcador incluye un material genético marcador, por ejemplo, un locus cromosómico, ADN, ARN o proteína cuya mutación o característica, por ejemplo, tamaño, secuencia, composición o cantidad es indicativa de un paciente cuya enfermedad tiene un TTP largo. Además en otro ejemplo, un marcador incluye un material génico marcador, por ejemplo, un locus cromosómico, ADN, ARN o proteína cuya mutación o característica, por ejemplo, tamaño, secuencia, composición o cantidad es indicativa de un paciente cuya enfermedad tiene una supervivencia a corto plazo después del tratamiento; alternativamente, un marcador incluye un material genético marcador, por ejemplo, un locus cromosómico, ADN, ARN o proteína cuya mutación o característica, por ejemplo, tamaño, secuencia, composición o cantidad es indicativo de un paciente cuya enfermedad tiene una supervivencia a largo plazo. Así, como se usa en este documento, el marcador pretende incluir todas y cada una de estas posibilidades, y además puede incluir cada marcador individualmente como marcador; o alternativamente puede incluir una o más, o todas las características colectivamente cuando se hace referencia a "marcadores" o "conjuntos de marcadores".

Un marcador de locus cromosómico útil para medir la determinación del pronóstico o tratamiento o estrategia de gestión de la enfermedad se selecciona del grupo que consiste en el cromosoma 22q12.2 (NF2), por ejemplo, del par de bases 29999545 a 30094589, cromosoma 18q21.1-21.2 (SMAD4), por ejemplo, del par de bases 48556583 a 48611412, cromosoma Xp11.2 (KDM6A), por ejemplo, del par de bases 44732423 a 44971847, cromosoma 4q31.3 (FBXW7), por ejemplo, del par de bases 153242410-153456172, cromosoma 17p13.1 (TP53), por ejemplo, del par de bases 7571720 a 7590868, y 9p21 (CDKN2A y CDKN2A_p14), por ejemplo, del par de bases 21967751 a 21994490. Los números de locus cromosómicos se basan en el genoma humano de referencia Build 37.3 (actual el 5 de octubre, 2011) en la base de datos de genes NCBI. Un marcador de ADN, marcador de ARN o proteína marcadora pueden corresponder a pares de bases en un marcador de locus cromosómico. Por ejemplo, un ADN marcador puede incluir ADN genómico de un marcador de locus cromosómico, el ARN marcador puede incluir un polinucleótido transcrito a partir de un marcador de locus, y una proteína marcadora puede incluir un polipéptido resultante de expresión en un marcador de locus cromosómico en una muestra, por ejemplo, que comprende células tumorales.

Un "ácido nucleico marcador" es un ácido nucleico (por ejemplo, ADN genómico, ARNm, ADNc) codificado por o correspondiente a un gen marcador de la invención. Dichos ácidos nucleicos marcadores incluyen ADN, por ejemplo, hebras de ADN genómico sentido y antisentido (por ejemplo, incluyendo cualquier intrón que ocurra allí), que

comprende la secuencia completa o parcial, por ejemplo, uno o más de los exones del ADN genómico, hasta e incluyendo el marco de lectura abierto de cualquiera de los genes marcadores o el complemento de dicha secuencia. Los ácidos nucleicos marcadores también incluyen ARN que comprende la secuencia completa o parcial de cualquier marcador o el complemento de dicha secuencia, en donde todos los restos de timidina se reemplazan con restos de uridina, ARN generado por transcripción de ADN genómico (es decir, antes de empalme), ARN generado por empalme de ARN transcrito a partir de ADN genómico, y proteínas generadas por traducción de ARN empalmado (es decir, incluyendo proteínas tanto antes como después de la escisión de regiones normalmente escindidas como secuencias de señal transmembrana). Como se usa en este documento, un "ácido nucleico marcador" también puede incluir un ADNc hecho por transcripción inversa de un ARN generado por transcripción de ADN genómico (incluyendo ARN empalmado). Un ácido nucleico marcador también incluye secuencias que difieren, debido a la degeneración del código genético, de la secuencia de nucleótidos de ácidos nucleicos que codifican una proteína que corresponde a un marcador, por ejemplo, un marcador mutado, de la invención, y por tanto codifican la misma proteína, por ejemplo, proteína mutada. Como se usa en este documento, la expresión "variante alélica" se refiere a una secuencia de nucleótidos que ocurre en un locus dado o a un polipéptido codificado por la secuencia de nucleótidos. Dichas variaciones alélicas que ocurren naturalmente pueden dar como resultado una variación del 1-5 % en la secuencia de nucleótidos de un gen dado. Se pueden identificar alelos alternativos secuenciando el gen de interés en varios individuos diferentes, por ejemplo, en células, por ejemplo, células germinales, de individuos sin cáncer. Esto puede realizarse fácilmente mediante el uso de sondas de hibridación para identificar el mismo locus genético en una variedad de individuos. La detección de cualquiera o todas esas variaciones de nucleótidos y polimorfismos de aminoácidos resultantes o variaciones que son el resultado de la variación alélica natural y que no alteran la actividad funcional de un gen marcador de tipo silvestre está dentro del alcance de la versión de tipo silvestre de un marcador descrito en este documento. Una "proteína marcadora" es una proteína codificada por o correspondiente a un marcador, por ejemplo, un ácido nucleico mutante, de la invención. Los términos "proteína" y "polipéptido" se usan indistintamente. Una proteína de un marcador puede ser denominada específicamente por su nombre o secuencia de aminoácidos, pero los expertos en la materia entienden, que mutaciones, deleciones y/o modificaciones postraduccionales pueden afectar la estructura de la proteína, aspecto, ubicación celular y/o comportamiento. A menos que se indique otra cosa, tales diferencias no se distinguen en este documento, y un marcador descrito en este documento pretende incluir cualquiera o todas esas variedades.

Como se usa en este documento, un "gen marcador" se refiere a un gen que puede tener una mutación tal que su ADN, ARN y/o proteína tiene una característica, por ejemplo, tamaño, secuencia, composición o cantidad(es) que proporcionan información sobre el pronóstico (es decir, son "informativos") después del tratamiento. Los genes marcadores descritos en este documento como vinculados con el resultado después del inhibidor de NAE, como tratamiento con sulfamato de metilo 1-sustituido (por ejemplo, MLN4924) son ejemplos de genes dentro de los marcadores de locus cromosómicos descritos anteriormente y se proporcionan en la Tabla 1. Secuencias de ARNm, marcos de lectura abiertos y las proteínas correspondientes a los genes marcadores también se enumeran en la Tabla 1. Un gen marcador enumerado en la Tabla 1 puede tener isoformas que son ubicuas o tienen expresión restringida. Excepto por el listado por separado de la isoforma CDKN2A_p14, las SEQ ID NOs de ADN en la Tabla 1 se refieren al ARNm que codifica la isoforma mayor o más larga y las SEQ ID NOs de proteína representan al menos un precursor de dicha isoforma y no necesariamente la proteína madura. Estas secuencias no pretenden limitar la identidad del gen marcador a esa isoforma o precursor. Las isoformas adicionales y proteínas maduras son fácilmente recuperables y comprensibles para un experto en la materia al revisar la información proporcionada bajo el gen Entrez (base de datos mantenida por el Centro Nacional de Información Biotecnológica, Bethesda, MD) identificada por el número de identificación que figura en la Tabla 1.

Tabla 1 Descripción del gen marcador para tratamiento con inhibidor de NAE

ID del gen marcador	Nombre del gen marcador	ID del gen Entrez	Ubicación del cromosoma	Par de bases inicial	Par de bases final	SEQ ID NOs:
NF2	neurofibromina 2	4771	22q	29999545	30094589	1, 2, 3
SMAD4	madres contra homólogo 4 decapentapléjico	4089	18q	48556583	48611412	4, 5, 6
KDM6A	lisina desmetilasa específica 6A	7403	Xp	44732423	44971847	7, 8, 9, 10, 11
FBXW7	caja F y dominio de repetición WD que contiene 7	55294	4q	153242410	153456172	12, 13, 14
TP53	proteína tumoral p53	7157	17p	7571720	7590868	15, 16, 17
CDKN2A	inhibidor de quinasa dependiente de ciclina 2A	1029	9p	21967751	21994490	18, 19, 20
CDKN2A_p14	variante p14 del inhibidor de quinasa dependiente de ciclina 2A	1029	9p	21967751	21994490	21, 22, 23
APC	poliposis coli adenomatosa	324	5q	112043202	112181936	24, 25, 26

Como se usa en este documento, una característica "informativa", por ejemplo, tamaño, secuencia, composición o cantidad de un marcador se refiere a una característica, por ejemplo, tamaño, secuencia, composición o cantidad cuyo valor o diferencia se correlaciona con pronóstico o resultado. La característica informativa, por ejemplo, tamaño, secuencia, composición o cantidad de un marcador se puede obtener analizando cualquiera de los ácidos nucleicos, por ejemplo, ADN o ARN, o proteína correspondiente al gen marcador. La característica, por ejemplo, tamaño (por ejemplo, longitud o peso molecular), secuencia (por ejemplo, secuencia de ácidos nucleicos o secuencia de proteínas), composición (por ejemplo, composición de bases o aminoácidos o resumen de péptidos o patrón de fragmentos génicos) o cantidad (por ejemplo, número de copias y/o nivel de expresión) de un marcador, por ejemplo, un marcador de locus cromosómico o un marcador en una muestra de un paciente puede ser "informativo" si es diferente del tipo silvestre o la variante alélica de la sustancia que se analiza. En algunas realizaciones, una característica de un marcador es informativa si indica que el gen marcador es de tipo silvestre. En una realización donde se mide la cantidad de un marcador, una cantidad es "informativa" si es mayor o menor que una cantidad de referencia en un grado mayor que el error estándar del ensayo empleado para evaluar la expresión. El nivel de expresión informativa de un marcador se puede determinar mediante la correlación estadística del nivel de expresión medido y el resultado, por ejemplo, buena respuesta, mala respuesta, tiempo hasta progresión largo, tiempo hasta progresión corto, supervivencia a corto plazo o supervivencia a largo plazo. El resultado del análisis estadístico puede establecer un umbral para seleccionar marcadores para usar en los métodos descritos en este documento. Alternativamente, un marcador, por ejemplo, un marcador de locus cromosómico, o un gen marcador que tiene características diferenciales, por ejemplo, tamaño, secuencia, composición o cantidades tendrán intervalos de cantidades típicos que predicen el resultado. Una característica informativa, por ejemplo, tamaño, secuencia, composición o cantidad es una característica, por ejemplo, tamaño, secuencia, composición o cantidad que entra dentro del intervalo de características, por ejemplo, tamaño, secuencia, composición o cantidades determinadas para el resultado. Además, un conjunto de marcadores juntos pueden ser "informativos" si la combinación de sus características, por ejemplo, tamaños, secuencias, composiciones o cantidades satisface o está por encima o por debajo de una puntuación predeterminada para un marcador, por ejemplo, un marcador de locus cromosómico, o un gen marcador, establecido según lo determinado por los métodos proporcionados en este documento. Translocación génica, variación de empalme de transcripción, eliminación y truncamiento son ejemplos de sucesos que pueden cambiar el tamaño del marcador, secuencia o composición, además de mutaciones puntuales que pueden cambiar la secuencia o composición del marcador. La medición de una sola característica, por ejemplo, marcador, de un gen marcador (es decir, ADN, ARN o proteína) puede proporcionar un pronóstico, es decir, indicar resultado. La medición de más de una característica, por ejemplo, marcador, de un gen marcador puede proporcionar un pronóstico cuando las cantidades informativas de las dos características son coherentes entre sí, es decir, las biologías de los resultados no son contradictorias. Ejemplos de resultados coherentes a partir de medición de múltiples características de un gen marcador pueden ser la identificación de una mutación o deleción sin sentido en un ADN o ARN y una cantidad baja o de bajo peso molecular de proteína codificada, o una mutación en una región que codifica un bolsillo de unión o sitio activo de una proteína y baja actividad de la proteína codificada. Un ejemplo diferente puede ocurrir cuando una proteína está en una ruta con un circuito de retroalimentación que controla su síntesis en función de su nivel de actividad. En este ejemplo, una baja cantidad o actividad de proteína puede asociarse con una gran cantidad de su ARNm mutado como tejido, debido a la mutación del gen marcador, por tanto privado de alimento por la actividad de la proteína y señala repetidamente la producción de la proteína.

Como se usa en este documento, "deleción génica" se refiere a una cantidad de número de copias de ADN inferior a 2 y "amplificación" se refiere a una cantidad de número de copias de ADN superior a 2. Una cantidad "diploide" se refiere a un número de copias igual a 2. El término "diploide o amplificación" puede interpretarse como "no eliminación" de una copia génica. En un marcador cuya cantidad informativa alternativa es deleción génica, la amplificación generalmente no se vería. Por el contrario, el término "diploide o deleción" puede interpretarse como "no amplificación" del número de copias. En un marcador cuya cantidad informativa alternativa es amplificación, la deleción génica generalmente no se vería. Por cuestiones de claridad, la deleción de la secuencia puede ocurrir dentro de un gen como resultado de la mutación del gen marcador y puede dar como resultado la ausencia de proteína transcrita o un ARNm o proteína acortado. Tal eliminación puede no afectar el número de copias.

Las expresiones "supervivencia a largo plazo" y "supervivencia a corto plazo" se refieren al periodo de tiempo después de recibir una primera dosis de tratamiento que se prevé que viva un paciente con cáncer. Un "superviviente a largo plazo" se refiere a un paciente que se espera que tenga una tasa de progresión más lenta o muerte tardía por el tumor que aquellos pacientes identificados como supervivientes a corto plazo. La "supervivencia mejorada" o "una tasa de mortalidad más lenta" son determinaciones de vida estimadas basadas en características, por ejemplo, tamaño, secuencia, composición o cantidad de uno o más de los marcadores descritos en este documento, por ejemplo, en comparación con un patrón de referencia como el 70%, 80 %, 90 % o más de la población estará viva un periodo de tiempo suficiente después de recibir una primera dosis de tratamiento. Una "tasa de mortalidad más rápida" o un "tiempo de supervivencia más corto" se refieren a determinaciones estimadas de la duración de la vida basadas en características, por ejemplo, tamaño, secuencia, composición o cantidad de uno o más de los marcadores descritos en este documento, por ejemplo, en comparación con un patrón de referencia como el 50 %, 40 %, 30 %, 20 %, 10 % o menos de la población no vivirá un periodo de tiempo suficiente después de recibir una primera dosis de tratamiento. En algunas realizaciones, el periodo de tiempo suficiente es al menos 6, 12, 18, 24 o 30 meses medidos desde el primer día de recibir una terapia contra el cáncer.

Un cáncer es "sensible" a un agente terapéutico o hay una "buena respuesta" a un tratamiento si su tasa de crecimiento se inhibe como resultado del contacto con el agente terapéutico, en comparación con su crecimiento en ausencia de contacto con el agente terapéutico. El crecimiento de un cáncer se puede medir de varias maneras, por ejemplo, la característica, por ejemplo, se puede medir el tamaño de un tumor o la expresión de marcadores tumorales apropiados para ese tipo de tumor. Por ejemplo, se pueden usar las definiciones de respuesta utilizadas para apoyar la identificación de marcadores asociados con mieloma y su respuesta a un inhibidor de NAE, como una terapia de sulfamato de metilo 1-sustituido, los criterios del Southwestern Oncology Group (SWOG) como se describe en Blade *et al*. (1998) Br J Haematol. 102:1115-23. Estos criterios definen el tipo de respuesta medida en mieloma y también la caracterización del tiempo hasta la progresión de la enfermedad, que es otra medida importante de la sensibilidad de un tumor a un agente terapéutico. Para tumores sólidos, las directrices de Criterios de evaluación de respuesta en tumores sólidos (RECIST) (Eisenhauer *et al*. (2009) E. J. Canc. 45:228-247) pueden usarse para apoyar la identificación de marcadores asociados con tumores sólidos y la respuesta de tumores sólidos a un inhibidor de NAE. Los grupos de trabajo internacionales se reúnen periódicamente para establecer, actualizar y publicar criterios de respuesta para varios tipos de cánceres. Dichos informes publicados pueden seguirse para respaldar la identificación de marcadores de los tumores sujetos y su respuesta a inhibidores de NAE. Los ejemplos son criterios para leucemia mielógena aguda (AML, Cheson *et al*. (2003) J.Clin. Oncol. 21:4642-4649), linfomas, por ejemplo, linfoma no Hodgkin y Hodgkin (Cheson *et a*l. (2007) J.Clin. Oncol. 25:579-596). Los criterios tienen en cuenta métodos de análisis la tomografía por emisión de positrones (PET), por ejemplo, para identificar sitios con actividad metabólica alterada medible (por ejemplo, en sitios de tumores) o para rastrear marcadores específicos en tumores *in vivo,* inmunohistoquímica, por ejemplo, para identificar células tumorales mediante detección de unión de anticuerpos a marcadores tumorales específicos y citometría de flujo, por ejemplo, para caracterizar tipos de células mediante marcadores diferenciales y tinciones fluorescentes, además de métodos tradicionales como histología para identificar la composición celular (por ejemplo, recuentos de blastos en un frotis de sangre o una biopsia de médula ósea, presencia y número de figuras mitóticas) o estructura tisular (por ejemplo, arquitectura tisular desordenada o infiltración celular de la membrana basal). La calidad de responder a un inhibidor de NAE, como una terapia de sulfamato de metilo 1-sustituido puede ser variable, con diferentes tipos de cáncer que muestran diferentes niveles de "respuesta" a un agente terapéutico dado, bajo diferentes condiciones. Además, las medidas de respuesta pueden evaluarse utilizando criterios adicionales más allá del tamaño de crecimiento de un tumor, incluyendo calidad de vida del paciente, grado de metástasis, etc. Además, se pueden evaluar marcadores y variables de pronóstico clínico (por ejemplo, proteína M en mieloma, niveles de PSA en cáncer de próstata) en situaciones aplicables.

Un cáncer que "no responde" o tiene una "mala respuesta" a un agente terapéutico o hay una pobre respuesta a un tratamiento si su tasa de crecimiento no se inhibe o se inhibe en un grado muy bajo, como resultado del contacto con el agente terapéutico en comparación con su crecimiento en ausencia de contacto con el agente terapéutico. Como se ha indicado anteriormente, el crecimiento de un cáncer se puede medir de varias maneras, por ejemplo, se puede medir el tamaño de un tumor o la expresión de marcadores tumorales apropiados para ese tipo de tumor. Por ejemplo, las definiciones de respuesta utilizadas para apoyar la identificación de marcadores asociados con la no respuesta de tumores a agentes terapéuticos, pueden usarse pautas como las descritas anteriormente. La calidad de no responder a un agente terapéutico puede ser muy variable, con diferentes tipos de cáncer que muestran diferentes niveles de "falta de respuesta" a un agente terapéutico dado, bajo diferentes condiciones. Además, las mediciones de falta de respuesta pueden evaluarse utilizando criterios adicionales más allá del tamaño de crecimiento de un tumor, incluyendo calidad de vida del paciente, grado de metástasis, etc. Además, se pueden evaluar marcadores y variables de pronóstico clínico (por ejemplo, proteína M en mieloma, niveles de PSA en cáncer de próstata) en situaciones aplicables.

Como se usa en este documento, "tiempo hasta progresión largo, "TTP largo" y "tiempo hasta progresión corto", "TTP corto" se refieren al periodo de tiempo hasta que la enfermedad estable conseguida por el tratamiento se convierte en una enfermedad activa. En ocasiones, un tratamiento da como resultado una enfermedad estable que no es una respuesta buena ni mala, por ejemplo, MR, la enfermedad simplemente no empeora, por ejemplo, se convierte en una enfermedad progresiva, durante un periodo de tiempo. Este periodo de tiempo puede ser al menos 4-8 semanas, al menos 3-6 meses o más de 6 meses.

"Tratamiento" significará el uso de una terapia para prevenir o inhibir el crecimiento tumoral adicional, así como para causar contracción de un tumor y proporcionar tiempos de supervivencia más largos. El tratamiento también está destinado a incluir la prevención de metástasis tumorales. Un tumor se "inhibe" o "trata" si al menos un síntoma (según lo determinado por la capacidad de respuesta/no capacidad de respuesta, tiempo de progresión, o indicadores conocidos en la técnica y descritos en este documento) del cáncer o tumor se alivia, termina, ralentiza, minimiza o previene. Cualquier mejora de cualquier síntoma, física o de otra manera, de un tumor según el tratamiento utilizando un régimen terapéutico (por ejemplo, inhibidor de NAE, como un régimen de sulfamato de metilo 1-sustituido) como se describe adicionalmente en este documento, está dentro del alcance de la invención.

Como se usa en este documento, el término "agente" se define en términos generales como cualquier cosa a que las células cancerosas, incluyendo células tumorales, pueden exponerse en un protocolo terapéutico. En el contexto de la presente invención, dichos agentes incluyen, pero no se limitan a, un inhibidor de NAE, como agentes de sulfamato de metilo 1-sustituidos, así como agentes quimioterapéuticos como se conocen en la técnica y se

describen con más detalle en este documento.

El término "sonda" se refiere a cualquier molécula, por ejemplo, una molécula aislada, que es capaz de unirse selectivamente a una molécula diana específicamente pretendida, por ejemplo un marcador de la invención. Un experto en la materia puede sintetizar sondas o derivarlas de preparaciones biológicas apropiadas. Para fines de detección de la molécula diana, las sondas pueden diseñarse específicamente para ser etiquetadas, como se describe en este documento. Los ejemplos de moléculas que se pueden utilizar como sondas incluyen, pero no se limitan a, ARN, ADN, proteínas, anticuerpos, y monómeros orgánicos.

Una característica "normal", por ejemplo, tamaño, secuencia, composición o cantidad de un marcador puede referirse a la característica, por ejemplo, tamaño, secuencia, composición o cantidad en una "muestra de referencia". Una muestra de referencia puede ser una muestra normal coincidente, por ejemplo, línea germinal, del mismo paciente de quien se deriva el tumor, por ejemplo, con una mutación somática. Una muestra de referencia puede ser una muestra de un sujeto sano que no tiene la enfermedad asociada al marcador o una característica de referencia, por ejemplo, la característica promedio, por ejemplo, tamaño, secuencia, composición o cantidad del marcador de tipo silvestre en varios sujetos sanos. Una característica de muestra de referencia, por ejemplo, tamaño, secuencia, composición o cantidad puede estar compuesta de una característica, por ejemplo, tamaño, secuencia, composición o cantidad de uno o más marcadores de una base de datos de referencia. Alternativamente, una característica "normal", por ejemplo, tamaño, secuencia, composición o nivel de expresión de un marcador es la característica, por ejemplo, tamaño, secuencia, composición o cantidad del marcador, por ejemplo, gen marcador en células no tumorales en un entorno o situación de respuesta similar del mismo paciente del que se deriva el tumor. La cantidad normal de número de copias de ADN es 2 o diploide, a excepción de genes ligados a X en hombres, donde el número normal de copias de ADN es 1.

"Sobreexpresión" y "subexpresión" de un gen marcador, se refieren a expresión del gen marcador de un paciente en un nivel mayor o menor (por ejemplo, más de tres mitades, al menos dos veces, al menos tres veces, mayor o menor nivel, etc.), respectivamente, que el nivel normal de expresión del gen marcador, por ejemplo, medido por ARNm o proteína, en una muestra de ensayo que es mayor que el error estándar del ensayo empleado para evaluar la expresión. Un nivel de expresión "significativo" puede referirse a un nivel que satisface o está por encima o por debajo de una puntuación predeterminada para un conjunto de genes marcadores según lo determinado por los métodos proporcionados en el presente documento.

"Complementario" se refiere al concepto amplio de complementariedad de secuencias entre regiones de dos cadenas de ácidos nucleicos o entre dos regiones de la misma cadena de ácidos nucleicos. Se sabe que un resto de adenina de una primera región de ácido nucleico puede formar enlaces de hidrógeno específicos ("emparejamiento de bases") con un resto de una segunda región de ácido nucleico que es antiparalela a la primera región si el resto es timina o uracilo. Igualmente, se sabe que un resto de citosina de una primera cadena de ácidos nucleicos puede emparejarse con un resto de una segunda cadena de ácidos nucleicos que es antiparalela a la primera cadena si el resto es guanina. Una primera región de un ácido nucleico es complementaria a una segunda región del mismo o diferente ácido nucleico si, cuando las dos regiones están dispuestas de forma antiparalela, al menos un resto de nucleótido de la primera región es capaz de emparejarse con un resto de la segunda región. En una realización, la primera región comprende una primera porción y la segunda región comprende una segunda porción, de modo que, cuando las porciones primera y segunda están dispuestas de forma antiparalela, al menos aproximadamente 50 %, al menos aproximadamente 75 %, al menos aproximadamente 90 %, o al menos aproximadamente 95 % o todos los restos de nucleótidos de la primera porción son capaces de emparejar bases con restos de nucleótidos en la segunda porción.

"Homólogo" como se usa en este documento, se refiere a similitud de secuencia de nucleótidos entre dos regiones de la misma cadena de ácidos nucleicos o entre regiones de dos cadenas de ácidos nucleicos diferentes. Cuando una posición de resto de nucleótido en ambas regiones está ocupada por el mismo resto de nucleótido, entonces las regiones son homólogas en esa posición. Una primera región es homóloga a una segunda región si al menos una posición de resto de nucleótido de cada región está ocupada por el mismo resto. La homología entre dos regiones se expresa en términos de proporción de posiciones de restos de nucleótidos de las dos regiones que están ocupadas por el mismo resto de nucleótidos (es decir, por porcentaje de identidad). 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 homología con 50 % de identidad. En una realización, la primera región comprende una primera porción y la segunda región comprende una segunda porción, de modo que, al menos aproximadamente 50 %, al menos aproximadamente 75 %, al menos aproximadamente 90 %, o al menos aproximadamente 95 % de las posiciones de restos de nucleótidos de cada una de las porciones están ocupadas por el mismo resto de nucleótidos. En una realización del 100 % de identidad, todas las posiciones de restos de nucleótidos de cada una de las porciones están ocupadas por el mismo restos de nucleótidos.

A menos que en este documento se especifique lo contrario, los términos "anticuerpo" y "anticuerpos" abarcan en general formas de anticuerpos naturales, por ejemplo, anticuerpos policlonales (por ejemplo, IgG, IgA, IgM, IgE) y anticuerpos monoclonales y recombinantes como anticuerpos monocatenarios, proteínas de dos y múltiples cadenas, quimérico, injertado con CDR, anticuerpos humanos y humanizados y anticuerpos multiespecíficos, así

como fragmentos y derivados de todo lo anterior, fragmentos (por ejemplo, los dAb, scFv, Fab, F(ab)'2, Fab') y derivados que tienen al menos un sitio de unión antigénica. Los derivados de anticuerpos pueden comprender una proteína o resto químico conjugado con un anticuerpo. El término "anticuerpo" también incluye variantes sintéticas y genéticamente modificadas.

Un "kit" es cualquier artículo de fabricación (por ejemplo, un envase o recipiente) que comprende al menos un reactivo, por ejemplo, una sonda, para detectar específicamente un marcador o conjunto de marcadores de la invención. El artículo de fabricación puede ser promovido, repartido, vendido u ofrecido para venta como una unidad para realizar, por ejemplo, *in vitro,* los métodos de la presente invención, por ejemplo, en una muestra obtenida de un paciente. Los reactivos incluidos en dicho kit pueden comprender al menos una sonda de ácido nucleico y, opcionalmente, uno o más cebadores y/o anticuerpos para usar en la detección de características de marcadores, por ejemplo, tamaño, composición o cantidad de secuencia, por ejemplo, expresión. Además, un kit de la presente invención puede contener instrucciones que describen un ensayo de detección adecuado. Tal kit se puede usar convenientemente, por ejemplo, en un entorno de ensayos clínicos o contractuales, para generar información, por ejemplo, en niveles de expresión, característica, por ejemplo, tamaño, secuencia o composición de uno o más marcadores, para registrar, almacenar, transmitir o recibir para permitir el diagnóstico, evaluación o tratamiento de pacientes que presentan síntomas de cáncer, en particular pacientes que exhiben la posible presencia de un cáncer capaz de tratamiento con terapia de inhibición de NAE, incluyendo, por ejemplo, cánceres hematológicos por ejemplo, mielomas (por ejemplo, mieloma múltiple), linfomas (por ejemplo, linfoma no Hodgkin), leucemias (por ejemplo, leucemia mielógena aguda) y tumores sólidos (por ejemplo, tumores de piel, pulmón, mama, ovario, *etc.).*

Los presentes métodos y composiciones están diseñados para uso en diagnóstico y terapéutica para un paciente que padece cáncer. Un cáncer o tumor se trata o diagnostica según los métodos actuales. El término "cáncer" o "tumor" incluye cualquier crecimiento neoplásico en un paciente, incluyendo un tumor inicial y cualquier metástasis. El cáncer puede ser del tipo de tumor hematológico o sólido. Los tumores hematológicos incluyen tumores de origen hematológico, incluyendo, por ejemplo, mielomas *(es decir,* mieloma múltiple), leucemias (por ejemplo, Síndrome de Waldenstrom, leucemia linfocítica crónica, leucemia mielógena aguda, leucemia mielógena crónica, otras leucemias), linfomas (por ejemplo, linfomas de linfocitos B, linfoma no Hodgkin) y síndrome mielodisplásico. Los tumores sólidos pueden originarse en órganos e incluyen cánceres como en piel, pulmón, cerebro, mama, próstata, ovario, colon, riñón, páncreas, hígado, esófago, estómago, intestino, vejiga, útero, cuello del útero, cabeza y cuello, sistema nervioso central, hueso, testículos, glándula suprarrenal, etc. El cáncer puede comprender una célula en donde un gen marcador tiene una mutación. Como se usa en este documento, células cancerosas, incluyendo células tumorales, se refieren a células que se dividen a un ritmo anómalos (aumentado) o cuyo control de crecimiento o supervivencia es diferente al de las células en el mismo tejido donde surge o vive la célula cancerosa. Las células cancerosas incluyen, pero no se limitan a, células en carcinomas, como carcinoma de células escamosas, carcinoma de células basales, carcinoma de glándulas sudoríparas, carcinoma de las glándulas sebáceas, adenocarcinoma, carcinoma papilar, adenocarcinoma papilar, cistadenocarcinoma, carcinoma medular, carcinoma indiferenciado, carcinoma broncogénico, melanoma, carcinoma de células renales, hepatoma-carcinoma de células hepáticas, carcinoma del conducto biliar, colangiocarcinoma, carcinoma papilar, carcinoma de células transicionales, coriocarcinoma, seminoma, carcinoma embrionario, carcinomas mamarios, carcinoma gastrointestinal, carcinomas de colon, carcinoma de vejiga, carcinoma de próstata, y carcinoma de células escamosas de la región del cuello y cabeza; sarcomas, como fibrosarcoma, mixosarcoma, liposarcoma, condrosarcoma, sarcoma osteogénico, cordosarcoma, angiosarcoma, endoteliosarcoma, linfangiosarcoma, sinoviosarcoma y mesoteliosarcoma; cánceres hematológicos, como mielomas, leucemias (por ejemplo, leucemia mielógena aguda, leucemia linfocítica crónica, leucemia granulocítica, leucemia monocítica, leucemia linfocítica), y linfomas (por ejemplo, linfoma folicular, linfoma de células del manto, linfoma difuso de linfocitos B grandes, linfoma maligno, plasmocitoma, sarcoma de células reticulares, o enfermedad de Hodgkin); y tumores del sistema nervioso, incluyendo glioma, meningioma, meduloblastoma, schwannoma o epidimoma.

Como se usa en este documento, el término "no invasivo" se refiere a un procedimiento que inflige un daño mínimo a un sujeto. En el caso de aplicaciones clínicas, un procedimiento de muestreo no invasivo puede realizarse rápidamente, por ejemplo, en un entorno sin cita, típicamente sin anestesia y/o sin instrumentos quirúrgicos o sutura. Los ejemplos de muestras no invasivas incluyen sangre, suero, saliva, orina, frotis bucales, cultivos de garganta, muestras de heces y frotis cervicales. Los análisis de diagnóstico no invasivos incluyen radiografías, resonancia magnética, tomografía de emisión de positrones, etc.

En este documento se describe la evaluación del resultado para tratamiento de un tumor mediante medición de la cantidad de marcadores farmacogenómicos. También se describe la evaluación del resultado por medios no invasivos, convenientes o de bajo coste, por ejemplo, de muestras de sangre. Los métodos típicos para determinar la extensión del cáncer o el resultado de un tumor hematológico, por ejemplo, linfoma, leucemia, por ejemplo, leucemia mielógena aguda, mieloma (por ejemplo, mieloma múltiple) pueden emplear biopsia de médula ósea para recoger tejido para genotipo o fenotipo, por ejemplo, análisis histológico. La invención proporciona métodos para determinar, evaluar, asesorar o proporcionar un régimen de terapia apropiado para tratar un tumor o controlar una enfermedad en un paciente. La monitorización de un tratamiento utilizando los kits y métodos desvelados en este documento puede identificar el potencial de resultados desfavorables y permitir su prevención y, por tanto, un ahorro en morbilidad, mortalidad y costes de tratamiento mediante ajuste del régimen terapéutico, cese de terapia o uso de

terapia alternativa.

La expresión "muestra biológica" pretende incluir una muestra de paciente, por ejemplo, tejido, células, fluidos biológicos y sus aislados, aislado de un sujeto, así como tejidos, células y fluidos presentes en un sujeto y pueden obtenerse de un paciente o un sujeto normal. En tumores hematológicos de la médula ósea, por ejemplo, tumores de mieloma, el análisis primario del tumor se puede realizar en muestras de médula ósea. Sin embargo, algunas células tumorales, (por ejemplo, células tumorales clonotípicas, células endoteliales circulantes), son un porcentaje de la población celular en sangre completa. Estas células también pueden movilizarse a la sangre durante el tratamiento del paciente con factor estimulante de colonias de granulocitos (G-CSF) en preparación para un trasplante de médula ósea, un tratamiento convencional para tumores hematológicos, por ejemplo, leucemias, linfomas y mielomas. Ejemplos de células tumorales circulantes en mieloma múltiple han sido estudiados, por ejemplo, por Pilarski *et al*. (2000) Blood 95:1056-65 y Rigolin *et al*. (2006) Blood 107:2531-5. Así, muestras no invasivas, por ejemplo, para medición *in vitro* de marcadores para determinar el resultado del tratamiento, pueden incluir muestras de sangre periférica. En consecuencia, las células dentro de sangre periférica pueden analizarse para determinar la cantidad de marcadores. Para pacientes con tumores hematológicos, una muestra de referencia, control para características normales, por ejemplo, tamaño, secuencia, composición o cantidad se puede obtener de piel o un hisopo bucal del paciente. Para tumores sólidos, una muestra de tumor típica es una biopsia del tumor y, por tanto, comprende células tumorales sólidas. Alternativamente, una muestra de células tumorales desprendidas o raspadas del sitio tumoral se puede recoger no invasivamente, como en sangre, esputo, un aspirado de pezón, orina, deposiciones, frotis cervical, etc. Para tumores sólidos, una muestra de referencia de control para característica normal, por ejemplo, tamaño, secuencia, composición o cantidad se puede obtener de sangre del paciente.

Los recipientes de extracción de sangre pueden comprender un anticoagulante, por ejemplo, heparina o ácido etilendiamintetraacético (EDTA), citrato sódico o soluciones de citrato con aditivos para preservar la integridad de la sangre, como dextrosa o albúmina o tampones, por ejemplo, fosfato. Si la cantidad de marcador se mide con el nivel de su ADN en la muestra, se puede añadir un estabilizador de ADN, por ejemplo, un agente que inhibe la ADNsa, a la muestra. Si la cantidad de marcador se mide con el nivel de su ARN en la muestra, se puede añadir un estabilizador de ARN, por ejemplo, un agente que inhibe la ARNasa, a la muestra. Si la cantidad de marcador se mide con el nivel de su proteína en la muestra, se puede añadir un estabilizador de proteínas, por ejemplo, un agente que inhibe proteasas, a la muestra. Un ejemplo de un recipiente de recolección de sangre son los tubos PAXGENE® (PREANALYTIX, Valencia, CA), útiles para estabilizar ARN después de extracción de sangre. Las muestras de sangre periférica se pueden modificar, por ejemplo, fraccionar, clasificar o concentrar (por ejemplo, para dar como resultado muestras enriquecidas con tumor o agotadas de tumor (por ejemplo, para una muestra de referencia)). Los ejemplos de muestras modificadas incluyen células de mieloma clonotípico, que puede recogerse, por ejemplo, selección negativa, por ejemplo, separación de glóbulos blancos de glóbulos rojos (por ejemplo, centrifugación diferencial a través de una solución densa de azúcar o polímero (por ejemplo, solución FICOLL® (división Amersham Biosciences de GE healthcare, Piscataway, NJ) o solución HISTOPAQUE®-1077, Sigma-Aldrich Biotechnology LP y Sigma-Aldrich Co., St. Louis, MO)) y/o selección positiva mediante unión de linfocitos B a un agente de selección (por ejemplo, un reactivo que se une a una célula tumoral o marcador progenitor mieloide, como CD34, CD38, CD138, o CD133, para aislamiento directo (por ejemplo, la aplicación de un campo magnético a soluciones de células que comprenden perlas magnéticas (por ejemplo, de Miltenyi Biotec, Auburn, CA) que se unen a los marcadores de linfocitos B) o clasificación celular activada por fluorescencia).

Alternativamente, se puede analizar una línea celular tumoral, por ejemplo, OCI-Ly3, célula OCI-Ly10 (Alizadeh *et a*l. (2000) Nature 403:503-511), una célula RPMI 6666, una célula SUP-B15, una célula KG-1, una célula CCRF-SB, una célula 8ES, una célula Kasumi-1, una célula Kasumi-3, una célula BDCM, una célula HL-60, una célula Mo-B, una célula JM1, una célula GA-10 o un linfoma de linfocitos B (por ejemplo, BC-3) o una línea celular o una colección de líneas celulares tumorales (véase, por ejemplo, McDermott *et al*. (2007) PNAS 104:19936-19941 o identificación del perfil de células tumorales anticáncer ONCOPANEL™ (Ricerca Biosciences, Bothell, WA)). Un experto en la materia puede seleccionar y obtener fácilmente las células apropiadas (por ejemplo, de la Colección Americana de Cultivos Tipo (ATCC®), Manassas, VA) que se utilizan en el presente método. Si las composiciones o métodos se utilizan para predecir el resultado del tratamiento en un paciente o controlar la eficacia de un protocolo terapéutico, entonces una muestra de tejido o sangre obtenida del paciente que se está tratando es una fuente útil de células o gen marcador o productos genéticos para un ensayo.

La muestra, por ejemplo, tumor, por ejemplo, biopsia o médula ósea, sangre o sangre modificada, (por ejemplo, que comprende células tumorales) y/o la referencia, por ejemplo, control combinado (por ejemplo, línea germinal), la muestra puede someterse a una variedad de técnicas de preparación y almacenamiento postcolección bien conocidas (por ejemplo, extracción de ácido nucleico y/o proteína, fijación, almacenamiento, congelación, ultrafiltración, concentración, evaporación, centrifugación, *etc.)* antes de evaluar la cantidad del marcador en la muestra.

En una realización, el estado mutacional de un gen marcador, por ejemplo, una mutación en un marcador se puede identificar secuenciando un ácido nucleico, por ejemplo, un ADN, ARN, ADNc o una proteína correlacionada con el gen marcador. Existen varios métodos de secuenciación conocidos en la técnica para secuenciar ácidos nucleicos. Un cebador de ácido nucleico puede diseñarse para unirse a una región que comprende un sitio de mutación

potencial o puede diseñarse para complementar la secuencia mutada en lugar de la secuencia de tipo silvestre. Los pares de cebadores pueden diseñarse para agrupar una región que comprende una mutación potencial en un gen marcador. Se puede usar un cebador o un par de cebadores para secuenciar una o ambas cadenas de ADN correspondientes al gen marcador. Se puede usar un cebador junto con una sonda, por ejemplo, una sonda de ácido nucleico, por ejemplo, una sonda de hibridación, para amplificar una región de interés antes de la secuenciación para aumentar las cantidades de secuencia para detección de una mutación en un gen marcador. Los ejemplos de regiones que pueden secuenciarse incluyen un gen completo, transcripciones del gen y un fragmento del gen o la transcripción, por ejemplo, uno o más de exones o regiones no traducidas. Ejemplos de mutaciones para selección de cebadores y análisis de secuencia o composición se pueden encontrar en bases de datos públicas que recopilan información sobre mutaciones, como COSMIC y dbGaP. Algunas mutaciones de genes marcadores como NF2, SMAD, KDM6A o FBXW7 se enumeran en las Tablas 8-11 en los Ejemplos como ejemplos de mutaciones que pueden estar asociadas con sensibilidad a inhibición de NAE, por ejemplo, inhibición por sulfamatos de 1-metilo, por ejemplo, MLN4924.

Los métodos de secuenciación son conocidos por un experto en la materia. Ejemplos de métodos incluyen el método Sanger, el método SEQUENOM™ y métodos de secuenciación de próxima generación (NGS). El método Sanger, que comprende el uso de electroforesis, por ejemplo, electroforesis capilar para separar fragmentos de ADN marcados con cebador alargado, puede automatizarse para aplicaciones de alto rendimiento. La secuenciación de extensión del cebador puede realizarse después de amplificación por PCR de las regiones de interés. El software puede ayudar con llamadas de base de secuencia y con identificación de mutación. El análisis de secuenciación SEQUENOM™ MASSARRAY® (San Diego, CA) es un método de espectrometría de masas que compara la masa real con la masa esperada de fragmentos particulares de interés para identificar mutaciones. La tecnología NGS (también llamada "secuenciación masiva paralela" y "secuenciación de segunda generación") en general proporciona un rendimiento mucho mayor que los métodos anteriores y utiliza una variedad de enfoques (revisado en Zhang *et al*. (2011) J. Genet. Genomics 38:95-109 y Shendure y Hanlee (2008) Nature Biotech. 26:1135-1145). Los métodos NGS pueden identificar mutaciones de baja frecuencia en un marcador en una muestra. Algunos métodos NGS (véase, por ejemplo, secuenciador genómico GS-FLX (Roche Applied Science, Branford, CT), analizador genómico (Illumina, Inc. San Diego, CA) analizador SOLID™ (Applied Biosystems, Carlsbad, CA), Polonator G.007 (Dover Systems, Salem, NH), HELISCOPE™ (Helicos Biosciences Corp., Cambridge, MA)) usan secuenciación de matriz cíclica, con o sin amplificación clonal de productos de PCR separados espacialmente en una celda de flujo y varios esquemas para detectar el nucleótido modificado marcado que se incorpora por la enzima de secuenciación (por ejemplo, polimerasa o ligasa). En un método NGS, los pares de cebadores pueden usarse en reacciones de PCR para amplificar regiones de interés. Las regiones amplificadas se pueden ligar en un producto concatenado. Las bibliotecas clonales se generan en la celda de flujo a partir de la PCR o los productos ligados y se amplifican aún más (PCR "puente" o "agrupación") para secuenciación de un solo extremo a medida que la polimerasa añade una base terminada reversiblemente, etiquetada que se muestra en uno de los cuatro canales, dependiendo de la identidad de la base etiquetada y luego eliminada para el siguiente ciclo. El software puede ayudar en la comparación con secuencias genómicas para identificar mutaciones.

La composición de proteínas y ácidos nucleicos puede determinarse de muchas maneras conocidas en la técnica, como tratándolos de maneras que se escinden, degradan o ingieren y luego analizando los componentes. Espectrometría de masas, electroforesis y cromatografía pueden separar y definir componentes para comparación. Las mutaciones que causan deleciones o inserciones pueden identificarse por diferencias de tamaño o carga en estos métodos. La digestión de proteína o digestión de ácido nucleico con enzima de restricción puede revelar diferentes patrones de fragmentos después de algunas mutaciones. Los anticuerpos que reconocen aminoácidos mutantes particulares en sus contextos estructurales pueden identificar y detectar estas mutaciones en las muestras (véase a continuación).

En una realización, ADN, por ejemplo, el ADN genómico correspondiente al marcador de tipo silvestre o mutado puede analizarse en formatos tanto *in situ* como *in vitro* en una muestra biológica utilizando métodos conocidos en la técnica. El ADN puede aislarse directamente de la muestra o aislarse después de aislar otro componente celular, por ejemplo, ARN o proteína. Hay kits disponibles para aislamiento de ADN, por ejemplo, Kit Micro ADN QIAAMP® (Qiagen, Valencia, CA). El ADN también puede amplificarse usando tales kits.

En otra realización, el ARNm correspondiente al marcador puede analizarse en formatos tanto *in situ* como *in vitro* en una muestra biológica utilizando métodos conocidos en la técnica. En los Ejemplos se incluye un ejemplo de un método para medir el nivel de expresión. Por ejemplo, se puede usar una sonda de ácido nucleico para hibridarse con un marcador y se puede medir la cantidad de sonda hibridada. Muchos métodos de detección de expresión usan ARN aislado. Para métodos *in vitro*, para la purificación de ARN de las células tumorales se puede utilizar cualquier técnica de aislamiento de ARN que no seleccione contra aislamiento de ARNm (véase, por ejemplo*.,* Ausubel *et al*., ed., Current Protocols in Molecular Biology, John Wiley & Sons, Nueva York 1987-1999). Adicionalmente, se pueden procesar fácilmente grandes cantidades de muestras de tejido utilizando técnicas bien conocidas por los expertos en la materia, como, por ejemplo, el proceso de aislamiento de ARN de una sola etapa de Chomczynski (1989, Patente de Estados Unidos N.º 4.843.155). El ARN puede aislarse utilizando procedimientos convencionales (véase, por ejemplo, Chomczynski y Sacchi (1987) Anal. Biochem. 162:156-159), soluciones (por ejemplo, trizol, TRI REAGENT® (Molecular Research Center, Inc., Cincinnati, OH; véase Patente de Estados Unidos N.º 5.346.994) o

kits (por ejemplo, un kit de aislamiento QIAGEN® Group RNEASY® (Valencia, CA) o el sistema de aislamiento de ARN total LEUKOLOCK™, división Ambion de Applied Biosystems, Austin, TX).

Para eliminar el ADN de las muestras de ARN se pueden usan etapas adicionales. La lisis celular se puede lograr con un detergente no iónico, seguido de microcentrifugación para eliminar los núcleos y, por tanto, la mayor parte del ADN celular. Posteriormente, el ADN puede aislarse de los núcleos para el análisis de ADN. En una realización, el ARN se extrae de células de diversos tipos de interés usando lisis con tiocianato de guanidinio seguido de centrifugación con CsCl para separar el ARN del ADN (Chirgwin *et al*. (1979) Biochemistry 18:5294-99). El Poli(A)+ARN se selecciona mediante selección con oligo-dT celulosa (véase Sambrook *et al*. (1989) Molecular Cloning-A Laboratory Manual (2ª ed.), Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.). Alternativamente, la separación del ARN del ADN se puede lograr mediante extracción orgánica, por ejemplo, con fenol caliente o fenol/cloroformo/alcohol isoamílico. Si se desea, se pueden añadir inhibidores de ARNasa al tampón de lisis. Igualmente, para ciertos tipos de células, puede ser deseable añadir al protocolo una etapa de desnaturalización/digestión de proteínas. Para muchas aplicaciones, es deseable enriquecer el ARNm con respecto a otros ARN celulares, como ARN de transferencia (ARNt) y ARN ribosómico (ARNr). La mayoría de los ARNm contienen una cola de poli(A) en su extremo 3'. Esto les permite enriquecerse mediante cromatografía de afinidad, por ejemplo, usando oligo(dT) o poli(U) acoplado a un soporte sólido, como medio de celulosa o SEPHADEX.R™. (Véase Ausubel *et al*. (1994) Current Protocols In Molecular Biology, vol. 2, Current Protocols Publishing, Nueva York). Una vez unido, el poli(A)+ARNm se eluye de la columna de afinidad usando EDTA 2 mM/SDS al 0,1 %.

La característica de un marcador de la invención en una muestra biológica, por ejemplo, después de obtener una muestra biológica (por ejemplo, una muestra de médula ósea, una biopsia tumoral o una muestra de referencia) de un sujeto de ensayo, puede evaluarse mediante cualquiera de una amplia variedad de métodos bien conocidos para detectar o medir la característica, por ejemplo, de un ácido nucleico (por ejemplo, ARN, ARNm, ADN genómico o ADNc) y/o proteína traducida. Los ejemplos no limitantes de tales métodos incluyen métodos inmunológicos para detección de secreciones, superficie celular, proteínas citoplasmáticas o nucleares, métodos de purificación de proteínas, ensayos de función o actividad proteica, métodos de hibridación de ácido nucleico, métodos de transcripción inversa de ácido nucleico y métodos de amplificación de ácido nucleico. Estos métodos incluyen la tecnología de matriz genética/chip, RT-PCR, ensayos de expresión génica TAQMAN® (Applied Biosystems, Foster City, CA), por ejemplo, en condiciones de laboratorio aprobadas por GLP, hibridación *in situ*, inmunohistoquímica, inmunotransferencia, FISH (hibridación *in situ* con fluorescencia), análisis FACS, transferencia Northern, transferencia Southern, Chips de perlas para análisis de ADN INFINIUM® (Illumina, Inc., San Diego, CA), PCR cuantitativa, matrices de cromosomas artificiales bacterianos, matrices de polimorfismo de un solo nucleótido (SNP) (Affymetrix, Santa Clara, CA) o análisis citogenéticos. Los métodos de detección de la invención pueden usarse así para detectar ARN, ARNm, proteína, ADNc, o ADN genómico, por ejemplo, en una muestra biológica tanto *in vitro* como *in vivo.* Además, las técnicas *in vivo* para detección de un polipéptido o ácido nucleico correspondiente a un marcador de la invención incluyen introducir en un sujeto una sonda marcada para detectar el biomarcador, por ejemplo, un ácido nucleico complementario a la transcripción de un biomarcador o un anticuerpo marcado, receptor de Fc o antígeno dirigido contra el polipéptido, por ejemplo, marcador tipo silvestre o mutante. Por ejemplo, el anticuerpo puede marcarse con un isótopo radiactivo cuya presencia y ubicación en un sujeto puede detectarse mediante técnicas de imagen convencionales. Estos ensayos se pueden realizar de varias maneras. Un expertos puede seleccionar entre estos u otros métodos apropiados y disponibles basados en la naturaleza de los marcadores, muestra de tejido y mutación en cuestión. Algunos métodos se describen con más detalle en secciones posteriores. Diferentes métodos o combinaciones de métodos podrían ser apropiados en diferentes casos o, por ejemplo en diferentes tipos de tumores o poblaciones de pacientes.

Las técnicas *in vitro* para detección de un polipéptido correspondiente a un marcador de la invención incluyen ensayos inmunoabsorbentes ligados a enzimas (ELISA), transferencias Western, matriz de proteínas, inmunoprecipitaciones e inmunofluorescencia. En tales ejemplos, la expresión de un marcador se evalúa usando un anticuerpo (por ejemplo, un anticuerpo radiomarcado, marcado con cromóforo, marcado con fluoróforo o marcado con enzima), un derivado de anticuerpo (por ejemplo, un anticuerpo conjugado con un sustrato o con la proteína o ligando de un par proteína-ligando (por ejemplo, biotina-estreptavidina)), o un fragmento de anticuerpo (por ejemplo, un anticuerpo de cadena sencilla, un dominio hipervariable de anticuerpo aislado, *etc.)* que se une específicamente con una proteína marcadora o fragmento de la misma, por ejemplo, una proteína o fragmento que comprende una región que puede mutar o una parte que comprende una secuencia mutada, o un resto mutado en su contexto estructural, incluyendo una proteína marcadora que ha sufrido toda o una parte de su modificación postraduccional normal. Un anticuerpo puede detectar una proteína con una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NO: 3, 6, 10, 11, 14, 17, 20 y 23. Alternativamente, un anticuerpo puede detectar una proteína mutada con una secuencia de aminoácidos variante seleccionada del grupo que consiste en un mutante de SEQ ID NO: 3, 6, 10, 11, 14, 17, 20 y 23. Los restos enumerados como mutados en bases de datos públicas como COSMIC de dbGaP pueden prepararse en composiciones inmunogénicas para generación de anticuerpos que reconocerán y se unirán específicamente a los restos mutantes. Otro método puede emplear pares de anticuerpos, en donde uno del par se uniría a una proteína marcadora cadena arriba, es decir, N-terminal a la región de mutación esperada, por ejemplo, sin sentido o deleción y el otro del par se uniría a la proteína cadena abajo. La proteína de tipo silvestre se uniría a ambos anticuerpos del par, pero una proteína con una mutación sin sentido o deleción se uniría solo al anticuerpo N-terminal del par. Un ensayo como un ensayo ELISA en sándwich podría detectar una pérdida de

cantidad de la proteína de tipo silvestre en la muestra tumoral, por ejemplo, en comparación con la muestra de referencia, o un ELISA convencional compararía los niveles de unión de los anticuerpos para inferir que una mutación está presente en una muestra tumoral.

Los métodos indirectos para determinar la cantidad o funcionalidad de un marcador de proteína también incluyen la medición de la actividad de la proteína. Por ejemplo, una muestra, o una proteína aislada de la muestra o expresada a partir de ácido nucleico aislado, clonado o amplificado a partir de la muestra se puede evaluar para actividad de la proteína marcadora. La actividad de NF2 se puede medir por su capacidad para asociarse con compañeros de unión, por ejemplo, en un ensayo sin células o en un ensayo basado en células. En un ejemplo, se puede medir la capacidad de NF2 para unirse a membranas de glóbulos rojos o p55/MPPl (Seo *et al*. (2009) Exp. Biol. Med. 234:255-262). En otro ejemplo, la actividad de SMAD4 se puede medir por su actividad en transducción de señales, por ejemplo, en un ensayo sin células o en un ensayo basado en células. En un ejemplo, se puede medir el estado de fosforilación de SMAD4, la unión de SMAD4 al ADN en un elemento de unión a Smad, por ejemplo, en un ensayo de desplazamiento de gel o en un ensayo indicador (véase, por ejemplo, Kuang y Chen (2004) Oncogene 23:1021-1029), pueden medirse o la translocación de SMAD4 entre el núcleo y el citoplasma puede visualizarse y cuantificarse en imágenes celulares. En otro ejemplo, la actividad de KDM6A se puede medir por su capacidad para desmetilar proteínas, por ejemplo, histonas. Por ejemplo, un ensayo puede medir el nivel de desmetilación de la lisina 27 de la histona 3 (Hong *et al*. (2007) PNAS 104:18439-18444). En otro ejemplo, la actividad de FBXW7 se puede medir por su capacidad para unirse a la ciclina E o asociarse al complejo de Skp-culina-caja F ubiquitina ligasa. En otro ejemplo, la actividad de TP53 se puede medir por la capacidad de unirse al ADN o formar tetrámeros.

En una realización, la expresión de un marcador se evalúa preparando ARNm/ADNc (es decir, un polinucleótido transcrito) de las células en una muestra de paciente, e hibridando el ARNm/ADNc con un polinucleótido de referencia, por ejemplo, una sonda de ácido nucleico aislada, por ejemplo, una sonda de hibridación, que es un complemento de un ácido nucleico marcador, o un fragmento del mismo. El ADNc puede, opcionalmente, amplificarse usando cualquiera de una variedad de métodos de reacción en cadena de la polimerasa antes de la hibridación con el polinucleótido de referencia. La expresión de uno o más marcadores también se puede detectar utilizando PCR cuantitativa para evaluar el nivel de expresión del marcador(es). Un ejemplo del uso de medir los niveles de ARNm es que una mutación inactivadora en un gen marcador puede dar como resultado un nivel alterado de ARNm en una célula. El nivel puede sobrerregularse debido a la producción de proteínas de señalización de retroalimentación en vista de la proteína no funcional o ausente o desregularse debido a la inestabilidad de una secuencia de ARNm alterada. Alternativamente, cualquiera de los muchos métodos conocidos para detectar mutaciones o variantes (por ejemplo, polimorfismos de un solo nucleótido, deleciones, etc., discutido anteriormente) de un marcador de la invención puede usarse para detectar la aparición de una mutación en un gen marcador en un paciente.

Un ejemplo de medición directa es la cuantificación de transcripciones. Como se usa en este documento, el nivel o cantidad de expresión se refiere a la cantidad absoluta de expresión de un ARNm codificado por el marcador o la cantidad absoluta de expresión de la proteína codificada por el marcador. Como alternativa a hacer determinaciones basadas en la cantidad de expresión absoluta de los marcadores seleccionados, las determinaciones pueden basarse en cantidades de expresión normalizadas. La cantidad de expresión se puede normalizar corrigiendo el nivel de expresión absoluto de un marcador al comparar su expresión con la expresión de un marcador de control que no es un marcador, por ejemplo, en una función de mantenimiento que se expresa constitutivamente. Los marcadores adecuados para normalización también incluyen genes constitutivos, como el gen de actina o microglobulina beta-2. Los marcadores de referencia para fines de normalización de datos incluyen marcadores que se expresan ubicuamente y/o cuya expresión no está regulada por oncogenes. En la técnica se conoce genes expresados constitutivamente y pueden identificarse y seleccionarse según el tejido relevante y/o la situación del paciente y los métodos de análisis. Dicha normalización permite comparar el nivel de expresión en una muestra, a otra muestra, por ejemplo, entre muestras de diferentes tiempos o diferentes sujetos. Además, el nivel de expresión se puede proporcionar como nivel de expresión relativo. El valor de referencia de una muestra de ADN genómico, por ejemplo, número de copias diploides, puede determinarse midiendo cantidades en células de sujetos sin tumor o en células no tumorales del paciente. Para determinar una cantidad relativa de un marcador o conjunto de marcadores, la cantidad del marcador o conjunto de marcadores se determina para al menos 1, o 2, 3, 4, 5 o más muestras, por ejemplo, 7, 10, 15, 20 o 50 o más muestras para establecer un valor de referencia, antes de determinar el nivel de expresión para la muestra en cuestión. Para establecer una medición de referencia, se determina la cantidad o nivel medio de cada uno de los marcadores o conjuntos de marcadores analizados en el mayor número de muestras y esto se usa como nivel de expresión de referencia para biomarcadores o conjuntos de biomarcadores en cuestión. La cantidad del marcador o conjunto de marcadores determinada para la muestra de ensayo (por ejemplo, nivel absoluto de expresión) se divide entre el valor de referencia obtenido para ese marcador o conjunto de marcadores. Esto proporciona una cantidad relativa y ayuda a identificar niveles anómalos de actividad de la proteína marcadora.

Las sondas basadas en la secuencia de una molécula de ácido nucleico de la invención pueden usarse para detectar transcripciones o secuencias genómicas correspondientes a uno o más marcadores de la invención. La sonda puede comprender un grupo marcador unido a la misma, por ejemplo, un radioisótopo, un compuesto fluorescente, una enzima o un cofactor enzimático. Tales sondas se pueden usar como parte de un kit de ensayo de diagnóstico para identificar células o tejidos que expresan la proteína, como midiendo los niveles de una molécula de

ácido nucleico que codifica la proteína en una muestra de células de un sujeto, por ejemplo, detectar niveles de ARNm o determinar si un gen que codifica la proteína ha sido mutado o eliminado.

Además de las secuencias de nucleótidos descritas en los registros de bases de datos descritos en este documento, los expertos en la materia observarán que los polimorfismos de la secuencia de ADN que conducen a cambios en la secuencia de aminoácidos pueden existir dentro de una población (por ejemplo, la población humana). Tales polimorfismos genéticos pueden existir entre individuos dentro de una población debido a la variación alélica natural. Un alelo es uno de un grupo de genes que ocurren alternativamente en un locus genético dado. Además, se observará que también pueden existir polimorfismos de ADN que afectan los niveles de expresión de ARN que pueden afectar el nivel de expresión general de ese gen (por ejemplo, afectando regulación o degradación).

Los cebadores o sondas de ácido nucleico comprenden una secuencia de nucleótidos complementaria a un marcador específico o una región mutada del mismo y son de longitud suficiente para hibridarse selectivamente con un gen marcador o ácido nucleico asociado con un gen marcador, por ejemplo, pueden unirse al ácido nucleico con especificidad de secuencia de bases y permanecer unidos, después del lavado. Se pueden usar cebadores y sondas para ayudar en el aislamiento y secuenciación de ácidos nucleicos marcadores. En una realización, el cebador o la sonda de ácido nucleico, por ejemplo, un oligonucleótido sustancialmente purificado, un ácido nucleico aislado, comprende una región que tiene una secuencia de nucleótidos que se hibrida, por ejemplo, bajo condiciones estrictas, a aproximadamente 6, 8, 10, 12, 15, 20, 25, 30, 40, 50, 60, 75, 100, 200, 350, 500 o más nucleótidos consecutivos de un gen marcador o una región que comprende una mutación en un marcador gen o transcripción del mismo o un complemento. En otra realización, la sonda de cebador o ácido nucleico es capaz de hibridarse con un ácido nucleico marcador que comprende una secuencia de nucleótidos de cualquier secuencia establecida en cualquiera de SEQ ID NOs: 1, 2, 4, 5, 7, 8, 9, 12, 13, 15, 16, 18, 19, 21, 22, 24, 25 o una secuencia en el cromosoma 22q del par de bases 29999545 a 30094589, cromosoma 18q del par de bases 48556583 a 48611412, cromosoma Xp del par de bases 44732423 a 44971847, cromosoma 4q del par de bases 153242410 a 153456172, cromosoma 17p del par de bases 7571720 a 7590868, cromosoma 9p del par de bases 21967751 a 21994490, o un complemento de cualquiera de los anteriores. Por ejemplo, una sonda de cebador o ácido nucleico que comprende una secuencia de nucleótidos de al menos aproximadamente 10 nucleótidos consecutivos, al menos aproximadamente 15 nucleótidos consecutivos, al menos aproximadamente 25 nucleótidos consecutivos, al menos aproximadamente 35 nucleótidos consecutivos, al menos aproximadamente 50 nucleótidos consecutivos, o que tienen de aproximadamente 15 a aproximadamente 20 nucleótidos establecidos en cualquiera de SEQ ID NOs: 1, 2, 4, 5, 7, 8, 9, 12, 13, 15, 16, 18, 19, 21, 22, 24, 25 o una secuencia en el cromosoma 22q del par de bases 29999545 a 30094589, cromosoma 18q del par de bases 48556583 a 48611412, cromosoma Xp del par de bases 44732423 a 44971847, cromosoma 4q del par de bases 153242410 a 153456172, cromosoma 17p del par de bases 7571720 a 7590868, la invención proporciona el cromosoma 9p del par de bases 21967751 a 21994490, o un complemento de cualquiera de los anteriores. Los cebadores o sondas de ácido nucleico que tienen una secuencia de más de aproximadamente 25, 40 o 50 nucleótidos también están dentro del alcance de la invención. En otra realización, una sonda de cebador o ácido nucleico puede tener una secuencia de al menos 70 %, al menos 75 %, 80 % u 85 %, o al menos, 90 %, 95 % o 97 % idéntica a la secuencia de nucleótidos de cualquier secuencia establecida en cualquiera de SEQ ID NOs: 1, 2, 4, 5, 7, 8, 9, 12, 13, 15, 16, 18, 19, 21, 22, 24, 25 o una secuencia en el cromosoma 22q del par de bases 29999545 a 30094589, cromosoma 18q del par de bases 48556583 a 48611412, cromosoma Xp del par de bases 44732423 a 44971847, cromosoma 4q del par de bases 153242410 a 153456172, cromosoma 17p del par de bases 7571720 a 7590868, cromosoma 9p del par de bases 21967751 a 21994490, o un complemento de cualquiera de los anteriores. Los análogos de ácido nucleico pueden usarse como sitios de unión para hibridación. Un ejemplo de un análogo de ácido nucleico adecuado es el ácido nucleico peptídico (véase, por ejemplo, Egholm *et al*., Nature 363:566 568 (1993); Pat. de Estados Unidos N.º 5.539.083).

En algunas realizaciones, la sonda de ácido nucleico puede diseñarse para unirse a la secuencia de tipo silvestre, entonces la presencia de una mutación en esa región puede causar una disminución, por ejemplo, disminución mensurable, en unión o hibridación por esa sonda. En otra realización, una sonda de ácido nucleico puede diseñarse para unirse a una secuencia mutante, entonces la presencia de una mutación en esa región puede causar un aumento en la unión o hibridación por esa sonda. En otras realizaciones, se puede diseñar un conjunto de sonda y cebador o un par de cebadores para agrupar una región en un marcador que puede tener una mutación, de modo que la amplificación basada en ese conjunto o par puede dar como resultado ácidos nucleicos que se pueden secuenciar para identificar la mutación.

Los cebadores o las sondas de ácido nucleico se pueden seleccionar utilizando un algoritmo que tiene en cuenta las energías de unión, composición base, complejidad de secuencia, energías de unión de hibridación cruzada y estructura secundaria (véase Friend *et al.* publicaciónde patente internacional WO 01/05935, publicada el 25 de enero de 2001; Hughes *et al*., Nat. Biotech. 19:342-7 (2001). Los cebadores útiles o las sondas de ácido nucleico de la invención se unen a secuencias que son únicas para cada transcripción, por ejemplo, regiones mutadas diana y pueden usarse en PCR para amplificar, detectar y secuenciar solo ese ácido nucleico particular, por ejemplo, transcripción o transcripción mutada. Ejemplos de algunas mutaciones de genes marcadores, por ejemplo, NF2, SMAD4, KDM6A y FBXW7 se encuentran en las Tablas de los Ejemplos (Tablas 8-11). Otras mutaciones se describen en los artículos de referencia citados en este documento y en bases de datos públicas descritas en este grupo. Un experto puede diseñar cebadores y sondas de ácido nucleico para los marcadores desvelados este

documento o marcadores relacionados con características similares, por ejemplo,, marcadores en los loci cromosómicos, o mutaciones en diferentes regiones del mismo gen marcador descrito en este documento, usando la experiencia en la materia, por ejemplo, ajustar el potencial para unión de la sonda de cebador o ácido nucleico a secuencias convencionales, mutantes o variantes alélicas manipulando la degeneración o el contenido de GC en el cebador o sonda de ácido nucleico. Los programas informáticos que se conocen bien en la técnica son útiles para diseño de cebadores con la especificidad requerida y las propiedades óptimas de amplificación, como Oligo versión 5.0 (National Biosciences, Plymouth, MN). Aunque para detectar los marcadores descritos en el presente documento y mutantes se pueden usar sondas y cebadores de ácido nucleico perfectamente complementarios, polimorfismos o alelos de los mismos, se contemplan desviaciones de la complementariedad completa cuando tales desviaciones no impiden que la molécula se hibride específicamente con la región diana. Por ejemplo, un cebador oligonucleotídico puede tener un fragmento no complementario en su extremo 5', con el resto del cebador siendo complementario a la región diana. Alternativamente, los nucleótidos no complementarios se pueden intercalar en la sonda o cebador de ácido nucleico siempre que la sonda o cebador resultante todavía sea capaz de hibridarse específicamente con la región diana.

Una indicación del resultado del tratamiento se puede evaluar estudiando la cantidad de 1 marcador, 2 marcadores, 3 marcadores o 4 marcadores, o más, por ejemplo, 5, 6, 7, 8, 9, 10, 15, 20 o 25 marcadores, o partes mutadas de los mismos, por ejemplo, genes marcadores que participan o interactúan con la ruta de la ligasa culina-RING, por ejemplo, supresores tumorales, por ejemplo, que pueden inactivas se por mutación somática en cáncer. Los marcadores se pueden estudiar en combinación con otra medida del resultado del tratamiento, por ejemplo, marcadores bioquímicos (por ejemplo, proteína M en mieloma, marcador de salud renal como proteinuria, niveles séricos de proteína C reactiva o citoqueratina 19, fragmento de citoqueratina 21-1 (CYFRA21-1) para NSCLC, niveles de orina de fibrinógeno/productos de degradación de fibrinógeno para cáncer de vejiga, niveles de catecolaminas en orina o en sangre para neuroblastoma, niveles séricos de antígeno de carbohidrato 19-9 (CA 19-9) o perfil metabólico para cáncer de páncreas o niveles sanguíneos de péptidos solubles relacionados con mesotelina (SMRP) en mesotelioma) o evaluación histológica (por ejemplo, recuento de blastos, número de figuras mitóticas por unidad de área, medición de profundidad de invasión de tumores de melanoma, tumores esofágicos o tumores de vejiga).

Los métodos estadísticos pueden ayudar a determinar el resultado del tratamiento midiendo la cantidad de marcadores, por ejemplo, medición de ADN, ARN o proteína. La cantidad de un marcador se puede medir en múltiples puntos temporales, por ejemplo, antes del tratamiento, durante el tratamiento, después del tratamiento con un agente, por ejemplo, un inhibidor de NAE. Para determinar la progresión del cambio en la expresión de un marcador desde un valor de referencia, por ejemplo, con el tiempo, los resultados de la expresión pueden analizarse mediante un modelo de regresión lineal de medidas repetidas (Littell, Miliken, Stroup, Wolfinger, Schabenberger (2006) SAS para modelos mixtos, 2ª edición. SAS Institute, Inc., Cary, NC)):

Ecuación 1

$$Y_{ijk} - Y_{ij0} = Y_{ij0} + tratamiento_i + día_k + (\text{tratamiento} * día)_{ik} + \varepsilon_{ijk}$$

donde Y_{ijk} es la expresión transformada log_2 (normalizada para genes constitutivos) en el k^o día del j^o animal en el i^o tratamiento, Y_{ij0} es la expresión transformada log_2 del valor de referencia definido (normalizado a los genes constitutivos) del j^o animal en el i^o tratamiento, $día_k$ se trata como una variable categórica, y ε_{ijk} es el término de error residual. Una matriz de covarianza (por ejemplo, autorregresiva de primer orden, simetría compuesta, ley de potencia espacial) se puede especificar para modelar las mediciones repetidas en cada animal a lo largo del tiempo. Además, cada punto temporal de tratamiento se puede comparar con el mismo punto temporal en el grupo de vehículos para probar si el valor del tratamiento fue significativamente diferente del vehículo.

Para analizar los datos se pueden usar otros métodos. Por ejemplo, los valores de expresión relativos podrían analizarse en lugar del número de ciclo. Estos valores podrían examinarse como veces de cambio o como una diferencia absoluta con respecto al valor inicial. Adicionalmente, un análisis de varianza de medidas repetidas (ANOVA) se podría usar si las varianzas son iguales en todos los grupos y puntos temporales. El cambio observado desde el valor de referencia en el último (u otro) punto temporal podría analizarse utilizando un ensayo t pareado, un ensayo exacto de Fisher (valor p = ς P(X=x) de x = 1 al número de situaciones, por ejemplo, mutaciones, probó que muestra sensibilidad a la inhibición de NAE) para probar la importancia de los datos de tamaños de muestra pequeños, o un ensayo de rango con signo de Wilcoxon si los datos no se distribuyen normalmente, para comparar si un paciente con tumor era significativamente diferente de un sujeto normal.

Una diferencia en la cantidad de un punto temporal al siguiente o de la muestra tumoral a la muestra normal puede indicar el pronóstico del resultado del tratamiento. Un nivel de referencia midiendo la expresión en 1, 2, 3, 4 o más veces antes del tratamiento, por ejemplo, en el tiempo cero, un día, tres días, una semana y/o dos semanas o más antes del tratamiento. Alternativamente, Un nivel de referencia se puede determinar a partir de una serie de sujetos, por ejemplo, sujetos normales o pacientes con el mismo estado de salud o trastorno, que no se someten o aún no se han sometido al tratamiento, como se ha analizado anteriormente. Alternativamente, se pueden usar valores de expresión depositados con el programa Omnibus de expresión génica (GEO) en el Centro Nacional de Información

Biotecnológica (NCBI, Bethesda, MD). Por ejemplo, los conjuntos de datos de cantidades de expresión de ARNm de mieloma muestreados antes de la terapia de inhibición de proteasoma incluyen el número de acceso GEO GSE9782, también analizado en Mulligan, *et al*. (2006) Blood 109:3177-88 y GSE6477, también analizado por Chng *et al*. (2007) Cancer Res. 67:292-9. Para probar el efecto del tratamiento en el tumor, la expresión del marcador se puede medir en cualquier momento o varias veces después de algún tratamiento, por ejemplo, después de 1 día, 2 días, 3 días, 5 días, 1 semana, 2 semanas, 3 semanas, 4 semanas, 1 mes, 2 meses, 3 meses y/o 6 o más meses de tratamiento. Por ejemplo, la cantidad de un marcador se puede medir una vez después de algún tratamiento, o en intervalos múltiples, por ejemplo, 1 semana, 2 semanas, 4 semanas o 2 meses, 3 meses o intervalos más largos durante el tratamiento. En algunas realizaciones, la medición de un marcador durante el tratamiento se puede comparar con la misma medición del marcador al inicio del estudio. En otras realizaciones, la medición de un marcador durante el tratamiento se puede comparar con la misma medición de marcador en un punto temporal anterior. Por el contrario, para determinar el inicio de la enfermedad progresiva después de suspender la administración de un régimen terapéutico, la cantidad del marcador se puede medir en cualquier momento o varias veces después, *por ejemplo,* 1 día, 2 días, 3 días, 5 días, 1 semana, 2 semanas, 3 semanas, 4 semanas, 1 mes, 2 meses, 3 meses y/o 6 o más meses después del último tratamiento. La medición de un marcador después del tratamiento se puede comparar con la misma medición de marcador al final del tratamiento. Un experto en la materia determinará el punto temporal o los puntos temporales para evaluar la cantidad del marcador dependiendo de varios factores, *por ejemplo,* la farmacocinética del tratamiento, la duración del tratamiento, farmacodinámica del tratamiento, edad del paciente, la naturaleza del trastorno o mecanismo de acción del tratamiento. Una tendencia en la dirección negativa o una disminución en la cantidad relativa al valor de referencia o un patrón de expresión predeterminado de un marcador de sensibilidad a la terapia de inhibición de NAE, por ejemplo, una disminución en un marcador de sensibilidad identificado en la Tabla 3, puede indicar una disminución en la respuesta del tumor a la terapia, por ejemplo, aumento de la resistencia. Una tendencia hacia un resultado favorable en relación con el valor inicial o un patrón de expresión predeterminado de un marcador del resultado del tratamiento indica la utilidad del régimen terapéutico o el beneficio continuo de la terapia. Una tendencia hacia un aumento en un marcador de resistencia, por ejemplo, un aumento en un marcador de resistencia identificado en la Tabla 3, puede indicar un resultado desfavorable.

Cualquier marcador, por ejemplo, gen marcador o combinación de marcadores, por ejemplo*,* genes marcadores de la invención, o mutaciones de los mismos, así como cualquier marcador conocido en combinación con los marcadores, por ejemplo*,* genes marcadores de la invención, puede usarse en las composiciones, kits y métodos de la presente invención. En general, los marcadores se seleccionan por su capacidad tan grande como sea posible para juzgar el estado mutacional de un gen marcador para predecir el resultado del tratamiento con inhibidor de NAE. Por ejemplo, la elección de los marcadores se selecciona por la mayor diferencia posible entre la característica, por ejemplo, tamaño, secuencia, composición o cantidad del marcador en muestras que comprenden células tumorales y la característica, por ejemplo, tamaño, secuencia, composición o cantidad del mismo marcador en células de control. Aunque esta diferencia puede ser tan pequeña como el límite de detección del método para evaluar la cantidad del marcador, en otra realización, la diferencia puede ser al menos mayor que el error estándar del método de evaluación. En el caso de ARN o cantidad de proteína, una diferencia puede ser al menos 1,5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 100, 500, 1000 veces o superior. La cantidad "baja" de ARN o proteína puede ser esa expresión en relación con la media general de las muestras tumorales (por ejemplo, tumor hematológico, por ejemplo, mieloma) que es baja. En el caso de la cantidad de ADN, por ejemplo, número de copias, la cantidad es 0, 1, 2, 3, 4, 5, 6 o más copias. Una eliminación hace que el número de copias sea 0 o 1; una amplificación hace que el número de copias sea mayor que 2. La diferencia puede calificarse mediante un nivel de confianza, por ejemplo, $p < 0,05$, $p < 0,02$, $p < 0,01$ o valor p menor.

La medición de más de un marcador, por ejemplo, un conjunto de 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 20 o 25 o más marcadores puede proporcionar un perfil de expresión o una tendencia indicativa del resultado del tratamiento. En algunas realizaciones, el conjunto de marcadores comprende no más de 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 20 o 25 marcadores. En algunas realizaciones, el conjunto de marcadores incluye una pluralidad de loci de cromosomas,, una pluralidad de genes marcadores, o una pluralidad de marcadores de uno o más genes marcadores (por ejemplo, ácido nucleico y proteína, ADN genómico y ARNm, o varias combinaciones de marcadores descritos en este documento). El análisis del resultado del tratamiento mediante la evaluación de la cantidad de marcadores en un conjunto puede ir acompañado de un método estadístico, por ejemplo, un análisis de votación ponderado que tenga en cuenta las variables que pueden afectar la contribución de la cantidad de un marcador en el conjunto a la clase o tendencia del resultado del tratamiento, por ejemplo*,* la relación señal-ruido de la medición o la eficacia de hibridación para cada marcador. Un conjunto de marcadores, por ejemplo, un conjunto de 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 20 o 25 o más marcadores, puede comprender un cebador, sonda o cebadores para analizar al menos un marcador de ADN o ARN descrito en este documento, por ejemplo, un marcador en el cromosoma 22q del par de bases 29999545 a 30094589, cromosoma 18q del par de bases 48556583 a 48611412, cromosoma Xp del par de bases 44732423 a 44971847, cromosoma 4q del par de bases 153242410 a 153456172, cromosoma 17p del par de bases 7571720 a 7590868, cromosoma 9p del par de bases 21967751 a 21994490, NF2, SMAD4, KDM6A, FBXW7, TP53, CDKN2A, CDKN2A_p14, o un complemento de cualquiera de los anteriores. Un conjunto de marcadores, por ejemplo, un conjunto de al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 20 o 25 o más marcadores, puede comprender un cebador, sonda o cebadores para detectar al menos uno o al menos dos o más marcadores, o al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 20 o 25 o más mutaciones en el marcadores por ejemplo, de NF2, SMAD4, KDM6A, TP53,

CDKN2A, CDKN2A_p14 y/o FBXW7. En otra realización, un conjunto de marcadores puede comprender marcadores para evaluar las características de NF2, SMAD4 y/o KDM6A. En una realización, un conjunto de marcadores para cáncer de útero o cuello uterino comprende al menos un marcador para evaluar al menos una característica de FBXW7. En una realización, un marcador establecido para cáncer de intestino, mama, pulmón, cabeza y cuello, cuello uterino o piel comprende al menos un marcador para evaluar al menos una característica de TP53. En una realización, un conjunto de marcadores para cáncer de intestino comprende marcadores para evaluar al menos una característica de cada uno de TP53 y APC. En una realización, un conjunto de marcadores para cáncer de piel o sistema nervioso central comprende al menos un marcador para evaluar al menos una característica de CDKN2A_p14. En una realización, un conjunto de marcadores para cáncer de cabeza y cuello o piel comprende al menos un marcador para evaluar al menos una característica de CDKN2A. En una realización, un conjunto de marcadores para cáncer de cabeza y cuello comprende al menos un marcador para evaluar al menos una característica de SMAD4. Los conjuntos de marcadores seleccionados pueden ensamblarse a partir de los marcadores proporcionados en este documento o seleccionados entre marcadores utilizando métodos proporcionados en este documento y métodos análogos conocidos en la técnica. Una forma de calificar un nuevo marcador para usar en un ensayo de la invención es correlacionar el número de copias de ADN en una muestra que comprende células tumorales con diferencias en la expresión (por ejemplo, veces de cambio desde el inicio) de un marcador, por ejemplo, un gen marcador. Una forma útil de juzgar la relación es calcular el coeficiente de determinación r2, después de resolver para r, el coeficiente de correlación de momento del producto Pearson y/o la preparación de un gráfico de mínimos cuadrados, utilizando métodos estadísticos convencionales. Una correlación puede analizar el número de copias de ADN frente al nivel de expresión del marcador, por ejemplo, un gen marcador. Un producto génico se puede seleccionar como marcador si el resultado de la correlación (r2, por ejemplo, la pendiente lineal de los datos en este análisis), es al menos 0,1-0,2, al menos 0,3-0,5, o al menos 0,6-0,8 o superior. Los marcadores pueden variar con una correlación positiva a la respuesta, TTP o supervivencia (es decir, cambiar los niveles de expresión de la misma manera que el número de copias, por ejemplo, disminuir cuando disminuya el número de copias). Los marcadores que varían con una correlación negativa con el número de copias (es decir, cambian los niveles de expresión de manera opuesta a los niveles de números de copias, por ejemplo, aumentan cuando se reduce el número de copias) proporcionan una determinación incoherente del resultado.

Otra forma de calificar un nuevo marcador para uso en el ensayo sería analizar la expresión de un gran número de marcadores en varios sujetos antes y después del tratamiento con un agente de ensayo. Los resultados de la expresión permiten identificar los marcadores que muestran grandes cambios en una dirección dada después del tratamiento en relación con las muestras de pretratamiento. Un modelo de regresión lineal de medidas repetidas se puede construir para identificar los genes que muestran cambios o diferencias estadísticamente significativas. Para clasificar luego estos genes significativos, se puede calcular el área bajo el cambio de, por ejemplo, valor de referencia frente a curva de tiempo. Esto puede dar como resultado una lista de genes que mostrarían los mayores cambios estadísticamente significativos. Luego se pueden combinar varios marcadores en un conjunto mediante el uso de métodos tales como análisis de componentes principales, métodos de agrupamiento (por ejemplo, medias k, jerárquico), análisis de varianza multivariante (MANOVA), o técnicas de regresión lineal. Para usar dicho gen (o grupo de genes) como marcador, en el conjunto de marcadores se podrían incluir genes que muestran 2, 2,5, 3, 3,5, 4, 4,5, 5, 7, 10 veces, o más diferencias de expresión de la valor de referencia. Un perfil de expresión, por ejemplo, un compuesto de las diferencias de nivel de expresión con respecto al valor de referencia o referencia del conjunto de marcadores añadidos indicaría la tendencia, por ejemplo, si la mayoría de los marcadores muestran un resultado particular, por ejemplo, una diferencia significativa con respecto al valor de referencia o referencia, por ejemplo, 60 %, 70 %, 80 %, 90 %, 95 % o más marcadores; o más marcadores, por ejemplo, 10 % más, 20 % más, 30 % más, 40 % más, muestra un resultado significativo en una dirección que en la otra dirección.

En realizaciones cuando las composiciones, kits y métodos de la invención se usan para caracterizar el resultado del tratamiento en un paciente, el marcador o conjunto de marcadores de la invención se selecciona de modo que se obtenga un resultado significativo en al menos aproximadamente 20 %, al menos aproximadamente 40 %, 60 % u 80 %, o en prácticamente todos los pacientes tratados con el agente de ensayo. El marcador o conjunto de marcadores de la invención se puede seleccionar de modo que se obtenga un valor predictivo positivo (VPP) superior a aproximadamente 10 % para la población general y se pueda inferir una confianza adicional en un marcador cuando el PPV se combina con un ensayo especificidad superior al 80 %.

Agentes terapéuticos

Los marcadores y conjuntos de marcadores de la presente invención evalúan la probabilidad de un resultado favorable de la terapia (por ejemplo, sensibilidad a un agente terapéutico) en pacientes, por ejemplo, pacientes con cáncer, *por ejemplo,* pacientes con cáncer hematológico *(por ejemplo,* mieloma múltiple, leucemias, linfoma, etc.) o cáncer de tumor sólido (por ejemplo, cáncer de piel como melanoma, cáncer de cabeza y cuello, como cáncer de esófago, cáncer de vejiga, cáncer de pulmón, como cáncer de pulmón no microcítico (NSCLC), adenocarcinoma de pulmón, cáncer del sistema nervioso central, como metástasis pulmonares en el cerebro o neuroblastoma, cáncer de páncreas, cáncer de mama, mesotelioma, cáncer de cuello uterino o intestinal como adenocarcinoma de colon o recto), basándose en su capacidad de afectar la característica, por ejemplo, composición o cantidad de un marcador o marcadores de la invención. Usando esta predicción, las terapias contra el cáncer se pueden evaluar para diseñar un régimen terapéutico más adecuado para pacientes en cualquier categoría.

En particular, los métodos se pueden utilizar para predecir la sensibilidad del paciente a los inhibidores de NAE como se describe en secciones anteriores. Los agentes probados en los presentes métodos pueden ser un agente único o una combinación de agentes. Los métodos de la invención incluyen la combinación de terapia de inhibición de NAE con terapia de inhibición de proteasoma y/u otros agentes adicionales, por ejemplo, seleccionado del grupo que consiste en agentes quimioterapéuticos. Por ejemplo, los métodos actuales pueden usarse para determinar si un solo agente quimioterapéutico, como un inhibidor de NAE (por ejemplo, MLN4924) puede usarse para tratar un cáncer o si uno o más agentes deben usarse en combinación con el inhibidor de NAE (por ejemplo, MLN4924). Las combinaciones útiles pueden incluir agentes que tienen diferentes mecanismos de acción, por ejemplo, el uso de un agente antimitótico en combinación con un agente alquilante y un inhibidor de NAE.

Como se usa en este documento, la expresión "inhibidor del proteasoma" se refiere a cualquier sustancia que inhibe directamente la actividad enzimática del proteasoma 20S o 26S *in vitro* o *in vivo.* En algunas realizaciones, el inhibidor del proteasoma es un ácido peptidil borónico. En Adams *et al*., se describen ejemplos de inhibidores del proteasoma del ácido peptidil borónico adecuados para uso en los métodos de la invención, Patentes de Estados Unidos N.os 5.780.454 (1998), 6.066.730 (2000), 6.083.903 (2000); 6.297.217 (2001), 6.465.433 (2002), 6.548.668 (2003), 6.617.317 (2003) y 6.747.150 (2004). En algunas realizaciones, el inhibidor del proteasoma del ácido peptidil borónico se selecciona del grupo que consiste en: ácido N (4 morfolin)carbonil-β-(1-naftil)-L-alanina-L-leucina borónico; ácido N (8 quinolin)sulfonil-β-(1-naftil)-L-alanina-L-alanina-L-leucina borónico; ácido N (pirazin)carbonil-L-fenilalanina-L-leucina borónico, y ácido N (4 morfolin)-carbonil-[O-(2-piridilmetil)]-L-tirosina-L-leucina borónico. En una realización particular, el inhibidor del proteasoma es ácido N (pirazin)carbonil-L-fenilalanina-L-leucina borónico (bortezomib; VELCADE®; anteriormente conocido como MLN341 o PS-341). Las publicaciones describen el uso de los compuestos de éster borónico y ácido borónico desvelados para reducir la tasa de degradación de proteínas musculares, para reducir la actividad de NF-kB en una célula, para reducir la tasa de degradación de la proteína p53 en una célula, para inhibir la degradación de la ciclina en una célula, para inhibir el crecimiento de una célula cancerosa e inhibir la adhesión celular dependiente de NF-kB. Bortezomib inhibe específica y selectivamente el proteasoma al unirse fuertemente (Ki = 0,6 nM) a uno de los sitios activos de la enzima. Bortezomib es selectivamente citotóxico y tiene un nuevo patrón de citotoxicidad en ensayos del Instituto Nacional del Cáncer (NCI) *in vitro* e *in vivo*. Adams J, *et al*. Cancer Res 59:2615-22. (1999).

Adicionalmente, los inhibidores de proteasoma incluyen inhibidores de proteasoma de péptido aldehído (Stein *et al*., Patente de Estados Unidos N.º 5.693.617 (1997); Siman *et al*., publicación de patente internacional WO 91/13904; Iqbal *et al*., J. Med. Chem. 38:2276-2277 (1995); e Iinuma *et al.,* publicación de patente internacional WO 05/105826, inhibidores de proteasoma de peptidil epoxi cetona (Crews *et al.,* Patente de Estados Unidos N.º 6.831.099; Smyth *et al.,* publicación de patente internacional WO 05/111008; Bennett *et al.,* publicación de patente internacional WO 06/045066; Spaltenstein *et al.,* Tetrahedron Lett. 37:1343 (1996); Meng, Proc. Natl. Acad. Sci. 96: 10403 (1999); y Meng, Cancer Res. 59: 2798 (1999)), inhibidores del proteasoma de alfa-cetoamida (Chatterjee y Mallamo, patentes de Estados Unidos N.os 6.310.057 (2001) y 6.096.778 (2000); y Wang *et al.,* Patentes de Estados Unidos N.os 6.075.150 (2000) y 6.781.000 (2004)), inhibidores de proteasoma de peptidil vinil éster (Marastoni *et al.,* J. Med. Chem. 48:5038 (2005), y peptidil vinilsulfona e inhibidores del proteasoma 2-ceto-1,3,4-oxadiazol, como los descritos en Rydzewski *et al.,* J. Med. Chem. 49:2953 (2006); y Bogyo *et al.,* Proc. Natl. Acad. Sci. 94:6629 (1997)), azapeptoides y (Bouget *et al.,* Bioorg. Med. Chem. 11:4881 (2003); Baudy-Floc'h *et al.,* publicación de patente internacional WO 05/030707; y Bonnemains *et al.,* publicación de patente internacional WO 03/018557), oligopéptidos de efrapeptina (Papathanassiu, publicación de patente internacional WO 05/115431), lactacistina y salinosporamida y análogos de los mismos (Fenteany *et al*., patentes de Estados UnidosN.os 5.756.764 (1998), 6.147.223 (2000), 6.335.358 (2002) y 6.645.999 (2003); Fenteany *et al.,* Proc. Natl. Acad. Sci. USA (1994) 91:3358; Fenical *et al.,* publicación de patente internacional WO 05/003137; Palladino *et al.,* publicación de patente internacional WO 05/002572; Stadler *et al.,* publicación de patente internacional WO 04/071382; Xiao y Patel, publicación de patente de Estados Unidos 2005/023162; y Corey, publicación de patente internacional WO 05/099687).

Los agentes terapéuticos adicionales para usar en combinación con un inhibidor de NAE (por ejemplo, MLN4924) en los métodos de la invención pueden comprender una clase conocida de agentes terapéuticos que comprenden esteroides glucocorticoides. La terapia con glucocorticoides comprende generalmente al menos un agente glucocorticoide (por ejemplo, dexametasona). En ciertas aplicaciones de la invención, el agente utilizado en los métodos de la invención es un agente glucocorticoide. Un ejemplo de un glucocorticoide utilizado en el tratamiento de pacientes con mieloma múltiple y otras terapias contra el cáncer es dexametasona. Los glucocorticoides adicionales utilizados en el tratamiento de la terapia hematológica y combinada en tumores sólidos incluyen hidrocortisona, predisolona, prednisona, y triamcinolona.

Otros agentes terapéuticos para usar en combinación con la terapia de inhibición de NAE incluyen agentes quimioterapéuticos. Se pretende que un "agente quimioterapéutico" incluya reactivos químicos que inhiben el crecimiento de células o tejidos proliferantes en donde el crecimiento de tales células o tejidos es indeseable. Agentes quimioterapéuticos como agentes antimetabólicos, por ejemplo, Ara AC, 5-FU y metotrexato, agentes antimitóticos, por ejemplo, taxano, vinblastina y vincristina, agentes alquilantes, por ejemplo, melfanlán, Carmustina (BCNU) y mostaza nitrogenada, inhibidores de la topoisomerasa II, por ejemplo, VW-26, topotecán y bleomicina,

agentes de rotura de hebra, por ejemplo, doxorrubicina y mitoxantrona (DHAD), agentes de reticulación, por ejemplo, cisplatino y carboplatino (CBDCA), radiación y luz ultravioleta y se conocen bien en la técnica (véase por ejemplo, Gilman A.G., *et al*., The Pharmacological Basis of Therapeutics, 8ª edición, Sec 12:1202-1263 (1990)), y se usan típicamente para tratar enfermedades neoplásicas. Los ejemplos de agentes quimioterapéuticos empleados generalmente en tratamientos de quimioterapia se enumeran a continuación en la Tabla 2.

TABLA 2: Agentes quimioterapéuticos

CLASE	TIPO DE AGENTE	NOMBRES GENÉRICOS (OTROS NOMBRES)
Alquilante	Mostazas nitrogenadas	Mecloretamina (HN_2) Ciclofosfamida Ifosfamida Melfalán (L-sarcolisina) Clorambucilo
	Etileniminas Y metilmelaminas	Hexametilmelamina Tiotepa
	Sulfonatos de alquilo	Busulfán
Alquilante	Nitrosoureas	Carmustina (BCNU) Lomustina (CCNU) Semustina (metil-CCNU) Estreptozocina (estreptozotocina)
Alquilante	Triacenos	Decarbazina (DTIC; dimetiltriazenoimidazolcarboxamida)
	Alquilador	cis-diaminadicloroplatino II (CDDP)
Antimetabolitos	Análogos de ácido fólico	Metotrexato (ametoterina)
	Pirimidina Análogos	Fluorouracilo ('5-fluorouracilo; 5-FU) Floxuridina (fluorodesoxiuridina; FUdR) Citarabina (arabinósido de citosina)
	Análogos de purina e Inhibidores Relacionados	Mercaptopurina (6-mercaptopurina; 6-MP) Tioguanina (6-tioguanina; TG) Pentostatina (2' - desoxicoformicina)
Productos Naturales	Alcaloides de la Vinca	Vinblastina (VLB) Vincristina
	Topoisomerasa Inhibidores	Etopósido Tenipósido Camptotecina Topotecán 9-amino-campototecina CPT-11
	Antibióticos	Dactinomicina (actinomicina D) Adriamicina Daunorrubicina (daunomicina; rubindomicina) Doxorrubicina Bleomicina Plicamicina (mitramicina)
Productos Naturales		Mitomicina (mitomicina C) TAXOL Taxotere
	Enzimas	L-asparaginasa
	Modificadores Respuesta Biológica	Interferón alfa Interleuquina 2
	Complejos Coordinación Platino	cis-diaminadicloroplatino II (CDDP) Carboplatino
	Antracendiona	Mitoxantrona
	Urea sustituida	Hidroxiurea
Miscelánea Agentes	Metilhidracina Derivado Adrenocortical Supresor	Procarbazina (N-metilhidrazina, (MIH) Mitotano (o,p'-DDD) Aminoglutetimida
Hormonas y Antagonistas	Progestinas	Caproato de hidroxiprogesterona Acetato de medroxiprogesterona Acetato de megestrol
	Estrógenos	Dietilestilbestrol Etinilestradiol

(continuación)

CLASE	TIPO DE AGENTE	NOMBRES GENÉRICOS (OTROS NOMBRES)
	Antiestrógeno	Tamoxifeno
	Andrógenos	Propionato de testosterona Fluoximesterona
	Antiandrógeno	Flutamida
	Liberación de gonadotropina Análogo hormonal	Leuprolida

Los agentes probados en los presentes métodos pueden ser un agente único o una combinación de agentes. Por ejemplo, los métodos actuales pueden usarse para determinar si un solo agente quimioterapéutico, tales como metotrexato, se puede usar para tratar un cáncer o si se puede usar una combinación de dos o más agentes en combinación con un inhibidor de NAE (por ejemplo, MLN4924). Las combinaciones útiles pueden incluir agentes que tienen diferentes mecanismos de acción, por ejemplo, el uso de un agente antimitótico en combinación con un agente alquilante y un inhibidor de NAE.

Los agentes descritos en este documento pueden administrarse por cualquier vía, incluyendo por vía intradérmica, por vía subcutánea, por vía oral, intraarterial o intravenosa. En una realización, la administración se realizará por vía intravenosa. La administración parenteral se puede administrar en bolo o por infusión.

La concentración de un compuesto desvelado en una mezcla farmacéuticamente aceptable variará dependiendo de varios factores, incluyendo la dosis del compuesto a administrar, las características farmacocinéticas de los compuestos empleados y la vía de administración. El agente puede administrarse en una dosis única o en dosis repetidas. Los tratamientos pueden administrarse diariamente o con mayor frecuencia dependiendo de una serie de factores, incluyendo la salud general de un paciente, y la formulación y vía de administración del compuesto o compuestos seleccionados.

Detecciones para inhibidores de NAE

La divulgación proporciona métodos (también denominados en este documento "ensayos de detección") para identificar moduladores, *es decir,* compuestos o agentes candidatos o de ensayo (por ejemplo, proteínas, péptidos, peptidomiméticos, peptoides, moléculas pequeñas u otros fármacos) que se unen a NAE u otras proteínas variantes de la enzima E1, que tienen un efecto estimulante o inhibidor en, por ejemplo, NAE, u otra expresión enzimática E1 o actividad enzimática, o tienen un efecto estimulante o inhibidor en, por ejemplo, la expresión o actividad de un NAE u otro sustrato enzimático E1 o proteínas en la ruta enzimática E1, por ejemplo, en la ruta NAE, por ejemplo, con una relación con la actividad de una ligasa culina-RING. Los compuestos identificados así pueden usarse para modular la actividad de productos génicos diana (por ejemplo, NAE u otros genes enzimáticos E1) en un protocolo terapéutico, para elaborar la función biológica del producto génico diana, o para identificar compuestos que interrumpen NAE u otras interacciones de la ruta enzimática E1.

En una descripción, la divulgación proporciona un método para identificar un compuesto como inhibidor de NAE, por ejemplo, como agente que modula la resistencia farmacológica de una célula, contactando primero una célula que comprende al menos una mutación en al menos un gen marcador con un compuesto de ensayo y luego midiendo la viabilidad de la célula o la inhibición del crecimiento de la célula. En algunas descripciones, la célula comprende un gen de resistencia identificado en la Tabla 3. En otras descripciones, la célula comprende un gen de sensibilidad identificado en la Tabla 3. El efecto del inhibidor de NAE se puede comparar con una célula de control no expuesta al compuesto. En algunas descripciones, el efecto de un agente en una célula que comprende un gen marcador de sensibilidad se puede comparar con el efecto de un agente en una célula que comprende un gen marcador de resistencia (véase, por ejemplo, Tabla 3). El compuesto se identifica como modulador de la resistencia a fármacos o un agente inhibidor de NAE cuando la viabilidad celular o el crecimiento celular disminuye. Los compuestos identificados como inhibidor de NAE, por ejemplo, como resistencia moduladora, que se identifican en los métodos anteriores también se incluyen dentro de la invención.

Métodos de detección

Un principio general de los ensayos de pronóstico implica la preparación de una muestra o mezcla de reacción que puede contener un marcador y una sonda, en 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 eliminarse y/o detectarse en la mezcla de reacción. Estos ensayos se pueden realizar de varias maneras.

Por ejemplo, un método para realizar dicho ensayo implicaría anclar el marcador o la sonda en un soporte de fase sólida, también conocido como sustrato, y la detección de complejos de marcador/sonda diana anclados en la fase sólida al final de la reacción. En una realización de dicho método, una muestra de un sujeto, que se analizará para determinar la presencia y/o concentración del marcador, se puede anclar en un soporte o soporte de fase sólida. En

otra realización, la situación inversa es posible, en donde la sonda puede anclarse a una fase sólida y una muestra de un sujeto puede reaccionar como un componente no anclado del ensayo. Un ejemplo de tal realización incluye el uso de una matriz o chip que contiene un marcador predictivo o conjunto de marcadores anclado para el análisis de expresión de la muestra.

Existen muchos métodos establecidos para anclar los componentes del ensayo a una fase sólida. Estos incluyen, sin limitación, moléculas marcadoras o de sonda que se inmovilizan mediante conjugación de biotina y estreptavidina. Tales componentes de ensayo biotinilados pueden prepararse a partir de biotina-NHS (*N*-hidroxi-succinimida) utilizando técnicas conocidas en la técnica (por ejemplo, kit de biotinilación, Pierce Chemicals, Rockford, IL) e inmovilizados en los pocillos de placas de 96 pocillos recubiertas con estreptavidina (Pierce Chemical). En ciertas realizaciones, las superficies con componentes de ensayo inmovilizados pueden prepararse con antelación y almacenarse.

Otros vehículos adecuados o soportes de fase sólida para tales ensayos incluyen cualquier material capaz de unirse a la clase de molécula a la que pertenece el marcador o sonda. Los soportes o vehículos conocidos incluyen, pero no se limitan a, vidrio, poliestireno, nailon, polipropileno, nailon, polietileno, dextrano, amilasas, celulosas naturales y modificadas, poliacrilamidas, gabros y magnetita. Un experto en la materia conocerá muchos otros vehículos adecuados para unir anticuerpo o antígeno, y podrá adaptar dicho soporte para usar con la presente invención. Por ejemplo, la proteína aislada de las células puede ejecutarse en una electroforesis en gel de poliacrilamida e inmovilizarse en un soporte de fase sólida como nitrocelulosa. El soporte se puede lavar con tampones adecuados seguido de tratamiento con el anticuerpo marcado de forma detectable. El soporte de fase sólida se puede lavar con el tampón por segunda vez para eliminar el anticuerpo no unido. La cantidad de etiqueta unida en el soporte sólido se puede detectar por medios convencionales.

Para realizar ensayos con los enfoques mencionados anteriormente, el componente no inmovilizado se añade a la fase sólida en donde se ancla el segundo componente. Una vez completada la reacción, los componentes no complejados pueden eliminarse (por ejemplo, mediante lavado) en condiciones tales que cualquier complejo formado permanecerá inmovilizado en la fase sólida. La detección de complejos marcador/sonda anclados a la fase sólida se puede lograr en varios métodos descritos en este documento.

En una realización, la sonda, cuando es el componente de ensayo no anclado, puede marcarse con el fin de detectar y leer el ensayo, tanto directa como indirectamente, con etiquetas detectables discutidas en este documento y que son bien conocidas por un experto en la materia. El término "marcado", con respecto a la sonda (por ejemplo, ácido nucleico o anticuerpo), pretende incluir el marcado directo de la sonda mediante acoplamiento (es decir, enlace físico) de una sustancia detectable a la sonda, así como el marcado indirecto de la sonda por reactividad con otro reactivo que se marca directamente. Un ejemplo de marcado indirecto incluye detección de un anticuerpo primario usando un anticuerpo secundario marcado con fluorescencia. También es posible detectar directamente la formación de complejo marcador/sonda sin más manipulación o marcado de ninguno de los componentes (marcador o sonda), por ejemplo, utilizando la técnica de transferencia de energía de fluorescencia (FET, véase, por ejemplo, Lakowicz *et al*., Patente de Estados Unidos N.º 5.631.169; Stavrianopoulos, *et al*., Patente de Estados Unidos N.º 4.868.103). Se selecciona una etiqueta de fluoróforo en la primera molécula 'donante' de tal manera que, tras la excitación con luz incidente de longitud de onda apropiada, su energía fluorescente emitida será absorbida por una etiqueta fluorescente en una segunda molécula 'aceptora', que a su vez puede fluorescencia debido a la energía absorbida. Alternativamente, la molécula de proteína 'donante' puede utilizar simplemente la energía fluorescente natural de los restos de triptófano. Se eligen etiquetas que emiten diferentes longitudes de onda de luz, de modo que la etiqueta de la molécula 'aceptora' se pueda diferenciar de la del 'donante'. Dado que la eficacia de la transferencia de energía entre las etiquetas está relacionada con la distancia que separa las moléculas, se pueden evaluar las relaciones espaciales entre las moléculas. En una situación en donde se produce la unión entre las moléculas, la emisión fluorescente de la etiqueta de la molécula 'aceptora' en el ensayo debe ser máxima. Un suceso de unión a FET puede medirse convenientemente a través de medios de detección fluorométrica convencionales bien conocidos en la técnica (por ejemplo, usando un fluorímetro).

En otra realización, La determinación de la capacidad de una sonda para reconocer un marcador se puede lograr sin marcar ninguno de los componentes del ensayo (sonda o marcador) utilizando una tecnología como Análisis de interacción biomolecular en tiempo real (BIA) (véase, por ejemplo,, Sjolander, S. y Urbaniczky, C. (1991) Anal. Chem. 63:2338-2345 y Szabo *et al*. (1995) Curr. Opin. Struct. Biol. 5:699-705). Como se usa en este documento, "BIA" o "resonancia de plasmón superficial" es una tecnología para estudiar interacciones bioespecíficas en tiempo real, sin marcar ninguno de los interactuantes (por ejemplo, BIACORE™). Los cambios en la masa en la superficie de unión (indicativo de un suceso de unión) dan como resultado alteraciones del índice de refracción de la luz cerca de la superficie (fenómeno óptico de resonancia de plasmón superficial (SPR)), dando como resultado una señal detectable que puede usarse como indicación de reacciones en tiempo real entre moléculas biológicas.

Alternativamente, en otra realización, se pueden realizar ensayos de diagnóstico y pronóstico análogos con marcador y sonda como solutos en fase líquida. En dicho ensayo, el marcador complejo y la sonda están separados de los componentes no complejados por cualquiera de una serie de técnicas convencionales, incluyendo, pero sin limitación: centrifugación diferencial, cromatografía, electroforesis e inmunoprecipitación. En centrifugación

diferencial, los complejos marcador/sonda pueden separarse de los componentes de ensayo no complejados a través de una serie de etapas centrífugas, debido a los diferentes equilibrios de sedimentación de los complejos en función de sus diferentes tamaños y densidades (véase, por ejemplo, Rivas, G., y Minton, A.P. (1993) Trends Biochem Sci. 18:284-7). Las técnicas cromatográficas convencionales también se pueden utilizar para separar las moléculas complejas de las no complejas. Por ejemplo, la cromatografía de filtración en gel separa las moléculas en función del tamaño, y mediante la utilización de una resina de filtración en gel adecuada en un formato de columna, por ejemplo, el complejo relativamente más grande puede separarse de los componentes no complejados relativamente más pequeños. Igualmente, las propiedades de carga relativamente diferentes del complejo marcador/sonda en comparación con los componentes no complejados pueden explotarse para diferenciar el complejo de los componentes no complejados, por ejemplo mediante la utilización de resinas de cromatografía de intercambio iónico. Dichas resinas y técnicas cromatográficas son bien conocidas por un experto en la materia (véase, por ejemplo, Heegaard, N.H. (1998) J. Mol. Recognit. 11:141-8; Hage, D.S., y Tweed, S.A. (1997) J. Chromatogr. B. Biomed. Sci. Appl. 699:499-525). La electroforesis en gel también se puede emplear para separar componentes de ensayo complejos de componentes no unidos (véase, por ejemplo, Ausubel *et al*., ed., Current Protocols in Molecular Biology, John Wiley & Sons, Nueva York, 1987-1999). En esta técnica, los complejos de proteínas o ácidos nucleicos se separan según el tamaño o la carga, por ejemplo. En algunas realizaciones, los materiales y condiciones de la matriz de gel no desnaturalizante en ausencia de agente reductor se usan para mantener la interacción de unión durante el proceso electroforético. Las condiciones apropiadas para el ensayo particular y sus componentes serán bien conocidas por un experto en la materia.

El ARNm aislado puede usarse en ensayos de hibridación o amplificación que incluyen, pero no se limitan a, análisis de Southern o de Northern, reacción en cadena de la polimerasa y ensayos de expresión génica TAQMAN® (Applied Biosystems, Foster City, CA) y conjuntos de sondas. Un método de diagnóstico para detección de niveles de ARNm implica poner en contacto el ARNm aislado con una molécula de ácido nucleico (sonda, por ejemplo, una sonda de hibridación) que puede hibridarse con el ARNm codificado por el gen o mutante detectado. En algunas realizaciones, los ácidos nucleicos que comprenden mutaciones de genes marcadores pueden usarse como sondas o cebadores. Las sondas o cebadores de ácido nucleico de la invención pueden ser ADN monocatenario (por ejemplo, un oligonucleótido), ADN bicatenario (por ejemplo, oligonucleótido bicatenario) o ARN. Los cebadores de la invención se refieren a ácidos nucleicos que hibridan con una secuencia de ácidos nucleicos que es adyacente a la región de interés y puede extenderse sobre una región de interés, por ejemplo, en una reacción de extensión o amplificación del cebador, o que cubre la región de interés, por ejemplo, una región de ácido nucleico que comprende un gen marcador o mutación del mismo. Una sonda de ácido nucleico puede ser, por ejemplo, un ADNc de longitud completa o una parte del mismo, como un oligonucleótido de al menos 7, 15, 20, 25, 30, 50, 75, 100, 125, 150, 175, 200, 250 o 500 o más nucleótidos consecutivos de la secuencia de ácidos nucleicos marcadora o complemento de la misma y suficiente para específicamente hibridarse en condiciones estrictas con un ARNm o ADN genómico que codifica un marcador de la presente invención o un complemento del mismo. La longitud exacta de la sonda de ácido nucleico dependerá de muchos factores que los expertos en la materia consideran y practican habitualmente. Las sondas de ácido nucleico de la invención pueden prepararse mediante síntesis química usando cualquier metodología adecuada conocida en la técnica, pueden producirse mediante tecnología recombinante, o pueden derivarse de una muestra biológica, por ejemplo, por digestión de restricción. En este documento se describen otras sondas adecuadas para uso en ensayos de diagnóstico de la invención. La sonda puede comprender un grupo marcador unido a la misma, por ejemplo, un radioisótopo, un compuesto fluorescente, una enzima, un cofactor enzimático, un hapteno, una etiqueta de secuencia, una proteína o un anticuerpo. Los ácidos nucleicos pueden modificarse en el resto base, en el resto de azúcar, o en el esqueleto de fosfato. Se incorpora un ejemplo de etiqueta de ácido nucleico utilizando la tecnología de base modificada SUPER™ (Nanogen, Bothell, WA, véase Patente de Estados Unidos N.º 7.045.610). El nivel de expresión se puede medir como niveles generales de ácido nucleico, por ejemplo, después de medir los niveles de ADN amplificados (por ejemplo, usando un colorante intercalante de ADN, por ejemplo, el colorante verde SYBR (Qiagen Inc., Valencia, CA) o como ácidos nucleicos específicos, por ejemplo, utilizando un diseño basado en sonda, con las sondas marcadas. Los formatos de ensayo TAQMAN® pueden usar el diseño basado en sonda para aumentar la especificidad y la relación señal/ruido.

Tales cebadores o sondas se pueden usar como parte de un kit de ensayo de diagnóstico para identificar células o tejidos que expresan la proteína, como midiendo cantidades de una molécula de ácido nucleico transcrita en una muestra de células de un sujeto, por ejemplo, detectar transcripción, niveles de ARNm o determinar si un gen que codifica la proteína ha sido mutado o eliminado. La hibridación de un ARN o un ADNc con la sonda de ácido nucleico puede indicar que el marcador en cuestión se está expresando. La invención abarca además la detección de 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 (por ejemplo, proteína que tiene la secuencia de SEQ ID NOs: 3, 6, 10, 11, 14, 17, 20, 23 o 26), y codifican así la misma proteína. Los expertos en la materia observarán que los polimorfismos de la secuencia de ADN que conducen a cambios en la secuencia de aminoácidos pueden existir dentro de una población (por ejemplo, la población humana). Tales polimorfismos genéticos pueden existir entre individuos dentro de una población debido a la variación alélica natural. Un alelo es uno de un grupo de genes que ocurren alternativamente en un locus genético dado. Dichas variaciones alélicas naturales pueden dar como resultado una variación del 1-5 % en la secuencia de nucleótidos de un gen dado. Se pueden identificar alelos alternativos secuenciando el gen de interés en varios individuos diferentes, por ejemplo, muestras normales de individuos. Esto puede realizarse fácilmente mediante el uso de sondas de hibridación para

identificar el mismo locus genético en una variedad de individuos. Se pretende que la detección de cualquiera y todas esas variaciones de nucleótidos y los polimorfismos de aminoácidos resultantes o variaciones que son el resultado de la variación alélica natural y que no alteran la actividad funcional estén dentro del alcance de la invención. Además, se observará que también pueden existir polimorfismos de ADN que afectan los niveles de expresión de ARN que pueden afectar el nivel de expresión general de ese gen (por ejemplo, afectando regulación o degradación).

Como se usa en este documento, el término "hibrida" pretende describir condiciones la hibridación y lavado bajo las cuales las secuencias de nucleótidos que son significativamente idénticas u homólogas entre sí permanecen hibridadas entre sí. En algunas realizaciones, las condiciones son tales que las secuencias al menos aproximadamente 70 %, al menos aproximadamente 80 %, al menos aproximadamente 85 %, 90 % o 95 % idénticas entre sí permanecen hibridadas entre sí para su posterior amplificación y/o detección. Las condiciones estrictas varían según la longitud de la secuencia de nucleótidos involucrada, pero los expertos en la materia las conocen y pueden encontrarse o determinarse basándose en las enseñanzas de Current Protocols in Molecular Biology, Ausubel *et al.*, eds., John Wiley & Sons, Inc. (1995), secciones 2, 4 y 6. Las condiciones y fórmulas estrictas adicionales para determinar tales condiciones pueden encontrarse en Molecular Cloning: A Laboratory Manual, Sambrook et al., Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), capítulos 7, 9 y 11. Un ejemplo no limitante de condiciones de hibridación estrictas para híbridos que tienen al menos 10 pares de bases de longitud incluye hibridación en cloruro sódico 4X/citrato sódico (SSC), a aproximadamente 65-70 ºC (o hibridación en 4X SSC más 50 % de formamida a aproximadamente 42-50 ºC) seguido de uno o más lavados en IX SSC, a aproximadamente 65-70 ºC. Un ejemplo no limitante de condiciones de hibridación altamente estrictas para tales híbridos incluye hibridación en IX SSC, a aproximadamente 65-70 ºC (o hibridación en IX SSC más 50 % de formamida a aproximadamente 42-50 ºC) seguido de uno o más lavados en 0,3X SSC, a aproximadamente 65-70 ºC. Un ejemplo no limitante de condiciones de hibridación de rigurosidad reducida para tales híbridos incluye hibridación en 4X SSC, a aproximadamente 50-60 ºC (o alternativamente hibridación en 6X SSC más 50 % de formamida a aproximadamente 40-45 ºC) seguido de uno o más lavados en 2X SSC, a aproximadamente 50-60 ºC. También se pretende que los intervalos intermedios a los valores mencionados anteriormente, por ejemplo, a 65-70 ºC o a 42-50 ºC estén incluidos en la presente invención. Otro ejemplo de condiciones de hibridación rigurosas es la hibridación en cloruro sódico 6X/citrato sódico (SSC) a aproximadamente 45 ºC, seguido de uno o más lavados en 0,2x SSC, SDS al 0,1 % a 50-65 ºC. Otro ejemplo de tampón de hibridación riguroso es hibridación en NaCl 1 M, tampón de ácido 2-(N-morfolino)etanosulfónico (MES) 50 mM (pH 6,5), sarcosina sódica al 0,5 % y formamida al 30 %. SSPE (1xSSPE es NaCl 0,15 M, NaH_2PO_4 10 mM, y EDTA 1,25 mM, pH 7,4) puede ser sustituido por SSC (1xSSC es NaCl 0,15 M y citrato sódico 15 mM) en los tampones de hibridación y lavado; los lavados se realizan durante 15 minutos cada uno después de completar la hibridación. Se prevé que la temperatura de hibridación para híbridos con menos de 50 pares de bases de longitud sería 5-10 ºC inferior a la temperatura de fusión (T_m) del híbrido, donde T_m se determina según las siguientes ecuaciones. Para híbridos de menos de 18 pares de bases de longitud, T_m(ºC) = 2 (N.º de bases A + T) + 4 (N.º de bases G + C). Para híbridos entre 18 y 49 pares de bases de longitud, T_m(ºC) = 81,5 + 16,6($\log_{10}[Na^+]$) + 0,41 (% G + C) - (600/N), donde N es el número de bases en el híbrido y [Na+] es la concentración de iones sodio en el tampón de hibridación ($[Na^+]$ para 1xSSC = 0,165 M). El experto en la materia también reconocerá que se pueden añadir reactivos adicionales a la hibridación y/o tampones de lavado para disminuir la hibridación no específica de las moléculas de ácido nucleico a las membranas, por ejemplo, membranas de nitrocelulosa o nailon, incluyendo, pero sin limitarse a, agentes bloqueantes (por ejemplo, BSA o ADN vehículo de esperma de salmón o arenque), detergentes (por ejemplo, SDS), agentes quelantes (por ejemplo, EDTA), Ficoll, polivinilpirrolidona (PVP) y similares. Cuando se usan membranas de nailon, en particular, un ejemplo adicional no limitante de condiciones de hibridación rigurosas es la hibridación en NaH_2PO_4 0,25-0,5 M, SDS al 7 % a aproximadamente 65 ºC, seguido de uno o más lavados con NaH_2PO_4 0,02 M, SDS al 1 % a 65 ºC, véanse, por ejemplo, Church y Gilbert (1984) Proc. Natl. Acad. Sci. USA 81:1991-1995, (o alternativamente 0,2X SSC, SDS al 1 %). Una sonda de cebador o ácido nucleico se puede usar sola en un método de detección, o se puede usar un cebador junto con al menos otra sonda de cebador o ácido nucleico en un método de detección. Los cebadores también se pueden usar para amplificar al menos una parte de un ácido nucleico. Las sondas de ácido nucleico de la invención se refieren a ácidos nucleicos que hibridan con la región de interés y que no se extienden más. Por ejemplo, una sonda de ácido nucleico es un ácido nucleico que hibrida específicamente con una región mutante de un biomarcador, y que por hibridación o ausencia de hibridación con el ADN de un paciente o el tipo de híbrido formado puede ser indicativo de la presencia o identidad de la mutación del biomarcador o la cantidad de actividad del marcador.

En un formato, el ARN se inmoviliza sobre una superficie sólida y se pone en contacto con una sonda, por ejemplo, desarrollando el ARN aislado en un gel de agarosa y transfiriendo el ARN del gel a una membrana, tal como nitrocelulosa. En un formato alternativo, las sondas de ácido nucleico se inmovilizan sobre una superficie sólida y el ARN se pone en contacto con las sondas, por ejemplo, en una matriz de chip génico AFFYMETRIX® o un chip SNP (Santa Clara, CA) o un conjunto personalizado utilizando un conjunto de marcadores que comprende al menos un marcador indicativo del resultado del tratamiento. Un experto en la materia puede adaptar fácilmente métodos conocidos de detección de ARN y ADN para usar en la detección de la cantidad de marcadores de la presente invención. Por ejemplo, el análisis de micromatrices de alta densidad o ADN ramificado puede beneficiarse de una mayor concentración de células tumorales en la muestra, como una muestra que se había modificado para aislar células tumorales como se describió en secciones anteriores. En una realización relacionada, una mezcla de

polinucleótidos transcritos obtenidos de la muestra se pone en contacto con un sustrato que tiene fijado al mismo un polinucleótido complementario u homólogo con al menos una parte (por ejemplo, al menos 7, 10, 15, 20, 25, 30, 40, 50, 100, 500 o más restos de nucleótido) de un ácido nucleico marcador. Si los polinucleótidos complementarios u homólogos con el marcador son detectables diferencialmente en el sustrato (por ejemplo, detectables usando diferentes cromóforos o fluoróforos, o fijados a diferentes posiciones seleccionadas), entonces los niveles de expresión de una pluralidad de marcadores se pueden evaluar simultáneamente usando un solo sustrato (por ejemplo, una micromatriz de "chip génico" de polinucleótidos fijados en posiciones seleccionadas). En una realización cuando se usa un método para evaluar la expresión del marcador que implica la hibridación de un ácido nucleico con otro, la hibridación se puede realizar bajo condiciones de hibridación estrictas.

Un método alternativo para determinar la cantidad de ARN correspondiente a un marcador de la presente invención en una muestra implica el proceso de amplificación de ácido nucleico, por ejemplo, mediante RT-PCR (la realización experimental expuesta en Mullis, 1987, Patente de Estados Unidos N.º 4.683.202), reacción en cadena de ligasa (Barany, 1991, Proc. Natl. Acad. Sci. USA, 88:189-193), replicación de secuencia autosostenida (Guatelli *et al.*, 1990, Proc. Natl. Acad. Sci. USA 87:1874-1878), sistema de amplificación transcripcional (Kwoh *et al.*, 1989, Proc. Natl. Acad. Sci. USA 86:1173-1177), Q-Beta Replicasa (Lizardi *et al.*, 1988, Bio/Technology 6:1197), replicación de círculo rodante (Lizardi *et al.*, Patente de Estados Unidos N.º 5.854.033) o cualquier otro método de amplificación de ácido nucleico, seguido por la detección de las moléculas amplificadas usando técnicas bien conocidas por los expertos en la materia. Estos esquemas de detección son especialmente útiles para la detección de moléculas de ácido nucleico si tales moléculas están presentes en números muy bajos. Como se usa en este documento, Los cebadores de amplificación se definen como un par de moléculas de ácido nucleico que pueden hibridarse en regiones 5' o 3' de un gen (hebras sentido y antisentido, respectivamente, o viceversa) y contienen una región corta en medio. En general, los cebadores de amplificación tienen una longitud de aproximadamente 10 a aproximadamente 30 nucleótidos y flanquean una región de aproximadamente 50 a aproximadamente 200 nucleótidos de longitud. En condiciones apropiadas y con reactivos apropiados, dichos cebadores permiten la amplificación de una molécula de ácido nucleico que comprende la secuencia de nucleótidos flanqueada por los cebadores.

Para métodos *in situ*, el ARN no necesita aislarse de las células antes de la detección. En dichos métodos, Una muestra de células o tejidos se prepara/procesa utilizando métodos histológicos conocidos. La muestra se inmoviliza sobre un soporte, típicamente un portaobjetos de vidrio y luego se pone en contacto con una sonda que puede hibridarse con ARN que codifica el marcador.

En otra realización de la presente invención, se detecta un polipéptido correspondiente a un marcador. En algunas realizaciones, un agente para detectar un polipéptido de la invención es un anticuerpo capaz de unirse a un polipéptido correspondiente a un marcador de la invención. En realizaciones relacionadas, el anticuerpo tiene una etiqueta detectable. Los anticuerpos pueden ser policlonales o monoclonales. Se puede usar un anticuerpo intacto, o un fragmento del mismo (por ejemplo, Fab o $F(ab')_2$).

Para determinar si una muestra contiene una proteína que se une a un anticuerpo dado se puede usar una variedad de formatos. Ejemplos de tales formatos incluyen, pero no se limitan a, inmunoensayo enzimático (EIA), radioinmunoensayo (RIA), análisis de Western blot y ensayo de inmunoabsorción ligado a enzimas (ELISA). Un experto en la materia puede adaptar fácilmente métodos conocidos de detección de proteínas/anticuerpos para uso en la determinación de si los linfocitos B expresan un marcador de la presente invención.

Otro método para determinar el nivel de un polipéptido correspondiente a un marcador es la espectrometría de masas. Por ejemplo, proteínas o péptidos intactos, por ejemplo, péptidos trípticos pueden analizarse a partir de una muestra, por ejemplo, una muestra de sangre, una muestra de linfa u otra muestra, que contiene uno o más marcadores polipeptídicos. El método puede incluir además tratar la muestra para reducir las cantidades de proteínas abundantes, por ejemplo, albúmina sérica, para aumentar la sensibilidad del método. Por ejemplo, la cromatografía líquida se puede utilizar para fraccionar la muestra de modo que partes de la muestra se puedan analizar por separado mediante espectrometría de masas. Las etapas se pueden realizar en sistemas separados o en un sistema combinado de cromatografía líquida/espectrometría de masas (LC/MS, véase por ejemplo, Liao, *et al.* (2004) Arthritis Rheum. 50:3792-3803). El sistema de espectrometría de masas también puede estar en modo tándem (MS/MS). La distribución del estado de carga de la mezcla de proteínas o péptidos puede adquirirse en uno o múltiples barridos y analizarse mediante métodos estadísticos, por ejemplo, utilizando el tiempo de retención y la relación masa-carga (m/z) en el sistema LC/MS, para identificar proteínas expresadas a niveles estadísticamente significativos diferencialmente en muestras de pacientes sensibles o no sensibles a terapia de inhibición de NAE. Ejemplos de espectrómetros de masas que se pueden usar son un sistema de trampa iónica (ThermoFinnigan, San José, CA) o un espectrómetro de masas de tiempo de vuelo cuadrupolo (Applied Biosystems, Foster City, CA). El método puede incluir además la etapa de huellas de masa peptídica, por ejemplo, en un método de espectrometría de masas por ionización por desorción asistida por matriz con tiempo de vuelo (MALDI-TOF). El método puede incluir además la etapa de secuenciar uno o más de los péptidos trípticos. Los resultados de este método pueden usarse para identificar proteínas de bases de datos de secuencias primarias, por ejemplo, mantenido por el Centro Nacional de Información Biotecnológica, Bethesda, MD, o el Instituto Suizo de Bioinformática, Ginebra, Suiza, y basándose en picos de bases m/z de péptido tríptico por espectrometría de masas.

Matrices legibles con dispositivos electrónicos

Los dispositivos electrónicos, incluyendo matrices legibles que comprenden al menos un marcador predictivo de la presente invención también se contemplan para uso junto con los métodos de la invención. Como se usa en este documento, "medios legibles con dispositivos electrónicos" se refiere a cualquier medio adecuado para almacenar, retener o contener datos o información que se pueda leer y acceder directamente por un dispositivo electrónico. Como se usa en este documento, el término "dispositivo electrónico" pretende incluir cualquier dispositivo informático o de procesamiento adecuado u otro dispositivo configurado o adaptado para almacenar datos o información. Los ejemplos de dispositivos electrónicos adecuados para uso con la presente invención y la monitorización de la información registrada incluyen dispositivos informáticos independientes; redes, incluyendo una red de área local (LAN), una red de área amplia (WAN) de Internet, Intranet y Extranet; dispositivos electrónicos como asistentes digitales personales (PDA), teléfonos móviles, buscapersonas y similares; y sistemas de procesamiento local y distribuido. Como se usa en este documento, "registrado" se refiere a un proceso para almacenar o codificar información en el medio legible del dispositivo electrónico. Los expertos en la materia pueden adoptar fácilmente cualquiera de los métodos conocidos actualmente para registrar información en medios conocidos para generar fabricaciones que comprenden los marcadores de la presente invención.

Por ejemplo, los sistemas de micromatrices son bien conocidos y utilizados en la técnica para evaluación de muestras, ya sea por evaluación de la expresión génica (por ejemplo, detección de ADN, detección de ARN, detección de proteínas) o producción de metabolitos, por ejemplo. Las micromatrices para uso según la invención incluyen una o más sondas de marcador(es) predictivo de la invención característicos de respuesta y/o no respuesta a un régimen terapéutico como se describe en este documento. En una realización, la micromatriz comprende una o más sondas correspondientes a uno o más de los marcadores seleccionados del grupo que consiste en marcadores que demuestran una mayor expresión en supervivientes a corto plazo, y genes que demuestran una mayor expresión en supervivientes a largo plazo en pacientes. Los expertos en la materia conocen varias configuraciones y métodos de micromatrices diferentes para su producción, y se describen, por ejemplo, en las patentes de Estados Unidos N.os: 5.242.974; 5.384.261; 5.405.783; 5.412.087; 5.424.186; 5.429.807; 5.436.327; 5.445.934; 5.556.752; 5.405.783; 5.412.087; 5.424.186; 5.429.807; 5.436.327; 5.472.672; 5.527.681; 5.529.756; 5.545.531; 5.554.501; 5.561.071; 5.571.639; 5.593.839; 5.624.711; 5.700.637; 5.744.305; 5.770.456; 5.770.722; 5.837.832; 5.856.101; 5.874.219; 5.885.837; 5.919.523; 5981185; 6.022.963; 6.077.674; 6.156.501; 6261776; 6346413; 6440677; 6451536; 6576424; 6610482; 5.143.854; 5.288.644; 5.324.633; 5.432.049; 5.470.710; 5.492.806; 5.503.980; 5.510.270; 5.525.464; 5.547.839; 5.580.732; 5.661.028; 5.848.659; y 5.874.219; Shena, *et al*. (1998), Tibtech 16:301; Duggan *et al*. (1999) Nat. Genet. 21:10; Bowtell *et al*. (1999) Nat. Genet. 21:25; Lipshutz *et al*. (1999) Nature Genet. 21:20-24, 1999; Blanchard, *et al*. (1996) Biosensors and Bioelectronics, 11:687-90; Maskos, *et al*., (1993) Nucleic Acids Res. 21:4663-69; Hughes, *et al*. (2001) Nat. Biotechol. 19:342, 2001. Una micromatriz de tejido se puede usar para identificación de proteínas (véase Hans *et al*. (2004) Blood 103:275-282). Una micromatriz de fago-epítopo se puede usar para identificar una o más proteínas en una muestra en función de si la proteína o las proteínas inducen autoanticuerpos en el paciente (Bradford *et al*. (2006) Urol. Oncol. 24:237-242).

Por tanto una micromatriz comprende una o más sondas correspondientes a uno o más marcadores identificados en Este documento, por ejemplo, aquellas indicativas del resultado del tratamiento, por ejemplo, para identificar genes marcadores de tipo silvestre, variantes alélicas normales y mutaciones de genes marcadores. La micromatriz puede comprender sondas correspondientes a, por ejemplo, al menos 2, al menos 3, al menos 4, al menos 5, al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 35, al menos 40, al menos 45, al menos 50, al menos 75, o al menos 100, biomarcadores y/o mutaciones de los mismos indicativos del resultado del tratamiento. La micromatriz puede comprender sondas correspondientes a uno o más biomarcadores como se establece en este documento. Además, la micromatriz puede comprender conjuntos de marcadores completos como se establece en este documento y que se pueden seleccionar y agrupar según los métodos establecidos en este documento. La micromatriz puede usarse para analizar la expresión de uno o más marcadores predictivos o conjuntos de marcadores predictivos en la matriz. En un ejemplo, el conjunto puede usarse para analizar más de un marcador predictivo o expresión de conjunto de marcadores en una muestra para determinar un perfil de expresión de marcadores en el conjunto. De esta manera, se pueden analizar simultáneamente hasta 44.000 marcadores para determinar la expresión. Esto permite desarrollar un perfil de expresión que muestra una batería de marcadores expresados específicamente en una o más muestras. Además, esto permite desarrollar un perfil de expresión para evaluar el resultado del tratamiento.

La matriz también es útil para determinar patrones de expresión diferencial de uno o más marcadores en condiciones normales y anómalos (por ejemplo, muestra, por ejemplo, tumor). Esto proporciona una batería de marcadores que podrían servir como una herramienta para facilitar la identificación del resultado del tratamiento de los pacientes. Además, la matriz es útil para determinar la expresión de marcadores de referencia para niveles de expresión de referencia. En otro ejemplo, la matriz se puede usar para monitorizar el periodo de tiempo de expresión de uno o más marcadores en la matriz.

Además de dicha determinación cualitativa, la invención permite la cuantificación de la expresión del marcador. Así, los marcadores predictivos se pueden agrupar basándose en conjuntos de marcadores o indicaciones de resultado

por la cantidad del marcador en la muestra. Esto es útil, por ejemplo, para determinar el resultado de la muestra mediante la puntuación de las cantidades según los métodos proporcionados en este documento.

La matriz también es útil para determinar el efecto de la expresión de un marcador en la expresión de otros marcadores predictivos en la misma célula o en diferentes células. Esto proporciona, por ejemplo, una selección de dianas moleculares alternativas para intervención terapéutica si se predice que el paciente tendrá un resultado desfavorable.

Reactivos y Kits

La invención también incluye kits para detectar la presencia de un polipéptido o ácido nucleico correspondiente a un marcador de la invención en una muestra biológica (por ejemplo, una muestra de médula ósea, biopsia tumoral o una muestra de referencia). Dichos kits se pueden usar para determinar el estado mutacional de al menos un gen marcador para evaluar el resultado del tratamiento, por ejemplo, determinar si un sujeto puede tener un resultado favorable, por ejemplo, después del tratamiento con inhibidores de NAE. Por ejemplo, el kit puede comprender un compuesto o agente marcado capaz de detectar un segmento de ADN genómico, un polipéptido o un ARN transcrito correspondiente a un marcador de la invención o una mutación de un gen marcador en una muestra biológica y medios para determinar la cantidad del segmento de ADN genómico, el polipéptido o ARN en la muestra. Los reactivos adecuados para la unión con una proteína marcadora incluyen anticuerpos, derivados de anticuerpos, fragmentos de anticuerpos y similares. Las reactivos adecuados para unión con un ácido nucleico marcador (por ejemplo, un ADN genómico, un ARNm, un ARNm cortado y empalmado, un ADNc o similares) incluyen ácidos nucleicos complementarios. La etiqueta se puede unir directamente al agente de unión al marcador, por ejemplo, sonda, por ejemplo, reactivo de ácido nucleico como una sonda o cebador o reactivo de proteína, como un agente de unión o anticuerpo específico, o un reactivo secundario puede comprender una etiqueta para marcado indirecto. El kit también puede contener una muestra de control o de referencia o una serie de muestras de control o de referencia que se pueden analizar y comparar con la muestra de ensayo. Por ejemplo, el kit puede tener una muestra de control positivo, por ejemplo, incluyendo uno o más marcadores o mutaciones descritos en este documento, o marcadores de referencia, por ejemplo, marcadores constitutivos para estandarizar el ensayo entre muestras o puntos temporales o genomas de referencia, por ejemplo, de sujetos sin tumor por ejemplo, para establecer el valor de referencia del número de copias diploides o el nivel de expresión de referencia de un marcador. A modo de ejemplo, el kit puede comprender fluidos (por ejemplo, tampón) adecuado para hibridar ácidos nucleicos complementarios o para unir un anticuerpo con una proteína con la que se une específicamente y uno o más compartimentos de muestra. El kit de la invención puede comprender opcionalmente componentes adicionales útiles para realizar los métodos de la invención, por ejemplo, un recipiente de recogida de muestras, por ejemplo, un tubo, y opcionalmente, medios para optimizar la cantidad de marcador detectado, por ejemplo, si puede haber tiempo o condiciones adversas de almacenamiento y manipulación entre el momento del muestreo y el momento del análisis. Por ejemplo, el kit puede contener medios para aumentar el número de células tumorales en la muestra, como se ha descrito anteriormente, un agente tamponante, un conservante, un agente estabilizante o reactivos adicionales para preparación de material celular o sondas para usar en los métodos proporcionados; y etiqueta detectable, solo o conjugado o incorporado dentro de las sondas proporcionadas. En una realización de ejemplo, un kit que comprende un recipiente de recogida de muestras puede comprender por ejemplo, un tubo que comprende anticoagulante y/o estabilizador, por ejemplo, se puede añadir un estabilizador de ARN, como se describió anteriormente, o conocido por los expertos en la materia. El kit puede comprender además componentes necesarios para detectar la etiqueta detectable (por ejemplo, una enzima o un sustrato). Para conjuntos de marcadores, el kit puede comprender una matriz o chip de conjunto de marcadores para usar en la detección de biomarcadores. Los kits también pueden incluir instrucciones para interpretar los resultados obtenidos con el kit. El kit puede contener reactivos para detectar uno o más biomarcadores, por ejemplo, 2, 3, 4, 5 o más biomarcadores descritos en este documento.

En una realización, el kit comprende una sonda para detectar al menos un biomarcador, por ejemplo, un marcador indicativo del resultado del tratamiento (por ejemplo, en tratamiento con inhibidor de NAE). En una realización ejemplar, el kit comprende una sonda de ácido nucleico para detectar un gen marcador seleccionado del grupo que consiste en SEQ ID NO: 1, 2, 4, 5, 7, 8, 9, 12, 13, 15, 16, 18, 19, 21, 22, 24, 25 o una secuencia en el cromosoma 22q del par de bases 29999545 a 30094589, cromosoma 18q del par de bases 48556583 a 48611412, cromosoma Xp del par de bases 44732423 a 44971847, cromosoma 4q del par de bases 153242410 a 153456172, cromosoma 17p del par de bases 7571720 a 7590868, cromosoma 9p del par de bases 21967751 a 21994490, o un complemento de cualquiera de los anteriores o SEQ ID NO: 3, 6, 10, 11, 14, 17, 20, 23 y/o 26. En algunas realizaciones, el kit comprende una sonda para detectar un marcador seleccionado del grupo que consiste en NF2, SMAD4, KDM6A, FBXW7, TP53, CDKN2A, CDKN2A_p14 y APC. En otras realizaciones, el kit comprende una sonda para detectar una mutación en un gen marcador seleccionado del grupo que consiste en NF2, SMAD4, KDM6A, FBXW7, TP53, CDKN2A, CDKN2A_p14 y APC. En una realización, un kit comprende sondas para detectar un conjunto de marcadores que comprende dos o más marcadores del grupo que consiste en NF2, SMAD4, KDM6A, FBXW7, TP53, CDKN2A, CDKN2A_p14 y APC. En otra realización, un kit comprende una sonda para detectar FBXW7 en cáncer de útero o cuello uterino. En una realización, un kit comprende una sonda para detectar TP53 en cáncer de intestino, mama, pulmón, cabeza y cuello, cuello uterino o piel. En una realización, un kit comprende una sonda para detectar TPC y APC en cáncer de intestino. En una realización, un kit comprende una sonda para

detectar CDKN2A_p14 en cáncer de piel o sistema nervioso central. En una realización, un kit comprende una sonda para detectar CDKN2A en cáncer de cabeza, cuello o piel. En una realización, un kit comprende una sonda para detectar SMAD4 en el cáncer de cabeza y cuello. En realizaciones relacionadas, el kit comprende una sonda de ácido nucleico que comprende o deriva de (por ejemplo, un fragmento, mutante o variante (por ejemplo, homóloga o complementaria) de la misma) una secuencia de ácido nucleico seleccionada del grupo que consiste en SEQ ID NO: 1, 2, 4, 5, 7, 8, 9, 12, 13, 15, 16, 18, 19, 21, 22, 24, y 25. Para kits que comprenden sondas de ácido nucleico, por ejemplo, kits basados en oligonucleótidos, el kit puede comprender, por ejemplo: uno o más reactivos de ácido nucleico, como un oligonucleótido (marcado o no marcado) que hibrida con una secuencia de ácido nucleico correspondiente a un marcador de la invención, opcionalmente fijado a un sustrato; y opcionalmente puede comprender además oligonucleótidos marcados no unidos con un sustrato, un cebador, un par de cebadores de PCR, por ejemplo, útiles para amplificar una molécula de ácido nucleico correspondiente a un marcador de la invención, sondas de baliza molecular, un conjunto de marcadores que comprende oligonucleótidos que hibridan con al menos dos secuencias de ácido nucleico correspondientes a marcadores de la invención, y similares. El kit puede contener un agente estabilizador de ARN.

Para kits que comprenden sondas de proteínas, por ejemplo, kits basados en ligandos o anticuerpos, el kit puede comprender, por ejemplo: (1) un primer anticuerpo (por ejemplo, unido a un soporte sólido) que se une a un polipéptido correspondiente a un marcador de la invención; y, opcionalmente, (2) un segundo, anticuerpo diferente que se une al polipéptido o al primer anticuerpo y se conjuga con un marcador detectable. El kit puede contener un agente estabilizador de proteínas. El kit puede contener reactivos para reducir la cantidad de unión no específica de material no biomarcador de la muestra a la sonda. Los ejemplos de reactivos para reducir la unión inespecífica incluyen detergentes no ioinicos, soluciones que contienen proteínas no específicas, como los que contienen albúmina o caseína, u otras sustancias conocidas por los expertos en la materia.

Un polipéptido aislado correspondiente a un marcador predictivo de la invención, o un fragmento o mutante del mismo, puede usarse como inmunógeno para generar anticuerpos usando técnicas convencionales para preparación de anticuerpos policlonales y monoclonales. Por ejemplo, un inmunógeno se usa típicamente para preparar anticuerpos inmunizando un sujeto adecuado (es decir, inmunocompetente) como conejo, cabra, ratón u otro mamífero o vertebrado. En otro aspecto adicional, la invención proporciona anticuerpos monoclonales o fragmentos de unión a antígeno de los mismos, anticuerpos o fragmentos que se unen específicamente a un polipéptido que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en las secuencias de aminoácidos de la presente invención, una secuencia de aminoácidos codificada por el ADNc de la presente invención, un fragmento de al menos 8, 10, 12, 15, 20 o 25 residuos de aminoácidos consecutivos de una secuencia de aminoácidos de la presente invención, una secuencia de aminoácidos que es al menos 95 %, 96 %, 97 %, 98 % o 99 % idéntica a una secuencia de aminoácidos de la presente invención (en donde el porcentaje de identidad se determina usando el programa ALIGN del paquete de software GCG con una tabla de restos de peso PAM120, una penalización de longitud de espacio de 12, y una penalización de espacio de 4) y una secuencia de aminoácidos que está codificada por una molécula de ácido nucleico que se hibrida con una molécula de ácido nucleico que consiste en las moléculas de ácido nucleico de la presente invención, o un complemento del mismo, bajo condiciones de hibridación de 6X SSC a 45 °C y lavado en 0,2 X SSC, SDS al 0,1 % a 65 °C. Los anticuerpos monoclonales pueden ser anticuerpos humanos, humanizados, quiméricos y/o no humanos. Una preparación inmunogénica apropiada puede contener, por ejemplo, polipéptido recombinantemente expresado o sintetizado químicamente. La preparación puede incluir además un adyuvante, como adyuvante completo o incompleto de Freund, o un agente inmunoestimulador similar.

En la técnica se conocen métodos para fabricar anticuerpos humanos. Un método para hacer anticuerpos humanos emplea el uso de animales transgénicos, como un ratón transgénico. Estos animales transgénicos contienen una parte sustancial del genoma productor de anticuerpos humanos insertado en su propio genoma y la producción de anticuerpos endógenos propios del animal se vuelve deficiente en la producción de anticuerpos. En la técnica se conocen métodos para fabricar tales animales transgénicos. Tales animales transgénicos se pueden preparar usando la tecnología XENOMOUSE™ o usando un enfoque de "minilocus". Los métodos para preparar XENOMICE™ se describen en las patentes de Estados Unidos N.os 6.162.963, 6.150.584, 6.114.598 y 6.075.181. Los métodos para preparar animales transgénicos usando el enfoque de "minilocus" se describen en las patentes de Estados Unidos N.os 5.545.807, 5.545.806 y 5.625.825; véase también la publicación internacional número WO93/12227.

Los anticuerpos incluyen moléculas de inmunoglobulina y partes inmunológicamente activas de moléculas de inmunoglobulina, es decir, moléculas que contienen un sitio de unión a antígeno que se une específicamente a un antígeno, como un polipéptido de la invención, por ejemplo, un epítopo de un polipéptido de la invención. Una molécula que se une específicamente a un polipéptido dado de la invención es una molécula que se une al polipéptido, pero no une sustancialmente otras moléculas en una muestra, por ejemplo, una muestra biológica, que naturalmente contiene el polipéptido. Por ejemplo, los propios fragmentos de unión a antígeno, así como polipéptidos monoméricos, diméricos o triméricos de longitud completa derivados de los anticuerpos descritos anteriormente son útiles. Los homólogos de anticuerpos útiles de este tipo incluyen (i) un fragmento Fab, un fragmento monovalente que consiste en VL, VH, CL y CP1; (ii) un fragmento $F(ab')_2$, un fragmento bivalente que comprende dos fragmentos Fab unidos mediante un puente disulfuro en la región de bisagra; (iii) un fragmento Fd que consiste en los dominios

VH y CH1; (iv) un fragmento Fv que consiste en los dominios VL y VH de un único brazo de un anticuerpo; (v) un fragmento dAb (Ward *et al*., Nature 341:544-546 (1989)), que consiste en un dominio VH; (vii) un anticuerpo de cadena pesada funcional de dominio único, que consiste en un dominio VHH (conocido como nanocuerpo) véase por ejemplo, Cortez-Retamozo, *et al*., Cancer Res. 64: 2853-2857 (2004), y referencias citadas en este; y (vii) una región determinante de complementariedad aislada (CDR), por ejemplo, una o más CDR aisladas junto con un marco suficiente para proporcionar un fragmento de unión a antígeno. Además, aunque los dos dominios del fragmento Fv, VL y VH, están codificados por genes separados, se pueden unir, utilizando métodos recombinantes, mediante un enlazador sintético que les permite formarse como una sola cadena de proteína en la que las regiones VL y VH se aparean para formar moléculas monovalentes (conocidas como Fv de cadena única (scFv); véanse, por ejemplo, Bird *et al*. Science 242:423-426 (1988); y Huston *et al*. Proc. Natl. Acad. Sci. USA 85:5879-5883 (1988). Estos anticuerpos de cadena sencilla también se pretende que entren dentro del término "fragmento de unión a antígeno" de un anticuerpo. Estos fragmentos de anticuerpo se obtienen usando técnicas convencionales conocidas por aquellos expertos en la materia y se seleccionan en función de su utilidad del mismo modo que los anticuerpos intactos. Se pueden preparar fragmentos de anticuerpo, tal como Fv, $F(ab')_2$ y Fab por escisión de la proteína intacta, por ejemplo, por escisión con proteasa o química. La invención proporciona anticuerpos policlonales y monoclonales. La presente invención también contempla variantes sintéticas y modificadas genéticamente (véase Patente de Estados Unidos N.º 6.331.415) de cualquiera de los anteriores. Los anticuerpos policlonales y monoclonales se pueden producir mediante una variedad de técnicas, incluyendo la metodología convencional de anticuerpos monoclonales murinos, por ejemplo, la técnica convencional de hibridación de células somáticas de Kohler y Milstein, Nature 256: 495 (1975), la técnica del hibridoma de linfocitos B humanos (véase Kozbor *et al*., 1983, Immunol. Today 4:72), la técnica de EBV-hibridoma (véase Cole *et al*., pp. 77-96 en Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., 1985) o técnicas de trioma. Véase, en general, (Harlow, E. y Lane, D. (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; y Current Protocols in Immunology, Coligan *et al*. ed., John Wiley & Sons, Nueva York, 1994. Para aplicaciones de diagnóstico, los anticuerpos pueden ser anticuerpos monoclonales, por ejemplo, generados en ratón, rata o conejo. Adicionalmente, para uso en aplicaciones *in vivo*, los anticuerpos de la presente invención pueden ser anticuerpos humanos o humanizados. Las células de hibridoma que producen un anticuerpo monoclonal de la invención se detectan seleccionando los sobrenadantes de cultivo de hibridoma en busca de anticuerpos que se unan al polipéptido de interés, por ejemplo, usando un ensayo ELISA convencional.

Si se desea, las moléculas de anticuerpos se pueden cosechar o aislar del sujeto (por ejemplo, de la sangre o suero del sujeto) y purificar adicionalmente mediante técnicas bien conocidas, como cromatografía de proteína A para obtener la fracción de IgG. Alternativamente, se pueden seleccionar anticuerpos específicos para una proteína o polipéptido de la invención o (por ejemplo, parcialmente purificado) o purificar mediante, por ejemplo, cromatografía de afinidad para obtener anticuerpos sustancialmente purificados y purificados. Por composición de anticuerpo sustancialmente purificada se entiende, en este contexto, que la muestra de anticuerpos contiene como máximo solo el 30 % (en peso seco) de anticuerpos contaminantes dirigidos contra epítopos distintos de los de la proteína o polipéptido deseado de la invención, y como máximo el 20 %, como máximo el 10 %, o como máximo el 5 % (en peso seco) de la muestra son anticuerpos contaminantes. Una composición de anticuerpo purificada significa que al menos el 99 % de los anticuerpos en la composición están dirigidos contra la proteína o polipéptido deseado de la invención.

Un anticuerpo dirigido contra un polipéptido correspondiente a un marcador de la invención (por ejemplo, un anticuerpo monoclonal) puede usarse para detectar el marcador (por ejemplo, en una muestra celular) para evaluar el nivel y el patrón de expresión del marcador. Los anticuerpos también se pueden usar como diagnóstica para controlar los niveles de proteínas en tejidos o fluidos corporales (por ejemplo, en una muestra de sangre u orina) como parte de un procedimiento de ensayo clínico, por ejemplo, para, por ejemplo, determinar la eficacia de un régimen de tratamiento concreto. La detección puede facilitarse acoplando el anticuerpo a una sustancia detectable. Los ejemplos de sustancias detectables incluyen diversas enzimas, grupos prostéticos, materiales fluorescentes, materiales luminiscentes, materiales bioluminiscentes y materiales radiactivos. Los ejemplos de enzimas adecuadas incluyen peroxidasa de rábano picante, fosfatasa alcalina, β-galactosidasa o acetilcolinesterasa; los ejemplos de complejos con grupos prostéticos adecuados incluyen estreptavidina/biotina y avidina/biotina; los ejemplos de materiales fluorescentes adecuados incluyen umbeliferona, fluoresceína, isotiocianato de fluoresceína, rodamina, diclorotriazinilamina fluoresceína, cloruro de dansilo o ficoeritrina; un ejemplo de un material luminiscente incluye luminol; los ejemplos de materiales bioluminiscentes incluyen luciferasa, luciferina y aequorina, y ejemplos de material radiactivo adecuado incluyen ^{125}I, ^{131}I, ^{35}S o ^{3}H.

En consecuencia, en un aspecto, la invención proporciona anticuerpos sustancialmente purificados o fragmentos de los mismos, y anticuerpos no humanos o fragmentos de los mismos, anticuerpos o fragmentos que se unen específicamente a un polipéptido que comprende una secuencia de aminoácidos codificada por un marcador identificado en este documento. Los anticuerpos sustancialmente purificados de la invención, o fragmentos de los mismos, pueden ser anticuerpos humanos, no humanos, quiméricos y/o humanizados.

En otro aspecto, la invención proporciona anticuerpos no humanos o fragmentos de los mismos, anticuerpos o fragmentos que se unen específicamente a un polipéptido que comprende una secuencia de aminoácidos que está codificada por una molécula de ácido nucleico de un marcador predictivo de la invención. Tales anticuerpos no

humanos pueden ser anticuerpos de cabra, ratón, oveja, caballo, pollo, conejo o rata. Alternativamente, 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 o anticuerpos monoclonales.

Los anticuerpos sustancialmente purificados o sus fragmentos pueden unirse específicamente a un péptido señal, una secuencia secretada, un dominio extracelular, un dominio transmembrana o citoplasmático o un bucle citoplasmático de un polipéptido de la invención. Los anticuerpos sustancialmente purificados o fragmentos de los mismos, los anticuerpos no humanos o fragmentos de los mismos, y/o los anticuerpos monoclonales o fragmentos de los mismos, de la invención se unen específicamente a una secuencia secretada o un dominio extracelular de las secuencias de aminoácidos de la presente invención.

La invención también proporciona un kit que contiene un anticuerpo de la invención conjugado con una sustancia detectable, e instrucciones de uso. Además otro aspecto de la invención es una composición de diagnóstico que comprende una sonda de la invención y un vehículo farmacéuticamente aceptable. En una realización, la composición de diagnóstico contiene un anticuerpo de la invención, un resto detectable y un vehículo farmacéuticamente aceptable.

Ensayos de sensibilidad

Una muestra de células cancerosas se obtiene de un paciente. Un nivel de expresión se mide en la muestra para un marcador correspondiente a al menos uno de los marcadores descritos en este documento. Se puede utilizar un conjunto de marcadores que comprenda marcadores identificados escritos de este documento, y juntarlos en un conjunto de marcadores usando los métodos descritos en este documento. Tal análisis se utiliza para obtener un perfil de expresión tumoral en el paciente. La evaluación del perfil de expresión se utiliza para determinar si se espera que el paciente tenga un resultado favorable y se beneficie del tratamiento, por ejemplo, terapia de inhibición de NAE (por ejemplo, tratamientos con un inhibidor de NAE (por ejemplo, MLN4924) solo o en combinación con agentes adicionales) o un agente alternativo que se espera que tenga un efecto similar en la supervivencia. La evaluación del perfil de expresión también se puede usar para determinar si se espera que un paciente tenga un resultado desfavorable y se beneficiaría de una terapia contra el cáncer que no sea la terapia de inhibición de NAE o se beneficiaría de un régimen de terapia de inhibición de NAE alterado. La evaluación puede incluir el uso de un conjunto de marcadores preparado utilizando cualquiera de los métodos proporcionados u otros métodos de puntuación similares conocidos en la técnica (por ejemplo, votación ponderada, combinación de características de umbral (CTF), Análisis de riesgos proporcionales de Cox, puntuación de componentes principales, puntuación predictiva lineal, vecino más cercano a K, etc.), por ejemplo, utilizando valores de expresión depositados con el programa Gene Expresion Omnibus (GEO) en el Centro Nacional de Información Biotecnológica (NCBI, Bethesda, MD). Además, la evaluación puede comprender el uso de más de un conjunto de marcadores preparad. Una terapia de inhibición de NAE se identificará como apropiada para tratar el cáncer cuando el resultado de la evaluación demuestre un resultado favorable o se identifique un régimen de terapia más agresivo para un paciente con un resultado desfavorable esperado.

En un aspecto, la invención presenta un método para evaluar a un paciente, por ejemplo, un paciente con cáncer, por ejemplo, un cáncer hematológico (por ejemplo, mieloma múltiple, leucemias, linfoma, etc.) o cáncer de tumor sólido (por ejemplo, melanoma, cáncer de esófago o cáncer de vejiga) para el resultado del tratamiento. El método incluye proporcionar una evaluación de la expresión de los marcadores en un conjunto de marcadores en el paciente, en donde el conjunto de marcadores tiene las siguientes propiedades: incluye una pluralidad de genes, cada uno de los cuales se expresa diferencialmente entre pacientes con resultados identificados y sujetos no afectados y contiene un número suficiente de marcadores expresados diferencialmente de tal manera que la cantidad diferencial (por ejemplo, en comparación con un nivel en una muestra de referencia no afectada) de cada uno de los marcadores en el marcador establecido en un sujeto es predictivo del resultado del tratamiento con no más de aproximadamente el 15 %, aproximadamente 10 %, aproximadamente 5 %, aproximadamente 2,5 %, o aproximadamente 1 % de falsos positivos (en donde falso positivo significa predecir que un paciente responde o no responde cuando el sujeto no lo es); y proporcionar una comparación de la cantidad de cada uno de los marcadores en el conjunto del paciente con un valor de referencia, evaluando así al paciente.

Examinar la cantidad de uno o más de los marcadores o conjuntos de marcadores identificados en una muestra de tumor tomada de un paciente durante el curso de la terapia de inhibición de NAE, también es posible para determinar si el agente terapéutico continúa funcionando o si el cáncer no responde (refractario) al protocolo de tratamiento. Por ejemplo, a un paciente que recibe un tratamiento de MLN4924 se le extirparían las células tumorales y se controlaría la expresión de un marcador o conjunto de marcadores. Si el perfil de la cantidad de uno o más marcadores identificados en este documento tipifica más un resultado favorable en presencia del agente, por ejemplo, el inhibidor de NAE, el tratamiento continuaría. Sin embargo, si el perfil de la cantidad de uno o más marcadores identificados en este documento tipifica más resultados desfavorables en presencia del agente, entonces el cáncer puede haberse vuelto resistente a la terapia, por ejemplo, terapia de inhibición de NAE y otro protocolo de tratamiento debería iniciarse para tratar al paciente. Por ejemplo, el cáncer puede comprender una mutación en un gen marcador asociado con resistencia a la inhibición de NAE.

Importante, Estas determinaciones se pueden hacer paciente por paciente o agente por agente (o combinaciones de agentes). Así, se puede determinar si es probable que una terapia de inhibición de NAE particular beneficie a un paciente particular o grupo/clase de pacientes, o si se debe continuar un tratamiento particular.

Uso de información

En un método de divulgación, la información, por ejemplo, sobre el estado mutacional del tumor de un paciente, por ejemplo, la característica de marcador(es) del paciente, por ejemplo, tamaño, secuencia, composición o cantidad (por ejemplo, el resultado de evaluar un marcador o conjunto de marcadores descrito en este documento), o sobre si se espera que un paciente tenga un resultado favorable, se proporciona (por ejemplo, comunica, por ejemplo, se comunican electrónicamente) a un tercero, por ejemplo, un hospital, clínica, una entidad gubernamental, parte de reembolso o compañía de seguros (por ejemplo, una compañía de seguros de vida). Por ejemplo, la elección del procedimiento médico, si pagar por un procedimiento médico, pagar por una parte que reembolsa, o el coste de un servicio o seguro puede ser función de la información. Por ejemplo, el tercero recibe la información, toma una determinación basada al menos en parte en la información, y opcionalmente comunica la información o elige un procedimiento, pago, nivel de pago, cobertura, etc., basado en la información. En el método, se determina el nivel de expresión informativa de un marcador o un conjunto de marcadores seleccionado o derivado de la Tabla 1 y/o descrito en este documento.

En una descripción de la divulgación, se evalúa una prima para el seguro (por ejemplo, de vida o médicos) en función de la información sobre uno o más niveles de expresión de marcadores, por ejemplo, un marcador o conjunto de marcadores, por ejemplo, un nivel de expresión asociado con el resultado del tratamiento (por ejemplo, la cantidad informativa). Por ejemplo, las primas se pueden aumentar (por ejemplo, en un cierto porcentaje) si los genes marcadores de un paciente o el conjunto de marcadores de un paciente descrito en este documento tienen características diferentes, por ejemplo, tamaño, secuencia, composición o cantidad entre un candidato asegurado (o un candidato que busca cobertura de seguro) y un valor de referencia (por ejemplo, una persona no afectada) o una muestra de referencia, por ejemplo, control combinado. Las primas también se pueden dimensionar dependiendo del resultado de la evaluación de un marcador o conjunto de marcadores descrito en este documento. Por ejemplo, las primas se pueden evaluar para distribuir el riesgo, por ejemplo, en función del marcador, por ejemplo, el resultado de evaluar un marcador o conjunto de marcadores descrito en este documento. En otro ejemplo, las primas se evalúan en función de los datos actuariales que se obtienen de pacientes que tienen resultados de tratamiento conocidos.

La información sobre la característica del marcador, por ejemplo, tamaño, secuencia, composición o cantidad, por ejemplo, el resultado de evaluar un marcador o conjunto de marcadores descrito en este documento (por ejemplo, la cantidad informativa), puede usarse, por ejemplo, en un proceso de suscripción de seguros de vida. La información se puede incorporar a un perfil sobre un sujeto. Otra información en el perfil puede incluir, por ejemplo, fecha de nacimiento, el género, estado civil, información bancaria, información de crédito, hijos, etc. Una póliza de seguro se puede recomendar en función de la información sobre la característica del marcador, por ejemplo, tamaño, secuencia, composición o cantidad, por ejemplo, el resultado de evaluar un marcador o conjunto de marcadores descrito en este documento, junto con uno o más elementos de información en el perfil. Una prima de seguro o evaluación de riesgo también se puede evaluar en función de la información del marcador o conjunto de marcadores. En una implementación, los puntos se asignan en función del resultado esperado del tratamiento.

En una descripción de la divulgación, la información sobre la característica del marcador, por ejemplo, tamaño, secuencia, composición o cantidad, por ejemplo, el resultado de evaluar un marcador o conjunto de marcadores descrito en este documento, se analiza mediante una función que determina si autorizar la transferencia de fondos para pagar un servicio o tratamiento brindado a un sujeto (o tomar otra decisión a la que se hace referencia en este documento). Por ejemplo, los resultados de analizar una característica, por ejemplo, tamaño, secuencia, la composición o cantidad de un marcador o conjunto de marcadores descrito en este documento puede indicar que se espera que un sujeto tenga un resultado favorable, sugiriendo que se necesita un ciclo de tratamiento, desencadenando así un resultado que indica o causa autorización para pagar un servicio o tratamiento brindado a un sujeto. En un ejemplo, la característica informativa, por ejemplo, tamaño, secuencia, composición o cantidad de un marcador o un conjunto de marcadores seleccionados o derivados de la Tabla 1 y/o descritos en este documento se determinan y el pago se autoriza si la cantidad informativa identifica un resultado favorable. Por ejemplo, una entidad, por ejemplo, un hospital, cuidador, entidad gubernamental, o una compañía de seguros u otra entidad que paga o reembolsa gastos médicos, puede usar el resultado de un método descrito en este documento para determinar si una parte, por ejemplo, una parte que no sea el paciente objeto, pagará los servicios (por ejemplo, una terapia en particular) o el tratamiento brindado al paciente. Por ejemplo, una primera entidad, por ejemplo, una compañía de seguros, puede usar el resultado de un método descrito en este documento para determinar si se debe proporcionar un pago financiero a, o en nombre de, un paciente, por ejemplo, si reembolsar a un tercero, por ejemplo, un vendedor de bienes o servicios, un hospital, médico u otro cuidador, por un servicio o tratamiento proporcionado a un paciente. Por ejemplo, una primera entidad, por ejemplo, una compañía de seguros, puede usar el resultado de un método descrito en este documento para determinar si continúa, interrumpe, inscribe a un individuo en un plan o programa de seguro, por ejemplo, un seguro de salud o plan o programa de seguro de vida.

En una descripción, la divulgación presenta un método para proporcionar datos. El método incluye proporcionar

datos descritos en este documento, por ejemplo, generados por un método descrito en este documento, para proporcionar un registro, por ejemplo, un registro descrito en este documento, para determinar si se proporcionará un pago. En algunas descripciones, los datos son proporcionados por ordenador, disco compacto, teléfono, facsímil, correo electrónico o carta. En algunas descripciones, los datos son proporcionados por una primera parte a una segunda parte. En algunas descripciones, la primera parte se selecciona entre el sujeto, un proveedor de atención médica, un médico tratante, una organización de mantenimiento de la salud (HMO), un hospital, una entidad gubernamental, o una entidad que vende o suministra el medicamento. En algunas descripciones, la segunda parte es un tercer pagador, una compañía de seguros, empleador, plan de salud patrocinado por el empleador, HMO, o entidad gubernamental. En algunas descripciones, la primera parte se selecciona entre el sujeto, un proveedor de atención médica, un médico tratante, una HMO, un hospital, una compañía de seguros, o una entidad que vende o suministra el medicamento y la segunda parte es una entidad gubernamental. En algunas descripciones, la primera parte se selecciona entre el sujeto, un proveedor de atención médica, un médico tratante, una HMO, un hospital, una compañía de seguros, o una entidad que vende o suministra el medicamento y la segunda parte es una compañía de seguros.

En otra descripción, la divulgación presenta un registro (por ejemplo, registro legible por ordenador) que incluye un listado y el valor de la característica, por ejemplo, tamaño, secuencia, composición o cantidad para el marcador o conjunto de marcadores para un paciente. En algunas descripciones, el registro incluye más de un valor para cada marcador.

A continuación, se ilustrará la presente invención mediante los ejemplos siguientes, que no se pretende que sean limitantes de ningún modo.

Ejemplos

Ejemplo 1. Identificaciones de panel de línea celular

Para apoyar el desarrollo clínico e identificar biomarcadores potenciales de sensibilidad o resistencia tumoral, dos paneles grandes de líneas celulares de cáncer (Panel 1, N = 653 (McDermott *et al.* (2007) PNAS 104:19936-19941); Panel 2, N = 240 (O'Day *et al.* (2010) Fourth AACR International Conference on Molecular Diagnostics in Cancer Therapeutic Development)) fueron tratados con MLN4924 y se generaron datos de viabilidad celular (CI50, CE50 y POC - Porcentaje de control).

Panel 1 (McDermott *et al.*, *supra).* Las líneas celulares se expusieron a tres concentraciones de MLN4924 (20 nM, 200 nM y 2 µM) durante 72 horas. La viabilidad, es decir, índice celular, se cuantificó midiendo la fluorescencia de una tinción de ácido nucleico permeante celular. Los valores medios se tomaron por triplicado para cada muestra y se compararon con el control DMSO para calcular el porcentaje de control. En los resultados, al control o valor de ningún actividad se le da un valor de aproximadamente 1, la sensibilidad se indica con un valor inferior a 1, con 0 como muerte de toda la población celular y resistencia indicada por un valor superior a 1. Para cada concentración se obtuvo un continuo de valores de viabilidad, entonces algunos valores fueron seleccionados como puntos de corte para la determinación final de sensibilidad o resistencia. Por ejemplo, los valores medios de POC menores que el valor medio de todos los POC registrados en el panel (<0,34) indicaron sensibilidad, valores de 0,34 a 0,75 indicaron sensibilidad límite, valores mayores que el 3er cuartil de todos los POC en el panel (>0,75) insensibilidad o resistencia. En general, una relación dosis respuesta se observó con líneas celulares sensibles. El criterio final de la línea celular como sensible, insensible o resistente se determinó por su viabilidad a 2 µM.

Panel 2 (Ricerca Biosciences, Inc., Bothell, WA). Se añadió MLN4924 en diluciones semi-log para 10 concentraciones y se trató durante 72 horas. La identificación celular de alto contenido mediante microscopía de fluorescencia incluyó análisis de imágenes para generar varios tipos de datos. Los resultados incluyeron valores de CE50 (después de medición del número de células, la concentración de CE50 se calculó a partir del punto de inflexión de una curva de porcentaje de control (POC) frente al log de la concentración MLN4924), valores CI50 (del gráfico POC-log MLN4924, CI50 es la concentración al 50 % de respuesta máxima posible), apoptosis (medición de la activación de la caspasa 3 trazada contra el log de la concentración de MLN4924, determinado como la concentración para inducción >5 veces) y actividad mitótica (determinada midiendo el aumento de fosfohistona 3). Una comparación de CE50 con CI50 (Figura 3) permitió la asignación de líneas celulares a sensibles, insensibles o resistentes. La identificación final de una línea celular como sensible o resistente se basó en los valores de CE50. El límite de sensibilidad es una CE50 media menor que el valor medio de todos los POC registrados en el panel (<0,36), la sensibilidad límite se asoció con una CE50 media de 0,36 a 1,67 y las líneas celulares insensibles o resistentes fueron identificadas por CE50 mayor que el 3er cuartil de todos los POC en el panel (>1,67).

Las líneas celulares superpuestas entre los paneles 1 y 2 (114 superposiciones) mostraron efectos coherentes de inhibición del crecimiento (coeficiente de correlación de Spearman = 0,72). Además, el análisis de histología y mutación en cada panel de líneas celulares en su conjunto (no solo líneas celulares superpuestas), también generó observaciones coherentes entre los dos paneles (Figuras 4A y B).

El ensayo exacto de Fisher utilizando el porcentaje medio de los valores de control se utilizó para evaluar las asociaciones de mutaciones individuales en las líneas celulares a la sensibilidad o resistencia MLN4924. Véase la Tabla 3 para un resumen de los valores de p para genes seleccionados. Los genes cuyas mutaciones estaban relacionadas con la sensibilidad en el ensayo exacto de Fisher incluyen NF2, SMAD4, KDM6A, CDKN2A y CDKN2A_p14. RB1 y TP53 estaban relacionados con la insensibilidad.

Tabla 3. Confianza de la asociación de marcador mutado con respuesta a MLN4924

Gen mutado	Panel 1 N Total	Panel 1 sensible al número	Valor p del Panel 1	Panel 2 N Total	Panel 2 sensible al número	Valor p del Panel 2	Fenotipo
NF2	10	8	0,067	5	4	0,187	Sensible
SMAD4	24	15	0,197	11	6	0,509	Sensible
KDM6A	6	6	0,019	4	4	0,062	Sensible
RB1	32	12	0,062	14	17	1,00	Resistente*
TP53	235	112	0,013	98	45	0,107	Resistente
CDKN2A	153	94	0,001	68	38	0,146	Sensible
CDKN2A_ p14	114	70	0,01	59	33	0,178	Sensible
* indica fenotipo para asociación RB1 no concluyente en todos los tipos celulares. Algunos tipos de tumores se asociaron más con la resistencia que otros: tumores del cerebro, vejiga, hueso y pulmón (NSCLC) tienen valores de p de 0,102, 0,205, 0,226 y 0,281, respectivamente. El resultado de que RB1 no es un marcador de sensibilidad está de acuerdo con el resultado de Jia *et al.* ((2011) Neoplasia 13:561-569) que excluyó la participación de RB1 en el mecanismo MLN4924.							

Ejemplo 2. Análisis de asociaciones de mutación

Una dificultad para correlacionar mutaciones de genes en líneas celulares con sensibilidad a un agente terapéutico es que muchas líneas celulares tienen más de un gen mutado. Por ejemplo, la línea celular llamada 8505C del carcinoma de tiroides tiene mutaciones en BRAF, TP53, NF2 y CDKN2A. En particular, las mutaciones TP53 y CDKN2A se produjeron conjuntamente con otras mutaciones en las líneas celulares. Para saber qué mutante está asociado con sensibilidad o resistencia en los paneles de la línea celular, se realizaron subanálisis.

APC frente a TP53. En el panel 1 de línea celular, 23 líneas celulares tienen una mutación en APC. De estas, 18 líneas celulares también tienen una mutación en TP53. Fue difícil determinar si APC es un impulsor de resistencia a MLN4924, o simplemente una mutación pasajera que se encuentra con frecuencia en mutantes TP53. Un análisis adicional de las líneas celulares se realizó mediante sustracción de líneas celulares con mutantes dobles que incluían TP53 (Tabla 4).

Tabla 4. Comparación de líneas celulares mutantes TP53 con APC y otras líneas celulares mutantes en el panel 1

A.	Restar mutantes TP53		
Gen	sens(78)	res(51)	valor p
RB1	0	2	0,154433
NRAS	7	7	0,285024
APC	2	3	0,30657
SMAD4	2	3	0,30657
BRCA2	0	1	0,395349
FAM123B	0	1	0,395349
MAP2K4	0	1	0,395349
B.	Restar mutantes APC		
Gen	sens(184)	res(158)	valor p
TP53	108	110	0,0234828
RB1	12	19	0,0573067
MAP2K4	2	5	0,1665217
C.	Mutantes dobles		
Gen	sens(190)	res(175)	valor p
APC+TP53	4	14	0,00842

Como se puede ver en la Tabla 4A, la resta de los mutantes TP53 del panel de la línea celular deja un N demasiado pequeño para permitir la conclusión de la asociación de las mutaciones restantes con resistencia de las líneas celulares TP53 wt al tratamiento con MLN4924. Después de eliminar los 23 mutantes APC, are que TP53 está asociado con resistencia (Tabla 4B). No obstante, las líneas celulares con mutaciones APC y TP53 muestran una fuerte asociación con resistencia (Tabla 4C). Adicionalmente, la mayoría de las líneas celulares en el subgrupo mutante APC+TP53 provienen de muestras de tumores de cáncer intestinal (Tabla 5, que también incluye líneas

celulares y datos del panel 2 y seis líneas celulares del panel 1 que no se incluyeron en el análisis sustractivo original).

Tabla 5. Subconjunto de líneas celulares con mutaciones en APC y TP53

Estirpe celular	Fuente de tejido tumoral	Histología tumoral	Viabilidad a 2 µM Panel 1	CE50 Panel 2	Mutaciones simultáneas
HT55	intestino	carcinoma de colon	0,5143		APC:APC:TP53
SW 1116	intestino	adenocarcinoma de colon	1,1599	10	APC:APC:KRAS:SMAD4: TP53
COCM1	intestino	carcinoma de colon	1,1438		APC:APC:PIK3CA:SMAD4: TP53
LS1034	intestino	adenocarcinoma		3,76	APC:KRAS:TP53
SW 1463	intestino	adenocarcinoma rectal	0,8963		TP53:FBXW7:KRAS:APC
NCI-H630	intestino	adenocarcinoma colorrectal	0,75		APC:TP53
SW1417	intestino	adenocarcinoma de colon	0,6374	9,03	APC:TP53:BRAF:PIK3R1
SW837	intestino	adenocarcinoma rectal	0,702	0,694	APC: APC: FAM123B:TP53: FBXW7:KRAS
SW620	intestino	adenocarcinoma de colon	0,4928	0,387	APC:TP53:TP53:KRAS: MAP2K4:SM AD4
NCI-H1581	NSCLC pulmonar	carcinoma de células escamosas	0,7365		APC:TP53
HCT-15	intestino	adenocarcinoma de colon	0,6654	0,836	APC:BRCA2:FAM123B: KRAS:MSH6: PIK3CA: TP53
C2BBe1	intestino	adenocarcinoma colorrectal	0,7159		APC:TP53
T84	intestino	carcinoma de colon	0,5777	0,556	KRAS: PIK3CA:TP53:APC
SW626	ovario	adenocarcinoma	0,5541		TP53:KRAS:APC
HT29	intestino	adenocarcinoma colorrectal	0,4376	0,25	APC: PIK3CA:TP53: BRAF: SMAD4:AP C
LK-2	pulmón	NSCLC pulmonar	0,2893		TP53:APC:CDKN2A
MKN28	estómago	metástasis	0,2847		APC:TP53:NF1
COLO-205	intestino	adenocarcinoma de colon	0,144	0,0923	TP53:BRAF:APC:SMAD4
HGC-27	estómago	carcinoma gástrico	0,1291		TP53:APC:PTEN:PIK3CA
NCIH747	intestino	adenocarcinoma de ciego		2,07	APC:APC:KRAS:TP53
SKUT1	músculo piso	leiomiosarcoma		0,0856	APC:APC:PIK3CA:PTEN: PTEN:RB1:T P53:TP53
SK-MEL-30	piel	melanoma maligno	0,7728		APC:NRAS:TP53
COLO320 HSR	intestino	adenocarcinoma de colon		1,78	APC:TP53
NCI-H520	NSCLC pulmonar	carcinoma de células escamosas	0,5577	0,261	CDKN2A:APC:TP53
NCI-H1703	NSCLC pulmonar	carcinoma adenoescamoso	0,1101		TP53:APC:CDKN2A
NCI-H1975	pulmón	adenocarcinoma	0,3049		TP53:APC:CDKN2A:PIK3CA: EGFR
RCM-1	intestino	adenocarcinoma rectal	0,7666		TP53:KRAS:APC
SK-OV-3	ovario	adenocarcinoma	0,397	10	TP53:PIK3CA:CDKN2A:APC

Asociación con histología. Es posible que algunas mutaciones de TP53 estén asociadas con resistencia en algunos tipos de tumor más que en otros. Las líneas celulares mutantes de TP53 se analizaron dentro de los tipos de tejido mediante un ensayo de Mann-Whitney (no paramétrico). Como se puede ver en la Figura 5, las mutaciones de TP53 se correlacionan significativamente con la resistencia a MLN4924 en líneas celulares de cáncer de colon (valor p =

0,04022).

Un análisis similar de asociaciones en otros tejidos indica que la mutación TP53 está asociada con la resistencia MLN4924 en líneas celulares de cáncer de mama y de pulmón (NSCLC).

Se analizaron mutaciones adicionales para la posible asociación de la histología tumoral con resistencia o sensibilidad al tratamiento MLN4924. La Tabla 6 proporciona resultados del análisis del panel 1 y la Tabla 7 proporciona resultados del análisis del panel 2, cuyo tamaño más pequeño resultó un desafío para este tipo de análisis. El límite de asociación con los tejidos se eligió con valores de $p < 0,05$.

Tabla 6. Asociación de mutación con histología en el Panel 1

Fuente de tejido tumoral	Mutación	N mutantes este cáncer	N tipo silvestre este cáncer	N mutantes todos los cánceres	N tipo silvestre todos los cánceres	Asociación	valor p
Cabeza_Cuello	TP53	15	1	111	79	sensible	0,0001
Cabeza_Cuello	CDKN2A	12	4	94	96	sensible	0,0072
Cabeza_Cuello	SMAD4	3	13	15	175	sensible	0,0262
Cuello uterino	TP53	1	6	124	51	resistente	0,0027
Pulmón	RB1	2	4	19	156	resistente	0,0178
Hueso	RB1	1	5	12	178	sensible	0,0477

Tabla 7. Asociación de mutación con histología en el Panel 2

Fuente de tejido tumoral	Mutación	N mutantes este cáncer	N tipo silvestre este cáncer	N mutantes todos los cánceres	N tipo silvestre todos los cánceres	Asociación	valor p
piel	CDKN2A	8	1	30	43	resistente	0,0022
piel	CDKN2A.p14	6	3	26	47	resistente	0,0004
piel	TP53	2	7	53	20	resistente	0,0036
CNS	CDKN2A.p14	5	4	26	47	resistente	0,0466*
* valor p para CNS derivado del porcentaje de viabilidad de control. Para piel, valor p derivado de CE50.							

A diferencia de la asociación general de TP53 con resistencia a MLN4924, en cáncer de cabeza y cuello, La mutación TP53 está asociada con sensibilidad. Se encontró un contraste similar para mutaciones de CDKN2A en piel y sistema nervioso central (CNS). Las mutaciones CDKN2A o CDKN2A.p14 se asociaron con resistencia al tratamiento MLN4924 en la piel y las líneas celulares tumorales del CNS, a pesar de la asociación general con la mutación CDKN2A y la sensibilidad al tratamiento MLN4924. Las tablas 6 y 7 también muestran que las mutaciones de TP53 están significativamente asociadas con la resistencia en el cáncer cervical y cáncer de piel; Las mutaciones de CDKN2A están significativamente asociadas con sensibilidad en cáncer de cabeza y cuello; Las mutaciones de SMAD4 están significativamente asociadas con sensibilidad en cáncer de cabeza y cuello; las mutaciones de RB1 se asocian significativamente con resistencia en cáncer de pulmón; y las mutaciones RB1 están significativamente asociadas con sensibilidad en cáncer de hueso.

Ejemplo 3. Resultados de detección de líneas celulares individuales

Las siguientes tablas incluyen resultados de las identificaciones de líneas celulares individuales que llevaron a conclusiones sobre marcadores cuyas mutaciones confieren sensibilidad a MLN4924.
La anotación de las mutaciones y la explicación de la sintaxis de mutaciones se pueden encontrar en la base de datos COSMIC.

Tabla 8. Resultados de identificaciones de líneas celulares con mutaciones en NF2.

Estirpe celular	Mutación ORF (SEQ ID NO: 2)	Mutación de proteína (SEQ ID NO: 3)	Fuente de tejido tumoral	Tipo de tumor	Viabilidad a 2 µM Panel 1	CE50 Panel 2	Fenotipo	Mutaciones simultáneas
647-V	c.115-1G>C	¿p.?	tracto urinario	primario	0,038	0,283	sensible	MAP2K4:NF2:RB1:TP53:TP53
ACHN	c.169C>T	p.R57*	riñón	primario	0,0621	0,652	sensible	CDKN2A:CDKN2a(p14):N F2
CAL-62	c.643G>T	p.E215*	tiroides	primario	0,0676	0,0837	sensible	CDKN2A:CDKN2a(pl4):KRAS: NF2:TP53
NUGC-3	c.683de1A	p.K228fs*23	estómago	primario	0,098		sensible	NF2:TP53
SW1573	c.1_363del363	¿p.?	pulmón	primario	0,1114		sensible	CDKN2A:CDKN2a(p14): CTNNB1: KRAS:N F2: PIK 3CA: SMAD4
8505C	c.385G>T	p.E129*	tiroides	primario	0,1405		sensible	BRAF:CDKN2A:NF2:TP53
CaR-1	c. 115_1737del1623	¿p.?	grande_ intestino	primario	0,3244		sensible	CDKN2A:CDKN2a(p14):NF2: STK11:TP53
SN12C	c.115-1G>C	¿p.?	riñón	primario	0,3357		sensible	NF2:TP53
MDA-MB-231	c.691G>T	p.E231*	mama	primario	0,5284	0,871	insensible	BRAF:CDKN2A:CDKN2a(p14): KRAS:NF2:TP53
S-117	c.221G>A	p.W74*	Tejido_blando	primario	0,6412		resistente	CDKN2A:CDKN2a(p14):NF2: TP53
RL95-2	c.514delA c. 1084C>T	p.R172fs*2 p.Q362*	endometrio	primario		0,094	sensible	BRCA2:BRCA2:HRAS:NF2: NF2:PTEN:PTEN:TP5 3:TP53

Tabla 9. Resultados de identificaciones de líneas celulares con mutaciones en SMAD4.

Estirpe celular	Mutación ORF (SEQ ID NO: 5)	Mutación de proteína (SEQ ID NO: 6)	Fuente de tejido tumoral	Tipo de tumor	Viabilidad a 2 µM Panel 1	CE50 Panel 2	Fenotipo e	Mutaciones simultáneas
CAL-27	c.733C>T	p.Q245*	Superior_ Tracto_aerodigestivo	primario	0,0271	0,139	sensible	CDKN2A:SMAD4:TP53
KP-4	c.1_1659del1659	¿p.0?	páncreas	metástasis	0,0518		sensible	CDKN2A:CDKN2a(p14):KRAS: SMAD4
GAMG	c.1_249del249	¿p.?	sistema_nervioso_central	primario	0,0721		sensible	CDKN2A:CDKN2a(p14): SMAD4:TP53
PANC-03-27	c.905_1659del755	¿p.?	páncreas	primario	0,0724		sensible	CDKN2A:CDKN2a(p14):KRAS: SMAD4:TP53
FADU	c.1_1659del1659	¿p.0?	Superior_ Tracto_aerodigestivo	primario	0,073	0,261	sensible	CDKN2A:SMAD4:TP53:TP53
SW1573	c.1_1659del1659	¿p.0?	pulmón	primario	0,1114		sensible	CDKN2A:CDKN2a(p14): CTNNB1:KRAS:NF2:PIK3 CA: SMAD4
KYSE-150	c.788-1 G>A	¿p.?	esófago	primario	0,1279		sensible	SMAD4:TP53
COLO-205	c.1_667del667	¿p.?	grande_ intestino	primario	0,144	0,0923	sensible	APC:BRAF:SMAD4:TP53
CAL-33	c.766C>T	p.Q256*	Superior_ Tracto_aerodigestivo	primario	0,1445		sensible	CDKN2A:PIK3CA:SMAD4: TP53
NCI-N87	c.1_955del955	¿p.?	estómago	metástasis	0,1461		sensible	SMAD4:TP53
CAPAN-1	c.1028C>G	p.S343*	páncreas	metástasis	0,1473	0,365	¿sensible?	BRCA2:CDKN2A:CDKN2a (p14):KRAS:SMAD4:TP 53
YAPC	c.1543de1A	p.R515fs*22	páncreas	primario	0,1871	0,234	Sensible	CDKN2A:CDKN2a(p14):KRAS: SMAD4:TP53
NCI-H2405	c.1_1659del1659	¿p.0?	pulmón	metástasis	0,2554		Sensible	BRAF:CDKN2A:CDKN2a(p14): MAP2K4:SMAD4:T P53
CFPAC-1	c.1_1659del1659	¿p.0?	páncreas	metástasis	0,2822	6,37	¿sensible?	KRAS:SMAD4:TP53
MDA-MB-468	c.1_1659del16 59	¿p.0?	mama	metástasis	0,2965	0,0267	sensible	PTEN:RB1:SMAD4:TP53
MKN45	c.1_1659del16 59	¿p.0?	estómago	metástasis	0,3614		insensible	CDKN2A:CDKN2a(p14): SMAD4
BxPC-3	c.1_1659del16 59	¿p.0?	páncreas	primario	0,4186	0,251	insensible	CDKN2A:CDKN2a(p14):MAP2K4:SMAD4:TP53
HT-29	c.931C>T	p.Q311*	grande_ intestino	primario	0,4376	0,25	insensible	APC:APC:BRAF:PIK3CA: SMAD4:TP53
UMC-11	c.1606_1612de ICTAGACG	p.L536fs*14	pulmón	primario	0,4472		insensible	CDKN2A:CDKN2a(p14): SMAD4:STK11:TP53
SW620	c.955+5G>C	¿p.?	grande_ intestino	primario	0,4928	0,387	insensible	APC: KRAS:MAP2K4:SMAD4: TP53:TP53

(continuación)

Estirpe celular	Mutación ORF (SEQ ID NO: 5)	Mutación de proteína (SEQ ID NO: 6)	Fuente de tejido tumoral	Tipo de tumor	Viabilidad a 2 µM Panel 1	CE50 Panel 2	Fenotipo e	Mutaciones simultáneas
PANC-08-13	c.366_367insA	p.C123fs*2	páncreas	primario	0,7767		resistente	CDKN2A:CDKN2a(p14):KRAS: SMAD4
COLO-678	c.1_1659del1659	p.O?	grande_ intestino	metástasis	0,814		resistente	APC:CDKN2A:CDKN2a(p14): FAM123B:KRAS:SM AD4
CoCM-1	c.956_1659del704	¿p.?	grande_ intestino	primario	1,1438		resistente	APC:APC:PIK3CA:SMAD4: TP53
SW954	c.378_379delC T	p.V128fs*14	vulva	primario		0,324	sensible	SMAD4:TP53

Tabla 10. Resultados de identificaciones de líneas celulares con mutaciones en KDM6A.

Estirpe celular	Mutación ORF (SEQ ID NO: 8 o 9)	Mutación de proteínas (SEQ ID NO: 10 u 11)	Fuente de tejido tumoral	Tipo de tumor	Viabilidad a 2 µM Panel 1	CE50 Panel 2	Fenotipo	Mutaciones simultáneas
HCC1806	c.444_564del121	p.0	mama	primario	0,0116		sensible	CDKN2A:CDKN2a(p14): KDM6A:STK11:TP53
KU-19-19	c.2587C>T	p.Q863*	tracto urinario	primario	0,058		sensible	CDKN2A:CDKN2a(p14): KDM6A:NRAS
MIA-PaCa-2	c.1_4206del42 06	p.O?	páncreas	primario	0,1145	0,239	sensible	CDKN2A:CDKN2a(p14): KDM6A:KRAS:TP53
KYSE-450	c.385_654del270	p.O	esófago	primario	0,1426		sensible	CDKN2A:CDKN2a(p14): EGFR:KDM6A:NOTCH1: TP53:TP53
KYSE-180	c.997C>T	p.Q333*	esófago	primario	0,2154		sensible	CDKN2A:CDKN2a(p14): KDM6A:TP53
LS-174T	c.3945_3946ins A	p.E1316fs*17	intestino_grueso	primario	0,3262	0,137	sensible	CTNNB1:KDM6A:KRAS: PIK3CA
BV-173	c.226_384del159	p.O	hematopoyético, tejido_linfoide	primario		0,131	sensible	CDKN2A:CDKN2a(p14): KDM6A
THP-1	c.1_1923del1923	p.O	hematopoyético, tejido_linfoide	primario		0,148	sensible	CDKN2A:CDKN2a(p14): KDM6A:NRAS:TP53

Tabla 11. Muestreo de resultados de identificaciones de líneas celulares con mutaciones en FBXW7.

Estirpe celular	Mutación ORF (SEQ ID NO: 13)	Mutación de proteína (SEQ ID NO: 14)	Fuente de tejido tumoral	Tipo de tumor	Viabilidad a 2 µM Panel 1	CE50 Panel 2	Fenotipo	Mutaciones simultáneas
AN3CA	c. 1321C>T	p.R441W	Útero	metástasis	0,1944		sensible	PTEN: TP53: PIK3R1
ESS-1	c. 1393C>T	p.R465C	Útero	primario	0,0529		sensible	FBXW7: TP53: PIK3CA: RB1
C-33a	c. 1394G>A	p.R465H	Cuello uterino	primario	0,0822		sensible	RB1: TP53: MSH2: PIK3CA: PTEN
HuCCT1	c.881C>G	p.S294*	Hígado	primario	0,1262		sensible	TP53: KRAS
MKN1	c. 1393C>T	p.R465C	Estómago	metástasis	0,4631		insensible	TP53: PIK3CA
AsPC-1	c. 1393C>T	p.R465C	Páncreas	metástasis	0,9889		resistente	CDKN2A: MAP2K4: KRAS: TP53
RCM-1	c. 1513C>T	p.R505C	Intestino	primario	0,7666		resistente	TP53: KRAS
SW1463	c. 1436G>A	p.R479Q	Intestino	primario	0,8963		resistente	TP53: KRAS: APC

Ejemplo 4. Asociación de deleción de TP53 con resistencia

Otro enfoque para determinar el papel de TP53 en la respuesta a MLN4924 fue un estudio en donde se eliminó el gen TP53. En estudios anteriores, la importancia de p53 en la respuesta de re-replicación a MLN4924 parecía depender de la manipulación genética específica y se esperaba que reflejara estrechamente la de sobreexpresión de CDT1 (Cdtl es un sustrato de dos complejos CRL alternativos y MLN4924 lo estabiliza en muchas líneas celulares). En estudios de supresión, p53 parecía comportarse de manera similar a la supresión de CDT1 en los puntos temporales tempranos, pero no en puntos temporales posteriores, a menos que se usaran concentraciones más altas de MLN4924. La transferencia Western sugirió la eliminación eficaz de la proteína p53 por ARNsi SMART-combinación, aunque generalmente el ARNi no da como resultado la pérdida completa de proteína. Por tanto, la proteína residual aún puede afectar la respuesta a MLN4924, particularmente porque MLN4924 da como resultado la estabilización de p53 (Liao *et al.* (2011) Mol. Cell Proteomics 10: 10. 1074/mcp.Ml 11.009183). El efecto de viabilidad de MLN4924 se evaluó en células HCT-116 que se eliminaron genéticamente para p53 junto con su control parental.

Se sembraron líneas celulares HCT-116 isogénicas emparejadas que eran de tipo silvestre (+/+) o nulas (-/-) para la expresión de p53 (HD PAR-018 y HD 104-001, respectivamente, Horizon Discovery Ltd) en placas separadas de 384 pocillos y luego se trataron al día siguiente con una titulación de MLN4924 por triplicado, y se incubaron durante 24, 36, 48 o 72 h, con densidades de siembra de 1600, 1200, 800 y 400 células/pocillo, respectivamente. Después de incubar el compuesto, la viabilidad de las células HCT-116 se evaluó mediante el ensayo ATPlite (Perkin Elmer) según las instrucciones del fabricante utilizando el sistema de imágenes LEADseeker (GE Healthcare).

Las células HCT-116 *TP53* +/+ (MLN4924 CL_{50} = 21 ± 1 nM) demostraron una mayor sensibilidad MLN4924 a las 72 h que las células HCT-116 *TP53* -/- (MLN4924 CL_{50} = 74 ± 5 nM; Figuras 6A-D). Estos resultados sugieren que la deficiencia de p53 hace que HCT-116 sea menos sensible a MLN4924, sugiriendo que el papel general de p53 a las 72 h es proapoptótico. Los puntos temporales anteriores refuerzan esta interpretación, ya que las células *TP53* -/- tienen menos muerte celular a las concentraciones más altas de fármaco a las 24, 36 y 48 h. Las transferencias Western mostraron que p21 todavía estaba regulado por MLN4924 en células *TP53* -/- HCT-116. En células HCT-116, la estabilización de p21 puede ser un efecto directo de la inhibición de CRL4-Cdt2 (Nishitani *et al*., 2008; Abbas *et al*., 2008; Kim *et al*., 2008).

Este resultado es contrario a la conclusión de Lin *et al*. (2010) Cancer Res. 70:10310-10320) usando las células inactivadas HCT116 -/- p53. En ese estudio, se concluyó que las células inactivadas TP53 eran más susceptibles a la muerte celular general o la inhibición del crecimiento por MLN4924. La razón por la que el presente estudio llega a una conclusión diferente a la de Lin *et al.* es el periodo de tiempo que las células fueron tratadas con MLN4924. En Lin *et al*., las células fueron tratadas durante 8 horas antes de un lavado. En los estudios actuales y en los paneles de líneas celulares de los ejemplos anteriores, las células fueron tratadas con MLN4924 continuamente durante 72 horas. En el lavado, se permite que los niveles de p53 se estabilicen y aprovechen la activación de vías alternativas que las vías anteriores que inicialmente fueron inhibidas y condujeron a una susceptibilidad más temprana.

Ejemplo 5. Aislamiento de ácido nucleico y métodos de secuenciación de ácido nucleico

Aislamientos genómicos y secuenciación de ADN. El aislamiento de ADN de células y tumores se realiza utilizando el kit de aislamiento DNAEASY® (Qiagen, Valencia, CA). El aislamiento de ARN se realiza utilizando MegaMax (división Ambion de Applied Biosystems, Austin, TX). Los aislamientos genómicos se realizan siguiendo los protocolos recomendados por el fabricante.

Metodología de secuenciación SANGER. Las amplificaciones por PCR se realizan utilizando condiciones de ciclado optimizadas por gen-exón. La secuenciación de la extensión del cebador se realiza utilizando BigDye de Applied Biosystems versión 3.1. Las reacciones se ejecutan En el analizador de ADN 3730x1 de Applied Biosystem. La secuencia de llamadas base se realiza utilizando KBTM Basecaller (Applied Biosystems). Las llamadas de mutación somática son determinadas por Mutation Surveyor (SoftGenetics) y se confirman manualmente alineando los datos de secuenciación con la secuencia de referencia correspondiente utilizando Seqman (DNASTAR).

Metodología de secuenciación SEQUENOM. Los ensayos de Sequenom (San Diego, CA) se diseñan utilizando la química TypePLEX® con extensión de base única. Este proceso consta de tres etapas: 1) Un archivo de texto que contiene los SNP o mutaciones de interés y la secuencia de flanqueo se carga en mysequenom.com, donde se ejecuta con un programa basado en la web ProxSNP, 2) La salida de ProxSNP se ejecuta con PreXTEND y 3) la salida de PreXTEND se ejecuta con el diseño de ensayo que determina el peso de masa esperado de los productos extendidos para garantizar la separación entre todos los picos potenciales encontrados dentro de una reacción multiplexada.

Los cebadores de PCR se diseñan para agrupar la región identificada en las etapas de diseño del ensayo. La región de interés se amplía en reacciones de PCR utilizando los cebadores. Se aplican 15 nl de producto amplificado y extendido en un 384 SpectroCHIP II usando un Nanodispenser RS1000. Se añade un calibrador de 3 puntos a cada chip para garantizar el rendimiento adecuado del espectrómetro de masas compacto Sequenom Maldi-tof.

El SpectroCHIP II se coloca en el espectrómetro de masas compacto Sequenom MALDI-TOF. El espectrómetro de masas se configura para disparar un máximo de 9 adquisiciones por cada punto en el spectroCHIP de 384 pocillos. El kit TypePLEX Gold SpectroCHIP II de Sequenom (10142-2) se utiliza siguiendo los protocolos recomendados por el fabricante. El análisis se realiza utilizando el software de análisis Sequenom, MassARRAY® Typer Analyzer v4.

Metodología de SECUENCIA DE PRÓXIMA GENERACIÓN (NGS). NGS dirigida utilizando la plataforma Illumina (Illumina, Inc. San Diego, CA) se utiliza para confirmar e identificar mutaciones de baja frecuencia en un marcador. Los pares de cebadores se diseñan para amplificar los exones de codificación. Los productos de PCR se cuantifican usando un ensayo PicoGreen y se combinan en proporciones molares iguales para cada muestra. Los productos purificados se reparan y concatenan por ligamiento. Los productos concatenados se utilizan para preparar la biblioteca Hi-Seq 2000. Los productos de PCR concatenados se cortan y se usan para hacer bibliotecas Hi-Seq 2000 con código de barras que consisten en 12 muestras con código de barras por grupo multiplexado. Las bibliotecas agrupadas Hi-Seq 2000 experimentan amplificación clonal por generación de grupo en ocho carriles de una celda de flujo Hi-Seq 2000 y se secuencian usando secuenciación de un solo extremo 1x100 en un Hi-Seq 2000. La coincidencia de las lecturas de secuenciación primaria con la construcción del genoma humano Hg18, así como el análisis SNP se realizan utilizando el software CASAVA de Illumina versión 1.7.1.

Procedimientos generales

RT-PCR cuantitativa

La síntesis de ADNc y la RT-PCR cuantitativa se realizan usando ensayos de expresión génica ABI, reactivos y sistemas de detección de secuencia ABI PRISM® 7900HT (Applied Biosystems, Foster City, CA) usando las siguientes condiciones de ciclo: mantener a 50 ℃ durante 2 minutos para activar AmpErase UNG, luego 95,0 ℃ durante 10 minutos para activar la ADN polimerasa, luego desarrollar 40 ciclos de dos partes de 95,0 ℃ durante 15 segundos y 60,0 ℃ durante 1 minuto. El dCt se calcula utilizando el Ct promedio de los genes de control B2M (Hs99999907_ml) y RPLPO (Hs99999902_ml). La cuantificación de la expresión de ARNm relativa se deriva usando el Método Ct Comparativo (Applied Biosystems). Los valores de veces de cambio de expresión de ARNm se generan a partir de una muestra normal y la muestra de tumor correspondiente.

Manipulación de muestras para muestras de mieloma

Tras la recolección de aspirado de médula ósea del paciente, las células de mieloma se enriquecen mediante una rápida selección negativa. El procedimiento de enriquecimiento emplea un cóctel de anticuerpos específicos de tipo celular junto con un anticuerpo que se une a los glóbulos rojos RosetteSep (Stem Cell Technologies). El cóctel de anticuerpos tiene anticuerpos con la siguiente especificidad: CD14 (monocitos), CD2 (linfocitos T y NK), CD33 (progenitores mieloides y monocitos), CD41 (plaquetas y megacariocitos), CD45RA (linfocitos B y T sin tratar) y CD66b (granulocitos). Los anticuerpos reticulan los tipos de células que no son de mieloma con los glóbulos rojos en las muestras. Los tipos de células unidas se eliminan utilizando un gradiente de densidad de ficoll modificado. Las

células de mieloma se recogen y se congelan.

El ARN total se aísla utilizando un kit de aislamiento QIAGFN® Group RNEASY® (Valencia, CA) y se cuantifica por espectrofotometría.

El ADN se aísla del flujo mediante la fracción de la columna utilizada en el método de aislamiento de ARN.

Análisis de expresión génica delmieloma en una matriz

El ARN se convierte en ARNc biotinilado mediante un protocolo de amplificación basado en T7 convencional (AFFYMETRIX® Inc., Santa Clara, CA). También se marca un pequeño número de muestras con ≥0,5 - 2,0 µg y posteriormente se hibrida si se producen 6 µg de ARNc. Para el procedimiento automatizado de amplificación T7, el ADNc y el ADNc marcado con biotina se purifican usando el sistema de purificación por PCR AMPURF®, siguiendo el protocolo del fabricante (AGENCOURT® Bioscience Corporation, Beverly, MA). El rendimiento del ARNc se evalúa por espectrofotometría y se fragmentan 10 µg de ARNc y se procesan adicionalmente para hibridación por triplicado en las matrices AFFYMETRIX® Human Genome HG-U133A y HG-U133B GFNFCHIP®. En casos donde el rendimiento del ARNc varió entre 6 µg y 10 µg, la muestra de ARNc completa está fragmentada.

El ARNc para cada muestra se hibrida con las matrices U133A/B por triplicado; operadores, lotes de chip, sitios clínicos y escáneres (GENECHIP® Scanner 3000) se controlan en todo momento. La sustracción del fondo, ajuste de suavizado, correcciones de ruido y los cálculos de señal se realizan con AFFYMETRIX® MAS5.0. Las métricas de control de calidad incluyen: factor de escala (<11) de llamada presente porcentual (> 25), relación β-actina 3':5' (<15) y fondo (<120). Las muestras que quedan fuera de estas métricas se excluyen del análisis posterior.

La puntuación de pureza del mieloma examina la expresión de genes conocidos en la bibliografía que se expresan altamente en las células de mieloma (y sus células precursoras plasmáticas normales), para expresión de genes que se sabe que se expresan altamente en células eritroides, neutrófilos y linfocitos T - véase el listado de 14 marcadores a continuación). La puntuación de mieloma = expresión de marcadores de mieloma (N.º 1-4 a continuación)/eritroide (N.º 5-7) + neutrófilos (N.º 8-11) + linfocitos T (N.º 12-14):

1. 205692_s_en CD38 antígeno CD38 (p45) mieloma/células plasmáticas
2. 201286_en SDC1 sindecano-1 mieloma/células plasmáticas
3. 201891_s_en B2M beta-2 microglobulina mieloma/células plasmáticas
4. 211528_x_en B2M beta-2 microglobulina mieloma/células plasmáticas
5. 37986_en EpoR receptor de eritropoetina célula eritroide
6. 209962_en EpoR receptor de eritropoetina célula eritroide
7. 205838_en GYPA glicoforinA célula eritroide
8. 203948_s_en MPO mieloperoxidasa neutrófilos
9. 203591_s_en receptor del factor 3 estimulante de colonias CSFR3 (granulocitos) neutrófilos
10. 204039_en CEBPACCAAT/proteína de unión potenciadora (C/EBP), neutrófilos alfa
11. 214523_en CEBPECCAAT/proteína de unión potenciadora (C/EBP), neutrófilos épsilon
12. 209603_en GATA3 proteína 3 unión a GATA linfocitos T
13. 209604_s_en GATA4 proteína 4 unión a GATA linfocitos T
14. 205456_en antígeno CD3ECD3E, polipéptido épsilon linfocitos T

Las muestras con una puntuación de pureza de mieloma inferior a 10 se excluyen de análisis posteriores.

Equivalentes

Aunque las realizaciones de la invención se han descrito usando términos específicos, dicha descripción es solo para fines ilustrativos, y debe entenderse que pueden realizarse cambios y variaciones sin apartarse de la invención. Los expertos en la materia reconocerán o serán capaces de determinar, usando únicamente experimentación de rutina, muchos equivalentes de las realizaciones específicas de la invención descritas en el presente documento. Se pretende que tales equivalentes estén abarcados por las siguientes reivindicaciones.

LISTADO DE SECUENCIAS

<110> Smith, Peter G Blakemore, Stephen J. Li, Bin Mulligan, George J. Schu, Matthew C. Millennium Pharmaceuticals, Inc.

<120> BIOMARCADORES DE RESPUESTA A INHIBIDORES DE NAE

<130> MPI11-009P1RNWO2M

<150> US 61/552.686
<151> 28/10/2011

<160> 26

<170> FastSEQ para Windows versión 4.0

<210> 1
<211> 6046
<212> ADN
<213> *Homo sapiens*

<400> 1

tgcgcttccc gcgggcgcgc ggagtgagga cggtgacagc cacgcgcgcg cgtacgcgcc 60
cgatgcagcg cggccccgtg accctagtcg gccgctgaga ggcgcgcgga gtctgggccg 120
ctgccgtcta ggggtcccgt cccgaggcgt ccccggcatc tccggcccga atcccggagt 180
gccgggtcgc gcctgcaccg aaggtcccgg ctcctgtgcc ctccctgcag ccgtcagggc 240
ccgtccccca actccccttt ccgctcaggc agggtcctcg cggcccatgc tggccgctgg 300
ggacccgcgc agcccagacc gttcccgggc cgggcagccg gccaccatgg tggccctgag 360
gcctgtgcag caactccagg ggggctaaag ggctcagagt gcaggccgtg gggcgcgagg 420
gtcccgggcc tgagccccgc gccatggccg gggccatcgc ttcccgcatg agcttcagct 480
ctctcaagag gaagcaaccc aagacgttca ccgtgaggat cgtcaccatg gacgccgaga 540
tggagttcaa ttgcgagatg aagtggaaag ggaaggacct ctttgatttg gtgtgccgga 600
ctctggggct ccgagaaacc tggttctttg gactgcagta cacaatcaag gacacagtgg 660
cctggctcaa aatggacaag aaggtactgg atcatgatgt ttcaaaggaa gaaccagtca 720
cctttcactt cttggccaaa ttttatcctg agaatgctga agaggagctg gttcaggaga 780
tcacacaaca tttattcttc ttacaggtaa agaagcagat tttagatgaa aagatctact 840
gccctcctga ggcttctgtg ctcctggctt cttacgccgt ccaggccaag tatggtgact 900
acgaccccag tgttcacaag cggggatttt tggcccaaga ggaattgctt ccaaaaaggg 960
taataaatct gtatcagatg actccggaaa tgtgggagga gagaattact gcttggtacg 1020
cagagcaccg aggccgagcc agggatgaag ctgaaatgga atatctgaag atagctcagg 1080
acctggagat gtacggtgtg aactactttg caatccggaa taaaaagggc acagagctgc 1140
tgcttggagt ggatgccctg gggcttcaca tttatgaccc tgagaacaga ctgaccccca 1200
agatctcctt cccgtggaat gaaatccgaa acatctcgta cagtgacaag gagtttacta 1260
ttaaaccact ggataagaaa attgatgtct tcaagtttaa ctcctcaaag cttcgtgtta 1320
ataagctgat tctccagcta tgtatcggga accatgatct atttatgagg agaaggaaag 1380
ccgattcttt ggaagttcag cagatgaaag cccaggccag ggaggagaag gctagaaagc 1440
agatggagcg gcagcgcctc gctcgagaga agcagatgag ggaggaggct gaacgcacga 1500
gggatgagtt ggagaggagg ctgctgcaga tgaaagaaga agcaacaatg gccaacgaag 1560
cactgatgcg gtctgaggag acagctgacc tgttggctga aaaggcccag atcaccgagg 1620
aggaggcaaa acttctggcc cagaaggccg cagaggctga gcaggaaatg cagcgcatca 1680
aggccacagc gattcgcacg gaggaggaga agcgcctgat ggagcagaag gtgctggaag 1740
ccgaggtgct ggcactgaag atggctgagg agtcagagag gagggccaaa gaggcagatc 1800
agctgaagca ggacctgcag gaagcacgcg aggcggagcg aagagccaag cagaagctcc 1860
tggagattgc caccaagccc acgtacccgc ccatgaaccc aattccagca ccgttgcctc 1920
ctgacatacc aagcttcaac ctcattggtg acagcctgtc tttcgacttc aaagatactg 1980
acatgaagcg gctttccatg gagatagaga aagaaaaagt ggaatacatg gaaaagagca 2040
agcatctgca ggagcagctc aatgaactca agacagaaat cgaggccttg aaactgaaag 2100

agagggagac agctctggat attctgcaca atgagaactc cgacaggggt ggcagcagca 2160
agcacaatac cattaaaaag ctcaccttgc agagcgccaa gtcccgagtg gccttctttg 2220
aagagctcta gcaggtgacc cagccacccc aggacctgcc acttctcctg ctaccgggac 2280
cgcgggatgg accagatatc aagagagcca tccatagggа gctggctggg ggtttccgtg 2340
ggagctccag aactttcccc agctgagtga agagcccagc ccctcttatg tgcaattgcc 2400
ttgaactacg accctgtaga gatttctctc atggcgttct agttctctga cctgagtctt 2460
tgttttaaga agtatttgtc ttcctttgtc taatgtggga ttcctgactc ccttcgtcca 2520
aggcaccggt gtgtgtgtgt cttgcactcc agagctgacc tccaccgccc agcctgggaa 2580
gtcattgtag ggagtgagac actgaagccc tgagaagcca gtgccatcat ccccaccccg 2640
cccagggttc cggaacattc attcccccac cggtgaggac ctggcatgca gcgaagcagc 2700
ccagcccggc ggatcccagg ccagcacgcc tgccggcttc tcatcgtcag ggagcccgcc 2760
cagagctcgt gacgagcaag tgctgggtcc ccgccaggca ccccgaggcg gcgctctggc 2820
tggcagctgg tggggaatag gcagggcagc tgtggctggg gagagacttt aggcagaagc 2880
tgtgatgcag gctgactgcc agccgagggg ctgggtagtg ccgtgcggga gctgatggta 2940
cagggcactc gctgtccccc tccggccacc ctagaccagg gtccgagagg caggcaggag 3000
ccactcatgt cttccccatt gcccgacgcc catagacgct ccttcctgtg tggggctggg 3060
gtactccctg gtcgtactgc agtcagcacc cgtaacccgg ctatgaccag ggatctgtaa 3120
gccctgtggc tccacaggtg ctgcttctca ctggcccaga ctctgagctc caccggccca 3180
gtctgcacgg cccatctgct tcaccttccc tcccagccac gtgccagtgg ccacagccca 3240
cttcccaacc cactgttgta cccaggcctc actttgctgt tgcccttgtc cctcttcggg 3300
ccctgaattt tctgttccct gggggccagc cagggccctt tgtgcccctc ccagcacagg 3360
cctgatgcag gtgtccactc acaggtggcg ctcacctagg ctgtcacagg acccacctcc 3420
atgccaggca acagagggcc acagaaccac ccccacggct cactccttgg tctggggcca 3480
ccttcttgcc ctttcttttt ttttttttt ttttttttcc gagatggagt ctctctctgt 3540
cacccaggtg ggagtgtggt ggcacaatct cggctcactg caacctccac ctcctgagtt 3600
caagcaattc tcctgcctca gcctcccaag tagctgggac tacaggcatg caccaccaca 3660
cttggccaat gttctgtatt tttaatagggg acggggtttt gccatgttag ctaggctggt 3720
ctcaaactca cacctgggat tacaggcatg agccactgca cccagcccct tcttgccctt 3780
tcttttctcc atggctgatg ctgctgtggc cagccagggc ccttgagatc cttccagttt 3840
ggctgttatg caaagcaggt gatttgtctt aatcagataa aagatagagg ctatgggggc 3900
ctcaagattt ttggagagca gaggtggtct ctggcaattc catctggttt tgagaaactt 3960
agcagctcac agagcacaga gatcctgcct tcttcctact atcaggctga cctaatgggg 4020
ttgggctgct cggcaactgc ttgggtcacc ttgccccaag gaaaccagcc ctgggtgcca 4080
cccagccact tagggtctac agggtgggac tccagaccta gagcgtaagt atggatgttg 4140
tggccctgtg tcttcctagt gtgacccagc caggagcgga agcttcaggc gtttgtaaag 4200
tgaggtctgg ctctgcctcc tccgtttttt ttttttttct gtttctgttt ctgttttttt 4260
tttgagatgg agtctcgctc tgtcgcccag gctggagtgc agtgtcacga tctcagctca 4320
ctgcaacctc cgcctcccag gtacaagaga ttctcctgcc tcagcctccc gagtagctgg 4380
gactacaggc gtgtgccacc atgcctggct aatttttgta tttttagtag agatggggtt 4440
tcgccatgtt agccagactg gtctcgaact cctgacctca ggtgatcctc ccaccccggc 4500
ttcccaaagt tctgggatta taggcgtgag ccaccatgcc cggtctcttc tcagtcttga 4560
agcccatccc tggatttcca ccaggagtta ctttcctcct gacctgtaaa tttgttcttt 4620
aacaatggct gcaggtggga gcatatggtg gtttataaaa acgctgtcgg gctttgtttc 4680
cttctttagc tgccgtgtct acttctgaag tctgggaagt gccaagccac gcggcctcaa 4740
gggagctggc tggtgtttga gctgtggcag aagcacctgg ggctccaggg agcaggctgg 4800
gaactgcagg accttgctca gccaggagca cttccccctc cttgaggcag gaatactgag 4860
gtgcctcccc acagatggag aaggtggaga ggaggatggg cctcaggagc atctcaagcc 4920
ccagtagcag gagaaagaaa gaaagagatg cctggttttc acagactggt tcctgtggct 4980
gggatgactg catccttttt ttttttttt tgagacggag tttttgcctt tgtcgcccag 5040
gctggagtgc aatggcgtga tctcggctca ccgcaacctc cgcctcccgg attcaagcaa 5100
ttctcctgcc tcagcctccc gagtagctgg gattacaggc acgcacctcc acgtccggct 5160
aattttgtat ttttagtgga gacggggttt ctccatgtcg gtcaggctgg tctcgaactc 5220
ccgacctcag gtgatctgcc cacctcggcc tcccaaagtg ctgggattac aggcatgagc 5280
caccgcgctt ggccagactg cgtccttttt aagcgaacat tttagggcct gggagtttgt 5340
caagtaagga agtctcaagc ccaaagagca gcgtcctgac catggtggtt tcattacgag 5400
cccttctgct ggctctcagg cagaagcccc acagcaccgg gaccattcat gaggtcactg 5460
cccagctcat gatgtccgtg aggctgtcct tttggccagt agccgtgtgc agctgtgtgg 5520
cacagatggc ttcgttcatc ctgatcaagg ccccacctca gccacagcag tccccccaac 5580
ctgtgttgtc caccctatta ttcatgtacc tgccaggccc tgctagatag caccccgtgg 5640
cattacataa cacttcatga gtggctgtgt cttgtaattt tggggacagg tttctctctt 5700
tccctctctt ttttttgtca aaagcccaga gactgacaac cagctgcagt gtctaagtgt 5760
tcctcactga cagggtgggg cctcaccacc cctggaggga gcagcgttgg cagggagaca 5820
gcctggccca gtgaccctgg gcccaagcca gcccctccag ggctttcagg gaagcgccat 5880

ccattttcaa agatgtcaaa cgtcacttct tcctgtaggg cccgagtcct gcctcctatc 5940
agggccagat catagaaggc tattttctat tctggggaac gattataact taaatgattg 6000
ttttaataaa aattctaagc tggaaaataa aaaaaaaaaa aaaaaa 6046

<210> 2
<211> 1788
<212> ADN
<213> *Homo sapiens*

<400> 2

atggccgggg ccatcgcttc ccgcatgagc ttcagctctc tcaagaggaa gcaacccaag 60
acgttcaccg tgaggatcgt caccatggac gccgagatgg agttcaattg cgagatgaag 120
tggaaaggga aggacctctt tgatttggtg tgccggactc tggggctccg agaaacctgg 180
ttctttggac tgcagtacac aatcaaggac acagtggcct ggctcaaaat ggacaagaag 240
gtactggatc atgatgtttc aaaggaagaa ccagtcacct ttcacttctt ggccaaattt 300
tatcctgaga atgctgaaga ggagctggtt caggagatca cacaacattt attcttctta 360
caggtaaaga agcagatttt agatgaaaag atctactgcc ctcctgaggc ttctgtgctc 420
ctggcttctt acgccgtcca ggccaagtat ggtgactacg accccagtgt tcacaagcgg 480
ggatttttgg cccaagagga attgcttcca aaaagggtaa taaatctgta tcagatgact 540
ccggaaatgt gggaggagag aattactgct tggtacgcag agcaccgagg ccgagccagg 600
gatgaagctg aaatggaata tctgaagata gctcaggacc tggagatgta cggtgtgaac 660
tactttgcaa tccggaataa aaagggcaca gagctgctgc ttggagtgga tgccctgggg 720
cttcacattt atgaccctga gaacagactg acccccaaga tctccttccc gtggaatgaa 780
atccgaaaca tctcgtacag tgacaaggag tttactatta aaccactgga taagaaaatt 840
gatgtcttca agtttaactc ctcaaagctt cgtgttaata agctgattct ccagctatgt 900
atcgggaacc atgatctatt tatgaggaga aggaaagccg attctttgga agttcagcag 960
atgaaagccc aggccaggga ggagaaggct agaaagcaga tggagcggca gcgcctcgct 1020
cgagagaagc agatgaggga ggaggctgaa cgcacgaggg atgagttgga gaggaggctg 1080
ctgcagatga aagaagaagc aacaatggcc aacgaagcac tgatgcggtc tgaggagaca 1140
gctgacctgt tggctgaaaa ggcccagatc accgaggagg aggcaaaact tctggcccag 1200
aaggccgcag aggctgagca ggaaatgcag cgcatcaagg ccacagcgat tcgcacggag 1260
gaggagaagc gcctgatgga gcagaaggtg ctggaagccg aggtgctggc actgaagatg 1320
gctgaggagt cagagaggag ggccaaagag gcagatcagc tgaagcagga cctgcaggaa 1380
gcacgcgagg cggagcgaag agccaagcag aagctcctgg agattgccac caagcccacg 1440
tacccgccca tgaacccaat tccagcaccg ttgcctcctg acataccaag cttcaacctc 1500
attggtgaca gcctgtcttt cgacttcaaa gatactgaca tgaagcggct ttccatggag 1560
atagagaaag aaaaagtgga atacatggaa aagagcaagc atctgcagga gcagctcaat 1620
gaactcaaga cagaaatcga ggccttgaaa ctgaaagaga gggagacagc tctggatatt 1680
ctgcacaatg agaactccga caggggtggc agcagcaagc acaataccat taaaaagctc 1740
accttgcaga gcgccaagtc ccgagtggcc ttctttgaag agctctag 1788

<210> 3
<211> 595
<212> PRT
<213> *Homo sapiens*

<400> 3

```
Met Ala Gly Ala Ile Ala Ser Arg Met Ser Phe Ser Ser Leu Lys Arg
 1               5                  10                  15
Lys Gln Pro Lys Thr Phe Thr Val Arg Ile Val Thr Met Asp Ala Glu
            20                  25                  30
Met Glu Phe Asn Cys Glu Met Lys Trp Lys Gly Lys Asp Leu Phe Asp
        35                  40                  45
Leu Val Cys Arg Thr Leu Gly Leu Arg Glu Thr Trp Phe Phe Gly Leu
    50                  55                  60
Gln Tyr Thr Ile Lys Asp Thr Val Ala Trp Leu Lys Met Asp Lys Lys
65                  70                  75                  80
Val Leu Asp His Asp Val Ser Lys Glu Glu Pro Val Thr Phe His Phe
                85                  90                  95
Leu Ala Lys Phe Tyr Pro Glu Asn Ala Glu Glu Glu Leu Val Gln Glu
            100                 105                 110
Ile Thr Gln His Leu Phe Phe Leu Gln Val Lys Lys Gln Ile Leu Asp
        115                 120                 125
```

```
Glu Lys Ile Tyr Cys Pro Pro Glu Ala Ser Val Leu Leu Ala Ser Tyr
    130                 135                 140
Ala Val Gln Ala Lys Tyr Gly Asp Tyr Asp Pro Ser Val His Lys Arg
145                 150                 155                 160
Gly Phe Leu Ala Gln Glu Glu Leu Leu Pro Lys Arg Val Ile Asn Leu
                165                 170                 175
Tyr Gln Met Thr Pro Glu Met Trp Glu Glu Arg Ile Thr Ala Trp Tyr
            180                 185                 190
Ala Glu His Arg Gly Arg Ala Arg Asp Glu Ala Glu Met Glu Tyr Leu
        195                 200                 205
Lys Ile Ala Gln Asp Leu Glu Met Tyr Gly Val Asn Tyr Phe Ala Ile
    210                 215                 220
Arg Asn Lys Lys Gly Thr Glu Leu Leu Leu Gly Val Asp Ala Leu Gly
225                 230                 235                 240
Leu His Ile Tyr Asp Pro Glu Asn Arg Leu Thr Pro Lys Ile Ser Phe
                245                 250                 255
Pro Trp Asn Glu Ile Arg Asn Ile Ser Tyr Ser Asp Lys Glu Phe Thr
            260                 265                 270
Ile Lys Pro Leu Asp Lys Lys Ile Asp Val Phe Lys Phe Asn Ser Ser
        275                 280                 285
Lys Leu Arg Val Asn Lys Leu Ile Leu Gln Leu Cys Ile Gly Asn His
    290                 295                 300
Asp Leu Phe Met Arg Arg Arg Lys Ala Asp Ser Leu Glu Val Gln Gln
305                 310                 315                 320
Met Lys Ala Gln Ala Arg Glu Glu Lys Ala Arg Lys Gln Met Glu Arg
                325                 330                 335
Gln Arg Leu Ala Arg Glu Lys Gln Met Arg Glu Glu Ala Glu Arg Thr
            340                 345                 350
Arg Asp Glu Leu Glu Arg Arg Leu Leu Gln Met Lys Glu Glu Ala Thr
        355                 360                 365
Met Ala Asn Glu Ala Leu Met Arg Ser Glu Glu Thr Ala Asp Leu Leu
    370                 375                 380
Ala Glu Lys Ala Gln Ile Thr Glu Glu Glu Ala Lys Leu Leu Ala Gln
385                 390                 395                 400
Lys Ala Ala Glu Ala Glu Gln Glu Met Gln Arg Ile Lys Ala Thr Ala
                405                 410                 415
Ile Arg Thr Glu Glu Glu Lys Arg Leu Met Glu Gln Lys Val Leu Glu
            420                 425                 430
Ala Glu Val Leu Ala Leu Lys Met Ala Glu Glu Ser Glu Arg Arg Ala
        435                 440                 445
Lys Glu Ala Asp Gln Leu Lys Gln Asp Leu Gln Glu Ala Arg Glu Ala
    450                 455                 460
Glu Arg Arg Ala Lys Gln Lys Leu Leu Glu Ile Ala Thr Lys Pro Thr
465                 470                 475                 480
Tyr Pro Pro Met Asn Pro Ile Pro Ala Pro Leu Pro Pro Asp Ile Pro
                485                 490                 495
Ser Phe Asn Leu Ile Gly Asp Ser Leu Ser Phe Asp Phe Lys Asp Thr
            500                 505                 510
Asp Met Lys Arg Leu Ser Met Glu Ile Glu Lys Glu Lys Val Glu Tyr
        515                 520                 525
Met Glu Lys Ser Lys His Leu Gln Glu Gln Leu Asn Glu Leu Lys Thr
    530                 535                 540
Glu Ile Glu Ala Leu Lys Leu Lys Glu Arg Glu Thr Ala Leu Asp Ile
545                 550                 555                 560
Leu His Asn Glu Asn Ser Asp Arg Gly Gly Ser Ser Lys His Asn Thr
                565                 570                 575
Ile Lys Lys Leu Thr Leu Gln Ser Ala Lys Ser Arg Val Ala Phe Phe
            580                 585                 590
Glu Glu Leu
        595
```

<210> 4
<211> 8789
<212> ADN
<213> *Homo sapiens*

<400> 4

atgctcagtg gcttctcgac aagttggcag caacaacacg gccctggtcg tcgtcgccgc 60
tgcggtaacg gagcggtttg ggtggcggag cctgcgttcg cgccttcccg ctctcctcgg 120
gaggcccttc ctgctctccc ctaggctccg cggccgccca gggggtggga gcgggtgagg 180
ggagccaggc gcccagcgag agaggccccc cgccgcaggg cggcccggga gctcgaggcg 240
gtccggcccg cgcgggcagc ggcgcggcgc tgaggagggg cggcctggcc gggacgcctc 300
ggggcggggg ccgaggagct ctccgggccg ccggggaaag ctacgggccc ggtgcgtccg 360
cggaccagca gcgcgggaga gcggactccc ctcgccaccg cccgagccca ggttatcctg 420
aatacatgtc taacaatttt ccttgcaacg ttagctgttg tttttcactg tttccaaagg 480
atcaaaattg cttcagaaat tggagacata tttgatttaa aaggaaaaac ttgaacaaat 540
ggacaatatg tctattacga atacaccaac aagtaatgat gcctgtctga gcattgtgca 600
tagtttgatg tgccatagac aaggtggaga gagtgaaaca tttgcaaaaa gagcaattga 660
aagtttggta aagaagctga aggagaaaaa agatgaattg gattctttaa taacagctat 720
aactacaaat ggagctcatc ctagtaaatg tgttaccata cagagaacat tggatgggag 780
gcttcaggtg gctggtcgga aaggatttcc tcatgtgatc tatgcccgtc tctggaggtg 840
gcctgatctt cacaaaaatg aactaaaaca tgttaaatat tgtcagtatg cgtttgactt 900
aaaatgtgat agtgtctgtg tgaatccata tcactacgaa cgagttgtat cacctggaat 960
tgatctctca ggattaacac tgcagagtaa tgctccatca agtatgatgg tgaaggatga 1020
atatgtgcat gactttgagg gacagccatc gttgtccact gaaggacatt caattcaaac 1080
catccagcat ccaccaagta atcgtgcatc gacagagaca tacagcaccc cagctctgtt 1140
agccccatct gagtctaatg ctaccagcac tgccaacttt cccaacattc ctgtggcttc 1200
cacaagtcag cctgccagta tactgggggg cagccatagt gaaggactgt tgcagatagc 1260
atcagggcct cagccaggac agcagcagaa tggatttact ggtcagccag ctacttacca 1320
tcataacagc actaccacct ggactggaag taggactgca ccatacacac ctaatttgcc 1380
tcaccaccaa aacggccatc ttcagcacca cccgcctatg ccgccccatc ccggacatta 1440
ctggcctgtt cacaatgagc ttgcattcca gcctcccatt tccaatcatc ctgctcctga 1500
gtattggtgt tccattgctt actttgaaat ggatgttcag gtaggagaga catttaaggt 1560
tccttcaagc tgccctattg ttactgttga tggatacgtg gacccttctg gaggagatcg 1620
cttttgtttg ggtcaactct ccaatgtcca caggacagaa gccattgaga gagcaaggtt 1680
gcacataggc aaaggtgtgc agttggaatg taaaggtgaa ggtgatgttt gggtcaggtg 1740
ccttagtgac cacgcggtct ttgtacagag ttactactta gacagagaag ctgggcgtgc 1800
acctggagat gctgttcata agatctaccc aagtgcatat ataaaggtct ttgatttgcg 1860
tcagtgtcat cgacagatgc agcagcaggc ggctactgca caagctgcag cagctgccca 1920
ggcagcagcc gtggcaggaa acatccctgg cccaggatca gtaggtggaa tagctccagc 1980
tatcagtctg tcagctgctg ctggaattgg tgttgatgac cttcgtcgct tatgcatact 2040
caggatgagt tttgtgaaag gctggggacc ggattaccca agacagagca tcaaagaaac 2100
accttgctgg attgaaattc acttacaccg ggccctccag ctcctagacg aagtacttca 2160
taccatgccg attgcagacc cacaaccttt agactgaggt cttttaccgt tggggccctt 2220
aaccttatca ggatggtgga ctacaaaata caatcctgtt tataatctga agatatattt 2280
cacttttgtt ctgctttatc ttttcataaa gggttgaaaa tgtgtttgct gccttgctcc 2340
tagcagacag aaactggatt aaaacaattt ttttttcct cttcagaact tgtcaggcat 2400
ggctcagagc ttgaagatta ggagaaacac attcttatta attcttcacc tgttatgtat 2460
gaaggaatca ttccagtgct agaaaattta gccctttaaa acgtcttaga gccttttatc 2520
tgcagaacat cgatatgtat atcattctac agaataatcc agtattgctg attttaaagg 2580
cagagaagtt ctcaaagtta attcacctat gttattttgt gtacaagttg ttattgttga 2640
acatacttca aaaataatgt gccatgtggg tgagttaatt ttaccaagag taactttact 2700
ctgtgtttaa aaagtaagtt aataatgtat tgtaatcttt catccaaaat atttttgca 2760
agttatatta gtgaagatgg tttcaattca gattgtcttg caacttcagt tttatttttg 2820
ccaaggcaaa aaactcttaa tctgtgtgta tattgagaat cccttaaaat taccagacaa 2880
aaaaatttaa aattacgttt gttattccta gtggatgact gttgatgaag tatacttttc 2940
ccctgttaaa cagtagttgt attcttctgt atttctaggc acaaggttgg ttgctaagaa 3000
gcctataaga ggaatttctt ttccttcatt catagggaaa ggttttgtat tttttaaaac 3060
actaaaagca gcgtcactct acctaatgtc tcactgttct gcaaaggtgg caatgcttaa 3120
actaaataat gaataaactg aatattttgg aaactgctaa attctatgtt aaatactgtg 3180
cagaataatg gaaacattac agttcataat aggtagtttg gatatttttg tacttgattt 3240
gatgtgactt tttttggtat aatgtttaaa tcatgtatgt tatgatattg tttaaaattc 3300
agtttttgta tcttggggca agactgcaaa ctttttttata tcttttggtt attctaagcc 3360
ctttgccatc aatgatcata tcaattggca gtgactttgt atagagaatt taagtagaaa 3420
agttgcagat gtattgactg taccacagac acaatatgta tgctttttac ctagctggta 3480

gcataaataa aactgaatct caacatacaa agttgaattc taggtttgat ttttaagatt 3540
ttttttttct tttgcacttt tgagtccaat ctcagtgatg aggtaccttc tactaaatga 3600
caggcaacag ccagttctat tgggcagctt tgtttttttc cctcacactc taccgggact 3660
tccccatgga cattgtgtat catgtgtaga gttggttttt ttttttttta atttttattt 3720
tactatagca gaaatagacc tgattatcta caagatgata aatagattgt ctacaggata 3780
aatagtatga aataaaatca aggattatct ttcagatgtg tttacttttg cctggagaac 3840
ttttagctat agaaacactt gtgtgatgat agtcctcctt atatcacctg gaatgaacac 3900
agcttctact gccttgctca gaaggtcttt taaatagacc atcctagaaa ccactgagtt 3960
tgcttatttc tgtgatttaa acatagatct tgatccaagc tacatgactt ttgtctttaa 4020
ataacttatc taccacctca tttgtactct tgattactta caaattcttt cagtaaacac 4080
ctaattttct tctgtaaaag tttggtgatt taagttttat tggcagtttt ataaaaagac 4140
atcttctcta gaaattgcta actttaggtc cattttactg tgaatgagga ataggagtga 4200
gttttagaat aacagatttt taaaaatcca gatgatttga ttaaaacctt aatcatacat 4260
tgacataatt cattgcttct ttttttgag atatggagtc ttgctgtgtt gcccaggcag 4320
gagtgcagtg gtatgatctc agctcactgc aacctctgcc tcccgggttc aactgattct 4380
cctgcctcag cctccctggt agctaggatt acaggtgccc gccaccatgc ctggctaact 4440
tttgtagttt tagtagagac ggggttttgc ctgttggcca ggctggtctt gaactcctga 4500
cctcaagtga tccatccacc ttggcctccc aaagtgctgg gattacgggc gtgagccact 4560
gtccctggcc tcattgttcc cttttctact ttaaggaaag ttttcatgtt taatcatctg 4620
gggaaagtat gtgaaaaata tttgttaaga agtatctctt tggagccaag ccacctgtct 4680
tggtttcttt ctactaagag ccataaagta tagaaatact tctagttgtt aagtgcttat 4740
atttgtacct agatttagtc acacgctttt gagaaaacat ctagtatgtt atgatcagct 4800
attcctgaga gcttggttgt taatctatat ttctatttct tagtggtagt catctttgat 4860
gaataagact aaagattctc acaggtttaa aattttatgt ctactttaag ggtaaaatta 4920
tgaggttatg gttctgggtg ggttttctct agctaattca tatctcaaag agtctcaaaa 4980
tgttgaattt cagtgcaagc tgaatgagag atgagccatg tacacccacc gtaagacctc 5040
attccatgtt tgtccagtgc ctttcagtgc attatcaaag ggaatccttc atggtgttgc 5100
ctttattttc cggggagtag atcgtgggat atagtctatc tcatttttaa tagtttaccg 5160
cccctggtat acaaagataa tgacaataaa tcactgccat ataaccttgc tttttccaga 5220
aacatggctg ttttgtattg ctgtaaccac taaataggtt gcctatacca ttcctcctgt 5280
gaacagtgca gatttacagg ttgcatggtc tggcttaagg agagccatac ttgagacatg 5340
tgagtaaact gaactcatat tagctgtgct gcatttcaga cttaaaatcc atttttgtgg 5400
ggcagggtgt ggtgtgtaaa ggggggtgtt tgtaatacaa gttgaaggca aaataaaatg 5460
tcctgtctcc cagatgatat acatcttatt atttttaaag tttattgcta attgtaggaa 5520
ggtgagttgc aggtatcttt gactatggtc atctggggaa ggaaaatttt acattttact 5580
attaatgctc cttaagtgtc tatggaggtt aaagaataaa atggtaaatg tttctgtgcc 5640
tggtttgatg gtaactggtt aatagttact caccatttta tgcagagtca cattagttca 5700
cacccttct gagagccttt tgggagaagc agttttattc tctgagtgga acagagttct 5760
ttttgttgat aatttctagt ttgctccctt cgttattgcc aactttactg gcattttatt 5820
taatgatagc agattgggaa aatggcaaat ttaggttacg gaggtaaatg agtatatgaa 5880
agcaattacc tctaaagcca gttaacaatt attttgtagg tggggtacac tcagcttaaa 5940
gtaatgcatt ttttttccc gtaaaggcag aatccatctt gttgcagata gctatctaaa 6000
taatctcata tcctcttttg caaagactac agagaatagg ctatgacaat cttgttcaag 6060
cctttccatt tttttccctg ataactaagt aatttctttg aacataccaa gaagtatgta 6120
aaaagtccat ggccttattc atccacaaag tggcatccta ggcccagcct tatccctagc 6180
agttgtccca gtgctgctag gttgcttatc ttgtttatct ggaatcactg tggagtgaaa 6240
ttttccacat catccagaat tgccttattt aagaagtaaa acgttttaat ttttagcctt 6300
tttttggtgg agttatttaa tatgtatatc agaggatata ctagatggta acatttcttt 6360
ctgtgcttgg ctatctttgt ggacttcagg ggcttctaaa acagacagga ctgtgttgcc 6420
tttactaaat ggtctgagac agctatggtt ttgaattttt agtttttttt ttttaaccca 6480
cttcccctcc tggtctcttc cctctctgat aattaccatt catatgtgag tgttagtgtg 6540
cctccttta gcattttctt cttctctttc tgattcttca tttctgactg cctaggcaag 6600
gaaaccagat aaccaaactt actagaacgt tctttaaaac acaagtacaa actctgggac 6660
aggacccaag acactttcct gtgaagtgct gaaaaagacc tcattgtatt ggcatttgat 6720
atcagtttga tgtagcttag agtgcttcct gattcttgct gagtttcagg tagttgagat 6780
agagagaagt gagtcatatt catattttcc cccttagaat aatattttga aaggtttcat 6840
tgcttccact tgaatgctgc tcttacaaaa actggggtta caagggttac taaattagca 6900
tcagtagcca gaggcaatac cgttgtctgg aggacaccag caaacaacac acaacaaagc 6960
aaaacaaacc ttgggaaact aaggccattt gttttgtttt ggtgtcccct ttgaagccct 7020
gccttctggc cttactcctg tacagatatt tttgacctat aggtgccttt atgagaattg 7080
agggtctgac atcctgcccc aaggagtagc taaagtaatt gctagtgttt tcagggattt 7140
taacatcaga ctggaatgaa tgaatgaaac tttttgtcct tttttttttct gttttttttt 7200
ttctaatgta gtaaggacta aggaaaacct ttggtgaaga caatcatttc tctctgttga 7260

```
tgtggatact tttcacaccg tttatttaaa tgctttctca ataggtccag agccagtgtt 7320
cttgttcaac ctgaaagtaa tggctctggg ttgggccaga cagttgcact ctctagtttg 7380
ccctctgcca caaatttgat gtgtgacctt tgggcaagtc atttatcttc tctgggcctt 7440
agttgcctca tctgtaaaat gagggagttg gagtagatta attattccag ctctgaaatt 7500
ctaagtgacc ttggctacct tgcagcagtt ttggatttct tccttatctt tgttctgctg 7560
tttgaggggg ctttttactt atttccatgt tattcaaagg agactaggct tgatatttta 7620
ttactgttct tttatggaca aaaggttaca tagtatgccc ttaagactta attttaacca 7680
aaggcctagc accaccttag gggctgcaat aaacacttaa cgcgcgtgcg cacgcgcgcg 7740
cgcacacaca cacacacaca cacacacaca cacaggtcag agtttaaggc tttcgagtca 7800
tgacattcta gcttttgaat tgcgtgcaca cacacacgca cgcacacact ctggtcagag 7860
tttattaagg ctttcgagtc atgacattat agcttttgag ttggtgtgtg tgacaccacc 7920
ctcctaagtg gtgtgtgctt gtaatttttt ttttcagtga aaatggattg aaaacctgtt 7980
gttaatgctt agtgatatta tgctcaaaac aaggaaattc ccttgaaccg tgtcaattaa 8040
actggtttat atgactcaag aaaacaatac cagtagatga ttattaactt tattcttggc 8100
tctttttagg tccattttga ttaagtgact tttggctgga tcattcagag ctctcttcta 8160
gcctaccctt ggatgagtac aattaatgaa attcatattt tcaaggacct gggagccttc 8220
cttggggctg ggttgagggt ggggggttgg ggagtcctgg tagaggccag ctttgtggta 8280
gctggagagg aagggatgaa accagctgct gttgcaaagg ctgcttgtca ttgatagaag 8340
gactcacggg cttggattga ttaagactaa acatggagtt ggcaaacttt cttcaagtat 8400
tgagttctgt tcaatgcatt ggacatgtga tttaagggaa aagtgtgaat gcttatagat 8460
gatgaaaacc tggtgggctg cagagcccag tttagaagaa gtgagttggg ggttggggac 8520
agatttggtg gtggtatttc ccaactgttt cctcccctaa attcagagga atgcagctat 8580
gccagaagcc agagaagagc cactcgtagc ttctgctttg gggacaactg gtcagttgaa 8640
agtcccagga gttcctttgt ggctttctgt atacttttgc ctggttaaag tctgtggcta 8700
aaaaatagtc gaacctttct tgagaactct gtaacaaagt atgtttttga ttaaaagaga 8760
aagccaacta aaaaaaaaaa aaaaaaaaa                                  8789
```

<210> 5
<211> 1659
<212> ADN
<213> *Homo sapiens*

<400> 5

atggacaata tgtctattac gaatacacca acaagtaatg atgcctgtct gagcattgtg 60
catagtttga tgtgccatag acaaggtgga gagagtgaaa catttgcaaa aagagcaatt 120
gaaagtttgg taaagaagct gaaggagaaa aaagatgaat tggattcttt aataacagct 180
ataactacaa atggagctca tcctagtaaa tgtgttacca tacagagaac attggatggg 240
aggcttcagg tggctggtcg gaaaggattt cctcatgtga tctatgcccg tctctggagg 300
tggcctgatc ttcacaaaaa tgaactaaaa catgttaaat attgtcagta tgcgtttgac 360
ttaaaatgtg atagtgtctg tgtgaatcca tatcactacg aacgagttgt atcacctgga 420
attgatctct caggattaac actgcagagt aatgctccat caagtatgat ggtgaaggat 480
gaatatgtgc atgactttga gggacagcca tcgttgtcca ctgaaggaca ttcaattcaa 540
accatccagc atccaccaag taatcgtgca tcgacagaga catacagcac cccagctctg 600
ttagccccat ctgagtctaa tgctaccagc actgccaact ttcccaacat tcctgtggct 660
tccacaagtc agcctgccag tatactgggg ggcagccata gtgaaggact gttgcagata 720
gcatcagggc ctcagccagg acagcagcag aatggattta ctggtcagcc agctacttac 780
catcataaca gcactaccac ctggactgga agtaggactg caccatacac acctaatttg 840
cctcaccacc aaaacggcca tcttcagcac cacccgccta tgccgcccca tcccggacat 900
tactggcctg ttcacaatga gcttgcattc cagcctccca tttccaatca tcctgctcct 960
gagtattggt gttccattgc ttactttgaa atggatgttc aggtaggaga gacatttaag 1020
gttccttcaa gctgccctat tgttactgtt gatggatacg tggacccttc tggaggagat 1080
cgcttttgtt tgggtcaact ctccaatgtc cacaggacag aagccattga gagagcaagg 1140
ttgcacatag gcaaaggtgt gcagttggaa tgtaaaggtg aaggtgatgt ttgggtcagg 1200
tgccttagtg accacgcggt ctttgtacag agttactact tagacagaga agctgggcgt 1260
gcacctggag atgctgttca taagatctac ccaagtgcat atataaaggt ctttgatttg 1320
cgtcagtgtc atcgacagat gcagcagcag gcggctactg cacaagctgc agcagctgcc 1380
caggcagcag ccgtggcagg aaacatccct ggcccaggat cagtaggtgg aatagctcca 1440
gctatcagtc tgtcagctgc tgctggaatt ggtgttgatg accttcgtcg cttatgcata 1500
ctcaggatga gttttgtgaa aggctgggga ccggattacc caagacagag catcaaagaa 1560
acaccttgct ggattgaaat tcacttacac cgggccctcc agctcctaga cgaagtactt 1620
cataccatgc cgattgcaga cccacaacct ttagactga 1659

<210> 6
<211> 552
<212> PRT
<213> *Homo sapiens*

<400> 6

```
Met Asp Asn Met Ser Ile Thr Asn Thr Pro Thr Ser Asn Asp Ala Cys
 1               5                  10                  15
Leu Ser Ile Val His Ser Leu Met Cys His Arg Gln Gly Gly Glu Ser
            20                  25                  30
Glu Thr Phe Ala Lys Arg Ala Ile Glu Ser Leu Val Lys Lys Leu Lys
        35                  40                  45
Glu Lys Lys Asp Glu Leu Asp Ser Leu Ile Thr Ala Ile Thr Thr Asn
    50                  55                  60
Gly Ala His Pro Ser Lys Cys Val Thr Ile Gln Arg Thr Leu Asp Gly
65                  70                  75                  80
Arg Leu Gln Val Ala Gly Arg Lys Gly Phe Pro His Val Ile Tyr Ala
                85                  90                  95
Arg Leu Trp Arg Trp Pro Asp Leu His Lys Asn Glu Leu Lys His Val
            100                 105                 110
Lys Tyr Cys Gln Tyr Ala Phe Asp Leu Lys Cys Asp Ser Val Cys Val
        115                 120                 125
Asn Pro Tyr His Tyr Glu Arg Val Val Ser Pro Gly Ile Asp Leu Ser
    130                 135                 140
Gly Leu Thr Leu Gln Ser Asn Ala Pro Ser Ser Met Met Val Lys Asp
145                 150                 155                 160
Glu Tyr Val His Asp Phe Glu Gly Gln Pro Ser Leu Ser Thr Glu Gly
                165                 170                 175
His Ser Ile Gln Thr Ile Gln His Pro Pro Ser Asn Arg Ala Ser Thr
            180                 185                 190
Glu Thr Tyr Ser Thr Pro Ala Leu Leu Ala Pro Ser Glu Ser Asn Ala
        195                 200                 205
Thr Ser Thr Ala Asn Phe Pro Asn Ile Pro Val Ala Ser Thr Ser Gln
    210                 215                 220
Pro Ala Ser Ile Leu Gly Gly Ser His Ser Glu Gly Leu Leu Gln Ile
225                 230                 235                 240
Ala Ser Gly Pro Gln Pro Gly Gln Gln Gln Asn Gly Phe Thr Gly Gln
                245                 250                 255
Pro Ala Thr Tyr His His Asn Ser Thr Thr Thr Trp Thr Gly Ser Arg
            260                 265                 270
Thr Ala Pro Tyr Thr Pro Asn Leu Pro His His Gln Asn Gly His Leu
        275                 280                 285
Gln His His Pro Pro Met Pro Pro His Pro Gly His Tyr Trp Pro Val
    290                 295                 300
His Asn Glu Leu Ala Phe Gln Pro Pro Ile Ser Asn His Pro Ala Pro
305                 310                 315                 320
Glu Tyr Trp Cys Ser Ile Ala Tyr Phe Glu Met Asp Val Gln Val Gly
                325                 330                 335
Glu Thr Phe Lys Val Pro Ser Ser Cys Pro Ile Val Thr Val Asp Gly
            340                 345                 350
Tyr Val Asp Pro Ser Gly Gly Asp Arg Phe Cys Leu Gly Gln Leu Ser
        355                 360                 365
Asn Val His Arg Thr Glu Ala Ile Glu Arg Ala Arg Leu His Ile Gly
    370                 375                 380
Lys Gly Val Gln Leu Glu Cys Lys Gly Glu Gly Asp Val Trp Val Arg
385                 390                 395                 400
Cys Leu Ser Asp His Ala Val Phe Val Gln Ser Tyr Tyr Leu Asp Arg
                405                 410                 415
Glu Ala Gly Arg Ala Pro Gly Asp Ala Val His Lys Ile Tyr Pro Ser
            420                 425                 430
Ala Tyr Ile Lys Val Phe Asp Leu Arg Gln Cys His Arg Gln Met Gln
        435                 440                 445
Gln Gln Ala Ala Thr Ala Gln Ala Ala Ala Ala Ala Gln Ala Ala Ala
    450                 455                 460
```

```
Val Ala Gly Asn Ile Pro Gly Pro Gly Ser Val Gly Gly Ile Ala Pro
465                 470                 475                 480
Ala Ile Ser Leu Ser Ala Ala Ala Gly Ile Gly Val Asp Asp Leu Arg
                485                 490                 495
Arg Leu Cys Ile Leu Arg Met Ser Phe Val Lys Gly Trp Gly Pro Asp
            500                 505                 510
Tyr Pro Arg Gln Ser Ile Lys Glu Thr Pro Cys Trp Ile Glu Ile His
        515                 520                 525
Leu His Arg Ala Leu Gln Leu Leu Asp Glu Val Leu His Thr Met Pro
    530                 535                 540
Ile Ala Asp Pro Gln Pro Leu Asp
545                 550
```

<210> 7
<211> 5778
<212> ADN
<213> *Homo sapiens*

<400> 7

gtgacacaat tacaacaact ttgtgctggt gccggggaag tttgtgtctc caacgaatcc 60
cctcagtgct ccccagcccc gcgcgctccg gccgttcccg ccgtccccgc ctgtggctgc 120
cccctgccca accccgcgat gtgaccctac agccgaaagc cgccgctgcc gacccggggg 180
ctccgcagcc cctgccgccg ccgccgccgc cttcaccgcc gccgcgttgg gatttttcgt 240
cgccgccgcc cgcggcggag gaggaggcgg cgataaagtt ggtgtgctgg tcccgcgcgc 300
agattggggg cgtcactgcg ggccccggtc cgaggggggg tgtcggcgtt ggagttgtga 360
attcgctgcg tttccatgaa atcctgcgga gtgtcgctcg ctaccgccgc cgctgccgcc 420
gccgctttcg gtgatgagga aaagaaaatg gcggcgggaa aagcgagcgg cgagagcgag 480
gaggcgtccc ccagcctgac agccgaggag agggaggcgc tcggcggact ggacagccgc 540
ctctttgggt tcgtgagatt tcatgaagat ggcgccagga cgaaggccct actgggcaag 600
gctgttcgct gctatgaatc tctaatctta aaagctgaag gaaaagtgga gtctgatttc 660
ttttgtcaat taggtcactt caacctctta ttggaagatt atccaaaagc attatctgca 720
taccagaggt actacagttt acagtctgac tactggaaga atgctgcctt tttatatggt 780
cttggtttgg tctacttcca ttataatgca tttcagtggg caattaaagc atttcaggag 840
gtgctttatg ttgatcccag cttttgtcga gccaaggaaa ttcatttacg acttgggctt 900
atgttcaaag tgaacacaga ctatgagtct agtttaaagc attttcagtt agctttggtt 960
gactgtaatc cctgcacttt gtccaatgct gaaattcaat ttcacattgc ccacttatat 1020
gaaacccaga ggaaatatca ttctgcaaaa gaagcttatg aacaactttt gcagacagag 1080
aatctttctg cacaagtaaa agcaactgtc ttacaacagt taggttggat gcatcacact 1140
gtagatctcc tgggagataa agccaccaag gaaagctatg ctattcagta tctccaaaag 1200
tccttggaag cagatcctaa ttctggccag tcctggtatt tcctcggaag gtgctattca 1260
agtattggga aagttcagga tgcctttata tcttacaggc agtctattga taaatcagaa 1320
gcaagtgcag atacatggtg ttcaataggt gtgctatatc agcagcaaaa tcagcccatg 1380
gatgctttac aggcctatat ttgtgctgta caattggacc atggccatgc tgcagcctgg 1440
atggacctag gcactctcta tgaatcctgc aaccagcctc aggatgccat taaatgctac 1500
ttaaatgcaa ctagaagcaa aagttgtagt aatacctctg cacttgcagc acgaattaag 1560
tatttacagg ctcagttgtg taaccttcca caaggtagtc tacagaataa aactaaatta 1620
cttcctagta ttgaggaggc gtggagccta ccaattcccg cagagcttac ctccaggcag 1680
ggtgccatga acacagcaca gcagaatact tctgacaatt ggagtggtgg acatgctgtg 1740
tcacatcctc cagtacagca acaagctcat tcatggtgtt tgacaccaca gaaattacag 1800
catttggaac agctccgcgc aaatagaaat aatttaaatc cagcacagaa actgatgctg 1860
gaacagctgg aaagtcagtt tgtcttaatg caacaacacc aaatgagacc aacaggagtt 1920
gcacaggtac gatctactgg aattcctaat gggccaacag ctgactcatc actgcctaca 1980
aactcagtct ctggccagca gccacagctt gctctgacca gagtgcctag cgtctctcag 2040
cctggagtcc gtcctgcctg ccctgggcag cctttggcca atggaccctt ttctgcaggc 2100
catgttccct gtagcacatc aagaacgctg ggaagtacag acactatttt gataggcaat 2160
aatcatataa caggaagtgg aagtaatgga aacgtgcctt acctgcagcg aaacgcactc 2220
actctacctc ataaccgcac aaacctgacc agcagcgcag aggagccgtg gaaaaaccaa 2280
ctatctaact ccactcaggg gcttcacaaa ggtcagagtt cacattcggc aggtcctaat 2340
ggtgaacgac ctctctcttc cactgggcct tcccagcatc tccaggcagc tggctctggt 2400
attcagaatc agaacggaca tcccaccctg cctagcaatt cagtaacaca gggggctgct 2460
ctcaatcacc tctcctctca cactgctacc tcaggtggac aacaaggcat taccttaacc 2520
aaagagagca agccttcagg aaacatattg acggtgcctg aaacaagcag gcacactgga 2580

gagacaccta acagcactgc cagtgtcgag ggacttccta atcatgtcca tcagatgacg 2640
gcagatgctg tttgcagtcc tagccatgga gattctaagt caccaggttt actaagttca 2700
gacaatcctc agctctctgc cttgttgatg ggaaaagcca ataacaatgt gggtactgga 2760
acctgtgaca aagtcaataa catccaccca gctgttcata caaagactga taactctgtt 2820
gcctcttcac catcttcagc catttcaaca gcaacacctt ctccaaaatc cactgagcag 2880
acaaccacaa acagtgttac cagccttaac agccctcaca gtgggctaca cacaattaat 2940
ggagaaggga tggaagaatc tcagagcccc atgaaaacag atctgcttct ggttaaccac 3000
aaacctagtc cacagatcat accatcaatg tctgtgtcca tataccccag ctcagcagaa 3060
gttctgaagg catgcaggaa tctaggtaaa aatggcttat ctaacagtag cattttgttg 3120
gataaatgtc cacctccaag accaccatct tcaccatacc ctcccttgcc aaaggacaag 3180
ttgaatccac ctacacctag tatttacttg gaaaataaac gtgatgcttt ctttcctcca 3240
ttacatcaat tttgtacaaa tccgaacaac cctgttacag taatacgtgg ccttgctgga 3300
gctcttaagt tagacctggg acttttctct actaaaactt tggtggaagc taacaatgaa 3360
catatggtag aagtgaggac acagttgttg cagccagcag atgaaaactg ggatcccact 3420
ggaacaaaga aaatctggca ttgtgaaagt aatagatctc atactacaat tgctaaatat 3480
gcacagtacc aggcctcctc attccaggaa tcattgagag aagaaaatga aaaaagaagt 3540
catcataaag accactcaga tagtgaatct acatcgtcag ataattctgg gaggaggagg 3600
aaaggaccct ttaaaaccat aaagtttggg accaatattg acctatctga tgacaaaaag 3660
tggaagttgc agctacatga gctgactaaa cttcctgctt ttgtgcgtgt cgtatcagca 3720
ggaaatcttc taagccatgt tggtcatacc atattgggca tgaacacagt tcaactatac 3780
atgaaagttc cagggagcag aacaccaggt catcaggaaa ataacaactt ctgttcagtt 3840
aacataaata ttggcccagg tgactgtgaa tggtttgttg ttcctgaagg ttactggggt 3900
gttctgaatg acttctgtga aaaaaataat ttgaatttcc taatgggttc ttggtggccc 3960
aatcttgaag atctttatga agcaaatgtt ccagtgtata ggtttattca gcgacctgga 4020
gatttggtct ggataaatgc aggcactgtt cattgggttc aggctattgg ctggtgcaac 4080
aacattgctt ggaatgttgg tccacttaca gcctgccagt ataaattggc agtggaacgg 4140
tacgaatgga acaaattgca aagtgtgaag tcaatagtac ccatggttca tctttcctgg 4200
aatatggcac gaaatatcaa ggtctcagat ccaaagcttt ttgaaatgat taagtattgt 4260
cttctaagaa ctctgaagca atgtcagaca ttgagggaag ctctcattgc tgcaggaaaa 4320
gagattatat ggcatgggcg gacaaaagaa gaaccagctc attactgtag catttgtgaa 4380
gtggaggttt ttgatctgct ttttgtcact aatgagagta attcacgaaa gacctacata 4440
gtacattgcc aagattgtgc acgaaaaaca agcggaaact tggaaaactt tgtggtgcta 4500
gaacagtaca aaatggagga cctgatgcaa gtctatgacc aatttacatt agctcctcca 4560
ttaccatccg cctcatcttg atattgttcc atggacatta aatgagacct tttctgctat 4620
tcaggaaata acccagttct gcaccactgg tttttgtagc tatctcgtaa ggctgctggc 4680
tgaaaactgt gtctatgcaa ccttccaagt gcggagtgtc aaccaactgg acgggagaga 4740
gtactgctcc tactccagga ctctcacaaa gctgatgagc tgtacttcag aaaaaaataa 4800
taatttccat gttttgtata tatctgacaa aactggcaac atcttacaga ctactgactt 4860
gaagacaacc tcttttatat ttctctattt ctgggctgat gaatttgttt tcatctgtct 4920
tttccccctt cagaattttc cttggaaaaa aaatactagc ctagctggtc atttctttgt 4980
aaggtagtta gcaattttaa gtctttcttt ggtcaacttt tttttaatgt gaaaagttag 5040
gtaagacact tttttactgc ttttatgttt ttctgtcttg ttttgagacc atgatggtta 5100
cacttttggt tcctaaataa aatttaaaaa attaacagcc aagtcacaaa ggtaatggat 5160
tgcacataga ctaaggaata aacttcagat ttgtgatttt tgtttctaat cttgatgtaa 5220
atttacacta tttataaata catatttatt gcttgaaaat atttgtgaat ggaatgctgt 5280
tatttttcc agatttacct gccattgaaa ttttaaggag ttctgtaatt tcaaacacta 5340
ctcctattac attttctatg tgtaaataaa actgcttagc attgtacaga aacttttatt 5400
aaaattgttt aatgtttaaa gagttttcta ttgtttgagt tttaaaaaag actttatgta 5460
cagtgcccag tttttgttca tttttgaaat ctgattatat atattttata tatacttatg 5520
tatgtatata taatatatat agaaatctgg atatatatgt ataaatcttt agaacttaaa 5580
tttttctcgt tttaagtttc acatctatgg tagatttttg aggtgtctac tgtaaagtat 5640
tgcttacaaa aagtatgatt atttttaaag aaatatatat ggtatgtatc ctcaagacct 5700
aaaatgtcag actggtttat tgttaagttg caattactgc aatgacagac caataaacaa 5760
ttgctgccaa aaaaaaaa 5778

<210> 8
<211> 4206
<212> ADN
<213> *Homo sapiens*

<400> 8

atgaaatcct gcggagtgtc gctcgctacc gccgccgctg ccgccgccgc tttcggtgat 60
gaggaaaaga aaatggcggc gggaaaagcg agcggcgaga gcgaggaggc gtcccccagc 120

ctgacagccg aggagaggga ggcgctcggc ggactggaca gccgcctctt tgggttcgtg 180
agatttcatg aagatggcgc caggacgaag gccctactgg gcaaggctgt tcgctgctat 240
gaatctctaa tcttaaaagc tgaaggaaaa gtggagtctg atttcttttg tcaattaggt 300
cacttcaacc tcttattgga agattatcca aaagcattat ctgcatacca gaggtactac 360
agtttacagt ctgactactg gaagaatgct gccttttttat atggtcttgg tttggtctac 420
ttccattata atgcatttca gtgggcaatt aaagcatttc aggaggtgct ttatgttgat 480
cccagctttt gtcgagccaa ggaaattcat ttacgacttg ggcttatgtt caaagtgaac 540
acagactatg agtctagttt aaagcatttt cagttagctt tggttgactg taatccctgc 600
actttgtcca atgctgaaat tcaatttcac attgcccact tatatgaaac ccagaggaaa 660
tatcattctg caaaagaagc ttatgaacaa cttttgcaga cagagaatct ttctgcacaa 720
gtaaaagcaa ctgtcttaca acagttaggt tggatgcatc acactgtaga tctcctggga 780
gataaagcca ccaaggaaag ctatgctatt cagtatctcc aaaagtcctt ggaagcagat 840
cctaattctg gccagtcctg gtatttcctc ggaaggtgct attcaagtat tgggaaagtt 900
caggatgcct ttatatctta caggcagtct attgataaat cagaagcaag tgcagataca 960
tggtgttcaa taggtgtgct atatcagcag caaaatcagc ccatggatgc tttacaggcc 1020
tatatttgtg ctgtacaatt ggaccatggc catgctgcag cctggatgga cctaggcact 1080
ctctatgaat cctgcaacca gcctcaggat gccattaaat gctacttaaa tgcaactaga 1140
agcaaaagtt gtagtaatac ctctgcactt gcagcacgaa ttaagtattt acaggctcag 1200
ttgtgtaacc ttccacaagg tagtctacag aataaaacta aattacttcc tagtattgag 1260
gaggcgtgga gcctaccaat tcccgcagag cttacctcca ggcagggtgc catgaacaca 1320
gcacagcaga atacttctga caattggagt ggtggacatg ctgtgtcaca tcctccagta 1380
cagcaacaag ctcattcatg gtgtttgaca ccacagaaat tacagcattt ggaacagctc 1440
cgcgcaaata gaaataattt aaatccagca cagaaactga tgctggaaca gctggaaagt 1500
cagtttgtct taatgcaaca acaccaaatg agaccaacag gagttgcaca ggtacgatct 1560
actggaattc ctaatgggcc aacagctgac tcatcactgc ctacaaactc agtctctggc 1620
cagcagccac agcttgctct gaccagagtg cctagcgtct ctcagcctgg agtccgtcct 1680
gcctgccctg ggcagccttt ggccaatgga ccctttttctg caggccatgt tccctgtagc 1740
acatcaagaa cgctgggaag tacagacact attttgatag gcaataatca tataacagga 1800
agtggaagta atggaaacgt gccttacctg cagcgaaacg cactcactct acctcataac 1860
cgcacaaacc tgaccagcag cgcagaggag ccgtggaaaa accaactatc taactccact 1920
cagggcttc acaaaggtca gagttcacat tcggcaggtc ctaatggtga acgacctctc 1980
tcttccactg ggccttccca gcatctccag gcagctggct ctggtattca gaatcagaac 2040
ggacatccca ccctgcctag caattcagta acacaggggg ctgctctcaa tcacctctcc 2100
tctcacactg ctacctcagg tggacaacaa ggcattacct taaccaaaga gagcaagcct 2160
tcaggaaaca tattgacggt gcctgaaaca agcaggcaca ctggagagac acctaacagc 2220
actgccagtg tcgagggact tcctaatcat gtccatcaga tgacggcaga tgctgtttgc 2280
agtcctagcc atggagattc taagtcacca ggtttactaa gttcagacaa tcctcagctc 2340
tctgccttgt tgatgggaaa agccaataac aatgtgggta ctggaacctg tgacaaagtc 2400
aataacatcc acccagctgt tcatacaaag actgataact ctgttgcctc ttcaccatct 2460
tcagccattt caacagcaac accttctcca aaatccactg agcagacaac cacaaacagt 2520
gttaccagcc ttaacagccc tcacagtggg ctacacacaa ttaatggaga agggatggaa 2580
gaatctcaga gccccatgaa aacagatctg cttctggtta accacaaacc tagtccacag 2640
atcataccat caatgtctgt gtccatatac cccagctcag cagaagttct gaaggcatgc 2700
aggaatctag gtaaaaatgg cttatctaac agtagcattt tgttggataa atgtccacct 2760
ccaagaccac catcttcacc ataccctccc ttgccaaagg acaagttgaa tccacctaca 2820
cctagtattt acttggaaaa taaacgtgat gctttctttc ctccattaca tcaattttgt 2880
acaaatccga acaaccctgt tacagtaata cgtggccttg ctggagctct taagttagac 2940
ctgggacttt tctctactaa aactttggtg gaagctaaca atgaacatat ggtagaagtg 3000
aggacacagt tgttgcagcc agcagatgaa aactgggatc ccactggaac aaagaaaatc 3060
tggcattgtg aaagtaatag atctcatact acaattgcta aatatgcaca gtaccaggcc 3120
tcctcattcc aggaatcatt gagagaagaa aatgaaaaaa gaagtcatca taaagaccac 3180
tcagatagtg aatctacatc gtcagataat tctgggagga ggaggaaagg acccttaaa 3240
accataaagt ttgggaccaa tattgaccta tctgatgaca aaaagtggaa gttgcagcta 3300
catgagctga ctaaacttcc tgcttttgtg cgtgtcgtat cagcaggaaa tcttctaagc 3360
catgttggtc ataccatatt gggcatgaac acagttcaac tatacatgaa agttccaggg 3420
agcagaacac caggtcatca ggaaaataac aacttctgtt cagttaacat aaatattggc 3480
ccaggtgact gtgaatggtt tgttgttcct gaaggttact ggggtgttct gaatgacttc 3540
tgtgaaaaaa ataatttgaa tttcctaatg ggttcttggt ggcccaatct tgaagatctt 3600
tatgaagcaa atgttccagt gtataggttt attcagcgac ctggagattt ggtctggata 3660
aatgcaggca ctgttcattg ggttcaggct attggctggt gcaacaacat tgcttggaat 3720
gttggtccac ttacagcctg ccagtataaa ttggcagtgg aacggtacga atggaacaaa 3780
ttgcaaagtg tgaagtcaat agtacccatg gttcatcttt cctggaatat ggcacgaaat 3840
atcaaggtct cagatccaaa gctttttgaa atgattaagt attgtcttct aagaactctg 3900

```
aagcaatgtc agacattgag ggaagctctc attgctgcag gaaaagagat tatatggcat 3960
gggcggacaa aagaagaacc agctcattac tgtagcattt gtgaagtgga ggttttgat 4020
ctgctttttg tcactaatga gagtaattca cgaaagacct acatagtaca ttgccaagat 4080
tgtgcacgaa aaacaagcgg aaacttggaa aactttgtgg tgctagaaca gtacaaaatg 4140
gaggacctga tgcaagtcta tgaccaattt acattagctc ctccattacc atccgcctca 4200
tcttga                                                            4206
```

<210> 9
<211> 4206
<212> ADN
<213> *Homo sapiens*

<400> 9

atgaaatcct gcggagtgtc gctcgctacc gccgccgctg ccgccgccgc tttcggtgat 60
gaggaaaaga aaatggcggc gggaaaagcg agcggcgaga gcgaggaggc gtcccccagc 120
ctgacagccg aggagaggga ggcgctcggc ggactggaca gccgcctctt tgggttcgtg 180
agatttcatg aagatggcgc caggacgaag gccctactgg gcaaggctgt tcgctgctat 240
gaatctctaa tcttaaaagc tgaaggaaaa gtggagtctg atttcttttg tcaattaggt 300
cacttcaacc tcttattgga agattatcca aaagcattat ctgcatacca gaggtactac 360
agtttacagt ctgactactg gaagaatgct gccttttttat atggtcttgg tttggtctac 420
ttccattata atgcatttca gtgggcaatt aaagcatttc aggaggtgct ttatgttgat 480
cccagctttt gtcgagccaa ggaaattcat ttacgagttg ggcttatgtt caaagtgaac 540
acagactatg agtctagttt aaagcatttt cagttagctt tggttgactg taatccctgc 600
actttgtcca atgctgaaat tcaatttcac attgcccact tatatgaaac ccagaggaaa 660
tatcattctg caaaagaagc ttatgaacaa cttttgcaga cagagaatct ttctgcacaa 720
gtaaaagcaa ctgtcttaca acagttaggt tggatgcatc acactgtaga tctcctggga 780
gataaagcca ccaaggaaag ctatgctatt cagtatctcc aaaagtcctt ggaagcagat 840
cctaattctg gccagtcctg gtatttcctc ggaaggtgct attcaagtat tgggaaagtt 900
caggatgcct ttatatctta caggcagtct attgataaat cagaagcaag tgcagataca 960
tggtgttcaa taggtgtgct atatcagcag caaaatcagc ccatggatgc tttacaggcc 1020
tatatttgtg ctgtacaatt ggaccatggc catgctgcag cctggatgga cctaggcact 1080
ctctatgaat cctgcaacca gcctcaggat gccattaaat gctacttaaa tgcaactaga 1140
agcaaaagtt gtagtaatac ctctgcactt gcagcacgaa ttaagtattt acaggctcag 1200
ttgtgtaacc ttccacaagg tagtctacag aataaaacta aattacttcc tagtattgag 1260
gaggcgtgga gcctaccaat tcccgcagag cttacctcca ggcagggtgc catgaacaca 1320
gcacagcaga atacttctga caattggagt ggtggacatg ctgtgtcaca tcctccagta 1380
cagcaacaag ctcattcatg gtgtttgaca ccacagaaat tacagcattt ggaacagctc 1440
cgcgcaaata gaaataattt aaatccagca cagaaactga tgctggaaca gctggaaagt 1500
cagtttgtct taatgcaaca acaccaaatg agaccaacag gagttgcaca ggtacgatct 1560
actggaattc ctaatgggcc aacagctgac tcatcactgc ctacaaactc agtctctggc 1620
cagcagccac agcttgctct gaccagagtg cctagcgtct ctcagcctgg agtccgtcct 1680
gcctgccctg ggcagccttt ggccaatgga ccctttctg caggccatgt tccctgtagc 1740
acatcaagaa cgcggggaag tacagacact attttgatag gcaataatca tataacagga 1800
aatggaagta atggaaacgt gccttacctg cagcgaaacg cactcactct acctcataac 1860
cgcacaaacc tgaccagcag cgcaaaggag ccgtggaaaa accaactatc taactccact 1920
caggggcttc acaaaggtca gagttcacat tcggcaggtc ctaatggtga acgacctctc 1980
tcttccactg ggccttccca gcatctccag gcagctggct ctggtattca gaatcagaac 2040
ggacatccca ccctgcctag caattcagta acacaggggg ctgctctcaa tcacctctcc 2100
tctcacactg ctacctcagg tggacaacaa ggcattacct taaccaaaga gagcaagcct 2160
tcaggaaaca tattgacggt gcctgaaaca agcaggcaca ctggagagac acctaacagc 2220
actgccagtg tcgagggact tcctaatcat gtccatcaga tgacggcaga tgctgtttgc 2280
agtcctagcc atggagattc taagtcacca ggtttactaa gttcagacaa tcctcagctc 2340
tctgccttgt tgatgggaaa agccaataac aatgtgggta ctggaacctg tgacaaagtc 2400
aataacatcc acccagctgt tcatacaaag actgataact ctgttgcctc ttcaccatct 2460
tcagccattt caacagcaac accttctcca aaatccactg agcagacaac cacaaacagt 2520
gttaccagcc ttaacagccc tcacagtggg ctacacacaa ttaatggaga agggatggaa 2580
gaatctcaga gccccatgaa aacagatctg cttctggtta accacaaacc tagtccacag 2640
atcataccat caatgtctgt gtccatatac cccagctcag cagaagttct gaaggcatgc 2700
aggaatctag gtaaaaatgg cttatctaac agtagcattt tgttggataa atgtccacct 2760
ccaagaccac catcttcacc ataccctccc ttgccaaagg acaagttgaa tccacctaca 2820
cctagtattt acttggaaaa taaacgtgat gctttctttc ctccattaca tcaattttgt 2880
acaaatccga acaaccctgt tacagtaata cgtggccttg ctggagctct taagttagac 2940
ctgggacttt tctctactaa aactttggtg gaagctaaca atgaacatat ggtagaagtg 3000

```
aggacacagt tgttgcagcc agcagatgaa aactgggatc ccactggaac aaagaaaatc 3060
tggcattgtg aaagtaatag atctcatact acaattgcta aatatgcaca gtaccaggcc 3120
tcctcattcc aggaatcatt gagagaagaa aatgaaaaaa gaagtcatca taaagaccac 3180
tcagatagtg aatctacatc gtcagataat tctgggagga ggaggaaagg acccttaaa 3240
accataaagt ttgggaccaa tattgaccta tctgatgaca aaaagtggaa gttgcagcta 3300
catgagctga ctaaacttcc tgcttttgtg cgtgtcgtat cagcaggaaa tcttctaagc 3360
catgttggtc ataccatatt gggcatgaac acagttcaac tatacatgaa agttccaggg 3420
agcagaacac caggtcatca ggaaaataac aacttctgtt cagttaacat aaatattggc 3480
ccaggtgact gtgaatggtt tgttgttcct gaaggttact ggggtgtttt gaatgacttc 3540
tgtgaaaaaa ataatttgaa tttcctaatg ggttcttggt ggcccaatct tgaagatctt 3600
tatgaagcaa atgttccagt gtataggttt attcagcgac ctggagattt ggtctggata 3660
aatgcaggca ctgttcattg ggttcaggct attggctggt gcaacaacat tgcttggaat 3720
gttggtccac ttacagcctg ccagtataaa ttggcagtgg aacggtacga atggaacaaa 3780
ttgcaaagtg tgaagtcaat agtacccatg gttcatcttt cctggaatat ggcacgaaat 3840
atcaaggtct cagatccaaa gctttttgaa atgattaagt attgtcttct aagaactctg 3900
aagcaatgtc agacattgag ggaagctctc attgctgcag gaaaagagat tatatggcat 3960
gggcggacaa aagaagaacc agctcattac tgtagcattt gtgaagtgga ggtttttgat 4020
ctgctttttg tcactaatga gagtaattca cgaaagacct acatagtaca ttgccaagat 4080
tgtgcacgaa aaacaagcgg aaacttggaa aactttgtgg tgctagaaca gtacaaaatg 4140
gaggacctga tgcaagtcta tgaccaattt acattagctc ctccattacc atccgcctca 4200
tcttga                                                            4206
```

<210> 10
<211> 1401
<212> PRT
<213> *Homo sapiens*

<400> 10

```
Met Lys Ser Cys Gly Val Ser Leu Ala Thr Ala Ala Ala Ala Ala Ala
 1               5                  10                  15
Ala Phe Gly Asp Glu Glu Lys Lys Met Ala Ala Gly Lys Ala Ser Gly
            20                  25                  30
Glu Ser Glu Glu Ala Ser Pro Ser Leu Thr Ala Glu Glu Arg Glu Ala
        35                  40                  45
Leu Gly Gly Leu Asp Ser Arg Leu Phe Gly Phe Val Arg Phe His Glu
    50                  55                  60
Asp Gly Ala Arg Thr Lys Ala Leu Leu Gly Lys Ala Val Arg Cys Tyr
65                  70                  75                  80
Glu Ser Leu Ile Leu Lys Ala Glu Gly Lys Val Glu Ser Asp Phe Phe
                85                  90                  95
Cys Gln Leu Gly His Phe Asn Leu Leu Leu Glu Asp Tyr Pro Lys Ala
            100                 105                 110
Leu Ser Ala Tyr Gln Arg Tyr Tyr Ser Leu Gln Ser Asp Tyr Trp Lys
        115                 120                 125
Asn Ala Ala Phe Leu Tyr Gly Leu Gly Leu Val Tyr Phe His Tyr Asn
    130                 135                 140
Ala Phe Gln Trp Ala Ile Lys Ala Phe Gln Glu Val Leu Tyr Val Asp
145                 150                 155                 160
Pro Ser Phe Cys Arg Ala Lys Glu Ile His Leu Arg Leu Gly Leu Met
                165                 170                 175
Phe Lys Val Asn Thr Asp Tyr Glu Ser Ser Leu Lys His Phe Gln Leu
            180                 185                 190
Ala Leu Val Asp Cys Asn Pro Cys Thr Leu Ser Asn Ala Glu Ile Gln
        195                 200                 205
Phe His Ile Ala His Leu Tyr Glu Thr Gln Arg Lys Tyr His Ser Ala
    210                 215                 220
Lys Glu Ala Tyr Glu Gln Leu Leu Gln Thr Glu Asn Leu Ser Ala Gln
225                 230                 235                 240
Val Lys Ala Thr Val Leu Gln Gln Leu Gly Trp Met His His Thr Val
                245                 250                 255
Asp Leu Leu Gly Asp Lys Ala Thr Lys Glu Ser Tyr Ala Ile Gln Tyr
            260                 265                 270
Leu Gln Lys Ser Leu Glu Ala Asp Pro Asn Ser Gly Gln Ser Trp Tyr
```

```
        275                 280                 285
Phe Leu Gly Arg Cys Tyr Ser Ser Ile Gly Lys Val Gln Asp Ala Phe
    290                 295                 300
Ile Ser Tyr Arg Gln Ser Ile Asp Lys Ser Glu Ala Ser Ala Asp Thr
305                 310                 315                 320
Trp Cys Ser Ile Gly Val Leu Tyr Gln Gln Gln Asn Gln Pro Met Asp
                325                 330                 335
Ala Leu Gln Ala Tyr Ile Cys Ala Val Gln Leu Asp His Gly His Ala
            340                 345                 350
Ala Ala Trp Met Asp Leu Gly Thr Leu Tyr Glu Ser Cys Asn Gln Pro
        355                 360                 365
Gln Asp Ala Ile Lys Cys Tyr Leu Asn Ala Thr Arg Ser Lys Ser Cys
    370                 375                 380
Ser Asn Thr Ser Ala Leu Ala Ala Arg Ile Lys Tyr Leu Gln Ala Gln
385                 390                 395                 400
Leu Cys Asn Leu Pro Gln Gly Ser Leu Gln Asn Lys Thr Lys Leu Leu
                405                 410                 415
Pro Ser Ile Glu Glu Ala Trp Ser Leu Pro Ile Pro Ala Glu Leu Thr
            420                 425                 430
Ser Arg Gln Gly Ala Met Asn Thr Ala Gln Gln Asn Thr Ser Asp Asn
        435                 440                 445
Trp Ser Gly Gly His Ala Val Ser His Pro Pro Val Gln Gln Gln Ala
    450                 455                 460
His Ser Trp Cys Leu Thr Pro Gln Lys Leu Gln His Leu Glu Gln Leu
465                 470                 475                 480
Arg Ala Asn Arg Asn Asn Leu Asn Pro Ala Gln Lys Leu Met Leu Glu
                485                 490                 495
Gln Leu Glu Ser Gln Phe Val Leu Met Gln Gln His Gln Met Arg Pro
            500                 505                 510
Thr Gly Val Ala Gln Val Arg Ser Thr Gly Ile Pro Asn Gly Pro Thr
        515                 520                 525
Ala Asp Ser Ser Leu Pro Thr Asn Ser Val Ser Gly Gln Gln Pro Gln
    530                 535                 540
Leu Ala Leu Thr Arg Val Pro Ser Val Ser Gln Pro Gly Val Arg Pro
545                 550                 555                 560
Ala Cys Pro Gly Gln Pro Leu Ala Asn Gly Pro Phe Ser Ala Gly His
                565                 570                 575
Val Pro Cys Ser Thr Ser Arg Thr Leu Gly Ser Thr Asp Thr Ile Leu
            580                 585                 590
Ile Gly Asn Asn His Ile Thr Gly Ser Gly Ser Asn Gly Asn Val Pro
        595                 600                 605
Tyr Leu Gln Arg Asn Ala Leu Thr Leu Pro His Asn Arg Thr Asn Leu
    610                 615                 620
Thr Ser Ser Ala Glu Glu Pro Trp Lys Asn Gln Leu Ser Asn Ser Thr
625                 630                 635                 640
Gln Gly Leu His Lys Gly Gln Ser Ser His Ser Ala Gly Pro Asn Gly
                645                 650                 655
Glu Arg Pro Leu Ser Ser Thr Gly Pro Ser Gln His Leu Gln Ala Ala
            660                 665                 670
Gly Ser Gly Ile Gln Asn Gln Asn Gly His Pro Thr Leu Pro Ser Asn
        675                 680                 685
Ser Val Thr Gln Gly Ala Ala Leu Asn His Leu Ser Ser His Thr Ala
    690                 695                 700
Thr Ser Gly Gly Gln Gln Gly Ile Thr Leu Thr Lys Glu Ser Lys Pro
705                 710                 715                 720
Ser Gly Asn Ile Leu Thr Val Pro Glu Thr Ser Arg His Thr Gly Glu
                725                 730                 735
Thr Pro Asn Ser Thr Ala Ser Val Glu Gly Leu Pro Asn His Val His
            740                 745                 750
Gln Met Thr Ala Asp Ala Val Cys Ser Pro Ser His Gly Asp Ser Lys
        755                 760                 765
Ser Pro Gly Leu Leu Ser Ser Asp Asn Pro Gln Leu Ser Ala Leu Leu
    770                 775                 780
```

```
Met Gly Lys Ala Asn Asn Asn Val Gly Thr Gly Thr Cys Asp Lys Val
785                 790                 795                 800
Asn Asn Ile His Pro Ala Val His Thr Lys Thr Asp Asn Ser Val Ala
                805                 810                 815
Ser Ser Pro Ser Ser Ala Ile Ser Thr Ala Thr Pro Ser Pro Lys Ser
            820                 825                 830
Thr Glu Gln Thr Thr Thr Asn Ser Val Thr Ser Leu Asn Ser Pro His
        835                 840                 845
Ser Gly Leu His Thr Ile Asn Gly Glu Gly Met Glu Glu Ser Gln Ser
    850                 855                 860
Pro Met Lys Thr Asp Leu Leu Leu Val Asn His Lys Pro Ser Pro Gln
865                 870                 875                 880
Ile Ile Pro Ser Met Ser Val Ser Ile Tyr Pro Ser Ser Ala Glu Val
                885                 890                 895
Leu Lys Ala Cys Arg Asn Leu Gly Lys Asn Gly Leu Ser Asn Ser Ser
            900                 905                 910
Ile Leu Leu Asp Lys Cys Pro Pro Pro Arg Pro Pro Ser Ser Pro Tyr
        915                 920                 925
Pro Pro Leu Pro Lys Asp Lys Leu Asn Pro Pro Thr Pro Ser Ile Tyr
    930                 935                 940
Leu Glu Asn Lys Arg Asp Ala Phe Phe Pro Pro Leu His Gln Phe Cys
945                 950                 955                 960
Thr Asn Pro Asn Asn Pro Val Thr Val Ile Arg Gly Leu Ala Gly Ala
                965                 970                 975
Leu Lys Leu Asp Leu Gly Leu Phe Ser Thr Lys Thr Leu Val Glu Ala
            980                 985                 990
Asn Asn Glu His Met Val Glu Val Arg Thr Gln Leu Leu Gln Pro Ala
        995                 1000                1005
Asp Glu Asn Trp Asp Pro Thr Gly Thr Lys Lys Ile Trp His Cys Glu
    1010                1015                1020
Ser Asn Arg Ser His Thr Thr Ile Ala Lys Tyr Ala Gln Tyr Gln Ala
1025                1030                1035                1040
Ser Ser Phe Gln Glu Ser Leu Arg Glu Glu Asn Glu Lys Arg Ser His
                1045                1050                1055
His Lys Asp His Ser Asp Ser Glu Ser Thr Ser Ser Asp Asn Ser Gly
            1060                1065                1070
Arg Arg Arg Lys Gly Pro Phe Lys Thr Ile Lys Phe Gly Thr Asn Ile
        1075                1080                1085
Asp Leu Ser Asp Asp Lys Lys Trp Lys Leu Gln Leu His Glu Leu Thr
    1090                1095                1100
Lys Leu Pro Ala Phe Val Arg Val Val Ser Ala Gly Asn Leu Leu Ser
1105                1110                1115                1120
His Val Gly His Thr Ile Leu Gly Met Asn Thr Val Gln Leu Tyr Met
                1125                1130                1135
Lys Val Pro Gly Ser Arg Thr Pro Gly His Gln Glu Asn Asn Asn Phe
            1140                1145                1150
Cys Ser Val Asn Ile Asn Ile Gly Pro Gly Asp Cys Glu Trp Phe Val
        1155                1160                1165
Val Pro Glu Gly Tyr Trp Gly Val Leu Asn Asp Phe Cys Glu Lys Asn
    1170                1175                1180
Asn Leu Asn Phe Leu Met Gly Ser Trp Trp Pro Asn Leu Glu Asp Leu
1185                1190                1195                1200
Tyr Glu Ala Asn Val Pro Val Tyr Arg Phe Ile Gln Arg Pro Gly Asp
                1205                1210                1215
Leu Val Trp Ile Asn Ala Gly Thr Val His Trp Val Gln Ala Ile Gly
            1220                1225                1230
Trp Cys Asn Asn Ile Ala Trp Asn Val Gly Pro Leu Thr Ala Cys Gln
        1235                1240                1245
Tyr Lys Leu Ala Val Glu Arg Tyr Glu Trp Asn Lys Leu Gln Ser Val
    1250                1255                1260
Lys Ser Ile Val Pro Met Val His Leu Ser Trp Asn Met Ala Arg Asn
1265                1270                1275                1280
Ile Lys Val Ser Asp Pro Lys Leu Phe Glu Met Ile Lys Tyr Cys Leu
```

```
                1285                1290                1295
Leu Arg Thr Leu Lys Gln Cys Gln Thr Leu Arg Glu Ala Leu Ile Ala
            1300                1305                1310
Ala Gly Lys Glu Ile Ile Trp His Gly Arg Thr Lys Glu Glu Pro Ala
        1315                1320                1325
His Tyr Cys Ser Ile Cys Glu Val Glu Val Phe Asp Leu Leu Phe Val
    1330                1335                1340
Thr Asn Glu Ser Asn Ser Arg Lys Thr Tyr Ile Val His Cys Gln Asp
1345                1350                1355                1360
Cys Ala Arg Lys Thr Ser Gly Asn Leu Glu Asn Phe Val Val Leu Glu
                1365                1370                1375
Gln Tyr Lys Met Glu Asp Leu Met Gln Val Tyr Asp Gln Phe Thr Leu
            1380                1385                1390
Ala Pro Pro Leu Pro Ser Ala Ser Ser
        1395                1400
```

<210> 11
<211> 1401
<212> PRT
<213> *Homo sapiens*

<400> 11

```
Met Lys Ser Cys Gly Val Ser Leu Ala Thr Ala Ala Ala Ala Ala Ala
 1               5                  10                  15
Ala Phe Gly Asp Glu Glu Lys Lys Met Ala Ala Gly Lys Ala Ser Gly
            20                  25                  30
Glu Ser Glu Glu Ala Ser Pro Ser Leu Thr Ala Glu Glu Arg Glu Ala
        35                  40                  45
Leu Gly Gly Leu Asp Ser Arg Leu Phe Gly Phe Val Arg Phe His Glu
    50                  55                  60
Asp Gly Ala Arg Thr Lys Ala Leu Leu Gly Lys Ala Val Arg Cys Tyr
65                  70                  75                  80
Glu Ser Leu Ile Leu Lys Ala Glu Gly Lys Val Glu Ser Asp Phe Phe
                85                  90                  95
Cys Gln Leu Gly His Phe Asn Leu Leu Leu Glu Asp Tyr Pro Lys Ala
            100                 105                 110
Leu Ser Ala Tyr Gln Arg Tyr Tyr Ser Leu Gln Ser Asp Tyr Trp Lys
        115                 120                 125
Asn Ala Ala Phe Leu Tyr Gly Leu Gly Leu Val Tyr Phe His Tyr Asn
    130                 135                 140
Ala Phe Gln Trp Ala Ile Lys Ala Phe Gln Glu Val Leu Tyr Val Asp
145                 150                 155                 160
Pro Ser Phe Cys Arg Ala Lys Glu Ile His Leu Arg Val Gly Leu Met
                165                 170                 175
Phe Lys Val Asn Thr Asp Tyr Glu Ser Ser Leu Lys His Phe Gln Leu
            180                 185                 190
Ala Leu Val Asp Cys Asn Pro Cys Thr Leu Ser Asn Ala Glu Ile Gln
        195                 200                 205
Phe His Ile Ala His Leu Tyr Glu Thr Gln Arg Lys Tyr His Ser Ala
    210                 215                 220
Lys Glu Ala Tyr Glu Gln Leu Leu Gln Thr Glu Asn Leu Ser Ala Gln
225                 230                 235                 240
Val Lys Ala Thr Val Leu Gln Gln Leu Gly Trp Met His His Thr Val
                245                 250                 255
Asp Leu Leu Gly Asp Lys Ala Thr Lys Glu Ser Tyr Ala Ile Gln Tyr
            260                 265                 270
Leu Gln Lys Ser Leu Glu Ala Asp Pro Asn Ser Gly Gln Ser Trp Tyr
        275                 280                 285
Phe Leu Gly Arg Cys Tyr Ser Ser Ile Gly Lys Val Gln Asp Ala Phe
    290                 295                 300
Ile Ser Tyr Arg Gln Ser Ile Asp Lys Ser Glu Ala Ser Ala Asp Thr
305                 310                 315                 320
```

```
Trp Cys Ser Ile Gly Val Leu Tyr Gln Gln Gln Asn Gln Pro Met Asp
                325                 330                 335
Ala Leu Gln Ala Tyr Ile Cys Ala Val Gln Leu Asp His Gly His Ala
            340                 345                 350
Ala Ala Trp Met Asp Leu Gly Thr Leu Tyr Glu Ser Cys Asn Gln Pro
        355                 360                 365
Gln Asp Ala Ile Lys Cys Tyr Leu Asn Ala Thr Arg Ser Lys Ser Cys
    370                 375                 380
Ser Asn Thr Ser Ala Leu Ala Ala Arg Ile Lys Tyr Leu Gln Ala Gln
385                 390                 395                 400
Leu Cys Asn Leu Pro Gln Gly Ser Leu Gln Asn Lys Thr Lys Leu Leu
                405                 410                 415
Pro Ser Ile Glu Glu Ala Trp Ser Leu Pro Ile Pro Ala Glu Leu Thr
            420                 425                 430
Ser Arg Gln Gly Ala Met Asn Thr Ala Gln Gln Asn Thr Ser Asp Asn
        435                 440                 445
Trp Ser Gly Gly His Ala Val Ser His Pro Pro Val Gln Gln Gln Ala
    450                 455                 460
His Ser Trp Cys Leu Thr Pro Gln Lys Leu Gln His Leu Glu Gln Leu
465                 470                 475                 480
Arg Ala Asn Arg Asn Asn Leu Asn Pro Ala Gln Lys Leu Met Leu Glu
                485                 490                 495
Gln Leu Glu Ser Gln Phe Val Leu Met Gln Gln His Gln Met Arg Pro
            500                 505                 510
Thr Gly Val Ala Gln Val Arg Ser Thr Gly Ile Pro Asn Gly Pro Thr
        515                 520                 525
Ala Asp Ser Ser Leu Pro Thr Asn Ser Val Ser Gly Gln Gln Pro Gln
    530                 535                 540
Leu Ala Leu Thr Arg Val Pro Ser Val Ser Gln Pro Gly Val Arg Pro
545                 550                 555                 560
Ala Cys Pro Gly Gln Pro Leu Ala Asn Gly Pro Phe Ser Ala Gly His
                565                 570                 575
Val Pro Cys Ser Thr Ser Arg Thr Arg Gly Ser Thr Asp Thr Ile Leu
            580                 585                 590
Ile Gly Asn Asn His Ile Thr Gly Asn Gly Ser Asn Gly Asn Val Pro
        595                 600                 605
Tyr Leu Gln Arg Asn Ala Leu Thr Leu Pro His Asn Arg Thr Asn Leu
    610                 615                 620
Thr Ser Ser Ala Lys Glu Pro Trp Lys Asn Gln Leu Ser Asn Ser Thr
625                 630                 635                 640
Gln Gly Leu His Lys Gly Gln Ser Ser His Ser Ala Gly Pro Asn Gly
                645                 650                 655
Glu Arg Pro Leu Ser Ser Thr Gly Pro Ser Gln His Leu Gln Ala Ala
            660                 665                 670
Gly Ser Gly Ile Gln Asn Gln Asn Gly His Pro Thr Leu Pro Ser Asn
        675                 680                 685
Ser Val Thr Gln Gly Ala Ala Leu Asn His Leu Ser Ser His Thr Ala
    690                 695                 700
Thr Ser Gly Gly Gln Gln Gly Ile Thr Leu Thr Lys Glu Ser Lys Pro
705                 710                 715                 720
Ser Gly Asn Ile Leu Thr Val Pro Glu Thr Ser Arg His Thr Gly Glu
                725                 730                 735
Thr Pro Asn Ser Thr Ala Ser Val Glu Gly Leu Pro Asn His Val His
            740                 745                 750
Gln Met Thr Ala Asp Ala Val Cys Ser Pro Ser His Gly Asp Ser Lys
        755                 760                 765
Ser Pro Gly Leu Leu Ser Ser Asp Asn Pro Gln Leu Ser Ala Leu Leu
    770                 775                 780
Met Gly Lys Ala Asn Asn Asn Val Gly Thr Gly Thr Cys Asp Lys Val
785                 790                 795                 800
Asn Asn Ile His Pro Ala Val His Thr Lys Thr Asp Asn Ser Val Ala
                805                 810                 815
Ser Ser Pro Ser Ser Ala Ile Ser Thr Ala Thr Pro Ser Pro Lys Ser
```

```
            820                 825                 830
Thr Glu Gln Thr Thr Thr Asn Ser Val Thr Ser Leu Asn Ser Pro His
        835                 840                 845
Ser Gly Leu His Thr Ile Asn Gly Glu Gly Met Glu Glu Ser Gln Ser
    850                 855                 860
Pro Met Lys Thr Asp Leu Leu Leu Val Asn His Lys Pro Ser Pro Gln
865                 870                 875                 880
Ile Ile Pro Ser Met Ser Val Ser Ile Tyr Pro Ser Ser Ala Glu Val
                885                 890                 895
Leu Lys Ala Cys Arg Asn Leu Gly Lys Asn Gly Leu Ser Asn Ser Ser
            900                 905                 910
Ile Leu Leu Asp Lys Cys Pro Pro Pro Arg Pro Pro Ser Ser Pro Tyr
        915                 920                 925
Pro Pro Leu Pro Lys Asp Lys Leu Asn Pro Pro Thr Pro Ser Ile Tyr
    930                 935                 940
Leu Glu Asn Lys Arg Asp Ala Phe Phe Pro Pro Leu His Gln Phe Cys
945                 950                 955                 960
Thr Asn Pro Asn Asn Pro Val Thr Val Ile Arg Gly Leu Ala Gly Ala
                965                 970                 975
Leu Lys Leu Asp Leu Gly Leu Phe Ser Thr Lys Thr Leu Val Glu Ala
            980                 985                 990
Asn Asn Glu His Met Val Glu Val Arg Thr Gln Leu Leu Gln Pro Ala
        995                 1000                1005
Asp Glu Asn Trp Asp Pro Thr Gly Thr Lys Lys Ile Trp His Cys Glu
    1010                1015                1020
Ser Asn Arg Ser His Thr Thr Ile Ala Lys Tyr Ala Gln Tyr Gln Ala
1025                1030                1035                1040
Ser Ser Phe Gln Glu Ser Leu Arg Glu Glu Asn Glu Lys Arg Ser His
                1045                1050                1055
His Lys Asp His Ser Asp Ser Glu Ser Thr Ser Ser Asp Asn Ser Gly
            1060                1065                1070
Arg Arg Arg Lys Gly Pro Phe Lys Thr Ile Lys Phe Gly Thr Asn Ile
        1075                1080                1085
Asp Leu Ser Asp Asp Lys Lys Trp Lys Leu Gln Leu His Glu Leu Thr
    1090                1095                1100
Lys Leu Pro Ala Phe Val Arg Val Val Ser Ala Gly Asn Leu Leu Ser
1105                1110                1115                1120
His Val Gly His Thr Ile Leu Gly Met Asn Thr Val Gln Leu Tyr Met
                1125                1130                1135
Lys Val Pro Gly Ser Arg Thr Pro Gly His Gln Glu Asn Asn Asn Phe
            1140                1145                1150
Cys Ser Val Asn Ile Asn Ile Gly Pro Gly Asp Cys Glu Trp Phe Val
        1155                1160                1165
Val Pro Glu Gly Tyr Trp Gly Val Leu Asn Asp Phe Cys Glu Lys Asn
    1170                1175                1180
Asn Leu Asn Phe Leu Met Gly Ser Trp Trp Pro Asn Leu Glu Asp Leu
1185                1190                1195                1200
Tyr Glu Ala Asn Val Pro Val Tyr Arg Phe Ile Gln Arg Pro Gly Asp
                1205                1210                1215
Leu Val Trp Ile Asn Ala Gly Thr Val His Trp Val Gln Ala Ile Gly
            1220                1225                1230
Trp Cys Asn Asn Ile Ala Trp Asn Val Gly Pro Leu Thr Ala Cys Gln
        1235                1240                1245
Tyr Lys Leu Ala Val Glu Arg Tyr Glu Trp Asn Lys Leu Gln Ser Val
    1250                1255                1260
Lys Ser Ile Val Pro Met Val His Leu Ser Trp Asn Met Ala Arg Asn
1265                1270                1275                1280
Ile Lys Val Ser Asp Pro Lys Leu Phe Glu Met Ile Lys Tyr Cys Leu
                1285                1290                1295
Leu Arg Thr Leu Lys Gln Cys Gln Thr Leu Arg Glu Ala Leu Ile Ala
            1300                1305                1310
Ala Gly Lys Glu Ile Ile Trp His Gly Arg Thr Lys Glu Glu Pro Ala
        1315                1320                1325
```

```
His Tyr Cys Ser Ile Cys Glu Val Glu Val Phe Asp Leu Leu Phe Val
    1330                1335                1340
Thr Asn Glu Ser Asn Ser Arg Lys Thr Tyr Ile Val His Cys Gln Asp
1345                1350                1355                1360
Cys Ala Arg Lys Thr Ser Gly Asn Leu Glu Asn Phe Val Val Leu Glu
                1365                1370                1375
Gln Tyr Lys Met Glu Asp Leu Met Gln Val Tyr Asp Gln Phe Thr Leu
            1380                1385                1390
Ala Pro Pro Leu Pro Ser Ala Ser Ser
        1395                1400
```

<210> 12
<211> 3896
<212> ADN
<213> *Homo sapiens*

<400> 12

ccttccgcag ctgccgcttc agtccgaagg aggaagggaa ccaacccact ttctcggcgc 60
cgcggctctt ttctaaaagt aatgtgaaaa cctttgcatc ttctgatagt ctagccaagg 120
tccaagaagt agcaagctgg cttttggaaa tgaatcagga actgctctct gtgggcagca 180
aaagacgacg aactggaggc tctctgagag gtaacccttc ctcaagccag gtagatgaag 240
aacagatgaa tcgtgtggta gaggaggaac agcaacagca actcagacaa caagaggagg 300
agcacactgc aaggaatggt gaagttgttg gagtagaacc tagacctgga ggccaaaatg 360
attcccagca aggacagttg gaagaaaaca ataatagatt tatttcggta gatgaggact 420
cctcaggaaa ccaagaagaa caagaggaag atgaagaaca tgctggtgaa caagatgagg 480
aggatgagga ggaggaggag atggaccagg agagtgacga ttttgatcag tctgatgata 540
gtagcagaga agatgaacat acacatacta acagtgtcac gaactccagt agtattgtgg 600
acctgcccgt tcaccaactc tcctccccat tctatacaaa aacaacaaaa atgaaaagaa 660
agttggacca tggttctgag gtccgctctt tttctttggg aaagaaacca tgcaaagtct 720
cagaatatac aagtaccact gggcttgtac catgttcagc aacaccaaca acttttgggg 780
acctcagagc agccaatggc caagggcaac aacgacgccg aattacatct gtccagccac 840
ctacaggcct ccaggaatgg ctaaaaatgt ttcagagctg gagtggacca gagaaattgc 900
ttgctttaga tgaactcatt gatagttgtg aaccaacaca agtaaaacat atgatgcaag 960
tgatagaacc ccagtttcaa cgagacttca tttcattgct ccctaaagag ttggcactct 1020
atgtgctttc attcctggaa cccaaagacc tgctacaagc agctcagaca tgtcgctact 1080
ggagaatttt ggctgaagac aaccttctct ggagagagaa atgcaaagaa gaggggattg 1140
atgaaccatt gcacatcaag agaagaaaag taataaaacc aggtttcata cacagtccat 1200
ggaaaagtgc atacatcaga cagcacagaa ttgatactaa ctggaggcga ggagaactca 1260
aatctcctaa ggtgctgaaa ggacatgatg atcatgtgat cacatgctta cagttttgtg 1320
gtaaccgaat agttagtggt tctgatgaca acactttaaa agtttggtca gcagtcacag 1380
gcaaatgtct gagaacatta gtgggacata caggtggagt atggtcatca caaatgagag 1440
acaacatcat cattagtgga tctacagatc ggacactcaa agtgtggaat gcagagactg 1500
gagaatgtat acacacctta tatgggcata cttccactgt gcgttgtatg catcttcatg 1560
aaaaaagagt tgttagcggt tctcgagatg ccactcttag ggtttgggat attgagacag 1620
gccagtgttt acatgttttg atgggtcatg ttgcagcagt ccgctgtgtt caatatgatg 1680
gcaggagggt tgttagtgga gcatatgatt ttatggtaaa ggtgtgggat ccagagactg 1740
aaacctgtct acacacgttg caggggcata ctaatagagt ctattcatta cagtttgatg 1800
gtatccatgt ggtgagtgga tctcttgata catcaatccg tgtttgggat gtggagacag 1860
ggaattgcat tcacacgtta acagggcacc agtcgttaac aagtggaatg gaactcaaag 1920
acaatattct tgtctctggg aatgcagatt ctacagttaa aatctgggat atcaaaacag 1980
gacagtgttt acaaacattg caaggtccca acaagcatca gagtgctgtg acctgtttac 2040
agttcaacaa gaactttgta attaccagct cagatgatgg aactgtaaaa ctatgggact 2100
tgaaaacggg tgaatttatt cgaaacctag tcacattgga gagtgggggg agtgggggag 2160
ttgtgtggcg gatcagagcc tcaaacacaa agctggtgtg tgcagttggg agtcggaatg 2220
ggactgaaga aaccaagctg ctggtgctgg actttgatgt ggacatgaag tgaagagcag 2280
aaaagatgaa tttgtccaat tgtgtagacg atatactccc tgcccttccc cctgcaaaaa 2340
gaaaaaaaga aaagaaaaag aaaaaaatcc cttgttctca gtggtgcagg atgttggctt 2400
ggggcaacag attgaaaaga cctacagact aagaaggaaa agaagaagag atgacaaacc 2460
ataactgaca agagaggcgt ctgctgtctc atcacataaa aggcttcact tttgactgag 2520
ggcagctttg caaaatgaga ctttctaaat caaaccaggt gcaattattt ctttattttc 2580
ttctccagtg gtcattgggc agtgttaatg ctgaaacatc attacagatt ctgctagcct 2640
gttcttttac cactgacagc tagacaccta gaaaggaact gcaataatat caaaacaagt 2700

actggttgac tttctaatta gagagcatct gcaacaaaaa gtcatttttc tggagtggaa 2760
aagcttaaaa aaattactgt gaattgtttt tgtacagtta tcatgaaaag ctttttttt 2820
tttttttttg ccaaccattg ccaatgtcaa tcaatcacag tattagcctc tgttaatcta 2880
tttactgttg cttccatata cattcttcaa tgcatatgtt gctcaaaggt ggcaagttgt 2940
cctgggttct gtgagtcctg agatggattt aattcttgat gctggtgcta gaagtaggtc 3000
ttcaaatatg ggattgttgt cccaaccctg tactgtactc ccagtggcca aacttattta 3060
tgctgctaaa tgaaagaaag aaaaaagcaa attatttttt tttatttttt ttctgctgtg 3120
acgttttagt cccagactga attccaaatt tgctctagtt tggttatgga aaaaagactt 3180
tttgccactg aaacttgagc catctgtgcc tctaagaggc tgagaatgga agagtttcag 3240
ataataaaga gtgaagtttg cctgcaagta aagaattgag agtgtgtgca aagcttattt 3300
tcttttatct gggcaaaaat taaaacacat tccttggaac agagctatta cttgcctgtt 3360
ctgtggagaa acttttcttt ttgagggctg tggtgaatgg atgaacgtac atcgtaaaac 3420
tgacaaaata ttttaaaaat atataaaaca caaaattaaa ataaagttgc tggtcagtct 3480
tagtgtttta cagtatttgg gaaaacaact gttacagttt tattgctctg agtaactgac 3540
aaagcagaaa ctattcagtt tttgtagtaa aggcgtcaca tgcaaacaaa caaaatgaat 3600
gaaacagtca aatggtttgc ctcattctcc aagagccaca actcaagctg aactgtgaaa 3660
gtggtttaac actgtatcct aggcgatctt ttttcctcct tctgtttatt tttttgtttg 3720
ttttatttat agtctgattt aaaacaatca gattcaagtt ggttaatttt agttatgtaa 3780
caacctgaca tgatggagga aaacaacctt taaagggatt gtgtctatgg tttgattcac 3840
ttagaaattt tattttctta taacttaagt gcaataaaat gtgtttttc atgtta 3896

<210> 13
<211> 2124
<212> ADN
<213> *Homo sapiens*

<400> 13

```
atgaatcagg aactgctctc tgtgggcagc aaaagacgac gaactggagg ctctctgaga 60
ggtaaccctt cctcaagcca ggtagatgaa gaacagatga atcgtgtggt agaggaggaa 120
cagcaacagc aactcagaca acaagaggag gagcacactg caaggaatgg tgaagttgtt 180
ggagtagaac ctagacctgg aggccaaaat gattcccagc aaggacagtt ggaagaaaac 240
aataatagat ttatttcggt agatgaggac tcctcaggaa accaagaaga acaagaggaa 300
gatgaagaac atgctggtga acaagatgag gaggatgagg aggaggagga gatggaccag 360
gagagtgacg attttgatca gtctgatgat agtagcagag aagatgaaca tacacatact 420
aacagtgtca cgaactccag tagtattgtg gacctgcccg ttcaccaact ctcctcccca 480
ttctatacaa aaacaacaaa aatgaaaaga aagttggacc atggttctga ggtccgctct 540
ttttctttgg gaaagaaacc atgcaaagtc tcagaatata caagtaccac tgggcttgta 600
ccatgttcag caacaccaac aacttttggg gacctcagag cagccaatgg ccaagggcaa 660
caacgacgcc gaattacatc tgtccagcca cctacaggcc tccaggaatg gctaaaaatg 720
tttcagagct ggagtggacc agagaaattg cttgctttag atgaactcat tgatagttgt 780
gaaccaacac aagtaaaaca tatgatgcaa gtgatagaac cccagtttca acgagacttc 840
atttcattgc tccctaaaga gttggcactc tatgtgcttt cattcctgga acccaaagac 900
ctgctacaag cagctcagac atgtcgctac tggagaattt tggctgaaga caaccttctc 960
tggagagaga aatgcaaaga agaggggatt gatgaaccat tgcacatcaa gagaagaaaa 1020
gtaataaaac caggtttcat acacagtcca tggaaaagtg catacatcag acagcacaga 1080
attgatacta actggaggcg aggagaactc aaatctccta aggtgctgaa aggacatgat 1140
gatcatgtga tcacatgctt acagttttgt ggtaaccgaa tagttagtgg ttctgatgac 1200
aacactttaa aagtttggtc agcagtcaca ggcaaatgtc tgagaacatt agtgggacat 1260
acaggtggag tatggtcatc acaaatgaga gacaacatca tcattagtgg atctacagat 1320
cggacactca aagtgtggaa tgcagagact ggagaatgta tacacacctt atatgggcat 1380
acttccactg tgcgttgtat gcatcttcat gaaaaaagag ttgttagcgg ttctcgagat 1440
gccactctta gggtttggga tattgagaca ggccagtgtt tacatgtttt gatgggtcat 1500
gttgcagcag tccgctgtgt tcaatatgat ggcaggaggg ttgttagtgg agcatatgat 1560
tttatggtaa aggtgtggga tccagagact gaaacctgtc tacacacgtt gcaggggcat 1620
actaatagag tctattcatt acagtttgat ggtatccatg tggtgagtgg atctcttgat 1680
acatcaatcc gtgtttggga tgtggagaca gggaattgca ttcacacgtt aacagggcac 1740
cagtcgttaa caagtggaat ggaactcaaa gacaatattc ttgtctctgg gaatgcagat 1800
tctacagtta aaatctggga tatcaaaaca ggacagtgtt tacaaacatt gcaaggtccc 1860
aacaagcatc agagtgctgt gacctgttta cagttcaaca agaactttgt aattaccagc 1920
tcagatgatg gaactgtaaa actatgggac ttgaaaacgg gtgaatttat tcgaaaccta 1980
gtcacattgg agagtggggg gagtgggggа gttgtgtggc ggatcagagc ctcaaacaca 2040
aagctggtgt gtgcagttgg gagtcggaat gggactgaag aaaccaagct gctggtgctg 2100
gactttgatg tggacatgaa gtga                                       2124
```

<210> 14
<211> 707
<212> PRT
<213> *Homo sapiens*

<400> 14

```
Met Asn Gln Glu Leu Leu Ser Val Gly Ser Lys Arg Arg Arg Thr Gly
 1               5                  10                  15
Gly Ser Leu Arg Gly Asn Pro Ser Ser Ser Gln Val Asp Glu Glu Gln
            20                  25                  30
Met Asn Arg Val Val Glu Glu Glu Gln Gln Gln Gln Leu Arg Gln Gln
        35                  40                  45
Glu Glu Glu His Thr Ala Arg Asn Gly Glu Val Val Gly Val Glu Pro
    50                  55                  60
Arg Pro Gly Gly Gln Asn Asp Ser Gln Gln Gly Gln Leu Glu Glu Asn
65                  70                  75                  80
Asn Asn Arg Phe Ile Ser Val Asp Glu Asp Ser Ser Gly Asn Gln Glu
                85                  90                  95
Glu Gln Glu Glu Asp Glu Glu His Ala Gly Glu Gln Asp Glu Glu Asp
            100                 105                 110
Glu Glu Glu Glu Glu Met Asp Gln Glu Ser Asp Asp Phe Asp Gln Ser
        115                 120                 125
Asp Asp Ser Ser Arg Glu Asp Glu His Thr His Thr Asn Ser Val Thr
    130                 135                 140
Asn Ser Ser Ser Ile Val Asp Leu Pro Val His Gln Leu Ser Ser Pro
145                 150                 155                 160
Phe Tyr Thr Lys Thr Thr Lys Met Lys Arg Lys Leu Asp His Gly Ser
                165                 170                 175
Glu Val Arg Ser Phe Ser Leu Gly Lys Lys Pro Cys Lys Val Ser Glu
            180                 185                 190
Tyr Thr Ser Thr Thr Gly Leu Val Pro Cys Ser Ala Thr Pro Thr Thr
        195                 200                 205
Phe Gly Asp Leu Arg Ala Ala Asn Gly Gln Gly Gln Gln Arg Arg Arg
    210                 215                 220
Ile Thr Ser Val Gln Pro Pro Thr Gly Leu Gln Glu Trp Leu Lys Met
225                 230                 235                 240
Phe Gln Ser Trp Ser Gly Pro Glu Lys Leu Leu Ala Leu Asp Glu Leu
                245                 250                 255
Ile Asp Ser Cys Glu Pro Thr Gln Val Lys His Met Met Gln Val Ile
            260                 265                 270
Glu Pro Gln Phe Gln Arg Asp Phe Ile Ser Leu Leu Pro Lys Glu Leu
        275                 280                 285
Ala Leu Tyr Val Leu Ser Phe Leu Glu Pro Lys Asp Leu Leu Gln Ala
    290                 295                 300
Ala Gln Thr Cys Arg Tyr Trp Arg Ile Leu Ala Glu Asp Asn Leu Leu
305                 310                 315                 320
Trp Arg Glu Lys Cys Lys Glu Glu Gly Ile Asp Glu Pro Leu His Ile
                325                 330                 335
Lys Arg Arg Lys Val Ile Lys Pro Gly Phe Ile His Ser Pro Trp Lys
            340                 345                 350
Ser Ala Tyr Ile Arg Gln His Arg Ile Asp Thr Asn Trp Arg Arg Gly
        355                 360                 365
Glu Leu Lys Ser Pro Lys Val Leu Lys Gly His Asp Asp His Val Ile
    370                 375                 380
Thr Cys Leu Gln Phe Cys Gly Asn Arg Ile Val Ser Gly Ser Asp Asp
385                 390                 395                 400
Asn Thr Leu Lys Val Trp Ser Ala Val Thr Gly Lys Cys Leu Arg Thr
                405                 410                 415
Leu Val Gly His Thr Gly Gly Val Trp Ser Ser Gln Met Arg Asp Asn
            420                 425                 430
Ile Ile Ile Ser Gly Ser Thr Asp Arg Thr Leu Lys Val Trp Asn Ala
        435                 440                 445
```

```
Glu Thr Gly Glu Cys Ile His Thr Leu Tyr Gly His Thr Ser Thr Val
    450                 455                 460
Arg Cys Met His Leu His Glu Lys Arg Val Val Ser Gly Ser Arg Asp
465                 470                 475                 480
Ala Thr Leu Arg Val Trp Asp Ile Glu Thr Gly Gln Cys Leu His Val
                485                 490                 495
Leu Met Gly His Val Ala Ala Val Arg Cys Val Gln Tyr Asp Gly Arg
            500                 505                 510
Arg Val Val Ser Gly Ala Tyr Asp Phe Met Val Lys Val Trp Asp Pro
        515                 520                 525
Glu Thr Glu Thr Cys Leu His Thr Leu Gln Gly His Thr Asn Arg Val
    530                 535                 540
Tyr Ser Leu Gln Phe Asp Gly Ile His Val Val Ser Gly Ser Leu Asp
545                 550                 555                 560
Thr Ser Ile Arg Val Trp Asp Val Glu Thr Gly Asn Cys Ile His Thr
                565                 570                 575
Leu Thr Gly His Gln Ser Leu Thr Ser Gly Met Glu Leu Lys Asp Asn
            580                 585                 590
Ile Leu Val Ser Gly Asn Ala Asp Ser Thr Val Lys Ile Trp Asp Ile
        595                 600                 605
Lys Thr Gly Gln Cys Leu Gln Thr Leu Gln Gly Pro Asn Lys His Gln
    610                 615                 620
Ser Ala Val Thr Cys Leu Gln Phe Asn Lys Asn Phe Val Ile Thr Ser
625                 630                 635                 640
Ser Asp Asp Gly Thr Val Lys Leu Trp Asp Leu Lys Thr Gly Glu Phe
                645                 650                 655
Ile Arg Asn Leu Val Thr Leu Glu Ser Gly Gly Ser Gly Gly Val Val
            660                 665                 670
Trp Arg Ile Arg Ala Ser Asn Thr Lys Leu Val Cys Ala Val Gly Ser
        675                 680                 685
Arg Asn Gly Thr Glu Glu Thr Lys Leu Leu Val Leu Asp Phe Asp Val
    690                 695                 700
Asp Met Lys
705
```

<210> 15
<211> 2591
<212> ADN
<213> *Homo sapiens*

<400> 15

```
gatgggattg gggttttccc ctcccatgtg ctcaagactg gcgctaaaag ttttgagctt 60
ctcaaaagtc tagagccacc gtccagggag caggtagctg ctgggctccg gggacacttt 120
gcgttcgggc tgggagcgtg ctttccacga cggtgacacg cttccctgga ttggcagcca 180
gactgccttc cgggtcactg ccatggagga gccgcagtca gatcctagcg tcgagccccc 240
tctgagtcag gaaacatttt cagacctatg gaaactactt cctgaaaaca acgttctgtc 300
ccccttgccg tcccaagcaa tggatgattt gatgctgtcc ccggacgata ttgaacaatg 360
gttcactgaa gacccaggtc cagatgaagc tcccagaatg ccagaggctg ctccccccgt 420
ggcccctgca ccagcagctc ctacaccggc ggcccctgca ccagcccccc cctggcccct 480
gtcatcttct gtcccttccc agaaaaccta ccagggcagc tacggtttcc gtctgggctt 540
cttgcattct gggacagcca agtctgtgac ttgcacgtac tcccctgccc tcaacaagat 600
gttttgccaa ctggccaaga cctgccctgt gcagctgtgg gttgattcca cacccccgcc 660
cggcacccgc gtccgcgcca tggccatcta caagcagtca cagcacatga cggaggttgt 720
gaggcgctgc ccccaccatg agcgctgctc agatagcgat ggtctggccc ctcctcagca 780
tcttatccga gtggaaggaa atttgcgtgt ggagtatttg gatgacagaa acacttttcg 840
acatagtgtg gtggtgccct atgagccgcc tgaggttggc tctgactgta ccaccatcca 900
ctacaactac atgtgtaaca gttcctgcat gggcggcatg aaccggaggc ccatcctcac 960
catcatcaca ctggaagact ccagtggtaa tctactggga cggaacagct ttgaggtgcg 1020
tgtttgtgcc tgtcctggga gagaccggcg cacagaggaa gagaatctcc gcaagaaagg 1080
ggagcctcac cacgagctgc ccccagggag cactaagcga gcactgccca acaacaccag 1140
ctcctctccc cagccaaaga agaaaccact ggatggagaa tatttcaccc ttcagatccg 1200
tgggcgtgag cgcttcgaga tgttccgaga gctgaatgag gccttggaac tcaaggatgc 1260

ccaggctggg aaggagccag gggggagcag ggctcactcc agccacctga agtccaaaaa 1320
gggtcagtct acctcccgcc ataaaaaact catgttcaag acagaagggc ctgactcaga 1380
ctgacattct ccacttcttg ttccccactg acagcctccc acccccatct ctccctcccc 1440
tgccattttg ggttttgggt ctttgaaccc ttgcttgcaa taggtgtgcg tcagaagcac 1500
ccaggacttc catttgcttt gtcccggggc tccactgaac aagttggcct gcactggtgt 1560
tttgttgtgg ggaggaggat ggggagtagg acataccagc ttagatttta aggtttttac 1620
tgtgagggat gtttgggaga tgtaagaaat gttcttgcag ttaagggtta gtttacaatc 1680
agccacattc taggtagggg cccacttcac cgtactaacc agggaagctg tccctcactg 1740
ttgaattttc tctaacttca aggcccatat ctgtgaaatg ctggcatttg cacctacctc 1800
acagagtgca ttgtgagggt taatgaaata atgtacatct ggccttgaaa ccacctttta 1860
ttacatgggg tctagaactt gacccccttg agggtgcttg ttccctctcc ctgttggtcg 1920
gtgggttggt agtttctaca gttgggcagc tggttaggta gagggagttg tcaagtctct 1980
gctggcccag ccaaaccctg tctgacaacc tcttggtgaa ccttagtacc taaaaggaaa 2040
tctcacccca tcccacaccc tggaggattt catctcttgt atatgatgat ctggatccac 2100
caagacttgt tttatgctca gggtcaattt ctttttttctt ttttttttt ttttttcttt 2160
ttctttgaga ctgggtctcg ctttgttgcc caggctggag tggagtggcg tgatcttggc 2220
ttactgcagc ctttgcctcc ccggctcgag cagtcctgcc tcagcctccg gagtagctgg 2280
gaccacaggt tcatgccacc atggccagcc aacttttgca tgttttgtag agatggggtc 2340
tcacagtgtt gcccaggctg gtctcaaact cctgggctca ggcgatccac ctgtctcagc 2400
ctcccagagt gctgggatta caattgtgag ccaccacgtc cagctggaag ggtcaacatc 2460
ttttacattc tgcaagcaca tctgcatttt caccccaccc ttcccctcct tctccctttt 2520
tatatcccat ttttatatcg atctcttatt ttacaataaa actttgctgc cacctgtgtg 2580
tctgaggggt g                                                      2591
```

<210> 16
<211> 1182
<212> ADN
<213> *Homo sapiens*

<400> 16

```
atggaggagc cgcagtcaga tcctagcgtc gagccccctc tgagtcagga aacattttca 60
gacctatgga aactacttcc tgaaaacaac gttctgtccc ccttgccgtc ccaagcaatg 120
gatgatttga tgctgtcccc ggacgatatt gaacaatggt tcactgaaga cccaggtcca 180
gatgaagctc ccagaatgcc agaggctgct ccccgcgtgg cccctgcacc agcagctcct 240
acaccggcgg cccctgcacc agccccctcc tggcccctgt catcttctgt cccttcccag 300
aaaacctacc agggcagcta cggtttccgt ctgggcttct tgcattctgg gacagccaag 360
tctgtgactt gcacgtactc ccctgccctc aacaagatgt tttgccaact ggccaagacc 420
tgccctgtgc agctgtgggt tgattccaca ccccgcccg gcacccgcgt ccgcgccatg 480
gccatctaca agcagtcaca gcacatgacg gaggttgtga ggcgctgccc ccaccatgag 540
cgctgctcag atagcgatgg tctggccct cctcagcatc ttatccgagt ggaaggaaat 600
ttgcgtgtgg agtatttgga tgacagaaac acttttcgac atagtgtggt ggtgccctat 660
gagccgcctg aggttggctc tgactgtacc accatccact acaactacat gtgtaacagt 720
tcctgcatgg gcggcatgaa ccggaggccc atcctcacca tcatcacact ggaagactcc 780
agtggtaatc tactgggacg gaacagcttt gaggtgcgtg tttgtgcctg tcctgggaga 840
gaccggcgca cagaggaaga gaatctccgc aagaaagggg agcctcacca cgagctgccc 900
ccagggagca ctaagcgagc actgcccaac aacaccagct cctctcccca gccaaagaag 960
aaaccactgg atggagaata tttcaccctt cagatccgtg ggcgtgagcg cttcgagatg 1020
ttccgagagc tgaatgaggc cttggaactc aaggatgccc aggctgggaa ggagccaggg 1080
gggagcaggg ctcactccag ccacctgaag tccaaaaagg gtcagtctac ctcccgccat 1140
aaaaaactca tgttcaagac agaagggcct gactcagact ga                    1182
```

<210> 17
<211> 393
<212> PRT
<213> *Homo sapiens*

<400> 17

```
Met Glu Glu Pro Gln Ser Asp Pro Ser Val Glu Pro Pro Leu Ser Gln
 1               5                  10                  15
Glu Thr Phe Ser Asp Leu Trp Lys Leu Leu Pro Glu Asn Asn Val Leu
            20                  25                  30
Ser Pro Leu Pro Ser Gln Ala Met Asp Asp Leu Met Leu Ser Pro Asp
        35                  40                  45
```

```
Asp Ile Glu Gln Trp Phe Thr Glu Asp Pro Gly Pro Asp Glu Ala Pro
    50                  55                  60
Arg Met Pro Glu Ala Ala Pro Arg Val Ala Pro Ala Pro Ala Ala Pro
65                  70                  75                  80
Thr Pro Ala Ala Pro Ala Pro Ala Pro Ser Trp Pro Leu Ser Ser Ser
                85                  90                  95
Val Pro Ser Gln Lys Thr Tyr Gln Gly Ser Tyr Gly Phe Arg Leu Gly
            100                 105                 110
Phe Leu His Ser Gly Thr Ala Lys Ser Val Thr Cys Thr Tyr Ser Pro
        115                 120                 125
Ala Leu Asn Lys Met Phe Cys Gln Leu Ala Lys Thr Cys Pro Val Gln
    130                 135                 140
Leu Trp Val Asp Ser Thr Pro Pro Pro Gly Thr Arg Val Arg Ala Met
145                 150                 155                 160
Ala Ile Tyr Lys Gln Ser Gln His Met Thr Glu Val Val Arg Arg Cys
                165                 170                 175
Pro His His Glu Arg Cys Ser Asp Ser Asp Gly Leu Ala Pro Pro Gln
            180                 185                 190
His Leu Ile Arg Val Glu Gly Asn Leu Arg Val Glu Tyr Leu Asp Asp
        195                 200                 205
Arg Asn Thr Phe Arg His Ser Val Val Val Pro Tyr Glu Pro Pro Glu
    210                 215                 220
Val Gly Ser Asp Cys Thr Thr Ile His Tyr Asn Tyr Met Cys Asn Ser
225                 230                 235                 240
Ser Cys Met Gly Gly Met Asn Arg Arg Pro Ile Leu Thr Ile Ile Thr
                245                 250                 255
Leu Glu Asp Ser Ser Gly Asn Leu Leu Gly Arg Asn Ser Phe Glu Val
            260                 265                 270
Arg Val Cys Ala Cys Pro Gly Arg Asp Arg Arg Thr Glu Glu Glu Asn
        275                 280                 285
Leu Arg Lys Lys Gly Glu Pro His His Glu Leu Pro Pro Gly Ser Thr
    290                 295                 300
Lys Arg Ala Leu Pro Asn Asn Thr Ser Ser Ser Pro Gln Pro Lys Lys
305                 310                 315                 320
Lys Pro Leu Asp Gly Glu Tyr Phe Thr Leu Gln Ile Arg Gly Arg Glu
                325                 330                 335
Arg Phe Glu Met Phe Arg Glu Leu Asn Glu Ala Leu Glu Leu Lys Asp
            340                 345                 350
Ala Gln Ala Gly Lys Glu Pro Gly Gly Ser Arg Ala His Ser Ser His
        355                 360                 365
Leu Lys Ser Lys Lys Gly Gln Ser Thr Ser Arg His Lys Lys Leu Met
    370                 375                 380
Phe Lys Thr Glu Gly Pro Asp Ser Asp
385                 390
```

<210> 18
<211> 1267
<212> ADN
<213> *Homo sapiens*

<400> 18

cgagggctgc ttccggctgg tgcccccggg ggagacccaa cctggggcga cttcaggggt 60
gccacattcg ctaagtgctc ggagttaata gcacctcctc cgagcactcg ctcacggcgt 120
ccccttgcct ggaaagatac cgcggtccct ccagaggatt tgagggacag ggtcggaggg 180
ggctcttccg ccagcaccgg aggaagaaag aggaggggct ggctggtcac cagagggtgg 240
ggcggaccgc gtgcgctcgg cggctgcgga gaggggggaga gcaggcagcg ggcggcgggg 300
agcagcatgg agccggcggc ggggagcagc atggagcctt cggctgactg gctggccacg 360
gccgcggccc ggggtcgggt agaggaggtg cgggcgctgc tggaggcggg ggcgctgccc 420
aacgcaccga atagttacgg tcggaggccg atccaggtca tgatgatggg cagcgcccga 480
gtggcggagc tgctgctgct ccacggcgcg gagcccaact gcgccgaccc cgccactctc 540
acccgacccg tgcacgacgc tgcccgggag ggcttcctgg acacgctggt ggtgctgcac 600
cgggccgggg cgcggctgga cgtgcgcgat gcctggggcc gtctgcccgt ggacctggct 660

gaggagctgg gccatcgcga tgtcgcacgg tacctgcgcg cggctgcggg gggcaccaga 720
ggcagtaacc atgcccgcat agatgccgcg gaaggtccct cagacatccc cgattgaaag 780
aaccagagag gctctgagaa acctcgggaa acttagatca tcagtcaccg aaggtcctac 840
agggccacaa ctgcccccgc cacaacccac cccgctttcg tagttttcat ttagaaaata 900
gagcttttaa aaatgtcctg ccttttaacg tagatatatg ccttccccca ctaccgtaaa 960
tgtccattta tatcattttt tatatattct tataaaaatg taaaaaagaa aaacaccgct 1020
tctgcctttt cactgtgttg gagttttctg gagtgagcac tcacgcccta agcgcacatt 1080
catgtgggca tttcttgcga gcctcgcagc ctccggaagc tgtcgacttc atgacaagca 1140
ttttgtgaac tagggaagct caggggggtt actggcttct cttgagtcac actgctagca 1200
aatggcagaa ccaaagctca aataaaaata aaataatttt cattcattca ctcaaaaaaa 1260
aaaaaaa 1267

<210> 19
<211> 471
<212> ADN
<213> *Homo sapiens*

<400> 19

atggagccgg cggcggggag cagcatggag ccttcggctg actggctggc cacggccgcg 60
gcccggggtc gggtagagga ggtgcgggcg ctgctggagg cgggggcgct gcccaacgca 120
ccgaatagtt acggtcggag gccgatccag gtcatgatga tgggcagcgc ccgagtggcg 180
gagctgctgc tgctccacgg cgcggagccc aactgcgccg accccgccac tctcacccga 240
cccgtgcacg acgctgcccg ggagggcttc ctggacacgc tggtggtgct gcaccgggcc 300
ggggcgcggc tggacgtgcg cgatgcctgg ggccgtctgc ccgtggacct ggctgaggag 360
ctgggccatc gcgatgtcgc acggtacctg cgcgcggctg cgggggggcac cagaggcagt 420
aaccatgccc gcatagatgc cgcggaaggt ccctcagaca tccccgattg a 471

<210> 20
<211> 156
<212> PRT
<213> *Homo sapiens*

<400> 20

```
Met Glu Pro Ala Ala Gly Ser Ser Met Glu Pro Ser Ala Asp Trp Leu
 1               5                  10                  15
Ala Thr Ala Ala Ala Arg Gly Arg Val Glu Glu Val Arg Ala Leu Leu
            20                  25                  30
Glu Ala Gly Ala Leu Pro Asn Ala Pro Asn Ser Tyr Gly Arg Arg Pro
        35                  40                  45
Ile Gln Val Met Met Met Gly Ser Ala Arg Val Ala Glu Leu Leu Leu
    50                  55                  60
Leu His Gly Ala Glu Pro Asn Cys Ala Asp Pro Ala Thr Leu Thr Arg
65                  70                  75                  80
Pro Val His Asp Ala Ala Arg Glu Gly Phe Leu Asp Thr Leu Val Val
                85                  90                  95
Leu His Arg Ala Gly Ala Arg Leu Asp Val Arg Asp Ala Trp Gly Arg
            100                 105                 110
Leu Pro Val Asp Leu Ala Glu Glu Leu Gly His Arg Asp Val Ala Arg
        115                 120                 125
Tyr Leu Arg Ala Ala Ala Gly Gly Thr Arg Gly Ser Asn His Ala Arg
    130                 135                 140
Ile Asp Ala Ala Glu Gly Pro Ser Asp Ile Pro Asp
145                 150                 155
```

<210> 21
<211> 1164
<212> ADN
<213> *Homo sapiens*

<400> 21

```
cgctcaggga aggcgggtgc gcgcctgcgg ggcggagatg ggcaggggcc ggtgcgtggg 60
tcccagtctg cagttaaggg ggcaggagtg gcgctgctca cctctggtgc caaagggcgg 120

cgcagcggct gccgagctcg gccctggagg cggcgagaac atggtgcgca ggttcttggt 180
gaccctccgg attcggcgcg cgtgcggccc gccgcgagtg agggttttcg tggttcacat 240
cccgcggctc acgggggagt gggcagcgcc aggggcgccc gccgctgtgg ccctcgtgct 300
gatgctactg aggagccagc gtctagggca gcagccgctt cctagaagac caggtcatga 360
tgatgggcag cgcccgagtg gcggagctgc tgctgctcca cggcgcggag cccaactgcg 420
ccgaccccgc cactctcacc cgacccgtgc acgacgctgc ccgggagggc ttcctggaca 480
cgctggtggt gctgcaccgg gccggggcgc ggctggacgt gcgcgatgcc tggggccgtc 540
tgcccgtgga cctggctgag gagctgggcc atcgcgatgt cgcacggtac ctgcgcgcgg 600
ctgcgggggg caccagaggc agtaaccatg cccgcataga tgccgcggaa ggtccctcag 660
acatccccga ttgaaagaac cagagaggct ctgagaaacc tcgggaaact tagatcatca 720
gtcaccgaag gtcctacagg gccacaactg cccccgccac aacccacccc gctttcgtag 780
ttttcattta gaaaatagag cttttaaaaa tgtcctgcct tttaacgtag atatatgcct 840
tcccccacta ccgtaaatgt ccatttatat cattttttat atattcttat aaaaatgtaa 900
aaaagaaaaa caccgcttct gccttttcac tgtgttggag ttttctggag tgagcactca 960
cgccctaagc gcacattcat gtgggcattt cttgcgagcc tcgcagcctc cggaagctgt 1020
cgacttcatg acaagcattt tgtgaactag ggaagctcag gggggttact ggcttctctt 1080
gagtcacact gctagcaaat ggcagaacca aagctcaaat aaaaataaaa taattttcat 1140
tcattcactc aaaaaaaaaa aaaa                                       1164
```

<210> 22
<211> 522
<212> ADN
<213> *Homo sapiens*

<400> 22

```
atgggcaggg ggcggtgcgt gggtcccagt ctgcagttaa gggggcagga gtggcgctgc 60
tcacctctgg tgccaaaggg cggcgcagcg gctgccgagc tcggccctgg aggcggcgag 120
aacatggtgc gcaggttctt ggtgaccctc cggattcggc gcgcgtgcgg cccgccgcga 180
gtgagggttt tcgtggttca catcccgcgg ctcacggggg agtgggcagc gccaggggcg 240
cccgccgctg tggccctcgt gctgatgcta ctgaggagcc agcgtctagg gcagcagccg 300
cttcctagaa gaccaggtca tgatgatggg cagcgcccga gtggcggagc tgctgctgct 360
ccacggcgcg gagcccaact gcgccgaccc cgccactctc acccgacccg tgcacgacgc 420
tgcccgggag ggcttcctgg acacgctggt ggtgctgcac cgggccgggg cgcggctgga 480
cgtgcgcgat gcctggggcc gtctgcccgt ggacctggct ga                    522
```

<210> 23
<211> 173
<212> PRT
<213> *Homo sapiens*

<400> 23

```
Met Gly Arg Gly Arg Cys Val Gly Pro Ser Leu Gln Leu Arg Gly Gln
 1               5                  10                  15
Glu Trp Arg Cys Ser Pro Leu Val Pro Lys Gly Gly Ala Ala Ala Ala
            20                  25                  30
Glu Leu Gly Pro Gly Gly Gly Glu Asn Met Val Arg Arg Phe Leu Val
        35                  40                  45
Thr Leu Arg Ile Arg Arg Ala Cys Gly Pro Pro Arg Val Arg Val Phe
    50                  55                  60
Val Val His Ile Pro Arg Leu Thr Gly Glu Trp Ala Ala Pro Gly Ala
65                  70                  75                  80
Pro Ala Ala Val Ala Leu Val Leu Met Leu Leu Arg Ser Gln Arg Leu
                85                  90                  95
Gly Gln Gln Pro Leu Pro Arg Arg Pro Gly His Asp Asp Gly Gln Arg
            100                 105                 110
Pro Ser Gly Gly Ala Ala Ala Ala Pro Arg Arg Gly Ala Gln Leu Arg
        115                 120                 125
Arg Pro Arg His Ser His Pro Thr Arg Ala Arg Arg Cys Pro Gly Gly
    130                 135                 140
Leu Pro Gly His Ala Gly Gly Ala Ala Pro Gly Arg Gly Ala Ala Gly
145                 150                 155                 160
Arg Ala Arg Cys Leu Gly Pro Ser Ala Arg Gly Pro Gly
                165                 170
```

<210> 24
<211> 10740
<212> ADN
<213> *Homo sapiens*

<400> 24

gtattggtgc agcccgccag ggtgtcactg gagacagaat ggaggtgctg ccggactcgg 60
aaatggggtc caagggtagc caaggatggc tgcagcttca tatgatcagt tgttaaagca 120
agttgaggca ctgaagatgg agaactcaaa tcttcgacaa gagctagaag ataattccaa 180
tcatcttaca aaactggaaa ctgaggcatc taatatgaag gaagtactta aacaactaca 240
aggaagtatt gaagatgaag ctatggcttc ttctggacag attgatttat tagagcgtct 300
taaagagctt aacttagata gcagtaattt ccctggagta aaactgcggt caaaaatgtc 360
cctccgttct tatggaagcc gggaaggatc tgtatcaagc cgttctggag agtgcagtcc 420
tgttcctatg ggttcatttc caagaagagg gtttgtaaat ggaagcagag aaagtactgg 480
atatttagaa gaacttgaga aagagaggtc attgcttctt gctgatcttg acaaagaaga 540
aaaggaaaaa gactggtatt acgctcaact tcagaatctc actaaaagaa tagatagtct 600
tcctttaact gaaaattttt ccttacaaac agatatgacc agaaggcaat tggaatatga 660
agcaaggcaa atcagagttg cgatggaaga acaactaggt acctgccagg atatggaaaa 720
acgagcacag cgaagaatag ccagaattca gcaaatcgaa aaggacatac ttcgtatacg 780
acagctttta cagtcccaag caacagaagc agagaggtca tctcagaaca agcatgaaac 840
cggctcacat gatgctgagc ggcagaatga aggtcaagga gtgggagaaa tcaacatggc 900
aacttctggt aatggtcagg gttcaactac acgaatggac catgaaacag ccagtgtttt 960
gagttctagt agcacacact ctgcacctcg aaggctgaca agtcatctgg gaaccaaggt 1020
ggaaatggtg tattcattgt tgtcaatgct tggtactcat gataaggatg atatgtcgcg 1080
aactttgcta gctatgtcta gctcccaaga cagctgtata tccatgcgac agtctggatg 1140
tcttcctctc ctcatccagc ttttacatgg caatgacaaa gactctgtat tgttgggaaa 1200
ttcccggggc agtaaagagg ctcgggccag ggccagtgca gcactccaca acatcattca 1260
ctcacagcct gatgacaaga gaggcaggcg tgaaatccga gtccttcatc ttttggaaca 1320
gatacgcgct tactgtgaaa cctgttggga gtggcaggaa gctcatgaac caggcatgga 1380
ccaggacaaa aatccaatgc cagctcctgt tgaacatcag atctgtcctg ctgtgtgtgt 1440
tctaatgaaa ctttcatttg atgaagagca tagacatgca atgaatgaac taggggggact 1500
acaggccatt gcagaattat tgcaagtgga ctgtgaaatg tatgggctta ctaatgacca 1560
ctacagtatt acactaagac gatatgctgg aatggctttg acaaacttga cttttggaga 1620
tgtagccaac aaggctacgc tatgctctat gaaaggctgc atgagagcac ttgtggccca 1680
actaaaatct gaaagtgaag acttacagca ggttattgcg agtgttttga ggaatttgtc 1740
ttggcgagca gatgtaaata gtaaaaagac gttgcgagaa gttggaagtg tgaaagcatt 1800
gatggaatgt gctttagaag ttaaaaagga atcaaccctc aaaagcgtat tgagtgcctt 1860
atggaatttg tcagcacatt gcactgagaa taaagctgat atatgtgctg tagatggtgc 1920
acttgcattt ttggttggca ctcttactta ccggagccag acaaacactt tagccattat 1980
tgaaagtgga ggtgggatat tacggaatgt gtccagcttg atagctacaa atgaggacca 2040
caggcaaatc ctaagagaga acaactgtct acaaacttta ttacaacact taaaatctca 2100
tagtttgaca atagtcagta atgcatgtgg aactttgtgg aatctctcag caagaaatcc 2160
taaagaccag gaagcattat gggacatggg ggcagttagc atgctcaaga acctcattca 2220
ttcaaagcac aaaatgattg ctatgggaag tgctgcagct ttaaggaatc tcatggcaaa 2280
taggcctgcg aagtacaagg atgccaatat tatgtctcct ggctcaagct tgccatctct 2340
tcatgttagg aaacaaaaag ccctagaagc agaattagat gctcagcact tatcagaaac 2400
ttttgacaat atagacaatt taagtcccaa ggcatctcat cgtagtaagc agagacacaa 2460
gcaaagtctc tatggtgatt atgtttttga caccaatcga catgatgata ataggtcaga 2520
caattttaat actggcaaca tgactgtcct ttcaccatat ttgaatacta cagtgttacc 2580
cagctcctct tcatcaagag gaagcttaga tagttctcgt tctgaaaaag atagaagttt 2640
ggagagagaa cgcggaattg gtctaggcaa ctaccatcca gcaacagaaa atccaggaac 2700
ttcttcaaag cgaggtttgc agatctccac cactgcagcc cagattgcca aagtcatgga 2760
agaagtgtca gccattcata cctctcagga agacagaagt tctgggtcta ccactgaatt 2820
acattgtgtg acagatgaga gaaatgcact tagaagaagc tctgctgccc atacacattc 2880
aaacacttac aatttcacta agtcggaaaa ttcaaatagg acatgttcta tgccttatgc 2940
caaattagaa tacaagagat cttcaaatga tagtttaaat agtgtcagta gtagtgatgg 3000
ttatggtaaa agaggtcaaa tgaaaccctc gattgaatcc tattctgaag atgatgaaag 3060
taagttttgc agttatggtc aatacccagc cgacctagcc cataaaatac atagtgcaaa 3120
tcatatggat gataatgatg gagaactaga tacaccaata aattatagtc ttaaatattc 3180
agatgagcag ttgaactctg gaaggcaaag tccttcacag aatgaaagat gggcaagacc 3240
caaacacata atagaagatg aaataaaaca aagtgagcaa agacaatcaa ggaatcaaag 3300

tacaacttat cctgtttata ctgagagcac tgatgataaa cacctcaagt tccaaccaca 3360
ttttggacag caggaatgtg tttctccata caggtcacgg ggagccaatg gttcagaaac 3420
aaatcgagtg ggttctaatc atggaattaa tcaaaatgta agccagtctt tgtgtcaaga 3480
agatgactat gaagatgata agcctaccaa ttatagtgaa cgttactctg aagaagaaca 3540
gcatgaagaa gaagagagac caacaaatta tagcataaaa tataatgaag agaaacgtca 3600
tgtggatcag cctattgatt atagtttaaa atatgccaca gatattcctt catcacagaa 3660
acagtcattt tcattctcaa agagttcatc tggacaaagc agtaaaaccg aacatatgtc 3720
ttcaagcagt gagaatacgt ccacaccttc atctaatgcc aagaggcaga atcagctcca 3780
tccaagttct gcacagagta gaagtggtca gcctcaaaag gctgccactt gcaaagtttc 3840
ttctattaac caagaaacaa tacagactta ttgtgtagaa gatactccaa tatgtttttc 3900
aagatgtagt tcattatcat ctttgtcatc agctgaagat gaaataggat gtaatcagac 3960
gacacaggaa gcagattctg ctaataccct gcaaatagca gaaataaaag aaaagattgg 4020
aactaggtca gctgaagatc ctgtgagcga agttccagca gtgtcacagc accctagaac 4080
caaatccagc agactgcagg gttctagttt atcttcagaa tcagccaggc acaaagctgt 4140
tgaattttct tcaggagcga aatctccctc caaaagtggt gctcagacac ccaaaagtcc 4200
acctgaacac tatgttcagg agaccccact catgtttagc agatgtactt ctgtcagttc 4260
acttgatagt tttgagagtc gttcgattgc cagctccgtt cagagtgaac catgcagtgg 4320
aatggtaagt ggcattataa gccccagtga tcttccagat agccctggac aaaccatgcc 4380
accaagcaga agtaaaacac ctccaccacc tcctcaaaca gctcaaacca agcgagaagt 4440
acctaaaaat aaagcaccta ctgctgaaaa gagagagagt ggacctaagc aagctgcagt 4500
aaatgctgca gttcagaggg tccaggttct tccagatgct gatactttat tacattttgc 4560
cacggaaagt actccagatg gattttcttg ttcatccagc ctgagtgctc tgagcctcga 4620
tgagccattt atacagaaag atgtggaatt aagaataatg cctccagttc aggaaaatga 4680
caatgggaat gaaacagaat cagagcagcc taaagaatca aatgaaaacc aagagaaaga 4740
ggcagaaaaa actattgatt ctgaaaagga cctattagat gattcagatg atgatgatat 4800
tgaaatacta gaagaatgta ttatttctgc catgccaaca aagtcatcac gtaaagcaaa 4860
aaagccagcc cagactgctt caaaattacc tccacctgtg gcaaggaaac caagtcagct 4920
gcctgtgtac aaacttctac catcacaaaa caggttgcaa ccccaaaagc atgttagttt 4980
tacaccgggg gatgatatgc cacgggtgta ttgtgttgaa gggacaccta taaacttttc 5040
cacagctaca tctctaagtg atctaacaat cgaatcccct ccaaatgagt tagctgctgg 5100
agaaggagtt agaggagggg cacagtcagg tgaatttgaa aaacgagata ccattcctac 5160
agaaggcaga agtacagatg aggctcaagg aggaaaaacc tcatctgtaa ccatacctga 5220
attggatgac aataaagcag aggaaggtga tattcttgca gaatgcatta attctgctat 5280
gcccaaaggg aaaagtcaca agcctttccg tgtgaaaaag ataatggacc aggtccagca 5340
agcatctgcg tcttcttctg cacccaacaa aaatcagtta gatggtaaga aaaagaaacc 5400
aacttcacca gtaaaaccta taccacaaaa tactgaatat aggacacgtg taagaaaaaa 5460
tgcagactca aaaaataatt taaatgctga gagagttttc tcagacaaca aagattcaaa 5520
gaaacagaat ttgaaaaata attccaaggt cttcaatgat aagctcccaa ataatgaaga 5580
tagagtcaga ggaagttttg cttttgattc acctcatcat tacacgccta ttgaaggaac 5640
tccttactgt ttttcacgaa atgattcttt gagttctcta gattttgatg atgatgatgt 5700
tgacctttcc agggaaaagg ctgaattaag aaaggcaaaa gaaaataagg aatcagaggc 5760
taaagttacc agccacacag aactaacctc caaccaacaa tcagctaata agacacaagc 5820
tattgcaaag cagccaataa atcgaggtca gcctaaaccc atacttcaga aacaatccac 5880
ttttccccag tcatccaaag acataccaga cagaggggca gcaactgatg aaaagttaca 5940
gaattttgct attgaaaata ctccggtttg cttttctcat aattcctctc tgagttctct 6000
cagtgacatt gaccaagaaa acaacaataa agaaaatgaa cctatcaaag agactgagcc 6060
ccctgactca cagggagaac caagtaaacc tcaagcatca ggctatgctc ctaaatcatt 6120
tcatgttgaa gataccccag tttgtttctc aagaaacagt tctctcagtt ctcttagtat 6180
tgactctgaa gatgacctgt tgcaggaatg tataagctcc gcaatgccaa aaaagaaaaa 6240
gccttcaaga ctcaagggtg ataatgaaaa acatagtccc agaaatatgg gtggcatatt 6300
aggtgaagat ctgacacttg atttgaaaga tatacagaga ccagattcag aacatggtct 6360
atcccctgat tcagaaaatt ttgattggaa agctattcag gaaggtgcaa attccatagt 6420
aagtagttta catcaagctg ctgctgctgc atgtttatct agacaagctt cgtctgattc 6480
agattccatc ctttccctga aatcaggaat ctctctggga tcaccatttc atcttacacc 6540
tgatcaagaa gaaaaaccct ttacaagtaa taaaggccca cgaattctaa aaccagggga 6600
gaaaagtaca ttggaaacta aaaagataga atctgaaagt aaaggaatca aaggaggaaa 6660
aaaagtttat aaaagtttga ttactggaaa agttcgatct aattcagaaa tttcaggcca 6720
aatgaaacag ccccttcaag caaacatgcc ttcaatctct cgaggcagga caatgattca 6780
tattccagga gttcgaaata gctcctcaag tacaagtcct gtttctaaaa aaggcccacc 6840
ccttaagact ccagcctcca aaagccctag tgaaggtcaa acagccacca cttctcctag 6900
aggagccaag ccatctgtga aatcagaatt aagccctgtt gccaggcaga catcccaaat 6960
aggtgggtca agtaaagcac cttctagatc aggatctaga gattcgaccc cttcaagacc 7020
tgcccagcaa ccattaagta gacctataca gtctcctggc cgaaactcaa tttcccctgg 7080

tagaaatgga ataagtcctc ctaacaaatt atctcaactt ccaaggacat catcccctag 7140
tactgcttca actaagtcct caggttctgg aaaaatgtca tatacatctc caggtagaca 7200
gatgagccaa cagaacctta ccaaacaaac aggtttatcc aagaatgcca gtagtattcc 7260
aagaagtgag tctgcctcca aaggactaaa tcagatgaat aatggtaatg gagccaataa 7320
aaaggtagaa ctttctagaa tgtcttcaac taaatcaagt ggaagtgaat ctgatagatc 7380
agaaagacct gtattagtac gccagtcaac tttcatcaaa gaagctccaa gcccaacctt 7440
aagaagaaaa ttggaggaat ctgcttcatt tgaatctctt tctccatcat ctagaccagc 7500
ttctcccact aggtcccagg cacaaactcc agttttaagt ccttcccttc ctgatatgtc 7560
tctatccaca cattcgtctg ttcaggctgg tggatggcga aaactcccac ctaatctcag 7620
tcccactata gagtataatg atggaagacc agcaaagcgc catgatattg cacggtctca 7680
ttctgaaagt ccttctagac ttccaatcaa taggtcagga acctggaaac gtgagcacag 7740
caaacattca tcatcccttc ctcgagtaag cacttggaga agaactggaa gttcatcttc 7800
aattctttct gcttcatcag aatccagtga aaaagcaaaa agtgaggatg aaaaacatgt 7860
gaactctatt tcaggaacca aacaaagtaa agaaaaccaa gtatccgcaa aaggaacatg 7920
gagaaaaata aaagaaaatg aattttctcc cacaaatagt acttctcaga ccgtttcctc 7980
aggtgctaca aatggtgctg aatcaaagac tctaatttat caaatggcac ctgctgtttc 8040
taaaacagag gatgtttggg tgagaattga ggactgtccc attaacaatc ctagatctgg 8100
aagatctccc acaggtaata ctccccggt gattgacagt gtttcagaaa aggcaaatcc 8160
aaacattaaa gattcaaaag ataatcaggc aaaacaaaat gtgggtaatg gcagtgttcc 8220
catgcgtacc gtgggtttgg aaaatcgcct gaactccttt attcaggtgg atgcccctga 8280
ccaaaaagga actgagataa aaccaggaca aaataatcct gtccctgtat cagagactaa 8340
tgaaagttct atagtggaac gtaccccatt cagttctagc agctcaagca aacacagttc 8400
acctagtggg actgttgctg ccagagtgac tccttttaat tacaacccaa gccctaggaa 8460
aagcagcgca gatagcactt cagctcggcc atctcagatc ccaactccag tgaataacaa 8520
cacaaagaag cgagattcca aaactgacag cacagaatcc agtggaaccc aaagtcctaa 8580
gcgccattct gggtcttacc ttgtgacatc tgtttaaaag agaggaagaa tgaaactaag 8640
aaaattctat gttaattaca actgctatat agacattttg tttcaaatga aactttaaaa 8700
gactgaaaaa ttttgtaaat aggtttgatt cttgttagag ggtttttgtt ctggaagcca 8760
tatttgatag tatactttgt cttcactggt cttattttgg gaggcactct tgatggttag 8820
gaaaaaaata gtaaagccaa gtatgtttgt acagtatgtt ttacatgtat ttaaagtagc 8880
atcccatccc aacttccttt aattattgct tgtcttaaaa taatgaacac tacagataga 8940
aaatatgata tattgctgtt atcaatcatt tctagattat aaactgacta aacttacatc 9000
agggaaaaat tggtatttat gcaaaaaaaa atgtttttgt ccttgtgagt ccatctaaca 9060
tcataattaa tcatgtggct gtgaaattca cagtaatatg gttcccgatg aacaagttta 9120
cccagcctgc tttgctttac tgcatgaatg aaactgatgg ttcaatttca gaagtaatga 9180
ttaacagtta tgtggtcaca tgatgtgcat agagatagct acagtgtaat aatttacact 9240
attttgtgct ccaaacaaaa caaaaatctg tgtaactgta aaacattgaa tgaaactatt 9300
ttacctgaac tagattttat ctgaaagtag gtagaatttt tgctatgctg taatttgttg 9360
tatattctgg tatttgaggt gagatggctg ctcttttatt aatgagacat gaattgtgtc 9420
tcaacagaaa ctaaatgaac atttcagaat aaattattgc tgtatgtaaa ctgttactga 9480
aattggtatt tgtttgaagg gtcttgtttc acatttgtat taataattgt ttaaaatgcc 9540
tcttttaaaa gcttatataa attttttttct tcagcttcta tgcattaaga gtaaaattcc 9600
tcttactgta ataaaaacaa ttgaagaaga ctgttgccac ttaaccattc catgcgttgg 9660
cacttatcta ttcctgaaat ttcttttatg tgattagctc atcttgattt ttaatatttt 9720
tccacttaaa ctttttttttc ttactccact ggagctcagt aaaagtaaat tcatgtaata 9780
gcaatgcaag cagcctagca cagactaagc attgagcata ataggcccac ataatttcct 9840
ctttcttaat attatagaat tctgtacttg aaattgattc ttagacattg cagtctcttc 9900
gaggctttac agtgtaaact gtcttgcccc ttcatcttct tgttgcaact gggtctgaca 9960
tgaacacttt ttatcaccct gtatgttagg gcaagatctc agcagtgaag tataatcagc 10020
actttgccat gctcagaaaa ttcaaatcac atggaacttt agaggtagat ttaatacgat 10080
taagatattc agaagtatat tttagaatcc ctgcctgtta aggaaacttt atttgtggta 10140
ggtacagttc tggggtacat gttaagtgtc cccttataca gtggagggaa gtcttccttc 10200
ctgaaggaaa ataaactgac acttattaac taagataatt tacttaatat atcttccctg 10260
atttgtttta aaagatcaga gggtgactga tgatacatgc atacatattt gttgaataaa 10320
tgaaaattta tttttagtga taagattcat acactctgta tttggggagg gaaaaccttt 10380
ttaagcatgg tggggcactc agataggagt gaatacacct acctggtgcc ttgaaaatca 10440
catcaagtag ttaattatct acccccttacc tgtgtttata acttccaggt aatgagaatg 10500
atttttttta aagctaaaat gccagtaaat aaaagtgcta tgacttgagc taagatattt 10560
gactccaatg cctgtactgt gtctactgca ccactttgta aacacttcaa tttactatct 10620
ttgaaatgat tgacctttaa atttttgcca aatgttatct gaaattgtct atgaatacca 10680
tctacttctg ttgttttccc aggcttccat aaacaatgga gatacatgca aaaaaaaaaa 10740

<210> 25
<211> 8532
<212> ADN
<213> *Homo sapiens*

<400> 25

atggctgcag cttcatatga tcagttgtta aagcaagttg aggcactgaa gatggagaac 60
tcaaatcttc gacaagagct agaagataat tccaatcatc ttacaaaact ggaaactgag 120
gcatctaata tgaaggaagt acttaaacaa ctacaaggaa gtattgaaga tgaagctatg 180
gcttcttctg gacagattga tttattagag cgtcttaaag agcttaactt agatagcagt 240
aatttccctg gagtaaaact gcggtcaaaa atgtccctcc gttcttatgg aagccgggaa 300
ggatctgtat caagccgttc tggagagtgc agtcctgttc ctatgggttc atttccaaga 360
agagggtttg taaatggaag cagagaaagt actggatatt tagaagaact tgagaaagag 420
aggtcattgc ttcttgctga tcttgacaaa gaagaaaagg aaaaagactg gtattacgct 480
caacttcaga atctcactaa aagaatagat agtcttcctt taactgaaaa tttttcctta 540
caaacagata tgaccagaag gcaattggaa tatgaagcaa ggcaaatcag agttgcgatg 600
gaagaacaac taggtacctg ccaggatatg gaaaaacgag cacagcgaag aatagccaga 660
attcagcaaa tcgaaaagga catacttcgt atacgacagc ttttacagtc ccaagcaaca 720
gaagcagaga ggtcatctca gaacaagcat gaaaccggct cacatgatgc tgagcggcag 780
aatgaaggtc aaggagtggg agaaatcaac atggcaactt ctggtaatgg tcagggttca 840
actacacgaa tggaccatga aacagccagt gttttgagtt ctagtagcac acactctgca 900
cctcgaaggc tgacaagtca tctgggaacc aaggtggaaa tggtgtattc attgttgtca 960
atgcttggta ctcatgataa ggatgatatg tcgcgaactt tgctagctat gtctagctcc 1020
caagacagct gtatatccat gcgacagtct ggatgtcttc ctctcctcat ccagctttta 1080
catggcaatg acaaagactc tgtattgttg ggaaattccc ggggcagtaa agaggctcgg 1140
gccagggcca gtgcagcact ccacaacatc attcactcac agcctgatga caagagaggc 1200
aggcgtgaaa tccgagtcct tcatcttttg gaacagatac gcgcttactg tgaaacctgt 1260
tgggagtggc aggaagctca tgaaccaggc atggaccagg acaaaaatcc aatgccagct 1320
cctgttgaac atcagatctg tcctgctgtg tgtgttctaa tgaaactttc atttgatgaa 1380
gagcatagac atgcaatgaa tgaactaggg ggactacagg ccattgcaga attattgcaa 1440
gtggactgtg aaatgtacgg gcttactaat gaccactaca gtattacact aagacgatat 1500
gctggaatgg ctttgacaaa cttgactttt ggagatgtag ccaacaaggc tacgctatgc 1560
tctatgaaag gctgcatgag agcacttgtg gcccaactaa aatctgaaag tgaagactta 1620
cagcaggtta ttgcaagtgt tttgaggaat ttgtcttggc gagcagatgt aaatagtaaa 1680
aagacgttgc gagaagttgg aagtgtgaaa gcattgatgg aatgtgcttt agaagttaaa 1740
aaggaatcaa ccctcaaaag cgtattgagt gccttatgga atttgtcagc acattgcact 1800
gagaataaag ctgatatatg tgctgtagat ggtgcacttg cattttggt tggcactctt 1860
acttaccgga gccagacaaa cactttagcc attattgaaa gtggaggtgg gatattacgg 1920
aatgtgtcca gcttgatagc tacaaatgag gaccacaggc aaatcctaag agagaacaac 1980
tgtctacaaa ctttattaca acacttaaaa tctcatagtt tgacaatagt cagtaatgca 2040
tgtggaactt tgtggaatct ctcagcaaga aatcctaaag accaggaagc attatgggac 2100
atgggggcag ttagcatgct caagaacctc attcattcaa agcacaaaat gattgctatg 2160
ggaagtgctg cagctttaag gaatctcatg gcaaataggc ctgcgaagta caaggatgcc 2220
aatattatgt ctcctggctc aagcttgcca tctcttcatg ttaggaaaca aaaagcccta 2280
gaagcagaat tagatgctca gcacttatca gaaacttttg acaatataga caatttaagt 2340
cccaaggcat ctcatcgtag taagcagaga cacaagcaaa gtctctatgg tgattatgtt 2400
tttgacacca atcgacatga tgataatagg tcagacaatt ttaatactgg caacatgact 2460
gtcctttcac catatttgaa tactacagtg ttacccagct cctcttcatc aagaggaagc 2520
ttagatagtt ctcgttctga aaaagataga agtttggaga gagaacgcgg aattggtcta 2580
ggcaactacc atccagcaac agaaaatcca ggaacttctt caaagcgagg tttgcagatc 2640
tccaccactg cagcccagat tgccaaagtc atggaagaag tgtcagccat tcatacctct 2700
caggaagaca gaagttctgg gtctaccact gaattacatt gtgtgacaga tgagagaaat 2760
gcacttagaa gaagctctgc tgcccataca cattcaaaca cttacaattt cactaagtcg 2820
gaaaattcaa ataggacatg ttctatgcct tatgccaaat tagaatacaa gagatcttca 2880
aatgatagtt taaatagtgt cagtagtagt gatggttatg gtaaaagagg tcaaatgaaa 2940
ccctcgattg aatcctattc tgaagatgat gaaagtaagt tttgcagtta tggtcaatac 3000
ccagccgacc tagcccataa aatacatagt gcaaatcata tggatgataa tgatggagaa 3060
ctagatacac caataaatta tagtcttaaa tattcagatg agcagttgaa ctctggaagg 3120
caaagtcctt cacagaatga aagatgggca agacccaaac acataataga agatgaaata 3180
aaacaaagtg agcaaagaca atcaaggaat caaagtacaa cttatcctgt ttatactgag 3240
agcactgatg ataaacacct caagttccaa ccacattttg gacagcagga atgtgtttct 3300
ccatacaggt cacggggagc caatggttca gaaacaaatc gagtgggttc taatcatgga 3360
attaatcaaa atgtaagcca gtctttgtgt caagaagatg actatgaaga tgataagcct 3420

accaattata gtgaacgtta ctctgaagaa gaacagcatg aagaagaaga gagaccaaca 3480
aattatagca taaaatataa tgaagagaaa cgtcatgtgg atcagcctat tgattatagt 3540
ttaaaatatg ccacagatat tccttcatca cagaaacagt cattttcatt ctcaaagagt 3600
tcatctggac aaagcagtaa aaccgaacat atgtcttcaa gcagtgagaa tacgtccaca 3660
ccttcatcta atgccaagag gcagaatcag ctccatccaa gttctgcaca gagtagaagt 3720
ggtcagcctc aaaaggctgc cacttgcaaa gtttcttcta ttaaccaaga aacaatacag 3780
acttattgtg tagaagatac tccaatatgt ttttcaagat gtagttcatt atcatctttg 3840
tcatcagctg aagatgaaat aggatgtaat cagacgacac aggaagcaga ttctgctaat 3900
accctgcaaa tagcagaaat aaaagaaaag attggaacta ggtcagctga agatcctgtg 3960
agcgaagttc cagcagtgtc acagcaccct agaaccaaat ccagcagact gcagggttct 4020
agtttatctt cagaatcagc caggcacaaa gctgttgaat tttcttcagg agcgaaatct 4080
ccctccaaaa gtggtgctca gacacccaaa agtccacctg aacactatgt tcaggagacc 4140
ccactcatgt ttagcagatg tacttctgtc agttcacttg atagttttga gagtcgttcg 4200
attgccagct ccgttcagag tgaaccatgc agtggaatgg taagtggcat tataagcccc 4260
agtgatcttc cagatagccc tggacaaacc atgccaccaa gcagaagtaa aacacctcca 4320
ccacctcctc aaacagctca aaccaagcga gaagtaccta aaaataaagc acctactgct 4380
gaaaagagag agagtggacc taagcaagct gcagtaaatg ctgcagttca gagggtccag 4440
gttcttccag atgctgatac tttattacat tttgccacgg aaagtactcc agatggattt 4500
tcttgttcat ccagcctgag tgctctgagc ctcgatgagc catttataca gaaagatgtg 4560
gaattaagaa taatgcctcc agttcaggaa aatgacaatg ggaatgaaac agaatcagag 4620
cagcctaaag aatcaaatga aaaccaagag aaagaggcag aaaaaactat tgattctgaa 4680
aaggacctat tagatgattc agatgatgat gatattgaaa tactagaaga atgtattatt 4740
tctgccatgc caacaaagtc atcacgtaaa gcaaaaaagc cagcccagac tgcttcaaaa 4800
ttacctccac ctgtggcaag gaaaccaagt cagctgcctg tgtacaaact tctaccatca 4860
caaaacaggt tgcaacccca aaagcatgtt agttttacac cgggggatga tatgccacgg 4920
gtgtattgtg ttgaagggac acctataaac ttttccacag ctacatctct aagtgatcta 4980
acaatcgaat cccctccaaa tgagttagct gctggagaag gagttagagg aggagcacag 5040
tcaggtgaat ttgaaaaacg agataccatt cctacagaag gcagaagtac agatgaggct 5100
caaggaggaa aaacctcatc tgtaaccata cctgaattgg atgacaataa agcagaggaa 5160
ggtgatattc ttgcagaatg cattaattct gctatgccca aagggaaaag tcacaagcct 5220
ttccgtgtga aaaagataat ggaccaggtc cagcaagcat ctgcgtcgtc ttctgcaccc 5280
aacaaaaatc agttagatgg taagaaaaag aaaccaactt caccagtaaa acctatacca 5340
caaaatactg aatataggac acgtgtaaga aaaaatgcag actcaaaaaa taatttaaat 5400
gctgagagag ttttctcaga caacaaagat tcaaagaaac agaatttgaa aaataattcc 5460
aaggacttca atgataagct cccaaataat gaagatagag tcagaggaag ttttgctttt 5520
gattcacctc atcattacac gcctattgaa ggaactcctt actgtttttc acgaaatgat 5580
tctttgagtt ctctagattt tgatgatgat gatgttgacc tttccaggga aaaggctgaa 5640
ttaagaaagg caaaagaaaa taaggaatca gaggctaaag ttaccagcca cacagaacta 5700
acctccaacc aacaatcagc taataagaca caagctattg caaagcagcc aataaatcga 5760
ggtcagccta aacccatact tcagaaacaa tccacttttc cccagtcatc caaagacata 5820
ccagacagag gggcagcaac tgatgaaaag ttacagaatt ttgctattga aaatactcca 5880
gtttgctttt ctcataattc ctctctgagt tctctcagtg acattgacca agaaaacaac 5940
aataaagaaa atgaacctat caaagagact gagccccctg actcacaggg agaaccaagt 6000
aaacctcaag catcaggcta tgctcctaaa tcatttcatg ttgaagatac cccagtttgt 6060
ttctcaagaa acagttctct cagttctctt agtattgact ctgaagatga cctgttgcag 6120
gaatgtataa gctccgcaat gccaaaaaag aaaaagcctt caagactcaa gggtgataat 6180
gaaaaacata gtcccagaaa tatgggtggc atattaggtg aagatctgac acttgatttg 6240
aaagatatac agagaccaga ttcagaacat ggtctatccc ctgattcaga aaattttgat 6300
tggaaagcta ttcaggaagg tgcaaattcc atagtaagta gtttacatca agctgctgct 6360
gctgcatgtt tatctagaca agcttcgtct gattcagatt ccatcctttc cctgaaatca 6420
ggaatctctc tgggatcacc atttcatctt acacctgatc aagaagaaaa acccttttaca 6480
agtaataaag gcccacgaat tctaaaacca ggggagaaaa gtacattgga aactaaaaag 6540
atagaatctg aaagtaaagg aatcaaagga ggaaaaaaag tttataaaag tttgattact 6600
ggaaaagttc gatctaattc agaaatttca ggccaaatga aacagcccct tcaagcaaac 6660
atgccttcaa tctctcgagg caggacaatg attcatattc caggagttcg aaatagctcc 6720
tcaagtacaa gtcctgtttc taaaaaaggc ccacccctta agactccagc ctccaaaagc 6780
cctagtgaag gtcaaacagc caccacttct cctagaggag ccaagccatc tgtgaaatca 6840
gaattaagcc ctgttgccag gcagacatcc caaataggtg ggtcaagtaa agcaccttct 6900
agatcaggat ctagagattc gacccccttca agacctgccc agcaaccatt aagtagacct 6960
atacagtctc ctggccgaaa ctcaatttcc cctggtagaa atggaataag tcctcctaac 7020
aaattatctc aacttccaag gacatcatcc cctagtactg cttcaactaa gtcctcaggt 7080
tctggaaaaa tgtcatatac atctccaggt agacagatga gccaacagaa ccttaccaaa 7140
caaacaggtt tatccaagaa tgccagtagt attccaagaa gtgagtctgc ctccaaagga 7200

```
ctaaatcaga tgaataatgg taatggagcc aataaaaagg tagaactttc tagaatgtct 7260
tcaactaaat caagtggaag tgaatctgat agatcagaaa gacctgtatt agtacgccag 7320
tcaactttca tcaaagaagc tccaagccca accttaagaa gaaaattgga ggaatctgct 7380
tcatttgaat ctctttctcc atcatctaga ccagcttctc ccactaggtc ccaggcacaa 7440
actccagttt taagtccttc ccttcctgat atgtctctat ccacacattc gtctgttcag 7500
gctggtggat ggcgaaaact cccacctaat ctcagtccca ctatagagta taatgatgga 7560
agaccagcaa agcgccatga tattgcacgg tctcattctg aaagtccttc tagacttcca 7620
atcaataggt caggaacctg gaaacgtgag cacagcaaac attcatcatc ccttcctcga 7680
gtaagcactt ggagaagaac tggaagttca tcttcaattc tttctgcttc atcagaatcc 7740
agtgaaaaag caaaaagtga ggatgaaaaa catgtgaact ctatttcagg aaccaaacaa 7800
agtaaagaaa accaagtatc cgcaaaagga acatggagaa aaataaaaga aaatgaattt 7860
tctcccacaa atagtacttc tcagaccgtt tcctcaggtg ctacaaatgg tgctgaatca 7920
aagactctaa tttatcaaat ggcacctgct gtttctaaaa cagaggatgt ttgggtgaga 7980
attgaggact gtcccattaa caatcctaga tctggaagat ctcccacagg taatactccc 8040
ccggtgattg acagtgtttc agaaaaggca aatccaaaca ttaaagattc aaaagataat 8100
caggcaaaac aaaatgtggg taatggcagt gttcccatgc gtaccgtggg tttggaaaat 8160
cgcctgaact cctttattca ggtggatgcc cctgaccaaa aaggaactga gataaaacca 8220
ggacaaaata atcctgtccc tgtatcagag actaatgaaa gttctatagt ggaacgtacc 8280
ccattcagtt ctagcagctc aagcaaacac agttcaccta gtgggactgt tgctgccaga 8340
gtgactcctt ttaattacaa cccaagccct aggaaaagca gcgcagatag cacttcagct 8400
cggccatctc agatcccaac tccagtgaat aacaacacaa agaagcgaga ttccaaaact 8460
gacagcacag aatccagtgg aacccaaagt cctaagcgcc attctgggtc ttaccttgtg 8520
acatctgttt aa                                                    8532
```

<210> 26
<211> 2843
<212> PRT
<213> *Homo sapiens*

<400> 26

```
Met Ala Ala Ala Ser Tyr Asp Gln Leu Leu Lys Gln Val Glu Ala Leu
 1               5                  10                  15
Lys Met Glu Asn Ser Asn Leu Arg Gln Glu Leu Glu Asp Asn Ser Asn
            20                  25                  30
His Leu Thr Lys Leu Glu Thr Glu Ala Ser Asn Met Lys Glu Val Leu
        35                  40                  45
Lys Gln Leu Gln Gly Ser Ile Glu Asp Glu Ala Met Ala Ser Ser Gly
    50                  55                  60
Gln Ile Asp Leu Leu Glu Arg Leu Lys Glu Leu Asn Leu Asp Ser Ser
65                  70                  75                  80
Asn Phe Pro Gly Val Lys Leu Arg Ser Lys Met Ser Leu Arg Ser Tyr
                85                  90                  95
Gly Ser Arg Glu Gly Ser Val Ser Ser Arg Ser Gly Glu Cys Ser Pro
            100                 105                 110
Val Pro Met Gly Ser Phe Pro Arg Arg Gly Phe Val Asn Gly Ser Arg
        115                 120                 125
Glu Ser Thr Gly Tyr Leu Glu Glu Leu Glu Lys Glu Arg Ser Leu Leu
    130                 135                 140
Leu Ala Asp Leu Asp Lys Glu Glu Lys Glu Lys Asp Trp Tyr Tyr Ala
145                 150                 155                 160
Gln Leu Gln Asn Leu Thr Lys Arg Ile Asp Ser Leu Pro Leu Thr Glu
                165                 170                 175
Asn Phe Ser Leu Gln Thr Asp Met Thr Arg Arg Gln Leu Glu Tyr Glu
            180                 185                 190
Ala Arg Gln Ile Arg Val Ala Met Glu Glu Gln Leu Gly Thr Cys Gln
        195                 200                 205
Asp Met Glu Lys Arg Ala Gln Arg Arg Ile Ala Arg Ile Gln Gln Ile
    210                 215                 220
Glu Lys Asp Ile Leu Arg Ile Arg Gln Leu Leu Gln Ser Gln Ala Thr
225                 230                 235                 240
Glu Ala Glu Arg Ser Ser Gln Asn Lys His Glu Thr Gly Ser His Asp
                245                 250                 255
Ala Glu Arg Gln Asn Glu Gly Gln Gly Val Gly Glu Ile Asn Met Ala
```

```
            260                 265                 270
Thr Ser Gly Asn Gly Gln Gly Ser Thr Thr Arg Met Asp His Glu Thr
        275                 280                 285
Ala Ser Val Leu Ser Ser Ser Ser Thr His Ser Ala Pro Arg Arg Leu
    290                 295                 300
Thr Ser His Leu Gly Thr Lys Val Glu Met Val Tyr Ser Leu Leu Ser
305                 310                 315                 320
Met Leu Gly Thr His Asp Lys Asp Asp Met Ser Arg Thr Leu Leu Ala
                325                 330                 335
Met Ser Ser Ser Gln Asp Ser Cys Ile Ser Met Arg Gln Ser Gly Cys
            340                 345                 350
Leu Pro Leu Leu Ile Gln Leu Leu His Gly Asn Asp Lys Asp Ser Val
        355                 360                 365
Leu Leu Gly Asn Ser Arg Gly Ser Lys Glu Ala Arg Ala Arg Ala Ser
    370                 375                 380
Ala Ala Leu His Asn Ile Ile His Ser Gln Pro Asp Asp Lys Arg Gly
385                 390                 395                 400
Arg Arg Glu Ile Arg Val Leu His Leu Leu Glu Gln Ile Arg Ala Tyr
                405                 410                 415
Cys Glu Thr Cys Trp Glu Trp Gln Glu Ala His Glu Pro Gly Met Asp
            420                 425                 430
Gln Asp Lys Asn Pro Met Pro Ala Pro Val Glu His Gln Ile Cys Pro
        435                 440                 445
Ala Val Cys Val Leu Met Lys Leu Ser Phe Asp Glu Glu His Arg His
    450                 455                 460
Ala Met Asn Glu Leu Gly Gly Leu Gln Ala Ile Ala Glu Leu Leu Gln
465                 470                 475                 480
Val Asp Cys Glu Met Tyr Gly Leu Thr Asn Asp His Tyr Ser Ile Thr
                485                 490                 495
Leu Arg Arg Tyr Ala Gly Met Ala Leu Thr Asn Leu Thr Phe Gly Asp
            500                 505                 510
Val Ala Asn Lys Ala Thr Leu Cys Ser Met Lys Gly Cys Met Arg Ala
        515                 520                 525
Leu Val Ala Gln Leu Lys Ser Glu Ser Glu Asp Leu Gln Gln Val Ile
    530                 535                 540
Ala Ser Val Leu Arg Asn Leu Ser Trp Arg Ala Asp Val Asn Ser Lys
545                 550                 555                 560
Lys Thr Leu Arg Glu Val Gly Ser Val Lys Ala Leu Met Glu Cys Ala
                565                 570                 575
Leu Glu Val Lys Lys Glu Ser Thr Leu Lys Ser Val Leu Ser Ala Leu
            580                 585                 590
Trp Asn Leu Ser Ala His Cys Thr Glu Asn Lys Ala Asp Ile Cys Ala
        595                 600                 605
Val Asp Gly Ala Leu Ala Phe Leu Val Gly Thr Leu Thr Tyr Arg Ser
    610                 615                 620
Gln Thr Asn Thr Leu Ala Ile Ile Glu Ser Gly Gly Gly Ile Leu Arg
625                 630                 635                 640
Asn Val Ser Ser Leu Ile Ala Thr Asn Glu Asp His Arg Gln Ile Leu
                645                 650                 655
Arg Glu Asn Asn Cys Leu Gln Thr Leu Leu Gln His Leu Lys Ser His
            660                 665                 670
Ser Leu Thr Ile Val Ser Asn Ala Cys Gly Thr Leu Trp Asn Leu Ser
        675                 680                 685
Ala Arg Asn Pro Lys Asp Gln Glu Ala Leu Trp Asp Met Gly Ala Val
    690                 695                 700
Ser Met Leu Lys Asn Leu Ile His Ser Lys His Lys Met Ile Ala Met
705                 710                 715                 720
Gly Ser Ala Ala Ala Leu Arg Asn Leu Met Ala Asn Arg Pro Ala Lys
                725                 730                 735
Tyr Lys Asp Ala Asn Ile Met Ser Pro Gly Ser Ser Leu Pro Ser Leu
            740                 745                 750
His Val Arg Lys Gln Lys Ala Leu Glu Ala Glu Leu Asp Ala Gln His
        755                 760                 765
```

```
Leu Ser Glu Thr Phe Asp Asn Ile Asp Asn Leu Ser Pro Lys Ala Ser
    770                 775                 780
His Arg Ser Lys Gln Arg His Lys Gln Ser Leu Tyr Gly Asp Tyr Val
785                 790                 795                 800
Phe Asp Thr Asn Arg His Asp Asp Asn Arg Ser Asp Asn Phe Asn Thr
                805                 810                 815
Gly Asn Met Thr Val Leu Ser Pro Tyr Leu Asn Thr Thr Val Leu Pro
            820                 825                 830
Ser Ser Ser Ser Ser Arg Gly Ser Leu Asp Ser Ser Arg Ser Glu Lys
        835                 840                 845
Asp Arg Ser Leu Glu Arg Glu Arg Gly Ile Gly Leu Gly Asn Tyr His
    850                 855                 860
Pro Ala Thr Glu Asn Pro Gly Thr Ser Ser Lys Arg Gly Leu Gln Ile
865                 870                 875                 880
Ser Thr Thr Ala Ala Gln Ile Ala Lys Val Met Glu Glu Val Ser Ala
                885                 890                 895
Ile His Thr Ser Gln Glu Asp Arg Ser Ser Gly Ser Thr Thr Glu Leu
            900                 905                 910
His Cys Val Thr Asp Glu Arg Asn Ala Leu Arg Arg Ser Ser Ala Ala
        915                 920                 925
His Thr His Ser Asn Thr Tyr Asn Phe Thr Lys Ser Glu Asn Ser Asn
    930                 935                 940
Arg Thr Cys Ser Met Pro Tyr Ala Lys Leu Glu Tyr Lys Arg Ser Ser
945                 950                 955                 960
Asn Asp Ser Leu Asn Ser Val Ser Ser Ser Asp Gly Tyr Gly Lys Arg
                965                 970                 975
Gly Gln Met Lys Pro Ser Ile Glu Ser Tyr Ser Glu Asp Asp Glu Ser
            980                 985                 990
Lys Phe Cys Ser Tyr Gly Gln Tyr Pro Ala Asp Leu Ala His Lys Ile
        995                 1000                1005
His Ser Ala Asn His Met Asp Asp Asn Asp Gly Glu Leu Asp Thr Pro
    1010                1015                1020
Ile Asn Tyr Ser Leu Lys Tyr Ser Asp Glu Gln Leu Asn Ser Gly Arg
1025                1030                1035                1040
Gln Ser Pro Ser Gln Asn Glu Arg Trp Ala Arg Pro Lys His Ile Ile
                1045                1050                1055
Glu Asp Glu Ile Lys Gln Ser Glu Gln Arg Gln Ser Arg Asn Gln Ser
            1060                1065                1070
Thr Thr Tyr Pro Val Tyr Thr Glu Ser Thr Asp Asp Lys His Leu Lys
        1075                1080                1085
Phe Gln Pro His Phe Gly Gln Gln Glu Cys Val Ser Pro Tyr Arg Ser
    1090                1095                1100
Arg Gly Ala Asn Gly Ser Glu Thr Asn Arg Val Gly Ser Asn His Gly
1105                1110                1115                1120
Ile Asn Gln Asn Val Ser Gln Ser Leu Cys Gln Glu Asp Asp Tyr Glu
                1125                1130                1135
Asp Asp Lys Pro Thr Asn Tyr Ser Glu Arg Tyr Ser Glu Glu Glu Gln
            1140                1145                1150
His Glu Glu Glu Glu Arg Pro Thr Asn Tyr Ser Ile Lys Tyr Asn Glu
        1155                1160                1165
Glu Lys Arg His Val Asp Gln Pro Ile Asp Tyr Ser Leu Lys Tyr Ala
    1170                1175                1180
Thr Asp Ile Pro Ser Ser Gln Lys Gln Ser Phe Ser Phe Ser Lys Ser
1185                1190                1195                1200
Ser Ser Gly Gln Ser Ser Lys Thr Glu His Met Ser Ser Ser Ser Glu
                1205                1210                1215
Asn Thr Ser Thr Pro Ser Ser Asn Ala Lys Arg Gln Asn Gln Leu His
            1220                1225                1230
Pro Ser Ser Ala Gln Ser Arg Ser Gly Gln Pro Gln Lys Ala Ala Thr
        1235                1240                1245
Cys Lys Val Ser Ser Ile Asn Gln Glu Thr Ile Gln Thr Tyr Cys Val
    1250                1255                1260
Glu Asp Thr Pro Ile Cys Phe Ser Arg Cys Ser Ser Leu Ser Ser Leu
```

1265 1270 1275 1280
Ser Ser Ala Glu Asp Glu Ile Gly Cys Asn Gln Thr Thr Gln Glu Ala
1285 1290 1295
Asp Ser Ala Asn Thr Leu Gln Ile Ala Glu Ile Lys Glu Lys Ile Gly
1300 1305 1310
Thr Arg Ser Ala Glu Asp Pro Val Ser Glu Val Pro Ala Val Ser Gln
1315 1320 1325
His Pro Arg Thr Lys Ser Ser Arg Leu Gln Gly Ser Ser Leu Ser Ser
1330 1335 1340
Glu Ser Ala Arg His Lys Ala Val Glu Phe Ser Ser Gly Ala Lys Ser
1345 1350 1355 1360
Pro Ser Lys Ser Gly Ala Gln Thr Pro Lys Ser Pro Pro Glu His Tyr
1365 1370 1375
Val Gln Glu Thr Pro Leu Met Phe Ser Arg Cys Thr Ser Val Ser Ser
1380 1385 1390
Leu Asp Ser Phe Glu Ser Arg Ser Ile Ala Ser Ser Val Gln Ser Glu
1395 1400 1405
Pro Cys Ser Gly Met Val Ser Gly Ile Ile Ser Pro Ser Asp Leu Pro
1410 1415 1420
Asp Ser Pro Gly Gln Thr Met Pro Pro Ser Arg Ser Lys Thr Pro Pro
1425 1430 1435 1440
Pro Pro Pro Gln Thr Ala Gln Thr Lys Arg Glu Val Pro Lys Asn Lys
1445 1450 1455
Ala Pro Thr Ala Glu Lys Arg Glu Ser Gly Pro Lys Gln Ala Ala Val
1460 1465 1470
Asn Ala Ala Val Gln Arg Val Gln Val Leu Pro Asp Ala Asp Thr Leu
1475 1480 1485
Leu His Phe Ala Thr Glu Ser Thr Pro Asp Gly Phe Ser Cys Ser Ser
1490 1495 1500
Ser Leu Ser Ala Leu Ser Leu Asp Glu Pro Phe Ile Gln Lys Asp Val
1505 1510 1515 1520
Glu Leu Arg Ile Met Pro Pro Val Gln Glu Asn Asp Asn Gly Asn Glu
1525 1530 1535
Thr Glu Ser Glu Gln Pro Lys Glu Ser Asn Glu Asn Gln Glu Lys Glu
1540 1545 1550
Ala Glu Lys Thr Ile Asp Ser Glu Lys Asp Leu Leu Asp Asp Ser Asp
1555 1560 1565
Asp Asp Asp Ile Glu Ile Leu Glu Glu Cys Ile Ile Ser Ala Met Pro
1570 1575 1580
Thr Lys Ser Ser Arg Lys Ala Lys Lys Pro Ala Gln Thr Ala Ser Lys
1585 1590 1595 1600
Leu Pro Pro Pro Val Ala Arg Lys Pro Ser Gln Leu Pro Val Tyr Lys
1605 1610 1615
Leu Leu Pro Ser Gln Asn Arg Leu Gln Pro Gln Lys His Val Ser Phe
1620 1625 1630
Thr Pro Gly Asp Asp Met Pro Arg Val Tyr Cys Val Glu Gly Thr Pro
1635 1640 1645
Ile Asn Phe Ser Thr Ala Thr Ser Leu Ser Asp Leu Thr Ile Glu Ser
1650 1655 1660
Pro Pro Asn Glu Leu Ala Ala Gly Glu Gly Val Arg Gly Gly Ala Gln
1665 1670 1675 1680
Ser Gly Glu Phe Glu Lys Arg Asp Thr Ile Pro Thr Glu Gly Arg Ser
1685 1690 1695
Thr Asp Glu Ala Gln Gly Gly Lys Thr Ser Ser Val Thr Ile Pro Glu
1700 1705 1710
Leu Asp Asp Asn Lys Ala Glu Glu Gly Asp Ile Leu Ala Glu Cys Ile
1715 1720 1725
Asn Ser Ala Met Pro Lys Gly Lys Ser His Lys Pro Phe Arg Val Lys
1730 1735 1740
Lys Ile Met Asp Gln Val Gln Gln Ala Ser Ala Ser Ser Ser Ala Pro
1745 1750 1755 1760
Asn Lys Asn Gln Leu Asp Gly Lys Lys Lys Lys Pro Thr Ser Pro Val
1765 1770 1775

```
Lys Pro Ile Pro Gln Asn Thr Glu Tyr Arg Thr Arg Val Arg Lys Asn
            1780                1785                1790
Ala Asp Ser Lys Asn Asn Leu Asn Ala Glu Arg Val Phe Ser Asp Asn
        1795                1800                1805
Lys Asp Ser Lys Lys Gln Asn Leu Lys Asn Asn Ser Lys Val Phe Asn
    1810                1815                1820
Asp Lys Leu Pro Asn Asn Glu Asp Arg Val Arg Gly Ser Phe Ala Phe
1825                1830                1835                1840
Asp Ser Pro His His Tyr Thr Pro Ile Glu Gly Thr Pro Tyr Cys Phe
                1845                1850                1855
Ser Arg Asn Asp Ser Leu Ser Ser Leu Asp Phe Asp Asp Asp Asp Val
            1860                1865                1870
Asp Leu Ser Arg Glu Lys Ala Glu Leu Arg Lys Ala Lys Glu Asn Lys
        1875                1880                1885
Glu Ser Glu Ala Lys Val Thr Ser His Thr Glu Leu Thr Ser Asn Gln
    1890                1895                1900
Gln Ser Ala Asn Lys Thr Gln Ala Ile Ala Lys Gln Pro Ile Asn Arg
1905                1910                1915                1920
Gly Gln Pro Lys Pro Ile Leu Gln Lys Gln Ser Thr Phe Pro Gln Ser
                1925                1930                1935
Ser Lys Asp Ile Pro Asp Arg Gly Ala Ala Thr Asp Glu Lys Leu Gln
            1940                1945                1950
Asn Phe Ala Ile Glu Asn Thr Pro Val Cys Phe Ser His Asn Ser Ser
        1955                1960                1965
Leu Ser Ser Leu Ser Asp Ile Asp Gln Glu Asn Asn Asn Lys Glu Asn
    1970                1975                1980
Glu Pro Ile Lys Glu Thr Glu Pro Pro Asp Ser Gln Gly Glu Pro Ser
1985                1990                1995                2000
Lys Pro Gln Ala Ser Gly Tyr Ala Pro Lys Ser Phe His Val Glu Asp
                2005                2010                2015
Thr Pro Val Cys Phe Ser Arg Asn Ser Ser Leu Ser Ser Leu Ser Ile
            2020                2025                2030
Asp Ser Glu Asp Asp Leu Leu Gln Glu Cys Ile Ser Ser Ala Met Pro
        2035                2040                2045
Lys Lys Lys Lys Pro Ser Arg Leu Lys Gly Asp Asn Glu Lys His Ser
    2050                2055                2060
Pro Arg Asn Met Gly Gly Ile Leu Gly Glu Asp Leu Thr Leu Asp Leu
2065                2070                2075                2080
Lys Asp Ile Gln Arg Pro Asp Ser Glu His Gly Leu Ser Pro Asp Ser
                2085                2090                2095
Glu Asn Phe Asp Trp Lys Ala Ile Gln Glu Gly Ala Asn Ser Ile Val
            2100                2105                2110
Ser Ser Leu His Gln Ala Ala Ala Ala Ala Cys Leu Ser Arg Gln Ala
        2115                2120                2125
Ser Ser Asp Ser Asp Ser Ile Leu Ser Leu Lys Ser Gly Ile Ser Leu
    2130                2135                2140
Gly Ser Pro Phe His Leu Thr Pro Asp Gln Glu Glu Lys Pro Phe Thr
2145                2150                2155                2160
Ser Asn Lys Gly Pro Arg Ile Leu Lys Pro Gly Glu Lys Ser Thr Leu
                2165                2170                2175
Glu Thr Lys Lys Ile Glu Ser Glu Ser Lys Gly Ile Lys Gly Gly Lys
            2180                2185                2190
Lys Val Tyr Lys Ser Leu Ile Thr Gly Lys Val Arg Ser Asn Ser Glu
        2195                2200                2205
Ile Ser Gly Gln Met Lys Gln Pro Leu Gln Ala Asn Met Pro Ser Ile
    2210                2215                2220
Ser Arg Gly Arg Thr Met Ile His Ile Pro Gly Val Arg Asn Ser Ser
2225                2230                2235                2240
Ser Ser Thr Ser Pro Val Ser Lys Lys Gly Pro Pro Leu Lys Thr Pro
                2245                2250                2255
Ala Ser Lys Ser Pro Ser Glu Gly Gln Thr Ala Thr Thr Ser Pro Arg
            2260                2265                2270
Gly Ala Lys Pro Ser Val Lys Ser Glu Leu Ser Pro Val Ala Arg Gln
```

```
        2275                2280                2285
Thr Ser Gln Ile Gly Gly Ser Ser Lys Ala Pro Ser Arg Ser Gly Ser
    2290                2295                2300
Arg Asp Ser Thr Pro Ser Arg Pro Ala Gln Gln Pro Leu Ser Arg Pro
2305                2310                2315                2320
Ile Gln Ser Pro Gly Arg Asn Ser Ile Ser Pro Gly Arg Asn Gly Ile
                2325                2330                2335
Ser Pro Pro Asn Lys Leu Ser Gln Leu Pro Arg Thr Ser Ser Pro Ser
            2340                2345                2350
Thr Ala Ser Thr Lys Ser Ser Gly Ser Gly Lys Met Ser Tyr Thr Ser
        2355                2360                2365
Pro Gly Arg Gln Met Ser Gln Gln Asn Leu Thr Lys Gln Thr Gly Leu
    2370                2375                2380
Ser Lys Asn Ala Ser Ser Ile Pro Arg Ser Glu Ser Ala Ser Lys Gly
2385                2390                2395                2400
Leu Asn Gln Met Asn Asn Gly Asn Gly Ala Asn Lys Lys Val Glu Leu
                2405                2410                2415
Ser Arg Met Ser Ser Thr Lys Ser Ser Gly Ser Glu Ser Asp Arg Ser
            2420                2425                2430
Glu Arg Pro Val Leu Val Arg Gln Ser Thr Phe Ile Lys Glu Ala Pro
        2435                2440                2445
Ser Pro Thr Leu Arg Arg Lys Leu Glu Glu Ser Ala Ser Phe Glu Ser
    2450                2455                2460
Leu Ser Pro Ser Ser Arg Pro Ala Ser Pro Thr Arg Ser Gln Ala Gln
2465                2470                2475                2480
Thr Pro Val Leu Ser Pro Ser Leu Pro Asp Met Ser Leu Ser Thr His
                2485                2490                2495
Ser Ser Val Gln Ala Gly Gly Trp Arg Lys Leu Pro Pro Asn Leu Ser
            2500                2505                2510
Pro Thr Ile Glu Tyr Asn Asp Gly Arg Pro Ala Lys Arg His Asp Ile
        2515                2520                2525
Ala Arg Ser His Ser Glu Ser Pro Ser Arg Leu Pro Ile Asn Arg Ser
    2530                2535                2540
Gly Thr Trp Lys Arg Glu His Ser Lys His Ser Ser Ser Leu Pro Arg
2545                2550                2555                2560
Val Ser Thr Trp Arg Arg Thr Gly Ser Ser Ser Ser Ile Leu Ser Ala
                2565                2570                2575
Ser Ser Glu Ser Ser Glu Lys Ala Lys Ser Glu Asp Glu Lys His Val
            2580                2585                2590
Asn Ser Ile Ser Gly Thr Lys Gln Ser Lys Glu Asn Gln Val Ser Ala
        2595                2600                2605
Lys Gly Thr Trp Arg Lys Ile Lys Glu Asn Glu Phe Ser Pro Thr Asn
    2610                2615                2620
Ser Thr Ser Gln Thr Val Ser Ser Gly Ala Thr Asn Gly Ala Glu Ser
2625                2630                2635                2640
Lys Thr Leu Ile Tyr Gln Met Ala Pro Ala Val Ser Lys Thr Glu Asp
                2645                2650                2655
Val Trp Val Arg Ile Glu Asp Cys Pro Ile Asn Asn Pro Arg Ser Gly
            2660                2665                2670
Arg Ser Pro Thr Gly Asn Thr Pro Pro Val Ile Asp Ser Val Ser Glu
        2675                2680                2685
Lys Ala Asn Pro Asn Ile Lys Asp Ser Lys Asp Asn Gln Ala Lys Gln
    2690                2695                2700
Asn Val Gly Asn Gly Ser Val Pro Met Arg Thr Val Gly Leu Glu Asn
2705                2710                2715                2720
Arg Leu Asn Ser Phe Ile Gln Val Asp Ala Pro Asp Gln Lys Gly Thr
                2725                2730                2735
Glu Ile Lys Pro Gly Gln Asn Asn Pro Val Pro Val Ser Glu Thr Asn
            2740                2745                2750
Glu Ser Ser Ile Val Glu Arg Thr Pro Phe Ser Ser Ser Ser Ser Ser
        2755                2760                2765
Lys His Ser Ser Pro Ser Gly Thr Val Ala Ala Arg Val Thr Pro Phe
    2770                2775                2780
```

```
Asn Tyr Asn Pro Ser Pro Arg Lys Ser Ser Ala Asp Ser Thr Ser Ala
2785                2790                2795                2800
Arg Pro Ser Gln Ile Pro Thr Pro Val Asn Asn Asn Thr Lys Lys Arg
                2805                2810                2815
Asp Ser Lys Thr Asp Ser Thr Glu Ser Ser Gly Thr Gln Ser Pro Lys
            2820                2825                2830
Arg His Ser Gly Ser Tyr Leu Val Thr Ser Val
        2835                2840
```

REIVINDICACIONES

1. Un método para identificar a un paciente con cáncer para tratamiento con un inhibidor de la enzima activadora de NEDD8 (NAE) que comprende:

a) determinar un estado mutacional de al menos un gen marcador en una muestra que comprende células tumorales obtenidas del paciente con cáncer, en donde el al menos un gen marcador es un supresor tumoral en una ruta de ligasa culina-RING;
b) determinar si el estado mutacional es indicativo de un resultado favorable del tratamiento con el inhibidor de NAE, en donde el inhibidor de NAE es sulfamato de ((1S,2S,4R)-4-{4-[(1S)-2,3-dihidro-1H-inden-1-ilamino]-7H-pirrolo[2,3-d]pirimidin-7-il}-2-hidroxiciclopentil)metilo; y
c) identificar al paciente para tratamiento con el inhibidor de NAE si el estado mutacional indica un resultado favorable de tratamiento con el inhibidor de NAE;

en donde el al menos un gen marcador se selecciona del grupo que consiste en *TP53, CDKN2A, CDKN2A_p14* y *APC.*

2. El método de la reivindicación 1, en donde el estado mutacional del al menos un gen marcador es mutante.

3. El método de la reivindicación 1, en donde el estado mutacional del al menos un gen marcador es de tipo silvestre.

4. El método de la reivindicación 2, en donde el mutante es un mutante inactivador.

5. El método de una cualquiera de las reivindicaciones precedentes, en donde la muestra comprende células tumorales hematológicas o células tumorales sólidas.

6. El método de la reivindicación 5, en donde la muestra que comprende células tumorales hematológicas es sangre.

7. El método de una cualquiera de las reivindicaciones 1-6, en donde el al menos un gen marcador es al menos dos genes marcadores.

8. El método de la reivindicación 2, en donde el al menos un gen marcador es CDKN2A o CDKN2A_p14.

9. El método de la reivindicación 3, en donde el al menos un gen marcador es TP53.

10. El método de la reivindicación 7, en donde los al menos dos genes marcadores son al menos TP53 y APC.

11. El método de una cualquiera de las reivindicaciones precedentes, en donde el estado mutacional se determina midiendo al menos una característica de al menos un marcador asociado con el al menos un gen marcador.

12. El método de la reivindicación 11, en donde la al menos una característica se selecciona del grupo que consiste en tamaño, secuencia, composición y cantidad.

13. El método de la reivindicación 11 o 12, en donde el al menos un marcador se selecciona del grupo que consiste en un locus cromosómico, un ácido nucleico y una proteína asociada con el al menos un gen marcador.

14. El método de la reivindicación 13, en donde el ácido nucleico es ADN genómico, ARNm o ADNc.

15. El método de una cualquiera de las reivindicaciones 11-14, en donde la al menos una característica es una secuencia y el al menos un marcador es un ácido nucleico.

16. El método de una cualquiera de las reivindicaciones 1 a 4, en donde el paciente con cáncer tiene un cáncer hematológico o un cáncer de tumor sólido.

17. El método de la reivindicación 16, en donde el cáncer hematológico es leucemia mielógena aguda o crónica.

18. El método de la reivindicación 16, en donde el cáncer hematológico es síndrome mielodisplásico.

19. El método de la reivindicación 16, en donde el cáncer de tumor sólido es cáncer de cabeza y cuello.

20. El método de una cualquiera de las reivindicaciones precedentes, que comprende adicionalmente:

a) identificar un estado mutacional del al menos un gen marcador en una segunda muestra del paciente, en donde el paciente ha sido tratado con el inhibidor de NAE antes de obtener la segunda muestra;
b) determinar si el estado mutacional es indicativo de un resultado favorable del tratamiento con el inhibidor de

NAE; y
c) determinar la continuación del tratamiento con el inhibidor de NAE si el estado mutacional de la segunda muestra indica un resultado favorable del tratamiento con el inhibidor de NAE.

21. Uso de un kit para identificar a un paciente con cáncer para tratamiento con un inhibidor de la enzima activadora de NEDD8 (NAE), comprendiendo el kit un reactivo para determinar un estado mutacional de al menos un gen marcador en una muestra que comprende células tumorales obtenidas del paciente con cáncer, en donde el inhibidor de NAE es un sulfamato de metilo 1-sustituido, en donde el al menos un gen marcador es un supresor tumoral en una resulta de ligasa culina-RING, y en donde el al menos un gen marcador se selecciona del grupo que consiste en *TP53, CDKN2A, CDKN2A_p14* y *APC.*

22. El uso del kit de la reivindicación 21, en donde el estado mutacional se determina por medición de al menos una característica de al menos un marcador del al menos un gen marcador, en donde el al menos un marcador se selecciona del grupo que consiste en ácido nucleico y proteína asociados con al menos un gen marcador.

23. El uso del kit de una cualquiera de las reivindicaciones 21-22, comprendiendo el kit además un estabilizador para añadir a la muestra.

24. El uso del kit de la reivindicación 22, en donde el al menos un marcador es ácido nucleico y el reactivo es al menos un cebador.

25. El uso del kit de la reivindicación 24, en donde el al menos un cebador se hibrida con una secuencia de ácidos nucleicos seleccionada del grupo que consiste en SEQ ID NOs: 1, 2, 4, 5, 7, 8, 9, 12, 13, 15, 16, 18, 19, 21, 22, 24, 25 o una secuencia en el cromosoma 22q del par de bases 29999545 a 30094589, cromosoma 18q del par de bases 48556583 a 48611412, cromosoma Xp del par de bases 44732423 a 44971847, cromosoma 4q del par de bases 153242410 a 153456172, cromosoma 17p del par de bases 7571720 a 7590868, cromosoma 9p del par de bases 21967751 a 21994490, y un complemento de cualquiera de los anteriores.

26. Un inhibidor de NAE para uso en un método de tratamiento de un paciente con cáncer, comprendiendo el método:

a) utilizar el estado mutacional de al menos un gen marcador en el tumor de un paciente, en donde el al menos un gen marcador se selecciona del grupo que consiste en *TP53, CDKN2A, CDKN2A_p14* y *APC,* seleccionar un paciente con cáncer que se espera que tenga un resultado favorable del tratamiento con el inhibidor de NAE, y
b) tratar al paciente con cáncer con el inhibidor de NAE, en donde el inhibidor de NAE es un sulfamato de metilo 1-sustituido.

27. El inhibidor de NAE para uso según la reivindicación 26, en donde el inhibidor de NAE es sulfamato de ((1S,2S,4R)-4-{4-[(1S)-2,3-dihidro-1H-inden-1-ilamino]-7H-pirrolo[2,3-d]pirimidin-7-il}-2-hidroxiciclopentil)metilo.

28. El inhibidor de NAE para uso de la reivindicación 26 o 27, en donde el paciente con cáncer tiene un cáncer hematológico o un cáncer de tumor sólido.

29. El inhibidor de NAE para uso de la reivindicación 28, en donde el cáncer hematológico es leucemia mielógena aguda o crónica.

30. El inhibidor de NAE para uso de la reivindicación 28, en donde el cáncer hematológico es síndrome mielodisplásico.

31. El inhibidor de NAE para uso de la reivindicación 28, en donde el cáncer de tumor sólido es cáncer de cabeza y cuello.

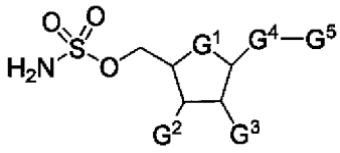


Figura 1

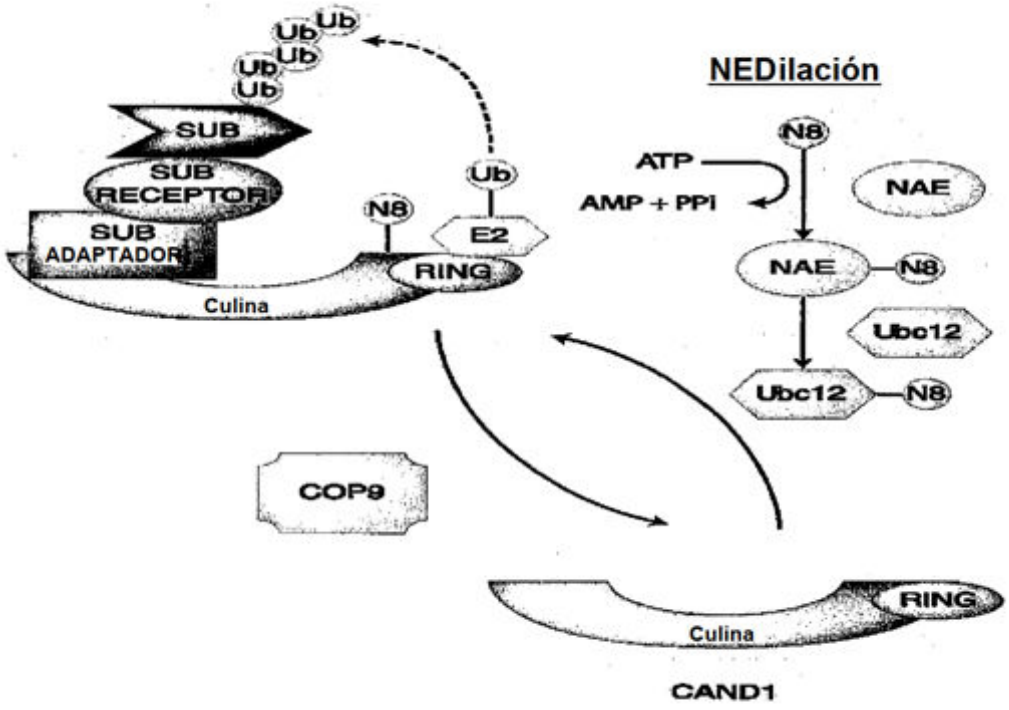

Ub
Ub
Ub
Ub
Ub
SUB
SUB RECEPTOR
SUB ADAPTADOR
Culina
N8
Ub
E2
RING
NEDilación
ATP
AMP + PPI
N8
NAE
NAE
N8
Ubc12
Ubc12
N8
COP9
RING
Culina
CAND1

FIG. 2

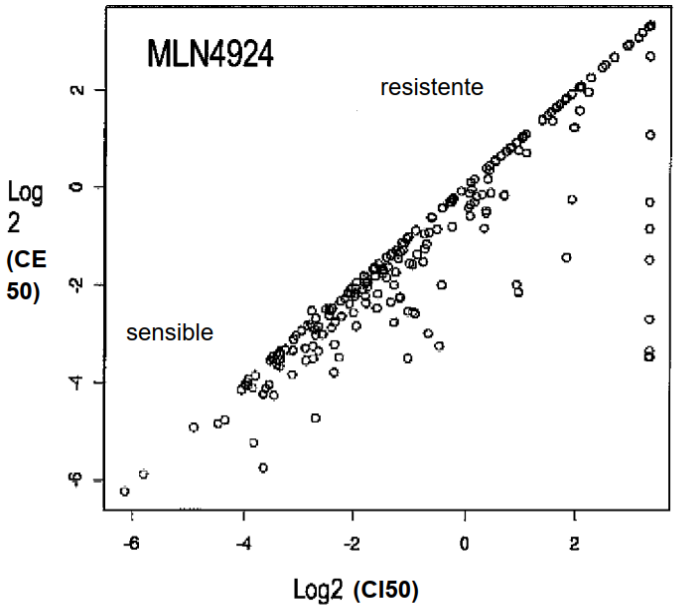

MLN4924
resistente
sensible
Log 2 (CE 50)
Log2 (CI50)
2
0
-2
-4
-6
-6
-4
-2
0
2

FIG. 3

FIG4A.

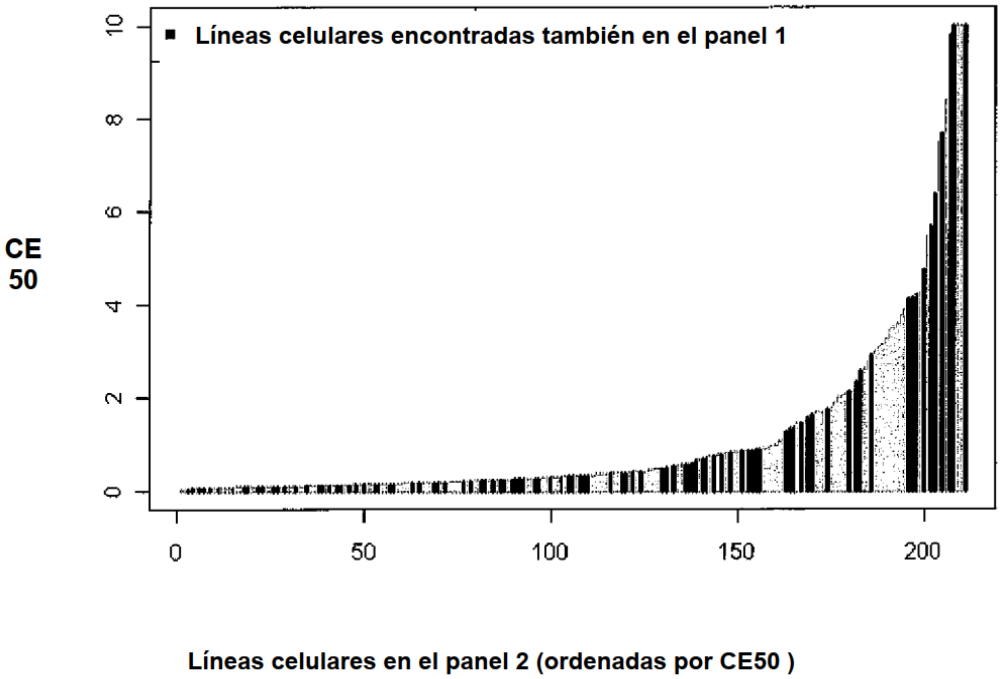


FIG4B.

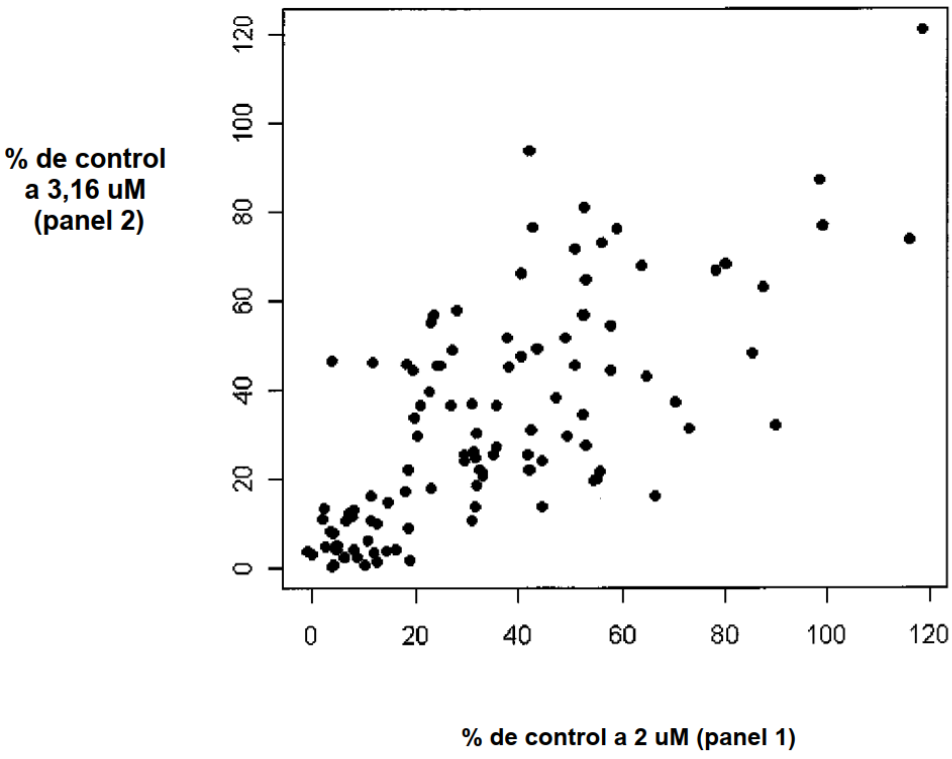


FIG. 4

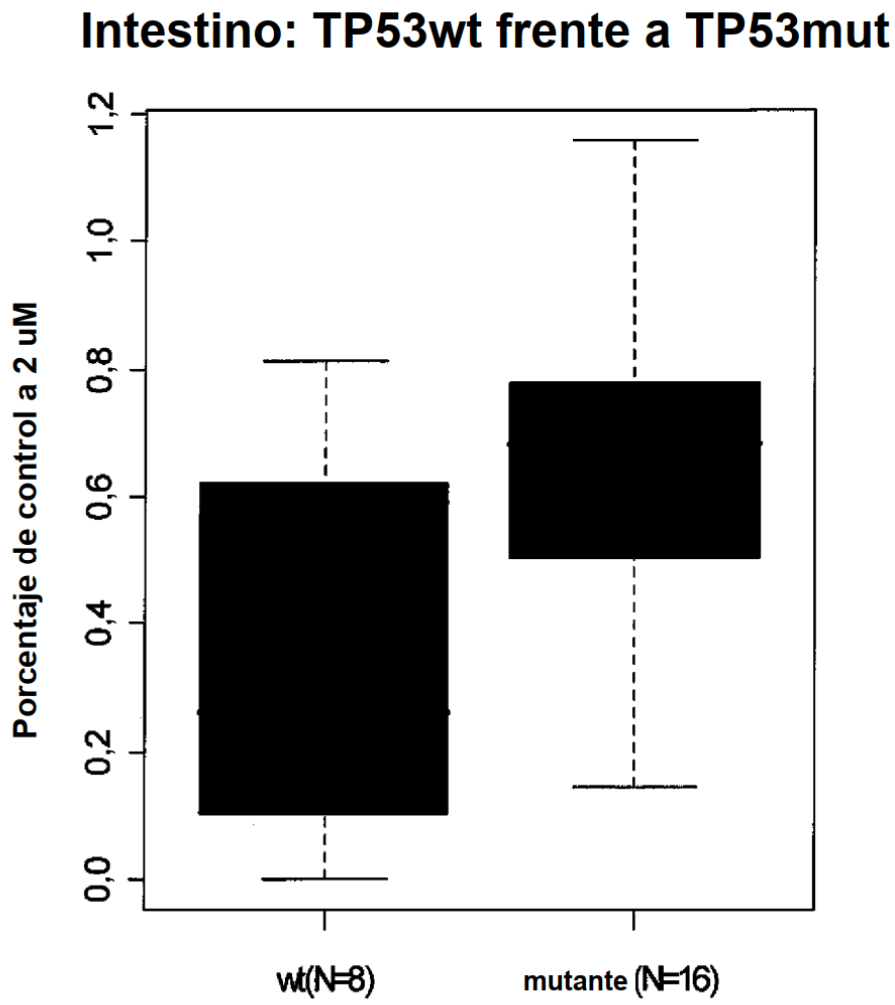

Intestino: TP53wt frente a TP53mut
Porcentaje de control a 2 uM
0,0
0,2
0,4
0,6
0,8
1,0
1,2
wt(N=8)
mutante (N=16)

FIG. 5

FIG. 6A

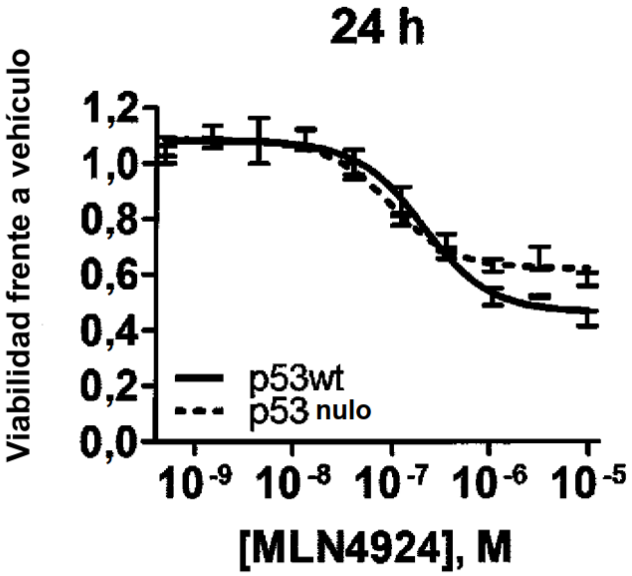

24 h
Viabilidad frente a vehículo
1,2
1,0
0,8
0,6
0,4
0,2
0,0
p53wt
p53 nulo
10-9 10-8 10-7 10-6 10-5
[MLN4924], M

FIG. 6B

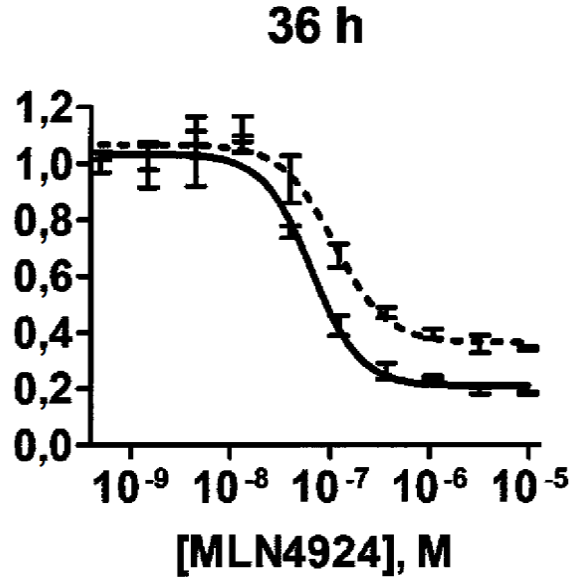

36 h
1,2
1,0
0,8
0,6
0,4
0,2
0,0
10-9 10-8 10-7 10-6 10-5
[MLN4924], M

FIG. 6C

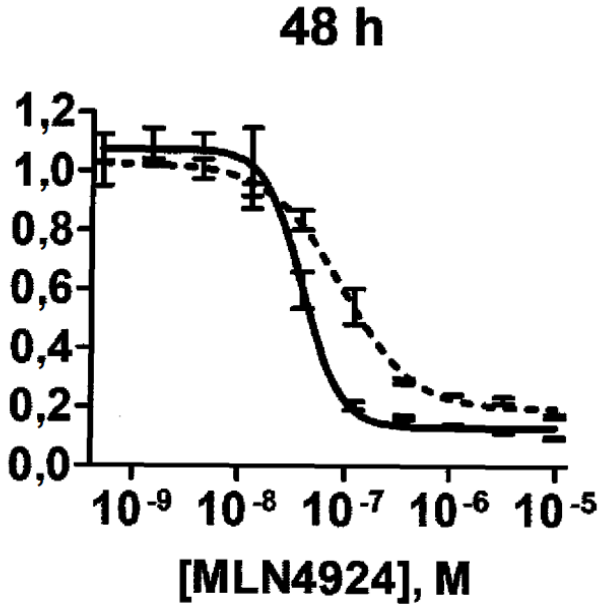
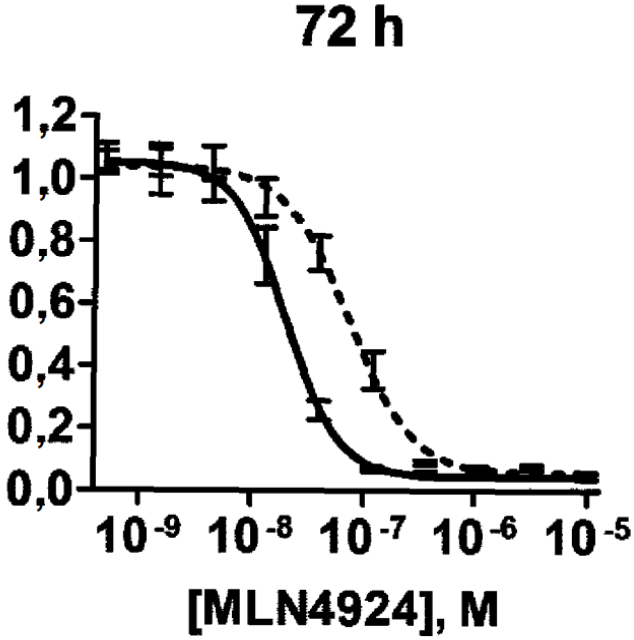

48 h
1,2
1,0
0,8
0,6
0,4
0,2
0,0
10-9 10-8 10-7 10-6 10-5
[MLN4924], M

FIG. 6D


72 h
1,2
1,0
0,8
0,6
0,4
0,2
0,0
10-9 10-8 10-7 10-6 10-5
[MLN4924], M