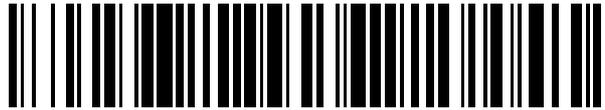


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 395 588**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **07.08.2008 E 08785427 (9)**

97 Fecha y número de publicación de la solicitud europea: **28.04.2010 EP 2179057**

54 Título: **Marcador predictivo para el tratamiento con inhibidor de EGFR**

30 Prioridad:

14.08.2007 EP 07114309

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

13.02.2013

73 Titular/es:

**F. HOFFMANN-LA ROCHE AG (100.0%)
GRENZACHERSTRASSE, 124
4070 BASEL, CH**

72 Inventor/es:

**DELMAR, PAUL;
KLUGHAMMER, BARBARA;
LUTZ, VERENA y
MCLOUGHLIN, PATRICIA**

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 395 588 T3

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

DESCRIPCIÓN

Marcador predictivo para el tratamiento con inhibidor de EGFR

5 La presente invención proporciona un biomarcador que es predictivo para el beneficio clínico del tratamiento con inhibidor de EGFR en pacientes con cáncer

10 Una serie de tumores malignos humanos se asocian con expresión aberrante o sobre-expresión del receptor del factor de crecimiento epidérmico (EGFR). EGF, factor de crecimiento transformante- α (TGF- α), y un número de otros ligandos que se unen al EGFR, estimulando la autofosforilación del dominio tirosina quinasa intracelular del receptor. Posteriormente se activan una serie de rutas intracelulares, y estos eventos corriente abajo resultan en la proliferación de células tumorales in vitro. Se ha postulado que la estimulación de las células tumorales a través de la EGFR pueden ser importante para el crecimiento del tumor y en la supervivencia del tumor in vivo.

15 Los primeros datos clínicos con Tarceva™, un inhibidor de la tirosina quinasa del EGFR, indican que el compuesto es seguro y bien tolerado en dosis que proporcionan la concentración diana efectiva (según se ha determinado en los datos preclínicos). Los ensayos clínicos de fase I y II en pacientes con enfermedad avanzada han demostrado que Tarceva™ ha presentado una actividad clínica prometedora en una serie de tumores epiteliales. En efecto, Tarceva™ ha demostrado que es capaz de inducir remisiones parciales duraderas en pacientes tratados
20 previamente para el cáncer de cabeza y cuello y NSCLC (cáncer de pulmón de células grandes) en una magnitud similar a la quimioterapia establecida de segunda línea, pero con el beneficio adicional de un mejor perfil de seguridad que la quimioterapia y una mayor comodidad (comprimidos en lugar de administración por vía intravenosa [iv]). Un ensayo recientemente finalizado, aleatorizado, doble ciego, controlado con placebo (BR.21) ha demostrado que el agente Tarceva™ por sí solo, prolonga significativamente y mejora la supervivencia de pacientes con NSCLC
25 en los que ha fallado el tratamiento estándar para la enfermedad avanzada.

Tarceva™ (erlotinib) es una molécula química pequeña, es un inhibidor, activo por vía oral, potente y selectivo de la tirosina quinasa de EGFR (EGFR-TKI).

30 El cáncer de pulmón es la principal causa de muerte relacionada con cáncer en América del Norte y Europa. En los Estados Unidos, el número de muertes secundarias al cáncer de pulmón supera las muertes totales combinadas de la segunda (colon), tercera (mama) y cuarta (próstata) de las causas principales de muerte por cáncer. Alrededor del 75% al 80% de todos los cánceres de pulmón son NSCLC, aproximadamente el 40% de los pacientes con enfermedad localmente avanzada y / o inextirpable. Este grupo suele incluir aquellos estadíos con masa IIIA y IIIB
35 de la enfermedad, con exclusión de los derrames pleurales malignos.

40 La incidencia bruta de cáncer de pulmón en la Unión Europea es del 52,5, la tasa de mortalidad de 48,7 casos/100000/año. Entre los hombres las tasas son de 79,3 y 78,3, entre las mujeres de 21,6 y 20,5, respectivamente. El NSCLC representa el 80% de todos los casos de cáncer de pulmón. Alrededor del 90% de la mortalidad por cáncer de pulmón entre los hombres y el 80% entre las mujeres, es atribuible al consumo de tabaco.

45 En los EE.UU., según la Sociedad Americana del Cáncer, en 2004, había aproximadamente 173.800 nuevos casos de cáncer de pulmón (93.100 en varones y 80.700 en mujeres) y representa alrededor del 13% de todos los nuevos casos de cáncer. La mayoría de los pacientes mueren como consecuencia de su enfermedad a los dos años desde el diagnóstico. Para muchos pacientes con NSCLC, el éxito del tratamiento sigue siendo difícil de alcanzar. Los tumores avanzados a menudo no son susceptibles de cirugía y también pueden ser resistentes a las dosis tolerables de la radioterapia y la quimioterapia. En los ensayos aleatorizados, las quimioterapias de combinación actuales más activas alcanzaron tasas de respuesta de aproximadamente 30% a 40% y una tasa de supervivencia a 1 año entre el 35% y el 40%. Esto es realmente un avance sobre la tasa de supervivencia del 10% a 1 año vista con solo
50 tratamiento de apoyo (Shepherd 1999).

Hasta hace poco, las opciones terapéuticas para los pacientes con recaída tras una recaída, se limitaban a la mejor atención de apoyo posible o paliativos. Un estudio reciente que comparó docetaxel (Taxotere) con el mejor tratamiento de apoyo, mostró que los pacientes con NSCLC podrían beneficiarse de la quimioterapia de segunda
55 línea después de que regímenes de primera línea a base de cisplatino hubieran fracasado. Pacientes de todas las edades y con estado funcional ECOG de 0, 1 ó 2 mostraron una mejora en la supervivencia con docetaxel, al igual que los que habían sido refractarios al tratamiento previo a base de platino. Los pacientes que no se beneficiaron de la terapia incluyen aquellos con pérdida de peso del 10%, niveles elevados de lactato deshidrogenasa, participación de múltiples órganos, o afectación hepática. Además, el beneficio de la monoterapia con docetaxel no se extiende más allá de la configuración de segunda línea. Los pacientes tratados con docetaxel como tratamiento de tercera
60 línea o más no mostraron prolongación de la supervivencia. Docetaxel como agente único se convirtió en un estándar de tratamiento de segunda línea para el NSCLC. Recientemente, otro ensayo aleatorizado de fase III para el tratamiento de segunda línea de NSCLC comparó pemetrexed (Alimta®) con docetaxel. El tratamiento con pemetrexed resultó en una eficacia clínicamente equivalente, pero con considerablemente menos efectos secundarios en comparación con docetaxel.
65

Desde hace tiempo se reconoce que existe una necesidad de desarrollar métodos de tratamiento individualizado contra el cáncer. Con el desarrollo de tratamientos dirigidos contra el cáncer, hay un interés particular en metodologías que podría proporcionar un perfil molecular de la diana tumoral, (es decir, aquellos que son predictivos para el beneficio clínico). Ya se han establecido pruebas del principio para perfiles de expresión génica en el cáncer con la clasificación molecular de los tipos de tumores que no son aparentes sobre la base de las pruebas morfológicas e inmunohistoquímicas actuales. Dos entidades patológicas diferentes se han diferenciado con pronósticos diferentes con la clasificación actual única de linfoma difuso de células B grandes utilizando perfiles de expresión génica.

Por lo tanto, es un objetivo de la presente invención proporcionar biomarcadores de expresión que son predictivos para el beneficio clínico del tratamiento con inhibidor de EGFR en pacientes con cáncer. Los siguientes documentos describen la predicción de la respuesta de un NSCLC al tratamiento con un inhibidor de EGFR, en su mayoría sobre la base de perfiles de expresión:

WO 2005/049829;

WO 2004/111273;

KAKIUCHI SOJI ET AL: "Prediction of sensitivity of advanced non-small cell lung cancers to gefitinib (Iressa, ZD1839)", HUMAN MOLECULAR GENETICS, vol. 13, n° 24, 15 Diciembre 2004, páginas 3029-3043;

OKANO TETSUYA ET AL: "Proteomic signature corresponding to the response to gefitinib (Iressa, ZD1839), an epidermal growth factor receptor tyrosine kinase inhibitor in lung adenocarcinoma", CLINICAL CANCER RESEARCH, vol. 13, n° 3, páginas 799-805.

En un primer objeto, la presente invención proporciona un método in vitro para predecir el beneficio clínico de un paciente con cáncer en respuesta al tratamiento con un inhibidor de EGFR que comprende los pasos de: determinar un nivel de expresión de un gen PTP4A1 en una muestra de tumor de un paciente y comparar el nivel de expresión del gen PTP4A1 con un valor representativo de un nivel de expresión del gen PTP4A1 en tumores de una población de pacientes que derivan de un tratamiento sin ningún beneficio clínico, donde un nivel menor expresión del gen PTP4A1 en la muestra de tumor del paciente es indicativo de un paciente que obtendrá beneficio clínico del tratamiento.

La abreviatura PTP4A1 significa una proteína tirosina fosfatasa de tipo IVA, miembro 1. El Id. De Sec. N ° 1 muestra la secuencia de nucleótidos de PTP4A1 humano.

El término "un valor representativo del nivel de expresión del gen PTP4A1 en tumores de una población de pacientes que no obtienen ningún beneficio clínico del tratamiento" se refiere a una estimación del nivel de expresión medio del gen marcador en tumores de una población de pacientes que no obtienen ningún beneficio clínico del tratamiento. Beneficio clínico se define como tener una respuesta objetiva o estabilización de la enfermedad durante 12 semanas.

En una realización preferida adicional, el gen PTP4A1 muestra entre 1,2 y 1,8 o más veces menor nivel de expresión en la muestra tumoral del paciente en comparación con un valor representativo de la población de pacientes que no obtienen ningún beneficio clínico del tratamiento.

En una realización preferida, el nivel de expresión del gen marcador se determina mediante la tecnología de microarrays u otras tecnologías que evalúan los niveles de expresión de ARN como RT-PCR cuantitativa, o mediante cualquier método que analiza el nivel de expresión de la proteína correspondiente, por ejemplo, inmunohistoquímica (IHC). La construcción y el uso de chips de genes son bien conocidos en la técnica, véase la Patente de EE.UU. N ° 5.202.231, 5.445.934, 5.525.464, 5.695.940, 5.744.305, 5.795.716 y 5.800.992. Véase también, Johnston, M. Curr. Biol. 8: R171-174 (1998); Iyer VR et al, Science 283:83-87 (1999).. Por supuesto, el nivel de expresión génica se puede determinar mediante otros métodos que son conocidos por un experto en la materia, tales como por ejemplo transferencia Northern, RT-PCR, PCR cuantitativa a tiempo real, extensión de cebadores, protección de RNasa, perfil del expresión de ARN.

El gen marcador de la presente invención se puede combinar con otros biomarcadores en grupos de biomarcadores. Los grupos de biomarcadores pueden construirse a partir de cualquier combinación de biomarcadores predictivos para hacer predicciones sobre el efecto del tratamiento con inhibidores de EGFR en pacientes con cáncer. Los biomarcadores y grupos de biomarcadores descritos aquí se pueden utilizar, por ejemplo, para predecir cómo los pacientes con cáncer responderán a la intervención terapéutica con un inhibidor de EGFR.

El término "gen" como se utiliza aquí comprende variantes del gen. El término "variante" se refiere a secuencias de ácido nucleico que son sustancialmente similares a las secuencias de ácidos nucleicos dadas por el número de acceso de GenBank. El término "sustancialmente similar" lo entiende bien un experto en la materia. En particular, una variante del gen puede ser un alelo que muestra intercambios de nucleótidos en comparación con la secuencia de ácido nucleico del alelo más frecuente en la población humana. Preferiblemente, dicha secuencia de ácido nucleico sustancialmente similar tiene una similitud de secuencia con el alelo más frecuente de al menos un 80%,

preferiblemente al menos un 85%, más preferiblemente al menos un 90%, lo más preferible al menos un 95%. El término "variantes" también se refiere a las variantes de corte y empalme.

5 El inhibidor de EGFR se puede seleccionar del grupo que consiste en gefitinib, erlotinib, PKI-166, EKB-569, GW2016, CI-1033 y un anticuerpo anti-erbB tal como trastuzumab y cetuximab.

En otra realización, el inhibidor de EGFR es erlotinib.

10 En aún otra realización, el cáncer es NSCLC.

Las técnicas para la detección y cuantificación de la expresión génica de los genes descritos por esta invención incluyen, pero no se limitan a transferencias Northern, RT-PCR, PCR cuantitativa a tiempo real, extensión de cebadores, protección de RNasa, perfiles de expresión de ARN y técnicas relacionadas. Estas técnicas son bien conocidas por los expertos en la técnica, véase, por ejemplo, Sambrook J et al, Molecular Cloning: A Laboratory Manual, Tercera Edición (Cold Spring Harbor Press, Cold Spring Harbor, 2000).

15 Las técnicas para la detección de la expresión de proteínas de los correspondientes genes descritos por esta invención incluyen, pero no se limitan a técnicas de inmunohistoquímica (IHC).

20 De acuerdo con la invención, las células de una muestra de tejido del paciente, por ejemplo, una biopsia de tumor o de cáncer, puede analizarse para determinar el patrón de expresión de uno o más biomarcadores. El éxito o el fracaso de un tratamiento contra el cáncer puede determinarse basándose en el patrón de expresión de biomarcadores de las células del tejido analizado (células de ensayo), por ejemplo, una biopsia de tumor o del cáncer, al ser relativamente similares o diferentes del patrón de expresión de un grupo control de uno o más de los biomarcadores. En el contexto de esta invención, se encontró que el gen de la tabla 3 está regulado negativamente, es decir, muestra un nivel de expresión más bajo, en los tumores de los pacientes que obtienen beneficio clínico del tratamiento inhibidor de EGFR en comparación con los tumores de los pacientes que no obtienen beneficio clínico del tratamiento con un inhibidor EGFR. Por lo tanto, si las células de ensayo muestran un perfil de expresión de biomarcadores que corresponde a la de un paciente que respondió al tratamiento del cáncer, es muy probable o predecible que el cáncer o tumor del individuo responderá favorablemente al tratamiento con el inhibidor de EGFR. Por el contrario, si las células de ensayo muestran un patrón de expresión de biomarcadores correspondiente a la de un paciente que no responde al tratamiento del cáncer, es muy probable o predecible que el cáncer o tumor del individuo no responda al tratamiento con el inhibidor de EGFR.

35 El biomarcador de la presente invención, es decir el gen listado en la tabla 3, es un primer paso hacia una terapia individualizada para los pacientes con cáncer, en particular para los pacientes con NSCLC refractario. Esta terapia individualizada permitirá a los médicos tratantes seleccionar el agente más apropiado de los medicamentos existentes para el tratamiento del cáncer, en particular para el NSCLC. El beneficio de la terapia individualizada para cada futuro paciente es: tasas de respuesta / número de pacientes que se benefician aumentará y el riesgo de efectos secundarios adversos debidos al tratamiento ineficaz se reducirá.

40 En un objeto adicional de la presente invención se proporciona un método terapéutico de tratamiento de un paciente de cáncer identificado por el método in vitro de la presente invención. Dicho método terapéutico comprende la administración de un inhibidor de EGFR para el paciente que ha sido seleccionado para el tratamiento basado en el patrón de expresión predictivo del gen de la tabla 3. Un inhibidor de EGFR preferido es erlotinib y un cáncer preferido para tratar es NSCLC.

Breve descripción de las figuras

50 La Figura 1 muestra el diseño del estudio;
La Figura 2 muestra el esquema de procesamiento de muestras;
La Figura 3a muestra los niveles de expresión de PTP4A1 en comparación con el resultado clínico del perfil de Genechip®;
La Figura 3b muestra los niveles de expresión de PTP4A1 en comparación con el resultado clínico de qRT-PCR y
55 La Figura 3c muestra la correlación entre las mediciones de Genechip® y qRT-PCR para PTP4A1.

Parte experimental

60 Bases del estudio y diseño del estudio

Se ha descrito recientemente mutaciones en el gen EGFR en el tejido tumoral de un subgrupo de pacientes con NSCLC y la asociación de estas mutaciones con la sensibilidad a erlotinib y gefitinib (Pao W, et al 2004; Lynch et al 2004; Páez et al. 2004). Para los pacientes combinados de dos estudios, se observó EGFR mutado en 13 de 14 pacientes que respondieron al gefitinib y en ninguno de los 11 pacientes tratados con gefitinib que no respondieron. La prevalencia de estas mutaciones fue del 8% (2 de 25) en pacientes no seleccionados con NSCLC. Estas mutaciones se encuentran más frecuentemente en adenocarcinomas (21%), en los tumores que afectan a mujeres

(20%), y en tumores de pacientes japoneses (26%). Estas mutaciones resultan en un aumento de la actividad in vitro de EGFR y aumento de sensibilidad a gefitinib. La relación de las mutaciones con enfermedad estable o duración de la supervivencia prolongada no ha sido evaluada prospectivamente.

5 Con base en los análisis exploratorios del estudio BR.21, parecía poco probable que el beneficio en la supervivencia observado sea sólo debido a las mutaciones del EGFR, ya que se mantiene un beneficio significativo en la supervivencia incluso cuando los pacientes con respuesta objetiva se excluyen del análisis (datos en el archivo). Otros mecanismos moleculares también deben contribuir al efecto.

10 Con base en el supuesto de que se produzcan cambios en los niveles de expresión de genes que son predictivos de la respuesta / beneficio para el tratamiento con Tarceva™, se utilizó el análisis de microarrays para detectar estos cambios.

15 Esto requirió una población de estudio claramente definida tratados con Tarceva en monoterapia™ tras el fracaso de la terapia de primera línea. Con base en la experiencia del estudio BR.21, la población beneficiaria se define como los que tienen respuesta objetiva, o estabilización de la enfermedad durante 12 semanas. Se analizaron los conjuntos de datos clínicos y de microarrays de acuerdo con un plan estadístico predefinido.

20 La aplicación de esta técnica requiere tejido fresco congelado (FFT). Por lo tanto tuvo que realizarse obligatoriamente una biopsia antes del inicio del tratamiento. El material recogido se congeló en nitrógeno líquido (N2).

25 Una segunda muestra de tumor se recogió al mismo tiempo y se almacenó en parafina (fijada en formol e incluida en parafina, FFPE). Esta muestra se analizó para detectar alteraciones en la ruta de señalización de EGFR.

La capacidad de realizar biopsias de tumores a través de broncoscopia es un requisito previo para este estudio. La broncoscopia es un procedimiento estándar para confirmar el diagnóstico de cáncer de pulmón. Aunque generalmente es segura, existe el riesgo remanente de complicaciones, por ejemplo, sangrado.

30 Este estudio fue un primer paso hacia una terapia individualizada para pacientes con NSCLC refractario. Esta terapia individualizada permitirá a los médicos tratantes seleccionar el agente más apropiado de los medicamentos existentes para esta indicación.

35 Una vez que esté disponible la terapia individualizada, el beneficio para cada paciente futuro será mayor que el riesgo que tienen que tomar los pacientes en el presente estudio: aumentará la tasa de respuesta / número de pacientes que se benefician, se reducirá el riesgo de efectos secundarios adversos debido a un tratamiento ineficaz.

Bases de la selección de la dosis

40 Tarceva™ se administra por vía oral una vez al día a una dosis de 150 mg hasta la progresión de la enfermedad, toxicidad intolerable o la muerte. La selección de esta dosis se basa en los parámetros farmacocinéticos, así como en el perfil de seguridad y tolerabilidad de esta dosis observado en los ensayos de Fase I, II y III en pacientes fuertemente pretratados con cáncer avanzado. Los niveles del medicamento observados en el plasma de pacientes con cáncer que reciben la 150 mg / día de dosis fueron consistentemente por encima de la concentración plasmática media de 500 ng / ml dirigido para la eficacia clínica. BR.21 mostró un beneficio de supervivencia con esta dosis.

Objetivos del estudio

50 El objetivo primario fue la identificación de genes expresados diferencialmente que son predictivos de beneficio (RC, RP o EE 12 semanas) del tratamiento con Tarceva™. Es un objetivo adicional importante la identificación de genes expresados diferencialmente predictivos de "respuesta" (RC, RP) para el tratamiento con Tarceva™.

55 Los objetivos secundarios fueron evaluar las alteraciones en las vías de señalización del EGFR con respecto al beneficio del tratamiento.

Diseño del estudio

Descripción del diseño del estudio y régimen de dosificación

60 Este fue un estudio de Fase II abierto, de identificación de marcadores predictivos. El estudio se llevó a cabo en aproximadamente 26 centros de cerca de 12 países. Se reclutaron 264 pacientes con NSCLC avanzado tras el fracaso de al menos un régimen de quimioterapia en un período de 12 meses. Se administró Tarceva™ oral de forma continua a una dosis de 150 mg / día. Se permitió la reducción de dosis en base a la tolerabilidad al tratamiento farmacológico. Los parámetros clínicos y de laboratorio se analizaron para evaluar el control de la enfermedad y la toxicidad. El tratamiento continuó hasta la progresión de la enfermedad, toxicidad inaceptable o la muerte. El diseño del estudio se representa en la figura 1.

Se obtuvieron muestras de tejido tumoral y de sangre para los análisis moleculares para analizar los efectos de Tarceva™ y para identificar subgrupos de pacientes que se benefician de la terapia.

5 Análisis de los marcadores predictivos

Las biopsias del tumor fueron tomadas dentro de las 2 semanas antes del inicio del tratamiento. Se recogieron dos muestras diferentes:

10 Se tomaron muestras del tumor a las 2 semanas antes del inicio del tratamiento. Se recogieron dos muestras diferentes:

La primera muestra se congeló siempre inmediatamente en N₂ líquido.

15 La segunda muestra se fijó en formol e incluida en parafina.

El tejido congelado rápidamente tuvo la más alta prioridad en este estudio.

20 La Figura 2 muestra un esquema del procesamiento de la muestra.

Análisis de microarray

25 Las muestras congeladas rápidamente se utilizaron para microdissección de captura por láser (LCM) de las células tumorales para extraer el ARN del tumor y el ARN del tejido circundante al tumor. El ARN se analizó en chips de microarrays Affymetrix (HG-U133A) para establecer el perfil de expresión de los genes tumorales de los pacientes. El control de calidad de los chips Affymetrix se utilizó para seleccionar aquellas muestras con una calidad adecuada para la comparación estadística.

30 Análisis de biomarcadores únicos de tejido fijado en formol incluido en parafina

La segunda biopsia del tumor, la muestra FFPE, se utilizó para realizar el análisis de mutaciones del ADN, IHC y ISH como se describe a continuación. Se realizaron análisis similares en el tejido recogido en el diagnóstico inicial.

35 El estado de las mutaciones del ADN de los genes que codifican EGFR y de otras moléculas implicadas en la ruta de señalización de EGFR se analizaron por secuenciación de ADN. La amplificación génica de EGFR y de los genes relacionados se estudiaron mediante FISH.

40 Los análisis de expresión de proteínas incluyeron análisis inmunohistoquímico [IHC] de EGFR y de otras proteínas dentro de la ruta de señalización de EGFR.

Análisis de respuesta

45 Se utilizaron los criterios RECIST (Medición unidimensional del tumor) para evaluar la respuesta. Estos criterios se pueden encontrar en el siguiente enlace: <http://www.eortc.be/recist/>

A tener en cuenta que: Para asignar un estado de RC o RP, los cambios en las medidas del tumor deben confirmarse mediante evaluaciones repetidas de al menos 4 semanas de diferencia, en cualquier momento durante el período de tratamiento.

50 En el caso de EE, las medidas de seguimiento deben haber cumplido los criterios de EE por lo menos una vez después de entrar en el estudio con un intervalo mínimo de 6 semanas.

En el caso de mantenerse la EE, las medidas de seguimiento deben cumplir con los criterios de EE al menos una vez después de entrar en el estudio con una duración de mantenimiento de al menos 12 semanas.

55 Análisis de la supervivencia

60 Se llevó a cabo una revisión regular del paciente cada 3 meses ya sea en la visita del paciente a la clínica o por teléfono. Se registraron todas las muertes. Al final del estudio fue necesaria la confirmación definitiva de la supervivencia de cada paciente.

Métodos

65 Preparación de muestras de ARN y control de calidad de las muestras de ARN

Todo el procesamiento de muestras de biopsias se realizó por un laboratorio de referencia de patologías; las

- 5 muestras frescas de tejido congeladas fueron enviadas desde los centros de investigación a las instalaciones de operaciones de muestras clínicas en Roche Basilea y desde allí al laboratorio de patología para su posterior procesamiento. Se utilizó microdissección mediante captura con láser para seleccionar las células tumorales del tejido circundante. Después de LCM, el ARN se purificó a partir del material tumoral enriquecido. El laboratorio de patología a continuación, llevó a cabo una serie de pasos para hacer una estimación de la concentración y la calidad del ARN.
- 10 Las RNAsas son enzimas que degradan el ARN y se encuentran en todas partes y todos los procedimientos en los que el ARN se utiliza deben estar estrictamente controlados para minimizar la degradación del ARN. La mayoría de las especies de ARNm tienen vidas medias más bien cortas y por lo tanto se consideran bastante inestables. Por lo tanto es importante realizar comprobaciones de integridad del ARN y cuantificación antes de realizar cualquier ensayo.
- 15 La concentración de ARN y el perfil de calidad puede analizarse usando un instrumento de Agilent (Agilent Technologies, Inc., Palo Alto, CA) llamado 2100 Bioanalyzer®. El software del instrumento genera un número de integridad del ARN (RIN), una estimación de cuantificación (Schroeder, A., et al, El RIN: Un número de integridad del ARN para la asignación de valores de integridad para mediciones de RNA. BMC Mol Biol, 2006 7: P. 3), y calcula las proporciones ribosómicas del ARN total de la muestra. El RIN se determina a partir del rastro electroforético total de la muestra de ARN, y por lo tanto incluye la presencia o ausencia de productos de degradación.
- 20 La calidad del ARN se analizó mediante un 2100 Bioanalyzer®. Sólo las muestras con al menos un pico de rRNA por encima del ruido añadido de poli-I y con suficiente ARN se seleccionaron para su posterior análisis en la plataforma Affymetrix. El ARN purificado se trasladó al Centro de Roche de Genómica Médica (RCMG, Basilea, Suiza) para el análisis de microarrays. Se recibieron 122 muestras de ARN desde el laboratorio de patología para su posterior procesamiento.
- 25 **Marcaje de dianas de muestras de tejido de ARN**
- 30 El marcaje de las dianas se llevó a cabo de acuerdo con el protocolo de amplificación del marcaje de dianas en dos ciclos de Affymetrix (Affymetrix, Santa Clara, California), según las instrucciones del fabricante.
- El método se basa en el procedimiento de amplificación lineal estándar Eberwine pero utiliza dos ciclos de este procedimiento para generar suficiente ARNc marcado para la hibridación en un microarray.
- 35 La entrada de ARN total usado en la reacción de marcaje fue de 10 ng para aquellas muestras en las que estaba disponible más de 10 ng de RNA, si estaba disponible menos de esta cantidad o si no había datos cuantitativos disponibles (debido a una concentración de ARN muy baja), se utilizó la mitad de la muestra total en la reacción. Los rendimientos de las reacciones de marcaje osciló entre 20-180 µg de ARNc. Un paso de normalización se introdujo al nivel de hibridación donde se utilizó 15 µg de ARNc para cada muestra.
- 40 Se utilizó ARN de referencia humano (Stratagene, Carlsbad, CA, USA) como muestra control en el flujo de trabajo con cada lote de muestras. Se utilizó 10 ng de este ARN como entrada junto con las muestras de prueba para verificar que los reactivos de marcaje y la hibridación trabajaban como se esperaba.
- 45 **Hibridaciones de microarray**
- Los microarrays HG-U133A de Affymetrix contienen más de 22.000 grupos de sondas dirigidas a aproximadamente 18.400 transcritos y a las variantes que representan unos 14.500 genes bien caracterizados.
- 50 La hibridación para todas las muestras se llevó a cabo de acuerdo con las instrucciones de Affymetrix (Affymetrix Inc., Expression Analysis Technical Manual, 2004). Brevemente, para cada muestra, 15 mg de ARNc marcado con biotina se fragmentó en presencia de cationes divalentes y se calentó e hibridó durante la noche en microarrays HG-U133A de Affymetrix de oligonucleótidos de genoma completo. Los días siguientes se tiñeron los microarrays con estreptavidina-ficoeritrina (Molecular Probes; Eugene, OR) según las instrucciones del fabricante. Los microarrays se escanearon utilizando un escáner GeneChip 3000 (Affymetrix), y las intensidades de señal se calcula ron automáticamente mediante el programa de funcionamiento GeneChip (SMOC) Versión 1.4 (Affymetrix).
- 55 **Análisis estadístico**
- 60 El análisis de los datos de Affymetrix™ constaba de cinco pasos principales.
- Paso 1 de control de calidad. El objetivo fue identificar y excluir del análisis datos de la matriz con un perfil de calidad por debajo del estándar.
- 65 Paso 2 de pre-procesamiento y normalización. El objetivo fue crear un "conjunto de datos para análisis" escalado y normalizado susceptibles de comparación entre chips. Comprendía una estimación y sustracción del ruido de fondo,

el resumen de la sonda y escalado.

5 Paso 3 de exploración y la descripción. El objetivo fue identificar los posibles sesgos y las fuentes de variabilidad. Consistió en la aplicación de técnicas descriptivas de análisis multivariados y univariados para identificar covariables influyentes.

10 Paso 4 de modelado y análisis. El objetivo fue identificar una lista de marcadores candidatos en base a la evaluación estadística de la diferencia en el nivel de expresión medio entre los pacientes con "beneficio clínico" y "sin beneficio clínico". Consistía en ajuste de un modelo estadístico adecuado para cada grupo de sondas y obtener una medida de significación estadística.

15 Paso 5 de análisis de la robustez. El objetivo fue generar una lista cualificada de marcadores candidatos que no dependen en gran medida de los métodos de pre-procesamiento y supuestos estadísticos. Consistía en reiterar el análisis con diferentes enfoques metodológicos y que intersecciona la lista de candidatos. Todos los análisis se realizaron utilizando el paquete de programas R.

Paso 1: Control de calidad

20 La evaluación de la calidad de los datos se basó en la comprobación de varios parámetros. Estos incluían parámetros de calidad estándar de Affymetrix GeneChip™, en particular: factor de escalado, porcentaje de la presente convocatoria y ruido de fondo medio. Este paso también incluye la inspección visual de las imágenes virtuales del chip para la detección de problemas de hibridación localizados, y la comparación de cada chip con una media de chip virtual para detectar cualquier desplazamiento atípico del comportamiento medio. También se realizó un análisis de correlación entre chips para detectar valores atípicos en las muestras. Además, fueron tomadas en consideración las medidas auxiliares de calidad de ARN obtenidas a partir del análisis de muestras de ARN con el Bioanalyzer Agilent 2100™.

30 Basado en estos parámetros, se excluyeron del análisis los datos de 20 microarrays. Así, se incluyeron en el análisis los datos de un total de 102 microarrays que representan 102 pacientes. La descripción clínica de este conjunto de 102 muestras se presentan en la tabla 1.

Tabla 1: Descripción de las características clínicas de los pacientes incluidos en el análisis

Variable	Valor	n=102 n (%)
Mejor respuesta	N/A	16 (15,7%)
	EP	49 (48,0%)
	EE	31 (30,4%)
	RP	6 (5,9%)
Beneficio clínico	NO	81 (79,4%)
	SI	21 (20,6%)
SEXO	MUJER	25 (24,5%)
	HOMBRE	77 (74,5%)
ETNIA	CAUCASIANA	65 (63,7%)
	ORIENTAL	37 (36,3%)
Histología	ADENOCARCINOMA	35 (34,3%)
	ESCAMOSO	53 (52,0%)
	OTROS	14 (13,7%)
FUMADOR	NO	20 (19,6%)
	SI	82 (80,4%)

35 Paso 2: Datos de pre-procesamiento y normalización

40 Se utilizó el algoritmo Therma (Irizarry, RA, et al, Summaries of Affymetrix GeneChip probe level data. Nucl Acids Res, 2003 31 (4): p. e15) para pre-procesamiento y normalización. Se utilizó el algoritmo mas5 (Affymetrix, GeneChip® Expression: Data Analysis Fundamentals 2004, Affymetrix) para realizar la detección para los grupos de sondas individuales. Los grupos de sondas llamados "ausente" o "marginal" en todas las muestras se extrajeron del posterior análisis, 5930 grupos de sondas se retiraron del análisis según en este criterio. El grupo de datos para el análisis consistió por lo tanto de una matriz con 16353 grupos de sondas (de 22283) medidas en 102 pacientes.

Paso 3: Descripción y exploración de los datos

45 Se realizó un análisis descriptivo exploratorio para identificar los posibles sesgos y las principales fuentes de variabilidad. Se seleccionó un conjunto de covariables con un impacto potencial sobre los perfiles de expresión génica. Comprendía ambas variables técnicas y clínicas. Las covariables técnicas incluyeron: fecha de procesamiento del ARN (referido más adelante como lote), RIN (como medida de la calidad / integridad del ARN), y

operador y centro de recogida de muestras. Las covariables clínicas incluyeron: tipo histológico, tabaquismo, estadio tumoral, puntuación de rendimiento (Oken, MM, et al, Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol, 1982. 5 (6): p. 649 -55), datos demográficos, estado del respondedor y el estado de beneficio clínico.

5 Las herramientas de análisis incluyen ANOVA univariante y análisis de componentes principales. Para cada una de estas covariables, se aplicó ANOVA univariante de forma independiente para cada grupo de sonda.

10 Se identificó un efecto significativo de la variable de lotes. En la práctica, la variable de lotes capturó diferencias entre las fechas de procesamiento de la muestra y el lote del chip Affymetrix. Después de comprobar que la variable de lote fue casi independiente de las variables de interés, el efecto lote se corrigió usando el método descrito en Johnson, WE, Li C., y Rabinovic A., Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostat 2007. 8 (1): p. 118-127.

15 Los grupos de datos normalizados establecido después de la corrección del efecto de los lotes sirvieron como datos de análisis establecido en los análisis posteriores.

Histología y RIN dos variables adicionales importantes destacados mediante el análisis descriptivo.

20 Paso 4: Modelado de datos y análisis.

25 Se ajustó un modelo lineal de forma independiente para cada grupo de sondas. Las variables incluidas en el modelo se describen en la tabla 2. Los parámetros del modelo se estimaron mediante la técnica de máxima probabilidad. El parámetro correspondiente a la variable de "beneficio clínico" (X1) se utilizó para evaluar la diferencia en el nivel de expresión entre el grupo de pacientes con beneficio clínico y el grupo sin beneficio clínico.

Tabla 2: Descripción de las variables incluidas en el modelo lineal.

Variable	Tipo	Valor
expresión génica	Dependiente (Y_{ip})	\log_2 de la intensidad del grupo de sondas i en el paciente p .
Intercepción	Media general (μ)	
Beneficio Clínico	Predicador de interés (X1)	SI / NO
Histología	Ajuste de covarible (X2)	ADENO. / ESCAM. / OTRO
RAZA	Aj. Cov. (X3)	ORIENT. / CAUCAS.
SEXO	Aj. Cov. (X4)	MUJER / HOMBRE
RIN	Aj. Cov. (X5)	[2,...,7.9]
FUMADOR	Aj. Cov. (X6)	ACTUAL/EX-FUMADOR/NUNCA
Estadio	Aj. Cov. (X7)	INEXTIRPABLE.III / IV

30 Para cada grupo de sondas i , el objetivo de la prueba estadística fue rechazar la hipótesis de que los niveles de expresión media en pacientes con beneficio clínico y pacientes sin beneficio clínico son iguales, teniendo en cuenta el ajuste de otras covariables enumeradas en la tabla 2. Formalmente, la hipótesis nula de igualdad se analizó contra una alternativa de dos colas. Los correspondientes valores de p se indican en la tabla 3.

35 La elección del modelo lineal fue motivada por dos razones. En primer lugar, el modelado lineal es un método versátil, bien caracterizado y sólido que permite el ajuste de las variables de confusión cuando se estima el efecto de la variable de interés. En segundo lugar, dado el tamaño de muestra de 102, y la normalización y escalado del conjunto de datos, la asunción de distribución normal era razonable y justificada.

40 Para cada grupo de sondas, la asunción de homogeneidad de varianza se evaluó utilizando pruebas Fligner-Killeen basadas en la residualidad del modelo. El análisis consistió en 3 pasos:

1. Analizar cada una de las variables categóricas para la homogeneidad de la varianza residual
2. Tener en cuenta la variable V con el menor valor de p
3. Si al menos el valor p es menor que 0,001, reajustar el modelo permitiendo diferente nivel de variables V para tener una varianza diferente.

45 Paso 5: Robustez

50 El objetivo del análisis de la robustez fue reducir el riesgo de que los resultados del análisis pudieran ser un artefacto y un resultado de los pasos de pre-procesamiento o suposiciones subyacentes en el análisis estadístico. Se consideraron los siguientes tres aspectos: a) inclusión o exclusión de algunas chips adicionales en el paso de control de calidad; b) algoritmo de pre-procesamiento y normalización, y c) supuestos estadísticos y pruebas de aproximación.

La lista de marcadores candidatos se define como el subconjunto de genes declarado consistentemente como significativo con diferentes ajustes de análisis. Las diferentes opciones de análisis aplicados han sido los siguientes:

- 5 a) Un subgrupo adicional de 8 chips se identificó en base a criterios de control de calidad más estrictos. Un "grupo de datos reducido" se definió mediante la exclusión de estos 8 chips.
- b) MAS5 se identificó como una alternativa a rma para el pre-procesamiento y normalización. MAS5 utiliza diferentes métodos para la estimación de ruido de fondo, el resumen de la sonda y la normalización.
- c) Se utilizaron dos pruebas estadísticas adicionales.

- 10 a. Una prueba de Wilcoxon para la diferencia entre beneficio clínico y sin beneficio clínico y
- b. una prueba de razón de verosimilitud (LRT) para el modelo de regresión logística donde el beneficio clínico se tomó como la variable de respuesta y la expresión de genes como covariable. Estas dos pruebas adicionales se basan en un conjunto diferente de supuestos estadísticos subyacentes. Para cada grupo de sondas, la LRT estaba siguiendo un Chi-cuadrado con 1 grado de libertad.

15 En resumen, se consideraron dos conjuntos de muestras (el conjunto de datos "completo" y el conjunto de datos "reducido"), y 2 algoritmos de pre-procesamiento (mas5 y rma); esto resultó en cuatro diferentes conjuntos de análisis de datos. Para cada uno de estos cuatro conjuntos de datos, se aplicaron tres pruebas estadísticas diferentes. Por lo tanto, para cada grupo de sondas, se calcularon tres valores de p. En cada conjunto de análisis de datos, se aplicó un criterio compuesto para identificar la lista de genes regulados diferencialmente. Este criterio compuesto se define como: el máximo valor p es inferior a 0,05 y los mínimos valores de p son inferiores a 0,001. El análisis de robustez utilizando un criterio de 1 para la identificación de genes marcadores proporcionó PTP4A1 como marcador predictivo para el tratamiento inhibidor de EGFR.

25 Tabla 3:

Gen marcador para un beneficio clínico basado en el análisis de la robustez después de la aplicación del criterio compuesto. La columna 1 es el identificador de Affymetrix del grupo de sondas. La columna 2 es el número de acceso de GenBank de la correspondiente secuencia del gen. La columna 3 es el correspondiente nombre oficial del gen. La columna 4 es el correspondiente cambio medio ajustado de número de veces en el nivel de expresión entre beneficio clínico y ningún beneficio clínico al paciente, estimado a partir del modelo lineal. La columna 5 es el valor p para la prueba de la diferencia en el nivel de expresión entre pacientes con beneficio clínico y sin beneficio clínico a, derivado del modelo lineal. La columna 6 es el intervalo de confianza del 95% para el cambio medio ajustado de veces en el nivel de expresión.					
ID del grupo de sonda Affymetrix	GenBank	Gen	cambio medio ajustado	Valor p	IC 95 %
200732_s_at	NM_003463 (Id. De Sec. N° 1)	PTP4A1	-1,44	3,2x1E-3	-1,8 , -1,2

Análisis estadístico adicional

30 Para el marcador candidato PTP4A1 seleccionado, se realizaron los siguientes análisis adicionales en un entorno validado por estadísticos independientes:

- Regresión de Cox univariante para la SLP (supervivencia libre de progresión) a partir del análisis primario de Affymetrix,
- Regresión logística univariante para el beneficio clínico a partir del análisis primario de Affymetrix, y
- 35 • Regresión de Cox univariante para la supervivencia a partir del análisis primario de Affymetrix

Los resultados de estos análisis se presentan a continuación. Estos son consistentes con los resultados del análisis principal y confirma la elección del marcador seleccionado.

40 Resultados: Regresión de Cox univariante para la SLP (supervivencia libre de progresión) a partir del análisis primario de Affymetrix:

Gen	Nº de pacientes	Proporción de riesgo	IC 95 % CI para la proporción de riesgo	valor p
PTP4A1	102	1,69	1,23; 2,33	0,0011

45 Resultados: Regresión logística univariante para el beneficio clínico a partir del análisis primario de Affymetrix:

Gen	Nº de pacientes	Razón de probabilidades	IC 95 % CI para la razón de probabilidades	valor p
PTP4A1	102	0,22	0,08; 0,57	0,0018

Resultados: Regresión de Cox univariante para la supervivencia a partir del análisis primario de Affymetrix:

Gen	Nº de pacientes	Proporción de riesgo	IC 95 % para la proporción de riesgo	valor p
PTP4A1	102	1,63	1,15; 2,30	0,0055

qRT-PCR

5 Se sintetizó ADNc utilizando SuperScript™ III First-strand Synthesis SuperMix para qRT-PCR (Invitrogen, CA, EE.UU.) según las instrucciones del fabricante, pero sin la inclusión de una digestión de RNasa H.

10 La PCR cuantitativa se realizó utilizando ensayos de expresión génica TaqMan® en un sistema de detección de secuencias ABI PRISM® 7900HT de acuerdo con las recomendaciones del fabricante (Applied Biosystems, CA, EE.UU.). Todos los ensayos se realizaron por triplicado.

15 Los cebadores y sondas utilizados sobrepasaron los límites del exón o estaban dentro de la secuencia de la sonda de interés Affymetrix GeneChip®. Dos genes de mantenimiento se incluyeron como controles endógenos: beta-2-microglobulina (B2M; Ensayo Hs9999907_ml) e hipoxantinafosforibosil transferasa (HPRT; Ensayo Hs9999909_ml).

20 Todas los análisis incluyeron un calibrador de muestra (ARN total MVPTM de pulmón humano adulto; Stratagene, CA, EE.UU.) y una curva estándar. Se utilizó como plantilla ARN total de referencia universal de humanos (Stratagene, CA, EE.UU.) para las curvas estándar PTPRF. Todas las muestras se midieron por triplicado. La cuantificación relativa se realizó usando el método $-\Delta Ct$.

Resultados

25 Tal como se describió anteriormente, se determinaron los perfiles de expresión génica Affymetrix GeneChip® para 102 pacientes incluidos en este estudio. Entre estos pacientes, se obtuvieron los resultados de qRT-PCR para 75 (tabla 4). Los datos demográficos y las características clínicas de los pacientes con resultados de qRT-PCR fueron similares a los de toda la población (n = 264) y de los pacientes con perfiles de expresión génica GeneChip® disponibles.

30 Tabla 4: Características basales: pacientes con análisis de qRT-PCR (n=75)

Característica	
Edad (mediana, rango)	62 (39-85)
Género; n (%)	
Hombre	19 (25)
Mujer	56 (75)
Escala de calidad de vida ECOG; n (%)	
0	7 (9)
1	45 (60)
2	23 (31)
Histología; n (%)	
Adenocarcinoma	27 (36)
Carcinoma de células escamosas	34 (45)
Carcinoma de células grandes	2 (3)
Otro	12 (16)
Estadío de la enfermedad; n (%)	
IIIB	22 (29)
IV	53 (71)
Número de regímenes previos de quimioterapia; n (%)	
0	19 (25)
1	36 (48)
≥2	20 (27)
Etnia; n (%)	
Caucasiana	51 (68)
Asiática	24 (32)
Tabaquismo; n (%)	
No fumador	12 (16)
Fumador	24 (32)
Ex - fumador	39 (52)

De los 75 pacientes con resultados de qRT-PCR, 4 (5%) presentaron una respuesta parcial (RP), 23 (31%)

presentaron EE, 39 (52%) presentaron EP, y 9 (12%) no pudieron evaluarse. Estos resultados fueron muy similares a los observados en la población total del estudio (n=264).

5 La Figura 3 muestra los niveles relativos de ARNm para PTP4A1 en cada paciente, según la evaluación del perfil de GeneChip® de Affymetrix y qRT-PCR. La Figura 3a muestra los niveles de expresión frente a la evolución clínica de los perfiles de GeneChip® y la Figura 3b muestra los niveles de expresión de qRT-PCR.

10 Hubo una buena correlación entre las mediciones de GeneChip® y qRT-PCR de la transcripción del mRNA de PTP4A1 (Figura 3c, p de Pearson = 0,6, p <0,01). Según se observó con los perfiles de GeneChip®, los niveles de mRNA de PTP4A1 analizados mediante qRT-PCR parecen correlacionarse con la respuesta a erlotinib, con niveles más altos observados en los respondedores frente a los no respondedores.

Discusión

15 Mediante el análisis de muestras de tejido con tecnología de microarrays de oligonucleótidos de alta densidad, y la aplicación de modelos estadísticos para los datos, fuimos capaces de identificar los genes cuyos niveles de expresión puede ser predictivos de pacientes que obtienen un beneficio clínico del tratamiento con erlotinib.

20 Se aplicó un criterio compuesto (definido anteriormente). Resultó en PTP4A1 como marcador predictivo para el tratamiento inhibidor de EGFR.

25 PTP4A1 es el segundo miembro de la familia de la proteína tirosina fosfatasa. Se encuentra en el cromosoma 6q12 y codifica una proteína que pertenece a una pequeña clase de proteína tirosina fosfatasas (PTP) prenilados, que contiene un dominio PTP y un motivo de prenilación C-terminal. Esta tirosina fosfatasa es una proteína nuclear, pero puede estar principalmente asociada a la membrana plasmática.

30 Varios informes han asociado la alta expresión de PTP4A1 con la actividad oncogénica, específicamente el aumento de la tasa de proliferación y motilidad celular e invasividad. Un estudio reciente en células grandes de cáncer de pulmón ha demostrado un papel funcional de PTP4A1 en los procesos de migración celular e invasión.

En este estudio, se encontró que PTP4A1 estaba regulada negativamente en los pacientes que obtienen un beneficio clínico del tratamiento con erlotinib. La función oncogénica de este gen es un argumento claro a favor de su utilidad como marcador tumoral.

35 Listado de secuencias

<110> F. Hoffmann-La Roche AG

<120> Marcador predictivo para el tratamiento con inhibidor de EGFR

<130> 24411WO

40 <150> EP07114309.3

<151> 2007-08-14

<160> 1

<170> PatentIn versión 3.4

<210> 1

45 <211> 5099

<212> DNA

<213> Homo sapiens

<400> 1

ES 2 395 588 T3

gatgaccgtg	gttaaggcag	tagaaagcag	gctcgacact	gagacgcggt	tccagcccta	60
gaaggattgc	attttacgat	tcaggcaaac	ttcgagtctc	ctcatgtgac	atgtgtgcac	120
tcctggaccg	ttttaattat	gagaaacgtg	gtttgcagca	tgattccttc	cagtcgataa	180
atcggaatct	ctctcgtccc	caccccttct	taacttcagg	cttcctgcat	cccggagcac	240
tcccggcagc	cccttcctc	ccccgccccg	gggatgctcc	gactcggcgc	ttagccattc	300
atcaaccggt	tcacaccggc	ggcggccgcc	gcggagtgac	gtccggaggg	ggcgggcctc	360
cgccccgcc	tgctggctcc	tggccccgcg	ttccaggccg	cgattggtgg	ctggaggggt	420
gcacgtcgcg	ccggctataa	aggggagggc	ttgtgacgca	agggcgcctc	ggcgcgtgta	480
ttggctcctt	cggctgcggg	ccggctcggc	tacgcgctct	gctccgagcc	gctcactgca	540
tggtagagtc	tggtgcccc	gccgccgctt	gcatcgccgc	caccgccgct	ccgccacgac	600
caccgccgcc	tcctgccctg	cagccaccgc	caccgcctgt	gtcggccgcc	cctcgggacc	660
ggctgtatga	ttaggccaca	atcttcaatg	agtaaacata	ttcctcaatt	ctgtggtggt	720
cttggtcaca	catttatgga	gtttctgaag	ggcagtggag	attactgcca	ggcacagcac	780
gacctctatg	cagacaagtg	aactgtagaa	actgattact	gctccaccaa	gaagcccca	840
taagagtggg	tatcctggac	acagaagtgt	tgaattgaaa	tccacagagc	attttacaag	900
agttctgacc	tggtatgggt	aaacctcagt	gcacttcttt	tctgttggcc	tcagtattac	960
tggtattgaag	aattgctgct	tcttgttagg	aggttcattt	cacttatcat	tacttacaac	1020
ttcatactca	aagcactgag	aatttcaagt	ggagtatatt	gaagtagact	tcagtttctt	1080
tgcatcattt	ctgtattcaa	tttttttaat	tatttcataa	ccctattgag	tgttttttaa	1140
ctaaattaac	atggctcgaa	tgaaccgcc	agctcctgtg	gaagtcacat	acaagaacat	1200
gagatttctt	attacacaca	atccaaccaa	tgcgacctta	aacaaattta	tagaggaact	1260
taagaagtat	ggagttacca	caatagtaag	agtatgtgaa	gcaacttatg	acactactct	1320
tgtggagaaa	gaaggtatcc	atgttcttga	ttggcctttt	gatgatgggtg	caccaccatc	1380
caaccagatt	gttgatgact	ggttaagtct	tgtgaaaatt	aagtttcgtg	aagaacctgg	1440

ttgttgatt	gctgttcatt	gcgttgcagg	ccttgggaga	gctccagtac	ttgttgcctt	1500
agcattaatt	gaaggtggaa	tgaatacga	agatgcagta	caattcataa	gacaaaagcg	1560
gcgtggagct	tttaacagca	agcaacttct	gtatttggag	aagtatcgtc	ctaaaatgcg	1620
gctgcgtttc	aaagattcca	acggtcatag	aaacaactgt	tgattcaat	aaaattgggg	1680
tgccaatgc	tactggaagt	ggaacttgag	atagggccta	atttgttata	catattagcc	1740
aacatgttgg	cttagtaagt	ctaatgaagc	ttccatagga	gtattgaaag	gcagttttac	1800
caggcctcaa	gctagacaga	tttggcaacc	tctgtatttg	ggttacagtc	aacctatttg	1860
gatacttggc	aaaagattct	tgctgtcagc	atataaaatg	tgcttgtcat	ttgtatcaat	1920
tgacctttcc	ccaaatcatg	cagtattgag	ttatgacttg	ttaaacttat	tcccatgcca	1980
gaatcttatc	aatacataag	aaatttagga	agattagggtg	ccaaaatacc	cagcacaata	2040
cttgtatatt	tttagtacca	tacagaagta	aaatcccagg	aactatgaac	actagacctt	2100
atgtggttta	ttccttcaat	catttcaaac	attgaaagta	gggcctacat	ggttatttgc	2160
ctgctcactt	tatgtttaca	tctcccat	tcataccaat	atacgtcagg	tttgcttaac	2220
cattgatttt	ttttttttt	taccaagtct	tacagtgatt	attttacgtg	tttccatgta	2280
tctcactttg	tgctgtatta	aaaaaacctc	cattttgaaa	atctacgttg	tacagaagca	2340
catgtcttta	atgtcttcag	acaaaaaagc	cttacattaa	tttaatgttt	gcactctgag	2400
gtgcaactta	acagggaggg	cctgagaaaa	gaatgggagg	gggctattaa	ttatttttag	2460
caaatgttg	cctttgtctt	gtgcaaacat	gtagaatatg	ctctttaatt	tagtaaaata	2520
tttttttaa	aggtagagat	gctttgttat	tgtaatcata	aacttcctga	aattcttgta	2580
atttttttcc	catacttatc	agaagtgtgt	ttaccaactt	atttttgttt	gaaagtgtga	2640
ttttttttt	ccttccaac	ctctcttgca	aaaaaagaaa	tgggtttctg	ctaatgaatt	2700
gagcagacat	ctaatatttt	atatgccttt	tgagctgtgt	aacttaatat	ttggatactt	2760
gacaatttgt	tttattatgt	aattgataaa	atggtgatgt	gtattaatgt	tagttcaacc	2820
atatatttat	actgtctggg	gatgtgtggt	tatagttctg	tgggagaaat	aattttgtca	2880
gtgttcacca	gcttgtaaaa	acttagtgcg	agagctgaaa	catctaaata	aataatgaca	2940
tgatttatc	atcattgaga	ttggtttctg	taaaattaac	ttattttgta	gaagacaaaa	3000
tgaattgcac	ttcacttaat	gtgtgtcctc	atctttttac	aaataaatga	aggattataa	3060
atgatgacag	catttttagta	aacttttaga	caaatattgt	tagggtcatt	catgaaaact	3120
ttaatactaa	aagcactttc	cattatatac	tttttaaagg	tctagataat	tttgaaccaa	3180
tttattattg	tgtactgagg	agaaataatg	tatagtagag	gacagccttg	gtttgtaaag	3240
ctcagttcca	ctagttcatg	gttttctgca	acttctgagc	ctcagttttc	tcctttgcaa	3300
attaataatt	acataccttt	atagattttg	aaattaattt	aaatattagt	atttggtaca	3360
tgaaggctta	atgttaagtt	tcctttaatg	atccacaata	atccctttga	tcacgttaat	3420
ctaaatctag	atgtctttgt	ctaatttttt	ttgaatagca	gttataaatg	taaaggactc	3480

ES 2 395 588 T3

aaagtttaag taaaaagtga tactccacct tgtgtttcaa agaatttagt tccacctctt 3540
cataaccagtt taacacttaa tatatttcat tggatttttag acagggcaaa aggaagaaca 3600
ggggcctctg gaggcccttg gttattttaa tcttggatta tttgtgatag taatcacaaa 3660
tttttgctta atttttaacc tgaggttttg tttttttttt aaaggaaatg cagcctagtc 3720
ttgagaacat aattttatat aatcaattac taaatgttaa actattacca cacagcccat 3780
aaaacagcat ttgcgtttat tgagagagag gatgtgccat catgattaat gaaaactatc 3840
ttttgagttt gaaaagaaat taatttgag tgtttgatt gtatatatgg tgctaaaaat 3900
aaattaattt actttataaa ccttatctgt acattatacg atgtgatgaa atttgctttt 3960
tatccaaata ttttgtatct tgtaaatatg gctaattata ggaatgccta taatacatct 4020
tagattcctt atatctaata agagttcaaa gagttatgag ttgaagtctt gaatgcagga 4080
aactatctga tagtgttcta aaatttggtt acttgggttt ggataccctt agtgggatga 4140
tgtaaataga ggctagctac ctaggcttgt ctatagcaac cataatgttg atgtaagtaa 4200
tgcggttact gaatcataag aaaatgccat ctcttttttag ttgaaggaaa actctggaag 4260
taggtgccat tggtcattct gcagtgcact gcaaccattg tttcccctag tgccctcttt 4320
tccctagggc attgctctcc tattcccacg ccttaacaca gctctatacc tagaagcagc 4380
cagcccaggc atgcagtcac atttaatcac atcccccttc tagagtgcct caaaatgatg 4440
tagtccctca acttggctaa agaatctcaa tctcttgaaa tttatttttt taatgtcata 4500
ttcatctggt aaatatctac tgtttgccag gcatttaaga atatggcaaa gaacataaaa 4560
gatggtgtca ccagattttg gtcaccaatg agtaccgcac ccgttgccat gattaagaga 4620
gaatgctttc tattggagtt tcaggaaata taatttgaga atactttaaa ggggaagtga 4680
agtataagtg aatgatattt ttcttttaca tgtaaacat gaagttattt caaagttaag 4740
ttttaaaca aatacatgaa gtagtgctg ccatacatgt taatattcta cattcttgct 4800
tccttaaatt aatatgtttg tgtgtatata tgtgcctcac acctgaattg aaaattaaag 4860
actggtttaa aagtggttta aaagtgacat ttaatgttc tccattacgt ttggggtaac 4920
cagcctaagt ggaatcttgg aaggaaagta agggaaaaac ttgtatttgc cttcaatgaa 4980
ttaaaccagt gatatgtgtt aacgtatgaa tgaaaggatt gatggtgatt ttataattat 5040
atatattgcc gcagtaacca gttaataaat tgatagctac caaaaaaaaa aaaaaaaaaa 5099

REIVINDICACIONES

- 5 1. Un método in vitro para predecir la respuesta de un paciente con NSCLC al tratamiento con erlotinib que comprende:
- 10 determinar un nivel de expresión de un gen PTP4A1 en una muestra de tumor de un paciente y comparar el nivel de expresión del gen PTP4A1 con un valor representativo de un nivel de expresión del gen PTP4A1 en los tumores de una población de pacientes que no obtienen ningún beneficio clínico del tratamiento, en el que un nivel de expresión más bajo del gen PTP4A1 en la muestra de tumor del paciente es indicativo de un paciente que obtendrá un beneficio clínico del tratamiento.
- 15 2. El método de la reivindicación 1, en el que el nivel de expresión se determina mediante tecnología de microarrays.
3. El método de la reivindicación 1 o 2, en el que el gen PTP4A1 muestra entre 1,2 y 1,8 o más veces menor nivel de expresión en la muestra de tumor del paciente en comparación con el valor representativo de un nivel de expresión génica de PTP4A1 en tumores de una población de pacientes que no obtienen ningún beneficio clínico del tratamiento.
- 20 4. El uso de un gen PTP4A1 para predecir la respuesta de un paciente con NSCLC al tratamiento con inhibidor de EGFR, en el que el inhibidor de EGFR es erlotinib.

Fig. 1

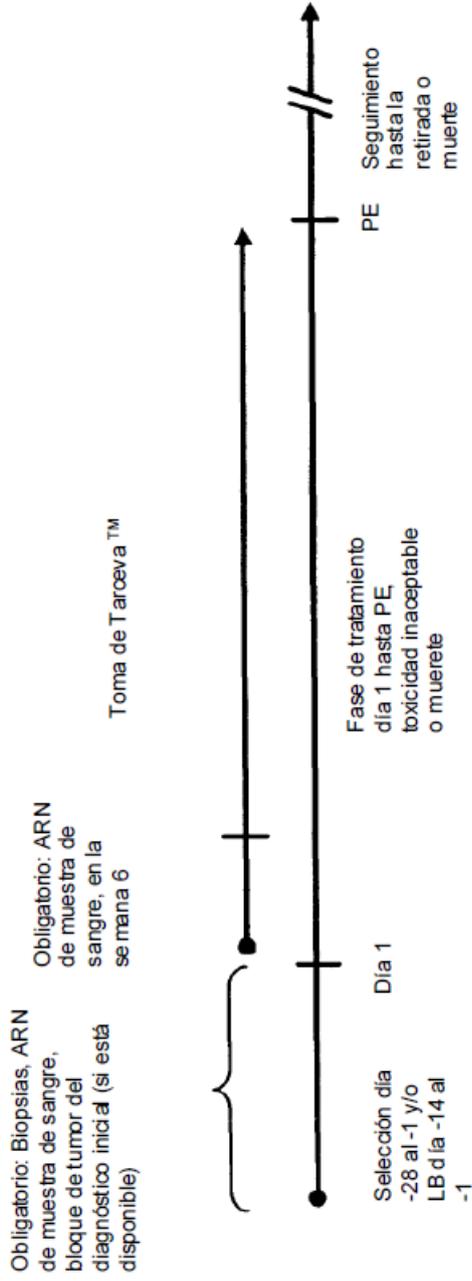


Fig. 2

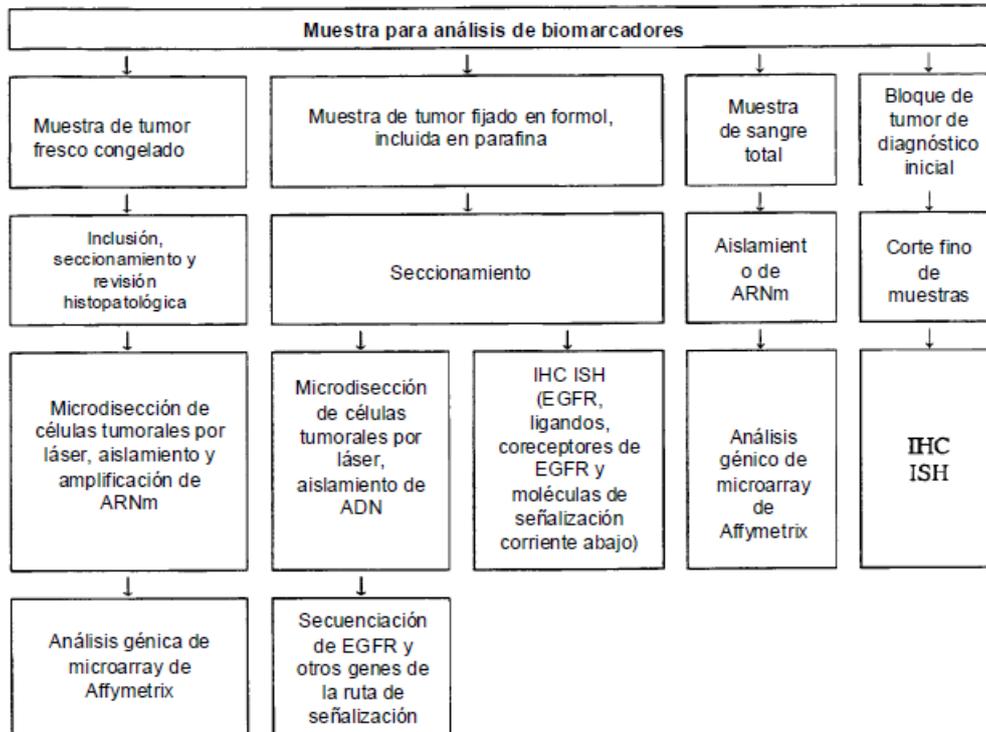


Fig. 3a

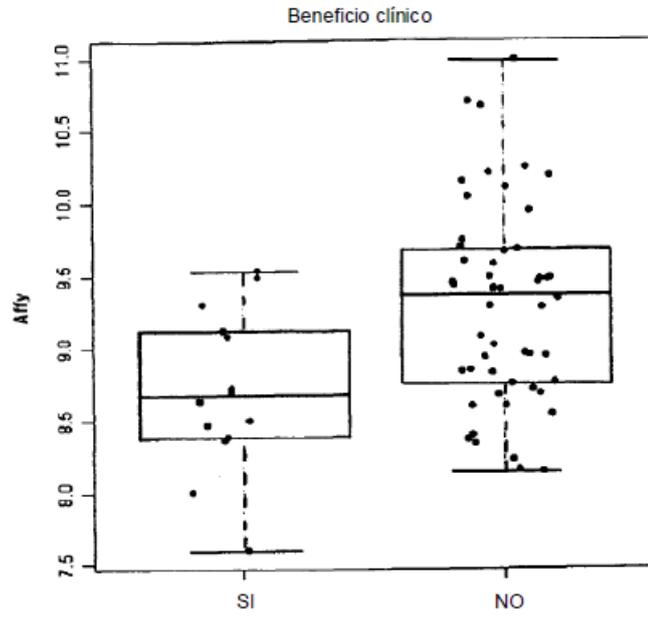


Fig. 3b

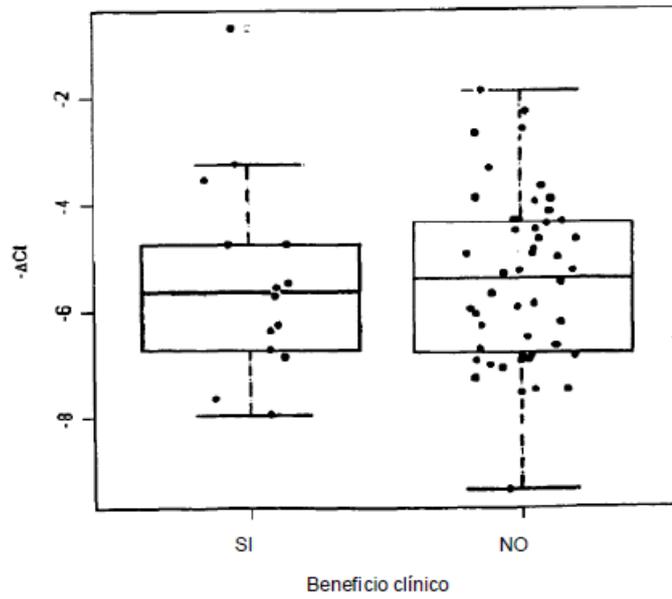


Fig. 3c

