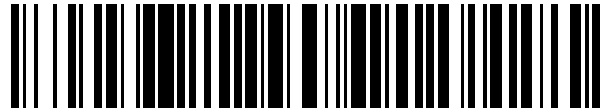


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 395 881**

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 08785429 (5)**

97 Fecha y número de publicación de la concesión europea: **19.09.2012 EP 2176430**

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

30 Prioridad:

14.08.2007 EP 07114302

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

15.02.2013

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 2 395 881 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 de inhibidor de EGFR

- 5 La presente invención proporciona un biomarcador que es predictivo del beneficio clínico del tratamiento con erlotinib en los pacientes con NSCLC.

10 Varios tumores malignos humanos se asocian a la expresión aberrante o a la sobreexpresión del receptor del factor de crecimiento epidérmico (EGFR). El EGF, el factor-? de crecimiento transformante (TGF-?) y varios otros ligandos se unen al EGFR, estimulando la autofosforilación del dominio intracelular de tirosina quinasa del receptor. Posteriormente se activa una diversidad de rutas intracelulares, y estos sucesos cadena abajo resultan en la proliferación de las células tumorales *in vitro*. Se ha propuesto que la estimulación de las células tumorales mediante el EGFR podría ser importante tanto para el crecimiento como para la supervivencia tumorales *in vivo*.

15 Los primeros datos clínicos obtenidos con TarcevaTM, un inhibidor de la EGFR tirosina quinasa, indican que el compuesto es seguro y que generalmente resulta bien tolerado a dosis que proporcionan la concentración efectiva pretendida (según se ha determinado a partir de datos preclínicos). Los ensayos clínicos de fase I y de fase II en
 20 pacientes con enfermedad avanzada han demostrado que TarcevaTM presenta una actividad clínica prometedora en un abanico de tumores epiteliales. En efecto se ha demostrado que TarcevaTM es capaz de inducir remisiones parciales duraderas en pacientes previamente tratados que presentan cáncer de cabeza y cuello y NSCLC (cáncer pulmonar de células no pequeñas) de un orden similar a la quimioterapia de segunda línea convencional, aunque con el beneficio añadido de un mejor perfil de seguridad que la quimioterapia y una mayor comodidad (tableta en lugar de administración intravenosa [i.v.]). Un ensayo aleatorizado de doble ciego controlado con placebo completado recientemente (BR.21) ha demostrado que el agente TarcevaTM por sí solo prolonga significativamente y
 25 mejora la supervivencia de los pacientes con NSCLC para los que ha fracasado la terapia estándar para la enfermedad avanzada.

30 TarcevaTM (erlotinib) es una molécula química pequeña; es un inhibidor selectivo potente, oralmente activo, de la tirosina quinasa del EGFR (EGFR-TKI).

El cáncer de pulmón es la causa principal de muertes relacionadas con el cáncer en Norteamérica y Europa. En los Estados Unidos, el número de muertes secundarias al cáncer de pulmón excede el número total combinado de muertes debidas a la segunda (colon), tercera (mama) y cuarta (próstata) causas principales de muerte por cáncer. Aproximadamente 75% a 80% de todos los cáncer de pulmón son NSCLC, presentándose aproximadamente 40%
 35 de los pacientes con una enfermedad localmente avanzada y/o no resecable. Este grupo típicamente incluye aquellos con enfermedad voluminosa de estadio IIIA y IIIB, excluyendo las efusiones pleurales malignas.

40 La incidencia bruta del cáncer de pulmón en la Unión Europea es de 52,5 y la tasa de muertes de 48,7 casos/100.000/año. Entre hombres las tasas son de 79,3 y 78,3, y entre mujeres, de 21,6 y 20,5, respectivamente. El NSCLC supone aproximadamente 80% de todos los casos de cáncer de pulmón. Aproximadamente 90% de la mortalidad por cáncer de pulmón en hombres, y 80% en mujeres, es atribuible al tabaquismo.

45 En los EUA, según la American Cancer Society, durante 2004 se produjeron aproximadamente 173.800 nuevos casos de cáncer de pulmón (93.100 en hombres y 80.700 en mujeres), representando aproximadamente 13% de todos los nuevos cánceres. La mayoría de pacientes muere como consecuencia de la enfermedad durante los dos años posteriores al diagnóstico. Para muchos pacientes con NSCLC, el tratamiento con éxito sigue siendo inalcanzable. Los tumores avanzados con frecuencia no admiten la cirugía y también pueden ser resistentes a dosis tolerables de radioterapia y quimioterapia. En ensayos aleatorizados, las quimioterapias de combinación actualmente más activas consiguieron tasas de respuesta de aproximadamente 30% a 40% y una tasa de supervivencia a 1 año de entre 35% y 40%. Esto constituye un avance respecto a la tasa de supervivencia a 1 año del 10% observada con sólo cuidados de apoyo.
 50

55 Hasta recientemente las opciones terapéuticas para los pacientes después de la recaída se limitaban al mejor cuidado de apoyo disponible o a la paliación. Un ensayo reciente que compara el docetaxel (Taxotere) con los mejores cuidados de apoyo ha demostrado que los pacientes con NSCLC podrían beneficiarse de la quimioterapia de segunda línea tras el fracaso de los regímenes de primera línea basados en cisplatino. Los pacientes de todas las edades y con estado funcional ECOG de 0, 1 ó 2 mostraron una supervivencia mejorada con el docetaxel, al igual que aquellos refractarios al tratamiento previo basado en platino. Entre los pacientes que no se beneficiaron de la terapia se incluían aquellos con una pérdida de peso del 10%, niveles elevados de lactato deshidrogenasa, compromiso multiorgánico o compromiso hepático. Además, el beneficio de la monoterapia con docetaxel no se extendió más allá del tratamiento de segunda línea. Los pacientes que recibieron docetaxel como tratamiento de
 60 tercera línea o posteriormente no mostraron ninguna prolongación de la supervivencia. El docetaxel como único agente se ha convertido en una terapia de segunda línea estándar para el NSCLC. Recientemente, otro ensayo de

fase III aleatorizado en la terapia de segunda línea del NSCLC ha comparado el pemetrexed (Alimta®) con el docetaxel. El tratamiento con pemetrexed resultó en una eficacia clínicamente equivalente aunque con significativamente menos efectos secundarios que el docetaxel.

5 Se admite desde hace mucho que existe una necesidad de desarrollar métodos para individualizar el tratamiento del cáncer. Con el desarrollo de los tratamientos personalizados del cáncer, existe un particular interés en las metodologías que podrían proporcionar un perfil molecular de la diana tumoral (es decir, aquéllas que son predictivas de beneficio clínico). Ya se ha establecido la prueba de principio para el perfilado de la expresión génica en el cáncer, con la clasificación molecular de tipos tumorales que no resultan evidentes basándose en los ensayos morfológicos e inmunohistoquímicos actuales. Se han diferenciado dos entidades patológicas separadas de diferente pronóstico, a partir de la única clasificación actual del linfoma de células B grandes difusas mediante la utilización del perfilado de expresión génica.

10 Por lo tanto, es un objetivo de la presente invención proporcionar biomarcadores de expresión que sean predictivos del beneficio clínico del tratamiento con erlotinib en los pacientes con NSCLC.

15 En un primer objetivo, la presente invención proporciona un método *in vitro* para predecir el beneficio clínico de un paciente de cáncer en respuesta al tratamiento con erlotinib, que comprende las etapas siguientes: determinar un nivel de expresión de un gen PTPRF en una muestra tumoral de un paciente y comparar el nivel de expresión del gen PTPRF con un valor representativo de un nivel de expresión del gen PTPRF en tumores de una población de pacientes que no deriva ningún beneficio clínico del tratamiento, en el que un nivel de expresión más alto del gen PTPRF en la muestra tumoral del paciente es indicativo de un paciente que obtendrá un beneficio clínico del tratamiento, en el que se ha definido el beneficio clínico como manifestar una respuesta objetiva o la estabilización de la enfermedad durante ≥ 12 semanas.

20 El acrónimo PTPRF se refiere a la proteína tirosina fosfatasa, tipo de receptor F. La Sec. Id. n° 1 muestra la secuencia de nucleótidos de la PTPRF, transcrito variante 1 y la Sec. Id. n° 2 muestra la secuencia de nucleótidos de la PRPRF humana transcrito variante 2.

25 La expresión "un valor representativo de un nivel de expresión de PTPRF en tumores de una población de pacientes que no obtiene ningún beneficio clínico del tratamiento" se refiere a una estimación del nivel de expresión medio del gen PTPRF en una población de pacientes que no obtienen un beneficio clínico del tratamiento. Se ha definido beneficio clínico como mostrar una respuesta objetiva o la estabilización de la enfermedad durante ≥ 12 semanas.

30 En una realización preferente adicional, el gen PTPRF muestra un nivel de expresión entre 1,1 y 1,8, preferentemente 1,1 y 1,6 ó más veces superior en la muestra tumoral del paciente que un valor representativo de la población de pacientes que no obtiene ningún beneficio clínico del tratamiento.

35 En una realización preferente adicional, el gen PTPRF muestra un nivel de expresión entre 1,2 y 1,8 ó más veces superior en la muestra tumoral del paciente que un valor representativo de la población de pacientes que no obtiene ningún beneficio clínico del tratamiento.

40 En una realización preferente, el nivel de expresión del gen marcador se determina mediante tecnología de micromatrices u otras tecnologías que evalúan los niveles de expresión del ARN, tales como la RT-PCR cuantitativa o mediante cualquier método que analice el nivel de expresión de la proteína respectiva, por ejemplo la inmunohistoquímica (IHC). La construcción y utilización de chips génicos son bien conocidos de la técnica. Ver las patentes US n° 5.202.231, n° 5.445.934, n° 5.525.464, n° 5.695.940, n° 5.744.305, n° 5.795.716 y n° 5.800.992. Ver también Johnston M., Curr. Biol. 8:R171-174, 1998; Lyer V.R. *et al.*, Science 283:83-87, 1999. Evidentemente el nivel de expresión génica puede determinarse mediante otros métodos que son conocidos por el experto en la materia, tales como, por ejemplo, transferencias northern, RT-PCR, PCR cuantitativa en tiempo real, extensión de cebadores, protección frente a ARNasa y perfilado de expresión del ARN.

45 El gen marcador utilizado en la presente invención puede combinarse con otros biomarcadores formando juegos de biomarcadores. Pueden construirse juegos de biomarcadores a partir de cualquier combinación de biomarcadores predictivos con el fin de realizar predicciones sobre el efecto del tratamiento con erlotinib en los pacientes con NSCLC. Los biomarcadores y juegos de biomarcadores indicados en la presente invención pueden utilizarse para, por ejemplo, predecir cómo los pacientes con NSCLC responderán a la intervención terapéutica con erlotinib.

50 El término "gen" tal como se utiliza en la presente memoria comprende variantes del gen. El término "variante" se refiere a secuencias de ácidos nucleicos que son sustancialmente similares a las secuencias de ácidos nucleicos indicadas mediante el número de acceso de GenBank. La expresión "sustancialmente similar" es perfectamente conocida por el experto en la materia. En particular, una variante génica puede ser un alelo que muestra intercambios de nucleótidos en comparación con la secuencia de nucleótidos del alelo más prevalente en la

población humana. Preferentemente, dicha secuencia de ácidos nucleicos sustancialmente similar presenta una similitud de secuencias respecto al alelo más prevalente de por lo menos 80%, preferentemente de por lo menos 85%, más preferentemente de por lo menos 90% y todavía más preferentemente de por lo menos 95%. El término "variantes" también pretende referirse a las variantes de procesamiento.

5 Entre las técnicas para la detección y cuantificación de la expresión génica de los genes indicados en la presente invención se incluyen, aunque sin limitarse a ellas, las transferencias northern, la RT-PCR, la PCR cuantitativa en tiempo real, la extensión de cebadores, la protección frente a ARNasa, el perfilado de expresión del ARN y técnicas relacionadas. Estas técnicas son bien conocidas por el experto en la materia; ver, por ejemplo, Sambrook J. *et al.*,
10 Molecular Cloning: A Laboratory Manual, tercera edición (Cold Spring Harbor Press, Cold Spring Harbor, 2000).

Entre las técnicas para la detección de la expresión de proteínas a partir de los genes respectivos descritos por la presente invención se incluyen, aunque sin limitarse a ellos, la inmunohistoquímica (IHC).

15 Las células de una muestra de tejido del paciente, por ejemplo de un tumor o biopsia de cáncer, pueden someterse a ensayo para determinar el patrón de expresión de uno o más biomarcadores. El éxito o fracaso de un tratamiento del NSCLC puede determinarse basándose en el patrón de expresión de biomarcadores de las células procedentes del tejido de ensayo (células de ensayo), por ejemplo del tumor o biopsia del cáncer, siendo relativamente similar o
20 diferente del patrón de expresión de un juego de control de uno o más biomarcadores. En el contexto de la presente invención, se ha encontrado que el gen de la Tabla 3 se encuentra regulado positivamente, es decir, muestra un nivel de expresión más alto, en tumores de pacientes que obtuvieron un beneficio clínico del tratamiento con erlotinib, en comparación con tumores de pacientes que no obtuvieron ningún beneficio clínico del tratamiento con erlotinib. De esta manera, en el caso de que las células de ensayo muestren un perfil de expresión de los biomarcadores que corresponda al de un paciente que respondió al tratamiento del NSCLC, es altamente probable o
25 predecible que el NSCLC o tumor del individuo responderá favorablemente al tratamiento con erlotinib. En contraste, en el caso de que las células de ensayo muestren un perfil de expresión de los biomarcadores que corresponda al de un paciente que no respondió al tratamiento del NSCLC, es altamente probable o predecible que el NSCLC o tumor del individuo no responderá al tratamiento con erlotinib.

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

Breve descripción de las figuras

40 La figura 1 muestra el diseño del estudio;
la figura 2 muestra un esquema del procesamiento de las muestras;
la figura 3a muestra los niveles de expresión de PTPRF frente al resultado clínico para el perfilado Genechip®;
la figura 3b muestra los niveles de expresión de PTPRF frente al resultado clínico para qRT-PCR, y
la figura 3c muestra la correlación entre las mediciones de GeneChip® y de qRT-PCR de la PTPRF.

45 Parte experimental

Justificación del estudio y del diseño del mismo

50 Recientemente, se han descrito mutaciones dentro del gen EGFR en el tejido tumoral de un subconjunto de pacientes con NSCLC y la asociación de estas mutaciones con la sensibilidad al erlotinib y al gefitinib (Pao W. *et al.*, 2004; Lynch *et al.*, 2004; Paez *et al.*, 2004). Para los pacientes agrupados de ambos estudios, se observó EGFR mutado en 13 de los 14 pacientes que respondieron al gefitinib y en ninguno de los 11 pacientes tratados con gefitinib que no respondieron. La prevalencia publicada de estas mutaciones es de 8% (2 de 25) en pacientes con NSCLC no seleccionados. Estas mutaciones se encontraron más frecuentemente en adenocarcinomas (21%), en
55 tumores procedentes de mujeres (20%) y en tumores procedentes de pacientes japoneses (26%). Estas mutaciones resultan en una actividad *in vitro* incrementada de EGFR y en una sensibilidad incrementada al gefitinib. La relación de las mutaciones con una enfermedad estable prolongada o la duración de la supervivencia no ha sido evaluada prospectivamente.

60 Basándose en análisis preliminares del estudio BR.21, aparentemente resultaba improbable que el beneficio de supervivencia observado se debiese únicamente a las mutaciones de EGFR ya que se mantuvo un beneficio de supervivencia significativo incluso tras excluir de los análisis a los pacientes que presentaban una respuesta objetiva (datos en archivo). Otros mecanismos moleculares también deben contribuir al efecto.

Basándose en la premisa de que existen cambios de los niveles de expresión génica que son predictivos de la respuesta/beneficio al tratamiento con Tarceva™, se utilizó el análisis de micromatrices para detectar dichos cambios.

5 Lo anterior requirió una población de estudio claramente definida tratada con monoterapia de Tarceva™ tras el fracaso de la terapia de primera línea. Basándose en la experiencia del estudio BR.21, se definió la población beneficiada como aquella que presentaba una respuesta objetiva o la estabilización de la enfermedad durante 12 semanas. Se analizaron bases de datos clínicos y de micromatrices según un plan estadístico predefinido.

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

15 Se recogió una segunda muestra tumoral simultáneamente y se almacenó en parafina (fijada con formalina e incluida en parafina, FFIP). Esta muestra se analizó para alteraciones en la ruta de señalización de EGFR.

20 La capacidad de llevar a cabo biopsias tumorales mediante broncoscopia era un requisito previo para este estudio. La broncoscopia es un procedimiento estándar para confirmar el diagnóstico del cáncer de pulmón. Aunque generalmente seguro, sigue existiendo riesgo de complicaciones, por ejemplo de sangrado.

Este estudio fue una primera etapa hacia una terapia individualizada para pacientes con NSCLC refractario. Esta terapia individualizada permitirá a los médicos tratantes seleccionar el agente más apropiado de entre los fármacos existentes para dicha indicación.

25 Tras hacerse disponible una terapia individualizada, el beneficio para cada paciente en el futuro superará el riesgo que deben correr los pacientes en el presente estudio:

- se incrementarán las tasas de respuesta / el número de pacientes beneficiados,
- se reducirá el riesgo de efectos adversos debido al tratamiento inefectivo.

30 Justificación de la selección de dosis

35 Se administró Tarceva™ por vía oral una vez al día a una dosis de 150 mg hasta la progresión de la enfermedad, toxicidades intolerables o muerte. La selección de dicha dosis se basó en parámetros farmacocinéticos, así como en el perfil de seguridad y tolerabilidad de dicha dosis observado en ensayos de fase I, II y III en pacientes fuertemente pretratados con cáncer avanzado. Los niveles de fármaco observados en el plasma de pacientes con cáncer que recibían la dosis de 150 mg/día eran consistentemente superiores a la concentración plasmática media de 500 ng/ml que era el objetivo para la eficacia clínica. BR.21 mostró un beneficio de supervivencia con dicha dosis.

40 Objetivos del estudio

45 El objetivo principal era la identificación de genes expresados diferencialmente que fuesen predictivos de beneficio (CR, RP o EE? 12 semanas) del tratamiento con Tarceva™. La identificación de genes expresados diferencialmente predictivos de "respuesta" (CR, RP) al tratamiento con Tarceva™ era un objetivo adicional importante.

Los objetivos secundarios eran evaluar las alteraciones en las rutas de señalización de EGFR con respecto al beneficio del tratamiento.

50 Diseño del estudio

Descripción general del diseño del estudio y del régimen de dosificación

55 Estudio de fase II de etiqueta abierta de identificación de marcadores predictivos. El estudio se llevó a cabo en aproximadamente 26 sitios en aproximadamente 12 países. Se incluyeron durante un periodo de 12 meses 264 pacientes con NSCLC avanzado tras el fracaso de por lo menos un régimen anterior de quimioterapia. Se administró Tarceva™ oral continuo a una dosis de 150 mg/día. Se permitieron reducciones de las dosis basándose en la tolerabilidad a la terapia farmacológica. Se evaluaron parámetros clínicos y de laboratorio con el fin de evaluar el control de la enfermedad y la toxicidad. El tratamiento se prolongó hasta la progresión de la enfermedad, toxicidad inaceptable o muerte. Se ilustra en la figura 1 el diseño del estudio.

60 Se obtuvieron muestras de tejido tumoral y de sangre para los análisis moleculares destinados a evaluar los efectos de Tarceva™ y para identificar los subgrupos de pacientes que se beneficiarán de la terapia.

Evaluaciones de los marcadores predictivos

Se obtuvieron biopsias de tumores en las 2 semanas anteriores al inicio del tratamiento. Se recogieron dos muestras diferentes:

5 la primera muestra en todos los casos se congeló inmediatamente en N₂ líquido.

La segunda muestra se fijó en formalina y se incluyó en parafina.

10 El tejido congelado instantáneamente se consideró prioritario en el presente estudio.

La figura 2 muestra un esquema del procesamiento de las muestras.

Análisis de micromatrices

15 Las muestras congeladas instantáneamente se utilizaron para la microdissección por captura con láser (MCL) de células tumorales con el fin de extraer el ARN tumoral y el ARN del tejido circundante al tumor. El ARN se analizó en chips micromatrices Affymetrix (HG-U133A) para establecer el perfil de expresión génica del tumor del paciente. Se utilizó un control de calidad de los chips Affymetrix para seleccionar aquellas muestras de calidad adecuada para la comparación estadística.

Análisis de biomarcadores individuales en tejido fijado en formalina e incluido en parafina

25 Se utilizó la segunda biopsia tumoral, la muestra FFIP, para llevar a cabo análisis de mutaciones del ADN, IHC e ISH, tal como se describe posteriormente. Se llevaron a cabo análisis similares de tejido recogido en el momento del diagnóstico inicial.

30 Se analizaron mediante secuenciación del ADN el estado de mutación del ADN de los genes codificantes de EGFR y de otras moléculas participantes en la ruta de señalización del EGFR. La amplificación génica del EGFR y genes relacionados se estudiaron mediante FISH.

Los análisis de expresión de proteínas incluyeron análisis inmunohistoquímicos [IHC] del EGFR y de otras proteínas en la ruta de señalización del EGFR.

35 Evaluaciones de la respuesta

Se utilizaron los criterios RECIST (medición unidimensional de los tumores) para evaluar la respuesta. Estos criterios pueden encontrarse en el enlace siguiente:

40 <http://www.eortc.be/recist/>

Observar que:

45 Para asignar un estatus de CR o de PR, deben confirmarse cambios en las mediciones de los tumores en evaluaciones repetidas separadas por lo menos por 4 semanas en cualquier momento durante el periodo de tratamiento.

50 En el caso de una EE, las mediciones de seguimiento deben cumplir los criterios de EE por lo menos en una ocasión después del inicio del estudio a un intervalo mínimo de 6 semanas.

En el caso de una EE mantenida, las mediciones de seguimiento deben cumplir los criterios de EE por lo menos en una ocasión después del inicio en el estudio con una duración del mantenimiento de por lo menos 12 semanas.

Evaluación de la supervivencia

55 Se llevó a cabo una comprobación periódica del estado cada 3 meses mediante una visita del paciente a la clínica o telefónicamente. Se registraron todas las muertes. Al final del estudio fue necesaria una confirmación definitiva de la supervivencia de cada paciente.

60 Métodos

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

5 Todo el procesamiento de las muestras de biopsia fue realizado por un laboratorio patológico de referencia; se enviaron muestras de tejido fresco congelado procedente de los sitios de investigación a las instalaciones de manipulación de muestras clínicas en Roche Basel y de ahí al laboratorio de patología para el procesamiento posterior. Se utilizó la microdissección con captura con láser para seleccionar las células tumorales del tejido circundante. Tras la MCL, se purificó el ARN a partir del material tumoral enriquecido. A continuación, el laboratorio de patología llevó a cabo varias etapas para realizar una estimación de la concentración y calidad del ARN.

10 Las ARNasas son enzimas degradadores del ARN y se encuentran en todas partes y por lo tanto todos los procedimientos en los que se utilice ARN deben controlarse estrictamente para minimizar la degradación del ARN. La mayoría de las especies de ARNm presentan vidas medias bastante cortas y por lo tanto se consideran bastante inestables. Por lo tanto, resulta importante llevar a cabo comprobaciones de integridad y cuantificaciones del ARN antes de cualquier ensayo.

15 Puede evaluarse la concentración y perfil de calidad del ARN utilizando un instrumento de Agilent (Agilent Technologies, Inc., Palo Alto, CA) denominado 2100 Bioanalyzer[®]. El software del instrumento genera un número de integridad del ARN (RIN), una estimación cuantitativo (Schroeder A. *et al.*, The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol. 7: 3, 2006) y calcula las proporciones ribosómicas de la muestra del ARN total. El RIN se determina a partir del registro electroforético completo de la muestra de ARN, que de esta manera incluye la presencia o ausencia de productos de degradación.

20 Se analizó la calidad del ARN utilizando un 2100 Bioanalyzer[®]. Sólo se seleccionaron muestras con por lo menos un pico de ARNr sobre el ruido de poli-I añadido y suficiente ARN para el análisis posterior en la plataforma Affymetrix. Se envió el ARN purificado al Roche Centre for Medical Genomics (RCMG, Basel, Suiza) para el análisis mediante micromatrices. Se recibieron 122 muestras de ARN del laboratorio de patología para el procesamiento posterior.

25 **Marcaje de la diana de muestras de ARN de tejido**

30 El marcaje de la diana se llevó a cabo siguiendo el protocolo de amplificación de marcaje de la diana en dos ciclos de Affymetrix (Affymetrix, Santa Clara, California) siguiendo las instrucciones del fabricante.

El método se basa en el procedimiento estándar de Eberwine de amplificación lineal, aunque utiliza dos ciclos de este procedimiento par agenerar suficiente ARNc marcado para la hibridación con una micromatriz.

35 La cantidad total de ARN utilizada en la reacción de marcaje fue de 10 ng para aquellas muestras en las que se disponía de más de 10 ng de ARN; en caso de disponerse de menos de esta cantidad, o si no se disponía de datos cuantitativos (debido a que la concentración de ARN era muy baja), se utilizó la mitad de la muestra total en la reacción. Los rendimientos de las reacciones de marcaje fueron de entre 20 y 180 µg de ARNc. Se introdujo una etapa de normalización al nivel de la hibridación, en la que se utilizaron 15 µg de ARNc para cada muestra.

40 Se utilizó ARN de referencia humano (Stratagene, Carlsbad, CA, USA) a modo de muestra de control en el flujo de trabajo con cada lote de muestras. Se utilizaron 10 ng de dicho ARN como entrada, conjuntamente con las muestras de ensayo con el fin de verificar que los reactivos de marcaje y de hibridación funcionaban tal como se esperaba.

45 **Hibridaciones con micromatrices**

Las micromatrices Affymetrix HG-U133A contienen más de 22.000 juegos de sondas con diana en aproximadamente 18.400 transcritos y variantes, representando aproximadamente 14.500 genes bien caracterizados.

50 La hibridación para todas las muestras se llevó a cabo siguiendo las instrucciones de Affymetrix (Affymetrix Inc., manual técnico de análisis de la expresión, 2004). Brevemente, para cada muestra se fragmentaron 15 µg de ARNc marcado con biotina en presencia de cationes divalentes y calor, y se hibridaron durante la noche con matrices Affymetrix HG-U133A de oligonucleótidos del genoma completo. Al día siguiente las matrices se tiñeron con estreptavidina-ficoeritrina (Molecular Probes, Eugene, OR) siguiendo las instrucciones del fabricante. Seguidamente se escanearon las matrices utilizando un escáner GeneChip 3000 (Affymetrix) y se calcularon automáticamente las intensidades de las señales con el software operativo GeneChip (GCOS) versión 1.4 (Affymetrix).

55 **Análisis estadístico**

60 El análisis de los datos de Affymetrix[™] consistió en cinco etapas principales.

La etapa 1 fue la de control de calidad. El objetivo era identificar y excluir del análisis datos de las matrices con un perfil de calidad deficiente.

La etapa 2 fue de preprocesamiento y normalización. El objetivo era crear una "base de datos para el análisis" normalizada y escalada que permitiese la comparación entre matrices. Comprendía la estimación y sustracción del ruido de fondo, y el agrupado ("sumarización") y escalado de las sondas.

5 La etapa 3 fue de exploración y descripción. El objetivo era identificar los potenciales sesgos y las fuentes de variabilidad. Consistía en la aplicación de técnicas de análisis descriptivo multivariante y univariante para identificar las covariables con influencia.

10 La etapa 4 fue de modelaje y ensayo. El objetivo era identificar una lista de los marcadores candidatos basándose en la evaluación estadística de las diferencias en el nivel medio de expresión entre los pacientes "con beneficio clínico" y "sin beneficio clínico". Consistía del ajuste a un modelo estadístico adecuado para cada juego de sondas y la obtención de una medida de la significación estadística.

15 La etapa 5 fue un análisis de robustez. El objetivo era generar una lista cualificada de marcadores candidatos que no dependiesen fuertemente de los métodos de preprocesamiento y de las premisas estadísticas. Consistía en reiterar el análisis utilizando diferentes enfoques metodológicos y extraer una intersección de las listas de candidatos obtenidas.

Todos los análisis se llevaron a cabo utilizando el paquete de software R.

20 Etapa 1: control de calidad

25 La evaluación de la calidad de los datos se basó en la comprobación de varios parámetros. Entre ellos se incluían parámetros de calidad de GeneChip™ de Affymetrix, en particular: el factor de escala, el porcentaje de expresión presente ("present call") y el nivel medio de fondo. Esta etapa también incluía la inspección visual de imágenes virtuales de los chips para la detección de problemas localizados de hibridación y la comparación de cada chip con un chip virtual en la mediana para la detección de cualquier desviación no habitual respecto a la mediana del comportamiento. También se llevó a cabo un análisis de correlación entre chips para detectar las muestras atípicas. Además, se consideraron medidas complementarias de la calidad del ARN obtenidas a partir del análisis de muestras de ARN con el Bioanalyzer™ 2100 de Agilent.

30 Basándose en estos parámetros, se excluyeron del análisis los datos procedentes de 20 matrices. De esta manera, se incluyeron en el análisis datos procedentes de un total de 102 matrices, que representaban 102 pacientes. Se proporciona en la Tabla 1 la descripción clínica de este conjunto de 102 muestras.

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%)
	SÍ	21 (20,6%)
SEXO	FEMENINO	25 (24,5%)
	MASCULINO	77 (74,5%)
ETNICIDAD	CAUCÁSICA	65 (63,7%)
	ORIENTAL	37 (36,3%)
Histología	ADENOCARCINOMA	35 (34,3%)
	ESCAMOSO	53 (52,0%)
	OTRAS	14 (13,7%)
¿Ha fumado alguna vez?	NO	20 (19,6%)
	SÍ	82 (80,4%)

Etapa 2: Preprocesamiento y normalización de los datos

5 Se utilizó el algoritmo rma (Irizarry R.A. *et al.*, Summaries of Affymetrix GeneChip probe level data. Nucl. Acids Res. 31(4): e15, 2003) para el preprocesamiento y la normalización. Se utilizó el algoritmo mas5 (AFFYMETRIX, GeneChip® Expression: Data Analysis Fundamentals. 2004, AFFYMETRIX) para realizar análisis de detección para los diferentes juegos de sondas. Los juegos de sondas con resultados de “ausente” o “marginal” en todas las muestras se eliminaron de los análisis posteriores; se eliminaron 5.930 juegos de sondas del análisis basándose en este criterio. Por lo tanto, la base de datos para el análisis consistió de una matriz con 16.353 (de un total de 22.283) juegos de sondas medidos en 102 pacientes.

Etapa 3: Descripción y exploración de los datos

15 Se llevó a cabo un análisis exploratorio descriptivo para identificar sesgos potenciales y las fuentes principales de variabilidad. Se cribó un conjunto de covariables con un impacto potencial sobre los perfiles de expresión génica. Comprendía variables tanto técnicas como clínicas. Entre las covariables técnicas se incluían: fecha del procesamiento del ARN (posteriormente denominado “lote”), RIN (como medida de la calidad/integridad del ARN), operador y centro de recolección de las muestras. Entre las covariables clínicas se incluían: tipo histológico, situación de tabaquismo, grado del tumor, puntuación de rendimiento (Oken M.M. *et al.*, Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am. J. Clin. Oncol. 5(6): 649-55, 1982), datos demográficos, estatus del respondedor y estatus de beneficio clínico.

Entre las herramientas de análisis se incluían el ANOVA univariante y el análisis de componentes principales. Para cada una de dichas covariables, se aplicó el ANOVA variante independientemente a cada juego de sondas.

25 Se identificó un efecto significativo de la variable lote. En la práctica, la variable lote captura diferencias entre fechas de procesamiento de las muestras y del gráfico del chip Affymetrix. Tras comprobar que la variable lote era prácticamente independiente de las variables de interés, se corrigió el efecto de lote utilizando el método descrito en Johnson W.E., C. Li y A. Rabinovic, Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostat. 8(1): 118-127, 2007.

La base de datos normalizados tras la corrección del efecto de lote sirvió como base de datos para el análisis en los análisis posteriores.

35 La histología y el RIN fueron dos variables importantes adicionales subrayadas en el análisis descriptivo.

Etapa 4: Modelaje y ensayo de los datos

40 Se ajustó independientemente un modelo lineal a cada juego de sondas. Las variables incluidas en el modelo se informan en la Tabla 2. Se estimaron los parámetros del modelo mediante la técnica de máxima probabilidad. El parámetro correspondiente a la variable “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 de la expresión génica	Dependiente del tipo (Y_{ip})	Valor log2 intensidad del juego de sondas i en el paciente p .
Punto de corte	Media global (μ)	
Beneficio clínico	Factor predictivo de interés (X1)	SÍ/NO
Histología	Covariable de ajuste (X2)	ADENO. / ESCAM. OTROS
ETNIA	Adj. Cov. (X3)	ORIENT. / CAUCAS.
SEXO	Adj. Cov. (X4)	FEMENINO / MASCULINO
RIN	Adj. Cov. (X5)	[2,...,7.9]
FUMADOR	Adj. Cov. (X6)	ACTUAL/ANTERIOR/NUNCA
Estadio	Adj. Cov. (X7)	NO RESECC.III / IV

45 Para cada juego de sondas i , el objetivo de la prueba estadística era rechazar la hipótesis de que los niveles medios de expresión en pacientes con beneficio clínico y en pacientes sin beneficio clínico eran iguales, considerando las

otras covariables de ajuste listadas en la Tabla 2. Formalmente, la hipótesis nula de igualdad se sometió a ensayo frente a una alternativa de dos colas. Bajo la hipótesis nula, la distribución del estadístico t para este ensayo sigue una distribución t de Student con 92 grados de libertad. Se proporcionan los valores de p correspondientes en la Tabla 3.

5 La elección de modelo lineal se basó en dos motivos. En primer lugar, el modelaje lineal es un enfoque versátil, bien caracterizado y robusto que permite el ajuste de las variables de confusión al estimar el efecto de la variable de interés. En segundo lugar, dado el tamaño muestral, de 102, y la normalización y escalado de la base de datos, la premisa de la distribución normal era razonable y estaba justificada.

10 Para cada juego de sondas se evaluó la premisa de homogeneidad de la varianza utilizando pruebas de Fligner-Killeen basadas en los residuales del modelo. El análisis consistió de 3 etapas:

- 15
1. Ensayo de cada variable categórica para la homogeneidad de la varianza residual
 2. Identificar la variable V con el valor de p más bajo
 3. En el caso de que el valor de p más bajo sea inferior a 0,001, reajuste del modelo que permita que las variables V de diferente nivel presenten una varianza diferente

20 Etapa 5: Robustez

El objetivo del análisis de robustez fue reducir el riesgo de que los resultados del análisis fuesen artefactuales y un resultado de las etapas de preprocesamiento o de las premisas subyacentes al análisis estadístico. Se consideraron los tres aspectos siguientes: a) inclusión o exclusión de unos cuantos chips adicionales en la etapa de control de la calidad, b) algoritmo de preprocesamiento y normalización, c) premisas estadísticas y enfoque de los ensayos.

25 Se definió la lista de marcadores candidatos como el subconjunto de genes declarados consistentemente como significativos en diferentes contextos de análisis. Las diferentes opciones de análisis aplicadas fueron las siguientes:

30 a) Se identificó un subconjunto adicional de 8 chips basándose en criterios de control de calidad más restrictivos. Se definió una "base de datos reducida" mediante la exclusión de dichos 8 chips. b) Se identificó MAS5 como alternativa a rma para el preprocesamiento y la normalización. MASS utiliza diferentes métodos para la estimación del fondo, el agrupado de las sondas y la normalización c) Se utilizaron dos ensayos estadísticos adicionales.

- 35
- a. Una prueba de Wilcoxon para la diferencia entre la presencia y la ausencia de beneficio clínico, y
 - b. un ensayo de cocientes de probabilidad (ECP) que somete a ensayo el modelo de regresión logística, en el que el beneficio clínico era la variable de respuesta y la expresión génica, la covariable. Dichos dos ensayos adicionales se basan en un conjunto diferente de premisas estadísticas subyacentes. Para cada juego de sondas se realizó un ECP tras una prueba Chi-cuadrado con 1 grado de libertad.

40 En resumen, se consideraron dos conjuntos de muestras (la base de datos "completa" y la base de datos "reducida") y 2 algoritmos de preprocesamiento (mas5 y rma); esto resultó en cuatro bases de datos de análisis diferentes. A cada una de dichas cuatro bases de datos se aplicaron tres pruebas estadísticas diferentes. Por lo tanto, para cada juego de sondas se calcularon tres valores de p. En cada base de datos para el análisis, se aplicó un criterio compuesto para identificar la lista de genes regulados diferencialmente. Dicho criterio compuesto se definió como:

45 valor de p máximo inferior a 0,05 y valor de p mínimo inferior a 0,001. El análisis de robustez utilizando el criterio 1 para identificar los genes marcadores rindió PTPRF como marcador predictivo para el tratamiento con erlotinib.

Tabla 3: Marcador génico de beneficio clínico basado en el análisis de robustez tras la aplicación del criterio compuesto

La columna 1 es el identificador de Affymetrix del juego de sondas. La columna 2 es el número de acceso de GenBank de la secuencia génica correspondiente. La columna 3 es el nombre génico oficial correspondiente. La columna 4 es el factor medio ajustado de cambio correspondiente del nivel de expresión entre el paciente con y sin beneficio clínico, según estimaciones del modelo lineal. La columna 5 es el valor de p para la prueba de las diferencias en el nivel de expresión en pacientes con y sin beneficio clínico derivadas del modelo lineal. La columna 6 es el intervalo de confianza al 95% para el factor medio ajustado de cambio del nivel de expresión.					
ID de juego de sondas Affymetrix	GenBank	Gen	Factor medio ajustado de cambio	valor de P	IC al 95%

200637_s_at	NM_002840 (Sec. Id. nº 1) NM_130440 (Sec. Id. nº 2)	PTPRF	1,35	1,2E-3	1,1 , 1,6
200635_s_at	NM_002840 (Sec. Id. nº 1) NM_130440 (Sec. Id. nº 2)	PTPRF	1,49	1,7E-4	1,2 , 1,8

Análisis estadísticos adicionales

5 Para el marcador candidato seleccionado PTPRF, se llevaron a cabo los análisis adicionales siguientes en un entorno validado por expertos estadísticos independientes:

- Regresión Cox univariante para SSP (supervivencia sin progresión) del análisis Affymetrix primario,
- Regresión logística univariante para beneficio clínico del análisis Affymetrix primario, y
- Regresión Cox univariante para la supervivencia del análisis Affymetrix primario

10 Se presentan los resultados de estos análisis posteriormente. Son consistentes con los resultados del análisis primario y confirman la elección del marcador seleccionado.

15 Resultados: Regresión Cox univariante para SSP (supervivencia sin progresión) del análisis Affymetrix primario:

Gen	nº de pacientes	cociente de riesgos	IC al 95% para el cociente de riesgos	Valor de p
PTPRF	102	0,5	0,34; 0,73	0,004

Resultados: Regresión Cox univariante para beneficio clínico del análisis Affymetrix primario:

Gen	nº de pacientes	Proporción de probabilidades	IC al 95% para la proporción de probabilidades	Valor de p
PTPRF	102	5,01	1,89; 13,33	0,0012

20 Resultados: Regresión Cox univariante para la supervivencia del análisis Affymetrix primario:

Gen	nº de pacientes	cociente de riesgos	IC al 95% para el cociente de riesgos	Valor de p
PTPRF	102	0,62	0,39; 0,97	0,0377

25 qRT-PCR

Se sintetizó ADNc utilizando la mezcla de síntesis de primera cadena SuperScript™ III para qRT-PCR (Invitrogen, CA, USA) siguiendo las instrucciones del fabricante aunque sin inclusión de un digerido de ARNasa H.

30 Se llevó a cabo una PCR cuantitativa utilizando ensayos de expresión génica TaqMan® en un sistema de detección de secuencias ABI PRISM® 7900HT siguiendo las recomendaciones del fabricante (Applied Biosystems, CA, USA). Todos los ensayos se llevaron a cabo por triplicado.

35 Los cebadores y sondas utilizados cruzaban los límites de los exones o se encontraban dentro de la secuencia de interés de la sonda Affymetrix Genechip®. Se incluyeron dos genes de mantenimiento como controles endógenos: β -2-microglobulina (B2M; ensayo Hs99999907_m1) e hipoxantinafosforibosil transferasa (HPRT; ensayo Hs99999909_m1).

40 Todos los análisis incluían una muestra calibradora (ARN total MVP™ procedente de pulmón humano adulto; Stratagene, CA, USA) y una curva estándar. Se utilizó ARN humano total de referencia universal (Stratagene, CA, USA) como molde para las curvas estándares de PTPRF. Todas las muestras se midieron por triplicado.

Se llevó a cabo la cuantificación relativa utilizando el método $-\Delta\text{Ct}$.

Resultados

5 Tal como se ha informado anteriormente, se determinaron los perfiles de expresión génica Affymetrix Genechip® para los 102 pacientes incluidos en el presente estudio. De entre estos pacientes, se obtuvieron resultados de qRT-PCR para 75 (Tabla 4). Las características demográficas y clínicas de los pacientes con resultados de qRT-PCR eran similares a las de la población entera (n=264) y a las de los pacientes para los que se disponía de un perfil de expresión génica Genechip®.

10 Tabla 4: Características de línea base: pacientes con análisis de qRT-PCR (n=75)

Característica	
Edad (mediana, rango)	62 (39-85)
Sexo, n (%)	
Masculino	19 (25)
Sexo, n (%)	
Femenino	56 (75)
Estado funcional 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)
Estadio de la enfermedad, n (%)	
IIIB	22 (29)
IV	53 (71)
Número de regímenes de quimioterapia anteriores, n (%)	
0	19 (25)
1	36 (48)
≥2	20 (27)
Etnicidad, n (%)	
Caucásica	51 (68)
Asiática	24 (32)
Historia de tabaquismo, n (%)	
Nunca	12 (16)
Actual	24 (32)
Anterior	39 (52)

15 De los 75 pacientes con resultados de qRT-PCR, 4 (5%) presentaba una respuesta parcial (RP), 23 (31%) presentaba una EE, 39 (52%) presentaba una EP y 9 (12%) no eran evaluables. Estos resultados son muy similares a los observados en la población de estudio completa (n=264).

20 La figura 3 muestra los niveles relativos de ARNm para PTPRF en pacientes individuales, evaluados mediante el perfilado Affymetrix Genechip® y qRT-PCR. La figura 3a muestra los niveles de expresión frente al resultado clínico para el perfilado Genechip® y la figura 3b muestra los niveles de expresión frente a la qRT-PCR.

25 Se observó una buena correlación entre las mediciones de Genechip® y de qRT-PCR del transcrito de ARNm de PTPRF (figura 3c, p de Pearson=0,76, p<0,01). Tal como se observó en el perfilado de Genechip®, los niveles de ARNm de PTPRF evaluados utilizando la qRT-PCR aparentemente se correlacionaban con la respuesta al erlotinib, observándose niveles más altos en los respondedores que en los no respondedores.

Comentario

30 Mediante el análisis de muestras de tejido con tecnología de micromatrices de oligonucleótidos de alta densidad y aplicando el modelaje estadístico a los datos, los presentes inventores pudieron identificar los genes cuyos niveles de expresión podían ser predictivos de los pacientes que obtendrían un beneficio clínico del tratamiento con erlotinib.

Se aplicó un compuesto criterio (definido anteriormente). Resultó en que PTPRF era un marcador predictivo para el tratamiento con erlotinib.

- 5 El gen PTPRF, situado en el cromosoma 1p34, codifica un miembro proteína de la familia de la proteína tirosina fosfatasa (PTP). Presenta una región extracelular, una sola región transmembranal y dos dominios catalíticos intracitoplasmáticos en tándem y de esta manera representa una PTP de tipo receptor. La región extracelular contiene tres dominios de tipo Ig y nueve dominios no de tipo Ig, de manera similar a la de la molécula de adhesión a células neurales.
- 10 En el presente estudio, se encontró que PTPRF presentaba una regulación relativamente positiva en pacientes que obtenían un beneficio clínico del tratamiento con erlotinib. Este resultado puede interpretarse en el contexto de los informes publicados que demuestran el papel potencial de este gen en diferentes mecanismos importantes de la tumorigénesis.
- 15 En primer lugar, se estableció claramente que EGFR es un sustrato de la PTPRF. En una investigación detallada, se caracterizó adicionalmente la interacción entre EGFR y PTPRF y se demostró que era compleja y que se encontraba estrechamente controlada. Estas observaciones condujeron a los presentes inventores a proponer que la PTPRF desempeña un papel importante y directo en el control de la señalización cadena abajo a partir del receptor de EGFR. En otra línea de evidencia, se observó que la PTPRF presentaba una actividad supresora tumoral que actuaba a través de un efecto de inhibición de la migración celular y posiblemente la inducción de apoptosis. Se indujo adicionalmente el mecanismo por el que la PTPRF controla el proceso de la migración celular. Dos estudios han demostrado que dicha proteína funciona mediante una interacción compleja con el complejo de la cadherina-E, mediada por una regulación directa de la actividad de la β -catenina.
- 20
- 25 La interacción directa con EGFR y una actividad supresora tumoral bien caracterizada son dos características prominentes que convierten a la PTPRF en un marcador particularmente convincente de la respuesta al erlotinib.

LISTADO DE SECUENCIAS

- 5 <110> GENETHON Centre National de la Recherche Scientifique (CNRS)
 <120> VECTORES VÍRICOS ADENOASOCIADOS PARA LA EXPRESIÓN DE LA DISFERLINA
 <130> G143-B-23719 PCT
- 10 <150> FR07.05479
 <151> 2007-07-26
 <160> 16
- 15 <170> PatentIn versión 3.3
 <210> 1
 <211> 29
 <212> ADN
 20 <213> secuencia artificial
 <220>
 <223> cebador NcoI exones 1-28
- 25 <400> 1
 ttccatgggc atgctgaggg tcttcatcc 29
 <210> 2
 <211> 42
 30 <212> ADN
 <213> secuencia artificial
 <220>
 <223> cebador HindIII MfeI exones 1-28
- 35 <400> 2
 ttcaattggg aagcttgccc accttgctca tcgacagccc gg 42
 <210> 3
 40 <211> 40
 <212> ADN
 <213> secuencia artificial
 <220>
 45 <223> cebador SpeI MluI exón 29-55
 <400> 3
 ttactagtgg acgctccag gctgggagta tagcatcacc 40
- 50 <210> 4
 <211> 45
 <212> ADN
 <213> secuencia artificial
- 55 <220>
 <223> cebador NotI exón 29-55
 <400> 4
 ttgcggccgc ctacagggca ggagagtctt cagctgaagg gcttc 45
- 60 <210> 5
 <211> 24
 <212> ADN

<213> secuencia artificial
 <220>
 <223> cebador HindIII-SD5'
 5 <400> 5
 ttaagcttag catgtggaac ctgg 24
 <210> 6
 10 <211> 27
 <212> ADN
 <213> secuencia artificial
 <220>
 15 <223> cebador MfeI-SD3'
 <400> 6
 ttcaattgag ctggagtgg ggggtgc 27
 <210> 7
 20 <211> 30
 <212> ADN
 <213> secuencia artificial
 <220>
 25 <223> cebador SpeI-SA5'
 <400> 7
 30 ttactagtgc aaattaggac cgagagtcag 30
 <210> 8
 <211> 28
 <212> ADN
 <213> secuencia artificial
 35 <220>
 <223> cebador MluI-SA
 <400> 8
 40 ttacgcgtgg gagggggaac cggtcact 28
 <210> 9
 <211> 18
 <212> ADN
 45 <213> secuencia artificial
 <220>
 <223> cebador Exon28.f
 50 <400> 9
 ctcaaccggg ctgtcgat 18
 <210> 10
 <211> 26
 55 <212> ADN
 <213> secuencia artificial
 <220>
 <223> cebador Exon29.r
 60 <400> 10
 gtcggtgtgt gtagtacatc ttctca 26

ES 2 395 881 T3

<210> 11
 <211> 12
 <212> ADN
 <213> secuencia artificial
 5
 <220>
 <223> cebador Exons2829.s
 <400> 11
 10 caaggctggg ag 12
 <210> 12
 <211> 264
 <212> ADN
 15 <213> secuencia artificial
 <220>
 <223> intrón 28 del gen humano de la disferlina
 20 <400> 12
 gtgggcagca tgtggaacct ggcgagcccc atccccggca agctctcaag ccatgctggt 60
 ggggacgact gaatgccagg gcccttcaact gggctatttc acccagggac gcttcttgaa 120
 ggcaccccc actccaagct gcaaattagg accgagagtc agtggccgct caagagtctg 180
 tgaccatgcc ccaaattcag agatgggtccc aggagagatg gggggaactg ccaagcaatg 240
 agtgaccggt tccccctccc ccag 264
 <210> 13
 25 <211> 275
 <212> ADN
 <213> secuencia artificial
 <220>
 30 <223> intrón 28 modificado
 <400> 13
 gtgggcaagc ttagcatgtg gaacctggcg agccccatcc cggcaagct ctcaagccat 60
 gctggtgggg acgactgaat gccagggccc ttcactgggc tatttcaccc agggacgctt 120
 cttgaaggca cccccactc caagctcaaa ttaggaccga gagtcagtgg ccgctcaaga 180
 35 gtctgtgacc atgccccaaa ttcagagatg gtcccaggag agatgggggg aactgccaaag 240
 caatgagtga ccggttcccc ctcccacgcg tccag 275
 <210> 14
 <211> 146
 40 <212> ADN
 <213> secuencia artificial
 <220>
 <223> intron 28 modificado-vector 1
 45 <400> 14

ES 2 395 881 T3

```

gtgggcaagc ttagcatgtg gaacctggcg agcccatcc cggcaagct ctcaagccat      60
gctggtgggg acgactgaat gccagggccc ttaactgggc tatttcaccc agggacgctt    120
cttgaaggca cccccactc caagct                                           146
<210> 15
<211> 129
5 <212> ADN
   <213> secuencia artificial
<220>
<223> intrón 28 modificado-vector 2
10 <400> 15
   caaattagga cggagagtca gtggccgctc aagagtctgt gaccatgccc caaatcaga      60
   gatggtccca ggagagatgg ggggaactgc caagcaatga gtgaccgggt cccctccca    120
   cgcgtccag                                                                129
15 <210> 16
   <211> 2080
   <212> PRT
   <213> disferlina humana
20 <400> 16
   Met  Leu  Arg  Val  Phe  Ile  Leu  Tyr  Ala  Glu  Asn  Val  His  Thr  Pro  Asp
   1          5          10          15
   Thr  Asp  Ile  Ser  Asp  Ala  Tyr  Cys  Ser  Ala  Val  Phe  Ala  Gly  Val  Lys
           20          25          30
   Lys  Arg  Thr  Lys  Val  Ile  Lys  Asn  Ser  Val  Asn  Pro  Val  Trp  Asn  Glu
           35          40          45
   Gly  Phe  Glu  Trp  Asp  Leu  Lys  Gly  Ile  Pro  Leu  Asp  Gln  Gly  Ser  Glu
           50          55          60
   Leu  His  Val  Val  Val  Lys  Asp  His  Glu  Thr  Met  Gly  Arg  Asn  Arg  Phe

```


ES 2 395 881 T3

Glu Thr Lys Leu Val Glu His Ser Glu Gln Lys Val Glu Asp Leu Pro
 565 570 575

Ala Asp Asp Ile Leu Arg Val Glu Lys Tyr Leu Arg Arg Arg Lys Tyr
 580 585 590

Ser Leu Phe Ala Ala Phe Tyr Ser Ala Thr Met Leu Gln Asp Val Asp
 595 600 605

Asp Ala Ile Gln Phe Glu Val Ser Ile Gly Asn Tyr Gly Asn Lys Phe
 610 615 620

Asp Met Thr Cys Leu Pro Leu Ala Ser Thr Thr Gln Tyr Ser Arg Ala
 625 630 635 640

Val Phe Asp Gly Cys His Tyr Tyr Tyr Leu Pro Trp Gly Asn Val Lys
 645 650 655

Pro Val Val Val Leu Ser Ser Tyr Trp Glu Asp Ile Ser His Arg Ile
 660 665 670

Glu Thr Gln Asn Gln Leu Leu Gly Ile Ala Asp Arg Leu Glu Ala Gly
 675 680 685

Leu Glu Gln Val His Leu Ala Leu Lys Ala Gln Cys Ser Thr Glu Asp
 690 695 700

Val Asp Ser Leu Val Ala Gln Leu Thr Asp Glu Leu Ile Ala Gly Cys
 705 710 715 720

Ser Gln Pro Leu Gly Asp Ile His Glu Thr Pro Ser Ala Thr His Leu
 725 730 735

Asp Gln Tyr Leu Tyr Gln Leu Arg Thr His His Leu Ser Gln Ile Thr
 740 745 750

Glu Ala Ala Leu Ala Leu Lys Leu Gly His Ser Glu Leu Pro Ala Ala
 755 760 765

Leu Glu Gln Ala Glu Asp Trp Leu Leu Arg Leu Arg Ala Leu Ala Glu
 770 775 780

Glu Pro Gln Asn Ser Leu Pro Asp Ile Val Ile Trp Met Leu Gln Gly
 785 790 795 800

ES 2 395 881 T3

Asp Lys Arg Val Ala Tyr Gln Arg Val Pro Ala His Gln Val Leu Phe
 805 810 815
 Ser Arg Arg Gly Ala Asn Tyr Cys Gly Lys Asn Cys Gly Lys Leu Gln
 820 825 830
 Thr Ile Phe Leu Lys Tyr Pro Met Glu Lys Val Pro Gly Ala Arg Met
 835 840 845
 Pro Val Gln Ile Arg Val Lys Leu Trp Phe Gly Leu Ser Val Asp Glu
 850 855 860
 Lys Glu Phe Asn Gln Phe Ala Glu Gly Lys Leu Ser Val Phe Ala Glu
 865 870 875 880
 Thr Tyr Glu Asn Glu Thr Lys Leu Ala Leu Val Gly Asn Trp Gly Thr
 885 890 895
 Thr Gly Leu Thr Tyr Pro Lys Phe Ser Asp Val Thr Gly Lys Ile Lys
 900 905 910
 Leu Pro Lys Asp Ser Phe Arg Pro Ser Ala Gly Trp Thr Trp Ala Gly
 915 920 925
 Asp Trp Phe Val Cys Pro Glu Lys Thr Leu Leu His Asp Met Asp Ala
 930 935 940
 Gly His Leu Ser Phe Val Glu Glu Val Phe Glu Asn Gln Thr Arg Leu
 945 950 955 960
 Pro Gly Gly Gln Trp Ile Tyr Met Ser Asp Asn Tyr Thr Asp Val Asn
 965 970 975
 Gly Glu Lys Val Leu Pro Lys Asp Asp Ile Glu Cys Pro Leu Gly Trp
 980 985 990
 Lys Trp Glu Asp Glu Glu Trp Ser Thr Asp Leu Asn Arg Ala Val Asp
 995 1000 1005
 Glu Gln Gly Trp Glu Tyr Ser Ile Thr Ile Pro Pro Glu Arg Lys
 1010 1015 1020
 Pro Lys His Trp Val Pro Ala Glu Lys Met Tyr Tyr Thr His Arg
 1025 1030 1035
 Arg Arg Arg Trp Val Arg Leu Arg Arg Arg Asp Leu Ser Gln Met

ES 2 395 881 T3

1040		1045		1050
Glu Ala Leu Lys Arg His Arg Gln Ala Glu Ala Glu Gly Glu Gly				
1055		1060		1065
Trp Glu Tyr Ala Ser Leu Phe Gly Trp Lys Phe His Leu Glu Tyr				
1070		1075		1080
Arg Lys Thr Asp Ala Phe Arg Arg Arg Arg Trp Arg Arg Arg Met				
1085		1090		1095
Glu Pro Leu Glu Lys Thr Gly Pro Ala Ala Val Phe Ala Leu Glu				
1100		1105		1110
Gly Ala Leu Gly Gly Val Met Asp Asp Lys Ser Glu Asp Ser Met				
1115		1120		1125
Ser Val Ser Thr Leu Ser Phe Gly Val Asn Arg Pro Thr Ile Ser				
1130		1135		1140
Cys Ile Phe Asp Tyr Gly Asn Arg Tyr His Leu Arg Cys Tyr Met				
1145		1150		1155
Tyr Gln Ala Arg Asp Leu Ala Ala Met Asp Lys Asp Ser Phe Ser				
1160		1165		1170
Asp Pro Tyr Ala Ile Val Ser Phe Leu His Gln Ser Gln Lys Thr				
1175		1180		1185
Val Val Val Lys Asn Thr Leu Asn Pro Thr Trp Asp Gln Thr Leu				
1190		1195		1200
Ile Phe Tyr Glu Ile Glu Ile Phe Gly Glu Pro Ala Thr Val Ala				
1205		1210		1215
Glu Gln Pro Pro Ser Ile Val Val Glu Leu Tyr Asp His Asp Thr				
1220		1225		1230
Tyr Gly Ala Asp Glu Phe Met Gly Arg Cys Ile Cys Gln Pro Ser				
1235		1240		1245
Leu Glu Arg Met Pro Arg Leu Ala Trp Phe Pro Leu Thr Arg Gly				
1250		1255		1260
Ser Gln Pro Ser Gly Glu Leu Leu Ala Ser Phe Glu Leu Ile Gln				
1265		1270		1275

ES 2 395 881 T3

Arg Glu Lys Pro Ala Ile His His Ile Pro Gly Phe Glu Val Gln
 1280 1285 1290

Glu Thr Ser Arg Ile Leu Asp Glu Ser Glu Asp Thr Asp Leu Pro
 1295 1300 1305

Tyr Pro Pro Pro Gln Arg Glu Ala Asn Ile Tyr Met Val Pro Gln
 1310 1315 1320

Asn Ile Lys Pro Ala Leu Gln Arg Thr Ala Ile Glu Ile Leu Ala
 1325 1330 1335

Trp Gly Leu Arg Asn Met Lys Ser Tyr Gln Leu Ala Asn Ile Ser
 1340 1345 1350

Ser Pro Ser Leu Val Val Glu Cys Gly Gly Gln Thr Val Gln Ser
 1355 1360 1365

Cys Val Ile Arg Asn Leu Arg Lys Asn Pro Asn Phe Asp Ile Cys
 1370 1375 1380

Thr Leu Phe Met Glu Val Met Leu Pro Arg Glu Glu Leu Tyr Cys
 1385 1390 1395

Pro Pro Ile Thr Val Lys Val Ile Asp Asn Arg Gln Phe Gly Arg
 1400 1405 1410

Arg Pro Val Val Gly Gln Cys Thr Ile Arg Ser Leu Glu Ser Phe
 1415 1420 1425

Leu Cys Asp Pro Tyr Ser Ala Glu Ser Pro Ser Pro Gln Gly Gly
 1430 1435 1440

Pro Asp Asp Val Ser Leu Leu Ser Pro Gly Glu Asp Val Leu Ile
 1445 1450 1455

Asp Ile Asp Asp Lys Glu Pro Leu Ile Pro Ile Gln Glu Glu Glu
 1460 1465 1470

Phe Ile Asp Trp Trp Ser Lys Phe Phe Ala Ser Ile Gly Glu Arg
 1475 1480 1485

Glu Lys Cys Gly Ser Tyr Leu Glu Lys Asp Phe Asp Thr Leu Lys
 1490 1495 1500

ES 2 395 881 T3

Val Tyr Asp Thr Gln Leu Glu Asn Val Glu Ala Phe Glu Gly Leu
 1505 1510 1515

Ser Asp Phe Cys Asn Thr Phe Lys Leu Tyr Arg Gly Lys Thr Gln
 1520 1525 1530

Glu Glu Thr Glu Asp Pro Ser Val Ile Gly Glu Phe Lys Gly Leu
 1535 1540 1545

Phe Lys Ile Tyr Pro Leu Pro Glu Asp Pro Ala Ile Pro Met Pro
 1550 1555 1560

Pro Arg Gln Phe His Gln Leu Ala Ala Gln Gly Pro Gln Glu Cys
 1565 1570 1575

Leu Val Arg Ile Tyr Ile Val Arg Ala Phe Gly Leu Gln Pro Lys
 1580 1585 1590

Asp Pro Asn Gly Lys Cys Asp Pro Tyr Ile Lys Ile Ser Ile Gly
 1595 1600 1605

Lys Lys Ser Val Ser Asp Gln Asp Asn Tyr Ile Pro Cys Thr Leu
 1610 1615 1620

Glu Pro Val Phe Gly Lys Met Phe Glu Leu Thr Cys Thr Leu Pro
 1625 1630 1635

Leu Glu Lys Asp Leu Lys Ile Thr Leu Tyr Asp Tyr Asp Leu Leu
 1640 1645 1650

Ser Lys Asp Glu Lys Ile Gly Glu Thr Val Val Asp Leu Glu Asn
 1655 1660 1665

Arg Leu Leu Ser Lys Phe Gly Ala Arg Cys Gly Leu Pro Gln Thr
 1670 1675 1680

Tyr Cys Val Ser Gly Pro Asn Gln Trp Arg Asp Gln Leu Arg Pro
 1685 1690 1695

Ser Gln Leu Leu His Leu Phe Cys Gln Gln His Arg Val Lys Ala
 1700 1705 1710

Pro Val Tyr Arg Thr Asp Arg Val Met Phe Gln Asp Lys Glu Tyr
 1715 1720 1725

ES 2 395 881 T3

Ser Ile Glu Glu Ile Glu Ala Gly Arg Ile Pro Asn Pro His Leu
 1730 1735 1740

Gly Pro Val Glu Glu Arg Leu Ala Leu His Val Leu Gln Gln Gln
 1745 1750 1755

Gly Leu Val Pro Glu His Val Glu Ser Arg Pro Leu Tyr Ser Pro
 1760 1765 1770

Leu Gln Pro Asp Ile Glu Gln Gly Lys Leu Gln Met Trp Val Asp
 1775 1780 1785

Leu Phe Pro Lys Ala Leu Gly Arg Pro Gly Pro Pro Phe Asn Ile
 1790 1795 1800

Thr Pro Arg Arg Ala Arg Arg Phe Phe Leu Arg Cys Ile Ile Trp
 1805 1810 1815

Asn Thr Arg Asp Val Ile Leu Asp Asp Leu Ser Leu Thr Gly Glu
 1820 1825 1830

Lys Met Ser Asp Ile Tyr Val Lys Gly Trp Met Ile Gly Phe Glu
 1835 1840 1845

Glu His Lys Gln Lys Thr Asp Val His Tyr Arg Ser Leu Gly Gly
 1850 1855 1860

Glu Gly Asn Phe Asn Trp Arg Phe Ile Phe Pro Phe Asp Tyr Leu
 1865 1870 1875

Pro Ala Glu Gln Val Cys Thr Ile Ala Lys Lys Asp Ala Phe Trp
 1880 1885 1890

Arg Leu Asp Lys Thr Glu Ser Lys Ile Pro Ala Arg Val Val Phe
 1895 1900 1905

Gln Ile Trp Asp Asn Asp Lys Phe Ser Phe Asp Asp Phe Leu Gly
 1910 1915 1920

Ser Leu Gln Leu Asp Leu Asn Arg Met Pro Lys Pro Ala Lys Thr
 1925 1930 1935

Ala Lys Lys Cys Ser Leu Asp Gln Leu Asp Asp Ala Phe His Pro
 1940 1945 1950

Glu Trp Phe Val Ser Leu Phe Glu Gln Lys Thr Val Lys Gly Trp

REIVINDICACIONES

1.Método *in vitro* para predecir la respuesta de un paciente con NSCLC al tratamiento con erlotinib, que comprende:

5 determinar un nivel de expresión de un gen PTPRF en una muestra tumoral de un paciente y comparar el nivel de expresión del gen PTPRF con un valor representativo de un nivel de expresión del gen PTPRF en tumores de una población de pacientes que no deriva ningún beneficio clínico del tratamiento, en el que un nivel de expresión más alto del gen PTPRF en la muestra tumoral del paciente es indicativo de un paciente que obtendrá un beneficio clínico del tratamiento, en el que se ha definido el beneficio clínico como manifestar una respuesta objetiva o la estabilización de la enfermedad durante ≥ 12 semanas.
10

2.Método según la reivindicación 1, en el que se determina el nivel de expresión mediante tecnología de micromatrices.

15 3.Método según la reivindicación 1 ó 2, en el que el gen PTPRF muestra un nivel de expresión entre 1,1 y 1,8, preferentemente entre 1,1 y 1,6 ó más veces superior en la muestra tumoral del paciente que el valor representativo de un nivel de expresión del gen PTPRF en tumores de una población de pacientes que no obtiene ningún beneficio clínico del tratamiento, en el que se define el beneficio clínico como presentar una respuesta objetiva o la estabilización de la enfermedad durante ≥ 12 semanas.
20

4.Método según las reivindicaciones 1 a 3, en el que el gen PTPRF muestra un nivel de expresión entre 1,2 y 1,8 ó más veces superior en la muestra tumoral del paciente que el valor representativo de la población de pacientes que no obtiene ningún beneficio clínico del tratamiento, en el que se define el beneficio clínico como presentar una respuesta objetiva o la estabilización de la enfermedad durante ≥ 12 semanas.
25

Fig. 1

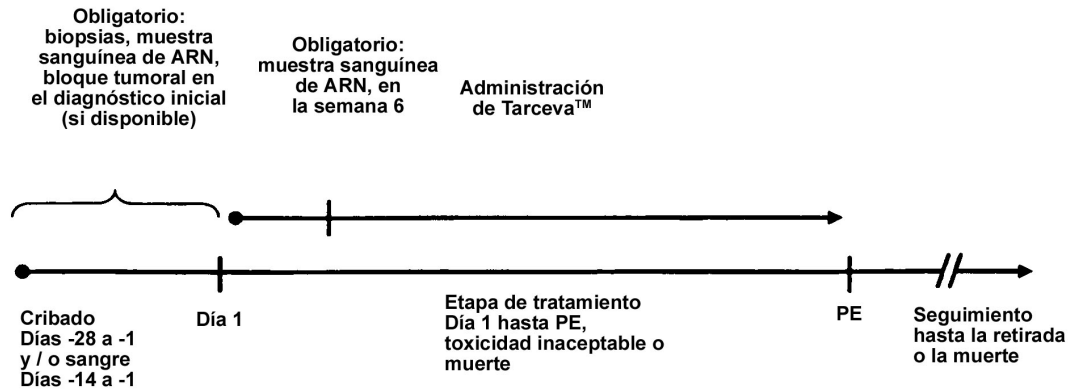


Fig. 2

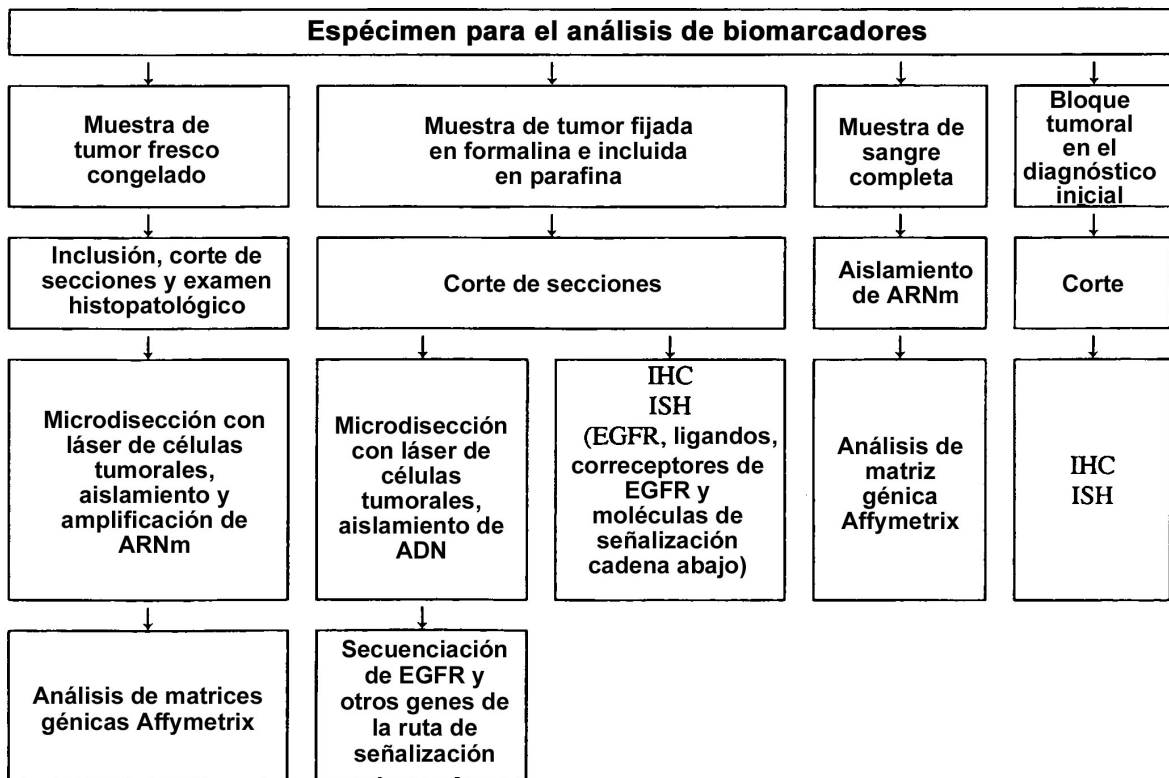


Fig. 3a

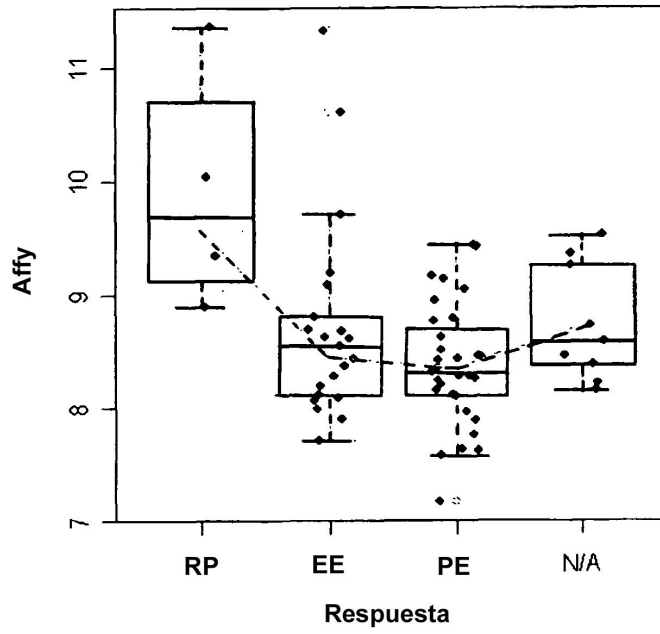


Fig. 3b

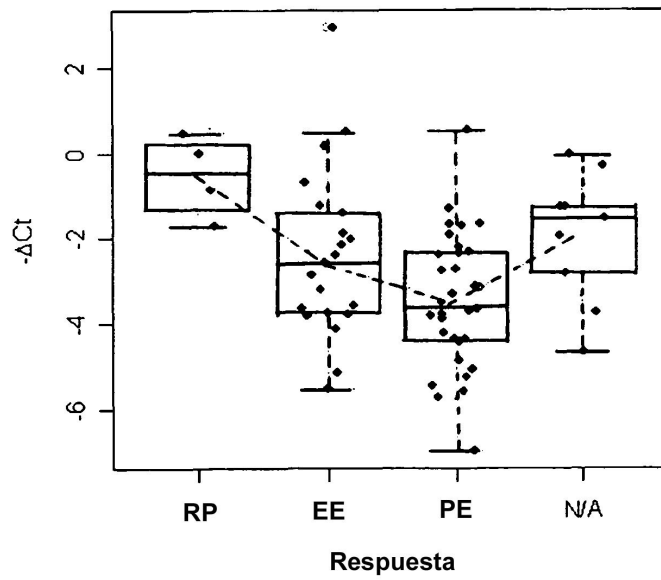


Fig. 3c

