

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 437 122**

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 08785425 (3)**

97 Fecha y número de publicación de la concesión europea: **18.09.2013** **EP 2188393**

54 Título: **Marcador predictivo en el tratamiento inhibidor del EGFR**

30 Prioridad:

14.08.2007 EP 07114314

20.05.2008 EP 08156516

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
09.01.2014

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 2 437 122 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 en el tratamiento inhibidor del EGFR

- 5 La presente invención proporciona un biomarcador predictivo de la respuesta al tratamiento con un inhibidor del EGFR en pacientes con cáncer

10 Muchos tumores malignos humanos están asociados a la expresión aberrante o a la sobreexpresión del receptor del factor de crecimiento epidérmico (EGFR). El EGF, el factor de crecimiento transformante α (TGF- α), y varios otros ligandos que se unen al EGFR, estimulan la autofosforilación del dominio tirosina quinasa intracelular del receptor. Subsiguientemente, se activan varias vías intracelulares y estos eventos secundarios resultan en la proliferación celular tumoral in vitro. Se ha postulado que la estimulación de las células tumorales mediante el EGFR puede ser importante tanto para el crecimiento tumoral como para la supervivencia del tumor in vivo.

15 Los datos clínicos precoces con Tarceva™, un inhibidor de la tirosina quinasa del EGFR, indican que el compuesto es seguro y generalmente bien tolerado a dosis que proporcionan la concentración efectiva deseada (tal y como se determina mediante los datos preclínicos). Los ensayos clínicos en fases I y II en pacientes con enfermedad avanzada han demostrado que Tarceva™ presenta una actividad clínica prometedora en varios tumores epiteliales. Además, se ha demostrado que Tarceva™ es capaz de inducir remisiones parciales duraderas en pacientes
20 previamente tratados de cáncer de cabeza y cuello, y NSCLC (cáncer pulmonar de células no pequeñas) de una magnitud similar a la establecida por la quimioterapia de segunda línea, pero con el beneficio añadido de un perfil de seguridad mejor que la quimioterapia y una mejor conveniencia (tabletas frente a la administración intravenosa [i.v.]). Un ensayo controlado con placebo, a doble ciego, aleatorizado y completado recientemente (BR.21) ha demostrado que el agente Tarceva™ en sí mejora y prolonga significativamente la supervivencia de los pacientes con NSCLC en
25 los que la terapia estándar para la enfermedad avanzada ha fallado.

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

30 El cáncer pulmonar es la principal causa de muerte relacionada con el cáncer en América del Norte y Europa. En los Estados Unidos, el número de muertes secundarias al cáncer pulmonar excede las muertes totales combinadas de la segunda (colon), tercera (mama), y cuarta (próstata) causas de muerte por cáncer. Aproximadamente entre el 75% y el 80% de todos los cánceres de pulmón son cánceres pulmonares de células no pequeñas (NSCLC), y aproximadamente el 40% de los pacientes se diagnostican con enfermedad localmente avanzada y/o irreseccable.
35 Habitualmente, este grupo incluye las enfermedades en estadios voluminosos IIIA y IIIB, y se excluye la extensión pleural.

La incidencia cruda de cáncer pulmonar en la Unión Europea es de 52,5 y la tasa de muerte de 48,7
40 casos/100.000/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 acumula el 80% de todos los casos de cáncer pulmonar. Aproximadamente, el 90% de la mortalidad por cáncer pulmonar entre los hombres, y el 80% entre las mujeres, se puede atribuir al tabaco.

Durante el 2004 en los EE.UU., de acuerdo con la American Cancer Society, hubo aproximadamente 173.800
45 nuevos casos de cáncer pulmonar (93.100 en hombres y 80.700 en mujeres) y contabilizaron aproximadamente el 13% de todos los nuevos casos de cáncer. La mayoría de los pacientes mueren a consecuencia de su enfermedad a los dos años del diagnóstico. Para la mayoría de los pacientes con NSCLC es difícil de lograr un tratamiento eficaz. Con frecuencia, los tumores avanzados no son tributarios de cirugía y también pueden ser resistentes a dosis tolerables de radioterapia y quimioterapia. En ensayos aleatorizados, la combinación de quimioterapias más activas logró unas tasas de respuesta de aproximadamente entre el 30% y el 40% y una tasa de supervivencia al primer año
50 de entre el 35% y el 40%. Realmente, esto supone un avance frente a la tasa de supervivencia al año del 10% que se obtiene únicamente con el tratamiento de soporte (Shepherd 1999).

Hasta recientemente, las opciones terapéuticas para los pacientes que padecían una recaída se limitaban al
55 tratamiento de soporte o al tratamiento paliativo. Un ensayo reciente que compara docetaxel (Taxotere®) con el mejor tratamiento de soporte demostró que los pacientes con NSCLC se podían beneficiar del tratamiento de segunda línea tras el fallo de los regímenes de primera de línea basados en cisplatino. Se demostró la mejora de la supervivencia en pacientes de todas las edades con puntuaciones en la escala ECOG de 0, 1 o 2 que siguieron tratamiento con docetaxel, así como lo hicieron aquellos pacientes refractarios al tratamiento previo basado en platino. Los pacientes que no se beneficiaron de la terapia incluyeron a aquellos con una pérdida de peso corporal
60 de > 10%, niveles elevados de lactato deshidrogenasa, afectación multiorgánica o afectación hepática. Además, el beneficio de la monoterapia con docetaxel monoterapia no se extendió más allá de la segunda línea de tratamiento. Los pacientes que recibieron docetaxel como tratamiento de tercera línea o posterior no mostraron un aumento de supervivencia. El agente docetaxel por sí solo se convirtió en un terapia de segunda línea estándar para el NSCLC. Recientemente, otro ensayo aleatorizado en fase III sobre la terapia de segunda línea del NSCLC comparó pemetrexed (Alimta®) con docetaxel. El tratamiento con pemetrexed dio como resultado una eficacia clínica
65 equivalente pero unos efectos secundarios significativamente menores en comparación al docetaxel.

Desde hace tiempo se sabe que existe la necesidad de desarrollar métodos para individualizar el tratamiento del cáncer. Con el desarrollo de tratamientos contra el cáncer dirigidos, surge un interés particular en las metodologías que puedan proporcionar un perfil molecular de la diana tumoral (es decir, las que predicen un beneficio clínico). Se ha establecido una constatación preliminar del perfilado de expresión génica en el cáncer con la clasificación molecular de tipos tumorales que no son aparentes en base a los ensayos morfológicos e inmunohistoquímicos actuales.

La patente WO2006/101925 (28/9/2006) describe métodos para determinar qué pacientes se beneficiarán más del tratamiento con un inhibidor de la quinasa del receptor del factor de crecimiento epidérmico (EGFR).

Por consiguiente, un objetivo de la presente invención es la proporción de biomarcadores de expresión que predicen la respuesta al tratamiento con un inhibidor del EGFR en pacientes con cáncer.

En un primer objetivo, la presente invención proporciona un método in vitro para predecir la respuesta de un paciente con NSCLC al tratamiento con erlotinib, e incluye los pasos de: determinar un nivel de expresión de mRNA de un gen RAPGEF5 en una muestra tumoral de un paciente y comparar el nivel de expresión del gen RAPGEF5 con un valor representativo del nivel de expresión del gen RAPGEF5 en tumores de una población de pacientes no respondedores, en el que un nivel de expresión mayor de, al menos, un gen en la muestra tumoral del paciente indica que un paciente responderá al tratamiento.

La abreviación RAPGEF5 significa factor intercambiador de nucleótidos de guanina Rap (GEF) 5. La secuencia de nucleótidos del gen humano RAPGEF5 y de sus variantes de transcripción 1, 6 y 11 se muestran en las SEC. ID. N° 1-4. Los números de acceso al banco de genes del gen RAPGEF5 y de sus variantes y los números de identificación de secuencia correspondientes (SEC. ID. N°) se enumeran en la tabla 4.

El término "un valor representativo de un nivel de expresión del gen RAPGEF5 en tumores de una población de pacientes no respondedores" hace referencia a la estimación de un nivel medio de expresión del gen marcador en tumores de una población de pacientes que no responde al tratamiento con el inhibidor del EGFR.

En una realización preferible, el gen marcador RAPGEF5 muestra una expresión habitualmente entre 1,5 y 2,7 o más veces mayor en la muestra tumoral del paciente respondedor en comparación con el valor representativo de un nivel de expresión del gen RAPGEF5 en tumores de una población de pacientes no respondedores.

En una realización preferible, el nivel de expresión del gen RAPGEF5 se determina mediante tecnología de microarrays u otras tecnologías que evalúan los niveles de expresión de RNA, como las RT-PCR cuantitativas o, pese a que no forman parte de la invención, mediante cualquier método que mida el nivel de expresión de la proteína respectiva, por ejemplo la inmunohistoquímica (IHC). En la materia se conoce la construcción y utilización de los chips génicos. Véanse las patentes de EE.UU N° 5.202.231; 5.445.934; 5.525.464; 5.695.940; 5.744.305; 5.795.716 y 1 5.800.992. Véanse 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 puede determinarse mediante otros métodos que los expertos en la materia conocen, tal como por ejemplo los Northern blot, la RT-PCR, la PCR cuantitativa a tiempo real, la extensión de cebadores, la protección de RNasa o el perfilado de expresión de RNA.

El gen de la presente invención puede combinarse en conjuntos de biomarcadores con otro biomarcadores predictivos. Los conjuntos de biomarcador pueden formarse mediante cualquier combinación de biomarcadores predictivos para predecir aproximadamente el efecto del tratamiento con un inhibidor del EGFR en pacientes con cáncer. Los diferentes biomarcadores y conjuntos de biomarcadores que aquí se describen pueden utilizarse, por ejemplo, para predecir cómo responderán los pacientes con cáncer a la intervención terapéutica con un inhibidor del EGFR.

El término "gen", tal y como se utiliza aquí, incluye variantes del gen. El término "variante" hace referencia a secuencias de ácidos nucleicos sustancialmente similares a las secuencias de ácidos nucleicos proporcionadas por el número de acceso al GenBank. Los expertos en la materia comprenden el término "sustancialmente similar". En particular, una variante de un gen puede ser un alelo que muestra intercambios de nucleótido comparables a la secuencia de ácido nucleico del alelo más prevalente en la población humana. Preferiblemente, tal secuencia de ácido nucleico sustancialmente similar tiene una semejanza de secuencia con el alelo más prevalente de al menos un 80%, preferiblemente al menos un 85%, más preferiblemente al menos un 90%, los más preferible al menos un 95%. El término "variantes" también se relaciona con las variantes de empalme.

En algunos aspectos, el inhibidor del EGFR puede seleccionarse a partir del grupo que incluye gefitinib, erlotinib, PKI-166, EKB-569, GW2016, CI-1033 y un anticuerpo anti-erbB tal como trastuzumab y cetuximab.

En realizaciones de la invención, el inhibidor del EGFR es erlotinib y el cáncer es el NSCLC.

Las técnicas para la detección de la expresión génica de los genes descritos en esta invención incluyen, pero no se limitan a, Northern blot, RT-PCR, PCR cuantitativa a tiempo real, extensión de cebadores, protección de RNasa, perfilado de la expresión de RNA y técnicas relacionadas. Los expertos en la materia conocen bien estas técnicas. Véase, por ejemplo, Sambrook J et al., Molecular Cloning: A Laboratory Manual, tercera edición (Cold Spring Harbor Press, Cold Spring Harbor, 2000).

Las técnicas para la detección de la expresión de proteínas de los genes respectivos descritos mediante esta descripción incluyen, pero no se limitan a, la inmunohistoquímica (IHC).

De acuerdo con la invención, se pueden evaluar las células de una muestra de tejido de un paciente, por ejemplo, una biopsia tumoral o de cáncer, para determinar el patrón de expresión de uno o más biomarcadores. Se puede determinar el éxito o el fracaso de un tratamiento contra el cáncer en base al patrón de expresión del biomarcador en las células del tejido del ensayo (células del ensayo), por ejemplo, una biopsia tumoral o de cáncer, y se compara si es relativamente similar o diferente al patrón de expresión de un conjunto control de uno o más biomarcadores. En el contexto de esta invención, se observó que el gen enumerado en la tabla 3 está regulado positivamente, es decir, muestra un nivel de expresión mayor en los tumores de pacientes que respondieron al tratamiento con un inhibidor del EGFR en comparación a los tumores de pacientes que no respondieron al tratamiento con un inhibidor del EGFR. Por consiguiente, si las células del ensayo muestran un perfil de expresión del biomarcador que se corresponde con el de un paciente que respondió al tratamiento contra el cáncer, es muy probable o predecible que el cáncer o tumor del individuo responda favorablemente al tratamiento con el inhibidor del EGFR. Por el contrario, si las células del ensayo muestran el patrón de expresión del biomarcador correspondiente al de un paciente que no respondió al tratamiento del cáncer, es muy probable o predecible que el cáncer o tumor del individuo no responda al tratamiento con el inhibidor del EGFR.

El gen marcador de la presente invención puede combinarse con otros biomarcadores para formar conjuntos de biomarcadores. Los conjuntos de biomarcadores pueden formarse mediante cualquier combinación de biomarcadores predictivos para predecir aproximadamente el efecto del tratamiento con un inhibidor del EGFR en pacientes con cáncer. Los biomarcadores y conjuntos de biomarcadores aquí descritos pueden utilizarse, por ejemplo, para predecir cómo responderán los pacientes con cáncer a la intervención terapéutica con un inhibidor del EGFR.

En un objetivo adicional, la presente invención proporciona erlotinib para su utilización en un paciente con NSCLC identificado mediante el método in vitro de la presente invención. Dicho método terapéutico incluye la administración de erlotinib al paciente seleccionado para el tratamiento en base al patrón de expresión predictivo del gen RAPGEF5 enumerado en la tabla 3. El inhibidor del EGFR es erlotinib y el cáncer a tratar es el NSCLC.

Breve descripción de las figuras

La figura 1 muestra el diseño del estudio;

La figura 2 muestra el esquema del procesamiento de la muestra;

La figura 3a muestra los niveles de expresión de RAPGEF5 frente al resultado clínico del perfilado con GeneChip®;

La figura 3b muestra los niveles de expresión de RAPGEF5 frente al resultado clínico con qRT-PCR y

La figura 3c muestra la correlación entre las mediciones con Genechip® y qRT-PCR para RAPGEF5.

Parte experimental

Justificación del estudio y el diseño del estudio

Recientemente, se han descrito mutaciones del gen EGFR en el tejido tumoral de un subgrupo de pacientes con NSCLC, así como la asociación de estas mutaciones con la sensibilidad a erlotinib y gefitinib (Pao W, et al. 2004; Lynch et al. 2004; Paez et al. 2004). En los pacientes combinados a partir de dos estudios, se observó el EGFR mutado en 13 de 14 pacientes que respondieron a gefitinib y no se observó en ninguno de los 11 pacientes tratados con gefitinib que no respondieron. La prevalencia descrita de estas mutaciones fue del 8% (2 de 25) en pacientes con NSCLC no seleccionados. Se observó que estas mutaciones son más frecuentes en los adenocarcinomas (21 %), en los tumores en mujeres (20%), y en los tumores de pacientes japoneses (26%). Estas mutaciones resultaron en una actividad in vitro aumentada del EGFR y una sensibilidad incrementada al gefitinib. La relación de las mutaciones con la prolongación de la enfermedad estable o de la duración de la supervivencia no se ha evaluado prospectivamente.

En base a los análisis exploratorios a partir del estudio BR.21, pareció improbable que el beneficio sobre la supervivencia observado se debiera solamente a las mutaciones del EGFR, dado que se mantiene un beneficio

significativo sobre la supervivencia incluso cuando los pacientes con una respuesta objetiva se excluyen de los análisis (datos en el documento). Otros mecanismos también deben contribuir al efecto.

En base a la asunción de que existen cambios en los niveles de expresión génica que pueden predecir la respuesta/beneficio con el tratamiento con Tarceva™, se utilizaron análisis con microarrays para detectar estos cambios.

Esto requirió una población de estudio claramente definida y tratada en monoterapia con Tarceva™ tras el fracaso de la terapia de primera línea. En base a la experiencia adquirida con el estudio BR.21, se definió a la población que puede beneficiarse por la presentación de una respuesta objetiva o la estabilización de la enfermedad durante ≥12 semanas. Los conjuntos de datos clínicos y de los microarrays se analizaron de acuerdo con un plan estadístico predefinido.

La aplicación de esta técnica requiere tejido congelado fresco (FFT). Por consiguiente, se debe llevar a cabo una biopsia obligatoria antes de empezar el tratamiento. El material recogido se congeló en nitrógeno líquido (N₂).

Se recogió una segunda muestra tumoral al mismo tiempo y se almacenó en parafina (embebido en parafina y fijado con formalina, FFPE). Esta muestra se analizó en busca de alteraciones en la vía de señalización del EGFR.

La capacidad de llevar a cabo las biopsias mediante la broncoscopia fue un prerrequisito para este estudio. La broncoscopia es un procedimiento estándar que se lleva a cabo para confirmar el diagnóstico del cáncer pulmonar. Pese a que generalmente es segura, existe un riesgo de complicaciones, por ejemplo de sangrado.

Este estudio fue un primer paso hacia la terapia individualizada para los pacientes con NSCLC refractario. Esta terapia individualizada permitirá que los médicos encargados del tratamiento seleccionen el agente más apropiado de entre los fármacos existentes para esta indicación.

Una vez esté disponible la terapia individualizada, el beneficio de cada futuro paciente compensará el riesgo que los pacientes tienen que asumir en el presente estudio:

La relación tasa de respuesta/número de pacientes beneficiados aumentará y el riesgo de efectos secundarios debidos a un tratamiento inefectivo disminuirá.

Justificación de la selección de la dosis

Se administró Tarceva™ por vía oral una vez al día a una dosis de 150 mg hasta que la enfermedad progresó, se presentaron toxicidades intolerables o sobrevino la muerte. La selección de esta dosis se llevó a cabo en base a parámetros farmacocinéticos, así como en base al perfil de seguridad y tolerabilidad observado para esta dosis en los ensayos de fase I, II y III en pacientes con cáncer avanzado y tratados previamente de forma agresiva. Los niveles de fármaco observados en el plasma de los pacientes con cáncer que recibieron la dosis de 150 mg/día estuvieron de forma consistente por encima de la concentración plasmática media de 500 ng/ml que se busca para la eficacia clínica. El estudio BR.21 mostró un beneficio en la supervivencia con esta dosis.

Objetivos del estudio

El objetivo principal fue la identificación de genes expresados de forma diferencial que puedan predecir el beneficio (respuesta completa (CR), respuesta parcial (PR) o enfermedad estable (SD) ≥12 semanas) con el tratamiento con Tarceva™. La identificación de genes expresados de forma diferencial que puedan predecir la "respuesta" (CR, PR) al tratamiento con Tarceva™ fue un objetivo adicional importante.

Los objetivos secundarios se relacionaron con la evaluación de alteraciones en las vías de señalización del EGFR con relación al beneficio del tratamiento.

Diseño del estudio

Descripción general del diseño del estudio y del régimen de dosificación

Este fue un estudio abierto de fase II para la identificación de un marcador predictivo. El estudio se llevó a cabo en aproximadamente 26 lugares de aproximadamente 12 países. Se enrolaron 264 pacientes con NSCLC avanzado tras el fracaso de al menos un régimen de quimioterapia previo, durante un periodo de 12 meses. Se administró Tarceva™ oral de manera continua a una dosis de 150 mg/día. Las reducciones de la dosis se permitieron en base a la tolerabilidad a la terapia con el fármaco. Se evaluaron parámetros clínicos y de laboratorio para evaluar el control de la enfermedad y la toxicidad. Se continuó el tratamiento hasta que se presentó progresión, toxicidad inaceptable o la muerte. El diseño del estudio se ilustra en la figura 1.

Se obtuvieron muestras de sangre y de tejido tumoral para el análisis molecular que se realiza para evaluar los efectos de Tarceva™ y para identificar subgrupos de pacientes que se benefician de la terapia.

Evaluación del marcador predictivo

Se tomaron biopsias del tumor 2 semanas antes del comienzo del tratamiento. Se recogieron dos muestras diferentes:

- La primera muestra siempre se congeló inmediatamente en N₂ líquido.
- La segunda muestra se fijó con formalina y se embebió en parafina.
- El tejido congelado instantáneamente tuvo la prioridad más alta en este estudio.
- La figura 2 muestra un esquema del procesamiento de la muestra.

Análisis de microarrays

Las muestras congeladas instantáneamente se utilizaron para la microdissección por captura láser (LCM) de las células tumorales para extraer el RNA tumoral y el RNA del tejido que rodea el tumor. El RNA se analizó en chips de microarray Affymetrix (HG-U133A) para establecer el perfil de la expresión génica del tumor del paciente. Se utilizó el control de calidad de los chips Affymetrix para seleccionar las muestras con una calidad adecuada para la comparación estadística.

Análisis de biomarcador único en tejido fijado con formalina y embebido en parafina

La segunda biopsia tumoral, la muestra FFPE, se utilizó para realizar análisis IHC, ISH y de mutación de DNA tal y como se describe más adelante. Se realizaron análisis similares en el tejido recogido en el momento inicial del diagnóstico.

Se analizó el estado de mutación del DNA de los genes que codifican por el EGFR y otras moléculas involucradas en la vía de señalización del EGFR mediante la secuenciación del DNA. La amplificación del EGFR y de genes relacionados se estudió mediante FISH.

El análisis de expresión de proteínas incluyó los análisis inmunohistoquímicos [IHC] del EGFR y otras proteínas de la vía de señalización del EGFR.

Evaluación de la respuesta

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

Nótese que: para que se pueda asignar un estado de CR o PR se deben confirmar los cambios en las mediciones del tumor mediante evaluaciones repetidas, al menos con 4 semanas de separación, en cualquier momento del periodo de tratamiento.

En el caso de la SD, las mediciones de seguimiento deben haber cumplido los criterios de SD al menos una vez tras la entrada en el estudio en un intervalo mínimo de 6 semanas.

En el caso de una SD mantenida, las mediciones de seguimiento deben haber cumplido los criterios de SD al menos una vez tras la entrada en el estudio con una duración de mantenimiento de al menos 12 semanas.

Evaluación de la supervivencia

Se llevó a cabo la comprobación regular del estado cada 3 meses, ya fuera mediante una visita del paciente a la clínica o mediante una llamada telefónica. Se registraron todas las muertes. Al final del estudio se requirió la confirmación definitiva de la supervivencia por parte de cada paciente.

Respuesta al tratamiento con erlotinib

En el estudio se enrolaron un total de 264 pacientes procedentes de 12 países y 26 centros. El 26% se presentaron en estadio IIIB y el 24% en estadio IV de NSCLC. El 13,6% (n=36) de los pacientes lograron una respuesta objetiva mientras que el 31,4% (n=83) obtuvo un beneficio clínico (definido como una respuesta objetiva o una enfermedad estable durante 12 semanas o más). La supervivencia media global fue de 7,6 (IC de 7-9) meses y la supervivencia libre de progresión fue de 11,3 (IC de 8-12) semanas. Los detalles sobre los datos clínicos se muestran en la tabla 1.

Se recogieron biopsias broncoscópicas congeladas frescas de todos los sujetos pero, o bien no todas las muestras tenían un contenido tumoral suficiente antes de la microdissección (LCM) o no tenía suficiente RNA evaluable tras la LCM para poder ejecutar un análisis con microarrays, por lo que sólo se obtuvo material tumoral adecuado de 125

pacientes; 122 de estos presentaron un RNA evaluable. Otro conjunto de 20 muestras no pasó la evaluación del control de calidad de los datos de los microarrays. En la tabla 1 se muestran las características clínicas de los 102 conjuntos de datos de microarray que fueron adecuados para el análisis estadístico. Mientras que en el estudio global 36 pacientes lograron una respuesta objetiva, sólo 6 de estos tuvieron datos de microarray; de manera similar en los que lograron un beneficio clínico, el número de sujetos con datos de microarray fue sólo de 21 en comparación con los 83 del conjunto de datos completo. Se catalogó a 6 de estos como respondedores parciales (PR), 31 presentaron una SD y 49 una PD (enfermedad progresiva); de los 6 pacientes con una PR, 5 presentaron un adenocarcinoma y uno un carcinoma de célula escamosa. Ningún paciente logró una CR en el conjunto de datos.

Métodos

Preparación de muestras de RNA y control de calidad de las muestras de RNA

El procesamiento de todas las muestras de biopsia se llevó a cabo por un laboratorio de referencia en patología; las muestras tisulares congeladas frescas se enviaron desde las sedes de los investigadores hasta las instalaciones de Clinical Sample Operations de Roche Basilea, y de ahí al laboratorio de patología para su procesamiento posterior. Se utilizó la microdissección por captura láser para seleccionar y separar las células tumorales del tejido circundante. Tras la LCM, se purificó el RNA del material tumoral enriquecido. Entonces, el laboratorio de patología llevó a cabo una serie de pasos para obtener una estimación de la concentración y calidad del RNA.

Las RNasas son enzimas que degradan RNA y que se encuentran de forma ubicua, por lo que todos los procedimientos en los que se debe utilizar RNA deben controlarse de forma estricta para minimizar la degradación del RNA. La mayoría de especies de mRNA presentan de por sí semividas relativamente cortas y, por consiguiente, se consideran bastante inestables. Así, es importante realizar comprobaciones de la integridad del RNA y cuantificarlo antes de cualquier ensayo.

La concentración del RNA y el perfil de calidad pueden evaluarse mediante la utilización de un instrumento de Agilent (Agilent Technologies, Inc., Palo Alto, CA) llamado 2100 Bioanalyzer®. El programa informático del instrumento genera un número de integridad del RNA (RIN), una cuantificación estimada (Schroeder, A., et al., The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol, 2006. 7: pág. 3), y calcula las proporciones ribosomales de la muestra de RNA total. El RIN se determina a partir del rastro electroforético completo de la muestra de RNA y, por tanto, incluye la presencia o ausencia de productos de degradación.

La calidad del RNA se analizó mediante el 2100 Bioanalyzer®. Sólo se seleccionaron muestras con suficiente RNA y con, al menos, un pico de rRNA superior al ruido añadido de poli-I para el análisis en la plataforma Affymetrix. El RNA purificado se envió al Roche Centre for Medical Genomics (RCMG; Basilea, Suiza) para su análisis mediante microarrays. Se recibieron 122 muestras de RNA desde el laboratorio de patología para su procesamiento posterior.

Marcaje de dianas de muestras de RNA tisular

El marcaje de dianas se llevó a cabo de acuerdo con el protocolo de amplificación de marcaje de dianas en dos ciclos de Affymetrix (Affymetrix, Santa Clara, California), tal y como se especifica en las instrucciones del fabricante.

El método se basa en el procedimiento de amplificación lineal estándar de Eberwine pero utiliza dos ciclos de este procedimiento para generar suficiente cRNA marcado para la hibridación en un microarray.

La entrada de RNA total utilizado en la reacción de marcado fue de 10 ng para las muestras que disponían de más de 10 ng de RNA; si se disponía de menos de tal cantidad o si no se disponía de datos relacionado con la cantidad (debido a una concentración muy baja de RNA), se utilizó la mitad de la muestra total en la reacción. El rendimiento de las reacciones de marcado osciló entre 20 y 180 µg de cRNA. Se introdujo un paso de normalización en el nivel de hibridación en el que se utilizaron 15 µg de cRNA para cada muestra.

Se utilizó RNA humano de referencia (Stratagene, Carlsbad, CA, EE.UU.) como muestra control en el flujo de trabajo con cada lote de muestras. Se utilizaron 10 ng de este RNA como entrada junto a las muestras del ensayo para verificar que el marcador y los reactivos de hibridación funcionaban tal y como se esperaba.

Hibridaciones de microarrays

Los microarrays HG-U133A de Affymetrix incluyen más de 22.000 conjuntos de sondas que se dirigen frente a aproximadamente 18.400 transcritos y variantes que representan aproximadamente 14.500 genes bien caracterizados.

Se llevó a cabo la hibridación de todas las muestras de acuerdo con las instrucciones de Affymetrix (Affymetrix Inc., Expression Analysis Technical Manual, 2004). En resumen, para cada muestra, se fragmentaron 15 µg de cRNA marcado con biotina en presencia de cationes divalentes y calor, y se hibridaron durante la noche con arrays de

oligonucleótidos del genoma completo HG-U133A de Affymetrix. Al día siguiente, los arrays se tñeron con estreptavidina-ficoeritrina (Molecular Probes; Eugene, OR) de acuerdo con las instrucciones del fabricante. Entonces se escanearon los arrays mediante un GeneChip Scanner 3000 (Affymetrix) y se calcularon automáticamente las intensidades de señal mediante el GeneChip Operating Software (GCOS) Versión 1.4 (Affymetrix).

5 Análisis estadístico

El análisis de los datos de Affymetrix™ incluyó cuatro pasos principales.

10 El paso 1 incluyó el control de calidad. La finalidad fue la identificación y exclusión del análisis de los datos con un perfil de calidad subóptimo.

15 El paso 2 incluyó la normalización y el procesamiento previo. La finalidad fue crear un “conjunto de datos de análisis” normalizado, escalado y susceptible a una comparación entre chips. Esto incluyó la estimación y sustracción del ruido basal, el escalado y la sumariación de las sondas.

20 El paso 3 incluyó la exploración y la descripción. La finalidad fue la identificación de sesgos y fuentes de variabilidad potenciales. Esto consistió en la aplicación de técnicas de análisis descriptivo multivariante y univariante para identificar las covariables influyentes.

25 El paso 4 incluyó el modelado y la evaluación. La finalidad fue la identificación de una lista de marcadores candidatos en base a la evaluación estadística de la diferencia en el nivel de expresión medio entre “respondedores” (pacientes con “respuesta parcial” o “respuesta completa” como mejor respuesta) y “no respondedores” (pacientes con “enfermedad progresiva” como mejor respuesta). Esto incluyó el ajuste de un modelo estadístico adecuado para cada conjunto de sondas y la derivación de una medida de significación estadística. Todos los análisis se llevaron a cabo utilizando el paquete de programas informáticos R.

Paso 1: control de calidad

30 La evaluación de la calidad de los datos se llevó a cabo en base a la comprobación de varios parámetros. Estos incluyeron los parámetros de calidad estándar de Affymetrix GeneChip™, en particular: factor de escalado, porcentaje de la convocatoria actual y ruido basal medio. Este paso también incluyó la inspección visual de las imágenes de chip virtuales para la detección de problemas de hibridación localizados, y la comparación de cada chip con un chip virtual medio para la detección de cualquier desviación inusual del comportamiento medio. El análisis de correlación entre chips también se llevó a cabo para detectar muestras atípicas. Además, se consideraron medidas auxiliares de la calidad del RNA obtenidas a partir del análisis de las muestras de RNA con el Bioanalyzer™ 2100 Agilent.

40 En base a estos parámetros, se excluyeron del análisis los datos de 20 arrays. Por consiguiente, en el análisis se incluyeron los datos de un total de 102 arrays que representaron a 102 pacientes. La descripción de este conjunto de 102 pacientes se incluye 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%)
	PD	49 (48,0%)
	SD	31 (30,4%)
	PR	6 (5,9%)
Beneficio clínico	No	81 (79,4%)
	Sí	21 (20,6%)
Sexo	Mujer	25 (24,5%)
	Hombre	77 (47,5%)
Etnia	Caucásica	65 (63,7%)
	Oriental	37 (36,6%)
Histología	Adenocarcinoma	35 (34,3%)
	Escamoso	53 (52,0%)
	Otros	14 (13,7%)
Fumador en algún momento	No	20 (19,6%)
	Sí	82 (80,4%)

Paso 2: procesamiento previo de los datos y normalización

El algoritmo rma (Irizarry, R.A., et al., Summaries of Affymetrix GeneChip probe level data. Nucl. Acids Res., 2003. 31(4): página e15) se utilizó para el procesamiento previo y la normalización. El algoritmo mas5 (AFFYMETRIX, GeneChip® Expression: Data Analysis Fundamentals. 2004, AFFYMETRIX) se utilizó para llevar a cabo decisiones de detección para cada conjunto individual de sondas. Los conjuntos de sondas etiquetados como "ausentes" o "marginales" en todas las muestras se eliminaron de los análisis posteriores; se eliminaron 5930 conjuntos de sondas del análisis en base a este criterio. Por consiguiente, el conjunto de datos del análisis incluye una matriz con 16.353 (de 22.283) conjuntos de sondas medidos en 102 pacientes.

Paso 3: descripción y exploración de los datos

El análisis exploratorio descriptivo se llevó a cabo para identificar sesgos potenciales y fuentes principales de variabilidad. Se cribó un conjunto de covariables con un impacto potencial en los perfiles de expresión génica. Esto incluyó tanto variables técnicas como clínicas. Las covariables técnicas incluyeron: la fecha del procesamiento del RNA (a lo que más adelante se hace referencia como lote), el RIN (como medida de relación calidad/integridad del RNA), y el operario y el centro de recogida de la muestra. Las covariables clínicas incluyeron: el tipo histológico, el consumo de tabaco, el grado tumoral, la puntuación en la escala de actividad (Oken, M.M., et al., Toxicity and Response Criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol, 1982. 5(6): páginas 649-55), los datos demográficos, el estado de respuesta y el estado de beneficio clínico.

Las herramientas de análisis incluyeron el ANOVA univariante y el análisis del componente principal. Para cada una de estas covariables, se aplicó un ANOVA univariante independientemente de cada conjunto de sondas.

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

El conjunto de datos normalizado tras la corrección del efecto lote sirvió como conjunto de datos del análisis en los análisis subsiguientes.

La histología y el RIN fueron dos variables adicionales importantes que el análisis descriptivo remarcó.

Paso 4: evaluación y modelado de datos

Se ajustó un modelo lineal de manera independiente a cada conjunto 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 verosimilitud. El parámetro correspondiente a la variable "respuesta" (X1) se utilizó para evaluar la diferencia en el nivel de expresión entre el grupo "respondedor" y los pacientes "no respondedores".

Tabla 2: descripción de las variables incluidas en el modelo lineal. El modelo lineal definido por estas variables, incluyendo un término de error de distribución normal, se ajustó a cada conjunto de sondas.

Variable	Tipo	Valores
Expresión génica	Dependiente (Y_{ip})	Intensidad log2 normalizada del conjunto de sondas i. en el paciente p.
Intercepción	Media global (μ)	
Respuesta	Predictor de interés (X1)	Sí/No
Histología	Covariable de ajuste (X2)	Adenocarcinoma/escamoso/otros
Raza	Covariable de ajuste (X3)	Oriental/caucásico
Sexo	Covariable de ajuste (X4)	Mujer/hombre
RIN	Covariable de ajuste (X5)	[2,...7,9]

En este modelo, la respuesta variable se definió de la siguiente manera:

- Respuesta = Sí: pacientes con respuesta parcial como su mejor respuesta (n=6)
- Respuesta= No: pacientes con enfermedad progresiva (PD) como su mejor respuesta o sin evaluación tumoral (n=65)

Se excluyeron los pacientes con un estado de enfermedad estable (n=31). La justificación fue que, centrándose en los pacientes con la respuesta más marcada a la terapia, el grupo no respondedor se haría más homogéneo.

Para cada conjunto de sondas, el objetivo del ensayo estadístico fue rechazar la hipótesis de que los niveles medios de expresión en los pacientes con respuesta al tratamiento y en los pacientes sin respuesta al tratamiento son

iguales, teniendo en cuenta las otras covariables de ajuste enumeradas en la tabla 2. Formalmente, la hipótesis nula de igualdad se evaluó frente a una alternativa de doble cara. Bajo la hipótesis nula, la distribución del estadístico Wald para esta evaluación sigue una distribución chi-cuadrado con 64 grados de libertad. Los valores de p correspondientes se describen en la tabla 3.

La elección del modelo lineal estuvo motivada por dos razones. La primera es que el modelo lineal es un enfoque versátil, bien caracterizado y robusto que permite el ajuste de variables de confusiones cuando se estima el efecto de la variable de interés. La segunda es que, con un tamaño muestral de 71 y la normalización y escalado del conjunto de datos, la asunción de la distribución normal es razonable y está justificada.

El asunto de la evaluación múltiple se abordó mediante la utilización de un criterio de tasa de falsos descubrimientos (FDR) para identificar la lista de genes expresados de forma diferencial (Benjamini et al., Journal of the Royal Statistical Society Series B-Methodological, 1995. 57(1): páginas 289-300). Los conjuntos de sondas con una FDR inferior al límite de 0,3 se consideran significativos. El punto de corte de 0,3 se escogió como un compromiso razonable entre una corrección rigurosa de la evaluación múltiple con un control estricto del riesgo de falsos positivos y el riesgo de no identificar marcadores verdaderamente diferenciales. El gen marcador identificado se describe en la tabla 3.

Tabla 3: marcadores basados en la comparación de "respondedores" con "progresores". Los respondedores se definieron como pacientes con una "respuesta parcial" (PR) como mejor respuesta. Los "progresores" se definieron como pacientes con una "enfermedad progresiva" (PD) o sin una evaluación disponible como mejores respuestas.

Los pacientes sin evaluación tumoral se incluyeron en el grupo de "progresores" porque en la mayoría de casos no se disponía de evaluación dada la pérdida temprana del paciente a causa de la progresión de la enfermedad o la muerte.

La columna 1 incluye el identificador Affymetrix del conjunto de sondas. La columna 2 incluye el número de acceso de GenBank de la secuencia génica correspondiente. La columna 3 incluye el nombre oficial del gen correspondiente. La columna 4 incluye el grado medio correspondiente y ajustado del cambio del nivel de expresión entre "respondedores" y "progresores", tal y como se estima con el modelo lineal. La columna 5 incluye el valor de p para la evaluación de la diferencia en el nivel de expresión entre "respondedores" y "progresores", tal y como se deriva del modelo lineal. La columna 6 incluye el intervalo de confianza del 95% para el cambio medio ajustado del nivel de expresión.

ID del conjunto de sondas Affymetrix	GenBank	Gen	Grado de cambio medio ajustado	Valor de p	IC 95%
204681_s_at	NM_012294	RAPGEF5	2	2,10E-05	1,5; 2,7

Para cada conjunto de sondas se evaluó la asunción de homogeneidad de varianzas mediante la utilización de ensayos Fligner-Killeen en base a los residuos del modelo. El análisis incluyó tres pasos:

Evaluar la igualdad de todas las variables categóricas en relación a la variancia residual entre sus niveles.

Indicar la variable V con el menor valor de p.

Si el valor menor de p es inferior a 0,001, reajustar el modelo para permitir que los diferentes niveles de las variables V tengan una variancia diferente.

Análisis estadístico adicional

Los siguientes análisis se llevaron a cabo con el marcador candidato seleccionado RAPGEF5 en un ambiente validado, por parte de estadísticos independientes:

- Regresión de Cox univariante para la PFS (supervivencia libre de progresión) a partir del análisis primario de Affymetrix,
- Regresión logística univariante para la respuesta a partir del análisis primario de Affymetrix, y
- Regresión de Cox univariante para la supervivencia a partir del análisis primario de Affymetrix

Los resultados de estos análisis se presentan más adelante. Estos son consistentes con los resultados del análisis primario y confirman la elección del marcador seleccionado.

Resultados: regresión de Cox univariante para la PFS (supervivencia libre de progresión) a partir del análisis primario de Affymetrix:

Gen	Nº de pacientes	Tasa de riesgo	IC del 95% para la tasa de riesgo	Valor de p
-----	-----------------	----------------	-----------------------------------	------------

PSPH 102 0,44 0,28; 0,69 0,0003

Resultados: regresión de Cox univariante para la respuesta a partir del análisis primario de Affymetrix:

Gen	Nº de pacientes	Odds Ratio	IC del 95% para el Odds Ratio	Valor de p
PSPH	102	95,77	5,84; >1000	0,0014

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

Gen	Nº de pacientes	Tasa de riesgo	IC del 95% para la tasa de riesgo	Valor de p
PSPH	102	0,47	0,28; 0,78	0,0040

qRT-PCR

10 Se sintetizó cDNA mediante la utilización de SuperScript™ III First-strand Synthesis SuperMix para qRT-PCR (Invitrogen, CA, EE.UU.) de acuerdo con las instrucciones del fabricante pero sin incluir la digestión con una RNasa H.

15 Se realizó una PCR cuantitativa mediante la utilización de ensayos de expresión génica TaqMan® en un sistema de detección de secuencia ABI PRISM® 7900HT de acuerdo con las recomendaciones del fabricante (Applied Biosistemas, CA, EE.UU.). Todos los ensayos se realizaron por triplicado.

20 Los ensayos de expresión génica utilizaron Hs00920287_ml [RAPGEF5] y se escogieron para que los cebadores y las sondas se cruzaran con los extremos del exón o se encontraran en la secuencia de la sonda de interés Affymetrix Genechip®. Se incluyeron dos genes constitutivos como controles endógenos: beta-2-microglobulina (B2M; Ensayo Hs99999907_ml) y hipoxantinfosforibosil transferasa (HPRT; Ensayo Hs99999909_ml).

25 Todas las ejecuciones incluyeron una muestra calibradora (RNA total MVP™ de pulmón de adulto humano; Stratagene, CA, EE.UU.) y una curva estándar. Se utilizó RNA total humano de referencia universal (Stratagene, CA, EE.UU.) como plantilla para las curvas estándar de EGFR y PSPH. El mRNA de RAPGEF5 estaba presente a un nivel inaceptablemente bajo en el RNA total humano de referencia universal, por lo que se utilizó el RNA total MVP™ de pulmón de adulto humano para construir las curvas estándar para este gen. Todas las muestras se midieron por triplicado. La cuantificación relativa se llevó a cabo mediante la utilización del método -ΔCt.

30 Resultados

35 Tal y como se describe con anterioridad, se determinaron los perfiles de expresión génica de Affymetrix Genechip® de los 102 pacientes incluidos en este estudio. De entre estos pacientes, se obtuvieron resultados de qRT-PCR de 75 pacientes (tabla 4). Las características demográficas y clínicas de los pacientes con resultados de qRT-PCR fueron similares a los de la población total (n=264) y a los pacientes con perfiles de expresión génica Genechip® disponibles.

Característica	
Edad (media, intervalo)	62 (39-85)
Género; n (%)	
Hombre	19 (25)
Mujer	56 (75)
Estado funcional ECOG; n (%)	
0	7 (9)
1	45 (60)
2	23 (31)
Histología; n (%)	
Adenocarcinoma	27 (36)
Carcinoma escamoso	34 (45)
Carcinoma de célula grande	2 (3)
Otros	12 (16)
Estadio de enfermedad	
IIIB	22 (29)
IV	53 (71)
Número de regímenes de quimioterapia previos; n (%)	
0	19 (25)
1	36 (48)
≥2	20 (27)
Etnia; n (%)	
Caucásico	51 (68)

Asiático	24 (32)
Historia de consumo de tabaco; n (%)	
Nunca	12 (16)
Actual	24 (32)
Ex fumador	39 (52)

De los 75 pacientes con resultados de qRT-PCR, 4 (5%) obtuvieron una respuesta parcial (PR), 23 (31%) presentaron una SD, 39 (52%) presentaron una PD, y 9 (12%) no se evaluaron. Estos resultados fueron muy similares a los observados en la población completa del estudio (n=264).

Correlación entre los datos de Affymetrix Genechip®, los datos de la qRT-PCR y los resultados clínicos

La figura 3 muestra niveles relativos de mRNA del RAPGEF5 en pacientes individuales, tal y como se evalúa mediante el perfilado de Affymetrix GeneChip® y la qRT-PCR. La figura 3a muestra los niveles de expresión frente al resultado clínico del perfilado de Genechip® y la figura 3b muestra los niveles de expresión frente al resultado clínico de la qRT-PCR.

La figura 3c muestra la correlación entre las mediciones de Genechip® y de qRT-PCR para RAPGEF5.

No se observó correlación alguna entre las mediciones de Genechip® y de qRT-PCR para RAPGEF5. De algún modo, la asociación entre el resultado clínico con erlotinib y los niveles del mRNA de RAPGEF5 evaluados mediante qRT-PCR fue más débil de la observada con el perfilado de Genechip® (figura 3b).

Identificación de genes asociados con la respuesta a erlotinib

Los respondedores se definieron como pacientes cuya mejor respuesta fue la respuesta parcial, mientras que los "progresores" se definieron como pacientes con una enfermedad progresiva o en los que no se realizó la evaluación (en la mayoría de casos a causa de una pérdida precoz debida a la progresión de la enfermedad o a la muerte). Por consiguiente, en este modelo se compararon 6 "respondedores" con 65 "progresores".

Se ajustó de forma independiente un modelo lineal a cada uno de los 16.353 conjuntos de sondas restantes utilizados en el análisis tras la extracción de aquellos conjuntos de sondas que no estuvieron presentes en ninguna de las muestras del total de 22.283 del microarray HG-U133A. Se calculó un valor de p para la diferencia de expresión entre la respuesta y la ausencia de respuesta con cada conjunto de sondas. Se aplicó una tasa de falsos descubrimientos (FDR) de 0,3 para corregir la evaluación múltiple. El marcador identificado en este análisis se muestra en la tabla 3.

Discusión

Se propuso la focalización en el receptor del factor de crecimiento epidérmico (EGFR) como medio de terapia contra el cáncer en base a su expresión aberrante ubicua en varios cánceres epiteliales (Davies & Chamberlin 1996, Mendelsohn, 2002). El EGFR está implicado en la patogénesis y la progresión de muchos tumores, entre los que se incluye al 40-80% de los tumores el NSCLC (Baselga & Albanell, 2002; Normanno et al, 2006), como resultado de mutaciones activadoras del dominio tirosina quinasa y/o su amplificación. Tras la activación, el receptor se somete a su dimerización, lo cual resulta en la fosforilación de dianas descendentes con roles en la proliferación celular, la metástasis, la inhibición de apoptosis y la neoangiogénesis (Mendelsohn & Baselga, 2003).

Se han desarrollado dos clases principales de inhibidores del EGFR, anticuerpos monoclonales dirigidos contra el dominio extracelular del receptor, y moléculas pequeñas inhibidoras de la tirosina quinasa del dominio catalítico del receptor. Estas últimas incluyen el erlotinib, que compite con el ATP por el lugar de unión intracelular.

En los últimos años se ha visto que varios factores juegan un rol en la sensibilidad al erlotinib, entre los que se incluye el género femenino, el estado de no fumador, el origen asiático y la histología de adenocarcinoma; dado que las tasas de respuesta potenciada son evidentes en tales subgrupos clínicos de pacientes, se están focalizando muchos esfuerzos en el descubrimiento de marcadores moleculares predictivos para la estratificación de pacientes. Las mutaciones en el EGFR, la amplificación del locus del gen EGFR y la sobreexpresión de EGFR sobre el nivel de la proteína se han asociado a la variación en los grados de respuesta, pese a que estos no son los únicos determinantes moleculares de respuesta.

Mediante el análisis de las muestras tisulares con la tecnología de microarrays de oligonucleótidos de alta densidad y la aplicación del modelado estadístico de los datos hemos sido capaces de identificar un gen cuyos niveles de expresión son predictivos de la respuesta al erlotinib (comparación de PR frente a PD) (tabla 3). Este gen es RAPGEF5, que se localiza en el cromosoma 7p15.3, también se sobreexpresa en los respondedores (regulación positiva 2,0 veces mayor; p=0,000021).

Se observó que RAPGEF5 se encuentra regulado positivamente aproximadamente 2 veces más en los pacientes con PR en comparación sólo con los pacientes con PD ($p=0,000021$) de nuestro análisis. RAPGEF5 es un factor de intercambio de nucleótidos de guanina (GEF) para Rap1, un miembro de la subfamilia RAS de GTPasas que funcionan en la transducción de señales. Los GEF proporcionan el enlace crucial entre los receptores de la superficie de la célula y la activación de la GTPasa mediante el ciclado entre los estados de unión a GDP inactivos y los de unión a GDP activos. Rap1 se relaciona estrechamente con Ras y está implicado en la regulación de varios procesos celulares que incluyen el control de la activación plaquetaria, la anergia de linfocitos T, la activación de linfocitos B, la diferenciación neuronal y está involucrado en el control de la adhesión celular y la regulación de la activación de integrinas. También es posible que alguno de estos procesos también sea relevante en la progresión del NSCLC. RAPGEF5 puede tener un papel indirecto en la angiogénesis mediante su control de Rap1, el que recientemente se ha identificado como regulador esencial de la quimiotaxis de células vasculares. La promoción de la neoangiogénesis es una característica prominente del crecimiento neoplásico; recientemente, se ha observado que bevacizumab, un agente que se dirige específicamente contra uno de los actores principales de la angiogénesis, el factor de crecimiento endotelial vascular (VEGF), alarga la supervivencia en combinación con la quimioterapia en el NSCLC no escamoso. Aún no se ha determinado el papel preciso del RAPGEF5 en la carcinogénesis, pero puede que proporcione un enlace adicional hacia la vía MAPK a través de la cual se sabe que señala el EGFR.

Tabla 5: lista de variantes de RAPGEF5 y sus correspondientes números de identificación de secuencia

Número de acceso al GenBank	Gen	Número de identificación de la secuencia
NM_012294	RAPGEF5	SEC. ID. N° 1
XM_928602	RAPGEF5 variante_1	SEC. ID. N° 2
XM_935655	RAPGEF5 variante_6	SEC. ID. N° 3
XM_937844	RAPGEF5 variante_11	SEC. ID. N° 4

LISTADO DE SECUENCIAS

<110> F. Hoffmann-La Roche AG

<120> Marcador predictivo para el tratamiento con un inhibidor del EGFR

<130> 24413WO

<150> EP07114314.3

<151> 14/08/2007

<150> EP08156516.0

<151> 20/05/2008

<160> 4

<170> PatentIn versión 3.4

<210> 1

<211> 6629

<212> DNA

<213> Homo sapiens

<400> 1

```

gctgcgccg cggctgaggg acctgcccgc gctgctgctg agcgggctca cgctgcgag      60
gaagcggagc gccgctgggg gccggactct atcaagaaga atatctaata cgtaccttga      120
acatacgccct tcccagattt atggagagaa ttcttcttgt gcaggaagag cattgaggaa      180
tattattatc gttcaagcag ctgacctgat aaaggacaga gtgaacctca aggggtttta      240
caggaggagc tgcgttgggt cagagctggt agactggctt ctagaacact gtcctttcgt      300
ccagtgcaga tctatggcca taggagtctg gcaactccta ctggacatgg gaattatggt      360
atcagtggac cagcatctat actttcaaga tacttatggt ttctaccagt tttcctctga      420
tgaatgtagc tacttgtact gtgaatttga aagagaagaa gaatggcaaa atggtgtcaa      480
gcttttactg caacttgtgc ctctcattcc tgccagaggt ggcattctgt aactgtctca      540
tcagaaaatt gaagactccg aagaaagcag tgatgaaatt cttgtgcgtc taacatctgc      600
gggtgcagaga gagctagcag ctgttattgc tttgaaagca aggaagtctg caattgaaca      660
agatgaagaa aacaacgaca aacatgtagc tgtaacagaa gccgaaagtg ttccagattc      720
tcaggcaggg gtgatgtgca agctccagga aagagatgaa atcggacgaa ttgaactagt      780
ccagaagctg gcaaaagaaa actatcagtt tttgcagacg gacaaaaaag aacaggagaa      840
gtctgaacac caagatgatg aagtgcgcgc tgttcagggt aaagagcaag accagagcgt      900
cctgggtgctg aagaaagtgc agtgctgtgg cccagcccc acagctggga gtgctggagag      960
ccattggaga tatgtggtgg tgtccgggac cccggagaag attttgagc accttttgaa     1020
tgacttgcac ctggaagaag tccaggacaa agaaacagag accctcctgg atgacttcct     1080
tctcacgtac actgtcttca tgacaactga tgacttgctg caggctctgt taaggcacta     1140
ttctgctaag aagtatcaag gcaaagagga aaactcagac gttccgcgta ggaaacgtaa     1200
agtcttgcac cttgtttccc agtggattgc tctgtacaaa gactggttac ctgaagatga     1260
acattcaaaa atgtttttta agaccatata taggaatgta ctggatgatg tttatgaata     1320

```

tccaatactt	gaaaaagaat	tgaaagaatt	tcaaaagata	cttggaatgc	accgtcgtca	1380
cactgtagat	gaatattcac	cacaaaaaaa	gaataaagcc	cttttccacc	aattcagtct	1440
taaggagaac	tggtccagc	atagaggaac	tgtgactgaa	acggaggaaa	ttttctgcca	1500
cgtgtatata	acagagcact	cctatgtcag	tgtgaaggca	aaagtttcca	gtatagccca	1560
agagatccta	aaagtcgtgg	cagaaaagat	ccagtatgca	gaagaggatc	tggtcttggg	1620
ggccatcaca	ttctctgggg	aaaagcatga	acttcagcca	aatgacttag	tcattctcaa	1680
atccctcgag	gcatctgggc	gaatatatgt	ctaccggaaa	gacctggcgg	acactttgaa	1740
cccatttgca	gaaaatgagg	aatcacagca	aaggctcgatg	aggattttgg	gaatgaacac	1800
ttgggatctt	gctctggaat	taatgaatgt	tgattggagt	ctattcaatt	caattcacga	1860
gcaagagctg	atctacttca	cgttcagcag	acaggggaagt	ggggaacaca	ctgcaaattct	1920
cagccttctg	ctccagagat	gcaatgaggt	ccagctttgg	gtggccacgg	agattctgct	1980
ctgcagccag	ctgggcaagc	gagtgacgct	ggtgaaaaaa	ttcatcaaaa	ttgcggctca	2040
ctgcaaagcc	cagagaaaacc	tgaattcttt	ctttgccatt	gtgatgggtc	tcaacactgc	2100
ttctgtcagt	cgactgtcgc	agacctggga	gaaaatccct	gggaagttaa	agaaactttt	2160
ctctgaactt	gaaagttaa	cagatccttc	cctaaatcac	aaagcctaca	gagatgcatt	2220
caaaaagatg	aagccaccaa	aaatcccttt	catgccctta	ttgcttaaag	atgtaacatt	2280
tattcatgaa	ggaaataaaa	cttttttggg	taatcttgtc	aattttgaaa	agctgcatat	2340
gatcgagac	actgtccgaa	ccctgagaca	ctgcaggact	aaccagtttg	gtgacctgtc	2400
tccaaaagag	catcaagagt	taaagtccta	tgtaaatcac	ctgtatgtca	ttgacagcca	2460
gcaggctctg	tttgagctct	cacacaggat	cgagcctcgg	gtgtgagccc	cactgcctca	2520
cctccctgtg	atctgcagca	ctttgagcta	cgggaatgtc	tatgccaagc	acgttgcttt	2580
cctgtgagaa	aagaagttgc	tgagttttat	cagtataacc	caagacattc	acaggaaagc	2640
cagccaaagc	gtgttcagga	agtgatgtca	gccaccagag	agggggagag	gtttctccat	2700
gctactctcg	ggacaagaag	gcagaaggag	agtcagaagc	attcttgaga	tggaagaaggc	2760
tggtttctta	tgatcacatt	gttgatccag	tccagttttc	aatatgagat	gtgccagcat	2820
caagacaaga	caacgtcttg	acatgcaatg	accaaataat	tcattaagag	cgtgcatgaa	2880
acaggaagga	gtttttactt	tgccatagttt	tagattactg	tccataagct	gtcaaagaag	2940
tcattctttt	gaacacctga	tgacagagac	agcatctcta	gatctccagg	gaggagaggt	3000
ttctgttgat	acaacctgtg	acatcaccaa	aagccacttg	tgtctagggg	gttagtgagg	3060
actgcagcta	gcatccatgc	tctgatgggc	agatgaacaa	tgtcaagggtg	tgcatcactt	3120
tgaccacaa	tcaactattg	acacatgctt	gcagggtgaaa	ttagtttctg	tacaactgat	3180
ttgcagctat	aggcaaggta	gatgaagttg	ctttgccagt	aaggaaaaat	agtaatcttt	3240
aagaaattga	ctcattgttt	aatttctggg	gattttcttt	atacttctaa	gcaggctctt	3300
atctttttatt	ggacataata	tgattttgaa	aaagcacagt	gcctgacaca	ttgcaaacac	3360

tcaccaactg	cttgcctgagg	tgacagagtc	acaaaagtct	gcattcttgt	gcctgatgat	3420
gcatttttgcg	tacctcatat	aggctccttg	cccacactat	ggaatgacag	cagccagtgc	3480
agggagggtta	agtgacattt	aatgagtgaa	gcacttagca	ctctctaggt	aataagatag	3540
tggttaattac	tagtgttttg	gcaaatgaaa	aatgccctga	aatagccaaa	tgtctgatta	3600
atgttggcaa	cttagaagtc	ctataatcca	actaccagcc	aaagcaggga	gcctttctat	3660
aatttgcctt	tttttttttt	tttttcaaaa	tctgagtctt	ctaaaatctt	attattccca	3720
tttttaccaa	ttgaggctcc	tgtagcaaat	aagacctctt	gatattttca	aggactgggt	3780
agaggatttc	tttcaacctt	cacatgaaca	aaacagccta	tgggtcaaaa	taatgaaatc	3840
caccttgcgc	tgctagatac	ttgtcacctt	gctaaaatgc	aagggcctgg	tccattcatt	3900
ttccaaatgc	aggagtcttg	gtgcacttct	cactcttcct	gcctgttcat	ctctttcatg	3960
cccacacaga	cctgtttcct	ttttgtctca	tcaacgcctc	attcatcctc	attactgagg	4020
cgtgtccaat	gctttttgac	atctttatag	cagtgtgtgt	tcctgggctc	aggaaccaca	4080
ctgagcttga	gatactgctg	gaaggaacca	tgtggagaga	aggtttggga	gaactttgag	4140
agagacttag	tttggcccag	catgtaaaac	ttcagtcctg	aacatttata	gggttttata	4200
gaagggcatc	ctccagggct	ggtccattca	gagaaatgct	gcatgctgcc	gtcatggaat	4260
gtggcccaca	ggacaccaga	gccgtgagaa	ccggagagca	gacttccttc	acggctgggc	4320
tgagcaaacc	ctccaaagcc	ctcctcacgc	agttactaac	aatagcatgg	gcttacagca	4380
caagcacgtg	ttctcacctt	tttcctatgc	cctggactaa	ggtttggcca	gtgtaatcat	4440
ataaggccat	cctgacattg	tttctgtgtt	tcaaaatttg	gatttttatt	tacattagaa	4500
ctacattgct	cctagtagaa	cattaccttt	aggggactaa	ttttccatgg	agaactattt	4560
cagcatattg	catgctgctc	agaccccaag	tcagatatgc	ccaccaagcc	agatgaagct	4620
acacaaatgt	ggtattttaa	tgcattttgt	acagtgactt	cagagtatct	cacatgacat	4680
gggtgtaaac	tggctgggga	gaaaatgatg	cttggttcacc	tcttcctcca	gccgtgggta	4740
ggtggtccta	ggggtagcag	aggggaaggga	ggattttgtg	cagtcaagat	ttgcttttcc	4800
atccttgtct	tctgaatgct	taaaatctct	gcatctttct	gaagttaaac	aactgtctcc	4860
agaggtttgc	caggcagcag	ctctcagaag	tttccaaagc	tttgcagaat	cttagatctg	4920
gaattaaaga	attcaagccc	gaattgtgag	aaccagatat	tactcaacag	aaagctcttt	4980
ctaaggaatc	tgagctgttc	actggtggac	agtgggtggg	cttgagtgtc	ccttggttaat	5040
aggatgggcc	atgcaccctc	tctggatatt	caccaaggcc	tcttcagaat	agggtttgtt	5100
ctggctagaa	gcgtggtcta	gaagatggct	aagctctttg	ccagctctca	tttggagttt	5160
tattattgca	taaaatcttc	gtcactctg	caaactttac	gtaatctggc	accttcggca	5220
ccaggtgggtg	caggggcact	tctaagtggg	ctctttttgt	tacagcacia	ctctcagaca	5280
gtcctgtggg	tctttggatt	cgtcagcatt	ccagcaaact	agccctgctt	agaagttagc	5340
acaagacagc	agaatgcagg	accccgtagg	caaaatcaca	accttgctat	taaaaaaaat	5400


```

tttgttttac atacacattt gcaggtgttc cctagagtgt ggtgttttga atttgctctt 5460
tgtcatctgt ataattgcc aatgattata gtgatacaca tgacctgcat tcactttttt 5520
ctagtttcct taattatgtt tagaataaat tcatttcctt agaccgagaa ccacaaacag 5580
gtagtgtgga gcatacaccg aatttagaag catgtggata aggtcagtgc tcacactgcc 5640
tagtccacag ggagaggatg ctgcatgaat atatacttgc ctctgagtgg aggagaaatc 5700
gtggcatgaa agagagagta ccagtgtatga cttcttatcc ctggagctgg gctttcactg 5760
ctacccatat cccagccctg cgagtctgtt ctagccagca cagacaccgc agatccggaa 5820
ctgaatgttc ctaaatggcg cagccaatcc aggcctttca gaaactgggc aaaaacatta 5880
aaatggggac gatcgggtct tccgcagtgg tccaacacag gatttctttt aaatgtttca 5940
aaaacatgtc cttaaaattt cagcctgctt cttagcgagt gggccagttt tgcttaaaac 6000
tgggtggggg gcggggggga agtttttaaa aattgccaaa aagtttagatg caaatgtatt 6060
actgtataaa gcaaagctgt atatactaaa catttttttag cagagtaata tttatttgca 6120
tagtctatct attgtattcg tatcacactg ttattaaata ctgggatgaa atcagtgacc 6180
tgaagcaaga aatcttgcct tttaatgtat cattaattag ggctgctgtg atattgtcag 6240
cttgcattaa caattagaag atagagaacc cgccatcagg gtgtctacct aacttctcag 6300
ggactacact tggtagtttt ccaccattta aagaactggg aaatatgaaa catttggtga 6360
gttaccagaa ttgccattaa cagtgttttc tttcccatat tccatgcttt ctgcctctgt 6420
gtatatatat aatatatatg tatatgactg tgctgtgtat ttatcgaagc tagtaagcaa 6480
taatttatat gtaaaaatgg ccaagcaata taagggtaaa acttatataa gtaaccctta 6540
ccttatcttg tattttcaat ttttttttaa aactgctttt ccaaatatga gactatgtta 6600
aagacactat tcatgcctct taaaaaaaaa 6629

```

<210> 2

<211> 2619

<212> DNA

5 <213> Homo sapiens

<400> 2

```

atgaaagcaa agcagtataa gctgccagcc ctccctcaag gagcttatgg attccttggg 60
gcaaatggac acacagaaaa acagtcaaca atgctgagca cccaggtgct gcacctgaat 120
agtagcaaaa acagagctct aaagaggagc gtacccacagc aggaggaact tcatgcagga 180
gccgactcca caaaactgga cttcgagtct ggactagtca agtgtcagct tactcacacc 240
gcttcctcga tttatggaga gaattcttct tgtgcaggaa gagcattgag gaatattatt 300
atcgttcaag cagctgacct gataaaggac agagtgaacc tcaaggggtt ttacaggagg 360
agctgcgttg ggtagagct ggtagactgg cttctagaac actgtccttt cgtccagtgc 420
agatctatgg ccataggagt ctggcaactc ctactggaca tgggaattat gttatcagtg 480
gaccagcatc tatactttca agatacttat gttttctacc agttttcctc tgatgaatgt 540
agctacttgt actgtgaatt tgaaagagaa gaagaatggc aaaatgggtg caagctttta 600

```

```

ctgcaacttg tgcctctcat tcctgccaga ggtggcatct gtgaactgtc tcatcagaaa 660
attgaagact ccgaagaaaag cagtgatgaa attcctgtgc gtctaacatc tgcggtgcag 720
agagagctag cagctgttat tgctttgaaa gcaaggaaat ctgcaattga acaagatgaa 780
gaaaacaacg acaaacatgt agctgtgaaca gaagccgaaa gtgttcacaga ttctcaggca 840
ggggtgatgt gcaagctcca ggaaagagat gaaatcggac gaattgaact agtcacagaag 900
ctggcaaaaag aaaactatca gtttttgtag acggacaaaa aagaacagga gaagtctgaa 960
caccaagatg atgaagtgc gactgttcag gttaaagagc aagaccagag cgtcctggtg 1020
ctgaagaaaag tgcagtgtct tggcccagcc cccacagctg ggagtgcgga gagccattgg 1080
agatatgtgg tgggtgtccg gaccccgag aagatttttg agcacctttt gaatgacttg 1140
cacctggaag aagtcacaga caaagaaaca gagaccttc tggatgactt ccttctcacg 1200
tacactgtct tcatgacaac tgatgacttg tgccaggctc tgtaaggca ctattctgct 1260
aagaagtatc aaggcaaaga ggaaaactca gacgttcgc gtaggaaacg taaagtcttg 1320
catcttggtt cccagtggat tgctctgtac aaagactggg tacctgaaga tgaacattca 1380
aaaatgtttt taaagaccat atataggaat gtactggatg atgtttatga atatccaata 1440
cttgaaaaag aattgaaaga atttcaaaaag atacttgaa tgcaccgtcg tcacactgta 1500
gatgaatatt caccacaaaa aaagaataaa gcccttttcc accaattcag tcttaaggag 1560
aactggctcc agcatagagg aactgtgact gaaacggagg aaattttctg ccacgtgtat 1620
ataacagagc actcctatgt cagtgtgaag gcaaaagttt ccagtatagc ccaagagatc 1680
ctaaaagtcg tggcagaaaa gatccagtat gcagaagagg atctggctct ggtggccatc 1740
acattctctg gggaaaagca tgaacttcag ccaatgact tagtcatctc caaatccctc 1800
gaggcatctg gtcgaatata tgtctaccgg aaagacctgg cggacacttt gaaccattt 1860
gcagaaaaat aggaatcaca gcaaaggctg atgaggattt tgggaatgaa cacttgggat 1920
cttgctctgg aattaatgaa ttttgattgg agtctattca attcaattca cgagcaagag 1980
ctgatctact tcacgttcag cagacagga agtggggaac aactgcaaa tctcagcctt 2040
ctgctccaga gatgcaatga ggtccagctt tgggtggcca cggagattct gctctgcagc 2100
cagctgggca agcagagtga gctggtgaaa aaattcatca aaattgcggc tactgcaaa 2160
gcccagagaa acctgaattc tttctttgcc attgtgatgg gtctcaaac tgcttctgtc 2220
agtcgactgt cgcagacctg ggagaaaatc cctgggaagt ttaagaaact tttctctgaa 2280
cttgaaagtt taacagatcc ttccctaaat cacaagcct acagagatgc attcaaaaag 2340
atgaagccac caaaaatccc tttcatgccc ttattgctta aagatgtaac atttattcat 2400
gaaggaaata aaactttttt ggataatctt gtcaattttg aaaagctgca tatgatcgca 2460
gacactgtcc gaaccctgag aactgcagg actaaccagt ttggtgacct gtctccaaaa 2520
gagcatcaag agttaaagtc ctatgttaat cacctgtatg tcattgacag ccagcaggct 2580
ctgtttgagc tctcacacag gatcgagcct cgggtgtga 2619

```

<210> 3

<211> 2619

<212> DNA

<213> Homo sapiens

<400> 3

atgaaagcaa agcagtataa gctgccagcc ctccctcaag gagcttatgg attccttggg	60
gcaaatggac acacagaaaa acagtcaaca atgctgagca cccaggtgct gcacctgaat	120
agtagcaaaa acagagctct aaagaggagc gtacccagc aggaggaact tcatgcagga	180
gccgactcca caaaactgga cttcgagtct ggactagtca agtgtcagct tactcacacc	240
gcttcctcga tttatggaga gaattcttct tgtgcaggaa gagcattgag gaatattatt	300
atcgttcaag cagctgacct gataaaggac agagtgaacc tcaaggggtt ttacaggagg	360
agctgcgttg ggtcagagct ggtagactgg cttctagaac actgtccttt cgtccagtgc	420
agatctatgg ccataggagt ctggcaactc ctactggaca tgggaattat gttatcagtg	480
gaccagcatc tatactttca agatacttat gttttctacc agttttcctc tgatgaatgt	540
agctacttgt actgtgaatt tgaaagagaa gaagaatggc aaaatgggtg caagctttta	600
ctgcaacttg tgcctctcat tcctgccaga ggtggcatct gtgaactgtc tcatcagaaa	660
attgaagact ccgaagaaag cagtgatgaa attcttgtgc gtctaacatc tgcggtgcag	720
agagagctag cagctgttat tgctttgaaa gcaaggaagt ctgcaattga acaagatgaa	780
gaaaacaacg acaaacatgt agctgtaaca gaagccgaaa gtgttccaga ttctcaggca	840
ggggtgatgt gcaagctcca ggaaagagat gaaatcggac gaattgaact agtccagaag	900
ctggcaaaaag aaaactatca gtttttgtag acggacaaaa aagaacagga gaagtctgaa	960
caccaagatg atgaagtgaac gactgttcag gttaaagagc aagaccagag cgtcctggtg	1020
ctgaagaaaag tgcagtgtct tggcccagcc cccacagctg ggagtgcgga gagccattgg	1080
agatatgttg tgggtgtccg gaccccgag aagatttttg agcacctttt gaatgacttg	1140
cacctggaag aagtccagga caaagaaaca gagaccctcc tggatgactt ctttctcacg	1200
tacactgtct tcatgacaac tgatgacttg tgccaggctc tgttaaggca ctattctgct	1260
aagaagtatc aaggcaaaga ggaaaactca gatgttccgc gtaggaaacg taaagtcttg	1320
catcttgttt cccagtggat tgctctgtac aaagactggg tacctgaaga tgaacattca	1380
aaaatgtttt taaagaccat atataggaat gtactggatg atgtttatga atatccaata	1440
cttgaaaaag aattgaaaga atttcaaaag atacttgga tgcaccgtcg tcacactgta	1500
gatgaatatt caccacaaaa aaagaataaa gcccttttcc accaattcag tcttaaggag	1560
aactggctcc agcatagagg aactgtgact gaaacggagg aaattttctg ccacgtgtat	1620
ataacagagc actcctatgt cagtgtgaag gcaaaagttt ccagtatagc ccaagagatc	1680
ctaaaagtcg tggcagaaaa gatccagtat gcagaagagg atctggctct ggtggccatc	1740
acattctctg gggaaaagca tgaacttcag ccaaagact tagtcatctc caaatccctc	1800

gaggcatctg gtcgaatata tgtctaccgg aaagacctgg cggacacttt gaacccattt	1860
gcagaaaatg aggaatcaca gcaaagggtcg atgaggattt tgggaatgaa cacttgggat	1920
cttgctcttg aattaatgaa ttttgattgg agtctattca attcaattca cgagcaagag	1980
ctgatctact tcacgttcag cagacaggga agtggggaac aactgcaaa tctcagcctt	2040
ctgctccaga gatgcaatga ggtccagctt tgggtggcca cggagattct gctctgcagc	2100
cagctgggca agcgagtgca gctggtgaaa aaattcatca aaattgcggc tactgcaaa	2160
gcccagagaa acctgaattc tttctttgcc atttgtatgg gtctcaacac tgcttctgtc	2220
agtcgactgt cgcagacctg ggagaaaatc cctgggaagt ttaagaaact tttctctgaa	2280
cttgaaagtt taacagatcc ttccctaaat cacaaagcct acagagatgc attcaaaaag	2340
atgaagccac caaaaatccc tttcatgccc ttattgctta aagatgtaac atttattcat	2400
gaaggaaata aaactttttt ggataatctt gtcaattttg aaaagctgca tatgatcgca	2460
gacactgtcc gaaccctgag aactgcagg actaaccagt ttggtgacct gtctccaaaa	2520
gagcatcaag agttaagtc ctatgttaat cacctgtatg tcattgacag ccagcaggct	2580
ctgtttgagc tctcacacag gatcgagcct cgggtgtga	2619

<210> 4
 <211> 2619
 <212> DNA
 <213> Homo sapiens

5

<400> 4	
atgaaagcaa agcagtataa gctgccagcc ctccctcaag gagcttatgg attccttggg	60
gcaaatggac acacagaaaa acagtcaaca atgctgagca cccagggtgct gcacctgaat	120
agtagcaaaa acagagctct aaagaggagc gtacccagc aggaggaact tcatgcagga	180
gccgactcca caaaactgga cttcgagtct ggactagtca agtgtcagct tactcacacc	240
gcttcctcga tttatggaga gaattcttct tgtgcaggaa gagcattgag gaatattatt	300
atcgttcaag cagctgacct gataaaggac agagtgaacc tcaaggggtt ttacaggagg	360
agctgcgttg ggtcagagct ggtagactgg cttctagaac actgtccttt cgtccagtgc	420
agatctatgg ccataggagt ctggcaactc ctactggaca tgggaattat gttatcagtg	480
gaccagcatc tatactttca agatacttat gttttctacc agttttcctc tgatgaatgt	540
agctacttgt actgtgaatt tgaaagagaa gaagaatggc aaaatgggtg caagctttta	600
ctgcaacttg tgcctctcat tcctgccaga ggtggcatct gtgaactgtc tcatcagaaa	660
attgaagact ccgaagaaaag cagtgatgaa attcttgtgc gtctaacatc tgcggtgcag	720
agagagctag cagctgttat tgctttgaaa gcaaggaagt ctgcaattga acaagatgaa	780
gaaaacaacg acaaacatgt agctgtaaca gaagccgaaa gtgttccaga ttctcaggca	840
ggggtgatgt gcaagctcca ggaaagagat gaaatcggac gaattgaact agtccagaag	900
ctggcaaaag aaaactatca gtttttgtag acggacaaaa aagaacagga gaagtctgaa	960
caccaagatg atgaagtgac gactgttcag gttaaagagc aagaccagag cgtcctggtg	1020

ctgaagaaag tgcagtgtctg tggcccagcc cccacagctg ggagtgcgga gagccattgg	1080
agatatgtgg tgggtgtccg gaccccgag aagattttgg agcacctttt gaatgacttg	1140
cacctggaag aagtccagga caaagaaaca gagaccctcc tggatgactt ccttctcacg	1200
tacactgtct tcatgacaac tgatgacttg tgccaggctc tgttaaggca ctattctgct	1260
aagaagtatc aaggcaaaga ggaaaactca gatgttccgc gtaggaaacg taaagtcttg	1320
catcttgttt cccagtggat tgctctgtac aaagactggt tacctgaaga tgaacattca	1380
aaaatgtttt taaagaccat atataggaat gtactggatg atgtttatga atatccaata	1440
cttgaaaaag aattgaaaga atttcaaaag atacttggaa tgcaccgtcg tcacactgta	1500
gatgaatatt caccacaaaa aaagaataaa gcccttttcc accaattcag tcttaaggag	1560
aactggctcc agcatagagg aactgtgact gaaacggagg aaattttctg ccacgtgtat	1620
ataacagagc actcctatgt cagtgtgaag gcaaaagtgt ccagtatagc ccaagagatc	1680
ctaaaagtctg tggcagaaaa gatccagtat gcagaagagg atctggctct ggtggccatc	1740
acattctctg gggaaaagca tgaacttcag ccaaagact tagtcatctc caaatccctc	1800
gaggcatctg gtcgaatata tgtctaccgg aaagacctgg cggacacttt gaacccattt	1860
gcagaaaatg aggaatcaca gcaaaggctg atgaggattt tgggaatgaa cacttgggat	1920
cttgctctgg aattaatgaa ttttgattgg agtctattca attcaattca cgagcaagag	1980
ctgatctact tcacgttcag cagacaggga agtggggaac aactgcaaa tctcagcctt	2040
ctgctccaga gatgcaatga ggtccagctt tgggtggcca cggagattct gctctgcagc	2100
cagctgggca agcgagtga gctggtgaaa aaattcatca aaattgcggc tctctgcaaa	2160
gcccagagaa acctgaattc tttctttgcc attgtgatgg gtctcaacac tgcttctgtc	2220
agtcgactgt cgcagacctg ggagaaaatc cctgggaagt ttaagaaact tttctctgaa	2280
cttgaaagt ttaacagatcc tttcctaaat cacaagcct acagagatgc attcaaaaag	2340
atgaagccac caaaaatccc tttcatgccc ttattgctta aagatgtaac atttattcat	2400
gaaggaaata aaactttttt ggataatctt gtcaattttg aaaagctgca tatgatcgca	2460
gacactgtcc gaacctgag aactgcagg actaaccagt ttggtgacct gtctccaaaa	2520
gagcatcaag agttaagtc ctatgttaat cacctgtatg tcattgacag ccagcaggct	2580
ctgtttgagc tctcacacag gatcgagcct cgggtgtga	2619

REIVINDICACIONES

- 5 1. Un método in vitro para predecir la respuesta de un paciente con NSCLC al tratamiento con erlotinib que incluye: la determinación de un nivel de expresión del mRNA de un gen RAPGEF5 en una muestra tumoral de un paciente y la comparación del nivel de expresión del mRNA del gen RAPGEF5 con un valor representativo del nivel de expresión del mRNA del gen RAPGEF5 en tumores de una población de pacientes no respondedores, en el que un nivel de expresión mayor del mRNA del gen RAPGEF5 en la muestra tumoral del paciente indica que un paciente responderá al tratamiento.
- 10 2. El método de la reivindicación 1, en el que el nivel de expresión del mRNA se determina mediante la tecnología de microarrays o cualquier tecnología que mida la expresión del RNA.
- 15 3. El método de las reivindicaciones 1-2, en el que el gen RAPGEF5 muestra una expresión de 1,5 a 2,7 mayor (o superior) de expresión del mRNA en la muestra tumoral del paciente respondedor en comparación al valor representativo de un nivel de expresión del mRNA del gen RAPGEF5 en tumores de una población de pacientes no respondedores.
- 20 4. La utilización del gen RAPGEF5 para predecir la respuesta de un paciente con NSCLC al tratamiento con erlotinib.
5. Erlotinib para la utilización en el tratamiento de un paciente con NSCLC identificado mediante un método de las reivindicaciones 1-3 como paciente que responderá al tratamiento que incluye la administración de erlotinib al paciente.

Fig. 1

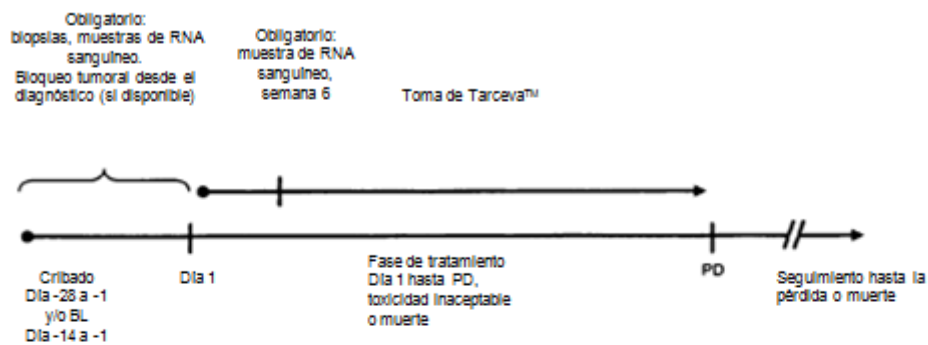


Fig. 2



Fig. 3a

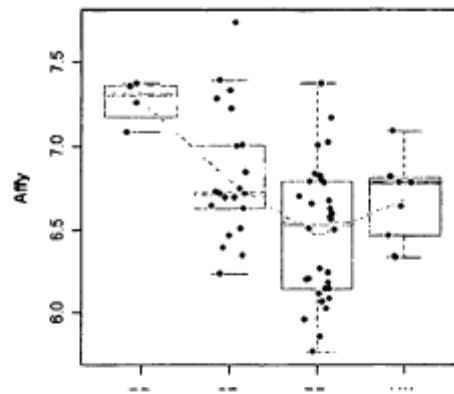


Fig. 3b

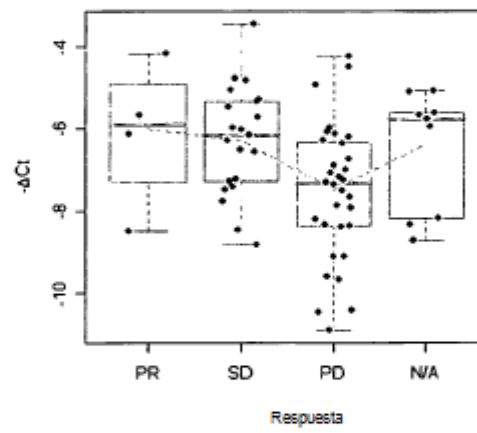


Fig. 3c

