

(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS
ESPAÑA



(11) Número de publicación: **2 433 012**

(51) Int. Cl.:

C12N 15/12 (2006.01)
C12N 15/63 (2006.01)
C07K 14/435 (2006.01)
C07K 16/18 (2006.01)
A61K 48/00 (2006.01)
A61K 39/395 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Fecha de presentación y número de la solicitud europea: **31.01.2002 E 02711634 (2)**

(97) Fecha y número de publicación de la concesión europea: **31.07.2013 EP 1364025**

(54) Título: **Un nuevo gen BNO1 cartografiado en el cromosoma 16q24.3**

(30) Prioridad:

31.01.2001 AU PP278301

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
05.12.2013

(73) Titular/es:

**BIONOMICS LIMITED (100.0%)
31 DALGLEISH STREET
THEbarton, SOUTH AUSTRALIA 5031, AU**

(72) Inventor/es:

**CALLEN, DAVID, FREDERICK;
POWELL, JASON, ANTHONY;
KREMMIDIOTIS, GABRIEL;
GARDNER, ALISON, ELAINE;
CRAWFORD, JOANNA;
BAIS, ANTHONY, JOHN y
KOCHETKOVA, MARINA**

(74) Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 433 012 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

Un nuevo gen BNO1 cartografiado en el cromosoma 16q24.3.

Campo técnico

La presente invención se refiere a un nuevo gen que se ha identificado en la punta distal del brazo largo del cromosoma 16 en 16q24.3. El gen BNO1 codifica un polipéptido que forma parte de un complejo de ubicitinina-ligasa implicado en el 5
direcciónamiento de proteínas mediante ubicitinación para una degradación en el proteasoma. Considerando que BNO1 está involucrado en la ubicitinación y la degradación de proteínas, la invención también se refiere a la terapia de trastornos asociados con este proceso, tales como el cáncer (en particular el carcinoma de mama y de próstata), la enfermedad inmune/inflamatoria y la enfermedad neurológica. Además, la invención se refiere al diagnóstico de trastornos 10
asociados con la ubicitinación y al escrutinio de fármacos para una intervención terapéutica en estos trastornos.

Técnica anterior

Se ha observado que el desarrollo de carcinomas humanos surge por una acumulación de cambios genéticos que implican agentes reguladores positivos de la función celular (oncogenes) y agentes reguladores negativos (genes supresores de tumores). Para que una célula somática normal se desarrolle en un tumor metastásico se requieren cambios a nivel 15
celular, tales como la inmortalización, la pérdida de inhibición por contacto y la capacidad de crecimiento invasivo, y cambios a nivel tisular, tales como evitar respuestas inmunes del hospedador y restricciones del crecimiento impuestas por las células circundantes, y la formación de un aporte de sangre para el tumor en crecimiento.

Estudios genéticos moleculares de carcinoma colorrectal han proporcionado pruebas sustanciales de que la generación 20
de malignidad requiere la acumulación secuencial de una serie de cambios genéticos dentro de la misma célula madre epitelial del colon. Para que una célula epitelial colónica normal se convierta en un adenoma benigno, progrese a adenoma intermedio y final, y finalmente se convierte en una célula maligna, son necesarias mutaciones inactivadoras en genes supresores de tumores y mutaciones activadoras en proto-oncogenes (Fearon y Vogelstein, 1990).

El empleo de una variedad de técnicas, tales como la pérdida de heterocigosis (LOH), la hibridación genómica comparativa (CGH) y estudios citogenéticos de tejidos cancerosos, beneficiándose todas ellas de anomalías cromosómicas 25
asociadas con la célula afectada, ha ayudado a identificar una variedad de genes supresores de tumores y de oncogenes asociados con una serie de tipos de tumores.

En un aspecto, estudios de cánceres tales como retinoblastoma y carcinoma de colon han respaldado el modelo de que 30
LOH es un episodio específico en la patogénesis del cáncer y se ha proporcionado un mecanismo para identificar los genes causantes de cáncer. Este modelo se pone de relieve aún más en el síndrome de Von Hippel-Lindau (VHL), un trastorno raro que predispone individuos a una variedad de tumores, tales como carcinomas de células claras del riñón y tumores de células de los islotes del páncreas. Tanto los casos esporádicos como los casos heredados del síndrome muestran LOH en el brazo corto del cromosoma 3 y translocaciones somáticas que implican 3p en tumores esporádicos, y también se ha observado un ligamiento genético con la misma región en familias afectadas. El gen supresor del tumor de VHL ya se ha identificado a partir de esta región del cromosoma 3 y se han detectado mutaciones en el mismo en el 35
100% de los pacientes que tienen un diagnóstico clínico de enfermedad de VHL. Además, el gen de VHL se inactiva en aproximadamente el 50-80% de la forma esporádica más común de carcinoma renal de células claras.

Los determinantes genéticos implicados en el cáncer de mama no están tan bien definidos como los del cáncer de colon, debido en parte a que los estadios histológicos del desarrollo del cáncer de mama están peor caracterizados. Sin embargo, igual que con el carcinoma de colon, se cree que una tienen que estar implicados variedad de genes en una 40
progresión por etapas, durante la génesis de tumores de mama.

Algunas mujeres parecen tener un mayor riesgo de desarrollar cáncer de mama. Análisis de ligamiento genético han mostrado que 5 a 10% de todos los cánceres de mama son debidos al menos a dos genes de susceptibilidad autosómicos dominantes. Generalmente, las mujeres portadoras de una mutación en un gen de susceptibilidad desarrollan cáncer de mama a una edad más temprana, en comparación con la población general, frecuentemente tienen tumores de mama bilaterales y tienen un mayor riesgo de desarrollar cánceres en otros órganos, particularmente carcinoma de ovario.

Un análisis de ligamiento genético en familias que muestran una alta incidencia de cáncer de mama de aparición temprana (antes de los 46 años de edad) fue un éxito en la cartografía del primer gen de susceptibilidad, *BRCA1*, en el 50
cromosoma 17q21 (Hall *et al.*, 1990). Con posterioridad a esto, se cartografió el gen *BRCA2* en el cromosoma 13q12-q13 (Wooster *et al.*, 1994), confirmando este gen una mayor incidencia de cáncer de mama masculino y una menor incidencia de cáncer de ovario, cuando se comparaba con *BRCA1*.

Tanto *BRCA1* como *BRCA2*, ya han sido clonados (Miki *et al.*, 1994; Wooster *et al.*, 1995) y se han identificado numerosas mutaciones en estos genes en individuos susceptibles con casos familiares de cáncer de mama.

Existen otros síndromes de cáncer de mama heredado, sin embargo, son raros. Las mutaciones hereditarias en el gen 55
TP53 se han identificado en individuos con síndrome de Li-Fraumeni, un cáncer familiar que produce neoplasias epite-

liales que tienen lugar en sitios múltiples, incluyendo la mama. Del mismo modo, mutaciones de la línea germinal en el gen *MMC4C1/PTEN* implicado en la enfermedad de Cowden y en el gen de la ataxia telangiectasia (AT), han mostrado que confieren un mayor riesgo de desarrollar cáncer de mama, entre otras manifestaciones clínicas, pero en conjunto solo representan un porcentaje bajo de las familias con una predisposición hereditaria al cáncer de mama.

- 5 Se ha mostrado que las mutaciones somáticas en el gen *TP53* se producen en un alto porcentaje de individuos con cáncer de mama esporádico. Sin embargo, a pesar de que se ha observado LOH en los loci de *BRCA1* y *BRCA2* con una frecuencia de 30 a 40% en los casos esporádicos (Cleton-Jansen *et al.*, 1995; Saito *et al.*, 1993), no hay prácticamente ninguna señal de mutaciones somáticas en el alelo conservado de estos dos genes en los cánceres esporádicos (Futreal *et al.*, 1994; Miki *et al.*, 1996). Datos recientes sugieren que la metilación del ADN de la secuencia del promotor 10 de estos genes puede ser un importante mecanismo de regulación a la baja. El uso de polimorfismos en la longitud de los fragmentos de restricción y de marcadores polimórficos de pequeñas repeticiones en tandem, ha identificado numerosas regiones con desequilibrio alélico en el cáncer de mama, lo que sugiere la presencia de genes adicionales que pueden estar implicados en el cáncer de mama. Datos compilados a partir de más de 30 estudios revelan la pérdida de ADN en al menos 11 brazos de cromosomas, con una frecuencia superior al 25%, estando regiones tales como 16q y 15 17p afectadas en más del 50% de los tumores (Devilee y Cornelisse, 1994; Brenner y Aldaz, 1995). Sin embargo, se conoce que solo algunas de estas regiones albergan genes supresores de tumores, los cuales se ha mostrado que están mutados en individuos con las dos formas de cáncer de mama, esporádica (genes *TP53* y *RB*) y familiar (genes *TP53*, *RB*, *BRCA1* y *BRCA2*).

20 Estudios citogenéticos han implicado la pérdida del brazo largo del cromosoma 16 como un episodio temprano en la carcinogénesis de mama, ya que se encuentra en tumores con pocas o ninguna otra anormalidad citogenética. Alteraciones en el cromosoma 1 y 16 también se han observado en varios casos de carcinoma ductal *in situ* (DCIS), el estadio preinvasivo de carcinoma de mama ductal. Además, estudios de LOH en muestras de DCIS identificaron la pérdida de marcadores de 16q en el 29 al 89% de los casos sometidos a ensayo (Chen *et al.*, 1996; Radford *et al.*, 1995). Además, el examen de tumores de otros tipos de tejido ha indicado que la LOH de 16q también se observa frecuentemente en los 25 carcinomas de próstata, hígado, ovario y neuroectodérmicos primitivos. En conjunto, estos hallazgos sugieren la presencia de un gen que se cartografía en el brazo largo del cromosoma 16 que está involucrado de manera decisiva en el desarrollo temprano de una gran proporción de cánceres de mama, así como de cánceres de otros tipos de tejidos, pero hasta la fecha no se ha identificado un gen de este tipo.

Descripción de la invención

- 30 La presente invención proporciona una molécula aislada de ácido nucleico de BNO1 que se cartografía en el cromosoma 16q24.3 que comprende la secuencia de nucleótidos expuesta en SEQ ID Números: 1 o 3.

También proporciona una molécula aislada de ácido nucleico de BNO1 que comprende la secuencia de nucleótidos expuesta en SEQ ID Números: 1 o 3, que codifica un polipéptido que forma parte de un complejo de ubictinina-ligasa implicado en el direccionamiento de proteínas mediante ubicitinación para una degradación en el proteasoma.

- 35 La invención también incluye una molécula aislada de ácido nucleico de BNO1 que es al menos 95% idéntica a una molécula de ADN que consiste en la secuencia de nucleótidos expuesta en SEQ ID Números: 1 o 3 y que codifica un polipéptido que forma parte de un complejo de ubictinina-ligasa implicado en el direccionamiento de proteínas mediante ubicitinación para una degradación en el proteasoma.

40 Una cualquiera de las variantes de polinucleótidos descritas anteriormente puede codificar una secuencia de aminoácidos que contiene al menos una característica funcional o estructural de BNO1.

Típicamente, la identidad de secuencia se calcula utilizando el algoritmo BLASTN con la matriz por defecto BLOSUM62.

45 La invención también incluye una molécula aislada de ácido nucleico de BNO1 que codifica un polipéptido que forma parte de un complejo de ubictinina-ligasa implicado en la degradación de proteínas mediante ubicitinación, y que se hibrida en condiciones rigurosas con una molécula de ADN que consiste en la secuencia de nucleótidos expuesta en SEQ ID Números : 1 o 3.

En condiciones rigurosas, la hibridación tendrá lugar lo más preferiblemente a 42°C en NaCl 750 mM, citrato trisódico 75 mM, 2% de SDS, 50% de formamida, 1X Denhart, 10% (v/v) de sulfato de dextrano y 100 µg/ml de ADN de esperma de salmón desnaturalizado. Variaciones útiles de estas condiciones serán claramente evidentes para los expertos en la 50 técnica. Las etapas de lavado que siguen a la hibridación se producen lo más preferiblemente a 65°C en NaCl 15 mM, citrato trisódico 1,5 mM y 1% de SDS. Variaciones adicionales de estas condiciones serán claramente evidentes para los expertos en la materia

La invención también proporciona una molécula aislada de ácido nucleico de BNO1 que codifica un polipéptido que tiene la secuencia de aminoácidos expuesta en SEQ ID Números: 2 o 4.

- 55 Preferiblemente, la identidad de la secuencia se determina utilizando el algoritmo BLASTP con la matriz por defecto BLOSUM62.

También describimos una molécula de ácido nucleico aislada que comprende los exones 1 a 9 o los exones 1, 2, 2.5 y 3 a 9, identificados en las secuencias de nucleótidos expuestas en SEQ ID Números: 1 y 3, respectivamente.

Aún más, se proporciona una molécula de ácido nucleico aislada que consiste en la secuencia de nucleótidos expuesta en SEQ ID Números: 1 o 3.

5 También describimos una molécula de ácido nucleico aislada que consiste en la secuencia de nucleótidos expuesta en SEQ ID NO: 1 desde la base 4 hasta la base 1.621 o expuesta en SEQ ID NO: 3 desde la base 4 hasta la base 1.708.

También describimos un gen aislado que comprende la secuencia de nucleótidos expuesta en SEQ ID Números: 1 o 3 y los elementos de control de **BNO1**.

10 Preferiblemente, los elementos de control de BNO1 son los que median en la expresión en el tejido de mama, próstata, hígado y ovario.

Las secuencias de nucleótidos de la presente invención se pueden modificar genéticamente usando métodos aceptados en la técnica para alterar las secuencias que codifican BNO1, para una variedad de objetivos. Estos incluyen, pero no se limitan a los mismos, la modificación de la clonación, el procesamiento y/o la expresión del producto génico. El reacomplamiento con PCR de fragmentos génicos y el uso de oligonucleótidos sintéticos permiten la modificación genética de secuencias de nucleótidos de BNO1. Por ejemplo, la mutagénesis dirigida al sitio mediada con nucleótidos puede introducir mutaciones que crean nuevos sitios de restricción, alterar los patrones de glicosilación y producir variantes por corte y empalme, etc.

20 Como resultado de la degeneración del código genético, se puede producir una serie de secuencias de polinucleótidos que codifican BNO1, algunas que pueden tener un mínimo de similitud con las secuencias de polinucleótidos de cualquier gen conocido y de origen natural. Por lo tanto, describimos todas y cada variación posible de la secuencia de polinucleótidos que se podría realizar mediante la selección de combinaciones basadas en posibles opciones de codones. Estas combinaciones se realizan de acuerdo con el código genético de tripletes convencional, tal como se aplica a la secuencia de polinucleótidos de BNO1 de origen natural, y todas las variaciones de este tipo se deben considerar como que se han descrito específicamente.

25 25 Los polinucleótidos de esta invención incluyen ARN, ADNc, ADN genómico, formas sintéticas y polímeros mixtos, cadenas codificantes y no codificantes, y se pueden modificar química o bioquímicamente, o pueden contener bases de nucleótidos no naturales o derivatizadas, tal como apreciarán los expertos en la técnica. Tales modificaciones incluyen marcadores, metilación, agentes intercalantes, alquilantes y enlaces modificados. En algunos casos, puede ser ventajoso producir secuencias de nucleótidos que codifican BNO1 o sus derivados que poseen un uso de codones sustancialmente diferente al de BNO1 de origen natural. Por ejemplo, se pueden seleccionar codones para aumentar la tasa de expresión del péptido en un hospedador procariótico o eucariótico particular que se corresponda con la frecuencia con la que codones particulares son utilizados por el hospedador. Otras razones para alterar la secuencia de nucleótidos que codifica BNO1 y sus derivados, sin alterar las secuencias de aminoácidos codificadas, incluyen la producción de transcritos de ARN que tienen propiedades más deseables, tales como una semivida mayor que los transcritos producidos a partir de la secuencia de origen natural.

35 La invención también incluye la producción de moléculas de ADN, que codifican BNO1 y sus derivados, o fragmentos de los mismos, enteramente mediante química sintética. Las secuencias sintéticas se pueden insertar en vectores de expresión y sistemas celulares que contienen los elementos necesarios para el control de la transcripción y la traducción de la secuencia codificadora insertada, en un hospedador adecuado. Estos elementos pueden incluir secuencias reguladoras, promotores, regiones 5' y 3' no traducidas y señales de iniciación específicas (tales como un codón de iniciación ATG y la secuencia de consenso Kozak) que permiten una traducción más eficaz de las secuencias que codifican BNO1. En los casos en los que se inserta la secuencia completa que codifica BNO1, incluyendo su codón de iniciación y secuencias reguladoras aguas arriba, en el vector de expresión apropiado, puede que no sean necesarias unas señales de control adicionales. Sin embargo, en los casos en que solo se inserta la secuencia codificadora o un fragmento de la misma, el vector tiene que proporcionar señales exógenas de control de la traducción, tal y como se ha descrito anteriormente. Tales señales pueden tener diversos orígenes, tanto naturales como sintéticos. La eficacia de la expresión se puede potenciar mediante la inclusión de potenciadores apropiados para el sistema particular de células hospedadoras empleado (Scharf *et al.*, 1994).

40 50 La presente invención permite la preparación de un polipéptido o una proteína purificada de BNO1, a partir de los polinucleótidos de la presente invención o de variantes de los mismos. Con el fin de realizar esto, las células hospedadoras se pueden transfectar con una molécula de ADN tal y como se ha descrito anteriormente. Típicamente, dichas células hospedadoras se transfecan con un vector de expresión que comprende una molécula de ADN de acuerdo con la invención. Una variedad de sistemas de vector de expresión/hospedador se puede utilizar para contener y expresar secuencias que codifican BNO1. Estos incluyen, pero no se limitan a los mismos, microorganismos tales como bacterias transformadas con vectores de expresión de ADN de plásmido o cósmido; levadura transformada con vectores de expresión de levadura, sistemas de células de insecto infectados con vectores de expresión vírica (por ejemplo, baculovirus); o sistemas de células tisulares de ratón o de otros animales o de seres humanos. Las células de mamífero también se pueden utilizar para expresar la proteína de BNO1 usando diversos vectores de expresión que incluyen sistemas de

plásmido, cósmido y víricos tales como sistemas de expresión de virus adenovíricos, retrovíricos o vaccinia. La invención no está limitada por la célula hospedadora empleada.

Las secuencias de polinucleótidos, o variantes de las mismas, de la presente invención se pueden expresar de manera estable en líneas celulares para permitir una producción a largo plazo de proteínas recombinantes en sistemas de mamíferos. Las secuencias que codifican BNO1 se pueden transformar en líneas celulares utilizando vectores de expresión que pueden contener orígenes de replicación víricos y/o elementos de expresión endógenos y un gen marcador seleccionable en el mismo vector o en un vector distinto. El marcador seleccionable confiere resistencia a un agente selectivo, y su presencia permite el crecimiento y la recuperación de células que expresan convenientemente las secuencias introducidas. Los clones resistentes de células transformadas de forma estable se pueden propagar utilizando técnicas de cultivo tisular apropiadas para el tipo de célula.

La proteína producida por una célula transformada se puede secretar o retener intracelularmente, dependiendo de la secuencia y/o del vector utilizado. Como bien entenderán los expertos en la técnica, los vectores de expresión que contienen polinucleótidos que codifican BNO1 se pueden diseñar para que contengan secuencias señal que dirigen la secreción de BNO1 a través de una membrana celular procariota o eucariota.

Además, una cepa de células hospedadoras se puede seleccionar por su capacidad para modular la expresión de las secuencias insertadas o para procesar la proteína expresada de la manera deseada. Tales modificaciones del polipéptido incluyen, pero no se limitan a, acetilación, glicosilación, fosforilación y acilación. La escisión posttraduccional de una forma "prepro" de la proteína, también se puede usar para direccionar, plegar y/o para la actividad de una proteína específica. Diferentes células hospedadoras que tienen una maquinaria celular específica y mecanismos característicos para las actividades posttraduccionales (por ejemplo, células HeLa o CHO), están disponibles en la American Type Culture Collection (ATCC) y se pueden seleccionar para asegurar la modificación y el procesamiento correctos de la proteína extraña.

Cuando se necesitan grandes cantidades de BNO1 como para la producción de anticuerpos, se pueden utilizar vectores que dirigen niveles elevados de expresión de BNO1 como los que contienen el promotor inducible del bacteriófago T5 y T7. La presente invención también incluye el uso de los sistemas de expresión descritos anteriormente para la generación y el aislamiento de proteínas de fusión que contienen dominios funcionales importantes de la proteína. Estas proteínas de fusión se utilizan para estudios de ligamiento, estructurales y funcionales, así como para la generación de anticuerpos apropiados.

Con el fin de expresar y purificar la proteína como una proteína de fusión, la secuencia apropiada de ADNc de BNO1 se inserta en un vector que contiene una secuencia de nucleótidos que codifica otro péptido (por ejemplo, la glutationina succinil transferasa). La proteína de fusión se expresa y se recupera a partir de células procariotas o eucariotas. La proteína de fusión se puede purificar a continuación mediante cromatografía de afinidad basada en la secuencia del vector de fusión y la proteína de BNO1 obtenida mediante escisión enzimática de la proteína de fusión.

Los fragmentos de BNO1 también se pueden producir por síntesis directa de péptidos utilizando técnicas de fase sólida. La síntesis automatizada se puede lograr mediante el uso del sintetizador de péptidos ABI 431A (Perkin-Elmer). Varios fragmentos de BNO1 se pueden sintetizar por separado y después se combinan para producir la molécula de longitud completa.

De acuerdo con la invención, se proporciona un polipéptido aislado que comprende la secuencia de aminoácidos expuesta en SEQ ID Números: 2 o 4.

De acuerdo con un aspecto adicional de la invención, se proporciona un polipéptido aislado que comprende la secuencia de aminoácidos expuesta en SEQ ID Números: 2 o 4, que forma parte de un complejo de ubicitina-ligasa involucrado en la degradación de proteínas a través de la ubicitinación.

La invención también incluye un polipéptido aislado que forma parte de un complejo de ubicitina-ligasa implicado en la degradación de proteínas a través de la ubicitinación que tiene al menos 95% de identidad con la secuencia de aminoácidos expuesta en SEQ ID Números: 2 o 4.

Preferiblemente, la identidad de secuencia se determina utilizando el algoritmo BLASTP con la matriz por defecto BLOSSUM62.

También está previsto un polipéptido aislado que consiste en la secuencia de aminoácidos expuesta en SEQ ID Números: 2 o 4.

En un aspecto adicional de la invención, se proporciona un método de preparación de un polipéptido tal y como se ha descrito anteriormente, que comprende las etapas de:

- (1) cultivar las células hospedadoras en condiciones eficaces para la producción del polipéptido; y
- (2) recoger el polipéptido.

La proteína de BNO1 sustancialmente purificada o fragmentos de la misma se pueden utilizar a continuación en otros

análisis bioquímicos para establecer la estructura secundaria y terciaria, por ejemplo, mediante cristalografía de rayos X de la proteína de BNO1 o mediante resonancia magnética nuclear (RMN). La determinación de la estructura permite el diseño racional de productos farmacéuticos que interaccionan con la proteína, alteran la configuración de la carga de la proteína o la interacción de la carga con otras proteínas, o modifican su función en la célula.

- 5 El gen BNO1 se ha identificado a partir de una región de LOH restringida, observada en el cáncer de mama y de próstata, y parece estar regulado a la baja en su expresión en líneas celulares de cáncer obtenidas a partir de estos tejidos. Además, existe una similitud estructural y química en el contexto de las secuencias y motivos, entre las regiones de BNO1 y las proteínas F-box. Las proteínas F-box son los componentes de reconocimiento del sustrato de una clase de las ligasas ubicuitina-E3, la clase denominada "SCF", que están implicadas en la degradación de proteínas a través de la ubicuitinación y la posterior proteólisis llevada a cabo por el proteasoma. Hasta la fecha, las proteínas que parecen estar reguladas por este mecanismo, incluyen oncogenes, genes supresores de tumores, factores de transcripción y otras moléculas de señalización. Estas proteínas influyen en muchos procesos celulares tales como la modulación de las respuestas inmunes e inflamatorias, el desarrollo y la diferenciación, así como procesos que están implicados en el desarrollo del cáncer, tales como la regulación del ciclo celular y la apoptosis. También se ha observado que BNO1 interacciona con Skp1, un componente esencial de las ligasas ubicuitina-E3 SCF.

20 Se ha proporcionado previamente un precedente convincente de una proteína supresora tumoral que pertenece al sistema de degradación ubicuitina-proteasoma, a través del gen VHL. Se ha demostrado que este gen se asocia con elongina C, elongina B y culina-2 en un complejo denominado VCB-CUL-2. Este complejo multiproteico muestra una similitud estructural y funcional con las ligasas de ubicuitina SCF y se ha mostrado que está involucrado en la ubicuitinación de sustratos de VHL.

25 En conjunto, esta información sugiere que BNO1 está implicado en los procesos que conducen al cáncer, especialmente carcinoma de mama y de próstata, lo más probable a través de su papel en la ubicuitinación de proteínas que participan en importantes funciones celulares, tales como la regulación del ciclo celular. Como BNO1 se expresa en muchos tipos de tejidos, alteraciones en la función de BNO1 también pueden causar patologías en esos tejidos a través de las siguientes alteraciones en el proceso de ubicuitinación.

30 Con la identificación de la secuencia de nucleótidos y proteínas de BNO1, se pueden utilizar sondas y anticuerpos dirigidos al gen en una variedad de ensayos de hibridación e inmunológicos, para escrutar y detectar la presencia de un gen normal o mutado o del producto génico. Además, la secuencia de nucleótidos y proteínas del gen BNO1 proporcionada en esta invención, permite métodos terapéuticos para el tratamiento de todas las enfermedades asociadas con anormalidades de la función de BNO1, incluyendo cáncer, enfermedad inmune/inflamatoria y trastornos neurológicos, y también permite métodos para el diagnóstico o el pronóstico de todas las enfermedades asociadas con anormalidades en la función de BNO1.

35 Ejemplos de tales trastornos incluyen, pero no se limitan a, cánceres, trastornos inmunes/inflamatorios y trastornos neurológicos. Los cánceres incluyen adenocarcinoma, leucemia, linfoma, melanoma, mieloma, sarcoma, teratocarcinoma y, en particular, cánceres de mama, próstata, hígado, ovario, cabeza y cuello, corazón, cerebro, páncreas, pulmón, músculo esquelético, riñón, colon, útero, testículos, glándula suprarrenal, sangre, células germinales, placenta, membrana sinovial, amígdalas, cuello uterino, tejido linfático, piel, vejiga, médula espinal, glándula tiroides y estómago. Otros tipos de cáncer pueden incluir los de huesos, médula ósea, vesícula biliar, ganglios, tracto gastrointestinal, paratiroides, pene, glándulas salivales, bazo y timo. Los trastornos inmunes/inflamatorios incluyen el síndrome de inmunodeficiencia adquirida (SIDA), la enfermedad de Addison, el síndrome de dificultad respiratoria de adulto, alergias, espondilitis anquilosante, amiloidosis, anemia, asma, aterosclerosis, anemia hemolítica autoinmune, tiroiditis autoinmune, polienodocrinopatía-candidiasis-distrofia ectodérmica autoinmune (APECED), bronquitis, colecistitis, dermatitis de contacto, enfermedad de Crohn, fibrosis quística, dermatitis atópica, dermatomiositis, diabetes mellitus, enfisema, linfopenia episódica con linfocitotoxinas, eritroblastosis fetal, eritema nodoso, gastritis atrófica, glomerulonefritis, síndrome de Goodpasture, gota, enfermedad de Graves, tiroiditis de Hashimoto, hipereosinofilia, síndrome de intestino irritable, esclerosis múltiple, miasenia grave, inflamación miocárdica o pericárdica, osteoartritis, osteoporosis, pancreatitis, polimiositis, psoriasis, síndrome de Reiter, artritis reumatoide, esclerodermia, síndrome de Sjogren, anafilaxia sistémica, lupus eritematoso sistémico, esclerosis sistémica, púrpura trombocitopénica, colitis ulcerosa, uveítis, síndrome de Werner, complicaciones de la cicatrización de heridas (por ejemplo, formación de cicatrices), cáncer, hemodiálisis y circulación extracorpórea, infecciones víricas, bacterianas, fúngicas, parasitarias, de protozoos y helmintos, y trauma. Los trastornos neurológicos pueden incluir la enfermedad de Parkinson y la enfermedad de Alzheimer.

40 En el tratamiento de enfermedades asociadas con una disminución de la expresión y/o de la actividad de BNO1, es deseable incrementar la expresión y/o la actividad de BNO1. En el tratamiento de trastornos asociados con el aumento de la expresión y/o de la actividad de BNO1, es deseable disminuir la expresión y/o la actividad de BNO1.

55 Mejora del gen BNO1 o de la función proteica de BNO1

Mejorar, estimular o reactivar el gen BNO1 o la función proteica de BNO1 se puede lograr en una variedad de maneras. También describimos la administración de una molécula de ADN aislada, tal y como se ha descrito anteriormente, a un sujeto que requiere que tal tratamiento se pueda iniciar.

Típicamente, BNO1 se administra a un sujeto para tratar o prevenir un trastorno asociado con la disminución de la actividad y/o la expresión de BNO1.

En un aspecto adicional, se proporciona el uso de una molécula de ADN aislada, tal y como se ha descrito anteriormente, en la preparación de un medicamento para el tratamiento de un trastorno asociado con una disminución de la actividad y/o la expresión de BNO1.

Típicamente, un vector capaz de expresar BNO1 o un fragmento o un derivado del mismo, se puede administrar a un sujeto para tratar o prevenir un trastorno asociado con una disminución de la actividad y/o la expresión de TSG18 incluyendo los descritos anteriormente, pero no limitado a los mismos. La transducción de vectores retrovíricos se utiliza frecuentemente para la terapia génica de células somáticas, debido a su elevada eficacia en la infección y la integración y la expresión estables. El gen BNO1 de longitud completa, o porciones del mismo, se pueden clonar en un vector retrovírico y se pueden dirigir desde su promotor endógeno o desde la repetición terminal larga retrovírica o desde un promotor específico para el tipo de célula diana de interés. Otros vectores víricos se pueden utilizar e incluyen, tal y como se conocen en la técnica, adenovirus, virus adenoasociado, virus vaccinia, papovavirus, lentivirus y retrovirus de origen aviar, murino y humano.

La terapia génica se llevaría a cabo según los métodos establecidos (Friedman, 1991; Culver, 1996). Se preparara un vector que contiene una copia del gen BNO1 ligada a elementos de control de la expresión y que es capaz de replicarse dentro de las células. Alternativamente, el vector puede tener una replicación incorrecta y puede requerir células auxiliares o un virus auxiliar para la replicación y la producción de virus y el uso en terapia génica.

La transferencia génica que emplea métodos no víricos de infección también se puede utilizar. Estos métodos incluyen la inyección directa de ADN, la captación de ADN desnudo en presencia de fosfato de calcio, la electroporación, la fusión de protoplastos o la entrega en liposomas. La transferencia génica también se puede lograr mediante la entrega como una parte de un cromosoma humano artificial o la transferencia génica mediada por receptores. Esto implica ligar el ADN a una molécula de direccionamiento que se unirá a receptores específicos de la superficie celular para inducir la endocitosis y la transferencia del ADN dentro de células de mamífero. Una de tales técnicas utiliza poli-L-lisina para enlazar la asialoglicoproteína al ADN. Un adenovirus también se añade al complejo para destruir los lisosomas y permitir de este modo que el ADN evite la degradación y pase al núcleo. La infusión de estas partículas por vía intravenosa ha dado lugar a la transferencia génica en hepatocitos.

En sujetos afectados que expresan una forma mutada de BNO1, se puede prevenir el trastorno introduciendo en las células afectadas una copia de tipo silvestre del gen, de tal manera que se recombinan con el gen mutante. Esto requiere un caso de doble recombinación para la corrección de la mutación del gen. Los vectores para la introducción de genes de esta forma son conocidos en la técnica, y se puede utilizar cualquier vector adecuado. Alternativamente, la introducción de otra copia del gen que es portadora de una segunda mutación en ese gen, se puede emplear con el fin de anular la mutación del gen original y bloquear cualquier efecto negativo.

En sujetos afectados que tienen una expresión reducida de BNO1, un mecanismo de regulación a la baja puede ser la metilación anormal de la isla CpG presente en el extremo 5' del gen. Por lo tanto, en un enfoque alternativo de la terapia, la administración de agentes que eliminan la metilación del promotor de BNO1, reactivará la expresión del gen BNO1 y puede suprimir el fenotipo asociado con la enfermedad.

En un aspecto adicional, un agonista adecuado también puede incluir una molécula o un péptido pequeño que puede imitar la función de BNO1 de tipo silvestre.

40 Inhibición del gen BNO1 o de la función proteica de BNO1

La inhibición de la función de un gen o una proteína mutada se puede lograr por una variedad de vías. También describimos un método para tratar un trastorno asociado con el aumento de la actividad y/o la expresión de BNO1, que comprende administrar un antagonista de BNO1 a un sujeto que requiera tal tratamiento.

En todavía otro aspecto de la invención, se proporciona el uso de un antagonista de BNO1 en la preparación de un medicamento para el tratamiento de un trastorno asociado con el aumento de la actividad y/o de la expresión de BNO1.

Tales trastornos pueden incluir los descritos anteriormente, pero no se limitan a los mismos. En un aspecto de la invención, una molécula de ADN aislada, que es el complemento de una cualquiera de las moléculas de ADN descritas anteriormente y que codifica una molécula de ARN que se hibrida con el ARNm codificado por BNO1, se puede administrar a un sujeto que requiera tal tratamiento.

En aún otro aspecto de la invención, se proporciona el uso de una molécula de ADN aislada que es el complemento de una molécula de ADN de la invención y que codifica una molécula de ARN que se hibrida con el ARNm codificado por BNO1, en la preparación de un medicamento para el tratamiento de un trastorno asociado con el aumento de la actividad y/o de la expresión de BNO1.

Típicamente, un vector que expresa el complemento del polinucleótido que codifica BNO1 se puede administrar a un sujeto para tratar o prevenir un trastorno asociado con el aumento de la actividad y/o la expresión de BNO1, que incluye

- los descritos anteriormente, pero no se limita a los mismos. Las estrategias no codificadoras (del inglés, "antisense strategies") pueden utilizar una variedad de enfoques incluyendo el uso de oligonucleótidos no codificadores, ribozimas, ADNzimas, inyección de ARN no codificador y transfección de vectores de expresión de ARN no codificador. Muchos métodos para introducir vectores en células o tejidos están disponibles y son igualmente adecuados para su uso *in vivo*, *in vitro* y *ex vivo*. Para la terapia *ex vivo*, los vectores se pueden introducir en células madre, excepto las células madre embrionarias humanas, tomadas del paciente y propagadas por clonación para un retrotransplante autólogo en el mismo paciente. La entrega mediante transfección, inyecciones de liposomas o mediante polímeros amino poliacilónicos, se puede lograr usando métodos que son bien conocidos en la técnica. (Por ejemplo, véase Goldman *et al.*, 1997).
- También describimos un método para tratar un trastorno asociado con un aumento de la actividad y/o de la expresión de BNO1 que comprende la administración de un antagonista de BNO1 a un sujeto que requiera tal tratamiento.
- En todavía otro aspecto de la invención, se proporciona el uso de un antagonista de BNO1 en la preparación de un medicamento para el tratamiento de un trastorno asociado con el aumento de la actividad y/o la expresión de BNO1.
- Tales trastornos pueden incluir, los descritos anteriormente, pero no se limitan a los mismos. En un aspecto, una proteína purificada de acuerdo con la invención se puede usar para producir anticuerpos que se unen específicamente a BNO1. Estos anticuerpos se pueden usar directamente como antagonistas o indirectamente como un mecanismo de direccionamiento o de entrega para llevar un agente farmacéutico a las células o tejidos que expresan BNO1. Tales anticuerpos pueden incluir, pero no se limitan a los mismos, anticuerpos policlonales, monoclonales, químéricos y de cadena sencilla, tal como lo entenderá una persona experta en la técnica.
- Para la producción de anticuerpos, diversos hospedadores que incluyen conejos, ratas, cabras, ratones, seres humanos y otros se pueden inmunizar mediante la inyección de una proteína de la invención o de cualquier fragmento u oligopéptido de la misma, que tiene propiedades inmunogénicas. Se pueden utilizar diversos adyuvantes para aumentar la respuesta inmunológica e incluyen, pero no se limitan a los mismos, adyuvante de Freund, geles minerales tales como hidróxido de aluminio y sustancias tensioactivas tales como lisolecitina. Los adyuvantes utilizados en los seres humanos incluyen BCG (bacilos de Calmette-Guerin) y Corynebacterium parvum.
- Se prefiere que los oligopéptidos, péptidos o fragmentos utilizados para inducir anticuerpos frente a BNO1, tengan una secuencia de aminoácidos que consiste en al menos aproximadamente 5 aminoácidos y, más preferiblemente, en al menos aproximadamente 10 aminoácidos. También es preferible que estos oligopéptidos, péptidos o fragmentos sean idénticos a una porción de la secuencia de aminoácidos de la proteína natural y que contengan la secuencia de aminoácidos completa de una pequeña molécula de origen natural. Tramos cortos de aminoácidos procedentes de estas proteínas se pueden fusionar con los de otra proteína, tal como KLH, y se pueden producir anticuerpos de la molécula químérica.
- Los anticuerpos monoclonales de BNO1 se pueden preparar usando cualquier técnica que proporcione la producción de moléculas de anticuerpo mediante líneas celulares continuas en cultivo. Estas incluyen, pero no se limitan a las mismas, la técnica del hibridoma, la técnica del hibridoma de linfocitos B humanos y la técnica de hibridoma de VEB. (Por ejemplo, véase Kohler *et al.*, 1975; Kozbor *et al.*, 1985; Cote *et al.*, 1983; Cole *et al.*, 1984).
- Los anticuerpos también se pueden producir induciendo la producción *in vivo* en la población de linfocitos o mediante el escrutinio de genotecas de inmunoglobulinas o paneles de reactivos que se unen de forma muy específica, tal y como se describe en la bibliografía. (Por ejemplo, véase Orlandi *et al.*, 1989; Winter *et al.*, 1991).
- Los fragmentos de anticuerpo que contienen sitios de unión específica a BNO1 también se pueden generar. Por ejemplo, tales fragmentos incluyen, fragmentos F(ab')₂ producidos por digestión con pepsina de la molécula de anticuerpo y fragmentos Fab generados por reducción de los puentes disulfuro de los fragmentos F(ab')₂. Alternativamente, las genotecas de expresión de Fab se pueden construir para permitir una rápida y fácil identificación de fragmentos Fab monoclonales con la especificidad deseada. (Por ejemplo, véase Huse *et al.*, 1989).
- Se pueden usar varios inmunoensayos para escrutar e identificar anticuerpos que tienen la especificidad deseada. Numerosos protocolos para ensayos de unión competitiva o inmunoradiométricos que utilizan anticuerpos policlonales o monoclonales con especificidades establecidas, son bien conocidos en la técnica. Tales inmunoensayos implican típicamente medir la formación de complejos entre una proteína y su anticuerpo específico. Se prefiere un inmunoensayo de dos sitios basado en anticuerpos monoclonales, que utiliza anticuerpos monoclonales que reaccionan con dos epítopos que no interfieren, pero también se puede emplear un ensayo de unión competitiva.
- También describimos un método para el tratamiento de un trastorno que se ha mostrado que está asociado con una actividad y/o una expresión anormal de BNO1, que comprende la administración de una molécula de ácido nucleico, un anticuerpo o un compuesto tal y como se ha descrito anteriormente, a un sujeto que requiera tal tratamiento.
- En otro aspecto, la invención proporciona el uso de una molécula de ácido nucleico, un anticuerpo o un compuesto tal y como se ha descrito anteriormente, en la preparación de un medicamento para el tratamiento de un trastorno que se ha mostrado que está asociado con una actividad y/o una expresión anormal de BNO1.
- En un aspecto adicional, se puede administrar una composición farmacéutica que comprende una molécula de ácido

nucleico, un anticuerpo o un compuesto tal y como se ha descrito anteriormente, y un vehículo farmacéuticamente aceptable.

- La composición farmacéutica se puede administrar a un sujeto para tratar o prevenir un trastorno asociado con una actividad y/o una expresión anormal de BNO1 que incluye los proporcionados anteriormente pero no se limitan a los mismos. Las composiciones farmacéuticas de acuerdo con la presente invención se preparan mezclando BNO1 o fragmentos activos de BNO1 o variantes de los mismos que tienen el grado de pureza deseado, con vehículos, excipientes o estabilizadores aceptables que son bien conocidos. Los vehículos, excipientes o estabilizadores aceptables no son tóxicos a las dosificaciones y concentraciones empleadas, e incluyen tampones tales como fosfato, citrato y otros ácidos orgánicos; antioxidantes que incluyen ácido ascórbico; polipéptidos de bajo peso molecular (menos de aproximadamente 10 residuos); proteínas, tales como albúmina de suero, gelatina o inmunoglobulinas; polímeros hidrófilos tales como polivinilpirrolidona; aminoácidos tales como glicina, glutamina, asparagina, arginina o lisina; monosacáridos, disacáridos y otros carbohidratos, que incluyen glucosa, manosa o dextrinas; agentes quelantes tales como EDTA; azúcar alcoholes tales como manitol o sorbitol; contraiones formadores de sales tales como sodio; y/o tensioactivos no iónicos tales como Tween, Pluronics o polietilenglicol (PEG).
- En otras realizaciones, cualquiera de los genes, péptidos, antagonistas, anticuerpos, secuencias complementarias o vectores de la invención se pueden administrar en combinación con otros agentes terapéuticos apropiados. La selección de los agentes adecuados la pueden realizar los expertos en la técnica, de acuerdo con principios farmacéuticos convencionales. La combinación de agentes terapéuticos puede actuar sinéricamente para efectuar el tratamiento o la prevención de los diversos trastornos descritos anteriormente. Con el uso de este enfoque, puede ser posible una eficacia terapéutica con dosificaciones más bajas de cada agente, reduciendo de este modo el potencial de efectos secundarios adversos.

Detección de fármacos

De acuerdo con todavía otro aspecto de la invención, los péptidos de la invención, particularmente polipéptidos de BNO1 purificados y células que los expresan, son útiles para la detección de agentes farmacéuticos candidatos en una variedad de técnicas para el tratamiento de trastornos asociados con una disfunción de BNO1. Tales técnicas incluyen, pero no se limitan a las mismas, la utilización de células hospedadoras eucariotas o procariotas que se transforman de forma estable con moléculas recombinantes que expresan el polipéptido de BNO1 o un fragmento del mismo, preferiblemente en ensayos de unión competitiva. Los ensayos de unión medirán la formación de complejos entre el polipéptido de BNO1, o fragmentos del mismo, y el agente que se está sometiendo a ensayo, o medirán el grado con el que un agente que se está sometiendo a ensayo va a interferir con la formación de un complejo entre el polipéptido de BNO1, o un fragmento del mismo, y un ligando conocido, en particular otros miembros del complejo SCF y sustratos de BNO1 direccionalmente para la ubicuitinación.

Otra técnica para escrutar fármacos proporciona una detección rápida de compuestos que tienen una afinidad de unión adecuada hacia el polipéptido de BNO1 (véase el documento de solicitud PCT publicada WO84/03564). En esta técnica mencionada, se puede sintetizar un gran número de pequeños compuestos peptídicos del ensayo sobre un sustrato sólido y se pueden analizar a través de la unión al polipéptido de BNO1 y lavar. El polipéptido de BNO1 unido se detecta a continuación por métodos bien conocidos en la técnica. En una variación de esta técnica, se pueden revestir placas directamente con los polipéptidos purificados para identificar los compuestos del ensayo que interaccionan.

Un método adicional para seleccionar fármacos implica el uso de líneas de células hospedadoras eucariotas que son portadoras de mutaciones en el gen BNO1. Las líneas de células hospedadoras también son defectuosas a nivel de polipéptido. Se pueden utilizar otras líneas celulares en las que la expresión génica de BNO1 se puede desconectar. Las líneas de células hospedadoras o las células se cultivan en presencia de diversos compuestos farmacéuticos y se mide la tasa de crecimiento de las células hospedadoras para determinar si el compuesto es capaz de regular el crecimiento de las células defectuosas.

El polipéptido de BNO1 también se puede usar para seleccionar compuestos obtenidos como resultado de la tecnología combinatoria de genotecas. Esta proporciona una forma de someter a ensayo un gran número de sustancias diferentes para estudiar su capacidad para modular la actividad de un polipéptido. Se prefiere el uso de bancos de péptidos (véase el documento de patente WO97/02048) con tales bancos y su uso conocido en la técnica.

Una sustancia identificada como moduladora de la función del polipéptido puede ser un péptido o un no péptido en la naturaleza. Las "moléculas pequeñas" no peptídicas se prefieren frecuentemente para muchas aplicaciones farmacéuticas *in vivo*. Además, un imitador o un agente mimético de la sustancia se puede diseñar para uso farmacéutico. El diseño de agentes miméticos basándose en un compuesto farmacéuticamente activo conocido (compuesto "guía") es un método común para el desarrollo de nuevos productos farmacéuticos. Esto es deseable frecuentemente cuando el compuesto activo original es difícil de sintetizar o costoso, o cuando se proporciona un método inadecuado de administración. En el diseño de un agente mimético, se identifican partes particulares del compuesto activo original que son importantes para la determinación de la propiedad diana. Estas partes o residuos que constituyen la región activa del compuesto se conocen como su farmacóforo. Una vez encontrado, la estructura del farmacóforo sirve de modelo de acuerdo con sus propiedades físicas, utilizando datos que proceden de una variedad de fuentes, incluyendo los datos de difracción de rayos X y RMN. A continuación, se selecciona una molécula molde, a la que se pueden añadir grupos químicos

que imitan el farmacóforo. La selección puede hacerse de manera que el agente mimético sea fácil de sintetizar, sea probable que sea farmacológicamente aceptable, no se degrade *in vivo* y conserve la actividad biológica del compuesto guía. Una mejora o una modificación adicional se puede llevar a cabo para seleccionar uno o varios agentes miméticos finales, útiles para el ensayo *in vivo* o clínico.

- 5 También es posible aislar un anticuerpo específico de una diana y después explicar su estructura cristalina. En principio, este método produce un farmacóforo a partir del cual se puede basar el diseño de fármacos posteriores, tal y como se describió anteriormente. Puede ser posible evitar por completo la cristalográfica de proteínas, generando anticuerpos anti-idiotípicos (anti-ids) de un anticuerpo funcional, farmacológicamente activo. Como imagen especular de una imagen 10 especular, se esperaría que el sitio de unión de los anti-ids fuera un análogo del sitio de unión original. El anti-id se podría usar entonces para aislar péptidos a partir de bancos de péptidos producidos química o biológicamente.

Cualquiera de los métodos terapéuticos descritos anteriormente se puede aplicar a cualquier sujeto que requiera una terapia de este tipo, incluyendo, por ejemplo, mamíferos tales como perros, gatos, vacas, caballos, conejos, monos y, lo más preferiblemente, seres humanos.

Aplicaciones para diagnóstico y pronóstico

- 15 Las secuencias de polinucleótidos que codifican BNO1 se pueden utilizar para el diagnóstico o el pronóstico *in vitro* de trastornos asociados con una disfunción de BNO1, o una predisposición a tales trastornos. Ejemplos de tales trastornos incluyen, pero no se limitan a los mismos, cánceres, trastornos del sistema inmune/inflamatorio y trastornos neurológicos. Los cánceres incluyen adenocarcinoma, leucemia, linfoma, melanoma, mieloma, sarcoma, teratocarcinoma y, en particular, cánceres de mama, próstata, hígado, ovario, cabeza y cuello, corazón, cerebro, páncreas, pulmón, músculo 20 esquelético, riñón, colon, útero, testículos, glándula suprarrenal, sangre, células germinales, placenta, membrana sinovial, amígdalas, cuello uterino, tejido linfático, piel, vejiga, médula espinal, glándula tiroideas y estómago. Otros cánceres pueden incluir los de hueso, médula ósea, vesícula biliar, ganglios, tracto gastrointestinal, paratiroides, pene, glándulas salivales, bazo y timo. Los trastornos inmuno/inflamatorios incluyen el síndrome de inmunodeficiencia adquirida (SIDA), la enfermedad de Addison, el síndrome de dificultad respiratoria de adulto, alergias, espondilitis anquilosante, amiloidosis, anemia, asma, aterosclerosis, anemia hemolítica autoinmune, tiroiditis autoinmune, polienodocrinopatía-candidiasis-distrofia ectodérmica autoinmune (APECED), bronquitis, colecistitis, dermatitis de contacto, enfermedad de Crohn, fibrosis quística, dermatitis atópica, dermatomiositis, diabetes mellitus, enfisema, linfopenia episódica con linfocitotoxinas, eritroblastosis fetal, eritema nodoso, gastritis atrófica, glomerulonefritis, síndrome de Goodpasture, gota, enfermedad de Graves, tiroiditis de Hashimoto, hipereosinofilia, síndrome de intestino irritable, esclerosis múltiple, miastenia grave, 25 inflamación miocárdica o pericárdica, osteoartritis, osteoporosis, pancreatitis, polimiositis, psoriasis, síndrome de Reiter, artritis reumatoide, esclerodermia, síndrome de Sjogren, anafilaxia sistémica, lupus eritematoso sistémico, esclerosis sistémica, púrpura trombocitopénica, colitis ulcerosa, uveítis, síndrome de Werner, complicaciones de la cicatrización de heridas (por ejemplo, formación de cicatrices), cáncer, hemodiálisis y circulación extracorpórea, infecciones víricas, bacterianas, fúngicas, parasitarias, de protozoos y helmintos, y trauma. Los trastornos neurológicos pueden incluir la 30 enfermedad de Parkinson y la enfermedad de Alzheimer.

El diagnóstico o el pronóstico se pueden usar para determinar la gravedad, el tipo o el estadio del estado de la enfermedad con el fin de iniciar una intervención terapéutica apropiada.

- En otra realización de la invención, los polinucleótidos que se pueden utilizar con fines de diagnóstico o pronóstico incluyen secuencias de oligonucleótidos, ADN genómico y moléculas de ADN y ARN complementarias. Los polinucleótidos 40 se pueden utilizar para detectar y cuantificar la expresión génica en tejidos de biopsias en los que mutaciones en BNO1 o una expresión anormal de BNO1 se pueden correlacionar con una enfermedad. El ADN genómico utilizado para el diagnóstico o el pronóstico se puede obtener a partir de células corporales, tales como las presentes en la sangre, biopsia de tejidos, muestra quirúrgica o material de autopsias. El ADN se puede aislar y usar directamente para la detección de una secuencia específica o se puede amplificar mediante la reacción en cadena de la polimerasa (PCR) antes del análisis. Del mismo modo, el ARN o ADNc también se pueden utilizar, con o sin amplificación mediante PCR. Para detectar una secuencia específica de ácido nucleico, se puede emplear la secuenciación directa de nucleótidos, la PCR con transcriptasa inversa (RT-PCR), la hibridación usando oligonucleótidos específicos, la digestión con enzimas de restricción y cartografía, la cartografía mediante PCR, la protección contra ARNasa y varios otros métodos. Los oligonucleótidos específicos de secuencias particulares se pueden sintetizar químicamente y marcar radiactivamente o no 45 radiactivamente e hibridar con muestras individuales, inmovilizadas sobre membranas u otros soportes sólidos o en solución. La presencia, la ausencia o el exceso de expresión de BNO1 se pueden visualizar a continuación usando métodos tales como autorradiografía, fluorometría o colorimetría.

- En un aspecto particular, las secuencias de nucleótidos que codifican BNO1 pueden ser útiles en ensayos que detectan 55 la presencia de trastornos asociados, en particular los mencionados anteriormente. Las secuencias de nucleótidos que codifican BNO1 se pueden marcar por métodos convencionales y añadir a una muestra de fluido o de tejido de un paciente, en condiciones adecuadas para la formación de complejos de hibridación. Después de un periodo de incubación adecuado, la muestra se lava y la señal se cuantifica y se compara con un valor estándar. Si la cantidad de señal en la muestra del paciente está significativamente alterada en comparación con una muestra testigo, entonces la presencia de niveles alterados de secuencias de nucleótidos que codifican BNO1 en la muestra indica la presencia del trastorno 60 asociado. Tales ensayos también se pueden usar para evaluar la eficacia de un régimen de tratamiento terapéutico

particular en estudios con animales, en ensayos clínicos o para vigilar el tratamiento de un paciente individual.

Con el fin de proporcionar una base para el diagnóstico o el pronóstico de un trastorno que se ha mostrado que está asociado con una mutación en BNO1, la secuencia de nucleótidos del gen BNO1 se puede comparar entre tejido normal y tejido enfermo con el fin de establecer si el paciente expresa un gen mutante.

- 5 Con el fin de proporcionar una base para el diagnóstico o el pronóstico de un trastorno que se ha mostrado que está asociado con una expresión anormal de BNO1, se establece un perfil normal o estándar de la expresión. Esto se puede lograr combinando fluidos corporales o extractos celulares tomados de sujetos normales, tanto animales como humanos, con una secuencia, o un fragmento de la misma, que codifica BNO1, en condiciones adecuadas para la hibridación o la amplificación. La hibridación convencional se puede cuantificar mediante la comparación de los valores obtenidos a partir de sujetos normales con los valores procedentes de un experimento en el que se utiliza una cantidad conocida de un polinucleótido sustancialmente purificado. Otro método para identificar un perfil normal o estándar de la expresión de BNO1 es a través de estudios con RT-PCR cuantitativa. El ARN aislado a partir de células corporales de un individuo normal, particularmente el ARN aislado a partir de células tumorales, se transcribe de forma inversa y se realiza una PCR en tiempo real, utilizando oligonucleótidos específicos del gen BNO1, para establecer un nivel normal de expresión del gen.

Los valores estándar obtenidos en estos dos ejemplos se pueden comparar con los valores obtenidos a partir de muestras de pacientes que son sintomáticos de un trastorno. Una desviación de los valores estándar se usa para establecer la presencia de un trastorno.

- 20 Una vez que se ha establecido la presencia de un trastorno y se inicia un protocolo de tratamiento, los ensayos de hibridación o los estudios con RT-PCR cuantitativa se pueden repetir regularmente para determinar si el nivel de expresión en el paciente comienza a aproximarse al que se observa en el sujeto normal. Los resultados obtenidos a partir de sucesivos ensayos se pueden usar para mostrar la eficacia del tratamiento durante un período que oscila desde varios días a meses.

- 25 En un aspecto, la hibridación con sondas de PCR que son capaces de detectar secuencias de polinucleótidos, incluyendo secuencias genómicas, que codifican BNO1 o moléculas estrechamente relacionadas, se puede utilizar para identificar secuencias de ácidos nucleicos que codifican BNO1. La especificidad de la sonda, tanto si se prepara a partir de una región altamente específica, por ejemplo, la región reguladora 5', o a partir de una región menos específica, por ejemplo, un motivo conservado, y el rigor de la hibridación o la amplificación, determinarán si la sonda identifica solo secuencias de origen natural que codifican BNO1, variantes alélicas o secuencias relacionadas.

- 30 Las sondas también se pueden utilizar para la detección de secuencias relacionadas y deben tener preferiblemente al menos 50% de identidad de secuencia con cualquiera de las secuencias que codifican BNO1. Las sondas para la hibridación de la presente invención pueden ser ADN o ARN, y se pueden obtener a partir de la secuencia de SEQ ID Números: 1 o 3, o a partir de secuencias genómicas que incluyen promotores, potenciadores e intrones del gen BNO1 (SEQ ID Números: 5-11).

- 35 Los medios para producir sondas de hibridación específicas para los ADNs que codifican BNO1, incluyen la clonación de secuencias de polinucleótidos que codifican BNO1 o derivados de BNO1 en vectores, para la producción de sondas de ARNm. Tales vectores son conocidos en la técnica y están disponibles comercialmente. Las sondas de hibridación pueden estar marcadas con radionucleidos tales como ^{32}P o ^{35}S , o con marcadores enzimáticos, tales como fosfatasa alcalina acoplada a la sonda a través de sistemas de acoplamiento de avidina/biotina u otros métodos conocidos en la técnica.

40 De acuerdo con un aspecto adicional de la invención, se proporciona el uso de un polipéptido tal y como se ha descrito anteriormente, en el diagnóstico o el pronóstico de un trastorno que se ha mostrado que está asociado con BNO1, o una predisposición a tales trastornos.

- 45 Cuando un ensayo de diagnóstico o pronóstico se ha de basar en la proteína de BNO1, son posibles una variedad de métodos. Por ejemplo, el diagnóstico o el pronóstico se pueden conseguir vigilando las diferencias en la movilidad electroforética de proteínas normales y mutantes. Este método será particularmente útil para identificar mutantes en los que están presentes sustituciones de la carga, o en los que inserciones, delecciones o sustituciones han dado como resultado un cambio significativo en la migración electroforética de la proteína resultante. Alternativamente, el diagnóstico se puede basar en diferencias en los patrones de escisión proteolítica de proteínas normales y mutantes, diferencias en las relaciones molares de los diversos residuos de aminoácidos o mediante ensayos funcionales que muestran la función alterada de los productos génicos.

- 55 En otro aspecto, anticuerpos que se unen específicamente a BNO1 se pueden usar para el diagnóstico o el pronóstico de trastornos caracterizados por una expresión anormal de BNO1, o en ensayos para vigilar pacientes que están siendo tratados con BNO1 o con agonistas, antagonistas o inhibidores de BNO1. Los anticuerpos útiles para fines de diagnóstico se pueden preparar de la misma manera que como se ha descrito anteriormente para los agentes terapéuticos. Los ensayos de diagnóstico o pronóstico de BNO1 incluyen métodos que utilizan el anticuerpo y un marcador para detectar BNO1 en fluidos corporales humanos o en extractos de células o tejidos. Los anticuerpos se pueden usar con o sin modificación y se pueden marcar mediante la fijación covalente o no covalente de una molécula informadora.

- Una variedad de protocolos para medir BNO1, que incluyen ELISAs, RIAs y FACS, son conocidos en la técnica y proporcionan una base para diagnosticar unos niveles alterados o anormales de la expresión de BNO1. Los valores normales o convencionales de expresión de BNO1 se establecen combinando fluidos corporales o extractos de células tomadas de sujetos mamíferos normales, preferiblemente humanos, con un anticuerpo de BNO1 en condiciones adecuadas para la formación de complejos. La cantidad de formación de complejo convencional se puede cuantificar mediante varios métodos, preferiblemente por medios fotométricos. Las cantidades de BNO1 expresadas en muestras de un sujeto, un testigo y en muestras de la enfermedad procedentes de tejidos de biopsias, se comparan con los valores estándar. La desviación entre los valores estándar y los del sujeto establece los parámetros para diagnosticar una enfermedad.
- 10 Una vez que a un individuo se le ha diagnosticado un trastorno, se pueden iniciar los tratamientos efectivos. Estos pueden incluir la administración de un agonista selectivo del BNO1 mutante, a fin de restablecer su función hasta un nivel normal, o la introducción de BNO1 de tipo silvestre, especialmente a través de métodos de terapia génica, tal y como se ha descrito anteriormente. Típicamente, se puede administrar un vector capaz de expresar el gen BNO1 de longitud completa apropiado o un fragmento o un derivado del mismo. En un método alternativo a la terapia, un polipéptido de BNO1 sustancialmente purificado y un vehículo farmacéuticamente aceptable se pueden administrar tal y como se ha descrito anteriormente o se pueden administrar fármacos que pueden sustituir la función de BNO1, o que imitan la acción de BNO1.

En el tratamiento de trastornos que se ha mostrado que están asociados con el aumento de la expresión y/o la actividad de BNO1, el individuo afectado se puede tratar con un antagonista selectivo, tal como un anticuerpo de la proteína correspondiente o una sonda no codificante (complementaria) para el gen correspondiente, tal y como se ha descrito anteriormente, o mediante el uso de fármacos que pueden bloquear la acción de BNO1.

Micromatrizes

En otra realización, ADNc completos, oligonucleótidos o fragmentos más largos obtenidos a partir de cualquiera de las secuencias de polinucleótidos descritas en este documento, se pueden utilizar como dianas en una micromatriz. La micromatriz se puede utilizar para controlar simultáneamente el nivel de expresión de un gran número de genes y para identificar variantes genéticas, mutaciones y polimorfismos. Esta información se puede utilizar para determinar la función génica, para comprender las bases genéticas de un trastorno, para diagnosticar o pronosticar un trastorno y para desarrollar y controlar las actividades de agentes terapéuticos. Las micromatrizes se pueden preparar, utilizar y analizar usando métodos conocidos en la técnica. (Por ejemplo, véase Schena *et al.*, 1996; Heller *et al.*, 1997).

Hospedadores transformados

La presente invención también proporciona la producción de modelos animales no humanos modificados genéticamente (con gen desactivado (del inglés "knock-out"), con inserción de secuencias génicas (del inglés "knock-in") y transgénicos) transformados con las moléculas de ADN de la invención. Estos animales son útiles para el estudio de la función del gen BNO1, para estudiar los mecanismos de una enfermedad en lo que se refiere al gen BNO1, para escrutar compuestos farmacéuticos candidatos, para la creación de cultivos de células de mamífero explantados que expresan la proteína o una proteína mutante y para la evaluación de posibles intervenciones terapéuticas.

El gen BNO1 se puede haber inactivado mediante una delección desactivadora del gen y, por lo tanto, se proporcionan animales no humanos modificados genéticamente con gen desactivado.

Las especies animales que son adecuadas para uso en los modelos animales de la presente invención son ratas, ratones, hámsteres, cobayas, conejos, perros, gatos, cabras, ovejas, cerdos y primates no humanos, tales como monos y chimpancés. Para los estudios iniciales, son muy deseables los ratones y las ratas genéticamente modificados, debido a su mantenimiento relativamente sencillo y a una vida más corta. Para ciertos estudios, levaduras o invertebrados transgénicos pueden ser adecuados y preferidos porque permiten una detección rápida y proporcionan un manejo mucho más fácil. Para estudios a largo plazo, los primates no humanos pueden ser deseables debido a su similitud con los seres humanos.

Para crear un modelo animal con BNO1 mutado, se pueden emplear varios métodos. Estos incluyen la generación de una mutación específica en un gen animal homólogo, la inserción de un gen humano de tipo silvestre y/o un gen animal humanizado por recombinación homóloga, la inserción de un gen humano mutante (única o múltiple) en forma de estructuras artificiales de ADNc genómico o minigén, usando elementos promotores de tipo silvestre, mutantes o artificiales, o la inserción de fragmentos modificados artificialmente del gen endógeno mediante recombinación homóloga. Las modificaciones incluyen la inserción de codones de detención mutantes, la delección de secuencias de ADN o la inclusión de elementos de recombinación (sitios lox p) reconocidos por enzimas tales como la recombinasa Cre.

Para crear un ratón transgénico, el cual se prefiere, una versión mutante de BNO1 se puede insertar en una línea germinal de ratón utilizando técnicas convencionales de microinyección de ovocitos o transfección o microinyección dentro de células madre embrionarias. Alternativamente, si se desea inactivar o reemplazar el gen BNO1 endógeno, se puede aplicar una recombinación homóloga utilizando células madre embrionarias murinas.

Para la inyección de ovocitos, una o varias copias del gen BNO1 mutante o de tipo silvestre se pueden insertar en el

pronúcleo de un ovocito de ratón recién fertilizado. El ovocito se reimplanta después en una madre adoptiva seudoprenada. Los ratones nacidos vivos se pueden seleccionar entonces según los integrantes, utilizando el análisis del ADN de la cola, buscando la presencia de secuencias de genes BNO1 humanos. El transgén puede ser una secuencia genómica completa inyectada como YAC, BAC, PAC o bien otro fragmento de ADN cromosómico, un ADNc ya sea con el promotor natural o con un promotor heterólogo, o un minigén que contiene toda la región codificadora y otros elementos que son necesarios para una expresión óptima.

De acuerdo con todavía otro aspecto de la invención, se proporciona el uso de animales no humanos modificados genéticamente, tal como se ha descrito anteriormente, para el escrutinio de compuestos farmacéuticos candidatos.

La identificación de la secuencia de nucleótidos y de aminoácidos de ambas isoformas de BNO1, permite la identificación de sustratos proteicos específicos de BNO1, utilizando estudios de interacción de proteínas, tales como el análisis del doble híbrido en levadura, como lo entenderán los expertos en la técnica. Estos sustratos proteicos serían dianas para la degradación a través de la ubicuitinación mediada por la ligasa ubicuitina-E3 que contiene BNO1. Cada isoforma de BNO1 puede compartir sustratos proteicos comunes o puede interaccionar con sustratos específicos de la isoforma.

También describimos un complejo de BNO1 de tipo silvestre y un sustrato específico de BNO1 que está destinado a una degradación mediante ubicuitinación.

También describimos un complejo de BNO1 y proteínas del complejo de ligasa ubicuitina-E3

También describimos un complejo de BNO1 de tipo silvestre y la proteína Skp1.

También describimos un ácido nucleico que codifica un polipéptido mutante de BNO1 que no puede formar un complejo con las proteínas de tipo silvestre con las que BNO1 de tipo silvestre forma un complejo. Típicamente, una de estas proteínas es Skp1, mientras que otras son sustratos proteicos específicos de BNO1 destinados a la degradación mediante ubicuitinación.

También describimos un polipéptido mutante de BNO1 que no puede formar un complejo con proteínas de tipo silvestre con las que BNO1 de tipo silvestre forma un complejo. Típicamente, una de estas proteínas es Skp1 mientras que otras son sustratos proteicos específicos de BNO1 destinados a la degradación mediante ubicuitinación.

También describimos el uso de complejos, tal y como se ha descrito anteriormente, en el escrutinio de compuestos farmacéuticos candidatos. También se puede escrutar en busca de un fármaco que sustituya la actividad de BNO1 en un paciente que carece de BNO1.

Se entenderá claramente que, aunque en el presente documento se hace referencia a un número de publicaciones de la técnica anterior, esta referencia no constituye una admisión de que cualquiera de estos documentos forme parte del conocimiento general común en la técnica, en Australia o en cualquier otro país. En toda esta memoria descriptiva y las reivindicaciones, las palabras "comprenden", "comprende" y "que comprende" se utilizan en un sentido no exclusivo, salvo cuando el contexto requiere lo contrario.

Breve descripción de los dibujos

Figura 1. Representación esquemática de tumores con pérdida alélica intersticial y terminal en el brazo del cromosoma 16q en las dos series de muestras tumorales. Los marcadores polimórficos se enumeran de acuerdo con su orden en 16q desde el centrómero al telómero y los marcadores utilizados para cada serie se indican con X. Los números de identificación de tumores se muestran en la parte superior de cada columna. A la derecha de la figura, se indican las tres regiones más pequeñas de pérdida de heterocigosisidad.

Figura 2. Análisis por transferencia Northern del gen BNO1. El tamaño del gen BNO1 en kilobases se indica mediante una flecha en el lado izquierdo del Northern. La transferencia contenía ARN de los siguientes tejidos: 1: Glándula mamaria; 2: Médula ósea; 3: Testículo; 4: Ovario; 5: Útero; 6: Próstata; 7: Estómago; 8: Vejiga; 9: Médula espinal; 10: Cerebro; 11: Páncreas; 12: Tiroides. Una banda aislada de aproximadamente 3,6 Kb se observó en todos los tejidos excepto en la médula ósea. La expresión más potente del gen se observó en el cerebro.

Figura 3. Alineación de la secuencia de F-box de BNO1 en comparación con la secuencia de consenso de F-box de acuerdo con la descripción de Kipreos y Pagano (2000). Se utiliza el código de aminoácidos de una sola letra. Las mayúsculas en negrita indican los residuos que se encuentran en más del 40% de las secuencias de F-box; las mayúsculas sin negrita indican residuos que se encuentran en el 20-40% de las secuencias de F-box; las letras minúsculas en negrita indican los residuos que se encuentran en el 15-19% de las F-box; las letras minúsculas sin negrita indican residuos encontrados en el 10-14% de las F-box. La línea superior representa el motivo de F-box de BNO1 que indica un alto grado de homología con el de consenso.

Figura 4. Análisis de la expresión mediante RT-PCR cuantitativa del gen BNO1 en líneas celulares de cáncer de mama. El número de copias de BNO1 en el ADNc de glándula mamaria normal (mama) se estableció arbitrariamente para un "valor de referencia" de 1.0e+06 (barra blanca). Los números de copias de ADNc de líneas celulares de cáncer de mama y de otros tejidos normales se calcularon en relación con el "valor de referencia". Las barras de color gris represen-

tan la regulación a la baja de la expresión en número de veces del amplicón, en comparación con la referencia del valor de referencia, mientras que las barras de color negro representan la regulación al alza de la expresión en número de veces del amplicón de la referencia del valor de referencia. Nota: las líneas celulares replicadas (a y b) representan cultivos celulares independientes, el aislamiento del ARN total y las reacciones de transcripción inversa. Las réplicas servían como otro nivel de control para vigilar la variabilidad de la expresión genética como resultado de las diferencias en las eficacias de la confluencia celular, la integridad del ARN total y la transcripción inversa.

Modos de realizar la invención

EJEMPLO 1: Recogida de material a partir de pacientes con cáncer de mama

Para este estudio se analizaron dos series de pacientes con cáncer de mama. La clasificación histopatológica de cada espécimen tumoral la realizaron nuestros colaboradores de acuerdo con los criterios de la Organización Mundial de la Salud (OMS, 1981). Los pacientes se catalogaron histopatológicamente según el método modificado de Bloom y Richardson (Elston y Ellis, 1990) y el material de los pacientes fue obtenido con la aprobación de los Comités locales de Ética Médica. El ADN del tejido tumoral y el ADN de sangre periférica de la misma persona se aislaron tal como se ha descrito anteriormente (Devilee *et al.*, 1991) utilizando protocolos convencionales de laboratorio.

- La serie 1 consistía en 189 pacientes operados entre 1986 y 1993 en tres hospitales holandeses, una universidad holandesa y dos centros periféricos. El tejido tumoral se congeló rápidamente a las pocas horas de la resección. Para el aislamiento del ADN, se seleccionó un bloque de tejido únicamente si contenía por lo menos 50% de células tumorales, tras el examen de secciones de tejido teñidas con hematoxilina y eosina, realizado por un patólogo. Los bloques de tejido que contenían menos de 50% de células tumorales se excluyeron de un análisis adicional.
- La serie 2 consistía en 123 pacientes operados entre 1987 y 1997 en el Centro Médico Flinders en Adelaida, Australia. De todos ellos, 87 se recogieron como especímenes frescos pocas horas después de la resección quirúrgica, confirmados como tejido maligno por un análisis patológico, se congelaron rápidamente en nitrógeno líquido y se almacenaron a -70°C. Las 36 muestras restantes de tejido tumoral se obtuvieron a partir de bloques de tumores de archivo embebidos en parafina. Antes del aislamiento del ADN, las células tumorales se microdisecionaron desde secciones de tejido montadas en portaobjetos de vidrio, con el fin de obtener al menos 80% de células tumorales. En algunos casos, no había sangre periférica disponible, de tal manera que en su lugar se utilizó tejido de ganglios linfáticos no malignos embebido en parafina, identificado patológicamente.

EJEMPLO 2: Análisis de LOH de marcadores del cromosoma 16q en muestras de cáncer de mama.

Con el fin de identificar la posición de los genes asociados con el cáncer de mama, se llevó a cabo el análisis de LOH de muestras tumorales. Un total de 45 marcadores genéticos que cartografián el cromosoma 16 se utilizaron para el análisis de LOH del tumor de mama, y se emparejaron con muestras normales de ADN recogidas para este estudio. La Figura 1 indica para qué serie tumoral se utilizaron y su posición citogenética. Los detalles relativos a todos los marcadores se pueden obtener a partir de la base de datos del genoma (GDB) en <http://www.gdb.org>. El orden físico de los marcadores con respecto uno al otro se determinó a partir de una combinación de información en GDB, cartografiando sobre un mapa híbrido celular somático del cromosoma 16 (Callen *et al.*, 1995) y a través de la información de la secuencia genómica.

Se emplearon cuatro métodos alternativos para el análisis de LOH:

- 1) Para los marcadores de RFLP y VNTR, se utilizó la transferencia Southern para someter a ensayo el desequilibrio alélico. Estos marcadores se utilizaron solo sobre un subconjunto de muestras. Los métodos utilizados eran tal y como se han descrito previamente (Devilee *et al.*, 1991).
- 2) Los marcadores de microsatélites fueron amplificados a partir de ADN tumoral y normal usando la reacción en cadena de la polimerasa (PCR) incorporando metodologías convencionales (Weber y May, 1989; Sambrook *et al.*, 1989). Una reacción típica consistía en 12 µl y contenía 100 ng de molde, 5 pmol de ambos cebadores, 0,2 mM de cada dNTP, 1 µCurie[α-³²P]dCTP, MgCl₂ 1,5 mM, 1,2 µl de tampón Supertaq (HT biotechnologies). Se utilizó un formador de imágenes de fósforo de tipo 445 SI (Molecular Dynamics, Sunnyvale, CA) para cuantificar los resultados ambiguos. En estos casos, el factor de desequilibrio alélico (FDA) se determinó como el cociente de las relaciones de altura de pico de la pareja de ADN normal y tumoral. El umbral de desequilibrio alélico se definió como una reducción del 40% de un alelo, de acuerdo con un FDA ≥1,7 o ≤0,59. Este umbral está de acuerdo con la selección de bloques de tejidos tumorales que contenían al menos 50% de células tumorales con un intervalo de error del 10%. El umbral de retención se ha determinado previamente dentro de un intervalo de 0,76 a 1,3 (Devilee *et al.*, 1994). Esto deja un intervalo de FDAs (0,58 - 0,75 y 1,31 - 1,69) para el que no se ha tomado ninguna decisión definitiva. Esta "zona gris" se indica por cuadros grises en la Figura 1 y los tumores que solo tienen valores en la "zona gris" se descartaron completamente del análisis.
- 3) El tercer método para determinar el desequilibrio alélico era similar al segundo método anterior, sin embargo, se omitió dCTP marcada radiactivamente. En su lugar, se realizó una PCR de marcadores de microsatélites polymórficos con uno de los cebadores de la PCR marcado con fluorescencia con FAM, TET o HEX. El análisis de los productos de la PCR generados se realizó en un secuenciador automático ABI 377 (PE Biosystems), utilizan

zando geles de poliacrilamida al 6% que contenían urea 8 M. Los valores de la altura de los picos y los tamaños de los picos se analizaron con el programa GeneScan (PE Biosystems). Se utilizaron los mismos umbrales para desequilibrio alélico, retención y zonas grises que en el análisis radiactivo.

5) También se utilizó un sistema basado en fluorescencia alternativa. En este caso, los cebadores de la PCR se marcaron con fluoresceína o hexaclorofluoresceína. Los volúmenes de la reacción de PCR eran de 20 µl e incluían 100 ng de molde, 100 ng de cada cebador, 0,2 mM de cada dNTP, MgCl₂ 1-2 mM, 1X tampón AmpliTaq Gold y 0,8 unidades de enzima AmpliTaq Gold (Perkin Elmer). Las condiciones de la ciclación eran 10 ciclos a 94°C durante 30 segundos, 60°C durante 30 segundos, 72°C durante 1 minuto, seguidos por 25 ciclos a 94°C durante 30 segundos, 55°C durante 30 segundos, 72°C durante 1 minuto, con una extensión final a 72°C durante 10 minutos. Los amplímeros de la PCR fueron analizados en un secuenciador automatizado ABI 373 (PE Biosystems), utilizando el programa GeneScan (PE Biosystems). El intervalo del umbral de FDA para la retención alélica se definió como 0,61 - 1,69, la pérdida alélica como ≤0,5 o ≥2,0, o la "zona gris" como 0,51 - 0,6 o 1,7 - 1,99.

10 Los tres primeros métodos se aplicaron a la primera serie de tumores mientras que el último método se adoptó para la segunda serie de muestras tumorales. Para el análisis estadístico, se realizó una comparación de los datos de desequilibrio alélico para la validación de los diferentes métodos de detección y de las diferentes series tumorales, utilizando la prueba de Chi-cuadrado.

15 La identificación de la región más pequeña de solapamiento (SRO) implicada en LOH es instrumental para reducir las posibilidades de la posición del gen que es diana de LOH. La Figura 1 muestra los resultados de LOH para muestras tumorales, que muestran pequeñas regiones de pérdida (es decir, LOH intersticial y telomérica) y no incluye muestras que mostraron LOH compleja (pérdida y retención alternas de marcadores). Al comparar los dos grupos de muestras, aparecen al menos tres regiones compatibles, estando dos en el telómero en la banda 16q24.3 y una en 16q22.1. La 20 región en 16q22.1 se define por los marcadores D16S398 y D16S301 y se basa en los casos de LOH intersticial observados en tres tumores de la serie 1 (239/335/478) y un tumor de la serie 2 (237). En el telómero (16q24.2 - 16q24.3), la 25 primera región se define por los marcadores D16S498 y D16S3407 y se basa en cuatro tumores de la serie 2 (443/75/631/408), mientras que la segunda región (16q24.3) se extiende desde D16S3407 al telómero y se basa en un tumor de la serie 1 (559) y tres de la serie 2 (97/240/466). La LOH limitada al telómero pero que implicaba ambas regiones identificadas en este sitio, se pudo encontrar en 17 muestras tumorales adicionales.

30 Otros estudios han mostrado que el brazo largo del cromosoma 16 también es una diana para la LOH en tumores de próstata, de pulmón, hepatocelular, de ovario, de rabdomiosarcoma y de Wilms. Un análisis detallado de los carcinomas de próstata ha revelado un solapamiento en las regiones más pequeñas de LOH observadas en este tipo de cáncer, con las observadas con cáncer de mama, lo que sugiere que 16q alberga un gen implicado en muchos tipos de tumores.

EJEMPLO 3: Construcción de un mapa físico de 16q24.3

35 Para identificar nuevos genes candidatos de cáncer de mama que se cartografián en las regiones de solapamiento más pequeñas en 16q24.3, se necesitaba un cóntigo basado en clones para el mapa físico que cubriera esta región. Al comienzo de esta fase del proyecto, los fragmentos de ADN genómico clonados, utilizados más comúnmente y de fácil acceso estaban contenidos en vectores lambda, cósvidos o YAC. Durante la construcción de todos los mapas físicos del cromosoma 16, clones procedentes de una variedad de genotecas de YAC se incorporaron en el mapa (Doggett et al., 1995). Estos incluían clones procedentes de una genoteca de YAC específica del cromosoma 16 clasificada por flujo 40 (McCormick et al., 1993), de las genotecas CEPH Mark I y MegaYAC y de una genoteca de YAC de medio-telómero (Riethman et al., 1989). Un análisis de STS y Southern detallado de los clones de YAC que cartografiaban 16q24.3, estableció que muy pocos estaban localizados entre el punto de ruptura del híbrido celular somático CY2/CY3 y el telómero del brazo largo. Sin embargo, aquellos que se habían localizado en esta región proporcionaron resultados incompatibles con la cartografía y se sospechó que se habían reorganizado o eliminado. El hecho de que los clones de YAC 45 producen sustratos de secuenciación débiles, junto con la dificultad para aislar el ADN humano clonado, la opción inicial preferida fue un mapa físico basado en clones de cósvidos.

50 Se había construido previamente una genoteca de cósvidos específicos del cromosoma 16 clasificada por flujo (Longmire et al., 1993), con clones de cósvidos individuales reticulados en matrices de alta densidad sobre membranas de nailon. Estos filtros contenían colectivamente ~15.000 clones que representaban una cobertura de aproximadamente 5,5 veces el cromosoma 16. Los cósvidos individuales que cartografiaban las regiones decisivas en 16q24.3, se identificaron mediante la hibridación de estas membranas con marcadores identificados por cartografiar la región mediante este estudio y estudios anteriores. La estrategia para alinear clones de cósvidos solapantes se basa en su contenido en STS y en el patrón de digestión con endonucleasas de restricción. Aquellos clones que se extendían lo más lejos posible de cada cóntigo inicial, se utilizaron después para recorrer el cromosoma mediante la hibridación de nuevo de los 55 extremos de estos cósvidos con las matrices de cósvidos de alta densidad. Este proceso continuó hasta que todos los cóntigos iniciales estaban unidos y, por lo tanto, la región que definía la posición de los genes supresores de tumores de cáncer de mama estaría contenida dentro del mapa. Los clones de cósvidos individuales que representaban el mínimo solapamiento posible en el cóntigo, se utilizaron a continuación para la identificación de secuencias transcritas mediante atrapamiento de exones y para la secuenciación genómica.

El cromosoma 16 se clasificó a partir del híbrido CY18 de células somáticas de ratón/humano, que contenía este cromosoma como el único ADN humano, y el ADN de CY18 parcialmente digerido con *Sau3A* se ligó en el sitio de clonación de *BamHI* del vector del cósmido sCOS-1. Todas las matrices se hibridaron y se lavaron usando métodos descritos en Longmire *et al.* (1993). Resumiendo, los 10 filtros se hibridaron previamente en 2 botellas grandes durante al menos 2 horas en 20 ml de una solución que contenía 6X SSC, EDTA 10 mM (pH 8,0); 10X solución de Denhardt, 1% de SDS y 100 µg/ml de ADN de esperma de salmón fragmentado desnaturalizado a 65°C. Las hibridaciones durante una noche con sondas marcadas con [α -³²P]dCTP se realizaron en 20 ml de solución de hibridación de nuevo aporte a 65°C. Los filtros se lavaron secuencialmente en soluciones de 2X SSC; 0,1% de SDS (lavado a temperatura ambiente), 2X SSC, 0,1% de SDS (temperatura ambiente durante 15 minutos), 0,1X SSC, 0,1% de SDS (temperatura ambiente durante 15 minutos) y 0,1X SSC, 0,1% de SDS (dos veces durante 30 minutos a 50°C si es necesario). Las membranas se expusieron a -70°C entre 1 y 7 días.

Los marcadores iniciales utilizados para escrutar la matriz de cósmidos eran los conocidos por estar localizados por debajo de los puntos de ruptura de las células híbridas somáticas CY2/CY3 y el telómero del brazo largo (Callen *et al.*, 1995). Estos incluían tres genes, *CMAR*, *DPEP1* y *MC1R*; el marcador de microsatélites D16S303; un fragmento del extremo del cósmido 317E5, que contenía el gen *BBC1*; y cuatro clones de ADNc, yc81e09, yh09a04, D16S532E y ScDNA-C113. El clon de ADNc del consorcio IMAGEN, yc81e09, se obtuvo a través del escrutinio de una genoteca de ADNc cebada con oligo-dT de cerebro infantil, normalizada (Soares *et al.*, 1994), con el inserto procedente del clon de ADNc ScDNA-A55. Tanto el clon ScDNA-A55 como el ScDNA-C113 se aislaron originalmente a partir de una genoteca de ADNc heteronuclear cebada con hexámero, construida a partir del híbrido CY18 de células somáticas de ratón/humano (Whitmore *et al.*, 1994). El clon de ADNc de IMAGEN yh09a04 se identificó a partir de una selección de ADNc directa del cósmido 37B2 que se había observado previamente que se cartografiaba entre el punto de ruptura CY18A(D2) y el telómero 16q. El EST, D16S532E, también se cartografiaba en la misma región. Con posterioridad a estos escrutinios iniciales, se utilizaron fragmentos de restricción que representaban los extremos de cósmidos para identificar clones solapantes adicionales.

El conjunto de cíntigos se basaba en métodos descritos previamente (Whitmore *et al.*, 1998). Más tarde, durante la construcción del mapa físico, se pusieron a disposición genotecas genómicas clonadas en vectores BAC o PAC (Genome Systems o Rosewell Park Cancer Institute). Estas genotecas se escrutaron para ayudar en el recorrido por el cromosoma o cuando se toparon con huecos que no se pudieron puentear usando los filtros de cósmidos. Todos los filtros de BAC y PAC se hibridaron y se lavaron de acuerdo con las recomendaciones del fabricante. Inicialmente, las membranas se hibridaron previamente por separado en botellas grandes de vidrio durante al menos 2 horas en 20 ml de 6X SSC; 0,5% de SDS; 5X Denhardt; 100 µg/ml de ADN de esperma de salmón desnaturalizado a 65°C. Las hibridaciones durante una noche con sondas marcadas con [α -³²P]dCTP se realizaron a 65°C en 20 ml de una solución que contenía 6X SSC; 0,5% de SDS; 100 µg/ml de ADN de esperma de salmón desnaturalizado. Los filtros se lavaron secuencialmente en soluciones de 2X SSC, 0,5% de SDS (temperatura ambiente, 5 minutos), 2X SSC, 0,1% de SDS (temperatura ambiente, 15 minutos) y 0,1X SSC, 0,5% de SDS (37°C 1 hora si es necesario). Los clones de PAC o BAC identificados se alinearon con el cíntigo existente basándose en su patrón de enzimas de restricción o formaban cíntigos únicos que se extendían mediante el escrutinio de filtros adicionales.

Cuando se observó que el microsatélite D16S303 era el marcador más telomérico en la región 16q24.3 (Callen *et al.*, 1995), se utilizó la hibridación fluorescente *in situ* (FISH) con cromosomas normales en metafase, utilizando cósmidos completos cartografiados en las proximidades de este marcador, para definir el límite telomérico del cíntigo. El ADN del cósmido completo se sometió a desplazamiento de mellas con biotina-14-dATP y se hibridó *in situ* con una concentración final de 20 ng/ μ l con metafases de 2 machos normales. El método FISH se había modificado en comparación con el previamente descrito (Callen *et al.*, 1990). Los cromosomas se tiñeron antes del análisis tanto con yoduro de propidio (como tinción de contraste) como con DAPI (para la identificación de cromosomas). Imágenes de preparaciones en metafase se capturaron por una cámara CCD refrigerada, utilizando la colección de imágenes CytoVision Ultra y un sistema de mejora (Applied Image Int. Ltd.). El cósmido 369E1 mostraba señales fluorescentes claras en el telómero del brazo largo del cromosoma 16. Sin embargo, esta sonda también proporcionó una señal clara en los telómeros de los brazos cromosómicos 3q, 7p, 9q, 11p y 17p. Por el contrario, el cósmido 439G8, que se cartografió próximo a D16S303, proporcionó señales fluorescentes solo en 16qter sin ninguna señal compatible detectada en otros telómeros. Estos resultados nos han permitido establecer el marcador de microsatélites D16S303 como el límite de la transición desde la eucromatina hasta las repeticiones subteloméricas, proporcionando un límite telomérico para el cíntigo (Whitmore *et al.*, 1998).

Se ha establecido un mapa físico de alta densidad que consistía en clones de cósmido, BAC y PAC, el cual se extiende aproximadamente 3 Mb desde el telómero del brazo largo del cromosoma 16. Este cíntigo se extiende más allá del punto de ruptura del híbrido celular somático CY2/CY3 e incluye las 2 regiones de LOH mínima, identificadas en la región 16q24.3 en muestras de cáncer de mama. Hasta la fecha, existe un solo hueco de tamaño desconocido en el cíntigo y se cerrará mediante experimentos adicionales de extensión del cíntigo. El grado de cobertura ha permitido la identificación del grupo de clones con mínimo solapamiento posible que se utilizaron posteriormente como moldes para métodos de identificación de genes, tales como la captura de exones y la secuenciación de ADN genómico.

EJEMPLO 4: Identificación de los genes candidatos de cáncer de mama mediante análisis de la secuencia de ADN genómico.

Clones de BAC y PAC seleccionados con un mínimo solapamiento a partir del cóntigo del mapa físico, se secuenciaron con el fin de ayudar en la identificación de los genes candidatos de cáncer de mama. Se preparó ADN a partir de clones seleccionados utilizando un kit para aislamiento de ADN de gran tamaño (Qiagen). Aproximadamente el 25-50 µg de ADN se cortó a continuación mediante nebulización (10 psi durante 45 segundos) y se establecieron extremos romos 5 utilizando metodologías convencionales (Sambrook *et al.*, 1989). A continuación, las muestras se aplicaron sobre un gel de agarosa con el fin de aislar el ADN en el intervalo de tamaño de 2-4 kb. Estos fragmentos se limpian eliminando la agarosa utilizando columnas QIAquick (Qiagen), se ligaron en puc18 y se utilizaron para transformar células de *E. coli* DH10B o DH5a competentes. Se aisló el ADN a partir de clones transformados y se secuenció utilizando cebadores 10 específicos del vector en un secuenciador ABI377. El análisis de la secuencia genómica se realizó empleando un programa informático PHRED, PHRAP y GAP4 en una estación de trabajo SUN. Para ayudar en la generación de cóntigos grandes de la secuencia genómica, la información presente en la base de datos de secuencias genómicas de gran productividad (htgs) en el NCBI, fue incorporada en la fase de ensamblado del análisis de la secuencia. En los cóntigos 15 resultantes de la secuencia genómica se enmascararon las repeticiones y se analizaron usando el algoritmo BLAST (Altschul *et al.*, 1997) para identificar la homología de nucleótidos y proteínas con secuencias en las bases de datos de GenBank no redundantes y de EST en el NCBI. La secuencia genómica también se analizó para estudiar la estructura génica prevista, utilizando el programa GENSCAN.

Los clones de ADNc homólogos del consorcio IMAGEN fueron adquiridos de Genome Systems y se secuenciaron. Estos tramos más largos de secuencia se compararon después con genes conocidos mediante comparaciones de secuencias de nucleótidos y aminoácidos, utilizando los procedimientos anteriores. Cualquier secuencia que se expresa en la mama se considera que es un gen candidato de cáncer de mama. Aquellos genes cuya función podría implicarlos en el proceso carcinogénico, tal y como se había previsto a partir de las búsquedas de homología con proteínas conocidas, fueron tratados como los candidatos más probables. La evidencia de que un candidato particular es el gen responsable, proviene de la identificación de alelos defectuosos del gen en individuos afectados o a partir del análisis de los niveles 20 de expresión de un gen candidato particular en muestras de cáncer de mama, en comparación con tejidos normales testigos. 25

EJEMPLO 5: Identificación de la secuencia de BNO1

Análisis de la secuencia genómica

Secuencias procedentes de clones BAC que se cartografiaban cerca del punto de ruptura de CY2/CY3, se ensamblaron y se utilizaron en búsquedas de homología con BLASTN en la base de datos dbEST del NCBI 30 (<http://www.ncbi.nlm.nih.gov>). Un gran número de clones de ADNc, se identificaron por formar parte de la secuencia en esta región y éstos se pudieron caracterizar adicionalmente en distintas agrupaciones de UniGene.

El clon de ADNc humano de IMAGE, 46795, correspondiente a la agrupación de UniGene Hs.7970, se secuenció y se utilizó en otras búsquedas de homología en bases de datos. Esto identificó un clon de ADNc solapante presente en la base de datos no redundante (GenBank número de orden AL117444) que extendía la secuencia del clon 46795 más allá de 5'. Como esta secuencia 5' adicional también estaba presente en la secuencia genómica situada en 5' de la secuencia del clon 46795, se confirmó que AL117444 lo más probablemente pertenecía al transcripto Hs.7970. Para verificar este hecho, se realizó una RT-PCR. 35

Brevemente, ARNm poliA⁺ procedente de una glándula mamaria normal (Clontech) se cebó inicialmente con un cebador oligo-dT y se transcribió de forma inversa utilizando el kit de RT de OmniScript (Qiagen) de acuerdo con las condiciones 40 del fabricante. Las reacciones de control se incluyeron para cada molde de ARN en las que se omitía la transcriptasa inversa de la etapa de síntesis del ADNc. Esto era para determinar la presencia de cualquier contaminación del ADN genómico en las muestras de ARN. La primera hebra de ADNc resultante se amplificó con PCR usando los cebadores AL-1 (específico para AL117444; SEQ ID NO: 20) y 7970-1 (específico para el extremo 3' de Hs.7970; SEQ ID NO: 21) utilizando el kit de HotStarTag (Qiagen) en un volumen de reacción de 10 µl durante 35 ciclos. Inicialmente, los cebadores 45 para el gen de mantenimiento testigo, Esterasa D (SEQ ID Números: 22 y 23) se utilizaron en una reacción distinta para confirmar la presencia de moldes de ADNc para cada reacción de transcripción inversa. Las secuencias de los cebadores se muestran en la Tabla 1. Estos experimentos confirmaron que el clon de ADNc AL117444 y de IMAGE 46795, pertenecían al transcripto Hs.7970.

Análisis Northern

50 Para determinar el tamaño del gen correspondiente a Hs.7970, una transferencia Northern poliA⁺ obtenida a partir de Clontech se investigó con una porción del gen que se había generado mediante PCR usando los cebadores BNO1-2 (SEQ ID NO: 24) y BNO1-3 (SEQ ID NO: 25). La Tabla 1 enumera las secuencias de los cebadores utilizados. Las hibridaciones se llevaron a cabo en 10 ml de solución ExpressHyb (Clontech) durante una noche a 65°C. Los filtros se lavaron de acuerdo con las condiciones del fabricante. La Figura 2 muestra los resultados de la hibridación. Una sola banda de aproximadamente 3,6 kb se detectó en la glándula mamaria, testículo, ovario, útero, próstata, estómago, vejiga, médula espinal, cerebro, páncreas y tiroides. La expresión más fuerte del gen se observó en el cerebro. El tamaño del ARNm correspondiente a Hs.7970 tal y como se determinó por la hibridación Northern, indicaba que se necesitaba obtener una secuencia 5' adicional para el gen. 55

Identificación de la secuencia 5'

Para identificar una secuencia 5' adicional para el transcripto His.7970, se utilizaron secuencias de ADNc presentes en dbEST correspondientes al ortólogo de ratón. El clon de ratón que se extendía más 5' (AU080856) incluía un sitio putativo de inicio de la traducción. La alineación de AU080856 con la secuencia genómica humana que contenía Hs.7970

5 definía la secuencia humana correspondiente de esta transcripción, hasta un sitio de inicio de la traducción idéntico. Se realizaron otros experimentos de RT-PCR que confirmaron la presencia de esta secuencia 5' en el transcripto Hs.7970 humano. Además, búsquedas adicionales en dbEST blast identificaron clones de ADNc humanos que contenían el extremo 5' del gen (por ejemplo, el clon de IMAGE 3958783).

10 Experimentos de RT-PCR indicaron también que Hs.7970 existe como una isoforma de corte y empalme alternativos. Esta variante se debe a la inclusión de un exón adicional en marco (exón 2.5) situado entre los exones 2 y 3.

Además, estos experimentos han establecido que el transcripto Hs.7970, denominado BNO1, existe como dos isoformas de corte y empalme alternativos. Una isoforma tiene 3.574 pb de longitud (SEQ ID NO: 1) y está compuesta por 9 exones que abarcan aproximadamente 55 kb de ADN genómico, mientras que la segunda forma de BNO1, que contiene el exón 2.5, tiene 3661 pb de longitud (SEQ ID NO: 3). La Tabla 2 muestra la estructura genómica del gen que indica el tamaño de los intrones y exones. El análisis de las isoformas de BNO1 indica que la isoforma 1 (menos el exón 2.5) tiene un marco de lectura abierto de 1.617 nucleótidos que codifica una proteína de 539 aminoácidos (SEQ ID NO: 2). La isoforma 2 (más el exón 2.5) de BNO1 tiene un marco de lectura abierto de 1.704 pb de longitud y codifica una proteína de 568 aminoácidos (SEQ ID NO: 4). Secuencias parciales de ADN genómico que indican uniones de exón/intrón para BNO1, se exponen en SEQ ID Números: 5-11.

20 **EJEMPLO 6:** Características de la secuencia de BNO1Secuencia de nucleótidos

Un gran número de clones de ADNc humano están presentes en dbEST, los cuales representan el gen BNO1. Una observación de los tejidos de estos clones de ADNc se obtuvo a partir de indicaciones de que el gen también se expresa en la glándula suprarrenal, la sangre, el colon, las células germinales, el corazón, el riñón, el hígado, el pulmón, el músculo, la placenta, la membrana sinovial, las amígdalas, el cuello uterino, el tejido linfático y la piel. Estos tejidos son adicionales a los que se ha mostrado que expresan BNO1 por el análisis Northern (por ejemplo, glándula mamaria, testículo, ovario, útero, próstata, estómago, vejiga, médula espinal, cerebro, páncreas y tiroides) y procedimientos de RT-PCR (por ejemplo, glándula mamaria humana).

30 La secuencia de nucleótidos de BNO1 humano también detecta un gran número de clones de ADNc de ratón, tal y como se ha mencionado anteriormente. El análisis de BLAST *in silico* de una secuencia de ADN genómico de ratón en la base de datos htgs en el NCBI usando la secuencia de nucleótidos de BNO1 humano, tuvo éxito en la identificación del nucleótido de BNO1 de ratón (SEQ ID NO: 12) y la correspondiente secuencia de aminoácidos (SEQ ID NO: 13). La homología de los aminoácidos entre los dos genes tiene un valor tan alto como 95% (desde el aminoácido 76 en el exón 1 hasta el aminoácido 369 en el exón 8), lo que sugiere que el gen está muy conservado entre las dos especies.

35 El análisis de la secuencia genómica humana situada en 3' con respecto al gen BNO1 identificó la presencia de una variedad de agrupaciones UniGene adicionales (Hs.130367, Hs.227170 y Hs.87068) que se ejecutaban en la misma orientación. Experimentos de RT-PCR que empleaban un cebador específico de Hs.130367 (130367-1; SEQ ID NO: 26) y Hs.87068 (87068-1; SEQ ID NO: 27) (véase la Tabla 1 para las secuencias de cebadores) indicaban que estas dos agrupaciones UniGene podrían estar ligadas. La secuenciación del producto de la RT-PCR también identificó la presencia de la agrupación Hs.227170. Otros experimentos de RT-PCR que empleaban un cebador específico de BNO1 (BNO1-1, SEQ ID NO: 28) en combinación con un cebador específico de Hs.130367 (130367-2; SEQ ID NO: 29) establecieron que Hs.130367 también podría estar ligado al gen BNO1 (véase la Tabla 1 para las secuencias de cebadores). Por lo tanto, las tres agrupaciones UniGene situadas en 3' de BNO1, representan lo más probablemente variantes de este gen que contienen secuencias 3' UTR adicionales. La ausencia de bandas de Northern que se corresponden con el tamaño de estas variantes de BNO1, sugiere que son formas raras del gen. SEQ ID Números: 14-19 representan las secuencias de nucleótidos de estas variantes.

Secuencia de aminoácidos

La secuencia de aminoácidos de BNO1 se utilizó en el análisis *in silico* para identificar proteínas homólogas con el fin de establecer la función del producto génico. El análisis de la proteína de BNO1 frente a las bases de datos Prosite y PfScan (http://www.isrec.isb-sib.ch/software/PFSCAN_form.html), mostró que las dos isoformas de corte y empalme de esta proteína (SEQ ID Números: 2 y 4) contienen un dominio F-box en el extremo amino terminal con un valor de esperanza altamente significativo de 5.6e-10. La Figura 3 muestra la secuencia de F-box de BNO1 en comparación con la secuencia de consenso de F-box.

55 F-box es un motivo proteico de aproximadamente 50 aminoácidos que define una familia en expansión de proteínas eucariotas. Las proteínas que contienen F-box son los componentes de reconocimiento del sustrato de los complejos SCF ubicuitina-ligasa. Estos complejos contienen cuatro componentes: Skp1, Culina, Rbx/Roc1/Hrt1 y una proteína F-box. El motivo de F-box enlaza la proteína F-box a otros componentes del complejo SCF mediante la unión del compo-

nente central de SCF, Skp1. Este motivo se encuentra generalmente en la mitad amino de las proteínas y frecuentemente se acopla con otros dominios proteicos en el extremo carboxi-terminal variable de la proteína. Los dominios carboxi terminales más comunes incluyen repeticiones ricas en leucina (LRRs) y dominios WD-40. Existen actualmente tres subdivisiones de la familia de proteínas F-box en función del tipo de motivos carboxi terminales presentes en las secuencias de proteínas. Siguiendo el modelo propuesto por Cencarelli et al. (1999) y Winston et al. (1999), la nomenclatura adoptada por la Organización del Genoma Humano especifica las F-box que contienen LRRs como FBXL, las que contienen repeticiones WD como FBXW y las que carecen de todos los dominios conocidos de interacción proteica, como FBXO. El análisis de la secuencia de BNO1 no pudo identificar motivos de proteínas adicionales presentes en el gen, lo que indica que BNO1 forma parte de la clase FBXO de las proteínas F-box.

La vía de degradación del proteasoma dependiente de ubicuitina es un mecanismo importante para regular la abundancia de proteínas en organismos eucariotas. Se ha mostrado que una amplia variedad de proteínas están reguladas por este mecanismo e incluyen oncogenes, genes supresores de tumores, factores de transcripción y otras moléculas de señalización (Hershko y Ciechanover, 1998; Baumeister et al., 1998). Estas proteínas influyen sobre una serie de procesos celulares importantes tales como la regulación del ciclo celular y la apoptosis, la modulación de las respuestas inmunes e inflamatorias, el desarrollo y la diferenciación. La amplia gama de proteínas y de procesos que están regulados por la ubicuitinación, sugiere que patologías derivadas de una interrupción del proceso de ubicuitinación también serán diversas. Por ejemplo, hay un precedente de esto en los trastornos neurodegenerativos. Parkin, una proteína mutada en formas hereditarias de la enfermedad de Parkinson, es una ligasa E3 de ubicuitina (Shimura et al., 2000) y en la enfermedad de Alzheimer se ha identificado una ubicuitinación inadecuada de las proteínas cerebrales (López Salón et al., 2000).

El proceso de ubicuitinación comienza con la adición de restos de ubicuitina (ubicuitinación) a proteínas diana y continúa con un proceso de múltiples etapas, cuyo punto final es la proteólisis de sustratos poliubicuitinados a través de un complejo multiproteico de 26S (Haas y Siepmann, 1997; Hochstrasser, 1996). La ubicuitinación de sustratos que son diana para la degradación requiere 3 clases de enzimas: las enzimas activadoras de ubicuitina (E1), las enzimas que se conjugan con ubicuitina (E2) y las ligasas de ubicuitina (E3). Las proteínas E3 tienen una función esencial en la progresión del ciclo celular. Se ha observado que los complejos SCF (una clase de ligasas E3) regulan la transición de la fase G1-S (revisado en Peters, 1998). Una amplia variedad de dianas de SCF se ha descrito e incluyen ciclinas de la fase G1, inhibidores de cinasas dependientes de ciclina, factores de replicación del ADN, factores de transcripción que favorecen la progresión del ciclo celular y otras proteínas celulares importantes. Las secuencias presentes en la región carboxi terminal variable de las proteínas F-box permiten, por lo tanto, la captación de sustratos específicos para la ubicuitinación y la degradación posterior.

Estudios recientes de la proteína supresora de tumores de Von Hippel-Lindau (VHL) han mostrado que forma parte de un complejo que actúa como una ligasa E3 de ubicuitina-proteína (Zaibo et al., 2001). La proteína VHL se une al complejo de la ligasa para dirigirse a proteínas que incluyen HIF α (factor inducible de hipoxia) (Ohh et al., 2000; Cockman et al., 2000) y VDU1 (enzima 1 de desubicitinación que interacciona con VHL) (Zaibo et al., 2001). HIF α se ha mostrado que regula genes implicados en la angiogénesis, un proceso decisivo para el crecimiento de tumores (Wang et al., 1995; Semenza, 2000), mientras que VDU1 tiene actividad desubicitinadora.

La función prevista para BNO1, basada en la presencia del dominio de F-box, indica que el gen puede estar implicado en una amplia gama de procesos celulares que incluyen la regulación del ciclo celular. Junto con el hecho que BNO1 se encuentra en una región de LOH observada en el cáncer de mama y otros tipos de tumores, esto sugiere que BNO1 es un gen candidato ideal para el cáncer de mama.

EJEMPLO 7: Examen del nivel de expresión de BNO1 en líneas celulares de cáncer de mama

Para investigar una posible función de BNO1 en el cáncer de mama, se comparó el nivel de expresión del gen en líneas celulares de cáncer de mama con testigos de tejido normal. El examen de la secuencia genómica que rodea BNO1 muestra que el extremo 5' que incluye el exón 1, es extremadamente rico en GC lo que sugiere la presencia de una isla de CpG. Sin pretender imponer ninguna teoría, se plantea la posibilidad de que existan mecanismos epigenéticos para inactivar la función de BNO1. Una metilación anormal en este sitio puede dar como resultado una regulación a la baja de la transcripción de BNO1 de la copia restante del gen. Estudios recientes han mostrado que este mecanismo ha sido responsable de la inactivación de otros genes supresores de tumores, tales como RB1 (Ohtani-Fujita et al., 1997), VHL (Prowse et al., 1997), MLH1 (Herman et al., 1998) y BRCA1 (Esteller et al., 2000).

Para detectar el nivel de expresión de BNO1 en muestras de cáncer en comparación con testigos normales, se realizó una RT-PCR cuantitativa utilizando cebadores específicos de BNO1. Esta implicaba inicialmente el aislamiento de ARN procedente de líneas celulares de cáncer de mama junto con testigos de líneas celulares apropiadas.

Líneas celulares de cáncer de mama/próstata y extracción de ARN

Las líneas celulares cancerígenas se adquirieron en la ATCC (EE.UU.) y se cultivaron en el medio de cultivo tisular recomendado. Se escogieron líneas celulares de cáncer de mama para el análisis con RT-PCR que mostraban homocigosisidad para una serie de marcadores que cartografiaban el cromosoma 16q, indicando un potencial de LOH para este brazo cromosómico. Las células se recogieron de cultivos confluentes y el ARN total se extrajo usando el kit RNAeasy

(Qiagen). Las líneas celulares de cáncer de mama obtenidas para la extracción de ARN eran BT549, MDA-MB-468, CAMA-1, ZR75-30, MDAMB-157, ZR75-1, SKBR3, MDA-MB-231, T47D y MDA-MB-436. La línea celular epitelial de mama normal MCF12A y la línea celular de próstata PC3 línea también se adquirieron. El ARNm poliA⁺ se aisló posteriormente a partir de todas las fuentes, utilizando el sistema de perlas de Oligotex (Qiagen). El ARNm poliA⁺ procedente de glándula mamaria, próstata, ovario e hígado normales fue adquirido comercialmente (Clontech, EE.UU.).

5

Transcripción inversa

El ARNm poliA⁺ se cebó con cebadores oligo-dT y se transcribió de forma inversa utilizando el kit de RT Omniscript (Qiagen) de acuerdo con las condiciones del fabricante. Las reacciones testigo se incluyeron para cada molde de ARN en las que se omitía la transcriptasa inversa de la etapa de síntesis del ADNc. Esto era para determinar la presencia de cualquier contaminación con ADN genómico en las muestras de ARN.

10

Normalización del ADNc

Se generaron amplicones internos con curva estándar a partir de una mezcla mixta de ADNc de tejido normal, usando el kit de polimerasa de ADN HotStarTaq® (Qiagen). Una mezcla de reacción suficiente para generar >1 µg de ADNc de amplicón contenía 10 µl de tampón 10X PCR (que contenía MgCl₂ 15 mM), 2 µl de mezcla de dNTPs 10 mM, 0,5 uM de cada cebador, 0,5 µl de 2,5 unidades de polimerasa HotStarTaq (Qiagen), 100 ng de molde de ADNc y agua tratada con DEPC hasta 100 µl. Los ciclos de la amplificación se realizaron del modo siguiente: 94°C durante 10 minutos, seguido de 35 ciclos a 93°C durante 20 segundos, 60°C durante 30 segundos y 70°C durante 30 segundos, con una extensión final a 72°C durante 4 minutos. Los amplicones se purificaron utilizando el kit de extracción en gel QIAquick (Qiagen) de acuerdo con las condiciones del fabricante y las concentraciones se midieron a A₂₆₀. Los amplicones purificados se diluyeron en serie 10 veces desde 10 ng/µl hasta 1 fg/µl. Estas diluciones sirvieron como patrones internos de concentración conocida para el análisis en tiempo real de amplicones específicos de BNO1, tal y como se describe a continuación.

15

PCR en tiempo real

Todos los moldes de ADNc se amplificaron utilizando el kit SYBR Green I PCR Master Mix (PE Biosystems, EE.UU.). Las reacciones de PCR tenían un volumen de 25 µl e incluían 12,5 µl de SYBR Green I PCR Master Mix, 0,5 µM de cada cebador, 2 µl de molde de ADNc normalizado (ver más abajo) y 9,5 µl de agua. El análisis con PCR en tiempo real se realizó utilizando el Rotor-Gene® 2000 (Corbett Research, AUS) con las siguientes condiciones de ciclos de amplificación: 94°C durante 10 minutos seguido por 45 ciclos a 93°C durante 20 s, 60°C durante 30 s y 70°C durante 30 s. Los datos de fluorescencia se adquirieron a 510 nm durante la fase de extensión a 72°C. El análisis de las curvas de fusión se realizó con una ciclación inicial de 99-50°C seguida de una monitorización de la fluorescencia durante el calentamiento a 0,2°C/segundo hasta 99°C. Antes de la cuantificación en tiempo real, el tamaño y la especificidad del producto se confirmaron mediante tinción con bromuro de etidio de los geles de agarosa al 2,5% después de una electroforesis de las PCRs completadas. Los cebadores específicos del testigo y de BNO1 utilizados para todas las aplicaciones de PCR en tiempo real se muestran en la Tabla 1 y se representan por las SEQ ID Números: 30-41.

20

Cuantificación de la PCR en tiempo real

Los análisis de cuantificación se realizaron en el sistema de análisis de muestras de ADN Rotor-Gene® (Versión 4.2, fabricado 96). Las curvas estándar se generaron amplificando diluciones en serie de 10 veces (1 µl de 10 pg/µl hasta 1 µl de 1 fg/µl por triplicado) del amplicón estándar interno durante la PCR en tiempo real de los amplicones de BNO1, procedentes de tejidos normales y de líneas celulares de cáncer de mama. Las concentraciones de los amplicones estándar internos se establecieron arbitrariamente en 1.0e+12 copias para patrones de 10 pg hasta 1.0e+08 copias para patrones de 1 fg. Los coeficientes de C_T (ciclo umbral) de variación para todas las diluciones estándar internas se promediaron en 2% entre las muestras por triplicado dentro de la misma ejecución y ejecuciones diferentes. El programa informático para la cuantificación RotorGene® generaba una línea de mejor ajuste en el parámetro C_T y determinaba el número de copias de amplicones de BNO1 de tejido normal desconocido y de la línea celular de cáncer de mama, mediante la interpolación de la intersección de ruido de fondo-banda de los amplicones de BNO1 frente a los patrones internos con un número de copias conocido.

30

Normalización y expresión relativa de los datos

Para tener en cuenta la variación de las concentraciones de partida del molde de muestra a muestra, se utilizó una cuantificación del ARN de Ribogreen® (Molecular Probes) para someter a ensayo de precisión 1 µg de ARN de tejido normal y ARN de líneas celulares de cáncer mama, para la síntesis de ADNc. Los niveles de expresión del gen de mantenimiento seleccionado se analizaron a continuación en todas las muestras para determinar el testigo endógeno más preciso para la normalización de datos. Los amplicones de mantenimiento incluían la Esterasa D (Número de orden M13450), la ciclofilina (número de orden X52851), APRT (número de orden M16446) y la ARN polimerasa II (número de orden Z47727). Como la ciclofilina mostraba el perfil de expresión menos variable, el número de copias calculado para BNO1 se dividió entre el número de copias respectivo de amplicones de ciclofilina, para cada línea celular de cáncer de mama y de tejido normal analizados. El número de copias de BNO1 en ADNc normalizado de mama normal se estableció arbitrariamente en un “valor de referencia” de 1.0e+06 copias. El número de copias de ADNc de líneas celulares de cáncer de mama y de otros tejidos normales se calculó en relación con el “valor de referencia”. Los datos se expresaron

como el log del número de copias relativo de ARNm. La Figura 4 muestra los resultados de estos experimentos.

El grado de variación en los niveles de expresión de ARNm para la ciclofilina, la subunidad de la ARN polimerasa II y APRT era relativamente uniforme entre los tejidos normales y las líneas celulares de cáncer. Combinaciones de tres vías para la normalización entre ciclofilina, subunidad de ARN polimerasa II y APRT mostraron una variación media de 7 veces y como máximo 50 veces en el nivel de expresión del ARNm entre las muestras. La significación de los niveles de expresión variables del ARNm dentro de un gen de interés se puede evaluar, por lo tanto, razonablemente basándose en estos resultados de normalización. Una disminución aberrante prevista en el número de copias de un ARNm de un gen de interés de ~100 veces en líneas celulares de cáncer de mama, en relación con un nivel de expresión en mama normal de "valor de referencia", se consideró, por lo tanto, que era significativamente anormal.

La Figura 4 indica que los amplicones de BNO1 específicos para el exón 5-7 y la isoforma 1 (menos el exón 2.5) muestran un patrón compatible de expresión de ARNm entre los tejidos normales y las líneas celulares de cáncer de mama. Para ambos amplicones analizados, las líneas celulares de cáncer de mama MDA-MB-468, SK-BR3, MDA-MB-231 y la línea celular de cáncer de próstata PC3, todas ellas mostraban un bajo nivel de expresión del ARNm con respecto al tejido de mama normal del "valor de referencia". Una reducción significativa de 725 veces en la expresión del ARNm del exón 5-7 de BNO1 se detectó en SK-BR3 con respecto a la expresión en tejidos normales de mama (equivalente a una regulación a la baja de aproximadamente 350.000-480.000 en la expresión de la molécula de ARNm). Se obtuvieron resultados similares para la isoforma 1 de BNO1 (menos el exón 2.5), con una reducción de 248 veces de la expresión del ARNm en SK-BR3 (equivalente a una regulación a la baja de aproximadamente 300.000-1.000.000 en la expresión de la molécula de ARNm). La isoforma 2 de BNO1 (más el exón 2.5) mostraba una expresión del ARNm significativamente baja en las líneas celulares MDA-MB-468, CAMA-1, SK-BR3 y MDA-MB-231, sin expresión detectada en ZR75-30. Estos resultados indican que ambas isoformas del gen BNO1 están reguladas a la baja en ciertas líneas celulares de cáncer de mama, así como en una línea celular de cáncer de próstata. El mecanismo exacto de esta regulación a la baja no se conoce en este estadio, pero puede ser el resultado de mecanismos tales como una mutación o una metilación del promotor. Partiendo de estos estudios de la expresión proponemos que BNO1 es una proteína responsable del desarrollo de cáncer de mama y de próstata. Debido a su amplio patrón de expresión tisular, BNO1 también puede estar implicada en cánceres que se originan en otros tejidos.

Se pueden utilizar otros métodos para detectar los niveles de expresión de BNO1. Estos incluyen la generación de anticuerpos policloniales o monoclonales, que son capaces de detectar cantidades relativas de las formas tanto normales como mutantes de BNO1, utilizando diversos inmunoensayos tales como los ensayos ELISA (véase el Ejemplo 11 y 12).

30 EJEMPLO 8: Análisis de tumores y de líneas celulares para estudiar mutaciones en BNO1

El gen BNO1 se escrutó mediante análisis de SSCP en ADN aislado a partir de tumores de la serie 1, así como de un subconjunto de tumores de la serie 2 (no se muestra en la Figura 1) que mostraban la pérdida de todo el brazo largo del cromosoma 16. Estas muestras de la serie 2 se utilizaron debido a que estaban disponibles mayores cantidades de ADN. En total se examinaron mutaciones en 45 tumores primarios de mama con LOH en 16q.

En una diversidad de líneas celulares también se escrutaron mutaciones. Estas incluían 22 líneas celulares de cáncer de mama (BT20, BT474, BT483, BT549, CAMA-1, DU4475, Hs578T, MCF7, MB157, MB231, MB361, MB415, MB436, MB453, MB468, SKBR3, T47D, UACC893, ZR75-1, ZR75-30, MB134 y MB175), 2 líneas celulares de cáncer de próstata (LNCAP y PC3), 2 líneas celulares de carcinoma gástrico (AGS y KATO), 1 línea celular de cáncer de hígado (HEP2) y 2 líneas de células epiteliales normales de mama (HBL100 y MCF12A). Todas las líneas celulares se adquirieron en la ATCC, se dejaron crecer de acuerdo con las condiciones de los fabricantes y el ADN se aisló a partir de las células cultivadas, utilizando protocolos convencionales (Wyman y White, 1980; Sambrook *et al.*, 1989).

Los exones de BNO1 se amplificaron mediante PCR usando cebadores intrónicos flankeantes que se marcaron en sus extremos 5' con HEX. Se hizo una excepción para el exón 1 y 8, ya que debido a su tamaño se tuvieron que dividir en 2 amplímeros solapantes. En la Tabla 3 se enumeran las secuencias de todos los cebadores utilizados para el análisis de SSCP, el tamaño esperado del amplímero y la concentración de MgCl₂ utilizada en la reacción de PCR. Las reacciones típicas de PCR se realizaron en placas de 96 pocillos con un volumen de 10 µl utilizando 30 ng de ADN molde. Las condiciones de ciclación fueron una etapa inicial de desnaturalización a 94°C durante 3 minutos, seguida por 35 ciclos a 94°C durante 30 segundos, 60°C durante 1½ minutos y 72°C durante 1½ minutos. Una etapa final de extensión a 72°C durante 10 minutos se realizó a continuación. Veinte µl de colorante de carga que comprendía 50% (v/v) de formamida, EDTA 12,5 mM y 0,02% (p/v) de azul de bromofenol se añadieron a las reacciones completadas, que se aplicaron posteriormente sobre geles de poliacrilamida al 4% y se analizaron en el sistema GelScan 2000 (Corbett Research, AUS) de acuerdo con las especificaciones del fabricante.

De los 12 amplicones sometidos a ensayo, solo en 2 se identificaron desplazamientos de bandas en SSCP. En el exón 2.5, se observaron desplazamientos de bandas idénticos en 2 muestras de tumores de la serie 1 (380 y 355) y en la línea celular de cáncer de mama MCF7. El análisis de SSCP del ADN normal correspondiente procedente de la muestra 380 y 355, identificaba el mismo desplazamiento de bandas, indicando que el cambio lo más probablemente no causaba la enfermedad. El análisis de la secuencia de este desplazamiento de bandas en todas las muestras mostró que un solo cambio de una base de nucleótidos (-5T→C) era responsable de este desplazamiento de bandas. Este cambio no afecta a la puntuación de sitios de consenso aceptores de corte y empalme para este exón y, por lo tanto, lo más probable

es que represente un polimorfismo. La incidencia de este cambio en la población general no se ha examinado aún. En el exón 8b, un desplazamiento de bandas fue identificado en una sola línea celular de cáncer (KATO). La secuenciación de este desplazamiento de bandas indica un cambio de C→T en la posición +10 de este amplicón que se encuentra en el sitio donante de corte y empalme (5' del sitio de corte y empalme). Este cambio de base se produce fuera de la secuencia de consenso de unión de corte y empalme y se prevé que la mutación no tenga efecto sobre el corte y empalme de este exón.

EJEMPLO 9: Inmunoprecipitación de BNO1 y Skp1

Para someter a ensayo si BNO1 contenía un motivo de F-box funcional, se empleó un ensayo de co-inmunoprecipitación. Este implicaba la clonación del marco de lectura abierto de longitud completa, marcado con Myc de BNO1 en los sitios Sall/Clal del vector de expresión retrovírico LNCX2 (Clontech), utilizando técnicas convencionales (Sambrook *et al.*, 1989). Después de esto, 10⁷ células 293T se transfecaron con 10 µg de la estructura artificial BNO1-LNCX2 o por separado con el vector LNCX2 solo como testigo usando Lipofectamina 2000 (Invitrogen), de acuerdo con las instrucciones del fabricante. Las células se recogieron 24 horas después de la transfección y se lisaron en 2 ml de tampón de lisis (Tris-HCl 50 mM [pH 7,5], NaCl 150 mM, 0,5% de Nonidet P-40 suplementado con PMSF 1 mM y 5 µg/ml de leupeptina, antiparina y aprotenina). Después de esto, se incubaron 0,5 ml de lisado celular con 2 µg de anticuerpo monoclonal anti-Myc (Roche) o un anticuerpo policlonal de conejo anti-p19^{Skp1} (Neo Markers, Fremont, CA) durante 1 hora y proteína A-Sefarosa durante 1 hora a 4°C. Los complejos inmunes se lavaron tres veces con 1 ml de tampón de lisis seguido de separación en 10% de SDS-PAGE e inmunotransferencia de acuerdo con técnicas convencionales (Sambrook *et al.*, 1989).

Los resultados de estos experimentos indicaron que BNO1 coprecipitaba específicamente con Skp1 endógena, lo que confirma tanto una asociación entre estas dos proteínas como la presencia de una F-box funcional dentro de BNO1. Esta interacción indica que BNO1 pertenece a un nuevo complejo de ligasas ubicuitina-E3 que puede ser decisivo para la degradación controlada de sustratos específicos de BNO1.

EJEMPLO 10: Análisis del gen BNO1

Los siguientes métodos se utilizan para determinar la estructura y la función de BNO1.

Estudios biológicos

Los vectores de expresión de mamíferos que contienen ADNc de BNO1 (que representa ambas isoformas de BNO1) se pueden transfecar en líneas celulares de carcinoma de mama, de próstata u otros que tienen lesiones en el gen. Se examina la inversión fenotípica en cultivos (por ejemplo, la morfología celular, el crecimiento de transformantes en agar blando, la tasa de crecimiento) y en animales no humanos (por ejemplo, la tumorigenidad en ratones sin pelo). Estos estudios pueden utilizar formas de BNO1 de tipo silvestre o mutante. La delección y los mutantes sustitutivos de BNO1, se pueden construir por mutagénesis *in vitro*.

Estudios biológicos moleculares

La capacidad de ambas isoformas de la proteína de BNO1 para unirse a proteínas conocidas y desconocidas, se puede examinar. Procedimientos tales como el sistema de doble híbrido en levadura se utilizan para descubrir e identificar cualquier ligando funcional, particularmente sustratos específicos de BNO1 o sustratos específicos de una isoforma que son dianas para la degradación por ubicuitinación. El principio que subyace al procedimiento de doble híbrido en levadura es que muchos activadores de la transcripción eucariota, que incluyen los de levadura, consisten en dos dominios modulares discretos. El primero es un dominio de unión al ADN que se une a una secuencia promotora específica y el segundo es un dominio de activación que dirige el complejo de la ARN polimerasa II para transcribir el gen aguas abajo del sitio de unión a ADN. Ambos dominios son necesarios para la activación de la transcripción ya que ninguno de los dos dominios puede activar la transcripción por sí mismo. En el procedimiento de doble híbrido en levadura, el gen de interés o partes del mismo (CEBO), se clona de tal manera que se expresa como una fusión con un péptido que tiene un dominio de unión al ADN. Un segundo gen, o una variedad de genes, tales como los de una genoteca de ADNc (DIANA), se clona de forma que se expresa como una fusión con un dominio de activación. La interacción de la proteína de interés con su ligando junta el péptido que se une a ADN con el dominio de activación e inicia la transcripción de los genes informadores. El primer gen informador seleccionará células de levadura que contienen proteínas que interaccionan (este informador es por lo general un gen nutricional requerido para el crecimiento en medios selectivos). El segundo informador se utiliza para la confirmación y aunque se expresa como respuesta a proteínas que interaccionan, generalmente no es necesario para el crecimiento.

La naturaleza de los genes y las proteínas que interaccionan con BNO1 también se puede estudiar de tal manera que estos ligandos también pueden ser dianas para el descubrimiento de fármacos. Son de particular interés las proteínas que interaccionan con BNO1 que se dirigen a una diana para la ubicuitinación y la posterior degradación mediante las ligasas ubicuitina-E3 que contiene BNO1.

Estudios estructurales

Proteínas recombinantes de BNO1 se pueden producir en células bacterianas, de levaduras, de insectos y/o de mamífe-

ro, y se utilizan en estudios cristalográficos y de RMN. Junto con la formación de modelos moleculares de la proteína, se puede facilitar el diseño de fármacos dirigido por la estructura.

EJEMPLO 11: Generación de anticuerpos policlonales contra BNO1

La información de la secuencia de nucleótidos y aminoácidos de BNO1 permite la producción de anticuerpos, que se unen selectivamente a la proteína de BNO1 o a fragmentos de la misma. Después de la identificación de mutaciones en el gen, los anticuerpos también se pueden preparar para que se unan selectivamente y distingan una proteína mutante de una proteína normal. Los anticuerpos específicos de epítopos mutagenizados son especialmente útiles en ensayos en cultivos celulares para la detección de células malignas en diferentes estadios del desarrollo tumoral. Estos anticuerpos también se pueden usar para detectar células malignas que se han tratado con agentes farmacéuticos, para evaluar el potencial terapéutico del agente.

Para preparar anticuerpos policlonales, se pueden diseñar péptidos cortos homólogos a la secuencia de aminoácidos de BNO1. Tales péptidos tienen típicamente una longitud de 10 a 15 aminoácidos. Estos péptidos se deben diseñar en regiones de menor homología con el ortólogo de ratón, para evitar interacciones cruzadas con especies en experimentos adicionales posteriores, tales como la producción de anticuerpos monoclonales. Los péptidos sintéticos se pueden conjugar entonces con biotina (Sulfo-NHS-LC biotina) usando protocolos convencionales suministrados con los kits disponibles comercialmente, tales como el kit de PIERCE® (PIERCE). Los péptidos biotinilados forman posteriormente complejos con avidina en solución y para cada complejo peptídico, se inmunizan 2 conejos con 4 dosis de antígeno (200 µg por dosis) a intervalos de tres semanas entre las dosis. La dosis inicial se mezcla con adyuvante completo de Freund, mientras que las dosis posteriores se combinan con inmuno-adyuvante de Freund. Después de completar la inmunización, a los conejos se les extrae sangre para el ensayo y la reactividad de los sueros se somete a ensayo mediante transferencia de manchas con diluciones en serie de los péptidos originales. Si los conejos muestran una reactividad significativa en comparación con el suero previo a la inmunización, son sacrificados posteriormente y la sangre se recoge de tal manera que se pueda separar el suero inmune para experimentos adicionales.

EJEMPLO 12: Generación de anticuerpos monoclonales específicos de BNO1

Los anticuerpos monoclonales se pueden preparar para BNO1 de la siguiente manera. Un inmunógeno que comprende proteína intacta de BNO1 o péptidos de BNO1 (de tipo silvestre o mutante) se inyecta en adyuvante de Freund a ratones, en donde cada ratón recibe cuatro inyecciones de 10 a 100 µg de inmunógeno. Después de la cuarta inyección, muestras de sangre tomadas de los ratones se examinan para estudiar la presencia de anticuerpos para el inmunógeno. Los ratones inmunes se sacrifican, se extirpan sus bazos y se preparan suspensiones de células individuales (Harlow y Lane, 1988). Las células del bazo sirven como fuente de linfocitos, que se fusionan a continuación con una célula asociada de mieloma de crecimiento permanente (Kohler y Milstein, 1975). Las células se siembran en placas con una densidad de 2×10^5 células/pocillo en placas de 96 pocillos y los pocillos individuales se examinan para estudiar el crecimiento. Estos pocillos se analizan a continuación para estudiar la presencia de anticuerpos específicos de BNO1 mediante ELISA o RIA, utilizando proteína diana de BNO1 de tipo silvestre o mutante. Las células en los pocillos positivos se expanden y se subclonian para establecer y confirmar la capacidad de monoclonación. Los clones con la especificidad deseada se expanden y se cultivan como ascitis en ratones seguido de purificación, usando cromatografía de afinidad utilizando proteína A Sefarosa, cromatografía de intercambio iónico o variaciones y combinaciones de estas técnicas.

Aplicabilidad industrial

El gen BNO1 está implicado en el cáncer y basándose en su función en el proceso de ubicitinación, BNO1 también puede estar implicado en mecanismos celulares que están regulados por este proceso. Las nuevas moléculas de ADN de la presente invención son por lo tanto útiles en métodos para la detección temprana de individuos susceptibles a una enfermedad, así como en procedimientos terapéuticos asociados con estos estados de la enfermedad.

TABLA 1

Cebadores empleados para el análisis de BNO1	
Nombre del cebador	Secuencia del cebador (5' → 3')
AL-1	GTG AAG AAG GAT GAG TTC TCC
7970-1	AGC TGA GCA TCA CAA TCT CC
ESTD-F	GGA GCT TCC CCA ACT CAT AAA TGC C
ESTD-R	GCA TGA TGT CTG ATG TGG TCA GTA A

Cebadores empleados para el análisis de BNO1	
Nombre del cebador	Secuencia del cebador (5' → 3')
BNO1-2	TGC GAA GCT GCT TCA CCG AT
BNO1-3	GGC CGT ACA TGC ACT CCA CTG
130367-1	GAG AAC CTG CAG TTG TGC TG
87068-1	ATG GTG CTG CTT GTA GCA AG
BNO1-1	TGC CCA TAT GAG ATG ACG AGG
130367-2	ACA CTC AGC AGT GGA CAC TTG
Ciclofilina-F ¹	GGC AAA TGC TGG ACC CAA CAC AAA
Ciclofilina-R ¹	CTA GGC ATG GGA GGG AAC AAG GAA
APRT-F ¹	GAC TGG GCT GCG TGC TCA TCC
APRT-R ¹	AGG CCC TGT GGT CAC TCA TAC TGC
ARN Polimerasa II-F ¹	AGG GGC TAA CAA TGG ACA CC
ARN Polimerasa II-R ¹	CCG AAG ATA AGG GGG AAC TAC T
BNO1 (Exón 5-7)-F ¹	CCG GCG GGA GGC AGG AGG AGT
BNO1 (Exón 5-7)-R ¹	GCG GCG GTA GGT CAG GCA GTT GTC
BNO1 (Isoforma 1)-F ¹	TGC GAA GCT GCT TCA CCG AT
BNO1 (Isoforma 1)-R ¹	GGC CGT ACA TGC ACT CCA CTG
BNO1 (Isoforma 2)-F ¹	GTG AAG TCG GGA CGT TTT GTG A
BNO1 (Isoforma 2)-R ¹	CCG TGG TGG GGC CCT TTG TGG

Nota: ¹ Estos cebadores se marcaron en sus extremos 5' con HEX. La isoforma 1 de BNO1 carece de exón 2.5 (SEQ ID NO: 1). La isoforma 2 de BNO1 contiene el exón 2.5 (SEQ ID NO: 3).

TABLA 2

Sitios de corte y empalme del gen BNO1						
Exón	Tamaño (pb)	Sitio 3' de corte y empalme (intrón/exón)	Potencia de consenso (%)	Sitio 5' de corte y empalme (intrón/exón)	Potencia de consenso (%)	Tamaño del intrón (pb)
1	343	5'UTR		TGCCGTGAGG/gtgagcgccg	83,03	23042

Sitios de corte y empalme del gen <i>BNO1</i>						
Exón	Tamaño (pb)	Sitio 3' de corte y empalme (intrón/exón)	Potencia de consenso (%)	Sitio 5' de corte y empalme (intrón/exón)	Potencia de consenso (%)	Tamaño del intrón (pb)
2	72	c t tg tac ag /AGTATGGTGT	94,28	TATGCGAAGC/ gtgagtgaat	75,36	1797
2.5	87	gtctg t tc a g/GTATAAACCC	90,0	TACACCTGCC/ gtatgtacct	66,97	11160
3	77	c t c c t g ta g /TGCTTCACCG	78,70	GAACGTGGTG/ gtaagtcccg	92,15	3408
4	168	c t c c t g ta g /GTGGACGGCC	84,95	CCACATCCAG/ gtgtgtcag	85,40	646
5	75	aacactga a g/ATTGTGAAGA	63,39	GAGGCAGGAG / gtgagccac	90,87	6612
6	110	c t ttt g ga a g/GAGTTTCGGA	85,65	GTCAGTACGA/ gtgagtgcgg	76,46	697
7	154	ctccccac a g/CAACTGCCTG	85,32	CAAGATCACG/ gtgagtggcg	88,50	1017
8	401	tgctccac a g/GGCGACCCCCA	89,22	GCAGGATGTG/ gtaaggatg	87,59	2375
9	2174	ttctg c t c ag/TTTTTATGGC	90,62	3'UTR		

TABLA 3

Cebadores empleados para el análisis de SSCP de <i>BNO1</i>				
Exón	Cebador 1 (5' → 3')	Cebador 2 (5' → 3')	[MgCl ₂]	Tamaño del producto (pb)
1a	GCGCTGGAGCGTGCGCACA	AGCTCGGGCGGCAGCTCCA	2,0 mM	269
1b	GGTCGGGGCGGGCTTG	GCCTCCACCTGGCAGGGA	2,0 mM	252
2	CTGTCGCGTTATGAGTTTG	GTACAAAGTTAACATGGATGGT	2,0 mM	168
2.5	AGGCATTGGGTCGTATTAC	AGAACGCCAAAGCTCGCAGGA	1,5 mM	198
3	GGCACGCTGGGTCTAACAC	CCTGCCCGTGCACAGACCT	1,5 mM	167
4	CTCATGGACCTTGCCCACATCT	GTCTGCAGCTGAGAACAGCAC	1,0 mM	290
5	GTGATGGACTCTGTTCTCAC	AGGTCCGCACCATATGAACAC	2,0 mM	170
6	CACAGCCTCCTGTCATATGGA	ACCCCAAGCACCGAGCAGGA	1,5 mM	187
7	GGCGTTCTCAGTCCTGCCT	CCCTGACTCCACAGCCCAC	1,5 mM	284
8a	CTGGCCTGAGCCCTGCTGA	ACCCTCTCGCGCACCTCCA	1,0 mM	171
8b	CAATGAGCTCTCCGCATC	CCATGCTGTCCCACCTCA	1,5 mM	354
9	AGAATGCTGTACGTGGCGTG	AGGAGGTGAGGGACTGAATG	1,0 mM	292

Cebadores empleados para el análisis de SSCP de <i>BNO1</i>				
Exón	Cebador 1 (5' → 3')	Cebador 2 (5' → 3')	[MgCl ₂]	Tamaño del producto (pb)
<i>Nota:</i> Todos los cebadores se marcaron en sus extremos 5' con HEX.				

Referencias

Las referencias citadas en esta memoria se enumeran en las páginas siguientes y se incorporan en esta memoria como referencia.

- 5 Altschul, SF. *et al.* (1997). Nucleic Acids Res. 25: 3389-3402.
 Brenner, AJ. y Aldaz CM. (1995). Cancer Res. 55: 2892-2895.
 Baumeister, W. *et al.* (1998). Cell 92: 367-380.
 Callen, DF. *et al.* (1990). Ann. Genet. 33: 219-221.
 Callen, DF. *et al.* (1995). Genomics 29: 503-511.
 10 Cenciarelli, C. *et al.* (1999). Curr. Biol. 9 : 1177-1179.
 Chen, T. *et al.* (1996). Cancer Res. 56: 5605-5609.
 Cleton-Jansen, A-M. *et al.* (1995). Br. J. Cancer 72: 1241-1244.
 Cockman, ME. *et al.* (2000). J. Biol. Chem. 275: 25733-25741.
 Cole, SP. *et al.* (1984). Mol. Cell Biol. 62: 109-120.
 15 Cote, RJ. *et al.* (1983). Proc. Natl. Acad. Sci. USA 80: 2026-2030.
 Culver, K. (1996). Gene Therapy : A Primer for Physicians. Segunda edición. (Mary Ann Liebert).
 Devilee, P. *et al.* (1991). Oncogene 6: 1705-1711.
 Devilee, P. y Cornelisse, CJ. (1994). Biochimica et Biophysica Acta 1198: 113-130.
 Doggett, NA. *et al.* (1995). Nature 377 Supl: 335-365.
 20 Elston, CW. y Ellis, IO. (1990). Histopathology 16: 109-118.
 Esteller, M. *et al.* (2000). J. Natl. Cancer Inst. 92: 564-569.
 Fearon, ER. y Vogelstein, B. (1990). Cell 61: 759-767.
 Friedman, T. (1991). In Therapy for Genetic Diseases. T Friedman (Ed). Oxford University Press, págs. 105-121.
 Futreal, PA. *et al.* (1994). Science 266: 120-122.
 25 Goldman, CK. *et al.* (1997). Nature Biotechnology 15: 462-466.
 Haas, AL, y Siepmann, TJ. (1997). FASEB 11: 1257-1268.
 Hall, JM. *et al.* (1990). Science 250: 1684-1689.
 Harlow, E. y Lane, D. (1988). Antibodies: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
 30 Heller, RA. *et al.* (1997). Proc. Natl. Acad. Sci. USA 94: 2150-2155.
 Herman, JG. *et al.* (1998). Proc. Natl. Acad. Sci. USA 95: 6870-6875.
 Hershko, A. y Ciechanover, A. (1998). Annu. Rev. Biochem. 67: 425-479.
 Hochstrasser, M. (1996). Ann. Rev. Gene. 30: 405-439.

- Huse, WD. *et al.* (1989). Science 246: 1275-1281.
- Kipreos, ET. y Pagano, M. (2000). Genome Biology 1: revisiones 3002.1-3002.7.
- Kohler, G. y Milstein, C. (1975). Nature 256: 495-497.
- Kozbor, D. *et al.* (1985). J. Immunol. Methods 81:31-42.
- 5 Longmire, JL. *et al.* (1993). GATA 10: 69-76.
- Lopez Salon, M. *et al.* (2000). J. Neurosci. Res. 62: 302-310.
- McCormick, MK. *et al.* (1993). Proc. Natl. Acad. Sci. USA 90: 1063-1067.
- Miki, Y. *et al.* (1994). Science 266: 66-71.
- Miki, Y. *et al.* (1996). Nature Genet. 13: 245-247.
- 10 Ohh, M. *et al.* (2000). Nat. Cell Biol. 2: 423-427.
- Ohtani-Fujita, N. *et al.* (1997). Cancer Genet. Cytogenet. 98:43-49.
- Orlandi, R. *et al.* (1989). Proc. Natl. Acad. Sci. USA 86: 3833-3837.
- Peters, JM. (1998). Curr. Opin. Cell Biology 10: 759-768.
- Prowse, AH. *et al.* (1997). Am. J. Hum. Genet. 60:765-771.
- 15 Radford, DM. *et al.* (1995). Cancer Res. 55: 3399-3405.
- Riethman, HC. *et al.* (1989). Proc. Natl. Acad. Sci. USA 86: 6240-6244.
- Saito, H. *et al.* (1993). Cancer Res. 53: 3382-3385.
- Sambrook, J. *et al.* (1989). Molecular cloning: a laboratory manual. Segunda edición. (Cold Spring Harbour Laboratory Press, Nueva York).
- 20 Scharf, D. *et al.* (1994). Results Probl. Cell Differ. 20: 125-162.
- Schena, M. *et al.* (1996). Proc. Natl. Acad. Sci. USA 93: 10614-10619.
- Semenza, GL. (2000). Gene Dev. 14: 1983-1991.
- Sharan, SK. *et al.* (1997). Nature 386: 804-810.
- Shimura, H. *et al.* (2001). Science 293: 263-269.
- 25 Soares, MB. *et al.* (1994). Proc. Natl. Acad. Sci. USA 91: 9228-9232.
- Wang, GL. *et al.* (1995). Proc. Natl. Acad. Sci. USA 92: 5510-5514.
- Weber, JL. y May, PE. (1989). Am. J. Hum. Genet. 44: 388-396.
- Whitmore, SA. *et al.* (1994). Genomics 20: 169-175.
- Whitmore, SA. *et al.* (1998). Genomics 50: 1-8.
- 30 WHO. (1981). Histological Typing of Breast Tumours. Segunda edición. (Geneva).
- Winston, JT. *et al.* (1999). Current Biology 9: 1180-1182.
- Winter, G. *et al.* (1991). Nature 349: 293-299.
- Wooster, R. *et al.* (1995). Nature 378: 789-791.
- Wooster, R. *et al.* (1994). Science 265: 2088-2090.
- 35 Wyman, A. y White, R. (1980). Proc. Natl. Acad. Sci. USA 77: 6754-6758.
- Zaibo, L. *et al.* (2001). J. Biol. Chem. (Artículos en prensa, Manuscrito M108269200).

LISTADO DE SECUENCIAS

<110> Bionomics Limited
<120> BNO1
<130> P3
5 <160> 41
<170> PatentIn versión 3.1
<210> 1
<211> 3574
<212> ADN
10 <213> Homo sapiens
<400> 1
ggcatggcgg tgtgtgctcg cctttgcccgt cgccgcggatg tcggcgccgc
60
cagcagcgcc ggggccccggc cgagacggcg gcggccgaca gcgagccgga cacagacccc
120
gaggaggagc gcattcgaggc tagcgccggg gtcggggcg gcttgtgcgc gggccctcg
180
ccgcgcgcggc cgcgctgctc gctgctggag ctgcgcggc agctgctggt ggagatcttc
240
gcgtcgctgc cgggcacgga cctaccaggc ttggccagg tctgcacgaa gttccggcgc
300
atccctccaca ccgacaccat ctggaggagg cgttgccgtg aggagtatgg tgtttgcgaa
360
aacttgcgga agctggagat cacaggcgtg tcttgcggg acgtctatgc gaagctgctt
420
caccgatata gacacatttt gggattgtgg cagccagata tcgggccata cggaggactg
480
ctgaacgtgg tggtgacgg cctgttcatac atcgggtgga tgtacctgcc tccccatgac
540
ccccacgtcg atgaccatat gagattcaag cctctgttca ggatccacct gatggagagg

600
aaggctgcca cagtggagtg catgtacggc cacaaagggc cccaccacgg ccacatccag
660
attgtgaaga aggatgagtt ctccaccaag tgcaaccaga cggaccacca caggatgtcc
720
ggcgggaggc aggaggagtt tcggacgtgg ctgagggagg aatggggcg cacgctggag
780
gacatcttcc acgagcacat gcaggagctc atcctgatga agttcatcta caccagtcag
840
tacgacaact gcctgaccta ccgcgcatac tacctgccgc ccagccccc cgacgacctc
900
atcaaggctg gcctttcaa aggtacctat ggcagccacg gcctggagat tgtatgctc
960
agcttccacg gccggcgtgc cagggcacc aagatcacgg gcgacccaa catccccgt
1020
gggcagcaga cagtggagat cgacctgagg catcgatcc agctgccga cctcgagaac
1080
cagcgcaact tcaatgagct ctccgcatac gtcctggagg tgcgcgagag ggtgcgcag
1140
gagcagcagg aaggcgggca cgaggcgggc gagggtcgtg gccggcaggg ccccccggag
1200
tcccagccaa gccctgcccc gcccaggca gaggcgcccc gcaaggccc agatggaca
1260
cctggtgagg atggtggcga gcctggggat gccgtagctg cggccgagca gcctgcccag
1320
tgtggcagg ggcagccgtt cgtgctgccc gtggcgtga gctccaggaa tgaggactac
1380

ccccgaacct gcaggatgtg ttttatggc acaggcctca tcgcgggcca cggcttcacc
1440
agccctgaac gcaccccccgg ggtttcatc ctcttcgatg aggaccgctt cgggttcgtc
1500
tggctggagc tgaaatcctt cagcctgtac agccgggtcc aggccacctt ccggAACGCA
1560
gatgcgcgt ccccacaggc cttcgatgag atgctcaaga acattcagtc cttcacctcc
1620
tgaccggcca catccttgcc gccacatccc gggtggctct ggggctctga actctgacct
1680
gtgaatagaa gcagcatgca ctttgaaat cggccctttt gaccagaacg cacacctcg
1740
cggggggccc agtccagcca ccccccagca ctttatgttag agagtgtgac atagacctgc
1800
atatttgtca gtgccatgat ggaagaagct gagcatgtct tacaaaaaac agagagaacg
1860
agcctgaata cagcagatgt aggggacacgc cgtgggaccg cgtgagaatt gaagcggtgg
1920
ggttcccgca ccctgggctg gctgggtgtt ttctcgaa gcaggaccct cctgactgg
1980
gctttcctg tgagcggata gagtgataga ctgggtcgatg tgtgagacgc atgtgctcca
2040
ccccactcct tttggggaa gccaggcaac agtggcctct gggaggggggt caggaagagg
2100
cgaacagctc aggcagcgca ggtgtatgg gcacagtacg cagagcaacg tcggaaagtt
2160
ggtaggatct caggcttggg gccgggactc tggagtgaat cccatttct ctacccggctt
2220

gcttggagtt tggacagaag catttcacct ctgatcttag cttccccacc tgtggagtgg
2280
gtttagtgac ctgagtcact aggaaatgtc acctgaatgc acagcccagc ccatgcacct
2340
gccccagccc ctccagcttt ggagccaagg ccatcggtcc agccacttga ctgtcctcga
2400
cggcctgttc cagacagggc gtttgttttg tccatgcctt cctccctgca cgcacacggc
2460
gtcaaaadca agctgccggc cactgtctcc agaacgcaag gctccaggcc cgtgtgtctg
2520
aagcagttag tggtccacac aggtgccagg agtgccata tgagatgacg aggaaacccc
2580
tttgcaggtg aggggacagc tttctagaaa agccacaccc gcatctgggg acacactttg
2640
gaaagtggga ccctccagcc tggagacccc atggactgtat gcctccactg ctgtgtgccc
2700
catgttgtt taacacctgc gtgtggggac cccatctgag gtcttggctg aggttggcat
2760
ctcctgaaga acagagagca cgggtgtccag agctggccct tccccagcc cacagccagc
2820
tccgtgcccc agtgggcgtc cccagcgagc cttccctctc tgccgcttgc ctttgtgtct
2880
gggctgctcc aagtccctgt gctgggcacc ctggacacgt cctgctggtg agggacctcg
2940
ggaaggtgac agtctgtgtg ctttgtgtg gagaccaacc tgaggatgtc ctggaaatg
3000
ttttcctgat gaatttctcc ttgactggcc tttaaagaac ataagaattc ccattgcccc

3060

gcctcagtgc atttggcaaa tgcttacttt gcttcccaga gtcagagaat tggcaaagg

3120

tcctaaatgg taatctggcc ggctgggag aaagactcac gagaaaagcc agtggagaaa

3180

gcccccttcc agggcggcag cagcgggaggc cacgcagacc ccgaggcgca cctgctggct

3240

cttgtgtgtg gccccagttt cttagggctt ttgcagcatt agcctacaag ctttgtcact

3300

ccctgccctc tgtggtggtc actgtttttc tctcttgcca aatgaggcag tctctgagtg

3360

acggtgactg tggccttgaa gcctggagga ctgtggca tgttagactgg caccttgaag

3420

attcaccatt gtttaataa aatcaagcaa atgctttttt accaagagcc cgagcctcgc

3480

tctaaggac gcagtcctag aggctgtcccc tttggggctt gaagagcaca ctgtggacg

3540

cacgtgcttc tgattaaagg aatctcagat ctca

3574

<210> 2

<211> 539

<212> PRT

5 <213> Homo sapiens

<400> 2

Met Ala Val Cys Ala Arg Leu Cys Gly Val Gly Pro Ser Arg Gly Cys
1 5 10 15Arg Arg Arg Gln Gln Arg Arg Gly Pro Ala Glu Thr Ala Ala Ala Asp
20 25 30

Ser Glu Pro Asp Thr Asp Pro Glu Glu Glu Arg Ile Glu Ala Ser Ala

ES 2 433 012 T3

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

ES 2 433 012 T3

Thr Leu Glu Asp Ile Phe His Glu His Met Gln Glu Leu Ile Leu Met
260 265 270

Lys Phe Ile Tyr Thr Ser Gln Tyr Asp Asn Cys Leu Thr Tyr Arg Arg
275 280 285

Ile Tyr Leu Pro Pro Ser Arg Pro Asp Asp Leu Ile Lys Pro Gly Leu
290 295 300

Phe Lys Gly Thr Tyr Gly Ser His Gly Leu Glu Ile Val Met Leu Ser
305 310 315 320

Phe His Gly Arg Arg Ala Arg Gly Thr Lys Ile Thr Gly Asp Pro Asn
325 330 335

Ile Pro Ala Gly Gln Gln Thr Val Glu Ile Asp Leu Arg His Arg Ile
340 345 350

Gln Leu Pro Asp Leu Glu Asn Gln Arg Asn Phe Asn Glu Leu Ser Arg
355 360 365

Ile Val Leu Glu Val Arg Glu Arg Val Arg Gln Glu Gln Gln Glu Gly
370 375 380

Gly His Glu Ala Gly Glu Gly Arg Gly Arg Gln Gly Pro Arg Glu Ser
385 390 395 400

Gln Pro Ser Pro Ala Gln Pro Arg Ala Glu Ala Pro Ser Lys Gly Pro
405 410 415

Asp Gly Thr Pro Gly Glu Asp Gly Gly Glu Pro Gly Asp Ala Val Ala
420 425 430

Ala Ala Glu Gln Pro Ala Gln Cys Gly Gln Gly Gln Pro Phe Val Leu
435 440 445

Pro Val Gly Val Ser Ser Arg Asn Glu Asp Tyr Pro Arg Thr Cys Arg
450 455 460

Met Cys Phe Tyr Gly Thr Gly Leu Ile Ala Gly His Gly Phe Thr Ser
465 470 475 480

ES 2 433 012 T3

Pro Glu Arg Thr Pro Gly Val Phe Ile Leu Phe Asp Glu Asp Arg Phe
485 490 495

Gly Phe Val Trp Leu Glu Leu Lys Ser Phe Ser Leu Tyr Ser Arg Val
500 505 510

Gln Ala Thr Phe Arg Asn Ala Asp Ala Pro Ser Pro Gln Ala Phe Asp
515 520 525

Glu Met Leu Lys Asn Ile Gln Ser Leu Thr Ser
530 535

<210> 3

<211> 3661

<212> ADN

5 <213> Homo sapiens

<400> 3
ggcatggcgg tgtgtgctcg cctttgcggc gtggggccgt cgcgccgatg tcggcgccgc

60

cagcagcgcc ggggccccggc cgagacggcg gcggccgaca gcgagccgga cacagacccc
120

gaggaggaggc gcatcgaggc tagcgccggg gtcggggggcg gtttgtgcgc gggccctcg
180

ccggccgcccc cgcgctgctc gctgctggag ctgcccggc agctgctggt ggagatcttc
240

gcgtcgctgc cgggcacgga cctaccagc ttggccagg tctgcacgaa gttccggcgc
300

atccctccaca ccgacacccat ctggaggagg cgttgccgtg aggagtatgg tgtttgcga
360

aaccttgcga agctggagat cacaggcgtg tcttgccgg acgtctatgc gaagcgtata
420

aaccttcgcg tgaagtccgg acgttttgcg aaaattctcc ctgattatga gcacatggcg
480

tacagagacg tttacacacctg cctgcttcac cgatatacgac acatTTTggg attgtggcag
540
ccagatatacg ggcatacgg aggactgctg aacgtggtgg tggacggcct gttcatcatc
600
gggtggatgt acctgcctcc ccatgacccc cacgtcgatg accctatgag attcaaggct
660
ctgttcagga tccacactgat ggagaggaag gctgccacag tggagtgcac gtacggccac
720
aaagggcccc accacggcca catccagatt gtgaagaagg atgagttctc caccaagtgc
780
aaccagacgg accaccacag gatgtccggc gggaggcagg aggagttcg gacgtggctg
840
agggaggaat gggggcgcac gctggaggac atcttccacg agcacatgca ggagctcatc
900
ctgatgaagt tcatactacac cagtcagtagc gacaactgcc tgacctaccc ccgcacatctac
960
ctgcgcggca gcccggccga cgacctcatc aagcctggcc tcttcaaagg tacctatggc
1020
agccacggcc tggagattgt gatgctcagc ttccacggcc ggcgtgccag gggcaccaag
1080
atcacgggcg accccaaacat ccccgctggg cagcagacag tggagatcga cctgaggcat
1140
cgatccagc tgcccgacct cgagaaccag cgcaacttca atgagctctc ccgcacatgtc
1200
ctggaggtgc gcgagagggt ggcgcaggag cagcaggaag gcgggcacga ggcggggcag
1260
ggtcgtggcc ggcaggggccc cggggagtcc cagccaaagcc ctgcccagcc cagggcagag
1320

ES 2 433 012 T3

gccccccagca agggcccaga tggacacct ggtgaggatg gtggcgagcc tggggatgcc
1380
gtagctgcgg ccgagcagcc tgcccagtgt gggcaggggc agccgttcgt gctgcccgta
1440
ggcgtgagct ccaggaatga ggactacccc cgaacctgca ggatgtgttt ttatggcaca
1500
ggcctcatcg cggccacgg ctaccaggc cctgaacgca ccccccgggt ctcatcctc
1560
ttcgatgagg accgcttgg gttcgtctgg ctggagctga aatccttcag cctgtacagc
1620
cgggtccagg ccaccttccg gaacgcagat gcgcgtccc cacaggcctt cgatgagatg
1680
ctcaagaaca ttcaagtccct cacccctga ccggccacat ctttgccgcc acatcccg
1740
tggctctggg gctctgaact ctgacacctg aatagaagca gcatgcactt tggaaatccg
1800
gcctttgac cagaacgcac acctcgctgg ggggcccagt ccagccaccc cccagcactt
1860
tatgttagaga gtgtgacata gacctgcata tttgtcagtgc ccatgtatggaa agaagctgag
1920
catgtcttac caaaaacaga gagaacgagc ctgaatacag cagatgttagg ggacagccgt
1980
gggaccgcgt gagaattgaa gcgggtgggt tcccgccaccc tgggctggct ggtggtttc
2040
tcgggaagca ggaccctcct gactggtgct cttcctgtga gcggatagag tgatagactg
2100
ggtcgtgtgt gagacgcatg tgctccaccc cactcccttt gggggaaagcc aggcaacagt

ES 2 433 012 T3

2160
ggcctctggg agggggtcag gaagaggcga acagctcagg cagcgcaggt gtgatggca
2220
cagtacgcag agcaagctcg ggaagtttgtt aggatctcag gcttggggcc gggactctgg
2280
agtgaatccc catttctcta ccggcttgct tggagttgg acagaagcat ttcacctctg
2340
atctcagctt cccccacctgt ggagtgggtt tagtgcacctg agtcactagg gaatgtcacc
2400
tgaatgcaca gcccagccca tgcacctgcc ccagccctc cagctttgga gccaaaggcca
2460
tcgttccagc cacttgactg tcctcgacgg cctgttccag acagggcggtt tgtttgc
2520
atgccttcct ccctgcacgc acacggcgtc aaaaccaagc tgccggccac tgtctccaga
2580
acgcaaggct ccaggcccggt gtgtctgaag cagttagtgg tccacacagg tgccaggagt
2640
gccccatatga gatgacgagg aaaccccttt gcaggtgagg ggacagcttt cttagaaaagc
2700
cacacctgca tctggggaca cactttggaa agtgggaccc tccagcctgg agaccccatg
2760
gactgatgcc tccactgctg tgtgccccat gttgtgttaa cacctgcgtg tggggacccc
2820
atctgaggtc ttggctgagg ttggcatctc ctgaagaaca gagagcacgg tgtccagagc
2880
tggcccttcc cccagcccac agccagctcc gtgcccaggt gggcgcccc agcgagcctt
2940
ccctctctgc cgcttgcct ttttgtgg ctgctccaag tcccttgtgtt gggcaccctg

3000

gacacgtcct gctggtgagg gacctcgaaa aggtgacagt ctgtgtgcct tggtgtggag

3060

accaacctga ggatgtcctg gaaaaatgttt tcctgatgaa ttttccttgc actggccctt

3120

aaagaacata agaattccca ttgcccccagcc tcagtcatt tggcaaatgc ttactttgt

3180

tcccaagatc agagaattgg caaagggtcc taaatggtaa tctggccggc ctgggagaaa

3240

gactcacgag aaaagccagt ggagaaagcg cccttcagg gcggcagcag cgggagccac

3300

gcagaccccg aggccgacact gctggcttt gtgtgtggcc ccagttctta gcggctttt

3360

cagcatttgc ctacaagctt tgtcaactccc tgccctctgt ggtggtcact gttttctct

3420

cttgccaaat gaggcagtct ctgagtgacg gtgactgtgg ctttgaagcc tggaggactg

3480

ttggcatgt agactggcac cttgaagatt caccattgtt taaataaaat caagcaaatg

3540

ctttttacc aagagccccga gcctcgctct aagggacgca gtccttagagg cgtgccctt

3600

ggggcttcaa gagcacactg tgggacgcac gtgcttctga ttaaaggaat ctcagatctc

3660

a

3661

<210> 4

<211> 568

<212> PRT

5 <213> Homo sapiens

<400> 4

ES 2 433 012 T3

Met Ala Val Cys Ala Arg Leu Cys Gly Val Gly Pro Ser Arg Gly Cys
1 5 10 15

Arg Arg Arg Gln Gln Arg Arg Gly Pro Ala Glu Thr Ala Ala Ala Asp
20 25 30

Ser Glu Pro Asp Thr Asp Pro Glu Glu Glu Arg Ile Glu Ala Ser Ala
35 40 45

Gly Val Gly Gly Leu Cys Ala Gly Pro Ser Pro Pro Pro Pro Arg
50 55 60

Cys Ser Leu Leu Glu Leu Pro Pro Glu Leu Leu Val Glu Ile Phe Ala
65 70 75 80

Ser Leu Pro Gly Thr Asp Leu Pro Ser Leu Ala Gln Val Cys Thr Lys
85 90 95

Phe Arg Arg Ile Leu His Thr Asp Thr Ile Trp Arg Arg Cys Arg
100 105 110

Glu Glu Tyr Gly Val Cys Glu Asn Leu Arg Lys Leu Glu Ile Thr Gly
115 120 125

Val Ser Cys Arg Asp Val Tyr Ala Lys Arg Ile Asn Pro Arg Val Lys
130 135 140

Ser Gly Arg Phe Val Lys Ile Leu Pro Asp Tyr Glu His Met Ala Tyr
145 150 155 160

Arg Asp Val Tyr Thr Cys Leu Leu His Arg Tyr Arg His Ile Leu Gly
165 170 175

Leu Trp Gln Pro Asp Ile Gly Pro Tyr Gly Leu Leu Asn Val Val
180 185 190

Val Asp Gly Leu Phe Ile Ile Gly Trp Met Tyr Leu Pro Pro His Asp
195 200 205

ES 2 433 012 T3

Pro His Val Asp Asp Pro Met Arg Phe Lys Pro Leu Phe Arg Ile His
210 215 220

Leu Met Glu Arg Lys Ala Ala Thr Val Glu Cys Met Tyr Gly His Lys
225 230 235 240

Gly Pro His His Gly His Ile Gln Ile Val Lys Lys Asp Glu Phe Ser
245 250 255

Thr Lys Cys Asn Gln Thr Asp His His Arg Met Ser Gly Gly Arg Gln
260 265 270

Glu Glu Phe Arg Thr Trp Leu Arg Glu Glu Trp Gly Arg Thr Leu Glu
275 280 285

Asp Ile Phe His Glu His Met Gln Glu Leu Ile Leu Met Lys Phe Ile
290 295 300

Tyr Thr Ser Gln Tyr Asp Asn Cys Leu Thr Tyr Arg Arg Ile Tyr Leu
305 310 315 320

Pro Pro Ser Arg Pro Asp Asp Leu Ile Lys Pro Gly Leu Phe Lys Gly
325 330 335

Thr Tyr Gly Ser His Gly Leu Glu Ile Val Met Leu Ser Phe His Gly
340 345 350

Arg Arg Ala Arg Gly Thr Lys Ile Thr Gly Asp Pro Asn Ile Pro Ala
355 360 365

Gly Gln Gln Thr Val Glu Ile Asp Leu Arg His Arg Ile Gln Leu Pro
370 375 380

Asp Leu Glu Asn Gln Arg Asn Phe Asn Glu Leu Ser Arg Ile Val Leu
385 390 395 400

Glu Val Arg Glu Arg Val Arg Gln Glu Gln Gln Glu Gly Gly His Glu
405 410 415

Ala Gly Glu Gly Arg Gly Arg Gln Gly Pro Arg Glu Ser Gln Pro Ser
420 425 430

ES 2 433 012 T3

Pro Ala Gln Pro Arg Ala Glu Ala Pro Ser Lys Gly Pro Asp Gly Thr
435 440 445

Pro Gly Glu Asp Gly Gly Glu Pro Gly Asp Ala Val Ala Ala Ala Glu
450 455 460

Gln Pro Ala Gln Cys Gly Gln Gly Gln Pro Phe Val Leu Pro Val Gly
465 470 475 480

Val Ser Ser Arg Asn Glu Asp Tyr Pro Arg Thr Cys Arg Met Cys Phe
485 490 495

Tyr Gly Thr Gly Leu Ile Ala Gly His Gly Phe Thr Ser Pro Glu Arg
500 505 510

Thr Pro Gly Val Phe Ile Leu Phe Asp Glu Asp Arg Phe Gly Phe Val
515 520 525

Trp Leu Glu Leu Lys Ser Phe Ser Leu Tyr Ser Arg Val Gln Ala Thr
530 535 540

Phe Arg Asn Ala Asp Ala Pro Ser Pro Gln Ala Phe Asp Glu Met Leu
545 550 555 560

Lys Asn Ile Gln Ser Leu Thr Ser
565

<210> 5

<211> 2340

<212> ADN

5 <213> Homo sapiens

<400> 5

tatttgtttt gtagacaggg tctcgctgta ttgcccaaggc cggcttcgag ctccctggcct

60

cgattgatac tccccctgg gcctccaaga gatgggtcc gagggcgagcc cacggcgacg

120

tgccggctg ctcaggttag aggacgcctt cgcggtcacc acccgccgac ctgggagacg

180

accccgctca gcggcctcg cgagccccag ctggagcagg cgtgcgcggc tcccagcagg
240
tgcaggaaca ggccccttt gggcgccgccc gcatggcagg gctcccttt ccagaccggg
300
cgcgcatccc cggatctttt gggccgcgccc ggccgcggcc cctccaagcc ctccccgggg
360
cttcgcagg gagctcgaaa tccccaaagg tccctgcaga gctccgcagg tcgggccttt
420
tggttaccat aaggcgaga cgttggaaacg cgcgttgtt caatggacaa aagggtttct
480
aggcgccctt tggggttctg gctgctgcct ctgtatttgg aggctgttaag gcgcatcttt
540
ctactcaccg gccggcgccgg cacagttcg ggccgcggaa gcgggacgca cgggcgcgag
600
ggcgaccccc tatctccaca aaagccgcgg cgcgaagtgg tcgcccggca gcctcgtag
660
cgcaatggc agcggttcag ttcataatc tgaaggctgt gagttcgagc ctcacacggg
720
gcagtctaact gtttgcact cggcatcacc actttctttt ctcatgcgg tcacggcg
780
cgccctcact ctggaggggaa gggcagcagt ggggggtctc tgagggtcgcc gccccgggg
840
gagggggtgg cgccggcgccc ggggagctct acgttaggggc ggggcttaggc tctccagggg
900
gcgtggcgag ctctgcggcg ggggcgtggc tggcgctgg cggggcgccc cggcgctgg
960
gcgtgcgcac aggcggcagc agtggccgtc actggcgcc atggcggtgt gtgcgcct

1020 ttgcggcgtg ggcccgtcgc gcggatgtcg gcgcgcggcag cagcgccggg gcccgccga
1080 gacggcggcg gccgacagcg agccggacac agaccccgag gaggagcgca tcgaggctag
1140 cgccggggtc gggggcggtct tgcgcggg cccctcgccg ccgcggccgc gctgctcgct
1200 gctggagctg ccgcggcggc tgctggtgga gatcttcgcg tcgctgccgg gcacggacct
1260 acccagcttgc cccaggtct gcacgaagt ccggcgcatc ctccacacccg acaccatctg
1320 gaggaggcgt tgccgtgagg gtgagcgccg gggggtggcg gggccgggag gggcggggg
1380 tccctgccag gtggaggcct cggagctggg agtggcgccc gcggtggccc cggccgggg
1440 ccaccagttgc ggccggggc ccggcgatgt ggtgtttgg gtgtgggtgg ggagcgcccg
1500 cggtgacacc acgttgagggg ggccagggag gtatttgagg cggtagggaa gggtccgagg
1560 ggtccaagag aggccagacgg ggttagggagg ggttgagggc gtcagggagg catcgaagag
1620 gccctgacgc gggcacggg acaccacggg gccgaggccg tgccgggagc tggggctgg
1680 atccctcgag gtctgcgcgg ggccataagct gacgcctggg ggccgcctcc tctgcctcg
1740 tcttgaaggc gagccgagggg tgccccgctg gcccctgacca gagacgaggt gaacttgaag
1800 aaatgggagc aggccggcg cggtttcacg cctgtaatct cagcacttag ggaggccgag

1860

gcgggggggt cacctgaggt caggagttcg agaccagcct ggtgaagatg gtaaaacccc

1920

gtctctactt acaaaaatac aaaaattagc cgggcgtggt agcgggcgcc tctaataccca

1980

gctactcggg agggtgaggc agaagaatgg cttgaacccg ggaggcggag gttgcactga

2040

gccaaacatct gggccattgc actccaccct aggcgacaga gtgcgacttg gtctcaaaca

2100

acagcaacaa aaaaaattac gggagtaggg gaagcccagt gtcggggct cctgatggtg

2160

ggggcgtaga gagacggatg gatcacagcg gtgcgggctg gactttgcc ttcagcaata

2220

ttggatgtaa cagatcacag gaggagatgt ttatataatc tggagttcaa ggctctctct

2280

gtttaaagggt tgacagctct ttgatgttca agcagctcat atttaggtaa aaaggacagg

2340

<210> 6

<211> 2072

<212> ADN

5 <213> Homo sapiens

<400> 6

aaggaagctt tttgatctag gtttcaattc ttgggtccag ctgaccataa gctccaggc

60

ctttgtggag tcatgtcacc tctctgcact tttttctct gttttatgg agatgacagt

120

ggtgactctc ttatggact gctgttaggt gcagtggat gatgccagc atggcttggc

180

atggtacaca gcacgtggaa agctcaatgc aggtgttgc agtagcagct cttggccat

240

gtcagttct gtagttgtgt ggattagtcc ctggcggtct ttccttcaa aggctagcgg

300

agacccaagc cgaggacact gggattttgc ttgggacgtg ctgggtgttg gtgattgtga

360

aaagtgcagc tgtgggtgggg tggagtagg ggacaaaagag gaaggtgcgt tcagcaggtg

420

agggtgtgag gacaggggtt gcgggaggtg tccagggccc tgcactggc cctggccaag

480

cctagccagt ggagaaggga caatgttcac ccctcccccc atgtcttgca cggccccctc

540

ttggccttgg gctgagttga acacacaggc agcacacaggga agtacatggg gtggactggc

600

ctctggcact gtctgaaccc taacaccagt ggtgaatttg tttccatgga aacatggcac

660

tgtgtccaga caactgaatt ctgcctcacc ttgttcataa actaggatt gtctgatatt

720

ggtttgttg gtaggtttc tagagcttat tagaatagac attgcagatt attatttgt

780

aaagggtgac attgactaaa atagaataat gtcttcatcg gtgaacaagg gtgtttactg

840

aatgtggaga agtcagtgaa atctccacag tgacagatgc actctggaga tggggctgag

900

gctaggtgtg cacccccc gccagccatc agcagcctgc ccacgtctgt cgcgttatga

960

gttggatc ttaaatttct gcaaatgttt cttgttacag agtatgggtt ttgcgaaaac

1020

ES 2 433 012 T3

ttgcggaaac tggagatcac aggcggtctc tgtcggacg tctatgcga gcgtgagtga
1080
atctatttgt taccatccat gattaacttt gtaccagaag cagacagtgc acatcaatga
1140
caaataatca aagtgattta gtccacactt ttgtttctc agacaccatc ttacagtcac
1200
atttgaata gaggactgg agtaacagca ctaaaattag ggaggggaca ccgtattctc
1260
ccattctggg catcgtagat actagtgtct tttaccaggc atcgaggccc ttttagagctg
1320
agaaaatgtag gctgcaccag agcgacttgg gtgttcctgg aggatgcctg tttctctggc
1380
ttctgggcct ccagccttca gnatggact gcctgtggc attggaaat gaagctgtgg
1440
tgccctctgt ccaagccccca ggtacggaga acagcccccc tgctagagtt ctccctccct
1500
ttagcattct gtgagcaggc tgagcccccc agcccgttt catgtttctc tggcaccc
1560
gcttccaga gtcaactgtg acactcaccc atcaggcaga ggctccttgg ccccaaggctg
1620
tcttcccgcg ggtgtttctt cagtgccagg gacttccca tttcttgctc actggacgaa
1680
gctcttggct ttttggatg gtcagagggt ttctgtgagt gtctcataacc ccacgtttc
1740
atcttctgac cattcttga gcttcttttc tttttcttt ttcatttctg caaggcagtt
1800
gctctttagag tttcttaggc catttgggtc atatcatctc aaagtgacat gaagacatat
1860

ttctctctaa gagttgtcag cacaaaaagt catgtcccg tagggacaca ggtgtttggc
1920
ctgcccaggt gtttagggttt cctgtaacag ccataggagt gtacacatgg atgctccttc
1980
cagctgtggg ggtcggaggt gttgcctgag cactgagtcc tcctgtcctg tgggagaggc
2040
catccgggca taggcaggca ggaggcagtg tg
2072
<210> 7
<211> 2087
<212> ADN
5 <213> Homo sapiens
<400> 7
agagttgtca gcaccaaaaag tcatgtcccg gtagggacac aggtgtttgg cctgcccagg
60
tgttagggtt tcctgtaaca gccataggag tgtacacatg gatgctcctt ccagctgtgg
120
gggtcggagg tgttgcctga gcactgagtc ctctgtcct gtgggagagg ccatccgggc
180
ataggcaggc aggaggcagt gtggccgagg gtagggcagg gggtgccac agaagagggg
240
ataggaggac agggcattgg gagatttctt ttcactgctg attcttgacc ctttcaaagt
300
gtttgcaaaa tgccggagaat tcacatgatc tgttccgac agtgtttatt ccccgcaatg
360
ttagtctgca gtggccaaag caaatgtcag tgttcatttt cacagcgcag caactgtgtt
420
tctgtgaatt tgctttaagg cttatccagg agaaaattac ttagctttgg gaacaggtgg
480

tggaagaaaa tcagccttag ctccagaaac gggtgtgt a gttgggcaac cttggatgac
540
ctgtgacgag gctggcctga gttagcaggc tggaaacgc caggtgggtg ctgggcaaag
600
tgagtatgtt agccgggta agtgttctt agacggtcat cccagatgtt cacctgccag
660
gacacctgcc ctcccaccc tcgagccccca gtgagctgtg gcctggggcc tgccggaggt
720
gtgctggccc cgggaggagt gtgatcaa atggaaagga tttccaagct tgctgccacc
780
gtgagttctt ggtggcacca ctgatggaag gaaaacgtgt cacagtgttgc ttctctccag
840
caggttctc agagccacct ctctgtgtcc tgggtggctg cataaacacg agggattgt
900
atcacactgg tgagggcag gcattgggtc gtattcacgc ttttcacaca gtcctgtaaa
960
gggagccatc cttagccctt tccctgtgt gtctgttcag gtataaacc tcgcgtgaag
1020
tcggacgtt ttgtgaaaat tctccctgat tatgaggcaca tggcgtaacag agacgtttac
1080
acctggccgt a tgtacccctt gcgagctttg gttctgcgg cagccagcac ggccaagaac
1140
tgcatgggaa gggccctctg attcacacat gggacagctt tggccctgga gcaagcggac
1200
ctcggtttt tttgtcatcc ttccacgttc acttgctgca gctggctctt gttcagctg
1260
tgacagagat cacacctgtg tggggaccc aggctgcatt tggcttaccc ttctgcagtg

1320

gtttctgggt actaccagtc gagatcactt taatgcacat tttcacatgt atcgtttatt

1380

gaggagctac tgcagacaca gagaggaacg aacagagtag aagctttaaa attttattct

1440

aaagtgaaat atgaatattt gacttctgac cttaattttgt tttaatcaca ttttaatagc

1500

ctatgtgaat taaaaatcac atatgcatac accctaata gacaaaggca ggtagcaaca

1560

gttgactctt ttttttttt tgaggcagag tctcactctg tcacccaagc tggagtgcag

1620

tgtcacaacc tctgcctcct cggttcaagt aatactgcca cctcagccctc ccaggttagct

1680

aggattacag gcatacacca ccatgcccag caaatttttt tgtatTTTA gtagagatgg

1740

ggtttcatca tgTTGGCCAG gctggtctca agtgatctgc ccatctcgcc ctccccagagt

1800

tctgggatta caggcatgag gtaccgtgcc tggctaacag ttgactctt agaacttaat

1860

tcttctttt ggccagacat ggtggctaat gcctgtaatc ccagtgcTTT gggaggccga

1920

ggcagggtgga tcatctgagg tcaggagttc gagacctgtc tggccaacat ggtgaaactc

1980

atctctacta aaagtacaaa aattagctgg gtgtggtggc gtgtgcttgt aatcccaggt

2040

acccagagggc tgaggtgggg gaatcgcttg aacccgggag gtgaagg

2087

<210> 8

<211> 2077

<212> ADN

5 <213> Homo sapiens

<400> 8

gctgatggag agcgggcagg tagggtggag aggcaagaag gttggtggca gccacatggc
60
tgagggccta gagcctggcc agggagtcgt ggagagaggc agtgggtggg ctgggggttc
120
aggcctgcct gaaggggagg cagctggtgc agtggccat accccatggg gtgaaggcct
180
gggcagggcc caggggcagc ttcgagggtg acctggagct gctcaggaag tgagatggcc
240
cagcctgacc tgaccattgg ctggcaagga acgggatgga gaagttgtgt cctgggcctt
300
cagcgagtgt gacattgtca ttgttggat agctttaaag atctgattgc ttatgacatg
360
cctttagcg ctaccagcat cttggcattt ggcaggtcta gtccagctcg ctgtttgcac
420
gtcttctgtc ttatccctag aagagagagt tcccagcctt gcttgcatttccccattga
480
tggaggctc atcacttcat gggagactca ttttacttag gccttctgag gatagttca
540
ttctgatagt ttttttttt tttttttttt ggagactgag tttccctctg tcgccccggc
600
tggagtgcag ttgtgtgatc tcggctcact gcaagctcca cctcccaggt tcatgccatt
660
ctcctgcctc agcctccaa gtagctggga ctacaggcac ccgccaccac acccggtaa
720
ttttttagt ttttagtaga gacggggttt caccgtgttta gccaggatgg tctgaatctc

780

ctgacccttgt aatccgccca aagtgctggg attaaaggcg tgagccaccc caccggcct

840

cattctgata catctttaga ggcttggtgt gcatgtgtt ggagggctct tgcaagcctt

900

ttgtgagtcc tcatggttcc ttcccccctc tcggggttcc tgggctccct gcggcacgct

960

gggtctaaca cagagacttg ctcttcctct cctcctgttag tgcttcaccc atatacacac

1020

attttggat tgtggcagcc agatatcggg ccatacggag gactgctgaa cgtgggtggta

1080

agtcccgag cctcgcgacg aggtctgtgc acgggcagga gtgggcctt acgtggagga

1140

atttgagggc ctcttctacc tgggtacaag ctggcccaga tgtgcgtttg aggttaattac

1200

caaaaaattc ttgctctgtc acttttctagg agccaatttt attttccaaa tgagcaaagg

1260

tttgcttaga gcaacacttc ggccttgtgg cccgtccttg gactctctgg tctgaggggg

1320

tggcttgta gggcctttgc ttccctggacg ctgaggccca tcgcgcacat gttggcgggg

1380

tctggagccc gccctcagcc aggtccttcc gttcctggtc tttgcccccc atgttccagc

1440

catggtgcac tctgacgtct gtttccttgg gttctgtgt ggacccgctc tgaccaccct

1500

gcctcccgcc ccctgtgccc tcctactgga gcctctcgaa cgaaaagccc tgctctgact

1560

ttgttgagct ctgggcaggt ggctggcccc tgctagaaag tgctagaacg ttgcaggcgg
1620
aagtccaggc tggtggcaga gtcgtctccg tttcccttct cccggcagcc tgcttgcgt
1680
tggggtctga ggacccgtga ggacatttct ttcctgttcg ttgggtctgc catggcagt
1740
cctgtctagt cctgggtact tcatcatggc tggagatgga agtctggatt tctttttat
1800
taaattcccg ttttcggtg aagtttcct tctctttgtt aatcttttg ggtacagtag
1860
tgaaagtttc ttgcagatct agccccggcc ttttcctctg atctccaggc gtgttggatg
1920
ccccccaccc ctcactgtgc gcctgggcac ttgctgggtg ccgtggctg gtctcctcaa
1980
gaggccctgg ccccgccctg cgggcaaagt ggagtcaggg cacctgctct ggcaggccc
2040
gggttatgcg cgctgtctcg gcttggggaa gcgtccc
2077
<210> 9
<211> 2889
<212> ADN
5 <213> Homo sapiens

<400> 9
ggatgtttcc tctgagagac cagccggca gtcacaggg tggcagatt gcttccttt
60
cttttagtgg aggaaggta attgtttac ttactggcag tgagtaatag agtaaaaaaaa
120
aaaaccccaa aaaacaaaaaa gcagtggata attggggaaa gtggagtctg tgagagaaca
180

gcaccagctc acgcttggacc caggctcggt agccacagaa atgatggaat tctctccttg
240
gttgctggag ctggcagtct catgggtccc agctctccca gaagggggtg aggctaggct
300
gtccttcctt ataagcatgt ggctgctgtg gccagagcgt gctggggccc acccgccctt
360
gcatttgcac gatcgccgcc cggacacgcg agatgcacg ggcaacgcac tcctcggttc
420
ttcccgctcc tgagcgcccc tcccggggtc ttggcgtctc tctcctccca tgcacttcgg
480
ttttgaagca aatcccaggc ccgtggcatt tcagctgtcc gtccctccgg gcagcgccct
540
aagaaaagggg gtgtttcctg catctccgtg tcactgccac acctgacaaa agtggcacac
600
ggtccttgc tgccctgttgc cccagggcat gggatgttc cctgatcate tcaaaggcctc
660
ctttgaatgg caatttgttc agctcaggac gcgaaacgagg cccttgact gctgttgttc
720
actctttctt gttggagat gcacccctt tccctctgtc tggteggcac ctcacccctgt
780
gtcactttctt ggattattct gcctcctctt ggtgtccacc tgcatactcag agctggaaga
840
tagtgctggg attcgaggca cctcgagggt gggcggtgcag gtcccatgtg ctggtctggc
900
aggagcttctt ctgtgccatt gcagtgcggg gagccgtgcg gggccggccct tgctctggcc
960
cgggagccac tcatggacct ttgcccatct cctcctgttag gtggacggcc tgttcatcat
1020

cgggtggatg tacctgcctc cccatgaccc ccacgtcgat gaccctatga gattcaagcc
1080
tctgttcagg atccacctga tggagaggaa ggctgccaca gtggagtgc tgtacggcca
1140
caaagggccc caccacggcc acatccaggt gtgtcagcg gcggggctgg gtcctcactg
1200
tcccaggcgt gtcctgtgtg ggctccagcc aggcctgccc tgctattctc agctgcagac
1260
ctgggctgta gcagatcggc ggggtggagg gaggctccgc cctgccctgc tgtgcattgt
1320
ttacgactcc gtggcgcagt ggactctgca gcggtcactt gctggacccc tcattgattct
1380
gctctcagtt agagccgctg tttcttgca acttcagttc ctctgtctt ttttcctgtt
1440
tgcaagaata tcagtgtgga atcaagtgcg ctctctgttc tctgatctgg tctggcagtg
1500
gcccccacgg tgagcacagc gtgtcatctg ccacaccttg tgtgtgagat gcagccctt
1560
tgctgctctg tgtcacgctg ggaaatgcaa acgccactct ctcagatgtg ccactgcctc
1620
ctgctcttgg gggagtgtttg ctcagggaga ctcagctccc tccgctggca ccgcgttgg
1680
cgcatcttgg gtgcctgcag tggggtttgg tgggtgtggt gagggagggg tcgtggccac
1740
ctgcctgggg tggggatgc tgtgatggac tctgttcctc actcttctt ttcccttgt
1800
tgcaaacact gaagattgtg aagaaggatg agttctccac caagtgcAAC cagacggacc

ES 2 433 012 T3

1860 accacaggat gtccggcggg aggcaggagg tgagccacc agccggccct gtgtcatat
1920 ggtgcggacc tttccttcc ccacggggaa agtacagacc catgcgggag agaagtca
1980 agcatgcgac ctgcatggtg gctgttcac atggttcagc cggattgtc agtgcagacc
2040 gtgcacccgg cacactcgag gaacagcatac gtgcggcgg gcagggtgcc ggtccccctcc
2100 ccctgcagaa cgaacgaacc acacaggccg tttcttcagt tagctctgct ttgcggctct
2160 ggcgggcaca tccacttgct gaggggttgc tgggaggctg cttctaggat gtgagctgca
2220 gggagacccg agggctgcac ccagagttcc tgtgttccc atccttgagc agaccgtgt
2280 gaggcttccg gatcgtgcca gtgcagcggg gaagcctgtg tgtgattgtt tgcctgagta
2340 tttaatatt gcccttgagt ttagcttcc aaggatctaa gtcttactgc cctctaaaa
2400 tactcttaag aaggaaggtg cggggctca cacctgtatt cccagcttt tgggaggctg
2460 aggcaagagg atcacttaaa gtcaggagtt caaaaccagc ctggcaaca tagtgagacc
2520 ctgtctctac aaaaacccaaa aatattagct gggatggtg gcatgccagc tactcaggag
2580 gctgaggtga gaggattgtc tgagcccagg aggtggagac tgcagtgagc tatgagcgca
2640 ccacccgact cccatgtggg tggcagagtg agaccctgcc tcaaaaggaa aaacaaaaaa

ES 2 433 012 T3

2700
ttagccaggc atgctaccat gctacgggt tatgccatgt tgcccaggct ggtcacaaac
2760
tcctgggctc aagccgtctt cccaccttga cctcccaaag tgctgggatt gcaggtgtga
2820
gccatcgcbc gcggccaggg tgtgctttc ttagctagtt tgagttgtgg tggtttcccc
2880
atcagccag
2889
<210> 10
<211> 4378
<212> ADN
5 <213> Homo sapiens
<400> 10
gtggtttca gaatagatct taaacgtatt ttgtaacatt tattcccaag tcttatgttt
60
tttgaggat tttgtacata gaattttttt aaattttatt tttatatttt gtgctgctag
120
gacataaaaaa tccaggtggc ttctgtgtt accatgtggt cctgccaggt gctcagtgc
180
gctcctgcag tgatttctgt ttcttgcgg agttggagc agtcatgtat gaaagcgatt
240
gaccatcgta tttcctcttt tttgtgagtt gcttatgcgt cccttcaga gtttaggtcta
300
gataatgact ggaaacactg tattgagctg ggaacttggaa agttgggtg aattctgtaa
360
attgtgtaga atcaaaccat taagcacgaa atatcctcag caccatgtga tggcagaatg
420
caggcggac tggtggggct gcggagttt tgggaaccgc cttaacatac ctgcctttcc

480

tttggaaaaag gaaaatccac ccagtcaactg gccgagggct cagtgcateg tccttggctg

540

actcgcacgg ctcaagggttg gtggtgaggg cttctggaac actcaggctc ccaggagcca

600

gagaagcagg ctgaggggct gtgctgtgca gggccagagt ctgtgttagat gctgttcctg

660

ccccaccggt gggcatggaa ggaggcccac ccctgagttc tggtcatggc tgtggccccc

720

gcctcacctg ggctccctgt agcccccgt gcctccgacc cagcagacca agggcccgaa

780

acccatccactt tttgttggca ggcacctctg acgtggggtc aaatccggca gctccccc

840

ctcttccctag cgccaagtcc atccctccccc accaagaggg ccctgcaggt gcctgtgcac

900

tgacaaagaa ggacacactg tgtccctctca gggcacacgc cggcagccctc ttcacagcc

960

cctgtcatat ggatattcat gacctgtgtt ctttttggaaag gagtttcgga cgtggctgag

1020

ggaggaatgg gggcgcacgc tggaggacat cttccacgag cacatgcagg agctcatcct

1080

gatgaagttc atctacacca gtcagtacga gtgagtgccg ctctgctcg gtgctgggg

1140

agccctctcc tggtgtgtgc tggggaccag ccctgtcccg gtgggtgctg gggcatagtc

1200

ctttgtctca gtgggtgctg gggttcagcc ctctctctcg ctgggggctg ggagtacagc

1260

ccctctctct ctgggtgctt ggagcatagt tcgctttc tgtggtaact tggaacacag
1320
ccccctcttc tgtgggtgct gggagtgcaag gctccctgt agctgctgtg gcacagcccc
1380
agaccccatc cttcacccgg gtgtctgttt ttatctgatt tcctgtcacc cctgccaaca
1440
ctggcactct agagctcagc cccgctcagc ccagtgcggg gatcccctcc atggtgactt
1500
gggctccctt ctccctggta atggagcaaa ctggcacttc ctttccctc ccacctaag
1560
ctctgccctg gcgctgatgc cccccttaac cccagaagtc ctgttgaaag cgtgcgtgct
1620
ggtgggtggc ggggtgagg gctccctgg gcgattgct gcccacaggg ctggtgaca
1680
cctggccctg gttgagctga gctgcccggcc ctacccggcc ctgcaaagct catccctggc
1740
accttggcg ttctcagtc tgcctgcccc acgaggtctc ctgtactgac ccgtccgctc
1800
cccacagcaa ctgcctgacc taccggccgca tctacctgcc gcccacggcc cccgacgacc
1860
tcatcaagcc tggcctcttc aaaggtacct atggcagcca cggcctggag atttgtatgc
1920
tcagcttcca cggccggcgt gccaggggca ccaagatcac ggtgagtggc ggctgacctg
1980
gttgggtggg ctctgggggc acctggccca agtggctgt ggagtcaagg aacttggca
2040
gacactgatc cccggcaccc ttgctcctca gagtggcctt gcacctgcag ccccgaggagc
2100

ttgggaaagc agattctcg ggcccccagcca gccccactga accagaattt catttccacg
2160
agaccctcag ggagtctgtg tgcttccaag gtggccttcc cacagcacccg ctcagccccg
2220
gccaccctgg cagccctgtg gggtttagga agtctggaca gcagacccca cacagaggct
2280
cagatthaac tggggacacgc ctgggaaagg ggcccccctcca gcagctgcac ggagaccctt
2340
gtgcttgcag acagctccca ccctgtctga gcttcagcag ctctctcttctt tggagtggcc
2400
tgagggtggg ctgtgtggtg agcccaagcag tgaacagggg ctcattgtcc catgccagg
2460
ctcctggcac atgctgccccg cctgcccctga atggttattt gggatcctgg cagacagcgc
2520
ctgaaggggc gaagcctatt ccccgagcct ccaggatca gggccacatc tggcataatg
2580
ggcattttccc tggaccccaag agttttggct cctggactcc tcacggagtt ttggggctga
2640
cgcatcttcc cagggctcgg ctccccgagct ctgcctttca ggcagacaga tgtgggctgc
2700
.cccccttgggc tgctgccccag ctctgggatgt ttctttaac ttctctgggc gcagttgagc
2760
aggtatgagc caccgtcctc aggggcactc ataggtgttg ctggccctgc cactttctaa
2820
gggattctga gacttcttct tctgttaagg acgttgctaa gtcagtagga gggctggc
2880
cagctcgctc agccagaggt ggatctgggc tggagccac acagtggtac tgctgctgct

2940

gctggcctga gccctgctga ctgccacactg ctccacaggg cgaccccaac atccccgctg

3000

ggcagcagac agtggagatc gacctgaggc atcgatcca gctgcccac ctcgagaacc

3060

agcgcaactt caatgagctc tccgcacatcg tcctggaggt ggcgagagg gtgcgccagg

3120

agcagcagga aggccccac gaggcccccg agggtcggtgg ccggcaggc ccccccggagt

3180

cccgccaag ccctgcccag cccagggcag aggccccag caagggccca gatggcacac

3240

ctggtgagga tggggcgag cctggggatg ccgtagctgc ggccgagcag cctgcccagt

3300

gtggcaggg gcagccgttc gtgctgccc tggtggtagtgc ctccaggaat gaggactacc

3360

cccgAACCTG CAGGATGTGG TAAGGATGCG GCGGGTACTG GGGCCTGAAG GTGGGACAGC

3420

atgggcttca gcgagggccc cagccccaca cctagcacag gcggagaggg cctgtgaccc

3480

cacagagggc ggcagccgt gctttgggac aggagtgcgg cctctgaccc cttgggccccat

3540

gttccccagc acctgagcaa gcggccgcgc agctgggtcc cgtcttggag gctcctgtcc

3600

ttccacccct tctgggtac ctcagagctg cagggcatg aggcttccag atgcctcaca

3660

tccctgcaat agtgcgcgtc ccccaaaaaa ttctaaagct acttggggc agtcaatcaa

3720

gtgaaatatac atgtaaactg tccagcagct ttgaaatgtt agaatgaaca agggcccttc

3780

cccacccacc ctgtggaaaag cccgtctgggt ttgggtgtctt cctggacacgc gtcttgccgg

3840

tcacaccttgg ccatctccccg gtgcgtggtt cagatgtggg tcctgctttc ctgccccctc

3900

cctcctctgt gcctgcctgc ctctgctgtg ccggggccagt gccttgggtgg ccagtggagt

3960

ggacaccaggc tgcgactgcg tgggagggggc tggcattgcc gctgccactg cagggcttgg

4020

gcgggctgaca tgggacgagg cttgcacaggc tgccagctcc tgtctcgctg actttttta

4080

tacagtttg tctggccac cgccttcagt gccacgggcc ccttgcgcgtt caggctgc

4140

ctcatagatg aacaaggccc tgccccgtgt ccttaccctt tagagctgtt taaattcaaa

4200

tgaactgaaa ctgaatatga aaaatccagg ccctcagccg cccaggccat gtttcaagtg

4260

ctccatggcc acatgtggct ggtggacagt gcagctctag aacattccat caccacagag

4320

ggttctgctg gacagtggcc ttgggggctg ttttgggggt ccgcctgtca gtctcctg

4378

<210> 11

<211> 3174

<212> ADN

5 <213> Homo sapiens

<400> 11

atagctaaaa aaggaaaaaag ctctgttacc cttcaaacat ctcagaactc ctaataccctc

60

acagtggct cagagccagt ttgcgagtca cagcacgtgc tgcagccact gaccaggaag

120

acctgtccag cctccgtccc cagctaacct gcaggacagt ccctgacgtg ctgaatatgg

180

gggcctccgg gcccttcct gtctagttt tggatctgca gggcctggcc acggcacctg

240

tacagggaaat gcctgttgct tttttctga ccacagtgcc gatggctgca cctgcctctg

300

ggccctggaa cttctccccg agcttgagaa gccctctgag gccaggccct ggtccacggg

360

gctgttcttt ccccccgcct ttggagcctc agtgggtgat tccaggccag ccccttactt

420

ctcgcatct gttggaagaa ttttagctgct tgcaagacag acataagtgt cttcctgtct

480

gatgttgacc tcaaagccat aaatgggtgt ttgcgacttc tgagttaatg tcagctgcag

540

gctgcctgta ttagagctaa ttgtatgggg acataactcc cagacattaa gatTTTTT

600

cacattgggc ctcccttatg aaatgtgtgt tttggAACAG agctctctgg gcctgcagag

660

acctcgTTT agttcagtgt ttcaGcattt tgcaGtcagt gggtGAATCC ctcacGGGTG

720

ctgcgagtca gCGCCCATCC CCCGAGCAGC CCCGAGCTCT GGCCTCTGCG TCCATCATCA

780

ggTgggctcg gcctgcGCC ttagctACCC CTTCAgAGAC aggctcAGCC cacACCCCCA

840

gctgcccctg cagAGACAGG CTGGCTCTGG gagtcAGCTT CTGGCTGATG AACAGTGGAT

900

ES 2 433 012 T3

gtggctttg cgccacagag cggggtcgca gaatgctgta cgtggcgtgc atttgactca
960
gccctcccc agtcacagt ttcccttttgc tttctgctca gtttttatgg cacaggcctc
1020
atcgcccccc acggcttcac cagccctgaa cgccaccccg gggctttcat cctcttcgat
1080
gaggaccgct tcgggttcgt ctggctggag ctgaaatcct tcagcctgta cagccgggtc
1140
caggccacct tccggaacgc agatgcgcgg tccccacagg ctttcgatga gatgtcaag
1200
aacattcagt ccctcacctc ctgaccggcc acatccttgc cgccacatcc cgggtggctc
1260
tggggctctg aactctgacc tgtgaataga agcagcatgc actttggaaa tccggcctt
1320
tgaccagaac gcacaccccg tcggggggcc cagtccagcc acccccccagc actttatgta
1380
gagagtgtga catagacctg catatttgc agtgcacatga tggaagaagc tgagcatgtc
1440
ttacaaaaaa cagagagaac gagcctgaat acagcagatg taggggacag ccgtgggacc
1500
gcgtgagaat tgaagcggtg gggttcccgcc accctgggt ggctgggtgt tttctcggt
1560
agcaggaccc tcctgactgg tgctttccct gtgagcggat agagtatacg actgggtcgt
1620
gtgtgagacg catgtgctcc accccactcc ttttggggaa agccaggcaa cagtggcctc
1680
tgggaggggg tcaggaagag gcgaacagct caggcagcgc aggtgtgatg ggcacagttac
1740

ES 2 433 012 T3

gcagagcaag ctgggaaat tggtaggatc tcaggcttgg ggccgggact ctggagtcaa
1800
tccccatttc tctaccggct tgcttggagt ttggacagaa gcatttcacc tctgatctca
1860
gctccccac ctgtggagtg ggttagtga cctgagtcac taggaaatgt cacctgaatg
1920
cacagccccag cccatgcacc tgccccagcc cctccagctt tgagccaag gccatcggtc
1980
cagccacttg actgtccctcg acggcctgtt ccagacaggg cgtttgttt gtccatgcct
2040
tcctccctgc acgcacacgg cgtaaaaacc aagctgccgg ccactgtctc cagaacgcaa
2100
ggctccaggc ccgtgtgtct gaagcagtga gtggccaca caggtgccag gagtgcggcat
2160
atgagatgac gaggaaaccc ctttgcaggt gaggggacag ctttctagaa aagccacacc
2220
tgcatctggg gacacacttt ggaaagtggg accctccagc ctggagaccc catggactga
2280
tgcctccact gctgtgtgcc ccatgttgtt ttaacacctg cgtgtggggaa ccccatctga
2340
ggtcttggct gaggttggca tctcctgaag aacagagagc acgggtgtcca gagctggccc
2400
ttccccccagc ccacagccag ctccgtgccc gagtggcggt ccccaagcgag cttccctct
2460
ctgccgcttg tccttgcgtc tgggtgtc caagtccttg tgctgggcac cttggacacg
2520
tcctgctggt gagggaccc gggaaaggta cagtctgtgt gccttgggtgt ggagaccaac

2580

ctgaggatgt cctggaaaat gtttcctga tgaatttctc cttgactggc ctttaagaa

2640

cataagaatt cccattgccc agcctcagtg catttggcaa atgcttactt tgcttcccag

2700

agttagagaa ttggcaaagg ttctaaatg gtaatctggc cggcctggga gaaagactca

2760

cgagaaaaagc cagtggagaa agcgcccttc cagggcggca gcagcgggag ccacgcagac

2820

cccgaggcgc acctgctggc tcttgtgtgt ggccccagtt tctagcggct tttgcagcat

2880

tagcctacaa gctttgtcac tcctgcct ctgtggtggt cactgtttt ctctttgcc

2940

aaatgaggca gtctctgagt gacggtgact gtggccttga agcctggagg actgttgggc

3000

atgttagactg gcaccttcaa gattcaccat tgtttaataaa aatcaagca aatgttttt

3060

taccaagagc ccgagcctcg ctctaaggga cgcagtccta gaggcgtgcc ctttgggct

3120

tgaagagcac actgtgggac gcacgtgctt ctgattaaag gaatctcaga tctc

3174

<210> 12

<211> 1767

<212> ADN

5 <213> Mus musculus

<400> 12

cgcgtccgcg tgcgcgcctc cgccgtggtg acgggcattgg cgggtgtgtgc tcggctctgc

60

ggcgtgggcc cgcgcgtgg gtgccgcgcg cgcgcgcgc gcccggccccc ggccgagact

120

gcggcggcgg acagtgaggc ggacacggac cccgaggagg agcgcatcga ggccccccg

180

gcgcggttgc ctctgctgga gctcccgcct gagctgctcg tggagatott cgcgtcgctg

240

ccggcaccg acctgcccag cctggctcag gtctgcagca ggttccgcgg aatcttgac

300

acggacacca tctggagacg gcgctgccgc gaggagtatg gcgtttgtga gaacctgcgg

360

aagctggaga tcacaggtgt gtctgccgg gacgtctatg caaaactgct tcaccgatac

420

agacacattt tggggctgtg gcagccagat atcggccgt acggaggatt gctgaacgta

480

gtggggacg gactgttcat cattggctgg atgtacctgc cacctcatga ccccacgtg

540

ggagaccca tgcggttcaa gccactgttt agaatccatc tgatggagag gaagtcggct

600

acagttggagt gtagtacgg ccacaaaaggc ccccacaaacg gccacatcca gattgtgaag

660

agggacgagt tctccaccaa gtgttaaccag acagatcacc acaggatgtc cggtgggagg

720

caggaggagt ttccggacgtg gctgaaggag gagtgccgc gcacgctgga agacatcttc

780

cacgagcaca tgcaggagct gattctgatg aagttcatct acaccagtcg gtacgacaac

840

tgcctgaccc accgacggat ctacccccc cccagccacc ctgacgaccc catcaagccc

900

ggcctttca agggcaccta tggcagccac gggctggaga ttgtatgtc cagttccac

960

ggctcacgca cctggggcac caagatcact ggcgacccca acatccccgc gggacagcag

1020

actgttagaga ttgacctgca gcgcgcacatc cagctgcgg acgtggagaa cctccgaaac

1080

ttcaacgagc tctccaggat tgtcctggag gtccggagc aggtgcggca ggagcaggag

1140

gccggcgagg ggcgcgcgc accccgggag cttcagcca aggccgctga tgggccacct

1200

gctaaggacg gcaaagagcc tggaggtgga gccgaggcag ctgagcagtc ggctcgct

1260

gggcaggggc agccgttgtt gtttcctgtg ggtgtcagct cgaggaacga ggattacccc

1320

cgcaactgccc gcctatgttt ctatggcaca ggcctcatcg ctggccacgg ctttaccagg

1380

cctgagcgca ccccccggagt cttcgctctg tttgtatgagg accgctttgg atttctgtgg

1440

ctggaattga agtccttcag cttgtacagg cgagtccagg ccaccttcca gaacgcccgc

1500

gcgcgcgcgc cgcaggcctt tgacgagatg ctcaggaaca tccagtctct cacccctgaa

1560

cctccgcata, gtggggcgag aggctgcacc ggggccccggg tgggggaagc atgcacttta

1620

gaaatgaacg cacacccctt cactgggttc cgggtcgccc cgggacgcctt cttgtatagt

1680

gtgttaacata ctcttgtaca tttgacttgt tggtagctat gaaggagaac gctaagcatg

1740

gtgagaaaaat aaacggagtt gagccag

1767

<210> 13

5 <211> 507

<212> PRT

<213> Mus musculus

<400> 13

ES 2 433 012 T3

Met Ala Val Cys Ala Arg Leu Cys Gly Val Gly Pro Ala Arg Gly Cys
1 5 10 15

Arg Arg Arg Gln Gln Arg Arg Gly Pro Ala Glu Thr Ala Ala Ala Asp
20 25 30

Ser Glu Ala Asp Thr Asp Pro Glu Glu Glu Arg Ile Glu Ala Gly Pro
35 40 45

Ala Arg Cys Ser Leu Leu Glu Leu Pro Pro Glu Leu Leu Val Glu Ile
50 55 60

Phe Ala Ser Leu Pro Gly Thr Asp Leu Pro Ser Leu Ala Gln Val Cys
65 70 75 80

Ser Arg Phe Arg Arg Ile Leu His Thr Asp Thr Ile Trp Arg Arg Arg
85 90 95

Cys Arg Glu Glu Tyr Gly Val Cys Glu Asn Leu Arg Lys Leu Glu Ile
100 105 110

Thr Gly Val Ser Cys Arg Asp Val Tyr Ala Lys Leu Leu His Arg Tyr
115 120 125

Arg His Ile Leu Gly Leu Trp Gln Pro Asp Ile Gly Pro Tyr Gly Gly
130 135 140

Leu Leu Asn Val Val Val Asp Gly Leu Phe Ile Ile Gly Trp Met Tyr
145 150 155 160

Leu Pro Pro His Asp Pro His Val Gly Asp Pro Met Arg Phe Lys Pro
165 170 175

ES 2 433 012 T3

Leu Phe Arg Ile His Leu Met Glu Arg Lys Ser Ala Thr Val Glu Cys
180 185 190

Met Tyr Gly His Lys Gly Pro His Asn Gly His Ile Gln Ile Val Lys
195 200 205

Arg Asp Glu Phe Ser Thr Lys Cys Asn Gln Thr Asp His His Arg Met
210 215 220

Ser Gly Gly Arg Gln Glu Glu Phe Arg Thr Trp Leu Lys Glu Glu Trp
225 230 235 240

Gly Arg Thr Leu Glu Asp Ile Phe His Glu His Met Gln Glu Leu Ile
245 250 255

Leu Met Lys Phe Ile Tyr Thr Ser Gln Tyr Asp Asn Cys Leu Thr Tyr
260 265 270

Arg Arg Ile Tyr Leu Pro Pro Ser His Pro Asp Asp Leu Ile Lys Pro
275 280 285

Gly Leu Phe Lys Gly Thr Tyr Gly Ser His Gly Leu Glu Ile Val Met
290 295 300

Leu Ser Phe His Gly Ser Arg Ala Trp Gly Thr Lys Ile Thr Gly Asp
305 310 315 320

Pro Asn Ile Pro Ala Gly Gln Gln Thr Val Glu Ile Asp Leu Gln Arg
325 330 335

Arg Ile Gln Leu Pro Asp Val Glu Asn Leu Arg Asn Phe Asn Glu Leu
340 345 350

Ser Arg Ile Val Leu Glu Val Arg Glu Gln Val Arg Gln Glu Gln Glu
355 360 365

Ala Gly Glu Gly Ala Ala Pro Pro Arg Glu Pro Ser Ala Lys Ala Ala
370 375 380

Asp Gly Pro Pro Ala Lys Asp Gly Lys Glu Pro Gly Gly Ala Glu
385 390 395 400

ES 2 433 012 T3

Ala Ala Glu Gln Ser Ala Ser Ser Gly Gln Gly Gln Pro Phe Val Leu
405 410 415

Pro Val Gly Val Ser Ser Arg Asn Glu Asp Tyr Pro Arg Thr Cys Arg
420 425 430

Leu Cys Phe Tyr Gly Thr Gly Leu Ile Ala Gly His Gly Phe Thr Ser
435 440 445

Pro Glu Arg Thr Pro Gly Val Phe Val Leu Phe Asp Glu Asp Arg Phe
450 455 460

Gly Phe Leu Trp Leu Glu Leu Lys Ser Phe Ser Leu Tyr Ser Arg Val
465 470 475 480

Gln Ala Thr Phe Gln Asn Ala Ala Ala Pro Ser Pro Gln Ala Phe Asp
485 490 495

Glu Met Leu Arg Asn Ile Gln Ser Leu Thr Ser
500 505

<210> 14

<211> 5106

<212> ADN

5 <213> Homo sapiens

<400> 14

ggcatggcgg tgtgtgctcg ccttgccgc gtggggccgt cgcgccgatg tcggcgccgc

60

cagcagcgcc ggggccccggc cgagacggcg gcggccgaca gcgagccgga cacagacccc

120

gaggaggagc gcatcgaggc tagcgccggg gtcggggcg gtttgtgcgc gggccctcg

180

ccgcgcgcgc cgcgctgctc gctgctggag ctgccgcgc agctgctggt ggagatcttc

240

gcgtcgctgc cggcacgga cttaccacgc ttggccagg tctgcacgaa gttccggcgc

300

ES 2 433 012 T3

atcctccaca ccgacaccat ctggaggagg cgttgccgtg aggagtatgg tgtttgcgaa
360
aacttgcgga agctggagat cacaggcgtg tcttgcggg acgtctatgc gaagctgctt
420
caccgatata gacacattt gggattgtgg cagccagata tcgggccata cggaggactg
480
ctgaacgtgg tggtggacgg cctgttcata atcgggtgga tgtacctgcc tccccatgac
540
ccccacgtcg atgaccctat gagattcaag cctctgttca ggatccaccc gatggagagg
600
aaggctgcca cagtggagtg catgtacggc cacaaagggc cccaccacgg ccacatccag
660
attgtgaaga aggatgagtt ctccaccaag tgcaaccaga cggaccacca caggatgtcc
720
ggcgggaggc aggaggagtt tcggacgtgg ctgagggagg aatggggggcg cacgctggag
780
gacatcttcc acgagcacat gcaggagctc atcctgatga agttcatcta caccagtca
840
tacgacaact gcctgaccta cgcgcgcata tacctgccgc ccagccgccc cgacgacctc
900
atcaaggctg gcctttcaa aggtacctat ggcagccacg gcctggagat tgtatgctc
960
agcttccacg gcccggcgtgc caggggcacc aagatcacgg gcgaccccaa catccccgt
1020
gggcagcaga cagtggagat cgacctgagg catcgatcc agctgcccga cctcgagaac
1080
cagcgcaact tcaatgagct ctccgcata gtcctggagg tgcgcgagag ggtgcggccag

1140

gagcagcagg aaggcgggca cgaggcgggc gagggtcgta gccggcaggg ccccccggag

1200

tcccagccaa gccctgcccc gcccaggca gagggcgcccc gcaagggccc agatgggaca

1260

cctggtgagg atggggcgca gcctggggat gccgtagctg cggccgagca gcctgcccag

1320

tgtggcagg ggcagccgtt cgtgctgccc gtggggcgtga gctccaggaa tgaggactac

1380

ccccgaacct gcaggatgtg ttttatggc acaggcccta tcgcgggcca cggcttcacc

1440

agccctgaac gcaccccccgg ggttttcatac ctcttcgtatg aggaccgtt cgggttcgtc

1500

tggctggagc tgaaaatcatt cagcctgtac agccgggtcc aggccacctt ccggaaacgca

1560

gatgcgcgt cccccacagggc ttccgtatgag atgctcaaga acattcagtc cctcacctcc

1620

tgaccggcca catccttgcc gccacatccc gggtggctct ggggctctga actctgaccc

1680

gtgaatagaa gcagcatgca ctttggaaat ccggccctttt gaccagaacg cacacctcg

1740

cggggggccc agtccagcca ccccccagca ctttatgttag agagtgtgac atagacctgc

1800

atatttgcata gtgccatgtat ggaagaagct gagcatgtct tacaaaaaac agagagaacg

1860

agcctgaata cagcagatgt aggggacagc cgtgggaccg cgtgagaatt gaagcggtgg

1920

ggttcccgca ccctgggctg gctgggtggtt ttctcgggaa gcaggaccct cctgactgg

ES 2 433 012 T3

1980

gctttcctg tgagcgata gagtataga ctgggtcgta tgtgagacgc atgtgctcca

2040

ccccacttct tttggggaa gccaggcaac agtggctct gggaggggt caggaagagg

2100

cgaacagctc aggccgcga ggtgtatgg gcacagtacg cagagcaagc tcggaaagtt

2160

ggtaggatct caggcttggg gccggactc tggagtgaat ccccattctt ctaccggctt

2220

gcttggagtt tggacagaag catttcacct ctgatctcag cttccccacc tgtggagtgg

2280

gttagtgac ctgagtcact agggaatgtc acctgaatgc acagcccagc ccatgcaccc

2340

gccccagccc ctccagcttt ggagccaagg ccattgttcc agccacttga ctgtcctcga

2400

cggcctgttc cagacagggc gtttgttttgc tccatgcctt ctccttgca cgcacacggc

2460

gtcaaaacca agctgccggc cactgtctcc agaacgcaag gtcaggcc cgtgtgtctg

2520

aagcagttag tggtccacac aggtgccagg agtgccata tgagatgacg agggaaacccc

2580

tttgcaggtg agggacacgc tttctagaaa agccacaccc gcatctgggg acacactttg

2640

gaaagtggga ccctccagcc tggagacccc atggactgat gcctccactg ctgtgtgcc

2700

catgttgtt taacacctgc gtgtggggac cccatctgag gtcttggtcg aggttggcat

2760

ctcctgaaga acagagagca cggtgtccag agctggccct tccccccagcc cacagccagc
2820
tccgtgccccg agtgggcgtc cccagcgagc cttccctctc tgccgcttgt ccttgtgtct
2880
gggctgctcc aagtcccttgt gctgggcacc ctggacacgt cctgctggtg agggacatcg
2940
ggaagggtgac agtctgtgtg ccttgggtgtg gagaccaacc tgaggatgtc ctggaaatg
3000
ttttcctgat gaatttcctcc ttgactggcc tttaaagaac ataagaattc ccattgccca
3060
gcctcagtgc atttggcaaa tgcttacttt gttcccaga gtcagagaat tggcaaaggt
3120
tcctaaatgg taatctggcc ggccctggag aaagactcac gaaaaagcc agtggagaaa
3180
gcgccttcc agggcggcag cagcgggagc cacgcagacc ccgaggcgca cctgctggct
3240
cttgcgtgtg gccccagttt ctagcggctt ttgcagcatt agcctacaag ctttgtcact
3300
ccctgccttc tgtggggcactgtttttc tctcttgcca aatgaggcag tctctgagtg
3360
acggtgactg tggccttgaa gcctggagga ctgtgggca tgttagactgg cacccttgaag
3420
attcaccatt gtttaaataa aatcaagcaa atgctttttt accaagagcc cgagcctcgc
3480
tctaaggac gcagtcctag aggctgtccc tttggggctt gaagagcaca ctgtggacg
3540
cacgtgcttc tgattaaagg aatctcagat ctcaattacg cttccagtgt ttgggtatag
3600

aaatagcttc cacccatcat gtctcagcca tgggctgtt gtcagttcat gtggctcctg
3660
gttctggtgt gtatgttggg ggcgggggtc tctccatggt ggtgacctgc agtgatgcca
3720
ggcagggcca gagccacaca gccaggaaag ggaggcctt ttggccgcac agccagtccc
3780
ttcagtcgtg actacaggta ttgtttttc cgctccgatg tgccttagc cagttcttgg
3840
ctccggttct gtagggacag gcactgaatc tgcgcgcctc aaaacagcag cttcccttcc
3900
gggggagggc atccaccctc tcaggggatc ctgcaggtgg cccatttcct gcaggtgaga
3960
actcggaagg gctgatgtcg tcatcagagg cctaaggca gctgagagtt ggataaaacc
4020
gtttccaagg aggaggctga gtaacccagt tcaggacagc caagcgcatt aggcttgatt
4080
ggggaaagggtg gcaggtggag ttgggagggtt gggactctcc atctttgca ccacggatgc
4140
ctttctgtcg ctgtctcaact ctggggcagg atcaagtctg ctctctggag tggggctgcc
4200
tgcaagtgcag ctctgcacac ctgaacgtgt tctttgtcac ttgtttggaa atgatgtgat
4260
tgaagatttc agagaggta ttggaggctt ttctgtgccg gcactgaatg ttcatttgc
4320
tgaggaagtt gcaaacgact tctgcaggct gagattcaag gcaggtggta ttggggtccc
4380
tcagcccacc tggccgtga cctcaagtgtt ccactgctga gtgtgagtttgcaggc

4440

ctgggtggtgg gagagcctca ggctccctcc ttcttcgttc ctgaccatgc cctggggccc

4500

ttcagtctgc ctgcggctct gtggcatccc tgccctgaca ctgggcacct gtgccccc tag

4560

caagcccacc tggcacacga ggagggaggg gtgggtggcc atgtccttcc tctagccaca

4620

tgccctgctg gcccgtccat tctgagctt gtgcagaacc gggctgtgagc tggagatttt

4680

tctctgagaa cctgcaggttg tgctgcagcc gcacgcaagg gcccttcagc cgctggctct

4740

ggcttccctg actcctcagg gcgtgttcac cccaggctt tctcacctgc acacggtag

4800

gccatttca gtgctcggtgg gcagtcacgg acagcagcag aaactcctca gccctttgt

4860

tacttcagaa cgcctgcccc catgcatttt ctgagctcgt gttgtcctca tggccgtggg

4920

gtcttgggtg cgaacaggag atgctgagct gtggtccacc gtccaagggtg ctgcagaaag

4980

cagctaggct cttttaggat gtttctattc tggttgctgc cttcgtggtg taactttaa

5040

gaacacttac gggaatgtgc tcatagaacc atcacctgtc ctgagaataa aactcctgga

5100

atcatg

5106

<210> 15

<211> 5193

<212> ADN

5 <213> Homo sapiens

<400> 15

ggcatggcgg tgtgtgtcg ccttgccgc gtggggccgt cgccgatg tcggcgccgc
60
cagcagcgcc ggggccccggc cgagacggcg gcggccgaca gcgagccgga cacagacccc
120
gaggaggagc gcatcgaggc tagcgccggg gtcggggcg gttgtgcgc gggccctcg
180
ccgcgcgcgc cgcgctgctc gctgctggag ctgcgcgcg agctgctggt ggagatcttc
240
gcgtcgctgc cggcacgga cctaccaggc ttggccagg tctgcacgaa gttccggcgc
300
atcctccaca ccgacaccat ctggaggagg cgttgccgtg aggagtatgg tgtttgcgaa
360
aacttgcgga agctggagat cacaggcgtg tcttgcggg acgtctatgc gaagcgtata
420
aacccctcgcg tgaagtcggg acgtttgtg aaaattctcc ctgattatga gcacatggcg
480
tacagagacg tttacacctg cctgcttcac cgatatacgc acatTTTggg atttgtggcag
540
ccagatatcg ggccatacgg aggactgctg aacgtggtgg tggacggcct gttcatcatc
600
gggtggatgt acctgcctcc ccatgacccc caegtcgatg accctatgag attcaaggct
660
ctgttcagga tccacctgat ggagaggaag gctgccacag tggagtgcgt gtacggccac
720
aaagggccccc accacggcca catccagatt gtgaagaagg atgagttctc caccaggc
780
aaccagacgg accaccacag gatgtccggc gggaggcagg aggagttcg gacgtggctg

840

agggaggaat gggggcgcac gctggaggac atcttccacg agcacatgca ggagctcatc

900

ctgatgaagt tcatactacac cagtcagtagc gacaactgcc tgacctaccg ccgcattctac

960

ctgcccggca gcccggccga cgacctcatc aagcctggcc tcttcaaagg tacctatggc

1020

agccacggcc tggagattgt gatgctcagg ttccacggcc ggcgtgccag gggcaccaag

1080

atcacggcg accccaacat ccccgctggg cagcagacag tggagatcga cctgaggcat

1140

cgatatccagc tgccccaccc cgagaaccag cgcaacttca atgagctctc ccgcattcgtc

1200

ctggaggtgc gcgagaggggt gcccaggag cagcaggaag gcgggcacga ggccggcgag

1260

ggtcgtggcc ggcaggggcc cccggagtcc cagccaagcc ctgcccagcc cagggcagag

1320

gcgcaggca agggcccaga tgggacaccc ggtgaggatg gtggcgagcc tggggatgcc

1380

gtatctgcgg ccgagcagcc tgcccaagtgt gggcaggggc agccgttcgt gctgcccgtg

1440

ggcgtgagct ccaggaatga ggactacccc cgaacctgca ggatgtgttt ttatggcaca

1500

ggcctcatcg cggggccacgg cttcaccaggc cctgaacgca cccccggggc cttcatcctc

1560

ttcgatgagg accgcttcgg gttcgtctgg ctggagctga aatccttcag cctgtacagc

1620

cgggtccagg ccacaccttccg gaacgcagat ggcgcgtccc cacaggcctt cgatgagatg
1680
ctcaagaaca ttcatgtccct cacctcctga cggccacat cttgccggc acatccccgg
1740
tggctctggg gctctgaact ctgacacctgt aatagaagca gcatgcactt tggaaatccg
1800
gcctttgac cagaacgcac acctcgtcgg ggggcccagt ccagccaccc cccagcactt
1860
tatgttagaga gtgtgacata gacctgcata tttgtcagtg ccatgtatggg agaagctgag
1920
catgtcttac caaaaacaga gagaacgagc ctgaatacag cagatgttagg ggacagccgt
1980
gggaccgcgt gagaattgaa gcgggtgggt tcccgaccc tgggtggct ggtgggtttc
2040
tcgggaagca ggacccttccct gactggtgct ctccctgtga gcggatagag tgatagactg
2100
ggtcgtgtgt gagacgcatg tgctccaccc cactcctttt gggggaaagcc aggcaacagt
2160
ggcctctggg aggggggtcag gaagaggcga acagctcagg cagcgcaggt gtgatggca
2220
cagtaacgcag agcaagctcg ggaagttggg aggatctcag gttggggcc gggactctgg
2280
agtgaatccc catttctcta ccggcttgct tggagttgg acagaagcat ttcacctctg
2340
atctcagctt ccccacctgt ggagtgggtt tagtgcaccc agtcaactagg gaatgtcacc
2400
tgaatgcaca gcccagccca tgcacctgcc ccagccctc cagtttggg gccaaggcca
2460

tcgttccagc cacttgactg tcctcgacgg cctgttccag acagggcggtt tgttttgtcc
2520
atgccttcct ccctgcacgc acacggcgtc aaaaccaagc tgccggccac tgtctccaga
2580
acgcaaggct ccaggccccgt gtgtctgaag cagttagtgg tccacacagg tgccaggagt
2640
gcccatatga gatgacgagg aaacccttt gcaggtgagg ggacagcttt ctagaaaagc
2700
cacacctgca tctggggaca cactttggaa agtgggaccc tccagcctgg agacccatg
2760
gactgatgcc tccactgctg ttgtccccat gttgtttaa cacctgcgtg tggggacccc
2820
atctgaggc ttggctgagg ttggcatctc ctgaagaaca gagagcacgg tgtccagagc
2880
tggcccttcc cccagccac agccagctcc gtgcccaggt gggcgcccc agcgagcctt
2940
ccctctctgc cgcttgcct ttgtctggg ctgctccaag tccttgtgt gggcaccctg
3000
gacacgtcct gctggtgagg gacctcgaaa aggtgacagt ctgtgtgact tggtgtggag
3060
accaacctga ggatgtcctg gaaaatgttt tcctgatgaa ttttccttg actggccttt
3120
aaagaacata agaattccca ttgcccagcc tcagtcatt tggcaaatgc ttactttgt
3180
tcccagagtc agagaattgg caaaggttcc taaaatggtaa tctggccggc ctgggagaaa
3240
gactcacgag aaaagccagt ggagaaagcg cccttccagg gcccagcag cgggagccac

3300
gcagaccccg aggccgacact gctggctctt gtgtgtggcc ccagtttcta gcggcttttg
3360
cagcattagc ctacaagett tgtcaactccc tgccctctgt ggtggtcact gtttttctct
3420
cttgccaaat gaggcagtct ctgagtgacg gtgactgtgg ctttgaagcc tggaggactg
3480
ttggcatgt agactggcac cttgaagatt caccattgtt taaataaaat caagcaaatg
3540
ctttttacc aagagcccgaa gcctcgctct aagggacgca gtccttagagg cgtgccctt
3600
ggggcttcaa gagcacactg tggacgcac gtgcttctga ttaaaggaat ctcagatctc
3660
aattacgctt ccagtgtttg ggtatagaaa tagcttccac ccatcatgtc tcagccatgg
3720
gctgttggtc agttcatgtg gtcctggtt ctggtgtgtt tggtggggc ggggtctct
3780
ccatggtgtt gacctgcagt gatgccaggc agggccagag ccacacagcc agaaaaggga
3840
ggcccttttg gccgcacagc cagtccttc agtcgtgact acaggtcttg tttttccgc
3900
tccgatgtgt ctttagccag ttcttggtc cggttctgtt gggacaggca ctgaatctgc
3960
gcgcctcaaa acagcagctt ccctccggg ggagggcatc caccctctca gggatcctg
4020
caggtggccc atttcctgca ggtgagaact cggaagggtc gatgtcgtca tcagaggct
4080
aaggcagct gagagttgga taaaaccgtt tccaaggagg aggctgagta acccagttca

ES 2 433 012 T3

4140

ggacagccaa gcgcattagg ctgttgggg gaagggtggca ggtggagttg ggagggttggg

4200

actctccatc ttttgcacca cggatgcctt tctgtcgctg tctcactctg gggcaggata

4260

aagtctgctc tctggagtgg ggctgcctgc agtgcagctc tgcacacacctg aacgtttct

4320

ttgtcacttg tttggaaatg atgtgattga agatttcaga gaggtcatttgg gaggttttc

4380

tgtgccggca ctgaatgttc atttgcattga ggaagttgca aacgaottct gcaggctgag

4440

attcaaggca ggtggatttg gggccctca gcccacctgg gccgtgaccc caagtgtcca

4500

ctgctgagtg tgagtggctt tgcaggctg gtgggtggag agcctcaggc tccctcccttc

4560

ttcgttcctg accatgcctt gggccccttc agtctgcctg cggctctgtg gcattccctgc

4620

cctgacactg ggcacctgtg cccctagcaa gcccacctgg cacacgagga gggaggggtg

4680

ggtgccatg tccttcctt agccacatgc cctgctggcc gctccattct gagctttgtg

4740

cagaaccggg tctgagctgg agatttttct ctgagaacct gcagttgtgc tgcagccgca

4800

cgcaagggcc cttcagccgc tggctctggc ttccctgact cctcaggccg tgttcaccctt

4860

caggctttct cacctgcaca cggtaggcc attttcagtg ctgtggca gtcacggaca

4920

gcagcagaaa ctcctcagcc cctttgttac ttcagaacgc ctgcccacat gcatcttctg
4980
agctcgtgtt gtcctcatgg ccgtggggtc ttgggtgcga acaggagatg ctgagctgtg
5040
gtccaccgtc caaggtgctg cagaaaggcag ctaggctatt ttaggatgtt tctattctgg
5100
ttgctgcatt cgtggtgtaa cttaaagaa cacttacggg aatgtgctca tagaaccatc
5160
acctgtcctg agaataaaaac tcctggaatc atg
5193
<210> 16
<211> 5924
<212> ADN
5 <213> Homo sapiens
<400> 16
ggcatggcggttgtgtgtcg ctttgcggc gtggggccgt cgcgccggatg tcggcgccgc
60
cagcagcgcc ggggccccggc cgagacggcg gcggccgaca gcgagccgga cacagacccc
120
gaggaggaggc gcatcgaggc tagcgccggg gtcggggcg gttgtgcgc gggccctcg
180
ccgcgcgcgc cgcgcgtgtc gctgctggag ctgcgcgcgc agctgctggt ggagatcttc
240
gcgtcgctgc cgggcacgga cttaccgcgttggccagg tctgcacgaa gttccggcgc
300
atcctccaca ccgacaccat ctggaggagg cgttgccgtg aggagtatgg tgtttgcgaa
360
aacttgcgga agctggagat cacaggcgtg tcttgcggg acgtctatgc gaagctgctt
420

caccgatata gacacatttt gggattgtgg cagccagata tcgggccata cgaggactg
480
ctgaacgtgg tggtggaacgg cctgttcatc atcgggtgga tgtacctgcc tccccatgac
540
ccccacgtcg atgaccctat gagattcaag cctctgttca ggatccacct gatggagagg
600
aaggctgcca cagtggagtg catgtacggc cacaaagggc cccaccacgg ccacatccag
660
attgtgaaga aggatgagtt ctccaccaag tgcaaccaga cggaccacca caggatgtcc
720
ggcgggaggc aggaggagtt tcggacgtgg ctgagggagg aatggggcg cacgctggag
780
gacatcttcc acgagcacat gcaggagctc atcctgatga agttcatcta caccagtcag
840
tacgacaact gcctgaccta cgcgcgcatt tacctgccgc ccagccgccc cgacgacctc
900
atcaaggctg gcctttcaa aggtacctat ggcagccacg gcctggagat tgttatgttc
960
agcttccacg gcccgcgtgc caggggcacc aagatcacgg gcgaccccaa catccccgt
1020
gggcagcaga cagtggagat cgacctgagg catcgatcc agctgccccga cctcgagaac
1080
cagcgcaact tcaatgagct ctccgcgcatt gtcctggagg tgcgcgagag ggtgcgcac
1140
gagcagcagg aaggcgggca cgaggcgggc gagggtcgtg gccggcaggg ccccccggag
1200
tcccagccaa gcccgcgcatt gcccaggca gaggcgcgcatt gcaaggccc agatggaca
1260

cctggtaagg atggtggcga gcctgggat gccgtagctg cggccgagca gcctgcccag
1320
tgtggcagg ggcagccgtt cgtgctgccc gtggcggtga gctccaggaa tgaggactac
1380
ccccgaacct gcaggatgtg ttttatggc acaggcctca tcgcgggcca cggtttcacc
1440
agccctgaac gcaccccccgg ggtcttcatac ctattcgatg aggaccgott cgggttcgtc
1500
tggctggagc tgaaaatcatt cagcctgtac agccgggtcc aggccacott ccggAACGCA
1560
gatgcgccgt cccccacagggc cttcgatgag atgctcaaga acattcagtc cctcacctcc
1620
tgaccggcca catccttgcc gccacatccc gggtggctct ggggctctga actctgaccc
1680
gtgaatagaa gcagcatgca cttggaaat ccggcctttt gaccagaacg cacacctcg
1740
cggggggccc agtccagcca ccccccagca ctttatgttag agagtgtgac atagacctgc
1800
atatttgtca gtgccatgat ggaagaagct gagcatgtct taccaaaaac agagagaacg
1860
agcctgaata cagcagatgt aggggacacgc cgtgggaccg cgtgagaatt gaagcggtgg
1920
ggttcccgca ccctgggctg gctgggtggtt ttctcgaa gcaggaccct cctgactgg
1980
gctcttcctg tgagcggata gagtgataga ctgggtcggt tgtagacgc atgtgctcca
2040
ccccactcct tttggggaa gccaggcaac agtggcctct gggaggggt caggaagagg

2100
cgaacagctc aggccggca ggtgtatgg gcacagtacg cagagcaagc tcggaaagg
2160
ggtaggatct caggcttggg gccgggactc tggagtgaat ccccatttct ctaccggctt
2220
gcttggagtt tggacagaag cattcacct ctgatctcag cttccccacc tgtggagtgg
2280
gttttagtgac ctgagtcact agggaatgtc acctgaatgc acagcccagc ccatgcacct
2340
gccccagccc ctccagctt ggagccaagg ccatacggtcc agccacttga ctgtcctcga
2400
cggcctgttc cagacagggc gtttgttttgc tccatgcctt cctccctgca cgcacacggc
2460
gtcaaaacca agctgccggc cactgtctcc agaacgcaag gctccaggcc cgtgtgtctg
2520
aagcagttagtgg tggccacac aggtgccagg agtgccata tgagatgacg agaaaacccc
2580
tttgcaggtg aggggacagc tttctagaaa agccacaccc gcatctgggg acacactttg
2640
gaaagtggga ccctccagcc tggagacccc atggactgat gcctccactg ctgtgtgccc
2700
catgttgtgt taacacctgc gtgtggggac cccatctgag gtcttggctg aggttggcat
2760
cteectgaaga acagagagca cggtgtccag agctggccct tccccagcc cacagccagc
2820
tccgtgcccc agtgggcgtc cccagcgagc ctccctctc tgccgcttgt ccttgtgtct
2880
gggctgctcc aagtcccttgt gctggccacc ctggacacgt cctgctggtg agggacctcg

ES 2 433 012 T3

2940

ggaaggtgac agtctgtgtg ccttgggtgtg gagaccaacc tgaggatgtc ctggaaatg

3000

tttcctgat gaatttctcc ttgactggcc tttaaagaac ataagaatc ccattgccca

3060

gcctcagtgc atttggcaaa tgcttacttt gcttcccaga gtcagagaat tggcaaaggt

3120

tcctaaatgg taatctggcc ggcctggag aaagactcac gagaaaagcc agtggagaaa

3180

gcgccttcc agggcgccag cagcgggagc cacgcagacc ccgaggcgca cctgctggct

3240

cttggtgtg gccccagttt ctagcggctt ttgcagcatt agcctacaag ctttgtcact

3300

ccctgccttc tgtggtggtc actgtttttc tctcttgcctt aatgaggcag tctctgagtg

3360

acggtgactg tggccttcaa gcctggagga ctgttggca tgttagactgg caccttgaag

3420

attcaccatt gtttaaataa aatcaagcaa atgctttttt accaagagcc cgagcctcgc

3480

tctaaggac gcagtcctag aggctgtccc tttggggctt gaagagcaca ctgtggacg

3540

cacgtgcttc tgattaaagg aatctcagat ctcaattacg cttccagtgt ttgggtatag

3600

aaatagcttc cacccatcat gtctcagcca tgggctgttg gtcagttcat gtggctcctg

3660

gttctgggtgt gtatgttggg ggcgggggtc tctccatgggt ggtgacctgc agtgtatgc

3720

ggcagggcc a gacccacaca gccaggaaag ggaggcctt ttggccgcac agccagtccc .
3780
ttcagtcgtg actacaggc tcgtttttc cgctccgatg tgccttagc cagttcttgg
3840
ctccggttct gtagggacag gcactgaatc tgcgccctc aaaacagcag ctcccttcc
3900
gggggagggc atccaccctc tcaggggatc ctgcaggtgg cccatttctc gcaggtgaga
3960
actcggaagg gctgatgtcg tcatcagagg cctaaggca gctgagagtt ggataaaaacc
4020
gtttccaagg aggaggctga gtaacccagt tcaggacago caagcgcatt aggcttgatt
4080
ggggaaagggtg gcaggtggag ttgggaggtt gggactctcc atctttgca ccacggatgc
4140
ctttctgtcg ctgtctcaact ctggggcagg atcaagtctg ctctctggag tggggctgcc
4200
tgcaagtgcag ctctgcacac ctgaacgtgt tcttgcac ttgtttggaa atgatgtgat
4260
tgaagatttc agagaggcata ttggaggctt ttctgtgccg gcactgaatg ttcatggca
4320
tgaggaagtt gcaaacgact tctgcaggct gagattcaag gcaggtggta ttggggtccc
4380
tcagccccacc tggccgtga cctcaagtgt ccactgctga gtgtgagtgg ctggcaggg
4440
ctgggtggtgg gagagcctca ggctccctcc ttcttcgttc ctgaccatgc cctggggcccc
4500
ttcagtcgtc ctgcggctct gtggcatccc tgccctgaca ctgggcacct gtgcggcttag
4560

caagcccacc tggcacacga ggagggaggg gtgggtggcc atgtccttcc tctagccaca
4620
tgccctgctg gccgctccat tctgagctt gtgcagaacc gggctgagc tggagatttt
4680
tctctgagaa cctgcagttg tgctgcagcc gcacgcaagg gcccttcagc cgctggctct
4740
ggcttcctg actcctcagg gcgtgttcac ccccaggctt tctcacctgc acacggttag
4800
gccatttca gtgctcggtgg gcagtcacgg acagcagcag aaactcctca gccctttgt
4860
tacttcagaa cgcctgcccc catgcattt ctgagctcgt gttgtcctca tggccgtggg
4920
gtcttgggtg cgaacaggag atgctgagct gtggccacc gtccaagggtg ctgcagaaag
4980
cagctaggct cttttaggat gtttctattt tggttgctgc cttcgtggtg taacttttaa
5040
gaacacttac gggaaatgtgc tcatagaacc atcacctgtc ctgagaataa aactcctgga
5100
atcatgatca agtccagtgt taacgtggcc caaccgtct gtacttctgg ggagagacca
5160
ggaacatcac tggactcctc atccccgtaa ttatttagag aagatgcaag cagcagatag
5220
tctccatgcg gctggacttt tttttgttgt tttttgagac agggcttgc tctgtcacct
5280
gggctggagt gcagagcggc gatcatggct ccctgaggcc tcaacctact aggctcaagc
5340
tgtctgcccc ccttagcctc ccaagtagct gggaccacag gcacccacca ccaccatgct

5400
tggcttaactt gttttttag agatggagtt ttgccatgtt gtcaggttg gtctcgaaact
5460
cccgatctca ggtgatccac ccgcctcgcc ctcccaaagt gctgggatta caggcgtgag
5520
cccctgcgcc ccagccttgg ggccctgtctt tgaatggaa tgagactgtg caaacccgtgg
5580
actaccctgt gtcacccaca gtcagtgcc ctgcctgccc gccctcaggg gctgctgacc
5640
gggagaccag ccagagcacg agggggtcag ggctgtgtgg gttttggcct gattctgcat
5700
ttggttgttt ctgggggcca tgtagcctgc ctgcattagg aaagcgctgt gccatctgat
5760
catgagcacc tctgcacccc ctggtaaggt gaccttgcag caggagctgt gccctgcctg
5820
ggtaggcacc cactaggtag gaccggagca atcctggcag ccgcacactg cacccgtgca
5880
cttggtttctc ctcacagttt caagtaaatc cgttttgaa ggct
5924
<210> 17
<211> 6011
<212> ADN
5 <213> Homo sapiens

<400> 17
ggcatggcggttgtgtgctcg cctttgcggc gtggggccgt cgcgccggatg tcggcgccgc
60
cagcagcgcc ggggccccggc cgagacggcg gcggccgaca gcgagccgga cacagacccc
120
gaggaggagc gcatcgaggc tagcgccggg gtcggggcg gcttgtgcgc gggcccccctcg

180

ccggccgcccc cgcgctgctc gctgctggag ctgccgcccc agctgctgggt ggagatcttc

240

gcgtcgctgc cgggcacgga cctaccaggc ttggcccagg tctgcacgaa gttccggcgc

300

atccctccaca ccgacaccat ctggaggagg cgttgccgtg aggagtatgg tgtttgcgaa

360

aacttgcgga agctggagat cacaggcgtg tcttgcggg acgtctatgc gaagcgtata

420

aaccctcgcg tgaagtgcgg acgtttgtg aaaatttcctc ctgattatga gcacatggcg

480

tacagagacg tttacacctg cctgcttcac cgatatacgac acatttggg attgtggcag

540

ccagatatcg ggccatacgg aggactgctg aacgtgggtgg tggacggcct gttcatcatc

600

gggtggatgt acctgcctcc ccatgaccccc cacgtcgatg accctatgag attcaagcct

660

ctgttcagga tccacctgat ggagaggaag gctgccacag tggagtgcgt gtacggccac

720

aaagggcccc accacggcca catccagatt gtgaagaagg atgagttctc caccaagtgc

780

aaccagacgg accaccacag gatgtccggc gggaggcagg aggagttcg gacgtggctg

840

agggaggaat gggggcgcac gctggaggac atcttccacg agcacatgca ggagctcatc

900

ctgatgaagt tcacatcac cagtcagtac gacaactgcc tgacctaccg ccgcacatctac

960

ctgccgcccc gcccggccga cgacctcatc aagcctggcc tcttcaaagg tacctatggc

ES 2 433 012 T3

1020

agccacggcc tggagattgt gatgctcagc ttccacggcc ggcgtgccag gggcaccaag

1080

atcacggcg accccaacat ccccgcgtggg cagcagacag tggagatcga cctgaggcat

1140

cggatccagc tgcccaccc cgagaaccag cgcaacttca atgagctctc ccgcacatgtc

1200

ctggaggtgc gcgagaggggt gcccgcaggag cagcaggaag gcgggcacga ggccccgcag

1260

ggtcgtggcc ggcaggccc cccggagtcc cagccaaagcc ctgcccagcc cagggcagag

1320

gcgcaggca agggcccaga tggcacacct ggtgaggatg gtggcgagcc tggggatgcc

1380

gtagctgcgg ccgagcagcc tgcccagtgt gggcaggggc agccgttgt gctgcccgtg

1440

ggcgtgagct ccaggaatga ggactacccc cgaacctgca ggatgtgttt ttatggcaca

1500

ggccatcg cggccacgg ctaccaggc cctgaacgca ccccccgggt ctatccctc

1560

tgcatgagg accgcttcgg gttcgtctgg ctggagctga aatccttcag cctgtacagc

1620

cgggtccagg ccaccttcgg gaacgcagat gcccgtccc cacaggcctt cgatgagatg

1680

ctcaagaaca ttcaatccct caccccttga cccgcacat cttggccgcc acatccccgg

1740

tggctctggg gctctgaact ctgacccgttg aatagaagca gcatgcactt tggaaatccg

1800

gcctttgac cagaacgcac acctcgtcgg ggggccagt ccagccaccc cccagcactt
1860
tatgttagaga gtgtgacata gacctgcata tttgtcagtgc ccatgtatggaa agaagctgag
1920
catgtcttac caaaaacaga gagaacgagc ctgaatacag cagatgttagg ggacagccgt
1980
gggaccgcgt gagaattgaa gcgggtgggt tccccaccc tgggctggct ggtggtttc
2040
tcgggaagca ggaccctcct gactggtgct cttccatgtga gcggatagag tgatagactg
2100
ggtcgtgtgt gagacgcatt tgctccaccc cactcctttt gggggaaagcc aggcaacagt
2160
ggcctctggg agggggtcag gaagaggcga acagctcagg cagcgcaggt gtatggca
2220
cagtagcag agcaagctcg ggaagttggt aggatctcag gcttggggcc gggactctgg
2280
agtgaatccc catttctcta ccggcttgct tggagttgg acagaagcat ttcacctctg
2340
atctcagctt ccccacctgt ggagtgggtt tagtgcacgt agtcactagg gaatgtcacc
2400
tgaatgcaca gcccagccca tgcacctgcc ccagcccctc cagctttgga gccaaaggcca
2460
tctgtccagc cacttgactg tcctcgacgg cctgtccag acagggcggt ttgtttgtcc
2520
atgccttcct ccctgcacgc acacggcgtc aaaaccaagc tgccggccac tgtctccaga
2580
acgcaaggct ccaggcccgt gtgtctgaag cagttagtgg tccacacagg tgccaggagt
2640

gccccatata gatgacgagg aaaccccttt gcaggtgagg ggacagctt ctagaaaagc
2700
cacacctgca tctggggaca cactttggaa ,agtgggaccc tccagcctgg agacccatg
2760
gactgatgcc tccactgctg tgtccccat gttgtgttaa cacctgcgtg tggggacccc
2820
atctgaggc ttggctgagg ttggcatctc ctgaagaaca gagagcacgg tgtccagagc
2880
tggcccttcc cccagccac agccagctcc gtgcccagt gggcgcccc agcgaggcctt
2940
ccctctctgc cgcttgcct tgtgtctgg ctgctccaag tccttgtct gggcacccctg
3000
gacacgtcct gctggtgagg gacctcgaaa aggtgacagt ctgtgtgcct tggtggtggag
3060
accaacctga ggatgtcctg ggaaatgttt tcctgatgaa ttttccttg actggcctt
3120
aaagaacata agaattccca ttgcccagcc tcagtgcatt tggcaaatgc ttactttgct
3180
tcccagagtc agagaattgg caaaggttcc taaatggtaa tctggccggc ctgggagaaaa
3240
gactcacgag aaaagccagt ggagaaagcg cccttcagg gggcagcag cgggagccac
3300
gcagaccccg aggccaccc gctggcttt gtgtgtggcc ccagtttcta gggcttttg
3360
cagcattagc ctacaagctt tgtcactccc tgccctctgt ggtggtcact gttttctct
3420
cttgccaaat gaggcagtct ctgagtgacg gtgactgtgg ctttgaagcc tggaggactg

3480

ttgggcatgt agactggcac cttgaagatt caccattgtt taaataaaat caagcaaatg

3540

cttttttacc aagagcccgaa gcctcgctct aagggacgca gtcctagagg cgtgccctt

3600

ggggcttgaa gaggcacactg tgggacgcac gtgcttctga ttaaaggaat ctcagatctc

3660

aattacgctt ccagtgtttg ggtatagaaa tagcttccac ccatcatgtc tcagccatgg

3720

gctgttggtc agttcatgtg gtcctggtt ctggtgtgta tggtggggc gggggctct

3780

ccatggtggt gacctgcagt gatgccaggc agggccagag ccacacagcc aggaaaggga

3840

ggcccttttg gccgcacagc cagtccttc agtcgtgact acaggtcttg tttttccgc

3900

tccgatgtgt ctttagccag ttcttggtc cggttctgta gggacaggca ctgaatctgc

3960

gcgcctcaaa acagcagctt cccttcggg ggagggcattt caccctctca ggggatcctg

4020

caggtggccc atttcctgca ggtgagaact cggaagggtt gatgtcgatca tcagaggct

4080

aaggcagct gagagttgga taaaaccgtt tccaaggagg aggctgagta acccagttca

4140

ggacagccaa gcgcatttgg cttgattggg gaaggtggca ggtggagttt ggaggttggg

4200

actctccatc ttttgcacca cggatgcctt tctgtcgctg tctcactctg gggcaggatc

4260

aagtctgctc tctggagtgg ggctgcctgc agtgcagctc tgcacacaccc aacgtttct

ES 2 433 012 T3

4320

ttgtcacttg tttggaaatg atgtgattga agatttcaga gaggtcattg gaggctttc

4380

tgtgccggca ctgaatgttc atttgcattga ggaagttgca aacgacttct gcaggctgag

4440

attcaaggca ggtggatttg gggccctca gcccacctgg gccgtgacct caagtgtcca

4500

ctgctgagtg tgagtggctt tgcaggcctg gtgggtggag agcctcaggc tccctccttc

4560

ttcggtccctg accatgcctt gggccccttc agtctgcctg cggctctgtg gcattccctgc

4620

cctgacactg ggcacctgtg cccctagcaa gcccacctgg cacacgagga gggaggggtg

4680

ggtgccatg tccttcctct accacatgc cctgctggcc gtcatttctt gagctttgtg

4740

cagaaccggg tctgagctgg agattttctt ctgagaacct gcagttgtgc tgcagccgca

4800

cgcaagggcc cttcagccgc tggctctggc ttccctgact ctcagggcg tgttcacccc

4860

caggctttctt cacctgcaca cggttaggcc atttcagtg ctctggca gtcacggaca

4920

gcagcagaaa ctcctcagcc cctttgttac ttcagaacgc ctgcccacat gcatcttctg

4980

agctcgtgtt gtcctcatgg ccgtggggtc ttgggtgcga acaggagatg ctgagctgtg

5040

gtccaccggtc caaggtgctg cagaaaggcag cttaggcttta ttaggatgtt tctattctgg

5100

ttgctgcctt cgtggtgtaa ctttaagaa cacttacggg aatgtgctca tagaaccatc
 5160
 acctgtcctg agaataaaaac tcctggaatc atgatcaagt ccagtgttaa cgtggcccaa
 5220
 cctgtctgtta cttctgggga gagaccagga acatcactgg actcctcatac cccgttaatta
 5280
 ttttagagaag atgcaagcag cagatagtct ccatgcggct ggtacttttt ttgttgtttt
 5340
 ttgagacagg gtcttgctct gtcacctggg ctggagtgca gagcggcgat catggctccc
 5400
 tgaggcctca acctactagg ctcaagctgt ctgccccct tagcctccca agtagctggg
 5460
 accacaggca cccacccacca ccatgttgg ctaacttttt ttttagaga tggagttttg
 5520
 ccatgttgct caggttggtc tcgaactccc gatctcaggt gatccaccccg cctcggcctc
 5580
 ccaaagtgct gggattacag gcgtgagccc ctgcgcacca gccttggggc ctgtctttga
 5640
 atggaatga gactgtgcaa accgtggact accctgtgtc acccacagct cagtggcctg
 5700
 cctgccggcc ctcaggggct gctgaccggg agaccagcca gggcacgagg gggtcagggc
 5760
 tgtgtgggtt ttggcctgat tctgcatttgc gttgtttctg gggccatgt agcctgcctg
 5820
 cattagaaaa gcgctgtgcc atctgatcat gagcacctct gcacccctg gtaaggtgac
 5880
 cttgcagcag gagctgtgcc ctgcctgggt aggcacccac tagtaggac cggagcaatc
 5940
 ctggcagccg ccacactgcac ccgtgcactt gtttctcctc acagttcaa gtaaatccgt
 6000
 ttttgaaggc t
 6011
 <210> 18
 5 <211> 7521
 <212> ADN
 <213> Homo sapiens
 <400> 18

ES 2 433 012 T3

ggcatggcgg tgtgtgctcg cctttgcggc gtggggccgt cgccgcggatg tcggcgccgc
60
cagcagcgcc ggggccccggc cgagacggcg gcggccgaca gcgagccgga cacagacccc
120
gaggaggagc gcatcgaggc tagcgccggg gtcggggcg gcttgtgcgc gggccctcg
180
ccgcgcgcggc cgcgctgctc gctgctggag ctgcgcgcgg agctgctggt ggagatctc
240
gcgtcgctgc cgggcacgga cctaccacgc ttggcccagg tctgcacgaa gttccggcgc
300
atccctccaca ccgacacccat ctggaggagg cgttgcgtg aggagtatgg tgtttgcgaa
360
aacttgcgga agctggagat cacaggcgtg tcttgcggg acgtctatgc gaagctgctt
420
caccgatata gacacatttt gggattgtgg cagccagata tcgggccata cggaggactg
480
ctgaacgtgg tggtgacgg cctgttcatc atcgggtgga tgtacctgcc tccccatgac
540
ccccacgtcg atgaccctat gagattcaag cctctgttca ggatccacct gatggagagg
600

aaggctgcca cagtggagtg catgtacggc cacaaaggc cccaccacgg ccacatccag
660
attgtgaaga aggatgagtt ctccaccaag tgcaaccaga cggaccacca caggatgtcc
720
ggcgggaggc aggaggagtt tcggacgtgg ctgagggagg aatggggcg cacgctggag
780
gacatcttcc acgagcacat gcaggagctc atcctgatga agttcatcta caccagtcag
840
tacgacaact gcctgaccta ccgcgcatac tacctgccgc ccagccgccc cgacgacctc
900
atcaaggctg gcctttcaa aggtacctat ggcagccacg gcctggagat tgtatgctc
960
agttccacg gccggcgtgc cagggcacc aagatcacgg gcgacccaa catccccgt
1020
ggcagcaga cagtggagat cgacctgagg catcgatcc agctgcccga cctcgagaac
1080
cagcgcaact tcaatgagct ctccgcatac gtcctggagg tgcgcgagag ggtgcgcag
1140
gagcagcagg aaggcgggca cgaggcgggc gagggctgtg gccggcaggg ccccccggag
1200
tcccagccaa gccctgcacca gcccaggcga gagggcgcacca gcaagggccc agatggac
1260
cctggtgagg atggtggcga gcctgggat gccgtagctg cggccgagca gcctgcccag
1320
tgtggcagg ggcagccgtt cgtgtgcaccc gtggggcgtga gctccaggaa tgaggactac
1380
ccccgaacct gcaggatgtg ttttatggc acaggcctca tcgcgggcca cggcttcacc

1440

agccctgaac gcaccccccgg ggttttcattt ctcttcgatg aggaccgctt cgggttcgtc

1500

tggctggagc tgaaatcctt caggctgtac agccgggtcc aggccacett ccggaacgca

1560

gatgcgcgt cccccacagggc cttcgatgag atgctcaaga acattcagtc ctcacacctcc

1620

tgaccggcca catccttgcc gccacatccc gggtggctct ggggctctga actctgacct

1680

gtgaatagaa gcagcatgca ctttgaaat ccggccctttt gaccagaacg cacacctcg

1740

cggggggccc agtccagcca ccccccagca ctttatgttag agagtgtgac atagacactgc

1800

atatttgtca gtgccatgat ggaagaagct gagcatgtct tacaaaaaac agagagaacg

1860

agcctgaata cagcagatgt aggggacagc cgtgggaccg cgtgagaatt gaagcggtgg

1920

ggttcccgca ccctgggctg gctgggtggtt ttctcggaa gcaggaccct cctgactgg

1980

gctttccctg tgagcggata gagtgataga ctgggtcgtg tgtgagacgc atgtgctcca

2040

ccccactcct tttggggaa gccaggcaac agtggcctct gggagggggg caggaagagg

2100

cgaacagctc aggcagcgca ggtgtgatgg gcacagtacg cagagcaac tcggaaagtt

2160

ggtaggatct caggcttggg gccgggactc tggagtgaaat ccccatttct ctaccggctt

2220

gcttggagtt tggacagaag catttcaccc ctgatctcag ctccccacc tggagtg

ES 2 433 012 T3

2280
gttagtgac ctgagtcact agggaatgtc acctgaatgc acagccccgc ccatgcaccc
2340
gccccagccc ctccagctt ggagccaagg ccatcggtcc agccacttga ctgtccctcg
2400
cggcctgttc cagacagggc gtttgttttgc tccatgcctt cctccctgca cgcacacggc
2460
gtcaaaacca agctgccggc cactgtctcc agaacgcaag gctccaggcc cgtgtgtctg
2520
aagcagttagtggccacac aggtgccagg agtgccata tgagatgacg aggaaacccc
2580
tttgcaggtg aggggacacgc ttcttagaaa agccacacct gcacactgggg acacactttg
2640
gaaagtggga ccctccagcc tggagacccc atggactgtat gcctccactg ctgtgtggcc
2700
catgttgtgt taacacctgc gtgtggggac cccatctgag gtcttggctg aggttggcat
2760
ctcctgaaga acagagagca cggtgtccag agctggccct tccccccagcc cacagccagc
2820
tccgtgcccgc agtggggcgtc cccagcgagc cttccctctc tgccgcttgt ctttgtgtct
2880
gggcgtgtcc aagtcccttgt gctgggcacc ctggacacgt cctgctggtg agggacctcg
2940
ggaaggtgac agtctgtgtg ctttgtgtg gagaccaacc tgaggatgtc ctggaaatg
3000
ttttcctgat gaatttctcc ttgactggcc tttaaagaac ataagaattc ccattgccc
3060

gcctcagtgc atttggcaaa tgcttacttt gcttcccaga gtcagagaat tggcaaagg
3120
tcctaaatgg taatctggcc ggcctggag aaagactcac gagaaaagcc agtggagaaa
3180
gcccccttcc agggcggcag cagcgggagc cacgcagacc ccgaggcgca cctgctggct
3240
cttgtgtgtg gccccagttt ctagcggctt ttgcagcatt agcctacaag ctttgtcact
3300
ccctgccttc tgtggtggtc actgtttttc tctcttgcca aatgaggcag tctctgagtg
3360
acggtgactg tggccttgaa gcctggagga ctgttggca tgttagactgg caccttgaag
3420
attcaccatt gtttaataaa aatcaagcaa atgctttttt accaagagcc cgagcctcgc
3480
tctaaggac gcagtcctag aggctgtccc tttggggctt gaagagcaca ctgtggacg
3540
cacgtgcattc tgattaaagg aatctcagat ctcaattacg cttccagtgt ttgggtata
3600
aaatagcttc cacccatcat gtctcagcca tgggctgttg gtcagttcat gtggctcctg
3660
gttctgggtgt gtatgttggg ggcgggggtc tctccatgggt ggtgacctgc agtgtatgcca
3720
ggcagggcca gagccacaca gccagggaaag ggaggccttt ttggccgac acgcaggccc
3780
ttcagtcgtg actacaggtc ttgtttttc cgctccgatg tgccttagc cagttcttgg
3840
ctccggttct gtagggacag gcactgaatc tgcgcgcctc aaaacagcag ctcccttcc
3900

gggggagggc atccaccctc tcaggggatc ctgcaggtgg cccatttcct gcaggtgaga
3960
actcggaagg gctgatgtcg tcatcagagg cctaaggca gctgagagtt ggataaaacc
4020
gttccaagg aggaggctga gtaacccagt tcaggacagc caagcgcatt aggcttgatt
4080
gggaaaggtg gcaggtggag ttggaggtt gggactctcc atctttgca ccacggatgc
4140
ctttctgtcg ctgtctcaact ctggggcagg atcaagtctg ctctctggag tggggctgcc
4200
tgcatgtcag ctctgcacac ctgaacgtgt tcttgcac ttgtttggaa atgatgtgat
4260
tgaagatttc agagaggcata ttggaggctt ttctgtgccg gcactgaatg ttcatttgca
4320
tgaggaagtt gcaaacgact tctgcaggct gagattcaag gcaggtggta ttggggteccc
4380
tcagcccacc tggccgtga cctcaagtgt ccactgctga gtgtgagtggtt ctttcaggg
4440
ctgggtgtgg gagagccctca ggctccctcc ttcttcgttc ctgaccatgc cctggggcccc
4500
ttcagtctgc ctgcggctct gtggcatccc tgccctgaca ctgggcacct gtgcggcttag
4560
caagcccacc tggcacacga ggagggaggg gtgggtggcc atgtccttcc tctagccaca
4620
tgccctgctg gccgctccat tctgag{ttt } gtgcagaacc gggctgagc tggagat{ttt }
4680
tctctgagaa cctgcagttt tgctgcagcc gcacgcaagg gcccttcagc cgctggctct

4740

ggctccctg actcctcagg gcgtgttcac ccccaggctt tctcacctgc acacggttag

4800

gccattttca gtgctcgtag gcagtcacgg acagcagcag aaactctca gccccttgt

4860

tactcagaa cgccctgccc catgcattt ctgagctgt gttgtctca tggccgtggg

4920

gtcttgggtg cgaacaggag atgctgagct gtggtccacc gtccaagggtg ctgcagaaag

4980

cagctaggct cttttaggat gtttctattt tggttgtgc cttagtggtg taactttaa

5040

gaacacttac gggaatgtgc tcataagaacc atcacctgtc ctgagaataa aactcctgga

5100

atcatgatca agtccagtgt taacgtggcc caacctgtct gtacttctgg ggagagacca

5160

ggaacatcac tggactcctc atccccgtaa ttattnagag aagatgcaag cagcagatag

5220

tctccatgct gctggtaactt ttttgttgt ttttgagac agggctttgc tctgtcacct

5280

gggctggagt gcagagcggc gatcatggct ccctgaggcc tcaacctact aggctcaagc

5340

tgtctgcccc ccttagcctc ccaagtagct gggaccacag gcacccacca ccaccatgct

5400

tggctaactt gttttgttag agatggagtt ttgccatgtt gctcagggtt gtctcgaact

5460

cccgatctca ggtgatccac ccgcctcggc ctccccaaagt gctgggatta caggcgtgag

5520

ccccctgcgcc ccagccttgg ggccctgttt tgaatggaa tgagactgtg caaaccgtgg

ES 2 433 012 T3

5580

actaccctgt gtcacccaca gctcagtggc ctgcctgcgg gccctcaggg gctgctgacc

5640

gggagaccag ccagagcacg agggggtcag ggctgtgtgg gttttggcct gattctgcat

5700

ttggttgttt ctgggggcca tgttagcctgc ctgcattagg aaagcgctgt gccatctgat

5760

catgagcacc tctgcacccc ctggtaaggt gaccttgcag caggagctgt gccatgcctg

5820

ggtaggcacc cactaggttag gacccggagca atcctggcag ccgccacctg caccctgtca

5880

cttgtttctc ctcacagttt caagtaaatc cgttttgaa ggcttgttgt gtgtttgtg

5940

atttctttgg gaatatgagt tggacggagg cgagacccctt aagccatgcg agctgtggc

6000

ctgggaaccc agacttccca gcttcttgag gaagtgtcag atttcccgcg ttgacagaag

6060

ggagcattga agggatgcct tggagccca agactggttt tccctgtgtc cttccctttg

6120

acctggcattc agaggtgtct cgagtcccta cccagggacc cagaggagtt cggggccccag

6180

tagatttct tagatttaaag ccaaagttagt ttgcattatac tgcaacgagg acagatatgg

6240

gagggaaatgt gctgagagcc aggcatgatga actgaggatc tcattgatct ttctttgtg

6300

tttactaaac tcataatgttc ttgttaaacag ttcttttagca tagacaatgtga aagtacccccc

6360

tgttctcatc ccagcctccc cgtgagtcac tgctgcta at taatgctgtt agcttggaaat
6420
tgtagaaaca ggatgttttc catggtaatgcactcaaagt acaccctcgaa ttggcagaaa
6480
ttggcaagtg tgattttcca agtgttggca gtgatgcagg ggaacaggaa cgccagggtgg
6540
gcagctgttt tggggacagc tggtaactgc tcataggcact aaggacacgg gcccaggac
6600
tggcatctgc atcctgaggt gtcaccctc gggcaacgcg agagcccagg catggccac
6660
gcagggatgt tcattgctac actgtgacaa ctgtcacagg ccggaaggag gcaggtggac
6720
tacgggtggag ccacccatgc tgcacactgg cagacggca cacagccttg ttccgttgca
6780
aaacaagtga gagatggtat tggtgtaaca tgtaaaaatg caaatactta atttttatca
6840
attcatgtgt gggaaaagc tgaagatacg cgtggaaatg gtgtggtcac ttctaggggt
6900
gtcggagggt agaacttcaa ctgtttgct ttaaaaagta aggatcgcat ggcagaacta
6960
gcacatgttc acctgttgat cctgataccg tggattacga gaccccccct ctttctgtg
7020
tggttcagaa acaagccccct cagacaggac acagtgccca gggcagtga cctgcaggcc
7080
cacccaactgc catctccgct ggtctcgaaaa ttgccacata gcctgccagc tgccgctgct
7140
tcctgggtgc cctccagggaa gagcagggaa tcgtgggtcc ccggcggtgg gtgtttcctt
7200

ctccggggag agcaggggat cgtgggtccc cggtggtggg tgtttccttc tctaagggttt
7260
gctgctgttt ccaggccctt ctgtggggcc tgggtcctgt cctggggcca agccacgggg
7320
tcatcctcag ctgcactggg cgtgccaacc acaaacgagt cacttgctac aagcagcacc
7380
atgcagcctc ctgtctggac gagaccctgc cccccacaga ctggagacgc accccgattt
7440
cccgagtcac aggggaaagt gtggatctga taaggacta aatgtggcgt ctttcatatg
7500
tttcttttac atatttttt t
7521
<210> 19
<211> 7608
<212> ADN
5 <213> Homo sapiens
<400> 19
ggcatggcgg tgtgtgctcg cctttgcggc gtggggccgt cgcgccgatg tcggcgccgc
60
cagcagcgcc ggggccccggc cgagacggcg gggccgaca gcgagccgga cacagacccc
120
gaggaggagc gcatcgaggc tagcgccggg gtcggggcgc gtttgtgcgc gggccctcg
180
ccgcgcgcgc cgcgcgtgctc gctgctggag ctgccgcgc agctgctggt ggagatctc
240
gcgtcgctgc cgggcacgga cttaccggc ttggccagg tctgcacgaa gttccggcgc
300
atcctccaca ccgacaccat ctggaggagg cgttgccgtg aggagtatgg tgtttgcgaa
360

aacttgcgga agctggagat cacaggcgtg tcttgcggg acgtctatgc gaagcgata
420
aacccctcgcg tgaagtgcggg acgttttgtg aaaattctcc ctgattatga gcacatggcg
480
tacagagacg tttacacactg cctgcttcac cgatatacac acatttggg attgtggcag
540
ccagatatcg ggccataacgg aggactgctg aacgtggtgg tggacggct gttcatcatc
600
gggtggatgt acctgcctcc ccatgaccgg cacgtcgatg accctatgag attcaagcct
660
ctgttcagga tccacctgat ggagaggaag gctgccacag tggagtgcac gtacggccac
720
aaagggcccc accacggcca catccagatt gtgaagaagg atgagttctc caccaagtgc
780
aaccagacgg accaccacag gatgtccggc gggaggcagg aggagttcg gacgtggctg
840
agggaggaat gggggcgcac gctggaggac atcttccacg agcacatgca ggagctcatc
900
ctgatgaagt tcacatcac cagtcagttac gacaactgccc tgacctaccg ccgcacatctac
960
ctgcccggccca gcccggccca cgacacctcata aagectggcc tcttcaaagg tacctatggc
1020
agccacggcc tggagattgt gatgctcagg ttccacggcc ggcgtgccag gggcaccaag
1080
atcacgggcg accccaacat ccccgctggg cagcagacag tggagatcga cctgaggcat
1140
cgatccagc tgcccgaccc cgagaaccag cgcaacttca atgagctctc ccgcacatcgac

1200
ctggaggtgc gcgagagggc gcccaggag cagcaggaag gcgggcacga ggccggcgag
1260
ggtcgtggcc ggcagggccc ccgggagtcc cagccaagcc ctgcccagcc cagggcagag
1320
gcgcaggca agggcccaga tggcacacct ggtgaggatg gtggcgagcc tggggatgcc
1380
gtagctgcgg ccgagcagcc tgcccagtgt gggcagggc agccgtttagt gctgcccgtg
1440
ggcgtgagct ccaggaatga ggactacccc cgaacctgca ggtatgtgttt ttatggcaca
1500
ggcctcatcg cgggccacgg ctaccaggc cctgaacgca cccccgggtt ctcatcctc
1560
ttcgatgagg accgcttcgg gttcgctctgg ctggagctga aatccttcag cctgtacagc
1620
cgggtccagg ccaccttcgg gaacgcagat ggcgtccc cacaggcatt cgatgagatg
1680
ctcaagaaca ttcagtcct cacctctga cggccacat cttgccgcc acatccggg
1740
tggctctggg gctctgaact ctgacctgtg aatagaagca gcatgcactt tggaaatccg
1800
gcctttgac cagaacgcac acctcgctgg gggcccaagt ccagccaccc cccagcactt
1860
tatgtagaga gtgtgacata gacctgcata tttgtcagtg ccatgatgga agaagctgag
1920
catgtcttac caaaaacaga gagaacgagc ctgaatacag cagatgtagg ggacagccgt
1980
gggaccgcgt gagaattgaa gcgggtgggt tcccgaccc tgggctggct ggtggtttc

2040

tcgggaagca ggaccctcct gactggtgct ctccctgtga gcggatagag tgatagactg

2100

ggtcgtgtgt gagacgcattg tgctccaccc cactcctttt gggggaaagcc aggcaacagt

2160

ggcctctggg aggggggtcag gaagaggcgaa acagctcagg cagcgcaggt gtgatggca

2220

cagtagcagcag agcaagctcg ggaagtttgtt aggatctcag gcttggggcc gggactctgg

2280

agtgaatccc catttctcta ccggcttgct tggagtttgg acagaagcat ttcacctctg

2340

atctcagctt cccccacctgt ggagtgggtt tagtgcacctg agtcaactagg gaatgtcacc

2400

tgaatgcaca gcccagecca tgcacctgcc ccagcccccc cagctttggaa gccaaaggcca

2460

tcgttccagc cacttgactg tcctcgacgg cctgttccag acagggcggtt tgttttgtcc

2520

atgccttcct ccctgcacgc acacggcgtc aaaaccaagc tgccggccac tgtctccaga

2580

acgcaaggct ccaggcccggt gtgtctgaag cagtgagtgg tccacacagg tgccaggagt

2640

gccccatatga gatgacgagg aaaccccttt gcaggtgagg ggacagcttt ctagaaaagc

2700

cacacctgca tctggggaca cactttggaa agtgggaccc tccagcctgg agaccccatg

2760

gactgatgcc tccactgctg tggcccccattt gttgtgttaa cacctgcgtg tggggacccc

2820

atctgaggtc ttggctgagg ttggcatctc ctgaagaaca gagagcacgg tgtccagagc
2880
tggcccttcc cccagccac agccagctcc gtgcccaggt gggcgcccc agcgagcctt
2940
ccctctctgc cgcttgcct tgcgtctggg ctgctccaag tccttgtct gggcaccctg
3000
gacacgtctt gctggtgagg gacctcgaaa aggtgacagt ctgtgtgcct tggtgtggag
3060
accaacctga ggatgtcctg ggaaatgttt tcctgatgaa tttctccttg actggccttt
3120
aaagaacata agaattccca ttgcccaagcc tcagtgcatt tggcaaatgc ttactttgt
3180
tcccagagtc agagaattgg caaaggttcc taaatggtaa tctggccggc ctgggagaaaa
3240
gactcacgag aaaagccagt ggagaaagcg cccttccagg gccggcagcag cgggagccac
3300
gcagaccccg aggccacact gctggctttt gtgtgtggcc ccagttctta gcggctttt
3360
cagcatttagc ctacaagctt tgcactccc tgccctctgt ggtggtcact gttttctct
3420
cttgc当地 gaggcagtct ctgagtgacg gtgactgtgg cttgaagcc tggaggactg
3480
ttggcatgt agactggcac cttgaagatt caccattgtt taaataaaat caagcaaatg
3540
ctttttacc aagagcccgaa gcctcgctct aagggacgca gtccttagagg cgtggcccttt
3600
ggggcttgaa gagcacactg tgggacgcac gtgcttctga ttaaaggaat ctcagatctc
3660

aattacgctt ccagtgtttg ggtatagaaaa tagcttccac ccatcatgtc tcagccatgg
3720
gctgttggtc agttcatgtg gctcctggtt ctggtgtgt a tgttgggggc gggggctct
3780
ccatggtggt gacctgcagt gatgccaggc agggccagag ccacacagcc aggaaaggga
3840
ggcctttttg gccgcacagc cagtccttc agtcgtact acaggtcttg tttttccgc
3900
tccgatgtgt ccttagccag ttcttggtc cggttctgt a gggacaggca ctgaatctgc
3960
gcccctcaaa acagcagctt ccctccggg ggagggcatac caccctctca gggatcctg
4020
caggtggccc atttcctgca ggtgagaact cggaaaggct gatgtcgta tcagaggct
4080
aaggcagct gagagttgga taaaaccgtt tccaaggagg aggctgagta acccagttca
4140
ggacagccaa gcgcattagg cttgattggg gaaggtggca ggtggagttt gggatgggg
4200
actctccatc ttttgcacca cggatgcctt tctgtcgctg tctcaactctg gggcaggatc
4260
aagtctgctc tctggagtgg ggctgcctgc agtgcagctc tgcacacaccc aacgtgttct
4320
ttgtcacttg tttggaaatg atgtgattga agattcaga gaggtcattt gaggctttc
4380
tgtgccggca ctgaatgttc atttgcatga ggaagttgca aacgacttct gcaggctgag
4440
attcaaggca ggtggatttgg ggtccctca gcccacctgg gccgtgaccc caagtgtcca

4500
ctgctgagtg tgagtggctt tgcaaggcctg gtgggtggag agcctcaggc tccctccttc
4560
ttcgttcctg accatgcctt gggcccccttc agtctgcctg cggctctgtg gcacccctgc
4620
cctgacactg ggcacctgtg cccttagcaa gcccacctgg cacacgagga gggaggggtg
4680
ggtgccatg tccttcctct agccacatgc cctgctggcc gctccattct gagctttgtg
4740
cagaaccggg tctgagctgg agattttct ctgagaacct gcagttgtgc tgcagccgca
4800
cgcaagggcc cttcagccgc tggctctggc ttccctgact cctcagggcg tgttcaccgg
4860
caggcttct cacctgcaca cggttaggcc atttcagtg ctcgtggca gtcacggaca
4920
gcagcagaaa ctcctcagcc cctttgttac ttcagaacgc ctgcccacat gcacattctg
4980
agctcgtgtt gtcctcatgg ccgtggggtc ttgggtgcga acaggagatg ctgagctgtg
5040
gtccaccgtc caaggtgctg cagaaagcag ctaggctatt ttaggatgtt tctattctgg
5100
ttgctgcctt cgtgggttaa cttaagaa cactacggg aatgtgctca tagaaccatc
5160
acctgtcctg agaataaaac tcctggaatc atgatcaagt ccagtgttaa cgtggcccaa
5220
cctgtctgtt cttctgggaa gagaccagga acatcactgg actcctcatc cccgttaatta
5280
tttagagaag atgcaagcag cagatagtct ccatgcggct ggtactttt ttgttgtttt

5340

ttgagacagg gtcttgcctc gtcacctggg ctggagtgc a g a c g g c g a t c a t g g c t c c c

5400

tgaggcctca acctactagg ctcaagctgt ctgcccgcct tagcctccca agtagctggg

5460

accacaggca cccaccacca ccatgcttgg ctaacttgtt tttgttagaga tggagtttg

5520

ccatgttgc t c a g g t t g g t c t c g a a c t c c c g a t c t c a g g t g a t c c a c c c g c t c

5580

ccaaagtgc t g g a t t a c a g g c g t g a g c c c t g c g c c c a g c t t g g g c c t g t c t t g a

5640

atggaatga gactgtgcaa accgtggact accctgtgtc acccacagct c a g t g g c t g

5700

c c t g c c g g c c c t c a g g g c t g a c c c g g g a g a c c a g c c a g c a c g a g g g c g g g c

5760

t g t g g g t t t t g g c c t g a t c t g c a t t t g g t t t c t g g g g c a t g t a g c c t g c c t g

5820

cattaggaaa g c g c t g t g c c a t c t g a t c a t g a c t g a g c a c c t c t g a c c c c c t g a a g g t g a c

5880

c t t g c a g c a g g a c t g t g c c t g c c t g g g t a g g c a c c c a c t a g g t a g g a c c g a a t c

5940

c t g g c a g c c g c c a c t g c a c c c t g c a c t t g t c c t c a c t a g t t c a a g t a a t c c g t

6000

t t t t g a a g g c t t g t g t g t t t g t g a t t t t g g g a a t a g a g t t g g a c g g a g g c g a

6060

g a g c c t t a a g c c a t g c g a g c t g c g g c c t g g a a c c c a g a c t t c c c a g c t t g a g g g a a

6120

gtgtcagatt tccccgttg acagaaggga gcattgaagg gatgccttgg agcccagaca
6180
gtggttgtcc ctgtgtcctt cccttgacc tggcatcaga ggtgtctcga gtcctaccc
6240
agggacccag aggagttcgg gccccagtag attttcttag atttaagcca aagtgagttg
6300
cattatctgc aacgaggaca gatatggag ggaatgtgct gagagccagg cagatgaact
6360
gaggatctca ttgatcttc ttttgtgtt actaaactca tatgttcttg taaacagttc
6420
tttagcatag acagtgaaag tacccctgt tctcatccca gcctccccgt gagtcactgc
6480
tgctaattaa tgctgttagc ttggaattgt agaaacagga tgtttccat ggtaatgcac
6540
tcaaagtaca ccctcgattt gcagaaattt gcaagtgttga ttttccaagt gttggcagtg
6600
atcagggga acaggaacgc aggtggggca gctgtttgg ggacagctgg tactagctca
6660
tggcactaag gacacgggcc cagggactgg catctgcate ctgaggtgtc caccctcgaa
6720
caacgcgaga gcccaggcat gggccacgca gggatgttca ttgctacact gtgacaactg
6780
tcacaggccg gaaggaggca ggtggactac ggtggagcca cccatgtgt cacctggcag
6840
acgggcacac agccttggc cgttcaaaa caagtgagag atggattttgg tgtaacatgt
6900
aaaaatgcaa atacttaatt tttatcaatt catgtgtggg gaaaagctga agatacgcgt
6960

gggaaatggtg tggtcacttc taggggtgtc ggagggtaga acttcaactg ttttgc
 7020 ttta
 aaaagtaagg atcgcatggc agaactagca tctgttcacc tggatcct gataccgtgg
 7080
 attacgagac cccccctttt ttctgtgtgg ttcagaaaca agcccccac agaggacaca
 7140
 gtgcccccagg gcagtgcacct gcaggccac ccactgccat ctccgcgtgg ctccgggttg
 7200
 ccacatagcc tgccagctgc ggctgcttcc tgggtgcctt ccagggagag cagggatcg
 7260
 tgggtccccg gcgggtgggtg tttccttctc cggggagagc aggggatcgt gggtccccgg
 7320
 tggtgtgggtgt ttccattctct aaggtttgcgt gctgttcca ggcctttctg tggggcctgg
 7380
 gtcctgtcct ggggccaagc cacggggtca tcctcagctg cactgggcgt gccaaccaca
 7440
 aacgagtcac ttgctacaag cagcaccatg cagcctcctg tctggacgag accctgcccc
 7500
 ccacagactg gagacgcacc ccgatttccc aggtcacagg gggaaatgttg gatctgataa
 7560
 gggactaaat gtggcgtctt tcatatgttt ctcttacata ttttattt
 7608
 <210> 20
 <211> 21
 <212> ADN
 5 <213> Homo sapiens
 <400> 20
 gtgaagaagg atgagttctc c
 21
 <210> 21
 <211> 20
 10 <212> ADN
 <213> Homo sapiens
 <400> 21
 agctgagcat cacaatctcc
 20
 <210> 22
 15 <211> 25
 <212> ADN
 <213> Homo sapiens
 <400> 22

ggagcttccc caactcataa atgcc
25
<210> 23
<211> 25
<212> ADN
5 <213> Homo sapiens

<400> 23
gcatgatgtc tgatgtggtc agtaa
25

<210> 24
<211> 20
10 <212> ADN
<213> Homo sapiens

<400> 24
tgcgaaagctg cttcacccat
20

<210> 25
15 <211> 21
<212> ADN
<213> Homo sapiens

<400> 25
ggccgtacat gcactccact g
21

20 <210> 26
<211> 20
<212> ADN
<213> Homo sapiens

<400> 26
gagaacctgc agtttgtgctg
25 20

<210> 27
<211> 20
<212> ADN
<213> Homo sapiens

30 <400> 27
atggtgctgc ttgttagcaag
20

<210> 28
<211> 21
<212> ADN
35 <213> Homo sapiens

<400> 28
tgcccatatg agatgacgag g
21

<210> 29
<211> 21
40 <212> ADN
<213> Homo sapiens

<400> 29

acactcagca gtggacactt g
 21
 <210> 30
 <211> 24
 <212> ADN
 5 <213> Homo sapiens
 <400> 30
 ggcaaatgct ggacccaaca caaa
 24
 <210> 31
 <211> 24
 10 <212> ADN
 <213> Homo sapiens
 <400> 31
 ctaggcatgg gagggAACAA ggaa
 24
 <210> 32
 15 <211> 21
 <212> ADN
 <213> Homo sapiens
 <400> 32
 gactgggctg cgtgctcatc c
 21
 20 <210> 33
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <400> 33
 aggccctgtg gtcactcata ctgc
 25 24
 <210> 34
 <211> 20
 <212> ADN
 <213> Homo sapiens
 30 <400> 34
 aggggctaAC aatggacacc
 20
 <210> 35
 <211> 22
 <212> ADN
 35 <213> Homo sapiens
 <400> 35
 ccgaagataa gggggacta ct
 22
 <210> 36
 <211> 21
 40 <212> ADN
 <213> Homo sapiens
 <400> 36

ccggcgggag gcaggaggag t
21
<210> 37
<211> 24
<212> ADN
5 <213> Homo sapiens
<400> 37
gcggcggtag gtcaggcagt tgtc
24
<210> 38
<211> 20
10 <212> ADN
<213> Homo sapiens
<400> 38
tgcgaaagctg cttcacccat
20
<210> 39
15 <211> 21
<212> ADN
<213> Homo sapiens
<400> 39
ggccgtacat gcactccact g
21
20 <210> 40
<211> 22
<212> ADN
<213> Homo sapiens
<400> 40
gtgaagtgcgg gacgttttgt ga
25 22
<210> 41
<211> 21
<212> ADN
<213> Homo sapiens
30 <400> 41
ccgtgggtggg gccctttgtg g
21

REIVINDICACIONES

1. Una molécula aislada de ácido nucleico de BNO1 que se cartografía en el cromosoma humano 16q24.3 y que comprende la secuencia de nucleótidos expuesta en SEQ ID Números: 1 o 3.
2. Una molécula aislada de ácido nucleico de BNO1 que comprende la secuencia de nucleótidos expuesta en SEQ ID Números: 1 o 3, que codifica un polipéptido capaz de formar parte de un complejo de ubicitina-ligasa implicado en el direccionamiento de proteínas mediante ubicitinación para una degradación en el proteasoma.
3. Una molécula aislada de ácido nucleico de BNO1 que es al menos 95% idéntica a una molécula de ADN que consiste en la secuencia de nucleótidos expuesta en SEQ ID Números: 1 o 3 y que codifica un polipéptido capaz de formar parte de un complejo de ubicitina-ligasa implicado en el direccionamiento de proteínas mediante ubicitinación para una degradación en el proteasoma.
4. Una molécula aislada de ácido nucleico de BNO1 que codifica un polipéptido que tiene la secuencia de aminoácidos expuesta en SEQ ID Números: 2 o 4.
5. Una molécula de ácido nucleico aislada que consiste en la secuencia de nucleótidos expuesta en SEQ ID Números: 1 o 3.
6. Un vector de expresión que comprende una molécula de ácido nucleico tal y como se ha definido en una cualquiera de las reivindicaciones 1 a 5, ligado funcionalmente a elementos de control adecuados.
7. Una célula aislada transformada con el vector de expresión según la reivindicación 6.
8. Una célula según la reivindicación 7, en la que la expresión de BNO1 recombinante se puede interrumpir.
9. Una célula según una cualquiera de las reivindicaciones 7 u 8, que es una célula eucariota.
10. Un método de preparación de un polipéptido codificado por cualquiera de los ácidos nucleicos según las reivindicaciones 1 a 5, que comprende las etapas de:
 - (1) cultivar una célula tal como se define en la reivindicación 7 o 9 en condiciones eficaces para la producción del polipéptido; y
 - (2) recoger el polipéptido.
11. Un polipéptido aislado de BNO1 que comprende la secuencia de aminoácidos expuesta en SEQ ID Números: 2 o 4.
12. Un polipéptido aislado de BNO1, que comprende la secuencia de aminoácidos expuesta en SEQ ID Números: 2 o 4, que es capaz de formar parte de un complejo de ubicitina-ligasa implicado en la degradación de proteínas a través de la ubicitinación.
13. Un polipéptido aislado de BNO1 capaz de formar parte de un complejo de ubicitina-ligasa implicado en la degradación de proteínas a través de la ubicitinación que tiene al menos 95% de identidad con la secuencia de aminoácidos expuesta en SEQ ID Números: 2 o 4.
14. Un polipéptido aislado de BNO1 que consiste en la secuencia de aminoácidos expuesta en SEQ ID Números: 2 o 4.
15. Un anticuerpo para la detección de BNO1 que es inmunológicamente reactivo con un polipéptido tal y como se define en una cualquiera de las reivindicaciones 11 a 14, preferentemente un anticuerpo monoclonal, un anticuerpo humanizado, un anticuerpo químico o un fragmento de anticuerpo que incluye un fragmento Fab, un fragmento F(ab')₂, un fragmento Fv, anticuerpos de cadena sencilla y anticuerpos de dominio único.
16. El uso de una molécula de ácido nucleico según una cualquiera de las reivindicaciones 1 a 5, en la preparación de un medicamento para el tratamiento de un trastorno asociado con una disminución de la expresión o de la actividad de BNO1.
17. El uso según la reivindicación 16, en el que la molécula de ácido nucleico es una parte de un vector de expresión que también incluye elementos de control adecuados.
18. El uso de un antagonista de BNO1 en la preparación de un medicamento para el tratamiento de un trastorno asociado con el aumento de la expresión o la actividad de BNO1, en donde dicho antagonista es un anticuerpo según la reivindicación 15.
19. El uso de una molécula de ácido nucleico aislada que es el complemento de una molécula de ácido nucleico tal y como se define en una cualquiera de las reivindicaciones 1 a 5, cuyo producto de transcripción es un ARNm que se hibrida con el ARNm codificado por BNO1, en la preparación de un medicamento para el tratamiento de un trastorno

asociado con el aumento de la actividad o de la expresión de BNO1.

20. Un método para escrutar un compuesto capaz de modular la actividad de BNO1 que comprende combinar un péptido según una cualquiera de las reivindicaciones 11 a 14, y un compuesto candidato, y determinar la unión de dicho compuesto candidato con dicho péptido.

5 21. Un método para escrutar candidatos a fármacos que comprende las etapas de:

- (1) proporcionar una célula según una cualquiera de las reivindicaciones 7 o 9;
- (2) añadir un candidato a fármaco a dicha célula; y
- (3) determinar el efecto de dicho candidato a fármaco sobre la expresión de BNO1 en dicha célula.

10 22. El uso de un ácido nucleico según una cualquiera de las reivindicaciones 1 a 5, para escrutar candidatos a fármacos.

23. El uso *in vitro* de un ácido nucleico según una cualquiera de las reivindicaciones 1 a 5, para el diagnóstico o el pronóstico de trastornos asociados con una disfunción de BNO1 o una predisposición a tales trastornos.

24. El uso *in vitro* de un polipéptido según cualquiera de las reivindicaciones 11 a 14, para el diagnóstico o el pronóstico de trastornos asociados con una disfunción de BNO1 o una predisposición a tales trastornos.

15 25. El uso *in vitro* de un anticuerpo según se ha definido en la reivindicación 15, para el diagnóstico o el pronóstico de un trastorno asociado con BNO1 o una predisposición a tales trastornos.

26. Un método *in vitro* para el diagnóstico o el pronóstico de un trastorno asociado con mutaciones en BNO1, o una predisposición a tales trastornos en un paciente, que comprende las etapas de:

20 comparar BNO1 o un ácido nucleico que codifica BNO1 procedente de una muestra que se debe obtener a partir de un paciente con BNO1 de tipo silvestre o un ácido nucleico que lo codifica con el fin de establecer si la persona expresa un BNO1 mutante.

27. Un método según la reivindicación 26, en el que se compara la secuencia de nucleótidos del ADN del paciente con la secuencia de ADN que codifica BNO1 de tipo silvestre.

25 28. Un método *in vitro* para el diagnóstico o el pronóstico de un trastorno asociado con una expresión o una actividad anormal de BNO1, o una predisposición a tales trastornos, que comprende las etapas de:

- (1) establecer un perfil para la expresión normal de BNO1 en sujetos no afectados;
- (2) medir el nivel de expresión de BNO1 en una persona sospechosa de tener una expresión o actividad anormal de BNO1; y
- (3) comparar el nivel de expresión medido con el perfil de expresión normal.

30 29. Un método según la reivindicación 28, en el que se emplea una transcriptasa inversa-PCR para medir los niveles de expresión.

30. Un método según la reivindicación 28, en el que un ensayo de hibridación que utiliza una sonda obtenida a partir de BNO1, o un fragmento del mismo, se emplea para medir los niveles de expresión.

35 31. Un método según la reivindicación 30, en el que la sonda tiene al menos 50% de identidad de secuencia con una secuencia nucleotídica que codifica BNO1, o un fragmento del mismo.

32. Un método *in vitro* para el diagnóstico o el pronóstico de un trastorno asociado con BNO1, o una predisposición a tales trastornos, que comprende las etapas de:

- (1) establecer una propiedad física de BNO1 de tipo silvestre;
- (2) medir la propiedad de un BNO1 expresado por una persona que es sospechosa de tener una anormalidad de BNO1; y
- (3) compararla con la propiedad establecida para BNO1 de tipo silvestre con el fin de establecer si la persona expresa un BNO1 mutante.

33. Un método según la reivindicación 32, en el que la propiedad es la movilidad electroforética.

34. Un método según la reivindicación 32, en el que la propiedad es el patrón de escisión proteolítica.

45 35. Un animal no humano modificado genéticamente seleccionado entre el grupo que consiste en ratas, ratones,

hámsteres, cobayas, conejos, perros, gatos, cabras, ovejas, cerdos y primates no humanos tales como monos y chimpancés, transformado con una molécula de ácido nucleico aislada tal y como se define en una cualquiera de las reivindicaciones 1 a 5.

5 36. Un animal no humano modificado genéticamente seleccionado entre el grupo que consiste en ratas, ratones, hámsteres, cobayas, conejos, perros, gatos, cabras, ovejas, cerdos y primates no humanos tales como monos y chimpancés, en el que el gen BNO1 homólogo y la función génica se han desactivado.

37. El uso de un animal no humano modificado genéticamente tal como se define en una cualquiera de las reivindicaciones 35 o 36, en el escrutinio en busca de compuestos farmacéuticos candidatos.

10 38. Una micromatriz para detectar BNO1 que comprende un ácido nucleico que codifica o bien una isoforma de BNO1, o un fragmento de la misma, o ácidos nucleicos que codifican ambas isoformas de BNO1, o fragmentos de las mismas.

Figura 1

Analizado en: serie 1: 189 casos totales

Serie 2: 123 casos totales

no realizado/no interpretable	
NI	no informativo
R	retención

Figura 2

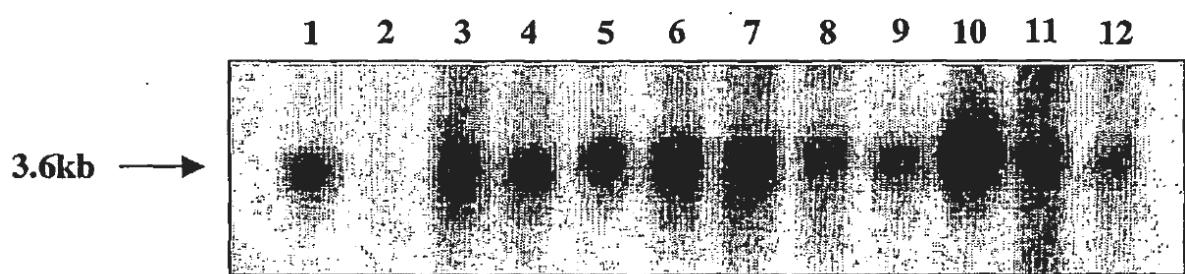


Figura 3

BN01 - R C S L E L P P E L L V E I F A S L P G T D L P S L A Q V C T K F R R I L H T D T I W R R - -
 k p F P L L R L p e E I L r K I L e k L D P i D L L r L R K V S K W R s L V D s 1 n i w f k f I e
 s s s . I s d m l K l i k e v f k h M p f k e R f n F S 1 t C R R F K r i i k k k f k i r k L 1
 r f n i d v n i r r s 1 i k k f 1 n 1 q 1 r d 1 f k d
 a s y e I v s t q r

Figura 4

