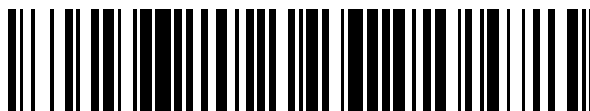


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 631 455**

51 Int. Cl.:

C12Q 1/68

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **07.12.2010 PCT/EP2010/069058**

87 Fecha y número de publicación internacional: **16.06.2011 WO11070005**

96 Fecha de presentación y número de la solicitud europea: **07.12.2010 E 10787761 (5)**

97 Fecha y número de publicación de la concesión europea: **08.02.2017 EP 2510113**

54 Título: **Alteraciones génicas de EGFR y PTEN predicen la supervivencia en pacientes con tumores cerebrales**

30 Prioridad:

07.12.2009 EP 09382271

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

31.08.2017

73 Titular/es:

**EUROPATH BIOSCIENCES, S.L. (100.0%)
Provença, 392 planta baja
08025 Barcelona, ES**

72 Inventor/es:

**DONOVAN, MICHAEL J.;
COLOMER VALERO, ANNA;
ERILL SAGALÉS, NADINA;
FERRER ABIZANDA, ISIDRE y
BOLUDA CASAS, SUSANA**

74 Agente/Representante:

ARIAS SANZ, Juan

Observaciones:

**Véase nota informativa (Remarks, Remarques o
Bemerkungen) en el folleto original publicado por
la Oficina Europea de Patentes**

ES 2 631 455 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

Alteraciones génicas de EGFR y PTEN predicen la supervivencia en pacientes con tumores cerebrales

Campo de la invención

- 5 La invención se refiere a los campos de diagnóstico y terapéutica, en particular a un método de proporcionar atención personalizada a pacientes con cáncer de cerebro basándose en la expresión de determinados genes en una muestra a partir de dichos pacientes, algunos de los cuales sirven como dianas de tratamiento.

Antecedentes de la invención

Gliomas: Diagnóstico y clasificación de la enfermedad

- 10 Un glioma es un tipo de cáncer que comienza en el cerebro o columna vertebral. Se denomina glioma porque surge de células gliales y/o sus precursores. El sitio más común de gliomas es el cerebro. Los gliomas se clasifican por tipo de célula, grado y localización. Los gliomas se nombran según el tipo específico de célula a la que más se asemejan. Los tipos principales de gliomas son:

- Ependimomas, gliomas derivados de células ependimarias.
- Astrocitomas, gliomas derivados de astrocitos; el glioblastoma multiforme (GBM) es el astrocitoma más común.
- 15 - Oligodendrogliomas, gliomas derivados de oligodendrocitos.
- Gliomas mixtos, tales como oligoastrocitomas, que contienen células de diferentes tipos de glía.

- 20 Los gliomas se clasifican además según su grado, que se determina mediante la evaluación patológica del tumor. Por tanto se puede distinguir entre gliomas de bajo grado que están bien diferenciados (no anaplásicos), benignos y presagian un mejor pronóstico para el paciente; y gliomas de alto grado, que no están diferenciados o anaplásicos, malignos y llevan a un peor pronóstico.

De los numerosos sistemas de calificación en uso, el más común es la Organización Mundial de la Salud (OMS) sistema de calificación para astrocitoma.

Tratamiento de gliomas cerebrales

- 25 El tratamiento para gliomas cerebrales depende de la localización, el tipo de célula y el grado de malignidad. A menudo, el tratamiento es un enfoque combinado, que usa cirugía, radioterapia y quimioterapia. La radioterapia es en forma de radiación con rayos externa o el enfoque estereotáctico usando radiocirugía. Los tumores de médula espinal pueden tratarse mediante cirugía y radiación. La temozolomida es un fármaco quimioterápico que puede cruzar la barrera hematoencefálica eficazmente y está usándose en terapia. A pesar de estos enfoques, la mayoría de los pacientes de glioma de alto grado sucumben a su enfermedad. Se necesitan nuevas intervenciones terapéuticas para dianas críticas para mejorar el desenlace en esta población de pacientes.

Glioblastoma multiforme (GBM)

- 35 El glioblastoma multiforme (GBM, grado IV de OMS) es un tumor cerebral altamente agresivo que se presenta como uno de los dos subtipos con perfiles moleculares y anamnesis distintos. El GBM primario se presenta completamente como una enfermedad de alto grado y el subtipo GBM secundario evoluciona desde la evolución lenta de una enfermedad de bajo grado.

Brown *et al.* (J Clin Oncology. 2008, 5603-5609) describen un ensayo de fase I/II de erlotinib combinado con temozolomida en pacientes que padecen GBM. Estos autores intentaron relacionar la respuesta de los pacientes con varios marcadores moleculares como EGFR, PTEN, P53, etc., pero dejaron de observar cualquier correlación entre supervivencia y niveles de expresión de dichos genes.

- 40 Mirimanoff *et al.* (J Clin Oncology. 2006, 2563-2569) describen un ensayo de fase III EORTC completado, en el que la metilación de promotor MGMT fue el factor pronóstico más fuerte para el desenlace y respuesta positiva a temozolomida.

- 45 Van den Bent *et al.* (J Clin Oncology. 2009, 27:1268-1274) describen un ensayo de fase II EORTC aleatorio reciente. Los pacientes con GBM progresiva tras asignarse aleatoriamente radioterapia anterior a o bien erlotinib o bien un brazo de control que recibió tratamiento con o bien temozolomida o bien carmustina (BCNU). El principal punto final fue supervivencia libre de progresión a los 6 meses (PFS). Se investigaron muestras de tumor obtenidas en la primera cirugía para la expresión de EGFR; mutantes de EGFRvIII; amplificación de EGFR; mutaciones de EGFR en exones 18, 19, y 21; y pAkt. No se identificó un marcador biológico claro asociado con desenlace mejorado a erlotinib.

- 50 Smith *et al.* (J. National Cancer Institute. 2001, 1246-1256) describen métodos para predecir la supervivencia de

pacientes con astrocitoma anaplásico y GBM. Estos autores identifican que la mutación de PTEN y amplificación de EGFR son marcadores de pronóstico independientes para pacientes con astrocitoma anaplásico y la amplificación de EGFR es un marcador de supervivencia para pacientes de mayor edad con GBM.

5 Umesh *et al.* (Clinical Neuropathology. Vol. 28 - No. 5/2009 (362-372)) describen un método para predecir el desenlace del paciente de glioma que comprende la detección de amplificación de EGFR y PTEN LOH mediante inmunohistoquímica. La conclusión descrita en dicho documento es que la amplificación de EGFR asociada a LOH del gen de PTEN tiende a una escasa supervivencia.

10 Prados *et al.* (J Clin Oncology. 2009, 27(4):579-84) describen un ensayo de fase II que evaluó erlotinib más temozolomida durante y después de la radiación, los pacientes tratados con terapia de combinación (es decir, erlotinib y temozolomida) tuvieron mejor supervivencia. Además, el estudio también evaluó varios marcadores biológicos y hallaron que la metilación del promotor MGMT junto con expresión de PTEN se asociaba a una supervivencia mejorada.

15 Sin embargo, existe todavía la necesidad de marcadores y combinaciones de marcadores adicionales útiles para predecir el desenlace clínico de los pacientes con glioma. Un área especial para el diagnóstico y pronóstico es el estudio de la muestra de biopsia. Un enfoque integrado que puede definir mejor el desenlace del paciente basándose en el tratamiento más apropiado usando perfiles tumorales será crítico, ofreciendo además al paciente con glioma una mejor calidad de vida.

Sumario de la invención

20 En un aspecto, la invención se refiere a un método para predecir el desenlace clínico de un sujeto que padece glioma que comprende determinar el nivel de expresión del gen de EGFR o los niveles de polisomía/amplificación del locus de EGFR en el cromosoma 7 y la pérdida de los niveles de heterocigosidad (LOH) del gen de PTEN en una muestra del mismo sujeto, y comparar dicho nivel de expresión o los niveles de polisomía/amplificación del gen de EGFR y el nivel LOH del gen de PTEN con valores de referencia convencionales, en el que el nivel LOH del gen de PTEN se mide mediante PCR, mediante un ensayo basado en hibridación, mediante secuenciación o mediante
25 análisis de SNP; y en el que un alto nivel LOH del gen de PTEN con respecto a dicho valor de referencia convencional y un alto nivel de expresión y/o altos niveles de polisomía/amplificación del gen de EGFR con respecto a dichos valores de referencia convencionales son indicativos de un buen desenlace clínico del sujeto.

30 En otro aspecto se da conocer también un método para predecir el desenlace clínico de un sujeto que padece glioma que comprende determinar el nivel LOH del gen de PTEN en una muestra del sujeto, y comparar dicho nivel LOH de gen de PTEN con un valor de referencia convencional, en el que el nivel LOH del gen de PTEN se mide mediante PCR, mediante un ensayo basado en hibridación, mediante secuenciación o mediante análisis de SNP; y en el que un alto nivel LOH del gen de PTEN con respecto a dicho valor de referencia convencional, es indicativo de un mal desenlace clínico del sujeto.

35 En otro aspecto, la invención se refiere al uso de un kit que comprende agentes que pueden detectar específicamente el nivel de expresión y/o la polisomía/amplificación del gen de EGFR y el LOH del gen de PTEN y, opcionalmente, un reactivo para detectar un gen constitutivo o la proteína codificada por dicho gen constitutivo y/o un reactivo para detectar los cromosomas 7 y 10, para predecir el desenlace clínico de un sujeto que padece glioblastoma multiforme, en el que el conjunto de agentes que pueden determinar específicamente el nivel LOH del gen de PTEN comprende un par de cebadores de oligonucleótidos adecuados para amplificar un fragmento
40 específico del gen de PTEN o una sonda de oligonucleótido marcada opcionalmente que se une selectivamente a una secuencia diana de polinucleótido sobre la región del cromosoma del gen de PTEN o reactivos adecuados para realizar una reacción de secuenciación o reactivos para realizar un análisis de SNP, y en el que si dichos agentes detectan altos niveles de expresión o un alto nivel de polisomía/amplificación del gen de EGFR y un alto nivel LOH del gen de PTEN, con respecto a valores de referencia, entonces el desenlace clínico del sujeto es bueno.

45 En otro aspecto, se da conocer también el uso de erlotinib y/o temozolomida en la fabricación de un medicamento para el tratamiento de un glioma en un sujeto que padece un glioma, en el que el medicamento es para un sujeto que tiene un alto nivel LOH del gen de PTEN, que se mide mediante PCR, mediante un ensayo basado en hibridación, mediante secuenciación o mediante un análisis de SNP, con respecto a un valor de referencia convencional y altos niveles de expresión y/o alta polisomía/amplificación del gen de EGFR con respecto a valores
50 de referencia convencionales.

55 Se da a conocer también el uso de radioterapia en un régimen para el tratamiento de un glioma en un sujeto que padece un glioma, en el que dicho sujeto tiene un alto nivel LOH del gen de PTEN, que se mide mediante PCR, mediante un ensayo basado en hibridación, mediante una tecnología de secuenciación, o mediante un análisis de SNP, con respecto a un valor de referencia convencional y altos niveles de expresión y/o alta polisomía/amplificación del gen de EGFR con respecto a valores de referencia convencionales.

Breve descripción de los dibujos

Figura 1. Curva de supervivencia de Kaplan-Meier que ilustra el tiempo desde la primera cirugía hasta la muerte /

final del seguimiento de solo pacientes GMB (puro), pacientes con astrocitoma y pacientes con oligodendrogliomas.

Figura 2. Curva de Kaplan-Meier que estima la supervivencia desde la primera cirugía hasta la muerte / recaída / progresión o final del seguimiento de solo pacientes GMB (puro), pacientes con astrocitoma y pacientes con oligodendrogliomas.

- 5 Figura 3. Curva de supervivencia de Kaplan-Meier que demuestra la estratificación de pacientes en el glioblastoma multiforme (GBM), astrocitoma anaplásico y grupo mixto con PTEN LOH ($p=0,04$).

Figura 4. Curva de supervivencia de Kaplan-Meier que demuestra la estratificación de pacientes en el glioblastoma multiforme (GBM), astrocitoma anaplásico y grupo mixto como que ambos tiene EGFR AMP/HP (amplificación - alta polisomía) y PTEN LOH ($p=0,034$).

- 10 Figura 5. Imágenes representativas de muestras de glioblastoma multiforme (GBM) teñidas con H&E (Hematoxilina y Eosina) (A). Ejemplos de experimentos FISH donde se ilustran AMP/HP EGFR DNA FISH (B) y monosomía PTEN (C).

Descripción detallada de la invención

- 15 Con el fin de facilitar el entendimiento de la invención descrito en esta solicitud de patente, se explican a continuación el significado de algunos términos y expresiones en el contexto de la invención.

El término "sujeto" se refiere a un miembro de una especie animal mamífera, e incluye, pero no está limitado al mismo, animales domésticos, primates y seres humanos; el sujeto es preferiblemente un ser humano, macho o hembra, de cualquier edad o raza. Alternativamente, el término "individuo" también se usa algunas veces en esta descripción para referirse a seres humanos.

- 20 El término "proteína" se refiere a una cadena molecular de aminoácidos, unida mediante enlaces covalentes o no covalentes. El término incluye todas las formas de modificaciones postraduccionales, por ejemplo, glicosilación, fosforilación o acetilación.

- 25 El término "anticuerpo" se refiere a una proteína con la capacidad de unirse específicamente a un antígeno. El término anticuerpo comprende anticuerpos recombinantes, anticuerpos monoclonales, o anticuerpos policlonales, intactos, o fragmentos de los mismos que mantienen la capacidad de unirse al antígeno, combicuerpos, etc., tanto de origen humano o humanizado como de origen no humano.

El término "oligonucleótido cebador", tal como se usa en la presente invención, se refiere a una secuencia de nucleótidos, que es complementaria a una secuencia de nucleótidos de un gen seleccionado. Cada oligonucleótido cebador hibrida con su secuencia diana de nucleótidos y actúa como un punto de inicio para polimerización de ADN.

- 30 Los inventores de la presente invención han hallado que el desenlace clínico de pacientes que padecen cáncer de glioma se correlaciona con niveles de expresión y/o los niveles de polisomía/amplificación del gen de EGFR y con el nivel LOH del gen de PTEN.

Por tanto, en un aspecto, la invención se refiere a un método para predecir el desenlace clínico de un sujeto que padece glioma, a continuación en el presente documento se denomina método [1] de la invención, que comprende:

- 35 a) determinar el nivel de expresión o el nivel de polisomía/amplificación del gen de EGFR y el nivel LOH del gen de PTEN en una muestra del mismo sujeto, y

- 40 b) comparar dicho nivel de expresión o el nivel de polisomía/amplificación del gen de EGFR y el nivel LOH del gen de PTEN con valores de referencia convencionales, en el que el nivel LOH del gen de PTEN se mide mediante PCR, mediante un ensayo basado en hibridación, mediante secuenciación o mediante un análisis de SNP; y en el que un alto nivel LOH del gen de PTEN con respecto a dicho valor de referencia convencional y un alto nivel de expresión y/o alto nivel de polisomía / amplificación del gen de EGFR con respecto a dichos valores de referencia convencionales son indicativos de un buen desenlace clínico del sujeto.

- 45 En la presente invención un "glioma" es un tipo de cáncer que comienza en el cerebro o columna vertebral. Se denomina glioma porque surge de células gliales y/o sus precursores. El sitio más común de gliomas es el cerebro. Los gliomas se clasifican por tipo de célula, grado, y localización. Los gliomas se nombran según el tipo específico de célula a la que más se asemejan. Los tipos principales de gliomas son:

- Ependimomas, gliomas derivados de células ependimarias.
- Astrocitomas, gliomas derivados de astrocitos; el glioblastoma multiforme (GBM) es el astrocitoma más común.
- Oligodendrogliomas, gliomas derivados de oligodendrocitos.
- 50 - Gliomas mixtos, tales como oligoastrocitomas, que contienen células de diferentes tipos de glía.

Los gliomas se clasifican además según su grado, que se determina mediante evaluación patológica del tumor. Por tanto, se puede distinguir entre (i) gliomas de bajo grado que están bien diferenciados (no anaplásicos), benignos y presagian un mejor pronóstico para el paciente; y (ii) gliomas de alto grado, que no están diferenciados o anaplásicos, malignos y llevan a un peor pronóstico.

- 5 En una realización preferida, el glioma es un glioblastoma multiforme (GBM) y más preferiblemente el glioblastoma es un glioblastoma temprano.

El glioblastoma multiforme (GBM) es la forma más común y maligna de tumores gliales y se compone de una mezcla heterogénea de malignos astrocitos y células endoteliales displásicas diferenciados escasamente. Afecta principalmente a adultos, implica los hemisferios cerebrales y tiene un desarrollo de la enfermedad rápido lo que a menudo conduce hasta la muerte.

- 10 En la presente invención "desenlace clínico" se entiende como el desarrollo esperado de una enfermedad. Denota el pronóstico del doctor de cómo una enfermedad del sujeto progresará, y si hay posibilidades de recuperación o recaída. El pronóstico del desenlace clínico puede realizarse usando cualquier medición de punto final en oncología y que conoce el médico. Los parámetros de punto final útiles para describir la progresión de una enfermedad incluyen:

- progresión libre de enfermedad que, tal como se usa en el presente documento, describe la proporción de pacientes en de pacientes en remisión que no han tenido recaída de la enfermedad durante el periodo de tiempo en estudio;

- 20 - respuesta objetiva, que, tal como se usa en el presente documento, describe la proporción de personas tratadas en las cuales se observa una respuesta completa o parcial;

- control de tumor, que, tal como se usa en el presente documento, se refiere a la proporción de personas tratadas en las que se observa una respuesta completa, respuesta parcial, respuesta menor o enfermedad estable en igual a o más de ($\geq$) 6 meses;

- 25 - supervivencia libre de progresión que, tal como se usa en el presente documento, se define como el tiempo desde el inicio de tratamiento hasta la primera medición de crecimiento de cáncer;

- tasa de supervivencia libre de progresión a los 6 meses o "PFS6" que, tal como se usa en el presente documento, se refiere al porcentaje de personas en el que están libres de progresión en los primeros seis meses tras el inicio del terapia; y

- 30 - supervivencia mediana que, tal como se usa en el presente documento, se refiere al tiempo en el que la mitad de los pacientes inscritos en el estudio están todavía vivos.

Un buen desenlace clínico se entiende como una situación en la que al menos el 10%, el 20%, el 30%, el 40%, el 50%, el 60%, el 70%, el 80%, el 90%, el 95% o incluso más de los pacientes tienen un resultado positivo con respecto a los parámetros de punto final descritos anteriormente.

- 35 El término "muestra" tal como se usa en el presente documento, se refiere a cualquier muestra que puede obtenerse del paciente. El método presente puede aplicarse a cualquier tipo de muestra biológica de un paciente, tal como una muestra de biopsia, tejido, célula o líquido (sangre completa, suero, saliva, semen, expectoración, orina, líquido cefalorraquídeo (LCR), lágrimas, moco, sudor, leche, extractos de cerebro y similares). En una realización particular, dicha muestra es una parte o muestra de tejido tumoral del mismo. En una realización más particular, dicha muestra de tejido de tumor es una muestra de tejido cerebral tumoral de un paciente que padece cáncer de cerebro. Dicha muestra puede obtenerse mediante métodos convencionales, por ejemplo, biopsia, usando métodos que conocen bien los expertos en la técnica en la técnica médica relacionada. Los métodos para obtener la muestra de la biopsia incluyen reparto bruto de una masa, o microdissección u otros métodos de separación de células conocidos en la técnica. Las células tumorales pueden obtenerse adicionalmente a partir de una citología de aspiración de aguja fina. Con el fin de simplificar la conservación y manejo de las muestras, estas pueden fijarse en formalina e
- 40 incrustarse en parafina o congelarse en primer lugar y se incrustó entonces en un medio con capacidad de criosolidificación, tal como compuesto OCT, a través de la inmersión en un medio altamente criogénico que permite una rápida congelación. En una realización particular, la muestra es una muestra de tumor que contiene un número sustancial de células tumorales.

- 45 Las muestras pueden obtenerse de sujetos diagnosticados previamente de glioma (pacientes), o de sujetos que no han sido diagnosticados previamente de glioma, o de pacientes diagnosticados de glioma Que se someten a tratamiento, o de pacientes diagnosticados de glioma que se han tratado previamente.

- 50 PTEN es la proteína homóloga de fosfatasa y tensina también conocida como BZS; MHAM; TEP1; MMAC1; PTEN1; 10q23del o MGC11227. PTEN es una proteína, que en seres humanos se codifica mediante el gen de PTEN (RefSeq ID NM_000314 SEQ ID NO: 1, proteína de referencia NP_000305.3 SEQ ID NO: 2) (Steck PA, *et al.* 1997 Nat. Genet. 15 (4): 356-62 2). PTEN actúa como un gen supresor tumoral a través de la acción de su producto de

proteína fosfatasa. Esta fosfatasa está implicada en la regulación del ciclo celular, evitando el crecimiento y división demasiado rápidamente de las células. Las mutaciones de este gen contribuye al desarrollo de determinados cánceres (Chu EC, *et al.* 2004 Med. Sci. Monit. 10 (10): RA235-41 3). Existe como homólogos otras especies, tales como ratones (NM_008960.2, SEQ ID NO: 3), rata (NM_031606.1, SEQ ID NO: 4), perro (NM_001003192, SEQ ID NO: 5), etc.

La proteína codificada por el gen de PTEN es una fosfatidilinositol-3,4,5-trisfosfato-3-fosfatasa. Contiene una tensina como dominio así como un dominio catalítico similar al de las proteínas tirosina fosfatasa de especificidad dual. A diferencia de mayoría de las proteínas tirosina fosfatasa, esta proteína desfosforila preferencialmente sustratos de fosfoinositida. Regula negativamente niveles intracelulares de fosfatidilinositol-3,4,5-trisfosfato en células y funciona como supresor tumoral mediante la ruta de señalización de Akt/PKB de regulación negativamente (Hamada K, *et al* 2005 Genes Dev 19 (17): 2054-65).

El receptor de factor de crecimiento epidérmico (EGFR; ErbB-1; HER1 en seres humanos) es el receptor de superficie celular para miembros de la familia de factor de crecimiento epidérmico (EGF-familia) de ligandos de proteína extracelulares (Herbst RS (2004). Int. J. Radiat. Oncol. Biol. Phys. 59 (2 Suppl): 21-6. 1) El EGFR es un miembro de la familia ErbB de receptores. El EGFR es una proteína que en seres humanos es codificada por diferentes isoformas: EGFR variante de transcripción 1 (NM_005228.3, SEQ ID NO: 6), variante de transcripción 2 (NM_201282.1, SEQ ID NO: 7), variante de transcripción 3 (NM_201283.1, SEQ ID NO: 8) y variante de transcripción 4 (NM_201284.1, SEQ ID NO: 9). Existe como homólogos otras especies, tales como ratón (NM_207655.2, SEQ ID NO: 10 y NM_007912.4, SEQ ID NO: 11), rata (NM_031507.1, SEQ ID NO: 12), perro (XM_533073.2, SEQ ID NO: 13), etc.

El método [1] de la invención comprende determinar el nivel de expresión del gen de EGFR. Como el experto en la técnica comprende, puede determinarse el nivel de expresión de un gen midiendo los niveles de ARNm codificados por dichos genes, midiendo ambos niveles de proteínas codificados por dichos genes y los niveles de variantes de los mismos, mediante el uso de sustitutos (número de copia de ADN) para asociar nivel de gen con ARNm y producto de proteína de dicho gen, etc.

Una variante de una proteína, por ejemplo, EGFR o PTEN, tal como se usa en el presente documento puede ser (i) una proteína en la que uno o más de los residuos de aminoácido se sustituye/n con un residuo de aminoácido conservado o no conservado (preferiblemente un residuo de aminoácido conservado) y tal residuo sustituido de aminoácido puede estar codificado o no por el código genético, (ii) una proteína que tiene uno o más residuos de aminoácido modificados, por ejemplo, residuos que se modifican mediante la unión de grupos sustituyentes, (iii) una proteína modificada siendo dicha proteína el resultado de un empalme alternativo del ARNm que codifica el EGFR o proteína PTEN, y/o (iv) un fragmento de la proteína. El término "fragmento" incluye también un péptido o proteína generada por medio de por medio de escisión proteolítica (incluyendo proteólisis multietapa) de una proteína original. Las variantes se consideran que están dentro del alcance de los expertos en la técnica de las enseñanzas en el presente documento.

Como se conoce en la técnica la "similitud" entre dos proteínas se determina comparando la secuencia de aminoácidos y sus sustitutos de aminoácidos conservados de una proteína con respecto a una secuencia de una segunda proteína. Las variantes según la presente invención incluyen péptidos o proteínas que tienen secuencias de aminoácidos que son al menos el 60%, el 65%, el 70%, el 72%, el 74%, el 76%, el 78%, el 80%, el 82%, el 84%, el 86%, el 88%, el 90%, el 95%, o incluso más, similar o idéntica a la secuencia de aminoácidos original. El grado de identidad entre dos proteínas puede determinarse usando algoritmos de ordenador y métodos que los expertos en la técnica conocen ampliamente. La identidad entre las dos secuencias de aminoácidos se determina preferiblemente usando el algoritmo BLASTP (BLASTManual, Altschul, S., *et al.*, NCBI NLM NIH Bethesda, Md. 20894, Altschul, S., *et al.*, J. Mol. Biol. 215: 403-410 (1990)).

Las proteínas pueden modificarse postraduccionalmente. Por ejemplo, modificaciones postraduccionales que se encuentran dentro del alcance de la presente invención incluyen escisión del péptido señal, glicosilación, acetilación, isoprenilación, proteólisis, miristoilación, plegamiento de proteínas y procesamiento proteolítico, etc. Adicionalmente, las proteínas pueden incluir aminoácidos no naturales formados mediante modificación postraducciona o introduciendo aminoácidos no naturales durante traducción.

En una realización preferida, la determinación de los niveles de expresión del gen de EGFR puede llevarse a cabo midiendo el nivel de expresión del ARNm codificado por el gen de EGFR, respectivamente. Para este fin, la muestra biológica puede tratarse para alterar físicamente o mecánicamente la estructura celular o tejido, para liberar componentes intracelulares en una disolución acuosa u orgánica para preparar ácidos nucleicos para análisis adicional. Los ácidos nucleicos se extraen de la muestra mediante procedimientos que conocen los expertos de la técnica y disponibles comercialmente. Entonces se extrae ARN de muestras congeladas o frescas mediante cualquiera de los métodos conocidos en la técnica, por ejemplo, Sambrook, Fischer y Maniatis, Molecular Cloning, un manual de laboratorio, (2ª ed.), Cold Spring Harbor Laboratory Press, Nueva York, (1989). Preferiblemente, se tiene cuidado para evitar la degradación del ARN durante el procedimiento de extracción.

En una realización particular, el nivel de expresión se determina usando ARNm obtenido de una muestra de tejido

fijado con parafina, incrustado en parafina. El ARNm puede aislarse de una muestra patológica o muestra de biopsia de archivo que se desparafina en primer lugar. Un método de desparafinación a modo de ejemplo implica lavado de la muestra parafinizada con un disolvente orgánico, tal como xileno, por ejemplo.

Las muestras desparafinizadas pueden rehidratarse con una disolución acuosa de un alcohol inferior. Los alcoholes inferiores adecuados, por ejemplo incluyen, metanol, etanol, propanoles y butanoles. Muestras desparafinizadas puede rehidratarse con lavados sucesivos con disoluciones de alcoholes inferiores de disminución de la concentración, por ejemplo. Alternativamente, la muestra se desparafina y se rehidrata simultáneamente. La muestra se lisa entonces y se extrae ARN de la muestra.

Mientras que todas las técnicas de elaboración de perfiles de expresión génica (RT-PCR, SAGE, o TaqMan) son adecuadas para su uso en la realización de los aspectos anteriores de la invención, los niveles de expresión del ARNm que codifica para EGFR o para PTEN se determinan a menudo mediante reacción de cadena de polimerasa de transcripción inversa (RT-PCR). La detección puede llevarse a cabo en muestras individuales o en microalineamientos de tejido.

Con el fin de normalizar los valores de la expresión de ARNm entre las diferentes muestras, es posible comparar los niveles de expresión del ARNm de interés en las muestras de prueba con la expresión de un ARN de control. Un "ARN de control" tal como se usa en el presente documento, se refiere a un ARN cuyos niveles de expresión no cambian o cambian solo en cantidades limitadas en células tumorales con respecto a células no tumorigénicas. Preferiblemente, un ARN de control es un ARNm derivado a partir de un gen constitutivo y que codifica para una proteína que se expresa constitutivamente y lleva a cabo funciones celulares esenciales. Genes constitutivos preferidos para su uso en la presente invención incluyen β -2-microglobulina, ubiquitina, proteína ribosomal 18-S, ciclofilina, GAPDH y actina. En una realización preferida, el ARN de control es ARNm de β -actina. En una realización, la cuantificación de la expresión génica relativa se calcula según el método Ct comparativo usando β -actina como un control endógeno y controles de ARN comercial como calibradores. Los resultados finales se determinan según la fórmula $2^{-(\Delta Ct \text{ de muestra} - \Delta Ct \text{ de calibrador})}$, en el que los valores de ΔCt del calibrador y muestra se determinan restando el valor CT del gen diana del valor del gen de β -actina.

La determinación del nivel de expresión del gen de EGFR necesita correlacionarse con los valores de referencia convencionales. Los valores de referencia convencionales corresponden al valor de la mediana de los niveles de expresión del gen de EGFR medidos en una colección de muestras de pacientes sanos. Una vez que se establece este valor de la mediana, el nivel de este marcador expresado en tejidos tumorales de pacientes puede compararse con este valor de la mediana, y por tanto asignarse a un nivel de "bajo", "normal" o "alto". La colección de muestras de cuyo nivel de referencia se deriva se constituirá preferiblemente de personas sanas de la misma edad como los pacientes. En cualquier caso puede contener un número diferente de muestras. En una realización más preferida, las muestras son muestras de cerebro de biopsia. Preferiblemente la colección debe ser suficiente para proporcionar un valor de referencia convencional preciso. Normalmente, el número de muestras usado para determinar un valor de referencia convencional es al menos 10, preferiblemente más de 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500 o incluso más muestras. El patrón puede incluir también células 'normales' presentes dentro del tejido enfermo y/o canceroso. Esto es particularmente cierto para extirpación de tumor cerebral y ocasionalmente muestras de biopsia que normalmente contienen un borde de tejido normal o tienen células inflamatorias, etc., que no contienen cambios de genes. Además, puede usarse un sistema de cultivo celular para evaluar contenido génico, como una manera de determinar la presencia o pérdida de genes individuales.

En una realización particular, un aumento en la expresión del gen de EGFR, como se determinó en la muestra anterior el valor de referencia convencional de al menos 1,1 veces, 1,5 veces, 5 veces, 10 veces, 20 veces, 30 veces, 40 veces, 50 veces, 60 veces, 70 veces, 80 veces, 90 veces, 100 veces o incluso más en comparación con el valor de referencia se considera como un nivel de expresión "alto" del gen de EGFR. En otra realización particular, una disminución en la expresión del gen de EGFR, como se determinará en la muestra siguiente el valor de referencia convencional de al menos 0,9 veces, 0,75 veces, 0,2 veces, 0,1 veces, 0,05 veces, 0,025 veces, 0,02 veces, 0,01 veces, 0,005 veces o incluso menor en comparación con el valor de referencia convencional se considera como un nivel de expresión "bajo" del gen de EGFR.

Alternativamente, en otra realización particular, el nivel de expresión del gen de EGFR puede determinarse midiendo tanto el nivel de la proteína codificada por dichos genes, es decir proteína EGFR, como los niveles de una variante de la misma. Aunque sería posible determinar el nivel de expresión del gen de PTEN midiendo tanto el nivel de la proteína codificada por dichos genes, es decir proteína PTEN, y los niveles de una variante de la misma, en la práctica, dicha opción no es adecuada para la realización de enseñanzas de la presente invención tal como tratará a continuación. El experto en la técnica conoce que la pérdida de expresión de proteína también puede ser el resultado de metilación del promotor de gen y no una pérdida del gen por sí mismo que afecta a otros genes asociados.

La determinación del nivel de expresión de una proteína, por ejemplo, proteína EGFR, o una variante de la misma puede llevarse a cabo mediante cualquier técnica convencional que conoce el experto en la técnica. En una realización particular, la determinación de los niveles de expresión de dicha proteína, por ejemplo, proteína EGFR, o una variante de la misma, se lleva a cabo mediante técnicas inmunológicas tales como por ejemplo, ELISA (Ensayo

por inmuoabsorción ligado a enzimas), inmunotransferencia de tipo Western, inmunofluorescencia (SI), análisis de inmunohistoquímica (IHC), etc. ELISA se basa en el uso de antígenos o anticuerpos marcados (por ejemplo, con enzimas) de modo que los conjugados formados entre el antígeno diana y el anticuerpo marcado da como resultado la formación de, por ejemplo, complejos enzimáticamente activos. Puesto que uno de los componentes (el antígeno o el anticuerpo marcado) se inmoviliza sobre un soporte, los complejos anticuerpo-antígeno se inmovilizan sobre el soporte y por tanto, puede detectarse mediante la adición de un sustrato que se convierte mediante la enzima en un producto que puede detectarse mediante, por ejemplo espectrofotometría, fluorometría, etc. Esta técnica no permite la localización exacta la proteína diana o la determinación de su peso molecular sin embargo permite una detección muy específica y altamente sensible de la proteína diana en una variedad de muestras biológicas (suero, plasma, homogeneizados de tejido, sobrenadantes posnucleares, ascitos y similares). La inmunotransferencia de tipo Western se basa en la detección de una proteína que se vuelve a disolver previamente mediante electroforesis en gel en condiciones desnaturalizantes e inmovilizada sobre una membrana, generalmente nitrocelulosa, mediante incubación con un anticuerpo específico para dicha proteína y un sistema de desarrollo (por ejemplo quimioluminiscencia, etc.). El análisis mediante inmunofluorescencia (SI) requiere el uso de un anticuerpo específico para la proteína diana para el análisis de la expresión y localización subcelular mediante microscopio. Generalmente, las células en estudio se fijan previamente con paraformaldehído y permeabilizan con un detergente no iónico. En una realización preferida, la proteína EGFR se detecta mediante un análisis inmunohistoquímica (IHC) usando secciones delgadas de muestra biológica inmovilizada sobre portaobjetos recubiertos. Las secciones se desparafinan entonces, si se derivaron de muestra de tejido parafinado, y se tratan de modo que se recupere el antígeno. La detección puede llevarse a cabo en muestras individuales o en microalineamientos de tejido. En otra realización el método usado para determinar el nivel de expresión de una proteína, por ejemplo, proteína EGFR, o una variante de la misma es un alineamiento proteómico.

Puede usarse prácticamente cualquier anticuerpo o reactivo conocido por unirse con alta afinidad a la proteína diana para detectar la cantidad de dicha proteína diana. Se prefiere no obstante el uso de un anticuerpo, por ejemplo sueros policlonales, sobrenadantes de hibridoma o anticuerpos monoclonales, fragmentos de anticuerpo, Fv, Fab, Fab' y F(ab')₂, ScFv, dicuerpos, tricuerpos, tetracuerpos y anticuerpos humanizados.

En aún otra realización, la determinación del nivel de expresión de una proteína, por ejemplo, EGFR, etc., puede llevarse a cabo construyendo una microalineación de tejido (TMA) que contiene las muestras de paciente ensambladas, y determinar los niveles de expresión de dicha proteína mediante técnicas IHC. La intensidad de inmunotinción puede evaluarse mediante dos bioquímicos clínicos diferentes y puntuarse usando criterios de punto de corte uniformes y claros, con el fin de mantener la reproducibilidad del método. Las discrepancias pueden resolverse mediante reevaluación simultánea. Brevemente, puede registrarse el resultado de inmunotinción como expresión negativa (0) frente a expresión positiva, como expresión baja (1+) expresión frente a moderada (2+) y como expresión alta (3+), teniendo en cuenta la expresión en células tumorales y el punto de corte específico para cada marcador. Como criterio general, se seleccionaron los puntos de corte con el fin de facilitar la reproducibilidad, y cuando sea posible, traducir acontecimientos biológicos.

La determinación del nivel de expresión del gen de EGFR necesita correlacionarse con los valores de referencia convencionales que corresponden al valor de la mediana de niveles de expresión del gen de EGFR medidos en una colección de muestras de tejido de cerebro de pacientes sanos. Preferiblemente la colección debe ser suficiente para proporcionar un nivel de referencia preciso. Normalmente el número de muestras usado para determinar los valores de referencia convencionales es al menos 10 muestras, preferiblemente más de 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, o incluso más.

En una realización preferida la muestra es una biopsia. Una vez que se establece este valor de la mediana, el nivel de este marcador expresado en tejidos tumorales de pacientes puede compararse con este valor de la mediana, y por tanto asignarse a nivel de "bajo", "normal" o "alto" tal como se definió anteriormente.

"Polisomía" / "amplificación" del locus de EGFR en el cromosoma 7 tal como se usa en la presente invención se entiende como la presencia de más de una copia del locus de EGFR en el cromosoma 7.

Además, tal como se usa en el presente documento, el término "LOH" dentro del contexto de la presente invención se refiere una "pérdida de heterocigosidad". LOH en una célula representa la pérdida de la función normal de un alelo de un gen en el que ya se inactivó el otro alelo. Este término se usa principalmente en el contexto de la oncogénesis; tras una mutación de inactivación en un alelo de un gen supresor tumoral que se produce en la célula germinal original, se transmite al cigoto que da como resultado un descendiente que es heterocigoto para este alelo. En oncología, la pérdida de heterocigosidad (LOH) se produce cuando el alelo funcional restante en una célula somática del descendiente llega a inactivarse mediante mutación. En el caso de PTEN LOH, esto da como resultado que no se produzca un supresor tumoral normal.

En una realización particular, pueden medirse la determinación del polisomía/amplificación nivel del gen de EGFR y la determinación de los niveles LOH del gen de PTEN, por ejemplo, en el ADN obtenido de las células tumorales según procedimientos convencionales tales como, por ejemplo, PCR cuantitativa, hibridación genómica comparativa (CGH) para tecnologías de microalineamientos, etc.; o en las células tumorales a partir la sección incrustada en parafina o de la preparación de citología mediante FISH usando sondas moleculares apropiadas, etc.

En una realización particular, la detección del polisomía/amplificación nivel del gen de EGFR y la determinación de los niveles LOH del gen de PTEN se lleva a cabo mediante a reacción en cadena de polimerasa (PCR).

En una realización particular, la detección del polisomía/amplificación nivel del gen de EGFR y la determinación de los niveles LOH del gen de PTEN se lleva a cabo mediante un ensayo basado en hibridación. En una realización particular, la etapa de detección del método [1] de la invención comprende poner en contacto la muestra de ácido nucleico con una o más sondas de ácido nucleico cada una de ellas se une selectivamente a una secuencia diana de polinucleótido sobre la región del cromosoma del loci de EGFR o PTEN, en condiciones en las que la sonda forma una complejo de hibridación estable con la secuencia diana de polinucleótido; y detectar el complejo de hibridación. En una realización particular, las sondas de ácido nucleico usadas en el método [1] de la invención se marcan con un fluoróforo. Alternativamente, en otra realización particular, la etapa de detección del complejo de hibridación comprende determinar el número de copia de la secuencia diana de polinucleótido, determinando de este modo la presencia de polisomía o LOH del gen diana.

En una realización preferida, dicho ensayo basado en hibridación se selecciona del grupo que consiste en transferencia Southern, hibridación *in situ* (ISH), fluorescencia ISH (FISH) e hibridación genómica comparativa (CGH). La transferencia Southern es un método usado rutinariamente en biología molecular par la detección de una secuencia de ADN específica en muestras de ADN; la transferencia Southern generalmente combina la transferencia de fragmentos de ADN separados por electroforesis a una membrana de filtración y detección de fragmento posterior mediante hibridación de sonda. La hibridación *in situ* (ISH) es un tipo de hibridación que usa una hebra de ADN o ARN complementaria marcada (es decir, sonda) para localizar una secuencia de ADN o ARN específica en una porción o sección de tejido (*in situ*), o, si el tejido es suficientemente pequeño, en el tejido completo; puede usarse ISH de ADN para determinar la estructura de los cromosomas. La *hibridación in situ* (FISH) con fluorescencia es una técnica citogenética usada para detectar y localizar la presencia o ausencia de secuencias de ADN específico en los cromosomas. FISH usa sondas fluorescentes que unen a solo algunas partes del cromosoma con los cuales muestran un alto grado de similitud de secuencia. La microscopía fluorescente puede usarse para averiguar donde se unió la sonda fluorescente a los cromosomas. FISH se usa a menudo para hallar características específicas en ADN para su uso en asesoramiento genético, medicina (por ejemplo, en diagnóstico médico para evaluar la integridad cromosómica), y identificación de especies; FISH también puede usarse para detectar y localizar ARNm específicos y otros transcritos en las secciones de tejido o cantidades completas (así, puede ayudar a definir los patrones espacio-temporales de expresión génica en células y tejidos). La hibridación genómica comparativa (CGH) o análisis de microalineamientos cromosómicos (CMA) es un método molecular-citogenético para el análisis de cambios de número de copias (ganancias/pérdidas) en el Contenido de ADN de un ADN de sujeto dado y a menudo en células tumorales; CGH detecta solo cambios cromosómicos desequilibrados. En una realización particularmente preferida, dicho ensayo basado en hibridación es un ensayo CGH.

En una realización particular, dicho ensayo basado en hibridación es un ensayo basado en alineamiento. En una realización particular, una vez que la muestra se ha obtenido y el ADN total se ha extraído, se lleva a cabo análisis de todo el genoma de cambios de número de copia de ADN mediante CGH. En general, para una medición de CGH típica, se aísla ADN genómico total de las poblaciones celulares de referencia y prueba, se marcan diferencialmente y se hibridaron a una representación del genoma que permite la unión de secuencias en diferentes localizaciones genómicas para que se distingan. Las reacciones de hibridación pueden realizarse en condiciones de rigurosidad diferentes. La rigurosidad de una reacción de hibridación incluye la dificultad con que dos moléculas de ácido nucleico cualquiera hibridarán entre sí. Preferiblemente, cada polinucleótido de hibridación se hibrida a su polinucleótido correspondiente en condiciones de rigurosidad reducida, más preferiblemente condiciones rigurosas, y lo más preferiblemente condiciones altamente rigurosas.

La cantidad de ADN de muestra es frecuentemente una restricción en medidas de CGH. Los procedimientos de CGH de alineamiento típico usan de desde 300 ng hasta 3 mg de ADN de muestra en la reacción de marcaje, equivalente a aproximadamente de 50.000 a 500.000 células de mamífero. Habitualmente, se emplean protocolos que marcan cebadores aleatorios, que amplifican también el ADN, de modo que se usan varios microgramos (µg) en la hibridación.

Se ha implementado alineamiento CGH usando una amplia variedad de técnicas. En una realización particular, se lleva a cabo alineamiento CGH usando alineamientos de clones genómicos de gran inserción tales como cromosomas artificiales bacterianos (BAC). Los principios y condiciones generales para la detección de ácidos nucleicos, tales como el uso de alineamiento CGH a microalineamientos BAC los conoce también el experto en la técnica. Esta técnica permite examinar el genoma completo para cambios de número de copia de ADN permitiendo por tanto la detección cuantitativa de la variación del número de copia de ADN en genomas tumorales con alta resolución (Pinkel D, *et al.* Nat Genet 1998; 20(2):207-11; Hodgson G, *et al.* Nat Genet 2001;29(4):459-64). Como ejemplo no limitativo ilustrativo, en el alineamiento CGH llevado a cabo por el método [1] de la invención pueden marcarse mediante tumores de prueba y ADN de referencia genómicos mediante cebado aleatorio usando fluoróforos Cy3 y Cy5. Entonces, pueden analizarse las imágenes de los alineamientos usando, por ejemplo, una cámara con dispositivo acoplado de carga (CCD) y software apropiado.

El principal reto técnico de alineamiento CGH es la generación de señales de hibridación que sean suficientemente intensas y específicas de modo que puedan detectarse cambios de número de copias. La intensidad de señal en un

elemento de alineamiento está afectada por varios factores incluyendo la composición de base, la proporción de contenido de secuencia repetitiva, y la cantidad de ADN en el elemento de alineamiento disponible para hibridación.

Los elementos de alineamiento elaborados a partir de clones BAC genómicos proporcionan normalmente señales más intensas que los elementos que emplean secuencias más cortas tales como ADNc, productos de PCR, y oligonucleótidos. Las señales más altas de los elementos de alineamiento más complejos dan como resultado mejor precisión de medición, permitiendo la detección de límites de transición de copia simple incluso en muestras con una alta proporción de células normales.

En otra realización preferida, dicho ensayo basado en hibridación es una hibridación in situ fluorescente (FISH) o FISH más carotipo de espectro (SKY) (Liehr T. et al 2008 Fluorescence In situ Hybridization (FISH) – Application Guide, Springer Berlin Heidelberg).

FISH permite detectar y localizar la presencia o ausencia de secuencias de ADN específico en los cromosomas, por ejemplo, FISH permite localizar la señal a un tipo de célula (tumoral) específica. FISH usa sondas fluorescentes que se unen solo a aquellas partes del cromosoma con las que muestran un alto grado de similitud de secuencia. La microscopía fluorescente puede usarse para averiguar donde se une la sonda fluorescente a los cromosomas.

El término “sonda” tal como se usa en el presente documento se refiere a cualquier ribopolinucleótido o desoxirribosecuencia de polinucleótido que se une específicamente a solo aquellas partes del cromosoma con que muestran un alto grado de similitud de secuencia. La sonda debe ser suficientemente grande para hibridar específicamente con su diana pero no tan larga como para impedir el proceso de hibridación. Existen muchas sondas de FISH diferentes que pueden usarse en la presente invención; los ejemplos ilustrativos, no limitativos de los mismos incluyen cromosomas artificiales bacterianos (BAC), sondas de oligonucleótidos de tipo mosaico (TOP), etc. El diseño de sondas para FISH lo conoce bien el experto en la técnica (por favor véanse Bayani J, Squire JA. Curr Protoc Cell Biol. 2004 Sep; capítulo 22: unidad 22.4; Bayani J, Squire J. Curr Protoc Cell Biol. 2004 Oct; capítulo 22: unidad 22.5; Navin, N. *et al.* Bioinformatics, volumen 22, número 19, 1 de octubre de 2006, págs. 2437-2438(2)) editorial: Oxford University Press). La sonda puede etiquetarse directamente con fluoróforos, con dianas para anticuerpos, con biotina, etc. El etiquetado puede realizarse de varias maneras, tales como mediante traslado de mella, mediante PCR usando nucleótidos etiquetados, etc.

La muestra puede fijarse e incrustarse en parafina, así puede realizarse una etapa adicionalmente de desparafinación.

Para la hibridación, se produce una preparación de cromosoma de interfase o metafase. Los cromosomas se adhieren firmemente a un sustrato, habitualmente, un vidrio. Las secuencias de ADN repetitivas deben bloquearse añadiendo fragmentos cortos de ADN a la muestra. La sonda se aplica entonces a el ADN de cromosoma y se incuban durante aproximadamente 12 horas mientras se hibrida. Las varias etapas de lavado retiran todas las sondas sin hibridar o parcialmente hibridadas. Tras lavados de poshibridación convencionales los portaobjetos se tiñen con la sonda de tinción de ADN tal como DAPI y se monta con un agente de montaje tal como Antifade.

Los resultados se visualizan y se cuantifican entonces usando, por ejemplo, un microscopio que puede excitar el colorante y registrar imágenes. Si la señal fluorescente es débil, la amplificación de la señal puede ser necesaria con el fin de exceder el umbral de detección del microscopio. La fuerza de señal fluorescente depende de muchos factores tales como eficacia de marcado de sonda, el tipo de sonda, y el tipo de colorante. Anticuerpos etiquetados fluorescentemente o estreptavidina se unen a la molécula de colorante. Estos componentes secundarios se seleccionan de modo que tengan una señal fuerte. En una realización preferida, antes de la obtención de imágenes el bioquímico clínico evaluó todos los portaobjetos y se identificaron regiones de interés basándose en criterios histopatológicos y calidad incluyendo, sin excluir otros, contenido tumoral, fijación apropiada, necrosis y vascularidad.

Los experimentos FISH diseñados para detectar o localizar expresión génica en células y tejidos se basan en el uso de un gen indicador, tal como uno que expresa proteína verde fluorescente (GFP) y similares, para proporcionar la señal fluorescente.

En una técnica alternativa para preparaciones de interfase o metafase, FISH de fibra, se adhieren cromosomas de interfase a un portaobjetos de tal modo que se estiran en una línea recta, en vez de enrollarse finamente, como en FISH convencional, o adoptar una conformación aleatoria, como en la interfase FISH. Esto se lleva a cabo aplicando accionamiento mecánico junto con la longitud del portaobjetos, o bien a células que se han fijado al portaobjetos y entonces se lisaron, o bien a una disolución de ADN purificado. La conformación de los cromosomas ampliada permite resolución drásticamente más alta - incluso bajando hasta unas pocas kilobases (kb).

En una realización particular aún más preferida, en paralelo a la detección de la polisomía/amplificación del locus de EGFR y el nivel LOH del gen de PTEN, se usan sondas de control FISH para cromosomas 10 y 7. Las “sondas de control FISH” son sondas que se unen específicamente para los cromosomas individuales, permitiendo así la determinación del número de cromosoma. En una realización preferida, las sondas de control FISH se refieren a secuencias satélite alfa. Secuencias satélite alfa, cuando son altamente repetitivas, son específicas de cada cromosoma individual. Estas secuencias flanquean los centrómeros y pueden presentar una diana medida en

megabases. En una realización preferida, estas sondas de control FISH se marcarían con colores diferentes que las sondas EGFR y PTEN.

En una realización preferida, la sonda FISH para PTEN es una sonda que se hibrida a la región 10q23 en el cromosoma 10 y contiene secuencias que flanquean tanto los extremos 5' y 3' del gen de PTEN; en una realización más preferida la sonda tiene entre 300 y 400 kb. En una realización más preferida, la sonda FISH para PTEN y la sonda de control FISH para el cromosoma 10 son la sonda de color dual LSI PTEN (10q23) / CEP 10 para PTEN (Vysis, Abbot Molecular) (Goberdhan, D., *et al.* Human Molecular Genetics 12 (2) (2003): 239-248; Eng, C., *et al.* Human Mutation 22 (2003): 183-198; Sasaki, H., *et al.* Am. J. de Pathology 159 (1) (2001): 359-367). Se describen otras sondas por Cairns *et al.* (Cairns *et al.* 1997. Cancer Res 57; 4997-5000) y por Hermans *et al.* (Hermans *et al.* 2004 Genes Chrom Cancer, 39; 171-184). El LSI PTEN (10q23) se marca con SpectrumOrange. La sonda SpectrumGreen CEP 10 se hibrida a secuencias satélite alfa específicas para cromosoma 10.

En una realización particular, la sonda FISH para EGFR es una sonda que se hibrida a la región 7p12 del cromosoma 7 y contiene el gen de EGFR completo. En una realización más preferida, la sonda FISH para el gen de EGFR y la sonda de control para el cromosoma 10 son la sonda de color dual LSI EGFR/CEP 7 (Vysis, Abbot Molecular). En una realización particular la sonda EGFR se marca con SpectrumOrange y se cubre una región de aproximadamente 300 kb de la región 7p12 del cromosoma 7 (Bredel, M., *et al.* 1999. Clin Cancer Res 5, 1786-92; Harris, A., *et al.* 1989. J Steroid Biochem 34, 123-31; Kitagawa, Y., *et al.* 1996. Clin Cancer Res 2, 909-14; Neal, D.E., *et al.* 1990. Cancer 65, 1619-25; Osaki, A., *et al.* 1992. Am J de Surg 164, 323-6; Pavelic, K., *et al.* 1993. Anticancer Res 13, 1133-7; Sauter, G., *et al.* 1996. Am J Pathol 148, 1047-53; Torregrosa, D., *et al.* 1997. Clin Chim Acta 262, 99-119). La sonda CEP 7, marcada con SpectrumGreen, se hibrida al ADN satélite alfa localizado en el centrómero del cromosoma 7 (7p11.1-q11.1).

Pueden usarse otros métodos conocidos en la técnica para determinar aberraciones en el número de copias; ejemplos ilustrativos, no limitativos de los mismos incluyen microalineamientos basados en oligonucleótidos (Lucito, *et al.* 2003. Genoma Res. 13:2291-2305; Bignell *et al.* 2004. Genoma Res. 14:287-295; Zhao, *et al.* 2004. Cancer Research, 64(9):3060-71).

En otra realización particular, el nivel de polisomía/amplificación del gen de EGFR y/o el nivel LOH del gen de PTEN se mide usando polimorfismos de longitud de secuencia simple (o microsatélites) (Virmani A.K. *et al.* Genes cromosomas Cancer 1998, 21 (4) 308-319) o los SNP como marcadores genéticos (Lindblad-toh K *et al.* Nat. Biotechnol. 2000, 18(9)1001-1005); en una realización particular, el nivel de polisomía/amplificación del gen de EGFR y/o el nivel LOH del gen de PTEN se mide usando un alineamiento SNP.

En otra realización particular, el nivel de polisomía/amplificación del gen de EGFR y/o el nivel LOH del gen de PTEN se miden mediante secuenciación de polinucleótidos, es decir, un método para determinar el orden de las bases de nucleótidos en una molécula de un polinucleótido. En una realización particular, la secuenciación de polinucleótidos se realiza usando una tecnología de secuenciación profunda. En una realización más particular, la tecnología de secuenciación usa un sistema de pirosecuenciación paralela a gran escala que puede secuenciar más o menos 400-600 megabases de ADN por serie con longitudes leídas de 400-500 pares de bases en un secuenciador adecuado (por ejemplo, secuenciador de genoma FLX con series de reactivos GS FLX Titanium). El sistema se base en fijar fragmentos de ADN nebulizados y ligados a adaptador a perlas de captura de ADN pequeñas en una emulsión de agua en aceite. El ADN fijado a estas perlas se amplifica entonces mediante PCR. Cada perla unida a ADN se coloca en un pocillo en una placa picotituladora, un chip óptico de fibra. También se empaquetan una mezcla de enzimas tales como ADN polimerasa, ATP sulfurilasa y luciferasa en el pocillo. Entonces se coloca la placa picotituladora en el sistema GS FLX para secuenciación. Esta tecnología de secuenciación puede secuenciar cualquier ADN de doble cadena y permitir una variedad de aplicaciones incluyendo secuenciación de genoma completo *de novo*, resecuenciación de genomas completos y regiones diana de ADN, metagenómica y análisis de ARN. En otra realización, la tecnología de secuenciación es la tecnología de secuenciación de amplicón, es decir, una secuenciación ultra profunda diseñada para permitir mutaciones que van a detectarse a niveles extremadamente bajos, y PCR amplifica regiones específicas, diana de ADN. Este método se usa para identificar mutaciones somáticas de baja frecuencia en muestras de cáncer o descubrir de las variantes raras en individuos infectados VIH.

La determinación del nivel del polisomía/amplificación de EGFR y el nivel de LOH de PTEN, necesita correlacionarse con un valor de referencia convencional. Dichos valores de referencia convencionales los genera el experto en la técnica en forma de una tabla que divide al paciente en número ascendente de copias del gen de EGFR o con respecto al nivel de LOH y número ascendente de copias del gen de PTEN.

Para EGFR, pueden generarse los valores de referencia, sin limitación, usando por ejemplo, la clasificación de Capuzzo *et al.* (Cappuzzo *et al.* JNCI 2005;4:643-55). En esta clasificación, se clasifican los pacientes en seis estratos con número ascendente de copias del gen de EGFR por célula según la frecuencia de células tumorales con número específico de copias del gen de EGFR por cromosoma 7. Tal como se mencionó anteriormente, están disponibles comercialmente sondas de control que detectan el cromosoma 7 para obtener un número de razón de copias de gen de EGFR/cromosoma 7. Por ejemplo, puede usarse la sonda SpectrumGreen CEP 7 (Vysis) que se hibrida a secuencias satélite alfa específicas para cromosoma 7.

Un ejemplo no limitativo de una tabla para clasificar un paciente que es responsable de la polisomía del gen de EGFR es:

1) disomía (D): ≤ 2 copias en $> 90\%$ de las células de la muestra;

5 2) baja trisomía (LT): ≤ 2 copias en $\geq 40\%$ de las células de la muestra o 3 copias en el $10\% - 40\%$ de las células de la muestra o ≥ 4 copias en $<10\%$ de las células de la muestra;

3) alta trisomía (HT): ≤ 2 copias en $\geq 40\%$ de las células de la muestra o 3 copias en $\geq 40\%$ de las células de la muestra o ≥ 4 copias en $<10\%$ de las células de la muestra;

4) baja polisomía (LP): ≥ 4 copias en el $10\% - 40\%$ de las células de la muestra; y

10 5) alta polisomía (HP): ≥ 4 copias en $\geq 40\%$ de las células de la muestra o la presencia de amplificación (presencia de agrupaciones de gen de EGFR ajustadas y una razón de gen de EGFR con respecto a cromosoma de ≥ 2 o ≥ 15 copias de EGFR por célula en $\geq 10\%$ de células analizadas de la muestra).

15 Como entenderá el experto en la técnica, el método [1] de la invención puede realizarse usando más de una muestra de un paciente. En tal caso, se considera que existe una "alto nivel de polisomía del gen de EGFR" cuando al menos el 50% , preferiblemente más del 50% , más del 60% , más del 70% , más del 80% , más del 90% , más del 95% , o incluso más, de las muestras del paciente se clasifican como "alta polisomía" (HP) según los criterios de clasificación descritos anteriormente.

20 Para la generación de una tabla para la clasificación de los pacientes que es responsable del nivel del LOH y número de copias del gen de PTEN, el experto en la técnica podría clasificar, por ejemplo, el pacientes en 4 estratos con número ascendente de copias del gen de PTEN por célula según la frecuencia de células tumorales con número específico de copias del gen de PTEN por cromosoma 10. Tal como se mencionó anteriormente, están disponibles comercialmente sondas de control que detectan el cromosoma 10 para obtener un número de razón de copias de gen de PTEN/cromosoma 10. Por ejemplo, puede usarse la sonda SpectrumGreen CEP 10 (Vysis) que se hibrida a secuencias satélite alfa específicas para cromosoma 10.

25 Un ejemplo no limitativo de una tabla para clasificar un paciente que es responsable de la polisomía / nivel de LOH del gen de PTEN es:

1) disomía (D): 2 copias de cada sonda en $>90\%$ de células;

2) LOH: <2 copias de PTEN sonda en $>10\%$ de células (incluye monosomía o disomía de cromosoma 10 con PTEN LOH, siempre en $>10\%$ de células);

30 3) polisomía (P): ≥ 3 copias de cada sonda (PTEN+ CEP 10) en $>10\%$ de células (no discrimina entre polisomía alta y baja, o trisomía de cromosoma 10); y

4) amplificado (AMP): definido por una razón del gen de PTEN con respecto a cromosoma 10 de ≥ 2 por célula en $\geq 10\%$ de células analizadas.

35 Como entenderá un experto en la técnica, el método [1] de la invención puede realizarse usando más de una muestra de un paciente. En tal caso, se considera que existe un "nivel LOH alto del gen de PTEN" cuando al menos el 50% , preferiblemente más del 50% , más del 60% , más del 70% , más del 80% , más del 90% , más del 95% , o incluso más, de las muestras del paciente se clasifican como "LOH" según los criterios de clasificación descritos anteriormente.

40 En la última etapa del método [1] de la invención, un alto nivel LOH del gen de PTEN con respecto a dichos valores de referencia convencionales y altos niveles de expresión y/o alta polisomía/amplificación del gen de EGFR son indicativos de un buen desenlace clínico del sujeto.

45 Como entenderá un experto en la técnica, las muestras pueden también considerarse no apropiadas para incluirse en cualquiera de los estratos de clasificación. Por tanto, en una realización preferida, adicionalmente, se tiñen secciones de muestras de tejido del paciente con colorantes como hematoxilina y eosina (H&E) y las revisan dos o más expertos en la técnica (es decir, bioquímicos clínicos) para evaluar contenido tumoral general, necrosis y calidad general (por ejemplo, efecto de cauterización térmica, fijación con morfología, etc.). En una realización preferida las muestras usadas para determinar los niveles de expresión y/o los niveles de polisomía del gen de EGFR y el nivel LOH del gen de PTEN tienen al menos el 20% , al menos el 30% , al menos el 40% , al menos el 50% , al menos el 60% , al menos el 70% , al menos el 80% , al menos el 90% , al menos el 95% de tumor tejido en la muestra, y tienen una fijación apropiada con morfología aceptable.

50 Las enseñanzas de la presente invención no están de acuerdo con las conclusiones alcanzadas por Umesh *et al.* (Clinical Neuropathology. Vol. 28 - No. 5/2009 (362-372)), quién describe un método para predecir el desenlace del paciente de glioma que comprende la detección de amplificación de EGFR y LOH de PTEN mediante inmunohistoquímica (IHC) en el que se afirma que el gen de EGFR amplificación asociada a LOH del gen de PTEN

tiende una escasa supervivencia.

Efectivamente, como se muestra en el ejemplo, las enseñanzas de la presente invención, que comprende combinar polisomía/amplificación de EGFR (es decir, expresión de EGFR) con LOH del gen de PTEN en el que el nivel LOH del gen de PTEN se mide mediante PCR, un ensayo basado en hibridación, secuenciación o alineamientos SNP [el nivel de polisomía/amplificación del gen de EGFR puede determinarse a nivel de proteína o, alternativamente, a nivel de ácido nucleico] muestran que una amplificación de EGFR asociada a LOH del gen de PTEN tiende a buena supervivencia.

Los ensayos realizados por los inventores han mostrado que, aparentemente, el uso de IHC para determinar el nivel del LOH del gen de PTEN no es adecuado para evaluar de manera precisa su perfil de expresión relevante que es especialmente importante en expresión baja pero sin pérdida de proteína PTEN, los tumores en los que una diferencia crítica de este tipo permitirá la evaluación verdadera de expresión de proteína en esta configuración. Además, pueden atribuirse diferentes resultados en un ensayo IHC a anticuerpos específicos y su sensibilidad para evaluar PTEN; existe bibliografía documentada para respaldar los retos de usar IHC para evaluar PTEN cuando se compara con FISH (por ejemplo, Reid *et al.*, British Journal of Cancer 2010, 102:678-684). Métodos FISH para ADN generalmente tiene impacto sobre la preparación de la muestra analítica previa, por el contrario, tales diferencias en fijación y manejo de muestras tienen impacto sobre el antígeno de PTEN en tejido. Una manera de abordar esto es usar modelado matemático con IF cuantitativa para evaluar expresión verdadera frente a unión no específica y 'ruido' en el sistema.

Por tanto, la invención proporciona un método para predecir de una manera más precisa el desenlace clínico de un sujeto que padece glioma.

Los hallazgos de los inventores permiten la determinación del desenlace clínico de un paciente que padece glioma midiendo el nivel LOH del gen de PTEN. Por tanto, se da conocer también un método para predecir el desenlace clínico de un sujeto que padece glioma, a continuación en el presente documento se denomina el método [2], comprendiendo dicho método:

a) determinar el nivel LOH del gen de PTEN en una muestra del sujeto, y

b) comparar dicho nivel LOH del gen de PTEN con un valor de referencia convencional,

en el que el nivel LOH del gen de PTEN se mide mediante PCR, o mediante un ensayo basado en hibridación, o mediante secuenciación o mediante un análisis de SNP; y

en el que un alto nivel LOH del gen de PTEN con respecto a dicho valor de referencia convencional, es indicativo de un mal desenlace clínico del sujeto.

El término glioma se ha definido previamente. En una realización preferida, el glioma es un glioblastoma multiforme (GMB) y más preferiblemente el glioblastoma es un glioblastoma temprano.

En una realización preferida, el desenlace clínico se mide como supervivencia.

El término "muestra" se ha definido previamente en relación con el método [1] de la invención y puede aplicarse a cualquier tipo de muestra biológica de un paciente, tal como una muestra de biopsia, tejido, célula o fluido (sangre completa, suero, saliva, semen, expectoración, líquido cefalorraquídeo (LCR), orina, lágrimas, moco, sudor, leche, extractos de cerebro y similares). En una realización particular, dicha muestra es una parte o muestra de tejido tumoral del mismo. En una realización más particular, dicha muestra de tejido de tumor es una muestra de tejido cerebral tumoral de un paciente que padece glioma o una muestra de tejido de cerebro incrustado en formalina.

Los métodos para determinar el nivel LOH del gen de PTEN se han descrito anteriormente así como los valores de referencia convencionales usados. En una realización preferida, el nivel LOH del gen de PTEN se determina mediante un ensayo basado en hibridación, por ejemplo, mediante FISH.

La invención también se refiere al uso de un kit útil en la implementación la metodología descrita en el presente documento. Por tanto, en otro aspecto, la invención se refiere al uso de un kit que comprende un conjunto de agentes que pueden determinar específicamente los niveles de expresión y/o la polisomía/amplificación de EGFR y el nivel LOH del gen de PTEN y, opcionalmente, un reactivo para detectar un gen constitutivo o la proteína codificada por dicho gen constitutivo y/o un reactivo para detectar los cromosomas 7 y 10, para predecir el desenlace clínico de un sujeto que padece glioblastoma multiforme, en el que el conjunto de agentes que pueden determinar específicamente el nivel LOH del gen de PTEN comprende un par de cebadores de oligonucleótidos adecuados para amplificar un fragmento específico del gen de PTEN y/o una sonda de oligonucleótido marcada opcionalmente que se une selectivamente a una secuencia diana de polinucleótido sobre la región del cromosoma del gen de PTEN y/o reactivos adecuados para realizar una reacción de secuenciación y/o reactivos para realizar un análisis de SNP.

En una realización particular del uso del kit de la invención, los agentes del kit pueden detectar específicamente los

niveles de ARNm de genes de EGFR y/o de PTEN o los niveles de las proteínas EGFR y/o PTEN, preferiblemente, los niveles de la proteína EGFR.

Los ácidos nucleicos que pueden hibridarse específicamente con los genes EGFR y/o de PTEN pueden ser uno o más pares de cebadores de oligonucleótidos para la amplificación específica de fragmentos de los ARNm (o sus correspondientes ADNc) de dichos genes y/o uno o más sondas para la identificación de uno o más genes seleccionados de dichos genes. Ácidos nucleicos que pueden hibridarse específicamente con los genes de EGFR y/o PTEN pueden ser también sondas para FISH.

Anticuerpos, o un fragmento del mismo, que pueden detectar un antígeno, que pueden unirse específicamente a proteínas EGFR y/o PTEN o a variantes de los mismos son, por ejemplo, anticuerpos monoclonales y policlonales, fragmentos de anticuerpo, Fv, Fab, Fab' y F(ab')₂, ScFv, dicuerpos, tricuerpos, tetracuerpos y anticuerpos humanizados.

Dichos reactivos, específicamente, las sondas y los anticuerpos, pueden fijarse sobre un soporte sólido, tal como una membrana, un plástico o un vidrio, tratado opcionalmente con el fin de facilitar la fijación de dichas sondas o anticuerpos sobre el soporte. Dicho soporte sólido, que comprende, al menos, un conjunto de anticuerpos que puede unirse específicamente a proteínas EGFR y/o PTEN o a variantes de los mismos, y/o sondas específicamente hibridadas con los genes de EGFR y/o PTEN, pueden usarse para la detección de los niveles de expresión por medio de tecnología de alineamiento.

Los kits para el uso de la invención comprenden opcionalmente reactivos adicionales para detectar un polipéptido codificado por un gen constitutivo o el ARNm codificado por dichos genes constitutivos. La disponibilidad de dicho reactivo adicional permite la normalización de mediciones tomadas en diferentes muestras (por ejemplo la muestra de prueba y la muestra de control) para excluir que las diferencias en la expresión de los diferentes marcadores biológicos son debidos a cantidades diferentes de proteína total en la muestra en vez de en diferencias reales en niveles de expresión relativos. Genes constitutivos, tal como se usa en el presente documento, se refieren a genes que codifican proteínas que se expresan constitutivamente y llevan a cabo funciones celulares esenciales. Los genes constitutivos preferidos para su uso en la presente invención incluyen β -2-microglobulina, ubiquitina, proteína ribosomal 18-S, ciclofilina, GAPDH y actina.

En una realización, el kit para su uso de la invención puede contener reactivos adecuados para realizar una reacción de secuenciación y/o reactivos para realizar un alineamiento SNP, por ejemplo, enzimas, nucleótidos, etc.

En otra realización, la invención se refiere al uso de un kit de la invención para predecir el desenlace clínico de un sujeto que padece glioblastoma multiforme, en el que si dichos agentes detectan un alto nivel de expresión y/o alto nivel de polisomía / amplificación del gen de EGFR y un alto nivel LOH del gen de PTEN, con respecto a valores de referencia convencionales, en una muestra a partir de dicho sujeto, entonces el desenlace clínico del sujeto es bueno.

Previamente se han descrito métodos para detectar los niveles de expresión de EGFR y los métodos para determinar la polisomía del gen de EGFR y el nivel LOH del gen de PTEN así como los valores de referencia convencionales (véase, por ejemplo, método [1] de la invención).

En otra realización, se da conocer también el uso de los niveles de expresión EGFR y/o niveles de polisomía/amplificación del gen de EGFR y el nivel LOH del gen de PTEN como marcador predictivo del desenlace clínico de un paciente con glioma. En una realización particular, el glioma es un glioblastoma multiforme (GMB). En otra realización particular, el desenlace clínico se mide como supervivencia.

En otra realización particular, se da a conocer también el uso de erlotinib y/o temozolomida en la fabricación de un medicamento para el tratamiento de un glioma en un sujeto que padece un glioma, en el que el medicamento es para un sujeto que tiene un alto nivel LOH del gen de PTEN, que se mide mediante PCR, mediante un ensayo basado en hibridación, mediante secuenciación o mediante un análisis de SNP, con respecto a un valor de referencia convencional y altos niveles de expresión y/o alta polisomía/amplificación del gen de EGFR con respecto a valores de referencia convencionales. Alternativamente, se da conocer también erlotinib y/o temozolomida para su uso en el tratamiento de un glioma en un sujeto que padece un glioma, en el que el medicamento es para un sujeto que tiene un alto nivel LOH del gen de PTEN, que se mide mediante PCR, mediante un ensayo basado en hibridación, mediante secuenciación o mediante un alineamiento SNP, con respecto a un valor de referencia convencional y altos niveles de expresión y/o alta polisomía/amplificación del gen de EGFR con respecto a valores de referencia convencionales. El glioma puede ser un glioblastoma multiforme (GMB). El desenlace clínico puede medirse como supervivencia.

En otra realización, se da a conocer también el uso de radioterapia en un régimen para el tratamiento de un glioma en un sujeto que padece un glioma, en el que dicho sujeto tiene un alto nivel LOH del gen de PTEN, que se mide mediante PCR, mediante un ensayo basado en hibridación, mediante secuenciación o mediante un análisis de SNP, con respecto a un valor de referencia convencional y altos niveles de expresión y/o alta polisomía/amplificación del gen de EGFR con respecto a valores de referencia convencionales. El glioma puede ser un glioblastoma multiforme (GMB). El desenlace clínico puede medirse como supervivencia.

EJEMPLOS

I. MATERIAL Y MÉTODOS

Pacientes y muestras

5 Se obtuvieron múltiples bloques incrustados en parafina fijados en formalina de cincuenta y seis (56) muestras de tumor cerebral primario del Hospital Universitari de Belvitge. Estos pacientes formaban parte de una cohorte longitudinal y se seleccionaron utilizando criterios de inclusión predeterminados que incluían un diagnóstico primario de uno cualquiera de los siguientes criterios: tumor astrocítico, oligodendrocítico y glial, con seguimiento continuo durante una mediana de 3 años y material tisular disponible. Se extrajeron datos clínicos de la historia clínica del paciente utilizando un conjunto de campos de datos predeterminados clínico-patológicos y de desenlace (véase la 10 tabla 1). El desglose histológico de las muestras de tumor era como sigue:

I. Grupo astrocitoma, total 44 pacientes: pilocítico (1), difuso (13), anaplásicos (16), y glioblastoma multiforme (14);

II. Grupo oligodendroglioma, total 6 pacientes: oligodendroglioma (6); y

III. Mixto, total 4 pacientes: oligoastrocitoma (3), y oligoastrocitomas anaplásicos (1).

15 Se realizó una revisión completa de los archivos de datos clínicos disponibles en todos los pacientes en que estaban disponibles bloques/muestras incrustados en parafina (FFPE) fijados en formalina para análisis adicional.

20 Se obtuvieron secciones de 5 µm de todos los bloques, se tiñeron con hematoxilina y eosina (H&E) y las revisaron dos bioquímicos clínicos para evaluar el contenido tumoral global (≥ 50% tumor en al menos un campo 200X), la necrosis y la calidad global (por ejemplo, efecto de cauterización térmica, fijación apropiada con aceptable morfología, etc). Entonces se obtuvieron secciones posteriores para ensayos FISH de ADN e inmunofluorescencia cuantitativa (SI) tal como se explica resumidamente a continuación. Se mantuvieron todos los bloques para estudios posteriores incluyendo RT-PCR.

Tabla 1. Campos de datos clínicos para la cohorte entera con desglose según grupo de diagnóstico			Glioblastoma N=15	Astrocitoma N=26	Mixto / otros N=5	No disponible N=10
Género del paciente	del	Masculino	8	17	4	5
		Femenino	7	9	1	3
		No disponible	0	0	0	2
Raza del paciente	del	Caucásica	15	25	5	8
		No disponible	0	1	0	2
Convulsiones sintomáticas		Si	4	13	2	5
		No	11	13	2	3
		No disponible	0	0	1	2
Presión intracraneal alta sintomática	alta	Si	7	5	1	1
		No	8	21	3	7
		No disponible	0	0	1	2
Número de lesiones pretratamiento	de en	Imagen 1	12	21	5	7
		Imagen 2	1	1	0	0
		Múltiples	2	3	0	1
		No disponible	0	1	0	2
1 ubicación de lesiones en imagen de pretratamiento	de de	Lóbulo frontal	10	10	3	5
		Lóbulo temporal	3	3	0	1
		Lóbulo parietal	1	6	1	0

	Opérculo	0	1	0	0
	Tronco encefálico	1	2	1	0
	Médula espinal	0	1	0	0
	Nervio óptico	0	0	0	1
	Cisura interhemisférica	0	1	0	0
	Cerebelo	0	0	0	1
		0	1	0	0
	No disponible	0	1	0	2
	Lóbulo temporal	1	1	0	1
	Lóbulo parietal	3	1	0	0
	Lóbulo occipital	2	1	0	1
2 ubicación de lesiones en imagen de pretratamiento	Cuerpo caloso	0	1	0	0
	Ganglios basales	2	0	0	0
	Tronco encefálico	0	1	0	0
	Tálamo	0	0	0	1
	No disponible	7	21	5	7
1 lado de las lesiones en imagen de pretratamiento	Derecho	9	12	2	6
	Izquierdo	4	11	3	2
	Desconocido	0	1	0	0
	No disponible	2	2	0	2
2 lado de las lesiones en imagen de pretratamiento	Izquierdo	1	1	0	1
	No disponible	14	25	5	9
	Biopsia	1	6	0	4
Tipo de muestra	Biopsia estereotáxica	2	4	0	1
	Resección	12	16	5	3
	No disponible	0	0	0	2
Tipo de eliminación post-quirúrgica	Completa	4	11	1	6
	Incompleta	10	12	4	2
	No disponible	1	3	0	2
	Si	7	12	3	1
Radioterapia	No	8	12	2	7
	No disponible	0	2	0	2
Tipo de	Focal	5	8	0	1

ES 2 631 455 T3

radiación	No disponible	10	18	5	9
Dosis total de radiación (**) (datos disponibles: n=20)		60,0 (60,0-60,0)	60,0 (40,0-60,0)	51,0 (48,0-54,0)	60,0 (60,0-60,0)
	Si	2	9	1	1
Quimioterapia	No	13	15	4	7
	No disponible	0	2	0	2
	BCNU	0	7	0	0
Tipo de primera quimioterapia	TMZ	2	1	0	0
	PCV	0	1	0	1
	No disponible	13	17	5	9
	Si	4	13	4	6
Recaída o progresión	No	9	10	1	1
	No disponible	2	3	0	3
	Focal	3	8	4	6
Tipo de primera recaída	Difusa	1	1	0	0
	Múltiple	0	2	0	0
	No disponible	11	15	1	4
	Si	3	10	4	5
Tratamiento de primera recaída	No	1	1	0	1
	No disponible	11	15	1	4
	Radiación	0	2	1	1
	Quimioterapia	2	3	1	0
1 tipo de tratamiento de primera recaída	Cirugía	1	4	2	4
	Radiocirugía	0	1	0	0
	No disponible	12	16	1	5
	Radiación	0	2	0	4
2 tipo de tratamiento de primera recaída	Quimioterapia	0	1	0	0
	No disponible	15	23	5	6
3 tipo de tratamiento de primera recaída	Quimioterapia	0	1	0	1
	No disponible	15	25	5	9
Valor de pretratamiento Karnovsky (**) (datos disponibles: n=49)		70,0(40,0-100)	90,0 (20,0-100)	95,0 (90,0-100)	100(60,0 100)
Valor de pretratamiento		60,0(30,0-90,0)	80,0(10,0-100)	95,0 (80,0-100)	100(30,0 100)

Karnovsky (**) (datos disponibles: n=49)

Muerte del paciente	Si	15	15	1	1
	No	0	7	3	5
	No disponible	0	4	1	4

ADN EGFR y PTEN FISH

Usando la sonda de color dual LSI EGFR/CEP 7 (Vysis, Abbot Molecular) para EGFR y la sonda de color dual LSI PTEN (10q23)/CEP 10 para PTEN (Vysis, Abbot Molecular), se realizó ADN FISH incluyendo hibridación, lavado y detección de fluorescencia en todas las muestras en la cohorte según protocolos estándar (Smith *et al.*, 2001. JNCI; 93:1246-1256). Brevemente, se desparafinaron secciones parafina en xilenos, se sometieron a microondas en disolución de citrato de sodio 10 mmol/l (pH 6.0) durante 5-10 minutos, enfriados hasta temperatura ambiente, se enjuagaron, y se trataron entonces con HCl pepsina durante 5 minutos a 37°C antes de enjuagarse y se deshidrataron. Se aplicó la mezcla de sonda precalentada a los portaobjetos y cubreobjetos sellados en el sitio con adhesivo de caucho. Los portaobjetos se desnaturalizaron entonces a 85°C durante de 4 a 6 minutos usando una cámara de hibridación automatizada y entonces se incubaron durante la noche a 37°C. Después de los lavados poshibridación convencionales, se tiñeron los portaobjetos con la tinción de ADN DAPI y se montaron con Antifade (Vectashield). Antes de la obtención de imágenes, un patólogo evaluó todos los portaobjetos H&E y se identificaron regiones de interés basándose en criterios histopatológicos y calidad incluyendo contenido tumoral, fijación apropiada, necrosis y vascularidad. Una imagen H&E de casos de GBM individuales se ilustra en la figura 5A. Un científico formado ciego para el desenlace sometió a obtención de imágenes todos los portaobjetos usando un microscopio inmunofluorescente Nikon. Se definieron criterios para anomalías FISH mediante el uso de muestras de cerebro histológicamente normales. Una ganancia simple requirió el 10% o más de núcleos con tres o más señales de sonda específicas para locus. La pérdida del brazo q del cromosoma 10 requirió que la razón PTEN/CEP10 media global fuese menos de 0.90. Se aplicó amplificación tanto para EGFR como para PTEN y requirió que la razón fuese de 2 o más y que el 10% o más de núcleos tuviera más de tres señales EGFR y/o PTEN. Se adquirieron regiones de interés representativas para documentación de señal en todas las muestras (ejemplos de imágenes representativas de FISH AMP/HP de EGFR y monosomía de PTEN se muestran en las figuras 5B y C, respectivamente). Para EGFR FISH se siguió el esquema de clasificación de Capuzzo *et al.* (Cappuzzo F, *et al.*, JNCI 2005;4: 643-55) y se identificaron 6 categorías para categoría EGFR en todas las muestras de tumor. Estas categorías individuales incluían: D=disomía, HT= alta trisomía, LT= baja trisomía, HP=alta polisomía, LP=baja polisomía, y AMP=amplificación. En este esquema tanto HP como AMP se consideran amplificación. Se caracterizó la categoría PTEN como D=disomía, LOH=>10%, P=polisomía, AMP=amplificado y NE=no evaluable. Un único observador revisó todos los casos y los registró de manera anónima con respecto a tipo de tumor o desenlace.

Inmunofluorescencia cuantitativa (SI)

Utilizando métodos establecidos previamente (Cordon Cardo *et al.*, 2007 JCI 2007; 117; 1876-83; Donovan *et al.*, 2008 JCO;26:3923-3929) se construyó el ensayo multiplex-1 (mplex1) que incorpora anticuerpos a: GFAP (proteína ácida fibrilar glial), PTEN (fosfatasa y homólogo de tensina), pAKT y Ki67 combinados con la tinción nuclear DAPI (4',6-diamidin-2-fenilindol) - véase la tabla 2 para una revisión completa de anticuerpos usados dentro del ensayo mplex1. Se evaluaron los anticuerpos individuales en ensayos IF simplex usando tejidos de control y líneas celulares para confirmar la sensibilidad a la expresión y la especificidad antes de la inclusión en el ensayo multiplex-1.

Tabla 2. Lista de reactivos

Anticuerpo	Proveedor	n.º catálogo	de	Dilución	Isotipo	Marcadores	Muestras
				1:30	M IgG1	2X M 488	Casos GBM, LNCaP cerebral NL
PTEN	Neomarkers						Tonsil
Ki67	Dako			1:100	M IgG1	2X M 555	Igual
pAKT	CST	n.º 3787		1:25	R IgG	2X R 594	Igual
GFAP	Dako			1:200	M IgG1	2X M 647	Igual

Brevemente, se desparafinizaron las muestras FFPE, se rehidrataron y se sometieron a un procedimiento de

recuperación de antígeno con el sistema de tampón de recuperación (Biocare Medical). Se combinó una serie de cuatro anticuerpos con DAPI en un ensayo de inmunofluorescencia multiplex. Se marcaron diferencialmente los reactivos enumerados en la tabla 2 y con los fluorocromos identificados y entonces se aplicaron de manera secuencial a secciones individuales usando un dispositivo de inmunotinción Nemesis 7200 (BioCare Medical). Un patólogo (ciego para el desenlace) adquirió un mínimo de tres imágenes o regiones de interés (ROI) usando un microscopio inmunofluorescente Nikon 90i equipado con un sistema de obtención de imágenes espectrales CRI (CRI). Utilizando bibliotecas sin mezclar desarrolladas previamente y el software CRI, se generaron imágenes en escala de grises individuales que representan el anticuerpo - combinación de filtro en investigación. Se evaluaron las imágenes individuales para determinar la razón señal:ruido, distribución celular y distribución conjunta con otros reactivos marcados con fluorocromos-anticuerpos. Usando el manual del software CRI existente, se crearon umbrales para que las combinaciones anticuerpo-fluorocromo individual se maximizaran la señal:ruido y conservaran la distribución. Se generaron métricas cuantitativas usando el software para identificar porcentajes de poblaciones celulares individuales que presentan una señal positiva (basándose en el umbral), normalizada con respecto a la región tumoral en evaluación. Se obtuvieron una serie de líneas celulares cáncer de próstata incluyendo LNCaP, PC3 y DU145 (ATCC), se hicieron crecer hasta confluencia, se cultivaron como aglomerados de células agar, se fijaron e incrustaron en parafina y entonces se colocaron las tres líneas celulares en un alineamiento celular para control de calidad durante tanto el desarrollo del ensayo como la investigación de casos de tumor cerebral.

Resultados y discusión

Pacientes y muestras

Había 54 pacientes con historias clínicas disponibles y muestras de tejido FFPE apropiadas para evaluación. El desglose de pacientes se enumera en la tabla 3 con la mayoría de pacientes en la categoría de diagnóstico de astrocitoma. Las características demográficas individuales incluyendo datos de sexo, localización de masa, tratamiento y desenlace se explica en detalle en la tabla 1.

Tabla 3. Categoría de diagnóstico de los pacientes

Códigos originales	GBM puro	Astrocitoma	Oligodendroglio ma	Mixto / otro
001: astrocitoma grado II	0	10	0	0
002: astrocitoma grado III	0	16	0	0
003: glioblastoma	15	0	0	0
004: oligoastrocitoma grado II	0	0	0	3
005: oligoastrocitoma grado III	0	0	0	2
006: oligodendroglioma grado II	0	0	4	0
007: oligodendroglioma grado III	0	0	1	0
008: astrocitoma pilocítico	0	2	0	0
009: gliomatosis cerebri	0	0	0	1
999: No disponible	0	0	0	2

Con el fin de entender las tendencias de supervivencia general dentro de la cohorte y en particular el grupo de glioblastoma frente a otros, se realizaron análisis de función de supervivencia Kaplan-Meier para estimar la supervivencia entre los tres grupos. La figura 1 es la curva de supervivencia de Kaplan-Meier que demuestra una supervivencia reducida en el grupo glioblastoma frente al astrocitoma anaplásico y categorías mixtas (prueba de log-rango $P=0,001$). La figura 2 es una segunda curva de función de supervivencia Kaplan-Meier que evalúa un punto final más compuesto que incluye supervivencia libre de progresión y global en las tres mismas categorías de diagnóstico (prueba log-rango $P<0,001$). Según la bibliografía, el grupo glioblastoma tiene una supervivencia global y libre de progresión reducida y en comparación con el astrocitoma y grupo oligodendroglioma puro que tiene el mayor tiempo de supervivencia de las tres categorías de diagnóstico. Como resulta evidente a partir de las curvas, parece que un subconjunto del grupo astrocitoma anaplásico se comporta como la categoría de glioblastoma.

EGFR y PTEN FISH

Dada la frecuencia de amplificación de EGFR (aproximadamente el 40%) y el papel notificado que desempeña EGFR en el desarrollo de glioblastoma, el primer estudio ADN FISH realizado fue sobre la caracterización de EGFR. Usando los métodos de puntuación de tanto Cappuzzo *et al.* (Cappuzzo *et al.*, 2005 JNCI; 4:643-55) como Smith *et al.* (Smith *et al.*, 2001 JNCI; 93:1246-1256), se investigaron los perfiles EGFR en todos los casos de tumor cerebral en

el grupo. Dado el tamaño global de la cohorte, se optó por clasificar adicionalmente a los pacientes como positivos si presentaban cualquier polisomía y/o amplificación del locus de EGFR o cromosoma 7. Esto es contraste con algunos estudios notificados que caracterizan independientemente EGFR ADN FISH o bien como polisomía o bien como amplificación. En el estudio, el 71% de los pacientes GBM tenía polisomía/amplificación EGFR en comparación con el 45% en el grupo astrocitoma anaplásico. No se encontró amplificación de EGFR o polisomía asociada a la supervivencia cuando se examinó la cohorte entera o subgrupos individuales (todos los pacientes de la prueba log-rango $P=0,8$; pacientes de solo glioblastoma $P=0,1$). También se realizó FISH con sondas de color dual para PTEN que utilizan un sistema de puntuación similar a EGFR; sin embargo además de disomía y polisomía, se incluyó la pérdida de PTEN como mayor del 10% en núcleos epiteliales tumorales. En los estudios, el 61% de muestras GBM tenía PTEN LOH en comparación con el 25% en el grupo astrocitoma anaplásico. Existe escasa bibliografía sobre la evaluación de PTEN mediante ADN FISH en muestras de tumor cerebral usando la mayoría de los estudios más técnicas moleculares tales como secuenciación, RT-PCR e inmunohistoquímica (IHC). De manera reseñable en la cohorte, PTEN LOH se asoció de manera invariable y estadística a una supervivencia reducida en el glioblastoma multiforme combinado, astrocitoma y grupos de paciente mixtos ($n=25$ sin LOH, supervivencia mediana 913D frente a $n=10$ con LOH, supervivencia mediana 174D; prueba log-rango $p=0,04$) [véase la figura 3].

Ni la polisomía/amplificación EGFR ni la pérdida de PTEN fueron factores pronóstico independientes, significativos para la supervivencia en el glioblastoma de solo el grupo ($P=0,1$); sin embargo, cuando se combinaron, se observó una tendencia a significativo con un aumento en la supervivencia global ($n=6$ con tanto EGFR/PTEN, supervivencia mediana 242D frente a $n=6$ sin, supervivencia mediana 71D; prueba log-rango $P=0,034$) (Figura 4). La hipótesis es que la combinación de PTEN LOH y amplificación de EGFR en glioblastoma puede representar un subconjunto de pacientes con un fenotipo tumoral más susceptible a radiación y/o tratamiento con temozolomida.

Tabla 4. Características clínicas de pacientes con condiciones de filtro previas

Idaithia	gender	age	sympseiz	sympctic	imagelesions	imagelocat_1	imagelocat_2	imageside_1	imageside_2	diagn
L08-00080	femenino	50	no	si	1	lóbulo frontal	NA	izquierda	NA	glioblastoma
L08-00084	femenino	65	no	si	1	lóbulo frontal	NA	derecha	NA	glioblastoma
L08-00085	masculino	71	no	si	1	lóbulo temporal	NA	derecha	NA	glioblastoma
L08-00087	masculino	66	no	no	1	lóbulo frontal	lóbulo parietal	derecha	NA	glioblastoma
L08-00090	femenino	74	no	no	1	lóbulo parietal	lóbulo occipital	derecha	NA	glioblastoma
L08-00091	masculino	68	no	no	1	lóbulo temporal	NA	derecha	NA	glioblastoma
L08-00094	femenino	66	si	no	1	lóbulo frontal	lóbulo temporal	derecha	NA	glioblastoma
L08-00097	masculino	50	si	si	1	lóbulo temporal	NA	derecha	NA	glioblastoma
L08-00098	femenino	49	si	no	1	lóbulo frontal	lóbulo parietal	izquierda	NA	glioblastoma
L08-00106	masculino	68	si	si	múltiples	lóbulo frontal	ganglio basal	NA	NA	glioblastoma
L08-00109	femenino	60	no	no	2	lóbulo frontal	lóbulo occipital	izquierda	NA	glioblastoma
L08-00128	masculino	37	no	si	1	lóbulo frontal	NA	derecha	izquierda	glioblastoma

NA: datos no disponibles

Idaithia	diagnspec	psiremoval	recur.progr	recurrttype	death
L08-00080	resección	completo	si	NA	si
L08-00084	resección	completo	si	focal	si
L08-00085	resección	completo	si	focal	si
L08-00087	resección	incompleto	no	NA	si
L08-00090	biopsia	incompleto	no	NA	si
L08-00091	resección	completo	no	NA	si
L08-00094	resección	NA	no	NA	si
L08-00097	resección	incompleto	no	NA	si
L08-00098	resección	incompleto	si	focal	si
L08-00106	biopsia esteotóxica	incompleto	no	NA	si
L08-00109	biopsia esteotóxica	incompleto	no	NA	si
L08-00128	resección	incompleto	si	difuso	si

NA: datos no disponibles

Información de variables de la tabla 4:

gender = "género de los pacientes"

age = "edad de los pacientes"

sympseiz = "convulsiones sintomáticas"

symptcp = "alta presión intracraneal sintomática"

imagelesions = " número de lesiones en imagen de pretratamiento"

imagelocat_1 = "localización de la 1ª lesión en imagen de pretratamiento"

imagelocat_2 = "localización de la 2ª lesión en imagen de pretratamiento"

5 imageside_1 = "lado de la 1ª lesión en imagen de pretratamiento"

imageside_2 = "lado de la 2ª lesión en imagen de pretratamiento"

diagn = "diagnóstico anatomopatológico"

diagnspec = " tipo de muestra"

psiremoval = "tipo de eliminación posquirúrgica"

10 recur.prog = "recaída o progresión"

recurrtype = "tipo de primera recaída"

death = "muerte del paciente"

15 En la tabla 4, se incluyeron 12 pacientes, 6 con el fenotipo EGFR AMP / PTEN LOH y 6 sin (véase columna 2), EGFR/Categoría PTEN (si / no con respecto a fenotipo). De los 6 pacientes con glioblastoma multiforme (GBM) con este fenotipo, 3 no mostraron evidencias de recaída clínica. Los 3 pacientes que recayeron en el grupo positivo EGFR AMP / PTEN LOH recibieron radiación y uno recibió tanto radiación como el agente quimioterapéutico temozolomida (TMZ). Aunque el número de pacientes es pequeño, los datos sugieren que los pacientes GBM con un tumor que presenta el EGFR AMP / PTEN LOH tienen un mejor desenlace clínico global que los pacientes sin este fenotipo independientemente del tratamiento actual incluyendo radiación y quimioterapia. Se concluye que el
20 fenotipo identificado que implica tanto señalización de EGFR como pérdida de PTEN pueden reflejar una buena actuación global de GBM, independientemente de las modalidades de tratamiento actuales y que las más nuevas terapias de investigación y ensayos clínicos beneficiarían el enfoque en la intersección de las dos rutas para mejorar adicionalmente el desenlace.

LISTA DE SECUENCIAS

25 <110> Europath Biosciences S.L.

<120> ALTERNACIONES GÉNICAS DE EGFR Y PTEN PREDICEN LA SUPERVIVENCIA EN PACIENTES CON TUMORES CEREBRALES

<130> P5001PC00

<150> EP 09382271.6

30 <151> 12-07-2009

<160> 13

<170> PatentIn versión 3.5

<210> 1

<211> 5572

35 <212> ADN

<213> *Homo sapiens*

<400>

cctccccctcg cccggcgcg	tcccgtccgc ctctcgctcg	cctcccgctt cccctcggtc	60
ttccgaggcg cccgggctcc	cggcgcggcg gcggaggggg	cgggcaggcc ggcgggcggt	120
gatgtggcgg gactctttat	gcgctgcggc aggatacgcg	ctcggcgctg ggacgcgact	180
gcgctcagtt ctctcctctc	ggaagctgca gccatgatgg	aagtttgaga gttgagccgc	240
tgtgaggcga ggccgggctc	aggcgaggga gatgagagac	ggcggcgggc ggcggccgga	300
gccccctctca gcgcctgtga	gcagccgcgg gggcagcgcc	ctcggggagc cggccggcct	360
gcggcgggcg cagcggcggc	gtttctcgcc tcctcttcgt	cttttctaac cgtgcagcct	420
cttcctcggc ttctcctgaa	agggaaggtg gaagccgtgg	gctcgggcgg gagccggctg	480
aggcgcgggc gcggcgggcg	cacctccgc tcctggagcg	ggggggagaa ggcggcgcg	540
cggcgggccgc ggcggctgca	gctccaggga gggggtctga	gtcgcctgtc accatctcca	600
gggctgggaa cgccggagag	ttggtctctc cccttctact	gcctccaaca cggcgggcgc	660
ggcgggcgga catccaggga	cccgggccgg ttttaaact	cccgtccgcc gccgccgcac	720
cccccgctggc ccgggctccg	gaggccgccg gcggaggcag	ccgttcggag gattattcgt	780
cttctcccca ttccgctgcc	gccgctgcc ggcctctggc	tgctgaggag aagcaggccc	840
agtcgctgca accatccagc	agccgccgca gcagccatta	cccggctgcg gtccagagcc	900
aagcgggcggc agagcgagg	gcatcagcta ccgccaagtc	cagagccatt tccatcctgc	960
agaagaagcc ccgccaccag	cagcttctgc catctctctc	ctcctttttc ttcagccaca	1020
ggctcccaga catgacagcc	atcatcaaag agatcgttag	cagaaacaaa aggagatatc	1080
aagaggatgg attcgactta	gacttgacct atatttatcc	aaacattatt gctatgggat	1140
ttcctgcaga aagacttgaa	ggcgtataca ggaacaatat	tgatgatgta gtaaggtttt	1200
tggattcaaa gcataaaaac	cattacaaga tatacaatct	ttgtgctgaa agacattatg	1260
acaccgccaa atttaattgc	agagttgcac aatatacttt	tgaagaccat aaccaccac	1320

agctagaact	tatcaaacc	ttttgtgaag	atcttgacca	atggctaagt	gaagatgaca	1380
atcatgttgc	agcaattcac	tgtaaagctg	gaaagggacg	aactggtgta	atgatatgtg	1440
catatattatt	acatcggggc	aaatttttta	aggcacaaga	ggccctagat	ttctatgggg	1500
aagtaaggac	cagagacaaa	aagggagtaa	ctattcccag	tcagaggcgc	tatgtgtatt	1560
attatagcta	cctgtttaaag	aatcatctgg	attatagacc	agtggcactg	ttgtttcaca	1620
agatgatgtt	tgaaactatt	ccaatgttca	gtggcggaac	ttgcaatcct	cagtttgtgg	1680
tctgccagct	aaaggtgaag	atatattcct	ccaattcagg	acccacacga	cggaagaca	1740
agttcatgta	ctttgagttc	cctcagccgt	tacctgtgtg	tggtgatatc	aaagtagagt	1800
tcttccacaa	acagaacaag	atgctaaaaa	aggacaaaat	gtttcacttt	tgggtaaata	1860
cattcttcat	accaggacca	gaggaaacct	cagaaaaagt	agaaaatgga	agtctatgtg	1920
atcaagaaat	cgatagcatt	tgcagtatag	agcgtgcaga	taatgacaag	gaatatctag	1980
tacttacttt	aacaaaaaat	gatcttgaca	aagcaaataa	agacaaagcc	aaccgatact	2040
tttctccaaa	ttttaaggtg	aagctgtact	tcacaaaaac	agtagaggag	ccgtcaaatc	2100
cagaggctag	cagttcaact	tctgtaacac	cagatgttag	tgacaatgaa	cctgatcatt	2160
atagatattc	tgacaccact	gactctgac	cagagaatga	accttttgat	gaagatcagc	2220
atacacaaat	tacaaaagtc	tgaatTTTTT	tttatcaaga	gggataaaac	accatgaaaa	2280
taaacttgaa	taaactgaaa	atggaccttt	TTTTTTTTta	tggcaatagg	acattgtgtc	2340
agattaccag	ttataggaac	aattctcttt	tcctgaccaa	tcttgTTTTa	ccctatacat	2400
ccacaggggt	ttgacacttg	ttgtccagtt	gaaaaaaggt	tgtgtagctg	tgtcatgtat	2460
atacctTTTT	gtgtcaaaag	gacattttaa	attcaattag	gattaataaa	gatggcactt	2520
tcccgtTTTT	ttccagtttt	ataaaaagtg	gagacagact	gatgtgtata	cgtaggaatt	2580
ttttcctttt	gtgttctgtc	accaactgaa	gtggctaaag	agctttgtga	tatactgggt	2640
cacatcctac	ccctttgcac	ttgtggcaac	agataagttt	gcagttggct	aagagagggt	2700
tccgaagggt	tttgctacat	tctaattgcat	gtattcgggt	taggggaatg	gagggaaatgc	2760
tcagaaagga	aataatttta	tgctggactc	tggaccatat	accatctcca	gctatttaca	2820
cacacctttc	tttagcatgc	tacagttatt	aatctggaca	ttcgaggaat	tggccgctgt	2880
cactgcttgt	tgtttgcgca	TTTTTTTTta	aagcatattg	gtgctagaaa	aggcagctaa	2940
aggaagtgaa	tctgtattgg	ggtacaggaa	tgaaccttct	gcaacatctt	aagatccaca	3000
aatgaaggga	tataaaaaata	atgtcatagg	taagaaacac	agcaacaatg	acttaaccat	3060
ataaatgtgg	aggctatcaa	caaagaatgg	gcttgaaaca	ttataaaaaat	tgacaatgat	3120
ttattaaata	tgttttctca	attgtaacga	cttctccatc	tcctgtgtaa	tcaaggccag	3180
tgctaaaatt	cagatgctgt	tagtacctac	atcagtcaac	aacttacact	tattttacta	3240

gttttcaatc ataatacctg ctgtggatgc ttcattgtgt gcctgcaagc ttcttttttc	3300
tcattaaata taaaatattt tgtaatgctg cacagaaatt ttcaatttga gattctacag	3360
taagcgtttt ttttctttga agatttatga tgcacttatt caatagctgt cagccgttcc	3420
acccttttga ccttacacat tctattacaa tgaattttgc agttttgcac attttttaaa	3480
tgtcattaac tgtagggaa ttttacttga atactgaata catataatgt ttatattaaa	3540
aaggacattt gtgttaaaaa ggaaattaga gttgcagtaa actttcaatg ctgcacacaa	3600
aaaaagaca tttgattttt cagtagaaat tgcctacat gtgctttatt gatttgctat	3660
tgaaagaata gggttttttt tttttttttt tttttttttt ttaaattgtgc agtggtgaat	3720
catttcttca tagtgctccc ccgagttggg actagggctt caatttcact tcttaaaaaa	3780
aatcatcata tatttgatat gccagactg catacgattt taagcggagt acaactacta	3840
ttgtaaagct aatgtgaaga tattattaaa aaggtttttt tttccagaaa tttggtgtct	3900
tcaaattata ccttcacctt gacatttgaa tatccagcca ttttgtttct taatggtata	3960
aaattccatt ttcaataact tattggtgct gaaattgttc actagctgtg gtctgaccta	4020
gttaatttac aaatacagat tgaataggac ctactagagc agcatttata gagtttgatg	4080
gcaaatagat taggcagaac ttcactctaaa atattcttag taaataatgt tgacacgttt	4140
tccatacctt gtcagtttca ttcaacaatt tttaaatttt taacaaagct cttaggattt	4200
acacatttat atttaaacat tgatatatag agtattgatt gattgctcat aagttaaatt	4260
ggtaaagtta gagacaacta ttctaaccac tcaccattga aatttatatg ccaccttgctc	4320
tttcataaaa gctgaaaatt gttacctaaa atgaaaatca acttcatgtt ttgaagatag	4380
ttataaatat tgttctttgt tacaatttcg ggcaccgcat attaaaacgt aactttattg	4440
ttccaatatg taacatggag ggccaggtca taaataatga cattataatg ggcttttgca	4500
ctgttattat ttttcctttg gaatgtgaag gtctgaatga gggttttgat tttgaatgtt	4560
tcaatgtttt tgagaagcct tgcttacatt ttatggtgta gtcattggaa atggaaaaat	4620
ggcattatat atattatata tataaatata tattatacat actctcctta ctttatttca	4680
gttaccatcc ccatagaatt tgacaagaat tgctatgact gaaagggttt cgagtcctaa	4740
ttaaaacttt atttatggca gtattcataa ttagcctgaa atgcattctg taggtaatct	4800
ctgagtttct ggaatatttt cttagacttt ttggatgtgc agcagcttac atgtctgaag	4860
ttacttgaag gcatcacttt taagaaagct tacagttggg ccctgtacca tcccaagtcc	4920
tttgtagctc ctcttgaaca tgtttgccat acttttaaaa gggtagttga ataaatagca	4980
tcaccattct ttgtgtggc acaggttata aacttaagtg gagtttaccg gcagcatcaa	5040
atgtttcagc tttaaaaaat aaaagtaggg tacaagttta atgtttagtt ctagaaattt	5100
tgtgcaatat gttcataacg atggctgtgg ttgccacaaa gtgcctcgtt tacctttaaa	5160
tactgttaat gtgtcatgca tgcagatgga aggggtggaa ctgtgcacta aagtgggggc	5220

ES 2 631 455 T3

tttaactgta gtatttggca gagttgcctt ctacctgcca gttcaaaagt tcaacctggt	5280
ttcatataga atatataatac taaaaaattt cagtctgtta aacagcctta ctctgattca	5340
gcctcttcag atactcttgt gctgtgcagc agtggctctg tgtgtaaatg ctatgcactg	5400
aggatacaca aaaataccaa tatgatgtgt acaggataat gcctcatccc aatcagatgt	5460
ccatttgta ttgtgtttgt taacaaccct ttatctctta gtgttataaa ctccacttaa	5520
aactgattaa agtctcattc ttgtcaaaaa aaaaaaaaaa aaaaaaaaaa aa	5572

<210> 2

<211> 403

<212> PRT

5 <213> Homo sapiens

<400> 2

ES 2 631 455 T3

Met	Thr	Ala	Ile	Ile	Lys	Glu	Ile	Val	Ser	Arg	Asn	Lys	Arg	Arg	Tyr	1	5	10	15
Gln	Glu	Asp	Gly	Phe	Asp	Leu	Asp	Leu	Thr	Tyr	Ile	Tyr	Pro	Asn	Ile	20	25	30	
Ile	Ala	Met	Gly	Phe	Pro	Ala	Glu	Arg	Leu	Glu	Gly	Val	Tyr	Arg	Asn	35	40	45	
Asn	Ile	Asp	Asp	Val	Val	Arg	Phe	Leu	Asp	Ser	Lys	His	Lys	Asn	His	50	55	60	
Tyr	Lys	Ile	Tyr	Asn	Leu	Cys	Ala	Glu	Arg	His	Tyr	Asp	Thr	Ala	Lys	65	70	75	80
Phe	Asn	Cys	Arg	Val	Ala	Gln	Tyr	Pro	Phe	Glu	Asp	His	Asn	Pro	Pro	85	90	95	
Gln	Leu	Glu	Leu	Ile	Lys	Pro	Phe	Cys	Glu	Asp	Leu	Asp	Gln	Trp	Leu	100	105	110	
Ser	Glu	Asp	Asp	Asn	His	Val	Ala	Ala	Ile	His	Cys	Lys	Ala	Gly	Lys	115	120	125	
Gly	Arg	Thr	Gly	Val	Met	Ile	Cys	Ala	Tyr	Leu	Leu	His	Arg	Gly	Lys	130	135	140	
Phe	Leu	Lys	Ala	Gln	Glu	Ala	Leu	Asp	Phe	Tyr	Gly	Glu	Val	Arg	Thr	145	150	155	160
Arg	Asp	Lys	Lys	Gly	Val	Thr	Ile	Pro	Ser	Gln	Arg	Arg	Tyr	Val	Tyr	165	170	175	

ES 2 631 455 T3

Tyr Tyr Ser Tyr Leu Leu Lys Asn His Leu Asp Tyr Arg Pro Val Ala
 180 185 190
 Leu Leu Phe His Lys Met Met Phe Glu Thr Ile Pro Met Phe Ser Gly
 195 200 205
 Gly Thr Cys Asn Pro Gln Phe Val Val Cys Gln Leu Lys Val Lys Ile
 210 215 220
 Tyr Ser Ser Asn Ser Gly Pro Thr Arg Arg Glu Asp Lys Phe Met Tyr
 225 230 235 240
 Phe Glu Phe Pro Gln Pro Leu Pro Val Cys Gly Asp Ile Lys Val Glu
 245 250 255
 Phe Phe His Lys Gln Asn Lys Met Leu Lys Lys Asp Lys Met Phe His
 260 265 270
 Phe Trp Val Asn Thr Phe Phe Ile Pro Gly Pro Glu Glu Thr Ser Glu
 275 280 285
 Lys Val Glu Asn Gly Ser Leu Cys Asp Gln Glu Ile Asp Ser Ile Cys
 290 295 300
 Ser Ile Glu Arg Ala Asp Asn Asp Lys Glu Tyr Leu Val Leu Thr Leu
 305 310 315 320
 Thr Lys Asn Asp Leu Asp Lys Ala Asn Lys Asp Lys Ala Asn Arg Tyr
 325 330 335
 Phe Ser Pro Asn Phe Lys Val Lys Leu Tyr Phe Thr Lys Thr Val Glu
 340 345 350
 Glu Pro Ser Asn Pro Glu Ala Ser Ser Ser Thr Ser Val Thr Pro Asp
 355 360 365
 Val Ser Asp Asn Glu Pro Asp His Tyr Arg Tyr Ser Asp Thr Thr Asp
 370 375 380
 Ser Asp Pro Glu Asn Glu Pro Phe Asp Glu Asp Gln His Thr Gln Ile
 385 390 395 400

Thr Lys Val

<210> 3

<211> 8229

<212> ADN

<213> *Mus musculus*

<400> 3

ES 2 631 455 T3

gcggcaggat acgcgcttgg gcgtcgggac ggggctgcgc tcagctctct cctctcgga	60
gctgcagcca tgatggaagt ttgagagttg agccgctgtg aggccaggcc cggcgcaggc	120
gagggagatg agagacggcg gcggccacgg cccagagccc ctctcagcgc ctgtgagcag	180
ccgcgggggc agcgcctcgc gggagccggc cgggcggcgg cggcggcagc ggcggcgggc	240
ctgcctcct cgtcgtctgt tctaaccggg cagcttctga gcagcttcgg agagagacgg	300
tggagaagc cgtgggctcg agcgggagcc ggcgcaggct cggcggctgc acctcccgct	360
cctggagcgg gggggagaag cggcggcggc ggcgcggcgt cgggggaggg ggtcggagtc	420
gcctgtcacc attgccaggg ctgggaacgc cggagagttg ctctctcccc ttctcctgcc	480
tccaacacgg cggcggcggc ggcggcacgt ccagggaccc gggccggtgt taagcctccc	540
gtccgcgcgc gccgcacccc ccctggcccc ggctccggag gccgcggag gaggcagccg	600
ctgcgaggat tatccgtctt ctccccattc cgctgcctcg gctgccaggc ctctggctgc	660
tgaggagaag caggcccagt ctctgcaacc atccagcagc cggcgcagca gccattaccc	720
ggctgcggtc cagggccaag cggcagcaga gcgaggggca tcagcgaccg ccaagtccag	780
agccatttcc atcctgcaga agaagcctcg ccaccagcag cttctgccat ctctctcctc	840
ctttttcttc agccacaggc tcccagacat gacagccatc atcaaagaga tcgttagcag	900
aaacaaaagg agatatcaag aggatggatt cgacttagac ttgacctata tttatccaaa	960
tattattgct atgggatttc ctgcagaaag acttgaaggt gtatacagga acaatattga	1020
tgatgtagta aggtttttgg attcaaagca taaaaacccat tacaagatat acaatctatg	1080
tgctgagaga cattatgaca ccgccaaatt taactgcaga gttgcacagt atccttttga	1140
agaccataac ccaccacagc tagaacttat caaacccctc tgtgaagatc ttgaccaatg	1200
gctaagtga gatgacaatc atgttgacgc aattcactgt aaagctggaa agggacggac	1260
tggtgtaatg atttgtgcat atttattgca tcggggcaaa tttttaaagg cacaagaggc	1320
cctagatttt tatggggaag taaggaccag agacaaaaag ggagtcacaa ttcccagtc	1380
gaggcgctat gtatattatt atagctacct gctaaaaaat cacctggatt acagaccctg	1440
ggcactgctg ttccacaaga tgatgtttga aactattcca atgttcagtg gcggaacttg	1500
caatcctcag tttgtggtct gccagctaaa ggtgaagata tattcctcca attcaggacc	1560
cacgcggcgg gaggacaagt tcatgtactt tgagttccct cagccattgc ctgtgtgtgg	1620
tgatatcaaa gtagagttct tccacaaaca gaacaagatg ctcaaaaagg acaaatgtt	1680
tcacttttgg gtaaatacgt tcttcatacc aggaccagag gaaacctcag aaaaagtgga	1740
aatggaagt ctttgtgac aggaaatcga tagcatttgc agtatagagc gtgcagataa	1800
tgacaaggag tatcttgtac tcacctaac aaaaaacgat cttgacaaag caaacaaga	1860
caaggccaac cgatacttct ctccaaattt taagggtgaaa ctatacttta caaaaacagt	1920

ES 2 631 455 T3

agaggagcca tcaaattccag aggcctagcag ttcaacttct gtgactccag atgttagtga	1980
caatgaacct gatcattata gatattctga caccactgac tctgatccag agaattgaacc	2040
ttttgatgaa gatcagcatt cacaaattac aaaagtctga tttttttttt cttatcaaga	2100
gggataaaat accatgaaaa aaaaaaaact tgaataaaact gaaatggacc tttttttttt	2160
tttttttttt ttaaattggca ataggacatt gtgtcagatt gcagttatag gaacaattct	2220
cttctcctga ccaatcttgt tttaccctat acatccacag ggttttgaca cttgttgtcc	2280
agttaaaaaa aggttgtgta gctgtgtcat gtatatacct ttttgtgtca aaaggacatt	2340
taaaattcaa ttaggataaa taaaagatgg cactttccca ttttattcca gttttataaa	2400
aagtggagac aggcctgatgt gtatacgcag gagtttttcc tttattttct gtcaccagct	2460
gaagtggctg aagagctctg attcccgggt tcacgtccta cccctttgca cttgtggcaa	2520
cagataagtt tgcagttggc taaggaagtt tctgcagggt tttgttagat tctaattgat	2580
gcacttgggt tgggaatgga gggaaatgctc agaaaggaaat gtttctacct gggctctgga	2640
ccatacacca tctccagctc cttagatgca cctttcttta gcatgctcca cttactaatc	2700
tggacatccg agagattggc tgctgtcctg ctgtttgttt gtgcatttta aagagcatat	2760
tggtgctaga caaggcagct agagtgagta tatttgtagt ggggtacagg aatgaacctat	2820
ctacagcatc ttaagaatcc acaaaggaag ggatataaaa aaagtggta tagatagata	2880
aaagacacag cagcaatgac ttaaccatac aaatgtggag gctttcaaca aaggatgggc	2940
tggaaacaga aaatttgaca atgatttatt cagtatgctt tctcagttgt aatgactgct	3000
ccatctccta tgtaatcaag gccagtgcata agagtcagat gctattagtc cctacatcag	3060
tcaacacctt acctttattt ttattaattt tcaatcatat acctactgtg gatgcttcat	3120
gtgctggctg ccagtttgtt tttctcctta aatattttat aattcttcac aggaaatttc	3180
aacttgagat tcaacagtaa gcaggttttg tttttttttt ttcctagaga ttgatgatgc	3240
gcgtcctcag tccagtggct gtcagacgtt cagccccttt gaccttacac attctattac	3300
aatgagtttt gcagttttgc acattttttt taaatgtcat taactgttag ggaattttac	3360
ttgaatactg aatacatata atgtgtatat taaaaaagtc attgtttgtg ttaaaaaaga	3420
aattagagtt gcagtaaatt tacagcactg cacgaataat aaggcattga agtttttcag	3480
tagaaattgt cctacagatg ctttatcgac ttgctattgg aagaatagat cttcttaaat	3540
gtgcagtgtt gagtcacttc gttatagtgg tagagttggg attagggtt caattttact	3600
tcttaaatat cattctatgt ttgatatgcc cagactgcat acaatttaaa gcaagagtac	3660
aactactatc gtaatggtaa tgtgaagatg ctattacaaa ggatctcctc ccaaccctc	3720
gggaatttgg tgtctttcaa attatatctt gaccttgaca tttgaatatc cagccattat	3780
tagatttctt aatggtgtga agtcccattt tcaataactt attggtgctg aaattgttca	3840
ctagctgtgg tctgacctag ttaatttaca agtacagatt gcataggacc cactagagaa	3900

ES 2 631 455 T3

gcatttatag	tttgatggta	agtagattag	gcagaacgcc	atctaaaata	ttcttagaaa	3960
ataatgttga	tgtattttcc	atacctcatc	agtttcactc	aaccaataaa	gtttttaaaa	4020
ttgtaacaaa	gctcttagga	tttacacatt	tatatattaa	cattgatata	tgaatattga	4080
ctgactgttg	ataaagtcag	agacaacttt	tcctgagatc	tcacatgga	aatctgtaca	4140
cccccttgc	tttcctaaaa	gctgaaagtg	gctgactaaa	atgcaaagca	gctgttgatg	4200
ttttgaagat	agtgataaac	actgttcttt	gttagttttg	ggcacagcat	gctaaactat	4260
aacttgtatt	gttccaatat	gtaacacaga	gggccaggtc	atgaataatg	acattacaat	4320
gggctgttgc	actgttaata	tttttccttt	ggaatgtgaa	ggtctgaatg	agggttttga	4380
ttttgaatgt	ttcagtgttt	ttgagaagcc	ttgcttacat	tttatgggtg	agtcattgga	4440
aatggaaaaa	tggcattata	tatatattat	atatatataa	atatatatat	tatacatact	4500
ctccttactt	tatttcagtt	accatcccca	tagaatttga	caagaattgc	tatgactgaa	4560
agggttttga	gtcctaattc	aaactttctt	tatgacagta	ttcacgatta	gcctgaagtg	4620
cattctgtag	gtgatctctc	ccgtgtttct	ggaatgcttt	cttagactct	tggatgtgca	4680
gcagcttatg	tgtctgaaat	gacttgaagg	catcaccttt	aagaaggctt	acagttgggc	4740
cccgtaacac	ccaagtcctc	tgtaattcct	cttgacatt	tttgccataa	ttgtaaaagg	4800
gtagttgaat	taaatagcgt	caccattctt	tgctgtggca	caggttataa	acttaagtgg	4860
agtttaccgg	cagcatcaaa	tgtttcagct	ttaaaaataa	aagtaggtta	caagttacat	4920
gttttagttt	agaaaatttg	tgcaatatgt	tcataacgat	ggctgtgggt	gccacaaagt	4980
gcctcgttta	cctttaaata	ctgttaatgt	gtcgtgcatg	cagacggaag	gggtggatct	5040
gtgcactaaa	cgggggggctt	ttactctagt	attcggcaga	gttgccttct	acctgccagc	5100
tcaaaagttc	gatctgtttt	catatagaat	atatatacta	aaaccatcca	gtctgtaaaa	5160
cagccttacc	ccgattcagc	ctcttcagat	actcttgtgc	tgtgcagcag	tggctctgtg	5220
tgtaaatgct	atgcactgag	gatacacaaa	tatgacgtgt	acaggataat	gcctcatacc	5280
aatcagatgt	ccatttggtta	ctgtgtttgt	taacaaccct	ttatctctta	gtgttataaa	5340
ctccacttaa	aactgattaa	agtctcattc	ttgtcattgt	gtgggtgttt	tattaaatga	5400
gagtatttat	aattcaaatt	gcttaaattc	attaaaatgt	tcagtaatgg	gcagccacat	5460
atgattacaa	agttcctgtg	cattttttota	tttttcccc	tccttgctat	cottccaagc	5520
aaagcatctt	tctgtcatct	tggtagacac	atacctgtct	actcatgggt	aagaagagca	5580
ctttaagcct	tagtcatcac	ttaataagtt	attccaggca	cagtaaaaag	ttcaaggttc	5640
ttggaaaacg	gtgcttattt	ctcttcttat	aagccagatg	tctgaagata	gccctaaccc	5700
caagaacggg	cttgatgtct	caggctctgt	ctgtggcttt	ctgttttttt	taacactgca	5760
gttggccatc	agcacatggg	aggtttcac	gggacttgtc	cagagtagta	ggctcaaata	5820

ES 2 631 455 T3

tactatctcc	tttctaatat	tcttaaaggc	taaggagtcc	tttcaatata	acagtaagat	5880
aacttgtgat	gttttagaag	taagcagacc	attaatgtca	atgtggagtc	ttaatgttac	5940
atgaagttga	tagtttctct	gtgacccatt	taaaaataca	aaccgagtag	catgcaatta	6000
tgtaaagaaa	tatgaagatt	atatgtagtc	acacattttc	tttagaattc	ttagtttggg	6060
gaaaacttga	atataaaggt	attttgattt	atatgacatt	ttgatgatat	ttgaaaaaaa	6120
ggaatttcct	gacattttgc	tttttagatca	tgtcccccat	tgtgctgtaa	tttaagccaa	6180
cttggttcag	tgaatgccat	caccattttc	attgagaatt	taaaactcac	cagtgtttaa	6240
catgcaggct	tctgagggct	cccggagaat	cagaccttaa	gccagttga	tttacttcta	6300
acgtgaaact	tcgagttcct	gtatactttg	ctagataaatt	tgtggtacat	ctaaagctta	6360
gtcttaagt	gcttgtgtgt	ggattttatt	caacattctt	gttgctaggg	tagagagaaa	6420
tgttgctgag	tagaaacaag	agtacccagt	tcaatgtggt	acagagagca	gtccctaaaa	6480
tctgtacaca	gtgtaatgga	ccacttttag	agtcaagagg	ctgatttttc	ctatgaaatt	6540
acattgcaac	aggaagcctt	ctagtatagt	tccttttact	gttagaatat	gtttttatgc	6600
atacgtata	gctgctttcc	catcttccaa	caacaggtat	caggatgtaa	gcaagcttta	6660
aacagtgtga	agatggcagg	atagtgtcat	cggtaacagt	cctctgactc	taaatgtagt	6720
tgctctgtaa	cactttgtga	atataacatc	acaattctca	tgtccttggg	gggggggggc	6780
ataccagta	ttagtatgtt	ttagtgacta	agcaatcatt	tttctgttta	ctcatgtaca	6840
ttttctcttt	aaaactaaaa	cctgtactgt	gtatgtctcc	aaagcctttt	agcttagttt	6900
ttaggaaatg	aacactgaat	ggatcacttt	ttagtgtagc	aggtatggga	tatgtgcatt	6960
atagagagac	cttgtcagct	ctctgggcct	atgtgaatgt	ttattgttgg	tgtgaggatg	7020
gtaggggaat	cagtaaatac	aagttacgtt	ggttttagcag	agcaagctca	gtgtgggtat	7080
ttctctttga	agcgtgggtc	gtgacgcact	gtgagtagag	aatttggtca	ccctttgagt	7140
cctcttgcat	tttgcaaact	tgctcagcaa	atgcgtacct	accttgcccc	ctaggtaaaa	7200
gcaggaaacta	ctactgattt	atctgtcact	cagctgtctt	tatatgtgtg	cttctgtgac	7260
ttgtatcaca	caagaatctt	aaagatttca	caaattgtta	ccttttagct	ctgaatgttg	7320
agtattctgg	tgggctaaca	acaagacaaa	ctcttgacag	tcatttgaga	attttcatga	7380
aacatttagc	tgaaaacatt	ttataattta	tgaaaaaaat	gtgttacctt	aaacttttac	7440
atatgtggga	gacattaact	gccatatttg	agcatactga	attttaaatt	taaaataaag	7500
ctgcatattt	ttaaatgaaa	tgtttaacaa	ggattcatat	tttttgtttt	ttaagattaa	7560
aaataattta	tgtcttctca	tgtggaacct	catctgtcac	aatggttaga	ttatacagaa	7620
tggagcaagg	cttgtagtgg	tttagcttac	agtaaaattc	ttaatgttta	gatgtgttta	7680
cttactggct	gttatgtata	cttttgagat	tttccacctg	ttctgtgtag	ttttctaaat	7740
gatactccta	cttaaaaaca	gcatttttagt	atctattttc	tgtctccatt	aaatggctct	7800

ES 2 631 455 T3

catttttctat tgagtttgga agtgtgcaca ttgtgtgtgt gtgtgtgtgt gtgtgtgtgc	7860
acacgtgtgc gcgcccgtgc gtgtgtctat ttgtggagtt tgtatgggag aattagtttt	7920
gaaagtgcta gaatagagat gaaatttggt tcaagtaaaa ttttcccact gggattttac	7980
agttttattgt aataaaatgt taatttttga tgaccttgaa tattaatgaa tttgttagcc	8040
tcttgatgtg tgcattaatg agatatatca aagttgtata ttaaaccaaa gttggagttg	8100
tggaagtgtt tttatgaagt tccgtttggc taccaatgga cataagacta gaaatacctt	8160
cctgtggaga atatttttcc tttaaacaat taaaaaggtt cattattttt gaaaaaaaaa	8220
aaaaaaaaa	8229

<210> 4

<211> 1212

<212> ADN

5 <213> *Rattus norvegicus*

<400> 4

atgacagcca tcatcaaaga gatcgtagc agaaacaaaa ggagatatca agaggatgga	60
ttcgacttag acttgaccta tatttatcca aatattattg ctatgggatt tcctgcagaa	120
agacttgaag gtgtatacag gaacaatatt gatgatgtag taagggtttt ggattcaaag	180
cataaaaacc attacaagat atacaatcta tgtgctgaga gacattatga caccgccaaa	240
tttaactgca gagttgcaca gtatcctttt gaagaccata acccaccaca gctagaactt	300
atcaaaccct tttgtgaaga tcttgaccaa tggctaagtg aagacgacaa tcatgttgca	360
gcaattcact gtaaagctgg gaaaggacgg actggtgtaa tgatttgtgc atatttattg	420
catcggggca agtttttaaa ggcacaagag gccctggatt tttatgggga agtaaggacc	480
agagataaaa agggagtaac tattcccagt cagaggcgct atgtatatta ttatagctac	540
ctgttaaaga atcacctgga ttacagacca gtggcactgt tgtttcacia gatgatgttt	600
gaaactattc caatgttcag tggcggaact tgcaatcccc agtttgtggt ctgccagcta	660
aaggtgaaga tctactcctc caactcagga cccacgcggc gggaggacaa gctcatgtac	720
tttgagttcc ctcagccatt gcctgtgtgt ggtgacatca aagtagagtt cttccacaaa	780
cagaacaaga tgctcaaaaa ggacaaaatg tttcactttt gggtaaatac gttcttcata	840
ccaggaccag aggaaacctc agaaaaagtg gaaaatggaa gtctttgtga tcaggaaatc	900
gatagcattt gtagtataga gcgtgcggat aatgacaagg agtatcttgt gctcaccctg	960
acaaaaaatg atcttgacaa agcaaacaaa gacaaggcca accgatactt ctctccaaat	1020
tttaaggtga agttatactt cacaaaaaca gtagaggagc catcaaatcc agaggctagc	1080
agttcaactt ctgtgactcc agacgttagt gacaatgaac ctgatcatta tagatattct	1140
gacaccactg actctgatcc agagaatgaa ccttttgatg aagatcagca ttcacaaatt	1200
acaaaagtct ga	1212

<210> 5

<211> 1396

<212> ADN

5 <213> *Canis familiaris*

<400> 5

ES 2 631 455 T3

```

ccgccccgcg ccaggccccg ggccgcctgc agcctgcgga ggaggccgcg ccgccccgcg      60
ctcctgccgt ctctctcctc cttcctctcc agccaccggc tcccagacat gacagccatc      120
atcaaggaga tcgtcagcag aaacaaaagg cgctaccagg aggatggggtt cgacttggac      180
ttgacctata tttatcccaa cattattgct atgggggtttc ctgcagaaag acttgaaggc      240
gtatacagga acaatattga tgatgtagta aggttttttg attcaaagca taaaaaccat      300
tacaagatat acaatctgtg tgctgaaaga cattatgata ccgccaaatt taactgcaga      360
gttgcacagt atccttttga agaccataat ccaccacagc tagaacttat caaacccctt      420
tgtgaagatc ttgaccaatg gctaagtga gatgacaatc atgttgcagc aattcactgt      480
aaagctggaa agggacgaac tgggtgtaat atttgtgcat atttattaca tcgggggcaa      540
tttctaaagg cacaagaggc cctagatttc tatggggaag taaggaccag agacaaaaag      600
ggagtaacta ttcccagtca gaggcgctat gtgtattatt atagctacct gttaaagaat      660
catctggatt atagaccagt ggcaactgtg tttcacaaga tgatgtttga aactattcca      720
atgttcagtg gcggaacttg caatcctcag tttgtggtct gccagctaaa ggtgaagatc      780
tattcctcca attcaggacc cacacgacgg gaagacaagt tcatgtactt tgagttccct      840
cagccattgc ctgtgtgcgg tgacatcaaa gtagagttct tccacaaaca gaacaagatg      900
ctaaaaaagg acaaaatgtt tcacttttgg gtaaacacat tcttcatacc aggaccagag      960
gaaacctcag aaaaagtaga aaatggaagt ctatgtgatc aagaaattga tagtatttgc      1020
agtatagaac gtgcagataa tgacaaggaa tatctagtac tcactttaac aaaaaatgat      1080
ctcgacaaag caaataaaga caaggccaac cgatattttt ctccaaattt taaggtgaag      1140
ctgtacttca caaaaactgt agaggagcca tcaaaccggg aggctagcag ttcaacttct      1200
gtgacgccag atgttagtga caatgaacct gatcattata gatattctga caccactgac      1260
tctgaccagc agaatgaacc ctttgatgaa gatcagcaca cacaatcac aaaagtctga      1320
atTTTTTTT atcaagaggg ataaaacacc atgaaaacaa acttgaataa actgaaattg      1380
gacctTTTTT ttttaa      1396

```

<210> 6

<211> 5616

<212> ADN

5 <213> Homo sapiens

<400> 6

```

ccccggcgca gcgcggccgc agcagcctcc gcccccgca cgggtgtgagc gcccgacgcg      60

```

ES 2 631 455 T3

gcccagggcgg ccggaggtccc gagctagccc cggcgggccgc cgccgcccag accggacgac 120

agggcacctc gtcggcgctcc gcccaggtcc ccgcctcgcc gccaacgccca caaccaccgc 180

gcacggcccc ctgactccgt ccagtattga tcgggagagc cggagcgagc tcttcgggga 240

gcagcgatgc gaccctccgg gacggccggg gcagcgctcc tggcgctgct ggctgcgctc 300

tgcccggcga gtcgggctct ggaggaaaag aaagtttgcc aaggcacgag taacaagctc 360

acgcagttgg gcacttttga agatcatttt ctcagcctcc agaggatgtt caataactgt 420

gaggtggtcc ttgggaattt ggaaattacc tatgtgcaga ggaattatga tctttccttc 480

ttaaagacca tccaggaggt ggctggttat gtcctcattg ccctcaacac agtggagcga 540

attccttttg aaaacctgca gatcatcaga ggaaatatgt actacgaaaa ttcctatgcc 600

ttagcagtct tatctaacta tgatgcaaat aaaaccggac tgaaggagct gcccatgaga 660

aatttacagg aaatcctgca tggcgccgtg cggttcagca acaaccctgc cctgtgcaac 720

gtggagagca tccagtggcg ggacatagtc agcagtgact ttctcagcaa catgtcgatg 780

gacttccaga accacctggg cagctgccaa aagtgtgatc caagctgtcc caatgggagc 840

tgctgggggtg caggagagga gaaactgccag aaactgacca aaatcatctg tgcccagcag 900

tgctccgggc gctgccgtgg caagtcccc agtgactgct gccacaacca gtgtgctgca 960

ggctgcacag gccccggga gagcgactgc ctggtctgcc gcaaattccg agacgaagcc 1020

acgtgcaagg acacctgccc cccactcatg ctctacaacc ccaccacgta ccagatggat 1080

gtgaaccccg agggcaaata cagctttggt gccacctgcg tgaagaagtg tccccgtaat 1140

tatgtggtga cagatcacgg ctcgctcgctc cgagcctgtg gggccgacag ctatgagatg 1200

gaggaagacg gcgtccgcaa gtgtaagaag tgcgaagggc cttgccgcaa agtgtgtaac 1260

ggaataggta ttggtgaatt taaagactca ctctccataa atgctacgaa tattaaacac 1320

ttcaaaaact gcacctccat cagtggcgat ctccacatcc tgccggtggc atttaggggt 1380

gactccttca cacatactcc tcctctggat ccacaggaac tggatattct gaaaaccgta 1440

aaggaaatca cagggttttt gctgattcag gcttggcctg aaaacaggac ggacctccat 1500

gcctttgaga acctagaaat catacgcggc aggaccaagc aacatggtca gttttctctt 1560

gcagtcgtca gcctgaacat aacatccttg ggattacgct ccctcaagga gataagtgat 1620

ggagatgtga taatttcagg aaacaaaaat ttgtgctatg caaatacaat aaactggaaa 1680

aaactgtttg ggacctccgg tcagaaaacc aaaattataa gcaacagagg tgaaaacagc 1740

tgcaaggcca caggccaggt ctgccatgcc ttgtgctccc ccgagggctg ctggggcccg 1800

gagcccaggg actgcgtctc ttgccggaat gtcagccgag gcagggaatg cgtggacaag 1860

tgcaaccttc tggaggggtga gccaaaggag tttgtggaga actctgagtg catacagtgc 1920

caccagagt gcctgcctca ggccatgaac atcacctgca caggacgggg accagacaac 1980

tgtatccagt gtgccacta cattgacggc cccactgcg tcaagacctg cccggcagga 2040

gtcatgggag	aaaacaacac	cctggtctgg	aagtacgcag	acgccggcca	tgtgtgccac	2100
ctgtgccatc	caaactgcac	ctacggatgc	actggggccag	gtcttgaagg	ctgtccaacg	2160
aatgggccta	agatcccgtc	catcgccact	gggatgggtg	gggccctcct	cttgctgctg	2220
gtggtggccc	tggggatcgg	cctcttcatg	cgaaggcgcc	acatcgttcg	gaagcgcacg	2280
ctgcgagggc	tgctgcagga	gagggagctt	gtggagcctc	ttacacccag	tggagaagct	2340
cccaaccaag	ctctcttgag	gatcttgaag	gaaactgaat	tcaaaaagat	caaagtgtg	2400
ggctccggtg	cgttcggcac	ggtgtataag	ggactctgga	tcccagaagg	tgagaaagtt	2460
aaaattcccg	tcgctatcaa	ggaattaaga	gaagcaacat	ctccgaaagc	caacaaggaa	2520
atcctcgatg	aagcctacgt	gatggccagc	gtggacaacc	cccacgtgtg	ccgcctgctg	2580
ggcatctgcc	tcacctccac	cgtgcagctc	atcacgcagc	tcatgccctt	cggtgcctc	2640
ctggactatg	tccgggaaca	caaagacaat	attggctccc	agtacctgct	caactggtgt	2700
gtgcagatcg	caaagggcat	gaactacttg	gaggaccgtc	gcttggtgca	ccgcgacctg	2760
gcagccagga	acgtactggt	gaaaacaccg	cagcatgtca	agatcacaga	ttttgggctg	2820
gccaaactgc	tgggtgcgga	agagaaagaa	taccatgcag	aaggaggcaa	agtgcctatc	2880
aagtggatgg	cattggaatc	aattttacac	agaatctata	cccaccagag	tgatgtctgg	2940
agctacgggg	tgaccgtttg	ggagttgatg	acctttggat	ccaagccata	tgaccggaatc	3000
cctgccagcg	agatctcctc	catcctggag	aaaggagaac	gcctccctca	gccaccata	3060
tgtaccatcg	atgtctacat	gatcatggtc	aagtgtctgga	tgatagacgc	agatagtgcg	3120
ccaaagttcc	gtgagttgat	catcgaattc	tccaaaatgg	cccagagacc	ccagcgctac	3180
cttgtcattc	agggggatga	aagaatgcat	ttgccaaagtc	ctacagactc	caacttctac	3240
cgtgccctga	tggatgaaga	agacatggac	gacgtgggtg	atgccgacga	gtacctcatc	3300
ccacagcagg	gcttcttcag	cagccccctc	acgtcacgga	ctccccctcct	gagctctctg	3360
agtgcaacca	gcaacaattc	caccgtgggt	tgcattgata	gaaatgggct	gcaaagctgt	3420
cccatcaagg	aagacagctt	cttgcagcga	tacagctcag	acccacagag	cgccttgact	3480
gaggacagca	tagacgacac	cttcctccca	gtgcctgaat	acataaacca	gtccgttccc	3540
aaaaggcccg	ctggctctgt	gcagaatcct	gtctatcaca	atcagcctct	gaacccccgcg	3600
cccagcagag	accacacta	ccaggacccc	cacagcactg	cagtgggcaa	ccccgagtat	3660
ctcaacactg	tccagcccac	ctgtgtcaac	agcacattcg	acagccctgc	ccactggggc	3720
cagaaaggca	gccaccaa	tagcctggac	aaccctgact	accagcagga	cttctttccc	3780
aaggaagcca	agccaaatgg	catctttaag	ggctccacag	ctgaaaatgc	agaataccta	3840
agggtcgctg	cacaaagcag	tgaatttatt	ggagcatgac	cacggaggat	agtatgagcc	3900
ctaaaaatcc	agactctttc	gatacccagg	accaagccac	agcaggtcct	ccatcccaac	3960

```

agccatgccc gcattagctc ttagaccac agactggttt tgcaacgttt acaccgacta 4020
gccaggaagt acttccacct cgggcacatt ttgggaagtt gcattccttt gtcttcaaac 4080
tgtgaagcat ttacagaaac gcatccagca agaattattgt ccctttgagc agaaatttat 4140
ctttcaaaga ggtatatattg aaaaaaaaaa aaagtatatg tgaggatttt tattgattgg 4200
ggatcttgga gtttttcatt gtcgctattg atttttactt caatgggctc ttccaacaag 4260
gaagaagctt gctggtagca cttgctaccc tgagttcatc caggcccaac tgtgagcaag 4320
gagcacaagc cacaagtctt ccagaggatg cttgattcca gtggttctgc ttcaaggctt 4380
ccactgcaaa aactaaaga tccaagaagg cttcatggc cccagcaggc cggatcggta 4440
ctgtatcaag tcatggcagg tacagtagga taagccactc tgtcccttcc tgggcaaaga 4500
agaaacggag gggatggaat tcttccttag acttactttt gtaaaaatgt cccacaggta 4560
cttactcccc actgatggac cagtggtttc cagtcatgag cgtagactg acttgtttgt 4620
cttccattcc attgttttga aactcagtat gctgccctg tcttgctgtc atgaaatcag 4680
caagagagga tgacacatca aataataact cggattccag cccacattgg attcatcagc 4740
atgtggacca atagcccaca gctgagaatg tggaatacct aaggatagca ccgcttttgt 4800
tctcgcaaaa acgtatctcc taatttgagg ctcatgaa atgcatcagg tcctttgggg 4860
catagatcag aagactacaa aaatgaagct gctctgaaat ctcttttagc catcacccca 4920
accccccaaa attagtttgt gttacttatg gaagatagtt ttctcctttt acttcaactc 4980
aaaagctttt tactcaaaga gtatatgttc cctccaggtc agctgcccc aaacccctc 5040
cttacgcttt gtcacacaaa aagtgtctct gccttgagtc atctattcaa gcacttacag 5100
ctctggccac aacagggcat ttacaggtg cgaatgacag tagcattatg agtagtgtgg 5160
aattcaggta gtaaatatga aactagggtt tgaaattgat aatgctttca caacatttgc 5220
agatgtttta gaaggaaaaa agttccttcc taaaataatt tctctacaat tggaagattg 5280
gaagattcag ctagttagga gccaccttt ttctctaatac tgtgtgtgcc ctgtaacctg 5340
actggttaac agcagtcctt tgtaaacagt gttttaaact ctctagtca atatccacc 5400
catccaattt atcaaggaag aaatggttca gaaaatattt tcagcctaca gttatgttca 5460
gtcacacaca catacaaaaat gttccttttg cttttaaagt aatttttgac tcccagatca 5520
gtcagagccc ctacagcatt gtaagaaag tatttgattt ttgtctcaat gaaaataaaa 5580
ctatattcat ttccactcta aaaaaaaaaa aaaaaa 5616

```

<210> 7

<211> 2239

<212> ADN

5 <213> *Homo sapiens*

ES 2 631 455 T3

<400> 7

ccccggcgca gcgcggccgc agcagcctcc gccccccgca cgggtgtgagc gcccgacgcg 60

ES 2 631 455 T3

gccgagggcg	ccggagtccc	gagctagccc	cgccggccgc	cgccgccag	accggacgac	120
agggcacctc	gtcggcgctc	gcccgagtc	ccgcctcgcc	gccaacgcc	caaccaccgc	180
gcacggcccc	ctgactccgt	ccagtattga	tcgggagagc	cggagcgagc	tcttcgggga	240
gcagcgatgc	gacctcccg	gacggccggg	gcagcgctcc	tggcgctgct	ggctgcgctc	300
tgcccggcga	gtcgggctct	ggaggaaaag	aaagtttgcc	aaggcacgag	taacaagctc	360
acgcagttgg	gcacttttga	agatcatttt	ctcagcctcc	agaggatgtt	caataactgt	420
gaggtggtcc	ttgggaattt	ggaaattacc	tatgtgcaga	ggaattatga	tctttccttc	480
ttaaagacca	tccaggaggt	ggctggttat	gtcctcattg	ccctcaacac	agtggagcga	540
attccttttg	aaaacctgca	gatcatcaga	ggaaatatgt	actacgaaaa	ttcctatgcc	600
ttagcagtct	tatctaacta	tgatgcaa	aaaaccggac	tgaaggagct	gcccattgaga	660
aatttacagg	aaatcctgca	tggcgccgtg	cggttcagca	acaaccctgc	cctgtgcaac	720
gtggagagca	tccagtggcg	ggacatagtc	agcagtgact	ttctcagcaa	catgtcgatg	780
gacttccaga	accacctggg	cagctgccaa	aagtgtgatc	caagctgtcc	caatgggagc	840
tgctgggggtg	caggagagga	gaactgccag	aaactgacca	aaatcatctg	tgcccagcag	900
tgctccgggc	gctgccgtgg	caagtcccc	agtgactgct	gccacaacca	gtgtgctgca	960
ggctgcacag	gccccggga	gagcgactgc	ctggtctgcc	gcaaattccg	agacgaagcc	1020
acgtgcaagg	acacctgccc	cccactcatg	ctctacaacc	ccaccacgta	ccagatggat	1080
gtgaaccccg	agggcaaata	cagctttggt	gccacctgcg	tgaagaagtg	tccccgta	1140
tatgtggtga	cagatcacgg	ctcgtgcgtc	cgagcctgtg	gggccgacag	ctatgagatg	1200
gaggaagacg	gcgtccgcaa	gtgtaagaag	tgcgaagggc	cttgccgcaa	agtgtgtaac	1260
ggaataggta	ttggtgaatt	taaagactca	ctctccataa	atgctacgaa	tattaaacac	1320
ttcaaaaact	gcacctccat	cagtggcgat	ctccacatcc	tgccggtggc	atttaggggt	1380
gactccttca	cacatactcc	tcctctggat	ccacaggaac	tggatattct	gaaaaccgta	1440
aaggaaatca	cagggttttt	gctgattcag	gcttggcctg	aaaacaggac	ggacctccat	1500
gcctttgaga	acctagaaat	catacgcggc	aggaccaagc	aacatggtca	gttttctctt	1560
gcagtcgtca	gcctgaacat	aacatccttg	ggattacgct	ccctcaagga	gataagtgat	1620
ggagatgtga	taatttcagg	aaacaaaaat	ttgtgctatg	caaatacaat	aaactggaaa	1680
aaactgtttg	ggacctcccg	tcagaaaacc	aaaattataa	gcaacagagg	tgaaaacagc	1740
tgcaaggcca	caggccaggt	ctgccatgcc	ttgtgctccc	ccgagggctg	ctggggcccg	1800
gagcccaggg	actgcgtctc	ttgccggaat	gtcagccgag	gcagggaatg	cgtggacaag	1860
tgcaaccttc	tggaggggtga	gccaaaggag	tttgtggaga	actctgagtg	catacagtgc	1920
caccagagt	gcctgcctca	ggccatgaac	atcacctgca	caggacgggg	accagacaac	1980
tgtatccagt	gtgccacta	cattgacggc	cccactgcg	tcaagacctg	cccggcagga	2040

ES 2 631 455 T3

```
gtcatgggag aaaacaacac cctggtctgg aagtacgcag acgccggcca tgtgtgccac 2100
ctgtgccatc caaactgcac ctacgggtcc taataaatct tcactgtctg actttagtct 2160
cccactaaaa ctgcatttcc tttctacaat ttcaatttct ccctttgctt caaataaagt 2220
cctgacacta ttcatttga 2239
```

<210> 8

<211> 1595

<212> ADN

5 <213> *Homo sapiens*

<400> 8

ES 2 631 455 T3

```

ccccggcgca gcgcggccgc agcagcctcc gccccccgca cgggtgtgagc gcccgacgcg      60
gccgaggcgg ccggagtcgc gagctagccc cggcggccgc cgccgcccag accggacgac      120
aggccacctc gtcggcgtcc gcccgagtcg ccgcctcgcc gccaacgcca caaccaccgc      180
gcacggcccc ctgactccgt ccagtattga tcgggagagc cggagcgagc tcttcgggga      240
gcagcgatgc gaccctccgg gacggccggg gcagcgctcc tggcgctgct ggctgcgctc      300
tgcccggcga gtcgggctct ggaggaaaag aaagtttgcc aaggcacgag taacaagctc      360
acgcagttgg gcacttttga agatcatttt ctcagcctcc agaggatggt caataactgt      420
gaggtggtcc ttgggaattt ggaaattacc tatgtgcaga ggaattatga tctttccttc      480
ttaaagacca tccaggaggt ggctgggtat gtcctcattg ccctcaacac agtggagcga      540
attccttttg aaaacctgca gatcatcaga ggaaatatgt actacgaaaa ttcctatgcc      600
ttagcagtct tatctaacta tgatgcaaat aaaaccggac tgaaggagct gcccatgaga      660
aatttacagg aaatcctgca tggcgccgtg cggttcagca acaaccctgc cctgtgcaac      720
gtggagagca tccagtggcg ggacatagtc agcagtgact ttctcagcaa catgtcgatg      780
gacttccaga accacctggg cagctgcca aagtgtgatc caagctgtcc caatgggagc      840
tgctgggggtg caggagagga gaactgccag aaactgacca aaatcatctg tgcccagcag      900
tgctccgggc gctgccgtgg caagtcccc agtgactgct gccacaacca gtgtgctgca      960
ggctgcacag gcccccgga gagcgactgc ctggtctgcc gcaaattccg agacgaagcc     1020
acgtgcaagg acacctgccc ccactcatg ctctacaacc ccaccacgta ccagatggat     1080
gtgaaccccg agggcaaata cagctttggt gccacctgcg tgaagaagtg tccccgtaat     1140
tatgtggtga cagatcacgg ctcgctgcgtc cgagcctgtg gggccgacag ctatgagatg     1200
gaggaagacg gcgtccgcaa gtgtaagaag tgcgaggggc cttgccgcaa agtgtgtaac     1260
ggaataggta ttggtgaatt taaagactca ctctccataa atgctacgaa tattaaacac     1320
ttcaaaaact gcacctccat cagtggcgat ctccacatcc tgccggtggc atttaggggt     1380
gactccttca cacatactcc tcctctggat ccacaggaac tggatattct gaaaaccgta     1440
aaggaaatca caggtttgag ctgaattatc acatgaatat aaatgggaaa tcagtgtttt     1500
agagagagaa cttttcgaca tatttcctgt tcccttgga taaaaacatt tcttctgaaa     1560
ttttaccgtt aaaaaaaaaa aaaaaaaaaa aaaaaa                                1595

```

<210> 9

<211> 2865

5 <212> ADN

<213> Homo sapiens

ES 2 631 455 T3

<400> 9

ccccggcgca gcgcggccgc agcagcctcc gccccccgca cggtgtgagc gcccgacgcg	60
gccgagggcg ccggagtccc gagctagccc cgggcgccgc cgccgccag accggacgac	120
agggcacctc gtcggcgctcc gcccgagtcc ccgcctcgcc gccaacgcca caaccaccgc	180
gcacggcccc ctgactccgt ccagtattga tcgggagagc cggagcgagc tcttcgggga	240
gcagcgatgc gacctcccg gacggccggg gcagcgctcc tggcgctgct ggctgcgctc	300
tgcccggcga gtcgggctct ggaggaaaag aaagtttgcc aaggcacgag taacaagctc	360
acgcagttgg gcacttttga agatcatttt ctgagcctcc agaggatgtt caataactgt	420
gaggtggtcc ttgggaattt ggaaattacc tatgtgcaga ggaattatga tctttccttc	480
ttaaagacca tccaggaggt ggctggttat gtcctcattg ccctcaacac agtggagcga	540
attccttttg aaaacctgca gatcatcaga ggaaatatgt actacgaaa ttcctatgcc	600
ttagcagctc tatctaacta tgatgcaaat aaaaccggac tgaaggagct gcccatgaga	660
aatttacagg aaatcctgca tggcgccgtg cggttcagca acaaccctgc cctgtgcaac	720
gtggagagca tccagtggcg ggacatagtc agcagtgact ttctcagcaa catgtcgatg	780
gacttccaga accacctggg cagctgcaa aagtgtgatc caagctgtcc caatgggagc	840
tgctgggggtg caggagagga gaactgccag aaactgacca aaatcatctg tgcccagcag	900
tgctccgggc gctgccgtgg caagtcccc agtgactgct gccacaacca gtgtgctgca	960
ggctgcacag gcccccgga gagcgactgc ctggtctgcc gcaaattccg agacgaagcc	1020
acgtgcaagg acacctgccc cccactcatg ctctacaacc ccaccacgta ccagatggat	1080
gtgaaccccc agggcaaata cagctttggt gccacctgcg tgaagaagtg tccccgtaat	1140
tatgtggtga cagatcacgg ctggtgcgtc cgagcctgtg gggccgacag ctatgagatg	1200
gaggaagacg gcgtccgcaa gtgtaagaag tgcgaagggc cttgccgcaa agtgtgtaac	1260
ggaataggta ttggtgaatt taaagactca ctctccataa atgctacgaa tattaacac	1320
ttcaaaaact gcacctccat cagtggcgat ctccacatcc tgccggtggc atttaggggt	1380
gactccttca cacatactcc tcctctggat ccacaggaac tggatattct gaaaaccgta	1440
aaggaaatca cagggttttt gctgattcag gcttggcctg aaaacaggac ggacctccat	1500
gcctttgaga acctagaaat catacgcggc aggaccaagc aacatggtca gttttctctt	1560
gcagtcgtca gcctgaacat aacatccttg ggattacgct ccctcaagga gataagtgat	1620

ES 2 631 455 T3

```

ggagatgtga taatttcagg aaacaaaaat ttgtgctatg caaatacaat aaactggaaa 1680
aaactgtttg ggacctccg tcaaaaaacc aaaattataa gcaacagagg tgaaaacagc 1740
tgcaaggcca caggccaggt ctgccatgcc ttgtgctccc ccgagggctg ctggggcccg 1800
gagcccaggg actgctctc ttgccggaat gtcagccgag gcagggaatg cgtggacaag 1860
tgcaaccttc tggaggggtga gccaaaggag tttgtggaga actctgagtg catacagtgc 1920
caccagagt gcctgcctca ggccatgaac atcacctgca caggacgggg accagacaac 1980
tgtatccagt gtgcccacta cattgacggc cccactgcg tcaagacctg ccgggcagga 2040
gtcatgggag aaaacaacac cctggtctgg aagtagcgag acgccggcca tgtgtgccac 2100
ctgtgccatc caaactgcac ctacgggcca ggaaatgaga gtctcaaagc catgttattc 2160
tgccttttta aactatcatc ctgtaatcaa agtaatgatg gcagcgtgtc ccaccagagc 2220
gggagcccag ctgctcagga gtcatgctta ggatggatcc cttctcttct gcgctcagag 2280
tttcagctgg gttgggggtg atgcagccac ctccatgcct ggccttctgc atctgtgatc 2340
atcacggcct cctcctgcca ctgagcctca tgccttcacg tgtctgttcc ccccgctttt 2400
cctttctgcc acccctgcac gtgggcccgc aggttcccaa gagtatccta cccatttctt 2460
tccttccact ccctttgcca gtgcctctca ccccaactag tagctaacca tcacccccag 2520
gactgacctc ttcctcctcg ctgccagatg attgttcaaa gcacagaatt tgtcagaaac 2580
ctgcagggac tccatgctgc cagccttctc cgtaattagc atggccccag tccatgcttc 2640
tagccttggg tccttctgcc cctctgtttg aaattctaga gccagctgtg ggacaattat 2700
ctgtgtcaaa agccagatgt gaaaacatct caataacaaa ctggctgctt tgttcaatgc 2760
tagaacaacg cctgtcacag agtagaaact caaaaatatt tgctgagtga atgaacaaat 2820
gaataaatgc ataataaata attaaccacc aatccaacat ccaga 2865

```

<210> 10

<211> 5983

<212> ADN

5 <213> *Mus musculus*

<400> 10

ES 2 631 455 T3

ctccccccagt cccgacccga gctaactaga cgtctgggca gcccagcgc aacgcgcagc	60
agcctccctc ctcttcttcc cgcactgtgc gctcctcctg ggctagggcg tctggatcga	120
gtcccggagg ctaccgcctc ccagacagac gacaggtcac ctggacgcga gcctgtgtcc	180
gggtctcgtc gttgccggcg cagtcactgg gcacaaccgt gggactccgt ctgtctcgga	240
ttaatcccg agagccagag ccaacctctc ccggtcagag atgcgaccct cagggaccgc	300
gagaaccaca ctgctggtgt tgctgaccgc gctctgcgcc gcaggtgggg cgttggagga	360
aaagaaagtc tgccaaggca caagtaacag gctcacccaa ctgggcactt ttgaagacca	420
ctttctgagc ctgcagagga tgtacaacaa ctgtgaagtg gtccttggga acttggaaat	480

ES 2 631 455 T3

tacctatgtg caaaggaatt acgacctttc cttcttaaag accatccagg aggtggccgg	540
ctatgtcctc attgccctca acaccgtgga gagaatccct ttggagaacc tgcagatcat	600
caggggaaat gctctttatg aaaacaccta tgccttagcc atcctgtcca actatgggac	660
aaacagaact gggcttaggg aactgcccac gcggaactta caggaaatcc tgattggtgc	720
tgtgcgattc agcaacaacc ccatcctctg caatatggat actatccagt ggagggacat	780
cgtccaaaac gtctttatga gcaacatgtc aatggactta cagagccatc cgagcagttg	840
ccccaaatgt gatccaagct gtcccaatgg aagctgctgg ggaggaggag aggagaactg	900
ccagaaattg accaaaatca tctgtgcccc gcaatgttcc catcgctgtc gtggcaggtc	960
ccccagtgc tgctgccaca accaatgtgc tgcgggggtg acagggcccc gagagagtga	1020
ctgtctggtc tgccaaaagt tccaagatga ggccacatgc aaagacacct gccaccact	1080
catgctgtac aacccaccca cctatcagat ggatgtcaac cctgaaggga agtacagctt	1140
tggtgccacc tgtgtgaaga agtgcccccg aaactacgtg gtgacagatc atggctcatg	1200
tgtccgagcc tgtgggcctg actactacga agtgggaagaa gatggcatcc gcaagtgtaa	1260
aaaatgtgat gggccctgtc gcaaagtttg taatggcata ggcatgtgtg aatttaaaga	1320
cacactctcc ataaatgcta caaacatcaa acacttcaaa tactgcactg ccatcagcgg	1380
ggaccttcac atcctgccag tggcctttta gggggattct ttcacgcgca ctccctctct	1440
agaccacga gaactagaaa ttctaaaaac cgtaaaggaa ataacaggct ttttgctgat	1500
tcaggcttgg cctgataact ggactgacct ccatgctttc gagaacctag aaataatacg	1560
tggcagaaca aagcaacatg gtcagttttc tttggcggtc gttggcctga acatcacatc	1620
actggggctg cgttccctca aggagatcag tgatggggat gtgatcattt ctggaaaccg	1680
aaatttgtgc tacgcaaaca caataaactg gaaaaaactc ttcgggacac ccaatcagaa	1740
aacccaaatc atgaacaaca gagctgagaa agactgcaag gccgtgaacc acgtctgcaa	1800
tcctttatgc tcctcggaag gctgctgggg ccctgagccc agggactgtg tctcctgcca	1860
gaatgtgagc agaggcaggg agtgcggtga gaaatgcaac atcctggagg gggaaaccaag	1920
ggagtttgtg gaaaattctg aatgcatcca gtgccatcca gaatgtctgc ccaggccat	1980
gaacatcacc tgtacaggca ggggaccaga caactgcac cagtgtgccc actacattga	2040
tggcccacac tgtgtcaaga cctgccagc tggcatcatg ggagagaaca aactctggt	2100
ctggaagtat gcagatgcca ataatgtctg ccacctatgc cagccaact gtacctatgg	2160
atgtgctggg ccaggtcttc aaggatgtga agtgtggcca tctgggcca agataccatc	2220
tattgccact gggattgtgg gtggcctcct cttcatagtg gtggtggccc ttgggattgg	2280
cctattcatg cgaagacgtc acattgttgc aaagcgtaca ctacgccgcc tgcttcaaga	2340
gagagagctc gtggaacctc tcacaccag cggagaagct ccaaaccaag ccacttgag	2400

ES 2 631 455 T3

gatattaaag	gaaacagaat	tcaaaaagat	caaagttctg	ggttcgggag	catttggcac	2460
agtgtataag	ggtctctgga	ttccagaagg	tgagaaagta	aaaatcccgg	tggccatcaa	2520
ggagttaaga	gaagccacat	ctccaaaagc	caacaaagaa	atccttgacg	aagcctatgt	2580
gatggctagt	gtggacaacc	ctcatgtatg	ccgcctcctg	ggcatctgtc	tgacctccac	2640
tgtccagctc	attacacagc	tcatgcccta	cggttgcctc	ctggactacg	tccgagaaca	2700
caaggacaac	attggctccc	agtacctcct	caactgggtg	gtgcagattg	caaagggcat	2760
gaactacctg	gaagatcggc	gtttggtgca	ccgtgacttg	gcagccagga	atgtactggt	2820
gaagacacca	cagcatgtca	agatcacaga	ttttgggctg	gccaaactgc	ttggtgctga	2880
agagaaagaa	tatcatgccg	aggggggcaa	agtgccctatc	aagtggatgg	ctttggaatc	2940
aattttacac	cgaatttata	cacaccaaag	tgatgtctgg	agctatgggtg	tcactgtgtg	3000
ggaactgatg	acctttgggt	ccaagcctta	tgatggaatc	ccagcaagtg	acatctcatc	3060
catcctagag	aaaggagagc	gccttccaca	gccacctatc	tgaccatcgc	atgtctacat	3120
gatcatggtc	aagtgtctga	tgatagatgc	tgatagccgc	ccaaagttcc	gagagttgat	3180
tcttgaattc	tccaaaatgg	cccagagacc	acagcgctac	cttggttatcc	agggggatga	3240
aagaatgcat	ttgccaagcc	ctacagactc	caacttttac	cgagccctga	tggatgaaga	3300
ggacatggag	gatgtagttg	atgctgatga	gtatcttata	ccacagcaag	gcttcttcaa	3360
cagcccgtcc	acgtcgagga	ctccccctct	gagttctctg	agtgcacta	gcaacaattc	3420
cactgtggct	tgcattaata	gaaatgggag	ctgccgtgtc	aaagaagacg	ccttcttgca	3480
gcggtacagc	tccgacccca	caggtgctgt	aacagaggac	aacatagatg	acgcattcct	3540
ccctgtacct	gaatatgtaa	accaatctgt	tcccaagagg	ccagcaggct	ctgtgcagaa	3600
ccctgtctat	cacaatcagc	ccctgcatcc	agctcctgga	agagacctgc	attatcaaaa	3660
tccccacagc	aatgcagtgg	gcaaccctga	gtatctcaac	actgccacgc	ctacctgtct	3720
cagtagtggg	tttaacagcc	ctgcactctg	gatccagaaa	ggcagtcacc	aaatgagcct	3780
agacaaccct	gactaccagc	aggacttctt	ccccaaaggaa	accaagccaa	atggcatatt	3840
taagggcccc	acagctgaaa	atgcagagta	cctacgggtg	gcacctcaa	gcagtgagtt	3900
tattggagca	tgacaagaag	gggcatcata	ccagctataa	aatgtctgga	ctttctagaa	3960
tcccaggacc	aactatggca	gcacctccac	ttctggtagc	catgcccacg	ctgtgtcaaa	4020
tgtcactcag	actggcttta	aagcataact	ctgatgggct	ttgtcactga	gccaaagaag	4080
gggcctctct	cctgatgcac	tttgggaagt	tgaaggtaga	tcaattgatc	ttcgaaactgt	4140
gaagattcca	caaaaaaggt	atccatcgag	aacattgtcc	attggaacag	aagtttgcct	4200
catgggtgag	tacatatggg	aaaaaaacag	acatatggag	cttatattta	gggaactttg	4260
ggattcttgt	ctttattgat	ttgattgatg	cactcttgta	gtctggtaga	cagagttgcc	4320
tggagccaac	tgaccagaca	gttggttcca	ccagctctgc	atcaagacac	ttccgtggca	4380

ES 2 631 455 T3

```

agacaactaa atgtataaga agtccatgga tgccttgagc aggccacact tgtacagcat      4440
taaaccatgg cagatacaat aggataagcc actttgttac ttactggggc tgggagaaga      4500
ggaatgacgg ggtagaattt tccctcagac gtacttttta tataaatatg tccctggcac      4560
ctaacacgcg ctagtttacc agtgttttct attagacttc cttctatggt ttctgtttca      4620
ttgttttgag ttgtaaatat gtgttcctgt cttcatttca tgaagtaaac aaacaaacaa      4680
aaaaccaggt attaagtatt atcaaagaac aaccatgatt ccacattcga acccattcaa      4740
accatcagta ttgtgaccaa aagcctttaa ctaagaagga gtaaccatgc aaaaatccat      4800
agaggaattt aacccaaaat tttagtctca gcattgtgtc tgctgaggtg tgtatatgag      4860
actacgaaag tgaactactc ttcaaatcca ctttgccttc actcctctat accctaaatc      4920
tagtgtaaac cacacatgga ggataacttt tttttttaat tttaaaagtg tttattagat      4980
atgtttttct tcctggtaaa ctgcagccaa acatcagtta agagccattt ttgataaaca      5040
ctatcacaat gatctcggga tccatccttt ccgatttacc aagtgatgga tagacgtgaa      5100
ctcataaaca ctaccataa gacaaaacaa tgagtgccag acaagacatc agccaggcac      5160
cagagcacag agcaggactg ggcaatctgt tggagatatc tagaaagttc acaaaggaaa      5220
caagattgtc cactaccttg tgagatctag cagtcataaa taccaggga atggaaagtg      5280
tgtttcctta cagcaccagg tcttcgatct tcctaatagtc gtgacccttt aatacagttt      5340
gccatgttgt ggtgaccccc aaccataaaa ttatttttgt tgctacttca taactgtaaa      5400
tttgctactc ttacagacca caatgtaaat atctgatatg ctatctgata tgcaggctat      5460
ctgacagagg tcgcaacccg cagggtgaga gccactgcct tcaaggcttt aatcaagaga      5520
gtagtgagct gagggcttta ctggtaagtc aggggcaagt ccaactcaat catcctcaca      5580
tactggctgc tccctcaggc ctgagaatga ggcttgagc atcctctggt ttcctaaccg      5640
ttatccatcc ctgactctca tctctgaaaa tagatgtcat ccatgaaatt aaggagtgag      5700
aatattaagc agcatttata gagctcaaaa ttccatgtca tcaccaggaa gtgccatggt      5760
gatcacagag aacacagagg agacatatag acagggtttt gctcaaaatt gggatataga      5820
atgagcctgt cagggtaccta tcaggagcgg taatccgtga gagagaaccg ttgcaagcca      5880
ctctaactgt agcaatgaaa ccctagtatt tttgtacttt gaaatacttt cttataacaa      5940
aataaagtag caaaaaaact gttcaaaaaa aaaaaaaaaa aaa                        5983

```

<210> 11

<211> 2678

<212> ADN

5 <213> *Mus musculus*

ES 2 631 455 T3

<400> 11

ctccccagtc cccgacccga gctaactaga cgtctgggca gcccagcgc aacgcgcagc 60

agcctccctc ctcttcttcc cgcactgtgc gctcctcctg ggctagggcg tctggatcga 120

ES 2 631 455 T3

gtcccgagg	ctaccgcctc	ccagacagac	gacaggtcac	ctggacgcga	gcctgtgtcc	180
gggtctcgtc	gttgccggcg	cagtcactgg	gcacaaccgt	gggactccgt	ctgtctcgga	240
ttaatcccgg	agagccagag	ccaacctctc	ccggtcagag	atgcgaccct	cagggaccgc	300
gagaaccaca	ctgctggtgt	tgttgaccgc	gctctgcgcc	gcaggtgggg	cgttggagga	360
aaagaaagtc	tgccaaggca	caagtaacag	gctcacccaa	ctgggcactt	ttgaagacca	420
ctttctgagc	ctgcagagga	tgtacaacaa	ctgtgaagtg	gtccttggga	acttggaat	480
tacctatgtg	caaaggaatt	acgacctttc	cttcttaaag	accatccagg	aggtggccgg	540
ctatgtcctc	attgccctca	acaccgtgga	gagaatccct	ttggagaacc	tgcagatcat	600
caggggaaat	gctctttatg	aaaacaccta	tgccttagcc	atcctgtcca	actatgggac	660
aaacagaact	gggcttaggg	aactgcccac	gcggaactta	caggaaatcc	tgattggtgc	720
tgtgcgattc	agcaacaacc	ccatcctctg	caatatggat	actatccagt	ggagggacat	780
cgtccaaaac	gtctttatga	gcaacatgtc	aatggactta	cagagccatc	cgagcagttg	840
ccccaaatgt	gatccaagct	gtcccaatgg	aagctgctgg	ggaggaggag	aggagaactg	900
ccagaaattg	acccaaatca	tctgtgccca	gcaatgttcc	catcgctgtc	gtggcaggtc	960
ccccagtgac	tgctgccaca	accaatgtgc	tgcggggtgt	acagggcccc	gagagagtga	1020
ctgtctggtc	tgccaaaagt	tccaagatga	ggccacatgc	aaagacacct	gcccaccact	1080
catgctgtac	aacccaccca	cctatcagat	ggatgtcaac	cctgaaggga	agtacagctt	1140
tggtgccacc	tgtgtgaaga	agtgcctccg	aaactacgtg	gtgacagatc	atggctcatg	1200
tgtccgagcc	tgtgggcctg	actactacga	agtggaagaa	gatggcatcc	gcaagtgtaa	1260
aaaatgtgat	gggccctgtc	gcaaagtttg	taatggcata	ggcattggtg	aatttaaaga	1320
cacactctcc	ataaatgcta	caaacatcaa	acacttcaaa	tactgcactg	ccatcagcgg	1380
ggaccttcac	atcctgccag	tggcctttta	gggggattct	ttcacgcgca	ctcctcctct	1440
agaccacga	gaactagaaa	ttctaaaaac	cgtaaaggaa	ataacaggct	ttttgctgat	1500
tcaggcttgg	cctgataact	ggactgacct	ccatgctttc	gagaacctag	aaataatacg	1560
tggcagaaca	aagcaacatg	gtcagttttc	tttgccggtc	gttggcctga	acatcacatc	1620
actggggctg	cgttccctca	aggagatcag	tgatggggat	gtgatcattt	ctggaaaccg	1680
aaatttgtgc	tacgcaaaca	caataaactg	gaaaaaactc	ttcgggacac	ccaatcagaa	1740
aacccaaatc	atgaacaaca	gagctgagaa	agactgcaag	gccgtgaacc	acgtctgcaa	1800
tcctttatgc	tcctcggaag	gctgctgggg	ccctgagccc	agggactgtg	tctcctgcc	1860
gaatgtgagc	agaggcaggg	agtgcgtgga	gaaatgcaac	atcctggagg	gggaaccaag	1920
ggagtttgtg	gaaaattctg	aatgcaccca	gtgccatcca	gaatgtctgc	cccaggccat	1980
gaacatcacc	tgtacaggca	ggggaccaga	caactgcac	cagtgtgccc	actacattga	2040

ES 2 631 455 T3

tggccacac	tgtgtcaaga	cctgcccagc	tggcatcatg	ggagagaaca	acactctggt	2100
ctggaagtat	gcagatgcca	ataatgtctg	ccacctatgc	cacgccaact	gtacctatgg	2160
atgtgctggg	ccaggtcttc	aaggatgtga	agtgtggcca	tctgggtacg	ttcaatggca	2220
gtggatctta	aagacctttt	ggatctaaga	ccagaagcca	tctctgactc	ccctctcacc	2280
ttccagtttc	ttccaaatcc	tctgggccag	ccagaggtct	cagattctgc	cctcttgccc	2340
tgtgcccacc	ttgttgacca	ctggacagca	tatgtgatgg	ctactgctag	tgccagcttc	2400
acaagagggt	aacactacgg	actagccatt	cttcctatgt	atctgtttct	gcaaatacag	2460
ccgctttact	taagtctcag	cacttcttag	tctcctcttt	tcctctcagt	agcccaaggg	2520
gtcatgtcac	aaacatggtg	tgaagggcta	ctttgtcaaa	tgaaaaggtc	tatcttgggg	2580
ggcatttttt	tcttttcttt	ttttcttgaa	acacattgcc	cagcaaagcc	aataaatttc	2640
tctcatcatt	ttgtttctga	taaattctta	ctattgat			2678

<210> 12

<211> 4161

<212> ADN

5 <213> *Rattus norvegicus*

<400> 12

ES 2 631 455 T3

ggaccgccac caagacaggc gacgggtcac ctggacgcga gtctgagtcc ggggtcccgtc	60
gtcgttgccg acgcagtcac cgggcacgac cgtgggactc agtctgactc ggattaatcc	120
cggagagcca gagccaacga ctgccgagcc gggatgcgac cctcagggac tgcgagaacc	180
aagctactgc tgctgctggc tgcgctctgc gccgcaggtg gggcgctgga ggaaaagaaa	240
gtttgccaag gcacaagtaa caggctcacc caactaggca cctttgaaga ccactttctg	300
agcctccaga ggatgttcaa caactgtgaa gtggtccttg gaaacttgga aatcacctat	360
gtgcaaagga attatgacct ttccttctta aagaccatcc aggaggtggc tggctatggt	420
ctcattgccc tgaacaccgt ggagagaatc cctttggaga acctgcagat catcagggga	480
aatgctctct acgaaaacac ctacgcctta gccgtcctgt ccaactatgg aaccaacaaa	540
actgggctta gggaactgcc catgcggaac ttacaggaaa ttctgatcgg tgctgtgcga	600
tttagcaaca accccatcct ctgcaatatg gagaccatcc agtggaggga catcgtccaa	660
gatgtctttc tgagcaacat gtcaatggac gtacagcgcc acctgacggg ctgcccgaaa	720
tgtgatccga gctgtcccaa tggaagctgc tggggaagag gagaggagaa ctgccagaaa	780
ttgacaaaaa tcatctgcgc ccagcaatgt tcccggcggt gtcgtggcag gtcccctagc	840
gactgctgcc acaaccagtg tgccgcaggg tgtacagggc ccagagagag tgactgtctg	900
gtctgccaca ggttccgaga tgaagccacg tgcaaagaca cctgccacc actcatgctg	960
tacaacccca ccacgtacca gatggatgtc aaccctgagg ggaagtacag ctttggtgcc	1020
acctgtgtga agaaatgccc cagaaactac gtggtgacag atcacggctc gtgtgtccgg	1080

ES 2 631 455 T3

gcctgtgggc	cagactacta	tgaagtagaa	gaagatggag	tcagcaagtg	taaaaaatgt	1140
gacggggccct	gccgcaaagt	ttgcaatggc	ataggcattg	gtgaatttaa	agacacactc	1200
tccataaatg	ctacaaacat	caaacacttc	aagtactgca	ctgccatcag	tggggacctc	1260
cacatcctgc	cagtggcctt	taagggggat	tctttcaccc	gcactcctcc	tctagaccca	1320
cgggaactag	aaattctcaa	aactgtgaag	gaaataacag	ggtttttgct	gattcaggct	1380
tggcctgaaa	actggactga	cctccatgct	tttgagaacc	tagaaataat	tcgtggcaga	1440
acaaagcaac	atggtcagtt	ttctctggcg	gttgtcggcc	tgaacataac	atcgctgggg	1500
ttgcgttccc	tcaaggagat	cagtgatggg	gatgtgatta	tttctgggaa	ccgaaatttg	1560
tgctacgcaa	acactataaa	ctggaaaaaa	ctcttcggga	cgcccaatca	aaagacccaa	1620
atcatgaaca	acagagctga	aaaggactgc	aaggccacga	accacgtctg	taatccttta	1680
tgctcctcgg	aaggctgctg	gggccctgag	cccacggact	gtgtctcctg	ccagaatgtg	1740
agcagaggca	gggagtgcgt	ggacaagtgc	aacatcctgg	agggggaacc	gagggagttt	1800
gtggaaaatt	ctgaatgcat	ccagtgccat	ccagaatgtc	tgccccagac	catgaacatc	1860
acctgtacag	gccggggggc	agacaaactgc	atcaagtgtg	cccactatgt	tgatgggtcc	1920
cactgtgtca	agacctgccc	ttcgggcctc	atgggggaga	acaacaccct	ggtctggaag	1980
tttgcatgatg	ccaataacgt	ctgccacctc	tgccatgcaa	actgtaccta	tggatgtgct	2040
gggccaggcc	ttaaaggatg	tcaacaacca	gaagggccaa	agatcccatc	catcgccact	2100
gggattgtgg	gtggcctcct	cttcatagta	gtgggtggcc	ttgggatcgg	cctcttcatg	2160
cgtcgacgtc	agcttgtccg	aaaacgtaca	ctacgccgcc	tgcttcaaga	gagagagctc	2220
gtggaacctc	tcacaccag	cggagaagct	ccgaaccaag	cccacttgag	gatattaaag	2280
gaaacagaat	tcaaaaagat	caaagttctg	ggttcaggag	catttggcac	agtgtataag	2340
ggtctctgga	tcccagaagg	cgagaaagtg	aaaatccctg	tggccatcaa	ggagttaaga	2400
gaagccacat	ctcccaaagc	caacaaggaa	atccttgatg	aagcctacgt	gatggccagt	2460
gtggacaacc	ctcatgtatg	ccgcctcctg	ggcatctgtc	tgacctccac	tgtccagctc	2520
attacacaac	tcatgcccta	tggttgcctc	ctggactatg	tccgagaaca	taaggacaac	2580
attggctccc	agtacctact	caactgggtg	gtgcagattg	caaagggcat	gaactacctg	2640
gaagaccggc	gttttggtaca	ccgtgacttg	gcagccagga	atgtactggt	aaagacacca	2700
cagcatgtca	agatcacaga	ttttggactg	gccaaactgc	ttggtgctga	ggagaaagaa	2760
taccatgcag	aggggggcaa	agtgcctatc	aagtggatgg	ctttggaatc	aattttacac	2820
cgaatttata	cacaccaaag	cgacgtctgg	agctatggag	tcaccgtgtg	ggaactgatg	2880
acctttgggt	ccaagcctta	tgatgggatc	cctgcaagtg	agatctcatc	catcctagag	2940
aaaggagagc	gccttcacac	gccacctatc	tgcaccatcg	acgtctacat	gatcatggtc	3000
aagtgcctga	tgatagatgc	tgatagccgc	ccaaagttcc	gagagttgat	tctcgaattc	3060

tccaaaatgg ccagagaccc acagcgctac cttgttatcc agggggatga aaggatgcat	3120
ttgccgagcc ctacagactc caactttttac cgagccctga tggaggagga ggacatggaa	3180
gacgtagttg atgctgatga atacctcatc ccacagcaag gcttcttcaa cagcccatcc	3240
acgtcacgga ctccactctt gagctctctg agtgcaaata gcaacagttc cactgtggct	3300
tgcattaata gaaatgggag ctgccgtgtc aaagaagacg ctttcttgca acggtatagc	3360
tccgatccca ccagcgtcct gacagaggac aacatagatg acacattcct tcccgtgcct	3420
gaatatataa accaatctgt tcccaagagg ccggctggct ctgtgcagaa cccagtctat	3480
cacaatcagc ccctgcatcc agctcctgga agagacctgc attatcaaaa tccccatagc	3540
aatgcggtga gcaaccctga gtatctcaac actgcccagc cgacctgcct cagtagtggg	3600
tttgacagct ctgccctctg gatccagaaa ggcagccacc aaatgagcct ggacaaccct	3660
gactaccagc aggacttctt tcccaaagaa gccaagccga atggcatctt taagggtccc	3720
acagctgaaa atgcagagta cctgcgggtg gcaccgcaa gcagtgagtt tagtggagca	3780
tgacattgaa gaggcattgt accagttaca aaccggactt tccagaggcc caggaccaag	3840
ccatggcagc acctctgctc ctgacagcca tgtccacatt gtgtcaaata tcaaaccctc	3900
agactggctt taaagcataa ctctgacggg ctttgtcact gagccaagaa gtgggtcccct	3960
cccctgatgc tctttgggaa gttgaaggta tatctattgg tcttcgaact gtgaagattc	4020
cactgaaagg tatccatcga gaacattgtc cttttggaac agaaggttgc ctcatggtga	4080
ggtacatagg gggaaaaaaaa acagacctat ggcgcttgct gacataggga actctgggat	4140
tcttgtcttt attgatttgc t	4161

<210> 13

<211> 4788

5 <212> ADN

<213> *Canis familiaris*

<400> 13

ES 2 631 455 T3

atgctttcac cacccaagtg tgagcaagtg caaggccgta gaggcaccgc aggaccggga	60
aaccgggaag agcagcggga tggggttgtc ctctggggg gtgtcgctgt gagcggaggc	120
cacacacgtg gaattcccga gagcgccgag gaggaggag acctggagca gcgtccccgg	180
gtgggaaaag caggggtgag aaaagcagtg ggaggacgca gtcattgtct ccactgtctg	240
cagggctgcg aggtcgagga gaagcaggtc gcatctctgt tcaccggag gaagtcatgg	300
ggggacaagc ttggggcccc taaagcacag ggtcatggga accctgatcc tatacttcca	360
gacctcctgt taggccttcc tctgaagcag agctcgggta agtacagccc gtggttcaaa	420
tccagcgact gcctgttttt taagccaagt ttctgggcac gagccggtgc caggccatt	480
cgtttggttc ttgtctgtgg ctgctttttg cacgctggcg gcggagcgga ggggctggga	540
cggacccttg gagaatgttc tgagcagcaa gctgggagtt ctctgcggg aattctgaaa	600

ES 2 631 455 T3

ggcccatccg aggtgtggcc ccctcagctt tcagtgcccg gcttgccccc cccacccccg	660
ctcaaatacca cccttgcccc ccgaggctcc ccgtggaagg gctctgatat ctgcctgcac	720
ctcattgccca ggcttgccctg cgctcctgca acaggcacct ttggtcttgg ggcagctgag	780
gctcgtaggc tagctgaggg gctggacaag ccttacctgg caggctgtaa aacagttctg	840
gaaggtcagg attatatttca gcatgcctac gacttgagga aagcgtgccg gggagagctt	900
agagaagtca ttccgggtca cgtcgtggct tattgtcgca accctgcacc tacatcagtg	960
gtggcatcgt tagggacctt gggattgggc aacgagctctg gccccatgtg catgcgttgc	1020
cgccagctgt gtacagatgt tccaagtga ggggtcaatc agagcagcct gtgcttccga	1080
gtgcgagcag cccagagctg tgaggccaca gaagagggaa atccccatag agttcaagag	1140
gatgtggcga ttgcttcata tttaatgtca gtgtcagtg atgggggtga tgatgtgggtg	1200
tctgaacatg tgctcgtgc cgtggcacca aaacgcaagg gcagacggtc agtctgccaa	1260
ggcacgagta accggctcac ccagttgggc acttttgaag accactttct gagccttcag	1320
aggatgttca ataactgtga ggtggtcctg ggggaatttg aaattaccta catgcaaagg	1380
aattatgacc tttccttctt aaagaccatc caggaagttg ccgggtatgt attgattgcc	1440
ctcaacacgg tggagaagat tcctttggaa aacctgcaga tcatccgagg aaatgtgctc	1500
tatgaaaaca cccatgcctt atccgtcctc tccaactatg gctcaaataa aactggactg	1560
caggagcttc ccctgagaaa cttgcatgaa atactacaag gcgctgtgcg cttcagcaat	1620
aaccctgtcc tctgcaacgt ggaaaccatc cagtggcggg acatcgttga caacgatttt	1680
ataagcaaca tgtcgatgga catccagaac caagcgggca ggtgccaaaa gtgtgatcca	1740
agctgtccca atggaagttg ctgggggtccg ggaaaggaga attgccagaa actgacccaa	1800
atcatctgtg cccagcagtg ttccggggcg tgccgcggca ggtccccag cgactgctgc	1860
cacaaccagt gtgccgccg ctgcacgggg ccccgcgaga gcgactgcct ggtctgccgc	1920
aagttccggg atgaagccac gtgcaaggac acctgcccgc ccctcatgct ctacaacccc	1980
accacctacc aaatggacgt caaccagag ggaaaataca gctttggtgc cacctgcgtg	2040
aagaaatgcc cccgcaacta cgtggtgaca gaccacggtt catgtgtccg cgctgcagc	2100
tctgacagct acgaggtgga ggaggatggt gtccgcaagt gtaagaagtg tgaggggcct	2160
tgtcgcaaag tttgtaatgg aatagggtt ggagagttca aagacacact ttccataaat	2220
gctacaaaca ttaaactctt caaaaactgc acgtcgatca gtggagacct tcatatcttg	2280
ccagtcgcat ttagaggtga ctccctcacg cataccctac ctctagatcc gaaggagctg	2340
gatatcctga aaactgtcaa ggaaataaca ggggttttgc tgattcaggc ctggcccga	2400
aataggactg acctccatgc tttcgagaac ctagaaatca tacgtggaag aacaaagcaa	2460
catggccagt tttctctggc ggtcgtcggc ctgaacataa cgtccttggg attacgctcc	2520

ctcaaggaga taagcgatgg agatgtgac atctctggaa accgaaaatt gtgctatgca	2580
aatacaataa actggaagaa actctttggg acttcaagtc agaaaaccaa aattataaac	2640
aacaaagatg aaaaagcctg caaggccata ggccacgtct gccatccatt gtgctcctca	2700
gaaggctgct gggggcccagg gccgagagac tgcgtgtcct gccgaaacgt cagccgtggc	2760
aaggaatgtg tggaaaagtg caacattctg gagggggagc caagggagtt tgtggagaat	2820
tgggagtgca tacagtgcc tccggaatgc ctgcccagg ccatgaacat aacctgcaca	2880
ggacgggggc cggacagctg catcaagtgc gccactaca tcgatggccc tcaactgcgtc	2940
aagacctgcc cggctggcat catgggagaa aacaacaccc tggctctgga gttttcggat	3000
ggcagccgca tgtgccacct gtgccatcca aactgcacct atggctgtga gggaccaggt	3060
cttgaaggct gtgcaaaacc tgggcccagg atcccatcca ttgctaccgg gattgtcggc	3120
ggcctcctct tgggtggtgg gtgtgccctt ggagtgggcc tctttttgcg ccggcgccac	3180
attgtccgga agcgcacgct tcgcagactg ctgcaagaaa gagagcttgt cgagcctctt	3240
acaccagcg gagaagctcc caaccaggct ctcttgagga tcttaaagga gacggagttc	3300
aaaaagatca aggtgctggg ctctggagca ttccggcacag tgtacaaggg actctggac	3360
ccagaaggcg agaaggttaa aattcccgtg gccatcaagg aattgagaga agccacgtct	3420
cccaaagcca acaaggaaat tcttgatgaa gcctacgtga tggccagtgt ggacaatccc	3480
cacgtgtgcc gcctcctggg catctgcctg acgtccacgg tgcagctcat cacgcagctc	3540
atgccctttg gctgcctcct ggactatgtc cgcgagcaca aggacaacat cggctcccag	3600
cacctgctca actggtgtgt gcagattgca aagggcatga actacctgga agaccggcgc	3660
ttggtgcacc gcgacctggc ggccaggaa gtcctggtga agaccccga gcacgtcaag	3720
atcacagatt ttgggctggc caaactgctg ggtgccgagg agaaagagta ccacgcggaa	3780
ggaggcaaag tgcctatcaa gtggatggct ttagaatcga ttttacaccg aatttataacc	3840
caccaaagcg atgtgtggag ctacggcgtc accgtgtggg agctgatgac cttcgggtcc	3900
aagccttacg acggtatccc tgcaagtgag atctccacca tccctggagaa gggagagcgc	3960
ctcccgcagc cgcocatatg caccatcgat gtctacatga tcatggtcaa gtgctggatg	4020
atagatgcag acagtcgccc aaaattccgt gagttgatca tcgaattctc caaaatggcc	4080
cgagaccgc agcgtacct tgtcatccag ggagatgaga ggatgcattt gccaaacct	4140
acagactcca atttttaccg cgcctgatg gacgaggagg acatggagga tgttgtggat	4200
gctgacgagt acctcatccc ccagcagggc ttcttccaca gccctccac ttcccggacc	4260
cccctcctaa gttctctgag cgcaccagc aacagttcca acgtggcttg catcgaccga	4320
aatgggacct gtcccctcaa agaagacagc ttcttgacgc ggtacagctc agaccccact	4380
ggcaccttga cggaggacaa catagatgac actttcctcc cagcaccoga atacataaac	4440
cagtctgttc ccaaaggcc tgcgggttct gtccagaatc ccgtctatca caatcagcct	4500

ES 2 631 455 T3

ctaaatccag ctctgccag agaccctcac taccaaaatc cccacagcaa cgcagtggac	4560
aaccctgagt atctcaacac ccaccccacc tgcgtcaaca gtgtcctcga caggcccagc	4620
ctctggaccc aggaggccaa ccaccaaatc agcctggaca accctgacta ccagcaggac	4680
ttctttccca aggaagccaa gtccaatggc atttttaagg gccctgcagc tgaaaatgca	4740
gactacctga gggtagcgcc accaagcagt gagtttattg gagcgtga	4788

REIVINDICACIONES

1. Método para predecir el desenlace clínico de un sujeto que padece glioblastoma multiforme (GBM) que comprende:
 - a) determinar el nivel de expresión o el nivel de polisomía/amplificación del gen de EGFR y la pérdida de nivel de heterocigosidad (LOH) del gen de PTEN en una muestra del mismo sujeto, y
 - b) comparar dicho nivel de expresión o el nivel de polisomía/amplificación del gen de EGFR y el nivel LOH del gen de PTEN con valores de referencia convencionales,
 en el que el nivel LOH del gen de PTEN se mide mediante PCR, mediante un ensayo basado en hibridación, mediante tecnología de secuenciación o mediante un análisis de SNP; y
- 10 en el que un alto nivel LOH del gen de PTEN con respecto a dicho valor de referencia convencional y un alto nivel de expresión y/o alto nivel de polisomía/amplificación del gen de EGFR con respecto a dichos valores de referencia convencionales son indicativos de un buen desenlace clínico del sujeto.
2. Método según la reivindicación 1, en el que el desenlace clínico se mide como supervivencia.
- 15 3. Método según una cualquiera de las reivindicaciones 1 ó 2, en el que la muestra es una muestra celular o tisular, preferiblemente una muestra de tejido tumoral.
4. Método según una cualquiera de las reivindicaciones 1 a 3, en el que el nivel de expresión del gen de EGFR se mide determinando el ARNm y/o el nivel de expresión de proteína de dicho gen.
5. Método según la reivindicación 1, en el que dicho ensayo basado en hibridación comprende una transferencia de tipo Southern, hibridación *in situ* (ISH), hibridación *in situ* fluorescente (ISH), o un ensayo de hibridación genómica comparativa (CGH).
- 20 6. Método según la reivindicación 1 y 5, en el que el nivel LOH del gen de PTEN se determina mediante FISH.
7. Método según una cualquiera de las reivindicaciones 1 a 6, en el que el glioblastoma es glioblastoma precoz.
8. Uso de un kit que comprende un conjunto de agentes que pueden determinar específicamente el nivel de expresión o el nivel de polisomía/amplificación del gen de EGFR y el nivel LOH del gen de PTEN para predecir el desenlace clínico de un sujeto que padece glioblastoma multiforme, en el que el conjunto de agentes que pueden determinar específicamente el nivel LOH del gen de PTEN comprende un par de cebadores de oligonucleótido para amplificar un fragmento específico del gen de PTEN, o una sonda de oligonucleótido marcada opcionalmente que se une selectivamente a una secuencia de polinucleótido diana sobre la región del cromosoma del gen de PTEN, o reactivos adecuados para realizar una reacción de secuenciación, o reactivos para realizar un análisis de SNP, y en el que si dichos agentes detectan altos niveles de expresión o un alto nivel de polisomía/ amplificación del gen de EGFR y un alto nivel LOH del gen de PTEN, con respecto a valores de referencia, entonces el desenlace clínico del sujeto es bueno.
- 25 9. Uso según la reivindicación 8, en el que el kit comprende además un reactivo para detectar un gen constitutivo o la proteína codificada por dicho gen constitutivo.
10. Uso según cualquiera de las reivindicaciones 8 ó 9, en el que el conjunto de agentes para determinar específicamente el nivel de expresión y/o el nivel de polisomía/amplificación del EGFR comprende un reactivo para detectar el nivel de ARNm del gen de EGFR o el nivel de la proteína EGFR.
- 30 11. Uso según la reivindicación 10, en el que el reactivo para detectar el nivel de ARNm del gen de EGFR o el nivel de la proteína EGFR comprende una sonda que se hibrida específicamente con el gen de EGFR, o un anticuerpo que se une a la proteína EGFR.
- 40

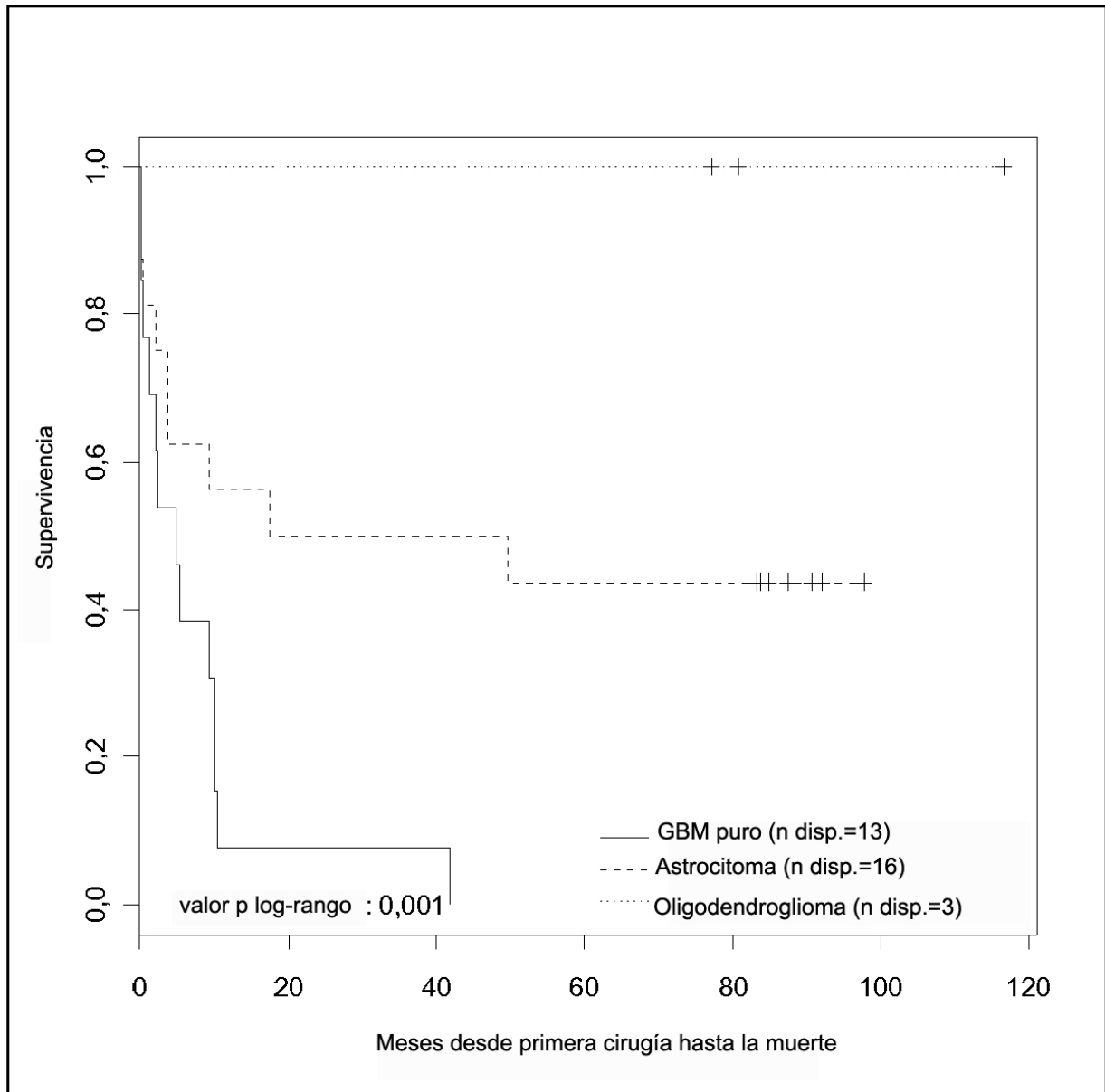
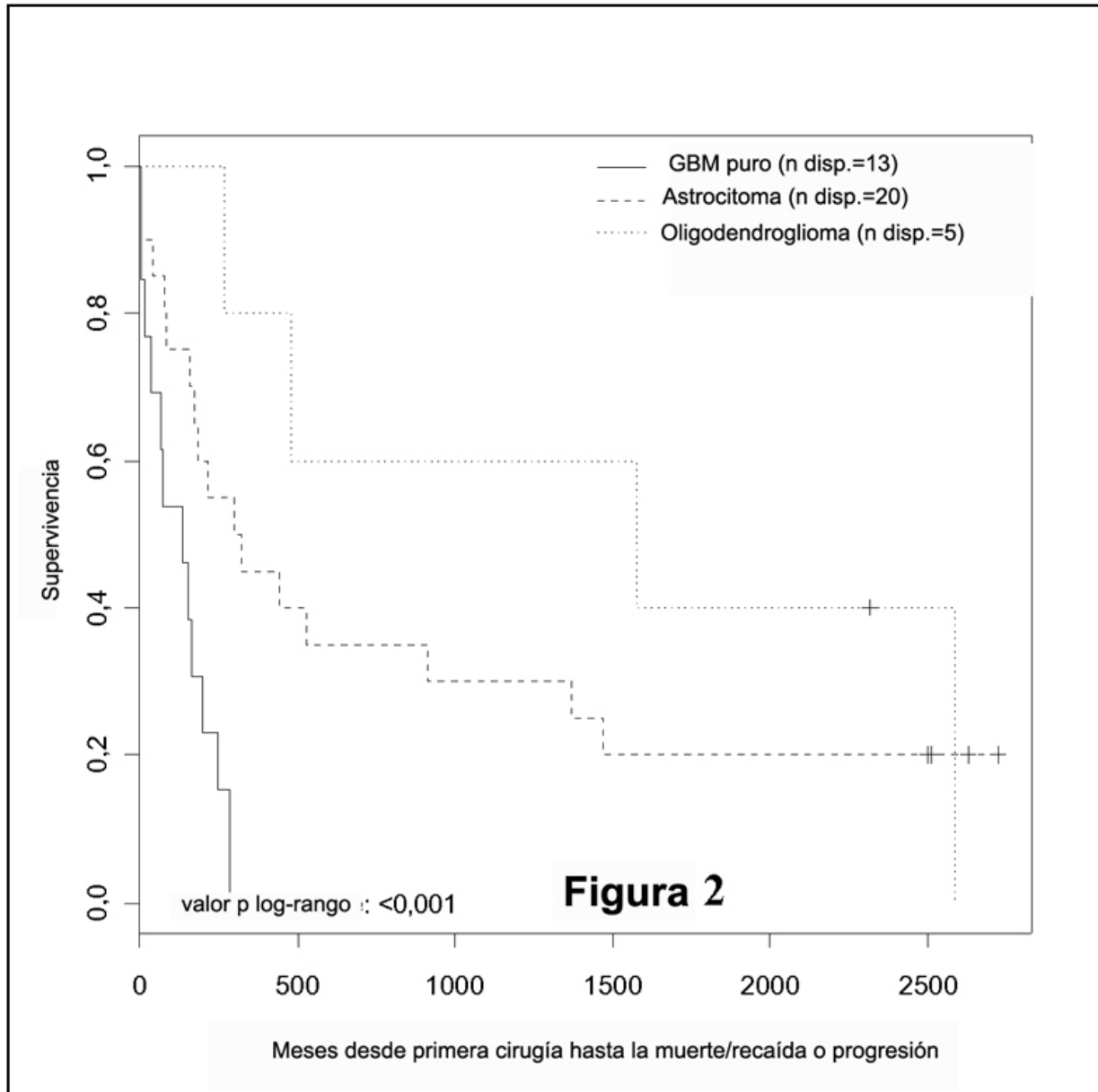


Figura 1



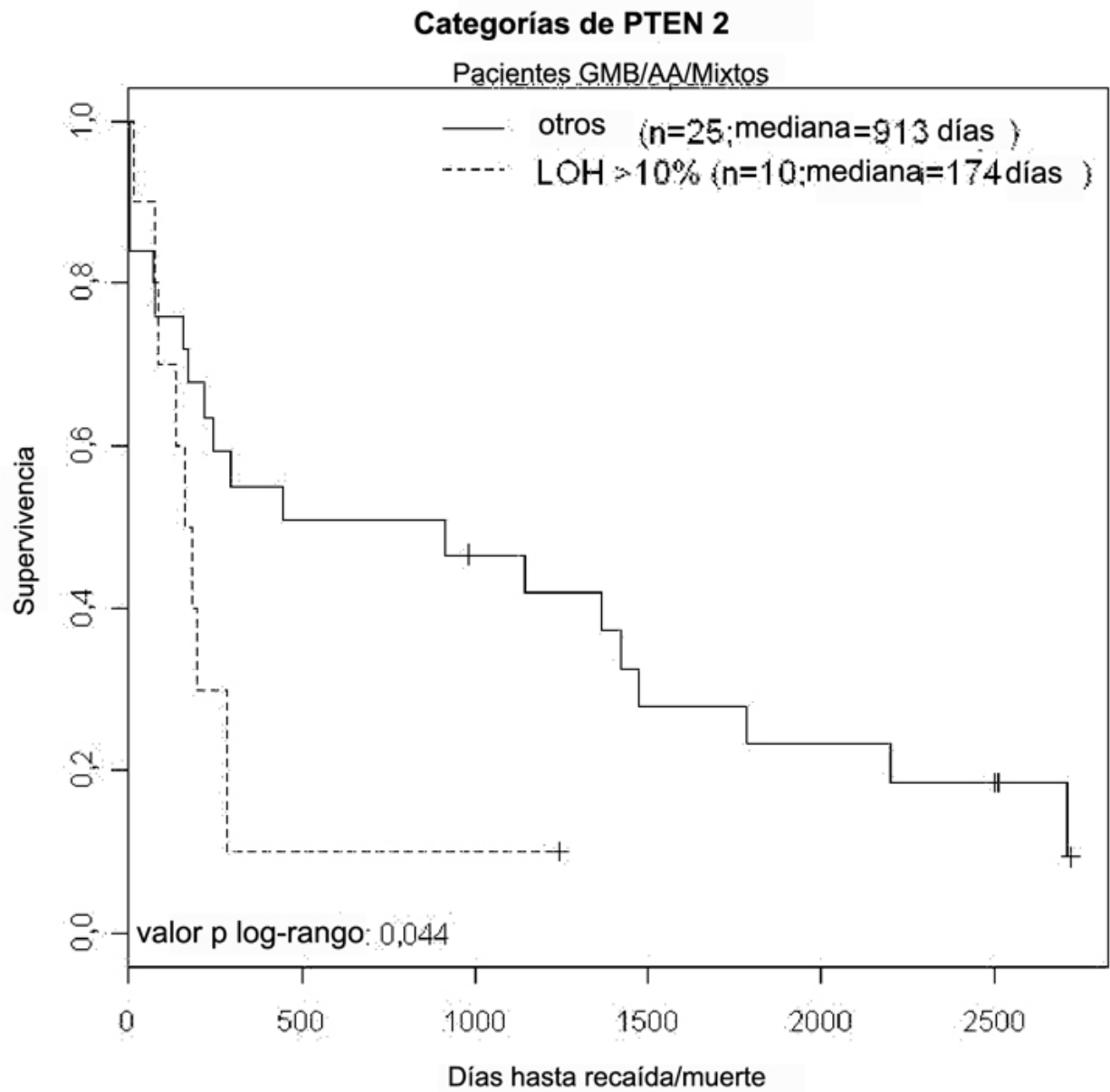


Figura 3

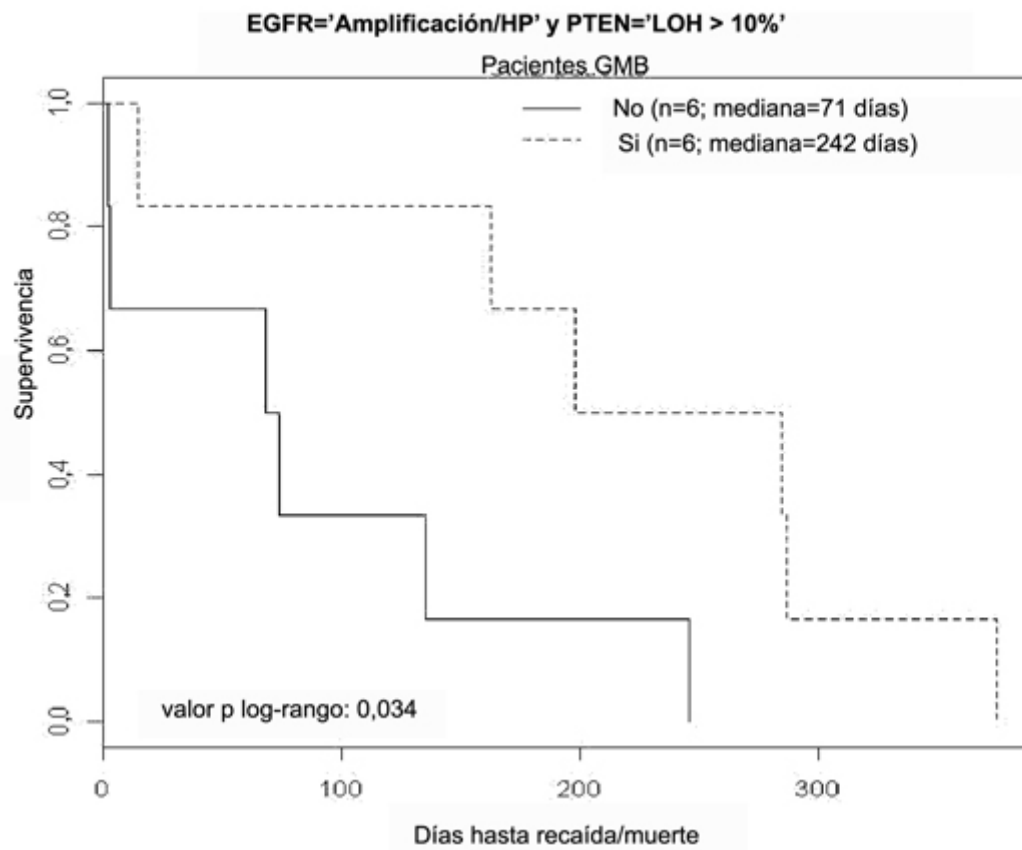


Figura 4

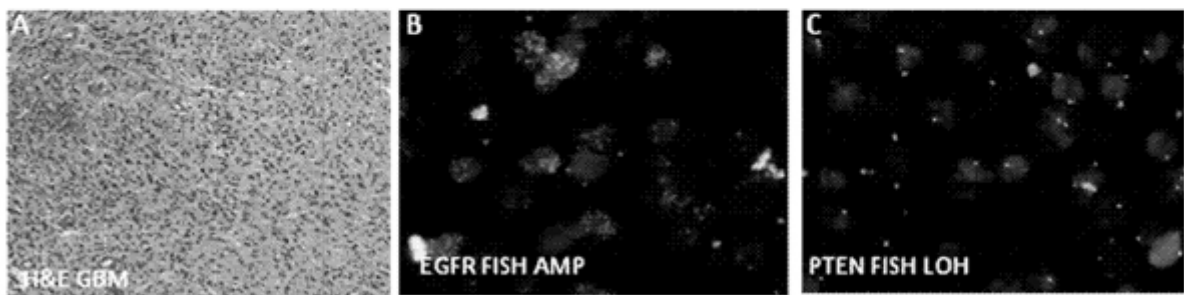


Figura 5